11
Forcing Powers

In this chapter, we propose a simple view of global input-output behavior of games through the lens of players’ strategic powers. These can be studied with earlier modal techniques, leading to systematic connections with the logics of actions in Parts I and II.

11.1 Forcing and strategic powers

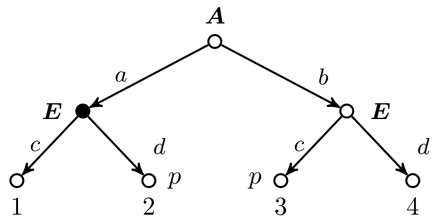
Social powers of agents are an important topic in their own right. In this chapter, we will ignore details of players’ moves, and concentrate on their powers for achieving outcomes of a game. Intuitive computations of this sort were already made in the Introduction and in Chapter 1 when discussing different views of games.

Definition 11.1 Forcing relations
The *forcing relation* $\rho^i_G s, X$ in a game tree holds if player i has a strategy for playing game G from state s onward whose resulting end states are always in the set X. When s is known in context (often it is the root of the game), the sets X are called the *powers* of player i. ■

In any finite game tree, we can find forcing relations by enumerating strategies, and sets of final states that can be reached by applying the strategy to all plays by the other player.[136] Determining powers can be done inductively using game algebra, but we defer this until Chapter 19. For now, we compute powers by hand.

136 The concept of powers also applies to infinite games, where outcomes can be identified with the total histories. Much of what we will have to say extends to this setting.

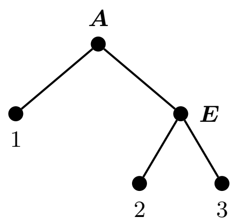
Example 11.1 Computing forcing relations
The following game displays the basic ingredients of strategies and outcome sets:

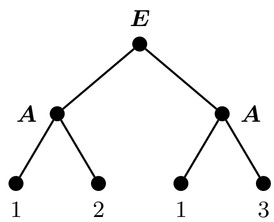


Here player $\boldsymbol{A}$ has two strategies, forcing the sets $\{1, 2\}$, $\{3, 4\}$, while $\boldsymbol{E}$ has four strategies, forcing the sets $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$. These sets encode players' powers over specific outcomes: a larger set says that the power is not strong enough for controlling a unique outcome, just keeping things within some upper range. ■

For a further perspective, we recall the earlier issue of game equivalence.

Example 11.2 Propositional distribution once again
Consider the following games, discussed in the Introduction and Chapter 1:





These games were not the same in their local action structure as studied in Chapter 1, but they are the same in terms of global powers. In the left-hand game, player $\boldsymbol{A}$ has two strategies, *left* and *right*, that guarantee outcomes in the sets $\{1\}$, $\{2, 3\}$. These are then player $\boldsymbol{A}$'s powers. Player $\boldsymbol{E}$ has two strategies as well, that guarantee outcomes in $\{1, 2\}$ and $\{1, 3\}$, making these sets $\boldsymbol{E}$'s powers. In the right-hand game, we found the same outcomes. First, player $\boldsymbol{E}$ has two strategies that force the sets $\{1, 2\}$ and $\{1, 3\}$. Next, $\boldsymbol{A}$ has four strategies LL, LR, RL, RR that yield, respectively, $\{1\}$, $\{1, 3\}$, $\{2, 1\}$, $\{2, 3\}$. Of these four, $\{1, 3\}$ and $\{2, 1\}$ can be dropped, however. For, on the above understanding of the notion, they represent weaker powers than $\{1\}$, and hence they are redundant. Thus player $\boldsymbol{A}$, too, has the powers $\{1\}$, $\{2, 3\}$, just as in the left-hand game. ■

This analysis presupposes some formal properties of players' powers, such as closure under supersets. We will now look at these constraints in some more generality, following a logical line linking up with other themes in this book.[137]

11.2 Formal conditions and game representation

Forcing relations as defined above are evidently closed under supersets:

$C1$ if $\rho^i_G s, Y$ and $Y \subseteq Z$, then $\rho^i_G s, Z$

This property is sometimes called Monotonicity. Another obvious constraint on these relations is Consistency: players cannot force the game into disjoint sets of outcomes, since playing strategies for these powers always produces an outcome:[138]

$C2$ if $\rho^i_G s, Y$ and $\rho^j_G s, Z$, then Y and Z overlap

Recall from Chapter 1 that all finite two-player games are determined: for any winning convention, one of the two players has a winning strategy. In the present abstract terms, this says the following. Let S be the total set of outcome states:

$C3$ if not $\rho^i_G s, Y$, then $\rho^j_G s, S - Y$

One can also think of this determinacy as a sort of Completeness.

It is easy to check these conditions for the earlier examples in this chapter. In fact they hold for players' general powers in our present setting.[139]

FACT 11.1 Any finite perfect information game G between players i and j yields powers in the root s satisfying conditions $C1$, $C2$, $C3$.

Conversely, these conditions are all that hold, by the following representation.

FACT 11.2 Any two families F_1, F_2 of subsets of some set S satisfying conditions $C1$, $C2$, and $C3$ are the root powers in some two-step game.

---

137 An early formal analysis of powers from a game-theoretic perspective was given in Bonanno (1992b).

138 Here and henceforth, we use i and j as generic names for the two players.

139 Interestingly, the stated conditions on powers may be viewed as a two-person version of the standard logical notion of an ultrafilter. This pulling apart of standard notions into many-agent notions will be very typical for the logic games to be studied in Part IV.

Proof Start with player i choosing between successors corresponding to all the inclusion-minimal sets in F_i. At these nodes, player j then gets to move, and can pick any member of the given set. Clearly then, player i has the powers specified in F_i. Now consider the powers of player j. In the game just defined, player j can force any set of outcomes that overlaps with each of the sets in F_i. But by the given constraints, these are precisely the sets in the initial family F_j. For instance, if some set of outcomes $\boldsymbol{A}$ overlaps with all the sets in F_i, then its complement, $S - A$, cannot occur in the latter family, and hence A itself must have been in F_j, by Completeness. ■

This argument gives a bit more than stated, that is to say, a normal form for games related to the usual game-theoretic strategic form that will be studied in Chapter 12. It just has two moves, and it does not matter to the construction which player begins. This is like the disjunctive and conjunctive normal forms of propositional logic. Indeed, the Boolean operations that produce such normal forms are part of a more general logical calculus of game equivalence that will be studied in Parts V and VI of this book.

Implementing social scenarios Our representation theorem may also be viewed in a different light. A list of powers is a specification for how much control we want to give to various agents in a given group, and a game is then the design of a simple social scenario implementing just those powers. For more complex power structures analyzed in logical terms, see Pauly (2001) and Goranko et al. (2013). Social scenarios where ways of giving agents different information may be part of the design are discussed in Papadimitriou (1996).

11.3 Modal forcing logic and neighborhood models

Forcing modalities One can introduce a matching modal language for games at this new level, with proposition letters, Boolean operators, and the following new modal operators:

$$\boldsymbol{M}, s \models \{G, i\}\varphi \quad \text{iff} \quad \begin{array}{l}\text{there exists a set power } X \text{ for player } i \text{ in } G \text{ such that} \\ \text{for all } t{\in}X\text{: } \boldsymbol{M}, t \models \varphi\end{array}$$

or more in line with the format of modal semantics:

$$\boldsymbol{M}, s \models \{G, i\}\varphi \quad \text{iff} \quad \text{there exists a set } X \text{ with } \rho^{\boldsymbol{i}}_G s, X \text{ and } \forall t \in X\text{: } \boldsymbol{M}, t \models \varphi$$

This is essentially the forcing modality $\{i\}\varphi$ that we have defined in Chapters 1 and 3 as a global perspective on some given extensive game G. In what follows, for convenience, we often write $\{G\}\varphi$ for the modality since the player i is understood from context.

This semantics involves a generalization of the modal models of Chapter 1, where accessibility took worlds to worlds, since we now work with world-to-set relations. The main effect of this change at the level of validities is reminiscent of the base logic of temporal forcing in Chapter 5.

FACT 11.3 Modal logic with the forcing interpretation satisfies all principles of the minimal modal logic of $\Diamond$ except for distribution of $\{\}$ over disjunctions.

That is, $G\varphi$ is upward monotonic:

$$\text{if} \models \varphi \to \psi, \text{ then} \models \{G\}\varphi \to \{G\}\psi$$

But the following is not valid:

$$\{G\}(\varphi \lor \psi) \;\to\; \{G\}\varphi \lor \{G\}\psi$$

The latter failure is precisely the point of forcing. Other players may have powers that keep me from determining results precisely. I may have a winning strategy, but it may still be up to you just where my victory will take place. For instance, in the very first game of this chapter, player $\boldsymbol{A}$ can force the set of outcomes $\{1, 2\}$, but neither $\{1\}$ nor $\{2\}$. It is also easy to check that the forcing modality $\{\}$ does not distribute over conjunction.

The basic modal neighborhood logic is decidable, with lower computational complexity than the minimal modal logic K over standard relational models.

Neighborhood modal logic Models with world-to-set accessibility relations for modal logics are called neighborhood models (Chellas 1980, van Benthem 2010b). The term reflects a connection with topology: see Section 11.5 below. These structures have many applications (Hansen et al. 2009), including fine-grained evidence models for the information dynamics of Chapter 7 (van Benthem & Pacuit 2011).

11.4 Bisimulation, invariance, and definability

Here is a key model-theoretic feature of neighborhood models that goes back to the issues of game equivalence discussed in the Introduction and later on in this book.

The invariance notion of bisimulation from Chapter 1 lifts to this setting without major effort. Consider any game model $\boldsymbol{M}$ plus forcing relations as defined above.

DEFINITION 11.2 Power bisimulation
A *power bisimulation* between two game models G, G' is a binary relation E between states in G, G' that satisfies the following two conditions:

(a) If xEy, then x and y satisfy the same proposition letters.
(b) If xEy and $\rho^{\boldsymbol{i}}_{G}x, U$, then there exists a set V with $\rho^{\boldsymbol{i}}_{G}y, V$ and $\forall v \in V \exists u \in U : uEv$; and vice versa. ■

Power bisimulation is a natural notion. It was proposed independently in concurrent dynamic logic (van Benthem et al. 1994), topological modal logics (Aiello & van Benthem 2002), and co-algebra (Baltag 2002). It also has a general model theory that is a natural extension of that in Chapter 1 for ordinary relational bisimulation. Without going into great detail, we give two illustrations.

FACT 11.4 The modal forcing language is invariant for power bisimulation.

