

Logics for Analyzing Games

Johan van Benthem and Dominik Klein

In light of logic’s historical roots in dialogue and argumentation, games and logic are a natural fit. Argumentation is a game-like activity that involves taking turns, saying the right things at the right time, and, in competitive settings, has clear pay-offs in terms of winning and losing. Pursuing this connection, specialized logic games have already been used in the middle ages as a tool for logic training (Hamblin, 1970). The modern area augmented this picture with formal dialogue games as foundation for logic, relating winning strategies in argumentation to cogent proofs (Kamlah and Lorenzen, 1984). Today, connections between logic and game theory span across a great number of different strands, involving the interface with game theory, but also linguistics, computer science, and further fields.

Themes from the extensive and growing area surrounding logic and games occur in various entries of this Encyclopedia, in particular on [uses of games in logic](#), [epistemic foundations of game theory](#), [formal approaches to social procedures](#), [logics for analyzing powers of agents](#), and [game semantics for programming and process languages](#). These entries differ in their emphasis, which may be on logic, game theory, or foundations of computer science. The present entry is concerned with logics for analyzing games, broadly speaking. It makes reference to other perspectives in this Encyclopedia where relevant.

Contents

1	Overview	3
1.1	Logic of games	3
1.2	Logic and game theory	4
1.3	Computation and agency	7
1.4	Games in logic	7
1.5	Probability	8
1.6	Zooming in	8
2	Game structure and game logics	9
2.1	Levels of representation	9
2.2	Invariance relations between games	10
2.3	Languages matching invariance relations	12
2.4	Modal logic of extensive games	13

2.5	Modal neighborhood logic for powers	14
2.6	Modal logic for strategic games	15
2.7	Strategies as logical objects	16
2.8	Simultaneous moves and imperfect information	17
2.9	Game algebra and dynamic logic of computation	17
2.10	Special topics	19
3	The Nature of Players	20
3.1	Preference and equilibria	20
3.2	Preference logics for games	20
3.3	Backward Induction in extensive form games	22
3.4	Iterated removal of dominated strategies	23
3.5	Goals	24
3.6	Knowledge, belief and limits of information	24
3.7	Higher order uncertainty and type spaces	29
3.8	Reasoning, bounded agency and player types	31
3.9	Thin and thick models for players	32
3.10	Special Topics	33
4	Analyzing Play	33
4.1	Game solution and pregame deliberation	33
4.1.1	Backward Induction via public announcement	34
4.1.2	Backward Induction via belief revision	35
4.1.3	Iterated removal of strictly dominated strategies	37
4.2	Information flow, knowledge, and belief during play	40
4.2.1	Epistemic update and imperfect information	40
4.2.2	Belief revision and Forward Induction	42
4.2.3	Post-game rationalization	44
4.2.4	Play in a long-term temporal perspective	44
4.3	Conclusion: theory of play	46
4.4	Further directions	46
5	Interfacing Logic and Probability in Games	48
5.1	Probabilities and beliefs	48
5.2	From logic to probability and back	49
5.3	Updating and tracking probabilities	50
5.4	Specializing to games	51
5.5	Further challenges for a logical approach	52
6	Gamification	53
6.1	Logic games	53
6.2	Recent logic-related games	54
6.3	Special topics	57
7	Summary and Further Directions	57
8	Bibliography	59

1 Overview

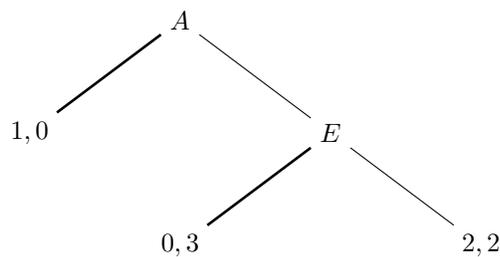
The present entry provides a comprehensive survey of logics for analyzing games, arranged under a number of unifying themes and perspectives. Also, occasional connections are made with other strands at the interface of logic and games covered elsewhere in this Encyclopedia. This overview section is a brief tour d’horizon for topics that will return in more detail later on.

1.1 Logic of games

Specific games stand for significant recurrent patterns of social interaction. In a perspective called ‘logic of games’, notions and results from logic are used to analyze the structure of various games. In fact, much classic reasoning about games involves notions that are familiar from logic.

Example Game solution reasoning.

Consider the following game tree for two players A, E , with turns marked, and with pay-offs written with the value for A first. Alternatively, these values can be interpreted as encoding players’ qualitative (or ordinal) preferences between outcomes.



Here is how players might reason. At her turn, E faces a standard decision problem, with two available actions and the outcome of action *left* better for her than that of *right*. So she will choose left. Knowing this, A expects that his choosing *right* will give him outcome 0, while going *left* gives him outcome 1, so he chooses left. As a result, both players are worse off than they would have been, had they played *right/right*. The reasoning in this scenario, in short, leads to an outcome that is not Pareto-optimal.

The example raises the question just why players should act this way, and whether, say, a more cooperative behavior could also be justified. An answer obviously depends on the players’ information and style of reasoning. Here it becomes of interest to probe the structure of the example. Looking more closely, many notions are involved in the above scenario: actions and their results, players knowledge about the structure of the game, their preferences about its results, but also how they believe the game will proceed. There are even counterfactual conditionals in the background, such as A ’s explaining his choice afterwards by saying that “if I had played *right*, then E would have played *left*”. These notions, moreover, are entangled in subtle ways. For instance, A does not choose *left* because it dominates *right* in the standard sense of always being better for him, but rather because *left* dominates *right*

according to his beliefs. How these beliefs are formed, in turn, depends on many other features of the game, including the nature of the players.

In short, even a very simple game like the one discussed brings together large parts of the agenda of philosophical logic in one very concrete setting. This entry will zoom in on the aspects mentioned here, with Section 3 dedicated to players' preferences and beliefs, while Section 4 addresses reasoning styles and the dynamics of attitudes as the game proceeds.

The analysis is structured by a few broad distinctions. Intuitively, games involve several phases that involve logic in different ways: deliberation prior to the game, as many game-theoretic solution concepts are in fact deliberation procedures that create initial expectations about how a game will go on. Observation and belief revision during game play, including reactions to deviations from prior expectations. And finally, post-game analysis, say, to settle what can be learnt from a defeat, or to engage in spin about one's performance. Moreover all this can be considered in two modes, assuming either a first-person participant or a third-person observer view of games and play.

1.2 Logic and game theory

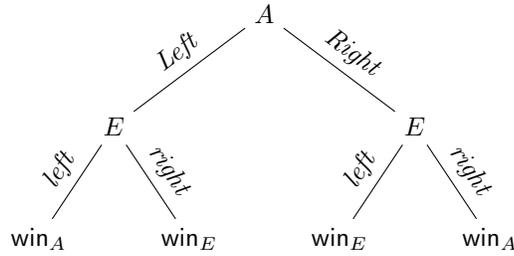
In many of the above topics, logics meets game theory. One such interface area is **epistemic game theory** where game play and solution concepts are analyzed and justified in light of various assumptions about players and their epistemic states, such as common knowledge or common belief in rationality. Epistemic game theory may be viewed as a joint offspring of logic and game theory, a form of progeny which constitutes a reliable sign of success of an interdisciplinary contact.

There are also other viable logical perspectives. In particular, one can look at game theory the way mathematical logicians look at any branch of mathematics. Following the style of the famous Erlangen Program, one can discuss the structures studied in that field and look for structural invariance relations and matching logical languages. Game theory is rich in structure, as it has several different natural notions of invariance. The tree format of extensive games offers a detailed view of what happens step by step as players make their moves, whereas the matrix format of strategic form games offers a high-level view that centers on outcomes. Yet other formats, to be discussed below, focus on players' control over the various outcomes. All these different levels of game structure come with their own logical systems, as will be detailed in Section 2. Moreover, these different logics do not just provide isolated snapshots: they can be related in a systematic manner.

In this way, the usual logical techniques can be brought to bear. For instance, formal languages can express basic properties of games, while model-checking techniques can determine efficiently whether these hold in given concrete games (cf. [Clarke et al., 1999](#)).

Example Winning strategies.

Consider the following game tree, with move relations for both players, and propositional letters win_i marking winning positions for player i .



Clearly, player E has a winning strategy against player A , i.e. a recipe that guarantees her to win, no matter what A does. This is expressed by a modal formula capturing exactly the right dynamics:

$$[\text{move}_A]\langle \text{move}_E \rangle \text{win}_E$$

Here $[\text{move}_A]$ is the universal modality “for all moves by player A ”, and $\langle \text{move}_E \rangle$ is the existential modality “for some move by player E ”. This two-step modality- or quantifier-based response pattern is typical for strategic powers of players in arbitrary games, as it captures the essence of sequential interaction. Crucially, logical laws can acquire game-theoretic import. For instance, the law of Excluded Middle applied to the above formula yields:

$$[\text{move}_A]\langle \text{move}_E \rangle \text{win}_E \vee \neg [\text{move}_A]\langle \text{move}_E \rangle \text{win}_E$$

or in a logically equivalent formulation:

$$[\text{move}_A]\langle \text{move}_E \rangle \text{win}_E \vee \langle \text{move}_A \rangle [\text{move}_E] \neg \text{win}_E$$

In two-step games like the above, where exactly one player wins (i.e. $\text{win}_A \leftrightarrow \neg \text{win}_E$), the latter formula expresses that either player E or player A has a winning strategy. More generally, this disjunctive assertion is a special case of Zermelo’s theorem, stating that every finite full information game is determined.

Having established the connection to logical languages, further model-theoretic themes can be applied fruitfully to games. Language based reasoning allows, for instance, to examine the preservation of properties between different games, based on the exact syntactic shape of their definition. Besides, logical syntax also supports logical proof theory. Hence, the latter’s rich pool of proof calculi may help to analyze basic results in game theory. This entry illustrates major recurring patterns of reasoning about interaction that come to light in this way.

Game theory also has a further natural level of representation, suppressing details of local moves and choices. The most familiar format for this are games in *strategic form*. In the simplest case of only two players, these correspond to a two-dimensional matrix, with rows standing for some player’s strategies, and columns for the other’s. Individual cells of such matrix hence correspond to the different possible *strategy profiles* of the game. Typically, all cells are labelled with information about the *outcomes* resulting from playing the corresponding strategies against one another. This labelling specifies players’ attitudes to outcome in terms of pay-offs, more abstract utilities, or simply markers for players’ preferences orders among outcomes.

Strategic form games, too, can model significant social scenarios. Here is an illustration from the philosophical literature on the evolution of social behavior.

Example The following game in matrix form is the *Stag Hunt* of Skyrms (2004), going back to ideas of David Hume. It serves as a metaphor for the social contract.

		E	
		H	S
A	H	1, 1	1, 0
	S	0, 1	2, 2

Each agent must decide between pursuing their own little project, hunting a hare, or joining in a larger collective endeavor, hunting stag. The former gives a moderate but guaranteed income, no matter what others do. The collective endeavor, on the other hand, can only succeed if all contribute, in which case everybody receives a high profit. If, however, some do not join, all contributions are lost and no contributor receives anything. In the corresponding strategic form game, all players have to decide on what to do in parallel, without knowing the actions of the others.

The Stag Hunt game has two pure strategy Nash equilibria: *every contributes*, and *nobody contributes*. Which of these stable outcomes ensues will crucially depend on the players' reasoning, their expectations about each other, and perhaps even further information stemming, for instance, from pre-game communication.

Clearly, analyzing strategic games involves agentive information, reasoning and expectations. All these aspects have tight connections to logic. Viewing outcomes as possible worlds, three relevant relations emerge between these. Within the matrix above, relating all cells in the same row fixes a unique choice already made by the row player A , while leaving E 's move completely open. In short, each horizontal row lists all possible choices of the column player E which A has to take into account. The corresponding modality may hence be said to describe A 's knowledge about the outcomes of the game given his choice. Still assuming the row player's perspective, relating cells vertically rather than horizontally corresponds to A 's freedom of choice among his available strategies. Of course, one could also assume player E 's perspective instead, viewing the horizontal direction as E 's freedom of movement, while the vertical directions captures her epistemic uncertainty.

Thus, a bimodal logic arises for matrix games with laws such as

$$K_E K_A \varphi \leftrightarrow K_A K_E \varphi,$$

capturing the grid structure of matrix games. For more than two players, this logic gains some additional options and subtleties to be discussed in Section 2.6.

The crucial third relation is that of player's preferences among outcomes. These, again, have matching modalities, now taken from preference logic (Hansson, 2002). With the help of some auxiliary devices, the three modalities can define the central game-theoretic notion of a Nash equilibrium (Harrenstein, 2004; van der Hoek and Pauly, 2007).

Logics for matrix games differ from those for extensive games, as grids behave quite differently from trees in terms of complexity. Yet, both fall under the same general methodology. Towards a common understanding, one might view the logic of matrix games as capturing the basic laws of parallel, rather than sequential action.

1.3 Computation and agency

Philosophical logic and mathematical logic are not the only illuminating perspectives on games. A third relevant viewpoint is that of computational logic. In modern computation, the paradigm is no longer single Turing machines but interacting systems of multiple processors. These processors may cooperate, but they might also compete for resources. In general, hence, it is useful to study multiple agents engaging in computation, be it within human, artificial or mixed societies. Though doing so, games become a natural model for computation, too. In fact, games are rich multi-agent systems where agents process information, communicate, and engage in actions, all driven by their respective preferences and goals. In the converse direction, computer science themes such as complexity and algorithmics have entered game theory, resulting in the area of computational game theory (Nisan et al., 2007). For a richer survey of computational logics of agency and games, see van der Hoek and Pauly (2007) and Shoham and Leyton-Brown (2008). The present entry contains occasional links to computation. These are especially prominent for reasoning about temporally extended games and their strategies (Sections 4.2, 4.4) and in the context of gamification (Section 6), where games are explored as a novel semantics for classical logical systems.

1.4 Games in logic

Finally, recall the start of this section, but with reverse perspective: instead of asking what logic can do for games, ask what games can do for logic. Argumentation and dialogue are basic notions for logic. Both can be studied using techniques and results from game theory (Lorenzen and Lorenz, 1978; Hamblin, 1970). In this perspective, logical validity of consequence rests on there being a winning strategy for a Proponent claiming the conclusion against an Opponent granting the premises in a game where moves are regulated by the logical constants. Many games have found uses in modern logic since the 1950s, with Ehrenfeucht-Fraïssé games for model comparison being a paradigmatic example. Besides these, also semantic verification or model construction can be cast as natural logic games.

This raises an intricate issue within in the philosophy of logic, concerning the nature of logic and in particular that of logical constants. A ‘weak thesis’ would hold that games constitute a natural technique for analyzing logical notions, as well as a didactic tool for teaching logic that appeals directly to vivid intuitions. Parts of the literature, however, also defend a ‘strong thesis’, suggesting that the primary semantics of certain logical systems may be procedural and game-theoretic, rather than denotational in a standard sense. This perspective, sometimes called ‘logic as games’, occurs in some attractive semantics for first-order languages (Hintikka and Sandu, 1997), as well as in game semantics for programming languages.

The theme of logic as games will appear only briefly in the present entry, which is mainly directed toward logics of games. Section 6 will discuss which questions arise from joining both perspectives on the interface of logic and games.

As it happens, the logic-as-games perspective is of broader relevance. Logic games were originally designed for particular tasks inside logic. Yet, taken to reality, they can help analyze or streamline actual lines of argumentation. As such they may be compared to designed parlor games that challenge reasoning skills. A game like *Clue* involves an intriguing mix of logical deduction, new information from drawing cards or public observation of moves, but also private communication acts by players (van Ditmarsch, 2000). Other parlor games, such as *Nine Men Morris* (Gasser, 1996) are graph games (Grädel et al., 2002) with added chance moves that serve to diminish the risk of finding a repeatable simple strategy on the fixed board. The logical study of playable designed games for bounded agents, and the design of new such games, is a natural sequel to this entry (cf. van Benthem and Liu, 2018).

1.5 Probability

Game theory may be understood as generalized interactive decision theory. A major vehicle for the latter, just as for standard decision theory, is probability theory. Within games, probability can assume many roles. It may, for instance, express players' degrees of belief quantitatively, but it can also enrich the space of actions with mixed strategies, thereby laying the ground for general equilibrium results. Probability can even play a role in the very definition of certain important games, especially in evolutionary game theory (Osborne and Rubinstein, 1994). In this entry, probability is only mentioned in passing. Section 5, however, maps some combinations of logic and probability that are suggested by the study of games.

1.6 Zooming in

Games have a natural interface with logic in all its varieties, including mathematical, philosophical, and computational logic. In one direction of contact, logic can provide new abstract notions underneath game theory. Conversely, game-theoretic notions can also serve to enrich logical analysis. The present entry mainly concentrates on the first of these directions, the use of logic for analyzing games. It does so mostly from a semantic perspective, the dominant paradigm so far in the area. Though proof-theoretic approaches will be mentioned occasionally. The sections to follow elaborate on this theme along several dimensions. Specific perspectives include logics for game structures (Section 2), logical analysis of the nature of players (Section 3) and of the process of game play (Section 4). Additional spotlights are put on the relationship between logic and probability in the context of games (Section 5) and the endeavor of Gamification (Section 6). Each section forms a free-standing exposition, which results in some unavoidable, and perhaps useful, overlap. Throughout the exposition, some familiarity is assumed with the basic concepts of logic and game theory. In particular, notions of game theory left unexplained here can be found in the corresponding entry and in Leyton-Brown and Shoham (2008).

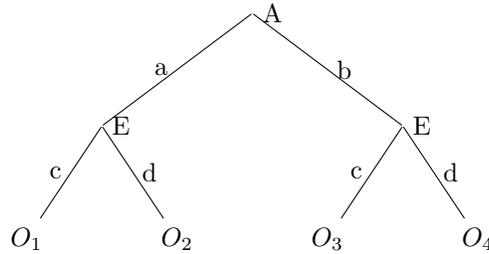
2 Game structure and game logics

This first zoom in section focuses on game structures in a narrow sense. *Game forms* leave aside agents and notions typical for these, such as preferences or information. Players, as well as the temporal progression of play will be added in later sections. Even so, there is a good deal of structure in game forms to be studied by logical techniques.

2.1 Levels of representation

The starting point of any logical analysis is to fix its perspective on games. This section will review several major candidates for doing so, starting with the two most prominent perspectives. The first of these makes the temporal structure of a game explicit, representing it as a tree in the standard mathematical sense.

Example A two-player extensive form game.



A game in *extensive form* is a tree where each non-terminal node or *state* specifies which player is to move next, while edges correspond to the players' possible moves. The leaves of the tree, finally, denote the possible outcomes O of game play. There are many possible variations on these stipulations for states and moves, but they do not affect the essentials of a logical analysis.

The second major perspective on games emphasizes the players' available strategies. Suppressing all information about temporal structure, a game in *strategic form* yields the matrix pictures known from game theory. In its classic interpretation, a game in strategic form represents a set of players that each select a complete strategy for the entire game without knowledge of the other players' choices. Each strategy profile, i.e. combination of one strategy per player, then induces an outcome O_i . The motivation for this structure might seem complex on first sight. Yet, it can also be viewed as something quite simple: a one-step game with parallel rather than sequential moves, which is the simplest case of simultaneous action.

Example A two-player strategic form game.

		E	
		a	b
A	c	O_1	O_2
	d	O_3	O_4

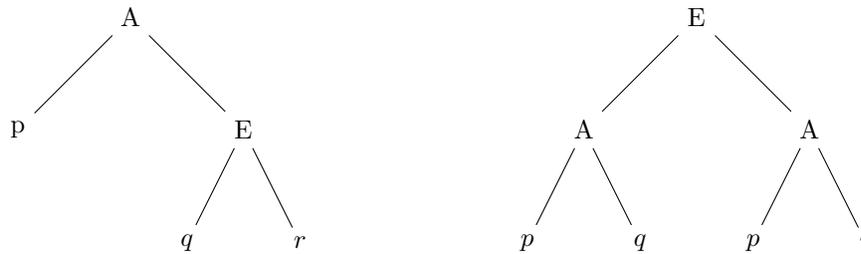
Extensive and strategic forms differ in their focus. The former emphasize the sequential temporal structure of a game, while the latter highlights strategy choice prior to play. One can freely switch between both when appropriate to the purpose at hand. Besides these two, there are other natural dimensions, highlighting players' powers for influencing outcomes (cf. Section 2.5) or players' information about the game (cf. Section 3.6).

Remark While all examples so far concerned two-player games, no such restriction is needed. Both extensive games and strategic form games work for any number of players, although occasional subtleties may occur. A few will be mentioned below. Moreover, with more players, the possible coalitions enter the picture, a topic that will not be treated in this entry. Finally, selected aspects of agency may sometimes enter through the back door. Many scenarios in real life contain external chance events outside of any player's control, such as a roll of a die, weather conditions, or technical malfunctions. Such factors can usually be incorporated into a logical analysis by admitting Nature as additional player.

2.2 Invariance relations between games

With different ways of representing a game at hand, there is a natural follow up question concerning equivalence. Given two game structures, when are they representations of the same underlying game? The answer is that it very much depends on what aspects one is interested in.

Example The same game, or not?



Consider the two game forms above. If one cares about exact sequences of moves or the choices players have along the way, these games are different. The game to the left has A move first, while E begins in the game on the right. In the game on the left, A may face a choice between p and q . This cannot happen on the right.

