Journal of Artificial Intelligence Research 49 (2014) 143-170 Submitted 10/13; published 02/14

A Procedural Characterization of Solution Concepts in Games

Joseph Y. Halpern HALPERN@CS.CORNELL.EDU
Computer Science Department
Cornell University
Ithaca, NY 14853, USA

Yoram Moses MOSES@EE.TECHNION.AC.IL
Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa, 32000, Israel

Abstract

We show how game-theoretic solution concepts such as Nash equilibrium, correlated equilibrium, rationalizability, and sequential equilibrium can be given a uniform definition in terms of a *knowledge-based program* with counterfactual semantics. In a precise sense, this program can be viewed as providing a procedural characterization of rationality.

1. Introduction

There is a general intuition that, in many situations, what players do depends on what they know. This leads to the hope that we can describe players' actions procedurally using *knowledge-based (kb) programs* (Fagin, Halpern, Moses, & Vardi, 1995, 1997) of the form "if I know (or do not know) X then I should do Y". For example, a kb program could say "If you don't know that Ann received the information, then send her a text message", which can be written

$$\textbf{if } \neg K_i(\text{Ann received info}) \textbf{ then } \text{send Ann a text message.}$$

This kb program has the form of a standard **if** ... **then** statement, except that the test in the **if** clause involves i's knowledge (expressed using the modal operator K_i).

Knowledge-based programs have been successfully applied in distributed computing, both to help in the design of new protocols and to clarify our understanding of existing protocols (see, e.g., Fagin et al., 1997; Dwork & Moses, 1990; Hadzilacos, 1987; Halpern, Moses, & Waarts, 2001; Halpern & Zuck, 1992; Mazer & Lochovsky, 1990; Mazer, 1990; Moses & Kislev, 1993; Moses & Tuttle, 1988; Neiger & Bazzi, 1992; Neiger & Toueg, 1993). They have also been applied successfully in planning (Brafman, Latombe, Moses, & Shoham, 1997; Lang & Zanuttini, 2012, 2013; Reiter, 2001). In this paper, we initiate a project on the use of kb programs in game theory. This seems like a particularly fruitful application area, since this seems to be just the kind of reasoning that people employ in games and decision-making problems.

We focus on one application of kb programs to game theory: characterizing solution concepts. Many solution concepts have been considered in the game-theory literature, ranging from Nash equilibrium and correlated equilibrium to refinements of Nash equilibrium such as sequential equilibrium and weaker notions such as rationalizability (see Osborne & Rubinstein, 1994, for an overview).

Typically, these solution concepts assume that players are rational, in the sense that each player's strategy represents a best response to the beliefs that he has about the strategies that other players are using. Indeed, a number of epistemic characterizations of various solution concepts have been provided in the literature. Of particular interest to us are characterizations in terms of *common knowledge* of rationality. These characterizations all have the following flavor: $\vec{\sigma}$ satisfies solution concept X iff there is a state ω in a model M where (a) it is common knowledge that players are rational, (b) the players are playing strategy profile $\vec{\sigma}$, and (possibly) (c) M satisfies some additional property. No additional properties are needed if X is rationalizability (Brandenburger & Dekel, 1987); we get correlated equilibrium if we assume that players have a common prior in M (Aumann, 1987); we get Nash equilibrium if we assume that this prior takes strategy choices to be uncorrelated (Aumann, 1987).[1] Similar results can be proved for sequential equilibrium and perfect equilibrium (Halpern & Moses, 2010).

The standard semantics of kb programs (see Section 2) essentially ensures that the kb programs being run by the players are common knowledge. Thus, we might hope that if we could find a kb program that captures rationality, then we could use it to give a procedural characterization of various solution concepts. In this paper we show that this goal is attainable.

Consider the following kb program $\mathbf{EQ}_i^\Gamma$ for player i, where $\mathcal{A}_i(\Gamma)$ denotes all the possible actions available to player i in game Γ. (This program applies both to normal-form and to extensive-form games. In an extensive-form game Γ, $\mathcal{A}_i(\Gamma)$ is the union of the actions available at each of i's information sets; we assume without loss of generality that the sets of actions at different information sets are disjoint.)

$$\begin{array}{l} \textbf{for each } \text{action } \mathsf{a} \in \mathcal{A}_i(\Gamma) \textbf{ do} \\ \quad \textbf{if } K_i(\mathit{intend}_i(\mathsf{a}) \ \wedge \ \bigwedge_{\mathsf{a}' \in \mathcal{A}_i(\Gamma)} \mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')) \\ \quad \textbf{then } \mathsf{play\ a}. \end{array}$$

Intuitively, we can think of this program as one that ensures that players are expected utility maximizers: Player i follows an action only if playing that action maximizes i's expected utility. More literally, $\mathbf{EQ}_i^\Gamma$ says that player i should play action a if she currently intends to do so (that is the intended semantics of $\mathit{intend}_i(\mathsf{a})$), and her utility from doing so is no less than her utility would be from playing an alternative action. Here $\mathrm{EU}_i(\mathsf{a}')$ represents i's expected utility if she were to play a' (conditional on i's beliefs at the current information set when we consider extensive-form games).[2] The utility is taken with respect to player i's current probability distribution (which can be viewed as a distribution over the strategy profiles of the other players, and thus represents i's beliefs about what other players are doing).

1. Although Aumann (1987) does not state the result for Nash equilibrium explicitly, it follows easily from his results on correlated equilibrium.
2. Note that $\mathrm{EU}_i(\mathsf{a}')$ incorporates a counterfactual. We are in a situation where i is actually intending to play a, but $\mathrm{EU}_i(\mathsf{a}')$ is what her expected utility would be were she to play a' instead. In the preliminary version of this paper (Halpern & Moses, 2007), we made this counterfactual reasoning explicit by having a counterfactual operator in the language. Following the suggestion of a reviewer, we have suppressed the counterfactual here, so as to focus on the issues of most interest to us.

Programs are generally viewed as describing when various actions are supposed to be performed, thereby providing a procedural specification of behavior. Knowledge-based programs are, in general, somewhat different. For example, consider the program $\mathbf{EQ}_i^\Gamma$. It is written assuming that each player has already chosen her strategy. One possible way to view $\overrightarrow{\mathbf{EQ}}^\Gamma$ (i.e., $(\mathbf{EQ}_1^\Gamma, \ldots, \mathbf{EQ}_n^\Gamma)$) is as specifying that, given a last-minute chance to change their minds, none of the (utility-maximizing) players would have a reason to deviate from their individual choices.

The test in $\mathbf{EQ}_i^\Gamma$ can be seen as embodying (one standard form of) rationality: it says that i should act so as to maximize i's expected utility. If it is common knowledge that each player i follows $\mathbf{EQ}_i^\Gamma$ (recall that the semantics of kb programs essentially ensures that the programs being run are commonly known), then, intuitively, the players should have common knowledge of rationality, and so be in equilibrium. Indeed, as we show, under appropriate assumptions, what is played when players act according to $\overrightarrow{\mathbf{EQ}}^\Gamma$ is an instance of a standard solution concept. $\overrightarrow{\mathbf{EQ}}^\Gamma$ captures many solution concepts in a uniform way. Solution concepts differ in the assumptions that are made about players' beliefs. The assumptions are essentially the ones that arose in the logical characterizations of solution concepts mentioned above. Thus, for example, for correlated equilibrium, rather than requiring common knowledge of rationality and that there is a common prior, we now require that players use $\overrightarrow{\mathbf{EQ}}^\Gamma$ and there is a common prior; similarly for the other solution concepts.

The semantics of kb programs allows us to capture these assumptions about players' beliefs by restricting to appropriate systems. The upshot is that we have a single kb program that arguably gives a procedural embodiment of rationality. Moreover, common knowledge that this program is being run provides us with a characterization of perhaps the most common solution concepts used in game theory: Nash equilibrium, correlated equilibrium, rationalizability, sequential equilibrium, and perfect equilibrium.

There has been considerable work in the last two decades focusing on the interplay between modal logic and game theory, and more specifically on epistemic logic and solution concepts in games (see, e.g., Aumann & Brandenburger, 1995; Benthem, 2007, 2010; Bruin, 2010; Bonanno, 2002; Harrenstein, Hoek, Meyer, & Witteveen, 2002; Lorini & Schwarzentruber, 2010). This paper differs from that line of work in that it relates equilibrium notions to knowledge-based programs, and shows the procedural commonality among equilibrium notions, as well as how slightly varying the epistemic assumptions gives rise to the different notions.

The rest of this paper is devoted to making these claims precise. In Section 2, we review the relevant background on game theory and knowledge-based programs. To give formal semantics to kb programs, we use the *runs-and-systems* framework (Fagin et al., 1995), which has been used in the computer science literature to represent complex systems. We specialize the framework so that it can represent the games that are of interest here. In Section 3, we show that $\mathbf{EQ}^\Gamma$ characterizes Nash equilibrium, correlated equilibrium, rationalizability, and sequential equilibrium in a game Γ, each in an appropriate context. We conclude in Section 4 with a discussion of our results, their implications and possible extensions, and a general discussion of the potential use of kb programs in game theory. In particular, we argue that despite the non-negligible overhead involved in dealing with knowledge-based programs, using them gives us a flexible and powerful tool for capturing intuitions such as best response.

2. Background and Definitions

In this section, we review the relevant background on games and knowledge-based programs, and define the semantics of the knowledge-based program $\mathbf{EQ}_i^\Gamma$ formally. We describe only what we need for proving our results. The reader is encouraged to consult a standard game theory text (e.g., Osborne & Rubinstein, 1994) for more on game theory, and the work of Fagin et al. (1995, 1997) for more on the runs-and-systems framework and on knowledge-based programs.

2.1 Games and Strategies

A game Γ in *extensive form* is described by a game tree $\mathcal{T} = \mathcal{T}_\Gamma$. A utility $u_i(h)$ is defined for each terminal history h in the game tree (where a *terminal history* in a game tree is just a path leading from the root to a leaf), specifying player i's utility when that history is played. Let Z_Γ denote the set of all terminal histories. (We omit the subscript Γ if it is clear from context.) Although it is typical to assume that only one player at a time moves in an extensive-form game, we allow arbitrary subsets of players to move. This added generality allows us to view normal-form games as a special case of extensive-form games. Thus, associated with each non-leaf node is the subset of players whose move it is at that node.[3] For each non-leaf node w of $\mathcal{T}$, there is a bijection between the possible sets of moves that can be played at w and the successors of w. If w' is the successor of w that corresponds to a particular set of moves, then w' can be thought of as the outcome of playing those moves at w. The nodes where a player i moves are partitioned into *information sets*.

