LOGIC IN CONFLICT

Logic in Conflict

Logical Explorations in Strategic Equilibrium

Logica in Conflict

Logische verkenningen in strategisch equilibrium

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. Dr. W.H. Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op maandag 6 september 2004 des middags te 12.45 uur door

Bernhard Paul Harrenstein

geboren op 27 juli 1970, te Bussum

promotoren: Prof. Dr. John-Jules Charles Meyer, Universiteit Utrecht
Prof. Dr. Wiebe van der Hoek, Liverpool University
co-promotor: Dr. Cees Witteveen, Technische Universiteit Delft

SIKS Dissertation Series No. 2004-14

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Graduate School for Information and Knowledge Systems.

The research reported in this thesis has been supported by the CABS project at Delft University of Technology.

Printed by Ponsen & Looijen B.V., Wageningen
ISBN 90-393-3753-5

To my parents

Contents

Preface

Game theory provides a framework for the mathematical analysis of interactive situations in which the individual expediency of each agent's choice of action may essentially depend on the choices the other agents make. For this reason it has sometimes been suggested that game theory had perhaps better been called 'interactive decision theory' or 'the analysis of conflict'. We take it that conflict, thus conceived, is a common and natural phenomenon inherent in social interaction.

One of the concerns of formal logic, traditionally the study of valid reasoning, is with formal languages and the extent to which they can be employed to describe and reason about abstract structures in a mathematically precise way. Within computer science logics are commonly used for the specification and verification of computer programs. The behavior of a complex computer system can in some cases be understood as the result of interaction between various autonomous processes and for its precise description an appeal to the conceptual apparatus of the social and economic science has proved to be fruitful.

This forms the background to this dissertation, which tells of an explorative investigation into logic and game theory. Central to its concerns is the game-theoretical concept of strategic equilibrium, which intuitively reflects a state from which no one wishes to deviate by unilaterally making another decision. In the first part of this thesis we give a logical analysis of this notion, using modal logic for the purposes of game theory. The second and third part a perspective is assumed in which game theory serves the purposes of logic. We argue how game-theoretical concepts, in particular notions of strategic equilibrium, can be invoked to enrich logical analyses. This leads up to a proposal for game-theoretical concept of logical consequence.

This dissertation recounts the culmination of these investigations, which were performed in the years 1999 to 2004 at the Institute of Information and Computing Sciences (ICS) at Utrecht University in the Intelligent Systems group of Prof. Dr. John-Jules Meyer. The research is part of by the Collective Agent Based Systems project (CABS) of the Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS) at the Delft University of Technology. The CABS project pursues the development of specification methods and algorithmic techniques for large-scale agent-based systems.

Numerous acknowledgements of support and encouragement during my writing this thesis are in order. First and for most, my gratitude is due to my supervisors John-

Jules Meyer, Wiebe van der Hoek and Cees Witteveen. John-Jules is rightly praised for his effervescence and boundless enthusiasm. His encouragement remained unabated even when, after a few months, my investigations led away from modal specification languages for multi-agent systems — which had originally been envisaged as the subject of my thesis — into the field of game theory. In Wiebe I have always found a person who always took great interest in my work. Our discussions on logic were always laced with humor; we have had many a good laugh over a proof or a counterexample together. I thank Cees for his confidence in my work — even if I could not always redeem my promises to involve coalitions in it — and for him letting me prepare a meal of escargots for his wife Marijke. I will not lightly forget the friendly spirit in which the meetings with the four of us were held.

Special thanks go to Johan van Benthem, who through the years has always encouraged and inspired me. I also like to thank him, Albert Visser, Harrie de Swart and Frank Veltman for their willingness to participate in the reading committee.

The text has also greatly benefitted from the exchange of ideas with colleagues and friends at various times and occasions. I am therefore indebted to Huib Aldewereld, Leon Aronson, Alexandru Baltag, Annette Bleeker, Frank de Boer, André Bos, Jan Broersen, Boudewijn de Bruin, Martin Caminada, Mehdi Dastani, Jurriaan van Diggelen, Frank and Virginia Dignum, Hans van Ditmarsch, Paul Dunne, Francien Deschesne, Rogier van Eijk, Jelle Gerbrandy, Robert Goijers, Davide Grossi, Remko Helms, Koen Hindriks, Joris Hulstijn, the members and PhD students of the ICS, Rosalie Iemhoff, Geert Jonker, Joost Joosten, Barteld Kooi, Meindert Kroese, Roman van der Krogt, Willem Labuschagne, Jeroen Lamers, Henk Jan Lebbink, Alessio Lomuscio, Sieuwert van Otterloo, the PhD students in the Dutch Research School in Logic (OzsL), Marc Pauly, Rohit Parikh, Cees Pierik, Henry Prakken, Roald Ramer, Birna van Riemsdijk, Robert van Rooy, Merlijn Sevenster, the PhD students in the Dutch Research School for Information and Knowledge Systems (SIKS), Richard Starmans, Allard Tamminga, those whom I forgot to mention here, Leon van der Torre, Hans Tonino, Jeroen Valk, Javier Vazquez Salceda, Yde Venema, Gerard Vreeswijk, Wieke de Vries, Mathijs de Weerdt, Marco Wiering and Jonne Zutt. The people in the ICS, in general and Intelligent Systems group in particular, made for a stimulating working environment and lively conversation; I thank them all. Special thanks are due to Wilke Schram, who secured me workplace in the Institute when my appointment expired and the ICS for letting him do so. Lydia Franken always extended the lending period of my books when they threatened to expire, thus making up for my habitual tardiness.

My friends in Amsterdam, Utrecht and elsewhere have played in their various ways played an essential part in the finalization of this thesis. *E.g.*, Vanessa Dirksen was great to share the PhD blues with, Valérie Boor, Arthur Gerla and Job Smeltink were incomparable as co-organizers of the *Whodommit*, Elena Brosio volunteered as my occasional sidekick, and Piter de Weerd has been inimitable in his suggestions for covers and mottos.

Finally I am most grateful to my parents and my sister Petra without whose love and support this book would not be.

Utrecht, July 2004

Introduction and Preliminaries

Chapter 1

Introduction

Conflict of interest is inherent in any social environment in which several volitional individuals interact. Such disputes are nothing to worry about or to be ashamed of, provided that no unseemly means are resorted to in resolving them. The obvious question to ask is how to make the best of such situations and how their excesses can be mitigated.

A situation of conflict may have various possible outcomes, with respect to which the individuals entertain particular and possibly divergent preferences. Moreover, each of the individuals exerts control over some of the variables that determine the eventual outcome and the individual decisions, taken together, determine the behavior of the collective. The way each individual's preferences relate to the individual's powers lends a formal structure to conflict situations that renders them amenable to mathematical investigation. Also other strategic aspects distinguishable in situations of conflict — such as, *e.g.*, the order in which the individuals are to move and the individuals' epistemic characteristics or their attitudes towards risk — allow for formalization. The mathematical analysis of conflict and other situations of social interaction belongs to the subject matter of *the theory of games* or simply *game theory*.

Games of strategy — in contrast to games in which an element of skill is predominant — present examples of conveniently delimitated conflict situations. Hence the name "game theory" as well as the typical accompanying terminology featuring players, wins and losses, strategies and moves. A *strategy* is here understood as a complete plan of playing a game, prescribing a move in every contingency. A *strategy profile* is a selection of strategies, for each player one, determining a unique outcome. These notions will be given mathematically precise definitions presently.

An issue that immediately suggests itself in this context concerns the judiciousness of the different courses of action that are open to the individuals in a particular conflict situation. This is by no means a trivial question and in the present formulation not a particularly clear one. There are various perspectives to take with respect to what judiciousness exactly comes down to and which aspects of conflict situations are to be taken into account. It is a well-known fact of everyday life that everyone pursuing brazenly his own ends may result in a social state for which there exists another that is more desirable for all. The *Tragedy of the Commons*, originally being the overgrazing of communal pastures as a result of each farmer unostentatiously exploiting them more intensively than his sustainable equal share permits, emblematizes a phenomenon that is all too familiar in social contexts. In the more formal setting of the theory of games, the incongruity of individual expediency and social desirability is concisely exemplified by the infamous *Prisoner's Dilemma* (*cf.* Section 1.1 of this introduction). An important task of the theory of games is to provide suitable and mathematically precise concepts to appraise the issue of the various ways in which behavior on the individual and the collective level are interrelated and to study their formal properties. Among these are the so-called *solution concepts*, which, drawing on its mathematical structure, associate with each game a collection of outcomes that are optimal or formally salient from a particular perspective of (individual) expediency.

When assessed from the formal point of view of game theory, conflict situations evince a formal similarity with structures that may be reencountered in domains of research that need not necessarily or exclusively concern *human* interaction as such. Conflict provides a fruitful metaphor for any situation that depends on various variables the control over which is distributed among different forces with individual ends or among processes designed for different purposes. This makes for a wide applicability of game theory also outside the social and economic sciences, for which it had originally been conceived. The employment of game-theoretical concepts and techniques are appropriate for any situation that can be conceived of as a system consisting of multiple active entities whose individually guided behavior determines, at least partly, the behavior of the system as a whole. Thus, game theory has proved itself relevant to such diverse areas of research as evolutionary biology, set theory, logic and, most recently, also to computer science and artificial intelligence.

This thesis is an exponent of the broadening scope of game theory. It concerns both logics designed to reason in a formally precise manner about games and the game-theoretical analysis of propositional logic.

Central to our logical investigations will be the game-theoretical notion of *Nash equilibrium*, it being one of the best-known and most widely applied solution concepts. In the most informal of terms, a strategy profile is a Nash equilibrium if no player benefits relative to her individual preference order by unilaterally deviating from it. As such Nash equilibrium captures a notion of stability for the possible outcomes of a game, though its precise significance continues to be a much discussed and disputed philosophical issue.

In Part I, we will argue that a strategy profile being a (subgame perfect) Nash equilibrium in a particular game reflects in a structural property of a Kripke frame associated with the game in question. We find that this structural property can be characterized by formula schemes in suitably chosen multi-modal logics.

In the latter two parts we come to construe propositional variables of propositional languages as decision variables, each one of which in the control of one of a number of players. This makes that formulas and theories impose a *game-theoretical* structure on logical space, *i.e.*, the set of valuations for the respective language. Part II concerns a class of strictly competitive games in which control over the propositional variables is distributed over two antagonistic players and which we find constitutes a Boolean algebra *modulo* a notion of strategic equivalence. In the final part, control over the propositional variables may be distributed over a larger number of players. Logical space then assumes the character of a non-strictly competitive multi-player game. A notion of strategic equilibrium closely related to Nash equilibrium is then deployed to formulate a game-theoretical concept of entailment, which generalizes classical consequence.

These researches are ultimately inspired by developments within the field of *Distributed Artificial Intelligence* (DAI), which concerns the design and study of environments in which automated intelligent systems or agents interact (multi-agent systems or MAS). This interaction may be between different agents each designed to achieve an individual goal or within a group of agents trying to solve a common problem together. Such distributed environments instance typically the kind of strategic situation that the theory of games is concerned with. Game theory provides Distributed Artificial Intelligence with suitable notions to conceptualize and describe distributed computational environments from a formal and strategic perspective. Reasoning about them formally, however, requires logic.

Before entering upon our logical explorations, some further reflection on the theory of games, the nature and role of its solution concepts and its relevance to logic and DAI is in order.

1.1 Game Theory and Solution Concepts

In their pioneering work *Theory of Games and Economic Behavior* von Neumann and Morgenstern maintained that the development of classical mathematics had to a great extent gone hand in hand with the modern advancement of the natural sciences (*cf.* von Neumann and Morgenstern (1944), pp. 6–7). In comparison, the formulation of proper mathematical concepts for the social sciences had been paid only scant attention to. Still, they claimed that the analysis of situations of conflict faces the mathematician with a conceptually new problem that had been "nowhere dealt with in classical mathematics" (*ibid.*, p. 11). There is no *prima facie* reason to suppose that the mathematical methods developed with a view on applications in the natural sciences would also be suitable for the social sciences. The theory of games was to provide formal and precise concepts to cope with this unfamiliar problem. Indeed, Luce and Raiffa state in their classic introduction to the field that:

> [G]ame theory is one of the first examples of an elaborate mathematical development centered solely in the social sciences. The conception derived from nonphysical problems, and the mathematics [...] was developed to deal with that conception. (Luce and Raiffa (1957), p. 11)

In a situation of conflict, the individuals entertain idiosyncratic preferences as to the possible outcomes of that situation. Still, the individuals exercise in general limited

Figure 1.1. Sherlock Holmes and Watson at Canterbury station in *The Final Problem*.

control over the variables that determine the eventual outcome. We refer to an individual player's choice for the values of the variables in her control as a *strategy*. If there is only one individual — even if she is not in control of all variables and some of them are left to chance — classical mathematics still provides the techniques to calculate which (randomized or mixed) choice of values for the variables in her control guarantee her an optimum outcome. This special case involving a single player only is studied by decision theory. However, things have been argued to change radically if there are multiple players involved. The problem is then that for each individual it may also depend on the strategies the other players adopt which outcome will come about. The best outcome an individual can achieve relative to a particular choice of strategy by the opponents may differ widely in desirability from the best he can achieve relative to another choice of strategy by the opponents. Moreover, which strategy guarantees an individual the best attainable outcome may depend on the strategies his opponents adopt. Waiting in front of the bank may be your best strategy for meeting a person if that person adopts the same strategy. Things, however, are quite different if the other person decides to search for you in the lobby of the hotel. Thus, for each player the optimality of playing a particular strategy may depend on the strategies his opponents choose. If, moreover, we assume that there be some proportionate correlation between the optimality of a player's strategy and that player adopting it, the optimality of the players' strategies may become mutually dependent and a circle becomes apparent (*cf.*. von Neumann (1928), p. 295, and Luce and Raiffa (1957), p. 61).

Consider, *e.g.*, the case of Sherlock Holmes, who, trainbound for the Continent, finds himself being pursued by his murderous adversary Moriarty, who happens to be on another train. If Moriarty gets off at Canterbury, Holmes' optimal strategy is to stay on the train. However, if Holmes acts accordingly, Moriarty's optimal strategy is to remain on his train as well. If, on the other hand, Moriarty decides to remain on the train, Holmes had better make an intermediate stop at Canterbury, jeopardizing the expediency of Moriarty's strategy. Which strategy is optimal for Holmes thus becomes dependent on which strategy is optimal for Moriarty and *vice versa*.[1]

Formally, a situation of conflict can be pictured as a collection of functions, one for each individual caught in the situation. The values these functions take are, moreover, dependent on the values of the same set of variables, the control over which has been distributed over the individuals (and possibly chance). Each of the individual endeavors to maximize his function with respect to his own preferences.[2] Von Neumann and Morgenstern write:

> Thus each participant attempts to maximize a function [...] of which he does not control all variables. This is certainly no maximum problem, but a peculiar and disconcerting mixture of several conflicting maximum problems. Every participant is guided by another principle and neither determines all variables which affect his interest. (von Neumann and Morgenstern (1944), p. 11)

Von Neumann and Morgenstern argue that, due to the mutual dependence of the optimality of the players's strategies, no player can treat the variables controlled by his opponents as statistical parameters, which assume values with a particular probability:

> Every participant can determine the variables which describe his own actions but not those of the others. Nevertheless, those "alien" variables cannot, from his point of view, be described by statistical assumptions. (*ibid.*, p. 11)

New mathematical concepts had to be developed to deal with this kind of problem and take over the role of the optimum, which was no longer thought to be feasible in this context (*cf. ibid.*, p. 39). Because of the formal and structural similarities between strategic parlour games and the more general situations of conflict, as studied by economics and other social sciences, the mathematical theory that had to achieve this was coined the *theory of games*. The solution concepts of game theory are to take over the role of the optimum in "solving" a game.

This view that the multi-player case is of an essentially different nature than the single-player case has met considerable opposition in the past few decades. The point von Neumann and Morgenstern make relies on their objective interpretation of probability. It has been argued that by reverting to a subjective conception of probabilities, as advanced by, *e.g.*, Savage (1954) and Ramsey (1926), it is possible to model an agent's expectation about the variables controlled by his opponents as statistical variables. Then each individual can calculate his optimal strategy in a game as were it a regular decision problem, with him in control of some variables and chance of the remaining ones. Conceived thus, decision theory and game theory are two manifestations

[1] The example is based on Conan Doyles' *The Final Problem*. Also compare Schelling (1960), p.87. Structurally the situation is much similar to the well-known game of *Matching Pennies* (*cf.*, page 121 of this thesis, below).

[2] In von Neumann (1928) a *Gesellschaftsspiel*, or parlour game, is expressly defined in these terms.

of a single theory of rational choice. This point has been emphasized and elaborated upon in Spohn (1982), Bernheim (1984), Pearce (1984), Brandenburger and Dekel (1987) and Tan and da Costa Werlang (1988).[3]

Notwithstanding the legitimacy and the technical and conceptual intricacies of this criticism on von Neumann and Morgenstern's exposition, there is a sense in which a situation of conflict *is* a 'disconcerting mixture of several conflicting maximum problems' to which convolutions no justice is done if construed as a mere optimization problem. Although each individual might be faced with the problem of optimizing a function, there is conceptually no unique principle of maximization for the collection of functions taken together. The fact that we are dealing with a collection of functions in a strategic conflict, makes that there are at least two perspectives, a collective and an individual one, from which to single out some outcomes as somehow optimal or otherwise significant. From the collective perspective, one could try to find a common principle of optimality for all individuals together, with respect to which the optimal outcomes can then be distinguished.[4] From the other, individual, point of view, the optimal outcomes of the combined problem could be taken as those outcomes that combine individually optimal strategies.

The two perspectives suggest different requirements for their respective notions of optimality to comply with. An outcome is called *Pareto efficient* if there is no other outcome in which all individuals are strictly better off. From the collective vantage point, one may look for outcomes that at least comply with this requirement. It be noted in passing that the notion of efficiency does not take into account the control of the individuals over the various variables. On the other hand, an individual's strategy is said to *(strictly) dominate* another of her strategies, if for each possible choice of strategy by her opponents, adopting the former invariably leads to (or is expected to lead to) an outcome that she values higher than the one that would result if she were to adopt the latter strategy. From the individual point of view, the optimal outcomes are to be sought among combinations of individuals' strategies none of which are dominated.

Although the Pareto efficient outcomes and the undominated ones coincide if there is only one individual, they may be even disjoint if there are multiple individuals involved. The latter phenomenon is epitomized by the familiar but illustrative Prisoner's Dilemma, attributed to A.W. Tucker. The story that goes with it is equally well-known; we give here the version of Luce and Raiffa.

> Two suspects are taken into custody and separated. The district attorney is certain that they are guilty of a specific crime, but he does not have adequate evidence to convict them at a trial. He points out to each prisoner that each has two alternatives: to confess to the crime the police are sure they have done, or not to confess. If they

[3] Also compare Franssen (1997), especially Chapter 3.

[4] The problem of combining individual preferences into an acceptable social preference order is a notoriously difficult one, which is studied by social choice theory. One of the most perplexing results of this respect is Arrow's famous theorem, which states the impossibility of a procedure to derive a social preference order from individual values, if the former is to satisfy certain intuitive properties (*cf.* Arrow (1963)). Relaxing the condition that a preference order, collective or individual, should be connected, as we will sometimes do in this thesis, however, sidesteps this issue.

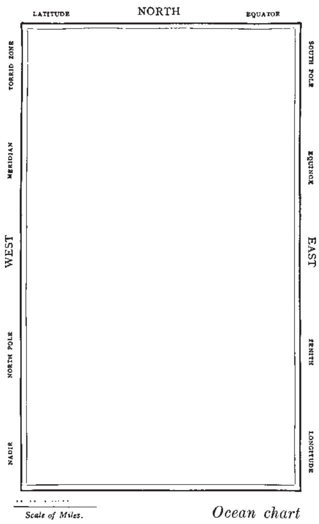
	Deny	*Confess*
Deny	2 2	3 0
Confess	0 3	1 1

Figure 1.2. The game matrix of the *Prisoner's Dilemma*. The figures in the top right corners of the cells indicate the ordinal preferences of the player who chooses columns. The figures in the bottom left corners of the cells those of the player who chooses rows. The outcome that would result if both players avoid playing dominated strategies is represented by the cell bottom right. Yet, both players are better off if the outcome in the top left cell came about. This is a manifestation of a vicious phenomenon inherent in social interaction.

> will both do not confess, then the district attorney states he will book them on some very minor trumped up charge such as petty larceny and illegal possession of a weapon, and they will both receive minor punishment; if they both confess they will be prosecuted, but he will recommend less than the most severe sentence; but if one confesses and the other does not, then the confessor will receive lenient treatment for turning state's evidence whereas the latter will get "the book" slapped on him. (Luce and Raiffa (1957), p. 95)

The awkward situation the prisoners are in is depicted in Figure 1.2, with the figures in the top-right corner of each cell indicating the ordinal preferences of the one prisoner, Bonnie, say, and those in the lower-left corner those of the other, Clyde. Here, all outcomes are Pareto efficient, except the one that results if they both confess. In contrast, for each of them to confess is the dominant strategy. If the one remains silent, other achieves a better outcome by confessing. Also if the one decides to squeal, the other had better do so as well. Still, if the two of them refuse to betray his or her partner in crime, and thus refrain from playing their dominant strategies, an outcome is achieved that is preferred by both.

The Prisoners' Dilemma shows that some outcomes may be salient from the collective perspective and others from the individual and only when considered in unison they may point at socially significant phenomena that would have escaped notice otherwise. Much of the fascination of game theory, methinks, derives from this tension between the collective and the individual level of analysis. Any scientific theory of conflict should provide apposite concepts that do justice to this distinction. Suppose that in the Prisoners' Dilemma, Bonnie and Clyde somehow achieve a Pareto efficient outcome. Any explanation of their behavior should also account for at least one of them playing a dominated strategy. Similarly, if they both play their undominated strategies, one should explain what there was in the situation that made them end up in an outcome that fails to be Pareto efficient.



The Bellman himself they all praised to the skies
Such a carriage, such ease and such grace!
Such solemnity too! One could see he was wise,
The moment one looked in his face!

He had bought a large map representing the sea,
Without the least vestige of land:
And the crew were much pleased when they found it to be
A map they could all understand.

Whats the good of Mercators North Poles and Equators,
Tropics, Zones, and Meridian Lines?
So the Bellman would cry: and the crew would reply,
They are merely conventional signs!

Other maps are such shapes, with their islands and capes!
But weve got our brave Captain to thank
(So the crew would protest) that hes bought us the best–
A perfect and absolute blank!

This was charming, no doubt: but they shortly found out
That the Captain they trusted so well
Had only one notion for crossing the ocean,
And that was to tingle his bell.

(from Lewis Caroll's *The Hunting of the Snark*)

Figure 1.3. The Bellman's ocean chart in *The Hunting of the Snark*.

Thus, the construal of a situation of conflict as a collection of functions each to be maximized according to a different principle, gives rise to different mathematical questions to be asked, depending on the perspective one takes. Appropriate concepts are called for to get to grips with this type of problem and the interaction between the collective and the individual level of analysis. Perhaps the situation can be compared with the sea captain who is used to get his bearings from the stars but now finds himself lost in a heavily forested and mountainous region. Although the stars may be of considerable help to him, he will also have to be proficient in the employment of such concepts as peaks and passes, glaciers and valleys as well as with that of the tree line.

This simile chimes in well with a general image of science propounded by Aumann (Aumann (1985)). What the sciences have in common is that ultimately they mean to improve our understanding of particular phenomena of our world in their abstract and concrete manifestations. At the most fundamental level, the sciences are to develop concepts that help us organize, systematize and reason about the phenomena belonging to a particular field of research and these concepts are to be judged by their success in doing so. The fundamental concepts of a science are not isolated. Rather, the way they are interconnected constitutes the scientific edifice. Moreover, if the concepts involved are of a formal nature, their mutual relationships can be analyzed using mathematical methods. By raising the analysis to a more abstract level, mathematical study of a science's concepts may bring to light expected or unexpected structural correlations with other sciences and open up new fields of application. These remarks hold in particular for the theory of games.

However sweeping these generalities may be, they point at a feature of the function of the game-theoretical solution concepts when actually put to use in a concrete field of application. Game theory provides concepts that help us to get to grips with situations in which a conflict of interests may arise. The various solution concepts facilitate this endeavor by summarizing information present in the formal description of a conflict situation as seen from a particular but strategic perspective. Each of them highlights different specific features and abstracts from others. Without such delimitative concepts we are lost like the Bellman's crew in the hunting for the Snark (*cf.* Figure 1.3).

On this conception, the primary task of game-theoretical solution concepts is not the prediction or description of actual or idealized behavior. Neither is it their role primarily of a normative nature, in the sense that they prescribe how people or rational self-interested agents *should* act in conflictual situations. Rather, they are *indicators*, each of them illuminating a situation from a particular angle and emphasizing some of its characteristics at the expense of disregarding others. They should not primarily be judged by their predictive or normative power but rather by how they help the scientist to get a firm hold on and organize his subject matter. This holds for the social scientists investigating human social behavior and the champions of DAI alike. Aumann puts it as follows:

> People ask, since game theory offers a multiplicity of solution concepts, what good can it be? Which solution notion is the right one? How do people 'truly' behave? [...] None of the solution notions tells us how people truly behave. [...] Rather, a solution notion is the scientists' way of organizing in a single framework many disparate phenomena and many disparate ideas. (Aumann (1985), pp. 34–35)

In another article (Aumann (1997), especially pp. 10–12 and p. 25), Aumann makes an apt comparison between game-theoretical solution concepts and statistical concepts as different as the mean and the median. In virtue of their clear intuitive content, in their own way, they help the statistician — or anybody employing statistical methods for that matter — to attain some kind of hold on various kinds of distribution.

In a similar way, the various solution concepts throw light on the social situations from different angles. The game-theorist develops appropriate solution concepts and investigates their properties and interrelationships. As the concepts are largely of a formal nature, game theory is to a great extent a mathematical affair and the methods employed cannot be too rigorous.

The significance of these concepts for the sciences applying game theory, however, should derive from somewhere else. A solution concept is significant if it helps the working scientist in the field to understand situations of social interaction. Which conclusions she is to draw from the way they are instantiated in a situation of conflict is ultimately up to her own scientific judgement and integrity.

Formal solution concepts single out outcomes that stand out from the others in a conflict situation in virtue of its mathematical description as a game. What conclusions to draw from the way they are instantiated in a particular situation of conflict is ultimately up to the scientist applying game theory. In particular, there need be *no* intrinsic connection between solution concepts and prediction, description or prescription of actual behavior. *E.g.*, in the Prisoner's Dilemma, the outcome with both Bonnie and Clyde speaking is the only outcome that is not Pareto efficient, still it is the only combination of the players' strategies that are not dominated. This says something significant about the situation and may help the scientist to assess the situation. However, what conclusions to draw from these data is not univocal and may depend on the application or kind of explanation the working scientist has in mind.

Game theory is often introduced as making particular *assumptions* as to the rationality of the players. In this context, *e.g.*, expected utility maximization is presented as an assumption that goes with a particular model of human behavior and the employment of a particular solution concept. This then raises the question as to the accuracy of this model and to what extent people are in fact expected utility maximizers. Although this question is interesting enough in itself, the significance of game theory does not hang on its fulfillment. What is at issue are the insights in actual human conduct afforded by the way it relates to the idealized behavior of a perfectly rational expected utility maximizer. This points at a considerably weaker connection. There is nothing wrong with making explicit, say, the epistemological conditions that ideal expected utility maximizing decision makers are to comply with if they are to arrive at an outcome distinguished by a particular solution concept. It is possible to establish such a relation between, *e.g.*, common knowledge of all players being expected utility maximizers and iterated strict dominance. However, it should be borne in mind that this says something about the perspective from which a solution concept assays a situation of distributed decision making, rather than about the assumptions it makes with respect to actual human behavior.

As indicators, the important thing for solution concepts is how they coherently relate to one another in different games and how they relate to other solution concepts in the same game. Moreover, their role in explanations requires there be some balance in how much detail of a situation of conflict they should bring to the fore. On the one hand, a solution concept should be able to make distinctions that are detailed enough to be of interest for the working scientist. On the other hand, by taking into account too many features that are specific to a situation, an explanation may become an *obscurum per obscurius*, explaining the obscure by the even more obscure. After all, one of the canons of explanation is to construe a particular phenomenon as a manifestation of a phenomenon on a more general, more comprehensive and more abstract level.

Solution concepts, in short, chart particular formally remarkable features of a situation of conflict. In the various fields of application of game theory, situations and environments are mathematically represented as games. Conceiving of solution concepts as indicators, they are stripped to their bare mathematical essentials. Emancipated thus from interpretations in terms of actual or ostensibly rational behavior sustains their application also in fields of research other than the social sciences.

1.2 Nash Equilibrium

In the previous section we championed the view of solution concepts as indicators, summarizing information about particular formally salient features of a game. It is, however, not easy to say how they achieve this apart from reiterating the definitions from which they derive their intuitive content and plausibility.

The definition of the pivotal game-theoretical concept in this thesis, *i.e.*, that of Nash equilibrium, has a seemingly clear intuitive content. A Nash equilibrium is a strategy profile from which none of the players of a game has an incentive to deviate unilaterally. An equivalent characterization can be given in terms of a best response of a player against a choice of strategies of his opponents. A Nash equilibrium is then a combination of players' strategies, each one of which constitutes a best response for the respective player against the combination of the other players' strategies it contains. As such it captures a notion of stability or that of a self-enforcing agreement. This way of articulating the perspective Nash equilibrium affords on conflict situations can be refined by specifying the exact epistemic properties of the agents sufficient for the outcome of the respective game to be a Nash equilibrium (*cf.* Brandenburger and Dekel (1987) and Aumann and Brandenburger (1995)).

In spite of these intricate results, however, it is not inherent in its definition what conclusions to draw from a particular strategy profile being a Nash equilibrium in a particular application. In his *Philosophical Investigations*, Wittgenstein compared the different uses of language with the diverse uses of the various tools in a toolbox:

> 11. Think of the tools in a toolbox: there is a hammer, pliers, a saw, a screw-driver, a ruler, a glue-pot, glue, nails and screw.—The functions of words are as diverse as the functions of these objects. [...] For their application is not presented to us so clearly. [...]
> 14. Imagine someone's saying: "All tools serve to modify something. Thus the hammer modifies the position of the nail, the saw the shape of the board, and so on."—And what is modified by the rule, the glue-pot, the nails?—"Our knowledge of thing's length, the temperature of the glue, and the solidity of the box."——Would anything be gained by this assimilation of expressions?—
> (Wittgenstein (1953))

Wittgenstein's musings led him to a philosophy of language that could be summarized by the slogan "meaning is use". We could apply a similar rationale to the game-theoretic solution concepts. Through experience we become conversant with their employment and may gain insight in their significance in and for different situations. Rather than searching for a generic meaning of (a strategy profile being a) Nash equilibrium — which may turn out to be quite spurious anyway — we had better investigate its *conditions of application*. The intuitive content of its definition may serve as a guide in its employment in many contexts, though in some contexts it may be a better guide than in others.

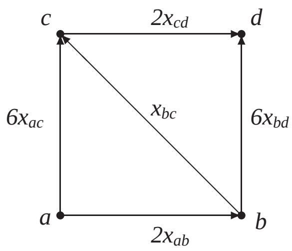
Keeping these remarks in mind, Nash equilibrium can in *normal circumstances* be used as an indicator of rational behavior in situations in which self-interested and utility maximizing agents interact. Especially in two-person strictly competitive games — in which a player benefits only if it goes to the detriment of the other player in an equal measure — Nash equilibrium could be taken to refer to "a kind of mathematical morality, or at least frugality, which claims that the sensible object of the player is to gain as much from the game as he can, safely, in the face of a skillful opponent who is pursuing an antithetical goal" (Williams (1954), p. 23). In the Nash equilibria of such games, both players choose their strategies as to maximize their respective security levels (*cf.* Osborne and Rubinstein (1994), Section 2.5). By the security level of a player's strategy we understand the level of preferability of the least payoff the player can guarantee himself by playing the strategy in question. This so-called maximin solution has got some particularly desirable formal properties. Its existence is guaranteed if the players are allowed to mix their strategies, *i.e.*, to play each of their strategies with a certain probability. Moreover, the equilibria are both *exchangeable* and *equivalent*, *i.e.*, if (α, β) and (α', β') are equilibria, with the first entry denoting the one player's strategy and the second that of the other player, so are (α, β') and (α', β) and, for each player, all equilibria are equally desirable.

In the more general setting in which multiple players interact who may have common as well as opposed interests, Nash equilibria in mixed strategies are still guaranteed to exists, but the other desirable properties no longer hold in general. It has been claimed, and rightly so I daresay, that it is situations that allow for both mutual dependence as well as reciprocal opposition that are of most interest to the student of interactive behavior. In such situations one may encounter Nash equilibria that are Pareto dominated by outcomes that are not in equilibrium. The following well-known example demonstrates in a dramatic fashion this fascinating though perhaps slightly disquieting phenomenon.

Consider the case of two players who are to divide a treasure of a number of gems and jewels. Assume further that they have settled on the following protocol. Alternately, each player has a choice to take either one or two gems. When a player opts for the two gems, the game stops immediately and each player keeps all the gems he has taken so far, with the remaining stones being lost forever. Otherwise, the game continues with the other player making his choice. If the number of gems and jewels is two, one may expect a self-interested player to take them both at the first opportunity. A similar argument holds if there are three gems, for if the first player were to take only one, the second also self-interested player is likely to take the two remaining ones, leaving no jewels for the first player to grab. In contrast, if the treasure is large, consisting of, say, ten thousand stones, one would expect players first to choose one gem for a number of rounds before terminating the game by taking two at a strategic moment. Nevertheless, in all Nash equilibria the first player to move takes two stones at his first opportunity no matter whether the treasure is large or small (and the second player is to take two jewels at his first opportunity). Observe that, no matter how many stones involved, in any possible outcome other than the one ensuing if the second player takes two stones at his first opportunity, both players are better off than when the first player immediately grabs two stones.[5]

---

[5]The reason for this is that in any given strategy of the second player in which he takes one stone until his *n*-th move and then takes two, the first player's best response is to take one stone until his *n*-th move and then take two. Similarly, the second player's best response to any strategy of the first player in which he takes one stone until the *n*-th move and then two, is to take one stone until the *n*-1-st move and then take two. Of course if the first player takes two at his first move, any choice of strategy will guarantee the second player



	acd	*abcd*	*abd*
acd	16, 16	10, 7	8, 8
abcd	7, 10	9, 9	7, 10
abd	8, 8	10, 7	16, 16

Figure 1.4. A simplified case of the Braess paradox. The left figure depicts a graph along which edges two players have to find a way from *a* to *d*. The labels yx_{ij} with which each edge is labelled denotes the cost incurred by each player travelling along it, with x_{ij} being the number of players travelling along the edge *ij*. In the matrix to the right the letter combinations indicate the different routes the players can decide to take. It can easily be established that without the edge *bc* equilibrium will ensue whenever one of the players takes the route *acd* and the other *abd*, both players incurring a cost of 8. However, if the edge *bc* is added, equilibrium results if both players take 'advantage' of this new opportunity at the cost of 9.

The guiding principle in the use of Nash equilibrium as an indicator of rational self-interested behavior seems to go awry in this example and the more so as the number of stones increases. But rather than dismissing Nash equilibrium on this basis as a solution concept, the question to ask here concerns its conditions of use and the idiosyncratic properties of this situation that seem to be strained to their limits. Focussing on this example only, however, will not do. Pareto dominated Nash equilibria pervade the realm of social interaction. It rears its not so pretty head, *e.g.*, also in the field of traffic control and operations research, where the Downs-Thomson and the Braess paradoxes show how increasing the capacity of a link in an (abstract) road network or even adding a new link may actually *increase* each road user's travel time in equilibrium (*cf.* Braess (1968), Arnott and Small (1994) and Figure 1.4). These abstract examples are thought to explain concrete cases of traffic congestion.

A comprehensive survey of these phenomena is beyond the scope of this thesis. There is, however, a trivial but remarkable detail to observe at this point. In situations in which each action that benefits one player is to the detriment of at least one other, all outcomes are Pareto efficient. This implies that the phenomenon of a Nash equilibrium being Pareto dominated can only occur in situations in which the players have some common interests. Moreover, the way the Nash equilibria relate to the Pareto efficient outcomes constitutes a significant feature of the game and to be an important part of

the same number of stones and any of his strategies is just as good a response as the next one. This example is an instance of the centipede game (*cf. e.g.*, Osborne and Rubinstein (1994), pp.106–107).

the conditions of application of Nash equilibrium as a solution concept.

In the tradition of game theory, the preferences of the players over the possible outcomes of a game are traditionally represented by numerical values.[6] This expedites the lavish and fruitful employment of probability theory and the methods of calculus in the interests of game theory. For the purposes of this thesis, however, we construe the players' preferences more generally as relations over the outcomes that satisfy transitivity and reflexivity. In Part I, these relations are moreover assumed to be connected (*viz.*, they are *total preorders* or *quasi-orders*) and the traditional notion of Nash equilibrium is still available. This is not the case for partial preorders, which we contemplate in Part III. To cope with these we come to consider two obvious generalizations of Nash equilibrium: *maximum* and *maximal equilibrium* (*cf.* page 28 for the definitions). In Section 2.1 we find that maximal equilibrium and a similar generalization of Pareto efficiency to partial preorders can be seen as two borderline cases of the same concept.

1.3 Solution Concepts in the Social Sciences

The remarks concerning solution concepts as indicators are especially importunate for the social sciences. The connections and relations established between the events and phenomena as investigated by the natural sciences are to a great extent of an extensional nature, *i.e.*, these relations are to hold between the events independently of the way these events are described. In contrast, the explanations the social sciences are after pertain to the actions people perform. We argue that actions are a special kind of event which require a kind of explanation different from the kind offered by the natural sciences. The explanation of an action is typically in terms of the *reasons* the agent had or may have had for its performance; we say a reason *rationalizes* an action. Whether a particular reason rationalizes a particular action, we claim, however, depends essentially on the way the action is described. Thus, the relation between actions and their reasons, *i.e.*, that of rationalization, is *intensional*. Davidson gives the following example:

> I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am at home. Here I need not have four things, but only one, of which four descriptions have been given. I flipped the switch because I wanted to turn on the light and by saying that I wanted to turn on the light I explain (give my reason for, rationalize) the flipping. But I do not, by giving this reason, rationalize my alerting of the prowler nor my illuminating of the room.
> (Davidson (1980), p.4–5)

What are the precise conditions for a reason to rationalize an action is — I presume — still very much an open question and also falls outside the scope of this thesis. This is not to say that social phenomena can impossibly be explained within a causal

[6] This is not to say that they are essentially quantitative in nature. The players' preferences are usually thought to ensue from their qualitative preferences over lotteries over the outcomes (*cf.*, *e.g.*, von Neumann and Morgenstern (1944), Myerson (1991) and Osborne and Rubinstein (1994)).

framework, but this is actually abstracting away from the features that distinguish it from a physical phenomenon, *viz.*, that actions can be described in terms of reasons.

On this view, an explanation of a particular course of action decided upon by an agent or a group of agents involves making explicit reasons that agent or that group of agents may have for that course of action. Due to the intensional nature of rationalization, an explanation of this type also involves finding appropriate descriptions of the actions. Yet, the adequacy of a description of an action may depend on its taking into account elements that are lost in the mathematical description of a situation. This point can even be put in slightly stronger terms, by stating that the reason rationalizing an action may very well involve very specific features of a situation, the formalization of which may be hard, arbitrary or even spurious. Schelling gives an example in which

> [...] husband and wife, separated in a department store, gaily traipse off to the Lost and Found by a tacit and jocular mutual appreciation that it is the "obvious" place to meet, [whereas] two mathematicians in the same situation — each aware that both are mathematicians — might look for a geometrically unique point rather than one that depended on a play on words. (Schelling (1960), p.114)

Assume for the sake of argument that both husband and wife and the mathematicians manage to meet at the supposed locations. Explanations of these occurrences should involve the reasons those separated had for their actions. Moreover, the reasons the former pair had for their actions are no inferior to those of the latter and, yet, all depended on specific peculiarities of the situation and the individuals. In a preceding paragraph Schelling writes:

> It is that the mathematical properties of a game, like the aesthetic properties, the historical properties, the legal and moral properties, and all the other suggestive and connotative details, can serve to focus the expectations of certain participants on certain solutions. (Schelling (1960), p.113)

If game theory is seen as primarily an economic framework, Aumann is in an important sense right in saying that "We must get used to the fact that economics is not astronomy, and game theory is not physics" (Aumann (1985), p.37).

These considerations are by no means meant to attenuate the role of formal game theory. They rather accentuate the role of the formal solution concepts as indicators, instead of as predictors. Consider phenomena as threats and promises, deterrence and inducement, coordination and commitment or relinquishing the initiative. If anything, these are interesting issues from a game-theoretical perspective. Both promises and threats are ways of an agent to conditionally commit herself to a particular course of action. Interestingly, the success of a threat does not depend on its being fulfilled; if a threat is efficacious, it deters the other party to take another course of action and threatener is no longer committed to carry out the threat. Typically, a threat deters "through its promise of mutual harm". In contrast, if a promise produces the desired behavior in the promisee, one remains committed to act in a particular way.

	Left	*Right*
Top	1, 0	3, 2
Bottom	0, 1	2, 3

	Left	*Right*
Top	1, 0	3, 1
Bottom	0, 3	2, 2

	Left	*Right*
Top	3, 0	1, 2
Bottom	2, 1	0, 3

Figure 1.5. Formal conditions for promises and threats. In the left matrix the column player may threaten to choose the left column, if the row player chooses *Top*. In the middle game, a similar threat is only likely to deter *Row* from playing *Top*, if it is accompanied by a conditional promise to play *Right* otherwise. In the situation on the right, neither playing can try to achieve a better outcome by posing a threat or making a promise.

In some situations an individual may have good reason to make a threat or a promise whereas in others it does not quite make so much sense.[7] The reasonableness of posing a threat or making a promise may very well depend on the mathematical structure of the game and the mathematical structure of the game can in turn be assessed using solution concepts. *E.g.*, when asked why she made a promise rather than a threat in a particular situation, an individual might reply that she expected, not knowing of an ulterior motive for him to do otherwise, her opponent to play his dominant strategy, that she could achieve a better outcome if he played another strategy and that, moreover, in virtue the mathematical structure of the situation a promise, in contrast to a treat, could be effective in this respect.

To illustrate this point consider the leftmost matrix in Figure 1.5. There *Top-Right* is a combination of the players dominant strategies. Still, the individual choosing columns, *Col*, could try to deter the other party from choosing the top row by threatening to play the left column in that case. If the threat is yielded to, *Col* is no longer committed to choose the left column and can thus achieve his best possible outcome, *viz.*, *Bottom-Right*. Observe that in this situation a promise would dissuade neither player from playing his dominant strategy. In the middle game, a similar threat by *Col* to choose the left column is likely to deter *Row* from playing *Top*, only if it is accompanied by a conditional *promise* to play *Right* otherwise. The cell *Top-Right* represents the outcome that will result if both players play their dominant strategies. Yet, neither promise nor threat will induce either player to play another strategy. In contrast, Bonnie and Clyde have no reason to threaten one another in the Prisoner's Dilemma (*cf.* Figure 1.2, above). Yet, they may achieve the outcome that Pareto dominates *Bottom-Right* if either of them can make a credible promise to remain silent if the other does

[7] The following examples are from Schelling (1960), to which the reader is referred for a more elaborate account of this kind of phenomenon.

so as well. Finally, the formal structure of the situation depicted by the matrix on the right in Figure 1.5 precludes the possibility of either player making an effective threat or promise.

These examples illustrate how the formal notion of dominance may help one to reason about such phenomena as threats and promises. In each case the argumentation points at the *formal* basis — or the lack thereof — for posing threats or making promises. Whether the agents actually do, will or should act accordingly, however, is at least contingent on the extent in which they can make credible their conditional commitment to a particular course of action. This, however, may very well depend on characteristics of the situation which one cannot or would not wish to account for in the mathematical structure of the game.

One of the aspects we have abstracted from in our exposition so far, however, is that of the sequential order in which the players are to act. Of course, a conditional threat is not likely to be effective if the threatener is to act first and the threatened party after. A similar thing holds for promises. Moreover, the sequential structure of a game could very well be accounted for in the mathematical representation of a game, giving rise to the notion of a game in *extensive form*. It is with this type of game the first part of this thesis is concerned with.

1.4 Game Theory, Logic and Artificial Intelligence

In the previous section we argued that the explanation of human behavior requires taking into account features of a situation that do not lend themselves for a sensible formalization. These reflections are meant neither to dispute the usefulness of the formal concepts of game theory to the social sciences nor to question the success of their application there. Rather they are meant to contrast the use of game theory in the social sciences with that in other fields of application that are amenable to a more complete formalization, as, *e.g.*, distributed computing. Whereas in the social sciences game-theoretical analyses reveal formal structures that may be invoked for a more comprehensive understanding of human behavior in conflict situations, there is a more direct match between the concepts of game-theory and the interactive behavior of computerized systems. The following quotation gives voice to this observation:[8]

> Most economic models assume idealized, rational decision makers interacting in narrow, precisely prescribed ways. These assumptions, while critical to the tractable exposition and implementation of any theory, often fail the test of descriptive adequacy. However, what may be unrealistic with respect to rich environments populated by imperfectly understood interacting human agents, may often provide adequate descriptions of restricted environments populated by formally specified interacting computational agents. (Boutilier, Shoham, and Wellman (1997), p.4)

[8]The passage also has an apparent critical undercurrent with respect to the employment of formal methods in the social sciences from which the author dissents.

Game-theoretical concepts may successfully be deployed in the design of autonomous computational systems that are to operate in interactive situations as well as in fully fledged multi-agent systems. In the latter case the issue is to create environments for multiple agents with given preferences to interact in such a way that one may expect the outcome of the interaction to possess certain desirable features (*mechanism design*). The specification of auction protocols in which none of the agents has an incentive to conceal the maximum price it is prepared to pay for a particular good is a case in point. The design of such and similar systems involving multiple agents requires suitable models of interaction. Moreover, their formal specification and verification call for logical frameworks enabling precise mathematical analysis of these systems with respect to their game-theoretical properties.

To illustrate this point consider, *e.g.*, a glutton and a gobbler about to dispatch a square cake that can be cut by them simultaneously. Both eat more cake rather than less. As they both like to have a chance of obtaining a large piece and do not wish to settle for half the cake from the outset, they agree to divide the cake according to the following protocol. The glutton and the gobbler cut the cake simultaneously, the latter vertically and the former horizontally. This results in the cake being divided in four, not necessarily equal, pieces. The glutton obtains the left upper and the lower right part of the cake and the gobbler whatever remains. The situation is depicted in Figure 1.6. A Nash equilibrium results if both cake aficionados separately settle on a strategy that may be expected to cut the cake in half. Both players may then look forward to half a cake, and observe that this will not change if one of them deviates unilaterally. To appreciate this, first observe that both of the two players can guarantee the cake to be divided equally, by cutting himself the cake in two equal parts, no matter what strategy his opponent adopts. Now suppose that one of the gorgers decides to play a strategy that cannot be expected to divide the cake evenly. Then, his adversary obtains more than half the cake if he also adopts a suitable strategy that divides the cake unequally. Suppose, *e.g.*, the gobbler cuts the cake vertically such that the left piece is larger than the right piece. If the glutton adopts a strategy that results in the gobbler obtaining less than half of the cake, the latter has reason to deviate unilaterally; he obtains a full half of the cake by cutting the cake vertically in two equal pieces. If, on the other hand, the gobbler's division together with the cut of the glutton apportions the glutton half or less of the cake, the glutton has good reason to deviate unilaterally. Given the gobbler's cut, he would have obtained a larger piece had he decided to make the upper part of the cake larger. This argument can be generalized as to apply to all courses of action in which one of the players fails to divide the cake evenly; in any such case no Nash equilibrium ensues. We can prove that in all Nash equilibria of this protocol the cake is divided equally and this can be taken as an indication that the protocol is fair.

A similar call for game-theoretical methods and concepts in Artificial Intelligence emerges if *constraint satisfaction problems* are considered in which the control over the relevant variables is distributed over multiple agents. In case the variables are binary, propositional logic can be employed to model such problems (*cf.* Yokoo, Durfee, Ishida, and Kuwabara (1998), Walsh, Yokoo, Hirayama, and Wellman (2001), Walsh and Wellman (2000)). On this view, the distribution of propositional variables obtains

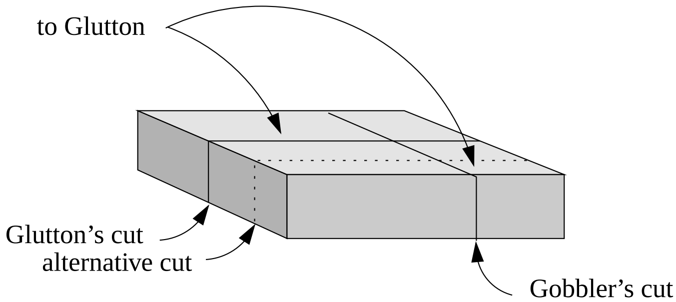


Figure 1.6. Division of the cake according to protocol. the glutton obtains the upper left and lower right pieces and th gobbler gets the remainder. Observe that given the gobbler's cut slightly off-center to the right, makes that the glutton had better make the alternative cut even though still better alternatives are available.

a logical significance and game theory may be invoked for a proper grasp of this phenomenon.

These reflections constitute background of the investigations presented in this thesis, in which both game theoretical concepts are subjected to logical analysis and logic is subjected to game-theoretical analysis.

1.5 Logic and Game Theory

Connections with games are by no means alien to the history of symbolic logic nor to its more recent developments. Pauly (2001) distinguishes two points of view in this line of research: the use of game theory for the purposes of logic and the deployment of logic for the purposes of game theory. The use of Ehrenfeucht-Fraïssé games in model theory (Hodges (1993, 1985); Doets (1996)) comes under the first heading. Lorenzen's dialogue games in constructive proof theory (Lorenzen and Lorenz (1978) and Hintikka's game-theoretical semantics (Hintikka (1983)) have exposed the allegedly fundamental interactive foundation of logic. Assuming the second perspective, modal logics have been employed to formally characterize the epistemic requirements on the part of the players for the outcome of the game to be guaranteed to satisfy a particular solution concept. Dynamic epistemic logics have been used for the analysis of knowledge and belief change in particular game settings (Baltag (2002); van Ditmarsch (2000)). In this context should also be mentioned Pauly's Coalitional Logic (Pauly (2001)), Parikh's Game Logic (Parikh (1984, 1985)) and Boudewijn de Bruin's analysis of the epistemic and rationality assumptions inherent in game-theoretical solution concepts (de Bruin (in preparation)). Their work develops modal logics with expressive power with respect to non-epistemic features of games. In the numerous papers by van Benthem (*cf.*, *e.g.*, van Benthem (to appear, 2001a, 2002)) neither perspective takes precedence and logical and game-theoretical structures are compared and analyzed on an equal footing.[9]

In this thesis both points of view are assumed. In the first part, we assume the first perspective and argue that extensive games — *i.e.,* games in which the sequential order of play has been made explicit — are relational structures and that modal languages can be employed to describe and reason about them. The main results comprise a modal characterization of Nash equilibrium (as well as that of a closely related solution concept called subgame perfect Nash equilibrium) and the soundness and completeness of an axiomatization of the accompanying logic.

The other point of view is predominant in the remaining two parts of this thesis. Propositional logic is subjected to a game-theoretical analysis, in which extensive use is made of equilibrium concepts. Here the underlying thought is to conceive of propositional variables as binary decision variables, the control over which is distributed among various decision making entities. Logical space, *i.e.,* the set of valuations, then assumes a game-theoretical and interactive character, spawning a number of logical issues. On this basis, game-theoretical extensions of the classical notions of validity and consequence are defined and studied.

In the second part, control over the propositional variables is distributed over two antagonists. The one aims to verify a formula by choosing appropriate values for the variables assigned to her, whereas the other endeavors to falsify the same formula by choosing values for his variables. This gives rise to the concept of a Boolean game and the related concept of relativized consequence. In this manner the concept of *control* is brought within the scope of (propositional) logic. These logical inquiries in this part are the preamble to the third part, in which the idea of distributed control is extrapolated to many-player environments. The reflections on distributed propositional control eventually lead up to the issue Chapter 9 is concerned with: *Which conclusions is one to draw from a family of theories, given that, for each of these theories, there is a player who controls a (disjoint) set of propositional variables and who seeks to satisfy his theory as well as he can by choosing appropriate values for the variables in his control?* To assess this problem, the game-theoretical concept of a *maximum equilibrium* — a generalization of Nash equilibrium to be introduced presently — is resorted to. We propose an accompanying notion of consequence, *game-theoretical consequence*, and study its formal properties in some detail. For the notion of game-theoretical consequence there are various possible definitions, involving different game-theoretical solution concepts. We have chosen for the option that is closest to classical logic, as to assure that the features that are specific to the framework can indeed be ascribed to the game-theoretical perspective taken and not so much to non-standard features of the underlying propositional logic.

The emphasis in logical investigations relating to game theory has traditionally been games in which only two antagonistic individuals figure, only one of which can win. These games constitute a proper subclass of two-person strictly competi-

[9] For more extensive and comprehensive overviews, the reader be referred to van Benthem (2001b) and Hintikka and Sandu (1997), Section 3.

tive games. We observed that this type of game has some particularly elegant and illuminating formal properties, granting them a prominent role in the development of game theory (*cf.*, page 14, above). Nevertheless, the strategic problem derives much of its significance and fascination from its ability to deal with situations in which the individuals have both common and opposed interests. Accordingly, in this thesis our logical studies will to focus largely on multi-player games and the accompanying solution concept of Nash equilibrium.

1.6 Overview

This dissertation is organized in three parts, in each of which logic and game theory are related in a different way.

The first part concerns extensive games with perfect information and a finite horizon, being a class of games that is associated with a proper subclass of Kripke structures for a specially designed multi-modal language. The focus is on the logical analysis of the notion of Nash equilibrium and its subgame perfect refinement. This part is a reorganized version of Harrenstein, van der Hoek, Meyer, and Witteveen (2003). Many of the underlying ideas stem from the earlier papers Harrenstein, van der Hoek, Meyer, and Witteveen (2000) and Harrenstein, van der Hoek, Meyer, and Witteveen (2002).

In Chapter 3, we prove that a strategy profile being a (subgame perfect) Nash equilibrium in a game is reflected by a particular structural property of the associated Kripke frame. This property is characterized by a formula scheme of the multi-modal language. We also show how this analysis can be executed using the language of propositional dynamic logic (PDL).

While Chapter 3 is mainly concerned with semantical issues related to the characterization of (subgame perfect) Nash equilibrium, Chapter 4 is devoted the completeness of an axiomatization for the ensuing multi-modal logic. A construction method is employed in the proof of this result. The main problem encountered in the completeness proof is to ensure that the model constructed belongs to the subclass of Kripke structures corresponding to the class of extensive games.

In the second and third part, the emphasis is shifted to the game-theoretical analysis of logic and the logical issues elicited by the particular game-theoretical view on logic taken. The thought underlying both parts is that the control over the values of the propositional variables of a propositional language can be thought of as being distributed among various individuals. The different choices an individual can make with respect to the variables in his control define a set of strategies he can choose from. The different sets of valuations he can thus guarantee the outcome of the game to end up in determine his manipulative powers. On this conception, the valuations used to interpret the propositional variables are construed as the strategy profiles of a strategic game. In other words, by distributing control over the variables logical space assumes a game-theoretical aspect.

In Chapter 5 we introduce a class of strictly competitive two-person games in which control over a set of binary decision variables is divided among two antagonistic players. The outcomes of any of these games are of two kinds only: victories for one player and victories for the other. Draws are not possible. Each set of binary decision variables thus defines a set of Boolean games.

We argue that Boolean games can be seen as representing the information structure of finite games of *imperfect information*. A Boolean game consists of a Boolean game and a distribution of the propositional variables over the two players, player 0 and player 1. We prove the Boolean forms to constitute a Boolean algebra, *modulo* a suitable notion of strategic equivalence. This Boolean algebra, moreover, is isomorphic to the Lindenbaum algebra of the propositional language with the binary decision variables as propositional variables. Each propositional formula then corresponds to a Boolean form.

The correspondence between propositional formulas and Boolean forms makes that control over propositional variables can be studied from a logical angle as well. This consideration engenders the notion of relativized logical consequence as advanced in Chapter 6, which defines a relation between propositional theories relative to each distribution of the propositional variables. This notion of relativized consequence generalizes the relativized concept of validity, which is such that, for each subset Δ of the propositional variables, a formula φ is Δ-valid if and only if the player 1 has a winning strategy in the Boolean game on the form corresponding to φ provided she has control over the propositional variables in Δ. Moreover, we find that the relativized notion of propositional consequence is a conservative extension of the classical notion of consequence. The work in Part II draws on material presented in Harrenstein, van der Hoek, Meyer, and Witteveen (2001).

In the remaining three chapters we pursue the idea of distributed control over propositional variables as a concept that is amenable to logical analysis. Boolean games are strategic games with the valuations of the respective propositional language as strategy profiles. As such a Boolean form, which corresponds to a propositional formula, and a distribution of the propositional variables impose a game-theoretical structure on logical space. In Part II the emphasis was on the logical properties of the formulas corresponding to the game-theoretical properties of the associated Boolean form, given a particular distribution of the propositional variables. The perspective taken in Part III is slightly different. The game-theoretical structure imposed by a Boolean game on logical space allows particular valuations to be singled out by means of game-theoretical solution concepts and one can investigate which formulas hold in the valuations that stand out in this way. In Chapter 7 it is argued that on this basis concepts of consequence can be defined. As a first example, the notion of *winning consequence* is advanced to illustrate the underlying thoughts. We find that also this notion conservatively extends classical consequence. Examining its formal properties in some detail, we eventually present a sound and complete Gentzen-style system for this concept.

The conceptualization of winning consequence, however, lends itself for generalization. It will be argued that, given a distribution of control over the propositional variables, theories and families of theories can be employed to define strategic games that have the valuations as strategy profiles. Moreover, these games may involve multiple players whose preferences need not be antagonistic. Then also more sophisticated game-theoretical solution concepts can be invoked to define consequence relations.

Chapter 8 initiates the notion of a *distributed evaluation game*, which constitutes the semantical basis for the concept of *game-theoretical consequence* advanced in the final chapter. A distributed evaluation game defines a strategic game on logical space on the basis of a distribution of the propositional variables over a number of players and a family of theories indexed by the set of players. The semantical ideas underlying this definition evince, at an abstract level, particular similarities with some proposals to formalize non-standard reasoning mechanisms in the field of philosophical logic and Artificial Intelligence. Veltman's proposal for an update semantics for defaults (Veltman (1996)) serves as a particular good example, in this respect. In Section 8.5 an effort is made to delimitate formally the class of distributed evaluation games within the more comprehensive class of strategic games with the valuations of a propositional language as strategy profiles. The material of this section, however, is inessential for a proper understanding of the subsequent chapter on game-theoretical consequence.

By now the stage has been set for the final chapter, in which the notion of game-theoretical consequence is introduced. Game-theoretical consequence relates families of theories relative to a distribution of control over the propositional variables among a number of individuals. The traditional problem of consequence can be understood as pertaining to the conclusions one may reasonably draw from a theory. Game-theoretical consequence bears on the more general issue which conclusions one may reasonably draw from a family of theories, given that for each theory there is an individual who strives to satisfy it by choosing appropriate values for the propositional variables he is assigned control over. Any such problem defines a strategic situation that can be modelled and evaluated as a distributed evaluation game.

Chapter 9 presents game-theoretical consequence as the logical offshoot of one of the possible ways to resolve this problem, *viz.*, the one by means of the game-theoretical solution concept of maximum equilibrium. Then, this notion is subjected to a formal analysis. Moreover, we find that game-theoretical consequence can be embedded in classical consequence and *vice versa*. The proof of this result relies on a semantical interpretation of game-theoretical consequence using the apparatus of rough set theory. The material of this chapter has been presented in a condensed form at the LOFT5 conference in 2002 (*cf.*, Harrenstein (2002)). A compendious statement of its main tenets can also be found in Harrenstein (to appear-a).

The three parts are largely self-contained and as such can be read independently of one another. The next chapter constitutes the preliminaries to the main body of work. They may be skipped on first reading and consulted when need be. Be that as it may, the results presented in the preliminaries may conduce to a better understanding of, in particular, the third part. After introducing the notions used in this thesis, strategic games and related concepts are defined. It be observed that our notion of a strategic game differs from tradition in that the players' preferences are not required to be a connected relation over the outcomes. This requires the generalization of the concept of Nash equilibrium, giving rise to the definition of maximum and maximal equilibrium. A similar remark concerns the concept of Pareto efficiency, which plays a lesser role in this thesis. Section 2.2 presents some elementary definitions and results of rough set theory and Section 2.3 concerns some basic facts of propositional logics. Section 2.4 deals with a semantical analysis of propositional logic using rough sets.

Chapter 2

Preliminaries

The material in this chapter is basically for referential purposes and had perhaps better be skipped on first reading. Unless otherwise stated, all proofs are by the author. Due to their elementary character, however, no originality can be claimed by him.

2.1 Strategic Games and Maximum Equilibria

We define a *strategic game* as a tuple $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$, where N is a countable non-empty set of players and for each player i in N and S_i is a non-empty set of strategies available to i. Accordingly, the generalized Cartesian product the various S_i, *i.e.*, $\prod_{i\in N} S_i$, is the set of *strategy profiles* of the game, which we also denote by S. The pair $\left(N, \{S_i\}_{i\in N}\right)$ we call the *frame* of the game $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$. For each $i \in N$, ρ_i is the empty relation or, otherwise, a reflexive and transitive, but not necessarily connected relation on the strategy profiles S. In this thesis, a relation that is either the empty relation or both reflexive and transitive we will also refer to as a *proto-order*.[1] We also use $\leqslant_i$ as the infix notion of ρ_i. Hence, S could considered to be an $\|N\|$-dimensional space with for each strategy profile s and each player i in N, s_i its i-th coordinate. We will adopt the notation (s_{-i}, s'_i) for the point that is like s except for the i-th coordinate, which is identical with the i-th coordinate of s'. Intuitively, each (s_{-i}, s'_i) denotes a strategy profile that player i can reach from s by unilaterally deviating.

At this point it should be emphasized that, although reflexive and transitive if not empty, the preference orders as defined by theories are not in general connected. This is at variance with the usual assumptions made in the theory of games. The game-theoretical solution concepts are likewise defined for connected preference relations and we see ourselves bound to generalize them in such a way that they apply to game

[1] Defined thus proto-orders satisfy transitivity and the condition that $(elementvar, y) \in \rho$ and $(y, z) \in \rho$ imply $(x, z) \in \rho$. Relations for which these two conditions hold are called *preorders* in Kuratowski and Mostowski (1976). In this thesis, however, we will reserve the concept of preorder for reflexive and transitive relations.

with partial preference orders as well.

The notion of a *Nash-equilibrium* in pure strategies is usually defined on games in which the preferences of the players are total pre-orders over the strategy profiles. Then, for a game $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ and any strategy profile s:

$$s \text{ is a \textit{Nash-equilibrium}} \quad \text{iff} \quad \text{for all } i \in N, \text{for all } s' \in S: \ (s_{-i}, s'_i) \leqslant_i s.$$

A strategy profile s is, or contains, *a best response for a player i* if for all strategy profiles s' in s, $(s_{-i}, s'_i) \leqslant_i s$. Obviously, the set of strategy profiles that contain a best response for each player coincides with the set of Nash-equilibria.

Since, however, our investigations concern games in which the players' preference relations are also allowed to be proto-orders over the strategy profiles, we are now confronted with at least two obvious conservative extensions of the notion of a Nash-equilibrium. On total pre-orders the notions of a maximal element (no other element is greater) and a maximum element (greater than any other element) coincide, but on partial pre-orders or the empty relation they may diverge. Similarly, we define for any player i and strategy profile s:

$$s \text{ is a \textit{maximal response for i}} \quad \text{iff} \quad \text{for no } s' \in S\colon s <_i (s_{-i}, s'_i),$$

$$s \text{ is a \textit{maximum response for i}} \quad \text{iff} \quad \text{for all } s' \in S: \ (s_{-i}, s'_i) \leqslant_i s.$$

Lacking connectivity, the set of maximal responses for a player i, however, may contain elements s and s' that are incomparable for i (on the i-th coordinate) but which are such that $s_j = s'_j$, for each $j \neq i$. This possibility is excluded for maximum response strategies. Accordingly, we introduce the concepts of a *maximal* and a *maximum* equilibrium as the intersections of the players' maximal and maximum response strategies, respectively. Both are (conservative) extensions the original definition of a Nash-equilibrium. Hence, for s a strategy profile in a game G, we define:

s is a *maximal equilibrium* in G iff s is a maximal response for all players i,

s is a *maximum equilibrium* in G iff s is a maximum response for all players i.

Observe that a strategy profile being a maximum equilibrium implies its being a maximal equilibrium but not in general the other way round. Observe further that by refining the preference orders of the players — *i.e.*, if the preference relations become smaller — the number of maximal equilibria may increase, this is impossible with maximum equilibria. Hence we have the following monotonicity property only for maximum equilibria.

Proposition 2.1.1 *(Monotonicity of maximum equilibria) Let G and G′ be the games $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ and $\left(N, \{S_i\}_{i\in N}, \{\rho'_i\}_{i\in N}\right)$, respectively. Let, further, for each player i, $\rho'_i \subseteq \rho_i$. Then:*

s is a maximum equilibrium in G′ implies s is a maximum equilibrium in G.

Proof: Consider an arbitrary strategy profile s which is *not* a maximum equilibrium in G. Then for some player i and for some strategy profile s', $\big((s_{-i}, s'_i), s\big) \notin \rho_i$. Since $\rho'_i \subseteq \rho_i$, $\big((s_{-i}, s'_i), s\big) \notin \rho'_i$. Hence, s is not a maximum equilibrium in G' either. $\dashv$

Pareto Efficiency

Other important economic concepts are those of *Pareto efficiency* and *strong Pareto efficiency.* Intuitively, a state s, or strategy profile, is *Pareto efficient* if it there is no other state, or strategy profile, which every individual strictly prefers to s. A state s is *strongly Pareto efficient*, if for every state s' that some player strictly prefers to s there is another player that strictly prefers s to s'.

These notions are usually defined for individual preference orders that are *total*, *i.e.*, for every pair of states s and s', a player either values s at least as high as s', or the other way round, *i.e.*, either $s' \leqslant_i s$ or $s \leqslant_i s'$. For such total individual orders it is in general the case that $s <_i s'$ if and only if $s' \nleqslant_i s$. Accordingly the following two definitions of Pareto efficiency are equivalent for total individual preference orders.

$$\begin{array}{lll} s \text{ is } \textit{Pareto}_1 \textit{ efficient} & \text{iff} & \text{for no } s' \in S, \text{ for all } i \in N\colon s <_i s', \\ s \text{ is } \textit{Pareto}_2 \textit{ efficient} & \text{iff} & \text{for all } s' \in S, \text{ for some } i \in N\colon s' \leqslant_i s. \end{array}$$

A similar remark applies to the following definitions of strong Pareto efficiency.

$$\begin{array}{lll} s \text{ is } \textit{strongly Pareto}_1 \textit{ efficient} & \text{iff} & \text{for all } s' \in S, \text{ for all } i \in N\colon \\ & & \qquad s <_i s' \text{ implies for some } j \in N,\ s \nleqslant_j s', \\ s \text{ is } \textit{strongly Pareto}_2 \textit{ efficient} & \text{iff} & \text{for all } s' \in S, \text{ for all } i \in N\colon \\ & & \qquad s' \nleqslant_i s \text{ implies for some } j \in N,\ s \nleqslant_j s', \\ s \text{ is } \textit{strongly Pareto}_3 \textit{ efficient} & \text{iff} & \text{for all } s' \in S, \text{ for all } i \in N\colon \\ & & \qquad s <_i s' \text{ implies for some } j \in N,\ s' <_j s, \\ s \text{ is } \textit{strongly Pareto}_4 \textit{ efficient} & \text{iff} & \text{for all } s' \in S, \text{ for all } i \in N\colon \\ & & \qquad s' \nleqslant_i s \text{ implies for some } j \in N,\ s' <_j s. \end{array}$$

Having assumed the number of players to be greater than zero, it can easily be verified that strong Pareto efficiency implies Parato efficiency.

If, however, the individual preference orders are allowed to be partial — *i.e.*, if it is not in general the case that $s <_i s'$ if and only if $s' \nleqslant_i s$ — the various definitions of Pareto efficiency diverge. Of all Pareto notions strong Pareto$_4$ efficiency is the strongest in the sense that any strategy being strongly Pareto$_4$ efficient implies that strategy to be Pareto efficient in any of the other five ways as well. The implications are strict, strong Pareto$_4$ efficiency is not in general implied by any of the other notions. Strong Pareto$_3$ efficiency implies both strong Pareto$_1$ and strong Pareto$_2$ efficiency, but not the other way round. Strong Pareto$_1$ and strong Pareto$_2$ efficiency are equivalent and both strictly

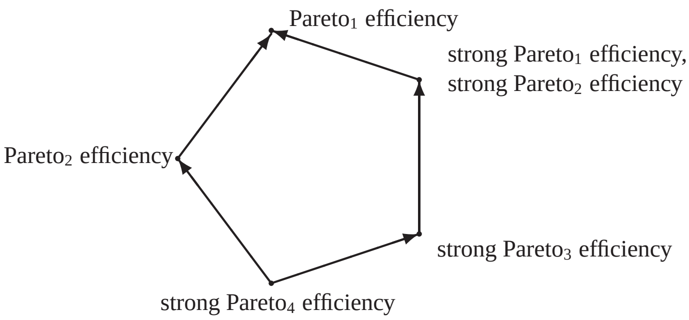


Figure 2.1. The interrelations between the various notions of Pareto efficiency for orders that are not necessarily connected.

imply Pareto$_1$ efficiency. Finally, Pareto$_2$ efficiency entails Pareto$_1$ efficiency, but not *vice versa*. Moreover, Pareto$_2$ efficiency entails none of the strong Pareto notions, and is implied by strong Pareto$_4$ efficiency only. How the various Pareto concepts relate is depicted in Figure 2.1. The following facts establish these observations.

Fact 2.1.2 *Let s be a strategy profile for some strategic game G. Then:*

$$s \text{ is strongly Pareto}_1 \text{ efficient} \quad \text{iff} \quad s \text{ is strongly Pareto}_2 \text{ efficient}$$

Proof: That strong Pareto$_2$ efficiency entails strong Pareto$_1$ efficiency is trivial, since $s <_i s'$ implies $s' \nleqslant_i s$. So assume some strategy profile s to be *not* strongly Pareto$_2$ efficient. For some strategy profile s' and some player i, then, $s' \nleqslant_i s$ and, moreover, $s \leqslant_j s'$, for all players j. Then in particular $s \leqslant_i s'$, and, therefore, also $s <_i s'$. We may conclude that s is not strongly Pareto$_1$ efficient. ⊣

Fact 2.1.3 *For every strategy profile s of a game G:*

(i)	*strong Pareto$_4$ efficiency*	*implies*	*strong Pareto$_3$ efficiency,*
(ii)	*strong Pareto$_3$ efficiency*	*implies*	*strong Pareto$_1$ efficiency,*
(iii)	*strong Pareto$_1$ efficiency*	*implies*	*Pareto$_1$ efficiency,*
(iv)	*strong Pareto$_4$ efficiency*	*implies*	*Pareto$_2$ efficiency,*
(v)	*Pareto$_2$ efficiency*	*implies*	*Pareto$_1$ efficiency.*

Moreover, none of the implications (i) through (v) hold in the opposite direction.

Proof: We prove each implication and its failure to hold in the opposite direction subsequently.

ad. (*i*) : The proof of the implication is straightforward as in general $s <_i s'$ implies $s' \nleqslant_i s$. As to the failure of the opposite direction, any game with more than one strategy profile and each player's preferences being defined by the identity relation over the strategy profiles will do as a counterexample.

ad. (*ii*) : The implication itself is again immediate since $s' <_i s$ implies $s \nleqslant_i s'$, by definition. For a counterexample for the converse direction, consider a game with at least two strategy profiles s and s' and only two players 1 and 2. Let 2's preference relation be given by the identity relation over the strategy profiles, and that of 1 the universal relation *minus* the pair (s', s). Then both s and s' are strongly Pareto$_1$ efficient. However, s is not strongly Pareto$_3$ efficient.

ad. (*iii*) : We may assume that N is non-empty. Assume further some strategy profile s not to be Pareto$_1$ efficient. Then, there is some strategy profile s' is such that $s <_i s'$ for all i in N. Hence, for all i in N also $s \leqslant_i s'$. With N not empty, it follows there is some i^* such that $s <_{i^*} s'$. Hence, s is not strongly Pareto$_1$ efficient either.

For an example refuting the opposite implication, consider a game with at least two strategy profiles s and s' and at least two players. Assume the players' preferences be given by the universal relation, except for those of one player, which are given by the universal relation *minus* (s', s). Then s is Pareto$_1$ efficient, but not strongly so.

ad. (*iv*) : First assume some strategy profile s to fail as a Pareto$_2$ efficient outcome. For some strategy profile s', then $s' \nleqslant_i s$ for all i in N. Then also $s' \nless_i s$ for all i in N. With N non-empty, there is, moreover, some player i^* such that $s' \nleqslant_{i^*} s$. Hence, s is not strongly Pareto$_4$ efficient either.

For a counterexample disproving the implication in the opposite direction to hold, the same counterexample as in (*iii*) will do. The strategy profile s is there Pareto$_2$ efficient, but not strongly Pareto$_4$ efficient.

ad. (*v*) : The implication is almost trivial, since in general $s' \leqslant_i s$ implies $s \nless_i s'$. A counterexample witnessing the failure of the opposite direction is provided by the same as that of (*i*). ⊣

Fact 2.1.4 *Let s be a strategy profile of a strategic game G. The following equivalences do **not** hold in either direction:*

strong Pareto$_3$ efficiency iff Pareto$_2$ efficiency,

strong Pareto$_2$ efficiency iff Pareto$_2$ efficiency,

Proof: We present a strategic game in which some strategy profile is strongly Pareto$_3$ efficient but not Pareto$_2$ efficient, as well as a game in which some strategy profile is Pareto$_2$ efficient but not strongly Pareto$_2$ efficient. The first counterexample is given by a game in which all players have the identity relation over the strategy profiles as their preferences and which has at least two strategy profiles. A game witnessing the second possibility is given by any game with more than one player, all of which preferences are given by the universal relation, except for one player i^*, whose preferences are given by the universal relation *minus* (s', s).

Hence, strong Pareto$_3$ efficiency does not in general imply Pareto$_2$ efficiency and Pareto$_2$ efficiency does not in general imply strong Pareto$_2$ efficiency. Now assume for a *reductio ad absurdum* that strong Pareto$_2$ efficiency implies Pareto$_2$ efficiency. By Fact 2.1.3, strong Pareto$_3$ efficiency implies strong Pareto$_2$ efficiency. But then strong Pareto$_3$ efficiency would also imply Pareto$_2$ efficiency, *quod non*. Similarly assume that Pareto$_2$ efficiency implies strong Pareto$_3$ efficiency. By Fact 2.1.3, strong Pareto$_3$ efficiency implies strong Pareto$_2$ efficiency and so Pareto$_2$ efficiency would imply strong Pareto$_2$ efficiency. Against the latter claim, however, we had found a counterexample. $\dashv$

Equilibria, Pareto Properties and Partial Preferences

An interesting topic is coalition formation in strategic games. A coalition has, in general, a greater enforcing power than each of its members on his own. The strategies of a coalition could be assumed to be given by the possible combinations of the strategies of its members. Intuitively, this corresponds to the assumption that the members of a coalition can coordinate their choice of strategy and, thus, the coalition as a whole gains greater control over the outcome of the game. This leaves the issue of how the coalitional preferences depend on the individual preferences of the members.

How to derive a preference order for a coalition from the preferences of its constituent members is a highly non-trivial issue and belongs to the field of social choice theory. Arrow's impossibility theorem (*cf.*, Arrow (1963)) states the impossibility of a general method to define coalitional preferences from the individual preferences — *i.e.*, of a social choice function — if this method is to comply with certain intuitive restrictions for each possible collection of individual preferences. One of these restrictions is that the coalitional preference order is to be a total preorder over the possible alternatives over which the individual preferences are defined.

In this thesis, however, we allow the individual preferences to be partial and the same lenient attitude is taken towards coalitional preferences. A coalition is then said to value one state s at least as much as another state s if and only if all players value s at least as much as s'. Formally, the coalitional preferences are obtained by simply taking the intersection of the preference relations of its constituent members. The coalitional preferences are then guaranteed to be reflexive and transitive, if all the individual preferences are. If one of the individual preference orders is empty, so is the coalitional preference order. However, the coalitional preference relation is not in general a total preorder, not even if all the individual preference relations are. This procedure complies with with the *strong Pareto property*, *i.e.*, if all coalition members value a state at least as much as another state, so does the coalition. *I.e.*, formally, for κ a coalition of players in a strategic game G, a coalitional preference relation ρ_κ complies with the strong Pareto property if and only if for all strategy profiles s and s' of G:

$$\text{if for all } i \in \kappa\colon s \leqslant_i s' \text{ then } s \leqslant_\kappa s'.$$

Let G be a strategic game. Each way in which coalitions can be formed fixes in a unique fashion another strategic game in which the coalitions are the players, whose preferences and powers are defined as above. Assuming that each player is a member of precisely one coalition, the possible ways in which coalition formation can take place is exhausted by the possible partitions of the players of the original game. These partitions of the players constitute a complete lattice, with as top the grand coalition, of which all players are a member, and as bottom the trivial partition, in which each coalition consists of exactly one player.

It now so happens that, in the game defined thus for the grand coalition, a strategy profile is a maximal equilibrium if and only if that strategy profile is strongly Pareto$_1$ efficient. Similarly, a strategy profile is a maximal equilibrium in the game for the trivial partition if and only if that strategy profile is also a maximal equilibrium in the original game. This shows that the notion of a maximal equilibrium and that of strong Pareto$_1$ efficiency are in an important sense extreme instances of one and the same concept.

The possible *coalition partitions* are given by the set of partitions $Part\,(N)$ over N, which constitutes a complete lattice under the refinement relation. For π and π' partitions of N let the refinement relation $\leqslant$ is formally defined as:

$$\pi \leqslant \pi' \quad \text{iff} \quad \text{for all } x \in \pi \text{ there is a } y \in \pi' \text{ such that } x \subseteq y.$$

Intuitively, $\pi \leqslant \pi'$ denotes that π at least as fine as π'. As such $\leqslant$ defines a partial order on $Part\,(S)$. Let $\boldsymbol{\kappa}_\top$ and $\boldsymbol{\kappa}_\perp$ denote respectively the top and bottom of this coalition lattice, *i.e.*, $\boldsymbol{\kappa}_\top =_{df.} \{N\}$ and $\boldsymbol{\kappa}_\perp =_{df.} \big\{\{i\} \ : \ i \in N\big\}$. We now define for each strategic game G each coalition partition $\boldsymbol{\kappa}$ of its players, the strategic game $G_{\boldsymbol{\kappa}}$, which intuitively is the game that results of the players of G join in the coalitions $\boldsymbol{\kappa}$ in a way that complies with the requirements formulated in the remarks above.

Definition 2.1.5 Let G be a strategic game given by $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ and let $\boldsymbol{\kappa}$ be a coalition partition of N. Define:

$$G_{\boldsymbol{\kappa}} \quad =_{df.} \quad \left(\boldsymbol{\kappa}, \{S_\kappa\}_{\kappa\in\boldsymbol{\kappa}}, \{\rho_\kappa\}_{\kappa\in\boldsymbol{\kappa}}\right),$$

where for each $\kappa \in \boldsymbol{\kappa}$:

$$S_\kappa \quad =_{df.} \quad \prod_{i\in\kappa} S_i \qquad \text{and} \qquad \rho_\kappa \quad =_{df.} \quad \bigcap_{i\in\kappa} \rho_i.$$

For each $\kappa \in \boldsymbol{\kappa}_\perp$ there is some $i \in N$ such that $\kappa = \{i\}$ and $(s_{-\kappa}, s'_\kappa) = (s_{-i}, s'_i)$. Similarly, we have N as the only coalition in $\boldsymbol{\kappa}_\top$ and, consequently, for all $\kappa \in \boldsymbol{\kappa}_\top$, $(s_{-\kappa}, s'_\kappa) = (s_{-N}, s'_N) = s'$. More in general, we may assume a natural isomorphism between $\prod_{\kappa\in\boldsymbol{\kappa}_\top} s\,(\kappa)$ and $\prod_{i\in N} S_i$ of the original game. We now have the following proposition.

Proposition 2.1.6 *Let G be the strategic game $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ and s a strategy* *profile. Then:*

$$\begin{array}{rcl} s \text{ is a maximum equilibrium } G_{\boldsymbol{\kappa}_\perp} & \textit{iff} & s \text{ is a maximum equilibrium in } G, \\ s \text{ is a maximal equilibrium } G_{\boldsymbol{\kappa}_\perp} & \textit{iff} & s \text{ is a maximal equilibrium in } G, \\ s \text{ is a maximal equilibrium } G_{\boldsymbol{\kappa}_\top} & \textit{iff} & s \text{ is strongly Pareto}_2 \text{ efficient in } G. \end{array}$$

Proof: The first two claims are trivial, given the remarks made in the text above. As to the third claim, observe for all strategy profiles s and s', that $(s_{-N}, s'_N) = s'$. Consider the the following equivalences:

$$\begin{array}{ll} & s \text{ is a maximal equilibrium } G_{\boldsymbol{\kappa}_\top} \\ \text{iff} & \text{for all } i \in \boldsymbol{\kappa}_\top : s \text{ is a maximal response for } i \\ \text{iff} & s \text{ is a maximal response for } N \\ \text{iff} & \text{for all } s' \in S\colon s \not<_N (s_{-N}, s'_i) \\ \text{iff} & \text{for all } s' \in S\colon s \not<_N s' \\ \text{iff} & \text{for all } s' \in S\colon s \not\leqslant_N s' \text{ or } s' \leqslant_N s \\ \text{iff} & \text{for all } s' \in S\colon \text{for some } i \in N,\ s \not\leqslant_i s' \text{ or for all } i \in N,\ s' \leqslant_i s \\ \text{iff} & \text{for all } s' \in S, \text{for all } i \in N\colon s' \not\leqslant_i s \ \text{ implies } \ \text{for some } j \in N,\ s \not\leqslant_j s' \\ \text{iff} & s \text{ is strongly Pareto}_2 \text{ efficient} \end{array}$$

This ends the proof. ⊣

It might seem that by appropriately defining coalitional preferences each Pareto concept could be seen as a borderline case of maximal or maximum equilibrium. Consider, *e.g.*, the case in which ρ_κ had been defined in such a way that for all s and s':

$$(s, s') \in \rho_\kappa \quad \text{iff} \quad \text{for all } i \in \kappa\colon s' \not<_i s.$$

Then it can in fact be proved that a strategy profile s is strongly Pareto$_3$ efficient in a game G if and only if s is a maximal equilibrium in $G_{\boldsymbol{\kappa}_\top}$. However, defined thus, the coalitional preferences are no longer in general guaranteed to be transitive.

We are now in a position to define the following conservative extensions of the concepts of maximum and maximal equilibrium.

Definition 2.1.7 *(Maximum $\boldsymbol{\kappa}$-equilibrium and maximal $\boldsymbol{\kappa}$-equilibrium)* A strategy profile s is a *maximum $\boldsymbol{\kappa}$-equilibrium* in a strategic game $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ if and only if s is a maximum equilibrium in the game $\left(\boldsymbol{\kappa}, \{S_\kappa\}_{\kappa\in\boldsymbol{\kappa}}, \{\rho_\kappa\}_{\kappa\in\boldsymbol{\kappa}}\right)$. Similarly, a strategy profile s is a *maximal $\boldsymbol{\kappa}$-equilibrium* in a strategic game $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ if and only if s is a maximal equilibrium in the game $\left(\boldsymbol{\kappa}, \{S_\kappa\}_{\kappa\in\boldsymbol{\kappa}}, \{\rho_\kappa\}_{\kappa\in\boldsymbol{\kappa}}\right)$.

	G	B
G	**1** **1** **1**	0 0 0
B	0 1 0	2 0 2

Matrix g

	G	B
G	2 0 2	0 1 0
B	0 0 0	**1** **1** **1**

Matrix b

Figure 2.2. The *Golden Heart* game without coalition formation. Abélard chooses rows, Eloïse chooses columns and the little sister Spoiler chooses matrices. The Nash equilibria are in boldface.

In virtue of Proposition 2.1.6, we then find that maximal equilibrium and strong Pareto$_2$ efficiency are extreme cases of one and the same concept, *viz.*, maximal $\boldsymbol{\kappa}$-equilibrium.

To conclude this section, we observe that coalition formation in a game results in at most a *reduction* of the maximum equilibria of a strategic game.

Proposition 2.1.8 *Let G a strategic game $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ and let $\boldsymbol{\kappa}$ and $\boldsymbol{\kappa}'$ be coalition partitions of N such that $\boldsymbol{\kappa} \leqslant \boldsymbol{\kappa}'$. Then the maximum equilibria in $G_{\boldsymbol{\kappa}'}$ are also maximum equilibria in $G_{\boldsymbol{\kappa}}$.*

Proof: Let s be a maximum equilibrium in $G_{\boldsymbol{\kappa}'}$, *i.e.*, for all $\kappa' \in \boldsymbol{\kappa}'$ and all strategy profiles s' we have $(s_{-\kappa'}, s'_{\kappa'}) \leqslant_{\kappa'} s$. We prove for an arbitrary coalition κ in $\boldsymbol{\kappa}$, an arbitrary player i in κ and an arbitrary strategy profile s' that $(s_{-\kappa}, s'_{\kappa}) \leqslant_i s$. Consider the unique $\kappa' \in \boldsymbol{\kappa}'$ such that $\kappa \subseteq \kappa'$ as well as the valuation $\left(s_{-\kappa'}, (s_{-\kappa}, s'_{\kappa})_{\kappa'}\right)$. Since, $i \in \kappa'$, then, $\left(s_{-\kappa'}, (s_{-\kappa}, s'_{\kappa})_{\kappa'}\right) \leqslant_i s$. Now observe that $(s_{-\kappa}, s'_{\kappa}) = \left(s_{-\kappa'}, (s_{-\kappa}, s'_{\kappa})_{\kappa'}\right)$, and we are done. $\dashv$

The inverse of Proposition 2.1.8, however, does not hold. Coalition formation may result in a decrease of the number of maximum equilibria, as the following example shows.

Example 2.1.9 Abélard and Eloïse are a young boy and girl very much in love and Spoiler is Eloïse's little sister. Abélard and Eloïse plan to go for a romantic ramble down town. The sister Spoiler would very much like to go as well, but she is allowed to only if accompanied by her elder sister. The two lovers, of course, would prefer to go just with the two of them and have the little sister stay behind. Nevertheless, they would rather have the sister join them than all of them staying at home with their parents. Moreover, both Abélard and Eloïse are indifferent between not going at all on the one hand and the two sisters going into the city without Abélard. The younger sister just wants to join her sister on what promises to be an exciting trip. It now so happens that if they are to go at all, they have to meet up with one another either in the

	$\{$G,$G\}$	$\{$B,$G\}$	$\{$G,$B\}$	$\{$B,$B\}$
g	1 1	0 0	0 1	2 0
b	2 0	0 1	0 0	1 1

Figure 2.3. Matrix of the *Golden Heart* game in which Abélard en Eloïse have teamed up. Now the little sister Spoiler chooses rows and Abélard and Eloïse jointly columns.

pub *The Golden Heart* or in front of the bank. Still, they have not yet made a specific appointment in this respect. If either the two sisters or the two lovers meet at the same place, they go down town together. If they all meet up at the same place, they go with the three of them. In any other case they will have to face a boring Saturday afternoon at home.

Observe that Abélard's and Eloïse's preferences coincide. The situation is summarized as a game in Figure 2.2, where Abélard chooses rows (G or B) Eloïse chooses columns (*G* or *B*) and Spoiler matrices (*Matrix g* or *Matrix b*). Abélard's ordinal preferences are represented by the figures bottom left in each box, those of Eloïse's by those top right and those of Spoiler's by those in the middle. We find that there are two maximum equilibria in this game, *viz.*, when they all meet at the same place and go down town with the three of them.

Abélard and Eloïse, however, are quite likely to form a coalition in an effort to withhold the little sister from spoiling all the romantic fun. As a coalition they can make sure to meet at the same place, still they cannot preclude the sister turning up there as well. The resulting situation is represented in Figure 2.3, this time Abélard and Eloïse jointly choosing columns and Spoiler choosing rows. In this game there are no maximum equilibria.

2.2 Rough Sets

In the last two parts, and then especially Part III, extensive use is made of the theory of rough sets. In this section the elementary concepts of rough set theory, *viz.*, the upper and lower approximations of a set, are introduced. In the next sections they are employed for some results for propositional logics. As such our employment of rough sets is divergent from normal use. First we will give some basic facts concerning upper and lower approximations.[2]

For S we have $Part\,(S)$ denote the set of partitions over S. Moreover, for π a partition in S and for x an element of S, $[x]_\pi$ is the unique block of π containing x.

[2]For a more extensive account the reader be referred to Pawlak (1991)

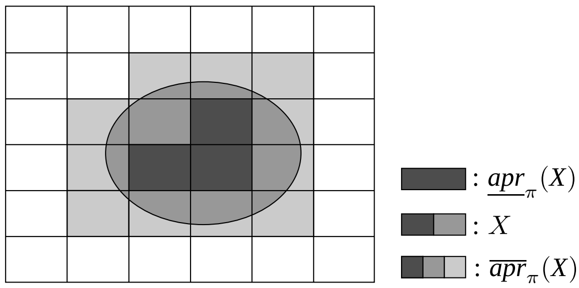


Figure 2.4. Rough sets in a set S partitioned by π. The oval represents the set X. The colored areas indicate the upper approximation and the darkly colored area the lower approximation.

Let S be a set and let π be a partition of S. Obviously, it is not in general the case that a subset X of S is identical to a union of a number of blocks in π. Still each subset can be characterized by two sets which do have this property. Define the *lower approximation* $\underline{apr}_\pi(X)$ and the *upper approximation* $\overline{apr}_\pi(X)$ by:

$$\begin{aligned}\underline{apr}_\pi(X) &=_{df.} \bigcup \{Y \in \pi : \ Y \subseteq X\},\\ \overline{apr}_\pi(X) &=_{df.} \bigcup \{Y \in \pi : \ Y \cap X \neq \emptyset\}.\end{aligned}$$

Let ε_π be the equivalence relation over S associated with the partition π. Then we also have the following equivalent characterizations of the lower and upper approximation of a subset X of S. For each $x \in S$:

$$\begin{aligned}x \in \underline{apr}_\pi(X) &\quad \text{iff} \quad \text{for all } s' \in S \text{ such that } s \sim_\pi s' : s' \in X,\\ x \in \overline{apr}_\pi(X) &\quad \text{iff} \quad \text{for some } s' \in S\colon s \sim_\pi s' \ \text{ and } \ s' \in X.\end{aligned}$$

Clearly, $\overline{apr}_\pi$ is a cylindrification operator on 2^S and $\underline{apr}_\pi$ its dual. As such they exemplify a more general mathematical concept that is also instanced by quantification and modality in logic. This observation has by no means escaped attention in the literature (*cf. e.g.*, Yao, Wong, and Lin (1997), Düntsch (1999) and Düntsch (no date)). Figure 2.4 illustrates the lower and upper approximations of a set X. Suppressing the subscript π, the approximation operations $\underline{apr}$ and $\overline{apr}$ satisfy the following elementary properties for subsets X and Y of a set S:[3]

$$
\begin{array}{rclcrcl}
\underline{apr}(\emptyset) & = & \emptyset & \qquad & \overline{apr}(\emptyset) & = & \emptyset \\
\underline{apr}(S) & = & S & & \overline{apr}(S) & = & S \\
\underline{apr}(X) & \subseteq & X & & \overline{apr}(X) & \supseteq & X \\
\underline{apr}(X) & \subseteq & \underline{apr}\left(\underline{apr}(X)\right) & & \overline{apr}(X) & \supseteq & \overline{apr}(\overline{apr}(X)) \\
X & \subseteq & \underline{apr}(\overline{apr}(X)) & & X & \supseteq & \overline{apr}\left(\underline{apr}(X)\right) \\
\overline{apr}(X) & \subseteq & \underline{apr}(\overline{apr}(X)) & & \underline{apr}(X) & \supseteq & \overline{apr}\left(\underline{apr}(X)\right) \\
\underline{apr}(X) & = & \overline{\overline{apr}(\overline{X})} & & \overline{apr}(X) & = & \overline{\underline{apr}(\overline{X})} \\
\underline{apr}(X \cap Y) & = & \underline{apr}(X) \cap \underline{apr}(Y) & & \overline{apr}(X \cap Y) & \subseteq & \overline{apr}(X) \cap \overline{apr}(Y) \\
\underline{apr}(X \cup Y) & \supseteq & \underline{apr}(X) \cup \underline{apr}(Y) & & \overline{apr}(X \cup Y) & = & \overline{apr}(X) \cup \overline{apr}(Y)\,.
\end{array}
$$

As two obvious consequences of these properties also both $\underline{apr}\left(\underline{apr}(X)\right) = \underline{apr}(X)$ and $\overline{apr}(\overline{apr}(X)) = \overline{apr}(X)$. Moreover, the latter four inequalities can be generalized to infinite sets of sets. Let $\boldsymbol{X} \subseteq 2^S$, then:

$$
\begin{array}{rclcrcl}
\underline{apr}\left(\bigcap \boldsymbol{X}\right) & = & \displaystyle\bigcap_{X \in \boldsymbol{X}} \underline{apr}(X) & \qquad & \overline{apr}\left(\bigcap \boldsymbol{X}\right) & \subseteq & \displaystyle\bigcap_{X \in \boldsymbol{X}} \overline{apr}(X) \\
\underline{apr}\left(\bigcup \boldsymbol{X}\right) & \supseteq & \displaystyle\bigcup_{X \in \boldsymbol{X}} \underline{apr}(X) & & \overline{apr}\left(\bigcup \boldsymbol{X}\right) & = & \displaystyle\bigcup_{X \in \boldsymbol{X}} \overline{apr}(X)\,.
\end{array}
$$

Both $\underline{apr}$ and $\overline{apr}$ satisfy upward monotonicity:

$$
\begin{array}{lll}
X \subseteq Y & \text{implies} & \underline{apr}(X) \subseteq \underline{apr}(Y)\,, \\
X \subseteq Y & \text{implies} & \overline{apr}(X) \subseteq \overline{apr}(Y)\,.
\end{array}
$$

We also have the following fact, which says that, given a partition, the fixed points of the upper and lower approximations coincide.

Fact 2.2.1 *Let S be a set, $X \subseteq S$ and $\pi \in Part(S)$. Then:*

$$X = \overline{apr}_\pi(X) \quad \textit{iff} \quad X = \underline{apr}_\pi(X)\,.$$

Proof: First assume $X = \overline{apr}(X)$. Observe that both $\overline{apr}(X) \subseteq \underline{apr}(\overline{apr}(X))$ and $\underline{apr}(\overline{apr}(X)) \subseteq \overline{apr}(X)$ are instances of rough set laws. Hence $\overline{apr}(X) = \overline{apr}\left(\underline{apr}(X)\right)$ and we may reason as follows:

$$X \;=_{\text{ass.}}\; \overline{apr}(X) \;=\; \underline{apr}(\overline{apr}(X)) \;=_{X = \overline{apr}(X)}\; \underline{apr}(X)\,.$$

The reasoning in the opposite direction is analogous.[4] ⊣

As a generalization of this fact, we also have the following.

[3] These inequalities are taken from Yao, Wong, and Lin (1997).
[4] For this elegant proof I am indebted to Boudewijn de Bruin.

Fact 2.2.2 *Let π be a partition of a set S. Let further $\boldsymbol{X}$ a set of subsets of S such that $\boldsymbol{X} \subseteq \pi$. Then:*

$$\overline{apr}_\pi(\bigcup \boldsymbol{X}) \;=\; \underline{apr}_\pi(\bigcup \boldsymbol{X}) \;=\; \bigcup \boldsymbol{X}.$$

Proof: Straightforward. $\dashv$

The partitions by means of which sets are approximated may be finer or coarser. The facts that follow concern the behavior of the approximation operations with respect to partitions of various degrees of coarseness. For π and π' partitions of a set S, *i.e.*, $\pi, \pi' \in Part\,(S)$, let $\pi \leqslant \pi'$ be formally defined as:

$$\pi \leqslant \pi' \quad \text{iff} \quad \text{for all } x \in \pi \text{ there is a } y \in \pi' \text{ such that } x \subseteq y.$$

Intuitively, $\pi \leqslant \pi'$ denotes that π at least as fine as π'. As such $\leqslant$ defines a partial order on $Part\,(S)$. Then $\leqslant$ defines a partial order on $Part\,(S)$. Rather, $Part\,(S)$ constitutes a complete lattice if ordered thus. We now also have the following monotonicity properties for the lower and upper approximation operations:

Fact 2.2.3 *Let π and π' be partitions of some set S. Then for all $X \subseteq S$:*

$$\begin{aligned} \pi \leqslant \pi' \quad &\textit{implies} \quad \underline{apr}_\pi(X) \supseteq \underline{apr}_{\pi'}(X)\,, \\ \pi \leqslant \pi' \quad &\textit{implies} \quad \overline{apr}_\pi(X) \subseteq \overline{apr}_{\pi'}(X)\,. \end{aligned}$$

Proof: Both cases are analogous; here we prove only the first. Assume $\pi \leqslant \pi'$ and consider an arbitrary $x \in \underline{apr}_{\pi'}(X)$. Consider the block $[x]_\pi$ of π; then, $[x]_{\pi'} \subseteq X$. By the assumption there is a block Y of π' such that $[x]_\pi \subseteq Y$. Then, $x \in Y$ and, therefore, $Y = [x]_{\pi'}$. Hence, $[x]_\pi \subseteq [x]_{\pi'} \subseteq X$. Since $x \in [x]_\pi$, we may conclude that $x \in \underline{apr}_\pi(X)$. $\dashv$

In words, the coarser the partition, the larger the upper approximation of a set and the smaller its lower approximation. The following fact conveys a stronger and closely related result.

Fact 2.2.4 *Let π and π' be partitions of some set S. Then:*

$$\begin{aligned} \pi \leqslant \pi' \quad &\textit{iff} \quad \textit{for all } X \subseteq S\colon\ \underline{apr}_\pi(X) \supseteq \underline{apr}_{\pi'}(X)\,, \\ \pi \leqslant \pi' \quad &\textit{iff} \quad \textit{for all } X \subseteq S\colon\ \overline{apr}_\pi(X) \subseteq \overline{apr}_{\pi'}(X)\,. \end{aligned}$$

Proof: The proofs of both cases run along analogous lines; we will here give that of the first. The left-to-right direction is immediate by Fact 2.2.3. For the opposite direction, assume that for all $X \subseteq S$ we have $\underline{apr}_\pi(X) \supseteq \underline{apr}_{\pi'}(X)$. Consider an arbitrary $X \in \pi$. By definition, X is a non-empty subset of S. As $S = \bigcup \pi'$, there is a $Y \in \pi'$ such that $X \cap Y \neq \emptyset$. By the assumption, then $\underline{apr}_\pi(Y) \supseteq \underline{apr}_{\pi'}(Y) =_{\text{Fact 2.2.2}} Y$. Because also $\underline{apr}_\pi(Y) \subseteq Y$, it follows that $Y = \underline{apr}_\pi(Y)$. By definition, $Y = \bigcup \{X' \in$

$\pi : \ X' \subseteq Y\big\}$. Since, $X \cap Y \neq \emptyset$, we have $X \cap X'' \neq \emptyset$, for some $X'' \in \big\{X' \in \pi : X' \subseteq Y\big\}$. With X and X'' blocks in the partition π, it follows that $X = X''$, and hence, $X \subseteq Y$. Having chosen X arbitrarily from π, we may conclude that $\pi \leqslant \pi'$. $\dashv$

In the sequel, we will mostly be interested in a particular class of partitions of a universe with respect to which the approximations are defined. If the universe set is the powerset of a set A, we define an equivalence relation holding between any two subsets of A if each element of a third subset of A is in the one subset if and only if it is an element of the other. Let A be a set and define for each $Z \subseteq A$ the equivalence relation ε_Z on 2^A such that for all $X, Y \subseteq A$:

$$(X,Y) \in \varepsilon_Z \quad \text{iff} \quad Z \cap X = Z \cap Y.$$

Sometimes we use the infix notation $X \sim_Z Y$ to convey that $(X,Y) \in \varepsilon_Z$. Observe that it is both a necessary and a sufficient condition for $X \sim_Z Y$ to hold that for all $z \in Z$, it is the case that $z \in X$ if and only if $z \in Y$. Note that $\varepsilon_\emptyset$ and ε_A are the universal relation and the identity relation on 2^A, respectively. More in general, we have $X \subseteq Y$ *if* and only if $\varepsilon_Y \subseteq \varepsilon_X$. The only-if direction is trivial. For the other direction assume for the contrapositive that $x \in X$ but $x \notin Y$. Then, $\emptyset \nsim_X \{x\}$ but $\emptyset \sim_Y \{x\}$, *i.e.*, the relations ε_X and ε_Y are distinct. Hence, the set $\{\varepsilon_X : \ X \subseteq A\}$ constitute a complete lattice with relation composition and intersection as join and meet, respectively. To appreciate this, consider the following fact.

Fact 2.2.5 *Let X and Y be subsets of a set A. Then:*

$$\varepsilon_{X\cap Y} \ = \ \varepsilon_X \circ \varepsilon_Y \qquad \textit{and} \qquad \varepsilon_{X\cup Y} \ = \ \varepsilon_X \cap \varepsilon_Y.$$

Proof: For the $\subseteq$-direction of the first claim, assume for arbitrary $s, s' \in 2^A$ that $(s,s') \in \varepsilon_{X\cap Y}$. Hence, $s \cap X \cap Y = s' \cap X \cap Y$. Define $s^* =_{df.} (s \cap X) \cup \big(s' \cap \overline{X}\big)$. Then:

$$s \cap X \ = \ (s \cap X \cap X) \cup \big(s' \cap \overline{X} \cap X\big) \ = \ \big((s \cap X) \cup (s' \cap \overline{X})\big) \cap X.$$

Hence, $(s, s^*) \in \varepsilon_X$. Also consider the following equalities:

$$\begin{aligned} s' \cap Y \ &= \ \big((s' \cap X) \cup (s' \cap \overline{X})\big) \cap Y \ = \ \big((s' \cap X \cap Y) \cup (s' \cap \overline{X} \cap Y)\big) \\ &= \ \big((s \cap X \cap Y) \cup (s' \cap \overline{X} \cap Y)\big) \ = \ \big((s \cap X) \cup (s' \cap \overline{X})\big) \cap Y. \end{aligned}$$

Accordingly also $(s^*, s') \in \varepsilon_Y$ and finally also $(s, s') \in \varepsilon_X \circ \varepsilon_Y$.

For the $\supseteq$-direction, assume $(s,s') \in \varepsilon_X \circ \varepsilon_Y$. So, for some $s'' \in 2^A$ both $(s, s'') \in \varepsilon_X$ and $(s'', s') \in \varepsilon_Y$. *I.e.*, both $s \cap X = s'' \cap X$ and $s'' \cap Y = s' \cap Y$. Consider the following equalities:

$$s \cap X \cap Y \ = \ s'' \cap X \cap Y \ = \ s'' \cap Y \cap X \ = \ s' \cap Y \cap X \ = \ s' \cap X \cap Y.$$

Hence, $(s, s') \in \varepsilon_{X\cap Y}$.

For the second claim, first assume for arbitrary $s, s' \in 2^A$ that $(s, s') \in \varepsilon_{X \cup Y}$. Then, $s \cap (X \cup Y) = s' \cap (X \cup Y)$. Since, $X, Y \subseteq X \cup Y$ then also both $s \cap X = s' \cap X$ and $s \cap Y = s' \cap Y$. We may conclude that $(s, s') \in \varepsilon_X \cap \varepsilon_Y$. For the opposite direction, assume $(s, s') \in \varepsilon_X \cap \varepsilon_Y$, *i.e.*, both $s \cap X = s' \cap X$ and $s \cap Y = s' \cap Y$. Now reason as follows:

$$s \cap (X \cup Y) = (s \cap X) \cup (s \cap Y) = (s' \cap X) \cup (s' \cap Y) = s' \cap (X \cup Y)\,.$$

We may conclude that $(s, s') \in \varepsilon_{X \cup Y}$. $\dashv$

The partition of 2^A as determined by ε_X, we denote by π_X. The notation $[x]_{\varepsilon_X}$ for the equivalence class under ε_X containing x we usually abbreviate to $[x]_X$. Obviously, π_A is the finest and $\pi_\varnothing$ the coarsest partition of 2^A. More in general we have that the larger the set X, the finer the partition π_X.

Fact 2.2.6 *Let X and Y be subsets of some set A. Then:*

$$X \subseteq Y \quad \textit{iff} \quad \pi_Y \leqslant \pi_X.$$

Proof: First assume $X \subseteq Y$ and consider an arbitrary $\boldsymbol{X} \in \pi_X$. We without loss of generality we may assume that $\boldsymbol{X} = [Z]_X$, for some $Z \subseteq A$. Now consider $[Z]_Y$ as well as an arbitrary $Z' \in [Z]_Y$. Then, $Z' \cap Y = Z \cap Y$. With the assumption that $X \subseteq Y$, then also $Z' \cap X = Z \cap X$. Hence, $Z' \in [Z]_X$. We may conclude that $\pi_Y \leqslant \pi_X$.

For the opposite direction, assume that $X \nsubseteq Y$, *i.e.*, that there be an $x \in X$ with $x \notin Y$. Consider this x along with the block $[\{x\}]_Y$ of π_Y. Observe that both $\{y\} \in [\{x\}]_Y$ and $\varnothing \in [\{x\}]_Y$. It suffices to show that for all $\boldsymbol{X} \in \pi_X$ if $\{x\} \in \boldsymbol{X}$ then $\varnothing \notin \boldsymbol{X}$. So consider an arbitrary $\boldsymbol{X} \in \pi_X$ with $\{x\} \in \boldsymbol{X}$ as well as an arbitrary $X' \in \boldsymbol{X}$. Then $\{x\} \sim_X X'$, and with $x \in \{x\}$ and $x \in X$ we may conclude that $x \in X'$. Hence, $X' \neq \varnothing$. $\dashv$

For X a subset of a set A we denote the approximation operators $\underline{apr}_{\pi_X}$ and $\overline{apr}_{\pi_X}$, as defined for subsets of 2^A, by $\underline{apr}_X$ and $\overline{apr}_X$, respectively. As an immediate result of the facts 2.2.4 and 2.2.6 we have the following corollary.