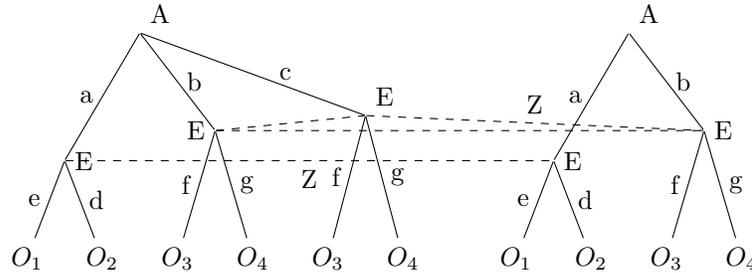
Caring about exact moves as done here constitutes a fine-grained perspective on games. There are others. When focusing on players' powers for bringing about certain outcomes, for instance, the analysis changes. In the game on the left, A has a strategy (playing *left*) that ensures the game to end up in an outcome satisfying p , and one (playing *right*) that restricts possible outcomes to those satisfying $q \vee r$. With this second strategy, the further choice which of q, r gets realized is left up

to player E . Also the second player, E , has two strategies in the game on the left, one (playing *left*) ensuring the outcome to satisfy $p \vee q$, the other (‘playing *right*’) guaranteeing that the outcome satisfies $p \vee r$.

Performing the same calculations for the game on the right, virtually the same player powers emerge. More precisely, A ’s uniform strategies *left-left* and *right-right* yield p and $q \vee r$ respectively, exactly the same powers as in the left game. The two remaining strategies *left-right* and *right-left* yield $p \vee q$ and $p \vee r$, both of which are mere weakenings of A ’s power to achieve p . Thus, at the level of players’ powers, the above two game forms should be considered the same.

As this example illustrates, there are several legitimate ways of comparing games. When taking a fine-grained focus on the internal structure of a game, a natural candidate is the notion of a *bisimulation* (cf. Blackburn et al., 2001). A bisimulation $Z \subseteq G_1 \times G_2$ relates states of two game forms G_1 and G_2 subject to four conditions: States m and n may only be related when *i*) the same player is to move in m and n , *ii*) m and n do not differ in any of their basic local properties, while *iii a*) whenever there is an available move of type a in G_1 leading to a state m' , there is a matching available move of type a in G_2 leading to a state n' with $m'Zn'$, and vice versa *iii b*) whenever there is a move in G_2 that leads to a state n' , there there is a move of the same type in G_1 leading to a state m' with $m'Zn'$.

Example A bisimulation between games.



This particular notion of bisimulation is not the only invariance that makes sense for games. A more coarse-grained perspective, for instance, might not distinguish moves by their particular action types, but merely by which player is to perform them. A corresponding bisimulation can be defined by omitting references to particular action types in conditions *iii a*) and *iii b*) above.

Further notions of bisimulation take an even coarser perspective on the games’ move structure, for instance by allowing to contract zones where the same player moves several times in a row. Finally, dropping all information about players and their choices, games can be compared by the sequences of moves they admit. This purely observational notion, known as *trace equivalence* in computation, may, however, be less relevant in the context of games. An alternative approach to coarsening focusses on the players’ powers to control outcomes, cf. Section 2.5.

While most notions of invariance discussed so far related to extensive form games, a similar style of analysis applies to games in strategic form. Pacuit et al. (2011)

define modal bisimulations that connect outcome states of different matrices, and apply bisimulation’s back-and-forth conditions to the relevant relations of players’ choice, freedom, and preference.

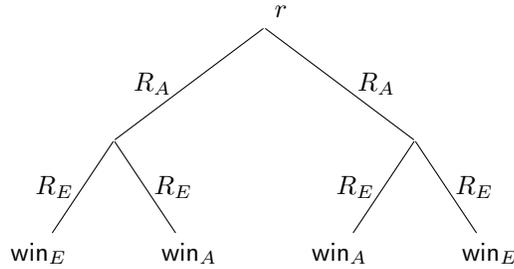
This may be a good point to stress once more that the present section is concerned with game forms only, omitting any player related aspect such as preferences between outcomes. When these are added, identifying appropriate notions of invariance becomes more challenging, as will be discussed in Section 3 below.

2.3 Languages matching invariance relations

The choice of invariance relations mirrors which structure is deemed relevant within a given perspective on games. A central tool for bringing out such relevant aspects is the existence of a logical language matching some invariance relation. In general, the more fine-grained the invariance perspective, the more distinctions a matching language should be able to make.

For a start, if one is interested in the properties a player can bring about through moves, a good choice of language is based on modalities $\langle \text{move}_i \rangle \varphi$, expressing that at least one of i ’s available moves leads to a next stage satisfying φ . The following illustrates how this language works in a given extensive form game.

Example Modal game language.



The modal formula $[\text{move}_A] \langle \text{move}_E \rangle \text{win}_E$, true at root r , expresses that E has a strategy that ensures her a win in two steps: whatever A does, E can react in such a way that she ends up in a node where she wins. In a more fine-grained perspective, the modal language could add expressions $[a], [b] \dots$ for specific move types a, b, \dots . In this language, the coarse-grained modality $\langle \text{move}_i \rangle \varphi$ is definable by the disjunction $\bigvee_{a \text{ is a move for } i} \langle a \rangle \varphi$, making the new language a refinement of the old.

In this way, general results of modal logic apply to games. For instance, take pointed models such as game trees with an indicator for the current moment. Whenever two such pointed models \mathcal{G}, m and \mathcal{G}', m' are bisimilar in the first sense defined above, the equivalence $\mathcal{G}, m \models \varphi$ iff $\mathcal{G}', m' \models \varphi$ holds for all formulas φ in a sufficiently rich modal language with modalities $[a]$ for each move label. Thus, one can switch between syntactic, language based perspectives and semantic invariance relations, depending on what is convenient for a given perspective on games. Entirely similar points hold for bisimulations and modal languages for power perspectives, or for strategic form games.

Finally, modal languages do not have exclusive rights. If still more fine-grained perspectives are needed, more expressive first-order or higher-order languages become serious contenders for describing games.

2.4 Modal logic of extensive games

A language for games facilitates both, defining properties of games and reasoning about them. An example are winning strategies for players in a two-step extensive game as just discussed. More generally, for any finite extensive game, there are formulas φ_j for each agent j that are true iff j has a winning strategy:

$$\varphi_j := [\text{move}_i]\langle \text{move}_j \rangle [\text{move}_i] \dots \text{win}_j$$

where the number of operators in the formula corresponds to the depth of the tree. Thus, logical laws governing reasoning with such formulas acquire game-theoretic content. For instance, the negation of the statement that one player, A , has a winning strategy is provably equivalent to saying that the other player, E , has a winning strategy, at least in those cases where A wins if and only if E does not:

$$\begin{aligned} \neg\varphi_A &= \neg\langle \text{move}_A \rangle [\text{move}_E] \langle \text{move}_E \rangle \dots \text{win}_A \\ &\leftrightarrow [\text{move}_A] \langle \text{move}_E \rangle [\text{move}_E] \dots \neg\text{win}_A \\ &\leftrightarrow [\text{move}_A] \langle \text{move}_E \rangle [\text{move}_E] \dots \text{win}_E = \varphi_E \end{aligned}$$

Hence, the logical law of excluded middle in its modal guise corresponds to Zermelo's theorem, stating determinacy for finite games.

Yet, there are limitations to such characterizations of game-theoretic properties in terms of logical laws. Formulas stating whether some player has a winning strategy change from model to model, as the number of modal operators depends on the size of the game tree. In fact, there is no uniform formula in the basic modal language expressing that player i can win in an arbitrary finite extensive form game. Such a formula can only be found in the modal μ -calculus (Venema, 2008), where the statement that i has winning strategy can be expressed with the fixed-point formula

$$\mu p. (\text{win}_i \vee (\text{turn}_i \wedge \langle i \rangle p) \vee \bigwedge_{j \neq i} (\text{turn}_j \wedge [j] p)$$

The more general point here is that the recursive nature of game-theoretic equilibria and solution concepts reflects naturally in logics with fixed-point operators for induction and recursion.

In this setting, known results about modal logic acquire a new significance. In the realm of finite models, for instance, having the same modal formulas true at two states is *equivalent* to there being a bisimulation connecting those two states (cf. Blackburn et al., 2001). Hence, whenever two finite games satisfy the same modal propositions in their respective roots they are equivalent in the sense of bisimulation. For infinite models, such results are less direct. A full equivalence between bisimulation and satisfying the same formulas, for instance, only holds for an extended modal language with infinite conjunctions and disjunctions. Other relevant results include

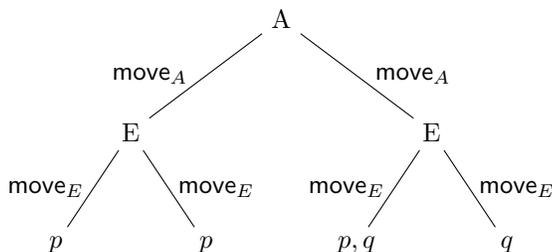
the existence of modal formulas that define given pointed models up to bisimulation. Such formulas sometimes exist in the basic modal language, sometimes in the μ -calculus, and always in the infinitary modal language. Applied to concrete games G , these modal definitions can be viewed as complete descriptions of all properties of G at the relevant level of invariance.

Finally, modal logic has many complete proof systems for capturing the valid consequences on various classes of models (Blackburn et al., 2001). These calculi of reasoning also apply to games, where they can capture aspects of specialized game-theoretic argumentation. Proof-theoretic perspectives are not the focus of this entry, but a number of strands will be mentioned where appropriate.

2.5 Modal neighborhood logic for powers

Besides extensive form games, standard modal logic is also suitable for the power perspective on game structure. Sometimes, one ignores the internal mechanisms of a game altogether, merely viewing it as a black box social mechanisms where players control outcomes to a certain extent. In this perspective, a player can *force* the outcome of the game to be in a some set X if she commands a strategy that ensures the game to end up in an outcome of X , no matter what the other players choose to do (van der Hoek and Pauly, 2007). Similarly, a player can force that some proposition φ holds if she has the power to enforce that the game ends in a φ state. The collection of all sets of outcomes an agent can force are often called her *forcing powers*. In classical game theory, these forcing powers sometimes go by the name of *effectivity functions* (Peleg, 1998), which are often also studied for coalitions of players (see Pauly, 2011; Goranko et al., 2013, and the entry on *coalitional powers*).

Example Powers in extensive games.



Notably, forcing powers are not closed under conjunction. In the game above, agent A can force p and q individually without being able to force $p \wedge q$. In modal logic terms, forcing powers give rise to a *neighborhood model* (Pacuit, 2017), where the neighborhood functions list the set of outcomes players can enforce from a given state. Reasoning about forcing powers can then employ a logical language with forcing modalities $\{i\}$ for each player:

$\{i\}\varphi$ agent i can force the outcome of the game to satisfy φ .

These modalities can be interpreted over the extended game forms with neighborhood functions described above. On the semantic side, a generalization of the neighborhood models defined above support a generalized notion of *power bisimulation*, see [Pacuit et al. \(2011\)](#).

The modal logic of powers allows to reason about games at a global level of description. The modal logic of neighborhood models validates the standard modal monotonicity principle

$$\{i\}\varphi \rightarrow \{i\}(\varphi \vee \psi),$$

as follows already from the truth definition of forcing modalities. However, as forcing powers are not closed under intersection, the aggregation law fails:

$$(\{i\}\varphi \wedge \{i\}\psi) \not\rightarrow \{i\}(\varphi \wedge \psi).$$

Instead, the logic contains new valid principles relating forcing modalities for different players. For instance, if i can force the truth of φ , then no other player j can force its falsity. Thus,

$$\{i\}\varphi \rightarrow \neg\{j\}\neg\varphi$$

is a valid principle of ‘consistency of powers’ in the logic of forcing powers. The converse of this principle for games with two players i, j

$$\neg\{j\}\neg\varphi \rightarrow \{i\}\varphi$$

expresses the notion of determinacy from the last section. This formula is not generally valid, but it is an axiom for the special class of determined games.

Finally, there also is an alternative, more algebraic perspective on powers, assuming an earlier-mentioned perspective of logic games. The two games depicted in the core example of Section 2.2 may be seen as evaluation games for propositional formulas

$$p \wedge (q \vee r) \quad \text{and} \quad (p \wedge q) \vee (p \wedge r).$$

Their equivalence qua powers, described earlier, then matches the standard propositional law of distribution. This algebraic perspective will return in Section 2.9.

More recent views of forcing and powers re-interpret the sets X of outcomes employed in the above definitions as referring to both players: one player restricts the total set of outcomes, while the other players can achieve all outcomes within that set. This variation significantly impacts the corresponding notions of game equivalence, as well as the modal languages used ([Bezhanishvili et al., 2018b](#)).

2.6 Modal logic for strategic games

In the strategic perspective on games, players select actions simultaneously, without having learned about their opponents’ choices of actions. This requires an additional level of analysis. Besides the various possible moves, an adequate representation must also track players’ uncertainty about how their opponents might act.

In terms of matching logical languages, this suggest a *multi-modal* approach, with $[\approx_i]$ ranging over i ’s possible choices, and $[\equiv_i]$ representing her uncertainty about the

opponents, see [Pacuit et al. \(2011\)](#). Moreover, when considering games rather than game forms, this picture needs to be enriched with a third feature, viz. preference modalities $[\preceq_i]$, see [Section 3](#).

Games in strategic form can be viewed naturally as models for a modal language of choice and uncertainty, where each *state* m consists of a strategy profiles, i.e., a sequence $(m_1, m_2 \dots)$ listing each player's choice of action. For convenience, the preference modality has been included:

$$\begin{aligned} \mathcal{G}, m \models [\approx_i]\varphi & \quad \text{Given the opponents' actions, } \varphi \text{ holds whatever } i \text{ does.} \\ \mathcal{G}, m \models [\equiv_i]\varphi & \quad \text{Given } i\text{'s choice, } \varphi \text{ holds whatever the opponents do.} \\ \mathcal{G}, m \models [\preceq_i]\varphi & \quad \varphi \text{ holds in all states at least as good as the current one.} \end{aligned}$$

This multi-modal language can express a variety of statements about strategic form games, such as:

$$\begin{aligned} \langle \approx_i \rangle \langle \equiv_i \rangle \varphi & \quad \varphi \text{ is a possible outcome of the game} \\ [\approx_i] [\equiv_i] \varphi & \quad \text{all outcomes of the game satisfy } \varphi \\ [\preceq_i] \langle \preceq_i \rangle \varphi & \quad \text{optimal states for some player satisfy } \varphi \end{aligned}$$

In the case of two players, one agent's choices corresponds to the other's uncertainty and vice versa. This shows in the validity of principles such as

$$[\equiv_i]\varphi \leftrightarrow [\approx_j]\varphi$$

More generally, the logic of matrix games includes the S5 axioms for both $[\approx_i]$ and $[\equiv_i]$, but also the commutation law

$$[\approx_i][\equiv_i]\varphi \leftrightarrow [\equiv_i][\approx_i]\varphi$$

expressing the grid-like structure of matrix games. This logic bears some resemblance to STIT-type logics of actions ([Herzig and Lorini, 2009](#)). Technically, a grid structure in models allows for encoding of undecidable computational problems ([Blackburn et al., 2001](#)), rendering it an open problem whether expressive modal logics of game matrices are decidable.

The step from two to more players, often routine in epistemic logics, can be delicate in the logic of matrix games. Accessibility relations of type $[\approx_i]$, interpreted as identity of profiles except for the i -coordinate, yield a product logic akin to the three-variable fragment of first-order logic which is known to be undecidable ([Bezhanishvili, 2006](#)). However, with only relations of identity at the i -coordinate, i.e. $[\equiv_i]$, the logic remains decidable ([Venema, 1998](#); [Van De Putte et al., 2017](#); [Lomuscio et al., 2000](#)).

2.7 Strategies as logical objects

There is further structure in extensive games than just single moves. In game trees, a player's *strategy* specifies what to do at each turn, whether this turn will ever be reached or not. An increasing body of work examines such strategies and their underlying formats. See [Gosh et al. \(2015\)](#) for an overview of various logical frameworks for reasoning about strategies.

In one concrete perspective, a strategy is akin to a program that instructs the agent on how to navigate a game tree. Hence, a natural logic of strategies uses the language of [propositional dynamic logic of programs](#) PDL, an approach that will return later. As programs are in general non-deterministic, such logics let a strategy recommend one or more actions the agent should take at each turn. In this perspective, strategies resemble plans that might remain partial.

In a program format, strategies start with basic actions, representing individual moves in a game tree. From there, complex programs π can be created using operations including sequential compositions $\pi_1 ; \pi_2$ (π_1 is to be performed followed by π_2), or choice $\pi_1 \cup_i \pi_2$ (agent i is to pick between actions π_1 and π_2). Moreover, a test operation $?\varphi$ for checking whether φ holds, enables strategies to react to properties of states or opponents' past actions. Finally, to describe continuous execution of a strategy along a game tree, it makes sense to have an operation π^* of program iteration, stating that π be executed arbitrarily often.

The language of PDL then has modal operators $[\pi]$ for every program π that can be defined from the basic actions and the operations just described. A simple such strategy advises player i to do a whenever it is her turn. The following formula states that this strategy ensures that φ holds throughout:

$$[[((?\text{turn}_i ; a) \cup (?\text{turn}_j ; \text{move}_j))^*]\varphi$$

Program definitions for strategies given here are closely related to the use of finite automata for defining strategies in computer science and game theory ([Osborne and Rubinstein, 1994](#); [Grädel et al., 2002](#); [Ramanujam and Simon, 2008](#)).

2.8 Simultaneous moves and imperfect information

In the extensive form games of [Section 2.1](#), players move in sequence and can base their decisions on full information of what has happened so far. The other extreme were games in strategic form, where agents move in parallel or, in the interpretation of strategy selection, have no means of picking up information during actual play. There are ample scenarios in between these extremes. Public good games with optional retribution against non-cooperators ([Andrighetto et al., 2013](#)), for instance, combine moments where some or all players make simultaneous moves with information collection along the way. Such parallel action can be mimicked in sequential games by limiting the information available to players at various states of the game. The resulting games of imperfect information will be discussed in [Section 3](#), alongside other sources of imperfect information.

Further well-known logical approaches to parallel action employ STIT logic ([Horty and Belnap, 1995](#); [Broersen, 2009](#)), and temporal logics such as ATL ([Alur et al., 2002](#)) or its epistemic variant ATEL ([van der Hoek and Wooldridge, 2003](#)).

2.9 Game algebra and dynamic logic of computation

So far, games were treated as monolithic entities that agents reason about in their entirety. This can be at odds with how real life agents conceptualize games. To

facilitate reasoning, games are often broken up into smaller tasks that are easier to handle separately. A chess player, for instance, may know how to solve different end games. Rather than reasoning about every possible situation until its end, she will evaluate different options in mid-play by considering which of these end games they will, most likely, lead up to. In this perspective, complex games are constructed out of simpler games that may profit from separate analysis. Games then form an algebra with operations that construct complex games from simpler ones. This style of thinking is reinforced when games are viewed as scenarios for interactive computation, where again algebraic methods are used widely (Bergstra et al., 2001).

Here is an illustration of this approach. For simplicity, consider only two players, A and E , the latter of which starts the game. One influential game algebra has the following operations, cf. Parikh (1985).

$G \cup G'$	Agent E has the choice between playing G and G' , i.e., represented by a choice node with two outcomes G and G'
$G; G'$	G is played first, followed by G'
$(\cdot)^d$	The roles of the players A and E are interchanged
$? \varphi$	Test game whether some property φ holds.

For instance, take a chess player in mid-game reasoning. For simplicity, restrict the possible end games to G_{E_1} and G_{E_2} . The player can then conceptualize mid-play as a game G_{mid} with end nodes labeled by propositions p_1 or p_2 , describing which of the two end games follows. The full remaining chess tree is then given by

$$G_{complete} = G_{mid}; ((?p_1; G_{E_1}) \cup (?p_2; G_{E_2})).$$

Equational axiomatizations for this game algebra can be found in Goranko (2003) and Venema (2003). However, following the analogy of propositional dynamic logic for an algebra of programs, there also is a *dynamic game logic* for this algebra of games, (Parikh, 1985). It adds a modality $\{G\}\varphi$ for each game G , with $\{G\}\varphi$ expressing that in game G , the first player E has a strategy to force the truth of φ . For the case of non-determined games, the language will be extended further to include separate modalities $\{G, i\}\varphi$, one for each player i . Dynamic game logic shows in a perspicuous manner how strategic abilities for complex games supervene on abilities in simpler games. This is done by means of reduction laws such as

$$\{G; G'\}\varphi \leftrightarrow \{G\}\{G'\}\varphi, \quad \{G \cup G'\}\varphi \leftrightarrow \{G\}\varphi \vee \{G'\}\varphi$$

For a complete list of reduction laws, as well as open problems in this dynamic game logic see Pauly (2011); van Benthem (2014). For other styles of game algebra, including also forms of parallel composition, cf. Abramsky (1997).

It should be said that imperfect information challenges this approach to game algebra. For instance, one may have to decompose a larger game into smaller subgames where agents need not know which of these subgames they are in. Game algebras with imperfect information have been studied in the context of Boolean Games (Harrenstein et al., 2001). A recent power-based game algebra with operations encoding imperfect information, showing some analogies with IF logic (Mann et al., 2011) can be found in Bezhanishvili et al. (2018b).

2.10 Special topics

Coalitions and Networks Nothing has been said so far about social or structural relations between players: they move individually and in interaction with all other players. However, in many games, groups of players can team up to jointly pursue goals, possibly in competition with other groups. Coalitions are a natural, but non-trivial extension of the logical frameworks introduced here, as strategic abilities of groups may exceed those of all members combined, see [Peleg \(1998\)](#) and the entry on [coalition powers in games](#). In other studies of social phenomena, the set of players is equipped with an additional network structure. An agents' outcome or behavior will then depend upon what network neighbors do ([Baltag et al., 2018](#); [Christoff, 2016](#)). Lastly, games on networks are closely related to information flows in social networks, as studied in depth by [Liu et al. \(2014\)](#) and [Seligman and Thompson \(2015\)](#) from a logical perspective.