A *behavioral strategy* σ for player i in an extensive-form game associates with each information set I a distribution $\sigma(I)$ over the actions that can be played at I. Thus, a strategy for player i tells player i what to do at each node in the game tree where i is supposed to move. The fact that a strategy determines actions as a function of the information sets captures the intuition that, at all the nodes that player i cannot tell apart, player i must do the same thing. A *pure* strategy S_i for i is deterministic, specifying a single action per information set. (Of course, a pure strategy can be viewed as a special case of behavioral strategy; it is a behavioral strategy that puts probability 1 on a particular action at each information set.) Since the game tree is assumed to be finite, there are only finitely many pure strategy profiles. A *mixed* strategy σ_i for i is a distribution over pure strategies.[4] Note that a player using a mixed strategy randomizes only once, at the beginning of the game; by way of contrast, a player using a behavioral strategy randomizes at each information set. A *pure (resp., mixed; behavioral) strategy profile* is a tuple $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ specifying a pure (resp., mixed; behavioral) strategy for each player. As usual, given a profile $\vec{x}$, we denote by $\vec{x}_{-i}$ the partial profile containing a component for all players

3. For ease of exposition, at this point we consider only games where there are no moves by nature. There is no difficulty in dealing with moves by nature; we discuss this in Section 4.
4. We consistently use S_i to denote a pure strategy and σ_i to denote a mixed strategy or a behavioral strategy.

other than i. We denote by $\mathcal{S} = \mathcal{S}(\mathcal{T})$ the set of all pure strategy profiles $\vec{S} = (S_1, \ldots, S_n)$ for the game tree $\mathcal{T}$.

A *normal-form game* can be viewed as a special case of an extensive-form game where each player makes only one move, and all players move simultaneously. The tree corresponding to such a game has a depth of one: all nodes other than the root are leaves.

2.2 The Runs-and-Systems Framework

To explain kb programs, we must first describe the runs-and-systems framework. We assume that, at any given point in time, a player in a game is in some *local state*. The local state could include the history of the game up to this point, the strategy being used by the player, and perhaps some other features of the player's type, such as beliefs about the strategies being used by other players. As we shall see, for the purposes of this paper, to model games, a player's local state will essentially consist of his strategy and his information set. A *global state* is a profile of local states: one local state for each player.

A *run* is a sequence of global states; formally, a run is a function from times to global states. Thus, $r(m)$ is the global state in run r at time m. For definiteness, we assume that time ranges over the natural numbers here. A *point* is a pair (r, m) consisting of a run r and time m. Let $r_i(m)$ be i's local state at the point (r, m); that is, if $r(m) = (s_1, \ldots, s_n)$, then $r_i(m) = s_i$. A *system* is a set of runs. A *probabilistic system* is a tuple $\mathcal{PS} = (\mathcal{R}, \vec{\mu})$, where $\mathcal{R}$ is a system and $\vec{\mu} = (\mu_1, \ldots, \mu_n)$ associates a probability μ_i on the runs of $\mathcal{R}$ with each player i. Intuitively, μ_i represents player i's prior beliefs. In the special case where $\mu_1 = \cdots = \mu_n = \mu$, the players have a *common prior* μ on $\mathcal{R}$. In this case, we write just $(\mathcal{R}, \mu)$.

2.3 Modeling a Game as a System

With each game Γ, we associate a system $\mathcal{R}^\Gamma$. In describing the system $\mathcal{R}^\Gamma$, we have to decide how to model the players' local states, that is, what they know at each point in the system? For the most part, the details do not matter for the analysis we do in this paper, but they can have a critical effect on other analyses. We believe that one of the advantages of the runs-and-systems approach is that it forces the modeler to think through carefully what the players' local states should be.

In the case of a normal-form game Γ, at time 0, we can take a player's local state to consist of the pure strategy that she is intending to play; the player's local state at times $m \geq 1$ consists of the strategy she played (which we take to be the same as the one that she was intending to play) and her utility. Even if we think of a player as having a mixed strategy, we can think of the pure strategy in her local state as the pure strategy that the player chooses after tossing her coin. We could, of course, also include the mixed strategy in the player's state. It turns out that doing so would make no difference; player i's strategy is encoded by the distribution μ_i on runs. The upshot of this approach to modeling things is that we can identify a run in $\mathcal{R}^\Gamma$ with a pure strategy profile in Γ; we denote by $r^{\vec{S}}$ the run coresponding to strategy profile $\vec{S}$.

To take a simple example, consider the normal-form game Γ_n in Figure 1, between two players, Alice (the row player) and Bob (the column player): There are four runs in $\mathcal{R}^{\Gamma_n}$, corresponding to the four strategy profiles in the game. In the run $r^{(T,L)}$, we have

	L	R
T	$(3,3)$	$(1,4)$
B	$(4,1)$	$(0,0)$

Figure 1: A simple 2-player game Γ_n.

$r^{(T,L)}(0) = (T, L)$ and $r^{(T,L)}(1) = ((T, 3), (L, 3))$. In this run, Alice's initial state is T, her intended strategy, and her state at all times $m \geq 1$ is $(T, 3)$, her strategy and utility.

In an extensive-form game, we take the runs in $\mathcal{R}^\Gamma$ to correspond to the terminal histories in the game tree; there is one run r^h corresponding to each history h. The points on a run r^h correspond to the nodes in the history h. That is, the global state $r^h(m)$ corresponds to the mth node w in history h. (If m is greater than the length of h, then $r^h(m) = r^h(|h|)$.) Suppose that i moves at node w. Then we can take i's local state at $r^h(m)$ to have the form (I_w, a), where I_w is i's information set at w, and a is the move that i makes at information set I_w in history h. Intuitively, this is saying that i knows his information set, and the move that he is intending to make. For the solution concepts we analyze in extensive-form games, we assume that i is using a behavioral strategy. Thus, we can think of a as the outcome of the coin toss at information set I in i's behavioral strategy. We still need to represent i's local state at points that correspond to nodes w where i does not move. The details of i's local state at points where i does not move do not matter much. For definiteness, if i does not move at the node w corresponding to (r^h, m), then we take i's local state to be $I_{w'}$, where w' is the most recent node preceding w where i does move; if i has not moved prior to w, then we take i's local state to be just $\langle \rangle$. If w is a final node in a terminal history, we also encode i's utility in i's local state, just as we did for normal-form games. Thus, i's state encodes i's information about what has happened thus far (this is the I_w component), whether or not it is i's move (this is captured by whether or not there is an action component for i in the local state), what i intends to do if it is his move, and (at points that correspond to the end of the game), each player's utility.

In the solution concepts we focus on here for extensive-form games, we think of the players as using a behavioral strategy. However, just as in normal-form games, we do not encode i's behavioral strategy in his local state. And again, it will turn out that the behavioral strategy is encoded in the probability distribution on runs.[5] We could have also included the game Γ itself in each player's local states, since we have implicitly assuming that the game Γ is common knowledge. Doing so would not change anything in the analysis; we have not done it simply to avoid cluttering the notation. Of course, this would be an appropriate thing to do in games where players are not fully aware of what game is being played (see, e.g., Halpern & Rêgo, 2013).

5. The alert reader may have spotted a potential problem here. Runs off the equilibrium path (and hence, information sets off the equilibrium path) get probability 0, so it may seem that we cannot use the probability on runs to infer i's behavioral strategy at such information sets. But, as we shall see, the probability on runs that we actually use is a nonstandard probability that is infinitesimally close to the actual probability generated by the strategy profile. This nonstandard probability gives positive probability to all runs, and hence can be used to infer i's behavioral strategy. But this is an artifact of our approach. In other contexts, we may well want to include i's behavioral strategy in i's local state.

Consider the 2-player extensive-form game Γ_e in Figure 2, where player 1 moves at w_1, and player 2 moves at the information set $\{w_2, w_3\}$:

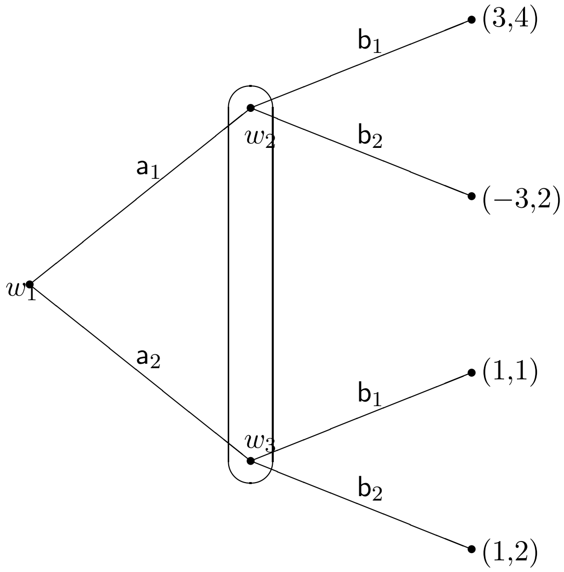

Figure 2: The extensive-form game Γ_e.

There are four runs in $\mathcal{R}^{\Gamma_e}$, corresponding to the four terminal histories. Call these histories h_1–h_4, going from top to bottom. Thus, in r^{h_1} (where the utility is (3,4)), 1's local state at $r^{h_1}(0)$ is $(\{w_1\}, \mathsf{a}_1)$, while 2's local state is $\langle\,\rangle$. At $r^{h_1}(1)$, 1's local state is $\{w_1\}$, while 2's is $(\{w_2, w_3\}, \mathsf{b}_1)$, and at $r^{h_1}(2)$, i's local state is $(\{w_1\}, 3)$, while 2's local state is $(\{w_2, w_3\}, 4)$.

Since there is a bijection between runs in $\mathcal{R}^{\Gamma}$ for a normal-form game Γ and pure strategy profiles, a distribution μ on pure strategies in a normal-form game Γ can be identified with a distribution on the runs in $\mathcal{R}^{\Gamma}$. Thus, we can associate with each mixed strategy profile $\vec{\sigma}$ in a normal-form game Γ the probabilistic system $(\mathcal{R}^{\Gamma}, \mu_{\vec{\sigma}})$, where $\mu_{\vec{\sigma}}$ is the distribution on strategy profiles (and hence also on runs) induced by $\vec{\sigma}$. Note that according to $\mu_{\vec{\sigma}}$, players' strategy choices are uncorrelated; the probability that player i chooses S_i and player j chooses S_j is just the product of the probability that i chooses S_i and the probability that j chooses S_j. Similarly, in an extensive-form game Γ, a behavioral strategy profile $\vec{\sigma}$ induces a distribution on histories, and hence also on the runs in $\mathcal{R}^{\Gamma}$. We again denote this distribution $\mu_{\vec{\sigma}}$.