Corollary 2.2.7 *Let X and Y be subsets of some set A. Then:*

$$X \subseteq Y \quad \textit{iff} \quad \textit{for all } \boldsymbol{X} \subseteq 2^A\colon\ \underline{apr}_X(\boldsymbol{X}) \subseteq \underline{apr}_Y(\boldsymbol{X})\,,$$
$$X \subseteq Y \quad \textit{iff} \quad \textit{for all } \boldsymbol{X} \subseteq 2^A\colon\ \overline{apr}_Y(\boldsymbol{X}) \subseteq \overline{apr}_X(\boldsymbol{X})\,.$$

Proof: Immediate from Fact 2.2.4 and Fact 2.2.6. $\dashv$

With respect to the behavior of lower and upper approximations of a set given partitions π_X and π_Y, we have the following two facts.

Fact 2.2.8 *Let A be some set, of which X and Y are subsets. Let, moreover, $\boldsymbol{X}$ be a subset of 2^A. Then:*

$$\begin{array}{rcl}\underline{apr}_{X\cap Y}(\boldsymbol{X}) & = & \underline{apr}_X\left(\underline{apr}_Y(\boldsymbol{X})\right),\\ \overline{apr}_{X\cap Y}(\boldsymbol{X}) & = & \overline{apr}_X\left(\overline{apr}_Y(\boldsymbol{X})\right),\\ \underline{apr}_{X\cup Y}(\boldsymbol{X}) & \supseteq & \underline{apr}_X(\boldsymbol{X})\cap\underline{apr}_Y(\boldsymbol{X}),\\ \overline{apr}_{X\cup Y}(\boldsymbol{X}) & \subseteq & \overline{apr}_X(\boldsymbol{X})\cap\overline{apr}_Y(\boldsymbol{X}).\end{array}$$

Proof: The proofs of the first two claims are analogous; here we only give that of the latter.

$$\begin{array}{lll} s\in\overline{apr}_{X\cap Y}(\boldsymbol{X}) & \text{iff} & s\sim_{X\cap Y}s' \text{ and } s'\in\boldsymbol{X},\ \text{for some } s'\in 2^A\\ & \text{iff}_{\text{Fact 2.2.5}} & s\sim_X s''\sim_Y s' \text{ and } s'\in\boldsymbol{X},\ \text{for some } s',s''\in 2^A\\ & \text{iff} & s\sim_X s'' \text{ and } s\in\overline{apr}_Y(\boldsymbol{X}),\ \text{for some } s''\in 2^A\\ & \text{iff} & s\in\overline{apr}_X\left(\overline{apr}_Y(\boldsymbol{X})\right).\end{array}$$

For the latter two claims merely observe that, in virtue of Coroll 2.2.7, both $\underline{apr}_X(\boldsymbol{X})\subseteq\underline{apr}_{X\cup Y}(\boldsymbol{X})$ and $\underline{apr}_Y(\boldsymbol{X})\subseteq\underline{apr}_{X\cup Y}(\boldsymbol{X})$, as well as both $\overline{apr}_{X\cup Y}(\boldsymbol{X})\subseteq\overline{apr}_X(\boldsymbol{X})$ and $\overline{apr}_{X\cup Y}(\boldsymbol{X})\subseteq\overline{apr}_Y(\boldsymbol{X})$. ⊣

Fact 2.2.9 *Let A a set and let I be a set of indices. Let further $\{X_i\}_{i\in I}$ and $\{\boldsymbol{X}_i\}_{i\in I}$ be indexed families of subsets of A and of subsets of 2^A, respectively. Then:*

$$\begin{array}{l}\bigcap_{i\in I}\underline{apr}_{X_i}(\boldsymbol{X}_i)\ \subseteq\ \underline{apr}_{\bigcup_{i\in I}X_i}\left(\bigcap_{i\in I}\boldsymbol{X}_i\right),\\ \bigcap_{i\in I}\overline{apr}_{X_i}(\boldsymbol{X}_i)\ \supseteq\ \overline{apr}_{\bigcup_{i\in I}X_i}\left(\bigcap_{i\in I}\boldsymbol{X}_i\right).\end{array}$$

Proof: First consider an arbitrary $Y\subseteq A$ and assume $Y\in\bigcap_{i\in I}\underline{apr}_{X_i}(\boldsymbol{X}_i)$, *i.e.*, for all $i\in I$, it is the case that $Y\in\underline{apr}_{X_i}(\boldsymbol{X}_i)$. Since, obviously, $X_i\subseteq\bigcup_{i\in I}X_i$, by Corollary 2.2.7 for all $i\in I$, also $Y\in\underline{apr}_{\bigcup_{i\in I}X_i}(\boldsymbol{X}_i)$. Therefore, $Y\in\bigcap_{i\in I}\underline{apr}_{\bigcup_{i\in I}X_i}(\boldsymbol{X}_i)$. Finally, by distribution of $\bigcap$ over $\underline{apr}$, we may conclude that $Y\in\underline{apr}_{\bigcup_{i\in I}X_i}\left(\bigcap_{i\in I}\boldsymbol{X}_i\right)$.

For the second claim, assume for an arbitrary $Y\subseteq A$, that $Y\in\overline{apr}_{\bigcup_{i\in I}X_i}\left(\bigcap_{i\in I}\boldsymbol{X}_i\right)$. Then there is some $Z\subseteq A$ such that $Y\sim_{\bigcup_{i\in I}X_i}Z$ and $Z\in\bigcap_{i\in I}\boldsymbol{X}_i$. It follows that for each $i\in I$, both $Y\sim_{X_i}Z$ and $Z\in\boldsymbol{X}_i$, *i.e.*, $Y\in\overline{apr}_{Y_i}(\boldsymbol{X}_i)$. We may conclude that $Y\in\bigcap_{i\in I}\overline{apr}_{X_i}(\boldsymbol{X}_i)$. ⊣

For the special partitions π_A and $\pi_\emptyset$, moreover the following equalities hold:

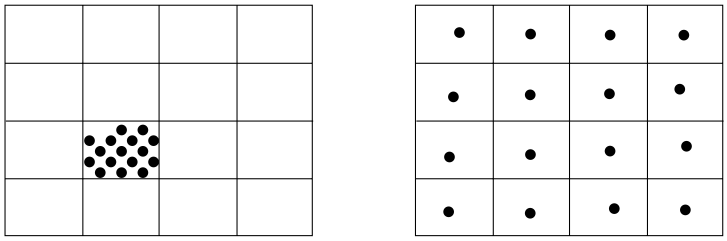


Figure 2.5. The figure on the left shows the partition π_X of a space 2^A, for some subset X of A. The figure in on the right shows the partition $\pi_{\overline{X}}$ on the same space. The elements of each block of π_X will be distributed over the blocks of the partition $\pi_{\overline{X}}$.

Fact 2.2.10 *Let A be a set and let $\boldsymbol{X} \subseteq 2^A$. Then:*

$$\begin{aligned}
\underline{apr}_A(\boldsymbol{X}) &= \overline{apr}_A(\boldsymbol{X}) = \boldsymbol{X} \\
\underline{apr}_\emptyset(\boldsymbol{X}) &= \begin{cases} \boldsymbol{X} & \textit{if } \boldsymbol{X} = 2^A \\ \emptyset & \textit{otherwise} \end{cases} \\
\overline{apr}_\emptyset(\boldsymbol{X}) &= \begin{cases} \boldsymbol{X} & \textit{if } \boldsymbol{X} = \emptyset \\ 2^A & \textit{otherwise.} \end{cases}
\end{aligned}$$

Proof: Observe that $\pi_A = \{\{a\} : a \in A\}$ and that $\pi_\emptyset = \{2^A\}$. Then the claims follow almost immediately from the definitions of upper and lower approximation. $\dashv$

For a subset X of A, the partitions π_X and $\pi_{\overline{X}}$ are closely related. Let the *generalized sum* or *the set of choice sets* over a family of sets $\boldsymbol{X} = \{X_i\}_{i \in I}$ be defined as:

$$\sum_{i \in I} X_i \quad =_{df.} \quad \left\{ \{f(i) : i \in I\} : f : I \to \bigcup_{i \in I} X_i \text{ and } \forall i \in I : f(i) \in X_i \right\}.$$

Omitting explicit reference to the index set, $\sum_{i \in I} X_i$ is also denoted by $\sum \boldsymbol{X}$. Then, the following proposition establishes that π_X is a set of choice sets of $\pi_{\overline{X}}$.

Proposition 2.2.11 *Let A be a set and let $X \subseteq A$. Consider the partitions π_X and $\pi_{\overline{X}}$ of 2^A. Then, π_X is a set of* choice sets *of $\pi_{\overline{X}}$, i.e., $\pi_X \subseteq \sum \pi_{\overline{X}}$.*

Proof: First observe that 2^A is not empty. It suffices to define for each $\pi_i \in \pi_X$ a function $f_{\pi_i} : \pi_{\overline{X}} \to 2^A$ such that (*a*) $f_{\pi_i}(\pi_j) \in \pi_j$, for each $\pi_j \in \pi_{\overline{X}}$ and (*b*) $\pi_i = \{f_{\pi_i}(\pi_j) : \pi_j \in \pi_{\overline{X}}\}$. So consider an arbitrary $\pi_i \in \pi_X$. Assuming the axiom of choice, there are choice functions $g : \pi_X \to 2^A$ and $g' : \pi_{\overline{X}} \to 2^A$ such that for each $\pi_k \in \pi_X$

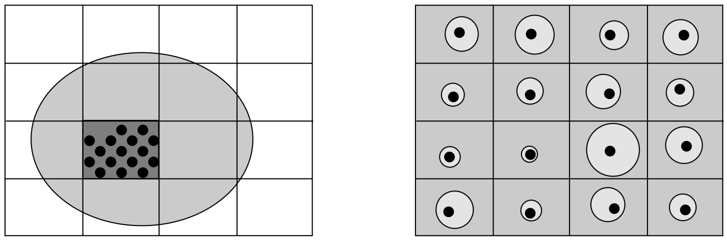


Figure 2.6. The left and right figure represent the partitions π_X and $\pi_{\overline{X}}$ as in Figure 2.5, above. The oval in the left figure and the lighter shaded areas in the right figure represent the subset $\boldsymbol{X}$ of 2^A. The darker areas are $\underline{apr}_X(\boldsymbol{X})$ and $\overline{apr}_{\overline{X}}(\boldsymbol{X})$ in the left and right figure, respectively. Because of the distribution of the elements of each block in π_X contained in $\boldsymbol{X}$ over the blocks of $\pi_{\overline{X}}$, it can now be recognized that $\underline{apr}_X(\boldsymbol{X})$ being non-empty implies $\underline{apr}_X(\boldsymbol{X}) = 2^A$.

and $\pi_j \in \pi_{\overline{X}}$ both $g(\pi_k) \in \pi_k$ and $g'(\pi_j) \in \pi_j$. Now define the function $f_{\pi_i} : \pi_{\overline{X}} \to 2^A$ such that for all $\pi_j \in \pi_{\overline{X}}$:

$$f_{\pi_i}(\pi_j) \quad =_{df.} \quad \big(g(\pi_i) \cap X\big) \cup \big(g'(\pi_j) \cap \overline{X}\big).$$

We show that $\pi_i = \big\{f_{\pi_i}(\pi_j) : \ \pi_j \in \pi_{\overline{X}}\big\}$.

The $\supseteq$-direction is almost immediate as from the definition almost immediately follows that $f_{\pi_i}(\pi_j) \sim_X g(\pi_i)$, for each $\pi_j \in \pi_{\overline{X}}$. Since g is a choice function, $g(\pi_i) \in \pi_i$. Hence, $f_{\pi_i}(\pi_j) \in \pi_i$, for each $\pi_j \in \pi_{\overline{X}}$. For the $\subseteq$-direction, consider an arbitrary $s \in \pi_i$. With 2^A non-empty, we may also assume the existence of s as well as that of the block $[s]_{\overline{X}}$ in $\pi_{\overline{X}}$. It suffices to show that $f_{\pi_i}([s]_{\overline{X}}) = s$. Observe that $\pi_i = [s]_X$. Since g and g' are choice functions, both $g([s]_X) \in [s]_X$ and $g'([s]_{\overline{X}}) \in [s]_{\overline{X}}$. Hence, both $s \sim_X g([s]_X)$ and $s \sim_{\overline{X}} g'([s]_{\overline{X}})$, *i.e.*, $s \cap X = g([s]_X) \cap X$ and $s \cap \overline{X} = g'([s]_{\overline{X}}) \cap \overline{X}$. We can now reason as follows:

$$s \;=\; (s \cap X) \cup \big(s \cap \overline{X}\big) \;=\; (g([s]_X) \cap X) \cup \big(g'([s]_{\overline{X}}) \cap \overline{X}\big) \;=\; f_{[s]_X}([s]_{\overline{X}}) \;=\; f_{\pi_i}([s]_{\overline{X}}).$$

This concludes the proof. ⊣

This proposition has the following corollary, which is also illustrated in Figure 2.6.

Corollary 2.2.12 *Let A be a set, $\boldsymbol{X} \subseteq 2^A$ and $X \subseteq A$. Then, $\underline{apr}_X(\boldsymbol{X}) = \emptyset$ or $\overline{apr}_{\overline{X}}(\boldsymbol{X}) = 2^A$.*

Proof: Assume $\underline{apr}_X(\boldsymbol{X}) \neq \emptyset$; it suffices to show that $\overline{apr}_{\overline{X}}(\boldsymbol{X}) = 2^A$. By the assumption, there is some $\pi_i \in \pi_X$ such that $\pi_i \subseteq \boldsymbol{X}$. Moreover, since 2^A is not empty, neither is π_i. Now consider an arbitrary $\pi_j \in \pi_{\overline{X}}$. By Proposition 2.2.11, then, $\pi_j = \{f(\pi_k) : \ \pi_k \in \pi_X\}$, for some choice function f mapping π_X on 2^A. Therefore, $f(\pi_i) \in \pi_j$. Moreover, $\pi_j \cap \boldsymbol{X} \neq \emptyset$ and so, $\pi_j \subseteq \overline{apr}_{\overline{X}}(\boldsymbol{X})$. With π_j having been chosen arbitrarily and the blocks of π_X exhausting 2^A, we may conclude that $\overline{apr}_{\overline{X}}(\boldsymbol{X}) = 2^A$. ⊣

2.3 Propositional Logics

A *propositional language* $L(A)$ consists of a set of formulas $\Phi(A)$ over some set of propositional variables A. Unless stated otherwise, we assume A to be countable. A *proper logic* for a propositional language $L(A)$ is a set of pairs of theories in $L(A)$, *i.e.*, a subset of $\Phi(A) \times \Phi(A)$.[5] For any logic Λ for $L(A)$ and any pair of theories Γ and Θ we say that Θ *logically follows from* Γ *in* Λ if $(\Gamma, \Theta) \in \Lambda$. In the sequel we will usually denote $(\Gamma, \Theta) \in \Lambda$ by $\Gamma \vdash^{\Lambda} \Theta$ and $(\Gamma, \Theta) \notin \Lambda$ by $\Gamma \nvdash^{\Lambda} \Theta$, also omitting the superscript whenever possible. At an intuitive level, it can be a help to read $\Gamma \vdash \Theta$ as signifying that some formulas in Θ hold whenever all formulas in Γ hold. Two formulas φ and ψ are *logically equivalent*, $\varphi \equiv_{\Lambda} \psi$ if both $\varphi \vdash^{\Lambda} \psi$ and $\psi \vdash^{\Lambda} \varphi$. The sets of *consequences* and *anti-consequences* of a theory Γ is then defined as:

$$
\begin{aligned}
Cn_{\Lambda}(\Gamma) &=_{df.} \left\{\varphi : \ \Gamma \vdash^{\Lambda} \varphi\right\} \\
An_{\Lambda}(\Gamma) &=_{df.} \left\{\varphi : \ \varphi \vdash^{\Lambda} \Gamma\right\}
\end{aligned}
$$

We say that a theory Γ is Λ*-closed* (or simply *closed*) if $\Gamma = Cn_{\Lambda}(\Gamma)$. A theory Δ is said to be a *set of axioms for* a theory Γ iff Γ and Δ have the same consequences. A theory is called *finitely axiomatizable* iff it has a finite set of axioms.

We take a very liberal attitude towards what to consider a logic and impose no further restrictions. Then the logics for a propositional language $L(A)$ can be partially ordered *via* set inclusion. Define $\Lambda \leqslant \Lambda'$ if and only if $\Lambda \subseteq \Lambda'$, for Λ and Λ' logics for $L(A)$. The set of logics for a propositional language $L(A)$ thus constitutes a field of sets[6] and as such a complete lattice and even a Boolean algebra.

We give a brief overview of the most common conditions usually imposed on logics.[7] A logic Λ is *reflexive* if $\Gamma \vdash \Gamma$, for all non-empty theories Γ. Very similar conditions are those of *diagonality* and *overlap*, which are satisfied if $\varphi \vdash \varphi$, for all formulas φ, and, respectively, if $\Gamma \vdash \Theta$, for all theories Γ and Θ such that $\Gamma \cap \Theta \neq \emptyset$. If $\varphi \vdash^{\Lambda} \varphi$, for some formula φ, we say that Λ is *diagonal for* φ. A logic is *monotonic* if for $\Gamma \subseteq \Gamma'$ and $\Theta \subseteq \Theta'$, $\Gamma \vdash \Theta$ implies $\Gamma' \vdash \Theta'$. For monotonic logics the conditions of reflexivity, diagonality and overlap coincide. We say a logic satisfies *cut* if $\Gamma \vdash \Theta \cup \{\varphi\}$ and $\Gamma' \cup \{\varphi\} \vdash \Theta'$ imply $\Gamma \cup \Gamma' \vdash \Theta \cup \Theta'$. Observe that this definition of cut is equivalent to any of the following two conditions for Ξ *finite*:[8]

$$(*) \qquad \Gamma \vdash \Theta \cup \Xi \ \text{ and } \ \Gamma' \cup \{\xi\} \vdash \Theta' \text{ for all } \xi \in \Xi \quad \text{imply} \quad \Gamma \cup \Gamma' \vdash \Theta \cup \Theta'$$

$$(**) \qquad \Gamma \cup \Xi \vdash \Theta \ \text{ and } \ \Gamma' \vdash \Theta' \cup \{\xi\} \text{ for all } \xi \in \Xi \quad \text{imply} \quad \Gamma \cup \Gamma' \vdash \Theta \cup \Theta'.$$

Also consider the following cut-like condition:

$$(***) \quad \varphi \vdash^{\Lambda} \psi, \ \Gamma \vdash^{\Lambda} \Theta \cup \{\varphi\} \ \text{ and } \Gamma' \cup \{\psi\} \vdash^{\Lambda} \Theta' \quad \text{imply} \quad \Gamma \cup \Gamma' \vdash^{\Lambda} \Theta \cup \Theta'.$$

[5] We follow Segerberg[1982] in this definition and the following remarks on logics.

[6] A *field of sets* S is a collection of subsets of a nonempty set X such that both the empty set $\emptyset$ and the set X are in S and S is closed under $\cap$, $\cup$ and $^{-}$ with respect to X (Chang and Keisler, 1973, p.39).

[7] Again we closely follow the exposition of Segerberg[1982], pp.34–39.

[8] Segerberg proposes a stronger and more general notion of cut defined as the conjunction of these two conditions with the restriction that Ξ be finite lifted (*cf.*, *ibid.*, p.37).

Fact 2.3.1 *Any logic Λ for $L(A)$ that satisfies cut also satisfies $(***)$. Moreover, $(***)$ implies cut, provided Λ satisfies diagonality.*

Proof: First assume $\varphi \vdash^{\Lambda} \psi$, $\Gamma \vdash^{\Lambda} \Theta \cup \{\varphi\}$ and $\Gamma' \cup \{\psi\} \vdash^{\Lambda} \Theta'$. From the former two assumptions and cut then $\Gamma \vdash^{\Lambda} \Theta \cup \{\psi\}$. Together with the third assumption and another application of cut, this yields $\Gamma \cup \Gamma' \vdash^{\Lambda} \Theta \cup \Theta$. For the second claim, assume $\Gamma \vdash^{\Lambda} \Theta \cup \{\varphi\}$ and $\Gamma' \cup \{\varphi\} \vdash^{\Lambda} \Theta'$. With diagonality of Λ, we have $\varphi \vdash^{\Lambda} \varphi$. Hence, $(***)$ entails $\Gamma \cup \Gamma' \vdash^{\Lambda} \Theta \cup \Theta'$ follows, which we had set out to prove. ⊣

Finally, a logic is *finite* (or *compact*) if $\Gamma \vdash \Theta$ if and only if there are finite $\Gamma' \subseteq \Gamma$ and $\Theta' \subseteq \Theta$ such that $\Gamma' \vdash \Theta'$. This notion is not to be confused with the notion of a finite semantics, to be introduced presently. We have the following fact:

Fact 2.3.2 *Let Λ be a reflexive, monotonic and finite cut logic for $L(A)$. Then for all theories Γ and Θ in $L(A)$:*

$$\Gamma \vdash^{\Lambda} \Theta \quad \textit{iff} \quad Cn\,(\Gamma) \vdash^{\Lambda} An\,(\Theta)\,.$$

Proof: For the left-to-right direction, observe that since Λ is reflexive, $\Gamma \subseteq Cn\,(\Gamma)$ as well as $\Theta \subseteq An\,(\Theta)$. Hence with monotonicity, $\Gamma \vdash^{\Lambda} \Theta$ implies $Cn\,(\Gamma) \vdash^{\Lambda} An\,(\Theta)$. For the opposite direction assume $Cn\,(\Gamma) \vdash^{\Lambda} An\,(\Theta)$. By finiteness, there are finite theories $\Gamma' \subseteq Cn\,(\Gamma)$ and $\Theta' \subseteq An\,(\Theta)$ such that $\Gamma' \vdash^{\Lambda} \Theta'$. By monotony of Λ then also $\Gamma \cup \Gamma' \vdash^{\Lambda} \Theta \cup \Theta'$. For each $\gamma \in \Gamma' - \Gamma$, we have $\gamma \in Cn\,(\Gamma)$, *i.e.*, $\Gamma \vdash^{\Lambda} \gamma$. Hence, by cut and its equivalence with $(**)$, $\Gamma \vdash^{\Lambda} \Theta \cup \Theta'$. Similarly, $\vartheta \in An\,(\Theta)$ — *i.e.*, it is the case that $\vartheta \vdash^{\Lambda} \Theta$ — for all $\vartheta \in \Theta'$. Therefore, by cut and its equivalence with $(*)$, eventually, $\Gamma \vdash^{\Lambda} \Theta$. ⊣

A logic Λ is *consistent* if $\Lambda \neq 2^{\Phi} \times 2^{\Phi}$. Obviously, there is only one *inconsistent* logic. For monotonic logics, this condition for consistency is equivalent with the pair $(\varnothing, \varnothing)$ not being an element of the logic, *i.e.*, $\varnothing \nvdash \varnothing$.

A *valuation-based semantics* (or just 'semantics') for a language $L(A)$ associates with each formula φ of the language a subset of 2^A which we call the *extension of* φ and denote by $[\![\varphi]\!]$. Here, 2^A is taken as the set of *valuations*, which will in the sequel frequently be referred to by S. Let $s \Vdash \varphi$ if $s \in [\![\varphi]\!]$ and $s \nVdash \varphi$ if $s \notin [\![\varphi]\!]$. The set of extensions of the formulas in a Γ we denote by $\mathscr{E}\,(\Gamma)$, *i.e.*, $\mathscr{E}\,(\Gamma) =_{df.} \left\{[\![\gamma]\!] : \ \gamma \in \Gamma\right\}$. The set of all formula extensions of a language $L(A)$, $\mathscr{E}\,(\Phi(A))$ we usually denote by simply $\mathscr{E}\,(A)$ or even just $\mathscr{E}$, if A is clear from the context. Let, furthermore, $[\![\Gamma]\!] =_{df.} \bigcap_{\gamma\in\Gamma} [\![\gamma]\!]$ and $\langle\!\langle\Gamma\rangle\!\rangle =_{df.} \bigcup_{\gamma\in\Gamma} [\![\gamma]\!]$. *Semantical consequence* is then defined as:

$$\Gamma \vDash \Theta \quad \text{iff} \quad [\![\Gamma]\!] \subseteq \langle\!\langle\Theta\rangle\!\rangle.$$

In a similar vein, a theory is said to be *satisfiable* if $[\![\Gamma]\!] \neq \varnothing$ and *valid* if $[\![\Gamma]\!] = 2^A$. A formula φ is *satisfiable* or *valid* if $\{\varphi\}$ is, respectively, satisfiable or valid. Any binary relation on the theories of a language $L(A)$ is said to be *sound* with respect to a logic Λ if it is a subset of Λ and *complete* whenever it is a superset of Λ.

We call a valuation-based semantics for a language $L(A)$ *finite* if for each formula φ of $L(A)$ there is a finite subset $X \subseteq_\omega A$ such that:

$$s \in [\![\varphi]\!] \text{ and } s \sim_X s' \quad \text{implies} \quad s' \in [\![\varphi]\!]$$

In terms of rough sets this means that for all formulas φ of $L(A)$ there is a finite set $X \subseteq A$ such that:

$$[\![\varphi]\!] \;=\; \overline{apr}_X([\![\varphi]\!])\,.$$

It now happens that if a semantics is finite, then for each formula φ in $L(A)$ there is a *smallest* finite set such that $[\![\varphi]\!] = \overline{apr}_X([\![\varphi]\!])$. This proposition is a corollary of the following lemma in rough set theory:

Lemma 2.3.3 *Let A be a countable set and $\boldsymbol{Z} \subseteq 2^A$. Assume further that there exists some finite $Z_0 \subseteq_\omega 2^A$ in $\boldsymbol{Z}$. Let $\boldsymbol{X} \subseteq 2^A$. Then:*

$$\overline{apr}_{Z_0}(\boldsymbol{X}) = \overline{apr}_Z(\boldsymbol{X}) \;\; \textit{for all } Z \in \boldsymbol{Z} \quad \textit{implies} \quad \overline{apr}_{Z_0}(\boldsymbol{X}) = \overline{apr}_{\bigcap \boldsymbol{Z}}(\boldsymbol{X})\,.$$

Proof: Assume for all $Z \in \boldsymbol{Z}$: $\overline{apr}_{Z_0}(\boldsymbol{X}) = \overline{apr}_Z(\boldsymbol{X})$. Since $\bigcap \boldsymbol{Z} \subseteq Z_0$, also $\overline{apr}_{Z_0}(\boldsymbol{X}) \subseteq \overline{apr}_{\bigcap \boldsymbol{Z}}(\boldsymbol{X})$, in virtue of Fact 2.2.7. So consider an arbitrary $s \in \overline{apr}_{\bigcap \boldsymbol{Z}}(\boldsymbol{X})$. We prove that $s \in \overline{apr}_{Z_0}(\boldsymbol{X})$. Then, for some $s_0 \in \boldsymbol{X}$, $s_0 \sim_{\bigcap \boldsymbol{Z}} s$. Since Z_0 is finite, so are $\bigcap \boldsymbol{Z}$ and $Z_0 - \bigcap \boldsymbol{Z}$. Let $Z_0 - \bigcap \boldsymbol{Z} = \{z_0, \ldots, z_n\}$. Observe that for each $z \in \{z_0, \ldots, z_n\}$, there is some $Z \in \boldsymbol{Z}$ such that $z \notin Z$. Assuming the axiom of choice, let $\{Z'_0, \ldots, Z'_n\} \subseteq \boldsymbol{Z}$ be such that $z_i \notin Z'_i$, for each $i \leqslant n$. For each $0 \leqslant i \leqslant n+1$, define s^*_i as follows:

$$\begin{aligned} s^*_0 \quad &=_{df.} \quad s_0 \\ s^*_{i+1} \quad &=_{df.} \quad \big(s^*_i - \{z_i\}\big) \cup \big(s \cap \{z_i\}\,\big) \end{aligned}$$

Since by definition $z_i \notin Z'_i$, and with s^*_i and s^*_{i+1} differing at most at z_i, it follows that for each $0 \leqslant i \leqslant n+1$, $s^*_i \sim_{Z'_i} s^*_{i+1}$. As a consequence $s^*_{i+1} \in \overline{apr}_{Z_0}(\boldsymbol{X})$, for each $i \leqslant n+1$. To appreciate this, observe that by assumption $s^*_0 \in \boldsymbol{X}$ and hence also $s^*_0 \in \overline{apr}_{Z_0}(\boldsymbol{X})$. Now assume $s^*_i \in \overline{apr}_{Z_0}(\boldsymbol{X})$. By the initial assumption $\overline{apr}_{Z_0}(\boldsymbol{X}) = \overline{apr}_{Z'_i}(\boldsymbol{X})$, so $s^*_i \in \overline{apr}_{Z'_i}(\boldsymbol{X})$. Since, moreover, $s^*_i \sim_{Z'_i} s^*_{i+1}$, $s^*_{i+1} \in \overline{apr}_{Z'_i}\left(\overline{apr}_{Z'_i}(\boldsymbol{X})\right)$. Then also $s^*_{i+1} \in \overline{apr}_{Z'_i}(\boldsymbol{X})$ and eventually $s^*_{i+1} \in \overline{apr}_{Z_0}(\boldsymbol{X})$.

We now prove by induction on i that for all $i \leqslant n+1$ (letting $\{z_k, \ldots, z_n\} = \varnothing$ if $k > n$):

$$s^*_i \sim_{Z_0 - \{z_i, \ldots, z_n\}} s.$$

For $i = 0$, recall that by assumption $s^*_0 \sim_{\bigcap \boldsymbol{Z}} s$, which is exactly what we have to prove considering that $\{z_0, \ldots, z_n\} = Z_0 - \bigcap \boldsymbol{Z}$. For the induction step, we may assume that $s^*_i \sim_{Z_0 - \{z_i, \ldots, z_n\}} s$. Now consider s_{i+1} as well as an arbitrary $z \in Z_0 - \{z_{i+1}, \ldots, z_n\}$. If $z \in Z_0 - \{z_i, \ldots, z_n\}$, just observe that $s^*_{i+1} \sim_{Z_0 - \{z_i, \ldots, z_n\}} s^*_i \sim_{Z_0 - \{z_i, \ldots, z_n\}} s$. If, however, the only remaining possibility obtains and $z = z_i$, then also $s^*_{i+1} \sim_{\{z_i\}} s$. Hence, we may conclude that $s^*_{i+1} \sim_{Z_0 - \{z_{i+1}, \ldots, z_n\}} s$.

In particular it holds that $s^*_{n+1} \in \overline{apr}_{Z_0}(\boldsymbol{X})$ and that $s^*_{n+1} \sim_{Z_0} s$. Hence, $s \in \overline{apr}_{Z_0}\big(\overline{apr}_{Z_0}(\boldsymbol{X})\big)$, which is so much as to say that $s \in \overline{apr}_{Z_0}(\boldsymbol{X})$. Wrapping things up, we recall that s had been chosen arbitrarily such that $s_0 \sim_{\bigcap \boldsymbol{Z}} s$. So, we may conclude that $\overline{apr}_{\bigcap \boldsymbol{Z}}(\boldsymbol{X}) \subseteq \overline{apr}_{Z_0}(\boldsymbol{X})$. ⊣

Observe that Lemma 2.3.3 *does not* hold in general if $\boldsymbol{Z}$ does not contain at least one finite element. For a counterexample, consider a countably infinite set A and let $\boldsymbol{X}$ be the set of infinite subsets of A. Hence $\boldsymbol{X} \neq 2^A$. Let further $a_0, \ldots, a_n, \ldots$ be an enumeration of A and set $\boldsymbol{Z} =_{df.} \big\{A - \{a_0, \ldots, a_n\} : \; n \in \omega\big\}$. Clearly, $\bigcap \boldsymbol{Z} = \varnothing$ and so $\overline{apr}_{\bigcap \boldsymbol{Z}}(\boldsymbol{X}) = 2^A$. However, since every $X \in \boldsymbol{X}$ has an infinite intersection with any $Z \in \boldsymbol{Z}$, we also have $\overline{apr}_Z(\boldsymbol{X}) = \overline{apr}_{Z'}(\boldsymbol{X})$ for all $Z, Z' \in \boldsymbol{Z}$. However, if $\boldsymbol{Z}$ is itself finite, the restriction of it containing a finite element can be dropped. Just recall that $\overline{apr}_Z(\overline{apr}_{Z'}(\boldsymbol{X})) = \overline{apr}_{Z \cap Z'}(\boldsymbol{X})$. We are now in a position to prove the following proposition.

Proposition 2.3.4 *For every finite semantics for $L(A)$ and each formula φ of $L(A)$ there is a (unique) smallest $X \subseteq A$ such that $[\![\varphi]\!] = \overline{apr}_X([\![\varphi]\!])$.*

Proof: Consider a finite semantics for $L(A)$ along with an arbitrary formula φ. Now consider the set $\boldsymbol{Z} =_{df.} \big\{Z \subseteq A : \; \overline{apr}_Z([\![\varphi]\!]) = [\![\varphi]\!]\big\}$. By finiteness, there is also a finite $Z_0 \in \boldsymbol{Z}$. Hence for all $Z \in \boldsymbol{Z}$, $\overline{apr}_Z([\![\varphi]\!]) = \overline{apr}_{Z_0}([\![\varphi]\!])$. By Lemma 2.3.3, $\overline{apr}_{\bigcap \boldsymbol{Z}}([\![\varphi]\!]) = \overline{apr}_{Z_0}([\![\varphi]\!]) = [\![\varphi]\!]$. Hence, by definition of $\boldsymbol{Z}$, also $\bigcap \boldsymbol{Z} \in \boldsymbol{Z}$, which proves the proposition. ⊣

This result warrants the definition of $\boldsymbol{A}(\varphi)$ as the *smallest* subset of propositional variables in A for which $[\![\varphi]\!] = \overline{apr}_X([\![\varphi]\!])$. More in general we set $\boldsymbol{A}(\Gamma) =_{df.} \bigcup_{\gamma \in \Gamma} \boldsymbol{A}(\gamma)$. Moreover we employ the notation $\varphi\,(a_0, \ldots, a_n)$ to indicate that $[\![\varphi]\!] = \overline{apr}_{\{a_0,\ldots,a_n\}}([\![\varphi]\!])$. Observe that $\boldsymbol{A}(\Gamma)$ does not in general denote the smallest subset X of propositional variables such that $[\![\Gamma]\!] = \overline{apr}_X([\![\Gamma]\!])$. We have the following two facts.

Fact 2.3.5 *Let φ be a formula of a propositional language $L(A)$. Then for all subsets Δ of A such that $\boldsymbol{A}(\varphi) \subseteq \Delta$:*

$$\overline{apr}_{\Delta}([\![\varphi]\!]) \;=\; \underline{apr}_{\Delta}([\![\varphi]\!]) \;=\; [\![\varphi]\!].$$

Proof: Merely consider the following equalities:

$$\begin{aligned}\overline{apr}_{\Delta}([\![\varphi]\!]) \;=\; \overline{apr}_{\Delta}\left(\overline{apr}_{\boldsymbol{A}(\varphi)}([\![\varphi]\!])\right) \;&=_{\text{Fact 2.2.8}} \\ \overline{apr}_{\Delta \cap \boldsymbol{A}(\varphi)}([\![\varphi]\!]) \;&=_{\boldsymbol{A}(\varphi) \subseteq \Delta}\; \overline{apr}_{\boldsymbol{A}(\varphi)}([\![\varphi]\!]) \;=\; [\![\varphi]\!].\end{aligned}$$

That then also $\underline{apr}_{\Delta}([\![\varphi]\!]) = [\![\varphi]\!]$, follows immediately from Fact 2.2.1. ⊣

Fact 2.3.6 *Let Γ be a theory in $L(A)$ and Δ a subset of A such that $\mathbf{A}(\Gamma) \subseteq \Delta$. Then:*

$$\underline{apr}_{\Delta}(\llbracket\Gamma\rrbracket) = \overline{apr}_{\Delta}(\llbracket\Gamma\rrbracket) = \llbracket\Gamma\rrbracket.$$
$$\underline{apr}_{\Delta}(\langle\!\langle\Gamma\rangle\!\rangle) = \overline{apr}_{\Delta}(\langle\!\langle\Gamma\rangle\!\rangle) = \langle\!\langle\Gamma\rangle\!\rangle.$$

Proof: Consider the following equalities:

$$\begin{array}{lllllll}\underline{apr}_{\Delta}(\llbracket\Gamma\rrbracket) & = & \bigcap_{\gamma\in\Gamma}\underline{apr}_{\Delta}(\llbracket\gamma\rrbracket) & =_{\text{Fact 2.3.5}} & \bigcap_{\gamma\in\Gamma}\llbracket\gamma\rrbracket & = & \llbracket\Gamma\rrbracket, \\ \overline{apr}_{\Delta}(\llbracket\Gamma\rrbracket) & = & \bigcup_{\gamma\in\Gamma}\overline{apr}_{\Delta}(\llbracket\gamma\rrbracket) & =_{\text{Fact 2.3.5}} & \bigcup_{\gamma\in\Gamma}\llbracket\gamma\rrbracket & = & \llbracket\Gamma\rrbracket.\end{array}$$

That also $\overline{apr}_{\Delta}(\llbracket\Gamma\rrbracket) = \llbracket\Gamma\rrbracket$ and $\underline{apr}_{\Delta}(\llbracket\Gamma\rrbracket) = \llbracket\Gamma\rrbracket$ then follows by Fact 2.2.1. $\dashv$

A *classical propositional language* $L(A)$ over a set of propositional variables A is a minimal set containing A as well as $\bot$ and for each formula φ contains another formula denoted by $(\neg\varphi)$ as well as for each pair of formulas φ and ψ a formula denoted by $(\varphi \vee \psi)$ and allows for a classical semantics. A *classical semantics* for a classical language $L(A)$ is such that for each propositional variable $a \in A$ and all formulas φ and ψ:

$$\begin{array}{rcl}\llbracket a\rrbracket & = & \{s \in S : \ a \in s\} \\ \llbracket\bot\rrbracket & = & \varnothing \\ \llbracket(\neg\varphi)\rrbracket & = & S - \llbracket\varphi\rrbracket \\ \llbracket(\varphi\vee\psi)\rrbracket & = & \llbracket\varphi\rrbracket \cup \llbracket\psi\rrbracket.\end{array}$$

The resulting logic we will refer to as *classical propositional logic* (CPC). Where possible we omit parentheses. We also have the usual abbreviations $\top$, $\varphi \wedge \psi$, $\varphi \rightarrow \psi$, $\bigwedge\Theta$ and $\bigvee\Theta$, where φ and ψ are formulas of $L(A)$ and Θ a finite and possibly empty sequences of formulas in $L(A)$. For each formula φ, the set of propositional variables occurring in φ is defined as usual and depicted by $A(\varphi)$. We use $A(\Gamma)$ to denote $\bigcup_{\gamma\in\Gamma} A(\gamma)$. The set algebra over the set of extension of a classical propositional language $L(A)$, *i.e.*, $\left(\mathscr{E}; \varnothing, 2^A, {}^-, \cup, \cap\right)$, we denote by $\mathfrak{E}_A$. For a classical propositional language $L(A)$, for each formula φ, define $[\varphi]_\equiv =_{df.} \{\psi : \ \varphi \equiv_A \psi\}$. Then $\left(\{\llbracket\varphi\rrbracket_\equiv : \ \varphi \in \Phi\}; [\bot]_\equiv, [\top]_\equiv, \neg, \vee, \wedge\right)$ is the *Lindenbaum algebra of* $L(A)$, also denoted by $\mathfrak{L}$. Here $\neg$, $\vee$ and $\wedge$ are the operations $\neg$, $\vee$ and $\wedge$ raised as to apply to equivalence classes of formulas.[9] Classical propositional logic is compact.

Fact 2.3.7 *(Compactness of* CPC*) Let $L(A)$ be a classical propositional language and Γ a theory in $L(A)$. Then:*

$$\Gamma \textit{ is satisfiable} \quad \textit{iff} \quad \textit{every finite subtheory } \Gamma' \textit{ of } \Gamma \textit{ is satisfiable}.$$

[9] *I.e.*, $\neg[\varphi] =_{df.} [\neg\varphi]$, $[\varphi] \vee [\psi] =_{df.} [\varphi \vee \psi]$ and $[\varphi] \wedge [\psi] =_{df.} [\varphi \wedge \psi]$. These definitions are representative independent.

Proof: *Cf.*, Barwise (1977), pages 26–28. $\dashv$

We state without proof that Fact 2.3.7 has the finiteness of CPC as a corollary.

Fact 2.3.8 CPC *is finite. I.e., for Γ and Θ theories in a classical propositional language $L(A)$:*

$$\Gamma \vDash^{\mathrm{CPC}} \Theta \textit{ implies there are finite subsets } \Gamma' \subseteq \Gamma \textit{ and } \Theta' \subseteq \Theta \textit{ such that } \Gamma' \vDash^{\mathrm{CPC}} \Theta'.$$

Classical semantics is also finite in the sense that for each formula φ there is a finite set Δ such that $[\![\varphi]\!] = \overline{apr}_\Delta([\![\varphi]\!])$.

Fact 2.3.9 *Classical semantics is finite.*

Proof: Consider an arbitrary classical propositional language $L(A)$ along with an equally arbitrary formula φ of $L(A)$. The proof is then by induction on φ.

First assume $\varphi = a$. Obviously $[\![a]\!] \subseteq \overline{apr}_{\{a\}}([\![a]\!])$; so, it suffices to show that $\overline{apr}_{\{a\}}([\![a]\!]) \subseteq [\![a]\!]$. Observe that for any $s \in \overline{apr}_{\{a\}}([\![a]\!])$, there is some $s' \in [\![a]\!]$ such that $s \sim_{\{a\}} s'$. Since by definition $a \in s'$, also $a \in s$. Hence, $s \in [\![a]\!]$.

Let $\varphi = \neg\psi$. In virtue of the induction hypothesis we may assume there to be a finite $\Delta \subseteq_\omega A$ such that $[\![\psi]\!] = \overline{apr}_\Delta([\![\psi]\!])$. Now consider the following equalities:

$$[\![\neg\psi]\!] \quad = \quad \overline{[\![\psi]\!]} \quad =_{\text{i.h.}} \quad \overline{\overline{apr}_\Delta([\![\psi]\!])} \quad = \quad \overline{\overline{\underline{apr}_\Delta\,(\overline{[\![\psi]\!]})}} \quad =_{\text{i.h.}} \quad \underline{apr}_\Delta([\![\neg\psi]\!])\,.$$

With Fact 2.2.1, we may conclude that $[\![\neg\psi]\!] = \overline{apr}_\Delta(\neg\psi)$.

In case $\varphi = \psi \vee \psi$, the induction hypothesis grants us there to be finite $\Delta, \Delta' \subseteq_\omega A$ such that $[\![\psi]\!] = \overline{apr}_\Delta([\![\psi]\!])$ and $[\![\chi]\!] = \overline{apr}_{\Delta'}([\![\chi]\!])$. Now consider the following equalities:

$$\begin{aligned}
\overline{apr}_{\Delta\cup\Delta'}([\![\psi \vee \chi]\!]) \quad &= \quad \overline{apr}_{\Delta\cup\Delta'}([\![\varphi]\!] \cup [\![\psi]\!]) \\
&= \quad \overline{apr}_{\Delta\cup\Delta'}([\![\psi]\!]) \cup \overline{apr}_{\Delta\cup\Delta'}([\![\chi]\!]) \\
&=_{\text{i.h.}} \quad \overline{apr}_{\Delta\cup\Delta'}(\overline{apr}_\Delta([\![\psi]\!])) \cup \overline{apr}_{\Delta\cup\Delta'}(\overline{apr}_{\Delta'}([\![\chi]\!])) \\
&= \quad \overline{apr}_{(\Delta\cup\Delta')\cap\Delta}([\![\psi]\!]) \cup \overline{apr}_{(\Delta\cup\Delta')\cap\Delta'}([\![\chi]\!]) \\
&= \quad \overline{apr}_\Delta([\![\psi]\!]) \cup \overline{apr}_{\Delta'}([\![\chi]\!]) \\
&=_{\text{i.h.}} \quad [\![\psi]\!] \cup [\![\chi]\!] \\
&= \quad [\![\psi \vee \chi]\!].
\end{aligned}$$

This concludes the proof. $\dashv$

For classical propositional logic we have in general $\boldsymbol{A}(\varphi) \subseteq A(\varphi)$. Hence the following fact.

Fact 2.3.10 *Let $L(A)$ be a classical propositional language, $\Delta \subseteq A$ and φ a formula of $L(A)$. Then:*

$$\begin{array}{rclcl} \underline{apr}_{\Delta}(\llbracket\varphi\rrbracket) & = & \underline{apr}_{\Delta\cap \boldsymbol{A}(\varphi)}(\llbracket\varphi\rrbracket) & = & \underline{apr}_{\Delta\cap A(\varphi)}(\llbracket\varphi\rrbracket) \\ \overline{apr}_{\Delta}(\llbracket\varphi\rrbracket) & = & \overline{apr}_{\Delta\cap \boldsymbol{A}(\varphi)}(\llbracket\varphi\rrbracket) & = & \overline{apr}_{\Delta\cap A(\varphi)}(\llbracket\varphi\rrbracket) \end{array}$$

Proof: Immediately from the Facts 2.2.8 and Fact 2.3.5 and the fact that for classical propositional in general $\boldsymbol{A}(\varphi) \subseteq A(\varphi)$. $\dashv$

In CPC also the following version of cut holds.

Fact 2.3.11 *Let Γ and Θ be theories in $L(A)$ and let $\Delta \subseteq A$. Then:*

$$\textit{if}\ \ \Gamma \cup \Delta' \vDash^{\text{CPC}} \Theta \cup (\Delta - \Delta')\,,\ \textit{for all}\ \Delta' \subseteq \Delta\ \ \textit{then}\ \ \Gamma \vDash^{\text{CPC}} \Theta.$$

Proof: By contraposition. Assume that $\Gamma \nvDash^{\text{CPC}} \Theta$. Then, there exists some valuation s such that $s \Vdash \gamma$ for all $\gamma \in \Gamma$ and $s \nVdash \vartheta$ for all $\vartheta \in \Theta$. Consider this valuation s along with the subsets of A given by $\Delta \cap s$ and $\Delta \cap \bar{s}$. Clearly, $\Delta - (\Delta \cap s) = \Delta \cap \bar{s}$. Then, $s \Vdash \gamma$ for all $\gamma \in \Gamma \cup (\Delta \cap s)$ and $s \nVdash \vartheta$ for all $\vartheta \in \Theta \cup (\Delta \cap \bar{s})$. Consequently, there is some $\Delta' \subseteq \Delta$ — *viz.*, $\Delta \cap s$ — such that $\Gamma \cup \Delta' \nvDash^{\text{CPC}} \Theta \cup (\Delta - \Delta')$. $\dashv$

We now introduce some terminology relating to formal systems, derivations and Gentzen-style sequent systems. A *formal system* for a formal language L is a number of *axioms* together with a number of *rules*. The axioms are sequences of formulas of L and the rules are relations between a finite number of sequences $\Sigma_0, \ldots, \Sigma_n$ and a finite sequence of formulas $T_0, \ldots, T_m$, allowing one to derive from all of $\sigma_0, \ldots, \sigma_n$, any one of $T_0, \ldots, T_m$. *I.e.*, for $\Sigma_0, \ldots, \Sigma_n$ and $T_0, \ldots, T_m$ sequences, a rule is denoted by:

$$\frac{\Sigma_0, \ldots, \Sigma_n}{T_0, \ldots, T_m}.$$

A *derivation* is then a finite sequence of sequences $\Sigma_0, \ldots, \Sigma_k$ of formulas such that each Σ_i ($0 \leqslant i \leqslant k$), is either an axiom or there is a rule $\dfrac{T_0, \ldots, T_n}{\Upsilon_0, \ldots, \Upsilon_m}$ such that $\Sigma_i = \Upsilon_j$ for some $0 \leqslant j \leqslant m$ and each T_i ($0 \leqslant i \leqslant n$) is identical to some Σ_{i_j} with $0 \leqslant i_j < i$.

For Σ and T finite sequences of formulas in $L(A)$, $\Sigma \Rightarrow T$ is called a *sequent*. A Gentzen-type system is a formal system containing a number of sequents as axioms and rules enabling one to derive a sequent, *the succedents of the rule*, from a finite number of other sequents, the *antecedents of the rule*. We use $\vdash \Sigma \Rightarrow T$ to denote the existence of a derivation of the sequent $\Sigma \Rightarrow T$. For Γ and Θ possibly infinite theories in $L(A)$, define $\Gamma \vdash \Theta$ if and only if $\vdash \Sigma \Rightarrow T$, for some $\Sigma \in \Gamma^*$ and $T \in \Theta^*$. The Gentzen-style systems GPC and GP, which are sound and complete with respect to CPC, are given in Table 2.4.

Axioms:

$$(0) \quad \bot \Rightarrow \epsilon \qquad (1) \quad \epsilon \Rightarrow \top \qquad (2) \quad a \Rightarrow a$$

Logical Rules:

$$\neg_L : \frac{\Sigma \Rightarrow T, \varphi}{\Sigma, \neg\varphi \Rightarrow T} \qquad \neg_R : \frac{\Sigma, \varphi \Rightarrow T}{\Sigma \Rightarrow T, \neg\varphi}$$

$$\wedge_L : \frac{\Sigma, \varphi, \psi \Rightarrow T}{\Sigma, \varphi \wedge \psi \Rightarrow T} \qquad \wedge_R : \frac{\Sigma \Rightarrow T, \varphi \qquad \Sigma \Rightarrow T, \psi}{\Sigma \Rightarrow T, \varphi \wedge \psi}$$

$$\vee_L : \frac{\Sigma, \varphi \Rightarrow T \qquad \Sigma, \psi \Rightarrow T}{\Sigma, \varphi \vee \psi \Rightarrow T} \qquad \vee_R : \frac{\Sigma \Rightarrow T, \varphi, \psi}{\Sigma \Rightarrow T, \varphi \vee \psi}$$

Structural Rules:

$$contr_L : \frac{\Sigma, \varphi, \varphi \Rightarrow T}{\Sigma, \varphi \Rightarrow T} \qquad contr_R : \frac{\Sigma \Rightarrow T, \varphi, \varphi}{\Sigma \Rightarrow T, \varphi}$$

$$perm_L : \frac{\Sigma, \varphi, \psi, P \Rightarrow T}{\Sigma, \psi, \varphi, P \Rightarrow T} \qquad perm_R : \frac{\Sigma \Rightarrow T, \varphi, \psi, \Upsilon}{\Sigma \Rightarrow T, \psi, \varphi, \Upsilon}$$

$$thin_L : \frac{\Sigma \Rightarrow T}{\Sigma, \varphi \Rightarrow T} \qquad thin_R : \frac{\Sigma \Rightarrow T}{\Sigma \Rightarrow T, \varphi}$$

$$cut : \frac{\Sigma \Rightarrow T, \varphi \qquad \Sigma, \varphi \Rightarrow T}{\Sigma \Rightarrow T}$$

Table 2.4. The System GPC. GP is like GPC except that GP lacks *cut*.

Fact 2.3.12 *(Soundness and completeness of* GPC*) The systems* GPC *and* GP *is complete with respect to* CPC*, i.e., for all theories Γ and Θ of a classical propositional language $L(A)$:*

$$\Gamma \models^{\mathrm{CPC}} \Theta \quad \textit{iff} \quad \Gamma \vdash_{\mathrm{GP}} \Theta \quad \textit{iff} \quad \Gamma \vdash_{\mathrm{GPC}} \Theta.$$

Sketch of proof: Soundness is by a straightforward induction on the length of the derivation in GP. Completeness is as in Barwise (1977), page 38–39. $\dashv$

2.4 Cylindrification in Propositional Logic

In the previous section we argued that $\overline{apr}_\pi$ can be construed as a cylindrification operation and $\underline{apr}_\pi$ as its dual. In first-order logic the quantifiers can be thought of in a similar manner. Quantification increases the expressive power of first-order logic tremendously and is responsible for Church's famous undecidability result. These phenomena connected with cylindrification, however, do not occur at all in classical propositional logic. As a matter of fact, the set of extensions $\mathscr{E}$ in classical propositional logic is closed under taking lower and upper approximations. Hence for each subset X of propositional variables we may assume the existence of formulas $\langle X\rangle\,\varphi$ and $[X]\,\varphi$ with the respective extensions $\overline{apr}_X(\llbracket\varphi\rrbracket)$ and $\underline{apr}_X(\llbracket\varphi\rrbracket)$. This we prove. Moreover we characterize the set of extensions of a classical propositional language as the set of fixed points of all operations $\overline{apr}_X$ and $\underline{apr}_X$ with X finite. Recall that $\mathscr{E}(A)$ denotes the set $\{\llbracket\varphi\rrbracket :\ \varphi \in \Phi(A)\}$.

Theorem 2.4.1 *Let $L(A)$ be a classical propositional language with a classical semantics and let $Fix\,(\overline{apr}_B) =_{df.} \{X \subseteq S :\ X = \overline{apr}_B(X)\}$. Then:*

$$\mathscr{E} \;=\; \bigcup_{B \subseteq_\omega A} Fix\,(\overline{apr}_B)\,.$$

Proof: The left-to-right direction is immediate by finiteness of classical semantics.

For the opposite direction consider an arbitrary $X \subseteq S$ such that $X = \overline{apr}_B(X)$, for some finite $B \subseteq A$. Now define for each $s \subseteq B$:

$$\beta_s \quad =_{df.} \quad \bigwedge\{b, \neg b' :\ b, b' \in B \text{ and } b \in s \text{ and } b' \notin s\}.$$

Obviously, $s \Vdash \beta_s$, for each $s \subseteq B$. Now set:

$$\beta \quad =_{df.} \quad \bigvee\{\beta_s :\ s \subseteq B \text{ and } s \in X\}.$$

We prove that $X = \llbracket\beta\rrbracket$. First assume, for an arbitrary $s \in S$, that $s \in \llbracket\beta\rrbracket$. Then $s \Vdash \beta_{s'}$, for some $s' \in X$ with $s' \subseteq B$. Some reflection reveals that $s \sim_B s'$ and subsequently $s \in \overline{apr}_A(X)$. With the assumption that $X = \overline{apr}_B(X)$, the latter is equivalent with $s \in X$.

Conversely, assume $s \in X$. Define $s^* =_{df.} s \cap B$; then $s^* \Vdash \beta_{s^*}$. Since, moreover, $s \sim_B s^*$ and β_{s^*} only depends on B we also have that $s \Vdash \beta_{s^*}$. It is equally clear that $s^* \subseteq B$. Since $s^* \sim_B s$, also $s^* \in \overline{apr}_B(X)$, *i.e.*, $s^* \in X$ by the assumption. *A fortiori*, also $s \Vdash \beta$, and we may conclude that $s \in \llbracket\beta\rrbracket$. $\dashv$

This result is very close to the more syntactically flavored fact of classical propositional logic that each formula is equivalent to a complete disjunctive normal form. Observe that if Δ comprises precisely the propositional variables occurring in a formula φ, then each disjunct of its complete disjunctive normal form characterizes a block in the partition π_Δ.

Corollary 2.4.2 *Let $L(A)$ be a classical propositional language. Then:*

$$\mathscr{E} \quad = \quad \left\{ \overline{apr}_B(X) : \ B \subseteq_\omega A \ \ and \ \ X \subseteq 2^A \right\}.$$

Proof: The inclusion of $\mathscr{E}$ in $\left\{\overline{apr}_B(X) : \ B \subseteq_\omega A \text{ and } X \subseteq 2^A\right\}$ is an immediate consequence of Theorem 2.4.1. For the opposite inclusion just observe that $\overline{apr}\,(\overline{apr}\,(X)) = \overline{apr}\,(X)$ is a law of rough set theory and again Theorem 2.4.1. $\dashv$

This corollary establishes classical propositional logic as the most expressive one with a finite semantics, in the sense that the extensions of its formulas exhaust the set of valuations that can be finitely approximated.

So, the set of extensions $\mathscr{E}\,(A)$ of formulas of a propositional language $L(A)$ is given by the fixed points of the approximation operations $\overline{apr}_B$ on sets of valuations with B finite. Obviously, $\mathscr{E}\,(A)$ does not exhaust in general the powerset of valuations 2^A. If A is countably infinite, so is the set of formulas of $L(A)$. The set of valuations, not to mention the set of sets of valuations, however, is uncountably infinite and so there can impossibly be a formula for each subset of valuations, or even for each valuation.

The set of theories of a propositional language $L(A)$ will be uncountable if A is countably infinite. Nevertheless, there will still be subsets of valuations that are not the extension of some theory. For an example consider $2^A - \{\emptyset\}$. For a *reductio ad absurdum* assume that $[\![\Gamma]\!] = 2^A - \{\emptyset\}$. Then there is at least one γ in Γ such that $[\![\gamma]\!]$ does not contain the empty set $\emptyset$. In virtue of Corollary 2.4.2, there is a finite subset B of propositional variables and some subset of valuations X such that $[\![\gamma]\!] = \overline{apr}_B(X)$. Now consider the *valuation* $\overline{B}$. Observe that with B finite and A infinite, $\overline{B}$ is not empty. Hence, $\overline{B} \in [\![\Gamma]\!]$ and *a fortiori* also $\overline{B} \in [\![\gamma]\!]$. Therefore, $\overline{B} \in \overline{apr}_B(X)$. Evidently, $\overline{B} \sim_B \emptyset$ and so $\emptyset \in \overline{apr}_B(\overline{apr}_B(X)) = \overline{apr}_{B\cap B}(X) = \overline{apr}_B(X) = [\![\gamma]\!]$, which is at variance with the assumption that $\emptyset \notin [\![\gamma]\!]$.

More important for our purposes, however, is that Theorem 2.4.1 and Corollary 2.4.2 pave the way for the following fact, which warrants the introduction of the 'rough-set' quantifiers to the propositional language.

Fact 2.4.3 *Let $L(A)$ be a classical propositional language with a classical semantics. Then for all $B \subseteq A$:*

$$X \in \mathscr{E} \quad implies \quad \overline{apr}_B(X) \in \mathscr{E}$$

Proof: At page 41 it was stated as a law of rough set theory that for any $B, C \subseteq A$, $\overline{apr}_B(\overline{apr}_C(X)) = \overline{apr}_{B\cap C}(X)$. Now consider an arbitrary $X \in \mathscr{E}$. By Theorem 2.4.1 there is a finite $C \subseteq A$ such that $X = \overline{apr}_C(X)$. Now consider an arbitrary $B \subseteq A$ and reason as follows:

$$\overline{apr}_B(X) \quad = \quad \overline{apr}_B(\overline{apr}_C(X)) \quad = \quad \overline{apr}_{B\cap C}(X)\,.$$

Since obviously $B \cap C$ is finite, with Corollary 2.4.2, $\overline{apr}_{B\cap C}(X) \in \mathscr{E}$ and we may conclude that $\overline{apr}_B(X) \in \mathscr{E}$. $\dashv$

From this fact follows that for each formula φ in $L(A)$ and $\Delta \subseteq A$ there are formulas ξ, ξ' in the language such that:

$$\begin{array}{rcl} \overline{apr}_{\Delta}(\llbracket\varphi\rrbracket) & = & \llbracket\xi\rrbracket \\ \underline{apr}_{\Delta}(\llbracket\varphi\rrbracket) & = & \llbracket\xi'\rrbracket \end{array}$$

We denote these formulas by, respectively, $\langle\Delta\rangle\,\varphi$ and $[\Delta]\,\varphi$. The following two clauses rephrase their semantics:

$$\begin{array}{rll} s \Vdash \langle\Delta\rangle\,\varphi & \text{iff} & \text{for some } s' \in S:\ s \sim_{\Delta} s' \ \text{ and } \ s' \Vdash \varphi, \\ s \Vdash [\Delta]\,\varphi & \text{iff} & \text{for all } s' \in S:\ s \sim_{\Delta} s' \ \text{ implies } \ s' \Vdash \varphi. \end{array}$$

Observe that $\langle\Delta\rangle$ is not a truth functional connective, witness the fact that $\varnothing \Vdash a \leftrightarrow b$ and $\varnothing \Vdash \langle\{b\}\rangle\, a$ but $\varnothing \nVdash \langle\{b\}\rangle\, b$. Still, from Theorem 2.4.1 follows that formulas of the form $\langle\Delta\rangle\,\varphi$ and $[\Delta]\,\varphi$ are equivalent to formulas expressible by means of the Boolean connectives only. The following reflections show how such formulas equivalent to $\langle\Delta\rangle\,\varphi$ and $[\Delta]\,\varphi$ can be obtained from φ.

For $L(A)$ a classical propositional language and $\Delta \subseteq A$, let Σ_{Δ} be the set of functions given by $\{\top,\bot\}^{\Delta}$. We extend each $\sigma \in \Sigma_{\Delta}$ to a function $\hat{\sigma}\colon \Phi(A) \to \Phi(A)$ which is defined inductively as follows:

$$\begin{array}{rcl} \hat{\sigma}\,(a) & =_{df.} & \begin{cases} \sigma\,(a) & \text{if } a \in \Delta, \\ a & \text{otherwise} \end{cases} \\ \hat{\sigma}\,(\bot) & =_{df.} & \bot \\ \hat{\sigma}\,(\top) & =_{df.} & \top \\ \hat{\sigma}\,(\neg\varphi) & =_{df.} & \neg\hat{\sigma}\,(\varphi) \\ \hat{\sigma}\,(\varphi \wedge \psi) & =_{df.} & \hat{\sigma}\,(\varphi) \wedge \hat{\sigma}\,(\psi) \\ \hat{\sigma}\,(\varphi \vee \psi) & =_{df.} & \hat{\sigma}\,(\varphi) \vee \hat{\sigma}\,(\psi)\,. \end{array}$$

Each $\hat{\sigma} \in \Sigma_{\Delta}$ replaces all occurrences of propositional variables Δ in a formula by either $\top$ or $\bot$. In the remainder we will confuse $\hat{\sigma}$ and σ. We will also write $\varphi\,(a_0,\ldots,a_n/\xi_0,\ldots,\xi_n)$ for $\hat{\sigma}\,(\varphi)$, if $\sigma \in \Sigma_{\{a_0,\ldots,a_n\}}$ and $\sigma\,(a_i) = \xi_i$, for each $0 \leqslant i \leqslant n$. We have the following pair of lemmas.

Lemma 2.4.4 *Let s be a valuation for a propositional language $L(A)$ and let Δ be a subset of A. Then, for each $\sigma \in \Sigma_{\overline{\Delta}}$ there is a valuation s' such that $s \sim_{\Delta} s'$ and $s \Vdash \sigma\,(\varphi)$ if and only if $s' \Vdash \varphi$, for all formulas φ.*

Proof: Consider an arbitrary $\sigma \in \Sigma_{\overline{\Delta}}$ and define s^*, such that for all $a \in A$:

$$s^*\,(a) \ =_{df.} \ \begin{cases} 1 & \text{if } a \notin \Delta \text{ and } \sigma\,(a) = \top, \\ 0 & \text{if } a \notin \Delta \text{ and } \sigma\,(a) = \bot, \\ s\,(a) & \text{otherwise.} \end{cases}$$

It can readily be established that $s \sim_\Delta s^*$. A straightforward inductive argument on an arbitrary formula φ then shows that $s \Vdash \sigma(\varphi)$ if and only if $s^* \Vdash \varphi$.

For $\varphi = a$, either $a \in \overline{\Delta}$ or $a \notin \overline{\Delta}$. If the latter, merely observe that $\sigma(a) = a$ and $s(a) = s^*(a)$. If the former, first assume that $\sigma(a) = \bot$. Then both $s \nVdash \sigma(a)$ and also $s^*(a)$, because $s^*(a) = 0$. Finally, assume $\sigma(a) = \top$. Then, $s \nVdash \sigma(a)$ and $s^* \Vdash a$. The latter because in this case $s^*(a) = 1$.

The case for $\varphi = \bot$ is trivial and that for $\varphi = \psi \vee \chi$ is immediate by the induction hypothesis. $\dashv$

Lemma 2.4.5 *Let s be a valuation for a propositional language $L(A)$ and let Δ be a subset of A. Then, for each valuation s' such that $s \sim_\Delta s'$, there is a $\sigma \in \Sigma_{\overline{\Delta}}$ such that $s \Vdash \sigma(\varphi)$ if and only if $s' \Vdash \varphi$, for all formulas φ.*

Proof: Consider an arbitrary valuation s' such that $s \sim_\Delta s'$. Define $\sigma^* \in \Sigma_{\overline{\Delta}}$ such that for all $a \in \overline{\Delta}$:

$$\sigma^*(a) \quad =_{df.} \quad \begin{cases} \top & \text{if } a \in s', \\ \bot & \text{otherwise.} \end{cases}$$

A straightforward inductive argument then shows that σ^* complies with the requirements as stated in the lemma. So consider an arbitrary formula φ. For the basis assume $\varphi = a$ for some $a \in A$. If $a \in \Delta$, then $\sigma^*(a) = a$ and:

$$s' \Vdash a \quad \text{iff} \quad a \in s' \quad \text{iff}_{s \sim_\Delta s'} \quad a \in s \quad \text{iff}_{\sigma^*(a) = a} \quad s \Vdash \sigma^*(a).$$

If, however, $a \notin \Delta$, first assume $s' \Vdash a$. Then, $\sigma^*(a) = \top$ and immediately $s \Vdash \top$. Now assume $s' \nVdash a$, then $a \notin s'$ and $\sigma^*(a) = \bot$. Observe that also $s \nVdash \bot$.

The inductive cases are either trivial or immediate by the induction hypothesis. $\dashv$

These lemmas prepare the ground for the following proposition.

Proposition 2.4.6 *Let φ be a formula in a classical propositional language $L(A)$ and Δ a subset of A. Then both:*

$$\underline{apr}_\Delta([\![\varphi]\!]) \;=\; \bigcap_{\sigma \in \Sigma_{\overline{\Delta}}} [\![\sigma(\varphi)]\!] \quad \textit{and} \quad \overline{apr}_\Delta([\![\varphi]\!]) \;=\; \bigcup_{\sigma \in \Sigma_{\overline{\Delta}}} [\![\sigma(\varphi)]\!].$$

Proof: As to the first claim, consider an arbitrary $s \in 2^A$. For the $\subseteq$-direction, assume that $s \notin \bigcap_{\sigma \in \Sigma_{\overline{\Delta}}} [\![\sigma(\varphi)]\!]$. Then, there is some $\sigma \in \Sigma_{\overline{\Delta}}$ such that $s \notin [\![\sigma(\varphi)]\!]$. By Lemma 2.4.4, there is some valuation s' such that $s \sim_\Delta s'$ and $s \Vdash \sigma(\varphi)$ if and only if $s' \Vdash \varphi$. Consider this s', then $s' \notin [\![\varphi]\!]$ and hence $s \notin \underline{apr}_\Delta([\![\varphi]\!])$.

For the $\supseteq$-direction, assume that $s \notin \underline{apr}_\Delta([\![\varphi]\!])$. Then, there is some valuation s' such that $s \sim_\Delta s'$ and $s' \notin [\![\varphi]\!]$. By Lemma 2.4.5, there is a $\sigma \in \Sigma_{\overline{\Delta}}$ such that $s \Vdash \sigma(\varphi)$ if and only if $s' \Vdash \varphi$. Hence, $s \notin [\![\sigma(\varphi)]\!]$ and *a fortiori*, $s \notin \bigcap_{\sigma \in \Sigma_{\overline{\Delta}}} [\![\sigma(\varphi)]\!]$.

For the second claim consider the following equalities:

$$\overline{apr}_{\Delta}(\llbracket\varphi\rrbracket) = \overline{\underline{apr}_{\Delta}(\llbracket\neg\varphi\rrbracket)} = \overline{\bigcap_{\sigma\in\Sigma_{\overline{\Delta}}} \llbracket\sigma(\neg\varphi)\rrbracket} = \bigcup_{\sigma\in\Sigma_{\overline{\Delta}}} \overline{\llbracket\neg\sigma(\varphi)\rrbracket} = \bigcup_{\sigma\in\Sigma_{\overline{\Delta}}} \llbracket\sigma(\varphi)\rrbracket.$$

This concludes the proof. ⊣

This result has the following corollary.

Corollary 2.4.7 *Let φ be a formula in a propositional language $L(A)$, let $\Delta \subseteq A$ and a a propositional variable not in Δ. Then:*

$$\begin{aligned}\underline{apr}_{\Delta}\left(\llbracket\varphi(a/\bot)\rrbracket \cap \llbracket\varphi(a/\top)\rrbracket\right) &= \underline{apr}_{\Delta}(\llbracket\varphi\rrbracket) \quad \textit{and}\\ \overline{apr}_{\Delta}\left(\llbracket\varphi(a/\bot)\rrbracket \cup \llbracket\varphi(a/\top)\rrbracket\right) &= \overline{apr}_{\Delta}(\llbracket\varphi\rrbracket).\end{aligned}$$

Proof: First consider the following equalities:

$$\underline{apr}_{\Delta}(\llbracket\varphi\rrbracket) \;=_{a\notin\Delta}\; \underline{apr}_{\Delta\cap\overline{\{a\}}}(\llbracket\varphi\rrbracket) \;=_{\text{Fact 2.2.8}}\; \underline{apr}_{\Delta}\left(\underline{apr}_{\overline{\{a\}}}(\llbracket\varphi\rrbracket)\right).$$

Then observe that $\underline{apr}_{\overline{\{a\}}}(\llbracket\varphi\rrbracket) = \llbracket\varphi(a/\bot)\rrbracket \cap \llbracket\varphi(a/\top)\rrbracket$, by Proposition 2.4.6. The proof of the second claim is analogous. ⊣

Observe that for each formula φ and each $\Delta \subseteq A$, the set $\{\sigma(\varphi) : \sigma \in \Sigma_{\Delta}\}$ is finite, because φ only contains a finite number of propositional variables. Hence, we may assume $\bigwedge_{\sigma\in\Sigma_{\Delta}} \sigma(\varphi)$ and $\bigvee_{\sigma\in\Sigma_{\Delta}} \sigma(\varphi)$ to be, respectively, a well-formed finite conjunction and a well-formed finite disjunction, even if Δ is infinite. On basis of Proposition 2.4.6 we may therefore assume $[\Delta]\,\varphi$ and $\langle\Delta\rangle\,\varphi$ to abbreviate the formulas $\bigwedge_{\sigma\in\Sigma_{\overline{\Delta}}} \sigma(\varphi)$ and $\bigvee_{\sigma\in\Sigma_{\overline{\Delta}}} \sigma(\varphi)$, respectively.

Part I

Nash Equilibria in Modal Logic

Chapter 3

A Modal Characterization of Nash Equilibrium

3.1 Introduction

With the advance of distributed and multi-agent systems there has been an increased interest in the relation between logic and game theory within the field of Artificial Intelligence (*cf.*, *e.g.*, van Benthem (2001b), Boutilier, Shoham, and Wellman (1997) and Pauly (2001)). In multi-agent environments, various decision making agents with various degrees of autonomy interact. The individual agents making up a multi-agent system may be designed for widely divergent and even conflicting tasks. Still, which actions are most conducive to an agent's ends in such situations, may well depend on the decisions of the other agents. The specification and verification of multi-agent systems calls for mathematically precise concepts that facilitate reasoning about such interactive strategic situations. Game theory is relevant to the field of *Artificial Intelligence* in that it provides an apposite conceptual framework in this respect.

The theory of games originated in the middle of the 20th Century with the recognition that, to that date, no theory in classical mathematics had dealt with social situations in which each individual tries to maximize a function according to an idiosyncratic principle without having control over all of the variables on which this function depends (*cf.*, von Neumann and Morgenstern (1944), p.11). Thus, game theory was developed as the mathematical study of game-like situations in which the eventual outcome depends on the individual choices of various agents, each of which has different preferences over the possible outcomes. In any such situation the traditional notions of optimality were thought no longer to suffice for a proper analysis and game-theoretical solution concepts were developed to take over their role. In this respect, the celebrated *Nash equilibrium* and its *subgame perfect* variety are archetypical in non-cooperative settings. Recall that, informally, a collective course of action, or a strategy profile, is said to be a *Nash equilibrium* if none of the participants has an incentive to deviate unilaterally from that course of action (*cf.*, page 28). Whether an agent has such an incentive depends on his individual preferences.

One of the guiding ideas of game theory is that situations of social interaction can fruitfully be compared with and analyzed as games by distinguishing players, their strategies and their interests. Games have proved to be an especially rewarding metaphor for social environments in which interacting agents are conceived of as players with individual preferences and powers of manipulation. This leaves the question how far the game metaphor goes and how far it should be carried. In order to arrive at a general theory of social interaction, specific and idiosyncratic aspects of games should be abstracted from, in favor of other, more generic features, which should duly be emphasized. Where the dividing line between the general and the specific should be drawn is not an objective matter and may very well depend on one's purposes. Still, it should always be borne in mind that:

> A model structure that is too simple may force us to ignore vital aspects of the real games we want to study. A model structure that is too complicated may hinder our analysis by obscuring fundamental issues. (Myerson (1991), p.37)

The order in which the players perform their actions in strategic situations has reasonably been argued to be a vital, rather than an obscuring, aspect in this sense. The models of strategic situations provided *by games in their extensive form* — or just *extensive games* — are especially designed to account for this type of sequential structure.

The extensive form of a game makes explicit the order in which the players are to choose among a number of alternative courses of action and how the alternatives available to a player are dependent on previous decisions. This makes that the extensive form of a game can be represented as a labelled tree, each subtree of which can be considered as an extensive game in its own right, *i.e.*, as a subgame of the game as a whole. An important solution concept that comes along with extensive games is that of *subgame perfect Nash equilibrium*. This ramification of the original concept selects among the Nash equilibria of the a game those strategy profiles that also qualify as a Nash equilibrium in each of its subgames.

In this chapter, we will give a logical analysis of extensive games and their (subgame perfect) equilibria. With an extensive game being introduced as a specific kind of relational structure, we employ multi-modal languages to this end. First we come to consider a multi-modal language which set of labels is assumed to possess relatively little structure. The last section of this chapter shows how this framework can be refined by deploying the language of Propositional Dynamic Logic (PDL).

Our approach is congenial to Bonanno's in Bonanno (1998), who used Computational Tree Logic (CTL) extended with a prediction relation to formalize the concept of backward induction, which is closely related to subgame perfect Nash equilibrium. Also the work of Baltag (Baltag (1999)) should be mentioned in this context.

Extensive games define a proper subclass of Kripke-frames for the special kind of multi-modal language we consider. Each strategy profile of an extensive game then corresponds to a subrelation in the frame and as such can be taken as the accessibility relation of a modal operator. Some strategy profiles qualify as a (subgame perfect)

	left	*right*
top	3 6	4 7
bottom	0 0	1 10

Figure 3.1.

Nash equilibrium and others do not. A strategy profile being a (subgame perfect) Nash equilibrium reflects in certain specific structural properties of the subrelation it defines on the corresponding frame. The result we are after is to characterize these structural properties by means of a multi-modal formula schema. So, for E an extensive game, $\mathfrak{F}_E$ its corresponding frame and s one if its strategy profiles, our quest is for a formula schema $\vartheta(s)$ such that:

$$\mathfrak{F}_E \Vdash \vartheta(s) \quad \text{iff} \quad s \text{ is a subgame perfect Nash equilibrium in } E.$$

Any such result would show that subgame perfect Nash equilibrium is a definable property of frames in appropriate multi-modal languages. Here it be emphasized that the frames in question belong to a special class of frames corresponding to extensive games.

In the next chapter, we propose a sound and complete axiomatization of a multi-modal logic which semantics is restricted to models based on this class of frames extensive games define. Remarkably, we will find that the axioms are nothing much out of the ordinary and can also be bestowed rather intuitive readings. The very austerity of the whole analysis we take as something speaking in its favor, as it shows how little is required of a modal language to be able to characterize (subgame perfect) Nash equilibrium.

3.2 Extensive Games with Perfect Information

Generally speaking, putting the egg in the pan first and then the butter does not work quite as well as putting in the butter first and then the egg. The order in which the actions are performed does matter in some cases. In strategic environments this is no different. What is more, the order in which agents can make their choices and moves often makes a *strategic difference* and as such is something the game-theorist had better not ignore entirely.

For an example, consider the strategic situation depicted in Figure 3.1. Here two players, *Row* and *Col*, choose between rows (*top* or *bottom*) and between columns (*left* or *right*), respectively. The matrix merely summarizes the payoffs to the players for the different possible choices of action. The figure bottom left in each quarter indicates the payoff to *Row*, the figure top right the one awarded to *Col*. The matrix is thought of as specifying no temporal structure whatsoever; as it is the players are not even assumed to move simultaneously.

In this situation *Row* may be tempted to play *bottom*, the idea being that this would leave *Col* the relatively unfavorable choice between an outcome of 1 and an outcome of zero. Anticipating that *Col* chooses the former, *Row* may expect a payoff of 10. In order to deter *Row* from taking this course of action, *Col* may threaten to play *left* if *Row* were to play *bottom*, resulting in the worst outcome for both players. This would force *Row* to play *top*, which guarantees a better outcome for *Col*. However, if the game structure were such that *Row* is (able) to move first, *Col*'s threats would be rendered void, provided he is not otherwise committed to choose the left column. If, in defiance of *Col*'s threats, *Row* were to play bottom anyway, *Col* would be presented with a *fait accompli* and *Col* had better make the best of a bad job and opt for the right colomn after all. Similarly, if *Col* were to move first a threat on his part does not make sense either. In that case, however, he can secure a more favorable outcome by choosing the left column. That would leave *Row* the easy choice between a payoff of 6 or one of zero. *Row* making the obvious decision would guarantee *Col* a payoff of 3, instead of the probable and miserly 1 he would have gotten had his initial choice been *right* with *Row* quite likely seizing the opportunity and playing *bottom*. This time, however, *Row* can try to achieve a better outcome for both by promising — and committing herself to fulfill this promise — to play *top* if *Col* decides on the right column.

Figure 3.1 as such leaves unspecified the sequential structure of how the game is played. Above we gave two possible interpretations, an obvious third would be to conceive Figure 3.1 as a fully-fledged game in strategic form, *e.g.*, by assuming the player to make their choices simultaneously or in ignorance of one another's. Be that as it may, the point of these reflections is that the order in which the players make moves does make a strategic difference. For one thing, the feasibility of making a threat or a promise may depend on it and in our example even the outcome of the game may as well. As such, the sequential structure of game has sensibly been made subject of game-theoretic study.

The sequential structure of a game is made mathematically precise in its *extensive form*, which represents the game as a labelled multi-player decision tree. Play commences at the root and each edge indicates a possible course of action for a player. At each node a player is to strike a decision how to act. The two temporal interpretations of Figure 3.1 can thus be represented as in Figure 3.2, below. The vectors at the leaves indicate the payoffs to the players, in both cases the first entry being the payoff *Row*, the second to *Col*. Observe that in both cases, there are four strategies for the second player that is to move. In the left picture, each of *Col*'s strategies has to specify whether to play *right* or *left* both if the first player plays *top* and if she plays *bottom*.

Similar concerns may drive the game-theorist to consider models of game-like situations with even more structure. Strategic reasoning may thus be argued to depend on epistemic features such as the players' knowledge of the situation they are in or the beliefs they entertain about it or about one another's beliefs, preferences and ra-

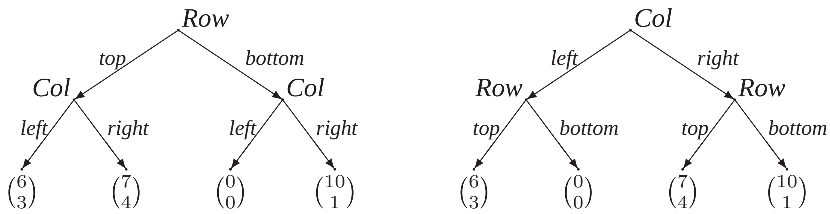


Figure 3.2.

tionality. Likewise, one may wish to consider players that randomize over their (pure) strategies. In an effort at keeping our logical analysis as perspicuous as possible, however, we will abstract from these issues and confine our attention to extensive games in pure strategies with perfect information, *i.e.*, the players are assumed to play a (pure) strategy with probability one or zero and they are assumed to be fully informed about the game's structure and the other players' preferences and powers. Moreover, we will assume that only one player can move at a time and that the games will eventually come to an end after a finite number of moves. With respect to the preferences of the players, we take into account the ordinal structure they determine over the possible outcomes only, as this suffices for our purposes. Concerns as to the intensity of preference as expressible by a specific rational or real number, do not enter the picture. Disregarding uncertainty on the part of the players as well as mixed (or randomized) strategies, our analyses are of a strictly qualitative nature. The following definition mathematically precise the notion of a *game in extensive form*, or just an *extensive game*.

Definition 3.2.1 *(Games in extensive form of perfect information)* A *game in extensive form E* is a tuple $\left(V, R, N, P, \{\rho_i\}_{i \in N}\right)$, where V is a set of vertices (or nodes) and R a relation on V such that (V, R) is a, possibly infinite, directed and irreflexive tree with a finite horizon, *i.e.*, (V, R) contains no infinite branches. The root node of (V, R) is usually denoted by v_ϵ. Furthermore, N is a non-empty but *finite* set of players. The function P assigns to each internal node in V the player in N that has to move at v. Finally, for each i in N, ρ_i is a total pre-order (a reflexive, transitive and connected relation) over the vertices in V, specifying i's preferences. Intuitively, $(v, v') \in \rho_i$ signifies that i values v' at least as high as v. A player i is called *indifferent* if $\rho_i = V \times V$ and *interested*, otherwise. Let Z denote the set of leaves of (V, R) and, for each player i, let V_i betoken the subset of vertices in which i is to move, *i.e.*, the set $\{v \in V : \; P(v) = i\}$.

This definition differs from more conventional ones in that the players' preferences are defined over all vertices rather than over the leaves only. Although for the relevant game-theoretical concepts the preferences over the leaf nodes suffice, we found that defining preferences over all vertices is more convenient for our logical analyses. Note further that the players' preferences over the internal nodes are independent of their preferences over the leaf nodes. In particular, they are not assumed to coincide with the preferences over the internal nodes that backwards induction would give rise to.

An extensive game is a labelled tree, the vertices of which represent the possible game positions and the edges (v, v') are possible actions for the player assigned to v. After a player has decided to play along a certain edge and acted accordingly, the game reaches a new game state. The position then reached is either a leaf, in which case the game terminates, or an internal node. In the latter case, the node reached can also be taken as the the root node of an extensive game, with the playing of which the game proceeds. This idea gives rise to the notion of a *subgame*. Let E be an extensive game given by the tuple $\left(V, R, N, P, \{\rho_i\}_{i\in N}\right)$. For each subtree (V', R') of (V, R) generated by some vertex v, another extensive game is obtained by appropriately restricting the assignment function P and each ρ_i to the vertices in V'. For each vertex v in V we define the subgame E_v as the tuple $\left(V', R', N, P', \{\rho_i'\}_{i\in N}\right)$, where $V' = \{v' \in V : \ (v, v') \in R^*\}$, $R' = \left\{(v', v'') \in V' \times V' : \ (v', v'') \in R\right\}$, $P' = P{\restriction} V'$ and for each $i \in N$, $\rho_i' = \rho_i \cap (V' \times V')$. Here, R^* denotes the reflexive transitive closure of R.

A *(pure) strategy* for a player in an extensive form is a complete plan for that player to play the respective game. As such a strategy has to account for a player's choices at all stages of the game in which that player is in control. A strategy even has to prescribe a player's actions in stages of the game it itself precludes from being reached. Intuitively, a *strategy profile* is then a combination of strategies, for each player one. The set of strategy profiles in an extensive game E is denoted by S_E, omitting the subscript where no ambiguity can arise. For our concerns the notion of a strategy profile is more fundamental than that of a strategy. We define a strategy profile s of an extensive game formally as a function mapping each *internal* vertex onto a vertex that succeeds it, *i.e.*, for each v in V, $(v, s(v)) \in R$. For any pair of strategy profiles s and s' and for each subset of players I we have $s^I_{s'}$ denote the strategy profile that is like s except on the vertices assigned to one of the players in I where it takes values from s'. *I.e.*, for all internal vertices v we have:

$$s^I_{s'}(v) \quad =_{df.} \quad \begin{cases} s'(v) & \text{if } P(v) \in I, \\ s(v) & \text{otherwise.} \end{cases}$$

We also have $s^i_{s'}$ abbreviate $s^{\{i\}}_{s'}$. A *strategy for a player i* is then the restriction of a strategy profile to the vertices in which i is in control. Accordingly the set of strategies for a player i is defined by $\{s{\restriction} V_i : \ s \in S\}$.

From each vertex v onwards a strategy profile s generates a path through the game-tree until a leaf node is reached. This path is given by the sequence $v_0, \ldots, v_n$ such that $v_0 = v$, v_n is a leaf and $v_{i+1} = s(v_i)$, for all $0 \leqslant i < n$, *i.e.*, by the sequence:

$$v,\ s(v),\ \ldots,\ s^n(v),$$

where $s^n(v)$ denotes the n-fold application of s to v. *E.g.*, $s^3(v) = s(s(s(v)))$. In this manner each strategy profile determines for each vertex a unique leaf node as outcome. With each strategy profile s we accordingly associate an *outcome function*, $\hat{s}$, which maps each vertex on the leaf it has as an outcome if the strategy profile s is followed iteratively. Formally we define for each strategy profile s the outcome function $\hat{s}$ inductively such that for each vertex v in V:

$$\hat{s}(v) \quad =_{df.} \quad \begin{cases} v & \text{if } v \text{ is a leaf,} \\ \hat{s}(s(v)) & \text{otherwise.} \end{cases}$$

On this basis, each extensive game E can be correlated with a strategic game G_E. The unique outcomes the strategy profiles determine for the root node can be represented in a matrix. Thus the various strategy profiles of extensive games can be compared with respect to the solution concepts available for strategic games, in particular that of Nash equilibrium. Exploiting the outcome function, the preference relation ρ_i for each player i can straightforwardly be raised as to apply to strategy profiles. Let E be the extensive game $\left(V, R, N, P, \{\rho_i\}_{i \in N}\right)$. The preferences of player i over the strategy profiles are made dependent on i's preferences over the outcomes these strategy profiles induce in the root node of the game-tree. Define the preference relation $\hat{\rho}_i$ over the strategy profiles of E for each player i such that for all strategy profiles s and s' for E:

$$(s, s') \in \hat{\rho}_i \quad \text{iff} \quad \left(\hat{s}(v_\epsilon), \hat{s}'(v_\epsilon)\right) \in \rho_i.$$

Let the tuple $\left(N, \{S_i\}_{i \in N}, \{\hat{\rho}_i\}_{i \in N}\right)$ define the strategic game G_E. A strategy profile s is then said to be a *Nash equilibrium* in an extensive game E if and only if s is a Nash equilibrium in the strategic game G_E. Intuitively, a strategy profile s is a Nash-equilibrium if none of the players benefit from unilaterally deviating from s. For s a strategy profile in an extensive game E given by $\left(V, R, N, P, \{\rho_i\}_{i \in N}\right)$ we obtain:

$$s \text{ is a } \textit{Nash equilibrium} \quad \text{iff} \quad \text{for all } i \in N, \text{and all } s' \in S_E \colon \left(\hat{s}^i_{s'}(v_\epsilon), \hat{s}(v_\epsilon)\right) \in \rho_i.$$

As an individual pendant of Nash-equilibrium we have the concept of a *best response for a player i* defined for a strategy profile s and a player i as:

$$s \text{ is a } \textit{best response for } i \quad \text{iff} \quad \text{for all } s' \in S_E \colon \left(\hat{s}^i_{s'}(v_\epsilon), \hat{s}(v_\epsilon)\right) \in \rho_i.$$

Obviously, a strategy profile is a Nash-equilibrium if and only if it is a best response for all players.

The notion of a Nash equilibrium entirely focusses on the outcomes the various strategy profiles determine from the root. Different strategy profiles may very well give rise to an identical path from root to leaf node — and as such determine the same outcome — and still differ widely on vertices off this path. The path a Nash equilibrium determines through the game tree is such that unilateral deviation from it will not benefit the defector. Yet, it has been argued that if the sequential structure of a game is taken into account, the notion of Nash equilibrium fails to make some important distinctions. Although one could accept — were it only for the sake of argument — the refusal to defect unilaterally from the equilibrium path as the very hallmark of game-theoretical level-headedness, *off* the equilibrium path a Nash equilibrium may strike

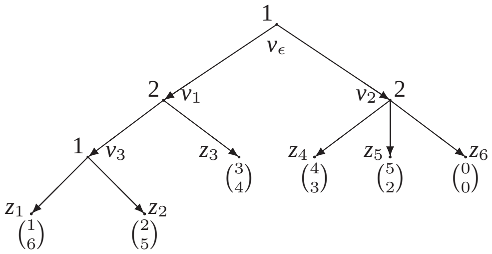


Figure 3.3. The extensive game of Example 3.2.2.

one as somewhat unsatisfactory. Consider once more the extensive game in Figure 3.2 in which *Row* is to move first. We have already argued that *Row* need not refrain from playing *bottom* even if *Col* were to threaten to choose the left column in that case. This is vindicated by all strategy profiles in which *Row* chooses *bottom* and *Col* subsequently playing *right* being Nash equilibria in this game. Strategies, however, determine choices for the players at all nodes where they are to play. The node that would have been reached had *Row* chosen *top* is no exception. At that node it would be slightly incomprehensible if *Col* were to choose the left column. Still the strategy profile in which *Row* plays *bottom* and *Col* plays *right* if *Row* were to play *bottom* and *left* otherwise, is nevertheless a Nash equilibrium.

As a refinement of Nash equilibrium that does do justice to the sequential structure of an extensive game, Selten (Selten (1965)) proposed the solution concept of a *subgame perfect Nash equilibrium*. Roughly speaking, a strategy profile is a subgame perfect Nash equilibrium in an extensive game E if it is a Nash equilibrium in all subgames of E. In the example above, any strategy profile that would prescribe *Col* to play *left* when *Row* has chosen the top row, would not qualify as a subgame perfect Nash equilibrium. Formally, for E the extensive game $\left(V, R, N, P, \{\rho_i\}_{i\in N}\right)$, define for each strategy profile s:

$$\begin{gathered} s \text{ is a } \textit{subgame perfect Nash equilibrium} \\ \text{iff} \\ \text{for all } v \in V, i \in N, \text{and } s' \in S_E : \left(\hat{s}^i_{s'}(v), \hat{s}(v)\right) \in \rho_i. \end{gathered}$$

As individual counterpart of this concept we also introduce *subgame perfect best responses for a player i* defined in a similar manner as a strategy profile that is a best response for i in all subgames. A formal definition of this concept is obtained by omitting the universal quantification over the players in the *definiens* of a subgame perfect Nash-equilibrium.

The following example illustrates the concepts that have been introduced so far.

Example 3.2.2 Figure 3.3 gives a graphical representation of a two-player game in

	ll	*lm*	*lr*	*rl*	*rm*	*rr*
LL	1, 6	1, 6	1, 6	3, 4	3, 4	3, 4
LR	2, 5	2, 5	**2**, **5**	3, 4	3, 4	3, 4
RL	**4**, **3**	5, 2	0, 0	**4**, **3**	5, 2	0, 0
RR	**4**, **3**	5, 2	0, 0	**4**, **3**	5, 2	0, 0

Figure 3.4. The strategic game associated with the extensive game of Example 3.2.2, with Player 1 choosing rows and Player 2 choosing columns.

extensive form. The preferences of the players over the leaves are represented by the vectors appended to the leaf nodes. The first entry indicates the preferences of Player 1 and the second those of Player 2. The higher the value, the more the outcome is preferred by the player. *E.g.*, the pair (z_5, z_3) is in the preference relation ρ_2, because 2 is smaller than 4. Player 1 has four strategies at her disposal and Player 2 six. Accordingly, there are twenty-four strategy profiles in all, each of which we indicate by a four letter subscript corresponding to the direction the players move at the vertices v_ϵ, v_3, v_1 and v_2, respectively. The choices of player 1 are denoted by capitals, those of player 2 by lower case letters. *E.g.*, the strategy profile s_{RLlr} is the functional relation given by $\{(v_\epsilon, v_2), (v_3, z_1), (v_1, v_3), (v_2, z_6)\}$. Starting from the root v_ϵ, it gives rise to the sequence v_ϵ, v_2, z_4 and, accordingly, we have:

$$\hat{s}_{RLlr}(v_\epsilon) = \hat{s}_{RLlr}(s_{RLlr}(v_\epsilon)) = \hat{s}_{RLlr}(v_2) = \hat{s}_{RLlr}(s_{RLlr}(v_2)) = \hat{s}_{RLlr}(z_6) = z_6.$$

This strategy profile, however, fails as a Nash equilibrium. Player 2 could deviate from s_{RLlr} at v_2 and play l there instead. This would make that s_{RLll} is played, yielding z_4 as outcome and guaranteeing him a payoff of 3 instead of zero. The corresponding strategic game is given in Figure 3.4. The Nash equilibria are given by the following relations on the vertices:

$$\begin{aligned}
s_{RLll} &= \{(v_\epsilon, v_2), (v_3, z_1), (v_1, v_3), (v_2, z_4)\}, \\
s_{RLrl} &= \{(v_\epsilon, v_2), (v_3, z_1), (v_1, z_3), (v_2, z_4)\}, \\
s_{RRll} &= \{(v_\epsilon, v_2), (v_3, z_2), (v_1, v_3), (v_2, z_4)\}, \\
s_{RRrl} &= \{(v_\epsilon, v_2), (v_3, z_2), (v_1, z_3), (v_2, z_4)\}, \\
s_{LRlr} &= \{(v_\epsilon, v_1), (v_3, z_2), (v_1, v_3), (v_2, z_6)\},
\end{aligned}$$

Of these only s_{RRll} is a subgame perfect Nash equilibrium as well. The strategy profile s_{LRlr}, *e.g.*, is excluded as a subgame perfect equilibrium since it is not a Nash-equilibrium in the subgame that has v_2 as root.

Obviously, every subgame perfect Nash equilibrium is also a Nash equilibrium. An important result known as Kuhn's theorem (*cf.*, Kuhn (1953)), establishes that every finite extensive game of perfect information has a subgame perfect Nash equilibrium in pure strategies. Closely related is the method of *backwards induction*, which is essentially an algorithm providing subgame perfect Nash equilibria which goes back to Zermelo (1913).

A strategy profile corresponds with a collection of paths through the tree and each of these paths starts at a different internal node. In particular, a strategy profile determines a path connecting the root with a leaf. Strategies can similarly be construed as subgraphs of (V, R). Another interesting subgraph results if one takes the union of a strategy profile s and the set of edges with the vertices possessed by a (sub-)set of players I as source. Intuitively, the significance of any such graph is that it reflects which outcomes a set of players can force to come about if they operate in coalition and the strategies of the other players are given. For the game of Example 3.2.2, this graph for Player 1 and the strategy profile s_{RLll} curbing Player 2's freedom of action is depicted in Figure 3.5.

To capture this notion formally we define for each strategy profile s and subset of players I a correspondence s_I on the vertices such that for all vertices v in V:

$$s_I(v) \quad =_{df.} \quad \begin{cases} \{w \in V : \ (v,w) \in R\} & \text{if } P(v) \in I, \\ \{s(v)\} & \text{otherwise.} \end{cases}$$

The correspondence s_I is obviously monotone in I, *i.e.*, $I' \subseteq I''$ implies $s_{I'} \subseteq s_{I''}$.

Each relation s_I, in turn, induces a correspondence on the vertices of the game, which value is a subset of the leaves of the tree. We define for each strategy profile s in S and each subset I of players in N, this correspondence $\hat{s}_I$ such that for each vertex v in V:

$$\hat{s}_I(v) \quad =_{df.} \quad \begin{cases} \{v\} & \text{if } v \text{ is a leaf,} \\ \bigcup \{\hat{s}_I(w) : \ w \in s_I(v)\} & \text{otherwise.} \end{cases}$$

We will write $\hat{s}_i$ for $\hat{s}_{\{i\}}$. The value the correspondence $\hat{s}_I$ takes at the root v_ϵ is the set of outcomes the players in I can force to come about by cooperating if the other players adhere to the strategy profile s. In our example, $s_1(v_\epsilon) = \{z_1, z_2, z_4\}$, where s represents s_{RLll}. Obviously, the more players in I the larger this set of forceable outcomes, *i.e.*, the monotonicity of s_I propagates to $\hat{s}_I$:

$$I \subseteq I' \quad \text{implies} \quad \hat{s}_I \subseteq \hat{s}_{I'}.$$

The following fact relates notations and will prove to be particularly convenient. Its proof is an easy check.

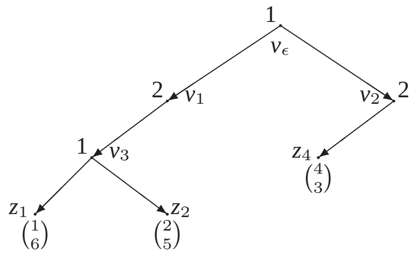


Figure 3.5. The graph of the correspondence $\hat{s}_1$ in the game of Example 3.2.2, where s represents s_{RLll}. Player 1 can force the game to terminate in either z_1, z_2 or z_4 if Player 2 adheres to the strategy profile s_{RLll}.

Fact 3.2.3 *Let s be a strategy profile of some extensive game E with v and v′ vertices therein and I a subset of its players. Then:*

$$v \in \hat{s}_I\,(v') \quad \textit{iff} \quad \textit{for some } s' \in S\colon\;\; \hat{s}^I_{s'}\,(v') = v.$$

Obviously as special case we have that $\hat{s}_\emptyset\,(v) = \{\hat{s}\,(v)\}$. The set of outcome nodes that can be reached by a strategy profile s with no player possibly deviating clearly contains as only the element the vertex that s determines as the unique outcome.

3.3 Describing and Reasoning about Extensive Games

Extensive games are based on trees and the players' preferences are defined as relations over the vertices in the previous section. Exploiting this relational structure, we propose a multi-modal language to describe extensive games and reason about them. In particular, we will argue that such a language can express whether a strategy profile of an extensive game is a (subgame perfect) Nash equilibrium.

Syntax and Semantics

Our formal researches are conducted within propositional multi-modal logic. A propositional multi-modal language $L(A, B)$ contains a non-empty but countable set of propositional variables A along with a countable set of labels B for monadic modalities. The formulas of $L(A, B)$ are thus given by the following BNF-grammar, with $a \in A$ and $\beta \in B$:

$$\varphi \quad ::= \quad a \;\big|\; \neg\varphi \;\big|\; \varphi_0 \wedge \varphi_1 \;\big|\; [\beta]\varphi$$

We assume the set of labels B to be the union of two disjoint sets and their Cartesian product, *i.e.*, $B = B_0 \cup B_1 \cup (B_0 \times B_1)$ with $B_0 \cap B_1 = \emptyset$. Moreover, B_0 will be assumed to be non-empty and finite. Multi-modal languages $L(A,B)$ with B structured thus we will refer to as *multi-modal matrix languages*.

Extensive games are taken as the basis of the frames any such multi-modal language describes. Truth-value assignments to the propositional variables at each vertex takes care of the interpretation of the propositional variables and the Boolean connectives are given their conventional interpretation. The labels in B_0 go proxy for the players of a game. For each $\beta \in B_0$, the accessibility relation R_β runs along the preference relation of one of the players of the game. This gives rise to the intuitive reading of $[\beta]\varphi$ as "φ *holds in all states at least as preferable to i as the present one,*" where i is the player associated with the label β. For convenience, the labels in B_0 are also called *player labels*. In contrast, the labels in B_1 stand for strategy profiles of the game $L(A,B)$ aims to describe and are therefore referred to as *strategy labels*. For each label $\beta \in B_1$ the accessibility relation R_β is defined as (the graph of) the function $\hat{s}$, where s is the strategy profile associated with β. As such R_β relates vertices to leaves only, being reflexive at the latter. Intuitively, $[\beta]\varphi$ then reads " *if, starting in the state of evaluation, all players choose their strategies as prescribed in s, the game ends in a situation in which* φ *holds*". Finally, let i be the player associated with the label β in B_0 and s the strategy profile associated with the label β' in B_1. Then, the accessibility relation $R_{(\beta,\beta')}$ connects each vertex v to the leaf nodes in $\hat{s}_i(v)$. Recall that $\hat{s}_i(v)$ collects the terminal nodes player i can force to come about, provided that the other players adhere to the strategy profile s. Then $[(\beta,\beta')]\varphi$ obtains the informal interpretation of "φ *holds in all outcome states that can be reached if at most player i deviates from s.*"

The frames and models for the multi-modal languages are also of a special kind. Rather than taking into account all relational structures, the formal semantics is defined on frames that are structurally closely related to extensive games. The notion of a *game-model* for $L(A,B)$ on such a *game-frame* is then introduced much in the usual fashion.

Definition 3.3.1 *(Game-frames and Game-models)* A *frame* for a multi-modal matrix language $L(A,B)$ is a tuple $\left(V, \{R_\beta\}_{\beta\in B}\right)$, where V is a set of vertices and $R_\beta \subseteq V \times V$, for each $\beta \in B$. A *label map* for $L(A,B)$ a multi-modal matrix language on an extensive game E is a function f mapping each label in B_0 onto a player in N and each label in B_1 onto a strategy profile of E. In the sequel we will usually tacitly assume such a label map f and denote the labels in B by their values under f, *i.e.*, if $f(\beta_0) = i$ and $f(\beta_1) = s$, we will write i, $\hat{s}_\emptyset$ and $\hat{s}_i$ for, respectively, β_0 in B_0, β_1 in B_1 and (β_0,β_1) in $B_0 \times B_1$. We say that the labels β_0 and β_1 *represent* the player i and the strategy profile s, respectively. A *game-frame* for $L(A,B)$ on an extensive game E is a tuple $\left(V, \{R_\beta\}_{\beta\in B}, f\right)$, where $\left(V, \{R_\beta\}_{\beta\in B}\right)$ is an frame frame for $L(A,B)$ and f a label map on E such that:

$$\begin{aligned} vR_i v' &\quad \text{iff} \quad (v,v') \in \rho_i \\ vR_{\hat{s}_\emptyset} v' &\quad \text{iff} \quad v' \in \hat{s}_\emptyset(v) \\ vR_{\hat{s}_i} v' &\quad \text{iff} \quad v' \in \hat{s}_i(v)\,. \end{aligned}$$

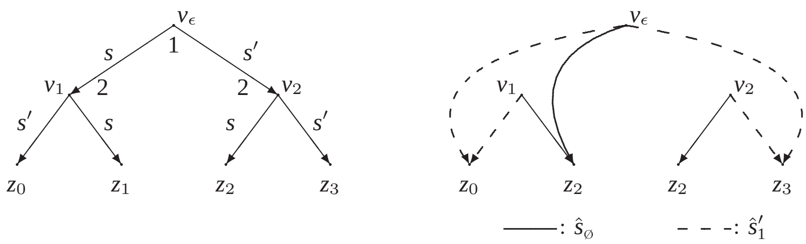


Figure 3.6. Transformation of an extensive game (left) to a game-frame (right) with respect to two strategy profiles s and s' and their corresponding accessibility relations $R_{\hat{s}_\emptyset}$ and $R_{\hat{s}'_1}$. In the righthand figure the reflexive arrows at the leaves are omitted.

We denote a game-frame on E by $\mathfrak{F}_E^{\mathrm{M}}$, tacitly assuming a label map f and usually omitting the superscript M. A frame $\mathfrak{F}$ is a game-frame for $L(A,B)$ *simpliciter* if there is some label map rendering it a game-frame on some extensive game E. A *game-model* $\mathfrak{M}$ for $L(A,B)$ is a pair $(\mathfrak{F},\mathfrak{V})$, where $\mathfrak{F}$ is a game-frame $\big(V,\{R_\beta\}_{\beta\in B}\big)$ for $L(A,B)$ and $\mathfrak{V}$ a function assigning to each vertex in V a subset of propositional variables in A, *i.e.*, $\mathfrak{V}\colon V\to 2^A$. Figure 3.6 illustrates the construction of a game-frame from an extensive game.

On this basis, multi-modal matrix languages are furnished with a standard modal semantics:

$$\begin{array}{lll}
\mathfrak{M},v\Vdash a & \text{iff} & a\in\mathfrak{V}(v)\\
\mathfrak{M},v\Vdash\neg\varphi & \text{iff} & \mathfrak{M},v\nVdash\varphi\\
\mathfrak{M},v\Vdash\varphi\wedge\psi & \text{iff} & \mathfrak{M},v\Vdash\varphi\ \text{ and }\ \mathfrak{M},v\Vdash\psi\\
\mathfrak{M},v\Vdash[\beta]\varphi & \text{iff} & \text{for all } v'\in V \text{ such that } vR_\beta v'\colon\ \mathfrak{M},v'\Vdash\varphi.
\end{array}$$

Furthermore, $\mathfrak{M}\Vdash\varphi$ denotes that for all vertices v in $\mathfrak{M}$ it is the case that $\mathfrak{M},v\Vdash\varphi$. $\mathfrak{M},v\Vdash\Gamma$ signifies that $\mathfrak{M},v\Vdash\gamma$, for all γ in Γ. Finally, $\mathfrak{F}\Vdash\varphi$ and $\mathfrak{F},v\Vdash\varphi$ denote that, respectively, $\mathfrak{M}\Vdash\varphi$ and $\mathfrak{M},v\Vdash\varphi$, for all models $\mathfrak{M}$ on $\mathfrak{F}$. We will use $\vDash_{\mathscr{C}}$ to symbolize *local semantical modal consequence* with respect to the class of *game frames* denoted by $\mathscr{C}$. *I.e.*, we have $\Gamma\vDash_{\mathscr{C}}\varphi$ if and only if $\mathfrak{M},v\Vdash\Gamma$ implies $\mathfrak{M},v\Vdash\varphi$, for all vertices v of all models $\mathfrak{M}$ on a *game-frame* in $\mathscr{C}$. In case $\mathscr{C}$ is the class of all game-frames, we write $\Gamma\vDash_{\mathrm{M}}\varphi$.

In the next chapter we will present a sound and complete axiomatization for this semantic modal consequence relation. The modal semantics of multi-modal matrix languages is confined to models on game-frames. This complicates the Henkin-style completeness proof to some degree as the model constructed should be guaranteed to be based on a game-frame.

Characterizing Subgame Perfect Nash Equilibria

A strategy profile being a Nash equilibrium in a particular extensive game, is a feature of that game which, moreover, is reflected in a particular structural property of game-frames defined on it. A condition for a game-frame to evince this feature is that the strategy profile in question and each interested player be represented by a label in the respective multi-modal matrix language. The aim of this section is to formulate this structural property of game-frames and to characterize it by means of a formula schema of the modal language. To make this idea somewhat more precise, let s be a strategy profile in some extensive game E. Let further $L(A,B)$ be a multi-modal matrix language and $\mathfrak{F}_E$ a game-frame on E in which s is represented by some strategy label in B. What we are after is a formula schema $\vartheta\,(s)$ such that:

$$\mathfrak{F}_E \Vdash \vartheta\,(s) \quad \text{iff} \quad s \text{ is a subgame perfect Nash equilibrium in } E.$$

It turns out that such a formula schema can be obtained as a special case of a familiar schema in standard modal correspondence theory. A frame $\left(V, \{R_\beta\}_{\beta\in B}\right)$ for a multi-modal language $L(A,B)$ — *i.e.*, not *per se* a multi-modal *matrix* language as introduced in the previous subsection — containing k, l, m and n as labels in B for $L(A,B)$ is said to have the (k,l,m,n)-*confluence property* if:

$$\text{for all } v,w,x \in V\colon\ vR_kw \text{ and } vR_mx \ \text{ imply } \ \text{for some } y \in V\colon\ wR_ly \text{ and } xR_ny.$$

Here the labels k, l, m and n need *not* necessarily be distinct. The following fact then holds. For a proof the reader be referred to Popkorn (1994).