Tracking This section contains a wide variety of perspectives on games. These differ in their invariance relations and their matching languages, offering different foci such as outcomes, powers, or the detailed temporal evolution of games. Even further perspectives will no doubt keep emerging. This diversity may seem overwhelming, making the field rather scattered. But here, another role of logic shows, by not just proliferating systems, but also as connecting them. Various logical translations exist between the languages and levels involved. Often, reasoning about games in a logic for some level can be mirrored precisely under translation into the logic of another level. Moreover, these translations can often keep track of changes in games under actions of information updates, a topic to be taken up in Section 3. Tracking of this kind is defined and studied in general logical terms in [van Benthem \(2016\)](#) and [Cinà \(2017\)](#).

Infinite games So far, games were tacitly assumed finite in length. This assumption is innocuous for many real life scenarios, yet there are notable exceptions. A prominent example are safety games, where one of the players, the guard, has to ensure a system to never leave a certain state, while the opponent attempts to deviate. Many technical tools for finite games also work for infinite games. There is, however, a number of conceptual and logical discontinuities. Since infinite games have no last moments, for instance, outcomes must be attached to complete histories of game play, rather than nodes of a tree. Reasoning about games then requires temporal modalities for a given history, but also modalities ranging over all open future histories. For analyzing powers, then, temporal versions of forcing modalities are needed. With these modifications, a logical style of analysis still applies. For instance, it is well-known that determinacy fails for infinite games ([Jech, 2006](#)). However, what holds for all games is a law of 'weak determinacy' stating that, if i has no strategy to force a set of histories satisfying φ , her opponent j can ensure that i will never obtain such a φ -strategy in the future. The difference between standard determinacy and weak determinacy is captured by the following two formulas, that are entirely in line with this section's style of analysis: $\{i\}\varphi \vee \{j\}\neg\varphi$ (determinacy) versus $\{i\}\varphi \vee \{j\}G\neg\{i\}\varphi$ (weak determinacy), where G is the temporal modality of 'always in the future on the current history'.

3 The Nature of Players

Game forms may be seen as spaces where players can operate. A game, however, is not fully determined by its game form alone. Rather, the players involved may import additional features relevant for game play. Players can, for instance, be limited in their powers of observation, either by aspects of the game structure or through cognitive limitations. The most striking added feature, however, is that players have preferences. Agents not only observe the world or act in it. While these describe mere kinematics of a game, agents also evaluate the present and various possible futures. Being driven by preferences, it is such evaluations that are the moving force behind player's choices. Preference, hence, take a prominent explanatory role for true game dynamics.

This section places its focus on the preferential and epistemic dimensions of players. Such factors are essential to notions of rationality where information, action, and preference are often entangled. In game theory, a harmony between these is often sought in notions of equilibrium for strategy profiles.

3.1 Preference and equilibria

Game trees and game matrices specify the moves available to players at different moments in time. They also indicate all possible outcomes, either as cells in a matrix, or as leaf nodes in an extensive game. However, to study what players *should* or *will* do in a game, a further component is needed: players' preferences. Such preferences need not only reflect material pay-offs or other features of outcome states. Rather, they may also relate to the process of play itself. Moreover, preferences may contain irreducibly subjective elements. Even when assuming the same role in a game, different players may disagree about the relative desirability of certain outcomes (Fehr and Schmidt, 1999).

Within a static, outcome-oriented perspective on games, a major emphasis is on equilibria: strategy combinations where all players do the best they can in light of their preferences and the opponents' strategies. A further, dynamic perspective focuses on how such equilibria relate to individual players' stepwise local reasoning on how to act in light of their beliefs and desires. This perspective is taken up in Section 4.

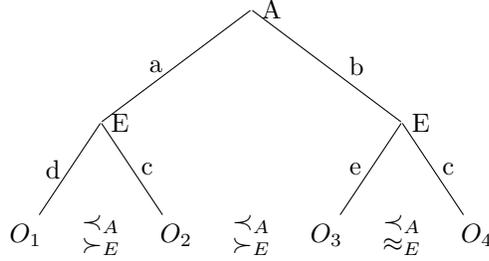
3.2 Preference logics for games

For reasoning about preferences, it must first be specified what it is that agents' preferences apply to. The orthodox account lets preferences exclusively range over possible outcomes (Osborne and Rubinstein, 1994). However, a growing trend in the logical literature assumes agents to rather care about the truth value of general propositions than can describe both the progression or outcome of a game. While not equivalent, the two perspectives are compatible. Both will be discussed in this section.

In the classical picture, player i 's preferences on a game tree are represented by a preference relation \prec_i ranging over the set of outcomes. Such a relation is usually

assumed transitive and reflexive, but need not be total.

Example A game tree with preferences.



Just as with the earlier modal logics for game forms, a relatively simple logical formalism can already express relevant aspects of agency in games. It offers a low-complexity language for expressing basic features of action and information, without going into details of the underlying quantitative mechanisms. More precisely, games with preferences naturally support a logic with modal operators $[\preceq_i]$ interpreted as follows:

$[\preceq_i]\varphi$ φ holds in all states at least as good as the current one to agent i .

Logics of this type can express various properties relevant to games. They can, for instance, say that all states better than the current one are φ states, making moving towards φ states a necessary condition for maximizing utility. They can also express, that all best states are φ states, with the formula

$$\langle \preceq_i \rangle [\preceq_i] \varphi$$

For more on modal preference logics, see [Hansson \(2002, 1990\)](#), [Girard \(2008\)](#) and [van der Torre \(1997\)](#).

Modal preference logic has further extensions with natural connections to games. In a refined perspective, for instance, preferences may derive from reasons, say, criteria or goals various agents want to achieve. This gives rise to a duality between preference relations among outcome states and priority orders over formulas, describing the agents' goals. Dynamic accounts, finally, track how preference can change under various input events. For more on both of these issues, see [Liu \(2011\)](#).

However, modal preference logics, construed either way, are not yet rich enough to express one of the essential notions of game theory. Further extensions are needed to deal with *best responses*, expressing that the current move of a player is the best she can do in light of her opponent's actions.

Best response moves are the main ingredient for game equilibria. Formally, a *Nash equilibrium* is a strategy profile, fixing a unique choice for each player, where nobody can improve by unilaterally changing strategy when all others maintain theirs. There are several ways of defining this property in extended modal preference languages. One possibility is to simply introduce a new atom b_i , stating that the current world

is the best player i could have achieved in light of the opponents' actions. In this language, Nash equilibrium are characterized by

$$\bigwedge_{i \in \text{Players}} b_i.$$

More explicit definitions exist, building on the strict preference modalities of [van Benthem et al. \(2009\)](#). Yet, perhaps the simplest illuminating approach uses an *intersection modality* from hybrid logic ([Areces and ten Cate, 2007](#)) combining the agent's preference relation with her uncertainty between the opponent's action to characterize best responses and Nash equilibria (cf. [Section 2.6](#)):

$$\bigwedge_{i \in \text{Players}} [\prec_i \cap \equiv_i] \perp$$

Expressing Nash equilibrium has served as a benchmark for logics of strategic games ([van der Hoek and Pauly, 2007](#)). Yet there are other desiderata, often connected to analyzing standard game-theoretic solution concepts for games. These are usually designed to find Nash equilibria or at least narrow down the strategy profiles to those compatible with certain requirements of rationality. Well-known methods of this kind are Backward Induction for extensive games and Iterated Removal of Strictly Dominated Strategies for strategic form games ([Osborne and Rubinstein, 1994](#)). These will be discussed now, as they raise intriguing further logical issues.

3.3 Backward Induction in extensive form games

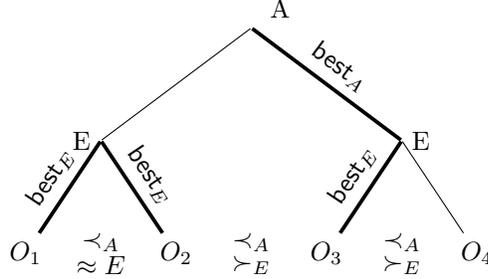
Here is a high-level description of Backward Induction. In extensive game forms, the aim is to introduce a new preference-based relation best_i , denoting that some move is the best a player can do at some given state. Thus, best_i is a subset of player i 's total move relation, to be defined in a suitable manner.

For final moves, standard decision theory suggests that a choice is best for the active player if no other move leads to a better outcome. When extending the analysis to earlier positions of the game, things depend crucially on players' expectations about their opponents' future behavior. Several possible policies exist, depending on the types of player involved. A widespread assumption in epistemic game theory is common belief in rationality, i.e., that all players involved are rational, believe their opponents to be rational, believe their opponents to believe that opponents are rational, and so on. In line with this assumption, the following algorithm extends the best_i relation recursively to non-terminal nodes:

Whenever player i is to move at state s , possible choices are assessed by comparing what would happen if, after that move, everybody followed their best relation. A possible move at s is included in i 's best relation if the best outcome of this move followed by repeated best moves by all players is at least as good as every other move i could make at s , followed by best moves by all players.

Here is how this bottom-up procedure works in practice.

Example Backward Induction.



This procedure is a qualitative version of classic game theory’s Backward Induction, which is based on utility values rather than preference relations (Leyton-Brown and Shoham, 2008).

Backward Induction and the resulting **best** relation is a prime example for the complex entanglement of preferences, information and action. A key modal axiom governing this relation was identified by Roy et al. (2006). **Best*** denotes here the transitive closure of the union of all best_i relations.

$$(\text{turn}_i \wedge \langle \text{best} \rangle [\text{best}^*](\text{end} \rightarrow p)) \rightarrow [\text{move}_i] \langle \text{best}^* \rangle (\text{end} \wedge \langle \preceq_i \rangle p)$$

Describing the limit of a dynamic process with a static property, this equivalence exemplifies a family of *characterization theorems* that play a crucial role within the logical analysis of games. Other dynamic perspectives can be analyzed in a similar logical style (Liu, 2011).

3.4 Iterated removal of dominated strategies

Iterative reasoning strategies akin to Backward Induction also exist for games in strategic form. Rather than defining a new unary move-predicate **best**, however, these procedures work by eliminating suboptimal actions. An action a is labeled suboptimal or *dominated* if there is some other available action, b , that guarantees a better result than a , no matter what the opponents do. In this case a rational player should drop a from her space of admissible acts, as she would never play it.

Just as Backward Induction, dominance reasoning has an iterative flavor. Assuming common belief of rationality, players can expect their opponents to also drop dominated actions from consideration. Doing so reduces the game and might render further moves dominated, as is illustrated in the following example. Within the left-to-right temporal progression, moves are greyed out as they get discarded. Players’ preferences are represented with numerical values with 1 the best and 4 the worst.

	c	d
a	1,1	4,3
b	2,2	3,4

	c	d
a	1,1	4,3
b	2,2	3,4

	c	d
a	1,1	4,3
b	2,2	3,4

The fact that further strategies may become dominated suggests to repeat the procedure, turning removal of dominated strategies into an iterated process. When games are finite, this process is guaranteed to converge in finite time. Iterated removal of dominated strategies on binary preference relations is a qualitative variant of the version employed in classical game theory, where cardinal utility values are assumed (Leyton-Brown and Shoham, 2008).

A closely related process is iterated removal of *weakly dominated* strategies, where some move a is deleted if there exists a b that outperforms a on some of the opponents' moves, while being at least as good on the remaining ones. Unlike its strict counterpart, iterated removal of weakly dominated strategies suffers from a number of technical and conceptual intricacies, such as order dependence of iterated deletion (Samuelson, 1992; Pacuit and Roy, 2011).

3.5 Goals

Generalizing the concept of winning or losing, agents can be assigned goals they pursue in a game. Restricting to a single goal per agent retains a binary perspective: A goal is reached or not. Goals, however, allow for additional flexibility. Besides pure competition as in win-lose games, these can also express pure coordination games, with everybody pursuing the same goal, or mixed motive games with partial overlap between different players' goals.

The concept of goal functions is particularly prominent in the logical framework of **Boolean games** (Harrenstein, 2004). There, each agent is given control over some atomic propositions, permitting her to freely decide on their truth value. Goals are then formulated as propositional formulas over the set of all players' atoms. Crucially, a player's goal formula might hence involve atoms that are not under her control. In iterated extensive Boolean games, goal formulas might also refer to properties of histories of play defined in temporal logics (Gutierrez et al., 2015)

3.6 Knowledge, belief and limits of information

There are various types of information players can possess or lack about a game. First and foremost, players can be uncertain about the types of opponents they are facing: their preferences, their reasoning about the game and how they expect the game to unfold. Second, agents' uncertainty can extend to the game itself. Players, of course' won't know their opponents choices in a simultaneous move game. Besides, agents might also have limited information about past moves and events. Such uncertainty can arise from the game structure eschewing certain observations, but also from failing to record past information properly. In yet more extreme cases, agents might even be unsure about the moves available to their opponents.

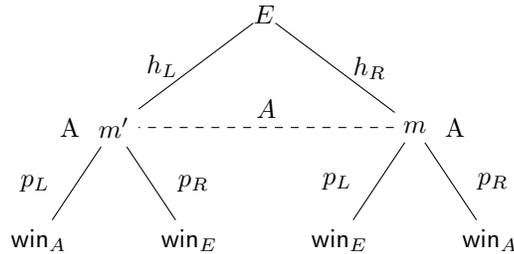
Given the various limitations to their knowledge, players may entertain beliefs to structure their uncertainty. Such beliefs may, naturally, change over time, as players communicate or observe the game unfold. The importance of beliefs in the logical analysis of games has been emphasized in Stalnaker (1998), who was the first to highlight the role of belief revision in analyzing reasoning about game solution.

Uncertainty about moves

In one sense of uncertainty, even highly idealized agents might have but limited information of what has happened so far. In certain cases, the game's structure may limit some players' observational powers of their opponents' moves. In other instances, agents may suffer under cognitive limitations restricting their perspective on the game. Or, sometimes, agents might simply fail to record some of the moves made by themselves or others.

Within *extensive games with imperfect information*, all such cases are represented by indistinguishability relations ---_A between states m, m' , expressing that agent A cannot distinguish between being at m and m' . Notably, this does not preclude the player from learning later on in the game whether he has been at m or m' .

Example A game with imperfect information.



While allowing agents to lack information of various kinds, the above analysis makes one structural assumption about players: they always know which moves are available to them at a given node. In extensive games with imperfect information, this translates to the requirement that whenever two states are indistinguishable to some agent, they coincide on the set of her possible actions.

The move from perfect to imperfect information has major implications for strategic reasoning. In the game depicted above, player A cannot distinguish between being at m and m' . When at the former, she may, for all she knows, be at m' instead. Hence A 's decision needs to account for both possibilities; she cannot base her choice on any property that holds at only of those locations. In particular, A has no available strategy that guarantees her ending up in a win_A node. Since E cannot ensure a win either, no player has a winning strategy. This is a central difference with finite perfect information games, where it is guaranteed that one of the players has a winning strategy, cf. Section 2.4.

The logic of imperfect information

Reasoning about imperfect information requires to extended languages for extensive form games with epistemic modalities. For each player i , modality $K_i\varphi$ represents i 's knowledge. The usual semantics of epistemic logic relates this to uncertainty, as encoded by the players's indistinguishability relation:

$$\mathcal{M}, m \models K_i\varphi \quad \text{all states } m' \text{ with } m \text{ ---}_i m' \text{ satisfy } \mathcal{M}, m' \models \varphi$$

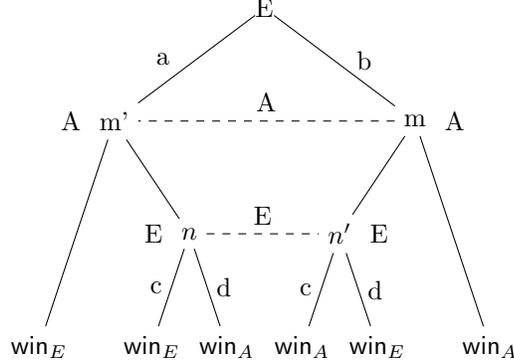
This language is best illustrated with the above game tree. To this end, interpret the tree as the classic children’s game where one player, A , has to guess in which hand her opponent, E , hides some little token. Once E has hidden the token in, say, her right hand (move h_R), the guessing player has a winning move *de re*: she should pick *right* (p_R). However, as the token was placed in secret, she may not know that picking *right* is a winning move: the player has no winning strategy *de dicto*. This is expressed by:

$$\mathcal{M}, m \models [p_R]win_A \wedge \neg K_A[p_R]win_A.$$

In a game theoretic setting, the *de re* vs. *de dicto* distinction has been studied by [Horty and Pacuit \(2017\)](#) and [van Benthem \(2001\)](#).

In light of these considerations, the notion of strategy change, and the logics for defining strategies involve epistemic elements. It seems reasonable to demand that agents cannot base their choice of strategy on everything that has happened before, but only on what they know, i.e., their current information ([Pacuit et al., 2006](#)). The resulting *uniform strategies* ([Maubert, 2014](#)), can be defined by the knowledge programs of [Fagin et al. \(1997\)](#). Further restrictions are possible, for instance granting agents limited memory that only reaches back a fixed number of moves ([Gutierrez et al., 2015](#)).

The epistemic action language can express many further phenomena in imperfect information games. The following game is an illustration.



Once player E arrives at node n , she cannot discern that actual situation from node n' . However, E must have possessed information earlier on that distinguishes n from n' : to arrive at n , she must have played a as her first choice, while n' can only be reached after playing b . Thus, E can only be uncertain between these two nodes if she forgot about her own previous actions.

The epistemic action language can distinguish between such scenarios with memory loss and those without. The property of *Perfect Recall* states that players retain full memory of all moves they observed. This can be expressed by the following axiom scheme ([Halpern and Vardi, 1986](#); [Bonanno, 2004](#))

$$K_i[a]\varphi \rightarrow [a]K_i\varphi.$$

Also the converse of this scheme admits a natural interpretation:

$$[a]K_i\varphi \rightarrow K_i[a]\varphi$$

This *No Miracles* property expresses that players can only learn by observing moves, not by any other methods extraneous to the game.

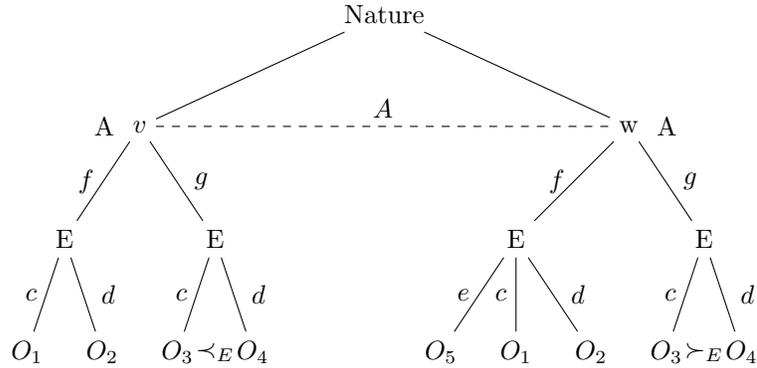
Of course, logic does not presuppose that all players have perfect memory, or that they cannot pick up any information outside the progression of play. Epistemic action language can equally well be employed to analyze more general scenarios where the above axioms do not hold. Especially within dynamic-epistemic versions, epistemic logics can produce modified versions that cover many more cases than those described here (van Benthem, 2014). Moreover, further modalities from epistemic logic make sense, in particular, those for common or distributed knowledge in groups of players (Fagin et al., 2004a; Meyer and van der Hoek, 1995).

An epistemic component with operators K_i fits with many logical perspectives on games. In particular, epistemic extensions are as compatible with coarse logics such as the earlier-mentioned $[\text{move}_i]$ -setting with a single move-modality per player, as with fine logics where each individual action type is represent by a distinct modality $[a]$. In fact, in Section 2.6, epistemic operators were used for analyzing games in strategic form, where modalities naturally relate to uncertainty about the other players' strategies.

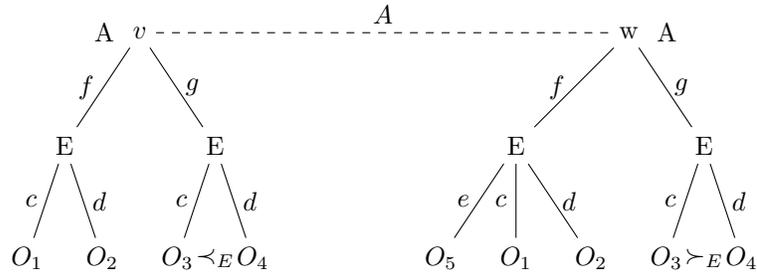
Uncertainty about options and preferences

In more general settings, uncertainty does not stop at the opponents' informational states. In international relations or economic bargaining, also the players' motivations and preferences are not fully known to all parties involved. Within the corresponding extended form games, players may be uncertain about their opponents' preferences and strategic options, whether they can afford a certain move or whether they actually possess the information they threatened to reveal. Obviously, uncertainty about preferences or available options will impact reasoning about the equilibria of a game. Strategic players might even try to exploit such uncertainties, for instance by pretending to have options they do not possess.

In a first pass, this type of uncertainty can be expressed by introducing nature as hypothetical player, with a first move that determines the preferences and available options for all players. A simple example is the game depicted below. At the start, A is uncertain whether E can reply to A 's move f by playing e . Likewise, she lacks information on whether E prefers O_3 over O_4 , or vice versa.