2.4 Knowledge-Based Programs

A *knowledge-based program* is a syntactic object. For our purposes, a knowledge-based program for player i is taken to have the form

if κ_1 **then** a_1
if κ_2 **then** a_2
$\ldots,$

where each a_j is an action for i, and each κ_j is a Boolean combination of formulas of the form $K_i\varphi$, in which the φ's can have nested occurrences of K_ℓ operators. We assume that the tests $\kappa_1, \kappa_2, \ldots$ are mutually exclusive, so that, in a kb program for player i, at most one of the tests evaluates to true in each local state for player i at which i moves. The program $\mathbf{EQ}_i^\Gamma$ can be written in this form by simply replacing the **for** ... **do** statement by one line for each possible action of i in the game Γ. That is, for each action $\mathsf{a} \in \mathcal{A}_i(\Gamma)$, there is a line in $\mathbf{EQ}_i^\Gamma$ of the form

$$\textbf{if } K_i(\mathit{intend}_i(\mathsf{a}) \wedge \bigwedge_{\mathsf{a}' \in \mathcal{S}_i(\mathcal{A})} \mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')) \textbf{ then } \mathsf{play\ a}.$$

Since a player i intends to play at most one action at a point (r, m) in $(\mathcal{R}^\Gamma, \vec{\mu})$, the tests in $\mathbf{EQ}_i^\Gamma$ are mutually exclusive. The tests are not necessarily exhaustive, since at a point in which the strategy S that player i uses is not a best response, no test of the form above is satisfied. Roughly speaking, if none of the tests is satisfied, then i does nothing (and performs the null action skip).

We want to define what it means for a (probabilistic) system $\mathcal{PS}$ to be *compatible* with a kb program. Intuitively, this is the case when all the moves made in $\mathcal{PS}$ are the ones recommended by the kb program. For simplicity, we give just enough of the required definitions here to be able to handle the case that $\mathcal{PS}$ has the form $(\mathcal{R}^\Gamma, \vec{\mu})$ and and the kb program is $\mathbf{EQ}_i^\Gamma$. For further details, the interested reader can consult Fagin et al. (1995, 1997).

As a first step to making this precise, for each standard system $\mathcal{PS} = (\mathcal{R}^\Gamma, \vec{\mu})$, we associate with each formula φ a set $[\![\varphi]\!]_{\mathcal{PS}}$ of points in $\mathcal{PS}$. Intuitively, $[\![\varphi]\!]_{\mathcal{PS}}$ is the set of points of $\mathcal{R}^\Gamma$ where φ is true. For $\mathit{intend}_i(\mathsf{a})$ this is easy:

- $[\![\mathit{intend}_i(\mathsf{a})]\!]_{\mathcal{PS}}$ is the set of points (r, m) of $\mathcal{PS}$ at which i moves at the node w in the game tree associated with (r, m) and a is the action encoded in i's local state.

Note that a kb program Pg_i for player i can attempt to "override" i's intentions; that is, the program can have a line of the form "**if** κ **then** $\mathsf{play\ a}'$" such that, at a point (r, m) in the probabilistic system $\mathcal{PS}$, κ is true (i.e., $(r, m) \in [\![\kappa]\!]_{\mathcal{PS}}$), but the action in i's local state is a, not a'. In this case, as we shall see, Pg would not be compatible with $\mathcal{PS}$.

The semantics of knowledge is defined as usual: the formula $K_i\varphi$ is true if φ is true at all the points that i considers possible. We view all of i's information at a point (r, m) as being encapsulated by i's local state at (r, m), which we denote $r_i(m)$. Thus, the set of points that i considers possible at a point (r, m) is $\mathcal{K}_i(r, m) = \{(r', m') : r'_i(m') = r_i(m)\}$; $\mathcal{K}_i(r, m)$ just consists of all the points where i has the same local state as at (r, m).

- $[\![K_i\varphi]\!]_{\mathcal{PS}}$ is the set of points (r, m) such that $\mathcal{K}_i(r, m) \subseteq [\![\varphi]\!]_{\mathcal{PS}}$.

It remains to give semantics to formulas of the form $\mathrm{EU}_i(\mathsf{a}') \geq \mathrm{EU}_i(\mathsf{a})$. Clearly the expected utility that i would obtain if i were to play a' at the point (r, m) depends on i's beliefs about what the other players are doing at (r, m). Roughly speaking, these beliefs are obtained by conditioning i's prior beliefs μ_i on $\mathcal{K}_i(r, m)$. But there is a small technical problem here. The probability distribution μ_i is a distribution on runs; $\mathcal{K}_i(r, m)$ is a set of points. We cannot condition μ_i on $\mathcal{K}_i(r,m)$. To enable conditioning, we first associate with $\mathcal{K}_i(r,m)$ the set $\mathcal{R}[\mathcal{K}_i(r,m)]$ of those runs that go through points in $\mathcal{K}_i(r,m)$; that is,

$$\mathcal{R}[\mathcal{K}_i(r,m)] = \{r' : (r',m) \in \mathcal{K}_i(r,m)).$$

We then define $\mu_{i,r,m} = \mu_i \mid \mathcal{R}[\mathcal{K}_i(r,m)]$. (For the purposes of this paper, we do not specify $\mu_{i,r,m}$ if $\mu_i(\mathcal{R}[\mathcal{K}_i(r,m)]) = 0$. It turns out to be irrelevant to our discussion.) Recall that there is a bijection between the runs in $\mathcal{R}^\Gamma$ and pure strategy profiles. Moreover, since player i knows his strategy, at all the runs in $\mathcal{R}[\mathcal{K}_i(r,m)]$, player i is using the same strategy. Thus, $\mu_{i,r,m}$ determines a distribution on $\vec{S}_{-i}$. We use this distribution to compute $\mathrm{EU}_i(\mathsf{a})$ and $\mathrm{EU}_i(\mathsf{a}')$. In the case of a normal-form game Γ, this suffices to compute i's expected utility if he were to play a' (although a' may not *in fact* be the strategy that i intends to use in the runs in $\mathcal{R}[\mathcal{K}_i(r,m)]$), under the assumption that all other players do use the strategy that they were intending to use; only i's strategy changes. In an extensive-form game, when considering a change from a to a' in a run r, we keep all other actions fixed (i.e., the actions of all other players throughout the run, and i's actions off the information set I), and consider the utility of the resulting run.

- If Γ is a normal-form game and $\mathcal{PS} = (\mathcal{R}^\Gamma, \vec{\mu})$, then $[\![\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')]\!]_{\mathcal{PS}}$ consists of those points (r,m) at which the expected utility for i of using a is at least as high as that of using a', where the expectation is taken with respect to the distribution on strategy profiles a_{-i} of the other players, as determined by $\mu_{i,r,m}$. If Γ is an extensive-form game, and at the point (r,m) player i intends to play action a, i is in information set I, and it is i's move. To compute $\mathrm{EU}_i(\mathsf{a}')$, for each run r' in $\mathcal{K}_i(r,m)$, let $h_{r'}[\mathsf{a}/\mathsf{a}']$ be the history in Γ where the same sequence of actions is played by each player as in r', except that at information set I, i plays a' rather than a. Then $\mathrm{EU}_i(\mathsf{a}') = \sum_{r' \in \mathcal{K}_i(r,m)} \mu_{i,r,m}(r') u_i(h_{r'}[\mathsf{a}/\mathsf{a}'])$. Again, $[\![\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')]\!]_{\mathcal{PS}}$ consists of all those points (r,m) where i moves and $\mathrm{EU}_i(\mathsf{a}')$ (computed as above) is no higher than $\mathrm{EU}_i(\mathsf{a})$.

Since player i knows his strategy (it appears in his local state), if $(r,m) \in [\![\mathit{intend}_i(\mathsf{a})]\!]_{\mathcal{PS}}$, then $\mathcal{K}_i(r,m) \subseteq [\![\mathit{intend}_i(\mathsf{a})]\!]_{\mathcal{PS}}$. Similarly, since $\mu_{i,r,m} = \mu_{i,r',m'}$ if $(r',m') \in \mathcal{K}_i(r,m)$, player i knows his probability distribution, so if $(r,m) \in [\![\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')]\!]_{\mathcal{PS}}$, then $\mathcal{K}_i(r,m) \subseteq [\![\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')]\!]_{\mathcal{PS}}$. Hence, the formula $\mathit{intend}_i(\mathsf{a}) \ \wedge\ \bigwedge_{\mathsf{a}' \in \mathcal{A}_i(\Gamma)} \mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')$ is equivalent to the epistemic formula $K_i(\mathit{intend}_i(\mathsf{a}) \ \wedge\ \bigwedge_{\mathsf{a}' \in \mathcal{A}_i(\Gamma)} \mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}'))$; that is, $\mathit{intend}_i(\mathsf{a}) \ \wedge\ \bigwedge_{\mathsf{a}' \in \mathcal{A}_i(\Gamma)} \mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')$ is true iff player i knows it. We have kept the K_i in the kb program just to emphasize that this is a formula whose truth depends only on what i knows and believes, and thus is a test that i can act on.

Intuitively, a system $\mathcal{PS}$ is compatible with a kb program profile $\overrightarrow{\mathsf{Pg}}$ if $\mathcal{PS}$ could have arisen if each player i uses Pg_i. We formalize this as follows.