Fact 3.3.2 *(Confluence) Let $L(A,B)$ be a multi-modal language containing k, l, m and n as labels. Then the formula schema $\langle k\rangle[l]\varphi \rightarrow [m]\langle n\rangle\varphi$ characterizes frames for $L(A,B)$ satisfying the (k,l,m,n)-confluence property.*

If R_n is taken to be the identity relation on the set of vertices (k,l,m,n)-confluence reduces to the following property, which for obvious reasons we dub (k,l,m)-*Euclidicity*:

$$\text{for all } v,w,x \in V\colon\ vR_kw \text{ and } vR_mx \ \text{ imply } \ wR_lx.$$

As a special case of Fact 3.3.2 we now obtain as a corollary the following fact, of which also the direct proof is elementary:

Corollary 3.3.3 *For $L(A,B)$ a multi-modal language containing k, l and m as labels, the formula schema $\langle k\rangle[l]\varphi \rightarrow [m]\varphi$ characterizes frames for $L(A,B)$ satisfying (k,l,m)-Euclidicity.*

By appropriately choosing k, l and m from the labels of a multi-modal matrix language $L(A,B)$ a strategy profile s being a subgame perfect best response for a player i in a game-frame $\mathfrak{F}$ can be characterized by the formula schema $\langle k\rangle[l]\varphi \rightarrow [m]\varphi$. Taking $\hat{s}_i$ for k, i for l and $\hat{s}_\emptyset$ for m, respectively, gives the desired result. Considering that this schema characterizes frames satisfying $(\hat{s}_i, i, \hat{s}_\emptyset)$-Euclidicity this makes informally sufficient sense:

$$\text{for all } v, v', v'' \in V: \ vR_{\hat{s}_i}v' \text{ and } vR_{\hat{s}_\emptyset}v'' \ \text{ imply } \ v'R_i v''.$$

In words this condition says that, if play commences at a vertex v, player i values the vertex v'' that the strategy profile s determines as an outcome at least as highly as any vertex v' that i can force to come about by unilaterally deviating from s. If this is the case, by deviating from s the player i will not be better off than by sticking to the strategy prescribed by s. The following proposition establishes this observation as an appropriate basis for the characterization of the game-theoretical property of a strategy profile being a subgame perfect response for a player in a game.

Proposition 3.3.4 *Let s be a strategy profile and i a player of an extensive game E. Let further $L(A,B)$ be a multi-modal matrix language, $\mathfrak{F}_E$ a game-frame for $L(A,B)$ on E in which s and i are represented by a label in B. Then:*

s is a subgame perfect best response for i in E iff $\mathfrak{F}_E$ is $(\hat{s}_i, i, \hat{s}_\emptyset)$-Euclidean.

Proof: For the left-to-right direction, assume the contrapositive, *i.e.*, that $\mathfrak{F}_E$ is *not* $(\hat{s}_i, i, \hat{s}_\emptyset)$-Euclidean. Then, there are vertices v, v' and v'' such that:

$$(a) \quad vR_{\hat{s}_i}v' \qquad\qquad (b) \quad vR_{\hat{s}_\emptyset}v'' \qquad\qquad (c) \quad \textit{not: } v'R_i v''.$$

The frame $\mathfrak{F}_E$ being a game-frame on E and i and s being represented by labels in B, these claims correspond to:

$$(a') \quad v' \in \hat{s}_i(v) \qquad\qquad (b') \quad v'' \in \hat{s}_\emptyset(v) \qquad\qquad (c') \quad (v', v'') \notin \rho_i.$$

With (a') and Fact 3.2.3 there is some s' such that $\hat{s}^i_{s'}(v) = v'$. Moreover, since $\hat{s}_\emptyset(v) = \{\hat{s}(v)\}$, also $\hat{s}(v) = v''$. Hence, with (c'), $\left(\hat{s}^i_{s'}(v), \hat{s}(v)\right) \notin \rho_i$, *i.e.*, s is *not* a subgame perfect Nash equilibrium in E.

For the right-to-left direction, assume that s is not a subgame perfect Nash equilibrium in E. Then for some vertex v, some player i and some strategy profile s', $\left(\hat{s}^i_{s'}(v), \hat{s}(v)\right) \notin \rho_i$. By definition of $\mathfrak{F}_E$ as a game-frame on E, however, both $vR_{\hat{s}_i}\hat{s}^i_{s'}(v)$ and $vR_{\hat{s}_\emptyset}\hat{s}(v)$. It follows that $\mathfrak{F}_E$ is not $(\hat{s}_i, i, \hat{s}_\emptyset)$-Euclidean. $\dashv$

Putting things together we obtain the following theorem, which lays down the results we set out to prove in this section. The reader recall that a player is called *interested* if he values some vertices strictly higher than other vertices.

Theorem 3.3.5 *Let $L(A,B)$ be a multi-modal matrix language, E an extensive game and $\mathfrak{F}_E$ a game-frame for $L(A,B)$ on E. Assume that I be a subset of the player labels B_0 and contain labels representing each* interested *player in E. For i a player and s a* *strategy profile in E both represented by a label in B, then:*[1]

$$\begin{array}{rll}
\mathfrak{F}_E, v_\epsilon \Vdash \langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi & \textit{iff} & \textit{s is a best response for i in E} \\
\mathfrak{F}_E \Vdash \langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi & \textit{iff} & \textit{s is a s.p. best response for i in E} \\
\mathfrak{F}_E, v_\epsilon \Vdash \bigwedge_{i \in I} \left(\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi \right) & \textit{iff} & \textit{s is a Nash equilibrium in E} \\
\mathfrak{F}_E \Vdash \bigwedge_{i \in I} \left(\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi \right) & \textit{iff} & \textit{s is a s.p. Nash equilibrium in E}.
\end{array}$$

Proof: As all claims rest on much the same principles, we only present the proof of the fourth claim here. For the right-to-left direction, first assume that the formula schema $\bigwedge_{i \in I} \left(\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi \right)$ is not valid in $\mathfrak{F}_E$. Hence, for some player i and some formula φ we have $\mathfrak{F}_E \nVdash \langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi$. Consequently, the formula schema $\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi$ is not valid in $\mathfrak{F}_E$ either. In virtue of Corollary 3.3.3, then, $\mathfrak{F}_E$ does not satisfy $(\hat{s}_i, i, \hat{s}_\emptyset)$-Euclidicity. With Proposition 3.3.4, then s is not a subgame perfect best response for i, and, *a fortiori*, neither a subgame perfect Nash equilibrium.

For the opposite direction, assume for some vertex v, some player i and some strategy profile s', that $\left(\hat{s}^i_{s'}(v), \hat{s}(v) \right) \notin \rho_i$. Observe that this renders i an interested player. An easy little inductive argument, which we will leave to the reader, establishes that $\hat{s}^i_{s'}(v) \in \hat{s}_i(v)$. Hence, by definition of $\mathfrak{F}_E$ both $vR_{\hat{s}_i}\hat{s}^i_{s'}(v)$ and $vR_{\hat{s}_\emptyset}\hat{s}(v)$. It follows that $\mathfrak{F}_E$ is not $(\hat{s}_i, i, \hat{s}_\emptyset)$-Euclidean. Hence, the formula schema $\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi$ is not valid on $\mathfrak{F}_E$ and *a fortiori* neither is the formula schema $\bigwedge_{i \in I} \left(\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi \right)$. This concludes the proof. $\dashv$

In the sequel, For each label $i \in B_0$ and each label $s \in B_1$, we refer to the axiom schema $\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi$ by $5_{s,i}$ and the axiom schema $\bigwedge_{i \in I} \left(\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi \right)$ by 5^I_s.

3.4 Characterizing Nash Equilibria in Dynamic Logic

In the previous section, we argued that the structural dependencies that obtain between the players' preferences and their strategies when a strategy profile is a (subgame perfect) Nash equilibrium can suitably be characterized in a multi-modal matrix language $L(A, B)$. Some of the labels of such a multi-modal matrix language represent a strategy profile s and are interpreted as the graph of $\hat{s}_\emptyset$. These labels, however, have no further internal structure and their accompanying accessibility relations are semantically primitive. As a consequence, in order to evaluate a formula of the form $[\hat{s}_\emptyset]\varphi$ at a vertex v, one needs to calculate, quite independently of the semantics, the value of $\hat{s}_\emptyset(v)$ in the game under scrutiny in order to identify the vertices reachable from v *via R*. A similar remark applies to the evaluation of formulas of the form $[\hat{s}_i]\varphi$ in a vertex v, which requires the calculation of the value of $\hat{s}_i(v)$. In the semantics of the multi-modal matrix

[1] Here $\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi$ and $\bigwedge_{i \in B_0} \left(\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\emptyset]\varphi \right)$ denote *formula schemas*, rather than formulas. Furthermore, 's.p.' abbreviates 'subgame perfect'.

languages much of the burden has thus been put on the the transformation of an extensive game to a game-frame, which requires reasoning of a game-theoretical rather than a logical nature.

For an illustration of this point consider once again the game of Example 3.2.2, above. Let $L(A,B)$ be a feasible multi-modal matrix language containing a label s for the strategy profile s_{RLll}. In order to evaluate, *e.g.*, the formula of the form $[\hat{s}_2]\varphi$ at a state v in a model on the corresponding frame, one should investigate whether φ holds in all vertices in $\hat{s}_2\,(v)$. The relation $R_{\hat{s}_2}$, however, is taken to be semantically primitive and to establish that, *e.g.*, $v_\epsilon R_{\hat{s}_2} z_5$ but that not $v_3 R_{\hat{s}_2} z_2$ the semantics is of no further help. These facts have to be obtained independently at the meta-level of reasoning.

The set of labels of the dynamic language of *Propositional Dynamic Logic* (PDL) has a richer structure, giving rise to a highly expressive modal logic. Exploiting this structure and expressive power some of the semantic burden can be shifted from the informal meta-level of reasoning about the model to the object-level of the logic. We will find that the relations corresponding to $\hat{s}_i$ and $\hat{s}_\emptyset$ are the accessibility relations associated with labels denoting complex programs, which allow for further semantical analysis. Also the way a frame for an appropriate dynamic language is constructed from an extensive game is more direct and preserves more of the treelike structure of an extensive game than was the case for multi-modal matrix languages (for an illustration of this point compare Figure 3.6, above, and Figure 3.7, below).

This section concerns a class of two-sorted multi-modal languages $L(A,B)$, where B is the union of two disjoint label sets B_0 and Π_{B_1}, where Π_{B_1} denotes the set of PDL-programs over a set B_1 of atomic programs. The set of formulas φ and the set of programs π of such a language $L(A,B)$ — which we will call *dynamic multi-modal languages* — are given by the following BNF-grammar, with $a \in A$, $\beta_0 \in B_0$ and $\beta_1 \in B_1$:

$$\begin{array}{lcl} \varphi & ::= & a \;\big|\; \neg\varphi \;\big|\; \varphi_0 \wedge \varphi_1 \;\big|\; [\pi]\varphi \;\big|\; [\beta_0]\varphi \\ \pi & ::= & \beta_1 \;\big|\; \pi_0;\pi_1 \;\big|\; \pi_0 \cup \pi_1 \;\big|\; \pi^* \;\big|\; \varphi? \end{array}$$

Extensive games are again used as the basis for the models on which such languages are interpreted. The propositional connectives and the program operators obtain their usual informal readings of negation ($\neg$), conjunction ($\wedge$), sequential composition (;), non-deterministic choice ($\cup$), iteration (*) and test ($?$). We also have the usual abbreviations, in particular that of "`while` φ `do` π `od`" for "$(\varphi?;\pi)^*\,;\neg\varphi$". The labels in B_0 go proxy for the players of a game, giving rise to the informal reading of $[i]\varphi$ as *"φ holds in all states at least as preferable to i as the state of evaluation"*, as before. The atomic programs in B_1 are interpreted as a subset of the edges of the game-tree. We assume this set of atomic programs B_1 to be the union of two disjoint sets B_{10} and B_{11}. Each atomic program $\beta \in B_{10}$ is associated with a player i and runs along those edges (v, v') of which v is assigned to the player i. Letting β_{10} be associated with player i, then, intuitively, $[\beta_{10}]\varphi$ reads *"if i is to move, then φ holds at the next stage of the game no matter which strategy i decides to act upon"*. The atomic program β_{10} in B_{10} associated with a player i will be denoted by $\pi(i)$. We will moreover assume that the set of players

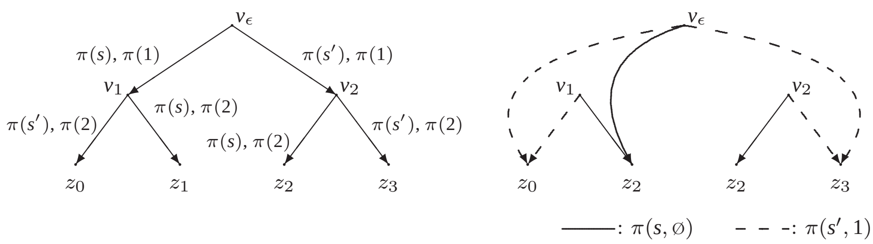


Figure 3.7. Transformation from the extensive game in Figure 3.6 to a dynamic game-frame (left) with respect to two strategy profiles s and s' and their corresponding atomic programs $\pi(s)$, $\pi(s')$, $\pi(1)$ and $\pi(2)$. The righthand figure shows the programs $\pi(s, \emptyset)$ and $\pi(s', 1)$. Note that these can be *"derived"* from the lefthand figure, whereas in the multi-modal framework these relations were primitive .

associated with the labels in B_0 is identical with the set of players associated with the labels in B_{10}. Each atomic program β_{11} in B_{11} is associated with a strategy profile s. Informally, $[\beta_{11}]\varphi$ holds at a vertex v, if φ holds at the next stage that the game will be in if the strategy profile s is adhered to. For s a strategy profile, $\pi(s)$ denotes the label in B_{11} it is thus associated with. We now define formally:

Definition 3.4.1 *(Dynamic game-frames and dynamic game-models)* A *frame* for a dynamic multi-modal language $L(A, B)$ is a tuple $\left(V, \{R_\beta\}_{\beta\in B}\right)$, where V is a set of vertices and $R_\beta \subseteq V \times V$, for each $\beta \in B$. A *label map* for $L(A, B)$ a dynamic multi-modal language on an extensive game E is a function f mapping each label in B_{11} onto a strategy profile of E and each label in B_0 and each label in B_{10} onto a player in N such that $f(B_0) = f(B_{10})$. In the sequel we will usually tacitly assume such a label map f and, if $f(\beta_0) = i$, $f(\beta_{10}) = j$ and $f(\beta_{11}) = s$, write i, $\pi(j)$ and $\pi(s)$ for, β_0 in B_0, β_{10} in B_{10} and β_{11} in B_{11}, respectively. A *dynamic game-frame* $\mathfrak{F}_E^{\mathrm{PDL}}$ for $L(A, B)$ on an extensive game E is a tuple $\left(V, \{R_\beta\}_{\beta\in B}, f\right)$, where $\left(V, \{R_\beta\}_{\beta\in B}\right)$ is an PDL-frame and f a label map on E such that:

$$\begin{aligned} vR_i v' &\quad \text{iff} \quad (v, v') \in \rho_i \\ vR_{\pi(i)} v' &\quad \text{iff} \quad P(v) = i \;\text{ and }\; (v, v') \in R \\ vR_{\pi(s)} v' &\quad \text{iff} \quad s(v) = v', \end{aligned}$$

If formal rigor permits we will often omit the superscript PDL for aesthetic reasons. A *dynamic game-model* $\mathfrak{M}$ for $L(A, B)$ is defined as usual as a pair $(\mathfrak{F}, \mathfrak{V})$, where $\mathfrak{F}$ is a dynamic game-frame for $L(A, B)$ and $\mathfrak{V}$ an interpretation function for the propositional variables in A, *i.e.*, $\mathfrak{V}\colon V \to 2^A$ as before. Figure 3.7 illustrates the transformation of an extensive game to a dynamic game frame.

The evaluation of the formulas of a dynamic multi-modal language $L(A, B)$ in a PDL-model is then as usual.

$$\begin{array}{lll}
\mathfrak{M}, v \Vdash a & \text{iff} & a \in \mathfrak{V}(v) \\
\mathfrak{M}, v \Vdash \neg\varphi & \text{iff} & \mathfrak{M}, v \nVdash \varphi \\
\mathfrak{M}, v \Vdash \varphi \wedge \psi & \text{iff} & \mathfrak{M}, v \Vdash \varphi \text{ and } \mathfrak{M}, v \Vdash \psi \\
\mathfrak{M}, v \Vdash [\beta]\varphi & \text{iff} & \text{for all } v' \in V \text{ such that } vR_\beta v' : \ \mathfrak{M}, v' \Vdash \varphi.
\end{array}$$

A PDL-model $\mathfrak{M}$ is said to be *regular* if program connectives ";", "$\cup$", "$*$" and "?" have their intuitive interpretations of *sequential composition*, *non-deterministic choice*, *iteration* and *test*, respectively, *i.e.*, if the following conditions are fulfilled:

$$\begin{aligned}
R_{\pi_1;\pi_2} &= R_{\pi_1} \circ R_{\pi_2} \\
R_{\pi_1 \cup \pi_2} &= R_{\pi_1} \cup R_{\pi_2} \\
R_{\pi^*} &= (R_\pi)^* \\
R_{\varphi?} &= \big\{(v, v) : \ \mathfrak{M}, v \Vdash \varphi\big\}.
\end{aligned}$$

Here $R_{\pi_1} \circ R_{\pi_2}$ denotes the relational composition of R_{π_1} and R_{π_2}, and $(R_\pi)^*$ is the transitive reflexive closure or ancestral of R_π. In the sequel we will assume PDL-models to be regular.

The important thing to observe in this definition is that the accessibility relations R_i, $R_{\pi(i)}$ and $R_{\pi(s)}$ can be read off from the extensive game specification almost immediately. In particular, the construction does not invoke the correspondences $\hat{s}_i$ and $\hat{s}_\varnothing$ for the interpretation of the atomic progams.

Theorem 3.3.5, above, showed that subgame perfect Nash equilibria are characterized in multi-modal matrix languages by the axiom schema $\bigwedge_{i \in N} \big(\langle \hat{s}_i \rangle [i]\varphi \rightarrow [\hat{s}_\varnothing]\varphi\big)$. The dynamic modal languages of this section do not possess the modalities $[\hat{s}_\varnothing]$ and $[\hat{s}_i]$ *explicitly*. However, for each dynamic modal language $L(A, B)$ they can be defined implicitly as molecular PDL-programs. Let s be a label in B_{11} representing a strategy profile and let $\{i_0, \ldots, i_m\}$ be a subset of labels in B_0 denoted by I. Then, introduce the following abbreviation:

$$\pi(s, I) \quad =_{df.} \quad \texttt{while } \langle \pi(s) \rangle \rightarrow \ \texttt{do } \pi(s) \cup \pi(i_0) \cup \ldots \cup \pi(i_m) \ \texttt{od}.$$

We will write $\pi(s, i)$ for $\pi(s, \{i\})$. The idea is then that the program $\pi(s, i)$ performs the same task in PDL as the label $\hat{s}_i$ in the multi-modal languages, and, similarly, $\pi(s, \varnothing)$ is the dynamic counterpart of the multi-modal label $\hat{s}_\varnothing$. Construed as a program, $\pi(s, I)$ performs non-deterministically one of the programs $\pi(i)$ or $\pi(s)$, as long as $\pi(s)$ is enabled. Given the informal readings of the atomic programs $\pi(s)$ and $\pi(i)$ have in the dynamic game-models, $\pi(s, I)$ also allows for a rather more game-theoretical interpretation. The accessibility relation $R_{\pi(s,I)}$ connects a vertex v with a leaf z of the game tree, whenever z is a possible outcome state if play is commenced in v and the strategy profile s is adhered to by all players, with the possible exception of the players in I. Formally, the following proposition vindicates this intuitive interpretation.

Proposition 3.4.2 *Let E denote the extensive game $\left(V,R,N,P,\{\rho_i\}_{i\in N}\right)$ and let be $\mathfrak{F}_E$ be a dynamic game-frame on E for a dynamic multi-modal language $L(A)$. Let furthermore I be a subset of players in N represented by labels in B_0 and s a strategy profile of E that is represented by a label in B_{11}. Then for all vertices v, v' in V:*

$$vR_{\pi(s,I)}v' \quad \textit{iff} \quad v' \in \hat{s}_I(v).$$

Proof: Consider an arbitrary model $\mathfrak{M}$ on $\mathfrak{F}_E$. Define the *height* of a vertex v in (V,R), denoted by $hgt(v)$, inductively as:

$$hgt(v) \quad =_{df.} \quad \begin{cases} 0 & \text{if } v \text{ is a leaf,} \\ 1 + max\left\{ hgt(v') : \ (v,v') \in R \right\} & \text{otherwise.} \end{cases}$$

The proof is then by induction on $hgt(v)$.

For the basis assume $hgt(v) = 0$. Then v is a leaf and we have $\hat{s}_I(v) = \{v\}$. Since v is a leaf there is no v' such that $s(v) = v'$ and accordingly, $\mathfrak{M}, v \not\models \langle\pi(s)\rangle\top$. Hence, the guard of $\pi(s,I)$ is not satisfied at v and $vR_{\pi(s,I)}v'$ if and only if $v' = v$, which proves the case.

For the induction step let $hgt(v) = n+1$. Then v is an internal node and by definition of a strategy profile there is some v' such that $s(v) = v'$, which makes that the guard of $\pi(s,I)$ is satisfied at v. Hence for all vertices v':

$$vR_{\pi(s,I)}v' \ \text{ iff } \ vR_{\pi(s)\cup\pi(i_0)\cup\ldots\cup\pi(i_m);\pi(s,I)}v' \ \text{ iff } \ vR_{\pi(s)\cup\pi(i_0)\cup\ldots\cup\pi(i_m)} \circ R_{\pi(s,I)}v'.$$

Now, either $P(v) \in I$ or $P(v) \notin I$. If the latter, for no $i \in I$ there is a v'' such that $vR_{\pi(i)}v''$. Hence, for an arbitrary vertex v'', we have $vR_{\pi(s)\cup\pi(i_0)\cup\ldots\cup\pi(i_m)}v''$ if and only if $vR_{\pi(s)}v''$. Consequently also, $vR_{\pi(s,I)}v''$ if and only if $vR_{\pi(s)} \circ R_{\pi(s,I)}v''$. Now consider the following equivalences:

$$\begin{aligned} vR_{\pi(s,I)}v' \quad & \text{iff}_{P(v)\notin I} \quad vR_{\pi(s)} \circ R_{\pi(s,I)}v' \quad \text{iff} \quad \text{for some } v'' : vR_{\pi(s)}v''R_{\pi(s,I)}v' \\ & \text{iff}_{(*)} \quad s(v)\,R_{\pi(s,I)}v' \quad \text{iff}_{i.h.} \quad v' \in \hat{s}_I(s(v)) \quad \text{iff}_{(**)} \quad v' \in \hat{s}_I(v). \end{aligned}$$

The induction hypothesis is applicable because obviously $hgt(s(v)) < hgt(v)$. Observe further that in virtue of Definition 3.4.1, $vR_{\pi(s)}v''$ if and only if $v'' = s(v)$; whence the equivalence marked with the asterisk. The inference step indicated with the double asterisk is valid in virtue of $s_I(v) = \{s(v)\}$, because $P(v) \notin I$, and, therefore, $\hat{s}_I(v) = \bigcup\left\{\hat{s}_I(w) : \ w \in s_I(v)\right\} = \hat{s}_I(s(v))$.

In the former case in which $P(v) \in I$, let i denote $P(v)$. Because $vR_{\pi(s)}v''$ implies $vR_{\pi(i)}v''$, then, also:

$$vR_{\pi(s)\cup\pi(i_0)\cup\ldots\cup\pi(i_m)}v'' \quad \text{iff} \quad vR_{\pi(s)\cup\pi(i)}v'' \quad \text{iff} \quad vR_{\pi(i)}v''.$$

Now consider the following equivalences:

$$
\begin{array}{lll}
vR_{\pi(s,I)}v' & \text{iff}_{P(v)=i} & vR_{\pi(i)} \circ R_{\pi(s,I)}v' \\
& \text{iff} & \text{for some } v'' : vR_{\pi(i)}v''R_{\pi(s,I)}v' \\
& \text{iff}_{(*)} & \text{for some } v'' \in s_I(v) : v''R_{\pi(s,I)}v' \\
& \text{iff}_{i.h.} & \text{for some } v'' \in s_I(v) : v' \in \hat{s}_I(v'') \\
& \text{iff} & v' \in \bigcup \left\{ \hat{s}_I(v'') : \ v'' \in s_I(v) \right\} \\
& \text{iff} & v' \in \hat{s}_I(v) .
\end{array}
$$

The induction hypothesis is applicable because for all $v'' \in s_I(v)$, it is the case that $hgt(v'') < hgt(v)$. Here, the inference step marked with the asterisk holds in virtue of Definition 3.4.1 and the definition of $s_I(v)$ on page 70, above. ⊣

The construction of the frames $\mathfrak{F}_E^{\mathrm{M}}$ and $\mathfrak{F}_E^{\mathrm{PDL}}$ from an extensive form E guarantees that if the one satisfies $(\hat{s}_i, i, \hat{s}_\varnothing)$-Euclidicity, the other satisfies $(\pi(s,i), i, \pi(s,\varnothing))$-Euclidicity and *vice versa*.

Corollary 3.4.3 *Let E be an extensive game. Consider a frame $\mathfrak{F}_E^{\mathrm{M}}$ for a multi-modal matrix language $L(A,B)$ that is a game-frame on E in virtue of a label map f. Let, further, $\mathfrak{F}_E^{\mathrm{PDL}}$ be a dynamic game-frame for a dynamic language $L(A',B')$ on E given by $\left(V, \{B_\beta\}_{B'}\right)$. Assume that $B_0 = B'_0$ and that $B_1 = B'_{11}$ and that $f(B_0) = f(B'_0) = f(B'_{10})$ and $f(B_1) = f(B_{11})$. Then, for each player i and each strategy profile s of E that are represented by labels in the respective languages, we have:*

$$\mathfrak{F}_E^{\mathrm{M}} \textit{ satisfies } (\hat{s}_i, i, \hat{s}_\varnothing)\textit{-Euclidicity} \ \textit{ iff } \ \mathfrak{F}_E^{\mathrm{PDL}} \textit{ satisfies } (\pi(s,i), i, \pi(s))\textit{-Euclidicity.}$$

Proof: Consider arbitrary vertices v and v' in the extensive game E. First observe that vR_iv' in $\mathfrak{F}_E^{\mathrm{M}}$ if and only if vR_iv' in $\mathfrak{F}_E^{\mathrm{PDL}}$ because $\mathfrak{F}_E^{\mathrm{M}}$ and $\mathfrak{F}_E^{\mathrm{PDL}}$ are a game-frame and a dynamic game-frame on E, respectively. Also for $X \subseteq \{i\}$:

$$vR_{\hat{s}_X}v' \text{ in } \mathfrak{F}_E^{\mathrm{M}} \quad \text{iff}_{\text{Def. 3.3.1}} \quad v' \in \hat{s}_X(v) \quad \text{iff}_{\text{Prop. 3.4.2}} \quad vR_{\pi(s,X)}s \text{ in } \mathfrak{F}_E^{\mathrm{PDL}}.$$

Hence, in particular, $R_{\hat{s}_\varnothing} = R_{\pi(s,\varnothing)}$ and $R_{\hat{s}_i} = R_{\pi(s,i)}$. The claim then follows immediately. ⊣

In virtue of this observation we now have the following result, which states that subgame perfect Nash equilibria can be characterized in dynamic multi-modal languages in much the same manner as that was the case for multi-modal matrix languages.

Corollary 3.4.4 *Let $L(A,B)$. Let $\mathfrak{F}_E^{\mathrm{PDL}}$ be a dynamic game frame on an extensive game E for $L(A,B)$ given by $\left(V, \{R_\beta\}_{\beta\in B}, f\right)$. Assume that I be a subset of the player labels B_0 containing labels representing each* interested *player in E. For i a player* *and s a strategy profile in E both represented by a label in, respectively,* B_0 *and* B_{11}*, then the following four equivalences hold:*

$$\mathfrak{F}_E^{\mathrm{PDL}}, v_\epsilon \Vdash \langle\pi(s,i)\rangle[i]\varphi \to [\pi(s,\varnothing)]\varphi \;\; \textit{iff} \;\; \textit{s is a best response for i in E}$$

$$\mathfrak{F}_E^{\mathrm{PDL}} \Vdash \langle\pi(s,i)\rangle[i]\varphi \to [\pi(s,\varnothing)]\varphi \;\; \textit{iff} \;\; \textit{s is a s.p. best response for i in E}$$

$$\mathfrak{F}_E^{\mathrm{PDL}}, v_\epsilon \Vdash \textstyle\bigwedge_{i\in I}\big(\langle\pi(s,i)\rangle[i]\varphi \to [\pi(s,\varnothing)]\varphi\big) \;\; \textit{iff} \;\; \textit{s is a Nash equilibrium in E}$$

$$\mathfrak{F}_E^{\mathrm{PDL}} \Vdash \textstyle\bigwedge_{i\in I}\big(\langle\pi(s,i)\rangle[i]\varphi \to [\pi(s,\varnothing)]\varphi\big) \;\; \textit{iff} \;\; \textit{s is a s.p. Nash equilibrium in E}.$$

Proof: Almost immediate from Theorem 3.3.5, Proposition 3.4.2 and the semantics of multi-modal matrix languages on game-frames. All cases run along analogous lines and the proof is confined to that of the fourth case. Consider the multi-modal matrix language $L(A, B')$ with $B' =_{df.} B_0 \cup B_{11} \cup B_0 \times B_{11}$. Then, $\big(V, \{R_\beta\}_{\beta\in B}, f\restriction B'\big)$ is a game-frame on E for $L(A, B')$ and let this game-frame be denoted by $\mathfrak{F}_E^{\mathrm{M}}$. Now consider the following equivalences:

s is a subgame perfect Nash equilibrium in E

iff$_{\text{Th. 3.3.5}}$	$\mathfrak{F}_E^{\mathrm{M}} \Vdash \bigwedge_{i\in I}\big(\langle\hat{s}_i\rangle[i]\varphi \to [\hat{s}_\varnothing]\varphi\big)$
iff	for all $i \in I$: $\mathfrak{F}_E^{\mathrm{M}} \Vdash \langle\hat{s}_i\rangle[i]\varphi \to [\hat{s}_\varnothing]\varphi$
iff	for all $i \in I$, $\mathfrak{F}_E^{\mathrm{M}}$ satisfies $(\hat{s}_i, i, \hat{s}_\varnothing)$-Euclidicity
iff$_{\text{Coroll. 3.4.3}}$	for all $i \in I$, $\mathfrak{F}_E^{\mathrm{PDL}}$ satisfies $(\pi(s,i), i, \pi(s))$-Euclidicity
iff	$\mathfrak{F}_E^{\mathrm{PDL}} \Vdash \bigwedge_{i\in I}\big(\langle\pi(s,i)\rangle[i]\varphi \to [\pi(s,\varnothing)]\varphi\big)$.

This concludes the proof. ⊣

A dynamic game-frame of Definition 3.4.1 reflects the structure of the underlying extensive game in considerably finer detail than the game-frame of Definition 3.3.1 does for the same game. This feature, however, comes with a vengeance in that it imposes heavier requirements on the models to be constructed in a Henkin-style completeness proof. The issue as to a complete axiomatization of the dynamic framework with respect to dynamic game-frames we leave as an open question.

3.5 Conclusions and Other Topics

In this chapter we proposed the use of multi-modal matrix languages for the formal description of a class of extensive games. The games in this particular class all had a finite horizon and assumed perfect information on the part of the players. By focussing on such a limited class of games, the correspondences between the games and the logic could be kept relatively simple. Independent issues were left out of the picture, so as to emphasize the fundamental idea of how modal languages can be used to describe extensive games. Thus, the analysis passed over fundamental game-theoretical topics such as coalition formation, randomization of strategies and other issues involving probabilities as well as over repeated games and games of infinite depth. Incorporation of these issues in the present framework warrants further investigation. Still, a proper treatment would quite likely demand considerable extensions of the languages presented in this chapter. Special mention should be made of imperfect information and related epistemic issues, as modal logics have prominent applications in the formal analysis of knowledge and belief (*cf.*, *e.g.*, Hintikka (1962), Fagin, Halpern, Moses, and Vardi (1995), Meyer and van der Hoek (1995) and Gerbrandy (1999)) and has been firmly established within the field of Artificial Intelligence. Moreover, such modal logics have been deployed in the analysis of the epistemic aspects of games (*cf.*, *e.g.*, Baltag (2002), Battigalli and Bonanno (1999) and van Ditmarsch (2000)). Incorporating features of these logics in the present framework may lead to a more comprehensive modal logical analysis. The concomitant complications should not be shunned.

The multi-modal matrix languages were especially designed to deal with (subgame perfect) Nash equilibrium in pure strategies. Its expressive power is limited to preferences and *individual* divergences from a strategy profile. The characterization of other game-theoretical notions — such as *Pareto efficiency*, *dominance* as well as the various refinements of Nash equilibrium as they have been suggested in the literature — may require more sophisticated concepts. More structure of the extensive games is preserved in the dynamic game-frames. Accordingly, we may expect more from the dynamic language of PDL as to expressiveness with respect to other game-theoretical concepts than Nash equilibrium alone.

These considerations put in perspective the multi-modal matrix languages as we proposed to use them in the description of extensive games. They should by no means be taken as a proposal for a comprehensive and ultimate logical language for the description of extensive games. Rather, we meant to expose some of the structural properties of extensive games which render some strategy profiles to be (subgame perfect) Nash equilibria. The fact that these properties are characterizable in quite an inelaborate formal language, says something fundamental about the elementary nature of Nash equilibria and the expressive requirements for a modal language to characterize them.

Chapter 4

Axiomatization of Extensive Game Logics

4.1 Introduction

In the previous chapter, multi-modal matrix languages were introduced in order to reason about particular features of extensive games with a finite horizon and in which all players have perfect information as to the structure of the game. Accordingly, the frames and models on which these multi-modal matrix languages were interpreted constitute special class of Kripke structures. The notion of modal consequence has been parameterized by a class of game-frames, *i.e.*, $\Gamma \vDash_{\mathscr{C}} \varphi$ was defined to hold if and only if for all vertices v of all models $\mathfrak{M}$ on a *game-frame* in $\mathscr{C}$, $\mathfrak{M}, v \Vdash \Gamma$ implies $\mathfrak{M}, v \Vdash \varphi$ (*cf.*, page 73). Each of these modal consequence relations defines a logic in the respective multi-modal matrix language. This chapter concerns the axiomatization of three of such logics: the one characterized by the class of *all* game-frames and those characterized by the class of game-frames in which a particular strategy profile is, respectively, a player's subgame perfect best response and a subgame perfect Nash equilibrium.

In the upcoming section we formulate a number of axiom schemas for multi-modal matrix languages, which are valid on all game-frames. We find that the minimal normal modal logic containing these axioms, denoted by M, is also complete with respect to the class of all game-frames. Furthermore, $\mathrm{M5}_{s,i}$ and $\mathrm{M5}_s^N$ are introduced as the extensive game logics that result if M augmented with, respectively, the axioms $5_{s,i}$ and 5_s^N, for a player label i and a strategy label s of $L(A,B)$ (*cf.*, page 76). The logic $\mathrm{M5}_{s,i}$ is proved to be complete with respect to the class of game-frames in which the strategy profile represented by the strategy label s is a subgame perfect best response for the player represented by the player label i. Similarly, $\mathrm{M5}_{s,i}$ coincides with the logic characterized by the class of frames based on extensive games in which the strategy label s represents a subgame perfect Nash-equilibrium (and in which there is a label for each interested player).

A Henkin-style construction method is employed to obtain these completeness results.

4.2 The Axioms

For a multi-modal matrix language $L(A,B)$ we have the following axioms. We assume β, β', β'', and β''' to range over the whole of B, β_0 and β_0' over the player labels in B_0, and β_1 and β_1' over the strategy labels in B_1.

$$\begin{aligned}
\textit{Taut.}:&\quad \text{any classical tautology.}\\
K:&\quad [\beta]\big(\varphi\to\psi\big)\to\big([\beta]\varphi\to[\beta]\psi\big)\\
T_{\beta_0}:&\quad [\beta_0]\varphi\to\varphi\\
4_{\beta_0}:&\quad [\beta_0]\varphi\to[\beta_0][\beta_0]\varphi\\
D!_{\beta_1}:&\quad [\beta_1]\varphi\leftrightarrow\langle\beta_1\rangle\varphi\\
E1_{(\beta_0,\beta_1),\beta_1}:&\quad [(\beta_0,\beta_1)]\varphi\to[\beta_1]\varphi\\
E2_{(\beta_0,\beta_1),(\beta_0',\beta_1')}:&\quad [(\beta_0,\beta_1)]\big([(\beta_0',\beta_1')]\varphi\leftrightarrow\varphi\big)\\
E3_{\beta,\beta',\beta'',\beta''',\beta_0}:&\quad [\beta][\beta']\big([\beta_0]\varphi\to\psi\big)\vee[\beta''][\beta''']\big([\beta_0]\psi\to\varphi\big).
\end{aligned}$$

The logic M is closed under the rules of *modus ponens* (*MP*) and *necessitation* (*Nec.*):

$$MP:\ \frac{\varphi\to\psi\qquad\varphi}{\psi}\qquad\qquad Nec.:\ \frac{\varphi}{[\beta]\varphi}.$$

The Hilbert-style axiom system given by these axioms and rules we will refer to as the normal modal logic M. Any multi-modal logic containing M we will refer to as an *extensive game logic*. Accordingly, M is the smallest extensive game logic, in a similar manner as K is the smallest normal modal logic.

Definition 4.2.1 *(Extensive game logics)* An *extensive game logic* Λ for a multi-modal matrix language $L(A,B)$ is any set of formulas of $L(A,B)$ closed under *MP* and *Nec.* and containing all instances of the axiom schemes of *Taut.*, *K*, 4_{β_0}, T_{β_0}, $D!_{\beta_1}$, $E1_{(\beta_0,\beta_1),\beta_1}$, $E2_{(\beta_0,\beta_1),(\beta_0',\beta_1')}$ and $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$. We will write $\Gamma\vdash_\Lambda\varphi$ if there exists a derivation of φ from the theory Γ in an extensive game logic Λ, as usual. The smallest extensive game logic we will refer to by M.

At the conclusion of this chapter we will come to review also the stronger extensive game logics than M, *viz.*, the logics $\mathrm{M5}_{s,i}$ and $\mathrm{M5}_s^I$. The former has $5_{s,i}$, for fixed s and i, as an additional axiom and the latter 5_s^I for a particular subset I of B_0.

Within the setting of extensive games and the intended interpretation of the multi-modal matrix languages, the axioms K through $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$ have quite intuitive readings. The axioms *Taut.* and *K* along with the two rules for *modus ponens* (*MP*) and necessitation (*Nec.*) guarantee extensive game logics to be normal logics. With the accessibility relations for the modal operators with labels in B_0 running over the preferences of players in an extensive game, T_{β_0} and 4_{β_0} warrant the players' preferences to be reflexive and transitive. The axiom $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$ reflects the players' preference relations being connected. Observe that in virtue of *Taut.*, T_{β_0} and $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$, we can derive the following axiom in each extensive game logic:

$$E4_{\beta,\beta',\beta_0}: \quad [\beta]\big([\beta_0]\varphi \to \psi\big) \vee [\beta']\big([\beta_0]\psi \to \varphi\big).$$

The labels in B_1 represent strategy profiles and in particular their outcome functions. For each label β_1 in B_1, the accessibility relation R_{β_1} connects, for some strategy profile s, any vertices v and v' such that $\hat{s}\,(v) = v'$ and as such is the graph of a function. Hence, Axiom $D!_{\beta_1}$, which characterizes functionality of the accessibility relation R_{β_1} in the general setting of modal correspondence theory. Axiom $E1_{(\beta_0,\beta_1),\beta_1}$ characterizes the inclusion of the accessibility relation R_{β_1} in $R_{(\beta_0,\beta_1)}$. The intuition behind this lies in the observation that any outcome that is determined by a strategy profile can also be reached if one of the players has the option to deviate from that strategy profile; the player in question may choose to adhere to the strategy profile after all. The labels in $B_0 \times B_1$ represent the correspondences $\hat{s}_i$ for strategy profiles s and players i. The value of any such correspondence is a set of leaf nodes, from each of which only the leaf itself can be reached. It is exactly this fact that $E2_{(\beta_0,\beta_1),(\beta_0',\beta_1')}$ conveys. Observe that as a consequence of $E2_{(\beta_0,\beta_1),(\beta_0',\beta_1')}$, $D!_{\beta_0}$ and $E1_{(\beta_0,\beta_1),\beta_1}$ we can derive the following more general axiom schema, in which both β and β' range over labels in either B_1 or $B_0 \times B_1$:

$$E5_{\beta,\beta'}: \quad [\beta]\left([\beta']\varphi \leftrightarrow \varphi\right).$$

Note that the scheme $E5_{\beta,\beta'}$ does not hold in general if β or β' are in B_0.

The cogency of these informal remarks are vindicated in the following proposition, which formally establishes the soundness on game-frames of the axioms in question.

Proposition 4.2.2 *(Soundness) Let $L(A,B)$ be a multi-modal matrix language. The axioms K through $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$ as well as $E4_{\beta,\beta',\beta_0}$ and $E5_{\beta,\beta'}$ are valid on all game-frames of $L(A,B)$. The rules MP and Nec., moreover, preserve validity on game-frames.*

Proof: For ordinary multi-modal frames the axioms T_β and 4_β characterize reflexivity and transitivity of R_β, respectively. Similarly $D!_\beta$ characterizes functionality of R_β and $E1_{\beta,\beta'}$ the inclusion of $R_{\beta'}$ in R_β. The axiom schema $E2_{\beta,\beta'}$ characterizes frames in which $R_{\beta'}$ is the identity at every vertex v that is reachable by R_β, *i.e.*, frames for which:

$$\text{for all } v, v' \in V: \quad vR_\beta v' \ \text{ implies } \ \text{for all } v'' \in V\colon v'R_{\beta'}v'' \ \text{ iff } \ v' = v''.$$

Finally, $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$ — $[\beta][\beta']\big([\beta_0]\varphi \to \psi\big) \vee [\beta''][\beta''']\big([\beta_0]\psi \to \varphi\big)$ — characterizes frames in which any two vertices v and v' are comparable with respect to R_{β_0}, whenever the one is reachable from from some third vertex v'' *via* $R_\beta \circ R_{\beta'}$ and the other from the same vertex v'' *via* $R_{\beta''} \circ R_{\beta'''}$. *I.e.*, more formally, $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$ characterizes frames in which for all vertices v, v', v'':

$$\text{if for some } w', w'' \colon vR_\beta w' R_{\beta'} v' \text{ and } vR_{\beta''} w'' R_{\beta'''} v'' \text{ then } v' R_{\beta_0} v'' \text{ or } v'' R_{\beta_0} v'.$$ [1]

All of the above are results of elementary modal correspondence theory.

Since the game-frames for $L(A,B)$ are special cases of ordinary multi-modal frames and the connectives obtain their usual Boolean interpretations *Taut.*, *K*, *MP* and *Nec.* hold without ado and it suffices to show that the properties T_{β_0} through $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$ characterize in ordinary frames are satisfied in the game-frames for $L(A,B)$.

The players' preferences were assumed to be reflexive, transitive and connected and *a fortiori* so are R_{β_0} for each $\beta_0 \in N$. This takes care of the soundness of T_{β_0}, 4_{β_0} and $E3_{\beta,\beta',\beta'',\beta''',\beta_0}$. Strategy profiles determine a unique leaf node as outcome. Formally, $\hat{s}_\varnothing(v) = \{\hat{s}(v)\}$, for each strategy profile s and each vertex v. *I.e.*, for each β_1 in B_1 the accessibility relation R_{β_1} is functional. Hence, $D!_{\beta_1}$ is valid in game-frames as well. In virtue of the monotonicity of $\hat{s}_I$ (*cf.*, page 70), we have in particular that $\hat{s}_\varnothing \subseteq \hat{s}_i$. Hence, also $R_{\beta_1} \subseteq R_{(\beta_0,\beta_1)}$ for all $\beta_0 \in B_0$ and $\beta_1 \in B_1$. The validity of $E1_{(\beta_0,\beta_1),\beta_1}$ follows. For $E2_{(\beta_0,\beta_1),(\beta_0,\beta_1)}$ it suffices to show that for all strategy profiles s and s' and for all vertices v and v' in a game-tree, $v' \in \hat{s}_i(v)$ implies $\hat{s}'_j(v') = \{v'\}$. Merely observe that in general $\hat{s}_i(v) \subseteq Z$ and that for all leaves $z \in Z$ we have that $\hat{s}'_j(z) = \{z\}$ by definition (*cf.* page 70).

Establishing that the rules *MP* and *Nec.* preserve validity on game-frames amounts to a routine check.

Finally, $E4_{\beta,\beta',\beta_0}$ and $E5_{\beta,\beta'}$ are valid on all game-frames of $L(A,B)$ because they are derivable in any extensive game logic Λ. $\dashv$

For easy reference Table 4.1 collects all axioms that have so far been dealt with; the labels are chosen in such a way as to reflect their intended game-theoretical readings as suggested in Definition 3.3.1, above. *I.e.*, typical elements of B_0, B_1 and $B_0 \times B_1$ are represented by, respectively, i, $\hat{s}_\varnothing$ and $\hat{s}_i$.

4.3 Completeness

This section concerns completeness results for the extensive game logics M, $\mathrm{M5}_{s,i}$ and $\mathrm{M5}_s^N$ in a multi-modal matrix language $L(A,B)$. *I.e.*, for Λ one of these logics and $\mathscr{C}$ the intended class of game-frames, we prove that:

$$\Gamma \vDash_{\mathscr{C}} \varphi \quad \text{implies} \quad \Gamma \vdash_\Lambda \varphi.$$

In order to prove completeness of an extensive game logic Λ with respect to a certain class of game-models $\mathscr{C}$, it suffices, to construct for each Λ-consistent theory Γ a game-model $\mathfrak{M}_\Gamma^\Lambda$ that satisfies Γ at some vertex and prove this model to be in the

[1] This property is a close multi-modal relative of that of *piecewise connectedness*. A Kripke frame is said to be piecewise connected if for all vertices v, v' and v'', vRv' and vRv'' implies either $v'Rv''$ or $v''Rv'$.

Taut.:	any classical tautology.
K:	$[\beta]\big(\varphi \to \psi\big) \to \big([\beta]\varphi \to [\beta]\psi\big)$
T_i:	$[i]\varphi \to \varphi$
4_i:	$[i]\varphi \to [i][i]\varphi$
$5_{s,i}$:	$\langle \hat{s}_i \rangle [i]\varphi \to [\hat{s}_\varnothing]\varphi$
5^I_s:	$\bigwedge_{i \in I} \big(\langle \hat{s}_i \rangle [i]\varphi \to [\hat{s}_\varnothing]\varphi\big)$
$D!_{\hat{s}_\varnothing}$:	$[\hat{s}_\varnothing]\varphi \leftrightarrow \langle \hat{s}_\varnothing \rangle \varphi$
$E1_{\hat{s}_i, \hat{s}_\varnothing}$:	$[\hat{s}_i]\varphi \to [\hat{s}_\varnothing]\varphi$
$E2_{\hat{s}_i, \hat{s}'_j}$:	$[\hat{s}_i]\big([\hat{s}'_j]\varphi \leftrightarrow \varphi\big)$
$E3_{\beta,\beta',\beta'',\beta''',i}$:	$[\beta][\beta']\big([i]\varphi \to \psi\big) \vee [\beta''][\beta''']\big([i]\psi \to \varphi\big)$
$E4_{\beta,\beta',i}$:	$[\beta]\big([i]\varphi \to \psi\big) \vee [\beta']\big([i]\psi \to \varphi\big)$
$E5_{\hat{s}_X, \hat{s}'_Y}$:	$[\hat{s}_X]\ ([\hat{s}'_Y]\varphi \leftrightarrow \varphi)$ where $X, Y \in N \cup \{\varnothing\}$

Table 4.1. List of axiom schemas for multi-modal matrix languages $L(A, B)$, where $B = N \cup S \cup (N \times S)$ and β and its primed varieties range over B.

class $\mathscr{C}$. We prove the contrapositive. Assume that $\Gamma \nvdash_\Lambda \varphi$, for an arbitrary theory Γ and an equally arbitrary formula φ. Then, $\Gamma \cup \{\neg\varphi\}$ is Λ-consistent and the model $\mathfrak{M}^\Lambda_{\Gamma \cup \{\neg\varphi\}}$ can be constructed. By assumption, $\mathfrak{M}^\Lambda_{\Gamma \cup \{\neg\varphi\}} \not\Vdash \Gamma \cup \{\neg\varphi\}$, and with $\mathfrak{M}^\Lambda_{\Gamma \cup \{\neg\varphi\}}$ in $\mathscr{C}$, this establishes that $\Gamma \nvDash_{\mathscr{C}} \varphi$.

The semantics for the multi-modal matrix languages is based on the notion of a game-frame. The hardest part in proving completeness for an extensive game logic Λ is thus in to guarantee that this model $\mathfrak{M}^\Lambda_\Gamma$ belongs to the appropriate class of *game-models*.

In this section we first show how for each theory consistent in an extensive game logic Λ a game model $\mathfrak{M}^\Lambda_\Gamma$ in the sense of Definition 3.3.1 can be constructed. *I.e.*, the frame underlying $\mathfrak{M}^\Lambda_\Gamma$ should demonstrably be based on an extensive game. To this end we adopt a Henkin-style construction (step-by-step) method (*cf.*, Blackburn, de Rijke, and Venema (2001), Section 4.6, pages 223–229). Although the axioms of M, M$5_{i,s}$ and M5^N_s are all of a standard nature, it is not obvious, however, whether a standard Henkin-style proof would produce a canonical model — *i.e.*, a model satisfying *each* Λ-consistent theory at the same time — that is based on a game-frame, or that can be transformed into a model that is. A proof of this is likely to become complicated because the structure of a canonical model is pretty much fixed. The construction method, in which for each Λ-consistent theory Γ separately a model $\mathfrak{M}^\Lambda_\Gamma$ satisfying Γ is defined, gives far more control over the structure of the model to be built. In particular, the process of constructing the model $\mathfrak{M}_\Gamma$ for a theory Γ can go hand in hand with the construction of an extensive game $\boldsymbol{E}_\Gamma$ underlying $\mathfrak{M}_\Gamma$. As such this method of proof is more natural for our purposes.

For M, the weakest extensive game logic, the construction of a game-model for each M-consistent theory suffices, as it is supposed to be complete with respect to the class of all game-frames for $L(A,B)$. For completeness of $\mathrm{M5}_{i,s}$ and $\mathrm{M5}_s^N$, however, it has additionally to be proved that the models this construction yields are in the appropriate class of models. *I.e.*, for $\mathrm{M5}_{i,s}$ it has to be shown that, for each $\mathrm{M5}_{i,s}$-consistent theory Γ the model $\mathfrak{M}_\Gamma^{\mathrm{M5}_{i,s}}$ is based on, an extensive game in which the strategy profile s contains a best response for player i. In the case of $\mathrm{M5}_s^N$, similarly, one should show that, in the extensive game underlying a model $\mathfrak{M}_\Gamma^{\mathrm{M5}_s^N}$, for each interested player for a $\mathrm{M5}_s^N$-consistent theory Γ, there be a label in N, and, moreover, the strategy profile denoted by the label s is a subgame perfect Nash equilibrium.

Before entering on the formal elaboration, we first devote some more or less informal remarks as to the structure of the proof.

The Structure of the Construction

For each extensive game logic Λ in a multi-modal matrix language and each theory Γ we construct a model $\mathfrak{M}_\Gamma^\Lambda$, omitting the superscript Λ when clear from the context. Let Λ be an extensive game logic and Γ be a Λ-consistent theory in a multi-modal matrix language $L(A,B)$. The main burden will be on guaranteeing that the model $\mathfrak{M}_\Gamma$ be an actual game model. For each Λ-consistent theory Γ, therefore, we construct a model $\mathfrak{M}_\Gamma$ along with an extensive game $\boldsymbol{E}_\Gamma$. A game-frame $\mathfrak{F}_\Gamma$ based on this game $\boldsymbol{E}_\Gamma$ then underlies $\mathfrak{M}_\Gamma$.

In the construction of the game-model $\mathfrak{M}_\Gamma$, we first define a labelled tree $\mathfrak{T}_\Gamma$ consisting of a tree $\boldsymbol{\Sigma}_\Gamma$ and a labelling function $\boldsymbol{\theta}_\Gamma$ assigning Λ-consistent theories to the vertices in $\boldsymbol{\Sigma}_\Gamma$. In particular, $\boldsymbol{\theta}_\Gamma$ assigns a maximal Λ-consistent extension of Γ to the root of $\boldsymbol{\Sigma}_\Gamma$. The set of vertices $\boldsymbol{\Sigma}_\Gamma$ is not entirely independent of the labelling function $\boldsymbol{\theta}_\Gamma$ since it may depend on the theory assigned to a particular vertex whether $\boldsymbol{\Sigma}_\Gamma$ should also contain another vertex. For this reason, induction is relied upon in the definition of $\mathfrak{T}_\Gamma$. This tree $\mathfrak{T}_\Gamma$ contains sufficient information for the definition of a fully-fledged extensive game denoted by $\boldsymbol{E}_\Gamma$ as well as that for the game model $\mathfrak{M}_\Gamma$ based on $\boldsymbol{E}_\Gamma$. The tree on which $\boldsymbol{E}_\Gamma$ is based is given by $\boldsymbol{\Sigma}_\Gamma$. The number of players in $\boldsymbol{E}_\Gamma$ turns out to be one greater than the number of player labels in $L(A,B)$. Also which player is to move at which node depends on the structure of $\boldsymbol{\Sigma}_\Gamma$. Finally, the players' preferences over the vertices of $\boldsymbol{\Sigma}_\Gamma$ are derived from the theories the labelling function $\boldsymbol{\theta}_\Gamma$ assign to the vertices of $\boldsymbol{\Sigma}_\Gamma$.

The vertices of the tree $\boldsymbol{\Sigma}_\Gamma$ are chosen in such a way that appropriate strategy profiles in $\boldsymbol{E}_\Gamma$ can easily be recovered to serve as the interpretations for the strategy labels in B_1 of $L(A,B)$. This furnishes us with a natural label map. Thus a frame $\mathfrak{F}_\Gamma$ is defined that is a *game-frame* on the extensive game $\boldsymbol{E}_\Gamma$ according to Definition 3.3.1. A suitable valuation function $\mathfrak{V}$ for $\mathfrak{F}_\Gamma$ is found by another appeal to the labelling

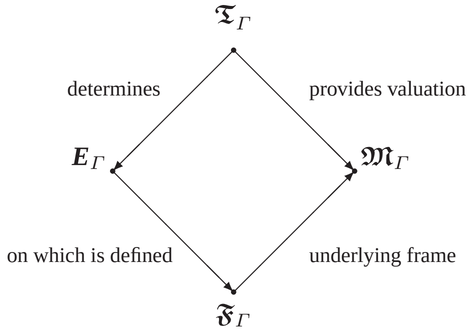


Figure 4.1. Structure of the construction of the game-model $\mathfrak{M}_\Gamma$, where $\mathfrak{T}_\Gamma = (\boldsymbol{\Sigma}_\Gamma, \boldsymbol{\theta}_\Gamma)$.

function $\boldsymbol{\theta}_\Gamma$ of $\mathfrak{T}_\Gamma$: let $\mathfrak{V}$ map each vertex v of $\mathfrak{F}_\Gamma$ onto the set of propositional variables contained in $\boldsymbol{\theta}_\Gamma(v)$, *i.e.*, onto the set $A \cap \boldsymbol{\theta}_\Gamma(v)$. Then define $\mathfrak{M}_\Gamma$ as the very model on $\mathfrak{F}_\Gamma$ with $\mathfrak{V}$ as valuation. The vertices of $\mathfrak{T}_\Gamma$ and $\mathfrak{M}_\Gamma$ coincide. We prove that any formula in the theory $\boldsymbol{\theta}_\Gamma(v)$ holds at v in $\mathfrak{T}_\Gamma$, *i.e.*, that $\mathfrak{M}_\Gamma, v \Vdash \boldsymbol{\theta}_\Gamma(v)$. Because in $\mathfrak{T}_\Gamma$ the root is decorated with a maximal Λ-consistent extension of Γ, we may eventually conclude that $\mathfrak{M}_\Gamma$ satisfies Γ at the root node. The dependencies of the various elements of the construction of $\mathfrak{M}_\Gamma$— *viz.*, $\mathfrak{T}_\Gamma$, $\boldsymbol{E}_\Gamma$, $\mathfrak{F}_\Gamma$ and $\mathfrak{M}_\Gamma$ itself — are as depicted in Figure 4.1.

Formal Exposition of the Construction

a In order to facilitate the proof, we first make some harmless but convenient assumptions. We assume to be working in a countable multi-modal matrix language $L(A, B)$ with $B = N \cup S \cup (N \times S)$. Let us further assume N to be given by a finite initial segment of the positive integers, *i.e.*, $N = \{1, \ldots, n\} \subseteq \omega$ with n the number of player-labels $\|N\|$. The game $\boldsymbol{E}_\Gamma$ to be constructed will comprise an additional *mystery player*, which will be denoted by 0. We will also assume an arbitrary but fixed enumeration $\varphi_0, \ldots, \varphi_n, \ldots$ of the formulas of $L(A, B)$. Moreover, the concept of a maximal Λ-consistent extension of a theories will be heavily relied upon. The Lindenbaum lemma for Λ, stating that any extensive game logic Λ-consistent theory can be extended to a maximal Λ-consistent theory, is reproduced without the routine proof.

Fact 4.3.1 *(Lindenbaum lemma) Every Λ-consistent theory in $L(A, B)$ can be extended to a maximal Λ-consistent theory.*

Having assumed a fixed enumeration of the formulas of $L(A, B)$ we may for each extensive game logic Λ assume a *closure operator* Cl_Λ mapping each Λ-consistent theory Γ

onto a *unique* maximal Λ-consistent extension of Γ. The subscript is usually omitted whenever the logic Λ is understood from the context. The construction of the models $\mathfrak{M}_\Gamma^\Lambda$ is uniform for all extensive game logics Λ *modulo* the notion of consistency involved.

The general idea of the proof is to start with an initial tree, the vertices of which are labelled with theories. In particular, the root will be associated with a maximal Λ-consistent extension of a Λ-consistent theory Γ. Then new vertices are introduced when necessary, *i.e.*, whenever "witness" states are required for formulas of the form $\neg[\hat{s}_i]\varphi$ or $\neg[i]\varphi$ occurring in the theories the vertices are labelled with. (There happens to be no such need for formulas of the form $\neg[\hat{s}_\emptyset]$.) This process should preserve the tree-like character of the structure and it should eventually culminate in the tree $\mathfrak{T}_\Gamma$, on basis of which the game-model $\mathfrak{M}_\Gamma$ can be defined. The nodes of the model coincide with the nodes of the eventual tree. Moreover, any maximal Λ-consistent theory associated with a vertex is to contain exactly those formulas that are satisfied at that vertex in the model. This makes that the theories associated with the vertices are subject to certain consistency constraints and these constraints should be vouchsafed in the process.

The vertices of the labelled tree $\mathfrak{T}_\Gamma$ are selected from the set $(T \cup S \cup \omega)^*$ of finite sequences over $T \cup S \cup \omega$ including the empty sequence ϵ. Henceforth we denote this set of strings by $\boldsymbol{\Sigma}$. The sets T, S and ω are assumed to be pairwise disjoint, T to be countably infinite, and S the set of strategy-labels of $L(A, B)$.[2] We also assume there to be a partition of T in a countably infinite number of countably infinite blocks. Hence, for each sequence σ in $\boldsymbol{\Sigma}$ we may assume there to be a unique countably infinite subset of T, denoted by T_σ and enumerated as $t_0^\sigma, \ldots, t_n^\sigma, \ldots$.

Suppressing the implicit ordering of the strings we will deliberately confuse the tree $\boldsymbol{\Sigma}_\Gamma$ and its set of vertices. For each natural number $n \in \omega$, we let x^n denote the string of n occurrences of x, *e.g.*, $x^3 = xxx$ and $yx^2z = yxxz$. Let furthermore $|\sigma|$ denote the *length* of a string σ. We have ϵ stand for the empty string. In the remainder strings are assumed to be ordered by the *immediate prefix relation* $\prec$, defined for strings σ and σ' over a set X in such a way that $\sigma \prec \sigma'$ if and only if there is some x in X with $\sigma x = \sigma'$. *E.g.*, the strings xy and xyz are thus related but yy and yxx are not.

Conceptually, in the game $\boldsymbol{E}_\Gamma$ to be constructed, the elements of $T \cup S \cup \omega$ could be seen as possible actions and each sequence σ as representing a *history of play* (*cf.*, Osborne and Rubinstein (1994), pp. 89–90). A sequence $ts's$ is then the vertex that will be reached if subsequently the 'actions' t, s' and s are performed. In the game $\boldsymbol{E}_\Gamma$, the strategy profile to be represented by the strategy label s can then easily be recovered as the function that maps each sequence σ onto σs. The sets T and ω are added in order to ensure that there be a sufficient number of vertices in $\boldsymbol{E}_\Gamma$, and eventually also a sufficient number of strategy profiles in the game $\boldsymbol{E}_\Gamma$. Roughly speaking, elements of T are used to introduce vertices falsifying φ as witnesses for formulas of the form $\neg[i]\varphi$. Similarly, the elements of ω are used to construct witness states for formulas of

[2]The first assumption requires one to distinguish an element x in $T \cup S \cup \omega$ from the sequence x of length one in $\boldsymbol{\Sigma}$. In the remainder we will generally assume a natural isomorphism between $T \cup S \cup \omega$ and the set $\{\sigma \in \boldsymbol{\Sigma} : |\sigma| = 1\}$.

the form $\neg[\hat{s}_i]\varphi$.

Now the stage has been set for the definition of the labelled tree $\boldsymbol{\mathfrak{T}}_\Gamma$. The fundamental idea is that for each formula of the form $\neg[\beta]\varphi$ in a theory associated with a vertex σ by $\boldsymbol{\theta}_\Gamma$, there should also be a vertex σ' that falsifies φ and, moreover, is reachable by the accessibility relation R_β in the model $\mathfrak{M}_\Gamma$. The theory associated with vertex σ' will thus contain $\neg\varphi$ and will also have to comply to certain consistency constraints. In particular, σ' should additionally satisfy $\{\psi : [\beta]\psi \in \boldsymbol{\theta}_\Gamma(\sigma)\}$. As, for different labels β and β' the theories $\{\psi : [\beta]\psi \in \Gamma\}$ and $\{\psi : [\beta']\psi \in \Gamma\}$ are not in general Λ-compatible, proper care should be taken that no leaf be reachable by two different accessibility relations R_β and $R_{\beta'}$. Simultaneously, it has to be ascertained that the tree constructed is a game-tree in accordance with the interpretation of strings as actions and the strategy profiles as described above. Thus, for each internal node σ of $\boldsymbol{\mathfrak{T}}_\Gamma$ and each strategy label in S of $L(A, B)$, we introduce a unique leaf σs^n, for some appropriate integer $n \in \omega$, serving as the outcome of the strategy profile the label s stands for when play is commenced in σ in the model $\mathfrak{M}_\Gamma$. Meanwhile, the tree $\boldsymbol{\mathfrak{T}}_\Gamma$ should also be guaranteed to have a finite horizon.

For each t in $\{\epsilon\} \cup T$, each theory Γ in $L(A, B)$ and each $n \in \omega$, we first define inductively a tree $\mathfrak{T}^n_{\Gamma,t}$ (no boldface!). The vertices of $\mathfrak{T}^n_{\Gamma,t}$ are decorated with theories in $L(A, B)$. Then the tree $\mathfrak{T}_{\Gamma,t}$ is defined as the limit of this induction. This latter kind of labelled tree will form the modules which eventually compose the tree $\boldsymbol{\mathfrak{T}}_\Gamma$ for Γ. The set of vertices of $\mathfrak{T}_{\Gamma,t}$ we denote by $\Sigma_{\Gamma,t}$ and the labelling function assigning theories of $L(A, B)$ to the vertices in $\Sigma_{\Gamma,t}$ by $\theta_{\Gamma,t}$. The root of $\mathfrak{T}_{\Gamma,t}$ is taken to be t and the other vertices and the theories assigned to them are chosen in accordance with the idea that Γ be eventually satisfiable at t.

At the basis of the induction, the tree $\mathfrak{T}^0_{\Gamma,t}$ is defined, with $\Sigma^0_{\Gamma,t}$ as vertices and $\theta^0_{\Gamma,t}$ as labelling function. The idea is that $\mathfrak{T}^0_{\Gamma,t}$ contain, for each strategy label s in S, a *unique* leaf that can be taken as the outcome of strategy profile corresponding to s when play is commenced in the root t. The design is such that along any such path each of the players in N is to move once. In general, player i is assumed to move at ts^i. In particular, the mystery player 0 makes a decision at the root t. Hence, $\Sigma^0_{\Gamma,t}$ contains t as well as each sequence ts^n with $n \leqslant \|N\| + 1$, the idea being that each player $i \in N$ is to move at ts^i and that $ts^{\|N\|+1}$ be the outcome s determines in t. Moreover, each strategy profile should prescribe a move at each internal node. Hence, for each ts^i with $i \leqslant \|N\|$ and each label s' in S different from s we also distinguish a leaf ts^is' in $\Sigma^0_{\Gamma,t}$. No further vertices are in $\Sigma^0_{\Gamma,t}$.

Figure 4.2 depicts the tree for $\Sigma^0_{\Gamma,t}$ in a language containing two labels for players and also two labels for (the outcome functions of) strategy profiles. The labelling function $\theta^0_{\Gamma,t}$ assigns the maximal Λ-consistent extension $Cl(\Gamma)$ of Γ to the root t and the empty theory to each of the internal vertices. However, arbitrary the latter may seem, it will prove to be convenient as the proof develops. If now the theory $Cl(\Gamma)$ is supposed to be satisfied at t in some game-model on $\Sigma^0_{\Gamma,t}$, then each leaf $ts^{\|N\|+1}$ should satisfy a formula ψ whenever $[\hat{s}_\emptyset]\psi$ is in $Cl(\Gamma)$. Similarly, if $Cl(\Gamma)$ contains a formula $[\hat{s}_i]\psi$, then each leaf ts^is' should satisfy ψ. The assignment function $\theta^0_{\Gamma,t}$ is

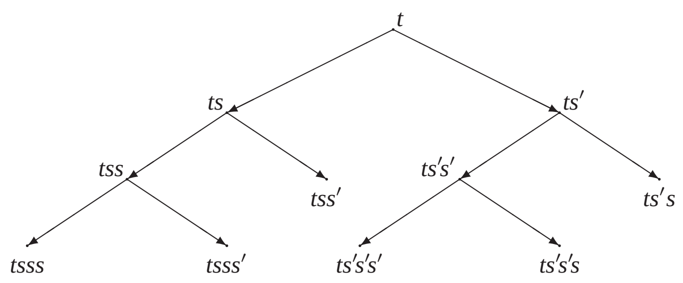


Figure 4.2. The tree $\Sigma^0_{\Gamma,t}$ for language $L(A,B)$ with $N = \{1,2\}$ and $S = \{s, s'\}$.

fixed accordingly.

The role of the mystery player can now also be revealed. Suppose t had been assigned to a player i^* with a label in N. Assume further that i^* be different from 1. Now consider strategy profiles s and s' with labels in S. Then, the leaf tss' would be reachable from t *via* the accessibility relation for the strategy label $\hat{s}_1$, as intended. However, the vertex tss' would also be reachable from the t by way of the accessibility relation for $\hat{s}'_{i^*}$! If so, the theory associated with tss' should then contain all formulas ψ for which $[\hat{s}_1]\psi$ is in Γ as well as all formulas ψ such that $[\hat{s}'_i]\psi$ is in Γ. The Λ-consistency of a such a theory, however, cannot in general be guaranteed. By assigning a player without a label in N to the root t this contingency does not occur.

Formally we define:

$$\mathfrak{T}^0_{\Gamma,t} \quad =_{df.} \quad \left(\Sigma^0_{\Gamma,t}, \theta^0_{\Gamma,t}\right),$$

where $\Sigma^0_{\Gamma,t}$ is a subset of $\boldsymbol{\Sigma}$ and $\theta^0_{\Gamma,t}$ a function assigning theories in $L(A,B)$ to the sequences σ in $\Sigma^0_{\Gamma,t}$ such that:

$$\Sigma^0_{\Gamma,t} \quad =_{df.} \quad \left\{t\right\} \cup \left\{ts^i s' : \ s, s' \in S \ \text{ and } \ 0 \leqslant i \leqslant \|N\| \right\}$$

$$\theta^0_{\Gamma,t}(\sigma) \quad =_{df.} \quad \begin{cases} Cl(\Gamma) & \text{if } \sigma = t \\ Cl\left(\left\{\psi : \ [\hat{s}_\emptyset]\psi \in Cl(\Gamma)\right\}\right) & \text{if } \sigma = ts^{\|N\|+1} \\ Cl\left(\left\{\psi : \ [\hat{s}_i]\psi \in Cl(\Gamma)\right\}\right) & \text{if } \sigma = ts^i s' \text{ and } s \neq s' \\ \emptyset & \text{otherwise.} \end{cases}$$

In the inductive step, defining $\mathfrak{T}^{n+1}_{\Gamma,t}$ from $\mathfrak{T}^n_{\Gamma,t}$, we check whether φ_n, *i.e.*, the $n-1$-st formula in the enumeration, is of the form $\neg[\hat{s}_i]\psi$. If it is and φ_n moreover occurs in $Cl(\Gamma)$, the string $ts^i n$ is added to the set of vertices and assigned the maximal Λ-extension of $\left\{\neg\psi\right\} \cup \left\{\chi : \ [\hat{s}_i]\chi \in \Gamma\right\}$. In any other case $\mathfrak{T}^{n+1}_{\Gamma,t}$ and $\mathfrak{T}^n_{\Gamma,t}$ are identical.

The idea behind this construction is that, if the root t is to satisfy a formula of the form $\neg[\hat{s}_i]\psi$, then a leaf v' should be reachable from t *via* the accessibility relation $R_{\hat{s}_i}$

and not satisfy ψ. Being reachable thus, the leaf v' should, in addition, also satisfy any formula χ such that t forces $[\hat{s}_i]\chi$. Note that a similar construction is unnecessary if φ_n is of the form $\neg[\hat{s}_\emptyset]\psi$. In virtue of Axiom $D!_{\hat{s}_\emptyset}$, the formula $\neg[\hat{s}_\emptyset]\psi$ is equivalent to $[\hat{s}_\emptyset]\neg\psi$ and the latter is thus an element of $Cl\,(\Gamma)$. Hence, this case has already been taken care of by the construction of $\mathfrak{T}^0_{\Gamma,t}$. A similar remark applies to the case in which a formula of the form $\neg[\hat{s}_\emptyset]\psi$ or $\neg[\hat{s}_i]\psi$ is contained in the theory assigned to a leaf σ in $\Sigma^n_{\Gamma,t}$ by $\theta^n_{\Gamma,t}$. Then, Axiom $E5_{\hat{s}_X,\hat{s}_Y}$ makes that $\neg\psi$ is already in the theory $\theta^n_{\Gamma,t}\,(\sigma)$. The theories assigned to the internal vertices contain no formulas, let alone formulas that require "*witness*" states. Formally define:

$$\mathfrak{T}^{n+1}_{\Gamma,t} \quad =_{df.} \quad \left(\Sigma^{n+1}_{\Gamma,t}, \theta^{n+1}_{\Gamma,t}\right),$$

where $\Sigma^{n+1}_{\Gamma,t}$ is a superset of the sequences in $\Sigma^n_{\Gamma,t}$ and $\theta^{n+1}_{\Gamma,t}$ is a function extending $\theta^n_{\Gamma,t}$ by associating with each sequence in $\Sigma^{n+1}_{\Gamma,t}$ a theory in $L(A,B)$:

$$\Sigma^{n+1}_{\Gamma,t} \quad =_{df.} \quad \begin{cases} \Sigma^n_{\Gamma,t} \cup \left\{ts^i n\right\} & \text{if } \varphi_n = \neg[\hat{s}_i]\psi \text{ and } \varphi_n \in Cl\,(\Gamma) \\ \Sigma^n_{\Gamma,t} & \text{otherwise.} \end{cases}$$

$$\theta^{n+1}_{\Gamma,t}\,(\sigma) \quad =_{df.} \quad \begin{cases} Cl\left(\left\{\neg\psi\right\} \cup \left\{\chi :\ [\hat{s}_i]\chi \in Cl\,(\Gamma)\,\right\}\right) & \text{if } \sigma = ts^i n, \varphi_n = \neg[\hat{s}_i]\psi \\ & \text{and } \varphi_n \in Cl\,(\Gamma) \\ \theta^n_{\Gamma,t}\,(\sigma) & \text{otherwise, } \textit{i.e.}, \text{ if } \sigma \in \Sigma^n_{\Gamma,t}. \end{cases}$$

Finally define $\mathfrak{T}_{\Gamma,t}$ as:

$$\mathfrak{T}_{\Gamma,t} \quad =_{df.} \quad \left(\bigcup_{n\in\omega} \Sigma^n_{\Gamma,t}\,,\ \bigcup_{n\in\omega} \theta^n_{\Gamma,t}\right).$$

The theories $\theta_{\Gamma,t}$ assigns to the vertices may also contain formulas of the form $\neg[i]\psi$. If this is the case for a vertex σ, then the construction should also contain a vertex associated with a maximal Λ-consistent theory containing $\neg\psi$ as well as any formula χ if $[i]\chi$ is in the theory associated with σ. To accommodate this type of formula, we push the construction one step further.

We now define inductively for each $n \in \omega$ a *collection of decorated trees of the form* $\mathfrak{T}_{\Gamma,t}$, from which we eventually manufacture the tree $\boldsymbol{\mathfrak{T}}_\Gamma$. At the basis, this collection consists of the tree $\mathfrak{T}_{\Gamma,\epsilon}$, which has Γ itself associated with its root ϵ. The empty sequence ϵ, being a prefix to any sequence, will also be the root of the tree to be constructed and, eventually, also of the game-model $\mathfrak{M}_\Gamma$. If any of the vertices σ of $\mathfrak{T}_{\Gamma,\epsilon}$ contains a formula φ_k of the form $\neg[i]\psi$, a new tree $\mathfrak{T}_{\Theta,t^\sigma_k}$ is added to the collection, on the understanding that the theory Θ equals $\left\{\neg\psi\right\} \cup \left\{\chi :\ [i]\chi \in \theta\,(\sigma)\,\right\}$. Then this process is repeated for the new collection, and so on. Thus define:

$$\begin{aligned} \boldsymbol{\mathfrak{T}}^0_\Gamma \quad &=_{df.} \quad \left\{\mathfrak{T}_{\Gamma,\epsilon}\right\} \\ \boldsymbol{\mathfrak{T}}^{n+1}_\Gamma \quad &=_{df.} \quad \bigcup_{(\Sigma,\theta)\in\boldsymbol{\mathfrak{T}}^n_\Gamma} \left\{\mathfrak{T}_{\Theta,t^\sigma_k} :\ \sigma \in \Sigma \text{ and } \varphi_k = \neg[i]\psi \text{ and } \varphi_k \in \theta\,(\sigma)\right\}, \end{aligned}$$

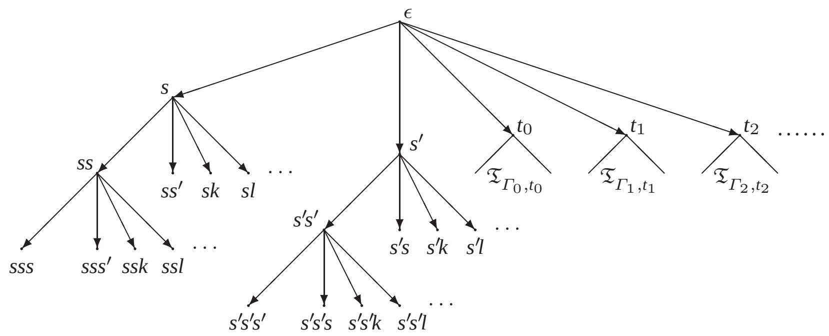


Figure 4.3. The tree $\mathfrak{T}_\Gamma$ for a theory Γ in a language $L(A,B)$ with $N=\{1,2\}$ and $S=\{s,s'\}$. Each subtree $\mathfrak{T}_{\Gamma_k,t_k}$ is introduced in virtue of a formula of the form $\neg[i]\psi$ being in the theory assigned to a vertex in another part of the tree.

where $\Theta = \{\neg\psi\} \cup \{\chi : \ [i]\chi \in \theta(\sigma)\}$. Then set:

$$\mathfrak{T}^\omega_\Gamma \ =_{df.} \ \bigcup_{n\in\omega} \mathfrak{T}^n_\Gamma .$$

We are now in a position to formally define the decorated tree $\mathfrak{T}_\Gamma$ as:

$$\mathfrak{T}_\Gamma \ =_{df.} \ \left(\boldsymbol{\Sigma}_\Gamma, \boldsymbol{\theta}_\Gamma\right),$$

where:

$$\begin{aligned} \boldsymbol{\Sigma}_\Gamma &=_{df.} \bigcup \left\{\Sigma : \ (\Sigma,\theta) \in \mathfrak{T}^\omega_\Gamma\right\} \\ \boldsymbol{\theta}_\Gamma &=_{df.} \bigcup \left\{\theta : \ (\Sigma,\theta) \in \mathfrak{T}^\omega_\Gamma\right\}. \end{aligned}$$

The tree $\mathfrak{T}_\Gamma$ having been defined thus for each theory Γ in $L(A,B)$, we are now almost in a position to define on its basis the game $\boldsymbol{E}_\Gamma$. First, however, some preliminary remarks and auxiliary results are in order. It should be checked that for each Λ-consistent theory Γ the tree $\mathfrak{T}_\Gamma$ is in fact a tree with its vertices $\boldsymbol{\Sigma}_\Gamma$ ordered by the immediate prefix relation $\prec$. Moreover, $\boldsymbol{\theta}_\Gamma$ should be ascertained to be a *function*. This is established by Fact 4.3.4. Since the construction of $\mathfrak{T}_\Gamma$ invokes the closure operator *Cl*, the function $\boldsymbol{\theta}_\Gamma$ should be guaranteed to assign a consistent theory — either empty or maximal consistent — to each vertex σ in $\boldsymbol{\theta}_\Gamma$. Fact 4.3.5 demonstrates that $\boldsymbol{\theta}_\Gamma$ is properly defined in this sense.

At each stage of the inductive construction of $\mathfrak{T}^\omega_\Gamma$ the set of vertices of each tree added to the collection is disjoint from any other set of vertices of a tree introduced thus, as well as from any set of vertices of a tree already present in the collection. Hence, the domains of the various functions θ remain separate as well. This observation is laid down formally in Fact 4.3.3. First we prove the following lemma; Figure 4.4 supports the underlying intuitive idea.

Lemma 4.3.2 *Let A be a countably infinite set. Let π be a countably infinite partition of A, i.e., $\pi \in Part\,(A)$, such that each $\pi_i \in \pi$ is countably infinite as well. The elements of each block π_i of π are enumerated as $a_0^{\pi_i}, \ldots, a_n^{\pi_i}, \ldots$. Let π^0 be a block in π and let f be an injective function mapping A on π, i.e., $f : A \to \pi$, such that π^0 is not in the image of A under f, i.e., $\pi^0 \notin f\,(A)$. For each $\sigma \in \omega^*$, let π_σ denote a block in π inductively as follows:*

$$\begin{aligned} \pi_\epsilon &=_{df.} \pi^0 \\ \pi_{\sigma n} &=_{df.} f\left(a_n^{\pi_\sigma}\right). \end{aligned}$$

Then, π_σ and $\pi_{\sigma'}$ are disjoint, for all distinct strings σ and σ' in ω^*.

Proof: By induction on $|\sigma|$, the length of σ. First assume $|\sigma| = 0$, *i.e.*, $\sigma = \epsilon$. Since σ' is distinct from σ there is some $\sigma'' \in \omega^*$ and some $k \in \omega$ such that $\sigma' = \sigma'' k$. Hence, $\pi_{\sigma'} = f\left(a_k^{\pi_{\sigma''}}\right)$ with $a_k^{\pi_{\sigma''}} \in \pi_{\sigma''}$. Then:

$$\pi_{\sigma'} \;=\; f\left(a_k^{\pi_{\sigma''}}\right) \;\neq\; \pi^0 \;=\; \pi_\epsilon \;=\; \pi_\sigma.$$

Now let $|\sigma| = n + 1$. Then $\sigma = \sigma''' l$ for some $\sigma''' \in \omega^*$ and some $l \in \omega$. This case is by induction on the length $|\sigma'|$ of σ'. If $|\sigma'| = 0$, the reasoning is analogous to the case in which $|\sigma| = 0$, above. So assume $|\sigma'| = m + 1$. Then, $\sigma = \sigma'' k$ for some $\sigma'' \in \omega^*$ and some $k \in \omega$. Then, both $\pi_\sigma = f\left(a_l^{\pi_{\sigma'''}}\right)$ and $\pi_{\sigma'} = f\left(a_k^{\pi_{\sigma''}}\right)$.

Either $\sigma'' = \sigma'''$ or $\sigma'' \neq \sigma'''$. If $\sigma'' = \sigma'''$, then $k \neq l$, since σ and σ' had been assumed to be distinct. Then, $a_l^{\pi_{\sigma'''}} \neq a_k^{\pi_{\sigma''}}$. If, on the other hand, $\sigma'' \neq \sigma'''$, then, by the induction hypothesis, $\pi_{\sigma'''}$ and $\pi_{\sigma''}$ may be assumed to be disjoint. With $a_l^{\pi_{\sigma'''}} \in \pi_{\sigma'''}$ and $a_k^{\pi_{\sigma''}} \in \pi_{\sigma''}$, again $a_l^{\pi_{\sigma'''}} \neq a_k^{\pi_{\sigma''}}$.

By injectivity of f, it follows that $\pi_\sigma = f\left(a_l^{\pi_{\sigma'''}}\right) \neq f\left(a_k^{\pi_{\sigma''}}\right) = \pi_{\sigma'}$. With π_σ and $\pi_{\sigma'}$ being blocks in the partition π, they are disjoint as well. $\dashv$

Fact 4.3.3 *Let (Σ, θ) and (Σ', θ') be distinct trees in $\mathfrak{T}_T^\omega$. Then, Σ and Σ' are disjoint, i.e., $\Sigma \cap \Sigma' = \emptyset$.*

Proof: Consider Cartesian product $\boldsymbol{\Sigma} \times \omega$. Define for each x in $S \cup T \cup \omega$ a subset Σ_x of $\boldsymbol{\Sigma} \times \omega$ as follows:

$$\Sigma_x \;=_{df.}\; \left\{\sigma : \text{ for some } \sigma' \in \boldsymbol{\Sigma},\; x\sigma' = \sigma\right\} \times \omega.$$

Obviously for distinct x and x' the sets Σ_x and $\Sigma_{x'}$ are disjoint. Moreover, if Σ_x and $\Sigma_{x'}$ are disjoint, so are $\bigcup_{n\in\omega} \left\{\sigma \in \boldsymbol{\Sigma} : \; (\sigma, n) \in \Sigma_x\right\}$ and $\bigcup_{n\in\omega} \left\{\sigma \in \boldsymbol{\Sigma} : \; (\sigma, n) \in \Sigma_{x'}\right\}$. Let Σ_ϵ be given by $\overline{\bigcup_{t\in T} \Sigma_t}$. Hence, $\{\Sigma_\epsilon\} \cup \{\Sigma_x : \; x \in T\}$ is a partition of $\boldsymbol{\Sigma} \times \omega$. Observe that for $x \in T \cup \{\epsilon\}$ and $\mathfrak{T}_{\Theta,x}$ given by (Σ, θ), we have $\Sigma \subseteq \bigcup_{n\in\omega} \left\{\sigma \in \boldsymbol{\Sigma} : (\sigma, n) \in \Sigma_x\right\}$. Let π be a partition of $\boldsymbol{\Sigma} \times \omega$ given by $\left\{\Sigma_x : \; x \in T\right\} \cup \left\{\Sigma_\epsilon\right\}$. Now define the function f mapping $\boldsymbol{\Sigma} \times \omega$ on π, such that for all $(\sigma, n) \in \boldsymbol{\Sigma} \times \omega$:

$$f\,(\sigma, n) \;=_{df.}\; \Sigma_{t_n^\sigma}$$

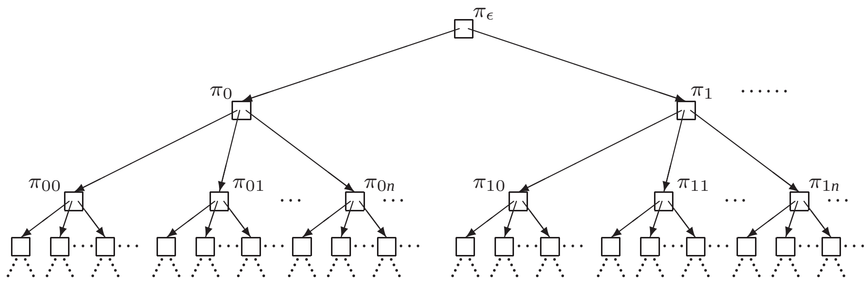


Figure 4.4. The development of the blocks as in Lemma 4.3.2. Each box represents a block π_σ of the partition π of A. The arrows indicate the function f, which maps elements of the blocks on blocks in π. By Lemma 4.3.2 these blocks are all pairwise disjoint.

On account of the assumptions made with respect to T and the definition of T_σ on page 92, the function f is clearly injective. Moreover, Σ_ϵ is not in the image of $\boldsymbol{\Sigma} \times \omega$ under f.

Assume each Σ_x to be enumerated in a particular way. For strings $\sigma \in \omega^*$, let π_σ denote a block in the partition π as defined in Lemma 4.3.2 *via* f with $\pi_\epsilon = \Sigma_\epsilon$. In virtue of Lemma 4.3.2, it then suffices to show that for all $n \in \omega$ and for each tree $\mathfrak{T}_{\Theta,x}$ in $\boldsymbol{\mathfrak{T}}^n_\Gamma$, there is some $\tau \in \omega^*$ such that $\pi_\tau = \Sigma_x$. The proof is by induction on $n \in \omega$. The basis is immediate as $\boldsymbol{\mathfrak{T}}^0_\Gamma = \left\{\mathfrak{T}_{\Gamma,\epsilon}\right\}$ and $\Sigma_\epsilon = \pi_\epsilon$. Now assume that $\mathfrak{T}_{\Theta,x} \in \boldsymbol{\mathfrak{T}}^{n+1}_\Gamma$. Then, $x = t^\sigma_k$, for some $\mathfrak{T}_{\Theta',x'} = (\Sigma, \theta) \in \boldsymbol{\mathfrak{T}}^n_\Gamma$, some $\sigma \in \Sigma$ and some $k \in \omega$. By the induction hypothesis there is some $\tau \in \omega^*$ such that $\pi_\tau = \Sigma_{x'}$. Let (σ, k) be the $m + 1$-st element in the enumeration of $\Sigma_{x'}$. Then $\Sigma_{t^\sigma_k} = \pi_{\sigma m}$, which concludes the proof. $\dashv$

Fact 4.3.4 *For each Λ-consistent theory Γ, the set $\boldsymbol{\Sigma}_\Gamma$ is a tree if ordered by the immediate prefix relation $\prec$, and $\boldsymbol{\theta}_\Gamma$ is a total function on $\boldsymbol{\Sigma}_\Gamma$.*

Sketch of proof: We first prove by that $\Sigma_{\Gamma,t}$ is a tree ordered by $\prec$, for all $t \in \{\epsilon\} \cup T$ and all theories Γ. For observe that $\Sigma^0_{\Gamma,x}$ is a tree ordered by $\prec$ by definition. Then observe that $\Sigma^{n+1}_{\Gamma,t}$ is obtained from $\Sigma^n_{\Gamma,t}$ by adding at most a fresh vertex ts^in. With ts^i already contained in $\Sigma^0_{\Gamma,t}$, we may by the induction hypothesis conclude that $\Sigma^{n+1}_{\Gamma,t}$ is a tree ordered by $\prec$ as well. It follows that $\Sigma_{\Gamma,t}$ is a tree ordered by $\prec$; otherwise there would be a smallest $n \in \omega$ such that $\Sigma^n_{\Gamma,t}$ is not a tree, *quod non*. Now consider $\boldsymbol{\mathfrak{T}}^\omega_\Gamma$. Observe that each tree $\mathfrak{T}_{\Theta,t}$ has t as root and that ϵ is an immediate prefix of t. This makes that $\boldsymbol{\Sigma}_\Gamma$ is a tree ordered by $\prec$. A similar argument shows that θ is functional for each (Σ, θ) in $\boldsymbol{\mathfrak{T}}^\omega_\Gamma$. By Fact 4.3.3 for any two (Σ, θ) and (Σ', θ') in $\boldsymbol{\mathfrak{T}}^\omega_\Gamma$, the sets Σ

and Σ' are disjoint. With each θ, for each $(\Sigma, \theta) \in \mathfrak{T}^{\omega}_{\Gamma}$, being defined precisely on Σ, we may conclude that $\boldsymbol{\theta}_{\Gamma}$, *i.e.*, $\bigcup \{\theta : \ (\Sigma, \theta) \in \mathfrak{T}^{\omega}_{\Gamma}\}$, is functional as well. $\dashv$

Figure 4.3 depicts the structure of $\mathfrak{T}_{\Gamma}$ for a language $L(A, B)$ with $N = \{1, 2\}$ and $S = \{s, s'\}$.

For technical convenience we distinguish particular subsets of vertices in $\mathfrak{T}_{\Gamma}$. First, the root node ϵ and the vertices t^{σ}_{n} that are the roots of subtrees $\mathfrak{T}_{\Gamma, t^{\sigma}_{n}}$ in $\mathfrak{T}_{\Gamma}$ are collected in $\boldsymbol{T}_{\Gamma}$, *i.e.*,

$$\boldsymbol{T}_{\Gamma} \quad =_{df.} \quad \{\epsilon\} \cup (T \cap \boldsymbol{\Sigma}_{\Gamma})\,.$$

For each t in $\boldsymbol{T}_{\Gamma}$ the set of internal vertices and the set of leaves of the respective subtree $\mathfrak{T}_{\Theta, t}$ of $\mathfrak{T}_{\Gamma}$ are denoted by $\boldsymbol{I}^{t}_{\Gamma}$ and $\boldsymbol{L}^{t}_{\Gamma}$, respectively. Obviously, $\boldsymbol{I}^{t}_{\Gamma}$ and $\boldsymbol{L}^{t}_{\Gamma}$ are disjoint and together exhaust the vertices in $\mathfrak{T}_{\Gamma', t}$. Finally, $\boldsymbol{W}^{t}_{\Gamma}$ comprises t together with the leaves $\boldsymbol{L}^{t}_{\Gamma}$. Some reflection reveals that $\boldsymbol{W}^{t}_{\Gamma}$ are exactly those vertices in $\mathfrak{T}_{\Gamma', t}$ which are labelled with a (non-empty) maximal consistent theory. Formally, for $\mathfrak{T}_{\Gamma}$ given by $(\boldsymbol{\Sigma}_{\Gamma}, \boldsymbol{\theta}_{\Gamma})$, define:

$$\begin{aligned}
\boldsymbol{I}^{t}_{\Gamma} \quad &=_{df.} \quad \{ts^{i} \in \boldsymbol{\Sigma}_{\Gamma} : \ s \in S \text{ and } i \leqslant \|N\|\},\\
\boldsymbol{L}^{t}_{\Gamma} \quad &=_{df.} \quad \{ts^{\|N\|+1}, ts^{i}s', ts^{i}n \in \boldsymbol{\Sigma}_{\Gamma} : \ n \in \omega \text{ and } s, s' \in S \text{ such that } s \neq s'\},\\
\boldsymbol{W}^{t}_{\Gamma} \quad &=_{df.} \quad \{t\} \cup \boldsymbol{L}^{t}_{\Gamma}.
\end{aligned}$$

On this basis also define:

$$\boldsymbol{I}_{\Gamma} \ =_{df.} \bigcup_{t \in \boldsymbol{T}_{\Gamma}} \boldsymbol{I}^{t}_{\Gamma}, \qquad \boldsymbol{L}_{\Gamma} \ =_{df.} \bigcup_{t \in \boldsymbol{T}_{\Gamma}} \boldsymbol{L}^{t}_{\Gamma}, \qquad \boldsymbol{W}_{\Gamma} \ =_{df.} \bigcup_{t \in \boldsymbol{T}_{\Gamma}} \boldsymbol{W}^{t}_{\Gamma}.$$

The internal vertices of $\mathfrak{T}_{\Gamma}$ are collected in $\boldsymbol{I}_{\Gamma}$. The set $\boldsymbol{W}_{\Gamma}$ contains precisely those vertices labelled with maximal consistent theories. Moreover, $\boldsymbol{L}_{\Gamma}$ contains the leaves of $\mathfrak{T}_{\Gamma}$. Finally, as some reflection reveals, we have that:

$$\boldsymbol{\Sigma}_{\Gamma} \quad = \quad \boldsymbol{I}_{\Gamma} \cup \boldsymbol{L}_{\Gamma}.$$

How the various "*types*" of sequence in the collection $\boldsymbol{\Sigma}_{\Gamma}$ relate to one another is illustrated in Figure 4.5.

Fact 4.3.5 *Let Γ be a Λ-consistent theory in $L(A, B)$. Let further $\mathfrak{T}_{\Gamma}$ be given by $(\boldsymbol{\Sigma}_{\Gamma}, \boldsymbol{\theta}_{\Gamma})$ and let $\sigma \in \boldsymbol{\Sigma}_{\Gamma}$. Then, $\boldsymbol{\theta}_{\Gamma}(\sigma)$ is a maximal Λ-consistent theory, if σ is in $\boldsymbol{W}_{\Gamma}$, and the empty theory, otherwise.*

Sketch of proof: By definition of each tree $\mathfrak{T}_{\Theta, x}$ given by (Σ, θ) we find that θ assigns the empty theory to the internal vertices other than the root (*cf.*, pages 94–94). Some reflection reveals that this makes that $\boldsymbol{\theta}_{\Gamma}$ assigns the empty theory to each vertex in $\boldsymbol{I}_{\Gamma}$ of $\mathfrak{T}_{\Gamma}$.

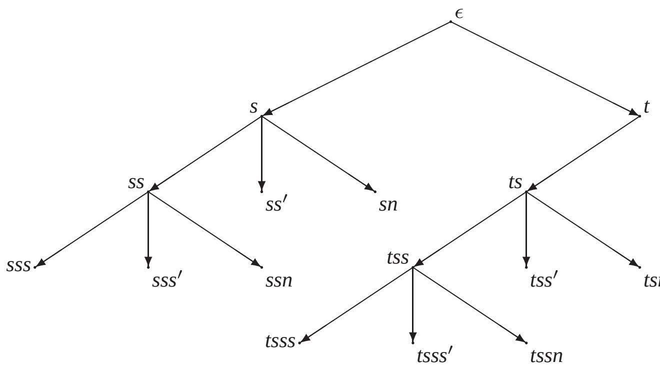


Figure 4.5. Types of vertex in $\mathfrak{T}_\Gamma$. The set $\boldsymbol{T}$ contains ϵ and (vertices of type) t. The vertices ϵ, s and ss make up $\boldsymbol{I}^\epsilon$, whereas t, ts and tss are the sole elements in $\boldsymbol{I}^t$. $\boldsymbol{L}^t$ is given by $\{tss', tsl, tsss, tsss', tssl\}$ and $\boldsymbol{W}$ contains all vertices except s, ss, ts and tss.

By definition, the root ϵ of $\mathfrak{T}_\Gamma$ is assigned $Cl\,(\Gamma)$, the maximum Λ-consistent closure of Γ. Assuming that Γ is Λ-consistent, so is $Cl\,(\Gamma)$. For the other possible cases, it suffices the prove the following equivalences:

$$
\begin{aligned}
\Gamma \nvdash \bot &\quad \text{implies} \quad \{\psi :\ [\hat{s}_i]\psi \in \Gamma\} \nvdash \bot,\\
\Gamma \nvdash \bot &\quad \text{implies} \quad \{\psi :\ [\hat{s}_\varnothing]\psi \in \Gamma\} \nvdash \bot,\\
\Gamma \cup \{\neg[i]\chi\} \nvdash \bot &\quad \text{implies} \quad \{\neg\chi\} \cup \{\psi :\ [i]\psi \in \Gamma\} \nvdash \bot,\\
\Gamma \cup \{\neg[\hat{s}_i]\chi\} \nvdash \bot &\quad \text{implies} \quad \{\neg\chi\} \cup \{\psi :\ [\hat{s}_i]\psi \in \Gamma\} \nvdash \bot.
\end{aligned}
$$

Each of these implications hold for any extensive game logic Λ. The logic Λ being a normal logic, the latter two can be proved by a standard argument (*cf.*, Blackburn, de Rijke, and Venema (2001), pp.198-9). The argument for the first and the second item is similar, although it essentially involves the axioms $D!_{\beta_1}$ and $E1_{(\beta_0,\beta_1),\beta_1}$. First assume the contrapositive that $\{\psi :\ [\hat{s}_i]\psi \in \Gamma\} \vdash \bot$. Then there is a finite number of formulas $\psi_0, \ldots, \psi_n \in \{\psi :\ [\hat{s}_i]\psi \in \Gamma\}$ such that $\psi_0, \ldots, \psi_n \vdash \bot$. Consider the following implications:

$$
\begin{aligned}
&\psi_0, \ldots, \psi_n \vdash \bot \quad \text{implies} \quad [\hat{s}_i]\psi_0, \ldots, [\hat{s}_i]\psi_n \vdash [\hat{s}_i]\bot \quad \text{implies} \quad \Gamma \vdash [\hat{s}_i]\bot\\
&\qquad \text{implies}_{\,E1_{\hat{s}_i,\hat{s}_\varnothing}} \quad \Gamma \vdash [\hat{s}_\varnothing]\bot \quad \text{implies}_{\,D!_{\hat{s}_\varnothing}} \quad \Gamma \vdash \langle\hat{s}_\varnothing\rangle\bot \quad \text{implies} \quad \Gamma \vdash \bot.
\end{aligned}
$$

The first and last implication are in virtue of Λ being a normal modal logic. The argument for the second item runs along analogous lines. From $\psi_0, \ldots, \psi_n \vdash \bot$ obtain $[\hat{s}_\varnothing]\psi_0, \ldots, [\hat{s}_\varnothing]\psi_n \vdash [\hat{s}_\varnothing]\bot$, which implies $\Gamma \vdash \langle\hat{s}_\varnothing\rangle\bot$ in virtue of $D!_{\hat{s}_\varnothing}$. Then, finally,

$\Gamma \vdash \bot$. This concludes the proof. $\dashv$

In the sequel we also find the following convention and fact useful of great usefulness. For each *leaf* σ in $\boldsymbol{L}_\Gamma$ and x in $\boldsymbol{T}_\Gamma$, we have β^x_σ denote a strategy label of $L(A,B)$ as follows:

$$\beta^x_\sigma \quad =_{df.} \quad \begin{cases} \hat{s}_i & \sigma = xs^i y \text{ for } i \in N,\, s \in S \text{ and } y \in \omega \cup (S - \{s\}). \\ \hat{s}_\emptyset & \text{otherwise, } i.e., \text{ if } \sigma = xs^{\|N\|+1}, \text{ for } s \in S. \end{cases}$$

We now have the following fact.

Fact 4.3.6 *Consider $\mathfrak{T}_\Gamma$ for Γ a Λ-consistent theory in $L(A,B)$. Let $x \in \boldsymbol{T}_\Gamma$ and σ a leaf in $\boldsymbol{L}^x_\Gamma$. Then:*

$$\left\{\varphi : \ [\beta^x_\sigma]\varphi \in \boldsymbol{\theta}_\Gamma(x)\right\} \ \subseteq \ \boldsymbol{\theta}_\Gamma(\sigma)\,.$$

Proof: Either $\sigma = xs^{\|N\|+1}$, $\sigma = xs^i s'$ or $\sigma = xs^i n$, for some $n \in \omega$ and some $s, s' \in S$, such that $s \neq s'$. In the first case, $\beta^x_\sigma = \hat{s}_\emptyset$. Assume for an arbitrary formula φ of $L(A,B)$ that $[\hat{s}]\varphi \in \boldsymbol{\theta}_\Gamma(x)$. By Fact 4.3.3, then $[\hat{s}_\emptyset]\varphi \in \theta_{\Theta,x}(x)$, where theory Θ such that $\mathfrak{T}_{\Theta,x}$, given by (Σ,θ), is included in $\mathfrak{T}^\omega_\Gamma$. Then, $[\hat{s}_\emptyset]\varphi \in Cl(\Theta)$. Consequently, $\varphi \in \theta_{\Theta,x}\left(xs^{\|N\|+1}\right)$ and by Fact 4.3.3 and with $\theta_{\Theta,x} \subseteq \boldsymbol{\theta}_\Gamma$, eventually, $\varphi \in \boldsymbol{\theta}_\Gamma\left(xs^{\|N\|+1}\right)$. The reasoning for the other two cases runs along analogous lines. $\dashv$

We are now in a position to define for each Λ-consistent theory Γ an extensive game $\boldsymbol{E}_\Gamma$ on basis of the labelled tree $\mathfrak{T}_\Gamma$.

Definition 4.3.7 *(Extensive games for Λ-consistent theories)* Let Λ be an extensive game logic in a multi-modal matrix language $L(A,B)$. Recall that B is given by $N \cup S \cup (N \times S)$ and N by $\{0,\ldots,n\}$ for some $n \in \omega$. Let Γ be a maximal Λ-consistent theory. We define: the extensive game $\boldsymbol{E}^\Lambda_\Gamma$ as:

$$\boldsymbol{E}^\Lambda_\Gamma \quad =_{df.} \quad \left(\boldsymbol{V}_\Gamma, \boldsymbol{R}_\Gamma, \boldsymbol{N}_\Gamma, \boldsymbol{P}_\Gamma, \{\boldsymbol{\rho}_i\}_{i\in \boldsymbol{N}_\Gamma}\right),$$

where $\boldsymbol{V}_\Gamma$ is defined as the set of sequences $\boldsymbol{\Sigma}_\Gamma$ in $\mathfrak{T}_\Gamma$ as above and $\boldsymbol{R}_\Gamma$ is given by the immediate prefix relation on $\boldsymbol{\Sigma}_\Gamma$, *i.e.*, for all $\sigma, \sigma' \in \boldsymbol{V}_\Gamma$:

$$\sigma \boldsymbol{R}_\Gamma \sigma' \quad \text{iff} \quad \text{for some } x \in S \cup T \cup \omega : \quad \sigma' = \sigma x.$$

The players of the game $\boldsymbol{E}_\Gamma$ are given by the player labels N of $L(A,B)$ together with a *mystery player* denoted by 0, *i.e.*,

$$\boldsymbol{N}_\Gamma \quad =_{df.} \quad N \cup \{0\}\,.$$

The player assignment function $\boldsymbol{P}_\Gamma$ is such that for each internal vertex σ of $\boldsymbol{V}_\Gamma$, *i.e.*, for each i in $\boldsymbol{N}_\Gamma$:

$$\boldsymbol{P}_\Gamma(\sigma) = i \quad \text{iff} \quad \sigma = xs^i, \text{ for some } x \in \boldsymbol{T}_\Gamma \text{ and some strategy label } s \in S.$$

Finally, the preferences of each player i in N are such that for all vertices $\sigma, \sigma' \in \boldsymbol{V}_\Gamma$:

$$(\sigma, \sigma') \in \boldsymbol{\rho}_i \quad \text{iff} \quad \text{for all formulas } \varphi\text{:} \quad [i]\varphi \in \boldsymbol{\theta}_\Gamma(\sigma) \ \text{ implies } \ \varphi \in \boldsymbol{\theta}_\Gamma(\sigma').$$

The mystery player 0 we assume to be entirely indifferent, *i.e.*, $\boldsymbol{\rho}_0 =_{df.} \boldsymbol{V}_\Gamma \times \boldsymbol{V}_\Gamma$. When no confusion is likely, we will omit the subscript Γ as well as the superscript Λ in $\boldsymbol{E}_\Gamma^\Lambda$.

The following fact establishes that, for each Λ-consistent theory Γ, the structure $\boldsymbol{E}_\Gamma$ is a properly defined extensive game according to Definition 3.2.1 on page 65.

Fact 4.3.8 *Let Γ be a maximal Λ-consistent theory Γ in $L(A,B)$. Then, $\boldsymbol{E}_\Gamma$ is a properly defined extensive game.*

Proof: Consider an arbitrary theory Γ in $L(A,B)$. By construction $(\boldsymbol{V}_\Gamma, \boldsymbol{R}_\Gamma)$ is a tree and some reflection reveals that the length of a string in $\boldsymbol{V}_\Gamma$ is no longer than $\|N\| + 2$. Hence, with N being finite, $(\boldsymbol{V}_\Gamma, \boldsymbol{R}_\Gamma)$ has a finite horizon. The set of players $\boldsymbol{N}_\Gamma$ of $\boldsymbol{E}_\Gamma$ is finite as well, since $\boldsymbol{N}_\Gamma$ contains just one element more than N, *viz.*, the mystery player. The player assignment function $\boldsymbol{P}_\Gamma$ total on the internal vertices by definition.

Finally, each of the players' preferences are reflexive, transitive and connected as required. For the mystery player 0 this is immediate, its preference relation being the universal relation over $\boldsymbol{V}_\Gamma$. So, for the remainder of the proof, consider an arbitrary player i in $\boldsymbol{N}_\Gamma$ other than 0, *i.e.*, i in N.

For reflexivity, consider an arbitrary vertex σ in $\sigma \in \boldsymbol{V}_\Gamma$. Then either σ in $\boldsymbol{W}_\Gamma$ or in $\boldsymbol{I}_\Gamma$. If the former, observe that, by Fact 4.3.5, the theory $\boldsymbol{\theta}(\sigma)$ is empty, and trivially $(\sigma, \sigma) \in \boldsymbol{\rho}_i$. In the latter case, also by Fact 4.3.5, the theory $\boldsymbol{\theta}(\sigma)$ is maximal Λ-consistent. Assume for an arbitrary formula φ that $[i]\varphi \in \boldsymbol{\theta}(\sigma)$. With axiom T_i, then also $\varphi \in \boldsymbol{\theta}(\sigma)$ and consequently, $(\sigma, \sigma) \in \boldsymbol{\rho}_i$.