From a logical point of view, this miraculous initial move by Nature is not needed. Standard epistemic models can represent the above scenario, and many more complex ones, by means of the indistinguishability relations introduced above. Technically, this requires to move beyond standard imperfect information trees to so-called *epistemic forests* (Gerbrandy et al., 2009a), sets of trees linked by epistemic relations. In particular, the above game tree transforms into

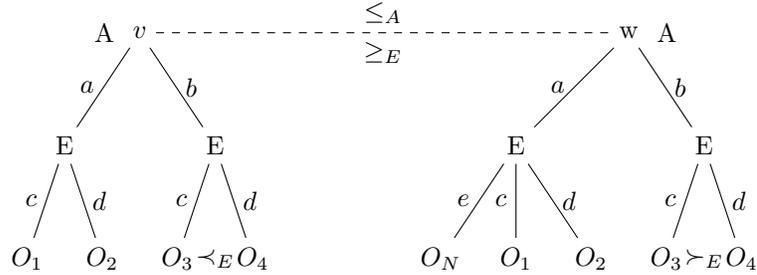


The epistemic action language fortrees works just as well on epistemic forests. However, in suitably expressive languages, the logic of forests is weaker than that of trees, as the set of validities on the class of n -player trees is a strict superset of the validities on n -player forests.

Imperfect information and beliefs

Further enrichments of the logical framework add semantic structure to the agents' uncertainty. When unable to determine the exact situation, players may classify options with respect to plausibility. To this end, epistemic models have been equipped with plausibility orderings \geq_i for players i (Boutillier, 1994; Stalnaker, 1968; Baltag and Smets, 2008).

In the preceding example, plausibility order might work as follows:



This richer structure is reflected by introducing new modalities for agents' beliefs, determined by the most plausible states:

$$\mathcal{M}, w \models B_i \varphi \quad \varphi \text{ holds in all } \geq_i\text{-maximal states in } i\text{'s epistemic range.}$$

Conditional belief, important to players' planning within a game, can be interpreted in the same style:

$$\mathcal{M}, w \models B_i^\psi \varphi \quad \varphi \text{ holds in all } \geq_i\text{-maximal } \psi\text{-states in } i\text{'s epistemic range.}$$

These clauses are intended to work in finite as well as in infinite settings. However, in the latter case minor modifications might be needed akin to those in [conditional logic](#). These have been proposed in various alternatives. Notably, this enriched epistemic-doxastic logic allows for further, less standard interpretations beyond those illustrated so far. Examples are 'strong belief', expressing that all relevant φ -states are more plausible than all relevant $\neg\varphi$ states, or 'safe belief' saying that φ holds at all states that are at least as plausible as the current one. See [van Benthem and Smets \(2015\)](#) for an overview of plausibility semantics and its connections to conditional logic, belief revision theory, dynamic-epistemic logic, and a wide range of philosophical and technical issues.

3.7 Higher order uncertainty and type spaces

In various scenarios, agents reason not only about the opponent's preferences or admissible moves, but also about their beliefs about the game and others' behavior therein. In fact, such higher-order reasoning can have a major impact on game play. A prime example is the Backward Induction procedure of [Section 3.3](#), where the construction of a best move relation crucially relied on common knowledge of rationality. More generally, agent's best moves frequently depend upon what they expect others to do. This phenomenon is especially prominent for simultaneous move games, where it occurs in both coordinative scenarios ([Skyrms, 2004](#); [Lewis, 2008](#)) as well as competitive ones ([Hotelling, 1990](#)). More details can be found in the entry on [epistemic game theory](#).

Arbitrary first and higher order levels of knowledge and belief can be represented with the above relational models, the standard tool in [epistemic](#) and doxastic logic. For information in extensive form games, the epistemic-doxastic perspective on states can be combined with move-relations in exactly the way described before. The result are *epistemic-doxastic trees* or *forests* that can represent most types of knowledge or

belief players might have about the game, including its exact shape, previous moves, opponents' preferences or opponents' first- and higher-order beliefs on any of these matters.

Outside of logic, higher-order information has also been modeled in classical game theory. Quantitative frameworks represent information as probability distributions over a given event space. In this setting, higher-order information corresponds to probability distributions over probability distributions of the right kind. More specifically, n^{th} order information corresponds to a probability distribution over the space of $(n-1)^{\text{th}}$ order beliefs. As shown by [Harsanyi \(1967\)](#), the limit of specifying higher and higher levels of information can be represented as a type space, where each agents' type is a probability distribution over states of nature and the other players' types. In an abstract sense to be discussed below, these types correspond to states in standard models for modal logic.

In addition to standard modal models, logic also has a straightforward analogue to probabilistic type spaces: *logical type spaces*. In a formal framework first introduced by [Fagin et al. \(1999\)](#), an n -type is a sequence $\mathbf{f}_n = \langle f_0, f_1, \dots, f_n \rangle$ where f_0 specifies the state of nature, i.e., a valuation recording which atomic propositions are true or false. f_1 lists for all players the states of nature they consider possible. f_m for $m \geq 0$ then specifies for all players which $(m-1)$ -types, i.e., sequences $\langle g_0, \dots, g_{m-1} \rangle$ they consider possible. In this way, an n -type fixes the player's higher-order beliefs up to level n . These types are, of course, subject to coherence conditions: the agents' k -types for different k must fit together. For instance, whenever some agent considers a k -type \mathbf{f}_k possible, she must also consider any initial segment $\mathbf{f}_{k'}$ for $k' < k$ possible. Conversely, any k' type the agent considers must be the initial segment of *some* k type the agent holds possible. Type spaces offer a semantics for the epistemic language: Inductively, for a formula φ of modal depth m , $K_i\varphi$ is true at a type \mathbf{f}_n if all $g_m \in f_{m+1}(i)$ satisfy φ or if $m > n$.

Alternatively, the set of n -types allows for a natural interpretation as relational models with accessibility relations defined by

$$\langle f_0, \dots, f_n \rangle R_i \langle g_0, \dots, g_n \rangle \quad \text{For all } m \leq n \text{ holds } g_{m-1} \in f_m(i)$$

Interpreting the set of n -types as a relational model yields a second way of evaluating the epistemic language on logical type spaces. For formulas of modal depth less than n the two interpretations coincide. Hence, up to finite depths, type spaces and their associated relational models are two perspectives on the same informational situation.

To fix all of the agents' beliefs, the analysis moves to types $\mathbf{f} = \langle f_0, f_1, \dots \rangle$, containing some f_n for every natural number n . In this extended framework the situation becomes more complicated. The space of all such types is universal in the following sense: every relational model can be mapped in a truth-preserving manner to the space of all types by sending each state to a full description of the agents' corresponding first- and higher-order informational attitudes. This map, however, is usually not a modal bisimulation. In fact, the process of type construction could be continued indefinitely, yielding a transfinite hierarchy of mutually non-bisimilar type spaces ([Heifetz and Samet, 1998](#)). Such transfinite types can become relevant when

the epistemic language is enriched with modalities for common group knowledge, in which case a full description of all expressible attitudes involves infinite hierarchies of higher-order information (Fagin et al., 1999). A recent logical study of type spaces including their probabilistic structure can be found in Bjorndahl and Halpern (2017).

The tight connection between type spaces and relational models is compatible with additional assumptions that might be imposed on the players' mental states. Fagin et al. (1999) characterizes when type spaces give rise to $S5$ models, while Galeazzi and Lorini (2016) do the same for multi-agent $KD45$ belief.

While relational models and logical type spaces represent exactly the same information, their main differences is in perspective. Relational models take a third person bird's eye view on possible worlds. Their starting point is a set of worlds rich enough to contain all states considered possible by the relevant agents, together with accessibility relations modeling players' information. From there, agents' first-order beliefs at the various worlds can be read off and, subsequently, also all higher levels of information. *Logical type spaces*, by contrast, assume a first person perspective. They take a full description of first and higher-order beliefs as primitive and treat indistinguishability as a derived relation.

Finally, it should be noted that type spaces assume a static perspective on games. No provisions are taken for representing moves or strategies explicitly, nor for incorporating updates of knowledge and belief that occur as a game in extensive form unfolds, cf. the discussion in Section 4. Thus, there is some distance between type spaces and the earlier epistemic-doxastic forest models for extensive games. As a first step towards filling this gap, it has been shown how type spaces can accommodate product updates from dynamic epistemic logic (Klein and Pacuit, 2014).

3.8 Reasoning, bounded agency and player types

Besides variations in preferences and beliefs, a third crucial aspect of players is their styles of information processing, decision making, and reasoning. Real cognitive agents are bounded in their information processing, as both their memory and reasoning capacities are limited. In particular, players may not be able to represent the entire game they are in, nor reason until the end of the game. This phenomenon of *short sight* has been studied in Grossi and Turrini (2012) and Turrini (2016). Moreover, in real-life iterated social interaction, payoffs are generated along the game, and may not be clear beforehand, (Axelrod and Hamilton, 1981). In such contexts, the best strategy in terms of short-term payoffs need not be optimal in the long run, but bounded agents may miss this longer horizon, (Klein et al., 2018).

Logical literature on bounded agency is too broad to be surveyed here. For a few research lines relevant to games, see Fagin and Halpern (1987); Heifetz et al. (2006) on epistemic logics with awareness, Artemov (2008) on *justification logics*, van Benthem and Pacuit (2011) on evidence logics, and Hansson (1998); Lorini (2018) on doxastic logics with computationally tractable belief bases.

In the game theoretic literature, bounded agents have often been represented as finite state machines (Gutierrez et al., 2015; Binmore and Samuelson, 1992). Limitations on reasoning capacities or memory size then translate into bounds on the machine

size. The resulting hierarchy allows for a fine grained analysis of information processing, reasoning, and thus bounded players of different types. This perspective fits well with the logical study of agency in computer science, (Grädel et al., 2002; Wooldridge, 2009).

In an integrated perspective, preferences, beliefs and reasoning styles can all be subsumed under the game-theoretic notion of a *player type*. For reasoning about the future course of a game, players will hence often entertain beliefs about each others' types. A simple example is Backward Induction, where players assume all opponents fully rational throughout. In more complex settings, individual actors may attempt to rationalize various moves observed and derive predictions about their opponents' future behavior by taking a broader range of options into account. Such players may start by assuming the counter-player to be a simple machine, and only move up to more complex views when required by evidence. In particular, there is no reason to assume uniformity of players or views. Within a given scenario, a diversity of player types might be present (Liu and Wang, 2013; Paul and Ramanujam, 2011; Ghosh and Verbrugge, 2017; Bergwerff et al., 2014). For some game-theoretic proposals concerning most frequently occurring player types, see Camerer (2011).

3.9 Thin and thick models for players

This section has outlined a variety of ways for incorporating players and agency into game forms. These come in a hierarchy of richness, ranging from annotated game trees to epistemic forests, type spaces, or yet more abstract models of games.

At the thinner end, the focus is on structural aspects of the game, incorporating players' preferences, but not necessarily their beliefs. Such frameworks are typically just rich enough to represent equilibria or backward induction paths, and to reason about these in logics of action and preference. Thin models leave much information about players' knowledge, beliefs, or their modus operandi unspecified, and put less emphasis on the actual dynamics of game play.

At the thicker end, models for games have lush worlds encoding players' preferences, information, beliefs, and perhaps even their complete types, including memory and reasoning capacities. Typical models of this kind are found in Stalnaker (1998) and Halpern (2001). When applied to extensive games rather than strategic-form games, thick models can anticipate anything that can happen in one large temporal universe, allowing to derive a full prediction about how play will proceed.

The distinction between thick and thin logical models seems folklore in applied logic. In fact, it occurred already in the earlier-mentioned choice between local logics with single step modalities versus temporal logics built over a complete universe of histories. Yet, there does not seem to be a unique best perspective. Rather, the choice between thick and thin models often depends on the exact goals pursued. A central consideration in this tradeoff is where the dynamics of stepwise play should be located: Thick models pre-encoded such dynamics, whereas thin models allow for an external dynamic logics for update (Baltag et al., 2009). The next section will highlight a number of ways for complementing a thin perspective with dynamic information on game play through representations of actions and updates.

3.10 Special Topics

Deontic reasoning Preference is closely related to obligation and permission. This shows in particular on the formal side, where preference logic in both static and dynamic variants (Hansson, 1990; Grossi et al., 2014) has clear analogies with **deontic logic**. Moreover, deontic and game-theoretic perspectives have given rise to many fruitful connections. In one direction, deontic notions may be seen as high-level descriptions of optimal actions given the information and obligations of agents (Kooi and Tamminga, 2008; Anglberger et al., 2015). Conversely, game solution procedures can enrich accounts of deontic notions (Parikh et al., 2011; Horty, 2016). Lastly, in artificial intelligence, deontic perspectives arise in tracking the behavior of a distributed system in relation to its goals, (Ågotnes and Wooldridge, 2010).

Mathematical foundations Incorporating players raises new questions about game equivalence. When agency matters, adequate notions of equivalence cannot stop at preserving properties of the underlying game form. Rather, *player-dependent equivalences* will also require preservation of players' beliefs, preferences or reasoning types. Incorporating such additional parameters makes it harder for two games to be equivalent, as new space for variation comes into play. On the other hand, agent limitations might also create new simpler game equivalences that can be studied by the tools of this entry.

4 Analyzing Play

The term 'game theory' suggests that everything of interest is captured in the format of a game with its moves and outcomes. The present entry reassembles this perspective, considering additional structure. A first extension was offered in Section 3, treating the nature of players as a topic in its own right. This section puts a spotlight on a second topic, game play in a wider sense.

Many themes in the literature on logic and games fall into three phases connected to play. Certain activities can already be conducted *before* the actual game. Examples are assessing the opponent or forming a plan. Most relevant choices and decisions, however, take place *during* the game - at least unless one thinks of players as automata blindly following preset strategies. Lastly, also *after* a game significant activities occur. These involve learning about opponent types, identifying crucial mistakes made or rationalizing the moves taken. In what follows, examples will be presented of each phase.

4.1 Game solution and pregame deliberation

When viewing games as static structures, rationality can be defined in terms of coherence between players' **beliefs, preferences and choices or intentions** (Elster, 1988). However, rationality also describes a quality of behavior, related to how players act or what they take advantage of when deliberating about a game. The relation between both perspectives can be made concrete when interpreting game solution procedures as styles of pregame deliberation. To follow is a dynamic analysis of Backward Induction (cf. section 3.3) that differs conceptually from **characterization theorems**

in terms of static properties such as common knowledge or common belief. In the dynamic analysis, these group properties are not assumed as preconditions. Rather, they are produced through the logic of deliberation.

4.1.1 Backward Induction via public announcement

The Backward Induction algorithm is usually presented in a quantitative setting, with player-utility values attached to outcomes (Leyton-Brown and Shoham, 2008). However, the same algorithm also works in a qualitative setting, with attitudes expressed by a preference relation between outcomes.

Backward Induction Backward Induction computes optimal moves for players. More specifically, at each choice node of an extensive form game, one or more of the available moves is labeled as optimal. For each player this set of optimal moves often forms a strategy in the usual game-theoretic sense, i.e. a function selecting a unique action to take at each of her choice nodes. Yet, there are degenerate cases where backward induction merely creates a relational strategy, restricting the available moves, while still leaving some choices to the player.

The principle driving the Backward Induction algorithm is that no player should ever select a move that is dominated by another move available at the same moment. Dominance here works in a recursive manner. Move a dominates move b if the corresponding player prefers each final outcome reachable from a by following Backward Induction moves to every outcome reachable from b by Backward Induction moves.

Public announcement of rationality In one perspective, Backward Induction can be understood as a process of prior-to-play deliberation, executed by players whose minds proceed in harmony. Deliberation steps are repeated public announcements (!rat) of rationality-at-nodes:

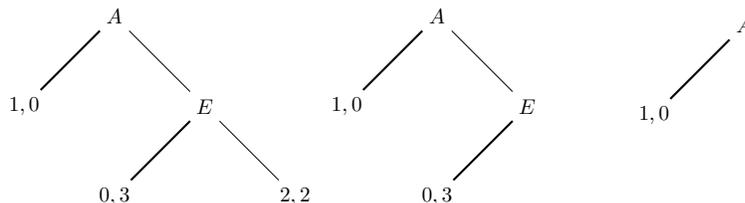
rat no player arrived at the current node via a strictly dominated move

Dominance here is a relation between the outcomes available after a certain move has been made. In one interpretation, some move a dominates another move b if every outcome that remains obtainable after a is preferred to any outcome reachable after a b move. However, there is a dynamic twist: crucially, the game tree considered, and hence the outcomes available, changes during the deliberation procedure.

The semantics of **announcement updates** works by trimming models. $!\varphi$ transforms a model M into a sub-model $M|_{\varphi}$ consisting of all those points in M that satisfy φ , while deleting all $\neg\varphi$ nodes. Relations on $M|_{\varphi}$ are those inherited from M . Crucially, deletion may change the truth value of formulas: after announcing $!\varphi$, some nodes in $M|_{\varphi}$ may satisfy $\neg\varphi$. In particular, with M the set of points in a game tree, the set of available histories may keep shrinking as successive announcements are made. Hence, repeated announcements of !rat make sense. In finite games, this process always reaches a limit, a smallest subgame where no move is dominated by another.

Example Solving games through iterated assertions of rationality.

Consider the following game, already introduced in Section 1.1. Iterated announcements of !rat removes nodes that can only be reached by dominated moves as long as this can be done. The trace of this procedure is:



Here, the Backward Induction solution emerges step by step. Stage 1 of the procedure rules out the leaf labeled with (2, 2) as the only point where **rat** fails. Stage 2 then rules out *E*'s choice node as new node where **rat** fails. In the resulting game tree, **rat** holds throughout.

More generally, let $(! \varphi, M)^\#$ be the limit (i.e. the first fixed point) of M under repeatedly announcing φ as long as it still true. In any game tree, the fixed-point $(! \text{rat}, M)^\#$ has **rat** true throughout. Its nodes contain the actual play computed by the Backward Induction algorithm (van Benthem, 2014).

Limit behavior Rationality is ‘self-fulfilling’ in the limit: if players commit to it in deliberation for long enough, they prune away all irrational moves and, a fortiori, all moves incompatible with common belief of rationality. The final outcome is a model with rational play at every point, a form of common knowledge of rationality. However, iterated announcements can also yield a different type of limit behavior: self-refutation. A prime example for this is the classic Muddy Children puzzle (Gierasimczuk and Szymanik, 2011) where repeatedly communicating ignorance leads to knowledge in the end. Also within game theory, a number of situations exist where (credible) announcements of future irrationality can leave some player better off than the Backward Induction solution (Leyton-Brown and Shoham, 2008).

4.1.2 Backward Induction via belief revision

Iterated belief revision. A different perspective of pre-game deliberation is couched in terms of belief instead of knowledge. The driving force here is rationality-in-belief:

rat* players never chose any move dominated by another in light of their beliefs how play will proceed from then on.

In this setting, the game tree itself remains invariant during deliberation: no histories are removed or ruled out. What may change is the relative plausibility of occurrence players assign to end nodes or, in infinite games, histories.

In plausibility semantics (briefly introduced in Section 3.6), an agent believes propositions that hold true in all those epistemically accessible worlds that are maximal in her plausibility order. The corresponding dynamics of deliberation does not proceed by point deletion, but by *soft updates* modifying the agents’ plausibility ordering. For Backward Induction, a ‘radical upgrade’ $\uparrow \varphi$ suffices, that moves all φ -worlds above

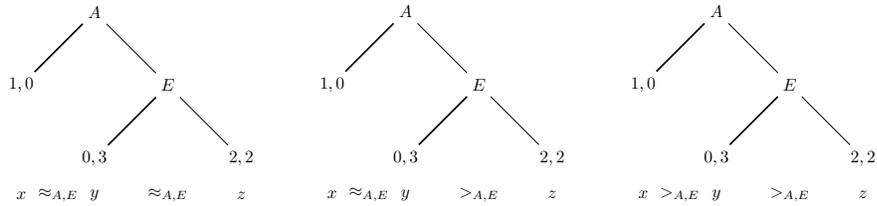
all $\neg\varphi$ -worlds states while maintaining the ordering within these two sets (Baltag and Smets, 2008).

Here is how this mechanism works in a game setting. Start with all end nodes equiplausible for all players. Since upgrades proceed by public announcement, all players will share the same beliefs throughout. In the procedure to follow, a move x is *dominated in belief* by a move y of the same choice node if, in the acting agent’s plausibility ordering, the most plausible end nodes reachable after y are all better for her than each most plausible end node compatible with move x . Now perform radical upgrades of type

$\uparrow \text{rat}^*$ If y is dominated in belief by x , make all end nodes following x more plausible than those after y .

Example Backward Induction, soft version.

Here are the stages for the new procedure in the preceding example, where the letters x, y, z stand for end nodes or histories of the game:



In the top node of the leftmost tree, going right is not dominated in beliefs for player A by going left. So, rat^* only affects E ’s turn, and radical upgrade with $\uparrow \text{rat}^*$ makes $(0,3)$ more plausible than $(2,2)$. After this change, going right has become dominated in beliefs in the top node, and a new upgrade takes place, making A ’s going left most plausible.

Iterated upgrade with rat^* always stabilizes to a fixed plausibility order, which is the same for all players. Identifying each history of a game with its end node allows for a belief analysis of Backward Induction (van Benthem and Gheerbrant, 2010). On finite trees, the histories emerging when all players resort to their Backward Induction strategies exactly correspond to the most plausible end nodes created by iterated radical upgrade with rationality-in-belief. An alternative dynamic-epistemic characterization of Backward Induction, using similar ideas in a different mix, can be found in Baltag et al. (2009).

Stabilization cannot be taken for granted. For other assertions φ , iterated upgrades $\uparrow \varphi$ can lead to oscillating or divergent plausibility orders. However, this divergence is limited. While cycles can occur for conditional beliefs, every truthful iterated sequence of radical upgrades eventually stabilizes all propositional beliefs (Baltag and Smets, 2009).