Proof Consider two models $\boldsymbol{M}$ and $\boldsymbol{N}$ with $\boldsymbol{M}, s \models \{\}\varphi$ and sEt. By the truth definition, there is a set U with $\rho^{\boldsymbol{i}}_{\boldsymbol{M}}s, U$ and for all $u \in U$: $\boldsymbol{M}, u \models \varphi$. Now by the back-and-forth clause (2), there is a set V in $\boldsymbol{N}$ with $\rho^{\boldsymbol{i}}_{\boldsymbol{N}}t, V$ and $\forall v \in V \exists u \in U : uEv$. So, every $v \in V$ is E-related to some $u \in U$, and by the inductive hypothesis, we then have that $\boldsymbol{N}, v \models \varphi$. But then by the truth definition, $\boldsymbol{N}, t \models \{\}\varphi$. ■

FACT 11.5 If finite models $\boldsymbol{M}, x$ and $\boldsymbol{N}, y$ satisfy the same forcing formulas, then there is a power bisimulation E between them with xEy.

Proof Define a relation E between states in the two models as follows:

uEv iff $\boldsymbol{M}, u$ and $\boldsymbol{N}, v$ satisfy the same modal forcing formulas.

CLAIM E is a power bisimulation.

The atomic clause is clear from the definition. Now, suppose that sRt, while also, for some subset U of $\boldsymbol{M}$, $\rho^{\boldsymbol{i}}_{\boldsymbol{M}}s, U$. We need to find a set V with

$$\rho^{\boldsymbol{i}}_{\boldsymbol{N}}t, V \text{ and } \forall v \in V \exists u \in U : uEv$$

Suppose that no such set exists. That is, for every set V in $\boldsymbol{N}$ with $\rho^{\boldsymbol{i}}_{\boldsymbol{N}}t, V$, there is a state $v^V \in V$ that is not E-related to any $u \in U$. Let us analyze the latter statement further. By the definition of the relation E, this means that for each

$u \in U$, v^V disagrees with u on some forcing formula ψ^u: say, it is true in u, and false in v^V. But then, the disjunction Ψ^V of all of these formulas is true in every member of U, and still false in v^V. Now let Ψ be the conjunction of all the latter formulas, where V runs over all sets satisfying $\rho^{\boldsymbol{i}}_{\boldsymbol{N}} t, V$. Evidently, we have

$$\boldsymbol{M}, u \models \Psi \text{ for each } u \in U$$

and hence also

$$\boldsymbol{M}, s \models \{\}\Psi$$

But then, by the above definition of E, also $\boldsymbol{N}, t \models \{\}\Psi$. This says that there is a set V with $\rho^{\boldsymbol{i}}_{\boldsymbol{N}} t, V$ each of whose members satisfies the formula Ψ. This contradicts the construction of Ψ, since v^V certainly does not satisfy its conjunct Ψ^V. ■

These results also hold over more general process models, not just game trees.[140]

11.5 Digression: Topological models and similarity games

Let us now take a more spatial look at the preceding notions, injecting some fresh air into what might otherwise be a rather dry exercise in generalized modal logic.

Topological semantics A special case of neighborhood models are *topological models* $M = (\boldsymbol{O}, V)$ consisting of a topological space $\boldsymbol{O}$ and a valuation V for proposition letters. The forcing relation $\rho\ x, U$ is now read as follows: x belongs to the open set U. The semantics for modal languages on these mathematical structures dates back to the 1930s.

Definition 11.3 Topological interpretation of modal logic
We say that $\Box\varphi$ is true at a point s in a topological model $\boldsymbol{M}$ in case s has an open neighborhood U all of whose points satisfy φ. ■

We refer to the literature for details on the topological reading of the modalities (see the survey van Benthem & Bezhanishvili 2007 or the textbook van Benthem 2010b). For instance, the base logic is $S4$, and there are many deep technical results tying further modal logics to significant mathematical structures.

140 More results lie down the road, generalizing those stated in Chapter 1 for standard modal logic. For instance, we can find power invariants for any game G by infinitary modal forcing formulas defining the class of all models power bisimilar with G.

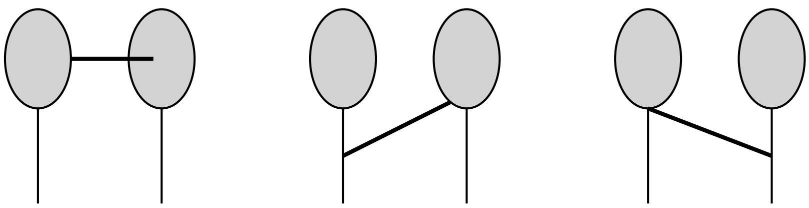
Here, however, our interest is in the relevant notion of structural invariance, connecting neighborhood bisimulation with the logic games of Part IV in this book. As we will see in Chapter 15, each notion of bisimulation invites a corresponding logic game for comparing models, where two players probe differences and similarities up to a specified number of rounds. This yields a fine structured version of bisimulation measuring degrees of similarity. Instead of doing this for power bisimulation in its abstract version on neighborhood models, we demonstrate the idea in a more concrete topological setting, taken from Aiello & van Benthem (2002).

Comparison games One can analyze the expressive power of our modal language with games between two players called *spoiler* (the difference player) and *duplicator* (the analogy player), comparing points in two given topological models.

DEFINITION 11.4 Topo-games
Rounds in *topo-games* proceed as follows, starting from some current match $s - t$. Spoiler takes one of these points, and chooses an open neighborhood U in its model. Duplicator responds with an open neighborhood V of the other current point. Still in the same round, Spoiler chooses a point $v \in V$, and then Duplicator chooses a point $u \in U$, making $u - v$ the new match. Duplicator loses if the two points differ in atomic properties. ■

This looks abstract, but it can be made very concrete.

EXAMPLE 11.3 Comparing points on spoons
In the spoons shown below, compare the following intuitively different points:



(a) (b) (c)

It is helpful to observe that, in these games, it does not matter whether players choose small or large open neighborhoods.[141]

141 To see this, recall the standard observation that, on relational models, evaluation of modal formulas at a current point s only needs to look at R-closed generated submodels

Case (a) If spoiler chooses a neighborhood to the left, then duplicator chooses a small interior disk to the right, and whatever spoiler chooses, there will be an inside point that duplicator can match in the open to the left. Therefore, this is a bad idea. If spoiler starts with a small disk on the right, however, then duplicator must respond with a disk on the edge to the left, which then allows spoiler to choose an object outside of the spoon, and every response by duplicator is losing, since it must be inside the spoon. Therefore, spoiler has a winning strategy in one round. This is reflected in the simple observation that there is a modal difference formula of operator depth 1 distinguishing the two positions. As a concrete example, $\Box p$ is true on the right, but not on the left.

Case (b) Spoiler's winning strategy starts by taking an open set on the handle to the left, after which duplicator must choose an open on the rim of the oval. Now spoiler chooses an object inside the spoon, and duplicator can only respond by either choosing outside of the spoon (an immediate loss), or on the handle. But the latter choice reduces the game to Case (a), which spoiler could already win in one round. A suitable difference formula this time has modal depth 2, say, $\Diamond\Box p$.

Case (c) is the most complex, as the point connecting rim and handle is much like an ordinary rim point. Spoiler has a winning strategy in three rounds, matching a modal difference formula of modal depth 3, say, $\Diamond(p \wedge \neg\Diamond\Box p) \wedge \Diamond\Box p$. ∎

The main result governing this game is similar to the "success lemma" for model comparison games found in Chapter 15.

FACT 11.6 Duplicator has a winning strategy in the comparison game over k rounds starting from two models $\boldsymbol{M}, s$ and $\boldsymbol{N}, t$ iff these two pointed models satisfy the same modal forcing formulas up to modal operator depth k.

The matching global relation for topological spaces plus propositional valuations is *topo-bisimulation*, that is, power bisimulation restricted to topological models.[142] Conversely, the topo-game is also easily adapted to general power bisimulation.

around s. Likewise, for any pointed topological model $\boldsymbol{M}, s$ and modal formula φ, and for any open neighborhood U of s, the following two assertions are equivalent: (a) $\boldsymbol{M}, s \models \varphi$, and (b) $\boldsymbol{M}|U, s \models \varphi$, where $\boldsymbol{M}|U$ is the model $\boldsymbol{M}$ restricted to the subset U.

142 The back-and-forth clauses resemble the topological definition of a continuous open function, for which both images and inverse images of open sets are open.

11.6 Compositional computation and game algebra

Powers in games can often be computed compositionally, as the reader may have done already when checking earlier examples. For instance, when a game starts with a move for player $\boldsymbol{E}$, the powers of $\boldsymbol{E}$ are just the union of the powers in the separate subgames that can be chosen. More abstractly, writing $G \cup H$ for an operation of choice between games G and H, we have the following valid equivalence:

$$\rho^{\boldsymbol{E}}_{G \cup H} s, Y \quad \text{iff} \quad \rho^{\boldsymbol{E}}_{G} s, Y \text{ or } \rho^{\boldsymbol{E}}_{H} s, Y$$

What we see here is an incipient algebra of game-forming operations. Natural operations forming new games are choice, but also role switch, sequential and parallel composition, that support identities of their own.[143] This algebra will emerge in Part IV of this book, and it will be explored further in Chapter 19, using logics with forcing modalities $\{G, i\}\varphi$ that can talk about composite game structure in their G argument.

11.7 Forcing intermediate positions

Powers tell us what players can achieve in the end. However, sometimes we also want to describe intermediate stages, getting closer to the earlier action level.

Example 11.4 Intermediate forcing
The local dynamics of the two distribution games in Example 11.2 are different. In the root on the left, but not on the right, player $\boldsymbol{A}$ can hand player $\boldsymbol{E}$ a choice between achieving q and achieving r. This might be expressed in a simple notation:

$$\{A\}^+(\{E\}^+q \wedge \{E\}^+r)$$

true in the left root, and false in the root on the right. The modified forcing modality $\{E\}^+\varphi$ in this formula says that player $\boldsymbol{E}$ has the power to take the game to some state, final or intermediate, where the proposition φ holds. ■

143 A simple example of a valid identity is that for role switch: $\rho^{\boldsymbol{E}}_{G^d} s, Y \leftrightarrow \rho^{\boldsymbol{A}}_{G} s, Y$.

This is a natural new intermediate level of game description: unlike bisimulation, intermediate forcing does not track specific actions, but it does care about choices of players on the move. It is easy to find a matching notion of bisimulation, and thus, we have one more level in the ladder of game equivalences discussed in Chapter 1.

11.8 Interfacing modal logics of forcing and action

Defining powers by actions The modal action languages of Chapter 1 were at least as expressive as the present forcing languages.

FACT 11.7 In any finite game model, $\{E\}\varphi$ is definable as a greatest fixed point of the recursion

$$\{E\}\varphi \leftrightarrow \big((\boldsymbol{end} \wedge \varphi) \vee (\boldsymbol{turn_E} \wedge \bigvee_a \langle a\rangle\{E\}\varphi) \vee (\boldsymbol{turn_A} \wedge \bigwedge_a [a]\{E\}\varphi)\big)$$

In specific finite models, we can even unwind $\{E\}\varphi$ as one modal action formula, whose depth depends on the size of the model. This is connected to the following general observation.