For transitivity a similar run-of-the-mill argument suffices. Assume that $(\sigma, \sigma'), (\sigma', \sigma'') \in \boldsymbol{\rho}_i$ for arbitrary vertices $\sigma, \sigma', \sigma'' \in \boldsymbol{V}_\Gamma$. Again if $\boldsymbol{\theta}(\sigma)$ is empty, immediately $(\sigma, \sigma'') \in \boldsymbol{\rho}_i$, as well. Otherwise, $\boldsymbol{\theta}(\sigma)$ is maximal Λ-consistent and assume for an arbitrary formula φ that $[i]\varphi \in \boldsymbol{\theta}(\sigma)$. By axiom 4_i, also $[i][i]\varphi \in \boldsymbol{\theta}(\sigma)$. By definition of $\boldsymbol{\rho}_i$, then subsequently $[i]\varphi \in \boldsymbol{\theta}(\sigma')$ and $\varphi \in \boldsymbol{\theta}(\sigma'')$. Hence, $(\sigma, \sigma'') \in \boldsymbol{\rho}_i$.

To prove that for each $i \in N$ the relation $\boldsymbol{\rho}_i$ is connected, we must show that either $(\sigma, \sigma') \in \boldsymbol{\rho}_i$ or $(\sigma', \sigma) \in \boldsymbol{\rho}_i$, for all $\sigma, \sigma' \in \boldsymbol{V}_\Gamma$. Consider arbitrary vertices $\sigma, \sigma' \in \boldsymbol{V}_\Gamma$. If either $\sigma \notin \boldsymbol{W}_\Gamma$ or $\sigma' \notin \boldsymbol{W}_\Gamma$, *i.e.*, if either $\boldsymbol{\theta}(\sigma)$ or $\boldsymbol{\theta}(\sigma')$ is empty, we are done immediately. Otherwise, *i.e.*, if both $\sigma \in \boldsymbol{W}_\Gamma$ and $\sigma' \in \boldsymbol{W}_\Gamma$, the axioms *E3* and *E4* are heavily relied upon.

First we introduce the auxiliary notion of a *connecting path*, which we define as a sequence of vertices $\tau_0, \upsilon_0, \ldots, \tau_n, \upsilon_n$, or $\tau_0, \upsilon_0, \ldots \tau_{n-1}, \upsilon_{n-1}, \tau_n$ in $\boldsymbol{V}_\Gamma$ such that $\tau_0 = \epsilon$, and for each $0 \leqslant m \leqslant n$ both $\upsilon_m \in \boldsymbol{W}_{\tau_m}$ and $\tau_{m+1} \in T_{\upsilon_m}$. Observe that each τ_i is in $\boldsymbol{T}_\Gamma$. The latter requirement guarantees that, given the construction of $\mathfrak{T}_\Gamma$, each vertex τ_{m+1} was introduced to $\boldsymbol{V}_\Gamma$ in virtue of some formula of the form $\neg[k]\chi$ being included in $\boldsymbol{\theta}(\upsilon_m)$. *I.e.*, we may assume that for each $0 < m \leqslant n+1$, there to be a $j \in N$ such that $\left\{\psi : \ [k]\psi \in \boldsymbol{\theta}(\upsilon_m)\right\} \subseteq \boldsymbol{\theta}(\tau_{m+1})$. Inspection of the various possible

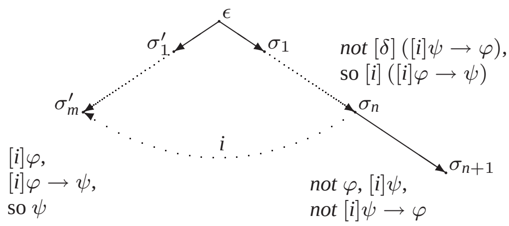


Figure 4.6.

cases along with an easy inductive argument reveal that for each vertex σ in $\boldsymbol{W}$, there is a connecting path of which σ is the last element.

The argument proceeds with a simultaneous induction on n and m, proving that for any two connecting paths $\sigma_0, \ldots, \sigma_n$ and $\sigma'_0, \ldots, \sigma'_m$ that either $(\sigma_n, \sigma'_m) \in \boldsymbol{\rho}_i$ or $(\sigma'_m, \sigma_n) \in \boldsymbol{\rho}_i$.

For $n = m = 0$, obviously, $\sigma_n = \sigma'_m = \epsilon$ and we are done immediately by reflexivity of $\boldsymbol{\rho}_i$. For the first inductive case — assuming the claim to hold for n and m and proving it also to hold for $n+1$ and m — consider two connecting paths $\sigma_0, \ldots, \sigma_n, \sigma_{n+1}$ and $\sigma'_0, \ldots, \sigma'_m$. If $n+1$ is odd, then $\sigma_n = \tau_{n/2}$ and $\sigma_{n+1} = \upsilon_{n/2}$. By definition of a connecting path then $\sigma_{n+1} \in \boldsymbol{W}^{\sigma_n}$. If in this case, $\sigma_{n+1} = \sigma_n$, then $\{\chi : \ [i]\chi \in \boldsymbol{\theta}(\sigma_n)\} \subseteq \boldsymbol{\theta}(\sigma_{n+1})$ as a consequence of axiom T_i and $\boldsymbol{\theta}(\sigma_{n+1})$ being maximal Λ-consistent. Otherwise, $\sigma_{n+1} \in \boldsymbol{L}^{\sigma_n}$ and $\{\chi : \ [\beta^{\sigma_n}_{\sigma_{n+1}}]\chi \in \boldsymbol{\theta}(\sigma_n)\} \subseteq \boldsymbol{\theta}(\sigma_{n+1})$, in virtue of Fact 4.3.6. If, on the other hand, $n+1$ is even, then $\sigma_{n+1} = \tau_{n/2}$ and $\sigma_n = \upsilon_{(n/2)-1}$. In this case, there is a $k \in N$ such that $(\sigma_n, \sigma_{n+1}) \in \boldsymbol{\rho}_k$, *i.e.*, $\{\chi : \ [k]\chi \in \boldsymbol{\theta}(\sigma_n)\} \subseteq \boldsymbol{\theta}(\sigma_{n+1})$. Let δ_{n+1} denote a label in B as follows:

$$\delta_{n+1} \quad =_{df.} \quad \begin{cases} i & \text{if } n+1 \text{ is odd and } \sigma_{n+1} \in \boldsymbol{T}, \\ \beta^{\sigma_n}_{\sigma_{n+1}} & \text{if } n+1 \text{ is odd and } \sigma_{n+1} \notin \boldsymbol{T}, \\ k & \text{otherwise, } i.e., \text{ if } n+1 \text{ is even.} \end{cases}$$

Then in general, we have that $\{\chi : \ [\delta_{n+1}]\chi \in \boldsymbol{\theta}(\sigma_n)\} \subseteq \boldsymbol{\theta}(\sigma_{n+1})$.

By the induction hypothesis, we may assume that $(\sigma_n, \sigma'_m) \in \boldsymbol{\rho}_i$ or $(\sigma'_m, \sigma_n) \in \boldsymbol{\rho}_i$. In the former case, assume that $(\sigma'_m, \sigma_{n+1}) \notin \boldsymbol{\rho}_i$, *i.e.*, for some formula φ, $[i]\varphi \in \boldsymbol{\theta}(\sigma'_m)$ but $\varphi \notin \boldsymbol{\theta}(\sigma_{n+1})$. We show that $(\sigma_{n_1}, \sigma'_m) \in \boldsymbol{\rho}_i$. To this end consider an arbitrary formula ψ such that $[i]\psi \in \boldsymbol{\theta}(\sigma_{n+1})$; we prove that $\psi \in \boldsymbol{\theta}(\sigma'_m)$. By maximal Λ-consistency of $\boldsymbol{\theta}(\sigma_{n+1})$, $[i]\psi \to \varphi \notin \boldsymbol{\theta}(\sigma_{n+1})$. This yields $[\delta_{n+1}]\big([i]\psi \to \varphi\big) \notin \boldsymbol{\theta}(\sigma_n)$. By maximal Λ-consistency and axiom $E4_{\delta_{n+1},i}$, then also that $[i]\big([i]\varphi \to \psi\big) \in \boldsymbol{\theta}(\sigma_n)$. Having assumed that $(\sigma_n, \sigma'_m) \in \boldsymbol{\rho}_i$, with the definition of $\boldsymbol{\rho}_i$, we have $[i]\varphi \to \psi \in \boldsymbol{\theta}(\sigma'_m)$. Subsequently, by maximal Λ-consistency and $[i]\varphi \in \boldsymbol{\theta}(\sigma_m)$, eventually, $\psi \in \boldsymbol{\theta}(\sigma'_m)$. This line of reasoning is illustrated in Figure 4.6.

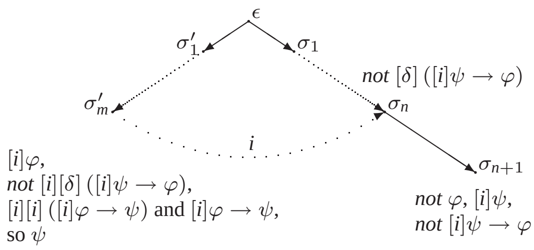


Figure 4.7.

Also if the other case obtains, *i.e.*, if $(\sigma'_m, \sigma_n) \in \boldsymbol{\rho}_i$, we can prove that $(\sigma'_m, \sigma_{n+1}) \notin \boldsymbol{\rho}_i$ implies $(\sigma_{n+1}, \sigma'_m) \in \boldsymbol{\rho}_i$. So assume $(\sigma'_m, \sigma_{n+1}) \notin \boldsymbol{\rho}_i$; then there is a formula φ such that $[i]\varphi \in \boldsymbol{\theta}\,(\sigma'_m)$ but $\varphi \notin \boldsymbol{\theta}\,(\sigma_{n+1})$. Consider an arbitrary formula ψ such that $[i]\psi \in \boldsymbol{\theta}\,(\sigma_{n+1})$; we prove that $\psi \in \boldsymbol{\theta}\,(\sigma_m)$. As in the previous case, $[i]\psi \to \varphi \notin \boldsymbol{\theta}\,(\sigma_{n+1})$ and consequently $[\delta_{n+1}]\big([i]\psi \to \varphi\big) \notin \boldsymbol{\theta}\,(\sigma_n)$. This time having assumed that $(\sigma'_m, \sigma_n) \in \boldsymbol{\rho}_i$, we obtain $[i][\delta_{n+1}]\big([i]\psi \to \varphi\big) \notin \boldsymbol{\theta}\,(\sigma'_m)$. By maximal Λ-consistency of $\boldsymbol{\theta}\,(\sigma'_m)$ and axiom $E3_{\delta_{n+1},i,i,i,i}$, then $[i][i]\big([i]\varphi \to \psi\big) \in \boldsymbol{\theta}\,(\sigma'_m)$. Two applications of axiom T_i, then give $[i]\varphi \to \psi \in \boldsymbol{\theta}\,(\sigma'_m)$. With $[i]\varphi$ having been assumed to be in $\boldsymbol{\theta}\,(\sigma'_m)$, the *desideratum* $\psi \in \boldsymbol{\theta}\,(\sigma'_m)$ follows. Figure 4.7 illustrates this argument.

Since the argument for the second inductive case — assuming the claim to hold for n and m and proving it also to hold for n and $m+1$ — runs along analogous lines, we may conclude that each player i's the preference relation $\boldsymbol{\rho}_i$ is connected. $\dashv$

A strategy profile of the extensive game $\boldsymbol{E}_\Gamma$ is then given by function mapping each internal vertex in $(\boldsymbol{V}_\Gamma, \boldsymbol{R}_\Gamma)$ onto a succeeding vertex. Not all strategy profiles of $\boldsymbol{E}_\Gamma$, however, are represented by a strategy label in S. We will assume the label s in S to represent the strategy profile $\boldsymbol{s}_s$ that maps each internal vertex σ onto σs. *I.e.*, for each internal vertex σ in $\boldsymbol{V}_\Gamma$ and each strategy label s in S we have:

$$\boldsymbol{s}_s\,(\sigma) \quad =_{df.} \quad \sigma s.$$

The following fact is an obvious consequence of this definition.

Fact 4.3.9 *Let $\boldsymbol{E}_\Gamma$ be the extensive game defined for a Λ-consistent theory Γ in $L(A, B)$. Let $x \in \boldsymbol{T}_\Gamma$, $i \in N$ and $s \in S$ and have $\boldsymbol{s}$ denote $\boldsymbol{s}_s$. Then:*

$$\begin{aligned}
\hat{\boldsymbol{s}}_\emptyset\,(x) &= \left\{xs^{\|N\|+1}\right\}, \\
\hat{\boldsymbol{s}}_i\,(x) &= \left\{xs^{\|N\|+1}, xs^i s', xs^i n \in \boldsymbol{\Sigma}_\Gamma : \ n \in \omega \text{ and } s' \in S \text{ such that } s' \neq s\right\}.
\end{aligned}$$

Sketch of proof: On basis of the definition of $\hat{s}_N$ on page 70, and inspection of the construction of $\mathfrak{T}_\Gamma$ and $\boldsymbol{E}_\Gamma$, the following can be established:

$$\hat{\boldsymbol{s}}_\varnothing(xs) = \left\{xs^{\|N\|+1}\right\}$$
$$\hat{\boldsymbol{s}}_i(xs) = \left\{xs^{\|N\|+1}, xs^i s', xs^i n \in \boldsymbol{\Sigma}_\Gamma : \ n \in \omega \text{ and } s' \in S \text{ such that } s' \neq s\right\}.$$

Now recall that the additional player 0 had been assumed to be different from any players in N. Since $\boldsymbol{P}(x) = 0$, for each $i \in N$ we have $\sigma \in \boldsymbol{s}_i(x)$ if and only if $\sigma = \boldsymbol{s}(x)$, *i.e.*, if $\sigma = xs$. Then, $\hat{\boldsymbol{s}}_i(x) = \hat{\boldsymbol{s}}_i(xs)$ and $\hat{\boldsymbol{s}}_\varnothing(x) = \hat{\boldsymbol{s}}_\varnothing(xs)$, which give the desired result. $\dashv$

The stage has now been set for the definition of the game-frame $\mathfrak{F}_\Gamma$ on $\boldsymbol{E}_\Gamma$ for $L(A,B)$. The set of vertices $\boldsymbol{V}_\Gamma$ is common to $\mathfrak{T}_\Gamma$, $\boldsymbol{E}_\Gamma$ and $\mathfrak{F}_\Gamma$. The function $\boldsymbol{\theta}_\Gamma$, assigning theories in $L(A,B)$ to vertices in $\boldsymbol{V}_\Gamma$, is invoked to define the game model $\mathfrak{M}_\Gamma$ on $\mathfrak{F}_\Gamma$. Then we prove that the maximal consistent theory the function $\boldsymbol{\theta}_\Gamma$ assigns to each vertex σ in $\boldsymbol{W}_\Gamma$, coincides precisely with the set of formulas that σ forces in $\mathfrak{M}_\Gamma$.

Definition 4.3.10 *(Game-frame and game-model on $\boldsymbol{E}_\Gamma$)* Let Λ be an extensive game logic for a multi-modal matrix language $L(A,B)$. Let $\boldsymbol{E}_\Gamma$ be the extensive game for a Λ-consistent theory in $L(A,B)$ given by $\left(\boldsymbol{V}_\Gamma, \boldsymbol{R}_\Gamma, \boldsymbol{N}_\Gamma, \boldsymbol{P}_\Gamma, \{\boldsymbol{\rho}_i\}_{i \in \boldsymbol{N}_\Gamma}\right)$. Define the label map $\boldsymbol{f}$ assigning players and strategy profiles in $\boldsymbol{E}_\Gamma$ to player labels, respectively, strategy labels in B:

$$\boldsymbol{f}(\beta) =_{df.} \begin{cases} i & \text{if } \beta \text{ is the player label } i \text{ in } B, \\ \boldsymbol{s}_s & \text{if } \beta \text{ is the strategy label } s \text{ in } B. \end{cases}$$

Then define the game-frame $\mathfrak{F}^\Lambda_\Gamma$ on $\boldsymbol{E}_\Gamma$ as the tuple $\left(\boldsymbol{V}_\Gamma, \{\boldsymbol{R}_\beta\}_{\beta \in B}, \boldsymbol{f}\right)$. The game model $\mathfrak{M}_\Gamma$ on $\mathfrak{F}^\Lambda_\Gamma$ is defined as the tuple $\left(\mathfrak{F}^\Lambda_\Gamma, \mathfrak{V}_\Gamma\right)$, where the interpretation function $\mathfrak{V}_\Gamma : \boldsymbol{V}_\Gamma \to 2^A$, is defined such that for each vertex σ in $\boldsymbol{V}_\Gamma$:

$$\mathfrak{V}_\Gamma(\sigma) =_{df.} \left\{a \in A : \ a \in \boldsymbol{\theta}(\sigma)\right\}.$$

Whenever no confusion is likely to arise the superscript Λ is omitted.

For each Λ-consistent theory Γ, the frame $\mathfrak{F}_\Gamma$ and the model $\mathfrak{M}_\Gamma$ are a properly defined game-frame and a properly defined game-model, complying with Definition 3.3.1 on page 72. The model $\mathfrak{M}_\Gamma$ can now be shown to satisfy the Λ-consistent theory Γ at the root node. In order to establish this, we prove something slightly stronger, *viz.*, that for each vertex σ of $\mathfrak{M}_\Gamma$ the theory $\boldsymbol{\theta}_\Gamma(\sigma)$, *i.e.*, the theory assigned to v by $\boldsymbol{\theta}_\Gamma$, contains exactly those formulas that are satisfied at v in $\mathfrak{M}_\Gamma$, provided that $\boldsymbol{\theta}_\Gamma(v)$ is non-empty.

Lemma 4.3.11 *(Truth Lemma)* *Let Γ be a Λ-consistent theory in $L(A,B)$. Consider both $\mathfrak{T}_\Gamma = (\boldsymbol{\Sigma}_\Gamma, \boldsymbol{\theta}_\Gamma)$ and $\mathfrak{M}_\Gamma$. Then for all vertices $\sigma \in \boldsymbol{W}_\Gamma$ and all formulas φ:*

$$\mathfrak{M}_\Gamma, \sigma \Vdash \varphi \quad \textit{iff} \quad \varphi \in \boldsymbol{\theta}_\Gamma(\sigma)$$

Proof: Consider an arbitrary $x \in \boldsymbol{T}_\Gamma$ as well as an arbitrary $\sigma \in \boldsymbol{W}_\Gamma^x$. The proof is then by induction on φ.

For φ a propositional variable we are done immediately by the definition of $\mathfrak{M}_\Gamma$. Similarly, if φ is a Boolean combination, *i.e.*, if $\varphi = \bot$, $\varphi = \neg\psi$ or $\varphi = \psi \wedge \chi$, maximal Λ-consistency of $\boldsymbol{\theta}(\sigma)$ takes care. Thus the modal cases remain.

Let $\varphi = [i]\psi$, for some $i \in N$. Assume φ to be the $n+1$-st formula in the enumeration, *i.e.*, $\varphi = \varphi_n$. First assume $[i]\psi \in \boldsymbol{\theta}(\sigma)$ and consider an arbitrary σ' such that $\sigma \boldsymbol{R}_i \sigma'$. Observe that in $\mathfrak{M}_\Gamma$, we have $\sigma \boldsymbol{R}_i \sigma'$ if and only if $[i]\varphi \in \boldsymbol{\theta}(\sigma)$ implies $\varphi \in \boldsymbol{\theta}(\sigma')$, for all formulas φ. Hence, $\psi \in \boldsymbol{\theta}(\sigma')$. Consequently, $\boldsymbol{\theta}(\sigma')$ is not empty and $\sigma' \in \boldsymbol{W}_\Gamma$. Therefore, the induction hypothesis is applicable, yielding $\mathfrak{M}_\Gamma, \sigma' \Vdash \psi$. Having chosen σ' arbitrarily, $\mathfrak{M}_\Gamma, \sigma \Vdash [i]\psi$ follows. For the opposite direction, assume $[i]\psi \notin \boldsymbol{\theta}(\sigma)$. With maximal Λ-consistency of $\boldsymbol{\theta}(\sigma)$, then $\neg[i]\psi \in \boldsymbol{\theta}(\sigma)$. Therefore, $\{\neg\psi\} \cup \{\chi : [i]\chi \in \boldsymbol{\theta}(\sigma)\} \subseteq \boldsymbol{\theta}(t_n^\sigma)$. By definition of $\boldsymbol{R}_i$ in $\mathfrak{M}_\Gamma$, we have $\sigma \boldsymbol{R}_i t_n^\sigma$. Since $t_n^\sigma \in \boldsymbol{W}_\Gamma$, then $\mathfrak{M}_\Gamma, t_n^\sigma \nVdash \psi$, in virtue of the induction hypothesis. Consequently, $\mathfrak{M}_\Gamma, \sigma \nVdash [i]\psi$.

Let $\varphi = [\hat{s}_i]\psi$, for some $i \in N$ and $s \in S$. We distinguish the case in which σ is in $\boldsymbol{T}_\Gamma$ from the one in which σ is a leaf in $\boldsymbol{L}_\Gamma^x$ for some $x \in \boldsymbol{T}_\Gamma$. If the latter, then σ is the only element of $\hat{\boldsymbol{s}}_i(\sigma)$, *i.e.*, $\hat{\boldsymbol{s}}_i(\sigma) = \{\sigma\}$, where $\boldsymbol{s}$ abbreviates $\boldsymbol{s}_s$. Being an instance of axiom $E2_{\beta_\sigma^x, \hat{s}_i}$, the formula $[\beta_\sigma^x]\big([\hat{s}_i]\psi \leftrightarrow \psi\big)$ is in $\boldsymbol{\theta}(x)$. With Fact 4.3.6, then, $[\hat{s}_i]\psi \leftrightarrow \psi \in \boldsymbol{\theta}(\sigma)$. Consider the following equivalences:

$$
\begin{array}{lll}
[\hat{s}_i]\psi \in \boldsymbol{\theta}(\sigma) & \text{iff}_{[\hat{s}_i]\psi \leftrightarrow \psi \in \boldsymbol{\theta}(\sigma)} & \psi \in \boldsymbol{\theta}(\sigma) \\
& \text{iff}_{i.h.} & \mathfrak{M}_\Gamma, \sigma \Vdash \psi \\
& \text{iff}_{\hat{\boldsymbol{s}}_i(\sigma) = \{\sigma\}} & \text{for all } \sigma' \text{ such that } \sigma' \in \hat{\boldsymbol{s}}_i(\sigma) : \mathfrak{M}_\Gamma, \sigma' \Vdash \psi \\
& \text{iff} & \text{for all } \sigma' \text{ such that } \sigma \boldsymbol{R}_{\hat{s}_i(\sigma)} \sigma' : \mathfrak{M}_\Gamma, \sigma' \Vdash \psi \\
& \text{iff} & \mathfrak{M}_\Gamma, \sigma \Vdash [\hat{s}_i]\psi.
\end{array}
$$

If, on the other hand, $\sigma = x$ for some $x \in \boldsymbol{T}_\Gamma$, assume $[\hat{s}_i]\psi \in \boldsymbol{\theta}(x)$. Then, by axiom $E1_{\hat{s}_i, \hat{s}_\emptyset}$, also $[\hat{s}_\emptyset]\psi \in \boldsymbol{\theta}(x)$. Consider an arbitrary σ' such that $\sigma \boldsymbol{R}_{\hat{s}_i} \sigma'$. Then, $\sigma' \in \hat{\boldsymbol{s}}_i(x)$ and $\sigma' \in \boldsymbol{L}^x$. Moreover, by Fact 4.3.9, either $\sigma = xs^{\|N\|+1}$ or $\sigma = xs^i y$, for $y \in \omega \cup (S - \{s\})$. If the former $\beta_{\sigma'}^x = \hat{s}_\emptyset$ and and if the latter, $\beta_{\sigma'}^x = \hat{s}_i$. In either case $[\beta_{\sigma'}^x]\psi \in \boldsymbol{\theta}(x)$. By Fact 4.3.6, then $\psi \in \boldsymbol{\theta}(\sigma')$. By the induction hypothesis follows that $\mathfrak{M}_\Gamma, \sigma \Vdash \psi$ and with σ' having been chosen arbitrarily, eventually, $\mathfrak{M}_\Gamma, x \Vdash [\hat{s}_i]\psi$.

For the opposite direction, assume that $[\hat{s}_i]\psi \notin \boldsymbol{\theta}(x)$. Then, $\neg[\hat{s}_i]\psi \in \boldsymbol{\theta}(x)$, by maximal Λ-consistency of $\boldsymbol{\theta}(x)$. Without loss of generality we may assume $\neg[\hat{s}_i]\psi$ to be the $n+1$-st element in the enumeration. Consider the vertex $xs^i n$ and let it be denoted by σ^*. Then, by Fact 4.3.9, $\sigma^* \in \hat{\boldsymbol{s}}_i(x)$ and *a fortiori* also $x \boldsymbol{R}_{\hat{s}_i} \sigma^*$. Moreover, $\neg\psi \in \boldsymbol{\theta}(\sigma^*)$, by construction of $\mathfrak{T}_\Gamma$. By the induction hypothesis, then, $\mathfrak{M}_\Gamma, \sigma^* \Vdash \neg\psi$ and, consequently, $\mathfrak{M}_\Gamma, \sigma^* \nVdash \psi$, which suffices for a proof.

Let $\varphi = [\hat{s}_\emptyset]$. Again we distinguish between σ being a leaf in $\boldsymbol{L}_\Gamma$ and σ in $\boldsymbol{T}_\Gamma$, dealing with the latter case first.

Let σ be denoted by x and the vertex $xs^{\|N\|+1}$ by σ^{**}. Then, $\beta^{x}_{\sigma^{**}} = \hat{s}_{\emptyset}$. First assume that $[\hat{s}_{\emptyset}]\psi \in \boldsymbol{\theta}(x)$, *i.e.*, $[\beta^{x}_{\sigma^{**}}]\psi \in \boldsymbol{\theta}(x)$. By Fact 4.3.6 then $\psi \in \boldsymbol{\theta}(\sigma^{**})$. The induction hypothesis subsequently yields $\mathfrak{M}_{\Gamma}, \sigma^{**} \Vdash \psi$. From Fact 4.3.9 we learn that σ^{**} is the only element in $\hat{\boldsymbol{s}}_{\emptyset}(x)$ and, hence, σ^{**} is also the only element such that $x\boldsymbol{R}_{\hat{s}_{\emptyset}}\sigma^{**}$. We may conclude that $\mathfrak{M}_{\Gamma}, x \Vdash [\hat{s}_{\emptyset}]\psi$.

For the opposite direction assume $[\hat{s}_{\emptyset}]\psi \notin \boldsymbol{\theta}(x)$. By maximal Λ-consistency of $\boldsymbol{\theta}(x)$, both $\neg[\hat{s}_{\emptyset}]\psi \in \boldsymbol{\theta}(x)$ and $\langle\hat{s}_{\emptyset}\rangle\neg\psi \in \boldsymbol{\theta}(x)$. In virtue of axiom $D!_{\hat{s}_{\emptyset}}$, then also $[\hat{s}_{\emptyset}]\neg\psi \in \boldsymbol{\theta}(x)$, *i.e.*, $[\beta^{x}_{\sigma^{**}}]\neg\psi \in \boldsymbol{\theta}(x)$. By Fact 4.3.6 then $\neg\psi \in \boldsymbol{\theta}(\sigma^{**})$ and, with maximal Λ-consistency of $\boldsymbol{\theta}(\sigma^{**})$, also $\psi \notin \boldsymbol{\theta}(\sigma^{**})$. The induction hypothesis is applicable and so $\mathfrak{M}_{\Gamma}, \sigma^{**} \nVdash \psi$. Fact 4.3.9 guarantees that $\sigma^{**} \in \hat{\boldsymbol{s}}_{\emptyset}(x)$ and so $x\boldsymbol{R}_{\hat{s}_{\emptyset}}\sigma^{**}$. Hence, eventually, $\mathfrak{M}_{\Gamma}, x \nVdash [\hat{s}_{\emptyset}]\psi$, which we had set out to prove.

Now suppose σ to be a leaf in $\boldsymbol{L}^{x}$ for some $x \in \boldsymbol{T}_{\Gamma}$. Then $\hat{\boldsymbol{s}}_{\emptyset}(\sigma) = \{\sigma\}$. In virtue of axiom $E5_{\beta^{x}_{\sigma}, \hat{s}_{\emptyset}}$ we have that $[\beta^{x}_{\sigma}]\big([\hat{s}_{\emptyset}]\varphi \leftrightarrow \varphi\big) \in \boldsymbol{\theta}(x)$. By Fact 4.3.6, also $[\hat{s}_{\emptyset}]\varphi \leftrightarrow \varphi \in \boldsymbol{\theta}(\sigma)$. Now consider the following, familiar looking, equivalences:

$$
\begin{array}{lll}
[\hat{s}_{\emptyset}]\psi \in \boldsymbol{\theta}(\sigma) & \text{iff}_{[\hat{s}_{\emptyset}]\psi \leftrightarrow \psi \in \boldsymbol{\theta}(\sigma)} & \psi \in \boldsymbol{\theta}(\sigma) \\
& \text{iff}_{i.h.} & \mathfrak{M}_{\Gamma}, \sigma \Vdash \psi \\
& \text{iff}_{\hat{\boldsymbol{s}}_{\emptyset}(\sigma) = \{\sigma\}} & \text{for all } \sigma' \text{ such that } \sigma' \in \hat{\boldsymbol{s}}_{i}(\sigma) : \mathfrak{M}_{\Gamma}, \sigma' \Vdash \psi \\
& \text{iff} & \text{for all } \sigma' \text{ such that } \sigma\boldsymbol{R}_{\hat{s}_{\emptyset}(\sigma)}\sigma' : \mathfrak{M}_{\Gamma}, \sigma' \Vdash \psi \\
& \text{iff} & \mathfrak{M}_{\Gamma}, \sigma \Vdash [\hat{s}_{\emptyset}]\psi.
\end{array}
$$

This concludes the proof. ⊣

Completeness

Definition 4.2.1 introduced M as the minimal extensive game logic. Proposition 4.2.2 proved its axioms to be sound with respect to the class of all game-frames. Completeness of M with respect to this comprehensive class of game-frames follows as a corollary of the Truth Lemma 4.3.11 and the fact that, for each M-consistent theory, the model $\mathfrak{M}_{\Gamma}$ is in fact a game-model.

Theorem 4.3.12 *(Completeness of* M*) Let Γ be a theory and φ a formula in a multi-modal matrix language $L(A,B)$. Then:*

$$\Gamma \vdash_{\mathrm{M}} \varphi \quad \textit{iff} \quad \Gamma \vDash_{\mathrm{M}} \varphi.$$

Proof: The left-to-right direction is taken care of by Proposition 4.2.2, above. For the right-to-left direction it suffices to prove that there is a model on a game-frame for each M-consistent theory Γ in $L(A,B)$. The construction of $\mathfrak{M}^{\mathrm{M}}_{\Gamma}$ for M, as defined above, provides such a model. For, proving the routine contrapositive, $\Gamma \nvdash_{\mathrm{M}} \varphi$ implies

$\Gamma \cup \{\neg\varphi\} \nvdash_{\mathrm{M}} \bot$. Then $\mathfrak{M}^{\mathrm{M}}_{\Gamma\cup\{\neg\varphi\}}$ exists and, by Lemma 4.3.11, $\mathfrak{M}^{\mathrm{M}}_{\Gamma\cup\{\neg\varphi\}}, \epsilon \Vdash \Gamma \cup \{\neg\varphi\}$. Then, also $\mathfrak{M}^{\mathrm{M}}_{\Gamma\cup\{\neg\varphi\}}, \epsilon \Vdash \Gamma$ and $\mathfrak{M}^{\mathrm{M}}_{\Gamma\cup\{\neg\varphi\}}, \epsilon \nVdash \varphi$, yielding $\Gamma \nvDash_{\mathrm{M}} \varphi$. $\dashv$

Completeness for the extensive game logics $\mathrm{M5}_{s,i}$ and $\mathrm{M5}^N_s$ (for particular labels s and i and N the whole set of player-labels) is be obtained in a similar fashion. The validity of the argument, however, depends on the models $\mathfrak{M}^{\mathrm{M5}_{s,i}}_{\Gamma}$ and $\mathfrak{M}^{\mathrm{M5}^N_i}_{\Gamma}$ belonging to the appropriate classes of game-models. *I.e.*, for each $\mathrm{M5}_{s,i}$-consistent theory Γ the model $\mathfrak{M}^{\mathrm{M5}_{s,i}}_{\Gamma}$ be based on an extensive game in which the strategy profile $\boldsymbol{s}_s$ represented by the label s in $L(A,B)$ contains a subgame perfect best response for player represented by the label i. Similarly, in the extensive game underlying a model $\mathfrak{M}^{\mathrm{M5}^N_s}_{\Gamma}$ for an $\mathrm{M5}^N_s$-consistent theory Γ, there be a label in N for each interested player and the strategy profile represented by the label s be a subgame perfect Nash equilibrium. For Γ an $\mathrm{M5}_{s,i}$-consistent theory, respectively, an $\mathrm{M5}^N_i$-consistent theory, we find that $\mathfrak{M}^{\mathrm{M5}_{s,i}}_{\Gamma}$ and $\mathfrak{M}^{\mathrm{M5}^N_s}_{\Gamma}$ do actually meet these requirements. The soundness of $\mathrm{M5}_{s,i}$ and $\mathrm{M5}^N_s$ with respect to these classes of frames being guaranteed by Proposition 4.2.2 and Theorem 3.3.5, we have the following results.

Theorem 4.3.13 *(Completeness of* $\mathrm{M5}_{s,i}$*) The logic* $\mathrm{M5}_{s,i}$ *is sound and complete with respect to the class of game-models built on game-frames in which s is a subgame perfect best response for player i.*

Proof: Soundness is a consequence of Proposition 4.2.2 and Theorem 3.3.5, on page 75 above. For completeness the proof is as that for M (Theorem 4.3.12), be it that it should also be shown that for any $\mathrm{M5}_{s,i}$-consistent theory, the strategy profile $\boldsymbol{s}_s$ is a subgame perfect best response for player i in the extensive game $\boldsymbol{E}^{\mathrm{M5}_{s,i}}_{\Gamma}$ that is defined in the course of the construction of the model $\mathfrak{M}^{\mathrm{M5}_{s,i}}_{\Gamma}$. In the remainder of the proof the subscript is omitted in $\boldsymbol{s}_s$. In virtue of Proposition 3.3.4 on page 75, it suffices to demonstrate that $\mathfrak{F}_{\Gamma}$ is $(\hat{\boldsymbol{s}}_i, i, \hat{\boldsymbol{s}}_{\emptyset})$-Euclidean. Consider an arbitrary $\mathrm{M5}_{s,i}$-consistent theory Γ and equally arbitrary vertices $\sigma, \sigma', \sigma'' \in \boldsymbol{V}_{\Gamma}$ such that $\sigma \boldsymbol{R}_{\hat{s}_i} \sigma'$ and $\sigma \boldsymbol{R}_{\hat{s}_{\emptyset}} \sigma''$. Then, $\sigma' \in \hat{\boldsymbol{s}}_i(\sigma)$ and $\sigma'' \in \hat{\boldsymbol{s}}_{\emptyset}(\sigma)$. We show that $\sigma' \boldsymbol{R}_i \sigma''$. *i.e.*, that $(\sigma', \sigma'') \in \boldsymbol{\rho}_i$, Consider an arbitrary formula φ and, proving the contrapositive, we assume $\varphi \notin \boldsymbol{\theta}_{\Gamma}(\sigma'')$ and demonstrate that $\varphi \notin \boldsymbol{\theta}_{\Gamma}(\sigma')$. Some reflection reveals that there be some $x \in \boldsymbol{T}_{\Gamma}$ such that $\sigma', \sigma'' \in \boldsymbol{L}^x$. Then, $\sigma'' = xs^{\|N\|+1}$. Moreover, $\sigma' = xs^{\|N\|+1}$ as well, or $\sigma' = xs^i y$, for some $y \in \omega \cup (S - \{s\})$. In either case $\sigma' \in \hat{\boldsymbol{s}}_i(x)$, by Fact 4.3.9, and so $x\boldsymbol{R}_{\hat{s}_i}\sigma'$. In virtue of Fact 4.3.6, furthermore, $[\beta^x_{\sigma''}]\varphi \notin \boldsymbol{\theta}_{\Gamma}(x)$. Since in this case $\beta^x_{\sigma''} = \hat{s}_{\emptyset}$, also $[\hat{s}_{\emptyset}]\varphi \notin \boldsymbol{\theta}_{\Gamma}(x)$. With $\langle \hat{s}_i\rangle[i]\varphi \to [\hat{s}_i]\varphi$ being an instance of axiom $5_{s,i}$ and by maximal $\mathrm{M5}_{s,i}$-consistency of $\boldsymbol{\theta}_{\Gamma}(x)$, then $\langle \hat{s}_i\rangle[i]\varphi \notin \boldsymbol{\theta}_{\Gamma}(x)$. By Lemma 4.3.11, we have $\mathfrak{M}_{\Gamma}, x \nVdash \langle \hat{s}_i\rangle[i]\varphi$ and hence $\mathfrak{M}_{\Gamma}, x \Vdash [\hat{s}_i]\neg[i]\varphi$. Having assumed that $x\boldsymbol{R}_{\hat{s}_i}\sigma'$, also $\mathfrak{M}_{\Gamma}, \sigma \nVdash [i]\varphi$. By another application of Lemma 4.3.11, eventually, $[i]\varphi \notin \boldsymbol{\theta}_{\Gamma}(\sigma')$. $\dashv$

Theorem 4.3.14 *The logic* $\mathrm{M5}^N_s$ *in* $L(A,B)$ *is sound and complete with respect to the class of game-frames built on games in which s is a subgame perfect Nash equilibrium and in which there is a label in N for each interested player.*

Proof: Soundness is again by Theorem 3.3.5 and Theorem 3.3.5, on page 75 above. For completeness, it suffices to show that, for each maximal $\mathrm{M5}_s^N$-consistent theory, the strategy profile $\boldsymbol{s}_s$ is a subgame perfect Nash equilibrium in the extensive game $\boldsymbol{E}_\Gamma^{\mathrm{M5}_s^N}$, on which the model $\mathfrak{M}_\Gamma^{\mathrm{M5}_s^N}$ is based. First observe that the mystery player 0 is an indifferent player, in the sense that his preference relation is universal by construction. As a consequence, each strategy profile is a best response for 0 and this holds in particular for $\boldsymbol{s}_i$. For any of the other players $i \in \boldsymbol{N}_\Gamma$, there is a player label in B_0. Hence, for any of them, the axiom $5_{\hat{s}_i, i, \hat{s}_\emptyset}$ is derivable in $\mathrm{M5}_s^N$. In a similar manner as in Theorem 4.3.13, it can thus be shown that for each *interested* player the strategy profile $\boldsymbol{s}_s$ is a best response for player i in $\boldsymbol{E}_\Gamma^{\mathrm{M5}_s^N}$. We may conclude the proof by observing that $\boldsymbol{s}_s$ is a subgame perfect Nash equilibrium in $\boldsymbol{E}_\Gamma^{\mathrm{M5}_s^N}$. ⊣

As an immediate consequence of these completeness results and the fact that in any derivation of a formula φ from a theory Γ in an extensive game logic Λ only a finite number of formulas can occur, we have that each of the logics M, $\mathrm{F5}_{s,i}$ and $\mathrm{F5}_s^N$ is *compact*. *I.e.*, for each theory Γ and each formula φ, if $\Gamma \vDash_\Lambda \varphi$ then there is a *finite* subtheory Γ_e of Γ such that $\Gamma_e \vDash_\Lambda \varphi$. We state this fact here as a corollary.

Corollary 4.3.15 *Let $L(A, B)$ be a multi-modal matrix language with N the set of player labels containing i and with s as a strategy label. Then, the extensive game logics* M, $\mathrm{M5}_{s,i}$ *and* $\mathrm{M5}_s^N$ *are compact.*

For Λ one of the extensive game logics M, $\mathrm{M5}_{s,i}$ and $\mathrm{M5}_s^N$ and Γ a M-consistent theory, the extensive game $\boldsymbol{E}_\Gamma^\Lambda$ has some noteworthy properties. In particular, the depth of the game $\boldsymbol{E}_\Gamma^\Lambda$ — *i.e.*, the length of the longest path in the game-tree connecting the root to a leaf — does not exceed the number of player labels in $L(A, B)$ plus two, *i.e.*, $\|N\| + 2$. Moreover, the players are assumed to play in a fixed order, and on each path in the game-tree from the root to a leaf, each player represented by a label in N moves at most once and any other player at most twice. Also, the number of players in each game $\boldsymbol{E}_\Gamma$ is always one greater than the number of labels in N.

Corollary 4.3.16 *Let Γ be a theory in a multi-modal matrix language $L(A, B)$ with N the set of player labels and let Λ one of the extensive game logics* M, $\mathrm{M5}_{s,i}$ *or* $\mathrm{M5}_s^N$. *For E an extensive game of perfect information E and $\mathfrak{M}_E$ a game-model on a game-frame $\mathfrak{F}_E$ for Λ on E. Assume that Γ is satisfiable in $\mathfrak{M}_E$. Then there is an extensive game of perfect information E' such that:*

1. *there is game-frame for Λ on E on which Γ is also satisfiable;*
2. *the game tree of E' is of a maximal depth of $\|N\| + 2$;*
3. *its players number $\|N\| + 1$ and move in a fixed order;*
4. *in each play of E', each player represented by a label in N moves at most once and any other player at most twice.*

Sketch of proof: Since Γ is satisfiable in a model on $\mathfrak{F}_E$, by Theorem 4.3.12, Λ-consistency of Γ follows. Construct the extensive game $\boldsymbol{E}^{\Lambda}_{\Gamma}$. The Truth Lemma 4.3.11 ensures that Γ is satisfied at the root node of $\boldsymbol{E}^{\Lambda}_{\Gamma}$. Moreover, $\boldsymbol{E}^{\Lambda}_{\Gamma}$ can be seen to possess the properties as stated in the corollary. $\dashv$

Part II

Boolean Games

Chapter 5

Boolean Games

5.1 Introduction

In game-like situations a player has to decide in the face of epistemic uncertainty. The structure of the game may be such that his information about the game is insufficient to distinguish one possible state of the game from the other. This may occur, *e.g.*, if she knows that one of her opponents has made a move, but not exactly which move. In such games of *imperfect information* (pure) strategies that prescribe different courses of action to a player in possible states of the game she cannot distinguish, are no longer available to her. The idea is, that due to her epistemic limitations, she would be unable to act upon such a strategy.

A method that is widely employed to make formally explicit this epistemic structure is by including *information sets* in the description of an extensive game. These information sets partition the set of internal vertices in such a way that in each block of this partition all vertices are assigned to one player. The respective player is thought to be unable to distinguish the different vertices in the information set. Consequently, any of his strategies will have to prescribe a similar course of action at each node in each information set. This similarity between different courses of action is commonly accounted for by introducing a one-one correspondence between the alternatives open to a player at one vertex and those of any other vertex in the same information set, *e.g.*, by labelling the vertices with actions. The strategies available to a player are then restricted to those that prescribe corresponding courses of action at different vertices at any two vertices in the same information set. Graphically, information sets consisting of more than one vertex may be depicted as dotted lines connecting the vertices it contains (*cf.*, the extensive game on the left in Figure 5.1). Those not connected thus are assumed to form an information set on their own.

In this chapter we consider two-person finite extensive games of imperfect information based on binary trees. The players are, moreover, assumed to be complete antagonists and the outcomes are of only two kinds: a win for the one player or a win

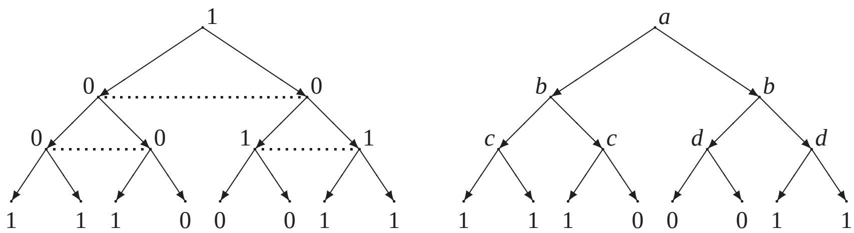


Figure 5.1. The figure on the left dipicts an extensive game of imperfect information in which figure two antagonists — player 0 and player 1. The outcomes at the leafs in which player 1 is victorious are labelled with 1; otherwise, *i.e.*, if player 0 wins, with 0. At each vertex the player to move has a binary choice between either going to the left or going to the right. The information sets are indicated by the dotted lines connecting vertices. The strategies of each player available to each player are assumed to be restricted in the sense that each of them should either prescribe to go left at all vertices in the information set or to move to the right at all of them. The picture on the right, is the Boolean form representing the epistemic structure of the game on the left. We find that this boolean form is represented by (equivalence class of) the propositional formula $(a \leftrightarrow b) \vee (\neg a \wedge \neg c)$. The corresponding Boolean game ensues if a and c are assigned to Player 1 and b and d to Player 0.

for the other player, without the possibility of a tie.

We may assume that from any vertex in such a game emanate exactly two edges of two different kinds. In accordance with the two-dimensional way we depict binary trees, these different kinds of vertices may be called *left moves* and *right moves*, or 0 *moves* and 1 *moves*, respectively. Without loss of generality it may also be assumed that, whenever two vertices are in the same information set, each strategy for the respective player prescribes a left move at the one vertex if and only if it prescribes a right move in the other. An alternative way of representing information sets is then by labelling the vertices by *binary decision variables*, in such a way that the vertices in the same information set are labelled by the same variable (*cf.*, the extensive game to the left in Figure 5.1). Control over the binary decision variables a vertex is labelled with is then be assigned to the player to play at that vertex. The strategies available to the players may then be represented as the different choices they can make with respect to the binary decision variables in their control. *I.e.*, a strategy for a player becomes a function mapping the decision variables in his control to the two values these variables may take. Represented thus, we refer to this class of games as *Boolean games*. Without the explicit assignment of the decision variables to the players, a Boolean game is called a *Boolean game form* or just a *Boolean form*.

A strategy profile for a Boolean is now an assignment of binary values to *all* of its decision variables. The same decision variables may occur different Boolean games. Moreover, we allow a strategy profile also to prescribe a value for decision variables that do not occur in the game and assume that the choices made for these "outside" variables do not affect the outcome of the game. This makes that the set of strategy profiles of distinct Boolean games may coincide and their strategic properties can be compared and assessed on this common basis. It also suggest a natural notion of equivalence of Boolean forms: two Boolean forms are said to be *equivalent* if each strategy profile determines the same outcome in both of them. Defining natural operations on Boolean game forms, we find that, *modulo* this notion of equivalence, Boolean forms constitute a Boolean algebra.

The principal observation of this chapter is that binary decision variables may be taken as the propositional variables of a propositional language. On this conception, each strategy profile, assigning binary values to all the decision variables, coincides with a valuation for that propositional language. We find that each Boolean form can be associated with a propositional formula, and *vice versa*, such that that two Boolean forms are equivalent if and only if the corresponding formulas are logically equivalent. This correspondence between Boolean forms and propositional variables, moreover, proves to determine an isomorphism between the Boolean algebra of Boolean forms and the Lindenbaum algebra of the respective propositional language. These algebraic considerations ensure that Boolean games can straightforwardly be related to classical propositional logic. The Boolean algebra of (equivalence classes of) games happens to be isomorphic to the Lindenbaum algebra of a suitable classical propositional language.

One should be careful, however, not to confuse the notion of equivalence between Boolean forms with notions of *game equivalence* as advanced in, *e.g.*, Thompson (1952) — in terms of congruence of reduced strategic form — or in Goranko (2001a), van Benthem (no date-b) and Pauly (2001) — based on the sets of outcomes the players can guarantee to end the game in. Equivalence of Boolean forms does not take into account the manipulative powers of the players. The strategic properties of a Boolean game — such as a player having a winning strategy or not — may depend on the way control over the decision variables over the players as well as on the structure of the underlying Boolean form.

The point is rather that the conception of Boolean forms as propositional formulas spawns a number of logical issues concerning distributed control over the propositional variables. Two strategic issues with respect to Boolean games present themselves. First, *given a Boolean form and a player, which distributions of the decision variables yield a Boolean game in which that player has a winning strategy?* Second, *given a distribution of the decision variables, in which Boolean games complying with this distribution has the one player a winning strategy, in which the other and in which neither of them?* In virtue of the isomorphism between the algebra of Boolean forms (*modulo* equivalence) and the Lindenbaum algebra of the corresponding propositional language, these questions have counterparts in propositional logic. Thus, distributed control of the propositional variables becomes a notion amenable to logical analysis.

A third issue concerns the properties of determined and indeterminate Boolean games. A two-player game of complete competition and which has wins and losses for the players as outcomes, such as Boolean games, is called *determined* if one of the players has a winning strategy. As a corollary of Kuhn's Theorem (*cf.*, Selten (1965) and page 70, above), any such game is determined, provided that it is finite and the players enjoy perfect information. Connections between logic and games with a pair of antagonists as participants, have frequently been pointed out. The employment of game-theory in *classical* logic has, however, generally been restricted to games of perfect information. Consequently, these games can generally be assumed to be determined as well. The assumption of perfect information games being determined has even been generalized to infinite games and as such it has been proposed as a rival of the *Axiom of Choice*. The use of games of imperfect information has generally been restricted to the semantical analysis of such abstruse phenomena as branching quantifiers and the independence-friendly interpretation of connectives (Hintikka, 1973; Hintikka and Sandu, 1997). Boolean games are not in general games of perfect information and neither are they generally determined. This eventuality, however, does occur only if none of the players has control over all propositional variables. Thus one may come to wonder about the distinguishing properties of determined Boolean games.

After having introduced Boolean games in this chapter, we will examine these three logical issues in the next chapter.

5.2 Boolean Games

Boolean games constitute a class of games of involving two antagonists, denoted by 0 and 1, for which there are only two outcomes: a win for the one player or a win for the other player. In both cases the player that fails to win loses. Moreover, at each stage of the game, one of the players has choice between two alternatives.

Borderline cases are the two atomic Boolean games denoted by **0** and **1**. The former player 0 wins without either of the players making a move, and the latter which is won by player 1 without having to act.

Furthermore, *complex* or *molecular* Boolean games are constructed recursively from these atomic games and a countable set of *binary decision variables*. For any two Boolean games g_0 and g_1 and any decision variable a, there is another Boolean game which we denote by $a\,(g_0, g_1)$. Each decision variable is assigned to the control of one of the two players. Moreover, each decision variable can take one of two values, represented by 0 and 1. In $a\,(g_0, g_1)$ it is up to the player to whom the decision variable a has been assigned, whether the game continues with g_0 or with g_1. Choosing the value 0 results in the game being continued with g_0. For the value 1 the game proceeds with playing the game g_1. Definition 5.2.1 formally defines the set of Boolean forms and Boolean games on a set of decision variables. The purpose of the distinction between Boolean games and Boolean forms is that in a later stage we will want to be able to compare Boolean games that only differ with respect to the assignment of the decision variables to the players.

Definition 5.2.1 *(Boolean Game Forms & Boolean Games)* Let A be a countable set that is disjoint from a two-element set $\{0, 1\}$ the latter set representing the two players. Define the set of *Boolean (game) forms over A* as the smallest set $B(A)$ such that:

$$\{0,1\} \subseteq B(A)$$
$$a \in A \text{ and } g, h \in B(A) \quad \text{imply} \quad (a, g, h) \in B(A)\,.$$

We usually depict (a, g, h) by $a\,(g, h)$, and when they occur as atomic games we usually write 0 and 1 in boldface, *i.e.*, as **0** and **1**, respectively. A *Boolean game on A* is a pair (g, π) consisting of a Boolean form g and a *control assignment function* $\pi\colon A \to \{0, 1\}$ assigning the decision variables in A to the players in $\{0, 1\}$. For $i \in \{0, 1\}$, the preimage of $\{i\}$ under π is the set of decision variables assigned to i and will usually be denoted by π_i.[1] If π is clear from the context we usually refer to (g, π) by simply g. The set of Boolean games over a set A given an assignment function π is denoted by $B\,(A, \pi)$ and the set of all Boolean games over A by $B\,(A)$. Again any of these parameters may be omitted if no confusion is likely. Let further a Boolean form h be called a *subform* of another Boolean form g if g and h are identical, or $g = a\,(h_0, h_1)$ and h is a subform of either h_0 or h_1.

Definition 5.2.1 can be understood as defining a class of *extensive* games. The recursion by means of which they are introduced suggests a sequential structure of play, which can be made explicit in a game tree. A molecular Boolean game $a\,(g_0, g_1)$ offers the player controlling the decision variable a the choice between two courses of action. After having made one choice, the game continues with playing g_0 and after

[1] Alternatively, π could be defined as a set of two disjoint subsets of A indexed by $\{0, 1\}$ and such that $\bigcup \pi = A$. If π does not contain the empty set, it is then a bi-partition indexed by $\{0, 1\}$.

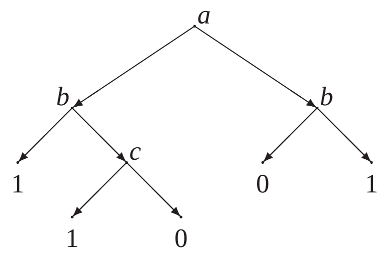


Figure 5.2. The Boolean form $a\big(b\big(\mathbf{1}, c\,(\mathbf{1}, \mathbf{0})\,\big), b\,(\mathbf{0}, \mathbf{1})\,\big)$.

having made the other the game continues with playing g_1. As such, Boolean games are based on finite binary trees, internal vertices of which are labelled with decision variables. We will assume that choosing 0 as the value for a decision variable takes one to the left and choosing 1 takes one to the right. Consider for instance the Boolean form $a\big(b\,(\mathbf{1}, c\,(\mathbf{1}, \mathbf{0}))\,, b\,(\mathbf{0}, \mathbf{1})\big)$. Play commences with the player having control over the decision variable a. If she chooses the value zero for a, the game continues with the Boolean game $b\,(\mathbf{1}, c\,(\mathbf{1}, \mathbf{0}))$, otherwise with $b\,(\mathbf{0}, \mathbf{1})$, and so on until an outcome is reached. Under these assumptions, the Boolean game $a\big(b\,(\mathbf{1}, c\,(\mathbf{1}, \mathbf{0}))\,, b\,(\mathbf{0}, \mathbf{1})\big)$ can be represented as in Figure 5.2. In much the same manner, the graph on the right in Figure 5.1 depicts the Boolean form:

$$a\big(b\,(c\,(\mathbf{1}, \mathbf{1})\,, c\,(\mathbf{1}, \mathbf{0}))\,, b\,(d\,(\mathbf{0}, \mathbf{0})\,, d\,(\mathbf{1}, \mathbf{1}))\big).$$

As pointed out in the introduction the binary decision variables are construed as identifying the players' information sets, requiring the players to choose strategies that assign a unique value to each decision variable. The decision variables — which label the nodes of the game-tree — indicate that these nodes are in the same information set. This interpretation of Boolean games as finite extensive games of imperfect information is enforced by an appropriate definition of a player's strategy. Thus, in the game of Figure 5.2, any strategy for the player with control over b should prescribe either a move to the right in both subgames $b\,(\mathbf{1}, c\,(\mathbf{1}, \mathbf{0}))$ and $b\,(\mathbf{0}, \mathbf{1})$, or a move to the left in both.

Accordingly, a *strategy* for a player in a Boolean game (g, π) in a class $B\,(A)$ is defined formally as a function that assigns one of the binary values 0 or 1 to the decision variables controlled by the player in question. Observe that this assignment is *not* restricted to the decision variables that occur in the Boolean form g, but that it is total on the whole set of decision variables assigned to a player. Because control over the decision variables is divided over the two players, a *strategy profile* is a function mapping the whole of A onto $\{0, 1\}$. As such, strategies and strategy profiles can be seen as (characteristic functions of) subsets of decision variables; in the sequel we will frequently exploit this equivocality. The set of strategy profiles for a Boolean game in $B(A, \pi)$, is thus given by 2^A and, as such, is *independent* of π. This is expressly not the case for strategies. The set of strategy profiles is denoted by S.

Each strategy profile determines a unique outcome for each Boolean form, *i.e.*, independently of the specific way control over the decision variables is divided over the two players. For **0** and **1** the outcome will invariably be 0 and 1, respectively. The outcome of a molecular game $a(g_0, g_1)$ will depend on the value assigned to a in s. In case a is assigned the value 1 in a strategy profile s, the outcome of the Boolean form $a(g_0, g_1)$ is identified with the outcome of g_1 given s. Otherwise, the outcome of s in $a(g_0, g_1)$ is identical to the outcome of g_0 given s. This gives rise to the definition of the *strategic form of a game g*, denoted by $\ulcorner g \urcorner$, as a function with the set of strategy profiles as domain and $\{0, 1\}$ as its range. Two Boolean forms are said to be *equivalent* if their strategic forms are identical. Observe that, with the set of strategy profiles being defined for a class of Boolean games $B(A)$, all games in this class share the same set of strategy profiles and, as such, they can be compared on this basis. This is even the case if the decision variables occurring in two Boolean forms do not coincide. Formally we have the following definition.

Definition 5.2.2 *(Strategic Forms and Their Equivalence)* Let $B(A)$ be the set of Boolean forms over the set A. For each $g \in B(A)$ we define its *strategic form* as a function $\ulcorner g \urcorner : S \to \{0, 1\}$, as follows:

$$\begin{aligned}
\ulcorner \mathbf{0} \urcorner (s) &=_{df.} 0 \\
\ulcorner \mathbf{1} \urcorner (s) &=_{df.} 1 \\
\ulcorner a(g_0, g_1) \urcorner (s) &=_{df.} \begin{cases} \ulcorner g_1 \urcorner (s) & \text{if } a \in s, \\ \ulcorner g_0 \urcorner (s) & \text{otherwise.} \end{cases}
\end{aligned}$$

The strategic form of a Boolean form, *c.q.* Boolean game, being a function from strategy profiles to a two-element set, can be taken as a characteristic function of a subset of the strategy profiles, *i.e.*, of $\{s \in S : \ulcorner g \urcorner (s) = 1\}$. The definition above then translates to:

$$\begin{aligned}
\ulcorner \mathbf{0} \urcorner &= \emptyset \\
\ulcorner \mathbf{1} \urcorner &= S \\
\ulcorner a(g_0, g_1) \urcorner &= \left(\ulcorner g_0 \urcorner \cap \{s \in S : a \notin s\} \right) \cup \left(\ulcorner g_1 \urcorner \cap \{s \in S : a \in s\} \right).
\end{aligned}$$

Two Boolean forms g and h are said to be *equivalent* — in symbols, $g \equiv h$ — if they have the same strategic form, *i.e.*:

$$g \equiv h \quad \text{iff} \quad \ulcorner g \urcorner = \ulcorner h \urcorner.$$

It is not hard to verify that the Boolean forms in Figures 5.2 and 5.1 are equivalent in this sense. Let A be given by $\{a, b, c, d\}$. We then find that both their strategic forms are the same set of sets of strategy profiles in 2^A, *viz.*:

$$\{\emptyset, \{b\}, \{c\}, \{d\}, \{a, b\}, \{b, d\}, \{c, d\}, \{a, b, c\}, \{a, b, d\}, A\}.$$

A Boolean game is defined as a Boolean form together with a function assigning control over the decision variables to the players. A player's strategies are given by the the different choices he can make with respect to the decision variables assigned to him. A Boolean game can thus be represented as a matrix, with the columns indicating the strategies for player 0 and the rows those of player 1. A combination of strategies then yields a strategy profile, each of which is represented by a cell in the matrix. The entries in the cells represents the values the strategic form takes for the respective strategy profiles.

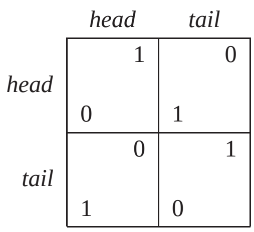
On this conception, Boolean games are ordinary strategic games, featuring two antagonists whose strategies are given by the various choices they can make with respect to the decision variables assigned to them. As Boolean games allow for two outcomes only — *viz.*, a win for the one player or win for the other — and assuming that both players (strictly) prefer winning to losing, a strategy profile is a Nash equilibrium if and only if it subsumes a strategy for one of the players that cannot fail to deliver victory, *i.e.*, no matter what strategy his opponent chooses to play. Thus, we say a strategy profile is a *winning strategy* for a player if it guarantees that player a win even if his opponent were to choose different values for the decision variables in her control. Define formally for s a strategy profile and i a player of a Boolean game (g, π):

$$s \text{ is a } \textit{winning strategy} \text{ for } i \quad \text{iff} \quad \text{for all } s' \in S\colon s \sim_{\pi_i} s' \text{ implies } \ulcorner g \urcorner (s') = i.$$

We say that a player i *has a winning strategy* if there is at least one strategy profile that is a winning strategy for i. Obviously, at most one of the players can have a winning strategy. Otherwise both players could play one of their respective winning strategies against one another. Then both players would secure a win, which is absurd given the assumption that one player's victory means the other's defeat. Assume for a *reductio ad absurdum*, that s be a winning strategy for 0 and s' a winning strategy for player 1. Let, furthermore, s^* be given (s_0, s'_1). Then, both $s \sim_{\pi_0} s^*$ and $s' \sim_{\pi_1} s^*$. Hence, both $\ulcorner g \urcorner (s^*) = 0$ and $\ulcorner g \urcorner (s^*) = 1$, *quod non*.

Games are called *determined* if one of the players has a winning strategy and *indeterminate*, otherwise. Not all Boolean games are determined; it is quite possible for neither of the players of a Boolean game to have a winning strategy. Typical examples of Boolean games for which this is the case are $a\,(b\,(\mathbf{0}, \mathbf{1})\,, b\,(\mathbf{1}, \mathbf{0}))$ and $a\,(b\,(\mathbf{1}, \mathbf{0})\,, b\,(\mathbf{0}, \mathbf{1}))$, if one player has control over a and the other over b. *E.g.*, strategy profile s grants player 1 a win in $a\,(b\,(\mathbf{1}, \mathbf{0})\,, b\,(\mathbf{0}, \mathbf{1}))$ if and only if a and b are either both present or both absent in s. Player 1, however, has no safeguard against player 0 assigning the 'opposite' value to b as he did to a, resulting in a win for player 0. Similarly, whatever value player 0 chooses to assign to b, player 1 might have chosen to set a to the same value and win. These games could be taken as the Boolean counterparts of the well-known game of *Matching Pennies*, in which two players toss a penny and the one player may keep them both if the upsides match; otherwise, the other player obtains the two pennies (*cf.*, Figure 5.3).

Although a player need not have a winning strategy at her disposal, she may have a *winning response*, *i.e.*, for each strategy her opponent may choose to play, she can



	$\emptyset$	$\{b\}$
$\emptyset$	0	1
$\{a\}$	1	0

Figure 5.3. On the left, the game *Matching Pennies*. The figures indicate victories and defeats, but not how many pennies are won or lost. On the right its representation by the a Boolean game $(a\,(b\,(\mathbf{0},\mathbf{1})\,,b\,(\mathbf{1},\mathbf{0})))$. Player 1 plays rows, player 0 columns.

always find an appropriate strategy in response that will secure her a win. Yet, for different strategies of the opponent the appropriate response may be a different strategy. As such winning responses are *not* strategies in the strict sense, but rather functions mapping each strategy of the opponent onto a strategy of the player herself. Formally, for i a player in a Boolean game (g,π):

$$\begin{gathered} i \text{ has a } \textit{winning response} \text{ in } (g,\pi) \\ \text{iff} \\ \text{for all } s\in S,\ \text{there is an } s'\in S \text{ such that } s\sim_{\pi_{1-i}} s' \text{ and } \ulcorner g\urcorner(s')=i. \end{gathered}$$

If a player has a winning strategy, he clearly has a winning response as well: against each strategy of the opponent he can play his winning strategy.

Conceiving of the strategic form of Boolean form as a subset of strategy profiles, the notions of a player having a winning strategy or a player having a winning response can conveniently be expressed using the apparatus of rough sets. As an immediate consequence of the definitions above, the set of winning strategies for Player 1 in g is given by $\underline{apr}_{\pi_1}\left(\ulcorner g\urcorner\right)$ and those of Player 0 likewise by $\underline{apr}_{\pi_0}\left(\overline{\ulcorner g\urcorner}\right)$ or, equivalently, by $\overline{\overline{apr}_{\pi_0}\left(\ulcorner g\urcorner\right)}$.[2] Then, it can easily be recognized that:

$$\begin{array}{lll} \text{Player 1 has a winning strategy} & \text{iff} & \underline{apr}_{\pi_1}\left(\ulcorner g\urcorner\right)\neq\emptyset, \\ \text{Player 0 has a winning strategy} & \text{iff} & \overline{apr}_{\pi_0}\left(\ulcorner g\urcorner\right)\neq S, \\ \text{Player 1 has a winning response} & \text{iff} & \overline{apr}_{\pi_0}\left(\ulcorner g\urcorner\right)=S, \\ \text{Player 0 has a winning response} & \text{iff} & \underline{apr}_{\pi_1}\left(\ulcorner g\urcorner\right)=\emptyset. \end{array}$$

[2] Recall that π_i here denotes a *set* of decision variables and as such determines a *partition* of the strategy profiles. It is with respect to this partition — which, if written out in full, would be denoted by the cumbersome π_{π_i} — that the rough set operators $\overline{apr}_{\pi_i}$ and $\underline{apr}_{\pi_i}$ approximate.

	$\emptyset$	$\{c\}$	$\{d\}$	$\{c,d\}$
$\emptyset$	1_{1}	1_{2}	1_{3}	1_{4}
$\{a\}$	0_{5}	0_{6}	0_{7}	0_{8}
$\{b\}$	1_{9}	0_{10}	1_{11}	0_{12}
$\{a,b\}$	1_{13}	1_{14}	1_{15}	1_{16}

	$\emptyset$	$\{b\}$	$\{c\}$	$\{b,c\}$
$\emptyset$	1_{1}	1_{9}	1_{2}	0_{10}
$\{a\}$	0_{5}	1_{13}	0_{6}	1_{14}
$\{d\}$	1_{3}	1_{11}	1_{4}	0_{12}
$\{a,d\}$	0_{7}	1_{15}	0_{8}	1_{16}

Figure 5.4. Two Boolean games on the Boolean form $a\big(b\,(\mathbf{1}, c\,(\mathbf{1},\mathbf{0}))\,, b\,(\mathbf{0},\mathbf{1})\big)$. In the game on the left control over the propositional variables a and b is assigned to player 1 and control over c and d to player 0. In the game on the right player 1 has control over a and d, and player 0 over b and c. The figures in the lower right corner of each cell indicate how the entries in the two matrices are correllated: in cells indicated by the same number the same strategy profile is played.

From these equivalences we can immediately read off that one player has a winning strategy if and only if the other has no winning response. The intuitively obvious observation that having a winning strategy implies having a winning response, essentially depends on the sets of decision variables π_0 and π_1 being disjoint. The assumption that, *e.g.*, Player 1 has a winning strategy furnishes one with a strategy profile s in $\underline{apr}_{\pi_1}\left(\ulcorner g \urcorner\right)$ that is such that for any strategy profile s', also $(s \cap \pi_1) \cup (s' \cap \pi_0)$ is in $\ulcorner g \urcorner$. Since clearly also $s' \sim_{\pi_0} (s \cap \pi_1) \cup (s' \cap \pi_0)$, we may conclude that Player 1 has a winning response as well. These remarks are summarized in the following proposition.

Proposition 5.2.3 *Let (g, π) be a Boolean game in $B\,(A, \pi)$ and let $i \in \{0, 1\}$. Then, at most one player has a winning strategy in (g, π), player i having a winning strategy implies i having a winning response, and player i has a winning strategy if and only if player $1 - i$ has no winning response.*

Proof: The third claim is immediate from the rough set characterization of a player having a winning strategy and a player having a winning response. The other two claims can be seen to follow as well if, in addition, Corollary 2.2.12 (page 44, above) is invoked. $\dashv$

5.3 Operations on Boolean Forms

In Definition 5.2.1, above, Boolean forms were introduced recursively, providing a way in which larger Boolean forms can be constructed from smaller ones. Here we introduce four operations on Boolean forms performing a similar task in a different manner. We prove that, *modulo* equivalence, the Boolean forms constitute a Boolean algebra with respect to these operations. Rather, this result forges a strong link between Boolean forms on the one hand and propositional formulas on the other.

The operations on Boolean forms, *complement* ($\overline{g}$), *sum* ($g+h$), *product* ($g \cdot h$) and *simultaneous sum-product* ($\circledast(g,h,k)$), are defined formally as follows.

Definition 5.3.1 Let A be a set. Define the set of four operations $(^{-}, +, \cdot, \circledast)$ of similarity type $(1,2,2,3)$ on the set of Boolean forms $B(A)$ inductively as follows, where a ranges over A:

$$
\begin{array}{lrcl}
(1) & \overline{\mathbf{0}} & =_{df.} & \mathbf{1} \\
 & \overline{\mathbf{1}} & =_{df.} & \mathbf{0} \\
 & \overline{a(g_0,g_1)} & =_{df.} & a(\overline{g_0},\overline{g_1}) \\
(2) & \mathbf{0}+h & =_{df.} & h \\
 & \mathbf{1}+h & =_{df.} & \mathbf{1} \\
 & a(g_0,g_1)+h & =_{df.} & a(g_0+h, g_0+h) \\
(3) & \mathbf{0}\cdot h & =_{df.} & \mathbf{0} \\
 & \mathbf{1}\cdot h & =_{df.} & h \\
 & a(g_0,g_1)\cdot h & =_{df.} & a(g_0\cdot h, g_0\cdot h) \\
(4) & \circledast(\mathbf{0},h,k) & =_{df.} & h \\
 & \circledast(\mathbf{1},h,k) & =_{df.} & k \\
 & \circledast(a(g_0,g_1),h,k) & =_{df.} & a(\circledast(g_0,h,k), \circledast(g_1,h,k)).
\end{array}
$$

Let the algebra $(B(A); \mathbf{0}, \mathbf{1}, ^{-}, +, \cdot)$ be denoted by $\mathfrak{B}_A$, suppressing the subscript A when clear from the context

These operations on Boolean forms have intuitive readings, which are best appreciated if Boolean games are thought of as trees. Taking the *complement* of a form g makes that all occurrences of the atomic games **1** and **0** are interchanged. The *sum* of two forms, $g+h$, is the result of replacing *any* occurrence of **0** in g by h. Addition yields the Boolean form in which the root of h is attached to any leaf node of g labelled with 0. The *product* of two forms, $(g \cdot h)$, is similar to their sum, be it that now it is every occurrence of **1** that is replaced by h. The operation $\circledast$ comes down to simultaneously adding one form and multiplying it with another simultaneously. Hence, $\circledast(g,h,k)$ yields the form that is like g except that each occurrence of **1** is replaced by an occurrence of h, and every occurrence of **0** by one of k. Figure 5.2 illustrates the workings of the operators.

With respect to the algebraic properties of these operations it is worth observing that $+$ and $\cdot$ are neither idempotent nor commutative. Neither do the absorption laws ($g+(g\cdot h)=g$ and $g\cdot(g+h)=g$) hold in general. Moreover, $+$ does not distribute over $\cdot$ and neither does $\cdot$ over $+$. Also the identity laws $g+\overline{g}=\mathbf{1}$, $g\cdot\overline{g}=\mathbf{0}$, $g+\mathbf{0}=\mathbf{1}$ and $g\cdot\mathbf{1}=\mathbf{0}$ fail to hold in general. Hence,$\mathfrak{B}$ is obviously *not* a Boolean algebra. However, Fact 5.3.2 summarizes, among other things, some of the Boolean properties that do hold for Boolean forms.

Fact 5.3.2 *Let g, h and k be Boolean forms in $B(A)$. Then:*

$$\begin{array}{rclcrcl}
\overline{\overline{g}} & = & g & & & & \\
g+\mathbf{0} & = & g & & g\cdot\mathbf{1} & = & g \\
g+(h+k) & = & (g+h)+k & & g\cdot(h\cdot k) & = & (g\cdot h)\cdot k \\
\overline{g+h} & = & \overline{g}\cdot\overline{h} & & \overline{g\cdot h} & = & \overline{g}+\overline{h} \\
\circledast(g,\mathbf{0},\mathbf{1}) & = & g & & \circledast(g,\mathbf{1},\mathbf{0}) & = & \overline{g} \\
\circledast(g,h,\mathbf{1}) & = & g+h & & \circledast(g,\mathbf{0},h) & = & g\cdot h \\
\circledast(g,\mathbf{1},h) & = & \overline{g}+h & & \circledast(g,h,\mathbf{0}) & = & \overline{g}\cdot h.
\end{array}$$

Proof: All proofs are straightforward, although it may strike the reader as slightly odd that the De Morgan laws hold, whereas, *e.g.*, commutativity and distributivity do not. Here we prove by induction on the complexity of g that $\overline{g+h}=\overline{g}\cdot\overline{h}$. For the basic case, *i.e.*, if $g=\mathbf{0}$ or $g=\mathbf{1}$, consider the following pair of equations:

$$\begin{array}{rcccccccl}
\overline{\mathbf{0}+h} & = & \overline{h} & = & \mathbf{1}\cdot\overline{h} & = & \overline{\mathbf{0}}\cdot\overline{h}, & & \\
\overline{\mathbf{1}+h} & = & \overline{\mathbf{1}} & = & \mathbf{0} & = & \mathbf{0}\cdot\overline{h} & = & \overline{\mathbf{1}}\cdot\overline{h}.
\end{array}$$

For the inductive case, *i.e.*, if $g=a(g_0,g_1)$, consider the following equations:

$$\begin{array}{rcl}
\overline{a(g_0,g_1)+h} & = & \overline{a(g_0+h,g_1+h)} \;=\; a\big(\overline{g_0+h},\overline{g_0+h}\big) \\
 & =_{i.h.} & a\big(\overline{g_0}\cdot\overline{h},\overline{g_0}\cdot\overline{h}\big) \;=\; a\big(\overline{g_0},\overline{g_1}\big)\cdot\overline{h} \;=\; \overline{a(g_0,g_1)}\cdot\overline{h}.
\end{array}$$

This ends the proof. ⊣

In addition, both of the following claims hold; the inductive proofs are elementary and duly omitted:

$$\begin{array}{rcl}
g+h=\mathbf{0} & \text{iff} & g=\mathbf{0} \text{ and } h=\mathbf{0}, \\
g\cdot h=\mathbf{1} & \text{iff} & g=\mathbf{1} \text{ and } h=\mathbf{1}.
\end{array}$$

Each Boolean form can be associated with a finite combination of forms in the subset $\{a(\mathbf{0},\mathbf{1}):\ a\in A\}$ together with $\mathbf{0}$ and $\mathbf{1}$ by means of a finite number applications of the operator $\circledast$. The atomic Boolean forms $\mathbf{0}$ and $\mathbf{1}$ are be associated with themselves and, inductively, each molecular game $a(g_0,g_1)$ by $\circledast(a(\mathbf{0},\mathbf{1}),h_0,h_1)$,

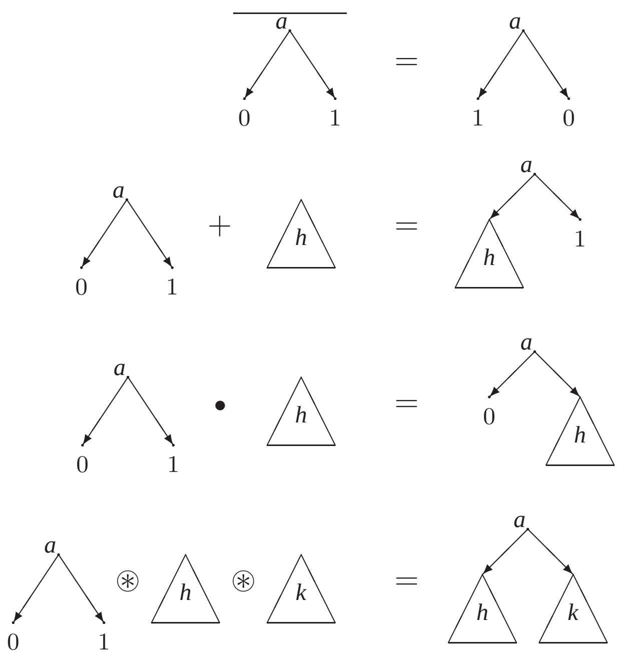


Table 5.2. Four operations on Boolean forms: complement, sum, product and simultaneous sum-product.

where h_0 and h_1 are the combinations associated with g_0 and g_1, respectively. Thus the Boolean form $a\big(b\,(\mathbf{1}, c\,(\mathbf{1},\mathbf{0}))\,, b\,(\mathbf{0},\mathbf{1})\big)$ of Figure 5.2 is identical to the combination:

$$\circledast\big(a\,(\mathbf{0},\mathbf{1}), \circledast\big(b\,(\mathbf{0},\mathbf{1}), \mathbf{1}, \circledast\,(c\,(\mathbf{0},\mathbf{1})\,, \mathbf{1}, \mathbf{0})\big), b\,(\mathbf{0},\mathbf{1})\big).$$

The following proposition lays down formally this general observation.

Proposition 5.3.3 *The algebra* $(B(A)\,; \mathbf{0}, \mathbf{1}, \circledast)$ *is generated by* $\{a\,(\mathbf{0},\mathbf{1}) : \ a \in A\}$.

Proof: Trivial. Consider an arbitrary g in $B(A)$. We prove by induction that g can be generated. For the basis, we are done immediately since $\mathbf{0}$ and $\mathbf{1}$ are nullary operations in $(B(A)\,; \mathbf{0}, \mathbf{1}, \circledast)$. For the induction step, let $a\,(h_0, h_1)$ and observe that $a\,(h_0, h_1) = \circledast\,(a\,(\mathbf{0},\mathbf{1})\,, h_0, h_1)$. With the induction hypothesis we are done. $\dashv$

A similar result cannot be obtained for the operations $^{-}$, $+$ and $\cdot$, not even if the set of molecular Boolean forms is extended to $\big\{a\,(\boldsymbol{i},\boldsymbol{j}) \in B(A) : \ \boldsymbol{i},\boldsymbol{j} \in \{\mathbf{0},\mathbf{1}\}\big\}$ together with $\mathbf{0}$ and $\mathbf{1}$. *E.g.,* the Boolean form $a\,(a\,(\mathbf{0},\mathbf{1})\,, a\,(\mathbf{1},\mathbf{0}))$ cannot be generated thus.

Proposition 5.3.4 *Let A be a non-empty set of binary decision variables. The algebra $\mathfrak{B}$ given by $(B(A)\,;\mathbf{0},\mathbf{1},^{-},+,\cdot)$ is* not *generated by any subset of $\left\{a\left(\boldsymbol{i},\boldsymbol{j}\right)\in B(A):\ \boldsymbol{i},\boldsymbol{j}\in\{\mathbf{0},\mathbf{1}\}\right\}$.*

Proof: Because A is non-empty, we may assume $a\in A$. Let g be an arbitrary Boolean form generated from $\left\{a\left(\boldsymbol{i},\boldsymbol{j}\right)\in B(A):\ \boldsymbol{i},\boldsymbol{j}\in\{\mathbf{0},\mathbf{1}\}\right\}$ by a finite number of applications of the operations $\mathbf{0}$, $\mathbf{1}$, $^{-}$, $+$ and $\cdot$. We show, for $\boldsymbol{i}$ and $\boldsymbol{j}$ *distinct* in $\{\mathbf{0},\mathbf{1}\}$, that:

$$a\left(\boldsymbol{i},\boldsymbol{j}\right) \text{ is a subform of } g \quad \text{implies} \quad a\left(\boldsymbol{j},\boldsymbol{i}\right) \text{ is no subform of } g.$$

This suffices for a proof, as both $a\left(\mathbf{0},\mathbf{1}\right)$ and $a\left(\mathbf{1},\mathbf{0}\right)$ are subforms of $a\left(a\left(\mathbf{0},\mathbf{1}\right),a\left(\mathbf{1},\mathbf{0}\right)\right)$ and $a\left(a\left(\mathbf{0},\mathbf{1}\right),a\left(\mathbf{1},\mathbf{0}\right)\right)$ is a Boolean form in $B(A)$. .

The proof is by induction on the number of occurrences of the operations $\mathbf{0}$, $\mathbf{1}$, $^{-}$, $+$ and $\cdot$ in g. So assume $a\left(\boldsymbol{i},\boldsymbol{j}\right)$ is a subform of g prove by induction on the number of operators occurring in g that for distinct $\boldsymbol{i}$ and $\boldsymbol{j}$ in $\{\mathbf{0},\mathbf{1}\}$.

The basis is trivial, since then $g\in\left\{a\left(\boldsymbol{i},\boldsymbol{j}\right)\in B(A):\ \boldsymbol{i},\boldsymbol{j}\in\{\mathbf{0},\mathbf{1}\}\right\}$. The inductive cases in which g is $\mathbf{1}$ or $\mathbf{0}$ are equally trivial.

Let $g=\overline{h}$. First observe that by a simple inductive argument, here omitted, shows that in general:

$$k \text{ is a subform of } h \quad \text{iff} \quad \overline{k} \text{ is a subform of } \overline{h}.$$

With $a\left(\boldsymbol{i},\boldsymbol{j}\right)$ a subform of $\overline{h}$, then, $\overline{a\left(\boldsymbol{i},\boldsymbol{j}\right)}$ is a subform in $\overline{\overline{h}}$. Observe that $\overline{a\left(\boldsymbol{i},\boldsymbol{j}\right)}=a\left(\boldsymbol{j},\boldsymbol{i}\right)$, since $\boldsymbol{i}$ and $\boldsymbol{j}$ were assumed to be distinct elements of $\{\mathbf{0},\mathbf{1}\}$. Moreover, $\overline{\overline{h}}=h$. By the induction hypothesis follows that $a\left(\boldsymbol{i},\boldsymbol{j}\right)$ is no subform of h. Hence, $\overline{a\left(\boldsymbol{i},\boldsymbol{j}\right)}$ is no subform of $\overline{h}$ either. Since, under the present assumptions $\overline{a\left(\boldsymbol{i},\boldsymbol{j}\right)}=a\left(\boldsymbol{j},\boldsymbol{i}\right)$, we are done.

Let $g=h+k$. This case is by induction on the complexity of k. First, assume $k=\mathbf{0}$. Then, $h+k=h+\mathbf{0}=h$, and we are done by the induction hypothesis. Second, let $k=\mathbf{1}$. An easy inductive argument establishes that $\mathbf{0}$ is no subform of $h+\mathbf{1}$, Consequently, neither is $a\left(\boldsymbol{i},\boldsymbol{j}\right)$, since $\boldsymbol{i}$ and $\boldsymbol{j}$ are distinct. Now consider the inductive case, in which $k=b\left(k_0,k_0\right)$. It suffices to show that, for distinct $\boldsymbol{m}$ and $\boldsymbol{n}$ in $\{\mathbf{0},\mathbf{1}\}$, if $a\left(\boldsymbol{m},\boldsymbol{n}\right)$ is no subform of k, it is no subform of $h+k$ either. So assume that $a\left(\boldsymbol{m},\boldsymbol{n}\right)$ be no subform of k; we prove then by an induction on the complexity of h that $a\left(\boldsymbol{m},\boldsymbol{n}\right)$ is no subform of $h+k$. If $h=\mathbf{0}$, then $h+k=\mathbf{0}+k=k$ and we are done by the assumption. If $h=\mathbf{1}$, then $h+k=\mathbf{1}+k=\mathbf{1}$, and the claim follows immediately. Finally, if $h=c\left(h_0,h_1\right)$, then $h+k=c\left(h_0,h_1\right)+k=c\left(h_0+k,h_1+k\right)$. By the induction hypothesis $a\left(\boldsymbol{m},\boldsymbol{n}\right)$ is neither a subform of h_0+k nor of h_1+k. Neither can it be the case that $h=a\left(\boldsymbol{i},\boldsymbol{j}\right)$. For either $\boldsymbol{m}=\mathbf{0}$ or $\boldsymbol{n}=\mathbf{0}$ and without loss of generality we may assume that $\boldsymbol{m}=\mathbf{0}$. Then, however, $\mathbf{0}=h_0+k$. Hence also $\mathbf{0}=k$, which is impossible with k being a molecular game.

As the argument for $g=h\cdot k$ runs along analogous lines, this concludes the proof.$\dashv$

The Boolean laws that fail to hold for $\mathfrak{B}$, however, are satisfied by the *quotient algebra* $\mathfrak{B}/_{\equiv}$, which is given by $\left\{B/_{\equiv}\,;[\mathbf{0}]_{\equiv},[\mathbf{1}]_{\equiv},^{-},+,\cdot\right\}$. Here $B/_{\equiv}$ is defined as

$\{[g]_\equiv : \ g \in B(A)\}$ and $\overline{}$, $+$ and $\cdot$ are the properly raised versions of complement, sum and product for Boolean forms, respectively. *I.e.*, we have in general that $\overline{[g]} =_{df.} [\overline{g}]$, $[g] + [h] =_{df.} [g+h]$ and $[g] \cdot [h] =_{df.} [g \cdot h]$. We first prove the following lemma as an intermediary result, which has as a corollary that $\equiv$ is a congruence relation. Hence, $\mathfrak{B}/_\equiv$ is properly defined in the first place.

Lemma 5.3.5 *For g and h Boolean forms in $B(A)$:*

$$\begin{aligned}
\ulcorner a\,(\mathbf{0},\mathbf{1})\urcorner &= \{s \in S : \ a \in s\} \\
\ulcorner \overline{g} \urcorner &= \overline{\ulcorner g \urcorner} \\
\ulcorner g + h \urcorner &= \ulcorner g \urcorner \cup \ulcorner h \urcorner \\
\ulcorner g \cdot h \urcorner &= \ulcorner g \urcorner \cap \ulcorner h \urcorner \\
\ulcorner \circledast\,(g,h,k)\urcorner &= \left(\overline{\ulcorner g \urcorner} \cap \ulcorner h \urcorner\right) \cup \left(\ulcorner g \urcorner \cap \ulcorner k \urcorner\right).
\end{aligned}$$

Proof: Throughout the proof we have s range over 2^A. For the first case consider the following equations:

$$\begin{aligned}
\ulcorner a\,(\mathbf{0},\mathbf{1})\urcorner &= \left(\ulcorner \mathbf{0} \urcorner \cap \{s\colon a \notin s\}\right) \cup \left(\ulcorner \mathbf{1} \urcorner \cap \{s\colon a \in s\}\right) \\
&= (\varnothing \cap \{s\colon a \notin s\}) \cup (S \cap \{s\colon a \in s\}) \\
&= \varnothing \cup \{s\colon a \in s\} \\
&= \{s \in S\colon a \in s\}\,.
\end{aligned}$$

The remaining cases are all by induction on g and all follow a similar pattern. We give here the proof of the third and the last case only.
For the basic cases we can reason as follows:

$$\begin{aligned}
\ulcorner \mathbf{0} + h \urcorner &= \ulcorner h \urcorner = \varnothing \cup \ulcorner h \urcorner = \ulcorner \mathbf{0} \urcorner \cup \ulcorner h \urcorner, \\
\ulcorner \mathbf{1} + h \urcorner &= \ulcorner \mathbf{1} \urcorner = S = S \cup \ulcorner h \urcorner = \ulcorner \mathbf{1} \urcorner \cup \ulcorner h \urcorner.
\end{aligned}$$

For the inductive case, in which $g = a\,(g_0, g_1)$ consider the following equalities:

$$\begin{aligned}
&\ulcorner a\,(g_0,g_1) + h \urcorner \\
= \quad & \ulcorner a\,(g_0 + h, g_1 + h)\urcorner \\
= \quad & \left(\ulcorner g_0 \urcorner + \ulcorner h \urcorner \cap \{s\colon a \notin s\}\right) \cup \left(\ulcorner g_1 \urcorner + \ulcorner h \urcorner \cap \{s\colon a \in s\}\right) \\
=_{\text{i.h.}} \quad & \left(\left(\ulcorner g_0 \urcorner \cup \ulcorner h \urcorner\right) \cap \{s\colon a \notin s\}\right) \cup \left(\left(\ulcorner g_1 \urcorner \cup \ulcorner h \urcorner\right) \cap \{s\colon a \in s\}\right) \\
=_{(*)} \quad & \left(\ulcorner g_0 \urcorner \cap \{s\colon a \notin s\}\right) \cup \left(\ulcorner g_1 \urcorner \cap \{s\colon a \in s\}\right) \cup \\
& \qquad \left(\ulcorner h \urcorner \cap \{s\colon a \notin s\}\right) \cup \left(\ulcorner h \urcorner \cap \{s\colon a \in s\}\right) \\
= \quad & \ulcorner a\,(g_0,g_1)\urcorner \cup \ulcorner h \urcorner.
\end{aligned}$$

The equation indicated with the asterisk is based on the Boolean laws governing the distribution of $\cap$ over $\cup$ and their respective commutativity.

Finally, for the last case, the following equations take care of the basis:

$$
\begin{aligned}
\ulcorner \circledast(\mathbf{0},h,k) \urcorner &= \ulcorner h \urcorner = \left(S \cap \ulcorner h \urcorner\right) \cup \varnothing = \left(\overline{\ulcorner \mathbf{0} \urcorner} \cap \ulcorner h \urcorner\right) \cup \left(\ulcorner \mathbf{0} \urcorner \cap \ulcorner k \urcorner\right),\\
\ulcorner \circledast(\mathbf{1},h,k) \urcorner &= \ulcorner k \urcorner = \varnothing \cup \left(S \cap \ulcorner k \urcorner\right) = \left(\overline{\ulcorner \mathbf{1} \urcorner} \cap \ulcorner h \urcorner\right) \cup \left(\ulcorner \mathbf{1} \urcorner \cap \ulcorner k \urcorner\right).
\end{aligned}
$$

For the inductive case we have the following:

$$
\begin{aligned}
&\ulcorner \circledast\left(a\left(g_0,g_1\right),h,k\right) \urcorner\\
= \quad & \ulcorner a\left(\circledast(g_0,h,k), \circledast(g_1,h,k)\right) \urcorner\\
= \quad & \left(\ulcorner \circledast(g_0,h,k) \urcorner \cap \{s\colon a \notin s\}\right) \cup \left(\ulcorner \circledast(g_1,h,k) \urcorner \cap \{s\colon a \in s\}\right)\\
=_{i.h.} \quad & \left(\left(\left(\overline{\ulcorner g_0 \urcorner} \cap \ulcorner h \urcorner\right) \cup \left(\ulcorner g_0 \urcorner \cap \ulcorner k \urcorner\right)\right) \cap \{s\colon a \notin s\}\right) \cup\\
& \qquad \left(\left(\left(\overline{\ulcorner g_1 \urcorner} \cap \ulcorner h \urcorner\right) \cup \left(\ulcorner g_1 \urcorner \cap \ulcorner k \urcorner\right)\right) \cap \{s\colon a \in s\}\right)\\
=_{\text{distr.}} \quad & \left(\overline{\ulcorner g_0 \urcorner} \cap \ulcorner h \urcorner \cap \{s\colon a \notin s\}\right) \cup \left(\ulcorner g_0 \urcorner \cap \ulcorner k \urcorner \cap \{s\colon a \notin s\}\right) \cup\\
& \qquad \left(\overline{\ulcorner g_1 \urcorner} \cap \ulcorner h \urcorner \cap \{s\colon a \in s\}\right) \cup \left(\ulcorner g_1 \urcorner \cap \ulcorner k \urcorner \cap \{s\colon a \in s\}\right)\\
=_{\text{comm.}} \quad & \left(\left(\overline{\ulcorner g_0 \urcorner} \cap \{s\colon a \notin s\} \cap \ulcorner h \urcorner\right) \cup \left(\overline{\ulcorner g_1 \urcorner} \cap \{s\colon a \in s\} \cap \ulcorner h \urcorner\right)\right) \cup\\
& \qquad \left(\left(\ulcorner g_0 \urcorner \cap \{s\colon a \notin s\} \cap \ulcorner k \urcorner\right) \cup \left(\ulcorner g_1 \urcorner \cap \{s\colon a \in s\} \cap \ulcorner k \urcorner\right)\right)\\
=_{\text{distr.}} \quad & \left(\left(\left(\overline{\ulcorner g_0 \urcorner} \cap \{s\colon a \notin s\}\right) \cup \left(\overline{\ulcorner g_1 \urcorner} \cap \{s\colon a \in s\}\right)\right) \cap \ulcorner h \urcorner\right) \cup\\
& \qquad \left(\left(\left(\ulcorner g_0 \urcorner \cap \{s\colon a \notin s\}\right) \cup \left(\ulcorner g_1 \urcorner \cap \{s\colon a \in s\}\right)\right) \cap \ulcorner k \urcorner\right)\\
= \quad & \left(\ulcorner a\left(\overline{g_0},\overline{g_1}\right) \urcorner \cap \ulcorner h \urcorner\right) \cup \left(\ulcorner a\left(g_0,g_1\right) \urcorner \cap \ulcorner k \urcorner\right)\\
= \quad & \left(\ulcorner \overline{a\left(g_0,g_1\right)} \urcorner \cap \ulcorner h \urcorner\right) \cup \left(\ulcorner a\left(g_0,g_1\right) \urcorner \cap \ulcorner k \urcorner\right).
\end{aligned}
$$

This concludes the proof. $\dashv$

Corollary 5.3.6 *The equivalence relation $\equiv$ on Boolean forms is a congruence relation on $\mathfrak{B}$ and, consequently, $\mathfrak{B}/_{\equiv}$ is a properly defined quotient algebra.*

Proof: Immediate from Lemma 5.3.5 and the fact that $\ulcorner g \urcorner$ is a set, for each Boolean form g. $\dashv$

On the basis of Lemma 5.3.5, the algebra of strategic forms *modulo* equivalence $\boldsymbol{\mathfrak{B}}_A$ is defined as follows:

$$
\boldsymbol{\mathfrak{B}}_A \quad =_{df.} \quad \left(\left\{\ulcorner g \urcorner :\ g \in B(A)\right\}; \varnothing, 2^A, {}^{-}, \cup, \cap\right).
$$

Suppressing the subscript A, the algebra $\mathfrak{B}_A$ is a *field of sets*[3] and, as such, also a Boolean algebra. Moreover, $\mathfrak{B}/_\equiv$ and $\mathfrak{B}$ are (trivially) isomorphic *via* the natural isomorphism, which maps each $[g]_\equiv$ onto $\ulcorner g \urcorner$. Let $B(A)$ be a set of Boolean forms on a set of binary decision variables A along with the classical propositional language $L(A)$. An inspection of Proposition 5.3.5 reveals that $\mathfrak{B}_A$ coincides with the extension algebra $\mathfrak{E}_A$ of the classical propositional language $L(A)$ (*cf.*, page 49). As the latter being isomorphic to the Lindenbaum algebra $\mathfrak{L}_A$ for $L(A)$ (*cf.*, page 49), so are both $\mathfrak{B}/_\equiv$ and $\mathfrak{B}_A$. Hence the following theorem.

Theorem 5.3.7 *Let A be a set and $L(A)$ and $B(A)$ be the classical propositional language over A and the set of Boolean forms on A, respectively. Then, $\mathfrak{B}_A$ coincides with the extension algebra $\mathfrak{E}_A$. Consequently, $\mathfrak{B}_A$, $\mathfrak{B}_A/_\equiv$, $\mathfrak{E}_A$ and the Lindenbaum algebra $\mathfrak{L}_A$ are pairwise isomorphic.*

Sketch of proof: Straightforward. The first claim is by a trivial inductive argument on the complexity of Boolean forms and propositional variables. Moreover, $\mathfrak{B}/_\equiv$ and $\mathfrak{B}$ are isomorphic *via* the isomorphism that maps $[g]_\equiv$ onto $\ulcorner g \urcorner$. Observe in this respect, that this map is bijective in virtue of the definition of $\equiv$. Finally, the algebras $\mathfrak{B}_A$, $\mathfrak{B}_A/_\equiv$, $\mathfrak{E}_A$ and the Lindenbaum algebra $\mathfrak{L}_A$ being pairwise isomorphic then follows immediately from $\mathfrak{E}_A$ and $\mathfrak{L}_A$ being identical. $\dashv$

In virtue of this theorem we can assume for each formula φ of a classical propositional language $L(A)$ there to be a Boolean form g_φ, and, *vice versa*, with each Boolean form g a formula φ_g such that:

$$[\![\varphi]\!] = \ulcorner g_\varphi \urcorner \qquad \text{and} \qquad \ulcorner g \urcorner = [\![\varphi_g]\!].$$

As an immediate consequence of Proposition 5.3.5, the following fact is obtained.

Fact 5.3.8 *Let A be a set, φ, ψ formulas in $L(A)$ and $g, h, k \in B$. Then:*

$$\begin{array}{rclrcl}
\ulcorner a\,(\mathbf{0},\mathbf{1})\urcorner & = & [\![a]\!] & \ulcorner \circledast\,(g,h,k)\urcorner & = & [\![\,(\neg\varphi_g \wedge \varphi_h) \vee (\varphi_g \wedge \varphi_k)\,]\!] \\
\ulcorner \mathbf{0} \urcorner & = & [\![\bot]\!] & \ulcorner \mathbf{1} \urcorner & = & [\![\top]\!] \\
\ulcorner \overline{g} \urcorner & = & [\![\neg\varphi_g]\!] & [\![\neg\varphi]\!] & = & \ulcorner \overline{g_\varphi} \urcorner \\
\ulcorner g+h \urcorner & = & [\![\varphi_g \vee \varphi_h]\!] & [\![\varphi \vee \psi]\!] & = & \ulcorner g_\varphi + g_\psi \urcorner \\
\ulcorner g \cdot h \urcorner & = & [\![\varphi_g \wedge \varphi_h]\!] & [\![\varphi \wedge \psi]\!] & = & \ulcorner g_\varphi \cdot g_\psi \urcorner.
\end{array}$$

We say that the formula φ_g represents *the Boolean form g and that the boolean form g_φ* represents *the formula φ.*

Proof: An immediate consequence of Proposition 5.3.5. $\dashv$

On the basis of this fact and Proposition 5.3.3, for each Boolean form g a propositional formula can straightforwardly be formulated that is equivalent to φ_g. On

[3] A *field of sets* S is a collection of subsets of a nonempty set X such that both the empty set $\emptyset$ and the set X are in S and S is closed under $\cap$, $\cup$ and $^-$ with respect to X (Chang and Keisler, 1973, p.39).

page 125 we observed that the game $a\big(b\,(\mathbf{1}, c\,(\mathbf{1}, \mathbf{0}))\,, b\,(\mathbf{0}, \mathbf{1})\big)$ of our example is identical to a combination of the Boolean forms $a\,(\mathbf{0}, \mathbf{1})$, $b\,(\mathbf{0}, \mathbf{1})$ and $b\,(\mathbf{0}, \mathbf{1})$ and the operations $\mathbf{0}$, $\mathbf{1}$ and $\circledast$. By Fact 5.3.8, we find in $\big(\neg a \wedge (\neg b \vee (b \wedge c))\,\big) \vee \big(a \wedge b\big)$ a formula that is equivalent to its logical representant. Some Boolean manipulations, gives the equivalent but simpler $(a \leftrightarrow b) \vee (\neg a \wedge \neg c)$ as a suitable logical representative of the Boolean form. The Boolean forms $a\,(b\,(\mathbf{0}, \mathbf{1})\,, b\,(\mathbf{1}, \mathbf{0}))$ and $a\,(b\,(\mathbf{1}, \mathbf{0})\,, b\,(\mathbf{0}, \mathbf{1}))$ of the typically indeterminate Boolean games can then be found to correspond to, respectively, the formulas $a \leftrightarrow b$ and $a \leftrightarrow \neg b$.

5.4 Evaluation Games

The considerations of the previous section give rise to the interpretation of Boolean games as a kind of evaluation game. Each Boolean form corresponds to a propositional formula, the decision variables to propositional variables and the strategy profiles to valuations for the respective propositional language. Furthermore, the roles of two players of a Boolean game can be construed as those of a *verifier* and of a falsifier *falsifier* of the formula representing the Boolean form in question. The verifier endeavors to satisfy the formula by finding appropriate values for the propositional variables assigned to her control, and the falsifier tries to make the formula false by choosing appropriate values for the propositional variables in his control. Whether a player has a winning strategy or a winning response given a particular formula then depends on the set of propositional variables assigned to her.

At this point a remark is in order with respect to the logical evaluation games advanced by Hintikka and Sandu (Hintikka, 1973 and Hintikka and Sandu, 1997) and their relation to Boolean games. They suggest a game-theoretical semantics for first-order logic, in line with their observation that "... mathematical logicians have spontaneously resorted to game-theoretical conceptualization practically every time they have had to deal wiht a kind of logic where Tarski-type truth definitions do not apply, including branching quantifiers languages, game quantifier languages and infinitely deep languages" (Hintikka and Sandu, 1997, p. 363).

Game-theoretical semantics (GTS) defines for each first-order formula φ, each first-order model $\mathfrak{A}$, and each assignment function f, a two-player game of strict competition — denoted by $G\,(\varphi, \mathfrak{A}, f)$. This definition is by recursion on the formula φ. The two players of the game play the antagonistic roles of verifier and falsifier of a formula. Let $R\,(t_0, \ldots, t_n)$, be an atomic formula. Then, the verifier wins the game $G\,(R\,(t_0, \ldots, t_n)\,, \mathfrak{A}, f)$ if $R\,(t_0, \ldots, t_n)$ is satisfied in $\mathfrak{A}$ with respect to the assignment function f. For molecular formulas the principal logical constant involved determines which player is to make a move. In the game $G\,(\varphi_0 \wedge \varphi_1, \mathfrak{A}, f)$ the falsifier chooses a conjunct φ_i, after which the game is continued playing the game $G\,(\varphi_i, \mathfrak{A}, f)$. The game $G\,(\varphi_0 \vee \varphi_1, \mathfrak{A}, f)$ is similar to that for conjunction, except that now it is up to the verifier to choose a disjunct. For games for the quantifiers follow the same pattern. Let Q a universal or an existential quantifier. Then, in a game $G\,((Qx)\,\varphi, \mathfrak{A}, f)$, one of the players selects an object a from the domain of $\mathfrak{A}$, after which the game is continued playing

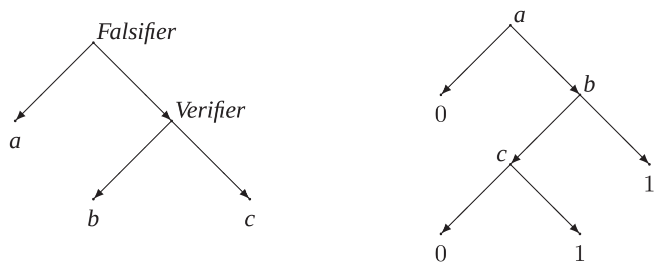


Figure 5.5. The propositional formula $a \wedge (b \vee c)$ represented as, repsectively, a Hintikka-Sandu evaluation game (left) and as a Boolean game (right). Whether *Verifier* has a winning strategy in the former depends on the background model underlying the game. *E.g.*, in the valuation $\{a, b\}$ she has, but in $\{a\}$ she has not. In contrast, whether a player has a winning strategy in the Boolean game, depends on the control over the propostional variables assigned to him. *E.g.*, if Player 1 has control over both a and b, she has a winning strategy. Not so, if she has control over b and c.

$G(\varphi, \mathfrak{A}, f[x/a])$. The verifier is to choose if Qx is the existential quantifier $\exists x$ and the falsifier otherwise, *i.e.*, if Qx is the universal quantifier $\forall x$. In the game $G(\neg\varphi, \mathfrak{A}, f)$ the players swap roles and play is continued with $G(\varphi, \mathfrak{A}, f)$. The result that forms the point of departure for the researches in GTS is then that for the classical first-order languages there is an intimate link with the Tarskian or truth-functional interpretation. For $\vDash$ the Tarskian satisfaction relation it is then the case that:

$$\mathfrak{A}, f \vDash \varphi \quad \text{iff} \quad \text{the player in the role of verifier has a winning strategy in } G(\varphi, \mathfrak{A}, f).$$

As a consequence a formula φ is logically valid if the verifier has a winning strategy for the evaluation game for φ on *all* models $\mathfrak{A}$.

Like in Boolean games, in the evaluation games of GTS a verifier and a falsifier vie for the truth-value of a formula. However, when restricted to a classical first-order language, the GTS evaluation games are of perfect information, and, hence, they are in general determined. The classical law of excluded middle is often seen as to reflect this fact. In the Boolean framework, a formula of the form $\varphi \vee \neg\varphi$ defines a game that will always be won by one of the players. We took, however, pains to point out that it is not in general the case that Boolean games are determined. For propositional languages, the GTS game determined by a formula differs structurally from the Boolean form defined by the same formula, as Figure 5.5 illustrates. Furthermore, the ‘powers’ of the players are determined by the logical constants in the GTS-framework. By contrast, what a player can achieve in a Boolean game, also depends on the set of propositional variables he has been assigned control over.

For languages other than that of classical first-order logic, however, the GTS-framework does allow for imperfect information. Due to the linear representation of formulas in the classical notation, the scopes of two quantifiers are forced to be either exclusive or nested. This restriction — which is equally arbitrary as spurious according to Hintikka and Sandu (Hintikka and Sandu 1997, p. 366) — is lifted, if quantifiers are allowed to "branch", as, *e.g.*, in:[4]

$$\begin{pmatrix} \forall x & \exists y \\ \forall z & \exists u \end{pmatrix} \varphi(x, y, z, u).$$

Here the quantifier $\exists y$ is thought to be in the scope of $\forall x$, but not in that of $\forall z$. Moreover, $\forall z$ is also thought to be independent of $\exists y$. This interdependence and independence of quantifiers can also be given a linear representation using the so-called *slash* notation. The formula above then becomes:

$$(\forall x)(\forall z)(\exists y/\forall z)(\exists u/\forall x)\,\varphi.$$

This notation can be generalized in that a quantifier can be "slashed" by any quantifier of the opposite type that occurs to the left of it in the formula in question. The evaluation game $G((\exists x/\forall y_0, \ldots, \forall y_n)\,\varphi, \mathfrak{A}, f)$ is then an imperfect information game, in which the verifier has to choose a value for the variable x *unbeknownst* of the values her opponent has chosen for the variables $y_0, \ldots, y_n$ earlier in the game.

If branching quantifiers are also allowed in quantified propositional logic, it turns out that for each Boolean game we can find a formula φ such that player 1 has a winning strategy in the Boolean game if and only if in the evaluation game $G(\varphi, s)$ the verifier has a winning strategy as well, where, s is any valuation of the propositional language. For (g, π) a Boolean game in the decision variables $a_0, \ldots, a_n, b_0, \ldots, b_m$ and π assigning control over $a_0, \ldots, a_m$ to player 0 and that over $b_0, \ldots, b_m$ to player 1, the corresponding formula in quantified propositional logic is obtained as:

$$\begin{pmatrix} \forall a_0, \ldots, \forall a_n \\ \exists b_0, \ldots, \exists b_m \end{pmatrix} \varphi_g(a_0, \ldots, a_n, b_0, \ldots, b_m).$$

We leave this claim here without a proof. In the quantified formula above, the prefix $\binom{\forall a_0, \ldots, \forall a_n}{\exists b_0, \ldots, \exists b_m}$ plays a similar role as the control assignment function in a Boolean game: both determine which player has control over which variables. However, for our purposes it turns out to be more convenient to deal with the distribution of control at the meta-level, not in the least because it facilitates the generalization of the concept of distributed control over propositional variables to situations in which multiple players interact. This issue will be addressed in Part III of this thesis.

In GTS the verifier having a winning strategy in the evaluation games $G(\varphi, \mathfrak{A}, f)$ for all models $\mathfrak{A}$ and all assignment functions f means that φ is valid. For Boolean games, likewise, there is a relation between a propositional formula φ being valid and player 1 having a winning strategy in a Boolean game on the Boolean form g_φ. This correspondence, however, only obtains in general if player 1 has control over no propositional variables occurring in φ. For the other distributions the propositional variables

[4] Branching quantifiers were proposed for the first time in Henkin (1961).

we find that the correspondence holds only with respect to a generalized notion of logical validity. In this way the notion of distributed control over propositional variables assumes a logical significance.

The next chapter will be devoted to the issue of how distribution of control over propositional variables relates to the logical properties of formulas. In particular the logical counterparts of the game-theoretical issues mentioned in the introduction of this chapter will be addressed. First, *given a Boolean form and a player, which distributions of the decision variables yield a Boolean game in which that player has a winning strategy?* Second, *given a distribution of the decision variables, in which Boolean games complying with this distribution has the one player a winning strategy, in which the other and in which neither of them?*

Chapter 6

Propositional Logic for Control

6.1 Introduction

In the previous chapter Boolean games based on a set of binary decision variables A were proved to entertain an intimate relation with the formulas of the propositional language with the same set A as propositional variables. The isomorphism between the algebra of (strategic forms of) Boolean forms in $B(A)$ and the Lindenbaum algebra of the corresponding propositional language $L(A)$ elicited the interpretation of Boolean games as a special kind of logical evaluation game. By choosing values for the decision variables assigned to them, the two players construct a valuation with respect to which the formula corresponding to the Boolean game in question should be evaluated. The one player strives for a valuation that verifies the formula, whereas her opponent aims at a valuation that renders it false. On this basis, game-theoretical and logical concepts can be matched.

Boolean forms correspond to propositional formulas, strategy profiles to valuations for the propositional variables and a win for player 1 in a Boolean form to the truth of the formula associated with that Boolean form. Furthermore, the propositional connectives obtain a game-theoretical significance as operations on Boolean forms. This kind of correspondence is not peculiar to Boolean games; congenial ones are central to the game-theoretical analyses of logical concepts in the framework of Hintikka and Sandu (Hintikka (1973), Hintikka and Sandu (1997)) and in that of Lorenzen's (Lorenzen and Lorenz (1978)). The conformity of game-theoretic and logical notions reappear at the level of the solution concepts. In a Hintikka-Sandu evaluation game for a formula φ and a background model $\mathfrak{A}$, *Verifier* having a winning strategy corresponds to φ being true in the model $\mathfrak{A}$. In Lorenzen's writings, a formula φ is derivable in a formal system if and only if the so-called *Proponent* has a winning strategy in the corresponding game.

The controversy between the two players of a Boolean game could be said to be over the truth value of a propositional formula where both players exercise control over disjoint sets propositional variables. Thus, in a Boolean game are distinguished a Boolean form and an assignment of the decision variables to the players. Generally, the manipulative powers of the players depend on both of these components. The Boolean form determines which player wins for each strategy profile and control over a greater number of propositional variables will usually help a player attain his goals.

The classical logical notions of validity and satisfiability are to be reencountered in the extreme cases where either the Boolean form assigns victory to one player for all strategy outcomes or where control over the propositional variables is concentrated in just one player.

In the first of these extreme cases, the Boolean form correspond to a propositional tautology or a contradiction. Hence, the validity of a propositional formula signifies that the corresponding game cannot otherwise but result in a victory for Player 1.

The other extreme is if one of the players has control over all propositional variables. If Player 1 disposes over *all* propositional variables, the problem she faces is that of classical satisfiability. If she has a winning strategy in such a game, values for the propositional variables that render the formula true can be found and the formula is classically satisfiable. If she is unable to find such values for the propositional variables, the formula is unsatisfiable. Similarly, the validity of a propositional formula signifies that in the corresponding Boolean game the player 1 has a winning strategy even if she has control over no propositional variable whatsoever. Hence, in these two extreme cases the game-theoretical concept of a winning strategy has logical counterparts in validity and satisfiability.

From a game-theoretical angle the interest of these extreme cases is only limited. If the Boolean form is tautology or a contradiction, the game hardly need to be played, as the outcome is fixed from the outset. On the other hand, if control over the propositional variables is concentrated in one player, the game reduces to a one-person game without any interaction to speak of.

Strategic and game-theoretical reasoning is rather about what a player can achieve relative to the powers and preferences of the opponent. The correspondence between Boolean forms and propositional formulas, as enunciated in the previous chapter, is one way of bringing strategic themes under the heading of classical propositional logic. We will argue that the issue of limited control over decision or propositional variables, as exemplified by Boolean games, motivates the study a generalized notion of logical consequence.

From this point of view, of special interest are the intermediate cases in which each player has control over a proper subset of the variables and in which the Boolean form corresponds to neither tautology nor contradiction. Then, however, the concept of a player having a winning strategy is no longer guaranteed to have a well-known counterpart in a traditional notion of classical logic. Taking seriously the game-theoretical perspective on logic as provided by Boolean games, this is an unsatisfactory state of affairs. In an effort filling up this lacuna between logical and game-theoretical concepts, one could parameterize the concepts of validity and satisfiability by a subset of propositional variables. Intuitively, a formula φ is valid relative to such a subset Δ, if there is a choice of values for the propositional variables in Δ such that φ holds in all valuations that comply with this choice. Dually, a formula φ is said to be satisfiable relative to a subset Δ if, for each assignment of values to the variables in Δ, there is a complementary choice of values for the variables outside Δ such that the resulting valuation satisfies φ. Hence, *relative to* $\{a\}$, the formula $a \vee b$ is both satisfiable and valid, $a \leftrightarrow b$ is satisfiable but not valid and $a \wedge b$ is neither satisfiable nor valid. Analogously, the concepts of Δ-refutability and Δ-unsatisfiability are introduced. In this manner, the idea of partial control over propositional variables is accounted for. We find that the classical notions of satisfiability, refutability and validity are all borderline cases of *both* relativized validity and relativized satisfiability. For the intermediate, *i.e.*, non-borderline, cases, they correspond to the game-theoretical concepts of a player having a winning strategy or a winning response in a Boolean game. *E.g.*, Player 1 turns out to have a winning strategy in a Boolean game (g, π) if and only if the formula corresponding with g is valid with respect to the set of propositional variables that π assigns to Player 1. Player 0 has a winning strategy in (g, π) if and only if the corresponding formula is unsatisfiable relative to the set of propositional variables π assigns to him.

In the theory of two-person games of pure conflict the concept of *determinacy* plays a central role. A game is said to be determined if one of the players has a winning strategy. One of the first game-theoretical results, due to Zermelo (1913), proved the determinacy of two-person games of perfect information. In the previous chapter, we argued that Boolean games are not determined in the above sense and had better be understood as games of imperfect information. Hence, the question which Boolean games are determined is not settled trivially by Zermelo's theorem. Via the correspondence between Boolean forms and propositional formulas, determinacy of Boolean games also has an immediate logical counterpart. For Δ a subset of propositional variables assigned to player 1 by the control assignment function π, we say a formula φ is Δ-*determined* if and only if the Boolean game (g_φ, π) is determined.

Boolean forms correspond to formulas and in the remarks above (relativized) validity and satisfiability were likewise thought of as merely applying to formulas. The central idea of distributed control over the variables of a propositional language can, however, be extrapolated so as to apply to properties of theories as well. In particular, logical consequence can also be relativized to a subset of propositional variables.

Rather than a binary relation between theories, relativized logical consequence is introduced as a ternary relation obtaining between two theories and a subset of propositional variables. Classical logical consequence is a borderline case: it coincides with the generalized notion if made relative to the empty set. Alternatively, the set of propositional variables could be seen as a *parameter*. Viewed thus, the relativizing of logical consequence is a *family* of consequence relations,[1] which can be ordered as a complete lattice. Classical consequence is then a limiting case, *viz.*, the bottom of the lattice.

The relativized concept of logical consequence is given a semantical definition in terms of valuations and subsets of propositional variables. Two propositional theories Γ and Θ thus connected relative to a subset of propositional variables Δ will be

[1]Recall that we took *any* relation between the theories of a propositional language as a consequence relation.

denoted by $\Gamma \models_\Delta \Theta$. The demand for sound and complete formal characterizations of this relativized consequence relation now pushes to the fore.

The issue of completeness, however, may be approached from two conceptually different angles. On the one hand one may emphasize the consequence relation each subset of propositional variables defines. The question is then, given a fixed subset of propositional variables Δ, between which pairs of theories the consequence relation parameterized by Δ holds. This issue is analogous to the classical problem of formal systems for propositional logic. As a matter of fact, any sound and complete formal systems for classical propositional logic may also be deployed as a calculus in which the desired results can be obtained. (*cf.*, Proposition 6.3.5, below). In Section 6.4 a Gentzen-style system for relativized consequence is presented.

On the other hand, one may focus on the subsets of propositional variables relative to which a particular theory follows from another. The problem is then to produce, for any given pair of theories Γ and Θ, the subsets of propositional variables Δ for which it is the case that $\Gamma \models_\Delta \Theta$.

The relevance of the latter system for Boolean games is that it provides a generalized answer to the question which are the minimal sets of decision variables over which one should have control in order to be able to win a Boolean game, *i.e.*, the minimal sets of decision variables that furnishes a player with a winning strategy. Moreover, the system also specifies the winning strategy itself, *i.e.*, not only does it give a set of decision variables control over which suffices to win the game, but it also produces actual values for those variables that win the game.

6.2 Relativized Validity and Satisfiability

A statement as to the classical validity of a formula makes a universal claim on the set of valuations: a formula is valid if and only if it is forced in *all* valuations. In a similar vein, the existential quantifier implicit in statements claiming the satisfiability of a formula likewise ranges over all valuations.

If it were someone's aim to construct a valuation that forces a particular a formula, it may suffice to be able to set the values of only some propositional variables, leaving the values of the other variables to the whims of Providence or even to the vindictiveness of an antagonist. One may need control over even fewer propositional variables, if in a similar situation one can make one's choice for the values dependent on those of one's adversary. In either case, no control over any variables at all is required if the formula in question is valid. By contrast, control over all variables may be needed if it is merely satisfiable in the classical sense. Of course, similar remarks are in order if it is someone's ambition to falsify a formula rather than to satisfy it. Again, his success in achieving his objective may depend on the set of propositional variables he has control over.

These considerations suggest a refinement of the classical quadripartite classification of formulas in terms of their being valid, refutable, satisfiable and unsatisfiable. These logical properties of formulas can be made dependent on a set of propositional variables.

$$\varphi \text{ is } \Delta\textit{-valid} \text{ iff for some } s \in S, \text{for all } s' \in S\colon s \sim_\Delta s' \text{ implies } s' \Vdash \varphi,$$

$$\varphi \text{ is } \Delta\textit{-unsatisfiable} \text{ iff for some } s \in S, \text{for all } s' \in S\colon s \sim_\Delta s' \text{ implies } s' \nVdash \varphi.$$

As dual notions we then obtain:

$$\varphi \text{ is } \Delta\textit{-satisfiable} \text{ iff for all } s \in S, \text{for some } s' \in S\colon s \sim_\Delta s' \text{ and } s' \Vdash \varphi,$$

$$\varphi \text{ is } \Delta\textit{-refutable} \text{ iff for all } s \in S, \text{for some } s' \in S\colon s \sim_\Delta s' \text{ and } s' \nVdash \varphi.$$

If one has control over no variables whatsoever, then one is entirely at the mercy of whether φ is valid or not. On the other hand, with total control over the propositional variables one can validate any formula provided that it be satisfiable. Similar remarks apply to the concept of refutability and unsatisfiability. The following proposition recapitulates these observations.

Proposition 6.2.1 *Let φ be a formula of some propositional language $L(A)$. Then:*

$$\begin{array}{rcccl} \varphi \textit{ is } \emptyset\textit{-valid} & \textit{iff} & \varphi \textit{ is } A\textit{-satisfiable} & \textit{iff} & \varphi \textit{ is classically valid}, \\ \varphi \textit{ is } \emptyset\textit{-unsatisfiable} & \textit{iff} & \varphi \textit{ is } A\textit{-refutable} & \textit{iff} & \varphi \textit{ is classically unsatisfiable}, \\ \varphi \textit{ is } \emptyset\textit{-satisfiable} & \textit{iff} & \varphi \textit{ is } A\textit{-valid} & \textit{iff} & \varphi \textit{ is classically satisfiable}, \\ \varphi \textit{ is } \emptyset\textit{-refutable} & \textit{iff} & \varphi \textit{ is } A\textit{-unsatisfiable} & \textit{iff} & \varphi \textit{ is classically refutable}. \end{array}$$

Proof: Immediately from the fact that ε_A is the identity relation and $\varepsilon_\emptyset$ the universal relation on 2^A, the set of valuations for $L(A)$. ⊣

The definitions of the relativized notions of validity and unsatisfiability of a formula evince a strong resemblance with the definitions of a player having a winning strategy in a Boolean game. This impression is vindicated in the following proposition.

Proposition 6.2.2 *Let φ be a formula in a propositional language $L(A)$ and let π be a partition of A, with π_1 the set of propositional variables assigned to player 1 and π_0 those to player 0. Then:*

$$\varphi \textit{ is } \pi_1\textit{-valid iff Player } 1 \textit{ has a winning strategy in } (g_\varphi, \pi),$$

$$\varphi \textit{ is } \pi_0\textit{-unsatisfiable iff Player } 0 \textit{ has a winning strategy in } (g_\varphi, \pi).$$

Proof: Almost immediately from the fact that $[\![\varphi]\!] = \ulcorner g_\varphi \urcorner$, Definition 6.3.1, the definition of a player having a winning strategy (*cf.*, page 120) and that of a player having a winning response (*cf.*, page 121) in a Boolean game. ⊣

As an immediate consequence of this proposition and the fact at most one player can have a winning strategy in a Boolean game, the following implications also hold.

$$\varphi \text{ is } \Delta\text{-valid implies } \varphi \text{ is } \overline{\Delta}\text{-satisfiable},$$

$$\varphi \text{ is } \Delta\text{-unsatisfiable implies } \varphi \text{ is } \overline{\Delta}\text{-refutable}.$$

Taking $a \leftrightarrow b$ or $a \leftrightarrow \neg b$ for φ and $\{a\}$ for Δ, moreover, provides a suitable counterexample against the inverse claims.

So, a formula is Δ-valid if and only if in the corresponding Boolean form player 1 has a winning strategy if assigned control over Δ. Similarly, a formula is Δ-unsatisfiable if player 0 has a winning strategy in the corresponding game, provided he decide over the values of the variables in Δ. This makes that the concept of a Boolean game being determined can also be expressed in logical terms. A Boolean game (g, π) is determined if one of the players has a winning strategy, *i.e.*, if the corresponding formula φ_g is either π_1-valid or π_0-unsatisfiable. It seems reasonable to extend the application of the concept of determinacy to formulas. Thus define for φ a formula of some propositional language $L(A)$:

$$\varphi \text{ is } \Delta\textit{-determined} \quad \text{iff} \quad \varphi \text{ is either } \Delta\text{-valid or } \overline{\Delta}\text{-unsatisfiable.}$$

It follows that a formula φ is Δ determined if and only if its being $\overline{\Delta}$-satisfiable implies its being Δ-valid and, equally, if and only if its being $\overline{\Delta}$-refutable implies its being Δ-unsatisfiable. A formula is called Δ-*indeterminate* if it is not Δ-determined. This then is a logical concept that is immediately inspired by the game-theoretical light that Boolean games shed on propositional logic and one of which the investigation is appropriate, if this perspective is taken seriously.

In the limiting cases in which Δ is either the whole set of propositional variables or empty, all formulas are determined. Because of Proposition 6.2.1, however, this is merely giving expression to the bland propositional facts that a formula is either valid or refutable, or that a formula is either satisfiable or unsatisfiable. For the intermediate cases, however, the set of formulas of a propositional language will never be exhausted by the set of determined formulas.

Proposition 6.2.3 *Let $L(A)$ be a classical propositional language and let $\Delta \subseteq A$. Then:*

$$\textit{the set of } \Delta\textit{-indeterminate formulas is non-empty} \quad \textit{iff} \quad \varnothing \subsetneq \Delta \subsetneq A.$$

Proof: The left-to-right direction is immediate by Proposition 6.2.1; merely consider the contrapositive. For the opposite direction, we may assume there to be propositional variables a and b in A such that $a \in \Delta$ and $b \notin \Delta$. Then, consider, *e.g.*, $a \leftrightarrow b$, which is neither Δ-valid nor $\overline{\Delta}$-unsatisfiable and hence Δ-indeterminate. $\dashv$

Thus for each non-empty proper subset Δ of propositional variables, we have as archetypical examples of indeterminate formulas $a \leftrightarrow b$ and $a \leftrightarrow \neg b$, provided that $a \in \Delta$ and $b \notin \Delta$.

6.3 Relativized Logical Consequence

The relativized concepts of validity and satisfiability of Section 6.2 can be seen as appropriate notions of validity and satisfiability for a common relativized notion of consequence. Classical consequence is defined as a relation between theories Γ and Θ, which, informally, holds whenever at least one formula in Θ holds whenever all of the formulas in Γ do. Formally, classical consequence was introduce as follows:

$$\Gamma \vDash^{\text{CPC}} \Theta \text{ iff for all } s \in S\colon \text{ if } \ s \Vdash \gamma, \text{ for all } \gamma \in \Gamma, \ \text{ then } \ s \Vdash \vartheta, \text{ for some } \vartheta \in \Theta.$$

This definition can also be given the more succinct formulation suggested on page 46:

$$\Gamma \vDash^{\text{CPC}} \Theta \quad \text{iff} \quad [\![\Gamma]\!] \subseteq \langle\!\langle \Theta \rangle\!\rangle.$$

We propose to make classical consequence dependent on a subset of propositional variables, writing $\Gamma \vDash_\Delta \Theta$ if the theory Θ follows from the theory Γ with respect to the set of propositional variables Δ. Intuitively, $\Gamma \vDash_\Delta \Theta$ holds if it is possible to assign values to the propositional variables in Δ such that any valuation complying with this assignment forces at least one formula in Θ whenever it forces all formulas in Γ as well. Formally we have the following definition:

Definition 6.3.1 *(Relativized consequence)* For Γ and Θ theories in a propositional language $L(A)$, define:

$$\Gamma \vDash_\Delta \Theta$$
$$\text{iff}$$
$$\exists s, \forall s' \text{ such that } s \sim_\Delta s'\colon \forall \gamma \in \Gamma : \ s' \Vdash \gamma \ \text{ implies } \ \exists \vartheta \in \Theta\colon s' \Vdash \vartheta.$$

Each subset Δ of propositional variables in A determines a proper logic Λ_Δ defined as $\Lambda_\Delta =_{df.} \{(\Gamma, \Theta) : \ \Gamma \vDash_\Delta \Theta\}$.

As an immediate consequence of this definition, a formula φ is Δ-valid if and only if $\varnothing \vDash_\Delta \varphi$ and φ is Δ-unsatisfiable if and only if $\varphi \vDash_\Delta \varnothing$. Similarly, φ is Δ-satisfiable if and only if $\varphi \nvDash_\Delta \varnothing$ and φ is Δ-refutable if and only if $\varnothing \nvDash_\Delta \varphi$.

Despite its rather tortuous formulation, relativized consequence can more intuitively be understood as a *localized* generalization of classical consequence. Rather than requiring $[\![\Gamma]\!]$ to be a subset of $\langle\!\langle \Theta \rangle\!\rangle$, *per se*, the relativized notion merely demands that this inclusion holds in one of a distinguished set of subsets of valuations. The relevant subsets of valuations for $\vDash_\Delta$ are given by the partition π_Δ. Recall that π_Δ is the partition of the set of valuations induced by the equivalence relation ε_Δ, *viz.*, the equivalence relation that holds between all valuations that agree on their values for the variables in Δ (*cf.*, page 40). A theory Γ entails another theory Θ relative to a subset of propositional variables Δ if there is some block of π_Δ within which the inclusion of $[\![\Gamma]\!]$ in $\langle\!\langle \Theta \rangle\!\rangle$ holds. The following is an equivalent characterization of relativized consequence (*cf.*, Figure 6.1).

$$\Gamma \vDash_\Delta \Theta \quad \text{iff} \quad \text{for some } X \in \pi_\Delta\colon \ [\![\Gamma]\!] \cap X \subseteq \langle\!\langle \Theta \rangle\!\rangle \cap X.$$

In the extreme case that Δ is empty, π_Δ is the trivial partition containing the whole set of valuations 2^A itself as the only block. Hence, the classical consequence relation can readily be seen to coincide with the $\Lambda_\varnothing$.

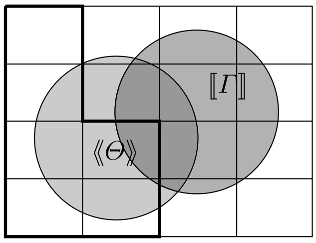


Figure 6.1. Logical space partitioned by ε_Δ. Here, $\Gamma \vDash_\Delta \Theta$, because in each of the blocks X within the area demarcated by the thick lines $[\![\Gamma]\!] \cap X \subset \langle\!\langle\Theta\rangle\!\rangle \cap X$.

Fact 6.3.2 *Let Γ and Θ be propositional theories in a propositional language $L(A)$. Then:*

$$\Gamma \vDash \Theta \quad \textit{iff} \quad \Gamma \vDash_\emptyset \Theta.$$

Proof: As for Proposition 6.2.1, merely observe that ε_A is the identity relation and $\varepsilon_\emptyset$ the universal relation on 2^A. ⊣

If Δ is a subset of Δ', the partition $\pi_{\Delta'}$ refines π_Δ. As a consequence, relativized consequence is upward monotonic in the set of propositional variables, in the sense that $\Delta \subseteq \Delta'$ implies $\Lambda_\Delta \subseteq \Lambda_{\Delta'}$. Also, since it is still the case that $[\![\Gamma]\!]$ shrinks and $\langle\!\langle\Gamma\rangle\!\rangle$ grows with Γ becoming larger, relativized propositional consequence is monotonic. The following proposition recapitulates this observation.

Proposition 6.3.3 *Let Γ, Γ', Θ and Θ be theories in $L(A)$ such that $\Gamma \subseteq \Gamma'$ and $\Theta \subseteq \Theta'$. Let further Δ' and Δ' be subsets of A such that $\Delta \subseteq \Delta'$. Then:*

$$\Gamma \vDash_\Delta \Theta \quad \textit{implies} \quad \Gamma' \vDash_{\Delta'} \Theta'.$$

Proof: Straightforward. Assume $\Gamma \vDash_\Delta \Theta$. Then, $X \cap [\![\Gamma]\!] \subseteq X \cap \langle\!\langle\Theta\rangle\!\rangle$, for some $X \in \pi_\Delta$. Consider this X. Since, $\Delta \subseteq \Delta'$, by Fact 2.2.6, also $\pi_{\Delta'} \leqslant \pi_\Delta$. Hence there is some $X' \in \pi_{\Delta'}$ with $X' \subseteq X$. Consequently also $X' \cap [\![\Gamma]\!] \subseteq X' \cap \langle\!\langle\Theta\rangle\!\rangle$. Because $\Gamma \subseteq \Gamma'$ and $\Theta \subseteq \Theta'$ also $[\![\Theta']\!] \subseteq [\![\Gamma]\!]$ and $\langle\!\langle\Theta\rangle\!\rangle \subseteq \langle\!\langle\Theta'\rangle\!\rangle$. We may conclude that $X' \cap [\![\Gamma']\!] \subseteq X' \cap \langle\!\langle\Theta'\rangle\!\rangle$ and therefore $\Gamma' \vDash_{\Delta'} \Theta'$. ⊣

Relativized consequence can alternatively be construed as a *family* $\{\Lambda_\Delta\}_{\Delta\subseteq A}$ of consequence relations, indexed by subsets of propositional variables. Define the ordering $\leqslant$ on $\{\Lambda_\Delta\}_{\Delta\subseteq A}$ by set-inclusion, *i.e.*, such that for all subsets Δ and Δ' of propositional variables:

$$\Lambda_\Delta \leqslant \Lambda_{\Delta'} \quad \text{iff} \quad \Lambda_\Delta \subseteq \Lambda_{\Delta'}.$$

We then obtain the following fact.

Fact 6.3.4 *Let $L(A)$ a propositional language with $\Delta, \Delta' \subseteq A$. Then:*

$$\Lambda_\Delta \leqslant \Lambda_{\Delta'} \quad \textit{iff} \quad \Delta \subseteq \Delta'.$$

Proof: The right-to-left direction is immediate by Proposition 6.3.3. For the opposite direction assume $\Delta \not\subseteq \Delta'$. Then, there exists some $a \in \Delta$ such that $a \notin \Delta'$. We then find that $\varnothing \vDash_\Delta a$, *i.e.*, $(\varnothing, a) \in \Lambda_\Delta$ but $\varnothing \nvDash_{\Delta'} a$, *i.e.*, $(\varnothing, a) \notin \Lambda_{\Delta'}$. $\dashv$

As an immediate corollary, then, $\left(\{\Lambda_\Delta\}_{\Delta \in 2^A}, \leqslant\right)$ is a complete lattice — even a Boolean algebra — with classical consequence as bottom element. The greatest lower bound and the least upper bound of a set $\{\Lambda_{\Delta_i}\}_{i\in I}$ are given by, respectively, $\Lambda_{\bigcap_{i\in I}\Delta_i}$ and $\Lambda_{\bigcup_{i\in I}\Delta_i}$, *i.e.*,

$$\bigwedge_{i\in I} \Lambda_{\Delta_i} = \Lambda_{\bigcap_{i\in I}\Delta_i} \qquad\qquad \bigvee_{i\in I} \Lambda_{\Delta_i} = \Lambda_{\bigcup_{i\in I}\Delta_i}.$$

Observe, however, that join and meet are not in general given by union and intersection, respectively. *I.e.*, it is not generally the case that $\Lambda_\Delta \wedge \Lambda_{\Delta'}$ equals $\Lambda_\Delta \cap \Lambda_{\Delta'}$ or that $\Lambda_\Delta \vee \Lambda_{\Delta'}$ coincides with $\Lambda_\Delta \cup \Lambda_{\Delta'}$. For a counterexample consider let a and b be distinct propositional variables. Then, $\varnothing \vDash_{\{a\}} a \vee b$ as well as $\varnothing \vDash_{\{b\}} a \vee b$. However, $\varnothing \nvDash_\varnothing a \vee b$, although clearly $\{a\} \cap \{b\} = \varnothing$.

Each block in a partition π_Δ, where Δ is a subset of propositional variables, can be characterized by a theory consisting of literals over A. This makes that each statement of the form $\Gamma \vDash_\Delta \Theta$ correspond to a statement $\Gamma' \vDash^{\mathrm{CPC}} \Theta$ in classical propositional logic.

Proposition 6.3.5 *Let Γ and Θ be theories of a propositional language $L(A)$ and let $\Delta \subseteq A$. Then:*

$$\Gamma \vDash_\Delta \Theta \quad \textit{iff} \quad \textit{for some } \Delta' \subseteq \Delta\colon\ \Gamma \cup (\Delta - \Delta') \vDash \Theta \cup \Delta'.$$

Proof: For the left-to-right direction assume $\Gamma \vDash_\Delta \Theta$. Then there is a valuation s such that for all valuations s' with $s \sim_\Delta s'$, either for some $\gamma \in \Gamma$, $s' \nVdash \gamma$ or for some $\vartheta \in \Theta$, $s' \Vdash \vartheta$. Consider this valuation s and define $\Delta' =_{df.} \Delta - s$. Then, $\Delta - \Delta' = \Delta \cap s$. Now consider an arbitrary valuation s^* and assume, for a *reductio ad absurdum* that both $s^* \in [\![\Gamma \cup (\Delta - \Delta')]\!]$ and $s^* \notin \langle\!\langle \Theta \cup \Delta' \rangle\!\rangle$. Observe that $s^* \Vdash d$ for all $d \in \Delta - \Delta'$, and $s^* \nVdash d$, for all $d \in \Delta'$. Hence, $s \sim_\Delta s^*$ and by the initial assumption, $s^* \nVdash \gamma$, for some $\gamma \in \Gamma$, or $s^* \Vdash \vartheta$, for some $\vartheta \in \Theta$. This, however, is at variance with the assumption that both $s^* \in [\![\Gamma \cup (\Delta - \Delta')]\!]$ and $s^* \notin \langle\!\langle \Theta \cup \Delta' \rangle\!\rangle$.

For the opposite direction, suppose there be some $\Delta' \subseteq \Delta$ such that $\Gamma \cup (\Delta - \Delta') \vDash \Theta \cup \Delta'$. Consider the valuation s^{**} defined by $s^{**} =_{df.} \Delta - \Delta'$. Now consider an arbitrary valuation s' such that $s^{**} \sim_\Delta s'$. Then, $s' \Vdash d$, for all $d \in \Delta - \Delta'$, and $s' \nVdash d$, for all $d \in \Delta'$. Still, by the initial assumption, either $s' \nVdash \varphi$, for some $\varphi \in \Gamma \cup \Delta - \Delta'$, or $s' \Vdash \varphi$, for some $\varphi \in \Theta \cup \Delta'$. In view of what must hold in s' for the propositional variables in $\Delta - \Delta'$ and those in Δ', this φ should be sought among Γ and Θ, *i.e.*, either $s' \nVdash \gamma$ for some $\gamma \in \Gamma$ or $s' \Vdash \vartheta$ for some $\vartheta \in \Theta$. We may conclude that $\Gamma \vDash_\Delta \Theta$. ⊣

In virtue of Proposition 6.3.5 many of the formal properties of classical propositional consequence are inherited by each of the relativized consequence relations. Thus we have the following corollaries.

Corollary 6.3.6 *Let Γ and Θ be theories and φ and ψ be formulas in a propositional language $L(A)$. Also let Δ be a subset of A. Then:*

$$\begin{aligned}
\Gamma \cup \{\varphi\} \vDash_\Delta \Theta \quad &\textit{iff} \quad \Gamma \vDash_\Delta \Theta \cup \{\neg\varphi\} \\
\Gamma \cup \{\neg\varphi\} \vDash_\Delta \Theta \quad &\textit{iff} \quad \Gamma \vDash_\Delta \Theta \cup \{\varphi\} \\
\Gamma \cup \{\varphi, \psi\} \vDash_\Delta \Theta \quad &\textit{iff} \quad \Gamma \cup \{\varphi \wedge \psi\} \vDash_\Delta \Theta \\
\Gamma \vDash_\Delta \Theta \cup \{\varphi, \psi\} \quad &\textit{iff} \quad \Gamma \vDash_\Delta \Theta \cup \{\varphi \vee \psi\}
\end{aligned}$$

Proof: Almost immediately from Proposition 6.3.5. Here, we only give the proof of the first. Consider the following equivalences:

$$\begin{aligned}
&\Gamma \cup \{\varphi\} \vDash_\Delta \Theta \\
&\text{iff}_{\text{Prop. 6.3.5}} \quad \text{for some } \Delta' \subseteq \Delta\colon \Gamma \cup \{\varphi\} \cup \Delta - \Delta' \vDash \Theta \cup \Delta' \\
&\text{iff}_{\text{CPC}} \quad \text{for some } \Delta' \subseteq \Delta\colon \Gamma \cup \Delta - \Delta' \vDash \Theta \cup \{\neg\varphi\} \cup \Delta' \\
&\text{iff}_{\text{Prop. 6.3.5}} \quad \Gamma \vDash_\Delta \Theta \cup \{\neg\varphi\}\,.
\end{aligned}$$

The other cases run along analogous lines. ⊣

Corollary 6.3.7 *Let $L(A)$ be a propositional language containing φ and ψ as formulas such $[\![\varphi']\!] \subseteq [\![\varphi]\!]$ and $[\![\psi]\!] \subseteq [\![\psi']\!]$ Let further Δ be any subset of A. Then for theories Γ and Θ:*

$$\Gamma \cup \{\varphi\} \vDash_\Delta \Theta \cup \{\psi\} \quad \textit{implies} \quad \Gamma \cup \{\varphi'\} \vDash_\Delta \Theta \cup \{\psi'\}\,.$$

Proof: Immediately from Proposition 6.3.5. ⊣

In a much similar fashion it can be argued that both of the following:

$$\begin{aligned}
&\Gamma \cup \{\varphi \vee \psi\} \vDash_\Delta \Theta \ \text{ implies } \ \Gamma \cup \{\varphi\} \vDash_\Delta \Theta \text{ and } \Gamma \cup \{\psi\} \vDash_\Delta \Theta, \\
&\Gamma \vDash_\Delta \Theta \cup \{\varphi \wedge \psi\} \ \text{ implies } \ \Gamma \vDash_\Delta \Theta \cup \{\varphi\} \ \text{ and } \Gamma \vDash_\Delta \Theta \cup \{\psi\}\,.
\end{aligned}$$

The converses of these latter two claims do not in general hold. Assume for instance that both $\Gamma \vDash_\Delta \Theta \cup \{\varphi\}$ and $\Gamma \vDash_\Delta \Theta \cup \{\psi\}$. Then for some blocks X and Y of the partition π_Δ, we have that $X \cap [\![\Gamma]\!] \subseteq X \cap \langle\!\langle \Theta \cup \{\varphi\} \rangle\!\rangle$ and $Y \cap [\![\Gamma]\!] \subseteq Y \cap \langle\!\langle \Theta \cup \{\psi\} \rangle\!\rangle$.

There is, however, no guarantee that the blocks X and Y are identical or, indeed, that, for any block Z in π_Δ, it is the case that both $Z \cap [\![\Gamma]\!] \subseteq Z \cap \langle\!\langle \Theta \cup \{\varphi\} \rangle\!\rangle$ and $Z \cap [\![\Gamma]\!] \subseteq Z \cap \langle\!\langle \Theta \cup \{\psi\} \rangle\!\rangle$. It is easy to find a direct counterexample. Observe that both $\varnothing \vDash_{\{a\}} a$ and $\varnothing \vDash_{\{a\}} \neg a$. Nevertheless, $\varnothing \nvDash_{\{a\}} a \wedge \neg a$. Using much the same example, the *cut* rule can also seen not to hold in general. Again, both $a \vDash_{\{a\}} \varnothing$ and $\varnothing \vDash_{\{a\}} a$, but $\varnothing \nvDash_{\{a\}} \varnothing$.

6.4 Formal Systems for Relativized Logical Consequence

The conclusion of the previous chapter held out the prospect of the resolution of three issues relating to Boolean games. The first of these concerns the sets of decision variables control over which suffices for a player to have a winning strategy in a particular Boolean game. The second issue concerns the Boolean games a player can win given control over a particular set of decision variables.

Proposition 6.2.2 on page 139 establishes the correspondence between a player having a winning strategy in a particular Boolean game and a formula being valid or unsatisfiable with respect to a subset of propositional variables. This makes that the issues mentioned above can be approached from a logical angle. The two issue can also more generally be formulated in terms of relativized consequence, *viz.*, *for fixed theories Γ and Θ, relative to which subsets Δ of propositional variables does the entailment $\Gamma \vDash_\Delta \Theta$ hold good?*, and *given a subset Δ of propositional variables, for which theories Γ and Θ is it the case that $\Gamma \vDash_\Delta \Theta$?* This section deals, in reverse order, with these two problems.

Sequent Calculus for Relativized Consequence

With for each subset Δ of propositional variables the proper logic Λ_Δ being semantically fixed, the question for which theories Γ and Θ it is the case that $\Gamma \vDash_\Delta \Theta$, has already been answered trivially. This leaves, however, the matter of a sound and complete syntactical characterization of the proper logics Λ_Δ. This issue we take up in this subsection, proposing for each logic Λ_Δ a sound and complete sequent calculus — denoted by GPC_Δ — which is very similar to the classical system GPC (*cf.*, page 52).

The axioms and rules for GPC_Δ are those of GPC (*cf.*, page 52, above) with, in addition, the following two axiom schemas:

$$(L_d) \quad d \Rightarrow \epsilon \qquad \text{and} \qquad (R_d) \quad \epsilon \Rightarrow d,$$

where d is assumed to be in Δ. Table 6.4 summarizes the system GPC_Δ.

For each subset of propositional variables Δ the system GPC_Δ inherits from GPC the left rule for disjunction ($\vee_L$), the right rule for conjunction ($\wedge_R$) as well as *cut*. This may seem odd as in the previous section we argued that the semantical counterpart of *cut* does not in general hold and that $\Gamma \cup \{\varphi\} \vDash_\Delta \Theta$ and $\Gamma \cup \{\psi\} \vDash_\Delta \Theta$ does not generally entail $\Gamma \cup \{\varphi \vee \psi\} \vDash_\Delta \Theta$. Neither do $\Gamma \vDash_\Delta \Theta \cup \{\varphi\}$ and $\Gamma \vDash_\Delta$

Axioms:

$$(0)\ \bot \Rightarrow \epsilon \quad (1)\ \epsilon \Rightarrow \top \quad (2)\ a \Rightarrow a \quad (L_d)\ d \Rightarrow \epsilon \quad (R_d)\ \epsilon \Rightarrow d$$

Provided that $d \in \Delta$ in (L_d) and (R_d).

Logical Rules:

$$\neg_L : \ \frac{\Sigma \Rightarrow T, \varphi}{\Sigma, \neg\varphi \Rightarrow T} \qquad \neg_R : \ \frac{\Sigma, \varphi \Rightarrow T}{\Sigma \Rightarrow T, \neg\varphi}$$

$$\wedge_L : \ \frac{\Sigma, \varphi, \psi \Rightarrow T}{\Sigma, \varphi \wedge \psi \Rightarrow T} \qquad \wedge_R : \ \frac{\Sigma \Rightarrow T, \varphi \qquad \Sigma \Rightarrow T, \psi}{\Sigma \Rightarrow T, \varphi \wedge \psi}$$

$$\vee_L : \ \frac{\Sigma, \varphi \Rightarrow T \qquad \Sigma, \psi \Rightarrow T}{\Sigma, \varphi \vee \psi \Rightarrow T} \qquad \vee_R : \ \frac{\Sigma \Rightarrow T, \varphi, \psi}{\Sigma \Rightarrow T, \varphi \vee \psi}$$

Structural Rules:

$$contr_L : \ \frac{\Sigma, \varphi, \varphi \Rightarrow T}{\Sigma, \varphi \Rightarrow T} \qquad contr_R : \ \frac{\Sigma \Rightarrow T, \varphi, \varphi}{\Sigma \Rightarrow T, \varphi}$$

$$perm_L : \ \frac{\Sigma, \varphi, \psi, P \Rightarrow T}{\Sigma, \psi, \varphi, P \Rightarrow T} \qquad perm_R : \ \frac{\Sigma \Rightarrow T, \varphi, \psi, \Upsilon}{\Sigma \Rightarrow T, \psi, \varphi, \Upsilon}$$

$$thin_L : \ \frac{\Sigma \Rightarrow T}{\Sigma, \varphi \Rightarrow T} \qquad thin_R : \ \frac{\Sigma \Rightarrow T}{\Sigma \Rightarrow T, \varphi}$$

$$cut : \ \frac{\Sigma \Rightarrow T, \varphi \qquad \Sigma, \varphi \Rightarrow T}{\Sigma \Rightarrow T}$$

Table 6.4. The System GPC$_\Delta$. In each derivation for each $d \in \Delta$ at most one of the axioms (L_d) and (R_d) may be used.

$\Theta \cup \{\psi\}$ imply $\Gamma \vDash_\Delta \Theta \cup \{\varphi \wedge \psi\}$. We find that a slight modification of the notion of a derivation in GPC$_\Delta$ keeps in check possible ill-effects of $\vee_L$, $\wedge_R$ and *cut*.

A *derivation* of a sequent $\Sigma \Rightarrow_{\Delta,\Delta'} T$ in GPC$_\Delta$ is defined as usual (*cf.*, page 51, above), be it that for each *d* in Δ at most one of the axioms (L_d) and (R_d) may be employed. The intuition behind the axioms (L_d) and (R_d) are that player 1 — who has control over all propositional variables in Δ — can set its value of *d* to either one or zero. This is precisely what it means to control a propositional variable. It is, however, impossible to set the value of Δ to both one and zero, at the same time. This is reflected in the restriction that only one of the axioms (L_d) and (R_d) may be employed in each derivation in GPC$_\Delta$. It can now also be understood why the presence of $\vee_L$, $\wedge_R$ and *cut* does not jeopardize the soundness of GPC$_\Delta$. The antecedents of these rules — *viz.*, $\Gamma \cup \{\varphi\} \vDash_\Delta \Theta$ and $\Gamma \cup \{\psi\} \vDash_\Delta \Theta$, $\Gamma \vDash_\Delta \Theta \cup \{\varphi\}$ and $\Gamma \cup \vDash_\Delta \Theta \cup \{\psi\}$ and $\Gamma \vDash_\Delta \Theta \cup \{\varphi\}$ and $\Gamma \cup \{\varphi\} \vDash_\Delta \Theta$ — may be valid and hence also derivable in GPC$_\Delta$. However, in the derivation of on of the members of any such pair an axiom (L_d) may (will) occur for some $d \in \Delta$ whereas (R_d) occurs in the derivation of the other member. Then, these derivations cannot be combined so as to obtain a derivation of the consequent of the rule.

In order to prove the soundness and completeness of GPC$_\Delta$ with respect to Λ_Δ we first introduce some notation and obtain an auxiliary result.

For Δ and Δ' *disjoint* subsets of propositional variables, *Δ-Δ'-consequence* is the relation $\vDash_{\Delta,\Delta'}$ such that for all theories Γ and Θ:

$$\Gamma \vDash_{\Delta,\Delta'} \Theta \quad \text{iff} \quad \Delta' \cup \Gamma \vDash^{\mathrm{CPC}} \Theta \cup \Delta.$$

Now consider the sequent system GPC$_{\Delta,\Delta'}$ as the classical sequent system GPC augmented with the axioms (L_d) for each *d* in Δ as well as the axioms (R_d) for each *d* in Δ'. A derivation in GPC$_{\Delta,\Delta'}$ is defined as it was for GPC on page 51 — *i.e.*, *without* the unusual restriction on the application of the axioms (L_d) and (R_d). We write $\Gamma \vdash^{\mathrm{GPC}}_{\Delta,\Delta'} \Theta$ for $\Gamma \vdash^{\mathrm{GPC}_{\Delta,\Delta'}} \Theta$.

The system GPC$_{\Delta,\Delta'}$ is sound with respect to Δ-Δ'-consequence. With respect to the rules also in GPC this is obvious. For the soundness of the two additional axiom schemas, merely observe that classically $\Delta' \cup \{a\} \vDash^{\mathrm{CPC}} \Delta$ and $\Delta' \vDash^{\mathrm{CPC}} \{b\} \cup \Delta$, if $a \in \Delta$ and $b \in \Delta'$. Therefore, by definition, also $\{a\} \vDash_{\Delta,\Delta'} \varnothing$ and $\varnothing \vDash_{\Delta,\Delta'} b$. We now have the following lemma.

Lemma 6.4.1 *Let Γ and Θ be theories in a propositional language $L(A)$. Let further Δ and Δ' be disjoint subsets of A. Then:*

$$\Delta' \cup \Gamma \vdash^{\mathrm{GPC}} \Theta \cup \Delta \quad \textit{iff} \quad \Gamma \vdash^{\mathrm{GPC}}_{\Delta,\Delta'} \Theta.$$

Sketch of proof: First assume $\Delta' \cup \Gamma \vdash^{\mathrm{GPC}} \Theta \cup \Delta$. Without loss of generality we may assume Δ, Γ, Θ and Δ' to be finite. Observe that each derivation in GPC is also a derivation in GPC$_{\Delta,\Delta'}$. Hence, $\Delta' \cup \Gamma \vdash^{\mathrm{GPC}}_{\Delta,\Delta'} \Theta \cup \Delta$. For any $E \cup \{d\} \subseteq \Delta$ and any $E' \cup \{d'\} \subseteq \Delta'$, we have $\{d\} \cup E' \cup \Gamma \vdash^{\mathrm{GPC}}_{\Delta,\Delta'} \Theta \cup E$ as well as $E' \cup \Gamma \vdash^{\mathrm{GPC}}_{\Delta,\Delta'}$

$\Theta \cup E \cup \{d'\}$. Both these claims hold in virtue of the axioms (L_d) and $(R_{d'})$. Since, $\mathrm{GPC}_{\Delta,\Delta'}$ contains *cut* we can show by a simple inductive argument, here omitted, that also $\Gamma \vdash^{\mathrm{GPC}}_{\Delta,\Delta'} \Theta$.

For the opposite direction, assume $\Gamma \vdash^{\mathrm{GPC}}_{\Delta,\Delta'} \Theta$. Then there is a derivation $\mathscr{D}$ in $\mathrm{GPC}_{\Delta,\Delta'}$ witnessing this fact. Replacing in $\mathscr{D}$ all axioms $\epsilon \Rightarrow a$ and $a \Rightarrow \epsilon$ by $a \Rightarrow a$ yields a derivation $\mathscr{D}^*$ in GPC. An easy inductive argument to the length of $\mathscr{D}^*$, here omitted, reveals that $\mathscr{D}^*$ witnesses $\Delta' \cup \Gamma \vdash^{\mathrm{GPC}} \Theta \cup \Delta$. ⊣

The ground has now been cleared for the following soundness and completeness result.

Theorem 6.4.2 *(Soundness and Completeness of* GPC_Δ*) Let Γ and Θ be theories in a propositional language $L(A)$ and let $\Delta \subseteq A$. Then,* GPC_Δ *is sound and complete with respect to Λ_Δ, i.e.:*

$$\Gamma \vdash^{\mathrm{GPC}_\Delta} \Theta \quad \textit{iff} \quad \Gamma \vDash_\Delta \Theta.$$

Proof: For soundness, assume that $\Gamma \vdash^{\mathrm{GPC}_\Delta} \Theta$ and let $\mathscr{D}$ be a derivation witnessing this fact. Let further $\Delta_0^{\mathscr{D}}$ be the set of propositional variables d such that the axiom (L_d) occurs in $\mathscr{D}$. Similarly, let $\Delta_1^{\mathscr{D}}$ contain precisely those propositional variables d such that the axiom (R_d) is employed in $\mathscr{D}$. By definition of a derivation in GPC_Δ both $\Delta_0^{\mathscr{D}}$ and $\Delta_1^{\mathscr{D}}$ are subsets of Δ. Moreover, by definition, for no propositional variable d the derivation $\mathscr{D}$ invokes both (L_d) and (R_d). Consequently, $\Delta_0^{\mathscr{D}}$ and $\Delta_1^{\mathscr{D}}$ are disjoint. It follows that $\mathscr{D}$ is a also derivation in $\mathrm{GPC}_{\Delta_0^{\mathscr{D}},\Delta_1^{\mathscr{D}}}$ witnessing $\Gamma \vdash^{\mathrm{GPC}}_{\Delta_0^{\mathscr{D}},\Delta_1^{\mathscr{D}}} \Theta$. By Lemma 6.4.1 then, $\Delta_1^{\mathscr{D}} \cup \Gamma \vdash^{\mathrm{GPC}} \Theta \cup \Delta_0^{\mathscr{D}}$, and by completeness of GPC with respect to CPC (Fact 2.3.12), also $\Delta_1^{\mathscr{D}} \cup \Gamma \vDash^{\mathrm{CPC}} \Theta \cup \Delta_0^{\mathscr{D}}$. In virtue of Proposition 6.3.5 and $\Delta_0^{\mathscr{D}}$ and $\Delta_1^{\mathscr{D}}$ being disjoint, $\Gamma \vDash_{\Delta_0^{\mathscr{D}} \cup \Delta_1^{\mathscr{D}}} \Theta$. Then by monotonicity of relativized consequence (Proposition 6.3.3), and $\Delta_0^{\mathscr{D}}$ and $\Delta_1^{\mathscr{D}}$ both being subsets of Δ, eventually, $\Gamma \vDash_\Delta \Theta$.

For completeness, assume $\Gamma \vDash_\Delta \Theta$. By Proposition 6.3.5, there exists some $\Delta' \subseteq \Delta$ such that $\Delta - \Delta' \cup \Gamma \vDash^{\mathrm{CPC}} \Theta \cup \Delta$. Hence, $\Delta - \Delta' \cup \Gamma \vdash^{\mathrm{GPC}} \Theta \cup \Delta$, because of completeness of GPC. By Lemma 6.4.1 also $\Gamma \vdash^{\mathrm{GPC}}_{\Delta',\Delta-\Delta'} \Theta$. Now observe that — with Δ' and $\Delta - \Delta'$ being disjoint subsets of Δ — each derivation in $\mathrm{GPC}_{\Delta',\Delta-\Delta'}$ is also a derivation in GPC_Δ. Consequently, $\Gamma \vdash^{\mathrm{GPC}_\Delta} \Theta$, which concludes the proof. ⊣

Minimal Propositional Bases

In the previous sections of this chapter the emphasis has been on the question which theory follows from another theory relative to some given subset of propositional variables. Here, the emphasis will be shifted to the question relative which subsets of propositional variables a given theory follows from another. Because of the monotonicity of relativized consequence, this issue can naturally be rephrased as which are the *minimal* subsets of propositional variables required to guarantee one theory follow from another. Such concerns are reasonable in contexts in which the variables are thought of as economic commodities, the acquisition of which might be expensive.

Moreover, having procured the necessary resources — possibly at great cost — little has been gained if one does not know how to deploy them.

A statement of the form $\Gamma \vDash_\Delta \Theta$ imparts the existence of a choice for the propositional variables in Δ that guarantees either one of the formulas in Γ to be false, or, otherwise, at least one of the formulas in Θ to be true. Yet, it is left entirely uncommented how this choice should be made. Similarly, the previous section was devoted to formally characterizing those formulas that are Δ-valid, Δ-unsatisfiable and Δ-determined. Game-theoretically, this could be interpreted as a singling out of Boolean games in which the player with control over Δ has a winning strategy or in which at least one of the players has a winning strategy. In the definition of a player *having* a winning strategy (*cf.*, page 120) the strategy that is actually winning is quantified away. Still, from the perspective of one of the players of a Boolean game, one might be more interested in the actual strategies that win a game than in the abstract existence of one. One could imagine a player getting a bit cranky at being told that there is a winning strategy for him, without being told what it looks like.[2]

These considerations are the informal background to the remaining part of this section. Here we will be concerned with an inductive definition that more generally associates each pair of theories with a set of *pairs* of subsets of propositional variables. Any such pair (Δ, Δ') in the set for the pair of propositional theories (Γ, Θ), is such that in all valuations that falsify all variables in Δ, verify all those in Δ' either one of the formulas in Γ is false or one of the formulas in Θ is true, *i.e.*, if it is *classically* the case that (*cf.*, Proposition 6.3.5):

$$\Delta' \cup \Gamma \vDash \Delta \cup \Theta.$$

Adopting the notation of the previous section, what we are after is, for each pair of theories Γ and Θ, the set $\left\{(\Delta, \Delta') : \ \Gamma \vDash_{\Delta,\Delta'} \Theta\right\}$.

As an auxiliary notion define, for Γ and Θ theories in a propositional language $L(A)$:

$$[\![\Gamma;\Theta]\!] \quad =_{df.} \quad \bigcup_{\gamma\in\Gamma} [\![\Gamma]\!] \cup \bigcup_{\vartheta\in\Theta} [\![\vartheta]\!].$$

Through writing out the definitions in full, we find for each valuation s in 2^A for $L(A)$ that:

$$s \in [\![\Gamma;\Theta]\!] \quad \text{iff} \quad \text{for all } \gamma \in \Gamma,\ s \Vdash \gamma \ \text{ implies } \ \text{for some } \vartheta \in \Theta,\ s \Vdash \vartheta.$$

[2]There is also another reason. Let $\boldsymbol{\Delta}(\Gamma,\Theta)$ be a temporary notation for set $\{\Delta \subseteq A : \ \Gamma \vDash_\Delta \Theta\}$, and with $\boldsymbol{\Delta}(\varphi)$ and $\boldsymbol{\nabla}(\varphi)$ for $\boldsymbol{\Delta}(\varnothing, \{\varphi\})$ and $\boldsymbol{\nabla}(\{\varphi\}, \varnothing)$, respectively. We find that it is impossible to provide $\boldsymbol{\Delta}(\varphi)$ with a neat compositional definition in φ. Then $\boldsymbol{\Delta}(a \wedge \neg a)$ coincides with the emptyset and as such is distinct from $\boldsymbol{\Delta}(a \wedge a)$, which is given by $\{\Delta \subseteq A : \ a \in A\}$. This difference, however, cannot be accounted for on the basis of $\boldsymbol{\Delta}(a)$, $\boldsymbol{\Delta}(\neg a), \boldsymbol{\nabla}(a)$ and $\boldsymbol{\nabla}(\neg a)$ alone, which all are identical to $\{\Delta \subseteq A : \ a \in A\}$. A similar argument shows that no compositional definition of $\boldsymbol{\Delta}(\Gamma,\Theta)$, depending only on the sets $\boldsymbol{\Delta}(\varphi)$ and $\boldsymbol{\nabla}(\varphi)$, for $\varphi \in \Gamma \cup \Theta$. To appreciate this consider $\boldsymbol{\Delta}(\{a\}, \{a\})$ and $\boldsymbol{\Delta}(\{a\}, \{\neg a\})$; the former is given by 2^A whereas the latter coincides with the distinct $\{\Delta \subseteq A : \ a \in A\}$. It should be remarked, however, that the method we employ is not compositional either.

Now define $[\![\Gamma;\Theta]\!]$ as the *smallest* subset $\boldsymbol{X}$ of $2^A \times 2^A$, such that:

$$\{(\overline{s}, s) : \ s \in [\![\Gamma;\Theta]\!]\} \subseteq \boldsymbol{X}, \quad \text{and}$$
$$\text{if } \ E' \subseteq E \ \text{ and } \ (\Delta \cup (E - E'), \Delta' \cup E') \in \boldsymbol{X}, \text{ for all } E' \subseteq E, \ \text{ then } \ (\Delta, \Delta') \in \boldsymbol{X}.$$

We then have the following proposition.

Proposition 6.4.3 *Let Γ and Θ be theories in a propositional language $L(A)$ and Δ and Δ'* disjoint *subsets of A. Then:*

$$(\Delta, \Delta') \in [\![\Gamma;\Theta]\!] \quad \textit{iff} \quad \Delta' \cup \Gamma \vDash^{\text{CPC}} \Theta \cup \Delta.$$

Proof: For the right-to-left direction assume $\Delta' \cup \Gamma \vDash^{\text{CPC}} \Theta \cup \Delta$ and consider an arbitrary $Y \subseteq \overline{\Delta \cup \Delta'}$. Let $s^* =_{df.} \Delta' \cup Y$. Observe that $\Delta' \subseteq s^*$ and that Δ and s^* are disjoint. We first prove that $s^* \in [\![\Gamma;\Theta]\!]$.

Assume that $s^* \Vdash \gamma$, for all $\gamma \in \Gamma$. With $\Delta' \subseteq s^*$ also $s^* \Vdash \gamma$, for all $\gamma \in \Delta' \cup \Gamma$. By the initial assumption then also $s^* \Vdash \vartheta$ for some $\vartheta \in \Theta \cup \Delta$. Since, s^* and Δ are disjoint, $s^* \Vdash \vartheta$, for some $\vartheta \in \Theta$. Hence, $s^* \in [\![\Gamma;\Theta]\!]$ and, *a fortiori*, $\left(\overline{s^*}, s^*\right) \in [\![\Gamma;\Theta]\!]$.

Then, with Y having been chosen arbitrarily, $\left(\overline{\Delta' \cup Y}, \Delta' \cup Y\right) \in [\![\Gamma;\Theta]\!]$, for all $Y \subseteq \overline{\Delta \cup \Delta'}$. Since $Y \subseteq \overline{\Delta \cup \Delta}$, also $\Delta \subseteq \overline{Y}$ and that, with Δ and Δ' disjoint, $\Delta \subseteq \overline{\Delta'}$. Hence, $\Delta \subseteq \overline{\Delta'} \cap \overline{Y}$ and with some Boolean reasoning:

$$\begin{aligned} \overline{\Delta' \cup Y} \ &= \ \overline{\Delta'} \cap \overline{Y} \ =_{\Delta \subseteq \overline{\Delta'} \cap \overline{Y}} \ \left(\Delta \cup \overline{\Delta'}\right) \cap \left(\Delta \cup \overline{Y}\right) \ = \ \Delta \cup \left(\overline{\Delta'} \cap \overline{Y}\right) \\ &= \ \left(\Delta \cup \overline{\Delta}\right) \cap \left(\Delta \cup \left(\overline{\Delta'} \cap \overline{Y}\right)\right) \ = \ \Delta \cup \left(\overline{\Delta} \cap \overline{\Delta'} \cap \overline{Y}\right) \ = \ \Delta \cup \left(\left(\overline{\Delta \cup \Delta'}\right) - Y\right). \end{aligned}$$

Accordingly, $\left(\Delta \cup \left(\left(\overline{\Delta \cup \Delta'}\right) - Y\right), \Delta' \cup Y\right) \in [\![\Gamma;\Theta]\!]$, for all $Y \subseteq \overline{\Delta \cup \Delta'}$, and so, eventually $(\Delta, \Delta') \in [\![\Gamma;\Theta]\!]$.

The opposite direction assume $(\Delta, \Delta') \in [\![\Gamma;\Theta]\!]$. We prove by induction on (Δ, Δ') that then also $\Delta' \cup \Gamma \vDash^{\text{CPC}} \Theta \cup \Delta$.

So, first assume $(\Delta, \Delta') = (\overline{s}, s)$, for some valuation $s \in [\![\Gamma;\Theta]\!]$ and consider an arbitrary valuation t such that $t \Vdash \gamma$, for all $\gamma \in \Delta' \cup s$. In case $t = s$, we are done immediately by definition. So assume t be distinct from s. Still, $t \Vdash \gamma$, for all $\gamma \in s \cup \Gamma$ and, hence, $s \subseteq t$. With t distinct from s there should be some $a \in \overline{s}$ such that $a \in t$. Therefore, $t \Vdash \vartheta$, for some $\gamma \in \Theta \cup \overline{s}$. With t having been chosen arbitrarily, we may conclude that $s \cup \Gamma \vDash^{\text{CPC}} \Theta \cup \overline{s}$. Hence, $\Delta' \cup \Gamma \vDash^{\text{CPC}} \Theta \cup \Delta$.

For the inductive step, assume that $(\Delta, \Delta') \in [\![\Gamma;\Theta]\!]$ in virtue of the existence of a subset E of A such that $(\Delta \cup (E - E'), \Delta' \cup E') \in [\![\Gamma;\Theta]\!]$, for all $E' \subseteq E$. By the induction hypothesis, also $\Delta' \cup E' \cup \Gamma \vDash^{\text{CPC}} \Theta \cup \Delta \cup (E - E')$, for all $E' \subseteq E$. Fact 2.3.11 on page 51, above, then yields $\Delta' \cup \Gamma \vDash^{\text{CPC}} \Theta \cup \Delta$. $\dashv$

6.5 Conclusion

The framework of Boolean games evoke a number of logical question *via* the correspondence between Boolean forms and propositional formulas. These considerations gave rise a generalization of classical consequence to a notion of consequence relativized by a subset of propositional variables. This subset of propositional variables was considered to be in the control of one player; the values of the remainder of the propositional variables were left to the whims of an opponent with antagonistic preferences. Thus the notion of distributed control over propositional variables has been central to our approach.

So far, control over the propositional variables has been thought of as being divided over two players. In the next part, we will consider the logical consequences of distributing control over multiple players. This will require further modifications of the logical framework. Moreover, it becomes natural to employ in the logical analysis game-theoretical solution concepts better suited for dealing with multi-player environments than that of a winning strategy.

Part III

Game Theoretical Consequence

Chapter 7

Winning Consequence

7.1 Introduction

Logical consequence is traditionally explained in terms of truth. We introduced classical consequence as a relation between theories. Intuitively, a theory Θ follows classically from another theory Γ if and only if the truth of all of the formulas in Γ implies the truth of at least one of the formulas in Θ. Each theory Γ is associated the set of extensions of the formulas it contains. This set of extensions is denoted by $\mathscr{E}(\Gamma)$ and formally defined as $\{[\![\gamma]\!] : \gamma \in \Gamma\}$, where $[\![\gamma]\!]$ denotes the set of valuations in which γ holds, for each γ. In terms of the sets of extensions classical propositional consequence a sound and complete semantics is obtained by defining:

$$\Gamma \vDash^{\mathrm{CPC}} \Theta \quad \text{iff} \quad \bigcap \{[\![\gamma]\!] : \gamma \in \Gamma\} \subseteq \bigcup \{[\![\vartheta]\!] : \vartheta \in \Theta\}.$$

What this characterization comes down to is that for Γ and Θ, two sets of valuations are singled out — by taking the intersection and union of $\mathscr{E}(\Gamma)$ and $\mathscr{E}(\Theta)$, respectively — and subsequently compared with respect to set inclusion.

This part concerns a type of consequence relation that can likewise be characterized as set inclusion between sets of valuations associated with the respective theories. These consequence relations, however, differ from the classical account in that the sets of valuations associated with the theories are essentially selected on the basis of a game-theoretical definition. The underlying idea is that, by distributing control over the propositional variables among a number of players, logical space assumes the structure of the frame of a strategic game, with the valuations as strategy profiles. We then argue that theories and formulas can be seen as providing an additional preferential structure, enabling us to use game-theoretical solution concepts to select sets of valuations. For different distributions of control and different theories these sets can be compared. The role of the solution concepts is thus analogous to that of union and intersection in the classical semantics.

In the informal interpretation that accompanies classical propositional logic, the variables are thought of as conveying information about the state the world is in. A valuation for the language could then be seen as a kind of oracle — used here in the informal sense of the word — yielding the values of the propositional variables for some possible state of affairs. As oracles befits, it is quite beside the point how they are possible.

Here, as in the previous part on Boolean games, we assume a different perspective. The world can very well be thought of as something we can change and manipulate. What is true of a certain situation then depends on the actions and decisions of the individuals that live and act in it. The relative manipulative powers of the individuals, however, may widely diverge. Some may be able to change the world in certain ways, others in other. These considerations give rise to the idea of propositional variables as binary decision variables controlled by individuals. A valuation is thus construed as resulting from particular choices individuals may make with respect to their decision variables, rather than as a record of some possible unalterable state of affairs. Assuming, moreover, that individuals in a social context do not usually operate in isolation, the decisions they make in this respect can be thought of as the result of an interactive process. Our point of departure is that this notion of *control* over propositional variables can be made subject to logical analysis and that the employment of game-theoretical techniques in this comes naturally. Since there is nothing in the notion of control over propositional variables that requires its distribution to be restricted to one or two players, our logical analyses need not be restricted to the two-person case. Eventually they well comprise the general case in which control over the propositional variables may be distributed over any countable number of players.

This way of viewing propositional variables as controllable by individuals has its precursors in the field of Artificial Intelligence. A good example is Boutilier's distinction between controllable and uncontrollable propositions (*cf.*, Boutilier (1994) and also Cholvy and Garion (2001)). Also in recent studies in distributed constraint satisfaction problems (Yokoo, Durfee, Ishida, and Kuwabara (1998), Walsh, Yokoo, Hirayama, and Wellman (2001)) the set of propositional variables is partitioned and the control over the values of the variables in each block is assigned to an agent. Their quest is for appropriate algorithms and protocols for groups of agents who jointly attempt to satisfy a propositional formula by choosing suitable values for the variables. If there is only one agent with control over all, these problems reduce to classical constraint satisfaction problems. This could be taken as an indication that in a sense the notion of control is not entirely foreign to classical logic.

By contrast, our concern is with the definition and investigation of consequence relations defined semantically by means of game-theoretical solution concepts. In our analyses we make the idealizing assumption that each variable is under the control of precisely one individual. If need be, an additional individual — *e.g.*, *Nature* or *Providence* — could be introduced, assigning values to variables that are normally thought to be beyond any individual's control. The different choices an individual can make with respect to the variables in her control then coincide with the strategies of some strategic game. A strategy profile then, collecting particular choices of the individuals, determines the values of all propositional variables and as such can be identified with the valuations of the respective propositional language.

A set of strategy profiles alone, however, does not define a game by itself. In order to conceive of the valuations not as mere the strategy profiles, but rather as the strategy profiles *of a specific strategic game,* also the preferences over the possible outcomes of the players should be specified. Formulas and theories provide this additional structure on logical space. The role of formulas and theories is thus analogous to the one they have in the classical semantics for classical propositional logic.

In a classical setting, the set of valuations constitute the logical space; it exhausts the possible ways in which the world can be fully described by means of a propositional language. Formulas and theories single out particular possible states of affairs, intuitively, by putting constraints on the possible ways the world looks like. Semantically, a formula demarcates those possible states of affairs in which that formula holds from those in which it does not.

In this part, however, we consider the valuations as the possible outcomes of games in which agents have control over the propositional variables. In this context, formulas and theories constrain the *game-theoretical* possibilities. Formulas and theories fix the preferences of the players over the possible outcomes and game-theoretical solution concepts are then applied to single out the valuations the outcomes that are game-theoretically likely, interesting or otherwise distinguished for particular social purposes. Pursuing this line of thought, we will eventually come to interpret theories and formulas semantically as relations over the valuations, rather than as sets of valuations.

In Boolean games, like in other game-theoretical approaches to logic, these ideas have already been present, be it perhaps in a rudimentary fashion. The interests of the two players are captured by the truth values a formula may take. The truth value of a formula, however, is no longer thought of as something that is somehow given independently; it is dependent on the decisions the players make. In a similar fashion, theories can be employed to define the players' preferences. *E.g.*, given a theory one could assume the one player to vie for its satisfaction, whereas the other rather saw at least one, or perhaps even all, of them false. There are, however, numerous ways in which theories can be used to define the preferences of the players. In the next chapter, we define a player's preferences on the basis of the relative logical strength of the formulas making up a theory.

In each Boolean game the players' preferences were assumed to be antagonistic. This made that a single formula sufficed to define the preferences of *both* players, *viz.*, the preferences of a verifier and those of a falsifier of that formula. In the interests of greater generality, we will come to lift the restriction that the players' preferences are necessarily related by a structural principle such as, *e.g.*, antagonism. The players' preferences are then specified by a theory for each individual player separately.

Whatever choice is made with respect to how the preferences of the players are extracted from theories or formulas, for each propositional language, any such choice defines a class of strategic games. Which strategies a player at his disposal has in such a game is determined by the propositional variables he controls. His preferences are fixed by formulas (or theories) of the language. The strategy profiles of any game in any such class are identified with the valuations of the respective propositional language. In this manner the valuations provide the basis on which different games in the same class can be compared. Moreover, given a suitable game-theoretical solution concept for a particular class of such games, the valuations are divided into those that comply with the solution concept and those that do not. As such, a solution concept singles out a set of valuations in much the same way as intersections and unions of extensions of formulas do in a Tarskian semantics for classical propositional logic. The question that pushes itself to the fore is then which formulas hold in the valuations thus singled out by a solution concept and which do not.

An issue that now suggests itself concerns the formulas that hold in the valuations that result if one of the players plays a winning strategy in a particular game. Suppose that a player has control over the propositional variable a and over a only. Assume further that all she wishes is the formula $a \vee b$ to be true. Then, setting the value of a to 1 is a winning strategy for her. Setting a to 0 is not, even though doing so does not entirely eliminate her chances of a favorable outcome. If her opponent happens to set b to 1, she still wins, but then without playing a winning strategy herself. Accordingly, in all strategy profiles in which she does play a winning strategy, obviously $a \vee b$ holds, but also the stronger formula a. At this point it be emphasized that this issue is different from the one that was addressed in the previous section on Boolean games. There the focus was not so much on the properties of winning strategies in a game as on the mere existence of a winning strategy for one of the players in Boolean games.

More in general, one could not so much be interested in the formulas that hold in the valuations singled out by a particular solution concept in a game of a particular class, as in how these valuations relate to those singled out by the same solution concept in *another* game in the same class. The strategy profiles a solution concept singles out for any two games in the same class, are drawn from the *same* set of valuations, *i.e.*, they constitute subsets of a common and more comprehensive set of strategy profiles. This makes that the set of strategy profiles a particular solution concept selects in one game can be compared with the set of strategy profiles thus distinguished by the same (or another, for that matter) solution concept in another game in a direct way. In particular, they can be compared with respect to set-theoretical properties, like set inclusion or disjointness.

For example, one could investigate whether the Nash equilibria of one game are disjoint from those of another. Figure 7.1 gives a graphical representation of three games for the propositional language containing a and b as only propositional variables. In these games each of two players is assigned control over one of the variables. Let the preferences of a player be given by a formula that player aims to satisfy. Thus, for the player with control over a these are given by a, $\neg(a \wedge b)$ and $\neg(b \rightarrow a)$, for the game on the left, the game in the middle and the game on the right, respectively. For the other player the preferences are then given by, respectively, $a \wedge b$, $\neg a$ and $a \rightarrow b$. The Nash equilibria of the game on the left and those of the game on the right are then disjoint. Relative to the assignment of the propositional variables to the players, this set-theoretical relation translates to a logical one between the pair of formulas a and $a \wedge b$, on the one hand, and the pair $\neg(a \wedge b)$ and $\neg a$, on the other. Similarly, the

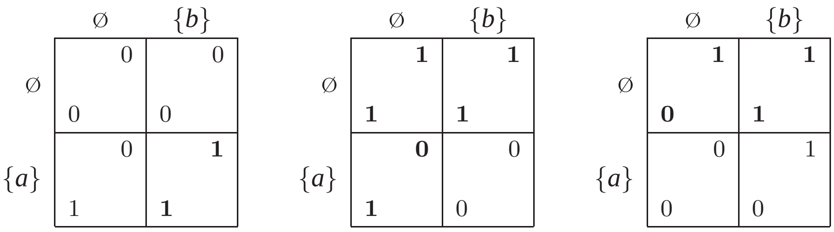


Figure 7.1. Three two-player games, in which the row player *Row* has control over the propositional variable a and the column player *Column* over b. In the leftmost game *Row*'s preferences are given by the formula a and *Column*'s by $a \wedge b$. In the game in the middle *Row* prefers valuations in which $\neg(a \wedge b)$ holds to those in which that is not the case and *Column* merely wishes a to be false. In the rightmost game, *Row*'s and *Column*'s preferences are given by, respectively, $\neg(b \rightarrow a)$ and $a \rightarrow b$. The Nash equilibria are in boldface.

Nash equilibria of game on the right are included in those of the middle game and says something different about the pairs of formulas $\neg(b \rightarrow a)$ and $a \rightarrow b$, and $\neg(a \wedge b)$ and $\neg a$.

Having assumed the games of each class being defined in a uniform fashion, each particular way of comparing the valuations complying with a particular solution concept can be elevated to a relation between pairs consisting a (collection of) theories or formulas and a distribution of the propositional variables. We propose to think of such relations between theories and formulas defined by game-theoretical solution concepts as consequence relations. In this manner the mutual dependencies between games with respect to a particular solution concept are studied through logic. This facilitates a purely formal treatment of the relations that hold between the games in question. The commensurability of the games such an analysis requires is guaranteed by the the fact that all games defined for one language share the same set of strategy profiles, *viz.*, the valuations of the respective propositional language.

The two fundamental ideas in the above — distributing control of the propositional variables over volitional agents and the interpretation of theories and formulas as preference relations — do not presuppose the logical analyses to be restricted to two-person games. Neither do they presuppose that the preferences of the players are of the binary kind that merely distinguishes wins from losses. These considerations have repercussions for the notion of logical consequence. We introduced classical consequence as a relation between theories. If distributed control over the variables is taken seriously, however, the relevant relations are of a more complicated nature. For one thing, the distribution of the variables itself should be accounted for. If, moreover, one does not assume the individual preferences of the players to be structurally interdependent, one may come to consider relations of a syntactically more complex kind.

In the next chapter, we will argue that theories can be employed to define a wide range of preference relations over the valuations. As such our proposal relates to a considerably more comprehensive class of games than that of two-player games of complete antagonism, which allow for only two different outcomes. The investigations of Chapter 9 below concern a relation between families of theories which is defined on the basis of these games and the solution concept of a *maximum equilibrium*. We will argue that this notion is in an important sense a conservative extension of classical propositional logic: classical consequence reduces to the game-theoretical relation if the control over the propositional variables is concentrated in one agent.

For the remainder of this chapter we adopt a more conventional course and investigate a logical consequence relation indexed by two subsets of propositional variables defined on basis of the notion of a *winning strategy*. Moreover, the focus will as yet be on one player only. The purpose of this exercise is to illustrate formally the general idea of a consequence relation defined in terms of a game-theoretical solution concept. The formal elaboration of this exercise relies on an extensive use of the machinery provided by *rough sets* as introduced in Section 2.2, as it will be in the formal analysis of *game-theoretical consequence* in Chapter 9, below. The players' preferences, however, are still thought of as distinguishing victories and defeats only.

7.2 Consequence Based on Winning Strategies

Giving content to the underlying ideas put forward in the introduction, we consider in this section a type of game in which a player is assigned control over a subset of propositional variables and in which the preferences of that player are defined by a theory of the respective language. The player is thought of as preferring valuations which satisfy the theory to those that do not, and being indifferent otherwise. Formulated thus, this gives only partial specification of a game, rather than a fully-fledged description of a game, as the preferences of the other players nor their manipulative powers are not specified. However, even so, the game-theoretical notion of the player having a *winning strategy* is applicable to these partial game-like structures. Moreover, without loss of generality a complete game description may be assumed by stipulating, *e.g.*, the existence of an antagonistic player who has control over all of the remaining propositional variables.

We introduce a family of consequence relations, *i.e.*, a family of relations between theories, in terms of these games and the concept of a winning strategy. This family we call *winning consequence* and its formal properties are investigated in the remainder of this chapter. Eventually, a Gentzen-style system is presented and proved to be sound and complete with respect to winning consequence.

Suppose a player has control over a subset of propositional variables Δ and aims at verifying a particular theory Γ. Then either there is a clever choice for the variables in Δ that renders the theory Γ true no matter what values are chosen for the propositional variables outside Δ, or there is no such choice. In the former case, the player in the possession of a *winning strategy* and in the latter she is not. Observe that whether a strategy for the player is winning or not, does not depend on how the propositional variables outside Δ obtain their values. These may be determined by *Providence,* by malicious or benevolent demons, by other players, or by whatever.

Formally, we define for each theory Γ and each subset of propositional variables Δ, a game $G(\Gamma, \Delta)$ with two players 1 and 0, of which the former has control over the propositional variables in Δ and the latter over the remaining ones. The strategies of player 1 are given by 2^{Δ} and those of Player 0, likewise, by $2^{\overline{\Delta}}$. The strategy profiles of any such game can thus be identified with the valuations for the language $L(A)$, as any strategy of Player 1 and any strategy of Player 0 taken together determine the values of *all* propositional variables. The winning conditions for any game $G(\Gamma, \Delta)$ grant Player 1 a win in all valuations s in the extension of Γ, *i.e.*, if $s \in [\![\Gamma]\!]$. In any other valuation Player 0 wins. A valuation s is said *to be a winning strategy for Player* 1 *in* $G(\Gamma, \Delta)$ if for all valuations s' that coincide with s on the values of all variables, player 1 wins the game, *i.e.*, in a game $G(\Gamma, \Delta)$:

$$s \text{ is a winning strategy for player 1 iff for all } s' \in S,\ s \sim_{\Delta} s' \text{ implies } s' \in [\![\Gamma]\!].$$

The solution concept of a winning strategy singles out a set of valuations in on the basis of a theory and a subset of propositional variables, in an analogous fashion as set-theoretical intersection did for each theory in the model-theoretical semantics for CPC. In the classical setting the focus was on the valuations in $[\![\Gamma]\!]$ for each theory Γ. Here, the selected valuations are the winning strategies of a game $G(\Gamma, \Delta)$ for each theory Γ and each subset of propositional variables Δ. A obvious question to ask is then which formulas hold in the winning strategies of a game $G(\Gamma, \Delta)$.

In the introduction to this chapter we saw that a would hold in all valuations that result if Player 1 has control over the variable a itself and if Γ is taken to be $\{a \vee b\}$. In the previous sentence the phrase 'hold in all valuations that ...', however, hints at a rudiment from the classical framework. Taking the game-theoretical point of view seriously, one could wish for a firm grip on how the valuations containing a winning strategy for Player 1 in one game $G(\Gamma, \Delta)$ relate to those containing a winning strategy for Player 1 in another game $G(\Theta, \Delta')$.

There are different ways in which this can be achieved. An obvious choice would be to compare the strategy profiles containing winning strategies for Player 1 in different games with respect to set-inclusion. This, however, would give rise to a rather inconvenient and lopsided formalism. For any pair of theories Γ and Θ and subsets of propositional variables Δ and Δ', we propose to compare the strategy profiles containing a winning strategy for Player 1 in the game $G(\Gamma, \Delta)$ with those strategy profiles that do *not* contain a winning strategy for Player 1 in the game $G\big(\{\neg\vartheta : \ \vartheta \in \Theta\}, \Delta'\big)$. However unnatural and contrived this definition may strike the reader at first sight, it succeeds in comparing a player's winning strategies in various games within a neat and symmetric formal framework. Moreover, it gives rise to a natural interpretation in terms of rough sets (*cf.*, Proposition 7.3.1 and Corollary 7.3.3, below). This characterization of the notion of winning consequence manifests its formal resemblance with the semantical definition of classical consequence, of which it happens to be a generalization (*cf.*, Corollaries 7.4.2 and 7.4.1, below). Accordingly, the relation of *winning consequence* is defined as follows:

Definition 7.2.1 *(Winning consequence)* For $L(A)$ a propositional language define the relation $\models^{\mathrm{W}}$ such that for all theories Γ and Θ and all subsets of propositional variables Δ and Δ':

$$\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$$
$$\text{iff}$$
$$\text{player 1's winning strategies in } G(\Gamma,\Delta) \text{ and } G\big(\{\neg\vartheta : \ \vartheta \in \Theta\}, \Delta'\big) \text{ are disjoint.}$$

Each pair of subsets Δ and Δ' of propositional variables in A determines a proper logic $\Lambda^{\mathrm{W}}_{\Delta,\Delta'}$ defined as $\Lambda^{\mathrm{W}}_{\Delta,\Delta'} =_{df.} \big\{(\Gamma,\Theta) : \ \Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta\big\}$.

For an example, consider once more Figure 7.1. Let in all games Player 1 be assigned control over a. Then, the matrix on the left represents $G(\{a\},\{a\})$ and the middle one the game $G(\{\neg(a \wedge b)\},\{a\})$. Observe that in $G(\{a\},\{a\})$ playing $\{a\}$ is a winning strategy for Player 1, whereas $\emptyset$ is a winning strategy in $G(\{\neg(a \wedge b)\},\{a\})$. Consequently, the sets of valuations containing a winning strategy for Player 1 in both games are disjoint and therefore $a \models^{\mathrm{W}}_{\{a\},\{a\}} a \wedge b$. Now assume that Player 1 is assigned control over b. Then, the righthand matrix depicts $G(\{\neg(b \to a)\},\{b\})$. Then, the sets of valuations containing a winning strategy for Player 1 in $G(\{a\},\{a\})$ and $G(\{\neg(b \to a)\},\{b\})$ overlap. Both sets contain the valuation $\{a,b\}$, in which both a and b are set to "true". Hence, $a \not\models^{\mathrm{W}}_{\{a\},\{b\}} b \to a$.

Some informal understanding of Definition 7.2.1 may also by gained by considering some extreme cases as to the choice of the parameters Δ and Δ'. The definition is chosen in such a way that the classical relation of consequence coincides with $\Lambda^{\mathrm{W}}_{A,A}$, where A denotes the full set of propositional variables of the language in question (*cf.*, Corollary 7.4.1, below). Hence, winning consequence could in a loose sense be said to be a conservative extension of the classical concept of consequence. The special cases in which Θ is a singleton and Δ' is taken to be A, moreover, have quite natural readings. It so happens that $\Gamma \models^{\mathrm{W}}_{\Delta,A} \{\varphi\}$ holds whenever φ is true in all strategy profiles of $G(\Gamma,\Delta)$ in which Player 1 plays a winning strategy.

7.3 Winning Consequence and Rough Sets

Winning consequence is defined in game-theoretical terms. An alternative characterization is possibly using rough set approximations of the extensions of formulas and theories. In this fashion, the firm set-theoretical grip propositional logic is regained that made the Tarskian account of classical logic so attractive. The formal development of the theory of winning consequence relies on its rough-set characterization.

The set of strategy profiles of a game $G(\Gamma,\Delta)$ in which Player 1 plays a winning strategy coincides with the lower approximation of the extension of Γ with respect to the equivalence relation π_Δ.

Proposition 7.3.1 *Let Γ be a theory in $L(A)$ and let Δ be a subset of A. Then the set of valuations containing a winning strategy for Player* 1 *in $G(\Gamma,\Delta)$ coincides with $\underline{apr}_\Delta(\llbracket\Gamma\rrbracket)$.*

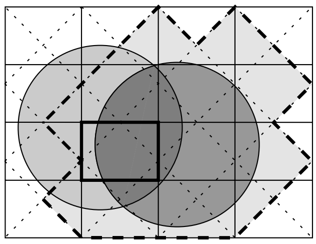


Figure 7.2. Logical space partitioned by π_Δ, indicated by the continuous lines, and by π_Δ, indicated by the dotted lines. The circle on the left depicts $[\![\Gamma]\!]$ and the one on the right $\langle\!\langle\Theta\rangle\!\rangle$. The boxed area then demarcates $\underline{apr}_\Delta([\![\Gamma]\!])$ and the dashed area $\overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle)$. Although obviously, $\Gamma \not\models^{\mathrm{CPC}} \Theta$, we find that $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$, because $\underline{apr}_\Delta([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle)$ (*cf.*. Proposition 7.3.2.)

Proof: Immediate from the definitions of a valuation containing a winning strategy for Player 1 in $G(\Gamma,\Delta)$ and that of the lower approximation of a set. Consider an arbitrary valuation s. Then, s contains a winning strategy for Player 1 in $G(\Gamma,\Delta)$ if an only if for all valuations s' with $s \sim_\Delta s'$: $s' \in [\![\Gamma]\!]$, *i.e.*, if and only if $s \in \underline{apr}_\Delta([\![\Gamma]\!])$. This concludes the proof. $\dashv$

This observation gives rise to the following characterization of $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ as the inclusion of the lower approximation of $[\![\Gamma]\!]$ with respect to π_Δ in the upper approximation of $\langle\!\langle\Theta\rangle\!\rangle$ with respect to $\pi_{\Delta'}$ (*cf.*. Figure 7.2).

Proposition 7.3.2 *Let Γ and Θ be theories of $L(A)$ and let Δ and Δ' be subsets of A. Then:*

$$\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{iff} \quad \underline{apr}_\Delta([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle)\,.$$

Proof: Proposition 7.3.1 establishes $\underline{apr}_\Delta([\![\Gamma]\!])$ and $\underline{apr}_{\Delta'}([\![\{\neg\vartheta:\ \vartheta\in\Theta\}]\!])$ as the set of valuations that contain a winning strategy for Player 1 in $G(\Gamma,\Delta)$ and in $G(\{\neg\vartheta:\ \vartheta\in\Theta\},\Delta')$, respectively. Observe that $[\![\{\neg\vartheta:\ \vartheta\in\Theta\}]\!] = \bigcap_{\vartheta\in\Theta}\overline{[\![\vartheta]\!]}$ and take notice of the following equalities:

$$\overline{\underline{apr}_{\Delta'}([\![\{\neg\vartheta:\ \vartheta\in\Theta\}]\!])} = \overline{\underline{apr}_{\Delta'}\left(\bigcap_{\vartheta\in\Theta}\overline{[\![\vartheta]\!]}\right)} =$$
$$\overline{apr}_{\Delta'}\left(\overline{\bigcap_{\vartheta\in\Theta}\overline{[\![\vartheta]\!]}}\right) = \overline{apr}_{\Delta'}\left(\bigcup_{\vartheta\in\Theta}[\![\vartheta]\!]\right) = \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle)\,.$$

This concludes the proof. $\dashv$

Recall that in case Γ is the empty theory, $[\![\Gamma]\!] = [\![\varnothing]\!] = \bigcap_{\gamma\in\varnothing}[\![\gamma]\!] = 2^A$. *I.e.*, it is *not* in general the case that $\varnothing \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$, for all theories Θ and all subsets $\Delta,\Delta' \subseteq A$. Because of the distribution of $\underline{apr}$ over $\bigcap$ and that of $\overline{apr}$ over $\bigcup$ (*cf.*, page 38, above), as an immediate result of Proposition 7.3.2 we also have the following.

Corollary 7.3.3 *Let Γ and Θ be theories of $L(A)$ and let Δ and Δ' be subsets of A. Then:*

$$\Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{iff} \quad \bigcap_{\gamma \in \Gamma} \underline{apr}_{\Delta}(\llbracket \gamma \rrbracket) \subseteq \bigcup_{\vartheta \in \Theta} \overline{apr}_{\Delta'}(\llbracket \vartheta \rrbracket)\,.$$

Proof: Immediate by Proposition 7.3.1 and the distribution of $\underline{apr}$ over $\bigcap$ and that of $\overline{apr}$ over $\bigcup$ (*cf.*, page 38. $\dashv$

7.4 Formal Development of Winning Consequence

The remainder of this chapter is devoted to the development of a Gentzen-style formal system for winning consequence in a propositional language $L(A)$. The system summarized in Table 7.5 on page 173 below, is proved sound and complete with respect to winning consequence. First, however, we review some of the formal properties of winning consequence, which form the basis of the soundness-direction of the above claim.

Properties of Winning Consequence

Proposition 7.3.2 and Corollary 7.3.3, above, have a number of other useful corollaries, the proofs of which almost invariably depend on the interaction between the laws of the theory of rough sets and the classical notion of the extension of a propositional formula. First and foremost, the claim that the relation of classical consequence coincides with $\Lambda^{\mathrm{W}}_{\Delta,\Delta'}$ if Δ and Δ' both equal A, as it was tentatively put forward in the previous section.

Corollary 7.4.1 *Let Γ and Θ be theories of $L(A)$. Then:*

$$\Gamma \vDash^{\mathrm{W}}_{A,A} \Theta \quad \textit{iff} \quad \Gamma \vDash^{\mathrm{W}}_{A(\Gamma),A(\Theta)} \Theta \quad \textit{iff} \quad \Gamma \vDash^{\mathrm{CPC}} \Theta.$$

Proof: Consider the following equivalences:

$$\begin{array}{lll} \Gamma \vDash^{\mathrm{W}}_{A,A} \Theta & \text{iff}_{\text{Prop. 7.3.2}} & \underline{apr}_{A}(\llbracket \Gamma \rrbracket) \subseteq \overline{apr}_{A}(\llbracket \Theta \rrbracket) \\ & \text{iff}_{\text{Fact 2.3.6}} & \underline{apr}_{A(\Gamma)}(\llbracket \Gamma \rrbracket) \subseteq \overline{apr}_{A(\Theta)}(\llbracket \Theta \rrbracket) \\ & \text{iff}_{\text{Prop. 7.3.2}} & \Gamma \vDash^{\mathrm{W}}_{A(\Gamma),A(\Theta)} \Theta. \end{array}$$

Similarly, also the following equivalences hold:

$$\begin{array}{lll} \Gamma \vDash^{\mathrm{W}}_{A,A} \Theta & \text{iff}_{\text{Prop. 7.3.2}} & \underline{apr}_{A}(\llbracket \Gamma \rrbracket) \subseteq \overline{apr}_{A}(\langle\!\langle \Theta \rangle\!\rangle) \\ & \text{iff}_{\text{Fact 2.3.6}} & \llbracket \Gamma \rrbracket \subseteq \langle\!\langle \Theta \rangle\!\rangle \\ & \text{iff} & \Gamma \vDash^{\mathrm{CPC}} \Theta. \end{array}$$

This concludes the proof. $\dashv$

Conversely, also each validity statement involving winning consequence of the form $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ has its counterpart in the classical notion of consequence. This phenomenon is due to the fact that in general $\underline{apr}_{\Delta}(\llbracket\varphi\rrbracket)$ is expressible in classical propositional logic (*cf.*, page 55).

Corollary 7.4.2 *Let Γ and Θ be theories in $L(A)$ and let Δ and Δ' be subsets of A. Then:*

$$\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{iff} \quad \{\, [\Delta]\,\gamma :\ \gamma \in \Gamma \} \models^{\mathrm{CPC}} \{\, \langle\Delta'\rangle\,\vartheta :\ \vartheta \in \Theta \}.$$

Proof: Consider the following equivalences:

$$\begin{aligned}
\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad & \text{iff}_{\text{Coroll. 7.3.3}} && \bigcap_{\gamma\in\Gamma} \underline{apr}_{\Delta}(\llbracket\gamma\rrbracket) \subseteq \bigcup_{\vartheta\in\Theta} \overline{apr}_{\Delta}(\llbracket\vartheta\rrbracket) \\
& \text{iff}_{\text{page 55}} && \bigcap_{\gamma\in\Gamma} \llbracket\, [\Delta]\,\gamma \rrbracket \subseteq \bigcup_{\vartheta\in\Theta} \llbracket\, \langle\Delta'\rangle\,\vartheta \rrbracket \\
& \text{iff} && \{\, [\Delta]\,\gamma :\ \gamma \in \Gamma \} \models^{\mathrm{CPC}} \{\, \langle\Delta'\rangle\,\vartheta :\ \vartheta \in \Theta \}.
\end{aligned}$$

This concludes the proof. $\dashv$

Recall that $[\Delta]\,\varphi$ and $\langle\Delta\rangle\,\varphi$ abbreviate the formulas $\bigwedge_{\sigma\in\Sigma_{\overline{\Delta}}} \sigma\,(\varphi)$ and $\bigvee_{\sigma\in\Sigma_{\overline{\Delta}}} \sigma\,(\varphi)$, respectively (*cf.*, page 57). Hence, we also have:

$$\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \text{iff} \quad \bigcup_{\gamma\in\Gamma} \{\sigma\,(\gamma) :\ \sigma \in \Sigma_{\overline{\Delta}}\} \models^{\mathrm{CPC}} \bigcup_{\vartheta\in\Theta} \{\sigma\,(\vartheta) :\ \sigma \in \Sigma_{\overline{\Delta'}}\}.$$

In virtue of Corollary 7.4.2, some important structural properties of classical consequence are inherited by winning consequence. Here we merely mention in this respect *compactness* and *consistency* — *i.e.*, in general, $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ implies there be *finite* $\Gamma' \subseteq \Gamma$ and $\Theta' \subseteq \Theta$ such that $\Gamma' \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta'$ and, respectively, $\emptyset \not\models^{\mathrm{W}}_{\Delta,\Delta'} \emptyset$.

Another property of classical consequence that also holds more in general for winning consequence is that of *overlap*, and *a fortiori*, also that of *reflexivity* and *diagonality*.

Proposition 7.4.3 *(Overlap) Let Γ and Θ be theories of $L(A)$ such that Γ and Θ are not disjoint, i.e., $\Gamma \cap \Theta \neq \emptyset$. Let further Δ and Δ' be subsets of A. Then:*

$$\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta.$$

Proof: Since Γ and Θ are not disjoint, there is some formula φ in $\Gamma \cap \Theta$; consider this φ. Then observe that in general $\underline{apr}_{\Delta}(\llbracket\varphi\rrbracket) \subseteq \llbracket\varphi\rrbracket \subseteq \overline{apr}_{\Delta'}(\llbracket\varphi\rrbracket)$. Hence, $\underline{apr}_{\Delta}(\llbracket\Gamma\rrbracket) \cap \underline{apr}_{\Delta}(\llbracket\varphi\rrbracket) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \cup \overline{apr}_{\Delta'}(\llbracket\varphi\rrbracket)$. Corollary 7.3.3 clinches the proof. $\dashv$

The concept of consequence as based on winning strategies is further upward monotonic in the sense that if $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ then also $\Gamma' \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta'$ for any theories Γ' and Θ' that include Γ and Θ, respectively.

Corollary 7.4.4 *(Monotonicity) Let Γ, Γ', Θ and Θ' be theories of $L(A)$ such that $\Gamma \subseteq \Gamma'$ and $\Theta \subseteq \Theta'$. Let further Δ and Δ' be subsets of A. Then:*

$$\Gamma \models^{W}_{\Delta,\Delta'} \Theta \quad \textit{implies} \quad \Gamma' \models^{W}_{\Delta,\Delta'} \Theta'.$$

Proof: Straightforward. Because $\Gamma \subseteq \Gamma'$ and $\Theta \subseteq \Theta'$, we have $[\![\Gamma']\!] \subseteq [\![\Gamma]\!]$ and $\langle\!\langle\Theta\rangle\!\rangle \subseteq \langle\!\langle\Theta'\rangle\!\rangle$. Hence also $\underline{apr}_{\Delta}([\![\Gamma']\!]) \subseteq \underline{apr}_{\Delta}([\![\Gamma]\!])$ and $\overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta'\rangle\!\rangle)$. Now the claim follows immediately from Proposition 7.3.2. Alternatively, the claim can be considered an immediate consequence of Corollary 7.4.2 and monotonicity of CPC. $\dashv$

By contrast, $\Lambda^{W}_{\Delta,\Delta'}$ is *downward* monotonic in Δ and Δ'. Informally, this is because the more propositional variables Player 1 has control over, the more likely she is to have a winning strategy available and the less likely it is that the set of strategy profiles containing one of her winning strategies to be included in another set.

Corollary 7.4.5 *Let Δ'' and Δ''' be subsets of A such that $\Delta'' \subseteq \Delta$ and $\Delta''' \subseteq \Delta'$. Then:*

$$\Gamma \models^{W}_{\Delta,\Delta'} \Theta \quad \textit{implies} \quad \Gamma \models^{W}_{\Delta'',\Delta'''} \Theta.$$

Proof: Observe that $\underline{apr}_{\Delta''}([\![\Gamma]\!]) \subseteq \underline{apr}_{\Delta}([\![\Gamma]\!])$ and that $\overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \subseteq \overline{apr}_{\Delta'''}(\langle\!\langle\Theta\rangle\!\rangle)$, are a special instances of Fact 2.2.4 on page 39, above. The proof is then immediate by Proposition 7.3.2. $\dashv$

Although upward monotonicity in Δ and Δ' fails for winning consequence in general, the validity of a statement $\Gamma \models^{W}_{\Delta,\Delta'} \Theta$ is not affected if Δ and Δ' are extended with propositional variables that do not occur in Γ and Θ, respectively. Hence, the following proposition.

Proposition 7.4.6 *Let Γ and Θ be theories of $L(\Delta)$ and let Δ and Δ' be subsets of A such that $\Delta'' \subseteq \Delta$, $\Delta''' \subseteq \Delta'$. Let further $\Delta'' \cap A(\Gamma) = \Delta \cap A(\Gamma)$ and $\Delta''' \cap A(\Theta) = \Delta' \cap A(\Theta)$. Then:*

$$\Gamma \models^{W}_{\Delta'',\Delta'''} \Theta \quad \textit{iff} \quad \Gamma \models^{W}_{\Delta,\Delta'} \Theta.$$

Proof: Consider the following equivalences:

$$\begin{array}{lll}
\Gamma \models^{W}_{\Delta,\Delta'} \Theta & \text{iff}_{\text{Prop. 7.3.1}} & \underline{apr}_{\Delta}([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \\
& \text{iff}_{\text{Fact 2.3.6}} & \underline{apr}_{\Delta}\left(\underline{apr}_{A(\Gamma)}([\![\Gamma]\!])\right) \subseteq \overline{apr}_{\Delta'}\left(\overline{apr}_{A(\Theta)}(\langle\!\langle\Theta\rangle\!\rangle)\right) \\
& \text{iff}_{\text{Prop. 2.2.8}} & \underline{apr}_{\Delta\cap A(\Gamma)}([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'\cap A(\Theta)}(\langle\!\langle\Theta\rangle\!\rangle) \\
& \text{iff}_{(*)} & \underline{apr}_{\Delta''\cap A(\Gamma)}([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'''\cap A(\Theta)}(\langle\!\langle\Theta\rangle\!\rangle) \\
& \text{iff}_{\text{Prop. 2.2.8}} & \underline{apr}_{\Delta''}\left(\underline{apr}_{A(\Gamma)}([\![\Gamma]\!])\right) \subseteq \overline{apr}_{\Delta'''}\left(\overline{apr}_{A(\Theta)}(\langle\!\langle\Theta\rangle\!\rangle)\right) \\
& \text{iff}_{\text{Fact 2.3.6}} & \underline{apr}_{\Delta''}([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'''}(\langle\!\langle\Theta\rangle\!\rangle) \\
& \text{iff}_{\text{Prop. 7.3.1}} & \Gamma \models^{W}_{\Delta'',\Delta'''} \Theta.
\end{array}$$

The equation marked with the asterisk holds, of course, in virtue of the assumptions that $\Delta'' \cap A(\Gamma) = \Delta \cap A(\Gamma)$ and that $\Delta''' \cap A(\Theta) = \Delta' \cap A(\Theta)$. $\dashv$

On basis of the Corollaries 7.4.4 and 7.4.5, we find that the proper logics $\Lambda^{\mathrm{W}}_{\Delta,\Delta'}$, for all subsets of propositional variables Δ and Δ', constitute a complete lattice ordered by the relation $\leqslant$ as defined on logics in general on page 45.

Fact 7.4.7 *Let $L(A)$ a propositional language and let $\boldsymbol{\Lambda}^{\mathrm{W}}$ denote the subset of proper logics defined by $\left\{\Lambda^{\mathrm{W}}_{\Delta,\Delta'} :\ \Delta, \Delta' \subseteq A\right\}$. Then, $\left(\boldsymbol{\Lambda}^{\mathrm{W}}, \leqslant\right)$ is a complete lattice.*

Proof: It suffices for a proof to show that for all subsets Δ, Δ', Δ'' and Δ''' of A:

$$\Lambda^{\mathrm{W}}_{\Delta,\Delta'} \quad \subseteq \quad \Lambda^{\mathrm{W}}_{\Delta'',\Delta'''} \quad \text{iff} \quad \Delta'' \subseteq \Delta \text{ and } \Delta''' \subseteq \Delta'.$$

Corollary 7.4.5 already takes care of the right-to-left direction. The opposite direction is proved by contraposition. Without loss of generality we may assume there to be some propositional variable a in A with $a \in \Delta''$ but $a \notin \Delta$. Observe that $\underline{apr}_{\Delta}(\llbracket a \rrbracket) =_{\text{Fact 2.3.10}} \underline{apr}_{\varnothing}(\llbracket a \rrbracket) =_{\text{Fact 2.2.10}} \varnothing$. By Corollary 7.3.3, then $a \vDash^{\mathrm{W}}_{\Delta,\Delta''} \varnothing$. However, $a \nvDash^{\mathrm{W}}_{\Delta'',\Delta'''} \varnothing$. To appreciate this last claim first observe that $\underline{apr}_{\Delta}(\llbracket a \rrbracket) =_{\text{Fact 2.3.10}} \underline{apr}_{\{a\}}(\llbracket a \rrbracket) =_{\text{Fact 2.3.5}} \underline{apr}_{A}(\llbracket a \rrbracket) =_{\text{Fact 2.2.10}} \llbracket a \rrbracket$. Also, evidently $\overline{apr}_{\Delta'''}(\langle\!\langle \varnothing \rangle\!\rangle) = \varnothing$. Hence, we are done by Corollary 7.3.3. $\dashv$

Corollary 7.4.1 establishes CPC as the bottom of the lattice $\left(\boldsymbol{\Lambda}^{\mathrm{W}}, \leqslant\right)$. The top $\Lambda_{\varnothing,\varnothing}$ of this lattice is not the inconsistent logic. Rather, it is characterized by the consequence relation that holds between any two theories Γ and Θ if and only if Γ containing tautologies only implies that Θ contains at least one satisfiable formula. Rephrased to some extent, this gives rise to the following fact.

Fact 7.4.8 *Let Γ and Θ be theories in a propositional language $L(A)$. Then:*

$$\Gamma \vDash^{\mathrm{W}}_{\varnothing,\varnothing} \Theta \quad \textit{iff} \quad \llbracket \Gamma \rrbracket \neq 2^A \textit{ or } \langle\!\langle \Theta \rangle\!\rangle \neq \varnothing.$$

Proof: Consider the following implications:

$$\begin{array}{lll} \llbracket \Gamma \rrbracket \neq 2^A \text{ or } \langle\!\langle \Theta \rangle\!\rangle \neq \varnothing & \text{implies}_{\text{ Fact 2.2.10}} & \underline{apr}_{\varnothing}(\llbracket \Gamma \rrbracket) = \varnothing \text{ or } \overline{apr}_{\varnothing}(\langle\!\langle \Theta \rangle\!\rangle) = 2^A \\ & \text{implies} & \underline{apr}_{\varnothing}(\llbracket \Gamma \rrbracket) \subseteq \overline{apr}_{\varnothing}(\langle\!\langle \Theta \rangle\!\rangle) \\ & \text{implies}_{\text{ Prop. 7.3.2}} & \Gamma \vDash^{\mathrm{W}}_{\varnothing,\varnothing} \Theta. \end{array}$$

For the opposite direction:

$$\begin{array}{lll} \Gamma \vDash^{\mathrm{W}}_{\varnothing,\varnothing} \Theta & \text{implies}_{\text{ Prop. 7.3.2}} & \underline{apr}_{\varnothing}(\llbracket \Gamma \rrbracket) \subseteq \overline{apr}_{\varnothing}(\langle\!\langle \Theta \rangle\!\rangle) \\ & \text{implies} & \underline{apr}_{\varnothing}(\llbracket \Gamma \rrbracket) \neq 2^A \text{ or } \overline{apr}_{\varnothing}(\langle\!\langle \Theta \rangle\!\rangle) \neq \varnothing \\ & \text{implies}_{\text{ Fact 2.2.10}} & \llbracket \Gamma \rrbracket \neq 2^A \text{ or } \langle\!\langle \Theta \rangle\!\rangle \neq \varnothing. \end{array}$$

This concludes the proof. ⊣

An important property of classical consequence that nevertheless fails to hold in general for winning consequence is that of *cut*. For a simple counterexample, observe that both $a \vDash^{\mathrm{W}}_{\varnothing,\{a\}} \varnothing$ and $\varnothing \vDash^{\mathrm{W}}_{\varnothing,\{a\}} a$. The former holds because $\underline{apr}_{\varnothing}(\llbracket a \rrbracket) = \varnothing$. By consistency of winning consequence, however, $\varnothing \nvDash^{\mathrm{W}}_{\varnothing,\{a\}} \varnothing$. Here, the reader be reminded that for the empty theory $\varnothing$, it is the case $\underline{apr}_{\varnothing}(\llbracket \varnothing \rrbracket) = \underline{apr}_{\varnothing}(\bigcap \varnothing) = \underline{apr}_{\varnothing}\big(2^A\big) = 2^A$ and that $\overline{apr}_{\{a\}}(\langle\!\langle \varnothing \rangle\!\rangle) = \overline{apr}_{\{a\}}(\bigcup \varnothing) = \overline{apr}_{\{a\}}(\varnothing) = \varnothing$.

Some form of transitivity, however, still holds for winning consequence. For any validity statement of the form $\Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta$, any formula in Γ may be replaced by a classically stronger one. Similarly, any formula in Θ may be replaced by a classically weaker one. As a special case, formulas that are logically equivalent in the classical sense may be substituted for one another in both Γ and Θ.

Corollary 7.4.9 *Let Γ and Θ be theories in $L(A)$ and let Δ and Δ' be subsets of A. Let further φ and ψ be formulas in $L(A)$ such that $\llbracket \varphi \rrbracket \subseteq \llbracket \psi \rrbracket$. Then:*

$$\Gamma, \psi \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{implies} \quad \Gamma, \varphi \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta,$$
$$\Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \varphi \quad \textit{implies} \quad \Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \psi.$$

Proof: Observe that $\llbracket \varphi \rrbracket \subseteq \llbracket \psi \rrbracket$ implies both $\underline{apr}_{\Delta}(\llbracket \varphi \rrbracket) \subseteq \underline{apr}_{\Delta}(\llbracket \psi \rrbracket)$ and $\overline{apr}_{\Delta'}(\llbracket \varphi \rrbracket) \subseteq \overline{apr}_{\Delta'}(\llbracket \psi \rrbracket)$. Then, $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \cap \underline{apr}_{\Delta}(\llbracket \psi \rrbracket) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle)$ implies $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \cap \underline{apr}_{\Delta}(\llbracket \varphi \rrbracket) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle)$. Similarly, $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \cup \overline{apr}_{\Delta'}(\llbracket \varphi \rrbracket)$ implies $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \cup \overline{apr}_{\Delta}(\llbracket \psi \rrbracket)$. The result then follows almost immediately from Proposition 7.3.2 and Corollary 7.3.3. ⊣

As to the interaction between the theories flanking the $\vDash^{\mathrm{W}}_{\Delta,\Delta'}$-sign and Δ and Δ' we have the following proposition. It captures the informal idea that a player cannot guarantee a propositional variable to hold if she fails to have control over it. Inspection of the proof will reveal that this holds in general for any *classically satisfiable* formula φ containing no variables in the player's control.

Proposition 7.4.10 *Let Δ and Δ' be subsets of propositional variables in A. Let further a be a propositional variable in A such that $a \notin \Delta$. Then both:*

$$a \vDash^{\mathrm{W}}_{\Delta,\Delta'} \varnothing \qquad \textit{and} \qquad \varnothing \vDash^{\mathrm{W}}_{\Delta',\Delta} a.$$

Proof: First observe that $A(a) = \{a\}$. This makes that under the conditions specified $\Delta \cap A(a) = \varnothing$. Also, $\llbracket a \rrbracket \neq \varnothing$ and $\llbracket a \rrbracket \neq 2^A$. Hence, $\underline{apr}_{\Delta}(\llbracket a \rrbracket) =_{\text{Prop. 2.3.10}} \underline{apr}_{\varnothing}(\llbracket a \rrbracket) =_{\text{Prop. 2.2.10}} \varnothing$. Similarly, $\overline{apr}_{\Delta}(\langle\!\langle \{a\} \rangle\!\rangle) = \overline{apr}_{\Delta}(\llbracket a \rrbracket) =_{\text{Prop. 2.3.10}} \overline{apr}_{\varnothing}(\llbracket a \rrbracket) =_{\text{Fact 2.2.10}} 2^A$. ⊣

The next number of propositions reflect the behavior of winning consequence with respect to the propositional connectives. These propositions are quite reminiscent of analogous ones holding for classical propositional logic. The latter cannot in general be extrapolated as to hold for winning consequence as well, unless certain conditions be observed. Yet, the first two of the following results are inherited without qualification. The first establishes that for winning consequence from absurdity anything follows and that anything entails triviality. The second concerns the introduction of conjunction in the antecedent and that of disjunction in the consequent. For the subsequent results for the remaining cases and connectives, however, the classical rules need to be modified to some extent, by imposing some constraints on their applicability.

Proposition 7.4.11 *Let Γ and Θ be theories in a propositional language $L(A)$ and Δ and Δ' subsets of A. Then:*

$$\Gamma, \bot \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta \qquad \textit{and} \qquad \Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \top.$$

Proof: Obviously $[\![\bot]\!] = \varnothing$ and $[\![\top]\!] = 2^A$. Hence, both $[\![\Gamma \cup \{\bot\}]\!] = \varnothing$ and $\langle\!\langle \Theta \cup \{\top\} \rangle\!\rangle = 2^A$. Now recall that for rough sets in general, both $\underline{apr}(\varnothing) = \varnothing$ and $\overline{apr}(S) = S$, where S is the universe (*cf.*, page 38, above) and we are done by Proposition 7.3.2. $\dashv$

The behavior of $\wedge$ at the lefthand side and that of $\vee$ at the righthand side of $\models^{\mathrm{W}}$ is as in classical propositional logic. This holds in virtue of the laws $\underline{apr}\left(\bigcap \boldsymbol{X}\right) = \bigcap_{X \in \boldsymbol{X}} \underline{apr}(X)$ and $\overline{apr}\left(\bigcup \boldsymbol{X}\right) = \bigcup_{X \in \boldsymbol{X}} \overline{apr}(X)$.

Proposition 7.4.12 *Let Γ and Θ be theories and φ and ψ formulas in $L(A)$. Let further Δ and Δ' be subsets of A. Then:*

$$\Gamma, \varphi, \psi \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{iff} \quad \Gamma, \varphi \wedge \psi \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta,$$
$$\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \varphi, \psi \quad \textit{iff} \quad \Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \varphi \vee \psi.$$

Proof: Straightforward by Proposition 7.3.2 and $\overline{apr}$ distributing over $\cup$ and $\underline{apr}$ over $\cap$. Observe the following equivalences:

$$\begin{array}{lll}
\Gamma, \varphi, \psi \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta & \text{iff}_{\text{Prop. 7.3.2}} & \underline{apr}_{\Delta}([\![\Gamma \cup \{\varphi, \psi\}]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \\
& \text{iff} & \underline{apr}_{\Delta}([\![\Gamma]\!]) \cap \underline{apr}_{\Delta}([\![\varphi]\!]) \cap \underline{apr}_{\Delta}([\![\psi]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \\
& \text{iff} & \underline{apr}_{\Delta}([\![\Gamma]\!]) \cap \underline{apr}_{\Delta}([\![\varphi]\!] \cap [\![\psi]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \\
& \text{iff} & \underline{apr}_{\Delta}([\![\Gamma]\!]) \cap \underline{apr}_{\Delta}([\![\varphi \wedge \psi]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \\
& \text{iff} & \underline{apr}_{\Delta}([\![\Gamma \cup \{\varphi \wedge \psi\}]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \\
& \text{iff}_{\text{Prop. 7.3.2}} & \Gamma, \varphi \wedge \psi \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta.
\end{array}$$

The other case goes by duality. $\dashv$

However, $\Gamma, \varphi \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ and $\Gamma, \psi \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ do not in general imply $\Gamma, \varphi \vee \psi \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$. Neither is it in general the case that $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \varphi$ and $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \psi$ imply $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'}$

$\Theta, \varphi \wedge \psi$. In view of Proposition 7.3.2, this corresponds to the failure of the inclusion of $\underline{apr}(X \vee Y)$ in $\underline{apr}(X) \cup \underline{apr}(Y)$ and that of $\overline{apr}(X) \cup \overline{apr}(Y)$ in $\overline{apr}(X \cap Y)$ to hold in general for rough sets. Nevertheless, Propositions 7.4.13 and 7.4.14 specify special conditions under which one may introduce a disjunction in the antecedent and conjunction in the consequent.

Proposition 7.4.13 *Let Γ and Θ be theories and φ and ψ be formulas of $L(A)$. Let Δ and Δ' be subsets of A. Then:*

$$\Gamma, \varphi \vDash^{W}_{\Delta \cup A(\varphi), \Delta'} \Theta \;\; \textit{and} \;\; \Gamma, \psi \vDash^{W}_{\Delta \cup A(\psi), \Delta'} \Theta \quad \textit{imply} \quad \Gamma, \varphi \vee \psi \vDash^{W}_{\Delta, \Delta'} \Theta,$$
$$\Gamma \vDash^{W}_{\Delta, \Delta' \cup A(\varphi)} \Theta, \varphi \;\; \textit{and} \;\; \Gamma \vDash^{W}_{\Delta, \Delta' \cup A(\psi)} \Theta, \psi \quad \textit{imply} \quad \Gamma \vDash^{W}_{\Delta, \Delta'} \Theta, \varphi \wedge \psi.$$

Proof: The proofs of both claims are analogous. Here we only give that of the former. Assume both $\Gamma, \varphi \vDash^{W}_{\Delta \cup A(\varphi), \Delta'} \Theta$ and $\Gamma, \psi \vDash^{W}_{\Delta \cup A(\psi), \Delta'} \Theta$. Then both:

$$\underline{apr}_{\Delta \cup A(\varphi)}(\llbracket \Gamma \rrbracket) \cap \underline{apr}_{\Delta \cup A(\varphi)}(\llbracket \varphi \rrbracket) \quad \subseteq \quad \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle), \text{ and}$$
$$\underline{apr}_{\Delta \cup A(\psi)}(\llbracket \Gamma \rrbracket) \cap \underline{apr}_{\Delta \cup A(\psi)}(\llbracket \psi \rrbracket) \quad \subseteq \quad \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle).$$

Since $\pi_{\Delta \cup A(\varphi)} \leqslant \pi_{\Delta}$ as well as $\pi_{\Delta \cup A(\psi)} \leqslant \pi_{\Delta}$, by Proposition 2.2.4 on page 39, both $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \subseteq \underline{apr}_{\Delta \cup A(\varphi)}(\llbracket \Gamma \rrbracket)$ and $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \subseteq \underline{apr}_{\Delta \cup A(\psi)}(\llbracket \Gamma \rrbracket)$. Because obviously also $\underline{apr}_{\Delta \cup A(\varphi)}(\llbracket \varphi \rrbracket) = \llbracket \varphi \rrbracket$ and $\underline{apr}_{\Delta \cup A(\psi)}(\llbracket \psi \rrbracket) = \llbracket \psi \rrbracket$, both:

$$\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \cap \llbracket \varphi \rrbracket \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle) \qquad \text{and} \qquad \underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \cap \llbracket \psi \rrbracket \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle).$$

Therefore, $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \cap \big(\llbracket \varphi \rrbracket \cup \llbracket \psi \rrbracket\big) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle)$, *i.e.*, $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \cap \llbracket \varphi \vee \psi \rrbracket \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle)$. Since, $\underline{apr}_{\Delta}(\llbracket \varphi \vee \psi \rrbracket) \subseteq \llbracket \varphi \vee \psi \rrbracket$, also $\underline{apr}_{\Delta}(\llbracket \Gamma \rrbracket) \cap \underline{apr}_{\Delta}(\llbracket \varphi \vee \psi \rrbracket) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle \Theta \rangle\!\rangle)$ and we may conclude that $\Gamma, \varphi \vee \psi \vDash^{W}_{\Delta, \Delta'} \Theta$. ⊣

Proposition 7.4.14 *Let Γ and Θ be theories and φ and ψ be formulas in $L(A)$ such that $A(\varphi)$ and $A(\psi)$ are disjoint. Then:*

$$\Gamma, \varphi \vDash^{W}_{\Delta, \Delta'} \Theta \;\; \textit{and} \;\; \Gamma, \psi \vDash^{W}_{\Delta, \Delta'} \Theta \quad \textit{imply} \quad \Gamma, \varphi \vee \psi \vDash^{W}_{\Delta, \Delta'} \Theta,$$
$$\Gamma \vDash^{W}_{\Delta', \Delta} \Theta, \varphi \;\; \textit{and} \;\; \Gamma \vDash^{W}_{\Delta', \Delta} \Theta, \psi \quad \textit{imply} \quad \Gamma \vDash^{W}_{\Delta', \Delta} \Theta, \varphi \wedge \psi.$$

Proof: It suffices to prove that under the conditions specified $\underline{apr}_{\Delta}(\llbracket \varphi \vee \psi \rrbracket) \subseteq \underline{apr}_{\Delta}(\llbracket \varphi \rrbracket) \cup \underline{apr}_{\Delta}(\llbracket \psi \rrbracket)$. By duality then also $\overline{apr}_{\Delta}(\llbracket \varphi \rrbracket) \cap \overline{apr}_{\Delta}(\llbracket \psi \rrbracket) \subseteq \overline{apr}_{\Delta}(\llbracket \varphi \wedge \psi \rrbracket)$. Consider an arbitrary valuation s and assume for contraposition both $s \notin \underline{apr}_{\Delta}(\llbracket \varphi \rrbracket)$ and $s \notin \underline{apr}_{\Delta}(\llbracket \psi \rrbracket)$. Hence there are valuations s' and s'' such that $s \sim_{\Delta} s'$ and $s' \notin \llbracket \varphi \rrbracket$, and $s \sim_{\Delta} s''$ and $s'' \notin \llbracket \psi \rrbracket$. Now define yet another valuation s^* such that for all $a \in A$:

$$s^*(a) \quad =_{df.} \quad \begin{cases} s'(a) & \text{if } a \in A(\varphi) - \Delta, \\ s''(a) & \text{if } a \in A(\psi) - \Delta, \\ s(a) & \text{otherwise.} \end{cases}$$

Note that s^* is well-defined in virtue of $A(\varphi)$ and $A(\psi)$ having been assumed to be disjoint. It can easily be established that $s' \sim_{A(\varphi)} s^*$ and $s'' \sim_{A(\psi)} s^*$. Hence, $s^* \notin [\![\varphi]\!]$ and $s^* \notin [\![\psi]\!]$. Therefore, $s^* \notin [\![\varphi]\!] \cup [\![\psi]\!]$ and $s^* \notin [\![\varphi \vee \psi]\!]$. Since, moreover, $s \sim_\Delta s^*$, we may conclude that $s \notin \underline{apr}_\Delta([\![\varphi \vee \psi]\!])$. ⊣

In classical propositional logic, formulas may be transposed to the other side of the turnstile provided that they are appended to a negation symbol, *i.e.*, $\Gamma, \varphi \vDash^{\mathrm{CPC}} \Theta$ implies $\Gamma \vDash^{\mathrm{CPC}} \Theta, \neg\varphi$. This holds on basis of the Boolean truism that $X \cap Y \subseteq Z$ implies $X \subseteq Z \cup \overline{Y}$. This principle, of course, also holds for rough sets and so we have in particular that $\underline{apr}_\pi(X) \cap \underline{apr}_\pi(Y) \subseteq \overline{apr}_{\pi'}(Z)$ implies $\underline{apr}_\pi(X) \subseteq \overline{apr}_{\pi'}(Z) \cup \overline{\underline{apr}_\pi(Y)}$ as well as $\underline{apr}_\pi(X) \subseteq \overline{apr}_{\pi'}(Z) \cup \overline{apr}_\pi(\overline{Y})$. If now π is finer than π', *i.e.*, if $\pi \leqslant \pi'$, then $\overline{apr}_\pi(\overline{Y}) \subseteq \overline{apr}_{\pi'}(\overline{Y})$. As another consequence, it then also holds that $\underline{apr}_\pi(X) \subseteq \overline{apr}_{\pi'}\left(Z \cup \overline{Y}\right)$. It is this principle of rough set theory which the next proposition invokes to account for the behavior of negation. The condition $A(\varphi) \cap \Delta' \subseteq A(\varphi) \cap \Delta$ enforces that the partitions involved in the approximations are suitably related as to coarseness.

Proposition 7.4.15 *Let Δ and Δ' be subsets of A and φ a formula in $L(A)$ such that $A(\varphi) \cap \Delta' \subseteq A(\varphi) \cap \Delta$. Then:*[1]

$$\begin{array}{lcl} \Gamma, \varphi \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta & \textit{implies} & \Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \neg\varphi, \\ \Gamma \vDash^{\mathrm{W}}_{\Delta',\Delta} \Theta, \varphi & \textit{implies} & \Gamma, \neg\varphi \vDash^{\mathrm{W}}_{\Delta',\Delta} \Theta. \end{array}$$

Proof: Consider the following implications:

$$\begin{array}{lll} \Gamma, \varphi \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta & \text{implies}_{\text{ Prop. 7.3.2}} & \underline{apr}_\Delta([\![\Gamma]\!]) \cap \underline{apr}_\Delta([\![\varphi]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \\ & \text{implies} & \underline{apr}_\Delta([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \cup \overline{\underline{apr}_\Delta([\![\varphi]\!])} \\ & \text{implies} & \underline{apr}_\Delta([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \cup \overline{apr}_\Delta\left(\overline{[\![\varphi]\!]}\right) \\ & \text{implies}_{\ (*)} & \underline{apr}_\Delta([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \cup \overline{apr}_{\Delta\cap A(\varphi)}\left(\overline{[\![\varphi]\!]}\right) \\ & \text{implies}_{\ (**)} & \underline{apr}_\Delta([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \cup \overline{apr}_{\Delta'\cap A(\varphi)}\left(\overline{[\![\varphi]\!]}\right) \\ & \text{implies}_{\ (*)} & \underline{apr}_\Delta([\![\Gamma]\!]) \subseteq \overline{apr}_{\Delta'}(\langle\!\langle\Theta\rangle\!\rangle) \cup \overline{apr}_{\Delta'}\left(\overline{[\![\varphi]\!]}\right) \\ & \text{implies} & \Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta, \neg\varphi. \end{array}$$

The implications indicated with $(*)$ are valid because in general $\overline{apr}_B([\![\varphi]\!]) = \overline{apr}_B\left(\overline{apr}_{A(\varphi)}([\![\varphi]\!])\right) = \overline{apr}_{B\cap A(\varphi)}([\![\varphi]\!])$. The implication indicated with $(**)$ holds in virtue of the assumption that $A(\varphi) \cap \Delta' \subseteq A(\varphi) \cap \Delta$ and hence $\overline{apr}_{\Delta\cap A(\varphi)}\left(\overline{[\![\varphi]\!]}\right) \subseteq \overline{apr}_{\Delta'\cap A(\varphi)}\left(\overline{[\![\varphi]\!]}\right)$. The argument for the second claim runs along analogous lines. ⊣

[1] Beware of the order of Δ and Δ'!

7.5 A Sequent System for Winning Consequence

The results of the previous section are the makings of a sound and complete formal Gentzen-type system W for winning consequence.

Definition 7.5.1 *(The System* W*)* Let $L(A)$ be a propositional language. For Σ and T finite sequences of formula of $L(A)$ and Δ and Δ' finite subsets of *A*, an expression of the form $\Sigma \Rightarrow_{\Delta,\Delta'} T$ is a *sequent* of W. The axioms and rules of W are given in Table 7.5. If a sequent $\Sigma \Rightarrow_{\Delta,\Delta'} T$ is derivable in W, this is denoted by $\vdash^{\mathrm{W}} \Sigma \Rightarrow_{\Delta,\Delta'} T$. For, possibly infinite, theories Γ and Θ of $L(A)$ and, possibly infinite, subsets of propositional variables Δ and Δ' we have:

$$\Gamma \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \text{iff} \quad \vdash^{\mathrm{W}} \Sigma \Rightarrow_{\Delta\cap A(\boldsymbol{\Sigma}),\Delta'\cap A(\boldsymbol{T})} T,$$

for Σ is a sequence of formulas in Γ^* and T a sequence of formulas in Θ^* and $\boldsymbol{\Sigma}$ and $\boldsymbol{T}$ denoting the sets of formulas occurring in Σ and T, respectively.

The soundness of W is established by a straightforward inductive argument.

Proposition 7.5.2 *(Soundness of* W*) Let Γ and Θ be theories in $L(A)$ and Δ and Δ' subsets of A. Then:*

$$\Gamma \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{implies} \quad \Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta.$$

Proof: Observe that in virtue of the definition of $\Gamma \vdash^{\mathrm{W}} \Theta$, Corollary 7.4.4 and Proposition 7.4.6 and , it suffices to prove that in general:

$$\vdash^{\mathrm{W}} \Sigma \Rightarrow_{\Delta\cap A(\boldsymbol{\Sigma}),\Delta'\cap A(\boldsymbol{T})} T \quad \text{implies} \quad \boldsymbol{\Sigma} \vDash^{\mathrm{W}}_{\Delta\cap A(\boldsymbol{\Sigma}),\Delta'\cap A(\boldsymbol{T})} \boldsymbol{T}.$$

We find that both the axioms (*0*) and (*1*), as well (*2*) are sound in this respect because of the Propositions 7.4.10 and 7.4.11. The logical, structural and replacement rules all preserve winning consequence. Propositions 7.4.12 through 7.4.15 prove the soundness of the left- and right introduction rules for the Boolean connectives. The contraction and permutation rules are valid in virtue of winning consequence being stated in terms of theories, *i.e.*, unordered sets of formulas. Monotonicity of winning consequence (*cf.*, Proposition 7.4.4) vindicate the rules $thin_L$ and $thin_R$. Corollary 7.4.5 grants the soundness of Δ-$slim_L$ and Δ-$slim_R$. Finally, the substitution rules are sound in virtue of the Corollaries 7.3.3 and 2.4.7, the latter stating the identity of $\underline{apr}_{\Delta}\left(\llbracket\varphi\,(a/\bot)\rrbracket \cap \llbracket\varphi\,(a/\top)\rrbracket\right)$ and $\underline{apr}_{\Delta}(\llbracket\varphi\rrbracket)$ and that of $\overline{apr}_{\Delta}\left(\llbracket\varphi\,(a/\bot)\rrbracket \cup \llbracket\varphi\,(a/\top)\rrbracket\right)$ and $\overline{apr}_{\Delta}(\llbracket\varphi\rrbracket)$, if $a \notin \Delta$ (*cf.*, page 57). ⊣

Inspection of the sequent system for W reveals that it contains obvious pendants for the axioms and structural rules of GP (*cf.*, Table 2.4 on page 52), above. The restrictions on the rules for the connectives are satisfied trivially if Δ, Δ', Δ^* and Δ^{**} are all taken to be identical to *A*.

Fact 7.5.3 *Let Γ and Θ be theories in $L(A)$. Then:*

$$\Gamma \vdash^{\mathrm{GP}} \Theta \quad \textit{iff} \quad \Gamma \vdash^{\mathrm{W}}_{A,A} \Theta.$$

Axioms:

$$(0)\ \bot \Rightarrow_{\Delta,\Delta'} \epsilon \qquad (1)\ \epsilon \Rightarrow_{\Delta,\Delta'} \top \qquad (2)\ a \Rightarrow_{\Delta,\Delta'} a$$

Logical Rules:

$$\neg_L : \quad \frac{\Sigma \Rightarrow_{\Delta;\Delta'} T, \varphi}{\Sigma, \neg\varphi \Rightarrow_{\Delta;\Delta'} T} \qquad \neg_R : \quad \frac{\Sigma, \varphi \Rightarrow_{\Delta';\Delta} T}{\Sigma \Rightarrow_{\Delta';\Delta} T, \neg\varphi}$$

Provided that $A(\varphi) \cap \Delta \subseteq A(\varphi) \cap \Delta'$, in $\neg_L$ and in $\neg_R$.

$$\wedge_L : \quad \frac{\Sigma, \varphi, \psi \Rightarrow_{\Delta;\Delta'} T}{\Sigma, \varphi \wedge \psi \Rightarrow_{\Delta;\Delta'} T} \qquad \vee_R : \quad \frac{\Sigma \Rightarrow_{\Delta;\Delta'} T, \varphi, \psi}{\Sigma \Rightarrow_{\Delta;\Delta'} T, \varphi \vee \psi}$$

$$\vee_L : \quad \frac{\Sigma, \varphi \Rightarrow_{\Delta^*;\,\Delta'} T \qquad \Sigma, \psi \Rightarrow_{\Delta^{**};\,\Delta'} T}{\Sigma, \varphi \vee \psi \Rightarrow_{\Delta;\Delta'} T}$$

$$\wedge_R : \quad \frac{\Sigma \Rightarrow_{\Delta';\,\Delta^*} T, \varphi \qquad \Sigma \Rightarrow_{\Delta';\,\Delta^{**}} T, \psi}{\Sigma \Rightarrow_{\Delta';\Delta} T, \varphi \wedge \psi}$$

Provided that in $\vee_L$ and $\wedge_R$ either (i) $\Delta^* = \Delta \cup A(\varphi)$ and $\Delta^{**} = \Delta \cup A(\psi)$, or (ii) $\Delta^* = \Delta^{**} = \Delta$ and $A(\varphi) \cap A(\psi) = \emptyset$.

Structural Rules:

$$contr_L : \quad \frac{\Sigma, \varphi, \varphi \Rightarrow_{\Delta;\Delta'} T}{\Sigma, \varphi \Rightarrow_{\Delta;\Delta'} T} \qquad contr_R : \quad \frac{\Sigma \Rightarrow_{\Delta;\Delta'} T, \varphi, \varphi}{\Sigma \Rightarrow_{\Delta;\Delta'} T, \varphi}$$

$$perm_L : \quad \frac{\Sigma, \varphi, \psi, P \Rightarrow_{\Delta;\Delta'} T}{\Sigma, \psi, \varphi, P \Rightarrow_{\Delta;\Delta'} T} \qquad perm_R : \quad \frac{\Sigma \Rightarrow_{\Delta;\Delta'} T, \varphi, \psi, \Upsilon}{\Sigma \Rightarrow_{\Delta;\Delta'} T, \psi, \varphi, \Upsilon}$$

$$thin_L : \quad \frac{\Sigma \Rightarrow_{\Delta;\Delta'} T}{\Sigma, \varphi \Rightarrow_{\Delta;\Delta'} T} \qquad thin_R : \quad \frac{\Sigma \Rightarrow_{\Delta;\Delta'} T}{\Sigma \Rightarrow_{\Delta;\Delta'} T, \varphi}$$

$$\Delta\text{-}slim_L : \quad \frac{\Sigma \Rightarrow_{\Delta\cup\{a\}\,;\,\Delta'} T}{\Sigma \Rightarrow_{\Delta;\Delta'} T} \qquad \Delta\text{-}slim_R : \quad \frac{\Sigma \Rightarrow_{\Delta\,;\,\Delta'\cup\{a\}} T}{\Sigma \Rightarrow_{\Delta;\Delta'} T}$$

Substitution Rules:

$$subst_L : \quad \frac{\Sigma, \varphi\,(a/\bot)\,, \varphi\,(a/\top) \Rightarrow_{\Delta,\Delta'} T}{\Sigma, \varphi \Rightarrow_{\Delta,\Delta'} T}$$

$$subst_R : \quad \frac{\Sigma \Rightarrow_{\Delta',\Delta} T, \varphi\,(a/\bot)\,, \varphi\,(a/\top)}{\Sigma \Rightarrow_{\Delta',\Delta} T, \varphi}$$

Provided that $a \notin \Delta$.

Table 7.5. The System W.

Proof: For the left-to-right direction, assume $\Gamma \vdash^{\mathrm{GP}} \Theta$. Then there is a derivation $\mathscr{D}$ of a sequent $\Sigma \Rightarrow T$ witnessing this fact. Let $\mathscr{D}^*$ be the sequence of sequents that results if each sequent $\Sigma' \Rightarrow T'$ in $\mathscr{D}$ is replaced by $\Sigma' \Rightarrow_{A(\boldsymbol{\Sigma}),A(\boldsymbol{T})} T$. Some reflection reveals that $\mathscr{D}^*$ is a derivation in W. In particular observe that the restrictions on the rules $\neg_L$, $\neg_R$, $\vee_L$ and $\wedge_R$ are complied with trivially, because we may assume $\mathscr{D}$ to have the subformula property. The opposite direction, follows by soundness of W, Corollary 7.4.2 and completeness of GP with respect to CPC. $\dashv$

As we present it, the proof of the completeness of W with respect to winning consequence parasitizes on the completeness of classical propositional logic with respect to GP. Proposition 7.4.2 establishes that for each statement of $\Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ there is a corresponding statement in CPC. In virtue of GP's completeness with respect to CPC, it thus suffices for completeness of W with respect to winning consequence to show that there is a derivation in W witnessing $\Gamma \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta$, for each derivation in GP witnessing $\bigcup_{\gamma\in\Gamma} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \vdash^{\mathrm{GP}} \bigcup_{\vartheta\in\Theta} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\}$. In demonstrating that this is indeed the case, we invoke the following lemma.

Lemma 7.5.4 *Let Γ and Θ be theories of a propositional language $L(A)$.*

$$\Gamma \cup \left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{implies} \quad \Gamma \cup \{\varphi\} \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta,$$
$$\Gamma \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \cup \left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \quad \textit{implies} \quad \Gamma \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \cup \{\varphi\}.$$

Sketch of proof: The proofs of both claims run along similar lines; here we only give that of the former. The proof is by induction on $\|\Delta \cap A(\varphi)\|$, *viz.*, the cardinality of $\Delta \cap A(\varphi)$. If $\|\Delta \cap A(\varphi)\| = 0$, the case is trivial because then $\left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} = \{\varphi\}$. For the induction step, let $\|\Delta \cap A(\varphi)\| = n + 1$ and let $\Delta \cap A(\varphi)$ be given by $\{a_0, \ldots, a_n\}$. Assume $\Gamma \cup \left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta$, then:

$$\vdash^{\mathrm{W}} \Sigma \Rightarrow_{\Delta\cap A(\boldsymbol{\Sigma}),\Delta'\cap A(\boldsymbol{T})} T,$$

for some sequences $\Sigma \in \left(\Gamma \cup \left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\}\right)^*$ and $T \in \Theta^*$. In virtue of *thin*$_L$ and the other structural rules, we may assume that $\Sigma = \Sigma', \psi_1, \ldots, \psi_{2^{n+1}}$, where $\left\{\psi_1, \ldots, \psi_{2^{n+1}}\right\} = \left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\}$ and none of $\psi_1, \ldots, \psi_{2^{n+1}}$ occurs in Σ'. Observe that $\left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} = \left\{\sigma(\varphi)(a_n/\bot), \sigma(\varphi)(a_n/\top) : \ \sigma \in \Sigma_{\overline{\Delta-\{a_n\}}}\right\}$. By 2^n applications of the substitution rule *subst*$_L$ — and a finite number of applications of the structural rules — then:

$$\vdash^{\mathrm{W}} \Sigma', \psi'_1, \ldots, \psi'_{2^n} \Rightarrow_{\Delta\cap A(\boldsymbol{\Sigma}),\Delta'\cap A(\boldsymbol{T})} T,$$

where $\left\{\psi'_1, \ldots, \psi'_{2^n}\right\} = \left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta-\{a_n\}}}\right\}$. Hence, $\Gamma \cup \left\{\sigma(\varphi) : \ \sigma \in \Sigma_{\overline{\Delta-\{a_n\}}}\right\} \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta$. By the induction hypothesis then eventually $\Gamma \cup \{\varphi\} \vdash^{\mathrm{W}} \Theta$. $\dashv$

We are now in a position to give the completeness proof for W with respect to winning consequence.

Theorem 7.5.5 *(Completeness of* W*) Let Γ and Θ be theories in $L(A)$ and Δ and Δ' subsets of A. Then:*

$$\Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta \quad \textit{implies} \quad \Gamma \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta.$$

Proof: Assume $\Gamma \vDash^{\mathrm{W}}_{\Delta,\Delta'} \Theta$. In virtue of Corollary 7.4.2 and the subsequent remark, then:

$$\bigcup_{\gamma\in\Gamma} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \vDash^{\mathrm{CPC}} \bigcup_{\vartheta\in\Theta} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\}.$$

By completeness of GP, then:

$$\bigcup_{\gamma\in\Gamma} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \vdash^{\mathrm{GP}} \bigcup_{\vartheta\in\Theta} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\},$$

and by Fact 7.5.3 also:

$$\bigcup_{\gamma\in\Gamma} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \vdash^{\mathrm{W}}_{A,A} \bigcup_{\vartheta\in\Theta} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\}.$$

Let $\mathscr{D}$ be a derivation of a sequent $\Sigma \Rightarrow_{A(\Sigma),A(T)} T$ witnessing this fact, for some sequences Σ and T in $\left(\bigcup_{\gamma\in\Gamma} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\}\right)^*$ and $\left(\bigcup_{\vartheta\in\Theta} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\}\right)^*$, respectively. Employing the rules Δ-*slim*$_L$ and Δ-*slim*$_R$, the derivation $\mathscr{D}$ can be extended to a derivation of the sequent $\Sigma \Rightarrow_{\Delta\cap A(\boldsymbol{\Sigma}),\Delta'\cap A(\boldsymbol{T})} T$. Hence:

$$\bigcup_{\gamma\in\Gamma} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\} \vdash^{\mathrm{W}}_{\Delta,\Delta'} \bigcup_{\vartheta\in\Theta} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\}.$$

Now let $\mathscr{D}'$ be a derivation of a a sequent $\Sigma \Rightarrow_{\Delta\cap A(\Sigma),\Delta'\cap A(T)} T$ witnessing this fact, for some sequences Σ and T in $\left(\bigcup_{\gamma\in\Gamma} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\}\right)^*$ and $\left(\bigcup_{\vartheta\in\Theta} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\}\right)^*$, respectively. In virtue of the weakening rules *thin*$_L$ and *thin*$_R$, we may assume that $\boldsymbol{\Sigma} = \bigcup_{\gamma\in\Gamma'} \left\{\sigma(\gamma) : \ \sigma \in \Sigma_{\overline{\Delta}}\right\}$ and $\boldsymbol{T} = \bigcup_{\vartheta\in\Theta'} \left\{\sigma(\vartheta) : \ \sigma \in \Sigma_{\overline{\Delta'}}\right\}$, for some finite $\Gamma' \subseteq \Gamma$ and $\Theta' \subseteq \Theta$. Accordingly, $\boldsymbol{\Sigma} \vdash^{\mathrm{W}}_{\Delta,\Delta'} \boldsymbol{T}$. By Lemma 7.5.4, then $\Gamma' \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta'$. Hence, also $\Gamma \vdash^{\mathrm{W}}_{\Delta,\Delta'} \Theta$, which concludes the proof. $\dashv$

7.6 Conclusion

In this chapter we took the view that distributed control over the values of propositional variables is a notion worthy of logical analysis. We came to regard the valuations of propositional languages as the strategy profiles of a strategic game. Thus a game-theoretical perspective on logical space was acquired, giving rise to new issues in propositional logic.

This chapter presented *winning consequence* to illustrate how these ideas can be elaborated formally in a relatively simple setting. From a game-theoretical point of view, moreover, the theory of winning consequence provides a formal framework in which a particular type of game can be studied with respect to the winning strategies they contain for one of the players. In short, every validity statement of the form $\Gamma \models^{\mathrm{W}}_{\Delta,\Delta'} \Theta$ can be interpreted in terms of the winning strategies of one of the players has in the games $G(\Gamma,\Delta)$ and $G(\Theta,\Delta')$. Moreover, a semantic interpretation of winning consequence is advanced, facilitating the formal development of the theory.

The notion of a winning strategy, however, takes into account precious little interaction between the players. Whether a particular strategy profile contains a winning strategy for a player only depends on the winning conditions and the power *of that player*. To which extent the other player or players can achieve their goals is quite irrelevant, in this respect.

In the next two chapters we continue the analysis of of distributed control over the variables in propositional logic. In doing so, however, we will come to consider a considerably more extensive class of games and a more sophisticated game-theoretical solution concept than that of a winning strategy. The preferences of the players in the games by means of which this analysis is performed define a finer-grained relation over the valuations. Moreover, the solution concept involved, *viz.*, maximum equilibrium, is of a more social and interactive nature.

In order to achieve this greater generality, we will come to revise the way theories determine the preferences of players. So far, outcomes have been divided between wins and losses and the players have tacitly been assumed to prefer the former to the latter. In the next chapter, we will argue how the notion of logical strength (in the classical sense) can be employed to define partial preorders as the players' preferences over the outcomes of a game.

The development of the logical framework will be analogous to that of winning consequence, yet the framework itself will be of a considerably wider scope because the games involved constitute a more comprehensive class of strategic games.

Chapter 8

Relational Semantics

8.1 Introduction

In the previous chapters we made a case for the conception of propositional variables as binary decision variables controlled by individuals. It then becomes natural to view a valuation for a propositional language as the combined result of the choices the individuals make with respect to their variables, rather than as a state of the world somehow given independently. If, moreover, the individuals are assumed to entertain individual preferences over the outcomes, some of the strategy profiles become salient from a social point of view. In particular, some can be distinguished from others by particular game-theoretical solution concepts, such as containing a winning strategy for one of the players or being a *maximum equilibrium*. Taking this perspective commands a predominantly game-theoretical view on the semantics of propositional logic.

In order to make plausible the view that valuations are the strategy profiles of some strategic game, players, their strategies and preferences should be specified. The players and their strategies are given by a partitioning of the propositional variables. A strategy for a player is then a choice for the values of the propositional variables assigned to her and the valuations can be thought about as strategy profiles. Thus, logical space assumes the structure of a frame of a strategic game (*cf.*, page 27).

In the classical setting, theories could be thought of as imparting information about the way the world is. On this conception, a theory demarcates the valuations that are consistent with the information it conveys from those that are not. If, however, we think about a logical possibility as a possible outcome of a decision making process, the classical image is less attractive. In interactive situations of which the outcome depends on the decisions of individual agents, the most relevant information concerns what makes the individuals decide in one way rather than in another. Thus, we come to view upon theories as imparting information about the players' preferences, *e.g.*, by reporting the goals they aim to achieve. Instead of interpreting a theory as the intersection or the union of the extensions of the formulas it contains, as in a traditional Tarskian setting, a theory will be thought of as determining a player's *preference relation* over the valuations. Conceiving of logical space as the frame of a strategic game, theories provide the complementary relational structure required for a fully fledged strategic game. In this manner, moreover, solution concepts become available to distinguish valuations on game-theoretical grounds.

Many, if not most, game-theoretical approaches to logic — such as Hintikka's Game Theoretical Semantics, Lorenzen's dialogue games and, in particular, the Boolean games as introduced in Part II of this thesis — concern two-player games in which players are, moreover, thought of as complete antagonists. One player strives for verification of a formula, the other for its falsification. It has been argued in a more general context that by imposing this restriction one passes by some of the most essential and vitalizing aspects of the situations of conflicting interests, witness Schelling in his *Strategy of Conflict*:

> But, in taking conflict for granted, and working with an image of participants who try to "win", a theory of strategy does not deny that there are common as well as conflicting interests among the participants. In fact the richness of the subject arises from the fact that [...] there is *mutual dependence as well as opposition*. Pure conflict, in which the interests of two antagonists are completely opposed, is a special case; it would arise in a war of complete extermination, otherwise not even in war. For this reason, "winning" in a conflict does not have a strictly competitive meaning; it is not winning relative to one's adversary. It means *gaining relative to one's own value system*; [...]. (Schelling (1960), p.4 (emphasis mine))

The antagonism in Boolean games — as well as in other game-theoretical analyses of classical logic — is due to the two players taking up the contrary roles of verifier and falsifier of one particular formula. Clearly, in a classical framework for a language containing negation, the falsifier could equivalently be understood as the verifier of the negation of the formula the verifier strives to bring about. Taking the game-theoretical perspective on logic as primary, an obvious generalization resolving the antagonism now suggests itself. Each player could be considered a verifier of a separate formula. It is then only a small step to lift the restriction that the formula of the one player be true whenever the one of the other is false. Rather, there is little need to assume these formulas to be related by any structural property whatsoever. Emancipated thus, both players acquire their own "value system". Within this setting mutual dependence can just as well be made sense of as antagonism. *E.g.*, a pure coordination problem, the extreme case of mutual dependence, arises if both players try to verify logically equivalent formulas. Also mixed forms of mutual dependence and opposition can be represented. Suppose that one player tries to verify $\neg(a \rightarrow b)$ and another $a \wedge b$. Then a being true furthers the interests of both players, but they are in conflict as to the truth-value of b.

However straightforward it may seem to lift the assumption of antagonism from a game-theoretical perspective, this move has some significant logical repercussions. With the preferences of the players being assumed to be independent of one another, a single formula no longer suffices to define the preferences of all players simultaneously. In the general case, for each player there may be need for a separate formula capturing his preferences. Once this has been conceded, however, there seems little point in limiting the number of players to two. One could distribute the propositional variables over any (countable) number of players, and consider any of these players the verifiers of separate formulas. Then, an assignment of the propositional variables over the various players together with a formula for each player defines a game.

This setup could be taken a step further. A formula on its own defines an order over the valuations of quite a rudimentary type. The player who is supposed to be its verifier may be thought of as preferring those valuations that satisfy the formula to those that do not, and being indifferent otherwise. This makes that a player either wins or loses, without the possibility of an intermediate outcome. However legitimate in itself, this is slightly unsatisfactory from the perspective of the game-theorist. An essential feature of social environments is that the eventual outcome depends on the choices of all players taken together. Each player has control over only a limited number of the relevant variables. It may very well depend on the choices of the other individuals, whether an individual is in a position to bring about an outcome that she prefers most. The achievement of the best possible outcome for a player may very well depend on the other players choosing particular values for their variables. As one cannot in general rely on one's opponents to be lenient in this way, sometimes a player will have to settle for a suboptimal outcome. A player's strategy may be optimal with respect to particular values for the other agents' variables, but be inferior to what she can achieve relative to other values for the other agents' variables. Still, such a relatively (or locally) optimal but absolutely (globally) lesser outcome constitutes an important game-theoretic datum. In this respect it be observed that, *e.g.*, the important solution concept of a Nash equilibrium is characterized as the combination of best-response strategies of all players, where a strategy is a best-response if it is the optimal choice *given a particular choice of strategy by the other players.* In social environments it is thus important to know, not only the most preferred outcomes of an agent, but also her preferences over the lesser preferred outcomes. In this chapter we will argue how theories together with the notion of logical strength of their constituent formulas can be employed to determine such finer-grained preferences over the valuations, *c.q.*, strategy profiles. Thus a more comprehensive class of games is brought within the scope of propositional logic.

This chapter leads up to a definition of the class of *distributed evaluation games* in Section 8.4. A distributed evaluation game is a strategic game specified by countable number of players, a function assigning propositional variables to the players — defining their strategies and manipulative powers — and a function assigning theories to players — defining their preferences over the valuations. In Section 8.5 an effort is made to demarcate the precise scope of the class of distributed evaluation games. We find that the class of distributed evaluation games for a propositional language $L(A)$ are those strategic games with the valuations as strategy profiles and for which each player's preference relation is the 'limit' of the finite approximations of a proto-order (*i.e.*, the empty relation or a reflexive and transitive relation) over the valuations (*cf.*, Theorem 8.5.15 on page 207).

Distributed evaluation games will form the semantical basis of the game-theoretical notion of consequence to be advanced and investigated in Chapter 9. This notion of consequence integrates the two main ideas on which this part pivots. First, it encapsulates the notion of distributed control over the propositional variables. Secondly, propositional theories are interpreted as reflexive and transitive relations over valuations rather than as mere sets thereof. These may seem rash departures from the traditional canons of logic, if no heed is taken. Nevertheless, the game-theoretical notion of consequence derives some of its respectability from the fact that classical consequence happens to be a special instance.

First, however, we propose a *relational semantics* for classical propositional consequence, which is phrased in terms of particular relations formulas define over the valuations rather than in terms of their extensions. From a classical point of view relational semantics has little to offer over and above a Tarskian semantics, in terms of extensions of formulas, as the latter is sound and complete with respect to classical propositional logic.[1] Still, it constitutes a natural starting point for the development of a game-theoretical concept of consequence.

The call for a richer structure on logical space has had many precursors in the field of artificial intelligence and philosophical logic. Semantical treatments of non-standard reasoning mechanisms often appeal to a richer ordinal structure on the models. Formal analyses of default reasoning (*e.g.*, Veltman (1996)) and studies in non-monotonic consequence relations (*cf. e.g.*, Shoham (1988), Kraus, Lehmann, and Magidor (1990) and Makinson (1994)) come under this heading. In this context, also qualitative decision theory (*e.g.*, Boutilier (1994)) and belief revision (*e.g.*, Gärdenfors (1988)) should be mentioned. In each of these cases the models that are somehow *optimal* with respect to these structures, play in one way or another a role in the definition of the key semantical concepts. Our proposal for a game-theoretical notion of consequence is in line with these researches, be it that the structure imposed on logical space is that of a distributed evaluation game and that the notion of optimality is understood in terms of compliance with a game-theoretical solution concept.

8.2 Relational Semantics for Propositional Logic

A propositional logic is introduced as a pair $(L, \vdash)$ where L is a propositional language over a countable set of propositional variables and $\vdash$ is a relation on theories of L. For classical propositional logic(CPC), $\Gamma \vdash^{\text{CPC}} \Theta$ informally reads "if all formulas in Γ are true, then so are some of Θ". This notion can be given a formal semantics in terms of valuations. Assuming classical consequence being given independently by $\vdash$, the following soundness and completeness result is obtained:

$$\Gamma \vdash^{\text{CPC}} \Theta \quad \text{iff} \quad \bigcap_{\gamma \in \Gamma} [\![\gamma]\!] \subseteq \bigcup_{\vartheta \in \Theta} [\![\vartheta]\!].$$

[1]Observe that on page 46 we defined soundness and completeness relative to an abstract relation between theories. As such neither a deductive system nor a semantics is primary in the definition of a logic and the notions of soundness and completeness apply to both.

Also other fundamental logical notions — such as validity, satisfiability and refutability — can be couched in terms of extensions of formulas or theories. In this manner, a formula φ is said to be valid if its extension coincides with the set of all valuations; otherwise φ is refutable.

The extension of a theory Γ could be seen as semantically summarizing the information contained in Γ. Yet in doing so, much structure of the set $\{[\![\gamma]\!] : \gamma \in \Gamma\}$ may be lost. In an attempt to retain more of the structure of the set $\{[\![\gamma]\!] : \gamma \in \Gamma\}$ we suggest an interpretation of a theory Γ as a relation over valuations, rather than a mere set thereof. For a natural definition the concept of *relative logical strength* is resorted to.

The relative logical strength of two formulas can also be captured in terms of the extensions of the formulas involved. Formally, a formula φ is said to be at least as strong as another formula ψ if any formula that follows from ψ is also a consequence of φ. In terms of sets of valuations, φ is then at least as strong as ψ if and only if the extension of φ is included in that of ψ. This definition of relative logical strength can straightforwardly be extrapolated as to hold between theories. A theory Θ is said to be logically at least as strong as another theory Γ if the consequences of Γ are contained in those of Θ. In this manner, relative logical strength induces a reflexive and transitive relation, *i.e.*, a preorder, on the subtheories of a theory.

For each theory Γ, this ordering in turn engenders an ordering over the valuations as follows. Let Γ_s be the subtheory of Γ containing exactly those formulas from Γ satisfied by s. Then, Γ_s is easily recognized as the (unique) logically strongest subtheory of Γ satisfied by s. On this basis, we might define a valuation s to be at least as strong as a valuation s' with respect to Γ if and only if Γ_s is logically at least as strong as $\Gamma_{s'}$. This ordering on the valuations is reflexive and transitive.

For an example, let Γ be the theory $\{a \vee b, \neg a, \neg a \wedge \neg b\}$ and consider the valuations $\{a\}$ and $\{b\}$. Then $\{a \vee b, \neg a\}$ is the strongest subtheory of Γ the valuation $\{b\}$ satisfies. The valuation $\{a\}$, on the other hand, satisfies no subtheory stronger than $\{a \vee b\}$. Because $\{a \vee b, \neg a\}$ is logically stronger than $\{a \vee b\}$, the valuation $\{b\}$ is ranked higher with respect to Γ than the valuation $\{a\}$. For a similar reason, the valuations $\{a\}$ and $\varnothing$ are incomparable with respect to Γ. Also consider Figure 8.1 for a pictorial illustration of these considerations.

Tarskian semantics for classical propositional logic disregards much of this ordinal structure a theory imposes on the valuations. This, of course, can be no censure of Tarskian semantics as a semantics for classical propositional logic, as its very soundness and completeness would belie this. However, recent semantical studies in nonstandard reasoning mechanisms had need for an ordinal structure on the set of valuations. A good example is Veltman's update semantics for default reasoning. Moreover, classical logic treats all inconsistent theories on a par. In particular, anything follows from an inconsistent theory. This distinguishes classical logic from paraconsistent logics. The orders two inconsistent theories induce over the set of valuations, however, may be very well be different. Let s be a valuation that forces b but not a and s' a valuation that forces both a and b. Consider again the theory $\{a \vee b, \neg a, \neg a \wedge \neg b\}$

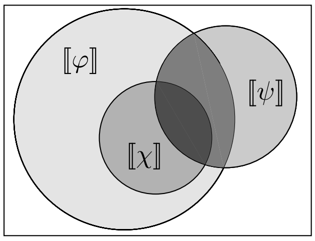


Figure 8.1. The extensions of three formulas φ, ψ and χ in logical space. The ordering on the valuations determined by the theory $\{\varphi, \psi, \chi\}$ on the basis of relative logical strength is indicated by the different shades of grey. The darker the area a valuation is in the higher that valuation is in the ordering on valuations determined by the three formulas, on the understanding that the valuations in $[\![\varphi]\!] - [\![\psi]\!]$ and those in $[\![\psi]\!] - [\![\varphi]\!]$ are incomparable. The valuations in the darkest area satisfy all of φ, ψ and χ and are ranked highest; those in the lighter areas satisfy no so strong a subtheory of $\{\varphi, \psi, \chi\}$ and are consequently ranked lower.

along with the theory $\{a, \neg a\}$. Both theories are classically inconsistent. However, with respect to the former s will be ranked higher than s'. Observe that any subtheory of $\{a \vee b, \neg a, \neg a \wedge \neg b\}$ satisfied by s will also be satisfied by s'. The valuation s, moreover, also satisfies the subtheory $\{a \vee b, \neg a\}$ whereas s' does not. However, with respect to the inconsistent theory $\{a, \neg a\}$, the valuations s and s' are incomparable, the former satisfying the subtheory $\{\neg a\}$ but not $\{a\}$ and the latter $\{a\}$ but not $\{\neg a\}$

The relation on the valuations based on the notion of logical strength as it was introduced above, however, has a drawback: it does not allow for a neat compositional definition. The malefactor is here the fact both theories containing merely contradictions and theories solely made up of tautologies induce the universal relation over the valuations.

In the next subsection we propose a relational semantics for classical propositional logic, in which formulas and theories are associated with *relations* over the valuations. The relation a theory is associated with is very similar to the relation its subtheories define on the valuations as based on their relative logical strength. It differs however from the latter in that it does allow for a compositional definition. The merits of this relational semantics are that it provides a natural point of departure for the formal analysis of game-theoretical consequence in Chapter 9. It provides a natural interpretation of theories if they are taken to reflect the interests, goals and preferences of individuals.

Set Induced Relations

In the next subsection a relational semantics for classical propositional logic is advanced. The set-theoretic basis for the semantics is provided by relations on a universe S that sets and sets of sets give rise to. With each subset X of a set S we associate a relation $\rho_0(X)$, which relates all elements outside X to any other element of S as well as all elements in X to one another. Intuitively, the objects in X are considered 'higher' than those outside X. Formally we define for each subset X of S and all elements x and x' of S:

$$(x,x') \in \rho_0(X) \quad \text{iff} \quad x \in X \ \text{ implies } \ x' \in X$$

On this basis we also define for each set $\boldsymbol{X}$ of subsets of S a relation $\boldsymbol{\rho}_0\,(\boldsymbol{X})$ on S:

$$\boldsymbol{\rho}_0\,(\boldsymbol{X}) \quad =_{df.} \quad \bigcap_{X\in\boldsymbol{X}} \rho_0(X)\,.$$

As can easily be checked, for each $X \subseteq S$, the relation $\rho_0(X)$ is a total pre-order, *i.e.*, it is reflexive, transitive and connected. For each set $\boldsymbol{X}$ of subsets of S, the relation $\boldsymbol{\rho}_0\,(\boldsymbol{X})$ is a *partial* pre-order over S, *i.e.*, it is both reflexive and transitive but not necessarily connected. As auxiliary notions we have $\bar{\rho}_0\,(X)$ and $\bar{\boldsymbol{\rho}}_0\,(\boldsymbol{X})$, defined as, respectively, $\rho_0\big(\overline{X}\big)$ and $\bigcap_{X\in\boldsymbol{X}} \bar{\rho}_0\,(X)$.

Of particular interest are the relations on the valuations induced by the extensions of formulas of a propositional language. Let φ be a formula and Γ a theory. We denote $\rho_0([\![\varphi]\!])$ by $\rho_0(\varphi)$ and $\bigcap_{\gamma\in\Gamma} \rho_0(\gamma)$ by $\boldsymbol{\rho}_0\,(\Gamma)$. Observe that, defined thus, $\boldsymbol{\rho}_0\,(\Gamma)$ *does not* in general coincide with $\rho_0([\![\Gamma]\!])$.

For Γ a theory, $\boldsymbol{\rho}_0\,(\Gamma)$ is exactly the relation the subtheories of Γ defines over the valuations relative to their respective logical strength. To appreciate this, define for each theory Γ the relation $\boldsymbol{\rho}_1(\Gamma)$ over the valuations, such that for all valuations s and s':

$$(s,s') \in \boldsymbol{\rho}_1(\Gamma) \quad \text{iff} \quad Cn\big(\{\gamma \in \Gamma :\ s \Vdash \gamma\}\big) \subseteq Cn\big(\{\gamma \in \Gamma :\ s' \Vdash \gamma\}\big).$$

Intuitively, $\boldsymbol{\rho}_1(\Gamma)$ relates a valuation s with another s' if the latter satisfies at least all those subtheories of Γ that s satisfies as well. As such, it is in effect the relation on the valuations based on the classical notion of logical strength that was proposed in the introduction to this chapter. We now have the following easy proposition.

Proposition 8.2.1 *Let Γ be a theory in a propositional language $L(A)$. Then the relations $\boldsymbol{\rho}_0\,(\Gamma)$ and $\boldsymbol{\rho}_1(\Gamma)$ coincide.*

Proof: First assume $(s,s') \in \boldsymbol{\rho}_0\,(\Gamma)$. Then, for all $\gamma \in \Gamma$, if $s \Vdash \gamma$ then $s' \Vdash \gamma$. Hence, $\{\gamma \in \Gamma :\ s \Vdash \gamma\} \subseteq \{\gamma \in \Gamma :\ s' \Vdash \gamma\}$. By monotonicity of Cn, immediately $Cn\,(\{\gamma \in \Gamma :\ s \Vdash \gamma\}) \subseteq Cn\,(\{\gamma \in \Gamma :\ s' \Vdash \gamma\})$, *i.e.*, $(s,s') \in \boldsymbol{\rho}_1(\Gamma)$. For the opposite direction, assume $Cn\,(\{\gamma \in \Gamma :\ s \Vdash \gamma\}) \subseteq Cn\,(\{\gamma \in \Gamma :\ s' \Vdash \gamma\})$ as well as for an arbitrary $\gamma \in \Gamma$ that $s \Vdash \gamma$. Then, $\gamma \in \{\gamma \in \Gamma :\ s \Vdash \gamma\}$ and by monotonicity of Cn also $\gamma \in Cn\,(\{\gamma \in \Gamma :\ s \Vdash \gamma\})$. By the assumption, $\gamma \in Cn\,(\{\gamma \in \Gamma :\ s' \Vdash \gamma\})$. Then $\{\gamma \in \Gamma :\ s' \Vdash \gamma\} \vDash^{\mathrm{CPC}} \gamma$, *i.e.*, for all valuations s'', if $s'' \Vdash \varphi$ for all $\varphi \in \{\gamma \in \Gamma :\ s' \Vdash \gamma\}$, then $s'' \Vdash \gamma$. Since trivially, $s' \Vdash \varphi$ for all $\varphi \in \{\gamma \in \Gamma :\ s' \Vdash \gamma\}$, in particular $s' \Vdash \gamma$. Therefore $(s,s') \in \boldsymbol{\rho}_0\,(\Gamma)$ and we are done. $\dashv$

For formulas φ, however, the relation $\rho_0(\varphi)$ does not have in general a neat compositional definition in the complexity of φ. To appreciate this, observe that both $\rho_0(\top)$

and $\rho_0(\bot)$ are the universal relation over the valuations. In contrast, for each propositional variable a, the relation $\rho_0(a)$ is not universal. Whenever a valuation s forces a but another valuation s' does not, the pair (s, s') will not be in $\rho_0(a)$. This is in particular the case for the valuations $\{a\}$ and $\emptyset$, which are guaranteed to exist for any language with a as a propositional variable. Now consider the formulas $\bot \wedge a$ and $\top \wedge a$. Since $[\![\top \wedge a]\!] = [\![a]\!]$, it should also be the case that the relations $\rho_0(\top \wedge a)$ and $\rho_0(a)$ coincide. So, with $\rho_0(a)$ not being the universal relation, neither is $\rho_0(\top \wedge a)$. However, $\rho_0(\bot \wedge a)$ is the universal relation on the valuations, in virtue of $\bot \wedge a$ and $\bot$ being logically equivalent, and as such having the same extension. Hence $\rho_0(\top \wedge a)$ is distinct from $\rho_0(\bot \wedge a)$. However, with $\rho_0(\top)$ and $\rho_0(\bot)$ being identical, this distinction cannot be made on the basis of the relations $\rho_0(\top)$, $\rho_0(\bot)$ and $\rho_0(a)$ alone.

The problem here of course is that $\rho_0(\emptyset)$ and $\rho_0(S)$ are the same relation. For any two non-empty subsets X and Y of S the reader can easily verify that $\rho_0(X)$ and $\rho_0(Y)$ coincide if and only if X and Y are identical (also compare Fact 8.5.1, below). By treating the empty set as a special case, many of the problems dissolve. So, define, for each subset X of a set S, the relation $\rho(X)$ on S as follows:[2]

$$\rho(X) \quad =_{df.} \quad \begin{cases} \{(x, x') : \ x \in X \text{ implies } x' \in X\} & \text{if } X \neq \emptyset, \\ \emptyset & \text{otherwise.} \end{cases}$$

Defined thus, $\rho(X)$ coincides with $\rho_0(X)$ for non-empty subsets X, and is empty otherwise (*cf.*, Fact 8.5.2 below). For A a set of propositional variables, let, $\mathscr{R}(A)$ denote the set $\{\rho(\varphi) : \ \varphi \text{ a formula in } L(A)\}$. Similarly, for each set $\boldsymbol{X}$ of subsets of S, we define the relation $\boldsymbol{\rho}(\boldsymbol{X})$ over S as:

$$\boldsymbol{\rho}(\boldsymbol{X}) \quad =_{df.} \quad \bigcap_{X \in \boldsymbol{X}} \rho(X).$$

Let further $\rho(\varphi)$ and $\boldsymbol{\rho}(\Gamma)$ denote $\rho([\![\varphi]\!])$ and $\bigcap_{\gamma \in \Gamma} \rho([\![\gamma]\!])$, respectively. Observe that $\boldsymbol{\rho}(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X})$ if and only if $\emptyset \notin \boldsymbol{X}$ (*cf.*, *idem*). For $\emptyset \in \boldsymbol{X}$, the relation $\boldsymbol{\rho}(\boldsymbol{X})$ is empty. As a dual notion we also introduce for each subset X of S the relation $\bar{\rho}(X)$ on S defined as $\rho(\overline{X})$. Also, for each set $\boldsymbol{X}$ of subsets of S, let $\bar{\rho}(\boldsymbol{X})$ denote the relation on S given by $\bigcap_{X \in \boldsymbol{X}} \bar{\rho}(X)$. We have $\bar{\rho}(\varphi)$ and $\bar{\boldsymbol{\rho}}(\Gamma)$ abbreviate $\bar{\rho}([\![\varphi]\!])$ and $\bigcap_{\gamma \in \Gamma} \bar{\rho}([\![\gamma]\!])$. We mention in passing that, in contradistinction to $\rho_0(\varphi)$, the relation $\rho(\varphi)$ does allow for a compositional definition in φ (*cf.*, Harrenstein (to appear-b)).

[2] It might seem that the identity relation *Id* would have been an equally suitable choice for $\rho(\emptyset)$, as there is no subset X of S such that $\rho_0(X) = Id$. Had the definition been chosen thus, however, an exception should be made in Proposition 8.2.3 below for propositional languages with no propositional variables. For such languages there is only one valuation, *viz.*, $\emptyset$, and again $\rho(\emptyset)$ would coincide with the universal relation over all valuations. Then it would have been the case that $max(\rho(\bot)) = \{\emptyset\}$ and $\overline{max(\bar{\rho}(\top))} = \emptyset$. Hence, $max(\rho(\bot)) \nsubseteq \overline{max(\bar{\rho}(\top))}$. In classical logic, however, $\bot \vdash \top$, even for languages lacking in propositional variables.

Relational Semantics

We are now in a position to furnish classical propositional logic with a relational semantics. With each formula φ we associate the relations $\rho(\varphi)$ and $\bar{\rho}(\varphi)$ over the valuations and, similarly, with each theory Γ the relations $\boldsymbol{\rho}(\Gamma)$ and $\bar{\boldsymbol{\rho}}(\Gamma)$. Denoting the set of *maximum* elements of a relation ρ by $max(\rho)$, we have the following proposition.

Proposition 8.2.2 *Let Γ be a theory in a propositional language $L(A)$. Then:*

$$[\![\Gamma]\!] \ = \ max(\boldsymbol{\rho}(\Gamma)) \qquad \textit{and} \qquad \langle\!\langle\Gamma\rangle\!\rangle \ = \ \overline{max(\bar{\boldsymbol{\rho}}(\Gamma))}.$$

Proof: First assume $[\![\Gamma]\!]$ to be empty. Assume further for a *reductio ad absurdum* that s is a maximum element of $\boldsymbol{\rho}(\Gamma)$ and consider an arbitrary $\gamma \in \Gamma$. Then, $(s', s) \in \rho(\gamma)$, for all valuations s'. So, in particular, $(s, s) \in \rho(\gamma)$ and from the definition of $\rho(\gamma)$ then follows that $[\![\gamma]\!] \neq \varnothing$. Hence, $s^* \in [\![\gamma]\!]$, for some s^*. Then also $(s^*, s) \in \rho(\gamma)$ and consequently $s \in [\![\gamma]\!]$ as well. With γ having been chosen as an arbitrary element of Γ, we have that $s \in [\![\Gamma]\!]$, which is at variance with the assumption that $[\![\Gamma]\!]$ be empty.

So, for the remainder of the proof we will assume $[\![\Gamma]\!]$ to be not empty. Consider an arbitrary valuation s. First assume that $s \notin [\![\Gamma]\!]$. Then $s \notin [\![\gamma]\!]$, for some $\gamma \in \Gamma$. With $[\![\Gamma]\!]$ not empty, we may assume there is some $s' \in [\![\gamma]\!]$. Then, however, $(s', s) \notin \rho(\gamma)$ and $(s', s) \notin \boldsymbol{\rho}(\Gamma)$. Hence, s is no maximum element of $\boldsymbol{\rho}(\Gamma)$. Finally, assume $s \in [\![\Gamma]\!]$. Now consider an arbitrary valuation s' along with an arbitrary $\gamma \in \Gamma$. Then, $s \in [\![\gamma]\!]$ and so $(s', s) \in \rho(\gamma)$. With γ having been chosen arbitrarily, also $(s', s) \in \boldsymbol{\rho}(\Gamma)$ and we may conclude that s is a maximum element of $\boldsymbol{\rho}(\Gamma)$. This concludes the first part of the proof

The second part of the proof can be obtained using the first one (duality). Merely consider the following equalities:

$$\begin{aligned} \langle\!\langle\Theta\rangle\!\rangle \ &= \ \bigcup_{\vartheta\in\Theta} [\![\vartheta]\!] \ = \ \overline{\bigcap_{\vartheta\in\Theta} \overline{[\![\vartheta]\!]}} \ = \ \overline{[\![\, \{\neg\vartheta : \ \vartheta \in \Theta\} \,]\!]} \\ &= \ \overline{max\left(\boldsymbol{\rho}\left(\{\neg\vartheta : \ \vartheta \in \Theta\}\right)\right)} \ = \ \overline{max\left(\bigcap_{\vartheta\in\Theta} \rho\left(\overline{[\![\vartheta]\!]}\right)\right)} \ = \ \overline{max\left(\bar{\boldsymbol{\rho}}(\Theta)\right)}. \end{aligned}$$

This concludes the proof. ⊣

As an immediate consequence of this result, we have the following corollary, which characterizes classical logical consequence in terms of the relations theories define. A theory Θ follows classically from another theory Γ if and only if the maximum elements of the relation $\boldsymbol{\rho}(\Gamma)$ are *no* maximum elements of the relation $\bar{\boldsymbol{\rho}}(\Theta)$.

Corollary 8.2.3 *Let Γ be a theory and φ a formula. Then:*

$$\Gamma \vdash^{\mathrm{CPC}} \Theta \quad \textit{iff} \quad max(\boldsymbol{\rho}(\Gamma)) \subseteq \overline{max(\bar{\boldsymbol{\rho}}(\Theta))}.$$

Proof: Immediate by Proposition 8.2.2. ⊣

As an alternative to classical consequence, one could define a consequence relation $\vdash^*$ as follows in terms of the maximum elements of the relations $\boldsymbol{\rho}_0(\Gamma)$ and $\bar{\boldsymbol{\rho}}_0(\Gamma)$:

$$\Gamma \vdash^* \Theta \quad \text{iff} \quad max(\boldsymbol{\rho}_0(\Gamma)) \subseteq \overline{max(\bar{\boldsymbol{\rho}}_0(\Theta))}.$$

This consequence relation is as the classical one, except for its behavior with respect to classical contradictions and tautologies. In virtue of Fact 8.5.2 below — which states that $\boldsymbol{\rho}(\Gamma) = \boldsymbol{\rho}_0(\Gamma)$ if and only if $\varnothing \notin \boldsymbol{X}$ — it can easily be appreciated that $\Gamma \vdash^* \Theta$ if and only if $\Gamma \vdash^{\mathrm{CPC}} \Theta$, provided that Γ contains no contradictions and Θ no tautologies. We have already seen, however, that $\rho_0(\bot)$ and $\rho_0(\top)$ are both interpreted as the universal relation over the valuations. Consequently, $\vdash^*$ treats classical contradictions on a par with classical tautologies. This makes that, *e.g.*, the classical rule *ex falso quod libet* fails for $\vdash^*$. For a counterexample, let a be a propositional variable of a language $L(A)$. Then observe that $\bot \nvdash^* a$, as $\mathit{max}(\boldsymbol{\rho}_0(\{\bot\})) = 2^A$ and $\overline{\mathit{max}(\bar{\boldsymbol{\rho}}_0(\{a\}))} = \overline{\mathit{max}(\boldsymbol{\rho}_0(\{\neg a\}))} = [\![a]\!]$. A similar remark would have applied, had $\vdash^*$ been defined in terms of the *maximal* elements of the relations $\boldsymbol{\rho}_0(\Gamma)$ and $\bar{\boldsymbol{\rho}}_0(\Theta)$. Observe in this respect that the maximal elements of $\boldsymbol{\rho}_0(\{\bot\})$ exhaust logical space just as well as the maximum elements of $\boldsymbol{\rho}_0(\{\bot\})$ do. The consequence relations defined in terms of maximal elements also exhibit non-monotonic features, but we will not pursue this issue here.

The advantage of the relational semantics is that it preserves more of the structure that formulas and theories impose on logical space. From the extension $[\![\Gamma]\!]$ of a theory Γ the extensions $[\![\gamma]\!]$ of the formulas γ in Γ cannot in general be recovered; this structure may have been lost beyond repair. In a strict sense a similar thing can be said of the relation $\boldsymbol{\rho}(\Gamma)$ and the relations $\rho(\gamma)$: it is not in general the case that from the relation $\boldsymbol{\rho}(\Gamma)$ the theory Γ can be reconstructed. Nevertheless, the relation $\boldsymbol{\rho}(\Gamma)$ can distinguish valuations s and s' even if neither of them is maximum in $\boldsymbol{\rho}(\Gamma)$, indicating that Γ contain a formula that is validated in the one but not in the other, or if Γ is inconsistent. *E.g.*, the valuation $\varnothing$ is strictly less than the valuation $\{a\}$ in the relation $\boldsymbol{\rho}(\{a, b\})$, yet neither of them is a maximum element in this respect. This feature of the relational semantics is especially serviceable when one is interested in the maximal or maximum elements of the relation determined by a theory as restricted to a subset of valuations, even if the maximal or maximum elements of the unrestricted relation are disjoint from that subset.

In our proposal for a game-theoretical notion of consequence, the theory induced relations are viewed upon as reflecting the preferences of a player. Moreover, we will be interested in the maximum elements of a relation induced by a theory within certain subsets of the valuations, *viz.*, those subsets that are still possible outcomes given a particular choice of strategy for all but one player. In view of Proposition 8.2.2, the extension of a theory, however, merely contains a player's most preferred outcomes, independently of her powers or the others players' preferences. In game-like situations, however, a player has generally control over only a limited number of the relevant variables. Whether she is able to achieve an outcome she prefers above all others, may well depend on the decisions of the other players. Moreover, even if a particular choice of values for the variables in an agent's control may achieve such a consummate outcome given certain partisan choices by the other players, it may have another, if not opposite, effect in case the other players decide differently. The best an agent can achieve relative to some fixed values for the other players' variables may be inferior

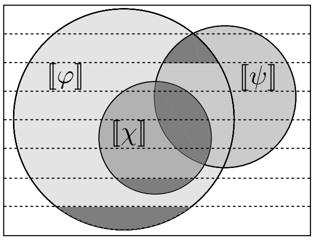


Figure 8.2. Let the preferences of a player *i* be captured in the theory $\{\varphi, \psi, \chi\}$. Player *i*'s preference relation over the valuations is then as in Figure 8.1, above. Here, each block represents a particular choice of strategy by *i*'s opponents. The darkest areas indicate where the maximum responses of *i* are to be found. Observe that a block may contain no maximum responses for *i* and also that a valuation may contain a maximum response for *i* even if it is outside the extension of the theory $\{\varphi, \psi, \chi\}$.

to what she can achieve relative to other values for the other players' variables. Such a locally optimal outcome is a significant detail from a game-theoretic point of view. Intuitively, what we are looking for are an agent's optimal outcomes given particular choices of strategies by the other agents. Let the preferences of an agent be represented by a theory Γ. Fixing the values of the variables outside the control of an agent, gives us a set of valuations, say X. If now the extensions of Γ and X are disjoint, the former provides us no information whatsoever as to which outcomes are most preferred by the agent *within* X. The relation $\boldsymbol{\rho}(\Gamma)$, however, does. In particular, it enables us to identify for each particular choice of strategy by the opponents, which are a player's maximum responses. In our proposal we will therefore refer to the maximum elements of the relation representing an agent *i*'s preferences in the subsets of valuations in which the values of all variables are fixed except for those which *i* controls. Figure 8.2 illustrates this point graphically.

8.3 Intermezzo: Veltman's Updates for Defaults

The additional ordinal structure the relational semantics for propositional logic engenders over the set of valuations, is quite superfluous if one's concerns are with classical consequence only. However, the semantics of a considerable number of non-standard variants or extensions of classical propositional logic appeal to a relational structure over the valuations or possible worlds. We have already mentioned qualitative decision theory (*e.g.*, Boutilier (1994)), belief revision (*e.g.*, Gärdenfors (1988)), and non-monotonic consequence relations (*e.g.*, Shoham (1988), Kraus, Lehmann, and Magidor (1990) and Makinson (1994)). The prime example in this respect is, of course, Kripke semantics for modal languages. As in Kripke semantics, this relational structure is often assumed to be given independently by the semantics, rather than induced by syntactic objects, such as formulas.

Veltman's analysis of defaults (Veltman (1996)), however, is different in this respect. There it is suggested that in a proper treatment of defeasible reasoning, some formulas are interpreted as imposing a relational structure on logical space. This enables one to distinguish among any subset of valuations those that are optimal with respect to this structure. In the semantics of formulas of another logical form these optimal valuations play a crucial role. There being a clear parallel between the concluding remarks of the previous section and Veltman's semantical ideas, we will here give a synopsis of the third section of 'Defaults in Update Semantics'.

Classical logic is monotonic in the sense that if a conclusion follows from a collection of premisses Γ, then the same conclusion also follows from any collection of premisses that includes Γ. If premisses are taken to represent the information available to an agent and conclusions the inferences that agent may reasonably draw from the premisses, it has been argued that much of human reasoning exhibits non-monotonic features. In the face of new evidence one may be happy to withdraw conclusions arrived at on the basis of information obtained previously. The new evidence is then said to defeat the conclusion and the conclusion itself is said to be defeasible. *E.g.*, if the only piece of information available is that it normally rains, one could arguably infer that it presumably rains. However, if one obtains as an additional piece of information that it as a matter of fact does not rain, one might be quite willing to retract the conclusion that it rains, as it does not.

Veltman gives a formal account of these and similar phenomena having to do with defeasible reasoning and the order in which information is received. Using a dynamic framework Veltman can account for the contrast between the *acceptability* of texts as (1) and (2), and the *unacceptability* of the sequence (3):

(1) "Normally, it rains. …Presumably, it rains."

(2) "Normally, it rains. …Presumably, it rains. …It does not rain."

(3) "Normally, it rains. …It does not rain. …Presumably, it rains."

Although many of the merits of Veltman's approach lie in its ability to deal with such examples using a dynamic framework, we concentrate on some of its static aspects. The intuition behind Veltman's approach is that a sentence like *"Presumably, it rains"* signifies that it rains in all of the most normal states of affairs that are consistent with the information available. This presupposes that the possible states of affairs can somehow be ordered with respect to normality. A distinguishing mark of Veltman's proposal is that this normality order over the possible states of affairs is determined by sentences like *"Normally, it rains"* that occurred earlier in the text, rather than merely fixed exogenously.

Veltman proposes a propositional modal language $L(A, \{\mathit{normally}, \mathit{presumably}\})$, where *normally* and *presumably* are modalities operating on formulas of the propositional language $L(A)$ only. *I.e.*, the formulas of $L(A, \{\mathit{normally}, \mathit{presumably}\})$ are given by the set $\{\varphi, \mathit{normally}\,\varphi, \mathit{presumably}\,\varphi : \ \varphi \text{ a formula of } L(A)\}$ and there is no nesting of the modalities. The intended readings of *normally* φ and *presumably* φ suggest themselves.

The formulas of $L(A, \{normally, presumably\})$ are interpreted in terms of states consisting of a so-called *expectation pattern* ρ and an *information set* X. An expectation pattern is a reflexive and transitive relation over the valuations for $L(A)$ and the information set X is a subset of valuations, intuitively, containing the possible states of affairs that are compatible with one's factual information about the world.Veltman distinguishes the minimal state **0** and the absurd state **1**, defined by $(S \times S, S)$ and $(Id, \varnothing)$, respectively. The valuations s that are minimal with respect to ρ — *i.e.*, such that $s' < s$ in ρ, for no valuation s' — are called *normal*. If the set of normal worlds in an expectation pattern ρ is not empty, ρ is said to be *coherent*. A state (ρ, X) is an *information state* if ρ is coherent and X non-empty, or if (ρ, X) is the absurd state **1**.

Semantically, formula φ in $L(A, \{normally, presumably\})$ is interpreted as a postfixed *operation* $[\varphi]$ on information states. For each formula φ in $L(A)$, the operation $[\varphi]$ performs an update on the information set of an information state, accommodating the information conveyed by φ without changing the expectation pattern. That is, provided that the update does not render the information set void, for then the absurd state results.

By contrast, if φ is of the form *normally* ψ and the extension $[\![\psi]\!]$ contains a normal world with respect to ρ, we have $[\varphi]$ operate on the expectation pattern ρ of an information state (ρ, X). It leaves the information set X as it was but ρ by removing from it all edges (s, s') with ψ holding in s but not in s'. This renders any valuation s that forces ψ strictly more normal than any valuation in which ψ does not hold but that was as normal as s in the original expectation pattern. As such $[normally\,\psi]$ imposes additional structure on logical space rendering ψ worlds more normal that non-φ worlds without affecting the agents factual information about the world. Rather, $[normally\,\psi]$ refines the expectation pattern by intersecting it with the inverse of the relation $\rho_0(\psi)$, as defined in the previous section. If, however, $[\![\psi]\!]$ fails to contain a normal world with respect to ρ, then updating (ρ, X) with $[normally\,\psi]$ will result in the absurd state **1**.

Finally, $[presumably\,\varphi]$ performs a test on information states. In case φ holds in all valuations that are minimal with respect to the expectation pattern of the information state, $[presumably\,\varphi]$ returns the original information state. Otherwise, it returns the absurd state. Let φ be a formula in $L(A)$. Then — employing notations used throughout this thesis — Veltman's formally definitions are given by:

$$(\rho, X)[\varphi] \;=_{df.}\; \begin{cases} (\rho, X \cap [\![\varphi]\!]) & \text{if } X \cap [\![\varphi]\!] \neq \varnothing, \\ \mathbf{1} & \text{otherwise.} \end{cases}$$

$$(\rho, X)[normally\,\varphi] \;=_{df.}\; \begin{cases} (\rho \cap \rho_0(\varphi)^{\smile}, X) & \text{if } [\![\varphi]\!] \text{ contains a normal world,} \\ \mathbf{1} & \text{otherwise.} \end{cases}$$

$$(\rho, X)[presumably\,\varphi] \;=_{df.}\; \begin{cases} (\rho, X) & \text{if } s \in [\![\varphi]\!], \text{ for all } s \text{ minimal in } X \text{ w.r.t. } \rho, \\ \mathbf{1} & \text{otherwise.} \end{cases}$$

(Here, $\rho_0(\varphi)^{\smile}$ denotes the *inverse* of the relation $\rho_0(\varphi)$.)

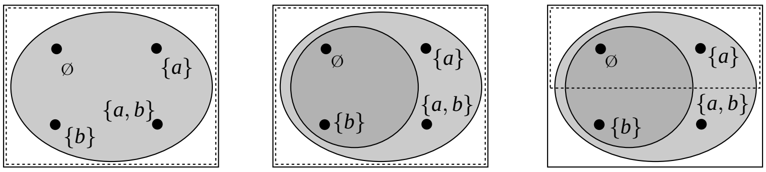


Figure 8.3. The figure on the left depicts $\mathbf{0}$ for a language with a and b as the only propositional variables. The dashed box and the grey balloons indicate the subset of valuations and the expectation patter, respectively. From left to right the figures depict the minimal state $\mathbf{0}$, $\mathbf{0}[normally\,a]$ and $\mathbf{0}[normally\,a][\neg b]$. The valuation $\{a\}$ is now minimal in $\mathbf{0}[normally\,a][\neg b]$, but for instance $\varnothing$ is not. Hence, $\mathbf{0}[normally\,a][\neg b] \Vdash presumably(a \wedge \neg b)$ but $\mathbf{0}[normally\,a][\neg b] \nVdash presumably(\neg(a \vee b))$.

For formulas φ of $L(A, \{normally, presumably\})$ and information states σ define $\sigma \Vdash \varphi$ if and only if $\sigma[\varphi] = \sigma$. Moreover, consequence $\vdash^{\mathrm{V}}$ for this particular system is defined as a relation between the *sequences* of formulas $\varphi_0, \ldots, \varphi_n$ and a formula ψ as follows:

$$\varphi_0, \ldots, \varphi_n \vdash^{\mathrm{V}} \psi \quad \text{iff} \quad \mathbf{0}[\varphi_0] \ldots [\varphi_n] \Vdash \psi.$$

The sequential order of the formulas $\varphi_0, \ldots, \varphi_n$ here makes a difference. *E.g.*, as the formal counterparts of (1), (2) and (3), above, we find:

(1′) $\mathbf{0}[normally\,a][presumably\,a] \neq \mathbf{1}$,

(2′) $\mathbf{0}[normally\,a][presumably\,a][\neg a] \neq \mathbf{1}$,

(3′) $\mathbf{0}[normally\,a][\neg a][presumably\,a] = \mathbf{1}$.

The reader be also referred to Figure 8.3 for further illustration.

The guiding principle behind Veltman's update semantics for defaults is that Boolean formulas φ and those of the form *normally* φ build up an information state. Suppose that $\mathbf{0}[\varphi_0] \ldots [\varphi_n]$ is an information state that is being constructed in the course of an update process and distinct from the absurd state. The constituent expectation pattern is then precisely the inverse of the relation $\boldsymbol{\rho}_0\,(\Theta)$, where Θ is given by exactly those formulas ψ such that the formula *normally* ψ is among $\varphi_0, \ldots, \varphi_n$. Formulas of the form *presumably* φ are then evaluated with respect to the information state constructed. A formula *presumably* φ holds in a non-absurd information state (ρ, X), *i.e.*, $\sigma \Vdash presumably\,\varphi$, if φ holds in all *optimal* states in ρ that are compatible with the factual information represented by X. Optimality is here taken as minimality with respect to the information pattern, but could equally well be defined as maximality with respect to its inverse.

In the next section distributed evaluation games are introduced as a special kind of strategic game. The strategy profiles of these games are the valuations of a propositional language, each player having control over a set of propositional variables. Each player *i* is associated with a theory Γ_i, which is interpreted as the relation $\boldsymbol{\rho}(\Gamma_i)$ on logical space and intuitively reflect the player's preferences. Thus the theory determines the player's preferences in much the same way as formulas of the form *normally* φ build up an expectation pattern in Veltman's default logic. The manipulative power of a player is relative to the propositional variables assigned to her. For each player, logical space is partitioned in blocks that contain valuations that coincide on the values of the variables assigned to her opponents. If it is given that the outcome of the game will be in one particular block, it is then up to her which of the outcomes in that block will prevail. To determine her best response strategies, she has thus to look for the valuations within each block that are optimal with respect to her preference order. As such, each of these blocks of a player's partition relates to her preference order in much the same way as the information set to the expectation pattern in Veltman's framework. Also in the game-theoretical setting, it is the optimal valuations that seem to be relevant.

In order to determine her best-response strategies, however, a player has to find the optimal strategies in *all* of the blocks. If one is interested in the Nash equilibria, one should, moreover, take into account the best-responses for all players. *I.e.*, one has to investigate all blocks in all of the players' partitions with respect to the players' individual preference orders. In Veltman's semantics for defeasible reasoning there is only one information set to be considered. However, in essence the principle remains the same.

8.4 Distributed Evaluation Games

In the Preliminaries a *strategic game* was introduced as a tuple $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$, with N as set of players, and for each player i in N a set of strategies S_i as well as a reflexive and transitive, or empty relation ρ_i over the strategy profiles $\prod_{i\in N} S_i$ which is usually denoted by S.

In this section we define, for each propositional language $L(A)$, a special class of strategic games, which we refer to as *distributed evaluation games*. Henceforth we will assume the set of propositional variables A to be non-empty. The distributed evaluation games provide a formalization of the interactive situations that result if the variables of the propositional language $L(A)$ are construed as binary decision variables the control over which is distributed over a number of individuals. The sets of propositional variables assigned to the players are assumed to be pairwise disjoint — *i.e.*, no joint control over a propositional variable occurs — and to exhaust the set of propositional variables A. If A is non-empty, moreover, each player controls at least one variable and each variable is controlled by one player. The strategies available to each player are given by the different binary choices he can make with respect to his propositional variables. *I.e.*, if A_i is the set of variables assigned to the control of player i, the set of strategies available to i is given by 2^{A_i}.

With each propositional variable controlled by precisely one player, each strategy profile of a distributed evaluation game determines an assignment of a binary value to each of the propositional variables. Thus each strategy profile can be seen as a valuation of $L(A)$ and each valuation as a strategy profile, there being no further restrictions on the strategies available to the players. This is the reason why having S denote both the set of valuations and the set of strategy profiles of a distributed evaluation game is a harmless ambiguity.

Each player i of a distributed evaluation game is thought of as the verifier of a separate theory Γ_i of $L(A)$ and to aim at an outcome of the game that satisfies as much as possible of Γ_i by choosing appropriate values for the propositional variables in his control. This leaves us the issue of a criterion to measure the degree to which a valuation satisfies a theory. The considerations of Section 8.2 concerning the relative logical strength of a theories enable us to be precise in this respect. Accordingly, the preferences of a player i in a distributed evaluation game are given by the relation $\boldsymbol{\rho}(\Gamma_i)$. Consequently, the preferences of i are considerably more gradated than merely distinguishing valuations that satisfy the whole of Γ_i from those those that do not. Formally, we have the following definition.

Definition 8.4.1 *(Distributed evaluation games)* Let $L(A)$ a propositional language on a non-empty set A of propositional variables. A strategic game $\left(N, \{S_i\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ is a *distributed evaluation game* for $L(A)$ if for each $i \in N$:

$$S_i \;=_{df.}\; 2^{A_i} \qquad \text{and} \qquad \rho_i \;=_{df.}\; \boldsymbol{\rho}(\Gamma_i)\,,$$

where Γ_i is a term of a family $\{\Gamma_i\}_{i\in N}$ of theories in $L(A)$ and A_i a term of a family $\{A_i\}_{i\in N}$ of non-empty and pairwise disjoint subsets of propositional variables that partitions A.[3]

For each distributed evaluation game $\left(N, \{2^{A_i}\}_{i\in N}, \{\rho_i\}_{i\in N}\right)$ for $L(A)$, a natural isomorphism exists between its strategy profiles $\prod_{i\in N} 2^{A_i}$ and 2^A, the set of valuations for $L(A)$. Accordingly, we will generally let the latter go proxy for the former.

In the context of distributed evaluation games, we will frequently identify the players with the propositional variables they control. Thus, a partition π of A is taken as the index set N of players and the family $\{A_i\}_{i\in N}$ itself, which assigns control over

[3]For the propositional language $L(\varnothing)$ without propositional variables this definition would leave the set of strategies for a player undefined and as such would not deliver well defined distributed evaluation games. One could, however, treat $L(\varnothing)$ as a special case and define strategic games with one player that has no control over any propositional variables at all. This player could be defined as having only one strategy at his disposal, *viz.*, $\varnothing$. Accordingly, any such game would have merely one strategy profile, *viz.*, the empty set $\varnothing$. Observe in this context that $\varnothing$ is the only element of *Part* $(\varnothing)$. The preferences of the player could be given by the relation $\boldsymbol{\rho}(\Gamma)$ induced by a theory Γ of $L(\varnothing)$ over the strategy profiles. There are then essentially two of such distributed evaluation games for $L(A)$. In the one game the player's preferences are given by the universal relation over the strategy profiles, *i.e.*, by $\{(\varnothing, \varnothing)\}$, in the other by the empty relation, depending on whether the theory defining the player's preferences is consistent or inconsistent. Neither of the two games is particularly interesting from our perspective. Most, if not all, of the subsequent results regarding distributed evaluation games and game-theoretical consequence would also hold if the notion of a distributed evaluation game were extended as to include these two games for $L(\varnothing)$ as well. Yet, including these games for $L(\varnothing)$ would complicate the formulation of the proofs, as each of them would have to treat $L(\varnothing)$ separately as a special case.

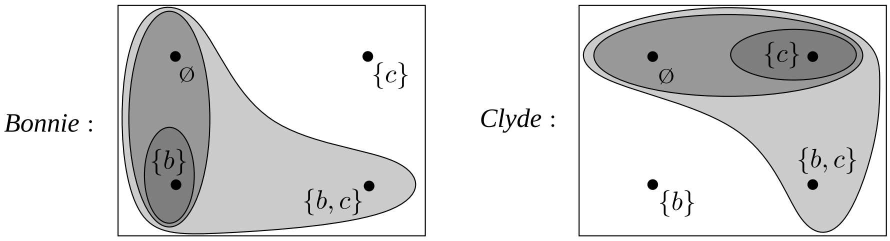


Figure 8.4. The extensions of the formulas in the theories by means of which *Bonnie*'s and *Clyde*'s preferences are defined.

propositional variables to the players, the identity function. *I.e.*, for each $X \in \pi$, we have $A_X = X$. Depending on whether the emphasis is on the player or the set of propositional variables he controls, we use i and, respectively, A_i or π_i to denote members of the partition π. Also, π_{-i} is short for the set $\bigcup_{j \neq i} \pi_j$. In the sequel we will use $\boldsymbol{\Gamma}_I$ to denote a family of theories indexed by I. Furthermore, $(\boldsymbol{\Gamma}_{-i}, \Gamma_i)_I$ is an abbreviation of the family of theories $\boldsymbol{\Gamma}_I^*$ such that for all $j \neq i$, $\boldsymbol{\Gamma}_j^* = \boldsymbol{\Gamma}_j$ and $\boldsymbol{\Gamma}_i^* = \Gamma$. For π a partition of the propositional variables and $\boldsymbol{\Gamma}_\pi$ a family of theories indexed by π, we denote by $G(\boldsymbol{\Gamma}_\pi)$ the distributed evaluation game with π the set of players. Each player i in π has control over the block π_i and her preferences are given by $\boldsymbol{\rho}(\boldsymbol{\Gamma}_i)$. In short, $G(\boldsymbol{\Gamma}_\pi)$ is the strategic game $\left(\pi, \{2^{A_i}\}_{i \in \pi}, \{\boldsymbol{\rho}(\boldsymbol{\Gamma}_i)\}_{i \in \pi}\right)$. We have $\overline{G}(\boldsymbol{\Gamma}_\pi)$ denote the game $G(\boldsymbol{\Gamma}_\pi^*)$, with $\boldsymbol{\Gamma}_i^* = \{\neg\gamma : \ \gamma \in \boldsymbol{\Gamma}_i\}$, for each $i \in \pi$.

Definition 8.4.1 is illustrated by the following example, which is the representation of the infamous Prisoner's Dilemma (*cf.*, page 8, above) as a strategic evaluation game.

Example 8.4.2 *(Prisoner's Dilemma)* Consider the language $L(A)$ with $A = \{b, c\}$ and N a set of players containing as sole elements *Bonnie* and *Clyde*. Let *Bonnie* be assigned control over the propositional variable b and *Clyde* over c, *i.e.*, $A_{Bonnie} = \{b\}$ and $A_{Clyde} = \{c\}$. Let further theory Γ_{Bonnie} be given by $\left\{c \to b, \neg c, \neg c \wedge b\right\}$ and Γ_{Clyde} by $\left\{b \to c, \neg b, \neg b \wedge c\right\}$. These stipulations define a distributed evaluation game, in which the strategies available to *Bonnie* are given by $2^{\{b\}} = \left\{\varnothing, \{b\}\right\}$ and those to *Clyde* by $2^{\{c\}} = \left\{\varnothing, \{c\}\right\}$. Intuitively, the binary decision variables b and c represent *Bonnie*'s and *Clyde*'s respective choices between denying and confessing. For both players, setting a decision variable to 0 means denying and setting it to 1 is to confess. The presence of the formula $c \to b$ in Γ_{Bonnie} conveys *Bonnie*'s preference to confess if *Clyde* does so as well and $\neg c \wedge b$ that she prefers above all the outcome in which she confesses but *Clyde* refrains from doing so. Taking the extensions of the formulas in the theories Γ_{Bonnie} and Γ_{Clyde} we obtain the following sets of sets of

	$\varnothing$	$\{c\}$
$\varnothing$	2 2	3 0
$\{b\}$	0 3	$\mathbf{1}$ $\mathbf{1}$

Figure 8.5. The game matrix of the *Prisoner's Dilemma*. The typical Pareto dominated Nash equilibrium is the outcome bottom right. The figures indicate the *ordinal* preferences of the players.

valuations:

$$\begin{aligned} \mathit{Bonnie}: &\quad \Big\{ \{ \{b\}, \{b,c\}, \varnothing \}, \{ \{b\}, \varnothing \}, \{ \{b\} \} \Big\}, \\ \mathit{Clyde}: &\quad \Big\{ \{ \{c\}, \{b,c\}, \varnothing \}, \{ \{c\}, \varnothing \}, \{ \{c\} \} \Big\}. \end{aligned}$$

Figure 8.4 serves as a graphical representation of how these extensions are related to one another with respect to set inclusion. The preferences of *Bonnie* and *Clyde* over the valuations $2^{\{b,c\}}$ are then given by the relations $\boldsymbol{\rho}\,(\mathit{\Gamma}_{\mathit{Bonnie}})$ and $\boldsymbol{\rho}\,(\mathit{\Gamma}_{\mathit{Clyde}})$, respectively, *i.e.*:

$$\begin{aligned} \mathit{Bonnie}: &\quad \{c\} < \{b,c\} < \varnothing < \{b\}, \\ \mathit{Clyde}: &\quad \{b\} < \{b,c\} < \varnothing < \{c\}. \end{aligned}$$

Thus, we obtain the strategic game depicted in Figure 8.5, with the characteristic (Pareto dominated) Nash equilibrium in bold face. Observe that this distributed evaluation game has a maximum equilibrium, although union of *Bonnie*'s and Clyde's theory, *i.e.*, the theory $\mathit{\Gamma}_{\mathit{Bonnie}} \cup \mathit{\Gamma}_{\mathit{Clyde}}$, is unsatisfiable.

Since distributed evaluation games are fully fledged strategic games, game-theoretical methods can be used to investigate them. Solution concepts may be employed to distinguish valuations that are somehow significant from a game-theoretical perspective. Each family $\boldsymbol{\Gamma}$ of theories and each partition π of propositional variables can thus be associated the set of valuations that comply with a particular solution concept in the distributed evaluation game $G(\boldsymbol{\Gamma}_\pi)$. The role of the solution concept can be seen as analogous to that of set-theoretic intersection in the definition of the extension of a theory on basis of the extensions of its constituent formulas. In the next chapter the notion of *maximal equilibrium* (*cf.*, page 28, above) is used in this manner to formulate a game-theoretical concept of consequence. Distributed evaluation games provide the semantical basis of this definition.

	$\emptyset$	$\{b\}$
$\emptyset$	0, 0	1, 1
$\{a\}$	$\mathbf{1}$, $\mathbf{1}$	2, 0

	$\emptyset$	$\{b\}$
$\emptyset$	0, 0	$\mathbf{1}$, $\mathbf{1}$
$\{a\}$	1, 1	0, 2

	$\emptyset$	$\{a\}$
$\emptyset$	0, 0	1, 1
$\{b\}$	$\mathbf{1}$, $\mathbf{1}$	2, 0

Figure 8.6. Let $\Gamma = \{a \to b, a \wedge b\}$ and $\Theta = \{a \leftrightarrow \neg b\}$. The left matrix represents the game $G(\Gamma_{\{a\}}, \Theta_{\{b\}})$. The matrix in the middle results if the preferences of the row and column player are interchanged. In the one on the right, the row and column players have exchanged the propositional variables they control. From our perspective this difference between the latter two games is immaterial and both are represented by $G(\Theta_{\{a\}}, \Gamma_{\{b\}})$. Their maximum equilibria (in boldface) differ from the game on the left.

The concept of maximum equilibrium pivots on the notion of unilateral deviation of a player from a strategy profile. In a distributed evaluation game $G(\boldsymbol{\Gamma}_\pi)$ it can be expressed in neat set-theoretic terms when a player i can achieve a strategy profile s' by unilaterally deviating from another strategy profile s. Since the strategy profiles of a distributed evaluation game are taken to be the valuations of the respective propositional language and because valuations stripped to their bare essentials are mere subsets of propositional variables, we have in general for any strategy profiles s and s' an any player i in a partition π:

$$(s_{-i}, s_i') \quad = \quad (s \cap \pi_{-i}) \cup (s' \cap \pi_i).$$

A game $G(\boldsymbol{\Gamma}_\pi)$ abstracts, as it were, from the identity of its players and only takes into account their relative powers and their preferences. The following fact states that such an abstraction is quite immaterial for our purposes, in which we focus on maximum and maximal responses and their equilibria. Since it is quite obvious that maximal and maximal equilibria are independent of the identity of the players, we leave the fact without its proof.

Fact 8.4.3 *Consider the distributed evaluation game* $(N, \{2^{A_i}\}_{i\in N}, \{\boldsymbol{\rho}(\Gamma_i)\}_{i\in N})$ *for a propositional language* $L(A)$. *Let* π *be the indexed set of the family* $\{A_i\}_{i\in N}$ *and let* $\{\Theta_X\}_{X\in\pi}$ *be the family of theories such that, for each* $X \in \pi$, *we have* $\Theta_X = \Gamma_i$ *if and only if* $X = A_i$. *Then the maximum (maximal) equilibria of* $(N, \{2^{A_i}\}_{i\in N}, \{\boldsymbol{\rho}(\Gamma_i)\}_{i\in N})$ *coincide with the maximum (maximal) equilibria of* $(\pi, \{2^X\}_{X\in\pi}, \{\boldsymbol{\rho}(\Theta_X)\}_{X\in\pi})$.

To illustrate this point, consider the language $L(\{a, b\})$ and let the partition π be given by $\{\{a\}, \{b\}\}$. Let the theories Γ and Θ be given by $\{a \to b, a \wedge b\}$ and $\{a \leftrightarrow \neg b\}$, respectively. Consider the distributed evaluation game $G(\Gamma_{\{a\}}, \Theta_{\{b\}})$. Indexing Γ with $\{b\}$ and Θ with $\{a\}$ gives rise to the game $G\big(\Theta_{\{b\}}, \Gamma_{\{a\}}\big)$. The maximum equilibria of these games differ (*cf.*, Figure 8.6), illustrating that control matters.

The game $G\big(\Theta_{\{a\}}, \Gamma_{\{b\}}\big)$ could be seen as the result of the players either adopting one another's preferences or swapping the propositional variable they have control over. Focussing on maximum equilibria as we do, however, both scenarios can be seen as different manifestations of the same phenomenon, witness Fact 8.4.3 on page 195. For our purposes, the important thing is the extend to which control over a set of propositional variables is conductive to the achievement of a desirable outcome relative to a preference order defined by a theory. The *identity* of the player who has control over those propositional variables and who entertains those preferences is quite immaterial. Distributed evaluation games precisely capture this formal dependency between preferences and control.

As presented here, the players' preferences in distributed evaluation games are fixed by the relations induced by theories. One could, however, prefer the image of theories *describing* the preferences, in a possibly partial fashion, rather than fixing them. On this conception, a distributed evaluation game would not so much *be* a strategic game, but rather it would *represent* a class of strategic games in which the preference relations comply with the constraints imposed by the theories. Theories could then be understood as partially specifying a preference relations. Making this idea precise, it seems reasonable to stipulate that a preference relation ρ complies with the constraints imposed by a theory Γ, if the relation ρ is included in $\boldsymbol{\rho}(\Gamma)$.[4] The following fact establishes that the two conceptions of distributed evaluation games make no difference with respect to the formulas that hold in all maximum equilibria of distributed evaluation game conceived as as strategic game and those that hold in all maximum equilibria of all strategic games in a distributed evaluation game, conceived of as a collection of strategic games.

Fact 8.4.4 *Let π be a partition of the propositional variables of a language $L(A)$. Let further φ be a formula and $\boldsymbol{\Gamma}_\pi$ a family of theories of $L(A)$. Then, φ holds in all maximum equilibria of $G(\boldsymbol{\Gamma}_\pi)$ iff φ holds in all maximum equilibria of each strategic game $\big(\pi, \{2^{A_i}\}_{i\in\pi}, \{\rho_i\}_{i\in\pi}\big)$ with $\rho_i \subseteq \boldsymbol{\rho}(\boldsymbol{\Gamma}_i)$, for each $i \in \pi$.*

Proof: From right to left the proof is trivial. The left-to-right direction follows immediately from Proposition 2.1.1 on page 28, above. $\dashv$

Both the extension of a theory and the maximum equilibria of a distributed evaluation game single out subsets valuations on the basis of information in the form of formulas. In the truth-theoretical semantics this information comes in the form of a single

[4] Construed thus, a distributed evaluation game is not a strategic game as such; it rather represents a collection of strategic games. In this thesis we employ a notion of a strategic games that is slightly more liberal than the usual notion in that the preference relations need not in general be connected and may even be empty. One may, of course, confine one's attention to, *e.g.*, the subclass of games in which all preference relations are total preorders. Each distributed evaluation game could then be understood as defining a set of such games. Restricting one's attention on such subclasses of strategic games, however, may have serious repercussions for the concept of game-theoretical consequence to be developed in the next chapter. Here we leave it as a subject for future research.

set of formulas, which, intuitively, describe a part of the world. In the game-theoretical case the information is couched in a family of theories, each member of which concerns the preferences of one player. The valuations that are singled out as maximum equilibria of a distributed evaluation game $G(\boldsymbol{\Gamma}_\pi)$ on the basis of a collection of formulas Ψ, *i.e.*, provided that $\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i$ contains formulas in Ψ only, will always include the valuations in the extension of Ψ. Intuitively, this vindicates formally the intuitive presumption that simultaneously accommodating all of each players' preferences is a sufficient condition for a strategy profile to qualify as a maximum equilibrium, though not a necessarily necessary one.

Proposition 8.4.5 *Let $\boldsymbol{\Gamma}_\pi$ be a family of theories in $L(A)$ indexed by a partition π of A. Then, $\bigcap_{i\in\pi} [\![\boldsymbol{\Gamma}_i]\!]$ is contained in the set of maximum equilibria of $G(\boldsymbol{\Gamma}_\pi)$.*

Proof: Straightforward. If $\bigcap_{i\in\pi} [\![\boldsymbol{\Gamma}_i]\!]$ is empty, the proof is trivial, so assume $\bigcap_{i\in\pi} [\![\boldsymbol{\Gamma}_i]\!]$ to be non-empty. Assume for an arbitrary valuation s that $s \in \bigcap_{i\in\pi} [\![\boldsymbol{\Gamma}_i]\!]$. Consider an arbitrary $i \in \pi$ and an equally arbitrary $\gamma \in \boldsymbol{\Gamma}_i$. Then, $s \in [\![\gamma]\!]$. Hence, $[\![\gamma]\!]$ is non-empty, and therefore $(s', s) \in \rho(\gamma)$, for all valuations s'. Hence, also $(s', s) \in \boldsymbol{\rho}(\boldsymbol{\Gamma}_i)$. It follows that s is a maximum response for i. With i having been chosen arbitrarily, s is a maximum equilibrium in $G(\boldsymbol{\Gamma}_\pi)$ as well. $\dashv$

The inverse of this claim, however, does not hold in general. Example 8.4.2 shows that a distributed evaluation game $G(\boldsymbol{\Gamma}_\pi)$ may have maximum equilibria although $\bigcap_{i\in\pi} [\![\Gamma_i]\!]$ is empty.

An interesting issue is the *formation of coalitions* in a distributed evaluation game $G(\boldsymbol{\Gamma}_\pi)$. Our concern is then which game results if the players of a distributed evaluation game join in coalitions and how its formal properties relate to those of the original game. Here we will make the natural but not necessary assumption that each coalition assumes control over the propositional variables that were previously controlled by its individual members. In virtue of Fact 8.4.3, the players of a distributed evaluation game may then be identified with the union of the propositional variables under the control of its members. This leaves the question of how the coalitional preferences relate to those of its constituent members.

On page 33 we described a particular way in which the players of a strategic game can join in coalitions. We had $G_{\boldsymbol{\kappa}}$ denote the game in which the players of the game G have formed the set of coalitions $\boldsymbol{\kappa}$. The coalitional preferences of each κ in $\boldsymbol{\kappa}$ were fixed as the intersection of the preference relations of its members, *i.e.*, as $\rho_\kappa =_{df.} \bigcap_{i\in\kappa} \rho_i$, for each κ in $\boldsymbol{\kappa}$. This way of combining individual preferences into a coalitional preference relation preserves the *strong Pareto property*, *i.e.*, if all players strictly prefer one outcome to another, the coalition as whole will do so as well.

This method of coalition formation is also applicable to distributed evaluation games. Let $G(\boldsymbol{\Gamma}_\pi)$ be a distributed evaluation game and suppose that a set of coalitions $\boldsymbol{\kappa}$ is formed in this manner. This results in a new distributed evaluation game with the coalitions as players. Each coalition κ in $\boldsymbol{\kappa}$ can then be identified with the set of propositional variables $\bigcup_{i\in\kappa} \pi_i$ and its preferences are given by $\boldsymbol{\rho}\left(\bigcup_{i\in\kappa} \boldsymbol{\Gamma}_i\right)$ (*cf.*, Figure 8.7). This observation is laid down formally in the following fact.

$$\underbrace{\overbrace{a_{0_0}, a_{0_1}, a_{0_2}, a_{0_3}}^{\Gamma_0}, \overbrace{a_{1_0}, a_{1_2}, a_{1_3}}^{\Gamma_1}, \overbrace{a_{2_0}, a_{2_2}}^{\Gamma_2}}_{\Gamma_0 \cup \Gamma_1 \cup \Gamma_2}, \underbrace{\overbrace{a_{3_0}, a_{3_1}}^{\Gamma_3}, \overbrace{a_{4_0}, a_{4_1}, a_{4_2}}^{\Gamma_4}}_{\Gamma_3 \cup \Gamma_4}, \underbrace{\overbrace{a_{5_0}, a_{5_1}}^{\Gamma_5}, \overbrace{a_{6_0}, a_{6_1}, a_{6_2}}^{\Gamma_6}}_{\Gamma_5 \cup \Gamma_6}$$

Figure 8.7. Illustration of coalition formation with each of the coalitional preferences the intersection of the preferences relations of its members. Each player i ($0 \leqslant i \leqslant 6$) has control over the variables a_{i_j} and i's preferences are captured by the theory Γ_i. Now suppose that 0, 1 and 2, 3 and 4, and 5 and 6 decide to join in coalitions, respectively. Then, *e.g.*, the coalition $\{3, 4\}$ obtains control over the propositional variables in $\{a_{3_0}, a_{3_1}, a_{4_0}, a_{4_1}, a_{4_3}\}$. The coalitional preferences are then given by the relations $\boldsymbol{\rho}(\Gamma_0 \cup \Gamma_1 \cup \Gamma_2)$, $\boldsymbol{\rho}(\Gamma_3 \cup \Gamma_4)$ and $\boldsymbol{\rho}(\Gamma_5 \cup \Gamma_6)$, respectively.

Fact 8.4.6 *For π a partition of the propositional variables of $L(A)$, let G be the distributed evaluation game $\left(\pi, \left\{2^{A_i}\right\}_{i\in\pi}, \left\{\boldsymbol{\rho}(\Gamma_i)\right\}_{i\in\pi}\right)$. Let $\boldsymbol{\kappa} \in Part(\pi)$ a coalition partition of π. Let further G^* be the game $\left(\pi^*, \left\{2^{A_k}\right\}_{k\in\pi^*}, \left\{\boldsymbol{\rho}(\Gamma_k^*)\right\}_{k\in\pi^*}\right)$, with $\pi^* =_{df.} \left\{\bigcup\kappa : \ \kappa \in \boldsymbol{\kappa}\right\}$ and for each $k \in \pi^*$:*

$$\Gamma_k^* =_{df.} \bigcup_{\substack{i\in\pi \\ i\subseteq k}} \Gamma_i.$$

Then, the maximum (maximal) equilibria of G^* coincide with those of $G_{\boldsymbol{\kappa}}$.

Sketch of proof: By definition 2.1.5, for each $\kappa \in \boldsymbol{\kappa}$ we have $S_\kappa = \prod_{i\in\kappa} S_i$. Because $S_i = 2^{A_i}$ and the propositional variables assigned to the players being pairwise disjoint, then there is a natural isomorphism between $\prod_{i\in\kappa} S_i$ and $2^{\bigcup_{i\in\kappa} A_i}$. Then the claim follows from Fact 8.4.3 as a corollary. ⊣

Thus, for distributed evaluation games $G(\boldsymbol{\Gamma}_\pi)$ and $G(\boldsymbol{\Theta}_{\pi'})$ with $\pi \leqslant \pi'$, we may say that in the latter players of former have joined in coalitions. Each $j \in \pi'$ then represents the coalition $\{i \in \pi : \ i \subseteq j\}$ have been formed. The way in which for each $j \in \pi'$ the theory $\boldsymbol{\Theta}_j$ relates to the theories in $\left\{\boldsymbol{\Gamma}_i : \ i \subseteq j\right\}$ reflects how the coalitional preferences depend on those of the members of the coalition. Letting $\boldsymbol{\Theta}_j$ be the union of the theories in $\left\{\boldsymbol{\Gamma}_i : \ i \subseteq j\right\}$ corresponds to intersecting the members' preference relations. In virtue of Fact 8.4.6, Corollary 8.4.7 presents a special case of Proposition 2.1.8 specifically for distributed evaluation games.

Corollary 8.4.7 *Let $L(A)$ be a propositional language and let π and π' be partitions of A such that $\pi \leqslant \pi'$. Let further $G(\boldsymbol{\Gamma}_\pi)$ and $G(\boldsymbol{\Theta}_{\pi'})$ be distributed evaluation games such that $\boldsymbol{\Theta}_k =_{df.} \bigcup_{\substack{i\in\pi \\ i\subseteq k}} \boldsymbol{\Gamma}_i$, for each $k \in \pi'$. Then, the maximum equilibria in $G(\boldsymbol{\Theta}_{\pi'})$ are also maximum equilibria in $G(\boldsymbol{\Gamma}_\pi)$.*

Sketch of proof: Observe that in general for each set of subsets $\boldsymbol{X}$ and each family of sets of subsets $\{\boldsymbol{Y}_i\}_{i\in I}$ such that $\boldsymbol{X} = \bigcup_{i\in I} \boldsymbol{Y}_i$ we have that $\boldsymbol{\rho}(\boldsymbol{X}) = \bigcap_{i\in I} \boldsymbol{\rho}(\boldsymbol{Y}_i)$.

Hence, in particular, $\boldsymbol{\rho}\,(\boldsymbol{\Theta}_k) = \bigcap_{\substack{i \in \pi \\ i \subseteq k}} \boldsymbol{\rho}\,(\boldsymbol{\Gamma}_i)$, for each $k \in \pi'$. The claim then follows immediately from Proposition 2.1.8 on page 35 and Fact 8.4.6. $\dashv$

Other ways of combining individual preferences have been studied within the field of *social choice theory*. It is an interesting issue how these correspond to ways of combining theories and the theory of game-theoretical consequence advanced in the next chapter may provide a suitable logical framework. Nevertheless, we will not pursue the matter here. The account of coalition formation presented in this section, however, will play a noticeable role in the formal analysis of game-theoretical consequence.

8.5 Formal Properties of Set-Induced Relations

In the previous section distributed evaluation were introduced as a class of strategic game. The strategy profiles of these games coincide with the valuations of a propositional language and the strategies of the players were given by the possible binary choices they could make with respect to the propositional variables they are assigned control over. The preferences of each players are specified by relation $\boldsymbol{\rho}\,(\Gamma)$ a theory determines over the valuations. In the next chapter a game-theoretical concept of consequence is proposed, which semantic formalization depends on the notion of a distributed evaluation game. There the relations $\boldsymbol{\rho}\,(\Gamma)$ play a much similar role in defining players' preferences as the the extensions of the formulas making up a theory did in the traditional semantic account of classical consequence.

This section concerns two formal issues relating to the class of relations induced by theories over the set valuations of a propositional language. The first pertains to the formal delineation of this class within the class of all reflexive and transitive, and otherwise empty, relations over the valuations. The second relates to closure properties of set-induced relations, *i.e.*, it concerns the problem to which sets $\boldsymbol{X}'$ a set $\boldsymbol{X}$ of subsets can be extended such that $\boldsymbol{\rho}\,(\boldsymbol{X})$ equals $\boldsymbol{\rho}\,(\boldsymbol{X}')$.

We first review some of the more elementary properties of the relations $\rho\,(X)$ and $\boldsymbol{\rho}\,(\boldsymbol{X})$. As a first fact we find that no two different subsets X and Y of a set S such that $\rho\,(X)$ and $\rho\,(Y)$ are identical relations on S.