Fixed-point logic In a more technical perspective, the Backward Induction strategy can be defined as largest subrelation of the total move relation that has at least one successor at each non-terminal node, while satisfying a confluence property between action and preference:

$$CF(s) \quad \forall x \forall y ((\text{turn}_i(x) \wedge xsy) \rightarrow \forall z (x \text{ move } z \rightarrow \exists u \exists v (\text{end}(u) \wedge \text{end}(v) \wedge ys^*v \wedge zs^*u \wedge u \leq_i v)))$$

This fact is the basis for proving that Backward Induction is definable in First Order Fix point logic $LFP(FFO)$ (van Benthem and Gheerbrant, 2010). Results in this line of research connect game solution and game-theoretic equilibria with fixed-point logics of computation. In simple settings, such as that of Zermelo’s Theorem mentioned earlier, modal fixed-point logics akin to μ -calculus suffice.

4.1.3 Iterated removal of strictly dominated strategies

Further game solution concepts can be analyzed with logics of iterated update. In particular, iterated updates are not restricted to extensive form games, but can also provide insights for games in strategic form. A paradigmatic algorithm is Iterated Removal of Strictly Dominated Strategies (SD^∞). In this setting, a strategy is considered dominated if there exists another strategy that yields a strictly higher payoff against any of the opponent’s actions.

Example Iterated removal of strictly dominated strategies (SD^∞).

Consider the following matrix. As usual, pairs list A ’s utility first, E ’s second.

		E		
		a	b	c
A	d	2, 3	2, 2	1, 1
	e	0, 2	4, 1	1, 0
	f	0, 1	1, 4	2, 0

First remove the right-hand column, i.e. E ’s action c which is dominated by a and b . With c being removed, A ’s action f has become strictly dominated. After its removal, E ’s action b becomes strictly dominated, and after that, A ’s action e . At the end of the process, iterated removal leaves nothing but the state (d, a) , the game’s unique Nash equilibrium. In general, the resulting game matrix after all removals is guaranteed to contain all Nash equilibria of the original game, but it may also contain further strategy combinations.

In this setting, the formal dynamic apparatus involves assertions appropriate to the matrix games of Section 2. In fact, various different types of rationality can be defined in logics for matrix games. Here is an illustration for two-player games, involving announcements of ‘Weak Rationality’:

WR Each player thinks that, compared with each of her available alternative actions, her current move might be at least as good for her.

This statement is the negation, for each player, of her current action being strongly dominated. Naturally, this property can be expressed formally with suitable epistemic action modalities. Yet, even as it stands, it is clear that Weak Rationality can be announced to prune away strategy profiles, and that in an iterated manner. The strategic game will change every time weak rationality is announced, initiating a stepwise process that resembles the earlier iterative announcements of rationality for Backward Induction. As observed there, limits of public announcements are

always reached eventually, as models can only get smaller. For announcing Weak Rationality, these limits match the outcome of SD^∞ precisely (van Benthem, 2007).

A similar style of analysis can be extended to other notions of rationality. For instance, taking B_i to stand for ‘the current action of player i is best for her against all actions of the opponent’, the following formula may be dubbed strong rationality (SR)

$$\langle E \rangle B_E \wedge \langle A \rangle B_A$$

Briefly, the formula expresses that both players have a reasonable hope of doing well. *Strong Rationality* in this sense is related to the rationalizability program for game solution (Pearce, 1984; de Bruin, 2005), where actions are discarded if a better response exists under all circumstances. Strong Rationality, too, drives a game solution method.

Example Updates with iterated announcements of strong rationality (SR).

Consider a slight variation on the previous example. Below is the sequence of updates for iterated announcements of strong rationality (SR^ω)

2, 3	2, 2	1, 1	2, 3	2, 2	2, 3	2, 3
0, 2	4, 1	1, 0	0, 2	4, 1	0, 2	0, 2
1, 1	3, 4	2, 0	1, 1	3, 4		

Each box may be viewed as an epistemic game model, as explained earlier. Again, every step of announcement increases players’ knowledge, until a fixed-point is reached, constituting an equilibrium where each player knows as much as they can.

Strong rationality is a more demanding condition than weak rationality. While SR implies WR , there can be moves that satisfy weak but not strong rationality. This shows in the following difference to the previous example. In the present matrix, announcing Weak Rationality stops after the first step elimination of action c . The reason is that, in the second matrix, the row player’s bottom move is not strictly dominated by any other action, so this row remains after re-announcing WR . However, under no possible circumstance is the row player’s bottom action best for her. This contradicts strong rationality and hence that row is eliminated by the next SR announcement. More generally, the game matrix resulting from $(SR)^\infty$ is a submatrix of that produced by $(WR)^\infty$. Notably, this is not entirely obvious, as the two update sequences may produce different epistemic models satisfying different formulas. A proof can be found in van Benthem (2014).

Just as with Backward Induction, there are connections between iterated announcements and fixed-point logics. The set of strategy profiles that survive iterated announcements of strong rationality can be defined in modal μ -calculi (Kozen, 1983; Venema, 2008). If announcements are generalized to arbitrary formulas, so-called deflationary fixed-point logics are needed for studying limit behavior (Ebbinghaus and Flum, 2005; Dawar et al., 2004).

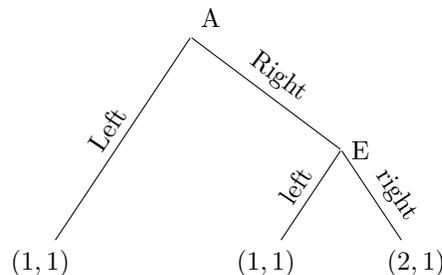
Further game solution concepts have been analyzed in a similar dynamic update style. The iterated regret minimization of Halpern and Pass (2012), for instance, has been captured in terms of iterated announcements (Bobbio and Cui, 2017).

It should be noted, that there are also more deductive takes on solution concepts, where successive inferences assume the role of the semantic updates above. A systematic proof-theoretic perspective on game-theoretic reasoning toward solution can be found in de Bruin (2005). Lastly, alternative analyses of Strong and Weak Rationality as well as other game solution concepts, in an abstract rewriting format of computational logic can be found in Apt (2005).

Digression: comparing across representation levels Different iterative announcement procedures lead to different analyses of games. Moreover when comparing various procedures across different frameworks, surprises may occur. For an illustration, take Backward Induction. Its dynamic analysis produced a new ‘best move’ relation or plausibility order on an extensive form game. The resulting strategy profiles may differ from a Nash equilibrium analysis of the associated strategic form game:

Example Backward Induction and Nash equilibria.

Consider the following game. E has no preferences between any outcomes, but A does, as marked by the utility values.



In the earlier BI analysis, neither move for A dominates the other in beliefs, so no move is eliminated. Now consider the two possible strategy profiles of each player and compute Nash equilibria:

$(Left, right)$ is not a Nash equilibrium, since A would do better by playing $Right$, but $(Left, left)$ is.

This illustrates differences in logical perspective on strategic and extensive form games. Primitive elements of the former, strategies, are complex objects in the latter’s game tree that cannot be identified completely at the level of individual nodes alone. A related point of connecting perspectives in classic game theory gave rise to the concept of subgame perfect Nash equilibria (Selten, 1975).

Further scenarios Casting game solution in terms of deliberation renders it an internal mental process: normally opponents do not sit down together to discuss their game play in advance, but reason about the opponents’ possible actions and

considerations. However, the deliberative techniques introduced above also apply to real conversational scenarios. An example related to game theory is the topic of disagreement, first introduced in an epistemic setting by [Aumann's 1976](#) seminal agreeing to disagree result. [Dégremont and Roy \(2012\)](#) investigate this topic with techniques of dynamic logic, building on classical results from [Geanakoplos and Polemarchakis \(1982\)](#). In this framework, any dialogue where agents keep stating whether or not they believe some formula φ leads to agreement in the limit model, where updates no longer have any effect. Briefly said, agents cannot disagree forever, at least when starting with different hard information, while sharing a well-founded plausibility order.

4.2 Information flow, knowledge, and belief during play

Game play is a dynamic process, where players repeatedly obtain new information about other players. Certain aspects of information collection are hard-wired into the game's structure, such as observing moves, or, in settings of imperfect observation, changing from one information state to another. Other updates may be extraneous, such as signals about the type of opponent one is dealing with. As of how, there is no general logical theory encompassing all these phenomena. Yet, instructive samples exist. The first topic to address concerns the players' knowledge, the second their belief.

4.2.1 Epistemic update and imperfect information

In one perspective, games annotated with imperfect information cells can be interpreted as recording a process of actual play. However, an imperfect information tree does not suffice to fully specify the trace of a real game. This raises the question on how to tease out what has really happened. One style of analysis involves techniques from [dynamic-epistemic logic](#). In this approach, players are assumed to have perfect recall, they do not forget anything they once knew, while also satisfying No Miracles: observation of actual game play is their only source of information, (cf. [Section 3.6](#)).

In a first approximation, every move triggers a public announcement, informing all players what just happened. Many games, however, include partially observable moves, where some players merely learn that an act has been performed, but not necessarily which. In this case, information processing requires *product updates* from dynamic-epistemic logic, allowing for an appropriate mixture of knowledge and uncertainty.

Example Decorating a game tree by updates.

The left-hand side of the following diagram displays the game's bare action structure, without any information on observability. However, when moving, players can distinguish their own actions, but not all moves of their opponents. Their precise observational powers are described by *event models* for the individual moves (cf. [van Ditmarsch et al., 2007](#)).

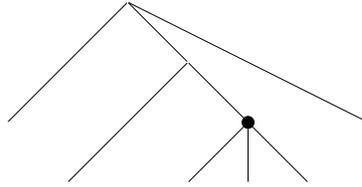
4.2.2 Belief revision and Forward Induction

Certain types of information may be judged inconclusive or not fully reliable. While unsuitable for advancing knowledge, such information may prompt agents to alter some of their beliefs. Such inconclusive evidence often concerns expectations concerning opponents' player types. Leaving aside the possibility of mere mistakes, all moves can be assumed to result from intentional, strategic considerations. By interpreting opponents' past moves, agents may hence infer about their beliefs, preferences, risk attitudes or reasoning types. Naturally, most such observations are not fully conclusive. The corresponding updates hence cannot delete any alternatives. Rather, they merely change the agent's plausibility ordering \leq_i among different options. Formally, this can be handled with the plausibility updates introduced for Backward Induction. However, the interpretation differs. Here, these updates do not represent steps in pregame deliberation, but result from actual moves during the game. Besides the radical upgrade introduced above, a number of further updating policies reflecting different attitudes to the information acquired are defined in [Baltag and Smets \(2008\)](#). The epistemic-plausibility patterns that can arise from systematic plausibility update in games have been identified in [Dégremont \(2010\)](#), using two counterparts of the earlier perfect recall and no miracles properties: 'Plausibility Revelation' and 'Plausibility Propagation'.

These results refer to but one aspect of belief in games. There are others. A further type of belief describes agents prior attitudes to the game, generated by past experience or deliberation. Another refers to the agents' beliefs about where they are located in the game tree, based on previous observations during play. To keep these notions separate, one might distinguish between more local 'beliefs' during play and future-oriented 'expectations' about the game's progression. Plausibility orders created by Backward Induction, for instance, describe expectations about future game play. These are not based on observations already made in the present game, and significantly, fail to satisfy the properties of plausibility revelation and propagation. Here is a particular strand of logical belief and its revision that is of independent game-theoretic interest.

Forward Induction Suppose some player has deviated from her Backward Induction strategy as computed in pre-game deliberation. What are others to make of this? Answers offered in the literature range from interpreting the deviation as an error without any future implications ([Aumann, 1995](#)) to treating it as significant in various ways ([Bicchieri, 1997](#)). In the latter vein, the deviation could be a signal for cooperation (believable or not), a sign of limited resources, or it could reveal other relevant information about the player's type.

More explicitly, the situation has the following aspects. At any stage of a game, players have several types of information, including their prior expectations of how the game would proceed and the perhaps surprising observations made along the way. If the game is to continue further, as in the state marked below, agents need to integrate both into expectations about the future course of the game.



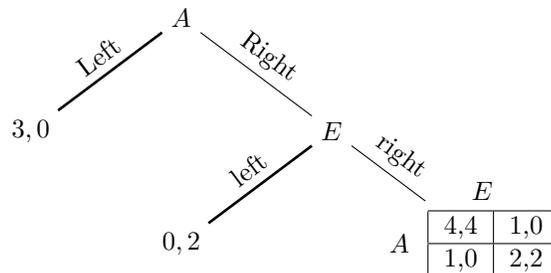
Rationalizing There is no unique best way of integrating information of the various kinds. Yet, a natural option is to maintain the assumption of opponents' rationality, taken in the earlier sense. Assuming preferences to be common knowledge, observed moves hence provide new information about the opponent's belief. More specifically, these beliefs have two components: expectations about what other players will do, and intentions about their own future actions. The driving principle will then be

Rationalizing By playing a move, a rational player communicates that this move is not strictly dominated-in-beliefs for her.

Clearly, rationalizing can only be maintained as long as the player does not choose a move that is strictly dominated under all circumstances. In that case, one must ascend a ladder of further hypotheses about the opponent, including the possibility of her making mistakes.

Reasoning policies of the above type are called *Forward Induction*. citetBSForward,BrandenburgerForward, analyze Forward Induction in extensive form games based on its known tight connection with Iterated Removal of Weakly Dominated Strategies in strategic form games. The following example involving explicit reasoning is from [Perea \(2012\)](#).

Example A Forward Induction scenario.



In the matrix game, no move dominates any other. Hence *E* should consider all outcomes possible. In this case, going *left* is safer for her than going *right*, and hence *A* should play *Left* at the start. However, if *E* rationalizes, and observes *A* going *Right*, she has extra information available at her choice node. Following the rationality assumption, *A* expects to do better than 3, which is only possible if he intends to play *Up* in the matrix game. Now this tells *E* to proceed to the matrix and play the *left* column therein. *E*'s results in a better payoff than the 2 of her original safe option.

From a logical perspective, a study of Forward Induction requires epistemic-doxastic models with ternary world-dependent plausibility relations, combined with the public announcement updates or plausibility upgrade described above (Section 4.1.2; see also van Benthem, 2014). No definitive logical analysis of Forward Induction has been published so far.

4.2.3 Post-game rationalization

Comparatively little attention has been paid in the literature to what players do after a game. Yet, these follow up-activities are often crucial, for instance to establish general lessons learned that may be valuable for future game play. Such interpretations are especially prominent in small or isolated groups, where the same opponent might be encountered again in the future.

Preference change after a game At a simple level, post-game activity can consist in setting, or altering, the second input parameter of rational choice beside belief: the players' preferences. Several folklore results relate to this option. For instance, when playing against a given strategy of another player with known preferences, any strategy can be rationalized by choosing suitable preferences among outcomes. Liu (2011) discusses several preference based rationalization algorithms using dynamic logics of preference change.

Preference change can also occur during a game. Players may receive new information about the game's end states and their properties. They may also follow a command or a suggestion from an authority, establishing a preference or reversing an earlier one. Relatedly, players may change their external goals pursued in the game, or they may adjust their preferences for more internal reasons, as in the phenomenon of 'sour grapes' (Elster, 2016).

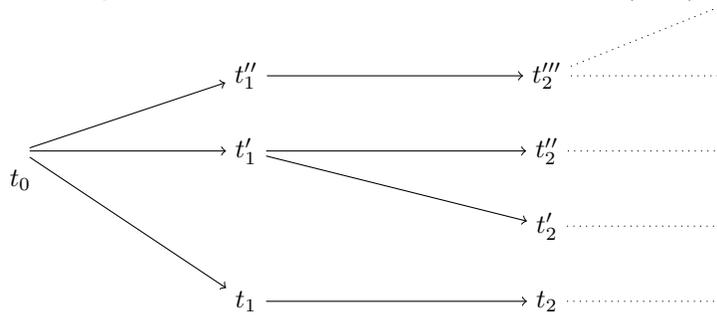
4.2.4 Play in a long-term temporal perspective

The main focus of this section is the local dynamics of what happens before, during and after a single game. There is also a broader perspective of time, in which all these activities are embedded in an extended temporal universe, large enough to include all possible trajectories of the game, finite or just as well infinite. In **evolutionary game theory**, in particular, infinite games typically arise from iterated play of finite games, with Lewis' signaling games (2008) as prominent example in philosophy.

Assuming an extended infinite temporal perspective raises additional questions about players' strategic foresight and adaptation (Christoff, 2016). Various long-term perspectives differ drastically in this respect, ranging from the minimal rationality assumptions typical of evolutionary game theory to high reasoning complexities of agents anticipating the long term effects of their choices.

Temporal logics While infinite game forms were briefly alluded to in Section 2.10, infinite play focusses on agency over time. A host of temporal logics for this end have been put forward, including interpreted systems (Fagin et al., 2004a), epistemic-temporal logic (Parikh and Ramanujam, 2003), STIT (Belnap and Perloff, 1988; Horty and Belnap, 1995), ATL (Alur et al., 2002; van der Hoek and Wooldridge,

2003), and others. Many of these systems combine multiple modalities. Consequentially, their complexity can be very high (undecidable, non-axiomatizable, or even non-arithmetical), as Halpern and Vardi (1986) show in a pioneering study for the case of combining time and knowledge. Surveying this area is beyond the scope of this entry, cf. the entry on temporal logics. A unifying view of connections between the various paradigms is presented in van Benthem and Pacuit (2006).



Just as with finite games, players’ preferences must be specified for studying equilibria in potentially infinite games. In lack of outcome nodes or final moments to attach preferences to, these are naturally thought of in terms of players’ goals, expressed as properties the game’s histories should satisfy. Such goals can be local propositional facts true at some particular moment. But goals can also concern global properties of histories such as avoiding or reaching the same position some specified number of times, or more abstractly, achieving safety or fairness in some appropriate sense. All such properties can be specified in temporal logics. For the case of Linear Temporal Logic (*LTL*), the ‘Boolean games’ of Gutierrez et al. (2015) have developed the temporal goal based approach in depth. Notably, this framework validates a logical version of the ‘folk theorem’ for iterated games, cf. Osborne and Rubinstein (1994): Under natural conditions on goals, iterated games can have novel equilibria not supervenient on the base game’s Nash equilibria. Further significant uses of temporal logics connect game theory with belief revision theory (Battigalli and Bonanno, 1999; Perea, 2012; Stalnaker, 1998).

Evolutionary game theory and dynamical systems A prominent application of iterated games occurs in evolutionary game theory (Maynard Smith, 1982; Hofbauer and Sigmund, 1998; Gintis, 2000), a framework that has many applications in biology, formal sociology, but also linguistics and philosophy (Lewis, 2008; Skyrms, 2010; Alexander, 2007; Clark, 2011).

Little work has been done so far on the logical analysis of evolutionary games along the dimensions of this entry. In fact, there are striking conceptual differences between evolutionary games and the style of analysis pursued here that might be dubbed ‘high rationality’-oriented. Rather than incorporating intentional, strategic actors, evolutionary games work by temporal progression of a dynamical system which is driven by individuals’ fitness values derived from game-like encounters with others. Within such systems, behavior is not driven by belief updates or complex strategic considerations. Rather, players typically display ‘low rationality, following certain

hard-wired strategies. Much of the evolutionary system’s dynamics is then driven by changes in the population’s composition of strategy types.

Even so, evolutionary game theory does invite connections to logic. The evolutionary success of simple strategies like Tit-for-Tat (Axelrod and Hamilton, 1981) raises the question of just when complex logically based high-rationality strategies can be replaced by equally efficient alternatives simple enough to be played by automata or similar models of bounded agents, (Grädel et al., 2002). At a higher level of abstraction, there also is an incipient line of research into the connection between logic and dynamical systems, a standard tool for analyzing evolutionary games. This strand includes a bimodal topological logic of time (Kremer and Mints, 2007), fixed-point logics of oscillation (van Benthem, 2015), and a systematic linkage between dynamic-epistemic update logics and dynamical systems over metric spaces (Klein and Rendsvig, 2017). At a much concreter level, one important species of evolutionary games are signaling games (Lewis, 2008; Cho and Kreps, 1987; Osborne and Rubinstein, 1994; Skyrms, 2010; van Rooij, 2004), where agents send and receive signals about the state of the world. Signaling games match up naturally with the earlier dynamics logic of information flow during play.

4.3 Conclusion: theory of play

Topics discussed in this section are less standard in the literature than those of the sections before. In an orthodox reading, various of the aspects addressed would not be considered part of game theory proper. The extended agenda followed here has been embraced by Pacuit et al. (2011) as a larger program for logic, going under the heading ‘Theory of Play’. The underlying line of reasoning is that games do not fully determine their outcomes, as they allow for various styles of play. Hence, it might be the process of play itself, including players’ types and how they change over time, that might be the best focus for understanding interaction, rather than mere games or game forms alone. Similar lines of argument can be found in the foundations of computation where it has been proposed that the essential topic of study should be *behavior* (Abramsky, 2008).

4.4 Further directions

Belief revision and learning theory Belief revision in repeated games bears natural resemblance to limit learning of formal learning theory (Kelly, 1996). Baltag et al. (2011) analyze learning in terms of initial epistemic-doxastic models over which finite histories of signals trigger the learner to revise beliefs, represented as changes in epistemic accessibility or plausibility order. It turns out that both iterated public announcement and iterated radical upgrade as discussed above are universal learning methods, though only the latter maintains this property in the presence of (finitely many) errors in the input stream.

Goal dynamics and intentions While preferences and goals have so far been assumed fixed and universally known, this is by no means necessary. van Otterloo (2005) presents a dynamic logic of strategic powers, where information about players’ intentions and preferences can be announced during play. Roy (2008) uses

announcements of intentions to obtain simplified solution procedures for strategic games. More concrete scenarios of extraneous information flow are found in [Parikh et al. \(2013\)](#), where agents manipulate the knowledge of others during play.