Definition 2.1 *$\mathcal{PS}$ is compatible with the kb program profile* $\vec{\mathsf{Pg}}$ if for all $r \in \mathcal{R}$ and $m \geq 0$, there is an action profile $\vec{\mathsf{a}}$ such that both (a) applying $\vec{\mathsf{a}}$ to (r,m) results in $(r,m+1)$ (see below) and (b) for each i such that $\mu_i(\mathcal{R}[\mathcal{K}_i(r,m)]) > 0$ and $\mu_i(r \mid \mathcal{R}[\mathcal{K}_i(r,m)]) > 0$, either there is a line **if** κ **then** a_i in Pg_i and $(r,m) \in [\![\kappa]\!]_{\mathcal{PS}}$, or there is no such line and a_i is the null move skip. ∎

We have not explained what it means to apply an action profile $\vec{\mathsf{a}}$ to a point (r, m). The general definition involves viewing action profiles as transformers of global states (see Fagin et al., 1995, 1997). Rather than going through the details of the general definition here, we just give the definition in the case that $\mathcal{PS}$ has the form $(\mathcal{R}^\Gamma, \mu)$ and $\overrightarrow{\mathsf{Pg}}$ is $\overrightarrow{\mathbf{EQ}}^\Gamma$. That is all we need for this paper, and in this case, the definition is quite simple. If Γ is a normal-form game, then the action profile $\vec{\mathsf{a}}$ applied to $r^{\vec{S}}(0)$ results in $r^{\vec{S}}(1)$ iff $\vec{\mathsf{a}} = \vec{S}$; if $m \geq 1$, then $\vec{\mathsf{a}}$ applied to $r^{\vec{S}}(m)$ results in $r^{\vec{S}}(m+1)$ iff $\mathsf{a}_1 = \cdots \mathsf{a}_n = \mathsf{skip}$. If Γ is an extensive-form game, then $\vec{\mathsf{a}}$ applied to $r^h(m)$ results in $r^h(m+1)$ iff $\mathsf{a}_i = \mathsf{skip}$ if i does not move at the information set I associated with $r^h(m)$ (which means that, in particular, if m is greater or equal to the length of h, then $\mathsf{a}_i = \mathsf{skip}$ for all players i) and, if i does move at I, then a_i is the move encoded in $r_i^h(m)$. Roughly speaking, this means that for a normal-form game Γ, $\mathcal{PS} = (\mathcal{R}^\Gamma, \mu)$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$ iff, for each player i and run r such that $\mu(r) > 0$, what does according to $\overrightarrow{\mathbf{EQ}}^\Gamma$ at $r(0)$ is what i intends to do according to his local state $r_i(0)$. Thus, if i intends to play strategy S, then it must be the case that $\mathrm{EU}_i(S) \geq \mathrm{EU}_i(S')$ for all other strategies S' for player i. The analogous statement is true in extensive-form games, although to compute whether $\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')$ at a point (r, m), we use the probability conditioned on $\mathcal{R}[\mathcal{K}_i(r, m)]$. That is, if a_i is the action played by i at (r, m), then a_i really is a best response for i, given i's beliefs at (r, m). While this observation makes all the proofs of the results relatively straightforward, it is important to note that this is really an instance of the general semantics of kb programs.

3. The Main Results

In this section, we show that $\mathbf{EQ}^\Gamma$ captures a number of standard solution concepts. We start by considering solution concepts in normal-form games, and then move to extensive-form games.

3.1 Capturing Solution Concepts in Normal-Form Games

We show $\mathbf{EQ}^\Gamma$ captures three of the most studied solution concepts in normal-form games: Nash equilibrium, correlated equilibrium, and rationalizability. The differences in how they are captured highlights the distinctions between the notions.

3.1.1 Nash Equilibrium

Recall that $(\mathcal{R}, \mu)$ is a probabilistic system in which all players have a common prior μ on runs.

Theorem 3.1 *The mixed strategy profile $\vec{\sigma}$ is a Nash equilibrium of the normal-form game Γ iff $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$.*

Proof First suppose that $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ is a Nash equilibrium of the game Γ. To see that $\mathcal{PS} = (\mathcal{R}^\Gamma, \mu_{\vec{\sigma}})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$, it suffices to show that if $\mu_{\vec{\sigma}}(r^{\vec{S}}) > 0$, then S_i is a best response with respect to $\mu_{\vec{\sigma}} \mid \mathcal{R}[\mathcal{K}_i(r^{\vec{S}}, 0)]$. (We need to consider only $(r^{\vec{S}}, 0)$ because Γ is a normal-form game, so there are no moves after time 0.)

Note that $\mu_{\vec{\sigma}} \mid \mathcal{R}[\mathcal{K}_i(r^{\vec{S}}, 0)] = \vec{\sigma}_{-i}$ (under the obvious identification of $\mu_{\vec{\sigma}} \mid \mathcal{K}_i(r, 0)$ with a distribution on $\mathcal{S}_{-i}$). Since $\vec{\sigma}$ is a Nash equilibrium, S_i must be a best response to $\vec{\sigma}_{-i}$. Thus, for all strategies $S' \in \mathcal{S}_i(\Gamma)$, we must have that $(r^{\vec{S}}, 0) \in [\![\mathrm{EU}_i(S) \geq \mathrm{EU}_i(S')]\!]_{\mathcal{PS}}$. It follows that $\mathcal{PS} = (\mathcal{R}^\Gamma, \mu_{\vec{\sigma}})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$.

For the converse, suppose that $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$. We want to show that $\vec{\sigma}$ is a Nash equilibrium. It suffices to show that each pure strategy S_i in the support of σ_i is a best response to $\vec{\sigma}_{-i}$. Let S_i be in the support of σ_i. Choose a strategy profile $\vec{S}_{-i}$ in the support of $\vec{\sigma}_{-i}$. Then $\mu_{\vec{\sigma}}(r^{\vec{S}}) > 0$. Moreover, $\mu_{\vec{\sigma}} \mid \mathcal{R}[\mathcal{K}_i(r^{\vec{S}}, 0)] = \vec{\sigma}_{-i}$. Since $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$ and $\mu_{\vec{\sigma}}(r^{\vec{S}}) > 0$, it must be the case that, for all strategies $S' \in \mathcal{S}_i(\Gamma)$, $(r^{\vec{S}}, 0) \in [\![\mathrm{EU}_i(S) \geq \mathrm{EU}_i(S')]\!]_{\mathcal{PS}}$. That is, S is indeed a best response to $\vec{\sigma}_{-i}$. ∎

Consider the game Γ_n described in Figure 1. It is easy to check that this game has three Nash equilibria: there are two equilibria in pure strategies: (B, L) and (T, R). There is also an equilibrium in mixed strategies where Alice randomizes (uniformly) between T and B, and Bob randomizes between L and R. That means that there are three probabilistic systems of the form $(\mathcal{R}^{\Gamma_n}, \mu)$ compatible with $\mathbf{EQ}^{\Gamma_n}$. In the first, μ puts probability 1 on $r^{(B,L)}$; in the second, μ puts probability 1 on $r^{(T,R)}$, and in the third, μ puts uniform probability on the four runs in the system.

3.1.2 Correlated Equilibrium

As is well known, players can sometimes achieve better outcomes than a Nash equilibrium if they have access to a helpful mediator. Consider the simple 2-player game Γ_n described in Figure 1. Recall that the total utility in each of the three Nash equilibria of the games (that is, the sum of the utilities of the two players) is at most 5. We get a higher total utility by using a trusted mediator, who makes a recommendation by choosing at random between (T, L), (T, R), and (B, L). This gives each player an expected utility of 8/3; thus, the total utility is 16/3. This is an example of a correlated equilibrium since, for example, if the mediator chooses (T, L), and thus sends recommendation T to Alice and L to Bob, then Alice considers it equally likely that Bob was told L and R, and thus has no incentive to deviate; similarly, Bob has no incentive to deviate. In general, a distribution μ over pure strategy profiles is a correlated equilibrium if players cannot do better than following a mediator's recommendation if the mediator makes recommendations according to μ. (Note that, as in our example, if a mediator chooses a (pure) strategy profile $(S_1, \ldots, S_n)$ according to μ, the mediator recommends S_i to player i; player i is told nothing about the strategy profile except for S_i.) Roughly speaking, a correlated equilibrium is a distribution η over (pure) strategy profiles in which every strategy T for player i that has a positive probability is a best response to the conditional probability $\eta \mid T$ projected onto $\mathcal{S}_{-i}$. (Note that the support of $\eta \mid T$ consists of strategy profiles $\vec{S}'$ such that $S'_i = T$, so $\eta \mid T$ can be viewed as a distribution on $\mathcal{S}_{-i}$. Intuitively, if player i knows that the prior probability on pure strategy profiles is η and is told to play T, then he believes that the probability on the strategy profiles in $\mathcal{S}_{-i}$ played by the other players is described by $\eta \mid T$ (projected onto $\mathcal{S}_{-i}$). Conversely, if η is a distribution over pure strategy profiles such that, for each player i and every strategy T for player i that is given positive probability by η is a best response to $\eta \mid T$, then η is a correlated equilibrium. Formally, we have the following definition.

Definition 3.2 (Aumann, 1974) A distribution η on pure strategy profiles is a *correlated equilibrium* if for each player i, each strategy S for player i such that $\eta(S) > 0$ (where we identify S with the set of pure strategy profiles $\vec{S}'$ such that $S'_i = S$), and each strategy S' for player i, we have

$$\sum_{S_{-i} \in \mathcal{S}_{-i}} u_i(S, S_{-i})\eta(S_{-i} \mid S) \geq \sum_{S_{-i} \in \mathcal{S}_{-i}} u_i(S', S_{-i})\eta(S_{-i} \mid S). \qquad \blacksquare$$

That is, η is a correlated equilibrium when, for every player i, if the mediator tells i to play a strategy S that has positive probability according to η, then i does not gain from switching to S', given his beliefs about what the other players will do, conditional on player i being told S.

Clearly a distribution η on strategy profiles in game Γ can be identified with a distribution on the runs in $\mathcal{R}^\Gamma$. We can easily capture correlated equilibrium using $\mathbf{EQ}^\Gamma$ in a way that generalizes Theorem 3.1. The only difference between Theorem 3.1 and Theorem 3.3 is that while $\mu_{\vec{\sigma}}$ in Theorem 3.1 is a product measure, the distribution η on runs in Theorem 3.3 is not necessarily a product measure (indeed, it is a product measure iff the correlated equilibrium is a Nash equilibrium).

Theorem 3.3 *The distribution η on strategy profiles is a correlated equilibrium of the (normal-form) game Γ iff $(\mathcal{R}^\Gamma, \eta)$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$.*

Proof The proof proceeds along lines similar to that of Theorem 3.1.

Suppose that η is a correlated equilibrium in Γ and $\eta(r^{\vec{S}}) > 0$. Again, we must show that S_i is a best response to $\eta \mid \mathcal{R}[\mathcal{K}_i(r^{\vec{S}}, 0)]$. But $\mathcal{R}[\mathcal{K}_i(r^{\vec{S}}, 0)]$ consists precisely of the runs where player i plays S_i. Thus, $\mu_{i,r^{\vec{S}},m} = \eta \mid S_i$. Since η is a correlated equilibrium, S_i is a best response to $\eta \mid S_i$; it thus easily follows that $(\mathcal{R}^\Gamma, \eta)$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$.