FACT 11.8 If there exists an action bisimulation between two models connecting points s and t, then there also exists a power bisimulation connecting s and t.

Games and game boards The two levels of description for games, in terms of moves and of powers, coexist when we make a natural distinction between games G themselves and associated structures that may be called "game boards" $\boldsymbol{M}(G)$ marking relevant external aspects of game states only. Game boards include the physical board states of a game like Chess, but they can also be abstract structures, such as the models where one plays the logical evaluation games of Chapter 14. Game boards will be studied in more depth in Chapters 18, 19, and 24, and a few simple observations will suffice for now.

Let G be a game of finite depth over a game board $\boldsymbol{M}(G)$, using moves available on the board, assigning turns to players, and determining a winning convention. This gives us an obvious projection map F from internal game states to external states on the board. Also, let the winning convention only use external facts on the game board, definable by a formula α:

$$\boldsymbol{win_E}(s) \text{ iff } \boldsymbol{M}, F(s) \models \alpha_{\boldsymbol{E}}$$

This sort of external winning convention holds for many games. Under these assumptions, the two levels are in harmony.

FACT 11.9 There is an effective equivalence between forcing statements about outcomes in a game G and modal properties of the associated game board $\boldsymbol{M}(G)$.

Proof If the game structure over atomic games is given by choices, switches, and compositions, we translate forcing statements at nodes into matching disjunctions, negations, and substitutions. At atomic games, we plug in the given definition. ■

We will see instances of this harmony in Chapters 14, 19, and 25.[144]

11.9 Powers in games with imperfect information

Having dealt with perfect information games, we now look at imperfect information. Chapter 3 studied such games at their action level, using a modal-epistemic language. Let us now consider the more global power level. Our treatment will be more sketchy than in preceding sections, showing merely how forcing views apply.

Powers First, the definition of players' powers is adapted as follows. Recall the basic notion of a uniform strategy assigning the same moves to nodes in the game tree that the relevant player cannot distinguish.

DEFINITION 11.5 Powers in imperfect information games
At each node of an imperfect information game, a player can *force* those sets of outcomes that are produced by following one of the player's uniform strategies. ■

EXAMPLE 11.5 Diminished powers in imperfect information games
Consider the two games drawn in Section 11.1. In the game of perfect information, player $\boldsymbol{A}$ had 2 strategies, and player $\boldsymbol{E}$ had 4, producing the following set powers:

$$\boldsymbol{A} \quad \{1,2\},\{3,4\} \qquad\qquad \boldsymbol{E} \quad \{1,3\},\{1,4\},\{2,3\},\{2,4\}$$

144 A useful example are graph games in computational logic (Libkin 2004), where players make alternating steps in their given move relations on a board. The first player to get stuck loses. There is an equivalence here between the assertions (a) player $\boldsymbol{E}$ has a winning strategy in the game at (G, s), and (b) $G, s \models (\langle R\rangle[S])^n\top$, with n the size of G.

With imperfect information added for $\boldsymbol{E}$ in the second game, $\boldsymbol{A}$'s powers are not affected, but $\boldsymbol{E}$'s are. $\boldsymbol{E}$'s two uniform strategies only yield two set powers:

$$\boldsymbol{A} \quad \{1,2\},\{3,4\} \qquad\qquad \boldsymbol{E} \quad \{1,3\},\{2,4\}$$

This loss of control may be seen as a weakness, but it really shows something else: the much greater power for representing varieties of social influence in imperfect information settings. ■

We can measure this power by extending the representation of Section 11.2.

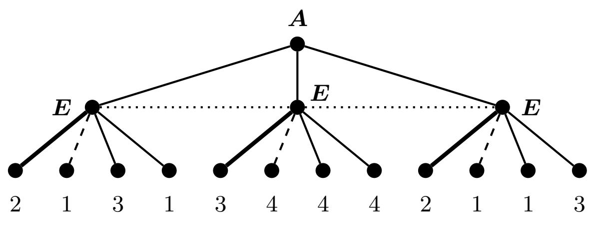
Representation theorem The earlier conditions of monotonicity $C1$ and consistency $C2$ still hold for powers in imperfect information games. What fails, however, is the completeness $C3$. As we just saw: player $\boldsymbol{A}$ lacks the power $\{1,4\}$, while $\boldsymbol{E}$ lacks its complement $\{2,3\}$.

THEOREM 11.1 Any two finite families of sets satisfying conditions $C1$ and $C2$ can be realized as the powers in a two-step imperfect information game.

Proof Instead of stating the procedure in complete formal detail, we go through an example displaying all the necessary tricks. We are given two lists:

minimal powers for player $\boldsymbol{A}$	$\{1,2,3\}$, $\{3,4\}$
minimal powers for player $\boldsymbol{E}$	$\{2,3\}$, $\{1,4\}$

Start with player $\boldsymbol{A}$ and put successor nodes for all of $\boldsymbol{A}$'s power sets, with possible duplications of these sets (as explained below), where player $\boldsymbol{E}$ gets to move. Here $\boldsymbol{E}$'s uniform strategy must prescribe the same move at each node, depicted below as lines with the same slope. $\boldsymbol{E}$'s actions, too, may involve duplications. First, take action types for each set of player $\boldsymbol{E}$, making sure all of these get represented via uniform strategies. There may still be excess outcomes in the powers for $\boldsymbol{E}$, and then dilute these by copying and permuting so that they end up in supersets of $\{2,3\}$, $\{1,4\}$. A picture conveys the idea better than words:



The reader may want to construct this representation by starting with just $\boldsymbol{A}$'s two left-most choices and three actions for $\boldsymbol{E}$, and then seeing why the additional actions have to be there. The sets for the third and fourth uniform strategy are all supersets of $\{2, 3\}$, $\{1, 4\}$.

A complication arises when $\boldsymbol{A}$'s sets involve outcomes not occurring in $\boldsymbol{E}$'s list. For concreteness, consider this case

minimal powers for player $\boldsymbol{A}$	$\{1, 2, 3\}$, $\{3, 4\}$
minimal powers for player $\boldsymbol{E}$	$\{2, 3\}$, $\{1, 4, 5\}$

We must dilute still more, adding $\boldsymbol{A}$-moves to cases where $\boldsymbol{E}$ can choose between 1, 4, and 5 so that the resulting powers of $\boldsymbol{E}$ are all extensions of $\{1, 2, 3\}$, $\{3, 4\}$.

The recipe A more precise description of the representation method is as follows. First, make sure that all atomic outcomes in players' sets occur for each of them, by adding redundant supersets if needed. Then create a preliminary branching for $\boldsymbol{A}$, making sure that all of $\boldsymbol{A}$'s sets are represented. Now represent each outcome set for $\boldsymbol{E}$ via a uniform strategy, by choosing enough branchings at midlevel, starting from the left. This step may involve some duplication of nodes, as illustrated by the case of initial lists $\boldsymbol{A}$ $\{1, 2\}$, $\boldsymbol{E}$ $\{1, 2\}$. Next, if there are still left over outcomes, repeat the following routine:

Suppose outcome i at midlevel point x was not used in our third step so far. Fix any outcome set for $\boldsymbol{E}$ produced by some uniform strategy σ. Say, σ chose outcome j at x. Duplicate this node x, and add two new branchings throughout (for all midlevel nodes considered so far) to get two further uniform strategies: one choosing j at x, and i in its duplicate, the other doing the opposite, and both following σ at all other nodes. This addition makes outcome j appear as it should for $\boldsymbol{A}$, while generating only a harmless superset of outcomes for $\boldsymbol{E}$.[145] ■

A variation One source of complexity in our construction is the fact that intersecting powers of the two players need not result in a singleton set with a unique outcome. If we add the latter condition, then we arrive at the matrix games of Chapter 12, where the forcing powers arise from simultaneous action, where players know their own choice, but can only observe the other player's choice afterward. We will discuss the latter scenario in connection with STIT logics of agency.

145 This method is very messy, and it would be nice to have a uniform construction, say, with pairs of actions and appropriate equivalence relations over these for the players.

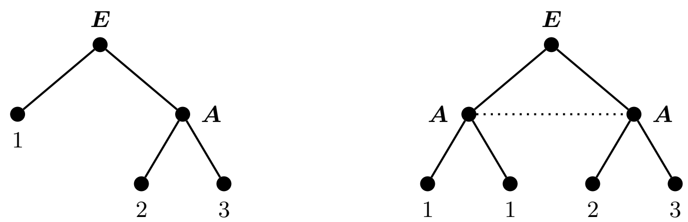
Game algebra As with powers in games of perfect information, the present setting suggests an algebra of game operations that respect power equivalence. For imperfect information games, however, compositionality with respect to subgames is much less obvious, and we postpone the topic until Chapters 20, 21, and 24.

11.10 Game transformations and extended logic

There are also direct links between our analysis and a classical topic in game theory. Let us take up a thread first started in Chapter 1.

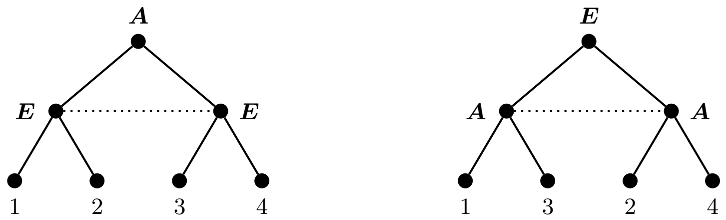
Game transformations and power invariance The four "Thompson transformations" of game theory (Thompson 1952, Osborne & Rubinstein 1994) turn normal form equivalent extensive games with imperfect information into each other. We show how this relates to our power analysis, returning to further logical background in Chapter 21.

EXAMPLE 11.6 Addition of a superfluous move
The first transformation, called addition of superfluous moves, tells us how to tolerate a new uncertainty by duplicating a move, switching game trees as follows:



Computing powers, nothing changes from left to right at the root. In the game to the right, the duplication adds no new outcomes for player $\boldsymbol{E}$, while the two uniform strategies for player $\boldsymbol{A}$ on the right capture the same outcome sets $\{1, 2\}$, $\{1, 3\}$ as on the left. ■

EXAMPLE 11.7 Interchanging moves
The key transformation interchanging moves of players occurred in the Introduction and elsewhere. We give a slight variant here, with some imperfect information added. In the following two games, the set powers for the players are the same:



Player $\boldsymbol{A}$ has powers $\{1, 2\}$, $\{3, 4\}$, while $\boldsymbol{E}$ has powers $\{1, 3\}$, $\{2, 4\}$. ■

This second Thompson transformation makes sequential play under imperfect information similar to simultaneous play, a topic that will return in Chapter 12.

The remaining two Thompson transformations are

(a) Inflation Deflation Adding a move that will not materialize in play.

(b) Coalescing Moves Re-encoding a zone where a player is to move successively as just one set of choices.

These reflect an emphasis on powers as the criterion of identity for a game, rather than imperfect information as such.

All of our observations so far are summarized in the following result.

THEOREM 11.2 Powers of players in imperfect information games remain the same under the Thompson transformations.’

This analysis suggests an analogue to Thompson’s result that his transformations are precisely those that leave the reduced form of a game unchanged. See Dechesne (2005) for further discussion, and connections with the independence-friendly logic of Chapter 21.

Game transformation as generalized logic A fascinating alternative interpretation is possible. The Thompson transformations induce a calculus for a generalized propositional logic. The above pictures of equivalent games relate to logical laws, the way we saw earlier with games and propositional distribution laws in the Introduction.

FACT 11.10 Addition of a Superfluous Move is propositional idempotence

$$A \wedge (B \vee C) \leftrightarrow (A \bigvee A) \wedge (B \bigvee C)$$

where the larger disjunctions indicate an uncertainty link.

We are into a nonstandard use of propositional logic now, where duplications matter, and operators can be linked. Another valid principle in this realm may be even more surprising.

FACT 11.11 Interchanging Moves is a propositional distribution law:

$$(A \bigvee B) \wedge (C \bigvee D) \leftrightarrow (A \bigwedge C) \vee (B \bigwedge D)$$

with large operators indicating dotted links, again a nonstandard feature.[146]

Thus, finite game trees with imperfect information are a natural extension of propositional formulas, leading to an intriguing calculus of equivalence in extended notation.[147] This theme will return in Parts IV and V of this book, showing how logical formulas themselves can act as game forms. But even without that more general background, it seems an interesting open problem to find a complete propositional logic of finite imperfect information games viewed as extended logical forms.

11.11 Forcing languages, uniform strategies, and knowledge

Again, there is a linguistic counterpart to power equivalence and outcome thinking. We need an epistemic version of the earlier forcing language (cf. Chapters 3 and 4), including knowledge operators, but now also with modified modalities

$\{G, i\}\varphi$ player i has a uniform strategy forcing only outcomes satisfying φ

This seems straightforward, but the epistemic setting hides a subtlety with outcomes. Indeed, the meaning of our modality in imperfect information games is not crystal clear.

EXAMPLE 11.8 Subtleties of interpretation
Suppose a player is in an end node x satisfying p, but cannot distinguish this from another end node y satisfying $\neg p$. Does the uniform strategy of doing nothing force p in the end node x, as we have always assumed so far? Yes, if we look at the

146 In the independence-friendly IF games of Chapter 21, this equivalence is expressed by $\forall x \exists y / \boldsymbol{x}\, Rxy \leftrightarrow \exists y \forall x / \boldsymbol{y}\, Rxy$, where slashes indicate informational independence.

147 Also striking is the occurrence character of the above rules. One occurrence of A need not have the same effect as several occurrences when dealing with uniform strategies. This is reminiscent of linear logic, whose game content will be explored in Chapter 20.

actual outcomes produced (just x), but not, if we think of the outcomes that are epistemically relevant to the player's deliberations, being both x and y. ■

Thus, before inquiring into this extended logic, we may have to worry about its intended interpretation. We will not propose an epistemic modal forcing logic here, but merely point to a connection with the logics of Chapter 3 for action and knowledge. If the truth of the assertion $\{G, i\}\varphi$ at game state s is to imply knowledge for player i at s that the forced proposition φ will hold, we need a broader notion of relevant outcomes. This also ties up with the earlier issue of knowing one's strategy and the interplay of know-how and know-that, as discussed in Chapter 4.

11.12 Conclusion

The power perspective on players of games has surfaced at several places in this book, and it will continue to do so. We have made this level of analysis precise, and we have shown how it can be studied by generalized modal logics over neighborhood models, lifting earlier notions such as bisimulation to power equivalence for games, and adding concrete spatial visualizations via topological games. Of course, there is an issue of adding further crucial game structure from Part I, such as knowledge and imperfect information. This can be done, and we found a strong connection between power invariance at this level and the classical Thompson transformations of game theory, although some non-trivial issues of interpretation arose with designing an epistemized logic. Another task that remains is adding preference logic as in Chapter 2, and studying a forcing language for what might be called players' optimal powers. Finally, moving beyond Part I, the logical dynamics of Part II applies just as well to the power view of games, but for evidence of this, we refer the reader to some recent literature on dynamic logics of acts that change neighborhood models.[148]

Even so, the style of analysis and the results in this chapter give ample reasons to conclude that our logical study in Parts I and II of this book generalizes smoothly to a natural global perspective on games that focuses on players' strategic powers.

148 In particular, we noted in Chapter 10 that neighborhood models can generalize dynamic-epistemic logic to deal with more fine-grained notions of evidence (van Benthem & Pacuit 2011). Another study of dynamics in neighborhood models is found in Zvesper (2010) for the purpose of analyzing epistemic game theory.

11.13 Literature

This chapter is based on van Benthem (2001b) and van Benthem (2002a). Use was also made of Aiello & van Benthem (2002).

We have pointed at the earlier work of Bonanno (1992b) on set-theoretic forms for extensive games. Also, we listed several good sources on neighborhood semantics. Finally, the results in Pauly (2001) on dynamic game logic and coalition logic are highly relevant, while Goranko et al. (2013) presents amendments and extensions.

12
Matrix Games and Their Logics

The best known format in game theory is different from both the extensive game trees that we have studied so far in this book and their associated power structures. The so-called *strategic form* gives players' complete strategies as single actions, where it is assumed that players choose these strategies simultaneously, resulting in a profile that can be evaluated at once. This is pictured in matrices for standard games such as Hawk versus Dove, or Prisoner's Dilemma, which were discussed briefly in our Introduction. While strategic forms are not a major focus in the book, this chapter and the next will show how they can be studied by means of the logical techniques at our disposal. Our main focus will be the additional perspectives on social action emerging in this way.

12.1 From trees and powers to strategic matrix forms

In Chapter 1, we defined extensive games as an object of study, and in Chapter 6, there was an extensive discussion of models for such games, ranging from game trees themselves to forest models and beyond to more abstract modal models for games. The following standard notion suffices for our purposes in what follows.

Definition 12.1 Strategic games
A *strategic game* G for a set of players N is a tuple $(N, \{A_i\}, \{\leq_i\})$ with a non-empty set A_i of actions for each $i \in N$, and a preference order $\leq_i$ for each player on the set of *strategy profiles*, tuples of actions for each player. As usual, given a strategy profile $\sigma = (a_1, \ldots, a_n)$, σ_i denotes the ith projection and σ_{-i} denotes the remaining profile after deleting player i's action σ_i. ■

Strategic games are given by the familiar matrix pictures of game theory books. Extensive game trees resemble existing models in computational or philosophical logic, but these matrix structures, too, provide a direct invitation for logical analysis of action, preference, knowledge, and freedom of choice. We will show how this can be done, using notions and techniques from Parts I and II, and then develop portions of this logic having to do with independent and correlated action, best response, equilibrium, and various rationality assertions that occur naturally. We will mainly discuss finite two-player games, but most results extend to more players.

Strategic games come with an extensive theory of their own (Osborne & Rubinstein 1994, Gintis 2000, Perea 2011, Brandenburger et al. 2014), and toward the end of the chapter, we briefly discuss some possible further interfaces, such as evolutionary games and the role of probability. We will also show how the logics of this chapter fit with existing topics in philosophical logic such as the STIT analysis of deliberate action, where STIT stands for "seeing to it that".

12.2 Models for strategic games

To a logician, a game matrix is a semantic model that invites the introduction of modal languages, with the strategy profiles themselves as possible worlds. As in Chapter 6, we could also use more abstract worlds, carrying strategy profiles without being identical to them. This is common in epistemic game theory, but it is not needed in what follows. Over strategy profiles, players have three natural kinds of structure, through preference, epistemic view, and action freedom.

Preference The definition of strategic games contained preference relations:

preference $\quad \sigma \leq_i \sigma'$ iff player i prefers profile σ' at least as much as σ

This is the standard notion of preference that we began studying using the logics of Chapter 2. It is usually taken to run between outcomes of a game, and this is what we will do. However, our strategy models would also support another, more ambitious generic use of preference, namely, that players prefer using one strategy over using another.

Epistemic view Matrices also support other natural relations. Here is another basic one from a logical perspective:

epistemic view $\quad \sigma \sim_i \sigma'$ iff $\sigma_i = \sigma'_i$

This epistemic relation represents player i's view of the game at the interim stage where i's choice is fixed, but the choices of the other players are unknown.[149]

Action freedom There is yet a third natural relation for players i in game models:

action freedom $\quad \sigma \approx_i \sigma'$ iff $\sigma_{-i} = \sigma'_{-i}$

This relation of freedom gives the alternative choices for player i when the other players' choices are fixed.[150] The three relations are entangled in various ways, a topic that we will return to.

This structure can all be packaged in the following notion.

Definition 12.2 Game models

Game models $\boldsymbol{M} = (S, N, \{\leq_i\}_{i \in N}, \{\sim i\}_{i \in N}, \{\approx_i\}_{i \in N}, V)$ are relational structures as described above, where a valuation V evaluates atomic propositions p on strategy profiles, viewed as special properties V(p) of these. The *full model* over a strategic game G is the game model $\boldsymbol{M}(G)$ whose worlds are all strategy profiles, with the three relations as defined above. Finally, a *general game model* is a submodel of a full game model. ■

We will find a logical and a game-theoretic use for general game models below.

What the models mean In this minimalist view, matrix games model a moment where players know their own action, but not that of the others, which happens at the time of decision when all the relevant available evidence is in. Richer models of games in the style of Chapter 6 would allow for knowledge or ignorance concerning other features, and also for more fine-grained beliefs that players can have about what will happen. Despite their intuitive importance, we will ignore these extra features at this point, as well as any issues having to do with probabilistic combinations of strategies.

Standard game matrices correspond directly to models in our style.

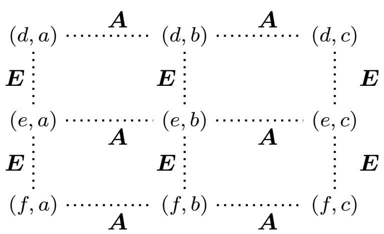
149 This is just one take on matrix games, albeit a common one. Bonanno (2012) provides an illuminating discussion of different construals of games in terms of deliberation versus actual choice. Likewise, Hu & Kaneko (2012) contains a careful disentanglement of various views of matrix games, in terms of predictions versus decisions.

150 The term freedom is taken from Seligman (2010), who points out that these three-relation structures make sense far beyond games. See Guo & Seligman (2012) for some further development.

EXAMPLE 12.1 Models from game matrices
Consider the following matrix, with this order of writing utility values for outcomes: (***A***-value, ***E***-value). It will be discussed in more detail in a later section.

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

The associated model is as follows with regard to epistemic view and freedom:



Here the uncertainty relation $\sim_{\boldsymbol{E}}$ runs inside columns, because ***E*** knows the action that ***E*** will take, but not that of ***A***. The uncertainty relation of ***A*** is symmetric, running inside the rows. ■

These concrete pictures will help understand later arguments. In a sense, the gestalt switch from matrix games to logical models is the main point of this chapter.

Alternative perspectives There are other ways of interpreting our pictures. We have already mentioned the diversity of interpretations in the game-theoretic literature. But also, one person's freedom can be another person's ignorance. Freedom is read in van Benthem (2007d) as "distributed knowledge" for player i in the sense of Fagin et al. (1995), stating what all other players, that is, the group $N - \{i\}$, know implicitly about the game.

12.3 Matching modal languages

Syntax The above three kinds of structure invite matching modalities. Game models obviously support the preference language of Chapter 2 with modalities $\langle pref_i \rangle$ or $\langle \leq_i \rangle$, since the latter can be interpreted on any model consisting of worlds with a betterness order. For matrix games, it also turns out to be useful to add modalities $\langle <_i \rangle$ for strict preference order as in van Benthem et al. (2009c). Likewise, game models interpret the epistemic language of Chapter 3 with knowledge operators K_i (or their existential variants $\langle know_i \rangle$) where a few special-purpose proposition letters may carry information about which player plays what action in a given world, or other relevant properties of strategy profiles. Finally, freedom can be described in the same style by modal operators $\langle free_i \rangle$.

Definition 12.3 Logical language for matrix games
The syntax of a combined *logic of matrix games* can be given as follows:

$$p \mid \neg\varphi \mid (\varphi \vee \psi) \mid \langle pref_i \rangle\varphi \mid \langle <_i \rangle\varphi \mid \langle know_i \rangle\varphi \mid \langle free_i \rangle\varphi$$

We will often write $\langle \leq_i \rangle\varphi$ for $\langle pref_i \rangle\varphi$, $\langle \sim_i \rangle\varphi$ for $\langle know_i \rangle\varphi$, and $\langle \approx_i \rangle\varphi$ for $\langle free_i \rangle\varphi$ to emphasize the underlying accessibility relations for these modalities. ■

Expressive power With a little expressive boost, languages such as these can define basic features of social scenarios and of games in particular. A benchmark is Nash equilibrium of profiles that has been much studied by logicians (cf. van der Hoek & Pauly 2006). As it happens, no formula of our basic modal language can define this notion. The reason is that it is not invariant under bisimulation (cf. Chapter 1, and van Benthem 2010b). But we can help ourselves to one more technical device, namely, modalities for intersections of relations (see Dégremont 2010, Zvesper 2010 for its uses in analyzing extensive games). Recall from our Introduction that action σ_i is a *best response* for player i against given behavior σ_{-i} of the others if no alternative action for i will lead to an outcome more preferred by i. Then, enough expressive power is available.

Fact 12.1 (a) The modal definition BR_i for best response is $\neg\langle <_i \cap \approx_i \rangle\top$. (b) The strategy profiles in Nash equilibrium are those satisfying the formula $\bigwedge_{i \in N} BR_i$.

Thus we see how game models suggest the use of mild extensions of modal logic, often from the area of hybrid logics (Areces & ten Cate 2006).[151]

151 Van Benthem (2004b) has a hybrid logic with preference modalities, nominals for specific worlds, a universal modality, and distributed group knowledge, that can formalize the rationality principles found in Section 12.7. Hybrid preference logics are explored in van Benthem et al. (2006b), Guo & Seligman (2012) have further game examples.

12.4 Modal logics for strategic games

Basic modal logics Next, we explore a calculus of reasoning for strategic games. For convenience, we restrict our attention to two-player games. First, given the nature of our three relations, the separate logics are standard systems: modal $S4$ for preference, and modal $S5$ for both epistemic outlook and action freedom. What is of greater interest is the interaction of the three modalities. In general, logics of knowledge and action need not support strong bridge principles, as we have seen in Chapter 3; or, at the very least, such bridge principles express non-trivial assumptions about agents. But our models for strategic games are full of mutual connections, so we expect much more to be valid.

One strong connection involves the following natural combination of modalities.

FACT 12.2 $[\sim_i][\approx_i]\varphi$ makes a proposition φ true in each world of a game model.

Proof On matrix game models, the sequential composition of the earlier relations for epistemic view and freedom is precisely the universal relation. ■

This gives the base language a universal modality $U\varphi$ (with an existential dual $E\varphi$), a device used widely in Part I. This modality can even be defined in two ways.

FACT 12.3 The equivalence $[\sim_i][\approx_i]\varphi \leftrightarrow [\approx_i][\sim_i]\varphi$ is valid in matrix game models.

Proof This validity depends on the geometrical confluence property of matrices that, if one can go $x \sim_i y \approx_i z$, then there exists a point u with $x \approx_i u \sim_i z$. To visualize this, compare the confluence property for perfect recall in Chapter 3. ■

This system can prove interesting principles, including one that will serve later as a law of independence for simultaneous actions in STIT logics.

FACT 12.4 The formula $(\langle\approx_i\rangle[\sim_i]\varphi \wedge \langle\approx_j\rangle[\sim_j]\psi) \to E(\varphi \wedge \psi)$ is derivable.

Proof First we prove an auxiliary formula in our logic:

$$\langle\approx_i\rangle[\sim_i]\varphi \to [\sim_i]\langle\approx_i\rangle\varphi$$

The steps are as follows:

(a)	$\varphi \to [\approx_i]\langle\approx_i\rangle\varphi$	(in modal $S5$)
(b)	$[\sim_i]\varphi \to [\sim_i][\approx_i]\langle\approx_i\rangle\varphi$	(from (a) in basic modal logic)
(c)	$[\sim_i]\varphi \to [\approx_i][\sim_i]\langle\approx_i\rangle\varphi$	(from (c) by the commutation axiom)
(d)	$\langle\approx_i\rangle[\sim_i]\varphi \to \langle\approx_i\rangle[\approx_i][\sim_i]\langle\approx_i\rangle\varphi$	(from (c) in basic modal logic)
(e)	$\langle\approx_i\rangle[\sim_i]\varphi \to [\sim_i]\langle\approx_i\rangle\varphi$	(from (d) in modal S5)

Now consider the formula $\langle\approx_i\rangle[\sim_i]\varphi \wedge \langle\approx_j\rangle[\sim_j]\psi$. The left-hand side implies $[\sim_i]\langle\approx_i\rangle\varphi$. Also, given that $\approx_j$ equals $\sim_i$ for opposite players, the right-hand side $\langle\approx_j\rangle[\sim_j]\psi$ is equivalent to $\langle\sim_i\rangle[\approx_i]\psi$. But then a standard formal proof in the basic modal logic derives $\langle\sim_i\rangle(\langle\approx_i\rangle\varphi \wedge [\approx_i]\psi)$, and this again implies $\langle\sim_i\rangle\langle\approx_i\rangle(\varphi \wedge \psi)$, that is, $E(\varphi \wedge \psi)$. ■

Complexity of grid structure Confluence may look like a pleasant feature of matrices, but its effects on logical systems can be dangerous. As we have seen in Chapters 2 and 3, bimodal logics on grid models are often not decidable, and not even axiomatizable, once they have a universal modality available. The reason was that such logics can encode tiling problems of high complexity, and hence our simple-looking strategic matrix games can have a Π^1_1-complete logical theory encoding a large amount of geometry.

Finite games These high-complexity results might be circumvented for finite games. The latter realm validates special laws that fail on general game models. For instance, finiteness implies upward well-foundedness of the two-step relation $\sim_A ; \sim_E$, i.e., it has only ascending sequences of finite length. Well-foundedness validates a Grzegorczyk axiom for the epistemic modalities (cf. Blackburn et al. 2001). The logic of finite matrix games may be less complex than for all games, as infinite tiling problems cannot be encoded.

12.5 General game models, dependence, and correlation

There is another, more rewarding avenue toward defusing the complexity danger. It is well known that the complexity of bimodal logics may go down drastically when we allow more models. In particular, the earlier general game models left out some strategy profiles to create less grid-like patterns. Then the logic changes drastically to something much simpler.

THEOREM 12.1 The complete epistemic logic of general game models is multi-S5.

Proof This result is from van Benthem (1996), which proves the key fact that every multi-$S5$ model has a bisimulation with a general game model.[152] ■

While this may look ad hoc, general game models have their uses. For instance, in Chapter 13, we will see how they arise in solving games when external information may come in ruling out strategy profiles in the manner of Chapter 7.

But there is a deeper interpretation of the change to general game models, having to do with *dependence* and independence of actions. Leaving out profiles from a full matrix model creates intuitive dependencies between actions. If we change the choice for one player, we may have to change that for the other to stay inside the given universe of available profiles.

Such models with gaps have been studied by logicians (van Benthem 1997, and in a different, much more systematic vein, Väänänen 2007). The motivation is that dependence is a key notion in many areas of reasoning (see the independence-friendly logics of Chapter 21), as well as in studies of information as correlation among situations (van Benthem & Martínez 2008).

Game theorists, too, have studied general game models through an interest in "correlated behavior" (Aumann 1987, Brandenburger & Friedenberg 2008). Once we omit strategy profiles, what one player does may have repercussions for another, and this leads to new views of games.[153] Thus, complexity of logics for games matches interesting decisions on how we view players: as acting independently, or as being correlated.

12.6 Special logics of best response

Let us now look more concretely at the logic of matrix game models, in both full and general versions. In what follows, to simplify matters, we will not use the full generality of our earlier three relations, choosing a different approach instead. We treat some basic game-theoretic properties involving preference and action as atomic propositions, to keep the logic simple. While somewhat nonstandard, this is an acceptable modus operandi. In our views on logical zoom, there is no moral duty to formalize everything beyond what is perspicuous and convenient.

152 It is an open problem to extend this to game models with all three base relations.

153 Sadzik (2009) analyzes games with correlated equilibria in terms of distributed knowledge and a notion of bisimulation. Isaac & Hoshi (2011) discuss game equivalence with correlations in play.

“lig-09-25” — 2013/10/29 — 9:44 — page 283 — #301

Absolute best For a start, observe that best response as defined earlier is an absolute property: the conjunction over alternatives runs over all actions available in the original strategic game G, whether these occur in the general game model $\boldsymbol{M}(G)$ or not. Because of absoluteness, B_j may be viewed as an atomic proposition that keeps its value when models change from larger to smaller domains of profiles, or vice versa. This context independence simplifies notation considerably.