Game change In many real life scenarios, players do not know the full game tree they are playing. Even if they did, it might change during play. Or, at least, players may attempt to change the game. A concrete example is provided by the game tree in Section [4.1.1](#). There, the inefficient Backward induction outcome $(1, 0)$ could be avoided by E promising not to go left. When made binding (for instance through imposing a fine) this announcement eliminates histories and, consequentially, a new backward Induction outcome of $(2, 2)$ results. Hence, both players can be made better off by restricting the freedom of one. Game theory has sophisticated analyses of such scenarios, including an analysis of ‘cheap talk’ ([Osborne and Rubinstein, 1994](#)), asking when such announcements are credible. On the logical side, this suggests an analysis of signaling games ([van Rooij, 2004](#)). We are not aware of any logical work done in this direction

Real games The discrepancy between specification of a game and the realities of play is especially striking in real game play, either of the ‘natural kind’ in common parlor games ([van Ditmarsch and Kooi, 2015](#)), or of the artificial kind found in the laboratory experiments of experimental game theory ([Camerer, 2011](#)). Little work has been done by logicians in this realm, though there is a broad tradition of computational analysis of games, [Schaeffer and van den Herik \(2002\)](#); [Kurzen \(2001\)](#). Any adequate logical analysis would clearly need to incorporate the considerations on bounded agency discussed in Section [3](#).

Mathematical foundations Logic of play as discussed here raises issues of how to interface local with global dynamics. This shows in particular with logical limit behavior, where observations and assertions are made repeatedly. Limit models of public announcement, as described earlier, can be ‘self-fulfilling’ or ‘self-refuting’. In the first case, the property asserted becomes common knowledge among all agents, whereas in the second it eventually becomes false at the actual world. With soft update on plausibility models, a third option arises, namely, infinite oscillation, or even divergence ([Baltag and Smets, 2009](#)). To date, there is no general logical theory of these phenomena, but see related work see [van Benthem \(2011\)](#) on the use of fixed-point logics for limit models, [Miller \(2005\)](#) on the high complexity of public announcement logic with finitely iterated announcements, and [Klein and Rendsvig \(2017\)](#) on limit behavior of product updates.

The topic of limit behavior also raises the issue of how local dynamic logics of agency relate to the global temporal logics discussed in Section [4.2.4](#). Towards clarifying the connection, [Gerbrandy et al. \(2009a\)](#), show how dynamic-epistemic logics can be seen as decidable fragments of more expressive temporal logics. [Baltag et al. \(2009\)](#) discuss the related theme of how dynamic representations can decrease complexity by shifting information from the temporal universe to dynamic events.

5 Interfacing Logic and Probability in Games

Probabilities are central to game theory, where they serve two prominent roles. First, they structure players' uncertainty about various aspects of the game, including the state of nature, the type of opponent faced, and past, present, and future moves of other players. Second, ever since the origins of Game Theory ([von Neumann and Morgenstern, 1944](#)), probabilistic randomization has served to expand the agents' space of possible actions. While the interpretation of such mixed strategies has been the subject of extensive debate ([Sugden, 1991](#)), it is uncontroversial that randomized moves add significant depth to the analysis of games. In fact, the concept of mixed strategies is vital for a number of seminal results in classic game theory including the existence of Nash equilibria in finite games of imperfect information.

Probability and its logic is a major topic in both mathematics and philosophy, as discussed in the entries on [interpretations of probability](#), and [logic and probability](#). The current section merely outlines a few key connections between probability and the logical analysis of games. The following presentation assumes that all state spaces are finite, ignoring important technical and conceptual issues around the transition from finite to infinite state spaces.

5.1 Probabilities and beliefs

Probabilistic methods are widely employed to represent beliefs of agents within and outside of games, witness [Bayesian epistemology](#). Typically, a probabilistic belief model consists of two components: a space of possible states that the agent's beliefs range over, plus a quantitative probability function denoting how probable the agent judges different propositions or states to be. In qualitative logical models, on the other hand, agentive belief is represented in varying degree of detail. The coarsest approach only distinguishes states the agent considers possible from those ruled out. This is the perspective of standard epistemic-doxastic logics, such as multi-agent S5 and KD45 discussed in [Section 3.6](#). More fine-grained perspectives are employed in the plausibility models of [Boutilier \(1994\)](#) and [Baltag and Smets \(2008\)](#), where the range of epistemic options is structured further by plausibility orderings encoding which options agents take to be less or more likely. See also [Sections 3.6](#) and [4.2](#).

Both probabilistic and plausibilistic perspectives can express that some alternative is more likely than another. However, there are also conceptual differences between the two frameworks. Probabilistic models can aggregate, allowing their logic to express, for instance, whether many low-probability worlds combined can outweigh even the highest-probability worlds. Such aggregated probabilities play a key role, for instance, in calculations of expected utility. Yet no such thing can be expressed in plausibility semantics. For another striking difference, plausibility models lead to a notion of belief that is closed under conjunction. This conjunction closure typically fails for probabilistic accounts of belief. See however [Leitgeb \(2017\)](#) for a sophisticated bridge between both types of modeling. This work is one instance of more general contacts between logic and probability that have revived recently.

5.2 From logic to probability and back

On a received view, logical and probabilistic frameworks emphasize different aspects of belief. Logic emphasizes coherence properties between the propositions believed, such as closure under logical implications or conjunction. Probabilistic reasoning, on the other hand, stresses graded information and attitudes towards uncertain events such as lotteries. Even so, there is a variety of approaches attempting to unify both types of reasoning by constructing bridges between the frameworks.

From qualitative to quantitative probability Early attempts in this direction go back to [De Finetti \(1974\)](#), striving for a purely qualitative axiomatization of probability theory. A step further towards logical reasoning are various theories of qualitative probability ([Kyburg, 1994](#)), often based on the assumption that classical quantitative notions of probability are too demanding for real-life agents. In this line of research agents need only reason with partial, comparative probability assessments, rather than having fully specified probabilities for all events. Logical frameworks for qualitative probability include [Seegerberg \(1971\)](#); [Fagin et al. \(1990\)](#) and [Delgrande and Renne \(2015\)](#). While differing in detail, all logics in this line have in common that they allow for expressions of the form $\varphi \preceq \psi$, indicating that ψ is judged at least as probable as φ . Recent frameworks add various additional refinements to this language. Similarly, [Heifetz and Mongin \(2001\)](#) expand the axiomatic analysis of probabilistic beliefs to higher order reasoning, working on probabilistic type space akin to those introduced in section 3.7.

Yet, the concept of qualitative probability is sometimes considered flawed: a complete set of probabilistic-logical principles guaranteeing that every complete description in the logical language corresponds to a unique probability measure turns out to require a complex calculus, involving an opaque infinite rule, ([Kraft et al., 1959](#); [Scott, 1964](#)). A new angle has been proposed in [Harrison-Trainor et al. \(2016\)](#) through axiomatizing a low-complexity qualitative probabilistic logic that emerges from relaxing the above unique correspondence requirement to merely requiring compatibility with all probability measures of a certain family.

However, none of the frameworks described here are specific for games. In fact, it remains to be seen whether qualitative probabilistic logics, old or recent, can be used for a qualitative analysis of game-theoretic solution concepts.

From quantitative to qualitative probability While the preceding line of research aims at recovering quantitative probability from qualitative notions, a converse project shows how ubiquitous qualitative patterns might arise naturally within a quantitative probabilistic setting. Building on what is sometimes dubbed the Lockean Thesis, *threshold approaches* connect probability to logic by stipulating that some φ is to be believed simpliciter if the probability of φ is above some appropriate numerical threshold t . For most choices of threshold t , such translations do not square well with standard logical desiderata, as beliefs will in general not be closed under conjunction. However, in recent work [Leitgeb \(2017\)](#) and [Lin and Kelly \(2012\)](#) have identified conditions under which one can do better, building on ideas of [Skyrms \(1977\)](#), [Leitgeb](#) identifies strong, context-dependent ‘robustness conditions’ on thresholds that guarantee the defined belief operator to satisfy the KD45 axioms after all. [Lin and Kelly](#), on the other hand, work with non-uniform thresholds for

transitioning between logical and probabilistic notions of belief, allowing to derive coherence between different forms of belief dynamics on the two sides. For a recent study of the mathematical foundations and limits of these approaches, as well as the conditional logics they generate see [Mierzewski \(2018\)](#).

5.3 Updating and tracking probabilities

In the dynamic perspective of game play, every move constitutes a new piece of information agents have to take into account. Besides, players may also change their beliefs about the game upon deliberation, through communication or any other signals they receive, be it more or less reliable (cf. Section 4). All such dynamic events raise the question of how new information is to be incorporated into the agents' beliefs and when probabilistic updates have corresponding logical revisions or vice versa.

If the new information is of the hard type, accepted as irrevocably true by all agents, the probabilistic counterpart of logical public announcements is Bayesian conditioning. Both notions track each other at a semantic level, meaning that their outputs amount to the same thing. Computing beliefs after a public announcement means recomputing in the submodel consisting of all states where the information received was true. This exactly is the same mechanism as for recalculating probabilities in Bayesian update. Moreover, for reasoning about updates, both approaches require conditional notions: conditional belief and conditional probability respectively. The quantitative notion of conditional belief relates to the logical notion of conditional belief $B(\varphi|\psi)$, with the slight caveat that the latter also allows for epistemic or doxastic operators inside either argument. More refined logical notions of conditional belief arise in the earlier-mentioned plausibility semantics of Section 4.1.2, ([Baltag and Smets, 2008](#)).

Given the co-existence of qualitative and quantitative perspectives, it makes sense to ask whether one can *track* the other. In a static sense, tracking asks whether different notions of belief can be translated into each other by omitting or transforming some of the semantic details involved. A dynamic interpretation expands on this by asking whether updating, either in the hard or soft varieties of Section 4, is compatible with these translations in a commutative diagram: Information update in a new perspective after translation should yield same result as first performing a matching update in the old perspective and then translating ([van Benthem, 2016](#)). In games, the topic of tracking may refer not just to information update, but also to solution concepts or moves in game play described at the various levels considered in Section 2.

The existence of tracking maps depends on the exact type of update considered. By now there is a wide variety of updating policies on plausibility models ([van Benthem and Smets, 2015](#)), not all of which have obvious probabilistic counterparts. Likewise, for well-known varieties of probabilistic update, such as Jeffrey update, where the probability of selected propositions can be reset at will, plausibility counterparts are not easy to find, though attempt [Gerbrandy et al. \(2009b\)](#) modifies dynamic-epistemic logic to allow for Jeffrey update and other generalized probabilistic policies.

5.4 Specializing to games

Virtually all aspects of game theory provide contacts between logical and probabilistic perspectives. Clearly, this is true for the different representations of knowledge, beliefs and their dynamics just discussed. Other contacts occur at the level of game forms, cf. Section 2, where probabilities enrich the space of strategies, and, correspondingly, where mixed moves require players to expand their preferences to mixed outcomes, cf. Section 3. Finally, at the level of reasoning about game play, cf. Section 4, probabilistic beliefs play a role in solution techniques such as dominance or expected utility based reasoning.

Available actions and mixed strategies Probabilistic mixtures of pure strategies are prominent in game theory, as they can secure outcomes and payoffs that no pure strategy alone could guarantee.

Example Matching Pennies.

Consider the well-known game of matching pennies in matrix form:

		Bob	
		x	y
Ann	a	1,-1	-1, 1
	b	-1, 1	1,-1

For Ann, a mixed strategy of playing a half of the time guarantees an expected outcome of 0, no matter what Bob does. No pure strategy could have achieved this.

From a logical point of view, mixed strategies can be treated as new primitive actions in the earlier logics of games (cf. Section 2). Yet, this treatment immediately renders the set of available actions infinite. A cautiously refined logical language, extending logical approaches to qualitative probability, can allow for expressions such as an agent playing action a with probability at least q , (Delgrande and Renne, 2015).

A more challenging general question is how to compare classic probabilistic approaches with their fixed-point results and ensuing equilibrium existence theorems, with the logical fixed-point approaches mentioned at several places in this entry. The latter operate with step-by-step ordinal iterations, as opposed to the gradual, approximative procedures that underlie the Brouwer or Kakutani fixed-point theorems relevant for classical game theory. A related question is just how much logic is needed to reproduce probabilistic existence theorems within a qualitative framework.

Adding players' preferences Once preferences are added, mixed strategies trigger additional intricacies. A strategy profile where some players pursue mixed strategies does not produce a unique outcome, but a weighted combination of outcomes. Thus, permitting mixed strategies requires lifting preference relations to probabilistic mixtures of outcomes or strategy profiles. Incorporating such mixtures may implicitly depart from the standard, purely qualitative perspective on outcomes (Ramsey, 2016; Savage, 1954).

Example Extended preference comparison.

The following two games are equivalent in terms of qualitative (i.e. ordinal) preferences between outcomes for both players. However, they differ in preferences between mixed outcomes, with $0.5(b, x) + 0.5(b, y) \succeq_A (a, x)$ holding in the game to the left, but not in that to the right.

	x	y
a	1,0	2,1
b	0,1	4,0

	x	y
a	1,0	2,1
b	-10,1	4,0

Going this way poses some logical challenges. For example, consider a preference relation over probabilistic mixtures of outcomes, where $m^t(a, b)$ stands for obtaining a with a probability of t , and b otherwise. This setting is in the scope of [von Neumann and Morgenstern's 1944](#) well-known ‘continuity axiom’ that is characterized by an implicit infinite disjunction:

$$a \preceq b \preceq c \Rightarrow \exists t \in (0, 1) : m^t(a, c) \preceq b \preceq m^{1-t}(a, c)$$

This seems well beyond the expressive power of standard probabilistic logics.

Solution concepts and game play Probabilization also impact the process of game play and its reasoning dynamics, for instance by changing the earlier-mentioned calculus of weak and strong dominance (Section 3.4). Consider the following game, cf. [de Bruin \(2005\)](#):

		B	
		x	y
A	a	0,5	5,0
	b	5,0	0,5
	c	1,1	1,1

None of A 's strategies are dominated in terms of pure strategies. However, in terms of expected outcomes, c is dominated by an equal mixture of a and b . Thus, solution procedures analyzed earlier such as iterated removal of weakly or strongly dominated strategies may provide different and incompatible outcomes, depending on whether mixed strategies are considered or not. No satisfactory logical analysis of the earlier kind seems to exist for this setting.

5.5 Further challenges for a logical approach

Further challenges to the interplay of logic and probabilistic reasoning abound. By way of conclusion, here is a dimension that seems hard to capture in purely qualitative logical terms. A characteristic feature of game- and decision-theoretic reasoning is that beliefs and preferences are *entangled* in various ways ([Liu, 2011](#)). For instance, the crucial notion of expected utility entangles probability, standing for beliefs, with utilities, representing preferences. Players faced with probabilistic uncertainty about the opponent's present and future actions are often advised to maximize expected

utility (von Neumann and Morgenstern, 1944; Savage, 1954). Hence, even if one has found qualitative counterparts for probabilistic belief and cardinal utility separately, entanglement poses the additional difficulty of merging these two qualitative analyses in a way that matches what the quantitative side achieves easily by forming some arithmetical combination of both components.

6 Gamification

The main topic of this entry is a logical approach to game theory, bringing classical notions and methods from logic to bear upon games. This project is sometimes called ‘logic of games’. There also is a converse direction of ‘logic as games’, where game theoretic concepts are employed to elucidate basic notions of logic. This section presents a brief discussion on this direction as a natural counterpoint to the main lines of the entry. For an extensive survey see the entries on [logic and games](#) and [games, abstraction and completeness](#).

6.1 Logic games

Many notions in logic have been analyzed in game-theoretic terms

Evaluation games There are well-known two-player games for evaluating a first-order formula φ within a given logical model. These games are played between Verifier and Falsifier, who can both test atomic assertions, and specify the value of variables from a given domain (Hintikka, 1973). The schedule of the game is determined from the syntactic structure of the formula φ . Disjunctions and existential quantifiers require choices for the Verifier, conjunctions and universal quantifiers for the Falsifier, and negations trigger a role switch between the two players. The result is the following match between winning strategies and the ordinary semantic notion of truth in the following sense:

Formula φ is true in model M under assignment s iff
the Verifier has a winning strategy in the associated game $game(M, s, \varphi)$.

Correspondingly, Falsifier has a winning strategy if the formula φ is false in the model. Evaluation games turn out to be an extremely flexible tool. By suitably modulating rules and winning conventions, adequate evaluation games can be found for most logical systems. Doing so, however, can be a highly non-trivial task, as witnessed by the intricate infinite ‘parity games’ corresponding to fixed-point logics such as the modal μ -calculus (Venema, 2008). For the present purpose, it should be noted that this style of analysis ties the very logical operations, conjunctions, disjunctions, modal operators, to natural moves in a game. Similarly, the notion of truth is linked to the fundamental game-theoretic notion of a strategy in an extensive form game (cf. Section 2): a complex, structured object which may here be understood as a reason or an explanation for the truth or falsity of the formula.

The links between both perspectives are so close, that valid principles of logic come to express game-theoretic facts. For instance, after a little analysis, the law of excluded middle implies that always either Verifier or Falsifier has a winning strategy. In

other words, logical evaluation games for classical logic are determined in the game-theoretic sense. In fact, This property extends to most games for non-classical logics.

Further logic games Logic games exist for many other purposes. Ehrenfeucht-Fraïssé games serve model comparison (Ehrenfeucht, 1961; Ebbinghaus and Flum, 2005), Lorenzen games perform proof analysis (Kamlah and Lorenzen, 1984) and tableau games execute model construction (Hodges, 1985). In each case, strategies in the game match important logical notions. In Lorenzen dialogue games, for instance, winning strategies for the Proponent of a claim correspond to proofs of that claim from premises granted by the Opponent, whereas winning strategies for the Opponent are constructions of counter-models. Thus, proofs and models, two quite distinct notions in logic, co-exist within a single game.

There exists an alternative, game-theoretic way of interpreting these connective results. Suppose the game under study is fixed, and associated with some sort of ‘game board’ representing major features of the game’s general state (think of Chess, though more abstract game boards may occur). Then the above equivalences suggest that winning strategies, i.e. a typical game-theoretic notion defined in terms of the complete extensive game tree, is equivalent to a simpler ‘invariant’ that can be defined entirely in terms of some game board associated with the tree’s nodes. Identifying useful such invariants is a well-known art in the analysis of concrete games. In terms of a main theme of this entry, invariants can live at different levels of representation associated with a given class of games.

Game semantics One can view logic games as mere didactic devices analyzing logical notions that were already well-understood. Or, in other terms, as offering a concrete way of teaching logic that draws on game-theoretic intuitions. However, logic games have more to offer. First of all, new logics are suggested by pursuing natural variations in winning conventions, moves, or scheduling within existing logic games. Moreover, viewing logical operations as game constructors suggests a new, refined view on logical constants. Conjunction, for instance, now splits naturally into a sequential and a parallel version. Similar examples of parallelism also exist in logics of computation. Moreover, associating quantifiers with object picking, as in evaluation games, turns quantifiers into special types of atomic games that connect to the following formula by an abstract operation of game composition. The general logic of this abstract composition operation combined with propositional operations of choice and switch has been shown decidable, providing a new decidable core logic inside first-order logic whose existence had not been suspected (van Benthem, 2014). Games, hence, can offer a fresh perspective on existing logical systems.

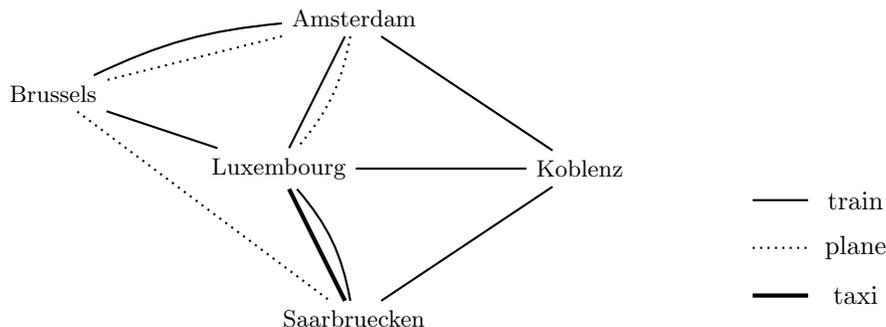
A major source of independent, game-theoretic perspectives on logic is the game semantics of computational logics. In this setting, the status of logic games may change. Rather than being a mere pedagogical or exploratory device, to some, these games are considered the true meaning of logical constants.

6.2 Recent logic-related games

The distinction between game logics and logic games is not always sharp. Recent literature has seen a number of games whose design is connected to logic, yet they are not meant to analyze logical notions per se.

Example Sabotage Games.

Sabotage games were proposed to analyze algorithmic tasks in adverse circumstances. Consider the below network between some European cities:



It is easy to travel either way between Amsterdam and the German town of Saarbruecken. Now, let a malevolent Demon start canceling connections in the network. At every stage, let the Demon take out one link, while the Traveler can afterwards follow one of the remaining links. This turns a one-agent planning problem into a two-player sabotage game. Zermelo-style reasoning shows that, from Saarbruecken, a German Traveler still has a winning strategy, while in Amsterdam, the Demon has the winning strategy against the Dutch Traveler, by first cutting a link close to Saarbruecken. The symmetry of the original search problem is broken.

The sabotage game has been applied to a variety of scenarios, including learning (Gierasimczuk et al., 2009), and communication networks (Aucher et al., 2018). On finite graphs, the game is clearly determined, with the computational complexity of identifying who has a winning strategy being Pspace-complete (Löding and Rohde, 2003).

The existence of this winning strategy can be expressed by a first-order formula. More specifically, winning conditions can be defined in a bimodal logic that combines a standard modality for travel steps with a new modality for one-step arrow deletion, interpreted in models $M = (W, R, V)$:

$$M, s \models [-]\varphi \quad \text{For each edge } (u, v) \text{ in } R : (W, R - \{(u, v)\}, V) \models \varphi$$

This logic fits the sabotage game closely. On top, it is a natural fragment of the first-order language of graphs. Surprisingly, this logic is undecidable (Löding and Rohde, 2003), making it one of the simplest examples of an undecidable modal logic over arbitrary models.