For the converse, suppose that η is a distribution on strategy profiles such that $(\mathcal{R}^\Gamma, \eta)$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$. We want to show that η is a correlated equilibrium. Suppose that T is a strategy for player i that has positive probability according to η. Thus, there is some run $r = r^{\vec{S}}$ such that $\eta(r) > 0$ and $S_i = T$. As we have seen, $\mu_{i,r,0} = \eta \mid T$. Since $(\mathcal{R}^\Gamma, \eta)$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$, it must be the case that T is a best response to $\eta \mid T$, which determines i's beliefs at $(r, 0)$. Thus, η is a correlated equilibrium. $\blacksquare$

Note that the fact that i's intended strategy is included in i's local state for normal-form games ensures that $\mu_{i,r^{\vec{S}},0} = \eta \mid S_i$. Intuitively, in a correlated equilibrium, the mediator tells i what strategy to follow, and i uses this information in determining a best response. Thus, i's local state should model this information. By way of contrast, Theorem 3.1 would hold even if the strategy were not part of i's local state. Since $\vec{\sigma}$ is a product measure in Theorem 3.1, it would still be the case that $\mu_{i,r,0} = \vec{\sigma}_{-i}$ for all runs r.

3.1.3 Rationalizability

Our characterization of both Nash equilibrium and correlated equilibrium involves a common prior on runs. Dropping this assumption gives rise to another standard solution concept: *rationalizability* (Bernheim, 1984; Pearce, 1984). Intuitively, a strategy for player i

is rationalizable if it is a best response to some beliefs that player i may have about the strategies that other players are following, assuming that each of these strategies is itself a best response to beliefs that one of the other players has about strategies that other players are following, and so on.

Following Osborne and Rubinstein (1994), say that a strategy S for player i in game Γ is *rationalizable* if, for each player j, there is a set $\mathcal{Z}_j \subseteq \mathcal{S}_j(\Gamma)$ and, for each strategy $T \in \mathcal{Z}_j$, a probability measure $\eta_{j,T}$ on $\mathcal{S}_{-j}(\Gamma)$ whose support is a subset of $\mathcal{Z}_{-j}$ such that

- $S \in \mathcal{Z}_i$; and
- for each player j and strategy $T \in \mathcal{Z}_j$, strategy T is a best response to (the beliefs) $\eta_{j,T}$.

Intuitively, the strategies in $\mathcal{Z}_i$ are the rationalizable strategies for player i. Player i can justify playing a strategy $T \in \mathcal{Z}_i$ because, by assumption, there is a distribution on $\mathcal{Z}_{-i}$ (representing i's beliefs about the strategies that other players are using) against which T is a best response. Moreover, each of the strategies to which i assigns positive probability are themselve justifiable, since they are in $\mathcal{Z}_{-i}$, and so are best responses to beliefs that place positive probability on strategies that are justifiable, and so on.

For ease of exposition, we consider only pure rationalizable strategies. This is essentially without loss of generality. It is easy to see that a mixed strategy σ_i for player i is a best response to some beliefs η_i of player i iff each pure strategy in the support of σ_i is a best response to η_i. Moreover, we can assume without loss of generality that the support of η_i consists only of pure strategy profiles.

Notice that in the game Γ_n of Figure 1, all strategies are rationalizable. Alice playing T is justified if Alice believes that Bob will play R; Bob playing R is justified if he believes that Alice will T; Alice playing B is justified if Alice believes that Bob will play L; and Bob playing L is justified if he believes that Alice will play R.

The following theorem characterizes rationalizability in our framework. Note that we now do not assume a common prior, so that there is a vector $\vec{\mu} = (\mu_1, \ldots, \mu_n)$, in which the μ_i's are not necessarily identical, describing the players' beliefs.

Theorem 3.4 *A pure strategy S for player i in the (normal-form) game Γ is rationalizable iff there exists a probabilistic system $\mathcal{PS} = (\mathcal{R}^\Gamma, \vec{\mu})$ and, for each player j, there exists a set $\mathcal{Z}_j \subseteq \mathcal{S}_j$ such that (a) μ_j gives every strategy in $\mathcal{Z}_j$ positive probability; (b) the support of μ_j is contained in $\mathcal{Z} = \mathcal{Z}_1 \times \cdots \times \mathcal{Z}_n$, (c) $S \in \mathcal{Z}_i$, and (d) $(\mathcal{R}^\Gamma, \vec{\mu})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$.*

Proof Suppose that $\mathcal{PS} = (\mathcal{R}^\Gamma, \vec{\mu})$ is a probabilistic system satisfying the four properties above. We want to show that S is rationalizable. Take the sets $\mathcal{Z}_i$ guaranteed to exist by the assumptions of the theorem to be the sets $\mathcal{Z}_i$ in the definition of rationalizability. For $T \in \mathcal{Z}_j$, let $\eta_{j,T}$ be $\mu_j \mid T$ projected onto $\mathcal{S}_{-j}$. Note that $\mu_j \mid T$ is well defined, since $\mu_j(T) > 0$. Moreover, the support of $\eta_{j,T}$ is contained in $\mathcal{Z}_{-j}$, since the support of μ_j is contained in $\mathcal{Z}$. We now show that every $T \in \mathcal{Z}_j$ is a best response to $\eta_{j,T}$. Since $\mu_j(T) > 0$, there must be a run $r^{\vec{S}}$ such that $S_j = T$ and $\mu_j(r^{\vec{S}}) > 0$. It is easy to see that $\mu_{j,r^{\vec{S}},0} = \mu_j \mid T = \eta_{j,T}$. Since $(\mathcal{R}^\Gamma, \vec{\mu})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$, it must be the case that T is a best response to $\eta_{j,T}$. Thus, S is rationalizable.

For the converse, suppose that $S \in \mathcal{S}_i$ is rationalizable. Thus, for each player j, there exist a set $\mathcal{Z}_j$ and, for each strategy $T \in \mathcal{Z}_j$, a measure $\eta_{j,T}$ on $\mathcal{S}_{-j}(\Gamma)$ such that T is a best response to $\eta_{j,T}$. Define μ_j by taking $\mu_j(\vec{S}) = \eta_{j,S_j}(\vec{S_{-j}})/|\mathcal{Z}_j|$ if $S_j \in \mathcal{Z}_j$, and taking $\mu_j(\vec{S}) = 0$ otherwise. First, observe that μ_j is a probability on $\mathcal{S}$: For each strategy $S_j \in \mathcal{Z}_j$, we have $\mu_j(S_j \times \mathcal{S}_{-j}) = \eta_{j,S_j}(\mathcal{S}_{-j})/|\mathcal{Z}_j| = 1/|\mathcal{Z}_j|$; the result easily follows. Moreover, for $S \in \mathcal{S}_j$, we have that $\mu_j(S) > 0$ iff $S \in \mathcal{Z}_j$ (of course, $\mu_j(S)$ is just $\mu_j(S \times \mathcal{S}_{-j})$. And since the support of $\mu_{j,T}$ is contained in $\mathcal{Z}_{-j}$ for each strategy $T \in \mathcal{S}_j$, it easily follows that the support of μ_j is contained in $\mathcal{Z}$. Finally, $S \in \mathcal{Z}_i$, by construction. Since T is a best response to $\eta_{j,T}$ for all $T \in \mathcal{Z}_j$, it easily follows that $\mathcal{PS} = (\mathcal{R}^\Gamma, \vec{\mu})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$. ∎

To see that all strategies in Γ_n are rationalizable, we can actually take $\mu_1 = \mu_2$ to be distributions that assign probability 1/2 to each of (T, R) and (B, L). It is easy to see that this satisfies the conditions of Theorem 3.4, taking $\mathcal{Z}_1 = \{T, B\}$ and $\mathcal{Z}_2 = \{L, R\}$. However, we do not have to take $\mu_1 = \mu_2$. As long as the support of both μ_1 and μ_2 is $\{(T, R), (B, L)\}$, any choice of μ_1 and μ_2 works.

Osborne and Rubinstein's definition of rationalizability allows $\mu_{j,T}$ to be such that j believes that other players' strategy choices are correlated. In most of the literature, players are assumed to believe that other players' choices are made independently. If we add the latter requirement, then we must impose the same requirement on the probability measures $\mu_1, \ldots, \mu_n$ in Theorem 3.4.

It is important in the characterization of rationalizability that i's strategy be part of i's local state. Intuitively, i's strategy together with his beliefs about the strategies of the remaining players determine i's *type*. In modeling rationalizability, it suffices to assume that i's strategy determines i's beliefs, so we can identify i's type with his strategy. By including the strategy in the local state, we are basically allowing different types of player i.

3.2 Capturing Solution Concepts in Extensive-Form Games

We now consider solution concepts in extensive-form games. Recall that, in this case, we assume that players are using behavioral strategies.

3.2.1 Nash Equilibrium

Here we get essentially the same result as Theorem 3.1: the behavioral strategy profile $\vec{\sigma}$ is a Nash equilibrium in the extensive-form game Γ iff $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}})$ is compatible with $\mathbf{EQ}^\Gamma$. However, a number of new subtleties arise in extensive-form games. First, Nash equilibrium in an extensive-form game does not require that players make a best response off the equilibrium path. This is dealt with in our definition of compatibility since the fact that we consider only points (r, m) such that $\mu_i(\mathcal{K}_i(r, m)) > 0$ means that we are considering only points on the equilibrium path.

Another subtlety arises from the fact that, in determining whether σ_i is a best response to σ_{-i} at a point (r, m), player i is allowed to change to a completely different strategy σ'_i. But the definition of $\mathbf{EQ}^\Gamma$ in extensive-form games only considers changing to a different *action*. To show that this suffices, we appeal to a result known in the literature as the *one-deviation* property (Osborne & Rubinstein, 1994). The one-deviation property holds if, in order to check that a behavioral strategy σ is a best response to σ_{-i}, it suffices to check local changes to σ; that is, it suffices to check behavioral strategies that differ from σ by just modifying what σ does at one information set.