Example 12.2 Expanded game models
Well-known games generate models of interest here. Consider the well known game of Battle of the Sexes with its two Nash equilibria (Osborne & Rubinstein 1994). The abbreviated diagram to the right has best-response atomic propositions at worlds where they are true:

		$\boldsymbol{E}$	
		a	b
$\boldsymbol{A}$	c	$2,1$	$0,0$
	d	$0,0$	$1,2$

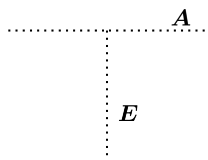
$$\begin{tikzcd}
(c,a) \arrow[r, dotted, no head, "\boldsymbol{A}"] \arrow[d, dotted, no head, "\boldsymbol{E}"'] & (c,b) \arrow[d, dotted, no head, "\boldsymbol{E}"] \\
(d,a) \arrow[r, dotted, no head, "\boldsymbol{A}"'] & (d,b)
\end{tikzcd}$$

$B_{\boldsymbol{A}}, B_{\boldsymbol{E}}$	–
–	$B_{\boldsymbol{A}}, B_{\boldsymbol{E}}$

Remark Mixed equilibria
This game also has an equilibrium in mixed strategies that we will not discuss here. This would require a probabilistic extension of the logics in this chapter that we do not provide, although we think it should be a straightforward generalization.

Likewise, Example 12.1 yields a full epistemic game model with nine worlds:

	a	b	c
d	$B_{\boldsymbol{A}}, B_{\boldsymbol{E}}$	–	–
e	$B_{\boldsymbol{E}}$	$B_{\boldsymbol{A}}$	–
f	–	$B_{\boldsymbol{E}}$	$B_{\boldsymbol{A}}$



As for the distribution of B_j atoms, by the above definition, every column in a full game model must have at least one occurrence of $B_{\boldsymbol{A}}$, and every row one of $B_{\boldsymbol{E}}$. ■

Example 12.3 Evaluating epistemic statements
(a) The following formula of our modal language expresses that all players think their current actions might be best for them:

$$\langle \boldsymbol{E} \rangle B_{\boldsymbol{E}} \wedge \langle \boldsymbol{A} \rangle B_{\boldsymbol{A}}$$

As we shall see in Section 12.7, this expresses a form of rationality related to those that were discussed in Chapters 3, 7, and 8. In the preceding model, this proposition is true in exactly the six worlds occurring in the a and b columns.

(b) The same model also highlights an important epistemic distinction. B_j states that j's current action is in fact a best response at ω. But j need not know that, as j need not know what the other player is doing. Indeed, the formula $K_{\boldsymbol{E}}B_{\boldsymbol{E}}$ is false throughout the above model, even though $B_{\boldsymbol{E}}$ is true at three worlds. A fortiori, then, common knowledge of rationality in its most obvious sense is often false in the full model of a game, even one with a unique Nash equilibrium. ■

With this enriched language, the logic of game models gains interest. An example are valid laws for best response such as the following principle, subtly different in syntax from the one discussed above.

FACT 12.5 The formula $\langle \boldsymbol{E} \rangle B_{\boldsymbol{A}} \wedge \langle \boldsymbol{A} \rangle B_{\boldsymbol{E}}$ is valid in all full game models.

Proof Use the fact that all rows and columns have entries with maximal values. ■

Relative best There are also alternative views on game models. In particular, the word "best" in the phrase best response is context-dependent. A natural relative version of best response in a general game model $\boldsymbol{M}$ looks only at the strategy profiles available inside $\boldsymbol{M}$. Inside a model, players know these are the only action patterns that will occur.

DEFINITION 12.4 Relative best response
The *relative best response* proposition B^*_j in a general game model $\boldsymbol{M}$ is true at only those strategy profiles where j's action is a best response to that of the opponent when the comparison set is all alternative strategy profiles inside $\boldsymbol{M}$. ■

With relative B^*_j, best profiles for j may change as a model changes. For instance, in a one-world model for a game, the single profile is relatively best for all players, although it may be absolutely best for none.[154]

REMARK What others know
Relative best response has independent interest. With two players, it says that the other player knows that j's current action is at most as good for j as j's action at ω. More generally, continuing with an earlier observation, B^*_j says that

154 For a similar idea in an abstract computational view on game solution, see Apt (2007).

"lig-09-25" — 2013/10/29 — 9:44 — page 285 — #303

the proposition "j's current action is at most as good for j as j's action at ω" is distributed knowledge at ω for the rest of the players $G - \{j\}$.[155]

Absolute best obviously implies relative best, but the converse is not true.

EXAMPLE 12.4 All models have relative best positions
To see the difference between the two notions, compare the two models

$1, 1(B_{\boldsymbol{A}})$	$0, 2(B_{\boldsymbol{E}})$
$0, 2(B_{\boldsymbol{E}})$	$1, 1(B_{\boldsymbol{A}})$

$1, 1(B_{\boldsymbol{A}}, B^*_{\boldsymbol{E}})$	
$0, 2(B_{\boldsymbol{E}})$	$1, 1(B_{\boldsymbol{A}})$

Removing one entry left no absolute $B_{\boldsymbol{E}}$ in the first row, but $B^*_{\boldsymbol{E}}$ adjusted. ■

It would be of interest to axiomatize the modal logic of best response completely. In this chapter, we will only pursue one particular form of reasoning with the notions introduced here, involving two basic notions of rationality.

12.7 A case study: Rationality assertions, weak and strong

The notion of rationality was studied extensively in Parts I and II of this book. In strategic games, rationality means playing a best response given what one knows or believes. But this is not the whole story. Our game models support further distinctions, such as absolute versus relative best response. Also, even when players in fact play their best action, they need not know that they are doing so. Thus, if rationality is a self-reflexive property, as commonly assumed, what can players know? This issue will return in our conversation scenarios for solving games in Chapter 13, where players can only communicate things they know to be true.

Weak rationality Players may not know that their action is best, but they can know *there is no alternative action that they know to be better*. They are no fools.

155 In the dynamic terms of Part II of this book, the other players might learn this fact about j by "pooling" their information. This observation is used in van Benthem et al. (2006b) for defining Nash equilibrium in an extended epistemic preference logic.

DEFINITION 12.5 Weak rationality
Weak rationality of player j at world ω in a model $\boldsymbol{M}$ is the assertion that, for each available alternative action, j thinks the current one may be at least as good:

$$WR_j \qquad \bigwedge_{a \neq \omega(j)} \langle j \rangle \text{ “}j\text{'s current action is at least as good for } j \text{ as } a\text{”}$$

The index set runs over worlds in the current model, as for relative best B_j^*.[156] ■

Weak rationality WR_j for j fails at those rows or columns in a two-player general game model that are strictly dominated for j. Unpacking quantifiers, $WR_{\boldsymbol{E}}$ says for a matrix column x that for each other column y, there is at least one row where $\boldsymbol{E}$'s value in x is at least as good as that in y. Such columns always exist, by a simple combinatorial argument.

FACT 12.6 Each finite general game model contains worlds where $\bigwedge_j WR_j$ is true.

Proof For convenience, look at games with just two players. We show something stronger, namely, that the model has "*WR* loops" of the form

$$s_1 \sim_{\boldsymbol{A}} s_2 \sim_{\boldsymbol{E}} \ldots \sim_{\boldsymbol{A}} s_n \sim_{\boldsymbol{E}} s_1 \qquad \text{with } s_1 \models B_{\boldsymbol{E}}^*,\ s_2 \models B_{\boldsymbol{A}}^*,\ s_3 \models B_{\boldsymbol{E}}^*,\ \ldots$$

By way of illustration, a Nash equilibrium by itself is a one-world *WR* loop.

First, taking maxima on the available positions (column, row) in the full game matrix, we see that the following two statements must hold everywhere:

$$\langle \boldsymbol{E} \rangle B_{\boldsymbol{A}}^* \qquad \langle \boldsymbol{A} \rangle B_{\boldsymbol{E}}^*$$

For example, the first says that, given a world with some action for $\boldsymbol{E}$, there must be some world in the model with that same action for $\boldsymbol{E}$ where $\boldsymbol{A}$'s utility is highest. (This need not hold with the above absolute $B_{\boldsymbol{A}}$, as its witness world may have been left out.) Repeating this, there is a never-ending sequence of worlds

$$B_{\boldsymbol{E}}^* \sim_{\boldsymbol{E}} B_{\boldsymbol{A}}^* \sim_{\boldsymbol{A}} B_{\boldsymbol{E}}^* \sim_{\boldsymbol{E}} B_{\boldsymbol{A}}^* \cdots$$

that must loop since the model is finite. Thus, some world in the sequence with, say, $B_{\boldsymbol{A}}^*$ is $\sim_{\boldsymbol{A}}$-connected to some earlier world w. Now, either w has $B_{\boldsymbol{E}}^*$, or w has a successor with $B_{\boldsymbol{E}}^*$ via $\sim_{\boldsymbol{A}}$ in the sequence. The former case reduces to the latter by the transitivity of $\sim_{\boldsymbol{A}}$. But then, looking backward along such a loop, using

156 An alternative version lets the index set run over all strategy profiles in the whole initial game, as with absolute best assertions B_j. It can be dealt with similarly.

the symmetry of the relations, we have a WR loop as defined above. Its worlds evidently validate weak rationality for both players: $\langle \boldsymbol{E} \rangle B^*_{\boldsymbol{E}} \wedge \langle \boldsymbol{A} \rangle B^*_{\boldsymbol{A}}$. ■

FACT 12.7 Weak rationality is epistemically introspective.

Proof By the accessibility in epistemic game models, if WR_j holds at some world ω in a model, it also holds at all worlds that j cannot distinguish from ω. Hence, the epistemic principle $WR_j \to K_j\, WR_j$ is valid on general game models. ■

Thus, $WR_j \to K_j\, WR_j$ is a logical law of game models with best response and rationality. This makes weak rationality suitable for public announcement, ruling out worlds on strictly dominated rows or columns every time when uttered.

Strong rationality Weak rationality is a logical conjunction of epistemic possibility operators: $\bigwedge \langle j \rangle$. A stronger form of rationality assertion inverts this order, expressing that players think that *their actual action may be best*. Instead of merely being no fools, they now reasonably hope they are being clever.

DEFINITION 12.6 Strong rationality
Strong rationality for player j at a world ω in a model $\boldsymbol{M}$ is the assertion that j thinks that the current action may be at least as good as all others:

$$SR_j \qquad \langle j \rangle \bigwedge_{a \neq \omega(j)} \text{“}j\text{'s current action is at least as good for } j \text{ as } a\text{”}$$

This time we use the absolute index set, running over all action profiles in the game. Hence, the assertion can be written equivalently as the modal formula $\langle j \rangle B_j$. Strong rationality for the whole group of players is the conjunction $\wedge_j SR_j$. ■

By the $S5$-law $\langle j \rangle \varphi \to K_j \langle j \rangle \varphi$, SR_j is something that players j will know if it is true. Thus, it behaves like WR_j. Strong and weak rationality are related as follows.

FACT 12.8 SR_j implies WR_j, but not vice versa.

Proof Consider the following game model, with B atoms indicated:

		$\boldsymbol{E}$: a	b	c
$\boldsymbol{A}$	d	1, 2	1, 0	1, 1
	e	0, 0	0, 2	2, 1

$B_{\boldsymbol{A}}, B_{\boldsymbol{E}}$	$B_{\boldsymbol{A}}$	–
–	$B_{\boldsymbol{E}}$	$B_{\boldsymbol{A}}$

No column or row dominates any other, and WR holds throughout for both players. But $SR_{\boldsymbol{E}}$ holds only in the two left-most columns. This is because it rejects actions that are never best, even though there need not be one alternative in the model that is better overall. ■

One advantage of SR_j over WR_j is the absoluteness of its proposition letters B_j. In the dynamics of Chapter 13, this feature underlies the monotonicity of the set transformation defined by announcing *SR*. Also, strong rationality has a direct game-theoretic meaning. It says that the current action of the player is a best response against at least one possible action of the opponent. This is the key assertion in rationalizability views of game solution, due to Bernheim and Pearce (cf. de Bruin 2004, Apt 2007), where one discards actions for which a better response exists under all circumstances.

Strong rationality need not be satisfiable in all general game models. But it is satisfiable in all full game models, thanks to the maximal utility values in rows and columns. The following result explains the logical import.

THEOREM 12.2 Each finite full game model has worlds verifying strong rationality.

Proof Much as with Theorem 12.2, there are "*SR* loops" of the form

$$s_1 \sim_{\boldsymbol{A}} s_2 \sim_{\boldsymbol{E}} \ldots \sim_{\boldsymbol{A}} s_n \sim_{\boldsymbol{E}} s_1 \qquad \text{with } s_1 \models B_{\boldsymbol{E}},\ s_2 \models B_{\boldsymbol{A}},\ s_3 \models B_{\boldsymbol{E}},\ \ldots$$

Instead of a proof, we give an illustration showing that each finite full game model has three-player *SR* loops. In such models, by the earlier observations about maxima on rows and columns, the following is true everywhere:

$$\langle \boldsymbol{B}, \boldsymbol{C} \rangle B_{\boldsymbol{A}}, \quad \langle \boldsymbol{A}, \boldsymbol{C} \rangle B_{\boldsymbol{B}}, \quad \langle \boldsymbol{A}, \boldsymbol{B} \rangle B_{\boldsymbol{C}}$$

Here the special modalities $\langle i, j \rangle$ have an accessibility relation $\sim_{\{i,j\}}$ keeping the coordinates for both i and j the same, that is, the intersection of $\sim_i$ and $\sim_j$.

But then, repeating this, by finiteness, we must have loops of the form $B_{\boldsymbol{A}} \sim_{\{\boldsymbol{A},\boldsymbol{C}\}} B_{\boldsymbol{B}} \sim_{\{\boldsymbol{A},\boldsymbol{B}\}} B_{\boldsymbol{C}} \sim_{\{\boldsymbol{B},\boldsymbol{C}\}} B_{\boldsymbol{A}} \cdots$ returning to the initial world with $B_{\boldsymbol{A}}$. Any world in such a loop satisfies $\langle \boldsymbol{A} \rangle B_{\boldsymbol{A}} \wedge \langle \boldsymbol{B} \rangle B_{\boldsymbol{B}} \wedge \langle \boldsymbol{C} \rangle B_{\boldsymbol{C}}$. For example, if the world itself has $B_{\boldsymbol{A}}$, by reflexivity, it satisfies $\langle \boldsymbol{A} \rangle B_{\boldsymbol{A}}$. Looking back at its parent $B_{\boldsymbol{C}}$ via $\sim_{\{\boldsymbol{B},\boldsymbol{C}\}}$, by symmetry, it has $\langle \boldsymbol{C} \rangle B_{\boldsymbol{C}}$. And looking at its grandparent $B_{\boldsymbol{B}}$ via $\sim_{\{\boldsymbol{B},\boldsymbol{C}\}}$ and $\sim_{\{\boldsymbol{A},\boldsymbol{B}\}}$, by transitivity, it also satisfies $\langle \boldsymbol{B} \rangle B_{\boldsymbol{B}}$. ■

In infinite game models, *SR* loops need not occur, and irrationality may prevail.

EXAMPLE 12.5 Irrationality in regress
Consider a grid of the form $N \times N$. Suppose that the best-response pattern has the predicates $B_{\boldsymbol{A}}$ and $B_{\boldsymbol{E}}$ occurring only on the two marked diagonals:

$$\begin{tikzcd}
B_{\boldsymbol{E}} \arrow[dr] & - & - & - \\
B_{\boldsymbol{A}} \arrow[dr] & B_{\boldsymbol{E}} \arrow[dr] & - & - \\
- & B_{\boldsymbol{A}} \arrow[dr] & B_{\boldsymbol{E}} \arrow[dr] & - \\
- & - & - & -
\end{tikzcd}$$

Every putative sequence $B_{\boldsymbol{E}} \sim_{\boldsymbol{A}} B_{\boldsymbol{A}} \sim_{\boldsymbol{E}} B_{\boldsymbol{E}} \sim_{\boldsymbol{A}} B_{\boldsymbol{A}} \cdots$ must break off. ■

Our case study will have shown how much surprising logical structure there is to reasoning with rationality once we look at models and syntax. It would be of interest to axiomatize the complete logic of general and full game models expanded with B_j, B_j^*, WR_j, and SR_j, either in our epistemic base language, or with the expressive additions from hybrid logic that we have mentioned occasionally. But we leave matters at this stage here, returning to rationality assertions only in the dynamic-epistemic setting of Chapter 13.

12.8 STIT logic and simultaneous action

The modal systems in this chapter may also be viewed as basic logics of simultaneous action, outside of the setting of games. In this coda to our main theme, we present an illustration, in the form of a brief excursion toward a well-known philosophical paradigm for agency called STIT (seeing to it that). For details of what follows here, the reader may consult van Benthem & Pacuit (2012). We also refer to classical sources for more extensive development of STIT, such as Horty & Belnap (1995) and Belnap et al. (2001), while Horty (2001) connects the framework with deontic logics of agency.

Models, language, and logic Models for STIT are branching time structures as defined in Chapter 5, with the following new feature. At each stage (h, t) of a point t on history h, each agent i is assigned a partition $C_i(h, t)$ of all future histories passing through t. Partition cells are choices that the agent has at time t. These choices satisfy one important constraint, expressing the independence of actions chosen by agents:

Independent action   Any two partition cells for different agents overlap.

These models support a temporal language as in Chapter 5, but we now add modalities for agents who see to it that some stated proposition is achieved:

STIT modality $\boldsymbol{M}, h, t \models [stit, i]\varphi$ iff there exists a partition cell X of $C_i(h, t)$ such that, for all histories h' in X, $\boldsymbol{M}, h', t \models \varphi$

The logic of this system is upward monotonic in these modalities, but not distributive over conjunction or disjunction: this can be seen by analogy to the forcing modalities of Chapter 11. The independence constraint validates the following:

Product Axiom $([stit, i]\varphi \wedge [stit, i]\psi) \to \Diamond(\varphi \wedge \psi)$

where $\Diamond$ is the existential modality over current branches introduced in Chapter 5.

Intuitively, one might think that actions achieve effects, not right now, but later in time. This can be incorporated using modal combinations $[stit, i]O\varphi$, where O stands for at the next moment (Broersen 2011) or $[stit, i]G\varphi$.

Two views of action STIT embodies a view of action as *choice* plus *control* of agents over outcomes, where it is crucial that actions may not give unique control over outcomes. This slack may be because other agents have control as well over what happens, or even with a single agent, multiple histories in a partition cell may arise from actions of the environment. Prima facie, this view of choosing an action looks different from the one in this book, which comes from the dynamic and temporal logics in Part I. There, as emphasized in Chapter 6, branchings between events or moves model possible choices for the agent, and labels provide explicit names for transitions.

The two styles of thinking are not mutually exclusive, however. Here is a way of combining insights from both. Despite the loose talk interchanging events, moves, and actions in this book (and most of the literature), modal logics of action, computation, and games really describe a structure of possible *events*. What STIT then adds is the idea of control that agents can have over events. Thus, a richer view of action doing justice to both would be this:

$$Action = Events + Control$$

Embedding STIT into matrix game logic One way of benefiting from this junction is by embedding the basic STIT logic in the logics of this chapter. The following result comes from van Benthem & Pacuit (2012), but Herzig & Lorini (2010) contains essentially the same insight.

Recall the matrix game models of Section 12.3, where we think of one player i choosing an action a while others choose theirs independently to create a complete strategy profile, that is, a tuple of individual actions. We can think of these composite simultaneous actions as labeling transitions in a temporal STIT model. The following connection then works at the level of languages.[157]

DEFINITION 12.7 Modal game translation
In terms of the modalities of Section 12.4, one can translate as follows

$$[stit, i]\varphi = [\sim_i]\varphi \qquad \Diamond\varphi = \langle\sim_i\rangle\langle\approx_i\rangle\varphi$$

Arbitrary complex formulas are then translated compositionally. ■

THEOREM 12.3 The preceding translation embeds the basic STIT logic faithfully into the modal logic of full matrix games.

Proof We look at only two agents. Our translation validates all STIT axioms, where the freedom modality refers to choices, and the action modality to simultaneous actions compatible with them. The Product Axiom holds on all game models, and indeed we derived it in modal game logic in Section 12.4.

To prove that the embedding is faithful, we must refute each non-valid STIT formula φ in our matrix models. To do so, consider any STIT temporal countermodel, and note that it suffices to look at the current moment where $\neg\varphi$ holds and its successor moments, since our basic STIT language does not contain temporal modalities. Thus, one can simply derive a two-agent basic STIT $S5$ model out of the temporal structure by letting histories be worlds, and defining agent's equivalence relations from their choice partitions. Now, the temporal box $\Box$ is the universal modality, while the two STIT modalities match the two equivalence relations. In principle, this might just lead to a general game model as defined earlier. However, the structure of the STIT model ensures that any two equivalence classes intersect.

Crucially, we can "unravel" this $S5$-model into a bisimilar model of a special form where the intersection of any two equivalence classes is a singleton. Then, a representation in matrix form is obvious: the actions are the equivalence classes for the agents, the unique outcomes are their intersections.[158]

---

157 One might want to prefix an existential modality over choices for i in the translation that follows, but the STIT modality is about the choice made on the actual history.

158 We suppressed important details here. A complete proof, including a bisimulation unraveling by means of a product construction, is in van Benthem & Pacuit (2012).

It follows in a straightforward manner that, under our translation, any STIT-satisfiable formula is also satisfiable in matrix games. Thus, our translation is both correct and faithful.[159] ■

The role of knowledge Our connection between STIT and matrix games introduced a notion of knowledge for agents that have decided, but do not know yet what the others have chosen. This is just one of several notions of knowledge that are important to games, as discussed in Chapters 3 and 6, when looking at information that players have about the past, present, and future of play. Knowledge is not made explicit in the STIT framework, but it is behind the scenes.[160] One can make this feature explicit by moving in the opposite direction of the above, involving the key system used in Part II.