Further graph games in a similar spirit exist, including the poison game of Duchet and Meyniel (1993), where the Demon poisons nodes, rather than deleting edges. Extensive studies of modal logics for changing graphs, and μ -calculi for defining generic solutions to graph games are given in Areces et al. (2011, 2015); Aucher et al. (2018). A classification of graph games including the effects of complex goal formulas and imperfect information is found in van Benthem and Liu (2018).

One perspective on such logics for reasoning about model change is the semantic games approach of Section 6.1. In standard evaluation games, the initial model does not change. Modalities for model change, however, require a process of formula evaluation where the model of evaluation changes as, say, witnesses for quantifiers are not replaced (unlike in standard semantics for first-order logic), or moves change facts by causing damage to accessibility relations. In other contexts, similar modalities are justified by physical measurements that change the phenomenon under investigation, (Hintikka, 2002; Renardel de Lavalette, 2001; Ågotnes and Wang, 2017). Such generalized form of semantics are of independent logical interest.

Example Knowledge Games.

New logical games also arise naturally within the dynamics of information, knowledge or belief as triggered by the process of game play (cf. Section 4). In particular, information update suggests conversation games between participants with similar or different goals. These games may be cooperative, with players aiming to pool their information, thereby turning distributed knowledge into common knowledge (Meyer and van der Hoek, 1995). But they can also be competitive, say, when players strive to be the first to know whether some relevant proposition holds. Mixtures between both modes also occur, for instance with some players aiming to communicate a fact that outsiders should not learn about (van Ditmarsch, 2003).

A concrete example are the ‘announcement games’ of Ågotnes and van Ditmarsch (2011). Players speak simultaneously and only once, while pursuing goal that are specified as epistemic formulas. Speaking is modeled by public announcement, and players preferences are binary. They prefer final models where their goal formula holds over those where it is false.

These games are conducted under imperfect information, as players may not know the true state of their epistemic model. Accordingly, the relevant strategies need to be uniform. Players must say the same thing in all states they cannot distinguish. In general, then, many solution concepts produce mixed strategy outcomes. In fact, it can be shown that there exists simple announcement games without any unique equilibrium in pure strategies. However, there is a relevant role for logic to play. Suppose that the goal statements are all ‘universal’, i.e. constructed from literals by applying only conjunction, disjunction, knowledge operators, and dynamic modalities with universal announcements. Truth of such formulas is preserved when transitioning from a model to a submodel. Consequently, epistemic uncertainty becomes less harmful and knowledge games with universal goals have equilibria in pure strategies. Recently, knowledge games have been expanded further to include both questions and answers as separate actions of issue change and information change (Ågotnes et al., 2011).

Example Boolean games.

A third example of game design in between game logics and logic games are the Boolean games of Harrenstein et al. (2001) and Gutierrez et al. (2015) that have been mentioned several times already. Each player is handed control over a subset of the propositional variables, and can pick truth values for these at will. Using goals specified in temporal logics, these games can model a surprising number of relevant

scenarios of agency. By now, a growing body of work addresses various aspects of Boolean games including their computational characteristics (both single-shot and iterated), their game-theoretic properties and equilibria (Gutierrez et al., 2015), and their connections with games played on social networks (Seligman and Thompson, 2015). Further discussion can be found in the entry on [coalitional powers](#).

6.3 Special topics

Back and forth between game logics and logic games The topic of this section suggests cycles between the two perspectives on logic and games. Given a logical system, one can design logical games for it, which can then again be studied using some appropriate game logic. Conversely, given a game, one can introduce a logic for describing it, and then introduce evaluation games for that logic, and so on. Sometimes these cycles reach fixed-points, where, say, the evaluation game for a formula describing some game is isomorphic to that game itself. But sometimes, the cycling continues. For discussion, see [Rebuschi \(2006\)](#); [van Benthem \(2014\)](#).

Imperfect information Logic games naturally support imperfect information, where players do not have complete access to what their opponents do. Epistemic variations can have far-reaching consequences for the corresponding logics. A particularly prominent framework among this lines is the [independence friendly logic](#) of [Hintikka and Sandu \(1989\)](#), see also [Hintikka and Sandu \(1997\)](#); [Mann et al. \(2011\)](#).

Argumentation games and graph games Another strand of game analysis with a connection to logic is the study of argumentation networks ([Dung, 1995](#); [Caminada and Gabbay, 2009](#)), with uses in AI and philosophy ([Grossi, 2013](#); [Shi, 2018](#))

Computational logic The material in this section is closely connected to games in computational logic, which serve to analyze expressive power of languages. For relevant results and connections with automata theory, see. [Grädel et al. \(2002\)](#) and [van Benthem \(2014\)](#).

Gaming and mechanism design Game design is a well-known aspect in the area of gaming ([Rouse and Ogden, 2000](#)). Likewise, mechanism design is an established topic in game theory ([Nisan and Ronen, 1999](#); [Osborne and Rubinstein, 1994](#)). For connections between between logic, game design and planning, see ([Löwe et al., 2011](#); [Löwe, 2008](#)).

7 Summary and Further Directions

This entry presents an overview of current work at the interface of logic and games. The topics surveyed fall in a number of strands including current logical analysis of games in the broadest sense, contacts between logic and classic game theory, connections with probability and with computation, and, lastly, the game theoretic content of logic itself. All this produced a perhaps bewildering variety of logical systems. Yet, this entry emphasized the coherence of different approaches to logical analysis, some ‘zooming in’ on particular aspects in detail, others ‘zooming out’, thereby focussing on general patterns. What has hopefully become clear in this way

is that the various topics span a highly interconnected field. For further details, see the various game-related entries in this Encyclopedia and the literature cited.

In terms of further desiderata, the coarse grained perspective of logical modeling may help discover new abstract, general patterns in social behavior beyond the details of games and game theory. A concrete example might be a description of general skills and insights that players acquire by playing a given type of games. A more foundational example would be a formalization of the insight that all concrete social interaction rest on underlying general notions of ‘dependence’ with their own high-level logic, cf. the semantic dependence logics of [Väänänen \(2007\)](#); [Baltag \(2018\)](#). For an alternative proof-theoretic approach in this style, relying on the general postulates for competitive games of [Johansen \(1982\)](#), see [Hu and Kaneko \(2012\)](#).

The interface area of logic and games still is in statu nascendi. Correspondingly, there are obvious gaps and desiderata on the logic side, which are reflected in the material in this entry.

In particular, a fundamental theme are syntactic perspectives on game-theoretic reasoning. Samples of a proof-theoretic style of analysis for play can be found in [de Bruin \(2005\)](#). More concretely, [Zvesper \(2010\)](#) analyzes classical results in epistemic game theory, (cf. [Tan and da Costa Werlang, 1988](#); [Aumann, 1999](#)), in terms of abstract modalities for belief and optimality, showing how a few simple proof rules from modal μ -calculus can capture the essence of famous results in epistemic game theory. Proof-theoretic aspects of logic have been so far overshadowed by semantic analyses, although this situation is changing slowly ([Artemov, 2014](#); [Kaneko, 2002](#); [Kaneko and Suzuki, 2003](#)). Model-based reasoning provides abstract semantic perspectives on games that can aid conceptual clarification, and the discovery of general laws. But it might turn out to be proof theory that governs the concrete reasoning used in the semantics, and that may be able to guide the context of justification in establishing general facts about games.

A further contact with logic that has been ignored in this entry is the rich interface between games and descriptive set theory ([Woodin, 2010](#); [Kanamori, 2008](#)).

It should be stressed once more that logic is not the only formal discipline that throws light on games. Quantitative probability enters the study of games in many ways, both in classical and in evolutionary game theory. The interface of logic and games may well profit from the many old and new contacts between logic and probability ([Leitgeb, 2017](#); [Lin and Kelly, 2012](#); [Harrison-Trainor et al., 2016](#)).

Another link that remained underrepresented in this entry are computational aspects. The study of games, play and players has natural connections with computation and agency in computer science and AI ([Grädel et al., 2002](#); [Abramsky, 2008](#); [Halpern, 2003](#); [Perea, 2012](#); [Brandenburger, 2014](#)). The proper perspective on what has been presented here may well turn out to be a triangle of interfaces between logic, games, and computation.

As for still broader connections, we have not done justice to all links between logic, games and philosophy, of which more are found in [Stalnaker \(1996, 1998, 1999\)](#). The same is true for links to linguistics and psychology ([Clark, 2011](#)). In this language-oriented connection, one should also mention the work of [Bjorndahl et al. \(2017\)](#)

on the natural language used in specifying games and reasoning about them, thus making game analysis more description-dependent.

Finally, the main thrust of this entry is theoretical and foundational. However, there also is a more practical aspect to logic and games. Logic plays a role in cognitive psychology and experimental game theory, if only to identify testable hypotheses related to Theory of Mind or strategic reasoning (Ghosh et al., 2014; Ghosh and Verbrugge, 2017; Bicchieri, 1997; Fagin et al., 2004a). Lastly, some work at the interface of logic and games suggests outreach to the world of actual parlor games (van Ditmarsch and Kooi, 2015; van Benthem and Liu, 2018).

All in all, the claim of this entry is a modest one. Logic and games form a natural combination, that may reveal interesting things when pursued explicitly. Even so, too much logic may import too much of a formal apparatus, which may end up strangling the games perspective: logical systems are infinite machineries that can easily overwhelm a concrete game of interest. In short, the contact has to be managed with care.

8 Bibliography

- Abramsky, S. (1997). Semantics of interaction: an introduction to game semantics. *Semantics and Logics of Computation*, 14(1).
- Abramsky, S. (2008). *Information, processes and games*, pages 483–549. Handbook in the Philosophy of Science. North Holland.
- Ågotnes, T., van Benthem, J., van Ditmarsch, H., and Minica, S. (2011). Question-answer games. *Journal of Applied Non-Classical Logics*, 21(3-4):265–288.
- Ågotnes, T. and van Ditmarsch, H. (2011). What will they say?—public announcement games. *Synthese*, 179(1):57–85.
- Ågotnes, T. and Wang, Y. (2017). Resolving distributed knowledge. *Artificial Intelligence*, 252:1 – 21.
- Ågotnes, T. and Wooldridge, M. (2010). Optimal social laws. In *Proceedings of the Ninth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, Toronto, Canada.
- Alexander, J. M. (2007). *The Structural Evolution of Morality*. Cambridge University Press.
- Alexander, J. M. (2009). Evolutionary game theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2009 edition.
- Alur, R., Henzinger, T. A., and Kupferman, O. (2002). Alternating-time temporal logic. *Journal of the ACM (JACM)*, 49(5):672–713.

- Andrighetto, G., Brandts, J., Conte, R., Sabater-Mir, J., Solaz, H., and Villatoro, D. (2013). Punish and voice: Punishment enhances cooperation when combined with norm-signalling. *PLOS ONE*, 8(6):1–8.
- Anglberger, A., Gratzl, N., and Roy, O. (2015). Obligation, free choice, and the logic of weakest permissions. *The Review of Symbolic Logic*, 8(4):807–827.
- Apt, K. (2005). Order independence and rationalizability. In *Proceedings of the 10th conference on Theoretical aspects of rationality and knowledge*, pages 22–38. National University of Singapore.
- Areces, C., Fervari, R., and Hoffmann, G. (2015). Relation-changing modal operators. *Logic Journal of the IGPL*, 23(4):601–627.
- Areces, C., Figueira, D., Figueira, S., and Mera, S. (2011). The expressive power of memory logics. *The Review of Symbolic Logic*, 4(2):290–318.
- Areces, C. and ten Cate, B. (2007). *Hybrid logics*, volume 3 of *Studies in Logic and Practical Reasoning*, pages 821–868. Elsevier.
- Arlo-Costa, H. and Egr, P. (2016). The logic of conditionals. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition.
- Artemov, S. (2008). The logic of justification. *The Review of Symbolic Logic*, 1(4):477–513.
- Artemov, S. (2014). On definitive solutions of strategic games. In *Johan van Benthem on Logic and Information Dynamics*, pages 487–507. Springer.
- Artemov, S. and Fitting, M. (2016). Justification logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition.
- Aucher, G., Benthem, J. v., and Grossi, D. (2018). Modal logics of sabotage revisited. *Journal of Logic and Computation*, 28(2):269–303.
- Aumann, R. (1976). Agreeing to disagree. *The annals of statistics*, pages 1236–1239.
- Aumann, R. (1995). Backward induction and common knowledge of rationality. *Games and Economic Behavior*, 8(1):6–19.
- Aumann, R. J. (1999). Interactive epistemology i: Knowledge. *International Journal of Game Theory*, 28(3):263–300.
- Axelrod, R. and Hamilton, W. (1981). The evolution of cooperation. *science*, 211(4489):1390–1396.
- Baltag, A. (2018). The logic of conditional dependency. ILLC Amsterdam.
- Baltag, A., Christoff, Z., Rendsvig, R. K., and Smets, S. (2018). Dynamic epistemic logics of diffusion and prediction in social networks. *Studia Logica*.

- Baltag, A., Gierasimczuk, N., and Smets, S. (2011). Belief revision as a truth-tracking process. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK XIII, pages 187–190. ACM.
- Baltag, A. and Renne, B. (2016). Dynamic epistemic logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition.
- Baltag, A. and Smets, S. (2008). A qualitative theory of dynamic interactive belief revision. In *Logic and the foundations of game and decision theory (LOFT 7)*, volume 3, pages 9–58. Amsterdam University Press.
- Baltag, A. and Smets, S. (2009). Group belief dynamics under iterated revision: fixed points and cycles of joint upgrades. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 41–50. ACM.
- Baltag, A., Smets, S., and Zvesper, J. A. (2009). Keep hoping for rationality: a solution to the backward induction paradox. *Synthese*, 169(2):301–333.
- Battigalli, P. and Bonanno, G. (1999). Recent results on belief, knowledge and the epistemic foundations of game theory. *Research in Economics*, 53(2):149–225.
- Battigalli, P. and Siniscalchi, M. (2002). Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106(2):356–391.
- Belnap, N. and Perloff, M. (1988). Seeing to it that: a canonical form for agentives. *Theoria*, 54(3):175–199.
- van Benthem, J. (2001). Games in dynamic-epistemic logic. *Bulletin of Economic Research*, 53(4):219–248.
- van Benthem, J. (2003). Logic games are complete for game logics. *Studia Logica*, 75(2):183–203.
- van Benthem, J. (2007). Rational dynamics and epistemic logic in games”. *International Game Theory Review*, 9(02):377–409.
- van Benthem, J. (2011). *Logical dynamics of information and interaction*. Cambridge University Press.
- van Benthem, J. (2014). *Logic in games*. MIT press.
- van Benthem, J. (2015). *Oscillations, logic, and dynamical systems*, pages 9–22. College Publications.
- van Benthem, J. (2016). *Tracking Information*, pages 363–389. Springer International Publishing, Cham.
- van Benthem, J. and Gheerbrant, A. (2010). Game solution, epistemic dynamics and fixed-point logics. *Fundamenta Informaticae*, 100(1-4):19–41.
- van Benthem, J., Girard, P., and Roy, O. (2009). Everything else being equal: A modal logic for ceteris paribus preferences. *Journal of philosophical logic*, 38(1):83–125.

- van Benthem, J. and Liu, F. (2004). Diversity of logical agents in games. *Philosophia Scientiae*, 8(2):163–178.
- van Benthem, J. and Liu, F. (2007). Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 17(2):157–182.
- van Benthem, J. and Liu, F. (2010). *Deontic logic and changing preferences*, pages 1–39. College Publications.
- van Benthem, J. and Liu, F. (2018). Graph games and logic design. *Journal of Philosophy and Social Sciences*, page to appear.
- van Benthem, J. and Pacuit, E. (2006). The tree of knowledge in action: Towards a common perspective. In *Proceedings of Advances in Modal Logic (AiML)*.
- van Benthem, J. and Pacuit, E. (2011). Dynamic logics of evidence-based beliefs. *Studia Logica*, 99(1-3):61.
- van Benthem, J. and Smets, S. (2015). Dynamic logics of belief change. *Handbook of epistemic logic*, pages 313–393.
- Benz, A., Jäger, G., Van Rooij, R., and Van Rooij, R. (2005). *Game theory and pragmatics*. Palgrave Macmillan.
- Bergstra, J., Ponse, A., and Smolka, S. (2001). *Handbook of process algebra*. Elsevier.
- Bergwerff, G., Meijering, B., Szymanik, J., Verbrugge, R., and Wierda, S. (2014). Computational and algorithmic models of strategies in turn-based games. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, pages 1778–1883.
- Bezhanishvili, N. (2006). *Lattices of intermediate and cylindric modal logics*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Bezhanishvili, N., Enqvist, S., and van Benthem, J. (2018a). A new game equivalence, its logic and algebra. *Journal of Philosophic Logic*.
- Bezhanishvili, N., Enqvist, S., and van Benthem, J. (2018b). A propositional dynamic logic for instantial neighborhood semantics. *Studia Logica*, pages 1–33.
- Bicchieri, C. (1997). *Rationality and coordination*. CUP Archive.
- Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.
- Binmore, K. and Samuelson, L. (1992). Evolutionary stability in repeated games played by finite automata. *Journal of economic theory*, 57(2):278–305.
- Bjorndahl, A. and Halpern, J. (2017). From type spaces to probability frames and back, via language. In *Proceedings of the 16th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK XVI, pages 75–87. EPTCS.
- Bjorndahl, A., Halpern, J., and Pass, R. (2017). Reasoning about rationality. *Games and Economic Behavior*, 104(3):146–164.

- Bjorndahl, A., Halpern, J. Y., and Pass, R. (2013). Language-based games. In *Proceedings of the 14th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK XIV, pages 39–48.
- Blackburn, P., Rijke, M. d., and Venema, Y. (2001). *Modal Logic*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.
- Bobbio, F. and Cui, J. (2017). A plausibility model for regret games.
- Bonanno, G. (2004). Memory and perfect recall in extensive games. *Games and Economic Behavior*, 47(2):237 – 256.
- Boutilier, C. (1994). Modal logics for qualitative possibility theory. *International Journal of Approximate Reasoning*, 10(2):173 – 201.
- Brandenburger, A. (2007). The power of paradox: some recent developments in interactive epistemology. *International Journal of Game Theory*, 35(4):465–492.
- Brandenburger, A. (2014). *The Language of Game Theory: Putting Epistemics into the Mathematics of Games*. World Scientific Publishing.
- Broersen, J. (2009). A complete stit logic for knowledge and action, and some of its applications. In Baldoni, M., Son, T. C., van Riemsdijk, M. B., and Winikoff, M., editors, *Declarative Agent Languages and Technologies VI*, pages 47–59, Berlin, Heidelberg. Springer Berlin Heidelberg.
- de Bruin, B. (2005). Game theory in philosophy. *Topoi*, 24(2):197–208.
- Camerer, C. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.
- Caminada, M. and Gabbay, D. (2009). A logical account of formal argumentation. *Studia Logica*, 93(2-3):109.
- Cardone, F. (2017). Games, full abstraction and full completeness. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2017 edition.
- Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.
- Christoff, Z. (2016). *Dynamic Logics of Networks Information Flow and the Spread of Opinion*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Cinà, G. (2017). *Categories for the working modal logician*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Clark, R. (2011). *Meaningful games: Exploring language with game theory*. MIT Press.
- Clarke, E., Grumberg, O., and Peled, D. (1999). *Model checking*. MIT press.

- Dawar, A., Grädel, E., and Kreutzer, S. (2004). Inflationary fixed points in modal logic. *ACM Transactions on Computational Logic*, 5:282 – 315.
- Dégremont, C. (2010). *The Temporal Mind: Observations on the logic of belief change in interactive systems*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Dégremont, C., Löwe, B., and Witzel, A. (2011). The synchronicity of dynamic epistemic logic. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 145–152. ACM.
- Dégremont, C. and Roy, O. (2012). Agreement theorems in dynamic-epistemic logic. *Journal of Philosophical Logic*, 41(4):735–764.
- Delgrande, J. and Renne, B. (2015). The logic of qualitative probability. In *IJCAI*, pages 2904–2910.
- Demey, L., Kooi, B., and Sack, J. (2017). Logic and probability. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2017 edition.
- van Ditmarsch, H. (2000). *Knowledge Games*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- van Ditmarsch, H. (2003). The russian cards problem. *Studia Logica*, 75(1):31–62.
- van Ditmarsch, H. and Kooi, B. (2015). One hundred prisoners and a light bulb. In *One Hundred Prisoners and a Light Bulb*, pages 83–94. Springer.
- van Ditmarsch, H., van Der Hoek, W., and Kooi, B. (2007). *Dynamic epistemic logic*, volume 337 of *Synthese Library*. Springer.
- Duchet, P. and Meyniel, H. (1993). Kernels in directed graphs: a poison game. *Discrete mathematics*, 115(1–3):273–276.
- Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n -person games. *Artificial intelligence*, 77(2):321–357.
- Ebbinghaus, H.-D. and Flum, J. (2005). *Finite model theory*. Springer Science & Business Media.
- Ehrenfeucht, A. (1961). An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49(129-141):13.
- van Eijck, J. and Verbrugge, R. L. (2017). Formal approaches to social procedures. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 edition.
- Elster, J. (1988). The nature and scope of rational-choice explanation. In *Science in reflection*, pages 51–65. Springer.
- Elster, J. (2016). *Sour grapes*. Cambridge University Press.