Let $\sigma[I/\mathsf{a}]$ be the behavioral strategy that is just like σ except that it assigns probability 1 to the action a at the information set I. Strategies of the form $\sigma[I/\mathsf{a}]$ are what we consider to show that the one-deviation property holds at information set I. But a best response at information set I allows changes not just at I, but at all information sets preceded by I. Our analysis considers only what are called games with perfect recall. Roughly speaking, in a game of perfect recall, all the players recall what moves they have made and what information sets they have passed through. This recollection is formalized by putting conditions on information sets. We omit the formal definition of perfect recall here (see Osborne & Rubinstein, 1994). In a finite extensive-form game Γ with perfect recall, for each player i, we can define a partial order $\preceq_i$ on player i's information sets such that $I \preceq_i I'$ if, for every history $h \in I$, there is a prefix h' of h in I'. Thus, $I \preceq_i I'$ if I' is preceded by I, or, equivalently, appears below I in the game tree. Given two information sets, we say that I precedes I', and write $I \preceq I'$, if $I = I'$ or I' comes after I in some history of the game (i.e., if some node in I' is preceded by a node in I in the game tree). In games of perfect recall, $\preceq$ is a partial order; there cannot be two distinct information sets I and I' such that $I \preceq I'$ and $I' \preceq I$. Given an information set I for player i, denote by $[\sigma_i', I, \sigma_i]$ the strategy for player i that agrees with σ_i on all information sets I' for player i such that $I \preceq I'$ and agrees with σ_i' on all other information sets.

We now recall the notion of best response for behavioral strategies given by Halpern (2013). A *belief system* (Kreps & Wilson, 1982) is a function μ that associates with each information set I a probability, denoted μ_I, on the histories in I. Given a behavioral strategy $\vec{\sigma}$ and a belief system μ in an extensive-form game Γ, let $\Pr_{\vec{\sigma}}$ denote the distribution on terminal histories induced by $\vec{\sigma}$ and define

$$\mathrm{EU}_i((\vec{\sigma},\mu) \mid I) = \sum_{h\in I}\sum_{z\in Z} \mu_I(h)\Pr_{\vec{\sigma}}(z \mid h)u_i(z).$$

Thus, the expected utility of $(\vec{\sigma},\mu)$ conditional on reaching I captures the expected payoff to player i if $\vec{\sigma}$ is played from information set I on, given that the relative likelihood of histories in I is determined by μ. Finally, if $\vec{\sigma}$ is a completely-mixed behavioral strategy profile, let $\mu^{\vec{\sigma}}$ be the belief system determined by $\vec{\sigma}$ in the obvious way:

$$\mu_I^{\vec{\sigma}}(h) = \Pr_{\vec{\sigma}}(h \mid I).$$

Definition 3.5 (Halpern, 2013) If $\varepsilon \geq 0$ and I is an information set for player i that is reached with positive probability by $\vec{\sigma}'$, then σ_i is an *ε-best response to $\vec{\sigma}'_{-i}$ for i conditional on having reached I using $\vec{\sigma}'$* if, for every strategy τ for player i, we have

$$\mathrm{EU}_i(((\sigma_i,\vec{\sigma}'_{-i}),\mu_I^{\vec{\sigma}'}) \mid I) \geq \mathrm{EU}_i(((\tau_i,\vec{\sigma}'_{-i}),\mu_I^{\vec{\sigma}'}) \mid I) - \varepsilon.$$

The strategy σ_i is an *ε-best response for i relative to $\vec{\sigma}'$* if σ_i is an ε-best response to $\vec{\sigma}'_{-i}$ for i conditional on having reached I using $\vec{\sigma}'$ for all information sets I for i that are reached with positive probability by $\vec{\sigma}'$. The strategy $\vec{\sigma}_i$ is a *best response for i relative to $\vec{\sigma}'$* (resp., *best response for i conditional on having reached I using $\vec{\sigma}'$*) if σ_i is a 0-best response for i relative to $\vec{\sigma}'$ (resp., 0-best response for i conditional on having reached I). ∎

Thus, σ_i is a best response for i relative to $\vec{\sigma}'$ if σ_i is a best response to σ'_{-i} at each information set I for player i that is reached with positive probability by $\vec{\sigma}'$, where we assume that $\vec{\sigma}'$ determines the probability of the histories in I and, in best responding, we allow player i to make arbitrary changes after I has been reached. determines the probability of reaching I.

The next result is basically the one-deviation property; it shows that a strategy that is optimal with respect to local changes is in fact a best response.

Theorem 3.6 *Let Γ be a game of perfect recall. The strategy σ_i is a best response response to σ_{-i} relative to $\vec{\sigma}$ iff* $\mathrm{EU}_i(\vec{\sigma}) \geq \mathrm{EU}_i(\sigma_i[I/\mathsf{a}], \vec{\sigma}_{-i})$ *for each information set I for player i that is reached with positive probability by $\vec{\sigma}$ and each action* a *that i can play at I.*

Proof This is essentially proved by Selten (1975), so we just briefly sketch the argument here. Clearly, if σ_i is a best response to σ_{-i} relative to $\vec{\sigma}$ and I is reached by $\vec{\sigma}$ with positive probability, then $\mathrm{EU}_i(((\vec{\sigma}, \mu_I^{\vec{\sigma}}) \mid I) \geq \mathrm{EU}_i((\sigma_i[I/\mathsf{a}], \vec{\sigma}_{-i}), \mu_I^{\vec{\sigma}}) \mid I)$ for all information sets I that are reached with positive probability according to $\mu^{\vec{\sigma}}$ at which player i moves, and all actions a it can take at I. For the converse, suppose that $\mathrm{EU}_i(((\vec{\sigma}, \mu_I^{\vec{\sigma}}) \mid I) \geq \mathrm{EU}_i((\sigma_i[I/\mathsf{a}], \vec{\sigma}_{-i}), \mu_I^{\vec{\sigma}}) \mid I)$ for all information sets I that are reached with positive probability according to $\mu^{\vec{\sigma}}$ at which player i moves, and all actions a it can take at I. By way of contradiction, suppose that σ_i is not a best response to σ_{-i} relative to $\vec{\sigma}$. Then there must be some information set I that is reached with positive probability by $\vec{\sigma}$ and strategy τ_i for player i such that $\mathrm{EU}_i(([\sigma_i, I, \tau_i], \vec{\sigma}_{-i}), \mu_I^{\vec{\sigma}}) \mid I) > \mathrm{EU}_i(((\vec{\sigma}, \mu_I^{\vec{\sigma}}) \mid I)$. We get an easy contradiction by considering a latest information set I for player i that is reached with positive probability by $\vec{\sigma}$ at which this inequality holds (so that at all information set $I' \neq I$ such that $I \preceq I'$, the inequality does not hold). (Since Γ is a game of perfect recall, the notion of "latest information set" is well defined.) ∎

Theorem 3.7 *The behavioral strategy profile $\vec{\sigma}$ is a Nash equilibrium of the extensive-form game Γ iff $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}})$ is compatible with* $\mathbf{EQ}^\Gamma$.

We omit this proof, since it is easier (and similar in spirit) to the the proofs presented later in this section.

3.2.2 Perfect Equilibrium

We start by considering (trembling-hand) perfect equilibrium (Selten, 1975). This is defined in both normal-form games and extensive-form games. For ease of exposition, we focus on the definition in extensive-form games, although essentially the same approach applies to normal-form games.

The idea is that $\vec{\sigma}$ is a perfect equilibrium if, not only is σ_i a best response to $\vec{\sigma}_{-i}$, but σ_i is a best response even if some players $j \neq i$ "tremble", and (with exceedingly small probability) play a strategy other than σ_j. To make this precise, define a *completely mixed (behavioral) strategy* for player i to be a strategy where, at each information set for player i, each action that can be played is played with positive probability. Observe that if $\vec{\sigma}$ is a completely mixed strategy in an extensive-form game Γ, then it will reach every information set I of each player in Γ with positive probability. For an extensive-form game Γ, the strategy profile $\vec{\sigma}$ is a *perfect equilibrium* in Γ iff there exists a sequence $\vec{\sigma}^n$ of completely mixed strategies such that $\vec{\sigma}^n \to \vec{\sigma}$ and, for all n and each information set I of player i, σ_i is a best response to $\vec{\sigma}^n_{-i}$ conditional on having reached I. Intuitively, $\vec{\sigma}^n_{-i}$ represents a "tremble"; $\vec{\sigma}^n_{-i}$, $n = 1, 2, 3, \ldots$ is a sequence of trembles converging to $\vec{\sigma}$. The strategy σ_i must be a best response to each tremble in this sequence. Thus, each strategy σ_i in a perfect equilibrium profile $\vec{\sigma}$ is not a best response to all possible trembles, but to the trembles along one particular path converging to $\vec{\sigma}$. (The definition of perfect equilibrium is essentially the same in normal-form games, except that there is no need to condition on the information set.)

Our result depends on a characterization of perfect equilibrium given by Halpern (2009, 2013) that uses nonstandard probabilities, which can assign infinitesimal probabilities to strategy profiles. A discrete *nonstandard probability distribution* on a set of runs is a discrete probability distribution that assigns nonstandard probabilities to runs so that the sum over all runs adds up to 1.

Halpern (2009) shows that by using nonstandard probability, we can capture Selten's intuition for trembling-hand equilibrium without needing to explicitly refer to sequences of strategy profiles, as is done in Selten's original definition. The idea is that the sequence converging to $\vec{\sigma}$ in Selten's original definition is replaced by a single completely mixed strategy profile that is infinitesimally close to $\vec{\sigma}$. To make this precise, we need a few definitions. It is well known that to every nonstandard real number ρ, there is a closest standard real number denoted $\mathsf{st}(\rho) \in I\!R$, and called "the standard part of ρ": the difference $|\rho - \mathsf{st}(\rho)|$ is an infinitesimal. Given a nonstandard probability measure ν, we can define the standard probability measure $\mathsf{st}(\nu)$ by taking $\mathsf{st}(\nu)(w) = \mathsf{st}(\nu(w))$ for all states $\omega \in \Omega$. Two possibly nonstandard distributions ν and ν' *differ infinitesimally* if $\mathsf{st}(\nu) = \mathsf{st}(\nu')$.

While, as Selten shows, a perfect equilibrium always exists in normal-form games, it does not necessarily exist in an arbitrary extensive-form game. However, it does exist in extensive-form games of perfect recall. We can now state Halpern's characterization of trembling-hand equilibrium. We say that two behavioral strategies σ_i and σ'_i for player i *differ infinitesimally* if the distributions $\sigma_i(I)$ and $\sigma'_i(I)$ differ infinitesimally for each information set I for player I. Two strategy profiles $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ and $\vec{\sigma}' = (\sigma'_1, \ldots, \sigma'_n)$ differ infinitesimally if σ_i differs infinitesimally from σ'_i at I, for every $i = 1, \ldots, n$.

Theorem 3.8 *(Halpern, 2009, 2013) The behavioral strategy profile $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ is a perfect equilibrium in an extensive-form game Γ of perfect recall iff there exists a nonstandard completely mixed behavioral strategy profile $\vec{\sigma}'$ that differs infinitesimally from $\vec{\sigma}$ such that σ_i is a best response to σ'_{-i} relative to $\vec{\sigma}'$ for each player i.*

When dealing with standard probabilities, in the definition of a probabilistic system $\mathcal{PS} = (\mathcal{R}, \vec{\mu})$ being compatible with knowledge-based program profile $\overrightarrow{\mathsf{Pg}}$, we required that the action played by Pg_i at (r, m) be the same as that played in the system $\mathcal{PS}$ at (r, m) only for runs r such that $\mu_i(r \mid \mathcal{R}[\mathcal{K}_i(r, m)]) > 0$. We now want to restrict not just to runs that have positive probability, but to runs that have "nontrivial" positive probability. The obvious choice would be to require that $\mathsf{st}(\mu_i(r \mid \mathcal{R}[\mathcal{K}_i(r, m)])) > 0$. In settings where all players are following a behavioral strategy, this requirement would would restrict to runs r where, at all times $m' > m$, if player j moves at the information set associated with (r, m'), then the move made by j is given positive standard probability by the behavioral strategy.

We would like this to be the case as well for the move made at (r,m). Since we encode i's intended move at (r,m) in i's local state $r_i(m)$ (recall that the local state models i's information *after* i has made the coin toss), conditional on $\mathcal{K}_i(r,m)$, player i's intended move has probability 1, even if it has infinitesimal probability according to his strategy.

For simplicity, we describe the requirement that we want only in systems of the form $\mathcal{R}^\Gamma$. In such systems, if I is an information set in Γ, let $\mathcal{R}[I]$ consist of all runs that go through information set I. Note that if i moves at I, then $\mathcal{R}[I]$ is the disjoint union of sets of the form $\mathcal{R}[\mathcal{K}_i(r,m)]$, for runs r where i's local state has the form (I,a). We now define what it means for $\mathcal{PS} = (\mathcal{R}^\Gamma, \vec{\mu})$, where $\vec{\mu}$ is a profile of nonstandard probability measures just as we did before, except that in clause (b), we replace the requirement that $\mu_i(r \mid \mathcal{R}[\mathcal{K}_i(r,m)]) > 0$ by $\mathsf{st}(\mu_i(r \mid \mathcal{R}[\mathcal{K}_i(r,m)])) > 0$, and if i moves at the information set I associated with (r,m), then we further strengthen this requirement to $\mathsf{st}(\mu_i(r \mid \mathcal{R}[I])) > 0$. Since $\mu(r \mid \mathcal{R}[I])$ is a convex combination of terms of the form $\mu(r \mid \mathcal{R}[\mathcal{K}_i(r',m)])$, where the sum is taken over the points (r',m) such that the node associated with (r',m) is in I and $\mu(\mathcal{R}[\mathcal{K}_i(r',m)]) > 0$, and $\mu(r \mid \mathcal{R}[\mathcal{K}_i(r',m)]) = 0$ if $(r,m) \notin \mathcal{K}_i(r',m)$, it easily follows that if $\mathsf{st}(\mu_i(r \mid \mathcal{R}[I])) > 0$ then $\mathsf{st}(\mu_i(r \mid \mathcal{R}[I])) > 0$.

For standard probability measures μ_i, it is easy to see that $\mu_i(\mathcal{R}[\mathcal{K}_i(r,m)]) > 0$ and that $\mu_i(r \mid \mathcal{R}[\mathcal{K}_i(r,m)]) > 0$ iff both $\mu_i(\mathcal{R}[\mathcal{K}_i(r,m)]) > 0$ and $\mu_i(r \mid \mathcal{R}[I])$, so this really is a generalization of the standard definitions.[6]

Roughly speaking, this says that compatibility is required only at points on which i places "significant" probability on the moves made by the strategy used by i. (This is made more precise in the proof of Theorem 3.9.)

Note that since only one player i moves at a node w in an extensive-form game (since we do not allow moves by nature), in the action profile $\vec{\mathsf{a}}$ such that applying $\vec{\mathsf{a}}$ to $r(m)$ results in $r(m+1)$, we have $\mathsf{a}_j = \mathsf{skip}$ for $j \neq i$. This is because $(r,m) \notin [\![\kappa]\!]_{\mathcal{PS}}$ for a test κ in $\overrightarrow{\mathbf{EQ}}_j^\Gamma$; these tests κ are true only at points where j moves.

Theorem 3.9 *The strategy profile $\vec{\sigma}$ is a perfect equilibrium of the extensive-form game Γ of perfect recall iff there exists a (possibly nonstandard) completely mixed behavioral strategy profile $\vec{\sigma}'$ such that σ_i differs infinitesimally from σ_i' and $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}'})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$.*

Proof Suppose that $\vec{\sigma} = (\sigma_1, \ldots, \sigma_n)$ is a perfect equilibrium of Γ. By Theorem 3.8, there exists a nonstandard completely mixed strategy profile $\vec{\sigma}'$ that differs infinitesimally from $\vec{\sigma}$ such that σ_i is a best response to σ_{-i}' relative to $\vec{\sigma}'$, for each player $i = 1, \ldots, n$. We show that $\mathcal{PS} = (\mathcal{R}^\Gamma, \mu_{\vec{\sigma}'})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$. Suppose that the information set $I = \mathcal{K}_i(r,m)$ associated with (r,m) is one where i moves, and that $\mathsf{st}(\mu_{\vec{\sigma}'}(r \mid \mathcal{R}[I])) > 0$. (Note that since $\vec{\sigma}'$ is completely mixed, it is guaranteed that $\mu_{\vec{\sigma}'}(\mathcal{R}[I]) > 0$.) Let a be

6. The reader may wonder why we did not just take $r_i(m)$ to be i's information set, rather than having it include the action that i plans to do. While the former choice would have simplified the discussion above, for the kb program $\mathbf{EQ}^\Gamma$ to be meaningful, i has to know what action he is about to do. Note that what we are doing here is considering i's information at two stages: before he has tossed the coin to determine his next action, and after he has tossed it. We are conditioning on his information before he tossed the coin, even though i's local state models only the situation after he has tossed the coin. Using this intuition we can extend the definition of compatibility beyond the scope of systems of the form $\mathcal{R}^\Gamma$, as long as the system is generated by players running randomized programs (like behavioral strategies), although making this precise is beyond the scope of this paper.

the action encoded in $r_i(m)$; that is, a is the action that i plans to play at (r,m). Since $\mathsf{st}(\mu_{\vec{\sigma}}(r \mid \mathcal{R}[I])) > 0$, it must be the case that the action a is given positive standard probability by the (completely mixed) distribution $\sigma'_i(I)$. Since $\vec{\sigma}$ is a perfect equilibrium, we have by Theorem 3.8 that σ_i is a best response to σ_{-i} relative to $\vec{\sigma}$. Let $\tau_i = [\sigma'_i, I, \sigma_i]$. By definition, τ_i and σ_i agree on their actions at I. Since σ_i gives a positive standard probability at I, so does τ_i. By Theorem 3.6, τ_i must be at least as good a response as $\tau_i[I/\mathsf{a}']$ to $\vec{\sigma}'_{-i}$, for any action a' that i can play at I. Since τ_i gives a positive standard probability, $\tau_i[I/\mathsf{a}]$ must be at least as good a response as $\tau_i[I/\mathsf{a}']$ to $\vec{\sigma}'_{-i}$ for any action a', conditional on reaching I using $\vec{\sigma}'$. It is easy to see that the expected utility of $\tau_i[I/\mathsf{a}]$ (resp., $\tau_i[I/\mathsf{a}']$) conditional on reaching I is just the value of $\mathrm{EU}_i(\mathsf{a})$ (resp., $\mathrm{EU}_i(\mathsf{a}')$) at the point (r,m). Thus, $(r,m) \in [\![\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')]\!]_{\mathcal{PS}}$, so $\mathcal{PS}$ is compatible with $\mathbf{EQ}^\Gamma$.

For the converse, suppose that $\vec{\sigma}'$ is a completely mixed behavioral strategy profile such that σ_i differs infinitesimally from σ'_i and $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}'})$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$. Let $\vec{\sigma} = \mathsf{st}(\vec{\sigma}')$. Again, let $\tau_i = [\sigma'_i, I, \sigma_i]$. By Theorems 3.8 and 3.6, it suffices to show that for each information set I for player i and each action a' that i can play at I, the strategy τ_i is at least as good a response as $\tau_i[I/\mathsf{a}']$ to $\vec{\sigma}'_{-i}$ conditional on having reached I using $\vec{\sigma}'$. To do this, it suffices to show that for each action a in the support of $\tau_i(I) = \sigma_i(I)$, strategy $\tau_i[I/\mathsf{a}]$ is at least as good a response as $\tau_i[I/\mathsf{a}']$ to $\vec{\sigma}'_{-i}$ conditional on having reached I using $\vec{\sigma}'$. So fix an information set I where player i moves and suppose that a is in the support of $\sigma_i(I)$. Let r be a history that reaches $I = \mathcal{K}_i(r,m)$ in which i plays a at I and all the players play an action that is given positive standard probability by $\vec{\sigma}$ (and hence also by $\vec{\sigma}'$) at all points preceding (r,m). Thus, $\mathsf{st}(\mu_{\vec{\sigma}}(r \mid \mathcal{R}[I])) > 0$. Since $\mathcal{PS}$ is compatible with $\mathbf{EQ}^\Gamma$, it must be the case that $(r,m) \in [\![\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')]\!]_{\mathcal{PS}}$. As in the first half of the proof, it now follows that $\tau_i[I/\mathsf{a}]$ is at least as good a response as $\tau_i[I/\mathsf{a}']$ to $\vec{\sigma}'_{-i}$ conditional on having reached I using $\vec{\sigma}'$. Thus, $\vec{\sigma}$ is a perfect equilibrium. ∎

3.2.3 Sequential Equilibrium

We next characterize sequential equilibrium in terms of $\mathbf{EQ}^\Gamma$. Recall that a *sequential equilibrium* (Kreps & Wilson, 1982) is an *assessment*, a pair $(\vec{\sigma}, \mu)$, where $\vec{\sigma}$ is a behavioral strategy profile and μ is a *belief system*, that is, a function that determines for every information set I a probability μ_I over the histories in I. Intuitively, if I is an information set for player i, then μ_I is i's subjective assessment of the relative likelihood of the histories in I. Roughly speaking, an assessment is a sequential equilibrium if both (a) at every information set where a player moves he chooses a best response given the beliefs he has about the histories in that information set and the strategies of other players, and (b) his beliefs are consistent with the strategy profile being played. We omit the formal definition here, and instead use a characterization of sequential equilibrium due to Halpern (2009).