Fact 8.5.1 *Let X and Y be subsets of some set S. Then:*

$$\rho\,(X) = \rho\,(Y) \quad \textit{iff} \quad X = Y.$$

Proof: The right-to-left direction is trivial. For the opposite direction suppose $X \neq Y$. Without loss of generality we may assume there be some $x \in X$ for which $x \notin Y$. In case Y is empty, $(x,x) \notin \rho\,(Y)$ but $(x,x) \in \rho\,(X)$ and *a fortiori* $\rho\,(X) \neq \rho\,(Y)$. If on the other hand Y is not empty there is some $y \in Y$. Then, $(y,x) \notin \rho\,(Y)$ and $(y,x) \in \rho\,(X)$. Again we may conclude that $\rho\,(X) \neq \rho\,(Y)$. $\dashv$

By contrast, $\rho_0(X)$ and $\rho_0(Y)$ may be identical even for distinct X and Y, be it only if either X or Y is the universe and the other the empty set. Both $\rho_0(X)$ and $\rho_0(Y)$

are then the universal relation. For X any subset other than the empty set, the relations $\rho_0(X)$ and $\rho(X)$ coincide. This observation also sustains a corresponding result for relations $\boldsymbol{\rho}(\boldsymbol{X})$ induced by sets of subsets $\boldsymbol{X}$.

Fact 8.5.2 *Let X be a subset of some non-empty set S. Then:*

$$\rho_0(X) = \rho(X) \quad \textit{iff} \quad X \neq \varnothing \qquad \textit{and} \qquad \boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}(\boldsymbol{X}) \quad \textit{iff} \quad \varnothing \notin \boldsymbol{X}.$$

Proof: For the first claim, the proof from right to left is trivial. So assume $X = \varnothing$. Then, $\rho(X) = \varnothing$ and $\rho_0(X) = S \times S$. Since S had been assumed to be non-empty, also $\rho_0(X) \neq \rho(X)$. For the second claim merely observe the following equalities:

$$\boldsymbol{\rho}_0(\boldsymbol{X}) \quad = \quad \bigcap_{X \in \boldsymbol{X}} \rho_0(X) \quad =_{\varnothing \notin \boldsymbol{X}} \quad \bigcap_{X \in \boldsymbol{X}} \rho(X) \quad = \quad \boldsymbol{\rho}(\boldsymbol{X}). \qquad \dashv$$

We also have the following equally easy fact.

Fact 8.5.3 *Let $\boldsymbol{X}$ be a set of subsets of a non-empty set S. Then:*

$$\boldsymbol{\rho}(\boldsymbol{X}) = \varnothing \quad \textit{iff} \quad \varnothing \in \boldsymbol{X}.$$

Proof: Straightforward. From right to left the proof is almost trivial. Merely observe that then $\rho(\varnothing) \in \{\rho(X) : X \in \boldsymbol{X}\}$ and, since $\rho(\varnothing) = \varnothing$, $\boldsymbol{\rho}(\boldsymbol{X}) = \bigcap_{X \in \boldsymbol{X}} \rho(X) = \varnothing$. For the opposite direction assume that $\varnothing \notin \boldsymbol{X}$. In virtue of Fact 8.5.2, then $\boldsymbol{\rho}(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X})$. With the latter begin reflexive, it follows that $\boldsymbol{\rho}(\boldsymbol{X})$ is reflexive as well. Having assumed S to be non-empty, we may conclude that $\boldsymbol{\rho}(\boldsymbol{X})$ is non-empty. $\dashv$

For any subset X, the relation $\rho(X)$ is *not* in general monotone in X. To appreciate this, let X and Y be two non-empty proper subsets of a set S such that Y is also a proper subset of X. Assume that $y \in Y$, $x \in X - Y$ and $z \notin X$. Then, $(y,x) \in \rho(X)$ but $(y,x) \notin \rho(Y)$. Moreover, $(x,z) \in \rho(Y)$ but $(x,z) \notin \rho(X)$. More in general we have the following fact.

Fact 8.5.4 *Let X and Y be* distinct *subsets of some set S. Then:*

$$\rho(X) \subseteq \rho(Y) \quad \textit{iff} \quad X = \varnothing \ \textit{ or } \ Y = S.$$

Proof: From right-to-left the claim is trivial. The opposite direction is by contraposition. So, assume $X \neq \varnothing$ and $Y \neq S$. Hence, $x \in X$ and $z \notin Y$, for some $x, z \in S$. In case Y is empty, we are done immediately, for then $(x,x) \in \rho(X)$ and $(x,x) \notin \rho(Y)$. So for the remainder of the proof we may assume there to be some $y \in S$ such that $y \in Y$. By Fact 8.5.2, moreover, both $\rho(X) = \rho_0(X)$ and $\rho(Y) = \rho_0(Y)$, which simplifies the reasoning.

With the assumption that X and Y be distinct, either $Y \nsubseteq X$, or $Y \subsetneq X$. In the former case, $y' \in Y$ and $y' \notin X$, for some $y' \in S$. Hence, $(y',z) \in \rho(X)$ and $(y',z) \notin \rho(Y)$. In the latter case, $x' \in X$ and $x' \notin Y$, for some $x \in S$. Because $Y \subseteq X$, also $y \in X$. Therefore, $(y,x') \in \rho(X)$ whereas $(y,x') \notin \rho(Y)$. $\dashv$

By contrast, both $\boldsymbol{\rho}_0(\boldsymbol{X})$ and $\boldsymbol{\rho}(\boldsymbol{X})$ are tidily downward monotone in $\boldsymbol{X}$.

Fact 8.5.5 *(Monotonicity) Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets of subsets of a set S. Then:*

$$\boldsymbol{X} \subseteq \boldsymbol{Y} \quad \textit{implies} \quad \boldsymbol{\rho}_0\,(\boldsymbol{Y}) \subseteq \boldsymbol{\rho}_0\,(\boldsymbol{X}) \;\; \textit{and} \;\; \boldsymbol{\rho}\,(\boldsymbol{Y}) \subseteq \boldsymbol{\rho}\,(\boldsymbol{X})\,.$$

Proof: Straightforward. Assume $\boldsymbol{X} \subseteq \boldsymbol{Y}$. Then also $\{\rho_0(X) : \; X \in \boldsymbol{X}\} \subseteq \{\rho_0(Y) : Y \in \boldsymbol{Y}\}$. Hence:

$$\boldsymbol{\rho}_0\,(\boldsymbol{Y}) \;=\; \bigcap \{\rho_0(Y) : \; Y \in \boldsymbol{Y}\} \;\subseteq_{\boldsymbol{X} \subseteq \boldsymbol{Y}}\; \bigcap \{\rho_0(X) : \; X \in \boldsymbol{X}\} \;=\; \boldsymbol{\rho}_0\,(\boldsymbol{X})\,.$$

The reasoning for $\boldsymbol{\rho}\,(\boldsymbol{Y}) \subseteq \boldsymbol{\rho}\,(\boldsymbol{X})$ runs along analogous lines. ⊣

The Scope of Distributed Evaluation Games

The main purpose of this section is to demarcate the class of distributed evaluation games for a propositional language $L(A)$ within the comprehensive class of strategic games that can be defined on the frames of distributed evaluation games. Thus, for each strategic game in the comprehensive class, the players and their strategies are those of some distributed evaluation game for $L(A)$. Moreover, the strategy profiles are given by the valuations for $L(A)$. Restricted thus, the issue boils down giving a precise characterization of the class of the players' preference relations that the definition of distributed evaluation games allows, *i.e.*, of the set $\{\boldsymbol{\rho}\,(\Gamma) : \; \Gamma \text{ is a theory in } L(A)\}$. For the propositional language $L(\varnothing)$, with no propositional variables, the issue is trivial. Then $\varnothing$ is the only valuation and the two preference relations that are possible, *viz.*, the empty relation and $\{(\varnothing, \varnothing)\}$, are represented by $\boldsymbol{\rho}\,(\{\bot\})$ and $\boldsymbol{\rho}\,(\{\top\})$. In case the number of propositional variables in A is finite, we find the set relations induced by the theories of $L(A)$ is complete with respect to the transitive and reflexive, or otherwise empty relations over the valuations. Matters are different if the set of propositional variables is infinite. Then the relations induced by theories on logical space constitute a proper subclass of all preference relations. This phenomenon is comparable with the fact that for a language with a countably infinite number of propositional variables, the extensions of theories do not exhaust the powerset of valuations.

The relations $\rho\,(X)$ have a neat characterization in terms of general properties of relations. We say that a relation ρ on a set S is *bisective* if it is transitive and moreover satisfies the following condition:

$$(*) \qquad \text{for all } x, x', x'' \in S : \;\; x \leqslant x' \;\text{ implies }\; x'' \leqslant x \text{ or } x' \leqslant x''.$$

Observe that the empty relation qualifies as bisective, as in that case both transitivity and $(*)$ are satisfied trivially in virtue of vacuous quantification. Any other bisective relation, however, is both reflexive in addition to being transitive.

Fact 8.5.6 *Any non-empty bisective relation over a set S is reflexive.*

Proof: Let ρ be a non-empty bisective relation over a set S. Then $x \leqslant x'$, for some $x, x' \in S$. Consider an arbitrary y in S. In virtue of ρ satisfying $(*)$, then, $y \leqslant x$ or $x' \leqslant y$. In both cases the reasoning runs along similar lines; here we deal with the former case only. If $y \leqslant x$, again in virtue of $(*)$, either $y \leqslant y$ or $x \leqslant y$. If $y \leqslant y$ we are done immediately. Otherwise, we have $y \leqslant x \leqslant y$ and by transitivity of ρ also $y \leqslant y$. $\dashv$

The following proposition characterizes a bisective relation over a universe S as one which coincides with $\rho(X)$ for some subset X of S. In its proof, as elsewhere in this section, $\uparrow_\rho x$ denotes the set $\{y \in S : \ (x,y) \in \rho\}$, for all elements x of and all relations ρ on a set S.

Proposition 8.5.7 *Let S be a set. Then the set $\left\{\rho(X) \subseteq S \times S : \ X \subseteq S\right\}$ coincides with the set of bisective relations on S.*

Proof: In case S is empty, the empty relation is the only (bisective) relation on S. Also, $\varnothing$ is the only subset of S. Now, observe that $\rho(\varnothing)$ is the empty relation as well. So, for the remainder of the proof we may assume S to be non-empty.

First consider an arbitrary $X \subseteq S$ along with equally arbitrary $x, x', x'' \in S$. We prove that $\rho(X)$ is bisective. For transitivity first assume that both (x,x') and (x',x'') are in $\rho(X)$ as well as that $x \in X$. Since $(x,x') \in \rho(X)$, also $x' \in X$ and because $(x',x'') \in \rho(X)$, moreover, $x'' \in X$. We may conclude that $(x,x'') \in \rho(X)$. To show that $\rho(X)$ satisfies condition $(*)$ as well, assume $(x,x') \in \rho(X)$. Either $x'' \in X$ or $x'' \notin X$. If the former, $(x',x'') \in \rho(X)$; if the latter $(x'',x) \in \rho(X)$. In both cases we are done.

To prove that for an arbitrary bisective relation ρ on S there is a subset X such that $\rho = \rho(X)$, assume ρ to be bisective and consider the set $\bigcap_{x\in S} \uparrow_\rho x$. Suppressing the subscript ρ in $\uparrow_\rho x$, we prove that $\rho = \rho\left(\bigcap_{x\in S} \uparrow x\right)$. In case ρ is empty, $\uparrow x$ is equally empty, for any $x \in S$. Having assumed S to be non-empty, $\bigcap_{x\in S} \uparrow x = \varnothing$. Hence, $\rho\left(\bigcap_{x\in S} \uparrow x\right) = \rho(\varnothing) = \varnothing$. For the remainder of the proof, we may accordingly assume ρ to be non-empty. In virtue of Fact 8.5.6, the relation ρ may be assumed to be reflexive as well.

For the $\subseteq$-inclusion, assume for arbitrary $y, y' \in S$ that $(y,y') \in \rho$. Assume further that $y \in \bigcap_{x\in S} \uparrow x$ and consider an arbitrary $x \in S$. Then $y \in \uparrow x$, *i.e.*, $(x,y) \in \rho$. By transitivity, also $(x,y') \in \rho$, *i.e.*, $y' \in \uparrow x$. With x having been chosen arbitrarily, $y \in \bigcap_{x\in S} \uparrow x$, and we are done.

For the $\supseteq$-inclusion, assume for arbitrary $y, y' \in S$ that $(y,y') \notin \rho$. Then, $y' \notin \uparrow y$ and, therefore, $y' \notin \bigcap_{x\in S} \uparrow x$. It suffices now to prove that $y \in \bigcap_{x\in S} \uparrow x$. So, consider an arbitrary $x \in S$; we prove that $y \in \uparrow x$. By reflexivity, $(y,y) \in \rho$. In virtue of $(y,y') \notin \rho$ and $(*)$, then $(y',y) \in \rho$. Again because of $(*)$, either (x,y') or $(y,x) \in \rho$. In the former case, $(x,y) \in \rho$ since ρ is transitive and $(y',y) \in \rho$. Also in the latter case we have $(x,y) \in \rho$, because $(y,y') \notin \rho$ and $(*)$. With x having been chosen arbitrarily, $y \in \bigcap_{x\in S} \uparrow x$. We may conclude that $(y,y') \notin \rho\left(\bigcap_{x\in S} \uparrow x\right)$. $\dashv$

For a classical propositional language with a countably infinite number of propositional variables, the relations $\rho(\varphi)$ for formulas φ of the language, exhaust the set of bisective relations on the valuations just as little as the extensions of the formulas exhaust the set of subsets of the valuations. Recall that the number of relations over the valuations of a countably infinite propositional language is uncountable, whereas the number of formulas remains countable. Corollary 2.4.2 on page 54 characterizes the set of extensions of a language as the approximations of the subsets of valuations by means of a finite subset of propositional variables. Proposition 8.5.10, gives a similar result for the relational semantics for classical propositional logic. Before getting there, however, we make some more general remarks concerning approximations of relations.

The approximation operators $\overline{apr}$ and $\underline{apr}$ on the powerset of some set S are relative to an equivalence relation ε on S. The coordinate-wise square[5] of an equivalence relation over S is again an equivalence relation over the Cartesian product of S. The coordinate-wise square of ε — denoted by $\varepsilon \otimes \varepsilon$ — can in turn be used to approximate *relations* on S by means of rough sets. Thus, we have for ρ a relation over a set S:

$$\begin{aligned} & (x,x') \in \overline{apr}_{\varepsilon\otimes\varepsilon}(\rho) \\ \text{iff} \quad & \text{for some } (y,y') \in S \times S\colon \big((x,x'),(y,y')\big) \in \varepsilon \otimes \varepsilon \text{ and } (y,y') \in \rho \\ \text{iff} \quad & \text{for some } y,y' \in S\colon (x,y),(x',y') \in \varepsilon \text{ and } (y,y') \in \rho. \end{aligned}$$

When no confusion is likely, we will denote the approximation operations on the relations of S relative to the squares of an equivalence relation ε_X also by $\overline{apr}_X$ and $\underline{apr}_X$. *I.e.*, we will write $\overline{apr}_X(\rho)$ for $\overline{apr}_{\varepsilon_X\otimes\varepsilon_X}(\rho)$.

The approximation operation $\overline{apr}_X$ on relations does not in general preserve transitivity. For a counterexample one consider a base set of three elements $\{a,b,c\}$. Let ρ be the smallest reflexive transitive relation on $2^{\{a,b,c\}}$ containing $(\varnothing,\{a\})$ and $(\{a,c\},\{a,b\})$. Transitivity fails for the relation $\overline{apr}_{\{a,b\}}(\rho)$. Observe in this respect that both $(\varnothing,\{a,c\})$ and $(\{a,c\},\{a,b\})$ are in $\overline{apr}_{\{a,b\}}(\rho)$. The latter because $\rho \subseteq \overline{apr}_{\{a,b\}}(\rho)$. For the former, observe that both $\varnothing \sim_{\{a,b\}}$ and $\{a\} \sim_{\{a,b\}} \{a,c\}$. As a consequence also $\big((\varnothing,\{a\}),(\varnothing,\{a,c\})\big) \in \varepsilon_{\{a,c\}} \otimes \varepsilon_{\{a,c\}}$. Nevertheless, $\big(\varnothing,\{a,b\}\big) \notin \overline{apr}_{\{a,b\}}(\rho)$. In a similar fashion it can be shown that the upper approximation operation does not preserve reflexivity.

Since, every relation $\rho(X)$ *is* transitive, witness Proposition 8.5.7, the set of bisective relations is not closed under taking approximations. The following proposition, however, establishes a general connection between bisective relations and their approximations.

[5] In Preliminaries. Let $\{S_i\}_{i\in I}$ be a family of sets. Let further for each $i \in I$, ρ_i be a relation on S_i. We define the *coordinate-wise product*, or *the product relation, of* $\{\rho_i\}_{i\in I}$ as the relation ρ^* on the generalized Cartesian order over the S_i such that for all $\vec{x},\vec{y} \in \prod_{i\in I} S_i$:

$$(\vec{x},\vec{y}) \in \rho^* \quad \text{iff} \quad \text{for all } i \in I\colon (x_i,y_i) \in \rho_i.$$

The *coordinate-wise square* $\rho \otimes \rho$ of a relation ρ on S is the coordinate-wise product relation of ρ with itself.

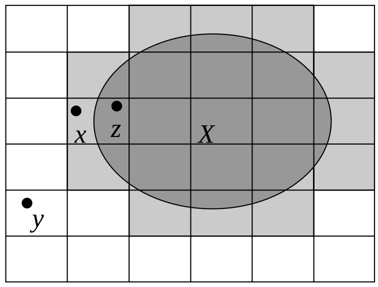


Figure 8.8. Counterexample against $\overline{apr}_{\varepsilon\otimes\varepsilon}\left(\rho(X)\right) \subseteq \rho\left(\overline{apr}_{\varepsilon}(X)\right)$. Let the partition be given by ε. Because $(x,y) \in \rho(X)$, immediately also $(x,y) \in \overline{apr}_{\varepsilon}(\rho(X))$. However, $x \in \overline{apr}_{\varepsilon}(X)$ and $y \notin \overline{apr}_{\varepsilon}(X)$. Therefore, $(x,y) \notin \rho(\overline{apr}_{\varepsilon}(X))$.

Proposition 8.5.8 *Let ε be an equivalence relation on some set S and $\varepsilon \otimes \varepsilon$ its coordinatewise square. Then for each $X \subseteq S$:*

$$\rho\left(\overline{apr}_{\varepsilon}(X)\right) \quad\subseteq\quad \overline{apr}_{\varepsilon\otimes\varepsilon}\left(\rho(X)\right)$$

Proof: In case X is empty, in general, both $\overline{apr}(X) = \overline{apr}(\emptyset) = \emptyset$ and $\rho(\overline{apr}(X)) = \rho(\overline{apr}(\emptyset)) = \rho(\emptyset) = \emptyset$. Therefore, in particular, both $\rho\left(\overline{apr}_{\varepsilon}(X)\right)$ and $\overline{apr}_{\varepsilon\otimes\varepsilon}\left(\rho(X)\right)$ are empty, and we are done immediately. Hence, for the remainder of the proof we may assume X to be non-empty.

Assume $(x,x') \in \rho(\overline{apr}_{\varepsilon}(X))$. If also $(x,x') \in \rho(X)$, then $(x,x') \in \overline{apr}_{\varepsilon\otimes\varepsilon}(\rho(X))$ follows immediately. So, assume that $(x,x') \notin \rho(X)$. Then $x \in X$ and $x' \notin X$. Hence, $x \in \overline{apr}_{\varepsilon}(X)$, and having assumed that $(x,x') \in \rho(\overline{apr}_{\varepsilon}(X))$, also $x' \in \overline{apr}_{\varepsilon}(X)$. Consequently there is an $x'' \in X$ such that $x' \sim_{\varepsilon} x''$. Then also $(x,x'') \in \rho(X)$. Since, both $x \sim x$ and $x' \sim x''$, that $(x,x') \in \overline{apr}_{\varepsilon\otimes\varepsilon}(\rho(X))$ follows, and we are done. $\dashv$

The opposite inclusion, however, does not hold in general. For a counterexample, consider a situation as pictured in Figure 8.8, in which there is a subset X of S and an equivalence relation ε such that there are some elements x and y of S such that neither x nor y are in X. Let there further be an equivalence relation ε with $(x,z) \in \varepsilon$ for some z in X and $(y,z) \in \varepsilon$ for *no* z in X. Accordingly, $x \in \overline{apr}_{\varepsilon}(X)$ and $Y \notin \overline{apr}_{\varepsilon}(X)$. Then, $(x,y) \in \overline{apr}_{\varepsilon\otimes\varepsilon}(\rho(X))$, because $(x,y) \in \rho(X)$ and $\varepsilon \otimes \varepsilon$ is reflexive. However, x is in $\overline{apr}_{\varepsilon}(X)$, whereas y is not and, therefore, $(x,y) \notin \rho(\overline{apr}_{\varepsilon}(X))$.

We say a relation ρ on 2^A is *of finite character* if and only if ρ is a fixed point of the operation $\overline{apr}_X$ on relations for some *finite* subset X of A, *i.e.*, if there is some finite $X \subseteq A$ such that $\rho = \overline{apr}_X(\rho)$. We find that the bisective relations of finite character on the set of valuations of a classical propositional language $L(A)$ coincide with the relations $\rho(\varphi)$, for formulas φ of $L(A)$. In analogy with Theorem 2.4.1 on page 53 we have the following theorem. Recall that $\mathscr{R}(A)$ denotes $\left\{\rho(\varphi) : \ \varphi \text{ a formula in } L(A)\right\}$.

Theorem 8.5.9 *Let $L(A)$ be a propositional language with A as propositional variables and S denote 2^A. For each finite subset $B \subseteq_\omega A$, let further $Fix\,(\overline{apr}_B)$ be defined as the set $\left\{\rho \in S \times S : \ \rho \text{ is bisective and } \rho = \overline{apr}_B(\rho)\right\}$. Then:*

$$\mathscr{R}\,(A) \quad = \quad \bigcup_{B \subseteq_\omega A} Fix\,(\overline{apr}_B)\,.$$

Proof: First consider an arbitrary relation ρ in $\mathscr{R}\,(A)$. Then there is some formula φ of $L(A)$ such that $\rho = \rho\,(\varphi)$. By Proposition 8.5.7, $\rho\,(\varphi)$ is bisective; we show that $\overline{apr}_{\boldsymbol{A}(\varphi)}\left(\rho\,(\varphi)\right) = \rho\,(\varphi)$ proving that it is of finite character as well. As $\rho\,(\varphi) \subseteq \overline{apr}_{\boldsymbol{A}(\varphi)}\left(\rho\,(\varphi)\right)$ is immediate, assume $(x,y) \in \overline{apr}_{\boldsymbol{A}(\varphi)}\left(\rho\,(\varphi)\right)$. Then, there are $x', y' \in 2^A$ such that $x \sim_{\boldsymbol{A}(\varphi)} x'$, $y \sim_{\boldsymbol{A}(\varphi)} y'$ and $(x',y') \in \rho\,(\varphi)$. Now assume $x \in [\![\varphi]\!]$. Then, $x' \in \overline{apr}_{\boldsymbol{A}(\varphi)}([\![\varphi]\!]) = [\![\varphi]\!]$. It follows that $y' \in [\![\varphi]\!]$ as well, and, hence, $y \in \overline{apr}_{\boldsymbol{A}(\varphi)}([\![\varphi]\!]) = [\![\varphi]\!]$. Therefore, $(x,y) \in \rho\,(\varphi)$.

For the opposite direction, consider an arbitrary bisective relation ρ of finite character. We may thus assume there to be a finite subset $B \subseteq_\omega A$ such that $\rho = \overline{apr}_B(\rho)$ as well as a subset $X \subseteq 2^A$ such that $\rho = \rho\,(X)$. If X is empty, we have $\rho = \rho\,(\varnothing) = \varnothing = \rho\,(\bot)$. So, for the remainder of the proof, we may assume X to be non-empty. We prove that $X = \overline{apr}_B(X)$. Then $\rho\,(X) = \rho\,(\overline{apr}_B(X))$. In virtue of Corollary 2.4.2 on page 54 and B being finite, there is also a formula φ such that $[\![\varphi]\!] = \overline{apr}_B(X)$ and, hence, $\rho\,(X) = \rho\,(\overline{apr}_B(X)) = \rho\,(\varphi)$.

Because $X \subseteq \overline{apr}_B(X)$ is immediate, we assume $s \in \overline{apr}_B(X)$, for an arbitrary s and prove that $s \in X$. Then there is some s' such that $s \sim_B s'$ and $s' \in X$. It follows that $(s,s') \in \rho\,(X)$. Moreover, since $s \sim_B s'$ and, trivially, both $s' \sim_B s'$ and $(s',s') \in \rho\,(X)$, also $(s',s) \in \overline{apr}_B(\rho\,(X))$. By the initial assumptions, $\overline{apr}_B(\rho\,(X)) = \rho\,(X)$ and therefore $(s',s) \in \rho\,(X)$. With $s' \in X$, finally, we may conclude that $s \in X$ as well. $\dashv$

The following corollary has a certain likeness with Corollary 2.4.2 on page 54 above, which characterized the extensions of the formulas of $L(A)$ in a much similar way.

Corollary 8.5.10 *Let $L(A)$ be a classical propositional language. Then $\mathscr{R}\,(A)$ coincides with the set of bisective relations of finite character on 2^A, the set of valuations, i.e.:*

$$\mathscr{R}\,(A) \quad = \quad \left\{\overline{apr}_B(\rho\,(X)) : \ X \subseteq 2^A \text{ and } B \subseteq_\omega A\right\}.$$

Proof: The inclusion of $\mathscr{R}\,(A)$ in $\left\{\overline{apr}_B(\rho\,(X)) : \ X \subseteq 2^A \text{ and } B \subseteq_\omega A\right\}$ follows immediate from Theorem 8.5.9. The inclusion in the opposite direction is an immediate consequence of $\overline{apr}_B(\rho\,(X)) = \overline{apr}_B(\overline{apr}_B(\rho\,(X)))$, which is an instance of a rough set law, and again Theorem 8.5.9. $\dashv$

As another corollary we find that, for propositional languages $L(A)$ on a *finite* set of propositional variables, $\mathscr{R}\,(A)$ is complete with respect to all bisective relations on 2^A.

Corollary 8.5.11 *Let A be a finite set of propositional variables on which $L(A)$ is defined. Then, $\mathscr{R}\,(A)$ is complete with respect to the bisective relations on 2^A.*

Proof: Immediate Theorem 8.5.10. ⊣

The relation a *theory* defines over the valuations, however, can be characterized as the limit of the *finite* approximations of a proto-order (*i.e.*, an empty or reflexive and transitive relation) over the valuations. With a finite approximation of a relation over the valuations we mean here the approximation of that relation relative to the equivalence relation defined over the valuations by a finite set of propositional variables, ie, relative to the coordinate square of a relation ε_B where B is a finite subset of propositional variables. First, we prove two preliminary facts.

Fact 8.5.12 *Let ρ be a relation over S. Assume ρ to be either reflexive and transitive or empty. Then:*

$$\rho \;=\; \boldsymbol{\rho}\left(\left\{\uparrow_\rho x : \; x \in S\right\}\right).$$

Proof: First assume that ρ be empty. Then, $\uparrow_\rho x = \varnothing$, for each $x \in S$. Hence, $\boldsymbol{\rho}\left(\left\{\uparrow_\rho x : \; x \in S\right\}\right) = \{\varnothing\}$ and $\boldsymbol{\rho}(\{\varnothing\}) = \rho(\varnothing) = \varnothing$. So, for the remainder of the proof we may assume ρ to be both reflexive and transitive.

First assume that $(y, y') \in \boldsymbol{\rho}(\{\uparrow x : \; x \in S\})$, for arbitrary $y, y' \in S$. Then, $y \in \uparrow x$ implies $y' \in \uparrow x$, for all $x \in S$. By reflexivity of ρ, we have $(y, y) \in \rho$, *i.e.*, $y \in \uparrow y$. With $y \in S$, then also $y' \in \uparrow y$, *i.e.*, $(y, y') \in \rho$.

For the opposite direction, assume that $(y, y') \in \rho$ as well as that $y \in \uparrow x$, for an arbitrary $x \in S$. Then, $(x, y) \in \rho$. By transitivity then also $(x, y') \in \rho$, *i.e.*, $y' \in \uparrow x$. Therefore, $(y, y') \in \rho(\uparrow x)$, and with x having been chosen arbitrarily, eventually, $(y, y') \in \boldsymbol{\rho}(\{\uparrow x : \; x \in S\})$. ⊣

This fact has the following corollary, which says that the class consisting of the reflexive and transitive relations over a set together with the empty relation can be characterized as intersections of bisective relations.

Corollary 8.5.13 *Let ρ be a relation over some set S. Then:*

$$\rho \textit{ is either transitive and reflexive or empty} \quad \textit{iff} \quad \textit{for some } \boldsymbol{X} \subseteq 2^S, \quad \rho = \boldsymbol{\rho}(\boldsymbol{X}).$$

Proof: The left-to-right direction is immediate by Fact 8.5.12. For the opposite direction, assume $\rho = \boldsymbol{\rho}(\boldsymbol{X})$, for some $\boldsymbol{X} \subseteq 2^S$. If $\boldsymbol{X}$ contains the empty set $\varnothing$, then $\left\{\rho(X) : X \in \boldsymbol{X}\right\}$ contains $\rho(\varnothing)$, *i.e.*, the empty relation. Hence, $\boldsymbol{\rho}(\boldsymbol{X}) = \bigcap_{X \in \boldsymbol{X}} \rho(X) = \varnothing$ and by the initial assumption, ρ is empty as well. So, henceforth $\boldsymbol{X}$ may be assumed not to contain the empty set. By Fact 8.5.6, $\rho(X)$ is both reflexive and transitive, for each $X \in \boldsymbol{X}$. An easy check then establishes that $\boldsymbol{\rho}(\boldsymbol{X}) = \bigcap_{X \in \boldsymbol{X}} \rho(X)$ is reflexive and transitive as well. ⊣

This result has as an immediate corollary that for a language $L(A)$ on a *finite* set of propositional variables, the class consisting of the relations $\boldsymbol{\rho}(\Gamma)$ for all theories Γ in $L(A)$ is also complete with respect to the reflexive and transitive and otherwise empty relations over the valuations.

Corollary 8.5.14 *Let $L(A)$ be propositional language on a finite set A of propositional variables. Then:*

$$\rho \text{ is either transitive and reflexive or empty} \quad \text{iff} \quad \text{for some } \Gamma \text{ of } L(A), \ \ \rho = \boldsymbol{\rho}(\Gamma)\,.$$

Proof: Immediately by the Corollaries 8.5.11 and 8.5.14. ⊣

The relations $\boldsymbol{\rho}(\Gamma)$, as defined by the theories of a language $L(A)$, however fail to exhaust the set of reflexive and transitive, or otherwise empty relations over the valuations of $L(A)$, if the set A of propositional variables is infinite. For, in any such case, the subsets of valuations of language $L(A)$ outnumber its formulas and for some $X \subseteq 2^A$ there is no formula φ in $L(A)$ such that $[\![\varphi]\!] = X$. We find that for any such X, there is no theory Γ of $L(A)$ such that relation $\boldsymbol{\rho}(\{X\})$ equals the relation $\boldsymbol{\rho}(\Gamma)$. To appreciate this observe that $\boldsymbol{\rho}(\{X\}) = \rho(X)$ and assume for a *reductio ad absurdum* that there be some Γ such that $\rho(X) = \boldsymbol{\rho}(\Gamma)$. Consider an arbitrary $\gamma \in \Gamma$. By choice of X, then $X \neq [\![\gamma]\!]$. Moreover, $\rho(X) \subseteq \rho(\gamma)$, for otherwise $\rho(X) \nsubseteq \rho(\gamma)$, which would be absurd because $\boldsymbol{\rho}(\Gamma) \subseteq \rho(\gamma)$. By Fact 8.5.4, then either $X = \varnothing$ or $[\![\gamma]\!] = 2^A$. Since $[\![\bot]\!] = \varnothing$ and $\bot$ is a formula of $L(A)$, the former cannot obtain by choice of X. Hence, $[\![\gamma]\!] = 2^A$ and, consequently, $\rho(\gamma)$ is the universal relation over the valuations. With γ having been chosen arbitrarily and the initial assumption, it follows both $\boldsymbol{\rho}(\Gamma)$ and $\rho(X)$ coincide with the universal relation as well. Hence, $\rho(X) = \rho(\top)$. This however yields a contradiction, because, by Fact 8.5.1, $X = [\![\top]\!]$ would follow, which is absurd with $\top$ being a formula of $L(A)$.

As the next best thing, the following theorem characterizes the set of relations defined over the valuations by the theories of a propositional language in the general case.

Theorem 8.5.15 *Let ρ be a reflexive and transitive relation or the empty relation over S, with $S = 2^A$ and A a set of propositional variables. Then:*

$$\rho = \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\rho) \quad \text{iff} \quad \text{for some theory } \Gamma \text{ in } L(A)\text{:} \quad \rho = \boldsymbol{\rho}(\Gamma).$$

Proof: If ρ is the empty relation, then so is any relation $\overline{apr}(\rho)$. Hence, $\rho = \varnothing = \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\rho)$. Now observe that for the theory $\{\bot\}$, the relation $\boldsymbol{\rho}(\{\bot\})$ is empty as well. Thus, for the remainder of the proof, we may assume ρ to be transitive and reflexive.

For the left-to-right direction, assume $\rho = \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\rho)$ and let $\boldsymbol{X} =_{df.} \{\uparrow_\rho s : s \in S\}$. (Henceforth in this proof we omit the subscript ρ in $\uparrow_\rho s$.) By Corollary 2.4.2 on page 54 above, there is a formula φ such that $[\![\varphi]\!] = \overline{apr}_B(X)$, for each $X \in \boldsymbol{X}$ and each finite $B \subseteq_\omega A$. Now let:

$$\Gamma^* \ =_{df.} \ \bigcup_{X \in \boldsymbol{X}} \{\varphi : \ \text{for some } B \subseteq_\omega A, \ [\![\varphi]\!] = \overline{apr}_B(X)\}.$$

We prove that $\rho = \boldsymbol{\rho}(\Gamma^*)$.

For the $\supseteq$-direction, assume an arbitrary pair (s, s') to be in $\boldsymbol{\rho}(\Gamma)$. Consider an arbitrary finite subset B of A; we show that $(s, s') \in \overline{apr}_B(\rho)$. Also consider $\uparrow_\rho s$. Then, there is some γ in Γ^* such that $[\![\gamma]\!] = \overline{apr}_B(\uparrow s)$. Since $(s, s') \in \boldsymbol{\rho}(\Theta)$, in particular $(s, s') \in \rho(\gamma)$ and so $s \in [\![\gamma]\!]$ implies $s' \in [\![\gamma]\!]$. By reflexivity of ρ, trivially, $s \in \uparrow s$ and *a fortiori* $s \in \overline{apr}_B(\uparrow s) = [\![\gamma]\!]$. Hence, also $s' \in [\![\gamma]\!] = \overline{apr}_B(\uparrow s)$. *I.e.*, for some $s'' \in S$, both $s' \sim_B s''$ and $s'' \in \uparrow s$, *i.e.*, $(s, s'') \in \rho$. As trivially, $s \sim_B s$, we may conclude that $(s, s') \in \overline{apr}_B(\rho)$.

For the $\subseteq$-direction, observe that with ρ reflexive and transitive and Fact 8.5.12 we have $\rho = \boldsymbol{\rho}\left(\{\uparrow s : \ s \in S\}\right)$. Hence:

$$\rho \;=\; \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\rho) \;=\; \bigcap_{B \subseteq_\omega A} \overline{apr}_B\big(\boldsymbol{\rho}\left(\{\uparrow s : \ s \in S\}\right)\big).$$

Assume for arbitrary valuations s and s' that $(s, s') \notin \boldsymbol{\rho}(\Gamma^*)$. Hence, for some $\gamma \in \Gamma^*$, both $s \in [\![\gamma]\!]$ and $s' \notin [\![\gamma]\!]$. By definition of Γ^*, there is some finite subset $B \subseteq_\omega A$ and some s_0 in S such that $[\![\gamma]\!] = \overline{apr}_B(\uparrow s_0)$. Then, for all valuations s'' such that $s' \sim_B s''$, $s'' \notin \uparrow s_0$. Also, by transitivity of ε_B, $s''' \in \uparrow s_0$, for all valuations s''' with $s \sim_B s'''$. Hence, $(s, s') \notin \overline{apr}_B(\rho(\uparrow s_0))$. Since $\boldsymbol{\rho}\left(\{\uparrow s : \ s \in S\}\right) \subseteq \rho(\{\uparrow s_0\})$, we obtain $(s, s') \notin \overline{apr}_B\big(\boldsymbol{\rho}\left(\{\uparrow s : \ s \in S\}\right)\big)$. Accordingly, $(s, s') \notin \bigcap_{B \subseteq_\omega A} \overline{apr}_B\big(\boldsymbol{\rho}\left(\{\uparrow s : \ s \in S\}\right)\big)$, *i.e.*, $(s, s') \notin \rho$.

For the opposite direction assume $\rho = \boldsymbol{\rho}(\Gamma)$ for some theory Γ of $L(A)$. Then $\rho \subseteq \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\rho)$ is immediate. Assume that $(s, s') \notin \rho$; we prove that $(s, s') \notin \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\rho)$. In virtue of the assumption, there is some γ such that $s \in [\![\gamma]\!]$ and $s' \notin [\![\gamma]\!]$. Now consider the finite set $\boldsymbol{A}(\gamma)$. For arbitrary valuations s'' and s''' such that $s \sim_{\boldsymbol{A}(\gamma)} s''$ and $s' \sim_{\boldsymbol{A}(\gamma)} s'''$, we have $s'' \in \overline{apr}_{\boldsymbol{A}(\gamma)}([\![\gamma]\!]) = [\![\gamma]\!]$ and $s''' \notin \overline{apr}_{\boldsymbol{A}(\gamma)}([\![\gamma]\!]) = [\![\gamma]\!]$. Therefore, $(s'', s''') \notin \rho(\gamma)$. It follows that $(s, s') \notin \overline{apr}_{\boldsymbol{A}(\gamma)}(\rho(\gamma))$. As $\boldsymbol{\rho}(\Gamma) \subseteq \rho(\gamma)$, also $\overline{apr}_{\boldsymbol{A}(\gamma)}(\boldsymbol{\rho}(\Gamma)) \subseteq \overline{apr}_{\boldsymbol{A}(\gamma)}(\rho(\gamma))$. Therefore, $(s, s') \notin \overline{apr}_{\boldsymbol{A}(\gamma)}(\boldsymbol{\rho}(\Gamma))$ and *a fortiori* $(s, s') \notin \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\boldsymbol{\rho}(\Gamma))$. Having assumed that $\rho = \boldsymbol{\rho}(\Gamma)$, eventually $(s, s') \notin \bigcap_{B \subseteq_\omega A} \overline{apr}_B(\rho)$. $\dashv$

Thus we find that the class of relations induced by theories of a propositional language contains the empty relation as well as those partial preorders over the valuations that can be considered the limit of their own finite approximations. Theorem 8.5.15 sets a bound on the preference relations that can be expressed as a relation $\boldsymbol{\rho}(\Gamma)$ for some propositional theory Γ and, indirectly, demarcates the class of distributed evaluation games from the more general class of games with 2^A as set of strategy profiles.

Closure Conditions for Set-Induced Relations

In classical logic a theory may be closed under its consequences without affecting its deductive properties. At a semantical level, this fact is reflected in that the extension of a theory Γ is identical to the extension of its closure under logical consequence, *i.e.*, in general $[\![\Gamma]\!] = [\![Cn(\Gamma)]\!]$.

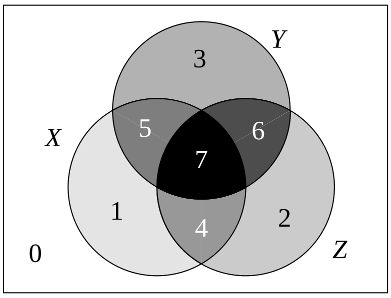


Figure 8.9. Three intersecting sets, X, Y and Z. A pair (x,y) is in the relation $\boldsymbol{\rho}(\{X,Y,Z\})$ the set of sets $\{X,Y,Z\}$ defines over the universe S whenever one can reach y from y without ever moving from an area to an area that is colored lighter. *E.g.*, $(x,y) \in \boldsymbol{\rho}(\{X,Y,Z\})$, for all elements x in area 1 and y in the darker colored area 5. But any element in area 2 and any element in area 5 are incomparable with respect to $\boldsymbol{\rho}(\{X,Y,Z\})$. Some reflection reveals that closing the set $\{X,Y,Z\}$ under intersections and unions would not distort this relation.

The relations over the valuations induced by theories, however, are more sensitive in this respect. In particular, it is not in general the case that the relations $\boldsymbol{\rho}(\Gamma)$ and $\boldsymbol{\rho}(Cn(\Gamma))$ are identical. For an easy counterexample consider a propositional language containing the propositional variables a and b. Obviously we have $a \vee b \in Cn(\{a\})$. For the valuations $\varnothing$ and $\{b\}$ clearly $(\{b\},\varnothing) \in \boldsymbol{\rho}(\{a\})$. However, $(\{b\},\varnothing) \notin \boldsymbol{\rho}(\{a, a \vee b\})$, because $\{b\} \Vdash a \vee b$ but $\varnothing \nVdash a \vee b$. Hence, $(\{b\},\varnothing) \notin \rho(a \vee b)$. The same argument holds for $\boldsymbol{\rho}_0$. It is obvious, however, that a theory Γ may generally be closed under formulas that are *logically equivalent* in the classical sense without affecting $\boldsymbol{\rho}(\Gamma)$.

At a set-theoretic level, a set of sets $\boldsymbol{X}$ cannot in general be closed under supersets without affecting the relation $\boldsymbol{\rho}(\boldsymbol{X})$. On the other hand, different sets of sets may very well induce the same relation on a universe, *i.e.*, $\boldsymbol{\rho}(\boldsymbol{X})$ and $\boldsymbol{\rho}(\boldsymbol{Y})$ may be identical even if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are distinct. This subsection aims at making precise the conditions on sets of sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ that have to be satisfied for the relations $\boldsymbol{\rho}(\boldsymbol{X})$ and $\boldsymbol{\rho}(\boldsymbol{Y})$ to be identical. Closure conditions on theories that preserve relations induced by theories then follow as a matter of course. We find that relations $\boldsymbol{\rho}_0(\boldsymbol{X})$ are slightly better behaved than relations $\boldsymbol{\rho}(\boldsymbol{X})$ and, therefore, we will focus on the former first.

As an example of monotonicity, $\boldsymbol{\rho}_0(\{X,Y\})$ includes $\boldsymbol{\rho}_0(\{X,Y,X\cap Y,X\cup Y\})$. The opposite inclusion, however, also holds in general. For a generalization of this fact, define for $\boldsymbol{X}$ a set of subsets of a set S:

$$\boldsymbol{X}^{\cup} =_{df.} \left\{\bigcup \boldsymbol{X}' : \boldsymbol{X}' \subseteq \boldsymbol{X}\right\} \qquad \boldsymbol{X}^{\cap} =_{df.} \left\{\bigcap \boldsymbol{X}' : \boldsymbol{X}' \subseteq \boldsymbol{X}\right\}.$$

The following proposition says in effect that, for any set of sets $\boldsymbol{X}$, relation $\boldsymbol{\rho}_0(\boldsymbol{X})$ is invariant under taking arbitrary intersections as well as under taking arbitrary unions.

Proposition 8.5.16 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets of subsets of a set S such that $\boldsymbol{X} \subseteq \boldsymbol{Y} \subseteq$*

$\boldsymbol{X}^{\cup} \cup \boldsymbol{X}^{\cap}$. *Then* $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{Y})$.

Proof: By monotonicity immediately $\boldsymbol{\rho}_0(\boldsymbol{Y}) \subseteq \boldsymbol{\rho}_0(\boldsymbol{X})$. Therefore, it suffices to prove that $\boldsymbol{\rho}_0(\boldsymbol{X}) \subseteq \boldsymbol{\rho}_0(\boldsymbol{Y})$. Consider arbitrary $x, x' \in S$ such that $(x, x') \in \boldsymbol{\rho}_0(\boldsymbol{X})$. Consider, furthermore, an arbitrary $Y \in \boldsymbol{Y}$ and assume that $x \in Y$. We prove that $x' \in Y$. Either $Y = \bigcap \boldsymbol{X}'$ or $Y = \bigcup \boldsymbol{X}'$, for some $\boldsymbol{X}' \subseteq \boldsymbol{X}$. Consider this $\boldsymbol{X}'$. In the former case, consider an arbitrary $X \in \boldsymbol{X}'$. Then both $X \in \boldsymbol{X}$ and $x \in X$. Since $(x, x') \in \boldsymbol{\rho}_0(\boldsymbol{X})$, in particular $(x, x') \in \rho_0(X)$. Hence, $x' \in X$. With X having been chosen as an arbitrary element of $\boldsymbol{X}'$, finally $x \in \bigcap \boldsymbol{X}'$, and we may conclude that $x \in Y$.

In the latter case — *i.e.*, if $Y = \bigcup \boldsymbol{X}'$ — we have $x \in X$, for some $X \in \boldsymbol{X}'$. As $\boldsymbol{X}' \subseteq \boldsymbol{X}$, also $X \in \boldsymbol{X}$ and with $(x, x') \in \boldsymbol{\rho}_0(\boldsymbol{X})$, in particular $(x, x') \in \rho_0(X)$. It follows that $x' \in X$ and, subsequently, $x' \in \bigcup \boldsymbol{X}'$, *i.e.*, $x' \in Y$. ⊣

Corollary 8.5.17 *Let* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *be sets of subsets of a set S. Then* $\boldsymbol{X} \subseteq \boldsymbol{Y} \subseteq \boldsymbol{X}^{\cup\cap}$ *implies* $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{Y})$. *Similarly, if* $\boldsymbol{X} \subseteq \boldsymbol{Y} \subseteq \boldsymbol{X}^{\cap\cup}$ *then* $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{Y})$.

Proof: In virtue of monotonicity it suffices to show that $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cup\cap})$ and that $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$. Evidently, $\boldsymbol{X} \subseteq \boldsymbol{X}^{\cap}$ as well as $\boldsymbol{X}^{\cap} \subseteq \boldsymbol{X}^{\cap\cup}$. With Proposition 8.5.16, $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap})$. And again with Proposition 8.5.16, also, $\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$. Hence, $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$. The proof for $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$ is fully analogous. ⊣

For relations induced by a theory over a set of valuations, this corollary means that a theory Γ may be closed under arbitrary conjunctions and disjunctions without affecting the relation $\rho_0(\Gamma)$.

The ground has now been cleared to formulate exact conditions under which the relations $\boldsymbol{\rho}_0(\boldsymbol{X})$ and $\boldsymbol{\rho}_0(\boldsymbol{Y})$ are identical, for possibly distinct sets of sets $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Proposition 8.5.18 *Let* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *be sets of subsets of a set S. Then:*

$$\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{Y}) \quad \textit{iff} \quad \boldsymbol{X}^{\cap\cup} = \boldsymbol{Y}^{\cap\cup}.$$

Proof: For the right-to-left direction, observe that $\boldsymbol{X}^{\cap\cup} = \boldsymbol{Y}^{\cap\cup}$ immediately implies $\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup}) = \boldsymbol{\rho}_0(\boldsymbol{Y}^{\cap\cup})$. Since, by Corollary 8.5.17, both $\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup}) = \boldsymbol{\rho}_0(\boldsymbol{X})$ and $\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup}) = \boldsymbol{\rho}_0(\boldsymbol{X})$, we are done.

The left-to-right direction is less straightforward. Assume the contrapositive $\boldsymbol{X}^{\cap\cup} \neq \boldsymbol{Y}^{\cap\cup}$. Without loss of generality we may assume there be an $X \subseteq S$ such that $X \in \boldsymbol{X}^{\cap\cup}$ and $X \notin \boldsymbol{Y}^{\cap\cup}$. Consider this X. Observe that trivially $\varnothing \subseteq \boldsymbol{Y}^{\cap}$ and $\bigcup \varnothing = \varnothing$. Hence, $\varnothing \in \boldsymbol{Y}^{\cap\cup}$. Moreover, since $\varnothing \subseteq \boldsymbol{Y}$ and $\bigcap \varnothing = S$, both $S \in \boldsymbol{Y}^{\cap}$ and $\{S\} \subseteq \boldsymbol{Y}^{\cap}$. Since $\bigcup \{S\} = S$, also $S \in \boldsymbol{Y}^{\cap\cup}$. It follows that $X \neq \varnothing$ and $X \neq S$. Now consider the set $\boldsymbol{Y}^*$, defined as:

$$\boldsymbol{Y}^* =_{df.} \left\{ Y \in \boldsymbol{Y}^{\cap} : \ Y \subseteq X \right\}.$$

Clearly, $\bigcup \boldsymbol{Y}^* \in \boldsymbol{Y}^{\cap\cup}$ and $\bigcup \boldsymbol{Y}^* \subseteq X$. Due to the assumption that $X \notin \boldsymbol{Y}^{\cap\cup}$, however, $X \neq \bigcup \boldsymbol{Y}^*$. Hence, there is some $x^* \in X$ such that $x^* \notin \bigcup \boldsymbol{Y}^*$. Consider this x^*. We prove that an x in S exists such that is not contained in X and for which it is moreover the case that, for all $Y \in \boldsymbol{Y}$, $x \in Y$, if $x^* \in Y$ as well. *I.e.*:

$(*)$ there is a $x \notin X$ such that for all $Y \in \boldsymbol{Y}$: $x^* \in Y$ implies $x \in Y$.

This suffices because, with $x^* \in X$ then $(x^*, x) \notin \rho_0(X)$. Then $(x^*, x) \notin \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$, because X had been assumed to be in $\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$. With Corollary 8.5.17, then $\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup}) = \boldsymbol{\rho}_0(\boldsymbol{X})$ and $(x^*, x) \notin \boldsymbol{\rho}_0(\boldsymbol{X})$ follows. Moreover, also $(x^*, x) \in \rho_0(Y)$, for each $Y \in \boldsymbol{X}$. Hence $(x^*, x) \in \boldsymbol{\rho}_0(\boldsymbol{Y})$, which would prove the proposition.

We prove $(*)$ by a *reductio ad absurdum*. So assume:

$(**)$ for all $x \notin X$ there is a $Y \in \boldsymbol{Y}$ such that both $x \notin Y$ and $x^* \in Y$.

Then, consider the set $\boldsymbol{Y}^{**}$, defined as:

$$\boldsymbol{Y}^{**} =_{df.} \bigcup_{x \notin X} \left\{ Y \in \boldsymbol{Y} : \ x \notin Y \text{ and } x^* \in Y \right\}.$$

By $(**)$ and the fact that $X \neq S$, we have $\boldsymbol{Y}^{**} \neq \varnothing$. Obviously, $\boldsymbol{Y}^{**} \subseteq \boldsymbol{Y}$ and so $\bigcap \boldsymbol{Y}^{**} \in \boldsymbol{Y}^{\cap}$. By construction, $x^* \in \bigcap \boldsymbol{Y}^{**}$. Moreover, by construction and $(**)$, also $\bigcap \boldsymbol{Z}^{**} \subseteq X$. It would follow that $\bigcap \boldsymbol{Y}^{**} \in \boldsymbol{Y}^*$ as well as that $x^* \in \bigcup \boldsymbol{Y}^*$, *quod non*. $\dashv$

Corollary 8.5.17 has as a special case that $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$, which signifies that closing a set of subsets $\boldsymbol{X}$ under arbitrary intersections and then arbitrary unions does not affect the relation induced on the universe. As a corollary of Proposition 8.5.18 we now find, moreover, that $\boldsymbol{X}^{\cap\cup}$ is also maximal in this respect, *i.e.*, that $\boldsymbol{X}$ can not be extended beyond $\boldsymbol{X}^{\cap\cup}$ without distorting the relation $\boldsymbol{\rho}_0(\boldsymbol{X})$.

Corollary 8.5.19 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets of subsets of S. Then:*

$$\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup} \cup \boldsymbol{Y}) \quad \textit{iff} \quad \boldsymbol{Y} \subseteq \boldsymbol{X}^{\cap\cup}.$$

Proof: From right to left the proof is trivial. So assume that $\boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup} \cup \boldsymbol{Y})$. It can easily be verified that $\boldsymbol{X} \cup \boldsymbol{Y} \subseteq \boldsymbol{X}^{\cap\cup} \cup \boldsymbol{Y} \subseteq (\boldsymbol{X} \cup \boldsymbol{Y})^{\cap\cup}$. By Proposition 8.5.16, then, $\boldsymbol{\rho}_0(\boldsymbol{X} \cup \boldsymbol{Y}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup} \cup \boldsymbol{Y})$. In virtue of the same proposition, also $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X}^{\cap\cup})$. With the initial assumption then it follows that $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X} \cup \boldsymbol{Y})$. Proposition 8.5.18 then gives $\boldsymbol{X}^{\cap\cup} = (\boldsymbol{X} \cup \boldsymbol{Y})^{\cap\cup}$. Because, $\boldsymbol{X}^{\cap\cup} \cup \boldsymbol{Y} \subseteq (\boldsymbol{X} \cup \boldsymbol{Y})^{\cap\cup}$, then $\boldsymbol{X}^{\cap\cup} \cup \boldsymbol{Y} \subseteq \boldsymbol{X}^{\cap\cup}$. We may conclude that $\boldsymbol{Y} \subseteq \boldsymbol{X}^{\cap\cup}$. $\dashv$

Similar results can be obtained for relations $\boldsymbol{\rho}(\boldsymbol{X})$ induced by sets of sets $\boldsymbol{X}$. Unfortunately, things are not as neat as for $\boldsymbol{\rho}_0(\boldsymbol{X})$. Because $\boldsymbol{\rho}(\boldsymbol{X})$ and $\boldsymbol{\rho}_0(\boldsymbol{X})$ are distinct only if $\boldsymbol{X}$ contains the empty set (Proposition 8.5.2), Proposition 8.5.18 also has the following corollary for $\boldsymbol{\rho}(\boldsymbol{X})$.

Corollary 8.5.20 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets of subsets of a set S. Then:*

$$\boldsymbol{\rho}(\boldsymbol{X}) = \boldsymbol{\rho}(\boldsymbol{Y}) \quad \textit{iff} \quad \boldsymbol{X}^{\cap\cup} = \boldsymbol{Y}^{\cap\cup} \ \textit{ or } \ \varnothing \in \boldsymbol{X} \cap \boldsymbol{Y}.$$

Proof: Immediately by the Facts 8.5.2 and 8.5.3 together with Proposition 8.5.18.

It is, however, not in general the case that for $\boldsymbol{X}$ not containing the empty set the relations $\boldsymbol{\rho}(\boldsymbol{X})$ and $\boldsymbol{\rho}(\boldsymbol{X}^{\cap\cup})$ coincide. Observe in this respect that $\boldsymbol{X}^{\cap}$ always contains the empty set, since $\bigcup \varnothing = \varnothing$ and $\varnothing \subseteq \boldsymbol{X}$. Moreover, $\boldsymbol{\rho}(\boldsymbol{X})$ need not be the empty relation, not even if $\boldsymbol{X}$ contains disjoint sets. In any such case, however, $\boldsymbol{X}^{\cap}$ will contain the empty set and $\boldsymbol{\rho}(\boldsymbol{X}^{\cap})$ will also end up empty. In order to obtain the desired closure properties, define for $\boldsymbol{X}$ a set of subsets of a set S:

$$\begin{aligned}
\boldsymbol{X}^{\sqcup} &=_{df.} \left\{ \bigcup \boldsymbol{X}' : \ \boldsymbol{X}' \subseteq \boldsymbol{X} \text{ and } \boldsymbol{X}' \neq \varnothing \right\} \\
\boldsymbol{X}^{\sqcap} &=_{df.} \boldsymbol{X} \cup \left\{ \bigcap \boldsymbol{X}' : \ \boldsymbol{X}' \subseteq \boldsymbol{X} \text{ and } \bigcap \boldsymbol{X}' \neq \varnothing \right\}.
\end{aligned}$$

The idea behind these definitions is essentially the same as those of $\boldsymbol{X}^{\cup}$ and $\boldsymbol{X}^{\cap}$, be it that they prevent the empty set to be included in $\boldsymbol{X}^{\sqcup}$ or $\boldsymbol{Y}^{\sqcap}$ if, and only if, $\boldsymbol{X}$ does not conclude the empty set. It is therefore not surprising that $\boldsymbol{X}^{\sqcup}$ and $\boldsymbol{X}^{\sqcap}$ are in extension very similar to $\boldsymbol{X}^{\cup}$ and $\boldsymbol{X}^{\cap}$, respectively.

Fact 8.5.21 *For $\boldsymbol{X}$ a set of subsets of some set S:*

$$\boldsymbol{X}^{\sqcup} = \begin{cases} \boldsymbol{X}^{\cup} - \{\varnothing\} & \textit{if } \varnothing \notin \boldsymbol{X}, \\ \boldsymbol{X}^{\cup} & \textit{otherwise} \end{cases} \qquad \boldsymbol{X}^{\sqcap} = \begin{cases} \boldsymbol{X}^{\cap} - \{\varnothing\} & \textit{if } \varnothing \notin \boldsymbol{X}, \\ \boldsymbol{X}^{\cap} & \textit{otherwise}. \end{cases}$$

Proof: For the first case, first assume $\varnothing \notin \boldsymbol{X}$. Observe that $\varnothing \notin \boldsymbol{X}^{\sqcup}$. For, assuming otherwise would that $\varnothing = \bigcup \boldsymbol{X}'$ for some non-empty $\boldsymbol{X}' \subseteq \boldsymbol{X}$. This would imply that $\boldsymbol{X}' = \{\varnothing\}$ and hence $\varnothing \in \boldsymbol{X}$, *quod non*. Hence, $\boldsymbol{X}^{\sqcup} \subseteq \boldsymbol{X}^{\cup} - \{\varnothing\}$ and it suffices to prove the opposite inclusion. Consider an arbitrary $X \in \boldsymbol{X}^{\cup} - \{\varnothing\}$. Then, $X \neq \varnothing$ and $X = \bigcup \boldsymbol{X}'$ for some $\boldsymbol{X}' \subseteq \boldsymbol{X}$. Moreover, $\boldsymbol{X}' \neq \varnothing$ by assumption, and so $\bigcap \boldsymbol{X}' = X \in \boldsymbol{X}^{\sqcup}$. Second, assume that $\varnothing \in \boldsymbol{X}$. Observe that trivially, $\boldsymbol{X}^{\sqcup} \subseteq \boldsymbol{X}^{\cup}$. Hence it suffices to prove the opposite inclusion. Consider an arbitrary $X \in \boldsymbol{X}^{\cup}$. If $X = \varnothing$, observe that by the assumption $\{\varnothing\} \subseteq \boldsymbol{X}$. Since $\{\varnothing\} \neq \varnothing$, it follows that $\bigcup \{\varnothing\} = \varnothing \in \boldsymbol{X}^{\sqcup}$. In case $X \neq \varnothing$ the proof is like the case in which $\varnothing \notin \boldsymbol{X}$.

For the second case, first assume that $\varnothing \notin \boldsymbol{X}$. Some reflection on the definitions reveals that then $\varnothing \notin \boldsymbol{X}^{\sqcap}$ and also $\boldsymbol{X}^{\sqcap} \subseteq \boldsymbol{X}^{\cap} - \{\varnothing\}$. Proving the opposite inclusion, consider an arbitrary $X \in \boldsymbol{X}^{\cap} - \{\varnothing\}$. Then, $X = \bigcap \boldsymbol{X}'$ for some $X' \subseteq \boldsymbol{X}$. It follows that $\bigcap \boldsymbol{X}' \neq \varnothing$ and so $\bigcap \boldsymbol{X}' = X \in \boldsymbol{X}^{\sqcap}$. Finally, let $\varnothing \in \boldsymbol{X}$ and observe that $\boldsymbol{X}^{\sqcap} \subseteq \boldsymbol{X}^{\cap}$. Hence, consider an arbitrary $X \in \boldsymbol{X}^{\cap}$. If $X = \varnothing$, we are done immediately by the assumption that $\varnothing \in \boldsymbol{X}$. Otherwise, the reasoning is like the case in which $\varnothing \notin \boldsymbol{X}$. $\dashv$

On basis of this fact and employing Proposition 8.5.16 the following closure properties for $\boldsymbol{\rho}(\boldsymbol{X})$ are obtained.

Proposition 8.5.22 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets of subsets of a set S such that $\boldsymbol{X} \subseteq \boldsymbol{Y} \subseteq \boldsymbol{X}^{\sqcup} \cup \boldsymbol{X}^{\sqcap}$. Then, $\boldsymbol{\rho}(\boldsymbol{X}) = \boldsymbol{\rho}(\boldsymbol{Y})$.*

Proof: First assume $\varnothing \in \boldsymbol{Y}$. Then, $\varnothing \in \boldsymbol{X}^{\sqcap}$ or $\varnothing \in \boldsymbol{X}^{\sqcup}$. In either case, $\varnothing \in \boldsymbol{X}$, by Fact 8.5.21. Then, $\boldsymbol{\rho}(\boldsymbol{X}) = \varnothing = \boldsymbol{\rho}(\boldsymbol{Y})$. So, for the remainder of the proof we may assume that $\varnothing \notin \boldsymbol{Y}$. Then, $\varnothing \notin \boldsymbol{X}$ and in virtue of Fact 8.5.21 both $\boldsymbol{X}^{\sqcup} = \boldsymbol{X}^{\cup}$ and $\boldsymbol{X}^{\sqcap} = \boldsymbol{X}^{\cap}$. Hence $\boldsymbol{X} \subseteq \boldsymbol{Y} \subseteq \boldsymbol{X}^{\cup} \cup \boldsymbol{X}^{\cap}$ and by Proposition 8.5.16, then $\boldsymbol{\rho}_0(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{Y})$. With the assumption that $\varnothing \notin \boldsymbol{X}$ and Fact 8.5.2, $\boldsymbol{\rho}(\boldsymbol{X}) = \boldsymbol{\rho}_0(\boldsymbol{X})$ and $\boldsymbol{\rho}(\boldsymbol{Y}) = \boldsymbol{\rho}_0(\boldsymbol{Y})$ and we may conclude that $\boldsymbol{\rho}(\boldsymbol{Y}) = \boldsymbol{\rho}(\boldsymbol{X})$. $\dashv$

As corollaries of Proposition 8.5.22 we find that a theory Γ may be closed under disjunctions and *consistent* conjunctions without this having consequences for the relation induced on the valuations.

Corollary 8.5.23 *Let Γ be a theory in a propositional language $L(A)$ Let φ and ψ be formulas in Γ. Then, $\boldsymbol{\rho}(\Gamma) = \boldsymbol{\rho}(\Gamma \cup \{\varphi \vee \psi\})$. Moreover, if $\{\varphi, \psi\}$ is classically satisfiable, then also $\boldsymbol{\rho}(\Gamma) = \boldsymbol{\rho}\big(\Gamma \cup \{\varphi \wedge \psi\}\big)$.*

Proof: Immediately by Proposition 8.5.22. $\dashv$

8.6 Conclusion

It becomes natural to conceive of the valuations of a propositional language as strategy profiles of some strategic game, if the control over the values of the propositional variables is distributed over different agents. From a game-theoretical perspective, there seems little reason to confine one's attention merely to games involving two antagonists. In the general case, control over the propositional variables may be distributed over any number of players.

Holding on to the principle of players as verifiers of a theory, the notion of relative logical strength can be invoked to fix the preferences of the players over the valuations. The general idea is that a player prefers a valuation to another if it satisfies a logically stronger subtheory of the theory he aims to verify than the other. Thus the players' preferences acquire a more gradated structure than those of the previous chapter. There they merely distinguished valuations that satisfy the whole theory from those that do not. The additional structure afforded thus, moreover, sustains the use of game-theoretical solution concepts in the selection of socially conspicuous valuations.

These considerations were the makings of the concept of a distributed evaluation game in this chapter. The next chapter they will be put to work. The issue we will be concerned with can informally be formulated as follows. *Which conclusions to draw from a family of theories, given that for each of these theories there is a player with control over a set of propositional variables who seeks to satisfy his theory as well as he can by choosing appropriate values for the variables in his control?* This is a logical problem phrased in game-theoretical terms and distributed evaluation games provide the game-theoretical structure needed for our proposal as to its resolution.

Chapter 9

Game-theoretical Consequence

9.1 Introduction

In their *Theory of Games and Economic Behavior* von Neumann and Morgenstern argued that situations of conflicting interests present a problem that had been "nowhere dealt with in classical mathematics" (von Neumann and Morgenstern (1944), p. 11). They maintained that, due to its interactive nature, a conflict situation could not be analyzed as a traditional optimization or decision problem. Rather, it is a "peculiar and disconcerting mixture of several maximum problems" (*ibid.*, p. 11). An optimization or decision problem for an individual can be represented formally as a function $f(\hat{x}_0, \ldots, \hat{x}_n)$. The individual's predicament is then to choose values for the variables $x_0, \ldots, x_n$ so as to maximize the value of $f(\hat{x}_0, \ldots, \hat{x}_n)$. The variables on which the function depends are regarded as decision variables that are in the control of the individual. Pursuing this conceptualization, a situation of conflict could in similar terms be understood as a *collection* of functions $g_i\,(\hat{x}_0, \ldots, \hat{x}_n)$, each one of which one of the participants tries to maximize by choosing suitable values for the variables in a way that furthers his idiosyncratic interests. Moreover, the variables on which these functions depend may overlap and the parties involved may have control over only some of the relevant variables. This makes that the optimal choices for an individual's variables, from his perspective, may be dependent on the very choices the other participants make in their effort to maximize their functions from their respective points of view. The issue may thus evoke a sense of immanent circularity.

The variety of interests as well as their interdependence make that there is no univocal principle as to what to consider a reasonable solution of a situation of conflict. Traditional notions of optimality were thought to be no longer adequate for such problems and new mathematical notions — *viz.*, game-theoretical solution concepts — had to be developed to take over their role (*ibid.*, page 39, also compare the introduction to this thesis). In non-cooperative settings Nash equilibrium is archetypical in this respect.

Having distinguished optimization problems and game-theoretical problems thus, the satisfiability problem for Classical Propositional Logic (CPC) could be classified as an optimization problem with respect to truth. A formula — or a theory, for that matter — is thought of as a function in the propositional variables that can take one of two values, true or false. The issue is then to choose values for the propositional variables, if that is possible, so as to satisfy the formula in question. Classical logical consequence can be understood in similar deliberative terms: a formula φ follows from a collection of premisses Γ if and only if, each choice for the truth values of the propositional variables (henceforth a *valuation*) that succeeds in satisfying all formulas in Γ, is a choice that makes φ hold as well.

As in this formulation there is present a definite element of choice with respect to the possible truth-assignments, we come to think of propositional variables as binary decision variables that are somehow controllable. The accompanying image of a logical possibility is that of a situation that obtains as the result of the decisions of an individual, rather than that of an unalterable state of affairs. This is the very view that was taken in the previous chapters of this thesis. In line with this, it also becomes natural to consider the case in which control over the propositional variables is distributed over multiple agents. Logical space then assumes a game-theoretical aspect with the valuations as strategy profiles. Pursuing this line of thought, each of the agents could be bestowed his own satisfiability problem, *i.e.*, a theory to satisfy.

In analogy with the relation between optimization and game-theoretical problems, these considerations give rise to the following issue, which can be regarded as the game-theoretical counterpart of the classical problem of logical consequence. *Which conclusions is one to draw from a family of theories, given that, for each of these theories, there is a player who controls a (disjoint) set of propositional variables and who seeks to satisfy his theory as well as he can by choosing appropriate values for the variables in his control?* This is a logical question, at the basis of which there is a game-theoretical problem. For its resolution we resort to the game-theoretical notion of a *maximum equilibrium*, introduced on page 28 of this thesis.

In the previous chapter we saw how each particular distribution of the propositional variables π and each particular family $\boldsymbol{\Gamma}_\pi$ of theories define a unique strategic game, *viz.*, the corresponding *distributed evaluation game* given by $G(\boldsymbol{\Gamma}_\pi)$. We propose to consider as the consequences of the family and the distribution, those formulas that are satisfied in maximum equilibria of the accompanying game. This defines a game-theoretical concept of consequence, which conservatively extends to a concept of consequence that relates pairs of families of theories. The following example illustrates the intuitions behind these considerations.

Example 9.1.1 Consider a situation involving the propositional language $L(\{a,b\})$ with only two variables, a and b. Let π be the partition $\{\{a\},\{b\}\}$ of this set of variables. Suppose further that the player with control over b wishes to satisfy $a \wedge \neg b$ and that the player with control over b desires the formula $\neg(a \vee b)$ to be true. The matrix of the ensuing distributed evaluation game is depicted in Figure 9.1. There are two maximum equilibria, *viz.*, the valuations $\{a\}$ and $\{a,b\}$. Since a satisfied by both equilibria, *i.e.*, in both the valuations $\{a\}$ and $\{a,b\}$, we find that, *e.g.*, a is

	$\varnothing$	$\{b\}$
$\varnothing$	1 0	0 0
$\{a\}$	**0** **1**	**0** **0**

Figure 9.1. The row player has control over *a*, and the column player assigns values to *b*. The numerical values merely represent the *ordinal* structure of the players' preferences. The two maximum equilibria are in boldface. Both satisfy *a*.

a game-theoretical consequence of $\left\{\{a \wedge \neg b\}_{\{a\}}, \{\neg(a \vee b)\}_{\{b\}}\right\}$. However, b does not follow game-theoretically because it not satisfied by the valuation $\{a\}$, although the latter is an equilibrium.

The distributed evaluation game imposes a game-theoretical structure on logical space for game-theoretical consequence in a similar way as the set of extensions of the formulas a theory consists of imposes a set-theoretical structure on the set of valuations in a classical setting. By means of a solution concept particular valuations can then be distinguished from others and investigated with respect to the formulas and theories they satisfy. The function of the solution concept, here maximum equilibrium, can thus be compared to that of the set-theoretic operation of intersection: singling out particular valuations on the basis of the structures induced by theories on logical space (*cf.*, Figure 9.2). The image these reflections are meant to evoke is that of a common pattern in which theories define a structure over the set of valuations and relative to which particular valuations are singled out as the more significant ones and studied with respect to, *e.g.*, the formulas and theories they satisfy. Commonly, it is the valuations that are in, a specified sense, optimal with respect to this structure that are semantically relevant in this way. *E.g.*, the extension $[\![\Gamma]\!]$ or the maximum valuations in the relation $\boldsymbol{\rho}(\Gamma)$ for classical logic, the most normal worlds within the information set relative to an expectation pattern for Veltman's update semantics for defaults (*cf.*, Section 8.3), and the maximum equilibria for game-theoretical consequence.

The additional structure of distributed evaluation games makes that the logical concept of game-theoretical consequence does sufficient justice to the interactive and interdependent nature of the underlying game-theoretical issues. In a conflict situation, a player may have to decide in the face of the possibility that whether he can achieve a most preferred outcome may depend on the choices of the other players. This contingency need not leave the player at a loss — nor the game-theorist examining the game from the outside. A particular choice of action may thus guarantee a player an optimal outcome relative to each possible choice of action of the opponents, without guaranteeing a most preferred outcome in all cases. For each course of action a player's opponents may decide upon, there is a set of outcomes that are still possible. The out-

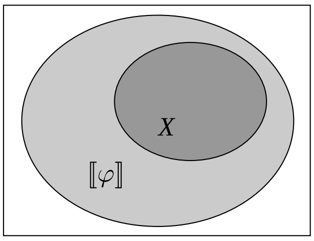


Figure 9.2. Let X be a set of valuations. If X coincides with $[\![\Gamma]\!]$ the formula φ is a classical consequence of Γ, because of the inclusion of X in $[\![\varphi]\!]$. If, however, the valuations are looked upon as the strategy profiles of some game, X could, *e.g.*, have been singled out the set of the maximum equilibria of a distributed evaluation game $G\big(\{\Gamma_i\}_{i\in\pi}\big)$. Then, φ is said to be a *game-theoretical consequence* of $\{\Gamma_i\}_{i\in\pi}$.

comes relevant for determining the maximum equilibria of a game are those that are optimal relative to the player's preferences within any such set, rather than those most preferred by a player within the whole set of outcomes.

9.2 Defining Game-Theoretical Consequence

The purport of the previous section is that game-theoretical consequence can be regarded as the game-theoretical counterpart of classical logical consequence, if the latter is understood in decision theoretical terms. We suggested to consider as game-theoretical consequences of a family of theories $\boldsymbol{\Gamma}_\pi$ indexed by a partition π of the propositional variables those formulas that hold in all maximum equilibria of the distributed evaluation game $G(\boldsymbol{\Gamma}_\pi)$. In this section we give a precise definition.

Just as the notion of classical consequence that relates theories and formulas extends to a relation between theories, the concept of game-theoretical consequence can conservatively be extended to a relation that connects families of theories. On a semantical level this extended relation of game-theoretical consequence compares the maximum equilibria of two distributed evaluation games. With each family of theories $\boldsymbol{\Gamma}_\pi$ we associate *two* games: $G(\boldsymbol{\Gamma}_\pi)$ as well as $\overline{G}(\boldsymbol{\Gamma}_\pi)$. Recall that the latter game is basically identical to the former, be it that the preferences of each player i are given by $\bar{\boldsymbol{\rho}}\,(\boldsymbol{\Gamma}_i)$ — *i.e.*, by the relation $\boldsymbol{\rho}\,(\{\neg\gamma\colon \gamma \in \boldsymbol{\Gamma}_i\})$ — rather than by $\boldsymbol{\rho}\,(\boldsymbol{\Gamma}_i)$ (*cf.*, page 193, above). This provides us with the appropriate dual notions to define game-theoretical consequence. Throughout this chapter, we will assume that the set of propositional variables of the propositional languages considered is not empty.

Definition 9.2.1 *(Game-theoretical consequence)* For partitions π and π' of a set of propositional variables A and families of theories $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ of a propositional language $L(A)$, define:

$$\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}$$
$$\text{iff}$$
$$\text{no maximum equilibrium of } G(\boldsymbol{\Gamma}_\pi) \text{ is a maximum equilibrium of } \overline{G}(\boldsymbol{\Theta}_{\pi'}).$$

For any pair π and π' of partitions of A, we denote by $\Lambda_{\pi,\pi'}$ the binary relation defined as $\left\{(\boldsymbol{\Gamma}_\pi, \boldsymbol{\Theta}_{\pi'}) : \ \boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}\right\}$. The set $\{\Lambda_{\pi,\pi'} : \ \pi, \pi' \in \mathit{Part}\,(A)\}$ we denote by $\boldsymbol{\Lambda}_A$, omitting the subscript when clear from the context.

Observe that this definition is quite in line with the truth-theoretical characterization of classical consequence. The standard semantics for classical propositional logic compares the intersection of the extensions of the formulas in the one theory with the union of the extensions of the formulas in the other theory with respect to set inclusion. Due to the duality of union and intersection, this warrants a well balanced and symmetric system as exemplified by its sound and complete sequent calculi. In the definition of game-theoretic consequence duality is likewise the guiding principle. Observe that a theory Θ follows classically from a theory Γ if and only if no valuation that is in the extension of Γ is in the extension of $\left\{\neg\vartheta : \vartheta \in \Theta\right\}$, witness the following equalities:

$$\langle\!\langle \Theta \rangle\!\rangle \quad = \quad \bigcup\left\{[\![\vartheta]\!] : \ \vartheta \in \Theta\right\} \quad = \quad \overline{\bigcap\left\{\overline{[\![\vartheta]\!]} : \ \vartheta \in \Theta\right\}} \quad = \quad \overline{[\![\left\{\neg\vartheta : \ \vartheta \in \Theta\right\}]\!]},$$

Hence, also:

$$[\![\Gamma]\!] \subseteq \langle\!\langle \Theta \rangle\!\rangle \quad \text{iff} \quad [\![\Gamma]\!] \subseteq \overline{[\![\,\{\neg\vartheta : \ \vartheta \in \Theta\}\,]\!]} \quad \text{iff} \quad [\![\Gamma]\!] \cap [\![\,\{\neg\vartheta : \ \vartheta \in \Theta\}\,]\!] = \varnothing.$$

However cumbersome, this paraphrasis of the semantical characterizition of classical consequence exposes its structural similarity with the formal definition of game-theoretical consequence.

For an example of the workings of this definition the reader consult Figure 7.1 on page 159, which shows that $\left\{\ \{a\}_{\{a\}}, \{a \wedge b\}_{\{b\}}\ \right\} \vDash \left\{\ \{a \wedge b\}_{\{a\}}, \{a\}_{\{b\}}\ \right\}$, as well as $\left\{\ \{a\}_{\{a\}}, \{a \wedge b\}_{\{b\}}\ \right\} \vDash \left\{\ \{b \to a\}_{\{a\}}, \{a \wedge \neg b\}_{\{b\}}\ \right\}$. Figure 9.3 provides a slightly more complicated example, in which two distributed evaluation games are compared that differ in the assignment of the propositional variables to the players.

In the remainder this chapter concerns this game-theoretical consequence relation as based on maximum equilibrium. In the next section we will investigate how it behaves with respect to various structural and logical properties such as monotony, reflexivity and cut. In this context its relation with classical propositional logic will come to be scrutinized as well. We also give a set-theoretical characterization of game-theoretical consequence using the machinery provided by rough sets.

9.3 Structural Properties

In Section 2.3, above, we introduced a propositional logic as a pair consisting of a language $L(A)$ on a set of propositional variables A and a consequence relation connecting

	$\varnothing$	$\{c\}$	$\{d\}$	$\{c,d\}$
$\varnothing$	**01** **10** $_{1}$	01 00 $_{2}$	00 10 $_{3}$	00 00 $_{4}$
$\{a\}$	01 00 $_{5}$	**11** **10** $_{6}$	00 00 $_{7}$	10 10 $_{8}$
$\{b\}$	00 10 $_{9}$	10 00 $_{10}$	01 11 $_{11}$	11 01 $_{12}$
$\{a,b\}$	00 00 $_{13}$	10 10 $_{14}$	01 01 $_{15}$	**11** **11** $_{16}$

	$\varnothing$	$\{b\}$	$\{d\}$	$\{b,d\}$
$\varnothing$	10 01 $_{1}$	10 00 $_{9}$	11 01 $_{3}$	**11** **11** $_{11}$
$\{a\}$	10 11 $_{5}$	00 01 $_{13}$	**11** **11** $_{7}$	01 01 $_{15}$
$\{c\}$	11 01 $_{2}$	01 11 $_{10}$	10 00 $_{4}$	00 10 $_{12}$
$\{a,c\}$	**11** **11** $_{6}$	01 01 $_{14}$	10 10 $_{8}$	00 00 $_{16}$

Figure 9.3. In the game on the left the preferences of the row player has control over a and b and entertains preferences that are captured by the formulas $a \leftrightarrow c$ and $b \wedge d$. The preferences of the column player, who has control over c and d, are summarized by the formulas $(a \vee b) \wedge c$ and $b \leftrightarrow d$. The payoff entries indicate the ordinal preferences assuming the product ordering over the strings 00, 01, 10 and 11, *i.e.*, 11 and 00 are top and bottom, respectively, and 01 and 10 mutually incomparable. The matrix on the right depicts the distributed evaluation game $\overline{G}\big(\big\{\, \{a \leftrightarrow b, c \wedge d\}_{\{a,c\}}\,, \{\neg\,(a \vee c) \wedge b, c \leftrightarrow d\}_{\{b,d\}}\,\big\}\big)$. The valuation $\{a,c\}$ (indicated by the number six in the right-bottom corner) is a maximum equilibrium in both games, and therefore $\big\{\, \{a \leftrightarrow b, c \wedge d\}_{\{a,c\}}\,, \{\neg\,(a \vee c) \wedge b, c \leftrightarrow d\}_{\{b,d\}}\,\big\}$ does not follow game-theoretically from $\big\{\, \{a \leftrightarrow c, b \wedge d\}_{\{a,b\}}\,, \{(a \vee b) \wedge c, b \leftrightarrow d\}_{\{c,d\}}\,\big\}$.

theories of $L(A)$. Game-theoretical consequence, by contrast, is defined between families of theories. This reflects exactly the disparity in the distributed and interactive nature of the problem underlying game-theoretical consequence and the decision theoretical character of the one at the basis of classical consequence.

We propose to extend the formal notion of a propositional logic for a propositional language $L(A)$ as consequence relation defined over the families of theories of $L(A)$ indexed by partitions of A. Although this might seem a radical departure from the original notion of a propositional logic, it is rather meant to conservatively extend it. Made explicit in this manner, the concept of control over propositional variables is rendered amenable to logical analysis.

Classical propositional logic reappears under a guise. The following proposition establishes game-theoretical consequence as a conservative extension of classical propositional logic. Intuitively, it says that the game-theoretical problem of consequence reduces to that of classical consequence if there is only one player who wields control over all propositional variables.

Proposition 9.3.1 *Let $\mathit{\Gamma}$ and $\mathit{\Theta}$ be theories of a propositional language $L(A)$. Then:*

$$\{\mathit{\Gamma}_A\} \vDash \{\mathit{\Theta}_A\} \quad \textit{iff} \quad \mathit{\Gamma} \vdash^{\mathrm{CPC}} \mathit{\Theta}.$$

Proof: Since, in both $G(\{\mathit{\Gamma}_A\})$ and $\overline{G}(\{\mathit{\Theta}_A\})$ in either game there is one player, here denoted by A, who has control over the full set of propositional variables. Accordingly, the set of maximum responses of player A is identical to the set of maximum equilibria in both games. Hence, in virtue of Proposition 8.2.3, it suffices to prove that the set of maximum responses of these players in $G(\{\mathit{\Gamma}_A\})$ and those in $\overline{G}(\{\mathit{\Theta}_A\})$ coincide with the sets of maximum elements of $\boldsymbol{\rho}\,(\mathit{\Gamma}_A)$ and $\bar{\boldsymbol{\rho}}\,(\mathit{\Theta}_A)$, respectively. First observe that for any two valuations s and s':

$$(s_{-A}, s'_A) \;=\; \big(s \cap \overline{A}\big) \cup (s' \cap A) \;=\; \varnothing \cup s' \;=\; s'.$$

Consequently, for any valuation s we may reason as follows:

$$\begin{aligned} s \in \mathit{max}\,(\boldsymbol{\rho}\,(\mathit{\Gamma}_A)) \quad &\text{iff} \quad \text{for all } s' : (s', s) \in \boldsymbol{\rho}\,(\mathit{\Gamma}_A) \\ &\text{iff} \quad \text{for all } s' : \big((s_{-A}, s'_A), s\big) \in \boldsymbol{\rho}\,(\mathit{\Gamma}_A) \\ &\text{iff} \quad s \text{ is a maximum response for } A \text{ in } G(\{\mathit{\Gamma}_A\}). \end{aligned}$$

In a similar fashion we can demonstrate that $s \in \mathit{max}\,(\bar{\boldsymbol{\rho}}\,(\mathit{\Theta}_A))$ if and only if s is a maximum response for A in $\overline{G}(\{\mathit{\Theta}_A\})$. ⊣

The next proposition also connects game-theoretical consequence and classical propositional logic. It guarantees the extrapolation of negative facts about the former to the latter.

Proposition 9.3.2 *Let π and π' be partitions of a set of propositional variables A and let $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories in $L(A)$. Then:*

$$\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'} \quad \textit{implies} \quad \bigcup_{i\in\pi} \boldsymbol{\Gamma}_i \vdash^{\mathrm{CPC}} \bigcup_{i\in\pi'} \boldsymbol{\Theta}_i.$$

Proof: It suffices to show that $\bigcap_{i\in\pi} [\![\boldsymbol{\Gamma}_i]\!]$ is included in the set of maximum equilibria of $G(\boldsymbol{\Gamma}_\pi)$ and that the every valuation s that is *not* a maximum equilibrium in $\overline{G}(\boldsymbol{\Theta}_{\pi'})$ is included in $\bigcup_{i\in\pi'} \langle\!\langle \boldsymbol{\Theta}_i \rangle\!\rangle$. Both claims follow from Proposition 8.4.5, the former immediately, the latter by some additional but straightforward set-theoretical reasoning. ⊣

In the opposite direction, the proposition does not hold in general. If some trivializing requirements are met, however, it does. The idea behind the next result is that if each of the players have control over all propositional variables on which the formulas of the theories representing their respective preferences depend, then the they can each achieve their individual ends independently of the decisions the other players make. In any such case there is no interesting interaction between the players. Each of them lives and acts as it were in his own compartment of the world demarcated by the propositional variables he controls.

Proposition 9.3.3 *Let π and π' be partitions of a set of propositional variables A and let $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories in $L(A)$. Assume that $\boldsymbol{A}(\gamma) \subseteq i$, for all $\pi_i \in \pi$ and all $\gamma \in \boldsymbol{\Gamma}_i$, and that $\boldsymbol{A}(\vartheta) \subseteq \pi_i$, for all $i \in \pi'$ and all $\vartheta \in \boldsymbol{\Theta}_i$. Then:*

$$\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i \vdash^{\text{CPC}} \bigcup_{i\in\pi'} \boldsymbol{\Theta}_i \quad \textit{iff} \quad \boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}.$$

Proof: The right-to-left direction is dealt with by Proposition 9.3.2. For the opposite direction, we reason as follows.

If $\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i$ contains a contradiction, $\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i \vdash^{\text{CPC}} \bigcup_{i\in\pi'} \boldsymbol{\Theta}_i$ holds trivially. In this case it is equally trivial that the game $G(\boldsymbol{\Gamma}_\pi)$ has no maximum equilibria, as $\bigcap_{i\in\pi} \boldsymbol{\rho}(\boldsymbol{\Gamma}_i) = \varnothing$. Similarly, if $\bigcup_{i\in\pi'} \boldsymbol{\Theta}_i$ contains a tautology, then also $\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i \vdash^{\text{CPC}} \bigcup_{i\in\pi'} \boldsymbol{\Theta}_i$. Observe that $\bar{\boldsymbol{\rho}}(\varphi)$ is the empty relation for tautologies φ, and hence the game $\overline{G}(\boldsymbol{\Theta}_{\pi'})$ has no maximum equilibria. Accordingly, also in this case, $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}$. So, for the remainder of the proof we may assume $\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i$ to contain no contradictions and $\bigcup_{i\in\pi'} \boldsymbol{\Theta}_i$ no tautologies.

Assume $\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i \vdash^{\text{CPC}} \bigcup_{i\in\pi'} \boldsymbol{\Theta}_i$. Assume further an arbitrary valuation s to be a maximum equilibrium of $G(\boldsymbol{\Gamma}_\pi)$. First observe that now $s \Vdash \gamma$ for all $\gamma \in \boldsymbol{\Gamma}_i$ for all $i \in \pi$. To appreciate this consider an arbitrary $i \in \pi$ and an equally arbitrary $\gamma \in \boldsymbol{\Gamma}_i$. In virtue of the opening remarks of this proof we may assume there to be a valuation s' such that $s' \Vdash \gamma$. Now consider the valuation (s_{-i}, s'_i). By definition $(s_{-i}, s'_i) \sim_i s'$ and with $\boldsymbol{A}(\gamma) \subseteq \pi_i$, we also have that $(s_{-i}, s'_i) \Vdash \gamma$. Moreover, with s being a maximum equilibrium, in particular $\big((s_{-i}, s'_i), s\big) \in \rho(\gamma)$. Hence, $s \Vdash \gamma$, and with i and γ having been chosen arbitrarily, also $s \Vdash \bigcup_{i\in\pi} \boldsymbol{\Gamma}_i$.

By the assumption, then, there is an $i \in \pi'$ as well as a $\vartheta \in \boldsymbol{\Theta}_i$ such that $s \Vdash \vartheta$. Since we could assume ϑ to be no tautology, there is some valuation s' such that $s' \nVdash \vartheta$. Now consider the valuation (s_{-i}, s'_i). Recall that we had assumed that $\boldsymbol{A}(\vartheta) \subseteq \pi_i$, and observing that $(s_{-i}, s'_i) \sim_i s'$ gives us $(s_{-i}, s'_i) \nVdash \vartheta$. Hence, $\big((s_{-i}, s'_i), s\big) \notin \bar{\boldsymbol{\rho}}(\vartheta)$. Eventually, we may conclude that s is *no* maximum equilibrium of the game $\overline{G}(\boldsymbol{\Theta}_{\pi'})$ and with s having been chosen as an arbitrary maximum equilibrium in $G(\boldsymbol{\Gamma}_\pi)$, we are done. $\dashv$

Observe that since the set $\boldsymbol{A}(\varphi)$ of propositional variables on which the interpretation of the formula φ depends is trivially a subset of the set of all propositional variables, Proposition 9.3.1 can also be obtained as a special case of Proposition 9.3.3.

Proposition 9.3.1, above, has as an immediate consequence that game-theoretical consequence is consistent in the sense that not every family of theories follows from any other. For one, since classically $\varnothing \nvdash^{\text{CPC}} \varnothing$, we have game-theoretically $\{\varnothing_A\} \nvDash \{\varnothing_A\}$. In a similar fashion Proposition 9.3.2 implies that each of the relations $\Lambda_{\pi,\pi'}$ is consistent as well.

Proposition 9.3.4 *(Consistency) For $L(A)$ a propositional language, let π and π' be partitions of A and let $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories in $L(A)$. Then $\boldsymbol{\Gamma}_\pi \nvDash \boldsymbol{\Theta}_{\pi'}$, if $\boldsymbol{\Gamma}_i = \varnothing$ and $\boldsymbol{\Theta}_i = \varnothing$, for each $i \in \pi$ and each $i \in \pi'$.*

Proof: Merely observe that $\bigcup_{i\in\pi}\boldsymbol{\Gamma}_i = \varnothing$ and $\bigcup_{i\in\pi'}\boldsymbol{\Theta}_i = \varnothing$ as well. Since classically $\varnothing \nvdash^{\mathrm{CPC}} \varnothing$, the claim follows immediately from Proposition 9.3.2 above. $\dashv$

Game-theoretical consequence as introduced in this chapter is based on the notion of maximum equilibrium rather than maximal equilibrium. In virtue of this choice game-theoretical consequence is monotonic. Adding more formulas to the theories constituting a family renders the preference relations in the corresponding games to be more refined. This observation together with Proposition 2.1.1 on page 28 secures monotonicity for game-theoretical consequence.

Proposition 9.3.5 *(Monotonicity) Let $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories in a language $L(A)$, indexed by the partitions π and π' of A. Let, furthermore, $\boldsymbol{\Gamma}'_\pi$ and $\boldsymbol{\Theta}'_\pi$ also be theories such that $\boldsymbol{\Gamma}_i \subseteq \boldsymbol{\Gamma}'_i$, for each $i \in \pi$, and $\boldsymbol{\Theta}_i \subseteq \boldsymbol{\Theta}'_i$, for each $i \in \pi'$. Then:*

$$\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'} \quad \textit{implies} \quad \boldsymbol{\Gamma}'_\pi \vDash \boldsymbol{\Theta}'_{\pi'}.$$

Proof: For each $i \in \pi$, $\boldsymbol{\Gamma}_i \subseteq \boldsymbol{\Gamma}'_i$, implies $\boldsymbol{\rho}\left(\boldsymbol{\Gamma}'_i\right) \subseteq \boldsymbol{\rho}\left(\boldsymbol{\Gamma}_i\right)$. Similarly, $\bar{\boldsymbol{\rho}}\left(\boldsymbol{\Theta}'_i\right) \subseteq \bar{\boldsymbol{\rho}}\left(\boldsymbol{\Theta}_i\right)$, for each $i \in \pi'$. By Proposition 2.1.1 on page 28, it follows that the maximum equilibria of $G\left(\boldsymbol{\Gamma}'_\pi\right)$ are included in those of $G(\boldsymbol{\Gamma}_\pi)$, and the maximum equilibria of $\overline{G}\left(\boldsymbol{\Theta}'_{\pi'}\right)$ in those of $\overline{G}(\boldsymbol{\Theta}_{\pi'})$. Now assume for contraposition that $\boldsymbol{\Gamma}'_\pi \nvDash \boldsymbol{\Theta}'_{\pi'}$. Then there is a strategy profile s that is a maximum equilibrium in both $G\left(\boldsymbol{\Gamma}'_\pi\right)$ and $G\left(\boldsymbol{\Theta}'_{\pi'}\right)$. Consequently, s is also a maximum equilibrium in both $G(\boldsymbol{\Gamma}_\pi)$ and $G(\boldsymbol{\Theta}_{\pi'})$ and so $\boldsymbol{\Gamma}_\pi \nvDash \boldsymbol{\Theta}_{\pi'}$. $\dashv$

As a consequence of the Propositions 9.3.3 and 9.3.5 we have that $\boldsymbol{\Theta}_{\pi'}$ follows game-theoretically from $\boldsymbol{\Gamma}_\pi$, if the family of theories $\boldsymbol{\Gamma}_\pi$ contains a contradiction or if the family of theories $\boldsymbol{\Theta}_{\pi'}$ contains a tautology. The following proposition captures this point.

Proposition 9.3.6 *Let π and π' be partitions of a set of propositional variables A and let $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories in $L(A)$. Then, $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}$, if either $\bigcup_{i\in\pi}\boldsymbol{\Gamma}_i$ contains a contradiction or $\bigcup_{i\in\pi'}\boldsymbol{\Theta}_i$ contains a tautology.*

Proof: First assume that for some $i \in \pi$, the theory $\boldsymbol{\Gamma}_i$ contains a contradiction γ^*. Then classically $\gamma^* \vdash^{\mathrm{CPC}} \varnothing$. Observe that $\boldsymbol{A}(\gamma^*) = \varnothing$, and so, trivially, $\boldsymbol{A}(\gamma^*) \subseteq \pi_i$. By Proposition 9.3.3, then $\boldsymbol{\Gamma}^*_\pi \vDash \boldsymbol{\Theta}^*_{\pi'}$, where $\boldsymbol{\Theta}^*_j = \varnothing$ for each $j \in \pi'$, $\boldsymbol{\Gamma}^*_j = \varnothing$ for each $j \in \pi$ distinct from i and $\boldsymbol{\Gamma}^*_i = \{\gamma^*\}$. By monotonicity of game-theoretical consequence (Proposition 9.3.5) then $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}$. If $\boldsymbol{\Theta}_i$ contains a tautology ϑ^* for some $i \in \pi'$, the argument runs along similar lines. $\dashv$

In stark contrast with these reassuring results, which point at important similarities between the classical and the game-theoretical notion of consequence, we find that diagonality of the consequence relation is no longer guaranteed. Diagonality is important property of classical consequence, *i.e.*, $\varphi \vdash \varphi$, for all formulas φ. However, for game-theoretical consequence it is not in general the case that $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Gamma}_\pi$, as the following example demonstrates.

	$\varnothing$	$\{b\}$
$\varnothing$	**0**, **0**	0, 0
$\{a\}$	0, 0	**1**, **1**

	$\varnothing$	$\{b\}$
$\varnothing$	**1**, **1**	**1**, **1**
$\{a\}$	**1**, **1**	0, 0

Figure 9.4. For $\boldsymbol{\Gamma}_\pi$ given by $\left\{\{a \wedge b\}_{\{a\}}, \{a \wedge b\}_{\{b\}}\right\}$, the matrix on the left represents the extensive game $G(\boldsymbol{\Gamma}_\pi)$ and the one to right the game $\overline{G}(\boldsymbol{\Gamma}_\pi)$.

Example 9.3.7 Again consider the classical propositional language containing a and b as only propositional variables. Let $\left\{\{a\}, \{b\}\right\}$ be the partition π and let $\boldsymbol{\Gamma}_\pi$ denote $\left\{\{a \wedge b\}_{\{a\}}, \{a \wedge b\}_{\{b\}}\right\}$. The matrices of the games $G(\boldsymbol{\Gamma}_\pi)$ and $\overline{G}(\boldsymbol{\Gamma}_\pi)$ are represented in Figure 9.4. Observe that the valuation $\varnothing$ is a maximum equilibrium in both $G(\boldsymbol{\Gamma}_\pi)$ and $\overline{G}(\boldsymbol{\Gamma}_\pi)$. Hence, $\left\{\{a \wedge b\}_{\{a\}}, \{a \wedge b\}_{\{b\}}\right\} \not\models \left\{\{a \wedge b\}_{\{a\}}, \{a \wedge b\}_{\{b\}}\right\}$.

However, it may still be the case that $\boldsymbol{\Gamma}_\pi \models \boldsymbol{\Gamma}_\pi$ for some partition π and family of theories $\boldsymbol{\Gamma}_\pi$. Similarly, for any partitions π and π' there may be formulas φ such that $\boldsymbol{\Gamma}_\pi \models \boldsymbol{\Theta}_{\pi'}$, if $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ are such that $\boldsymbol{\Gamma}_i = \boldsymbol{\Theta}_j = \{\varphi\}$ for all $i \in \pi$ and $j \in \pi'$. Whenever this is the case we say that that $\Lambda_{\pi,\pi'}$ is *diagonal with respect to* φ.

Some form of reflexivity can, of course, be rescued by imposing the trivializing restrictions on the families of theories of Proposition 9.3.3. This gives rise to the following proposition.

Proposition 9.3.8 *(Modified overlap) Let π and π' be partitions of a set of propositional variables A and let $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}'_{\pi'}$ be families of theories in $L(A)$. Assume that $\boldsymbol{A}(\gamma) \subseteq \pi_i$, for all $i \in \pi$ and all $\gamma \in \boldsymbol{\Gamma}_i$, and $\boldsymbol{A}(\gamma) \subseteq \pi_i$, for all $i \in \pi'$ and all $\gamma \in \boldsymbol{\Theta}_i$. Then, $\boldsymbol{\Gamma}_\pi \models \boldsymbol{\Theta}_{\pi'}$, if $\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i \cap \bigcup_{i\in\pi'} \boldsymbol{\Theta}_i \neq \varnothing$.*

Proof: Immediate consequence of the Propositions 9.3.2 and 9.3.3, above and the fact that *overlap* holds for classical propositional logic. $\dashv$

9.4 Control, Consequence, and Coalitions

A characteristic feature of game-theoretical consequence is the distribution of control over the propositional variables. The families of theories game-theoretical consequence connects may be indexed by different partitions of the propositional variables. A formal topic that suggests itself is how the set of game-theoretical consequences of a family of theories indexed by one partition relates to the set of game-theoretical consequences of a family of theories indexed by another partition. In order to assay this issue with some success, we need some understanding of the ways theories can systematically be combined into one theory and how a theory can be distributed over various theories. In Section 8.4 we have already touched upon this issue, addressing one particular way of combining theories, *viz.*, by simply taking their union.

Game-theoretical consequence brings within the scope of logic the notion of distributed control over the propositional variables. Classical consequence is the borderline case in which all control is concentrated in one player. A singleton collection of theories $\{\Gamma\}$ will invariably be indexed by the whole set of propositional variables. In general, however, a family of theories can be indexed by various partitions in a variety of ways. Moreover, a game-theoretical validity $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}$ depends both on the theories in $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ and on the way they are indexed by the partitions π and π'. An alternative indexing of the the same collection of theories by the same partition may have repercussions with respect to what follows game-theoretically from that collection of theories.

In the definition of classical consequence we distinguished a subrelation $\Lambda_{\pi,\pi'}$, for each pair of partitions π and π' of the propositional variables. Together these relations make up the set $\boldsymbol{\Lambda}_A$ and this section concerns its internal structure, *i.e.*, how the various relations $\Lambda_{\pi,\pi'}$ relate to one another.

Proposition 9.3.2 says that, in some some sense, the relation $\Lambda_{(\{A\},\{A\})}$ is stronger than any other relation $\Lambda_{(\pi,\pi')}$. Some caution is here in place as we have not made clear yet what 'in some sense' signifies in this context. Obviously, it is not in general the case that $\Lambda_{(\{A\},\{A\})}$ contains $\Lambda_{(\pi,\pi')}$ in the set-theoretical sense, as they relate families of theories indexed by different partitions. Still, every validity in $\Lambda_{(\pi,\pi')}$ has a counterpart in $\Lambda_{(\{A\},\{A\})}$ involving the same formulas. As the main result of this section, we find that this observation can be generalized as to hold between any two relations $\Lambda_{(\pi,\pi')}$ and $\Lambda_{(\pi'',\pi''')}$ such that $\pi \leqslant \pi''$ and $\pi' \leqslant \pi'''$. To this end we associate with each $\Lambda_{\pi,\pi'}$ in $\boldsymbol{\Lambda}_A$, the proper logic $\Lambda^*_{\pi,\pi'}$ — *i.e.*, a relation between theories of $L(A)$ — as follows:

$$\Lambda^*_{\pi,\pi'} \quad =_{df.} \quad \left\{ \left(\bigcup_{i\in\pi} \boldsymbol{\Gamma}_i, \bigcup_{j\in\pi'} \boldsymbol{\Theta}_j \right) : \; \boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'} \right\}.$$

Then define $\Lambda_{\pi,\pi'} \leqslant \Lambda_{\pi'',\pi'''}$ as $\Lambda^*_{\pi,\pi'} \subseteq \Lambda^*_{\pi'',\pi'''}$. Recall that the partitions over a set of propositional variables A possess a definite order-theoretic structure, as they are ordered as a complete lattice with respect to their coarseness as follows.

$$\pi \leqslant \pi' \quad \text{iff} \quad \text{for all } x \in \pi, \text{ there is a } y \in \pi' \text{ such that } x \subseteq y.$$

Intuitively, $\pi \leqslant \pi'$ denotes that π is at least as fine as π'. Since, ordered thus, the set of partitions $\mathit{Part}(A)$ over A constitutes a complete lattice, so does the direct product $\mathit{Part}(A) \times \mathit{Part}(A)$.

In section 8.4 we argued that each block in a partition represents the control over the propositional variables a coalition obtains some of players represented by the blocks of a finer partition decide to collaborate. If, moreover, the coalitional preference relation is given by the intersection of the preference relations of its members, then Proposition 2.1.8 on page 35 guarantees that there will be no increase of maximum equilibria as a result of this move. This phenomenon is the foundation of the following proposition about game-theoretical consequence.

Proposition 9.4.1 *Let π, π', π'' and π''' be partitions of some set of propositional variables A such that $\pi \leqslant \pi''$ and $\pi' \leqslant \pi'''$. Let further $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories in $L(A)$. Define the families of theories $\boldsymbol{\Gamma}^*_{\pi''}$ and $\boldsymbol{\Theta}^*_{\pi'''}$, such that for all $j \in \pi''$ and $k \in \pi'''$:*

$$\Gamma^*_j =_{df.} \bigcup_{\substack{i \in \pi \\ i \subseteq j}} \Gamma_i \qquad \text{and} \qquad \Theta^*_k =_{df.} \bigcup_{\substack{i \in \pi' \\ i \subseteq k}} \Theta_i .$$

Then, $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}$ implies $\boldsymbol{\Gamma}^*_{\pi''} \vDash \boldsymbol{\Theta}^*_{\pi'''}$.

Proof: Immediately from Proposition 2.1.8 on page 35 and Fact 8.4.6 on page 198. ⊣

In the previous section we found that game-theoretical consequence does not in general satisfy reflexivity; neither is each $\Lambda_{[,\pi,\pi']}$ diagonal with respect to all formulas. The following lemma and theorem show that nevertheless the latter property plays an important structural role in the anatomy of $\boldsymbol{\Lambda}_A$.

Lemma 9.4.2 *Let $L(A)$ be a propositional language and π, π', π'' and π''' be partitions of A. Then,*

$$\pi \leqslant \pi'' \text{ and } \pi' \leqslant \pi'''$$
$$\textit{iff}$$
$$\text{for each formula } \varphi\colon \big\{\{\varphi\}_i\big\}_{i\in\pi} \vDash \big\{\{\varphi\}_j\big\}_{j\in\pi'} \text{ implies } \big\{\{\varphi\}_i\big\}_{i\in\pi''} \vDash \big\{\{\varphi\}_j\big\}_{j\in\pi'''}.$$

Proof: The left-to-right direction follows from Proposition 9.4.1. Merely observe that $\bigcup_{\substack{i\in\pi \\ i\subseteq j}} \{\varphi\}_i = \{\varphi\}$ for each $j \in \pi'$ and, similarly, $\bigcup_{\substack{i\in\pi' \\ i\subseteq j}} \{\varphi\}_i = \{\varphi\}$, for each $j \in \pi'''$. The opposite direction is by contraposition. Assume that either $\pi \nleqslant \pi''$ or $\pi' \nleqslant \pi'''$. As the proof for both cases runs along analogous lines, we only give that of the former.

Let $\pi \nleqslant \pi''$. Then, π'' is different from the trivial partition $\{A\}$ and $A \neq \varnothing$. It also follows that there is some block π_0 of π for which there are two distinct blocks π''_0 and π''_1 of π'' such that both $\pi_0 \cap \pi''_0 \neq \varnothing$ and $\pi_0 \cap \pi''_1 \neq \varnothing$. So we may assume the existence of two propositional variables a_0 and a_1 such that $a_0 \in \pi_0 \cap \pi''_0$ and $a_1 \in \pi_0 \cap \pi''_1$. Consider these along with the formula $a_0 \wedge a_1$. First we prove the following two claims, of which $\big\{\{a_0 \wedge a_1\}_i\big\}_{i\in\pi} \vDash \big\{\{a_0 \wedge a_1\}_j\big\}_{j\in\pi'}$ is an immediate consequence:

$$s \text{ is a maximum equilibrium in } G\big(\big\{\{a_0 \wedge a_1\}_i\big\}_{i\in\pi}\big) \quad \text{iff} \quad s \in [\![a_0 \wedge a_1]\!]$$
$$s \text{ is a maximum equilibrium in } \overline{G}\big(\big\{\{a_0 \wedge a_1\}_j\big\}_{j\in\pi'}\big) \quad \text{iff} \quad s \notin [\![a_0 \wedge a_1]\!].$$

From right-to-left these claims hold in virtue of Proposition 8.4.5. For the opposite direction, first consider a valuation such that $s \notin [\![a_0 \wedge a_1]\!]$; we know that such an s

exists. Define $s^* =_{df.} s \cup \{a_0, a_1\}$. Then, obviously, $s^* \in [\![a_0 \wedge a_1]\!]$. Hence, $(s^*, s) \notin \boldsymbol{\rho}(\{a_0 \wedge a_1\})$. Consequently, $s^* \not\leqslant_i s$, for each player i of $G\big(\big\{\ \{a_0 \wedge a_1\}_i\ \big\}_{i \in \pi}\big)$. Moreover, with $a_0, a_1 \in \pi_i$, also $s^* = (s_{-i}, s_i^*)$ and s fails as a maximum response for any player i and, the set of players never being empty, also as a maximum equilibrium of $G\big(\big\{\ \{a_0 \wedge a_1\}_i\ \big\}_{i \in \pi}\big)$. This proves the first claim.

Now consider a valuation s such that $s \in [\![a_0 \wedge a_1]\!]$ as well as the unique $j \in \pi'$ such that $a_0 \in \pi'_j$. Define $s^{**} =_{df.} s - \{a_0\}$; then, $s^{**} \notin [\![a_0 \wedge a_1]\!]$. Moreover, $s^{**} = \big(s_{-j}, s_j^{**}\big)$, having assumed that $a_0 \in \pi'_j$. Hence, $(s^{**}, s) \notin \bar{\boldsymbol{\rho}}(\{a_0 \wedge a_1\})$ and $s^{**} \not\leqslant_j s$ for all players j of $\overline{G}\big(\big\{\ \{a_0 \wedge a_1\}_j\ \big\}_{j \in \pi'}\big)$ and in particular for j as above. Therefore, s is no maximum equilibrium in $\overline{G}\big(\big\{\ \{a_0 \wedge a_1\}_j\ \big\}_{j \in \pi'}\big)$ either.