- Fagin, R., Geanakoplos, J., Halpern, J. Y., and Vardi, M. Y. (1999). The hierarchical approach to modeling knowledge and common knowledge. *International Journal of Game Theory*, 28(3):331–365.
- Fagin, R. and Halpern, J. (1987). Belief, awareness, and limited reasoning. *Artificial intelligence*, 34(1):39–76.
- Fagin, R., Halpern, J., and Megiddo, N. (1990). A logic for reasoning about probabilities. *Information and computation*, 87(1-2):78–128.
- Fagin, R., Halpern, J., Moses, Y., and Vardi, M. (1997). Knowledge-based programs. *Distributed Computing*, 10(4):199–225.
- Fagin, R., Halpern, J., Moses, Y., and Vardi, M. (2004a). *Reasoning about knowledge*. MIT press.
- Fagin, R., Halpern, J., Moses, Y., and Vardi, M. (2004b). *Reasoning about knowledge*. MIT press.
- Fehr, E. and Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868.
- Fernández-Duque, D., Pacuit, E., and van Benthem, J. (2014). Evidence and plausibility in neighborhood structures. *Annals of Pure and Applied Logic*, 165(1):106–133. The Constructive in Logic and Applications.
- De Finetti, B. (1974). *Theory of probability*, volume 6. John Wiley & Sons.
- Galeazzi, P. and Lorini, E. (2016). Epistemic logic meets epistemic game theory: a comparison between multi-agent kripke models and type spaces. *Synthese*, 193(7):2097–2127.
- Galliani, P. (2017). Dependence logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 edition.
- Garson, J. (2016). Modal logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2016 edition.
- Gasser, R. (1996). Solving nine men’s morris. *Computational Intelligence*, 12(1):24–41.
- Geanakoplos, J. and Polemarchakis, H. (1982). We can’t disagree forever. *Journal of Economic Theory*, 28(1):192–200.
- Gerbrandy, J., Hoshi, T., Pacuit, E., and van Benthem, J. (2009a). Merging frameworks for interaction. *Journal of Philosophical Logic*, 38(5):491–526.
- Gerbrandy, J., Kooi, B., and van Benthem, J. (2009b). Dynamic update with probabilities. *Studia Logica*, 93(1):67–96.
- Ghosh, S., Meijering, B., and Verbrugge, R. (2014). Strategic reasoning: Building cognitive models from logical formulas. *Journal of Logic, Language and Information*, 23(1):1–29.

- Ghosh, S. and Verbrugge, R. (2017). Studying strategies and types of players: experiments, logics and cognitive models. *Synthese*.
- Gierasimczuk, N., Kurzen, L., and Velázquez-Quesada, F. (2009). Learning and teaching as a game: A sabotage approach. In He, X., Horty, J., and Pacuit, E., editors, *Logic, Rationality and Interaction (LORI 2009, Chongqing, China)*, pages 119–132. Springer.
- Gierasimczuk, N. and Szymanik, J. (2011). A note on a generalization of the muddy children puzzle. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge, TARK XIII*, pages 257–264. ACM.
- Gintis, H. (2000). *Game theory evolving: A problem-centered introduction to modeling strategic behavior*. Princeton university press.
- Girard, P. (2008). *Modal Logic for Belief and Preference Change*. PhD thesis, PhD thesis, Stanford University.
- Goranko, V. (2003). The basic algebra of game equivalences. *Studia Logica*, 75(2):221–238.
- Goranko, V. and Galton, A. (2015). Temporal logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2015 edition.
- Goranko, V., Jamroga, W., and Turrini, P. (2013). Strategic games and truly playable effectivity functions. *Autonomous Agents and Multi-Agent Systems*, 26(2):288–314.
- Gosh, S., Verbrugge, R., and van Benthem, J. (2015). *Models of Strategic Reasoning*. Springer.
- Grädel, E., Thomas, W., and Wilke, T., editors (2002). *Automata, logics, and infinite games*, volume 2500 of *Lecture Notes in Computer Science*. Springer.
- Grossi, D. (2013). Abstract argument games via modal logic. *Synthese*, 190(1):5–29.
- Grossi, D., Liu, F., and van Benthem, J. (2014). Priority structures in deontic logic. *Theoria*, 80(2):116–152.
- Grossi, D. and Turrini, P. (2012). Short sight in extensive games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 805–812. International Foundation for Autonomous Agents and Multiagent Systems.
- Gutierrez, J., Harrenstein, P., and Wooldridge, M. (2015). Iterated boolean games. *Information and Computation*, 242:53–79.
- Halpern, J. (2001). Substantive rationality and backward induction. *Games and Economic Behavior*, 37(2):425 – 435.
- Halpern, J. (2003). A computer scientist looks at game theory. *Games and Economic Behavior*, 45(1):114 – 131.

- Halpern, J. Y. and Pass, R. (2012). Iterated regret minimization: A new solution concept. *Games and Economic Behavior*, 74(1):184–207.
- Halpern, J. Y. and Vardi, M. Y. (1986). The complexity of reasoning about knowledge and time. In *Proceedings of the eighteenth annual ACM symposium on Theory of computing*, pages 304–315. ACM.
- Hamblin, C. L. (1970). *Fallacies*. Methuen.
- Hansson, S. O. (1990). Preference-based deontic logic (pdl). *Journal of Philosophical Logic*, 19(1):75–93.
- Hansson, S. O. (1998). *Revision of Belief Sets and Belief Bases*, pages 17–75. Springer Netherlands.
- Hansson, S. O. (2002). *Preference Logic*, pages 319–393. Springer Netherlands, Dordrecht.
- Harrenstein, P. (2004). *Logic in conflict*. PhD thesis, PhD thesis, Utrecht University.
- Harrenstein, P., van der Hoek, W., Meyer, J.-J., and Witteveen, C. (2001). Boolean games. In *Proceedings of the 8th conference on Theoretical aspects of rationality and knowledge*, pages 287–298. Morgan Kaufmann Publishers Inc.
- Harrison-Trainor, M., Holliday, W., and Icard, T. (2016). A note on cancellation axioms for comparative probability. *Theory and Decision*, 80(1):159–166.
- Harsanyi, J. (1967). Games with incomplete information played by bayesian players, i–iii. *Management science*, 14(3):159–182,320–334,486–502.
- Heifetz, A., Meier, M., and Schipper, B. C. (2006). Interactive unawareness. *Journal of economic theory*, 130(1):78–94.
- Heifetz, A. and Mongin, P. (2001). Probability logic for type spaces. *Games and Economic Behavior*, 35(1):31 – 53.
- Heifetz, A. and Samet, D. (1998). Knowledge spaces with arbitrarily high rank. *Games and Economic Behavior*, 22(2):260 – 273.
- Hendricks, V. and Symons, J. (2015). Epistemic logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2015 edition.
- van den Herik, J., Iida, H., and Plaat, A. (2011). *Computers and Games, 7th International Conference, Revised Selected Papers*. Springer.
- Herzig, A. and Lorini, E. (2009). A dynamic logic of agency i: Stit, capabilities and powers. *Journal of Logic, Language and Information*, 19(1):89–121.
- Herzig, A. and Schwarzentruher, F. (2008). Properties of logics of individual and group agency. *Advances in modal logic*, 7:133–149.
- Hintikka, J. (1962). *Knowledge and belief*. Cornell UP.

- Hintikka, J. (1973). *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic*. Oxford, Clarendon Press.
- Hintikka, J. (2002). Quantum logic as a fragment of independence-friendly logic. *Journal of Philosophical Logic*, 31(3):197–209.
- Hintikka, J. and Sandu, G. (1989). Informational independence as a semantical phenomenon. In *Studies in Logic and the Foundations of Mathematics*, volume 126, pages 571–589. Elsevier.
- Hintikka, J. and Sandu, G. (1997). Game-theoretical semantics. In van Benthem, J. and Ter Meulen, A., editors, *Handbook of logic and language*, pages 361–410. Elsevier.
- Hodges, W. (1985). *Building models by games*. Cambridge UP.
- Hodges, W. (2013). Logic and games. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2013 edition.
- van der Hoek, W. and Pauly, M. (2007). Modal logic for games and information. In Blackburn, P., van Benthem, J., and Wolter, F., editors, *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*, pages 1077 – 1148. Elsevier.
- van der Hoek, W. and Wooldridge, M. (2003). Cooperation, knowledge, and time: Alternating-time temporal epistemic logic and its applications. *Studia Logica*, 75(1):125–157.
- Hofbauer, J. and Sigmund, K. (1998). *Evolutionary games and population dynamics*. Cambridge university press.
- Horty, J. (2016). Epistemic oughts in stit semantics. In Broersen, J., Condoravdi, C., Nair, S., and Pigozzi, G., editors, *Deontic Logic and Normative Systems. Proceedings of DEON 2018*, pages 157–176.
- Horty, J. and Belnap, N. (1995). The deliberative stit: A study of action, omission, ability, and obligation. *Journal of philosophical logic*, 24(6):583–644.
- Horty, J. and Pacuit, E. (2017). Action types in stit semantics. *The Review of Symbolic Logic*, 10(4):617–637.
- Hotelling, H. (1990). Stability in competition. In *The Collected Economics Articles of Harold Hotelling*, pages 50–63. Springer.
- Hu, T.-w. and Kaneko, M. (2012). Critical comparisons between the nash noncooperative theory and rationalizability.
- Jech, T. (2006). *Set theory*. Springer.
- Johansen, L. (1982). On the status of the nash type of noncooperative equilibrium in economic theory. *The Scandinavian Journal of Economics*, pages 421–441.

- Kamlah, W. and Lorenzen, P. (1984). *Fallacies*. University Press of America.
- Kanamori, A. (2008). *The higher infinite: large cardinals in set theory from their beginnings*. Springer Science & Business Media.
- Kaneko, M. (2002). Epistemic logics and their game theoretic applications: Introduction. *Economic Theory*, 19(1):7–62.
- Kaneko, M. and Suzuki, N.-Y. (2003). Epistemic models of shallow depths and decision making in games: Horticulture. *The Journal of Symbolic Logic*, 68(1):163–186.
- Kelly, K. (1996). *The logic of reliable inquiry*. OUP.
- Klein, D., Marx, J., and Scheller, S. (2018). Rationality in context. *Synthese*.
- Klein, D. and Pacuit, E. (2014). Changing types: Information dynamics for qualitative type spaces. *Studia Logica*, 102(2):297–319.
- Klein, D. and Rendsvig, R. K. (2017). Convergence, continuity and recurrence in dynamic epistemic logic. In Baltag, A., Seligman, J., and Yamada, T., editors, *Logic, Rationality, and Interaction (LORI 2017, Sapporo, Japan)*, pages 108–122. Springer.
- Kooi, B. and Tamminga, A. (2008). Moral conflicts between groups of agents. *Journal of Philosophical Logic*, 37(1):1–21.
- Kozen, D. (1983). Results on the propositional μ -calculus. *Theoretical Computer Science*, 27(3):333 – 354. Special Issue Ninth International Colloquium on Automata, Languages and Programming (ICALP) Aarhus, Summer 1982.
- Kraft, C., Pratt, J., and Seidenberg, A. (1959). Intuitive probability on finite sets. *The Annals of Mathematical Statistics*, 30(2):408–419.
- Kremer, P. and Mints, G. (2007). Dynamic topological logic. In *Handbook of Spatial Logics*, pages 565–606. Springer.
- Kurzen, L. (2001). *Logic for social software*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Kyburg, H. (1994). Believing on the basis of the evidence. *Computational Intelligence*, 10(1):3–20.
- Leitgeb, H. (2017). *The stability of belief: How rational belief coheres with probability*. Oxford University Press.
- Lewis, D. (2008). *Convention: A philosophical study*. John Wiley & Sons.
- Leyton-Brown, K. and Shoham, Y. (2008). *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. Morgan and Claypool Publishers, 1st edition.
- Lin, H. and Kelly, K. (2012). Propositional reasoning that tracks probabilistic reasoning. *Journal of Philosophical Logic*, 41(6):957–981.

- Little, D. (2016). *New directions in the philosophy of social science*. Pickering & Chatto Publishers.
- Liu, F. (2011). *Reasoning about preference dynamics*, volume 354. Springer.
- Liu, F., Seligman, J., and Girard, P. (2014). Logical dynamics of belief change in the community. *Synthese*, 191(11):2403–2431.
- Liu, F. and Wang, Y. (2013). Reasoning about agent types and the hardest logic puzzle ever. *Minds and Machines*, 23(1):123–161.
- Löding, C. and Rohde, P. (2003). Solving the sabotage game is pspace-hard. In Rován, B. and Vojtáš, P., editors, *Mathematical Foundations of Computer Science 2003*, pages 531–540, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Lomuscio, A. and Ryan, M. (1998). Ideal agents sharing (some!) knowledge. In *Proceedings of the 13th European Conference on Artificial Intelligence*, pages 557–561. Wiley.
- Lomuscio, A., van der Meyden, R., and Ryan, M. (2000). Knowledge in multi-agent systems: Initial configurations and broadcast. *ACM Trans. Comput. Logic*, 1(2):247–284.
- Lorenzen, P. and Lorenz, K. (1978). *Dialogische logik*. Wissenschaftliche Buchgesellschaft.
- Lorini, E. (2018). In praise of belief bases: Doing epistemic logic without possible worlds. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2-7, 2018*.
- Löwe, B. (2008). Logic and the simulation of interaction and reasoning: Introductory remarks. In *Proceedings of the AISB 2008 Symposium on Logic and the Simulation of Interaction and Reasoning, Aberdeen 2008*, volume 1, pages iii–v.
- Löwe, B., Pacuit, E., and Witzel, A. (2011). Del planning and some tractable cases. In van Ditmarsch, H., Lang, J., and Ju, S., editors, *Logic, Rationality, and Interaction (LORI 2011, Guangzhou, China)*, pages 179–192. Springer.
- Mann, A., Sandu, G., and Sevenster, M. (2011). *Independence-Friendly Logic: A Game-Theoretic Approach*. Cambridge University Press.
- Maubert, B. (2014). *Logical foundations of games with imperfect information : uniform strategies. (Fondations logiques des jeux à information imparfaite : stratégies uniformes)*. PhD thesis, University of Rennes 1, France.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge university press.
- McNamara, P. (2014). Deontic logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2014 edition.

- Meyer, J.-J. and van der Hoek, W. (1995). *Epistemic Logic for AI and Computer Science*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.
- Mierzewski, K. (2018). *Probabilistic Stability: dynamics, nonmonotonic logics, and stable revision*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Miller, Joseph Sand Moss, L. (2005). The undecidability of iterated modal relativization. *Studia Logica*, 79(3):373–407.
- von Neumann, J. and Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton university press.
- Nisan, N. and Ronen, A. (1999). Algorithmic mechanism design. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 129–140. ACM.
- Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V., editors (2007). *Algorithmic game theory*. Cambridge University Press.
- Osborne, M. J. and Rubinstein, A. (1994). *A course in game theory*. MIT press.
- van Otterloo, S. (2005). *A Strategic Analysis of Multi-agent Protocols*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Pacuit, E. (2017). *Neighborhood semantics for modal logic*. Short Textbooks in Logic. Springer.
- Pacuit, E., Parikh, R., and Cogan, E. (2006). The logic of knowledge based obligation. *Synthese*, 149(2):311–341.
- Pacuit, E. and Roy, O. (2011). A dynamic analysis of interactive rationality. In van Ditmarsch, H., Lang, J., and Ju, S., editors, *Logic, Rationality, and Interaction*, pages 244–257, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Pacuit, E. and Roy, O. (2017). Epistemic foundations of game theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2017 edition.
- Pacuit, E., Roy, O., and van Benthem, J. (2011). Toward a theory of play: A logical perspective on games and interaction. *Games*, 2(1):52–86.
- Parikh, R. (1985). The logic of games and its applications. In Karplinski, M. and van Leeuwen, J., editors, *Topics in the Theory of Computation*, volume 102 of *North-Holland Mathematics Studies*, pages 111 – 139. North-Holland.
- Parikh, R., Loohuis, L., and Baskent, C. (2011). Epistemic norms. *Handbook on Deontic Logic and Normative Systems*, College Publications, London.
- Parikh, R. and Ramanujam, R. (2003). A knowledge based semantics of messages. *Journal of Logic, Language and Information*, 12(4):453–467.

- Parikh, R., Taşdemir, Ç., and Witzel, A. (2013). The power of knowledge in games. *International Game Theory Review*, 15(04):1340030.
- Paul, S. and Ramanujam, R. (2011). Neighbourhood structure in large games. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 121–130. ACM.
- Pauly, M. (2002). A modal logic for coalitional power in games. *Journal of logic and computation*, 12(1):149–166.
- Pauly, M. (2011). *Complexity in Interaction*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Pauly, M. and Parikh, R. (2003). Game logic: An overview. *Studia Logica: An International Journal for Symbolic Logic*, 75(2):165–182.
- Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050.
- Peleg, B. (1998). Effectivity functions, game forms, games, and rights. *Social Choice and Welfare*, 15(1):67–80.
- Perea, A. (2012). *Epistemic game theory: Reasoning and choice*. Cambridge University Press.
- Van De Putte, F., Tamminga, A., and Duijf, H. (2017). Doing without nature. In Baltag, A., Seligman, J., and Yamada, T., editors, *Logic, Rationality, and Interaction*, pages 209–223. Springer.
- Ramanujam, R. (2014). Logical player types for a theory of play. In Smets, S. and Baltag, A., editors, *Johan van Benthem on logic and information dynamics*, pages 509–528. Springer.
- Ramanujam, R. and Simon, S. (2008). A logical structure for strategies. In Bonanno, G., van der Hoek, W., and Wooldridge, M., editors, *Logic and the foundations of game and decision theory (LOFT 7)*, volume 3 of *Texts in Logic and Games*, pages 183–208.
- Ramsey, F. (2016). Truth and probability. In *Readings in Formal Epistemology*, pages 21–45. Springer.
- Rebuschi, M. (2006). If and epistemic action logic. In *The Age of Alternative Logics. Assessing Philosophy of Logic and Mathematics Today*, pages 261–281. Springer.
- Renardel de Lavalette, G. (2001). *A logic of modification and creation*, pages 197–219. CSLI.
- van Rooij, R. (2004). Signalling games select horn strategies. *Linguistics and philosophy*, 27(4):493–527.
- Ross, D. (2018). Game theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition.

- Rouse, R. and Ogden, S. (2000). *Game design theory and practice*. Wordware Publishing Inc.
- Roy, O. (2008). *Thinking before acting: intentions, logic, rational choice*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Roy, O., van Benthem, J., and van Otterloo, S. (2006). *Preference logic, conditionals and solution concepts in games*, pages 61–77. Uppsala Philosophical Studies 53.
- Samuelson, L. (1992). Dominated strategies and common knowledge. *Games and Economic Behavior*, 4(2):284–313.
- Savage, L. (1954). *The foundations of statistics*. Wiley & Sons.
- Schaeffer, J. and van den Herik, J., editors (2002). *Chips challenging champions*, volume 134. North Holland.
- Scott, D. (1964). Measurement structures and linear inequalities. *Journal of mathematical psychology*, 1(2):233–247.
- Seegerberg, K. (1971). Qualitative probability in a modal setting. In *Studies in Logic and the Foundations of Mathematics*, volume 63, pages 341–352. Elsevier.
- Seligman, J. and Thompson, D. (2015). Boolean network games and iterated boolean games. In van der Hoek, W., Holliday, W. H., and Wang, W.-f., editors, *Logic, Rationality, and Interaction*, pages 353–365. Springer.
- Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International journal of game theory*, 4(1):25–55.
- Shi, C. (2018). *Reason to Believe*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Shoham, Y. and Leyton-Brown, K. (2008). *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.
- Skyrms, B. (1977). Resiliency, propensities, and causal necessity. *The Journal of Philosophy*, 74(11):704–713.
- Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.
- Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford University Press.
- Stalnaker, R. (1996). Knowledge, belief and counterfactual reasoning in games. *Economics & Philosophy*, 12(2):133–163.
- Stalnaker, R. (1998). Belief revision in games: forward and backward induction. *Mathematical Social Sciences*, 36(1):31–56.
- Stalnaker, R. (1999). Extensive and strategic forms: Games and models for games. *Research in Economics*, 53(3):293–319.

- Stalnaker, R. C. (1968). A theory of conditionals. *American Philosophical Quarterly*, pages 98–112.
- Steele, K. and Stefansson, H. O. (2016). Decision theory. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition.
- Sugden, R. (1991). Rational choice: A survey of contributions from economics and philosophy. *The Economic Journal*, 101(407):751–785.
- Tan, T. C.-C. and da Costa Werlang, S. R. (1988). The bayesian foundations of solution concepts of games. *Journal of Economic Theory*, 45(2):370 – 391.
- van der Torre, L. (1997). *Reasoning about obligations: defeasibility in preference-based deontic logic*. PhD thesis, Erasmus University Rotterdam.
- Tulenheimo, T. (2017). Independence friendly logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 edition.
- Turrini, P. (2016). Computing rational decisions in extensive games with limited foresight. In *AAAI*, pages 630–636.
- Turrini, P. (2017). Logics for analyzing power in normal form games. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2017 edition.
- Väänänen, J. (2007). *Dependence logic: A new approach to independence friendly logic*, volume 70. Cambridge University Press.
- Venema, Y. (1998). Rectangular games. *Journal of Symbolic Logic*, 63(4):1549–1564.
- Venema, Y. (2003). Representation of game algebras. *Studia Logica*, 75(2):239–256.
- Venema, Y. (2008). Lectures on the modal μ -calculus. ILLC Amsterdam.
- Wang, Y. (2010). *Epistemic modelling and protocol dynamics*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.
- Woodin, W. H. (2010). *The axiom of determinacy, forcing axioms, and the nonstationary ideal*, volume 1. Walter de Gruyter.
- Wooldridge, M. (2009). *An introduction to multiagent systems*. John Wiley & Sons.
- Zvesper, J. (2010). *Playing with information*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam.