20

MODAL LOGIC FOR GAMES AND INFORMATION

Wiebe van der Hoek and Marc Pauly

1 INTRODUCTION

Game-theoretic ideas have a long history in logic. A game-theoretic interpretation of quantification goes back at least to C.S. Peirce, and game-theoretic versions for all essential logical notions (truth in a model, validity, model comparison) have been developed subsequently. The connections between game theory and *modal* logic, on the other hand, have been developed only more recently. At the time of writing, the area is still an active one, so active in fact, that one might even argue that it is too early to be included in a handbook of modal logic. In spite of this concern, we believe that the volume of research falling into the category of modal logic and game theory justifies a survey. Our attempt in this chapter is to put some structure on the various strands of research, to create an organisation which highlights what we consider to be the essential lines of research. As with all endeavours of this sort, one cannot include all the research considered valuable, and we are well aware that our choice of topics also reflects our own interests and expertise.

Game theory has developed a wealth of interesting ideas for describing interactions which may involve a conflict of interest. So far, the logic community has restricted its attention to relatively few of these, mainly studying 2-player extensive games of perfect information which are strictly competitive (even win/lose). In fact, even this restricted class of games has turned out to be extremely rich, as set theory and computer science can testify. Still, given this traditionally narrow focus of logic, it is encouraging to see that more recent work in logic has extended the game-theoretic toolbox considerably, introducing, e.g., cooperative game theory, imperfect information and games involving more than 2 players. While even this can only be a beginning, by now the game-theoretic ideas used in logic certainly go beyond the intuitively natural idea of a winning strategy, and hence we will start our chapter with a section explaining the necessary background in game theory.

Having packed our game-theoretic baggage, our tour starts in Section 3 with new sights of a familiar landscape, possibly the most natural way to link games to modal logic. Game trees can be viewed as Kripke models, where the possible moves are modeled by an accessibility relation and additional information about payoffs and turn taking are encoded by propositional atoms. Structural equivalence notions such as bisimulation then turn into game equivalence notions, and we can investigate extensions of the modal language which can capture game-theoretic solution concepts such as the subgame-perfect equilibrium.

Leaving the first three sections together with the concluding Section 12 aside, the rest of this chapter can be divided into three reading tracks, epistemic logics (Sections 4-8), game logic (Section 9), and coalition logics (Sections 10-11), which can be pursued independently. The epistemic logic track describes approaches based on (dynamic) epistemic logic for dealing with imperfect information in games. The section on game logic describes an extension of Propositional Dynamic Logic for reasoning about games, focusing on operations for combining games like programs. The coalition logic track, finally, discusses a range of logics developed for modeling coalitional power in games, possibly also adding temporal or epistemic operators.

In Section 4 we introduce a widely accepted logic for knowledge in the area of games, where the assumptions impose that players are fully introspective and their knowledge is veridical. The more interesting properties are to be found when studying knowledge of *groups* of players though, with *common knowledge* being the main and most intriguing notion in this palette. Section 5 introduces *interpreted systems*, a dominant paradigm in computer science to deal with knowledge and time, where time corresponds with steps in a protocol, or, for our purposes, indeed a game.

Such epistemic notions as introduced in Section 4 and 5 play an important role in games of imperfect information, and, as some major results in early game theory indicate, even beyond that (see Section 8). For instance, a procedure that yields a Nash equilibrium in extensive games, called *backward induction*, finds its justification in the assumption about common knowledge about rationality of the players. We will see, however, that nowadays epistemologists put the need for the inherently infinite conjunctions that come along with common knowledge into perspective and a non-trivial analysis of games can be given without falling back on such strong assumptions.

Where the emphasis in Section 4 is on the knowledge that the players have *about* the game, in Sections 7 and 8 the emphasis will be on how knowledge evolves *during* a game. Hiding one's knowledge can be beneficial for a player within a game, but revealing his ignorance can also be disastrous, and may benefit other players. Moreover, in certain games (like *Cluedo* and many card games) the winning conditions are purely epistemic: the game ends in a win for that player who is the first who happens to know some crucial information.

The dynamic logic of games discussed in Section 9 takes Propositional Dynamic Logic as its starting point. By a change in the underlying semantics, programs become 2-player games which can be combined using the old program operations of sequential composition, test, etc. Besides these program operations, a new duality operator is added which interchanges the roles of the players. Using this new operator, nondeterministic choice splits into two versions depending on which player makes the choice. A typical formula $[(a \cap b); (a \cup c)]p$, for instance, expresses that player 2 has a strategy for achieving p in the game where first, player 1 chooses between a and b, and then player 2 chooses between a and c (for details, see Section 9).

Section 10 introduces Coalition Logic, a basic modal logic for reasoning about the ability of groups in different kinds of games. For a set of individuals C, the formula $[C]p$ expresses that the members of C have a joint strategy for achieving p at the next stage of game. In Section 11, this language is extended to Alternating-time Temporal Logic (ATL) by adding operators for talking about the long-term future, where we can state, e.g., that a coalition can achieve p eventually. ATL is a game-theoretic generalisation of Computation Tree Logic (CTL), with applications in the formal verification of multi-agent systems. Further extensions of ATL to ATL*, the alternating μ-calculus and ATEL are presented. ATEL adds epistemic operators to ATL in order to express, e.g., that a coalition has a strategy for getting an agent to know something eventually.

While our focus in this chapter is on *modal logic for games*, there are also many *games for modal logic*. The reader interested in this reverse connection is referred to other chapters of this handbook (i.e., Chapter 12 and 17) for more details. The similarity between programs and games and the relevance of epistemic and temporal issues, on the other hand, suggest that modal logic may provide an interesting new perspective on games, and it is this perspective we would like to present in this chapter.