An argument analogous to that for the second claim above shows that also:

$$s \text{ is a maximum equilibrium in } \overline{G}\big(\big\{\ \{a_0 \wedge a_1\}_j\ \big\}_{j \in \pi'''}\big) \quad \text{iff} \quad s \notin [\![a_0 \wedge a_1]\!].$$

Hence, in particular, the valuation $\varnothing$ is a maximum equilibrium in $\overline{G}\big(\big\{\ \{a_0 \wedge a_1\}_j\ \big\}_{j \in \pi'''}\big)$.

We complete the proof by showing that the valuation $\varnothing$ is also a maximum equilibrium in $G\big(\big\{\ \{a_0 \wedge a_1\}_i\ \big\}_{i \in \pi''}\big)$, for then it also holds that $\big\{\{a_0 \wedge a_1\}_i\big\}_{i \in \pi''} \nvDash \big\{\{a_0 \wedge a_1\}_j\big\}_{j \in \pi'''}$.

Consider an arbitrary player $k \in \pi''$. Then, not both $a_0 \in \pi''_k$ and $a_1 \in \pi''_k$. Now consider an arbitrary valuation s. Then not both $a_0 \in (\varnothing_{-k}, s_k)$ and $a_1 \in (\varnothing_{-k}, s_k)$. Hence, $(\varnothing_{-k}, s_k) \notin [\![a_0 \wedge a_1]\!]$. Because $\varnothing \notin [\![a_0 \wedge a_1]\!]$ it follows that $\big((\varnothing_{-k}, s_k), \varnothing\big) \in \boldsymbol{\rho}(\{a_0 \wedge a_1\})$. With s having been chosen arbitrarily, $\varnothing$ is a maximum response for k. And with k having been chosen arbitrarily as well, it follows that the valuation $\varnothing$ is a maximum equilibrium in $G\big(\big\{\ \{a_0 \wedge a_1\}_j\ \big\}_{j \in \pi''}\big)$. $\dashv$

We conclude this section with the following theorem, which, in effect, says that the ordering on the set $\boldsymbol{\Lambda}_A$ of subrelations of game-theoretical consequence for $L(A)$ can be derived from the complete lattice *Part* $(A) \times$ *Part* (A).

Theorem 9.4.3 *(Characterization of $\Lambda_{\pi,\pi'} \leqslant \Lambda_{\pi'',\pi'''}$) For $L(A)$ a propositional language and π, π', π'' and π''' partitions over A:*

$$\Lambda_{\pi,\pi'} \leqslant \Lambda_{\pi'',\pi'''} \quad \textit{iff} \quad \pi \leqslant \pi'' \ \textit{ and } \ \pi' \leqslant \pi'''.$$

Proof: For the right-to-left direction assume that $\pi \leqslant \pi''$ and $\pi' \leqslant \pi'''$ and consider an arbitrary $(\Gamma, \Theta) \in \Lambda^*_{\pi,\pi'}$. Then, $\Gamma = \bigcup_{i \in \pi} \boldsymbol{\Gamma}_i$ and $\Theta = \bigcup_{j \in \pi'} \boldsymbol{\Theta}_j$, for some families of theories $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ such that $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}$. Now let $\boldsymbol{\Gamma}^*_k =_{df.} \bigcup_{\substack{i \in \pi \\ i \subseteq k}} \boldsymbol{\Gamma}_i$, for each $k \in \pi''$, and $\boldsymbol{\Theta}^*_k =_{df.} \bigcup_{\substack{j \in \pi' \\ j \subseteq k}} \boldsymbol{\Theta}_i$, for each $k \in \pi'''$. By the assumption that $\pi \leqslant \pi''$ and $\pi' \leqslant \pi'''$, then obviously, $\bigcup_{i \in \pi} \boldsymbol{\Gamma}_i = \bigcup_{k \in \pi''} \boldsymbol{\Gamma}^*_k$ and $\bigcup_{j \in \pi'} \boldsymbol{\Theta}_i = \bigcup_{k \in \pi'''} \boldsymbol{\Theta}^*_k$. Accordingly, $\bigcup_{k \in \pi''} \boldsymbol{\Gamma}^*_k = \Gamma$ and $\bigcup_{k \in \pi'''} \boldsymbol{\Theta}^*_k = \Theta$. Moreover, in virtue of Proposition 9.4.1, also $\boldsymbol{\Gamma}^*_{\pi''} \vDash \boldsymbol{\Theta}^*_{\pi'''}$. Hence, $(\Gamma, \Theta) \in \Lambda^*_{\pi'',\pi'''}$.

The opposite direction follows from Proposition 9.3.5 and Lemma 9.4.2. Assume, for contraposition, that either $\pi \nleqslant \pi''$ or $\pi' \nleqslant \pi'''$. By Lemma 9.4.2, there is a formula φ such that $\{\{\varphi\}_i\}_{i\in\pi} \vDash \{\{\varphi\}_i\}_{i\in\pi'}$ but $\{\{\varphi\}_i\}_{i\in\pi''} \nvDash \{\{\varphi\}_i\}_{i\in\pi'''}$. Then, $(\{\varphi\}, \{\varphi\}) \in \Lambda^*_{\pi,\pi'}$. Now consider arbitrary families of theories $\boldsymbol{\Gamma}^*_{\pi''}$ and $\boldsymbol{\Theta}^*_{\pi'''}$ such that $\bigcup_{i\in\pi''} \boldsymbol{\Gamma}^*_i = \bigcup_{i\in\pi} \{\varphi\}_i = \{\varphi\}$ and $\bigcup_{i\in\pi'''} \boldsymbol{\Theta}^*_i = \bigcup_{i\in\pi'} \{\varphi\}_i = \{\varphi\}$. Then, $\boldsymbol{\Gamma}^*_i \subseteq \{\varphi\}$, for each $i \in \pi''$, and $\boldsymbol{\Theta}^*_i \subseteq \{\varphi\}$ each $i \in \pi'''$. By Proposition 9.3.5, stating the monotonicity of game-theoretical consequence, it then follows that $\boldsymbol{\Gamma}^*_{\pi''} \nvDash \boldsymbol{\Theta}^*_{\pi'''}$. Therefore, we may conclude the proof observing that $(\{\varphi\}, \{\varphi\}) \in \Lambda^*_{\pi,\pi'}$ but $(\{\varphi\}, \{\varphi\}) \notin \Lambda^*_{\pi'',\pi'''}$. $\dashv$

As an immediate consequence, the following corollary also holds. It says that $\Lambda_{\pi,\pi'} \leqslant \Lambda_{\pi'',\pi'''}$ if and only if $\Lambda_{\pi'',\pi'''}$ preserves diagonality for all formulas φ for which $\Lambda_{\pi,\pi'}$ is diagonal.

Corollary 9.4.4 *Let $L(A)$ be a propositional language and π, π', π'' and π''' be partitions of A. Then, $\Lambda_{(\pi,\pi')} \leqslant \Lambda_{(\pi'',\pi''')}$ if and only if $\Lambda_{(\pi'',\pi''')}$ is diagonal with respect to all formulas with respect to which $\Lambda_{(\pi,\pi')}$ is diagonal as well.*

Proof: Immediately from Lemma 9.4.2 and Theorem 9.4.3. $\dashv$

9.5 Rough Set Characterization

In this section we show how game-theoretical consequence can be given a set-theoretical semantics. Employing the machinery of rough set theory, it is demonstrated how each family of theories $\boldsymbol{\Gamma}_\pi$ can be associated with a set of sets of valuations that precisely coincides with the maximum equilibria of the distributed evaluation game $G(\boldsymbol{\Gamma}_\pi)$. The significance of this result is that it forges a strong link with classical logic, facilitating the logical analysis of game-theoretical consequence.

The intuition behind this essentially technical result derives from the understanding of a maximum equilibrium as the intersection of the players' maximum responses. Each player has control over a subset of propositional variables and his opponents over the remaining ones. Each set of propositional variables partitions the valuations and this is in particular the case for the propositional variables controlled by the opponents of a particular player. *I.e.*, let π be partition of the propositional variables and let π_{-i} be the set of variables controlled by the opponents of player *i*. With this set π_{-i} is associated the equivalence relation $\varepsilon_{\pi_{-i}}$, which in turn induces the partition $\pi_{\pi_{-i}}$ over the valuations. A maximum response for player *i* is then a valuation that is a maximum with respect to his preferences within each block of this partition.

Let now the preference of player *i* be given by a relation $\boldsymbol{\rho}(\boldsymbol{X})$, where $\boldsymbol{X}$ is a set of subsets of valuations. We show that then the approximation operators $\underline{apr}_\Delta$ and $\overline{apr}_\Delta$ on sets of valuations — where Δ is a subset of propositional variables — are available to single out player *i*'s maximum responses on the basis of the set $\boldsymbol{X}$ alone. In particular, we invoke the approximation operations with respect to the empty set and with respect to the set of propositional variables controlled by the opponents of a player *i*.

Formally, we have the following definition.

Definition 9.5.1 Let π be a partition of the propositional variables of some propositional language $L(A)$. Let further $\boldsymbol{X}$ be a family $\{\boldsymbol{X}_i\}_{i\in\pi}$ of subsets of 2^{2^A}, *i.e.*, a family of sets of subsets of valuations. Define for each $i \in \pi$:

$$\begin{aligned}
\mathsf{N}(\boldsymbol{X}_i) &=_{df.} \bigcap_{X\in\boldsymbol{X}_i} \left(\overline{apr}_{\varnothing}(X) \cap \left(X \cup \underline{apr}_{\pi_{-i}}(\overline{X})\right)\right),\\
\mathsf{N}(\boldsymbol{X}) &=_{df.} \bigcap_{i\in\pi} \mathsf{N}(\boldsymbol{X}_i)\,.
\end{aligned}$$

We introduce dual concepts of these notions as follows:

$$\begin{aligned}
\mathsf{И}(\boldsymbol{X}_i) &=_{df.} \overline{\mathsf{N}\left(\{\overline{X} :\ X \in \boldsymbol{X}_i\}\right)}\\
\mathsf{И}(\boldsymbol{X}) &=_{df.} \bigcup_{i\in\pi} \mathsf{И}(\boldsymbol{X}_i)
\end{aligned}$$

Let $\boldsymbol{\Gamma}_\pi$ a family of theories in a language $L(A)$ indexed by a partition π of A. For each $i \in \pi$, we write $\mathsf{N}(\Gamma)$ for $\mathsf{N}(\{[\![\gamma]\!] :\ \gamma \in \Gamma\})$ and we have $\mathsf{N}(\boldsymbol{\Gamma}_\pi)$ denote $\bigcap_{i\in\pi} \mathsf{N}(\Gamma_i)$. Similarly, $\mathsf{И}(\Gamma)$ abbreviates $\mathsf{И}(\{[\![\gamma]\!] :\ \gamma \in \Gamma\})$ and $\mathsf{И}(\boldsymbol{\Gamma}_\pi)$ represents $\bigcup_{i\in\pi} \mathsf{И}(\Gamma_i)$.

The intuition underlying these definitions is as follows. Assuming that the preferences of the player i are given by $\boldsymbol{\rho}(\boldsymbol{X}_i)$, the set $\mathsf{N}(\boldsymbol{X}_i)$ is meant to collect the maximum elements of the relation $\boldsymbol{\rho}(\boldsymbol{X}_i)$ *within each block of the partition* $\pi_{\pi_{-i}}$. The relation $\boldsymbol{\rho}(\boldsymbol{X}_i)$ is given by $\bigcap_{X\in\boldsymbol{X}_i} \rho(X)$, *i.e.*, for each $X \in \boldsymbol{X}$, the relation $\rho(X)$ determines in part the relation $\boldsymbol{\rho}(\boldsymbol{X})$. In case $\boldsymbol{X}$ contains the empty set, the relation $\boldsymbol{\rho}(\boldsymbol{X})$ is empty as well (*cf.*, Fact 8.5.3 on page 200). Then the set of maximum responses of player i is empty as well. Observe that $\overline{apr}_{\varnothing}(X)$ is empty if X is empty, and the whole set of valuations, otherwise. As such, $\overline{apr}_{\varnothing}(X)$ constitutes, as it were, a test for X being non-empty. Thus we have:

$$\overline{apr}_{\varnothing}(X) \cap \left(X \cup \underline{apr}_{\pi_{-i}}(\overline{X})\right) \quad = \quad \begin{cases} X \cup \underline{apr}_{\pi_{-i}}(\overline{X}) & \text{if } X \neq \varnothing\\ \varnothing & \text{otherwise.}\end{cases}$$

Accordingly, in case $\boldsymbol{X}_i$ contains the empty set, the set of maximum responses for player i is empty as well. It is, however, equally true that in that case for some $X \in \boldsymbol{X}_i$, the set $\overline{apr}_{\varnothing}(X) \cap \left(X \cup \underline{apr}_{\pi_{-i}}(\overline{X})\right)$ is empty, and with that so is $\mathsf{N}(\boldsymbol{X}_i)$.

Now consider the case in which $\boldsymbol{X}$ does not contain the empty set. For each $X \in \boldsymbol{X}_i$ the relation $\rho(X)$ distinguishes the valuations contained in X from those that are not, intuitively ranking the former higher than the latter. Construed as part of the preference relation $\boldsymbol{\rho}(\boldsymbol{X}_i)$ of player i, the relation $\rho(X)$ excludes as maximum responses for i those valuations that are inferior in this sense *within a block of the partition* $\pi_{\pi_{-i}}$, a block of valuations in which all values for the propositional variables are fixed but those for those controlled by i. *I.e.*, if X has a non-empty intersection with a block Y of $\pi_{\pi_{-i}}$

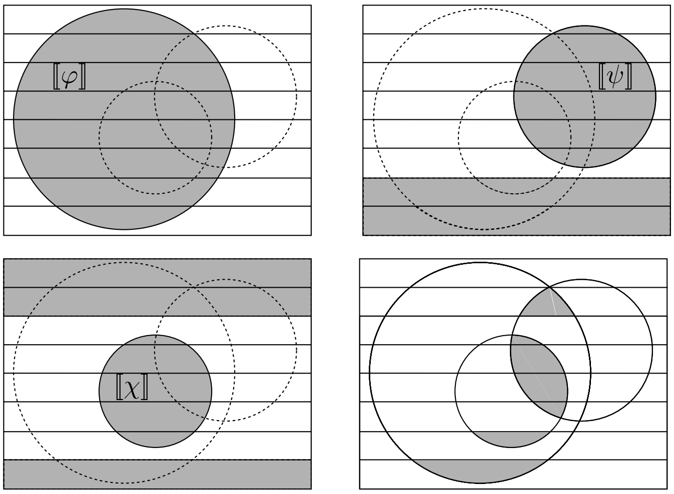


Figure 9.5. Let the preferences of a player i be given by three sets, say by $[\![\varphi]\!]$, $[\![\psi]\!]$ and $[\![\chi]\!]$ as in Figure 8.1, on page 182. Each block represents a particular choice of strategy by i's opponents. The upper left, upper right and lower left figure then represent $\overline{apr}_{\emptyset}([\![\varphi]\!]) \cap \left([\![\varphi]\!] \cup \underline{apr}_{\pi_{-i}}(\overline{[\![\varphi]\!]})\right)$, $\overline{apr}_{\emptyset}([\![\psi]\!]) \cap \left([\![\psi]\!] \cup \underline{apr}_{\pi_{-i}}(\overline{[\![\psi]\!]})\right)$ and $\overline{apr}_{\emptyset}([\![\chi]\!]) \cap \left([\![\chi]\!] \cup \underline{apr}_{\pi_{-i}}(\overline{[\![\chi]\!]})\right)$, respectively. The bottom right figure depicts $\mathsf{N}(\{\varphi, \psi, \chi\})$, *i.e.*, the maximum responses of player i. Also compare Figure 8.2, on page 187.

then all valuations outside $X \cap Y$ are disqualified as a maximum response for i. If a block Y of $\pi_{\pi_{-i}}$ and X are disjoint, however, $\rho\,(X)$ does not exclude any valuations in Y as a maximum response for player i. This is precisely what $X \cup \underline{apr}_{\pi_{-i}}\left(\overline{X}\right)$ achieves for non-empty subsets X of valuations. Doing this for each $X \in \boldsymbol{X}_i$ and intersecting the results delivers precisely the maximum responses for i. Also compare Figure 9.5 for further illustration of this point and consider the following proposition for a formal proof.

Proposition 9.5.2 establishes $\mathsf{N}(\boldsymbol{X}_i)$ as precisely the set of maximum responses for player i in a game $\left(\pi, \{2^{A_i}\}_{i\in\pi}, \{\boldsymbol{\rho}\,(\boldsymbol{X}_i)\}_{i\in\pi}\right)$. By taking the intersection of the maximum responses of all players, the set of maximum equilibria is obtained. In this manner we arrive at the characterization of maximum equilibria in rough set theory we were after.

Proposition 9.5.2 *Let A be a set, π a partition of A and $\boldsymbol{X}$ a family $\{\boldsymbol{X}_i\}_{i\in\pi}$ of subsets of 2^{2^A}. Let the game $\left(\pi, \{2^{A_i}\}_{i\in\pi}, \{\boldsymbol{\rho}\,(\boldsymbol{X}_i)\}_{i\in\pi}\right)$ be denoted by G. Then:*

$\mathsf{N}(\boldsymbol{X}_i)$ *is the set of maximum responses for i in $G(\boldsymbol{X})$.*

Proof: First assume $\boldsymbol{X}_i$ contain the empty set. Then, $\boldsymbol{\rho}\,(\boldsymbol{X}_i) = \emptyset$, by Fact 8.5.2 on page 200. Accordingly the set of maximum responses for i is empty as well. Also, $\overline{apr}_{\varnothing}(\varnothing) = \varnothing$ and hence $\overline{apr}_{\varnothing}(\varnothing) \cap \big(\varnothing \cup \underline{apr}_{\pi_{-i}}(\overline{\varnothing})\big) = \varnothing$. It follows that $\mathsf{N}(\boldsymbol{X}_i) = \varnothing$ as well. So, for the remainder of the proof we may assume that $\boldsymbol{X}_i$ does not contain the empty set.

Assume for some valuation s that $s \notin \mathsf{N}(\boldsymbol{X}_i)$. Then, $s \notin \overline{apr}_{\varnothing}(X) \cap \big(X \cup \underline{apr}_{\pi_{-i}}(\overline{X})\big)$, for some $X \in \boldsymbol{X}_i$. Having assumed that X is not empty, $\overline{apr}_{\varnothing}(X) = 2^A$, and it follows that both $s \notin X$ and $s \notin \underline{apr}_{\pi_{-i}}(\overline{X})$. Then, $s \in \overline{apr}_{\pi_{-i}}(X)$ and hence there exists some $s' \in X$ such that $s \sim_{\pi_{-i}} s'$. It follows that $(s', s) \notin \rho(X)$ and, subsequently, that $(s', s) \notin \boldsymbol{\rho}(\boldsymbol{X}_i)$. Observing that $s' = (s_{-i}, s'_i)$, we may conclude that s contains no maximum response for i in G.

For the opposite direction assume that some s in 2^A be no maximum response for i in G. Hence, there is some s' in 2^A such that $\big((s_{-i}, s'_i), s\big) \notin \boldsymbol{\rho}(\boldsymbol{X}_i)$. Accordingly, for some $X \in \boldsymbol{X}_i$ we have $(s_{-i}, s'_i) \in X$ but $s \notin X$. Since, $(s_{-i}, s'_i) \sim_{\pi_{-i}} s$, also $s \in \overline{apr}_{\pi_{-i}}(X)$ and so $s \notin \underline{apr}_{\pi_{-i}}(\overline{X})$. Hence $s \notin \overline{apr}_{\varnothing}(X) \cap \big(X \cup \underline{apr}_{\pi_{-i}}(\overline{X})\big)$ and so $s \notin \mathsf{N}(X)$. We may conclude that $s \notin \mathsf{N}(\boldsymbol{X}_i)$, and we are done. ⊣

This proposition has as an immediate consequence that $\mathsf{N}(\boldsymbol{X})$ precisely includes the maximum equilibria of the game $\big(\pi, \{2^{A_i}\}_{i\in\pi}, \{\boldsymbol{\rho}(\boldsymbol{X}_i)\}_{i\in\pi}\big)$.

Corollary 9.5.3 *Let A be a set, π a partition of A and $\boldsymbol{X}$ a family $\{\boldsymbol{X}_i\}_{i\in\pi}$ of subsets of 2^{2^A}. Let the game $\big(\pi, \{2^{A_i}\}_{i\in\pi}, \{\boldsymbol{\rho}(\boldsymbol{X}_i)\}_{i\in\pi}\big)$ be denoted by G. Then:*

$$\mathsf{N}(\boldsymbol{X}) \text{ is the set of maximum equilibria in } \big(\pi, \{2^{A_i}\}_{i\in\pi}, \{\boldsymbol{\rho}(\boldsymbol{X}_i)\}_{i\in\pi}\big).$$

Proof: Immediately from Proposition 9.5.2 and the definition of a maximum equilibrium as the intersection of the players maximum responses. ⊣

By straightforwardly applying the definitions we can also derive the corresponding clauses for the dual concepts $\mathsf{И}(\boldsymbol{X}_i)$ and $\mathsf{И}(\boldsymbol{X})$. Let, furthermore, the game $\overline{G}$ be given by $\big(\pi, \{2^{A_i}\}_{i\in\pi}, \{\bar{\boldsymbol{\rho}}(\boldsymbol{X}_i)\}_{i\in\pi}\big)$. Then:

$\mathsf{И}(\boldsymbol{X}_i)$ is the set of strategy profiles *not* containing a best response for i in $\overline{G}(\boldsymbol{X})$,

$\mathsf{И}(\boldsymbol{X})$ is the set of strategy profiles that are *no* maximum equilibrium in $\overline{G}(\boldsymbol{X})$.

These observations prepare the ground for the following theorem.

Theorem 9.5.4 *Let π and π' be partitions of A. Let further $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be theories in the language $L(A)$. Then:*

$$\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'} \quad \textit{iff} \quad \mathsf{N}(\boldsymbol{\Gamma}_\pi) \subseteq \mathsf{И}(\boldsymbol{\Theta}_{\pi'})\,.$$

Proof: Immediately from Proposition 9.5.3 and the subsequent remarks in the text. ⊣

Theorem 9.5.4 has an important consequence concerning the relation between classical propositional logic and game-theoretical consequence. On page 55 we found that each approximation by a subset of propositional variables of the extension of a formula in a language $L(A)$, is also the extension of a formula in $L(A)$. *I.e.*, propositional logic has expressive power with respect to approximations of extensions of formulas. This makes that for each family of theories $\boldsymbol{\Gamma}_\pi$, a theory Γ^* in $L(A)$ can be found such that the extension $[\![\Gamma^*]\!]$ and $\mathsf{N}(\boldsymbol{\Gamma}_\pi)$ are identical sets. Similarly, it is possible to find for each family of theories $\boldsymbol{\Theta}_{\pi'}$, a theory Θ^* such that $\langle\!\langle \Theta^* \rangle\!\rangle = \mathsf{И}(\boldsymbol{\Theta}_{\pi'})$.

To prove this, we define for each subset $X \subseteq A$ two injective functions f_X and g_X mapping formulas of $L(A)$ onto formulas of $L(A)$.

$$f_X(\varphi) \quad =_{df.} \quad \langle \emptyset \rangle \varphi \wedge \left(\langle \overline{X} \rangle \varphi \rightarrow \varphi \right),$$
$$g_X(\varphi) \quad =_{df.} \quad \neg f_X(\neg\varphi).$$

Observe that, for φ a formula, $f_X(\varphi)$ and $g_X(\varphi)$ are one another's duals. For any theory Γ and any $X \subseteq A$, let $f_X(\Gamma)$ and $g_X(\Gamma)$ stand for, respectively, $\{f_X(\gamma) : \gamma \in \Gamma\}$ and $\{g_X(\gamma) : \gamma \in \Gamma\}$. For any family of theories $\boldsymbol{\Gamma}_\pi$ indexed by a partition π of A, we have $f(\boldsymbol{\Gamma}_\pi)$ and $g(\boldsymbol{\Gamma}_\pi)$ denote the theories $\bigcup_{i\in\pi} f_{\pi_i}(\boldsymbol{\Gamma}_i)$ and $\bigcup_{i\in\pi} g_{\pi_i}(\boldsymbol{\Gamma}_i)$, respectively. We have the following fact.

Fact 9.5.5 *Let φ be a formula of a propositional language $L(A)$ let all X and Y subsets of A. Then:*

$$[\![g_X(\varphi)]\!] \quad \subseteq \quad [\![\varphi]\!] \quad \subseteq \quad [\![f_Y(\varphi)]\!].$$

Proof: Since $[\![\langle\emptyset\rangle \varphi]\!] = \overline{apr}_\emptyset([\![\varphi]\!])$, clearly $[\![\varphi]\!] \subseteq [\![\langle\emptyset\rangle \varphi]\!]$ because of the rough set law that $Z \subseteq \overline{apr}(Z)$. Obviously also $[\![\varphi]\!] \subseteq [\![\langle\overline{Y}\rangle \varphi \rightarrow \varphi]\!]$. Hence, $[\![\varphi]\!] \subseteq [\![f_Y(\varphi)]\!]$. For the inclusion of $[\![g_X(\varphi)]\!]$ in $[\![\varphi]\!]$, observe that $g_X(\varphi)$ is equivalent to $[\emptyset]\varphi \vee \left(\varphi \wedge \neg \left[\overline{X}\right] \varphi\right)$. Because in general $\underline{apr}(Z) \subseteq Z$, for any set Z, also $\overline{apr}_\emptyset([\![\varphi]\!]) \subseteq [\![\varphi]\!]$ and *a fortiori* $[\![[\emptyset]\varphi]\!] \subseteq [\![\varphi]\!]$. Clearly also $[\![\varphi \wedge \neg \left[\overline{X}\right] \varphi]\!] \subseteq [\![\varphi]\!]$. It follows that $[\![g_X(\varphi)]\!] \subseteq [\![\varphi]\!]$. ⊣

We then find that each valid instance of game-theoretical consequence has a counterpart in classical propositional logic (CPC).

Proposition 9.5.6 *Let π and π' be partitions of some set A of propositional variables. Let further $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories. Then:*

$$\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'} \quad \textit{iff} \quad f(\boldsymbol{\Gamma}_\pi) \vdash^{\mathrm{CPC}} g(\boldsymbol{\Theta}_{\pi'})$$

Proof: Observe that in general $[\![\langle\emptyset\rangle \varphi \wedge \left(\langle\overline{X}\rangle \varphi \rightarrow \varphi\right)]\!] = \overline{apr}_\emptyset([\![\varphi]\!]) \cap \left([\![\varphi]\!] \cup \underline{apr}_{\overline{X}}\left(\overline{[\![\varphi]\!]}\right)\right)$. Then the following equalities hold:

$$\left[\!\!\left[\bigcup_{i\in\pi}\{f_i(\gamma) : \gamma \in \boldsymbol{\Gamma}_i\}\right]\!\!\right] \;=\; \bigcap_{i\in\pi}\bigcap_{\gamma\in\boldsymbol{\Gamma}_i} [\![\langle\emptyset\rangle \varphi \wedge \left(\langle\overline{X}\rangle \varphi \rightarrow \varphi\right)]\!] \;=$$
$$\bigcap_{i\in\pi}\bigcap_{\gamma\in\boldsymbol{\Gamma}_i} \left(\overline{apr}_\emptyset([\![\gamma]\!]) \cap \left([\![\gamma]\!] \cup \underline{apr}_{-i}\left(\overline{[\![\gamma]\!]}\right)\right)\right) \;=\; \mathsf{N}(\boldsymbol{\Gamma}_\pi).$$

Using this result, the following equalities hold as well:

$$\left\langle\!\!\left\langle \bigcup_{i\in\pi'} \{g_i(\vartheta) : \vartheta \in \boldsymbol{\Theta}_i\} \right\rangle\!\!\right\rangle = \bigcup_{i\in\pi'} \bigcup_{\vartheta\in\boldsymbol{\Theta}_i} \overline{[\![f_i(\neg\vartheta)]\!]} =$$

$$\bigcup_{i\in\pi'} \overline{\bigcap_{\vartheta\in\boldsymbol{\Theta}_i} [\![f_i(\neg\vartheta)]\!]} = \bigcup_{i\in\pi'} \overline{\mathsf{N}\big(\{\overline{[\![\vartheta]\!]} : \vartheta \in \boldsymbol{\Theta}_i\}\big)} = \mathsf{И}(\boldsymbol{\Theta}_{\pi'}) .$$

Hence, we have the following equivalences:

$$\begin{array}{lll} \boldsymbol{\Gamma}_\pi \models \boldsymbol{\Theta}_{\pi'} & \text{iff}_{\text{Prop. 9.5.4}} & \mathsf{N}(\boldsymbol{\Gamma}_\pi) \subseteq \mathsf{И}(\boldsymbol{\Theta}_{\pi'}) \\ & \text{iff} & \left[\!\!\left[\bigcup_{i\in\pi} \{f_i(\gamma) : \gamma \in \boldsymbol{\Gamma}_i\} \right]\!\!\right] \subseteq \left\langle\!\!\left\langle \bigcup_{i\in\pi'} \{g_i(\vartheta) : \vartheta \in \boldsymbol{\Theta}_i\} \right\rangle\!\!\right\rangle \\ & \text{iff} & \bigcup_{i\in\pi} \{f_i(\gamma) : \gamma \in \boldsymbol{\Gamma}_i\} \vdash^{\text{CPC}} \bigcup_{i\in\pi'} \{g_i(\vartheta) : \vartheta \in \boldsymbol{\Theta}_i\} \\ & \text{iff} & f(\boldsymbol{\Gamma}_\pi) \vdash^{\text{CPC}} g(\boldsymbol{\Theta}_{\pi'}) \end{array}$$

This concludes the proof. ⊣

Proposition 9.5.6 has many corollaries, as it enables one to extrapolate results from classical propositional logic to game-theoretical consequence. Perhaps the most important is that game-theoretical consequence is decidable.

Corollary 9.5.7 *Let $L(A)$ be propositional language with A finite and let π and π' be partitions of A. Let further $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be finite families of theories of $L(A)$. Then, the problem whether $\boldsymbol{\Gamma}_\pi \models \boldsymbol{\Theta}_{\pi'}$ is decidable.*

Sketch of proof: Immediately from Proposition 9.5.6 and the decidability of CPC. For each $i \in \pi$ and each $j \in \pi'$, the functions f_{π_i} and $g_{\pi'_j}$ make that the problem $\boldsymbol{\Gamma}_\pi \models \boldsymbol{\Theta}_{\pi'}$ can be translated into the equivalent problem $f(\boldsymbol{\Gamma}_\pi) \vdash^{\text{CPC}} g(\boldsymbol{\Theta}_{\pi'})$, which we know is decidable. Observe in this respect that, since the functions f_{π_i} and $g_{\pi'_j}$ map formulas on formulas, the theories $f(\boldsymbol{\Gamma}_\pi)$ and $g(\boldsymbol{\Theta}_{\pi'})$ are finite. So it suffices to prove that the functions f_{π_i} and $g_{\pi'_j}$ are decidable. That this is indeed the case is revealed by some reflection on their definition and the fact that $[\Delta]\,\varphi$ and $\langle\Delta\rangle\,\varphi$ may be taken to abbreviate the formulas $\bigwedge_{\sigma\in\Sigma_{\overline{\Delta}}} \sigma(\varphi)$ and $\bigvee_{\sigma\in\Sigma_{\overline{\Delta}}} \sigma(\varphi)$, respectively (*cf.*, page 2.4) and it, moveover, being given that A is finite. ⊣

An important property of classical propositional logic that has so far been conspicuously absent from our analysis of game-theoretical consequence is that it satisfies *cut* (*cf.*, page 45, above). The principle of cut, however, needs some slight rephrasing before it sensibly said to hold for game-theoretical consequence. To illustrate this point consider the version of cut in which $\Gamma \cup \{\varphi\} \vdash \Theta$ and $\Gamma' \vdash \Theta' \cup \{\varphi\}$ imply $\Gamma \cup \Gamma' \vdash \Theta \cup \Theta'$. Cut, as it were, sets conditions for the 'removal' of a formula φ from theories if it occurs on the left of the turnstile in one validity statements and on the right in another. With game-theoretical consequence, however, it is families of theories rather than theories that flank the turnstile. For the game-theoretical consequence relation, we find that a formula may be removed in a similar fashion if it occurs in a theory in the family to the left of the turnstile in a validity statement and in a formula in another theory in the family to the right of the turnstile in another validity statement.

Proposition 9.5.8 *(Cut) Let φ be a formula in $L(A)$ and let π and π' be partitions of a set of propositional variables A. Let further Γ, Γ', Θ and Θ' be theories and $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}'$ be families of theories indexed by π an d $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}'$ be families of theories indexed by π'. Let $\boldsymbol{\Gamma}''$ and $\boldsymbol{\Theta}''$ be families of theories indexed by π and π' respectively such that $\boldsymbol{\Gamma}''_i = \boldsymbol{\Gamma}_i \cup \boldsymbol{\Gamma}'_i$, for each $i \in \pi$, and $\boldsymbol{\Theta}''_j = \boldsymbol{\Theta}_j \cup \boldsymbol{\Theta}'_j$, for each $j \in \pi'$. Then:*

$$(\boldsymbol{\Gamma}_{-i}, \Gamma_i)_\pi \vDash \left(\boldsymbol{\Theta}_{-j}, (\Theta \cup \{\varphi\})_j\right)_{\pi'} \quad \textit{and} \quad \left(\boldsymbol{\Gamma}'_{-i}, (\Gamma' \cup \{\varphi\})_i\right)_\pi \vDash \left(\boldsymbol{\Theta}'_{-j}, \Theta'_j\right)_{\pi'}$$

$$\textit{imply}$$

$$\left(\boldsymbol{\Gamma}''_{-i}, (\Gamma \cup \Gamma')_i\right)_\pi \vDash \left(\boldsymbol{\Theta}''_{-j}, (\Theta \cup \Theta')_j\right)_{\pi'}.$$

Proof: Assume that $(\boldsymbol{\Gamma}_{-i}, \Gamma_i)_\pi \vDash \left(\boldsymbol{\Theta}_{-j}, (\Theta \cup \{\varphi\})_j\right)_{\pi'}$ along with $\left(\boldsymbol{\Gamma}'_{-i}, (\Gamma' \cup \{\varphi\})_i\right)_\pi \vDash \left(\boldsymbol{\Theta}'_{-j}, \Theta'_j\right)_{\pi'}$. In virtue of Proposition 9.5.6 then both:

$$f\left((\boldsymbol{\Gamma}_{-i}, \Gamma_i)_\pi\right) \vdash^{\text{CPC}} g\left((\boldsymbol{\Theta}_{-j}, \Theta_j)_{\pi'}\right) \cup \left\{g_j(\varphi)\right\}$$
$$f\left(\left(\boldsymbol{\Gamma}'_{-i}, \Gamma'_i\right)_\pi\right) \cup \left\{f_i(\varphi)\right\} \vdash^{\text{CPC}} g\left(\left(\boldsymbol{\Theta}'_{-j}, \Theta'_j\right)_{\pi'}\right).$$

In virtue of Fact 9.5.5, also $g_j(\varphi) \vdash^{\text{CPC}} f_i(\varphi)$. Proposition 2.3.1 on page 46 then yields:

$$f\left((\boldsymbol{\Gamma}_{-i}, \Gamma_i)_\pi\right) \cup f\left(\left(\boldsymbol{\Gamma}'_{-i}, \Gamma'_i\right)_\pi\right) \vdash^{\text{CPC}} g\left((\boldsymbol{\Theta}_{-j}, \Theta_j)_{\pi'}\right) \cup g\left(\left(\boldsymbol{\Theta}'_{-j}, \Theta'_j\right)_{\pi'}\right).$$

We obtain $\left(\boldsymbol{\Gamma}''_{-i}, \Gamma \cup \Gamma'_i\right)_\pi \vDash \left(\boldsymbol{\Theta}''_{-j}, \Theta \cup \Theta'_j\right)_{\pi'}$ through another application of Proposition 9.5.6, which concludes the proof. ⊣

Proposition 9.5.6 established that every problem of game-theoretical consequence has an equivalent problem in classical propositional logic. The converse of this claim is of course also trivially true because of the congruence of $\Lambda_{\{A\},\{A\}}$ and CPC, witness Proposition 9.3.1. We conclude this section with a stronger result, stating that every problem of classical propositional logic has its counterpart in $\Lambda_{\pi,\pi'}$ in $\boldsymbol{\Lambda}_A$. The following lemma gives a preliminary result.

Lemma 9.5.9 *Let φ be a formula in $L(A)$ and π a partition of A. Then there are families of theories $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_\pi$, with $\boldsymbol{\Gamma}_i$ and $\boldsymbol{\Theta}_i$ finite for each $i \in \pi$, such that:*

$$[\![\varphi]\!] \;=\; \left[\!\left[f(\boldsymbol{\Gamma}_\pi) \right]\!\right] \;=\; \left\langle\!\left\langle g(\boldsymbol{\Theta}_\pi) \right\rangle\!\right\rangle.$$

Proof: Since $\boldsymbol{A}(\varphi)$ is finite, there is also a finite set Z of blocks in π such that $\boldsymbol{A}(\varphi) \subseteq \bigcup Z$. Let $\boldsymbol{X} =_{df.} \left\{\bigcup Y : Y \subseteq Z\right\}$. Obviously, $\boldsymbol{X}$ can be ordered as a lattice by set inclusion, with $\varnothing$ as bottom and $\bigcup Z$ as top. Define $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_\pi$ such that for each $i \in \pi$:

$$\boldsymbol{\Gamma}_i \;=_{df.}\; \left\{\langle X\rangle\,\varphi : X \in \boldsymbol{X}\right\} \qquad \text{and} \qquad \boldsymbol{\Theta}_i \;=_{df.}\; \left\{[X]\,\varphi : X \in \boldsymbol{X}\right\}.$$

Clearly, with Z finite, so are $\boldsymbol{X}$ and $\boldsymbol{\Gamma}_i$, for each $i \in \pi$. We prove that $[\![\varphi]\!] = [\![f(\boldsymbol{\Gamma}_\pi)]\!]$, as the claim that $[\![\varphi]\!] = \langle\!\langle g(\boldsymbol{\Theta}_\pi)\rangle\!\rangle$ follows by a similar argument.

For the inclusion of $[\![\varphi]\!]$ in $[\![f(\boldsymbol{\Gamma}_\pi)]\!]$, merely observe that for each $\pi_i \in \pi$ and each $X \in \boldsymbol{X}$, for each $X \in \boldsymbol{X}$, we have that $[\![\varphi]\!] \subseteq [\![\langle X\rangle\varphi]\!] \subseteq [\![f_{\pi_i}(\langle X\rangle\varphi)]\!]$. The last inclusion is in virtue of Fact 9.5.5.

For the inclusion of of $[\![f(\boldsymbol{\Gamma}_\pi)]\!]$ in $[\![\varphi]\!]$, first consider the case in which φ is a contradiction, *i.e.*, if $[\![\varphi]\!] = \emptyset$. Then, subsequently $\overline{apr}_X([\![\varphi]\!]) = \emptyset$ and $\overline{apr}_\emptyset(\overline{apr}_X([\![\varphi]\!])) = \emptyset$, for each $X \in \boldsymbol{X}$. This is because of the rough set law that $\overline{apr}(\emptyset) = \emptyset$. It follows that both $[\![\langle X\rangle\varphi]\!] = \emptyset$ and $[\![\langle\emptyset\rangle\langle X\rangle\varphi]\!] = \emptyset$. Hence, $[\![f_{\pi_i}(\langle X\rangle\varphi)]\!] = \emptyset$, for each $\pi_i \in \pi$, and, therefore, $[\![f(\boldsymbol{\Gamma}_\pi)]\!] = \emptyset$, as well. So, for the remainder of the proof we may assume $[\![\varphi]\!]$ to be inhabited.

Assume there be a valuation s such that $s \notin [\![\varphi]\!]$. As we could assume $[\![\varphi]\!]$ to be non-empty, with Fact 2.2.10 on page 43, above, and the definition of $\langle\emptyset\rangle\varphi$, it follows that $[\![\langle\emptyset\rangle\varphi]\!] = 2^A$. Thus, in particular, $s \in [\![\langle\emptyset\rangle\varphi]\!]$. With, $\boldsymbol{A}(\varphi) \subseteq \bigcup Z$, however, $[\![\varphi]\!] = \overline{apr}_{\bigcup Z}([\![\varphi]\!]) = [\![\langle\bigcup Z\rangle\varphi]\!]$, by Fact 2.3.10 on page 51, and so $s \notin [\![\langle\bigcup Z\rangle\varphi]\!]$. Because $\boldsymbol{X}$ is finite, there is an $X \in \boldsymbol{X}$ such that $s \notin [\![\langle X\rangle\varphi]\!]$ and for which it is moreover the case that $s \in [\![\langle X'\rangle\varphi]\!]$, for all $X' \in \boldsymbol{X}$ with $X' \subsetneq X$. We may moreover assume that X is not empty. Hence, $\pi_i \subseteq X$, for some $\pi_i \in Z$. Define $X^* =_{df.} X - \pi_i$; obviously $X^* \in \boldsymbol{X}$ and $X^* = \pi_{-i} \cap X^*$. Then, $s \in [\![\langle X^*\rangle\varphi]\!]$, *i.e.*, $s \in \overline{apr}_{X^*}([\![\varphi]\!])$. Now consider the following equalities:

$$\overline{apr}_{X^*}([\![\varphi]\!]) = \overline{apr}_{\pi_{-i}\cap X^*}([\![\varphi]\!]) =_{\text{Fact 2.2.8}} \overline{apr}_{\pi_{-i}}(\overline{apr}_{X^*}([\![\varphi]\!])) = [\![\langle\pi_{-i}\rangle\langle X^*\rangle\varphi]\!].$$

Hence $s \in [\![\langle\pi_{-i}\rangle\langle X^*\rangle\varphi]\!]$. It follows that $s \notin [\![\langle\pi_{-i}\rangle\langle X^*\rangle\varphi \to \langle X^*\rangle\varphi]\!]$. Accordingly, $s \notin [\![f_i(\langle X^*\rangle\varphi)]\!]$ and, *a fortiori*, $s \notin [\![f(\boldsymbol{\Gamma}_\pi)]\!]$. $\dashv$

The following theorem can now be established.

Theorem 9.5.10 *Let Γ and Θ be theories in $L(A)$ and let π and π' be partitions of A. Then there are families of theories $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ such that:*

$$\Gamma \vdash^{\text{CPC}} \Theta \quad \textit{iff} \quad \boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'}.$$

Proof: Almost immediately from Proposition 9.5.6 and Lemma 9.5.9 $\dashv$

Hence, for every classical statement $\Gamma \vdash^{\text{CPC}} \Theta$ there is an equivalent game-theoretical statement $\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_\pi$ in $\Lambda_{\pi,\pi'}$. Proposition 9.5.6 established that this claim also holds in the opposite direction. Hence, we obtain the following corollary, which states that for any partitions π, π', π'' and π''' the statements of any $\Lambda_{\pi,\pi'}$ have their counterparts in $\Lambda_{\pi'',\pi'''}$.

Corollary 9.5.11 *Let π, π', π'' and π''' be partitions of A and let $\boldsymbol{\Gamma}_\pi$ and $\boldsymbol{\Theta}_{\pi'}$ be families of theories in $L(A)$. Then there are families of theories $\boldsymbol{\Gamma}'_{\pi''}$ and $\boldsymbol{\Theta}'_{\pi'''}$ such that:*

$$\boldsymbol{\Gamma}_\pi \vDash \boldsymbol{\Theta}_{\pi'} \quad \textit{iff} \quad \boldsymbol{\Gamma}'_{\pi''} \vDash \boldsymbol{\Theta}'_{\pi'''}.$$

Proof: Almost immediately from Proposition 9.5.6 and Theorem 9.5.10. $\dashv$

9.6 Conclusion

In this chapter we advanced a concept of logical consequence based on the game-theoretical notion of maximum equilibrium. For the notion of game-theoretical consequence there are various possible definitions, involving different game-theoretical solution concepts. We have chosen for the option that is closest to classical logic, as to assure that the idiosyncratic features of the framework can indeed be ascribed to the game-theoretical angle we adopted and not so much to non-standard features of the underlying propositional logic. Other choices are, however, quite possible and worth investigating. In particular, one could define a game-theoretical notion of consequence using the solution concept of *maximal* equilibrium, instead of *maximum* equilibrium. This would give rise to a non-monotonic framework.

Classical consequence was proved to be a special case of game-theoretical consequence. From this perspective, it stands to reason to investigate game-theoretical consequence using the standard logical techniques and concepts. The issue of sound, complete and elegant formal systems for it is still very much open in this respect.

Game-theoretical consequence, however, also raises some issues of its own, for the proper treatment of which it would seem that concepts from other sciences should be employed. We have already mentioned social choice theory as a possible conceptual source to get a firm grasp of how to combine and distribute theories, if the latter are looked upon as representing preference orders.

In distributed evaluation games the players were identified with the variables they control. So far the emphasis has been on the set maximum equilibria given different theories defining the preferences of the players. We could also invert this image, and take the preferences of players as fixed and investigate the sets of maximum equilibria for varying assignments of the variables to the players. Game theory may here provide the apposite concepts.

Another issue is that of the existence of maximum equilibria in distributed evaluation games. This is the game-theoretical counterpart of the issue of satisfiability in classical logic. Maximum equilibria in pure strategies do not in general exist, and only pure strategies we considered. Lattice theoretic restrictions may be imposed on the strategies and preferences of players such that the existence of equilibria is guaranteed (cf. Topkis (1998), Fudenberg and Tirole (1991)). An example is the lattice-theoretic concept of (quasi-)supermodularity, which is closely related to economic notion of complementarity. These reflections, however, raise the question what these concepts correspond to on a logical level.

Game-theoretical consequence provides a generalization of classical logic, in the study of which we argued concepts from game-theory, economics and social choice theory become relevant and apposite.

Bibliography

Arnott, R. and Small, K. (1994), “The Economics of Traffic Congestion.” *American Scientist* **82**, 446–455.

Arrow, K. J. (1963), *Social Choice and Individual Values*. Wiley, New York, second edition.

Aumann, R. J. (1985), “What is Game Theory Trying To Accomplish?.” in: Arrow, K. and Honkapohja, S. (eds.), *Frontiers of Economics*. Basil Blackwell, London and New York. Includes comments by Reinhard Selten and Martin Shubik.

——— (1997), “Game Theory.” in: Eatwell, J., Milgate, M. and Newman, P. (eds.), *Game Theory*, The New Palgrave, pp. 1–54. Macmillan, London and Basingstoke.

Aumann, R. J. and Brandenburger, A. (1995), “Epistemic Conditions of Nash Equilibrium.” *Econometrica* **63**, 1161–1180.

Baltag, A. (1999), “A Logic for Games.” in: Pauly, M. and Baltag, A. (eds.), *Proceedings of the ILLC Workshop on Logic and Games, Held in Amsterdam, November 19-20, 1999*, ILLC Prepublications Series PP-1999-25, pp. 19–20. ILLC, Amsterdam.

——— (2002), “A Logic for Suspicious Players: Epistemic Actions and Belief-Updates in Games.” *Bulletin of Economic Research* **54** (1), 1–45.

Barwise, J. (1977), “An Introduction to First-Order Logic.” in: Barwise, J. (ed.), *Handbook of Mathematical Logic*, chapter 1. North-Holland, Amsterdam.

Battigalli, P. and Bonanno, G. (1999), “Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory.” *Research in Economics* **53**, 149–225.

Benthem, J. F. A. K. van (1996), *Exploring Logical Dynamics*. Studies in Logic, Language and Information. CSLI, Stanford, Ca.

——— (2001a), “Games in Dynamic Epistemic Logic.” *Bulletin of Economic Research* **53** (4), 219–248.

——— (2001b), *Logic and Games. Lecture Notes*. ILLC, Amsterdam.

——— (2002), “Extensive Games as Process Models.” *Journal of Logic, Language and Information* **11** (3), 289–313. Special issue on Logic and Games edited by P. Dekker and M. Pauly.

——— (2003a), “The Epistemic Logic of IF Games.” ILLC Prepublication Series

PP-2003-04, ILLC, Amsterdam.

——— (2003b), “Rational Dynamics and Epistemic Logic in Games.” ILLC Prepublication Series PP-2003-06, ILLC, Amsterdam.

——— (2003c), “What Logic Games are Trying to Tell Us.” ILLC Prepublication Series PP-2003-05, ILLC, Amsterdam.

——— (no date-a), “Logic and Game Theory: Close Encounters of the Third Kind.” in preparation.

——— (no date-b), “Logic Games are Complete for Game Logics.” in preparation.

——— (to appear), “Games and Strategies inside Elementary Logic.” in: *Proceedings of the 7th Asian Logic Conference, Taiwan, June, 1999*. to appear.

Bernheim, B. D. (1984), “Rationalizable Strategic Behavior.” *Econometrica* **52**, 1007–1028.

Binmore, K. (1992), *Fun and Games. A Text on Game Theory*. D.C. Heath and Company, Lexington, MA.

Blackburn, P., de Rijke, M. and Venema, Y. (2001), *Modal Logic*. Volume 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge U.P., Cambridge.

Bonanno, G. (1991), “The Logic of Rational Play in Games of Perfect Information.” *Economics and Philosophy* **7**, 37–65.

——— (1998), *Branching Time Logic, Perfect Information Games and Backward Induction*. Department of Economics, University of California.

Boutilier, C. (1994), “Towards a Logic for Qualitative Decision Theory.” in: *Proceedings of the Fourth International Conference on Knowledge Representation and Reasoning (KR’94)*, pp. 75–86. Morgan Kaufmann, San Francisco.

Boutilier, C., Shoham, Y. and Wellman, M. P. (1997), “Economic Principles of Multi-Agent Systems.” *Artificial Intelligence* **94**, 1–6.

Braess, D. (1968), “Über ein Paradoxon aus der Verkehrsplanung.” *Unternehmensforschung* **12**, 258–268.

Brandenburger, A. and Dekel, E. (1987), “Rationalizability and Correlated Equilibria.” *Econometrica* **55**, 1391–1402.

Bruin, B. P. de (1998), “Game Transformations and Game Equivalence.” Discussion Paper X-1999-01, ILLC, Amsterdam.

——— (2000), “Modelling Knowledge in Games, Topology and Logic.” Master’s Thesis X-1999-01, ILLC, Amsterdam.

——— (in preparation), “The Epistemic and Rationality Assumptions of Game Theoretic Solution Concepts (working title).” Ph.D. thesis, Institute for Logic, Language and Information, Amsterdam.

——— (to appear), “Game Theory in Philosophy.” *Topoi* **XXIV** (2). To be published in 2005.

Carnap, R. (1943), *Formalization of Logic*. Harvard U.P., New Haven.

Chang, C. C. and Keisler, J. K. (1973), *Model Theory*. Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam.

Cholvy, L. and Garion, C. (2001), "An Attempt to Adapt a Logic of Conditional Preferences for Reasoning with Contrary-To-Duties." *Fundamenta Informatica* **48**, 183–204.

Cleave, J. P. (1991), *A Study of Logics*. Oxford U.P., Oxford.

D'Agostino, M. (1999), "Tableau Methods for Classical Propositional Logic." in: D'Agostino, M., Gabbay, D. M., Hähnle, R. and Posegga, J. (eds.), *Handbook of Tableau Methods*. Kluwer, Dordrecht.

Davidson, D. (1980), *Essays on Actions and Events*. Oxford U.P., Oxford.

Ditmarsch, H. van (2000), "Knowledge Games." Ph.D. thesis, University of Groningen and Institute for Logic, Language and Information, Groningen and Amsterdam.

Doets, K. (1996), *Basic Model Theory*. Studies in Logic, Language and Information. CSLI Publications, Stanford, Ca.

Dunne, P. E. and van der Hoek, W. (to appear), "Representation and Complexity in Boolean Games." in: *Proceedings of the 9th European Conference Logics in Artificial Intelligence, September 27-30, Lisbon, Portugal*, Volume 3229 of *Lecture Notes on Artificial Intelligence*. Springer Verlag, Berlin.

Düntsch, I. (1999), *Rough Sets and Algebras of Relations*. University of Ulster.

——— (no date), *A Logic for Rough Sets*. University of Ulster.

Eatwell, J., Milgate, M. and Newman, P. (eds.) (1989), *Game Theory*. The New Palgrave. Macmillan, London and Basingstoke, 2nd edition.

Ehrenfeucht, A. (1962), "An Application of Games to the Completeness Problem for Formalized Theories." *Fundamenta Mathematicae* **49**, 129–141.

Epstein, R. L. (1995), *The Semantic Foundations of Logic: Propositional Logics*. Oxford U.P., Oxford, second edition.

Etchemendy, J. (1990), *The Concept of Logical Consequence*. Harvard U.P., Cambridge, Mass.

Fagin, R., Halpern, J., Moses, J. and Vardi (1995), *Reasoning about Knowledge*. MIT Press, Cambridge, Mass.

Fitting, M. (1999), "Introduction to Handbook of Tableau Methods." in: D'Agostino, M., Gabbay, D. M., Hähnle, R. and Posegga, J. (eds.), *Handbook of Tableau Methods*. Kluwer, Dordrecht.

Franssen, M. (1997), "Some Contributions to Methodological Individualism in the Social Sciences." Ph.D. thesis, University of Amsterdam, Amsterdam.

Fudenberg, D. and Tirole, J. (1991), *Game Theory*. The MIT Press, Cambridge, Mass.

Gale, D. and Steward, F. H. (1953), "Infinite Games with Perfect Information." in: Kuhn, H. W. and Tucker, A. W. (eds.), *Contributions to the Theory of Games, Volume II*, Annals of Mathematics Studies 28, pp. 1–54. Princeton U.P., Princeton, N.J.

Gärdenfors, P. (1988), *Knowledge in Flux*. The MIT Press, Cambridge, Mass.

Gentzen, G. (1934–35), "Untersuchungen über das logische Schließen." *Mathemat-*

ische Zeitschrift **39**, 176–210, 405–431. English translation in M.E. Szabo (ed.), *The Collected Papers of Gerhard Gentzen*, North-Holland, Amsterdam, 1969.

Gerbrandy, J. (1999), "Bisimulations on Planet Kripke." Ph.D. thesis, University of Amsterdam, Amsterdam. ILLC Dissertation Series DS-1999-01.

Goldblatt, R. (1992), *Logics of Time and Computation*. Volume 7 of *CSLI Lecture Notes*. CSLI Publications, Stanford, 2nd edition.

Goranko, V. (2001a), "The Basic Algebra of Game Equivalences." in: Pauly, M. and Sandu, G. (eds.), *ESSLLI Workshop on Logic and Games*.

——— (2001b), "Coalition Games and Alternating Temporal Logics." in: Benthem, J. van (ed.), *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Eighth Conference (TARK 2001)*. Morgan Kaufmann, San Francisco.

Grätzer, G. (1968), *Universal Algebra*. D. Van Nostrand, Princeton, N.J.

Groenendijk, J. and Stokhof, M. (1997), "Questions." in: Benthem, J. van and ter Meulen, A. (eds.), *Handbook of Logic and Language*. Elsevier, Amsterdam.

Halmos, P. R. (1960), *Naive Set Theory*. The University Series in Undergraduate Mathematics. D. van Nostrand Company, Inc., Princeton, N.J.

Harel, D., Kozen, D. and Tiuryn, J. (2001), *Dynamic Logic*. MIT Press, Cambridge, Mass. & London.

Harrenstein, B. P. (1998), "From a Modal Point of View. A Logical Investigations into Modalities De Dicto and De Re." Technical Note X-1998-02, ILLC, Amsterdam.

——— (2002), "A Game-theoretical Notion of Consequence." in: Bonanno, G., Colombatto, E. and van der Hoek, W. (eds.), *Fifth Conference on Logic and the Foundations of Game and Decision Theory (LOFT 5)*. ICER, Turin.

——— (to appear-a), "Logical Consequence and the Theory of Games." *Philosophia Scientiae.* to appear in the special issue on game theory and logic.

——— (to appear-b), "Set Induced Relations and Relational Semantics." Technical Report UU-CS-2004-032, ICS, Utrecht University, Utrecht.

Harrenstein, B. P., van der Hoek, W., Meyer, J.-J. C. and Witteveen, C. (1999), "Subgame Perfect Nash-Equilibria in Dynamic Logic." in: Pauly, M. and Baltag, A. (eds.), *Proceedings of the ILLC Workshop on Logic and Games, Held in Amsterdam, November 19-20, 1999*, ILLC Prepublications Series PP-1999-25, pp. 29–30. ILLC, Amsterdam.

——— (2000), "A Modal Interpretation of Nash-Equilibria and Some Related Concepts." in: Giacomo, B., Colombatto, E. and van der Hoek, W. (eds.), *Proceedings of the Fourth Conference on Logic and the Foundations of Game and Decision Theory (LOFT4), June 30–July 2, 2000*, pp. 1–21. ICER, Torino.

——— (2001), "Boolean Games." in: Benthem, J. van (ed.), *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Eighth Conference (TARK 2001)*, pp. 287–298. Morgan Kaufmann, San Francisco.

——— (2002), "On Modal Logic Interpretations of Games." in: Harmelen, F. van (ed.), *ECAI 2002, 15th European Conference on Artificial Intelligence, July 21–26*

2002, Lyon, France, Volume 77 of *Frontiers in Artificial Intelligence and Applications*, pp. 28–32. IOS Press, Amsterdam.

——— (2003), "A Modal Characterization of Nash Equilibrium." *Fundamenta Informatica* **57** (2–4), 281–321.

Henkin, L. (1961), "Some Remarks on Infinite Long Formulas." in: Bernays, P. (ed.), *Infinitistic Methods. Proceedings of the Symposium on Foundations of Mathematics*, pp. 167–183. Pergamon Press and PWN, Oxford and Warsaw.

Henkin, L., Monk, J. D. and Tarski, A. (1971), *Cylindric Algebras, Part I*. Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam.

Hintikka, J. (1962), *Knowledge and Belief*. Cornell U.P., New York.

——— (1973), *Logic, Language Games, and Information*. Oxford U.P., Oxford.

——— (1983), *The Game of Language. Studies in Game-Theoretical Semantics and Its Applications*. Volume 22 of *Synthese Library*. D. Reidel, Dordrecht.

——— (2002), "Hyperclassical Logic (A.K.A. IF Logic) and Its Implications for Logical Theory." *The Bulletin of Symbolic Logic* **3** (8), 404–423.

Hintikka, J. and Sandu, G. (1997), "Game-theoretical Semantics." in: Benthem, J. van and ter Meulen, A. (eds.), *Handbook of Logic and Language*. Elsevier, Amsterdam.

Hodges, W. (1985), *Building Models by Games*. Volume 2 of *London Mathematical Society Student Texts*. Cambridge U.P., Cambridge.

Hodges, W. (1993), *Model Theory*. Cambridge U.P., Cambridge.

Hughes, G. E. and Cresswell, M. J. (1968), *An Introduction to Modal Logic*. Methuen, London.

Jongh, D. de and Veltman, F. (1995), *Intensional Logics*. FdW, University of Amsterdam, Amsterdam.

Kelly, J. S. (1987), *Social Choice Theory: An Introduction*. Springer, New York.

Kleene, S. C. (1952), *Introduction to Metamathematics*. North-Holland, Amsterdam.

——— (1967), *Mathematical Logic*. John Wiley & Sons, New York.

Kooi, B. P. (2003), "Knowledge, Chance and Change." Ph.D. thesis, University of Groningen and Institute for Logic, Language and Information, Groningen and Amsterdam.

Kozen, D. (1983), "Results in the Propositional μ-Calculus." *Theoretical Computer Science* **27**, 333–354.

Kraus, S. (1997), "Negotiation and Cooperation in Multi-Agent Environments." *Artificial Intelligence* **94**, 79–97.

Kraus, S., Lehmann, D. and Magidor, M. (1990), "Nonmonotonic Reasoning, Preferential Models and Cumulative Logics." *Artificial Intelligence* **44**, 167–207.

Kreps, D. M. (1997), "Nash Equilibrium." in: Eatwell, J., Milgate, M. and Newman, P. (eds.), *Game Theory*, The New Palgrave, pp. 157–177. Macmillan, London and Basingstoke.

Kuhn, H. W. (1953), "Extensive Games and the Problem of Information." in:

Kuhn, H. W. and Tucker, A. W. (eds.), *Contributions to the Theory of Games, Volume II*, Annals of Mathematics Studies 28, pp. 1–54. Princeton U.P., Princeton, N.J. Reprinted in Kuhn (1997).

——— (ed.) (1997), *Classics in Game Theory*. Princeton U.P., Princeton, N.J.

Kuratowski, K. and Mostowski, A. (1976), *Set Theory. With and Introduction to Descriptive Set Theory*. Studies in Logic and the Foundations of Mathematics. North-Holland and PWN, Amsterdam and Warszawa, second edition.

Leonard, R. J. (1992), "Creating a Context for Game Theory." in: Weintraub, E. R. (ed.), *Toward a History of Game Theory*, pp. 29–76. Duke U.P., Durham, N.C. and London. History of Political History. Annual Supplement to Volume 24.

Lewis, D. (1969), *Convention: A Philosophical Study*. Harvard U.P., Cambridge, Mass.

——— (1973), *Counterfactuals*. Basil Blackwell, Oxford.

Lin, Y. Y. and Lin, T. Y. (n.d.), *Generalization of Rough Sets using Modal Logics*. Lakehead University and San Jose University.

Linder, B. van (1996), "Modal Logics for Rational Agents." Ph.D. thesis, Faculty of Mathematics and Computer Science, Utrecht University, Utrecht.

Lorenzen, P. and Lorenz, K. (1978), *Dialogische Logik*. Wissenschaftliche Buchgesellschaft, Darmstadt.

Luce, R. D. and Raiffa, H. (1957), *Games and Decisions. Introduction and Critical Survey*. John Wiley & Sons, New York.

Makinson, D. (1994), "General Patterns in Nonmonotonic Reasoning." in: Gabbay, D., Hogger, C. J. and Robinson, J. A. (eds.), *Nonmonotonic Reasoning and Uncertain Reasoning*, Volume 3 of *Handbook of Logic in Artificial Intelligence and Logic Programming*, pp. 35–110. Oxford U.P., Oxford.

Meyer, J.-J. C. and van der Hoek, W. (1995), *Epistemic Logic for AI and Computer Science*. Volume 41 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge U.P., Cambridge.

Mirkowski, P. (1992), "What Were von Neumann and Morgenstern Trying to Accomplish." in: Weintraub, E. R. (ed.), *Toward a History of Game Theory*, pp. 113–147. Duke U.P., Durham, N.C. and London. History of Political History. Annual Supplement to Volume 24.

Myerson, R. B. (1991), *Game Theory. Analysis of Conflict*. Harvard U.P., Cambridge, Mass.

Nasar, S. (1998), *A Beautiful Mind. A Biography of John Forbes Nash, Jr., Winner of the Nobel Prize for Economics 1994*. Simon and Schuster, New York.

Nash, J. (1950), "Equilibrium Points in n-Person Games." *PNAS* **36**, 48–49. Reprinted in Kuhn (1997).

——— (1951), "Non-Cooperative Games." *Annals of Mathematics* **54**, 286–295. Reprinted in Kuhn (1997).

Neumann, J. von (1928), "Zur Theorie der Gesellschaftsspiele." *Mathematische An-*

nalen **100**, 295–320.

Neumann, J. von and Morgenstern, O. (1944), *Theory of Games and Economic Behavior*. Princeton U.P., Princeton, N.J.

Osborne, M. J. and Rubinstein, A. (1994), *A Course in Game Theory*. MIT Press, Cambridge, Mass.

Otterloo, S. van , van der Hoek, W. and Wooldridge, M. (2004), "Preferences in Game Logics." in: *Proceedings of AAMAS 2004*. New York. Accepted for publication.

Parikh, R. (1984), "Logics of Knowledge, Games and Dynamic Logic." in: Joseph, M. and Shyamasundar, R. (eds.), *Foundations of Software Technology and Theoretical Computer Science*, Lecture Notes in Computer Science 181, pp. 202–222. Springer, New York.

——— (1985), "The Logic of Games and its Applications." in: Karpinski, M. and van Leeuwen, J. (eds.), *Topics in the Theory of Computation, FCT'83, Borgholm, Sweden, August 21-27, 1983*, Annals of Discrete Mathematics 14, pp. 111–140. North-Holland, Amsterdam.

——— (1998), "A Connection between Nash Equilibria and Propositional Logic." Unpublished two-page technical note.

Pauly, M. (2001), "Logic for Social Software." Ph.D. thesis, Institute for Logic, Language and Information, Amsterdam.

Pawlak, Z. (1991), *Rough Sets: Theoretical Aspects of Reasoning about Data*. Kluwer, Dordrecht.

Pearce, D. G. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection." *Econometrica* **52**, 1029–1050.

Popkorn, S. (1994), *First Steps in Modal Logic*. Cambridge U.P., Cambridge.

Ramsey, F. P. (1926), "Truth and Probability." in: Braithwaite, R. (ed.), *The Foundations of Mathematics and other Logical Essays*, pp. 156–198. Routledge and Kegan Paul, London.

Rapoport, A. (1997), "Prisoner's Dilemma." in: Eatwell, J., Milgate, M. and Newman, P. (eds.), *Game Theory*, The New Palgrave, pp. 199–204. Macmillan, London and Basingstoke.

Rosenschein, J. S. and Zlotkin, G. (1994), *Rules of Encounter. Designing Conventions for Automated Negotiation among Computers*. MIT Press, Cambridge, Mass.

Rosenthal, R. (1981), "Games of Perfect Information, Predatory Pricing and the Chain Store Paradox." *Journal of Economic Theory* **25**, 92–100.

Saarinen, E. (ed.) (1978), *Game Theoretical Semantics*. D. Reidel, Dordrecht.

Sandholm, T. W. and Lesser, V. R. (1997), "Coalitions among Computationally Bounded Agents." *Artificial Intelligence* **94** (1), 99–137. Special Issue on Economic Principles of Multiagent Systems.

Savage, L. J. (1954), *Foundations of Statistics*. John Wiley, New York.

Schelling, T. C. (1960), *The Strategy of Conflict*. Harvard U.P., Cambridge, Mass.

Scott, D. (1973), “Background to Formalization.” in: LeBlanc, H. (ed.), *Truth, Syntax and Modality*, Volume 68 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam and London.

——— (1974), “Rules and Derived Rules.” in: LeBlanc, H. (ed.), *Logical Theory and Semantic Analysis. Essays Dedicated to Stig Kanger on His Fiftieth Birthday*, Volume 63 of *Synthese Library*, pp. 147–161. D. Reidel, Dordrecht and Boston.

Segerberg, K. (1982), *Classical Propositional Operators*. Volume 5 of *Oxford Logic Guides*. Oxford U.P., Oxford.

Selten, R. (1965), “Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit.” *Zeitschrift für die gesamte Staatswissenschaft* **121**, 301–324.

Sevenster, M. (2004), “Complexity Results for an Independence Friendly Propositional Logic.” in: Otterloo, S. van , McBurney, P., van der Hoek, W. and Wooldridge, M. (eds.), *Proceedings of the First Knowledge and Games Workshop, Liverpool, 10th and 11th July 2004*, pp. 23–33. AgentLink and the University of Liverpool, Liverpool.

Shoham, Y. (1988), *Reasoning About Change*. MIT Press, Cambridge, Mass.

Smullyan, R. (1968), *First-Order Logic*. Springer Verlag, New York.

Spohn, W. (1982), “How To Make Sense of Game Theory.” in: Stegmüller, W., Balzer, W. and Spohn, W. (eds.), *Philosophy of Economics*, pp. 298–312. Springer, Berlin.

Stirling, C. (1996), “Games and Modal Mu-Calculus.” in: Margaria, T. and Steffen, B. (eds.), *Proceedings of Second International Workshop on Tools and Algorithms for the Construction and Analysis of Systems (TACAS’96)*, Lecture Notes in Computer Science 1055, pp. 298–312. Springer, New York.

Tamminga, A. M. (1994), “Logics of Rejection: Two Systems of Natural Deduction.” *Logique et Analyse* **146**, 169–208.

——— (2001), “Belief Dynamics. [Epistemo]logical Investigations.” Ph.D. thesis, Institute for Logic, Language and Information, Amsterdam.

Tan, T. C.-C. and da Costa Werlang, S. R. (1988), “The Bayesian Foundations of Solution Concepts of Games.” *Journal of Economic Theory* **45**, 370–391.

Tarski, A. (1936a), “Der Wahrheitsbegriff in den formalisierten Sprachen.” *Studia Philosophica* **1**, 261–405. Translated as ‘The concept of truth in formalized languages’ in: J.H. Woodger (ed.), *Logic, Semantics and Metamathematics*.

——— (1936b), “Über den Begriff der logischen Folgerung.” in: *Actes du Congrès International de Philosophie Scientifique VII*, Actualités Scientifiques et Industrielles 394, pp. 1–11. Hermann & Cie, Paris. Translated as ‘On the Concept of Logical Consequence’ in: J.H. Woodger (ed.), *Logic, Semantics and Metamathematics*.

——— (1944), “The Semantic Conception of Truth.” *Philosophy and Phenomenological Research* **4**, 341–375.

Thompson, F. B. (1952), “Equivalence of Games in Extensive Form.” Research Memorandum RM-759, RAND Corporation, Santa Monica, Ca. Reprinted in Kuhn (1997).

Topkis, D. M. (1998), *Supermodularity and Complementarity*. Princeton U.P., Princeton, N.J.

Veltman, F. (1996), “Defaults in Update Semantics.” *Journal of Philosophical Logic* **25**, 221–261.

Venema, Y. (2001), “Representing Game Algebras.” in: Pauly, M. and Sandu, G. (eds.), *ESSLLI Workshop on Logic and Games*.

Vos, M. de and Vermeir, D. (2001), “Decisions, Agents and Games.” in: Benthem, J. van (ed.), *Theoretical Aspects of Rationality and Knowledge, Proceedings of the Eighth Conference (TARK 2001)*, pp. 219–232. Morgan Kaufmann, San Francisco.

Walsh, W. E. (2001), “Market Protocols for Decentralized Supply Chain Formation.” Ph.D. thesis, University of Michigan.

Walsh, W. E. and Wellman, M. P. (2000), “MarketSAT: An Extremely Decentralized (but Really Slow) Algorithm for Propositional Satisfiability.” in: *Seventeenth National Conference on Artificial Intelligence*, pp. 303–309.

Walsh, W. E., Yokoo, M., Hirayama, K. and Wellman, M. P. (2001), “On Market-Inspired Approaches to Propositional Satisfiability.” in: *Seventeenth International Joint Conference on Artificial Intelligence*, pp. 1152–1158.

Weerdt, M. de (2003), “Plan Merging in Multi-Agent Systems.” Ph.D. thesis, Delft University of Technology, Delft.

Williams, J. D. (1954), *The Compleat Strategyst*. the RAND Series. McGraw Hill, New York.

Wittgenstein, L. (1953), *Philosophische Untersuchungen – Philosophical Investigations*. Basil Blackwell, Oxford. German-English edition.

Yao, Y. Y., Wong, M. and Lin, T. Y. (1997), “A Review of Rough Set Models.” in: Lin, T. Y. and Cercone, N. (eds.), *Rough Sets and Data Mining. Analysis of Imprecise Data*, pp. 47–75. Kluwer, Dordrecht.

Yokoo, M., Durfee, E. H., Ishida, T. and Kuwabara, K. (1992), “Distributed Constraint Satisfaction for Formalizing Distributed Problem Solving.” in: *International Conference on Distributed Computing Systems*, pp. 614–621.

——— (1998), “The Distributed Constraint Satisfaction Problem: Formalization and Algorithms.” *Knowledge and Data Engineering* **10** (5), 673–685.

Zermelo, E. (1913), “Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels.” in: Hobson, E. W. and Love, A. E. H. (eds.), *Proceedings of the Fifth International Congress of Mathematicians*, Volume II, pp. 501–4. Cambridge U.P., Cambridge.

Logica in conflict

(samenvatting in het Nederlands)

Speltheorie biedt een formeel raamwerk voor de strategische analyse van situaties van sociale interactie, verder ook wel conflictsituaties of spelen genoemd. Kenmerkend voor conflictsituaties is dat de uitkomst afhangt van de beslissingen die verschillende actoren kunnen maken ten aanzien van een individueel te volgen strategie en dat voor ieder individu de raadzaamheid van een bepaalde keuze essentieel kan afhangen van de beslissingen die de andere individuen nemen. Zo bezien is conflict een natuurlijk sociaal verschijnsel.

Naast haar voor de hand liggende relevantie voor de sociale, economische en politieke wetenschappen, heeft de speltheorie ook belangrijke toepassingen binnen zulke uiteenlopende disciplines als de evolutionaire biologie, wiskundige logica en verzamelingenleer. De afgelopen jaren heeft de speltheorie zich ook kunnen verheugen op een toegenomen interesse vanuit de informatica en de (gedistribueerde) *Artificiële Intelligentie*.

Binnen de informatica wordt formele logica toegepast bij de specificatie en verificatie van computerprogramma's en computationele systemen. Het gedrag van een complex en gedistribueerd systeem kan in sommige gevallen worden gezien als het resultaat van een interactie tussen verschillende (tot op zekere hoogte) autonome processen. Bij het redeneren over dergelijke computersystemen wordt met succes een beroep gedaan op concepten die afkomstig zijn uit de sociale en economische wetenschappen, waaronder de speltheorie. Deze ontwikkeling vormt de achtergrond van het onderzoek waarvan in deze dissertatie verslag wordt gedaan, een logische verkenning waarbij het speltheoretische begrip *strategisch equilibrium* centraal staat. De aanpak behelst zowel een logische analyse van speltheoretische concepten (Deel I) als het gebruik van speltheoretische concepten om logische analyses te verrijken (Delen II en III).

Speltheorie

Speltheoretisch onderzoek betreft situaties waarin meerdere actieve elementen (*spelers*) kunnen worden onderscheiden die ieder de keuze hebben tussen verschillende wijzen van handelen (*strategieën*). Een combinatie van keuzes waarbij iedere speler

	Duif	*Havik*
Duif	2 2	3 1
Havik	1 3	0 0

Figuur 1.

zijn strategie bepaalt (een *strategieprofiel*) resulteert in een unieke uitkomst van het spel. Iedere speler wordt bovendien geacht bepaalde voorkeuren te hebben ten aanzien van de mogelijke uitkomsten. Merk in dit verband op dat de individuen in een conflictsituatie zowel gemeenschappelijke als tegengestelde belangen kunnen hebben en dat puur antagonisme eerder uitzondering dan regel is. Het speltheoretisch vraagstuk betreft welke beslissing iedere speler het best kan nemen in het licht van zijn individuele belangen. Het punt is dat de optimaliteit van een bepaalde beslissing voor een speler kan afhangen van de beslissingen die de andere spelers nemen en een zekere circulariteit dient zich aan.

In het inleidende hoofdstuk van hun baanbrekende werk *Theory of Games and Economic Behavior*, dat in 1944 zijn eerste editie beleefde, betogen von Neumann en Morgenstern dat het speltheoretisch vraagstuk de wiskundige voor een nieuw probleem stelt waarvan niet op voorhand mag worden aangenomen dat traditionele wiskundige concepten volstaan voor een bevredigende analyse. In het bijzonder beargumenteerden zij dat een speltheoretische situatie niet zonder meer gerepresenteerd kan worden als een optimaliseringsvraagstuk, waarbij waarden voor variabelen $x_0, \ldots, x_n$ gevonden moeten worden teneinde de waarde van een functie $f(\hat{x}_0, \ldots, \hat{x}_n)$ te maximaliseren. Een speltheoretische situatie zou beter voorgesteld kunnen worden als een verzameling van functies $g_i(\hat{x}_0, \ldots, \hat{x}_n)$, waarvan iedere speler er één tracht te maximaliseren, met dien verstande dat de variabelen waarvan de verschillende functies afhankelijk zijn kunnen overlappen en iedere speler controle heeft over slechts een deel van die variabelen. De traditionele noties van optimaliteit zouden te kort schieten voor de analyse van dergelijke problemen en hun rol dient overgenomen te worden door concepten die specifiek zijn toegesneden op het interactieve en latent circulaire karakter van de materie.

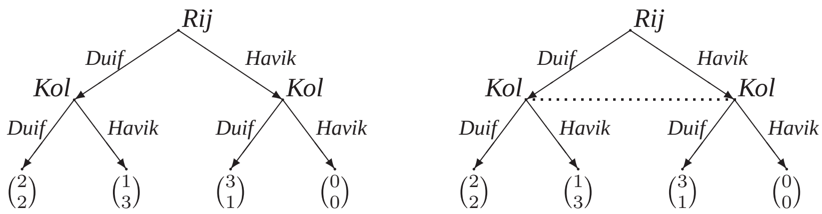
Ter illustratie diene het volgende voorbeeld. Beschouw de situatie waarin twee spelers, *Rij* en *Kol*, ieder de keuze hebben tussen twee mogelijk in te nemen houdingen: een agressieve, als een havik, en een inschikkelijke, als een vredesduif. Hierbij zij aangetekend dat iedere speler er het meeste baat bij heeft zich als een duif op te stellen wanneer hij zich geconfronteerd ziet met een havik, terwijl het beter is een havik te zijn ten opzichte van een duif. De minst nastrevenswaardige uitkomst voor beide spelers resulteert indien zowel de één als de ander zich agressief opstelt; voor beide zou het dan beter zijn geweest zich inschikkelijk te tonen. De situatie is samengevat in Figuur 1, waar *Rij* de keuze heeft tussen de rijen en *Kol* tussen de kolommen. De preferenties van de spelers *Kol* en *Rij* worden aangeduid door, respectievelijk, de getallen rechts boven en linksonder in iedere cel. Voor beide spelers is het afhankelijk van welke strategie de ander kiest of het beter is een inschikkelijke dan wel een agressieve koers te varen.

Speltheoretische oplossingsconcepten dienen ertoe wiskundig vat te krijgen op dergelijke conflictsituaties. Eén van de bekendste oplossingsconcepten is het *Nash equilibrium*, dat uitdrukking geeft aan een intuïtieve notie van strategisch evenwicht. Er is sprake van een Nash equilibrium indien geen van de spelers er voordeel bij heeft eenzijdig af te wijken van zijn gekozen strategie. In het voorbeeld zijn de Nash equilibria gegeven door die keuzes waarbij één zich als een duif en de ander zich als een havik opstelt.

In bepaalde klassen van spelen blijkt de notie van Nash equilibrium nauw verbonden te zijn met een andere belangrijke speltheoretische notie, namelijk de *winnende strategie*. Een handelingswijze geldt als een winnende strategie voor een speler, indien deze een overwinning garandeert ongeacht de handelwijze van eventuele tegenstanders. Beschouw de specifieke klasse van spelen waarin twee antagonisten figureren en iedere uitkomst kan worden geclassificeerd als een overwinning voor de ene dan wel een overwinning voor de andere speler. Het kan eenvoudig worden aangetoond dat een strategieprofiel een Nash equilibrium is dan en slechts dan als het een winnende strategie voorschrijft aan één van de twee spelers.

Voor speltheoretische analyses kunnen ook andere aspecten van een conflictsituatie in beschouwing worden genomen dan enkel de spelers, de hun ter beschikking staande strategieën en hun preferenties over de mogelijke uitkomsten. Zo kan bijvoorbeeld de sequentiële structuur van een spel — d.w.z. de volgorde waarin de verschillende spelers hun keuzes maken — expliciet worden gemaakt in de zogenaamde *extensieve vorm*. Formeel kan een spel dan worden gerepresenteerd als een gelabelde boom, waarbij de wortel, de interne knopen en de bladeren respectievelijk de begintoestand, de beslismomenten voor de spelers en de eindtoestanden vertegenwoordigen. De takken staan voor de keuzes die spelers kunnen maken op een bepaald beslismoment. De linker boom in Figuur 2 geeft de extensieve vorm van het spel van ons voorbeeld waarbij we als additioneel gegeven hebben dat eerst *Rij* en dan *Kol* hun opstelling in het conflict bepalen. Merk op dat *Kol* hier de beschikking heeft over *vier* strategieën: voor beide beslismomenten waar *Kol* een keuze moet maken heeft hij twee opties. Hierdoor kan *Kol* zijn handelingswijze af laten hangen van de beslissing die *Rij* neemt op een eerder moment.

In sommige gevallen is het evenwel onrealistisch aan te nemen dat een speler zijn handelswijze afhankelijk kan maken van eerder gemaakte keuzes, bijvoorbeeld indien de informatiestructuur van het spel zodanig is dat hij in het ongewisse blijft wat betreft de eerder genomen beslissingen. Hierdoor kan het voor een speler onmogelijk blijken om op basis van de hem ter beschikking staande informatie een onderscheid te maken tussen verschillende mogelijke toestanden waarin hij zijn keuze moet maken, hier gerepresenteerd door de beslisknopen. Aldus wordt het hem onmogelijk een strategie



Figuur 2.

te spelen die verschillende acties voorschrijft op ononderscheidbare beslisknopen. Dit epistemische aspect van spelsituaties wordt gerepresenteerd door equivalentieklassen van beslisknopen, *informatieverzamelingen*, aan de structuur van een extensief spel toe te voegen en te postuleren dat op alle beslisknopen in dezelfde informatieverzameling een speler dezelfde keuze dient te maken. In ons voorbeeld zouden aldus voor *Kol* de strategieën die verschillende handelswijzen voorschrijven op zijn twee beslisknopen uitgesloten worden, indien *Kol* wordt geacht onwetend te zijn wat betreft de eerder gemaakte strategiekeuze van *Rij*. Grafisch wordt een informatieverzameling gerepresenteerd als een stippellijn die precies die beslismomenten verbindt die de informatieverzameling bevat, zoals in de rechter boom in Figuur 2. Een extensief spel heet een spel van *volledige infomatie* indien iedere informatieverzameling slechts één beslisknoop bevat, en anders een spel van *onvolledige informatie*.

Deel I: Modale karakterisering van Nash equilibrium

De formele talen van modale logica's zijn bij uitstek geschikt om over relationele structuren te redeneren. Met hun onderliggende boomstructuur lenen extensieve spelen zich bij uitstek voor een modaal logische analyse. Met dit oogmerk wordt in Deel I van deze dissertatie een klasse van multi-modale talen voorgesteld — de klasse van de *multi-modale matrixtalen*. De semantiek van een dergelijke taal is beperkt tot de klasse van de zogenaamde spelframes. Ieder spelframe kan op een systematische manier geassocieerd worden met een extensief spel van volledige informatie en met een eindige horizon.

Het blijkt dat, indien een strategieprofiel van een spel een Nash equilibrium is, dit zijn weerslag vindt in een structurele eigenschap van het spelframe dat met het spel geassocieerd is. Wat meer zij, we tonen aan dat deze structurele eigenschap van frames gekarakteriseerd kan worden middels een formuleschema van de betreffende multimodale taal. Dat wil zeggen dat er een formuleschema $\vartheta(s)$ is, met een parameter s die varieert over strategieprofielen, zodanig dat voor ieder extensief spel E en haar geassocieerde frame $\mathfrak{F}_E$ geldt dat:

$\mathfrak{F}_E \Vdash \vartheta(s)$ dan en slechts dan als s een Nash equilibrium is van het spel E is.

Soortgelijke resultaten worden bewezen voor een verfijning van Nash equilibrium, namelijk het zogenaamde subspel-perfecte equilibrium. Ook tonen we aan dat de analyse kan worden volvoerd gebruikmakend van Propositionele Dynamische Logica (PDL), een bekende multi-modale logica specifiek ontwikkeld om over computer programma's te redeneren.

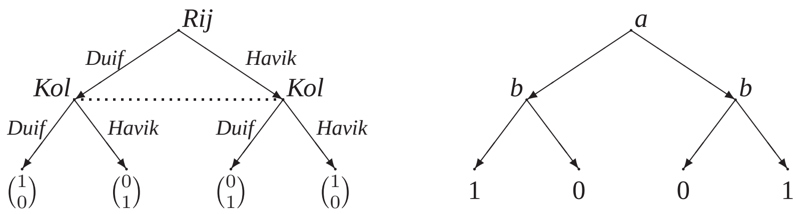
Deel I sluit af met een hoofdstuk dat geheel is gewijd aan de correctheid en volledigheid van een Hilbert-stijl axiomasysteem ten aanzien van de klasse van spelframes.

Deel II: Booleaanse Spelen

In de laatste twee delen wordt de aandacht verlegd naar speltheoretische analyses van propositielogica. Een propositionele taal bevat formules die volgens bepaalde syntactische regels zijn opgebouwd uit propositievariabelen en logische symbolen voor conjunctie ($\wedge$), disjunctie ($\vee$), negatie ($\neg$) en materiële implicatie ($\rightarrow$). Aan de basis van de semantiek voor klassieke propositielogica ligt de notie van een valuatie, een functie die één van de waarheidswaarden 'waar' of 'onwaar' toekent aan iedere propositievariabele en, *via* een inductieve definitie, eveneens aan iedere formule. Klassiek logisch gevolg kan dan worden gekarakteriseerd als een relatie die bestaat tussen twee verzamelingen formules Γ en Θ in geval de waarheid van ten minste één formule uit Θ afgedwongen wordt in alle valuaties waarin alle formules van Γ waar zijn. Dit is een conservatieve uitbreiding van het traditionele idee dat een conclusie logisch volgt uit een verzameling premissen indien de waarheid van de premissen de waarheid van de conclusie afdwingt. Een formule is logisch geldig, indien haar waarheid wordt afgedwongen in alle valuaties.

De bovenstaande informele presentatie kan de indruk wekken dat een valuatie de waarheidswaarde van een propositievariabele bepaalt als ware het een soort van onafhankelijk gegeven mogelijke stand van zaken. Een gedachte die aan de twee laatste delen van deze dissertatie ten grondslag ligt is dat propositievariabelen ook als binaire beslisvariabelen kunnen worden gezien waarvan de waarheidswaardes gecontroleerd worden door spelers met individuele preferenties ten aanzien van de waarheid van bepaalde formules. De valuaties, die tezamen de logische ruimte constitueren, kunnen dan worden beschouwd als de mogelijke uitkomsten van een interactief beslisproces dat geanalyseerd kan worden als een spel. Valuaties die voldoen aan speltheoretische oplossingsconcepten, zoals Nash equilibrium en winnende strategie, verkrijgen zo een logische significantie. Deel II en Deel III van deze dissertatie betreffen de logische consequenties van deze voorstelling van zaken.

In hoofdstuk 5, het eerste hoofdstuk van Deel II, introduceren we *Booleaanse spelen* als reprensentaties van de epistemische structuur van eindige extensieve spelen voor twee antagonistische spelers, 1 and 0, waarvan er maar één het spel kan winnen. Iedere interne knoop van een Booleaanse spel is gelabeld met een binaire beslisvariabele en wel zo dat twee knopen met dezelfde beslisvariabele worden gelabeld dan en slechts dan als zij element zijn van dezelfde informatieverzameling. De controle over de beslisvariabelen is verdeeld over de twee spelers. Een strategie voor een speler is dan een toewijzing van een binaire waarde aan iedere variabele in zijn beheer. Een strategiepro-



Figuur 3. Een extensief spel met onvolledige informatie en haar representatie als een Booleaanse spel, waarbij *Rij* controle heeft over de propositievariabele *a* en *Kol* over *b*.

fiel is dus een toewijzing van binaire waarden aan *alle* beslisvariabelen en bepaalt aldus een valuatie voor de propositievariabelen. Een Booleaans spel zonder de allocatie van controle over de variabelen noemen we een Booleaanse vorm.

Op deze wijze kan met iedere verzameling beslisvariabelen een verzameling Booleaanse spelen worden geassocieerd. Merk op dat de strategieprofielen van ieder Booleaans spel in zo'n verzameling gelijk zijn. Dit maakt het mogelijk een notie van equivalentie tussen Booleaanse vormen te definiëren: twee Booleaanse vormen zijn equivalent indien ieder strategieprofiel in beide spelen dezelfde uitkomst bepaalt.

We definiëren ook een aantal operaties op Booleaanse vormen en tonen aan dat de aldus gevormde algebra van Booleaanse vormen, *modulo* de bovenstaande notie van equivalentie, een Booleaanse algebra is. Deze algebra is bovendien isomorf met de Lindenbaum algebra van de klassieke propositionele taal waarvan propositievariabelen samenvallen met de beslisvariabelen die voorkomen in de Booleaanse vormen. Iedere Booleaanse vorm kan aldus worden geassocieerd met een unieke propositionele formule en *vice versa*.

Het wordt zo mogelijk de speltheoretische eigenschappen van Booleaanse vormen te vergelijken met de logische eigenschappen van de corresponderende propositionele formules. Zo blijkt de logische notie van geldigheid gerelateerd aan het speltheoretische concept van een winnende strategie. De notie van een winnende strategie heeft echter betrekking op Booleaanse spelen, waarbij de controle over beslisvariabelen is gespecificeerd, en niet zozeer op Booleaanse vormen als zodanig, terwijl logische geldigheid een eigenschap van formules is. Op natuurlijke wijze kan evenwel de notie van controle over variabelen verdisconteerd worden in de definitie van logische geldigheid. Laat Δ een deelverzameling propositievariabelen zijn. Een formule φ is dan Δ-geldig indien het mogelijk is de waarheid van φ af te dwingen door een keuze te maken voor de waarden voor de propositievariabelen in Δ. Traditionele logische noties blijken randgevallen van dit gerelativeerde concept; zo valt klassieke geldigheid samen met $\emptyset$-geldigheid, en klassieke vervulbaarheid met A-geldigheid, waar A de volledige verzameling propositievariabelen is. Meer in het algemeen hebben we de volgende correspondentie: een formule is Δ-geldig dan en slechts dan als de speler met controle over Δ een winnende strategie heeft in de corresponderende Booleaanse vorm.

De notie van Δ-geldigheid kan op natuurlijke wijze worden uitgebreid naar een gerelativeerd concept van gevolg. Dan geldt $\Gamma \vDash_\Delta \Theta$ indien het mogelijk is door waarheidswaarden aan de propositievariabelen in Δ toe te kennen, te garanderen dat tenminste één formule in Θ waar is als alle formules in Γ waar zijn. In het zesde hoofdstuk wordt deze gerelativeerde geldigheidsnotie bestudeerd, waarbij zich tenminste twee vragen aandienen. Enerzijds is er het probleem, voor welke theorieën Γ en Θ het het geval is dat $\Gamma \vDash_\Delta \Theta$, gegeven een verzameling propositievariabelen Δ. Voor iedere deelverzameling van propositievariabelen Δ wordt de notie van Δ-gevolg van een correcte en volledige sequenten-calculus voorzien. Anderzijds wordt de vraag onder de loep genomen voor welke verzamelingen van propositievariabelen Δ het zo is dat $\Gamma \vDash_\Delta \Theta$, voor gegeven theorieën Γ en Θ.

Deel III: Speltheoretisch gevolg

Deel III leidt tot een speltheoretische generalisering van de klassieke notie van logisch gevolg. Net als in Deel II, staat hier staat de gedachte centraal dat valuaties kunnen worden beschouwd als de uitkomsten van een strategisch spel, waarbij iedere speler waarden toekent aan een apart deel van de propositievariabelen waarover hij controle heeft. Het traditionele probleem van logisch gevolg betreft welke conclusies kunnen worden getrokken gegeven de waarheid van bepaalde premissen. We beargumenteren dat aan dit probleem een optimaliseringsprobleem ten grondslag ligt. De vraag is dan wat het analoge speltheoretische probleem is. In ons voorstel wordt de waarheid van de premissen bepaald door individuele keuzes van de spelers ten aanzien van de variabelen die ze controleren. Dit geeft aanleiding het volgende probleem dat de kern vormt van Deel III: *welke conclusies mag men trekken ten aanzien van de uitkomst van een spel waarbij iedere speler een individuele verzameling van premissen poogt waar te maken door een strategische keuze te maken ten aanzien van de propositievariabelen waarover hij controle heeft?* Dit is een logische vraag waaraan een speltheoretisch probleem ten grondslag ligt.

Bij het bepalen van de formules die klassiek volgen uit een theorie Γ zijn semantisch gezien sommige valuaties van groter belang dan andere, namelijk die waarin alle formules in Γ waar zijn. Dit beroep op waarheid is in overeenstemming met het beeld van valuaties als mogelijke standen van zaken. Worden valuaties daarentegen voorgesteld als de uitkomsten van een strategisch spel dan ligt het meer voor de hand de aandacht te richten op die valuaties die zich onderscheiden vanwege hun speltheoretische eigenschappen.

Hoofdstuk 7 is gewijd aan de logische analyse van een gevolgrelatie tussen theorieën die gebaseerd is op de speltheoretische notie van een winnende strategie. Laat voor iedere theorie Γ en voor iedere deelverzameling Δ van propositievariabelen, $G(\Gamma, \Delta)$ een spel zijn dat de speler met controle over de propositievariabelen in Δ wint indien de uitkomst een valuatie is die alle formules in Γ waar maakt. Met betrekking tot een verzameling Δ van propositievariabelen, geldt een formule φ dan als een gevolg van een theorie Γ dan en slechts dan als φ het geval is in alle valuaties die resulteren als de speler met controle over Δ een winnende strategie speelt in het spel $G(\Gamma, \Delta)$. Deze notie van gevolg wordt onder meer voorzien van een bewijstheoretische basis in de vorm van een sequentencalculus, waarvan we de correctheid en volledigheid aantonen.

Ten einde meer recht te doen aan het speltheoretische karakter van de materie, wordt in de laatste twee hoofdstukken de scope van de analyse verruimd door meer expliciet spelen in beschouwing te nemen waarin meerdere spelers deelnemen die bovendien zowel tegengestelde als gemeenschappelijke belangen kunnen hebben. De rol van het concept van een winnende strategie wordt bovendien overgenomen door een variant van het speltheoretische oplossingsconcept Nash equilibrium.

Dientenbehoeve worden in Hoofdstuk 8 de zogenaamde *gedistribueerd evaluatiespelen* geïntroduceerd. Een gedistribueerd evaluatiespel geeft formeel invulling aan de notie van een spel waarin iedere speler de waarheid van een individuele theorie “zoveel mogelijk” poogt te bewerkstelligen door een strategische keuze te maken voor de propositievariabelen die hij controleert. Wat het exact betekent meer of minder van een theorie waar te maken, maken we formeel precies door iedere theorie Γ op eenduidige wijze te associëren met een reflexieve en transitieve relatie $\boldsymbol{\rho}(\Gamma)$ op de valuaties. Voor π een partitie van de propositievariabelen en $\{\Gamma_i\}_{i\in\pi}$ een familie van theorieën met π als indexverzameling, is $G(\{\Gamma_i\}_{i\in\pi})$ het gedistribueerde evaluatiespel waarbij iedere speler controle heeft over precies één blok van π en de preferenties van de speler met controle over het blok i van π gegeven zijn door de relatie $\boldsymbol{\rho}(\Gamma_i)$. Bij de evaluatie van gedistribueerde evaluatie spelen hanteren wij *maximum equilibrium* als oplossingsconcept. Maximum equilibrium is een conservatieve uitbreiding van Nash-equilibrium die ook van toepassing is op partiële preferentierelaties.

Gedistribueerde evaluatiespelen vormen een omvangrijke doch stricte subklasse van de volledige klasse van strategische spelen met valuaties als strategieprofielen. Hoofdstuk 8 sluit af met een resultaat dat betrekking heeft op de formele karakterisering van gedistribueerde evaluatiespelen in dit verband.

Hoofdstuk 9 staat in het teken van de een speltheoretische gevolgtrekkingsrelatie tussen families van theorieën die partities van de propositievariabelen als indexverzameling hebben. Deze speltheoretische notie van speltheoretisch gevolg laat zich evenwel het best begrijpen als een relatie tussen een familie theorieën $\{\Gamma_i\}_{i\in\pi}$ en een formule φ. Dan geldt φ als een speltheoretisch gevolg van $\{\Gamma_i\}_{i\in\pi}$ indien φ het geval is in alle valuaties die een maximum equilibrium zijn in het gedistribueerde evaluatiespel $G(\{\Gamma_i\}_{i\in\pi})$.

Voor een eenvoudig voorbeeld beschouwe men een propositionele taal met slechts twee propositievariabelen, a en b, en de partitie $\{\{a\}, \{b\}\}$. Stel dat de speler met controle over a de waarheid van de formule $a \wedge \neg b$ als doel heeft, terwijl de andere speler, met controle over b, het liefst de formule $\neg(a \vee b)$ waar ziet. De spelmatrix van deze situatie is weergegeven in figuur 4. Er zijn twee equilibria in dit spel, namelijk de valuaties $\{a\}$ en $\{a, b\}$. Aangezien de formule a waar is in beide equilibria, geldt a als een speltheoretisch gevolg van de familie $\{\{a \wedge \neg b\}_{\{a\}}, \{\neg(a \vee b)\}_{\{b\}}\}$.

Klassiek logisch gevolg is het randgeval van speltheoretisch gevolg waarbij één speler controle heeft over alle propositievariabelen. Meer in het algemeen blijkt spel-

	$\varnothing$	$\{b\}$
$\varnothing$	0, 1	0, 0
$\{a\}$	**1**, **0**	**0**, **0**

Figuur 4. De speler die rijen kiest heeft controle over de propositievariabele a; de andere speler, met de keuze tussen de kolommen, bepaalt de waarheidswaarde van b. De dikgedrukte uitkomsten zijn de maximum equilibria.

theoretisch gevolg inbedbaar in klassiek logisch gevolg en *vice versa*. Deze resultaten worden verkregen dankzij een verzamelingtheoretische karakterisering van speltheoretisch gevolg. Hierbij wordt een beroep gedaan op benaderingsoperaties die bekend zijn uit de theorie van de zogenaamde *rough sets*. Dit maakt dat speltheoretisch gevolg formele eigenschappen als monotonie en beslisbaarheid overerft van klassiek gevolg.

Curriculum Vitae

Bernhard Paul Harrenstein werd op 27 juli 1970 te Bussum geboren. In juni 1988 behaalde hij het VWO diploma aan het Willem de Zwijger College eveneens te Bussum. Na een jaar aan de University of East Anglia (UEA) in Norwich in Engeland te hebben doorgebracht en het CPE (Certificate of Proficiency in English) te hebben behaald, begon hij in september 1989 aan de studie Wijsbegeerte aan de Universiteit van Amsterdam. Op 28 augustus 1998 studeerde hij daar *cum laude* af op een scriptie over modale predicaatlogica met de titel *From a Modal Point of View*.

Met het onderzoek waarvan deze dissertatie een verslag doet, werd in januari 1999 een aanvang genomen toen hij in dienst trad van de Universiteit Utrecht als Assistent in Opleiding bij het Instituut voor Informatica en Informatiekunde in groep Intelligente Systemen van Prof. Dr. John-Jules Meyer. Het onderzoek werd verricht in het kader van het Collective Agent-Based Systems (CABS) project van de Technische Universiteit Delft. In februari 2003 liep zijn assistentschap officieel af, waarna hij zich heeft toegelegd op het voltooien van zijn dissertatie aan het Instituut voor Informatica aan de Universiteit Utrecht.

Titles in the SIKS Dissertation Series

1998

Johan van den Akker, *DEGAS - An Active, Temporal Database of Autonomous Objects*, CWI, 1998-1

Floris Wiesman, *Information Retrieval by Graphically Browsing Meta-Information*, UM, 1998-2

Ans Steuten, *A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective*, TUD, 1998-3

Dennis Breuker, *Memory versus Search in Games*, UM, 1998-4

E.W. Oskamp, *Computerondersteuning bij Straftoemeting*, RUL, 1998-5

1999

Mark Sloof, *Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products*, VU, 1999-1

Rob Potharst, *Classification using decision trees and neural nets*, EUR, 1999-2

Don Beal, *The Nature of Minimax Search*, UM, 1999-3

Jacques Penders, *The practical Art of Moving Physical Objects*, UM, 1999-4

Aldo de Moor, *Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems*, KUB, 1999-5

Niek J.E. Wijngaards, *Re-design of compositional systems*, VU, 1999-6

David Spelt, *Verification support for object database design*, UT, 1999-7

Jacques H.J. Lenting, *Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation.*, UM, 1999-8

2000

Frank Niessink, *Perspectives on Improving Software Maintenance*, VU, 2000-1

Koen Holtman, *Prototyping of CMS Storage Management*, TUE, 2000-2

Carolien M.T. Metselaar, *Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief*, UVA, 2000-3

Geert de Haan, *ETAG, A Formal Model of Competence Knowledge for User Interface Design*, VU, 2000-4

Ruud van der Pol, *Knowledge-based Query Formulation in Information Retrieval*, UM, 2000-5

Rogier van Eijk, *Programming Languages for Agent Communication*, UU, 2000-6

Niels Peek, *Decision-theoretic Planning of Clinical Patient Management*, UU, 2000-7

Veerle Coup, *Sensitivity Analyis of Decision-Theoretic Networks*, EUR, 2000-8

Florian Waas, *Principles of Probabilistic Query Optimization*, CWI, 2000-9

Niels Nes, *Image Database Management System Design Considerations, Algorithms and Architecture*, CWI, 2000-10

Jonas Karlsson, *Scalable Distributed Data Structures for Database Management*, CWI, 2000-11

2001

Silja Renooij, *Qualitative Approaches to Quantifying Probabilistic Networks*, UU, 2001-1

Koen Hindriks, *Agent Programming Languages: Programming with Mental Models*, UU, 2001-2

Maarten van Someren, *Learning as problem solving*, UvA, 2001-3

Evgueni Smirnov, *Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets*, UM, 2001-4

Jacco van Ossenbruggen, *Processing Structured Hypermedia: A Matter of Style*, VU, 2001-5

Martijn van Welie, *Task-based User Interface Design*, VU, 2001-6

Bastiaan Schonhage, *Diva: Architectural Perspectives on Information Visualization*, VU, 2001-7

Pascal van Eck, *A Compositional Semantic Structure for Multi-Agent Systems Dynamics*, VU, 2001-8

Pieter Jan 't Hoen, *Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes*, RUL, 2001-9

Maarten Sierhuis, *Modeling and Simulating Work Practice BRAHMS: a multiagent modeling and simulation language for work practice analysis and design*, UvA, 2001-10

Tom M. van Engers, *Knowledge Management: The Role of Mental Models in Business Systems Design*, VUA, 2001-11

2002

Nico Lassing, *Architecture-Level Modifiability Analysis*, VU, 2002-01

Roelof van Zwol, *Modelling and searching web-based document collections*, UT, 2002-02

Henk Ernst Blok, *Database Optimization Aspects for Information Retrieval*, UT, 2002-03

Juan Roberto Castelo Valdueza, *The Discrete Acyclic Digraph Markov Model in Data Mining*, UU, 2002-04

Radu Serban, *The Private Cyberspace Modeling Electronic Environments inhabited by Privacy-concerned Agents*, VU, 2002-05

Laurens Mommers, *Applied legal epistemology; Building a knowledge-based ontology of the legal domain*, UL, 2002-06

Peter Boncz, *Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications*, CWI, 2002-07

Jaap Gordijn, *Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas*, VU, 2002-08

Willem-Jan van den Heuvel, *Integrating Modern Business Applications with Objectified Legacy Systems*, KUB, 2002-09

Brian Sheppard, *Towards Perfect Play of Scrabble*, UM, 2002-10

Wouter C.A. Wijngaards, *Agent Based Modelling of Dynamics: Biological and Organisational Applications*, VU, 2002-11

Albrecht Schmidt, *Processing XML in Database Systems*, UVA, 2002-12

Hongjing Wu, *A Reference Architecture for Adaptive Hypermedia Applications*, TUE, 2002-13

Wieke de Vries, *Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems*, UU, 2002-14

Rik Eshuis, *Semantics and Verification of UML Activity Diagrams for Workflow Modelling*, UT, 2002-15

Pieter van Langen, *The Anatomy of Design: Foundations, Models and Applications*, VU, 2002-16

Stefan Manegold, *Understanding, Modeling, and Improving Main-Memory Database Performance*, UVA, 2002-17

2003

Heiner Stuckenschmidt, *Onotology-Based Information Sharing In Weakly Structured Environments*, VU, 2003-1

Jan Broersen, *Modal Action Logics for Reasoning About Reactive Systems*, VU, 2003-02

Martijn Schuemie, *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*, TUD, 2003-03

Milan Petkovic, *Content-Based Video Retrieval Supported by Database Technology*, UT, 2003-04

Jos Lehmann, *Causation in Artificial Intelligence and Law - A modelling approach*, UVA, 2003-05

Boris van Schooten, *Development and specification of virtual environments*, UT, 2003-06

Machiel Jansen, *Formal Explorations of Knowledge Intensive Tasks*, UvA, 2003-07

Yongping Ran, *Repair Based Scheduling*, UM, 2003-08

Rens Kortmann, *The resolution of visually guided behaviour*, UM, 2003-09

Andreas Lincke, *Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture*, UvT, 2003-10

Simon Keizer, *Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks*, UT, 2003-11

Roeland Ordelman, *Dutch speech recognition in multimedia information retrieval*, UT, 2003-12

Jeroen Donkers, *Nosce Hostem - Searching with Opponent Models*, UM, 2003-13

Stijn Hoppenbrouwers, *Freezing Language: Conceptualisation Processes across ICT-Supported Organisations*, KUN, 2003-14

Mathijs de Weerdt, *Plan Merging in Multi-Agent Systems*, TUD, 2003-15

Menzo Windhouwer, *Feature Grammar Systems - Incremental Maintenance of Indexes to Digital Media Warehouses*, CWI, 2003-16

David Jansen, *Extensions of Statecharts with Probability, Time, and Stochastic Timing*, UT, 2003-17

Levente Kocsis, *Learning Search Decisions*, UM, 2003-18

2004

Virginia Dignum, *A Model for Organizational Interaction: Based on Agents, Founded in Logic*, UU, 2004-01

Lai Xu, *Monitoring Multi-party Contracts for E-business*, UvT, 2004-02

Perry Groot, *A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving*, VU, 2004-03

Chris van Aart, *Organizational Principles for Multi-Agent Architectures*, UVA, 2004-04

Viara Popova, *Knowledge discovery and monotonicity*, EUR, 2004-05

Bart-Jan Hommes, *The Evaluation of Business Process Modeling Techniques*, TUD, 2004-06

Elise Boltjes, *Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes*, UM, 2004-07

Joop Verbeek, *Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politiële gegevensuitwisseling en digitale expertise*, UM, 2004-08

Martin Caminada, *For the Sake of the Argument; explorations into argument-based reasoning*, VU, 2004-09

Suzanne Kabel, *Knowledge-rich indexing of learning-objects*, UVA, 2004-10

Michel Klein, *Change Management for Distributed Ontologies*, VU, 2004-11

The Duy Bui, *Creating emotions and facial expressions for embodied agents*, UT, 2004-12

Wojciech Jamroga, *Using Multiple Models of Reality: On Agents who Know how to Play*, UT, 2004-13

Paul Harrenstein, *Logic in Conflict. Logical Explorations in Strategic Equilibrium*, UU, 2004-14