Extending DEL with control Consider a dynamic-epistemic language as in Chapter 7 with knowledge modalities $K_i\varphi$ and dynamic event modalities $[\boldsymbol{\mathcal{E}}, e]\varphi$. Control fits very easily in this setting. First, we endow event models with equivalence relations for control by the relevant agents. Then we can define a dynamic STIT-style "control modality" as follows:

$$\begin{array}{l} \boldsymbol{M}, s \models [\boldsymbol{\mathcal{E}}, e, \mathit{control}_i]\varphi \quad \text{iff} \quad \text{in all product models } (\boldsymbol{M}, s) \times (\boldsymbol{\mathcal{E}}, f) \\ \text{for all events } f \text{ that are control equivalent in } \boldsymbol{\mathcal{E}} \text{ to } e \text{ for agent } i, \\ (\boldsymbol{M}, s) \times (\boldsymbol{\mathcal{E}}, f), (s, f) \models \varphi. \end{array}$$

This is like a DEL operator stating the knowledge that an agent has acquired after product update with the current event model $\boldsymbol{\mathcal{E}}$.[161] The complete dynamic logic of this expanded system lies embedded in DEL in an obvious manner. Its laws for the new control operator are essentially those of STIT. Thus, we get a combined logic of events and agent's choices that might enrich our analysis of games in Part II.[162]

---

159 If we make our modal language stronger, however, the singleton intersection property would give rise to a special axiom beyond basic STIT.

160 A recent proposal for adding knowledge to STIT while making a junction with the modal logic of games in this chapter is in Ciuni & Horty (2013).

161 Control in private-information scenarios is harder to interpret intuitively.

162 However, this richer logic of agent's choices has special features. It lacks the characteristic recursion laws of Chapter 7 directly for the STIT control modality, since this operator does not distribute over conjunction or disjunction. Also, unlike most DEL logics, this new system does not have a modality reflecting its dynamic control relations in the static epistemic base models. The latter feature makes event models really sui generis.

Simultaneous action in extensive games Let us now generalize to interactive behavior over time, as in the extensive games of Parts I and II. Our presentation of STIT emphasized single moments of simultaneous choice. Next, consider players' consecutive choices in longer games, represented by the strategic powers of Chapter 11. We cannot quite adopt the simplest set of constraints here, since determinacy is lost in a setting of simultaneous action. Yet the STIT constraints on choice seem close to those in the representation results for powers in imperfect information games in Section 11.9 that required only monotonicity and consistency.

Instead of pursuing this, we make an observation in the main line of this chapter. What happens to the key constraints on choice when we consider iterated simultaneous action? Most importantly, the crucial "partition property" for histories under single choices disappears. When we make consecutive choices, the space of available strategies grows. In a one-step game, agents can only choose one of their actions ab initio. But now, they can let the next action depend on observed behavior of other agents. A famous case is the Tit for Tat strategy in evolutionary game theory (see also our Introduction): it copies the opponent's preceding move. Hence, strategies available in extensive games need not just choose one action uniformly; they can depend on the behavior of others. It is easy to see that disjointness for sets of outcomes, that is, the powers matching players' strategies, may fail now, and what is left is just monotonicity and consistency.

Remark Public observation
This richer set of strategies does depend crucially on public observation of moves. Without this, players cannot make their choices dependent on what others have done, and we get a DEL product model of two consecutive actions that does satisfy the partition condition. In terms of Section 12.5, one-step simultaneous action does not allow for sequential dependence of actions, but it may allow for correlation.

Dynamifying STIT Finally, the DEL perspective of Part II also suggests a more radical move, reanalyzing the scenarios that motivated STIT in the first place. What are the main events that occur in a choice scenario? The main stages would seem to be: deliberation, decision, action, and observation. In a first stage, we analyze our options, and find optimal choices. Next, at the decision stage, we make up our mind and choose an action. Then everyone acts publicly, and this gets observed, something we can also model as a separate stage, although things happen simultaneously. All of these stages can be analyzed using the DEL-style models of Part II: Chapter 8 discussed deliberation, while Chapter 9 had many systems for mid-game scenarios. Taken together, this might give a much richer view of simultaneous action than those suggested above.

Further directions Much more can be said about merges of control-based action with preference structure and temporal evolution. We refer to the cited literature, as well as Broersen et al. (2006) and Xu (2010) for some recent developments.

12.9 Conclusion

In this chapter, we have connected the logical approach of this book to the standard format of game theory, that is, games in strategic form. The link arose by sensitizing the reader to a gestalt switch. The simple matrix pictures in any textbook on game theory are models for sophisticated modal logics of action, knowledge, and preference. We introduced these logics originally for extensive games in Part I, but they turn out to make equal sense for strategic games, although with a number of new twists. We have developed a bit of the resulting reasoning about knowledge, freedom, best response, and rationality, showing a number of interesting valid patterns, and finding sources of complexity in assumptions of dependence and independence.

Many new research problems come to light in this view, which has not been as well-investigated as the extensive games of Parts I and II. These have occurred at various places in this chapter, concerning complexity of validity, and complete axiomatizations of various logics, modal or extended hybrid, of full and general matrix game models. In addition, strategic form games raise issues beyond those investigated in earlier chapters, including correlations between behavior of players, linking with a current interest in dependence logics, and with reasoning about simultaneous action.

12.10 Literature

This chapter is based on van Benthem (2007d), van Benthem et al. (2011), and van Benthem & Pacuit (2012).

For further links between logic and classical game theory, offering a wide spectrum of motivations and tools, see Battigalli & Bonanno (1999a), Stalnaker (1999), Bonanno (2001), Halpern (2003a), Brandenburger & Keisler (2006), Kaneko (2002), Lorini et al. (2009), Kaneko & Suzuki (2003), and Fitting (2011). Bonanno (2012) surveys ways of construing logics for games under different scenarios of exploratory and in-game thinking, with an insightful analysis of the role of strategies. Surveys of recent contacts between modal logic and games are found in van der Hoek & Pauly (2006), de Bruin (2010), Dégremont (2010), and Zvesper (2010).

12.11 Further directions

This chapter proposed a modal logic perspective on games in strategic form, but many further directions remain to be explored. Here are a few.

From knowledge to belief The preceding discussion has been couched entirely in terms of knowledge. But as in Part II, agency is usually driven by beliefs, and so one would like to redo things in that style. One simple format would add the plausibility orderings of Chapter 7 to our game models, modeling players' beliefs as what is true in all most plausible epistemically accessible strategy profiles. Again, one could think of these expectations as produced by a dynamic process, either one of deliberation, as explored in Chapter 8, or of external information signals about behavior of other players, as discussed in Chapter 9. There is no obstacle in principle to such an extension of our logics for strategic games, and in fact, doing so may get closer to belief-based models of games used by game theorists (cf. Battigalli & Bonanno 1999a, Perea 2011, Brandenburger et al. 2014).

From games to general social action While our matrix logics stay close to strategic games, we can also view them as a way of analyzing more general social scenarios. Hu & Kaneko (2012) give a case study, tying logic to the general postulates for competitive games in Johansen (1982), that extract the basics of styles of social interaction at a higher abstraction level. Moving beyond games, the authors also make interesting connections with philosophical views on determinism and free will as guiding our intuitions about social action, using logics with simultaneous fixed points, such as those in Parts I and II of this book. Finding general postulates characterizing major styles of social interaction is an intriguing challenge for our logical approach as well.

Evolutionary game theory Matrix games are usually seen as one-shot scenarios, but, infinitely repeated, they are also the standard building blocks for evolutionary games (Maynard-Smith 1982, Hofbauer & Sigmund 1998, Gintis 2000). Evolutionary game theory is a major pillar of modern game theory, and in the form of signaling games whose equilibria set up matches between real world situations and code signals, it has found uses in fields such as linguistics and philosophy, witness the examples in our Introduction (cf. Skyrms 1996, van Rooij 2004, Clark 2011).

Some key notions of evolutionary game theory fall directly under the style of analysis in this chapter. For instance, Kooistra (2012) shows, following Gintis (2000), how to collapse players and strategies in evolutionary games, leading to matrix models with a relation $\tau \leq_\sigma \tau'$ saying that strategy τ' leads to better results than strategy τ from the perspective of the σ player. This reflects our earlier remark about preferences between strategy profiles rather than outcomes. Over these models, evolutionarily stable equilibria can be defined using the above modal logics plus some hybrid gadgets.

Another approach to evolutionary games would use the temporal logic of forcing in infinite games developed in Chapter 5, now in a version including simultaneous moves by all players. It seems quite feasible to generalize most of the results that we obtained for sequential action, perhaps exploiting analogies with alternating-time temporal logic (Alur et al. 2002).

However, the main feature of evolutionary games is the temporal progression of a dynamical system driven by recursion equations of fitness as affected by encounters described in a matrix game. The logical dynamics of such a system calls for a representation of the resulting *population changes* rather than belief changes for individual players. Therefore, we need systems in the spirit of Chapter 7 that update numbers of players following the given strategies. In addition to this local dynamics, evolutionary games typically show long-term temporal behavior that is sui generis. The logics of Chapters 4 and 5 may have something to contribute here – witness our brief discussion of typical evolutionary strategies such as Tit for Tat (Axelrod 1984) – and the same may be true for the learning scenarios via limits of updates introduced in Chapter 7 (cf. Leyton-Brown & Shoham 2008, Hutegger & Skyrms 2012 for evolutionary views on learning).

There have hardly been any logical studies of general dynamical systems per se (but see Kremer & Mints 2007 for a dynamic topological logic capturing some key recurrence patterns in limit behavior). It should be easy to add more expressive temporal languages capturing further features of infinite long-term behavior.

Evolutionary games are a natural extension of the concerns in this book, describing phenomena of mass behavior or public opinion that reign when details of deliberation have long been forgotten. They also pose a challenge of interfacing the dynamic logics in Part II, driven by discrete recursion laws for short episodes, with the differential equations driving the long-term behavior of dynamical systems. The current lack of a natural connection between these two worlds seems to be one of the major open ends in this book.

Probability, belief, and behavior Another major open end is the missing analysis of the use of probability of game theory. Probabilities enter in various guises in the study of agency. They may represent strengths of belief for individual agents, but they may also stand for frequencies of behavior in entire populations, or aggregated experiences in memory (Bod et al. 2003). There is no obstacle in principle to combining the logics in this book with either kind of probability. As noted in Chapter 7, van Benthem et al. (2009b) show how to extend update in dynamic-epistemic logics with both individual strengths of belief and more frequency-like process probabilities. But as pointed out in the Introduction, probabilities play one more important role in game theory, as they extend the set of possible behaviors making every matrix game have a Nash equilibrium in mixed strategies. The closest logical counterpart to such a construction in this book may be the imperfect information games of semantic evaluation in Chapter 21, where the logical universe of objects and functions can get probabilized. A lot more can and should be said on these interfaces of logic and probability in the study of social behavior, but, perhaps unwisely, this topic is outside the scope of this book.