Theorem 3.10 *(Halpern, 2009, 2013) An assessment $(\vec{\sigma}, \mu)$ is a sequential equilibrium in an extensive-form game Γ with perfect recall iff there exist an infinitesimal ε and a nonstandard completely mixed strategy profile $\vec{\sigma}'$ that differs infinitesimally from $\vec{\sigma}$ such that σ_i is an ε-best response to σ_{-i} relative to $\vec{\sigma}$, for each player i.*

The only difference between sequential equilibrium and perfect equilibrium in this characterization is that with perfect equilibrium σ_i must be a best response to σ_{-i} relative to

$\vec{\sigma}$ while with sequential equilibrium, it need only be an ε-best response for some infinitesimal ε. To capture this difference, when dealing with sequential equilibrium, we reinterpret the formula $\mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}')$ so as to ignore infinitesimal differences. Thus, the formula is true unless $\mathsf{st}(\mathrm{EU}_i(\mathsf{a}') - \mathrm{EU}_i(\mathsf{a})) > 0$ (or, equivalently, it is true if the standard part of i's expected utility using a is greater than or equal to the standard part of i's expected utility using a'). $\mathcal{PS}$ is *st-compatible with* $\overrightarrow{\mathbf{EQ}}^\Gamma$ (standing for compatible with respect to standard values) if $\mathcal{PS}$ is compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$ under this reinterpretation of $\overrightarrow{\mathbf{EQ}}^\Gamma$. Clearly, when all the probability distributions in $\mathcal{PS}$ are standard, the notions compatibility and st-compatibility of $\overrightarrow{\mathbf{EQ}}^\Gamma$ with $\mathcal{PS}$ coincide.

Theorem 3.11 *The assessment $(\vec{\sigma}, \mu)$ is a sequential equilibrium of the finite extensive-form game Γ of perfect recall iff there exists (possibly nonstandard) completely mixed strategy profile $\vec{\sigma}'$ such that σ_i differs infinitesimally from σ_i' and $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}'})$ is st-compatible with $\overrightarrow{\mathbf{EQ}}^\Gamma$.*

Proof The proof is almost identical to that of Theorem 3.9, replacing "best response" by "ε-best response", and "compatible" with "st-compatible". We leave details to the reader. ∎

3.2.4 Subgame-Perfect Equilibrium

Subgame-perfect equilibrium, defined by Selten (1965), is usually considered in games of perfect information, where all information sets are singletons. In games of perfect information, subgame perfection, sequential equilibrium, and trembling-hand perfect equilibrium all agree, so we do not need to provide a separate characterization. However, subgame perfection is actually defined for arbitrary games of perfect recall.

Given a game Γ of perfect information and a node w in Γ, the subtree of Γ rooted at w determines a *subgame* that we denote Γ^w if, for every information set I in Γ that includes a node w' at or below w in Γ, all the nodes in I are below w in Γ. For example, the subtree of the game Γ_e in Figure 2 does *not* determine a subgame, since the information set $\{w_2, w_3\}$ includes a node that is at or below w_2 (namely, w_2 itself), but w_3 is not below w_2. The strategy profile $\vec{\sigma}$ is a *subgame-perfect equilibrium* if, for every subgame Γ^w of Γ, $\vec{\sigma}$ restricted to the nodes in Γ^w is a Nash equilibrium in Γ^w. Note that subgame perfection places no requirements on the action played at nodes that do not determine subgames, beyond the fact that the action must be part of a Nash equilibrium at nodes higher in the tree that do determine subgames.

We can use a program much like $\mathbf{EQ}^\Gamma$ to characterize subgame-perfect equilibrium in arbitrary games of perfect recall. We need to make two changes to $\mathbf{EQ}^\Gamma$. First, we need to say that a best response is required only at points where *subgame* holds, where $(r, m) \in [\![subgame]\!]_{\mathcal{PS}}$ if the node w associated with (r, m) determines a subgame. (Note that $[\![subgame]\!]_{\mathcal{PS}}$ consists of all points in a game of perfect information.) Second, we need to say that there are no constraints at points where *subgame* does not hold. So, for each action $\mathsf{a} \in \mathcal{A}_i(\Gamma)$, we now have two lines of the form

if $K_i(\mathit{intend}_i(\mathsf{a}) \ \wedge \ \mathit{subgame} \wedge \bigwedge_{\mathsf{a}' \in \mathcal{S}_i(\mathcal{A})} \mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}'))$ **then** play a

if $K_i(\mathit{intend}_i(\mathsf{a}) \ \wedge \ \bigwedge_{\mathsf{a}' \in \mathcal{S}_i(\mathcal{A})} \mathrm{EU}_i(\mathsf{a}) \geq \mathrm{EU}_i(\mathsf{a}'))$ **then** play a.

Call the resulting program $\mathbf{SUBEQ}_i^\Gamma$.

These changes make it clear that subgame perfect is a somewhat awkward notion in games where players do not have perfect information. In any case, with this change, an analogue of Theorem 3.9 holds for subgame perfect equilibria.

Theorem 3.12 *The strategy profile $\vec{\sigma}$ is a subgame-perfect equilibrium of the extensive-form game Γ of perfect recall iff there exists a (possibly nonstandard) completely mixed behavioral strategy profile $\vec{\sigma}'$ such that σ_i differs infinitesimally from σ_i' and $(\mathcal{R}^\Gamma, \mu_{\vec{\sigma}'})$ is subgame compatible with $\mathbf{SUBEQ}^\Gamma$.*

We omit the proof here, which is similar in spirit to that of Theorem 3.9.

4. Discussion and Conclusions

The essential intuition in many solution concepts is that (it is common knowledge that) players are making a best response to their beliefs. We have shown that this "procedural" intuition can be captured by a single knowledge-based program, denoted $\overrightarrow{\mathbf{EQ}}^\Gamma$. The differences between these solutions concepts lies in differences in assumptions about players' beliefs and in what counts as a best response.

- In Nash equilibrium, players believe that a mixed strategy profile is being played (and have common belief about which one it is).
- In correlated equilibrium, the players believe that a correlated strategy profile is being played (and have common belief about which one it is).
- In perfect equilibrium, they can be viewed as believing that a nonstandard completely mixed strategy profile is being played (and having common belief about which one it is), and caring only about what happens in situations with positive standard probability.
- In sequential equilibrium, they can similarly be viewed as believing that a nonstandard completely mixed strategy profile is being played (and have common belief about which one it is), and caring only about what happens at states with positive standard probability and about best responses with respect to standard differences (an ε better response for some infinitesimal ε is not viewed as being better).
- In rationalizability, different players may hold different beliefs about the strategy profile being played.

While the unification given by kb programs arguably does give insight, there is clearly a significant amount of overhead in the kb program framework. It is certainly reasonable to ask whether it is worth dealing with the overhead just to get such a unification, given that the intuitions are certainly well understood in the game-theory literature.

If the sole advantage of using kb programs was to prove the theorems in this paper, then perhaps the answer is "no", but we believe that the kb program framework offers much more to game theory than just this unification. For one thing, kb programs can capture the intuition of best response more generally. We give a few examples here:

- Dealing with moves by nature: We have assumed for simplicity that there were no moves by nature in the extensive-form games being analyzed. To deal with moves by nature, we first expand the notion of a global state so that it includes the local state of "nature", not just the local states of the players. We can think of nature's local state as consisting of the current node in the game tree. In addition, we think of nature, just like the players, as following a behavioral strategy (where nature's move depends on its local state). With these minor changes, all our results still go through, with no change. In particular, all of the theorems in the paper continue to hold even for games where nature moves, with no change in the proof.

- Bayesian games: In a Bayesian game, players have *types*. We can think of a type as a description of a player's private information. There is assumed to be a commonly-known distribution over type profiles. A strategy can be viewed as a function from types to actions. A player's utility depends on both the action profile and the type profile. The standard solution concept considered in Bayesian games is a Bayes-Nash equilibrium. In a Bayes-Nash equilibrium, no player wants to switch to a different strategy, since doing so results in a lower expected utility (see Osborne & Rubinstein, 1994, for details). We can again capture a Bayes-Nash equilibrium in our framework. Now a player's local state would include the player's type, and a run can be characterized by the strategy profile and type profile. This means that the set of runs in $\mathcal{R}^\Gamma$ is larger. With these changes, an analogue of Theorem 3.1 holds for Bayes-Nash equilibrium.

- Beyond expected utility maximization: All of the solution concepts that we have considered in the paper are based on maximizing expected utility. But we can also consider solution concepts based on other decision criteria. For example, Boutilier and Hyafil (2004) consider *minimax-regret* equilibria, where each player uses a strategy that is a best-response in a minimax-regret sense to the choices of the other players. Similarly, we can use *maximin equilibria* (Aghassi & Bertsimas, 2006). As pointed out by Chu and Halpern (2003), all these decision rules can be viewed as instances of a generalized notion of expected utility, where (a) uncertainty is represented by a *plausibility measure*, a generalization of a probability measure, (b) utilities are elements of an arbitrary partially ordered space, and (c) plausibilities and utilities are combined using $\oplus$ and $\otimes$, generalizations of $+$ and $\times$. Just by interpreting "$\mathrm{EU}_i = u$" appropriately, we can capture these more exotic solution concepts as well. Moreover, applying the same ideas and essentially the same proof we can capture solution concepts in games in which the game itself is not common knowledge, or where players are not aware of all available moves, as discussed by Halpern and Rêgo (2013).

All the results mentioned up to now are straightforward, and much in the spirit of the results we have already shown. A more interesting situation arises when we consider games of imperfect recall. Part of the overhead in the framework is the need to specify exactly what the players' local states are, that is, what they know. In the context of games of perfect recall, this is perhaps not that important, but when we move to games of imperfect recall, this becomes highly significant. Consider the single-player game depicted in Figure 3, first introduced by Piccione and Rubinstein (1997).

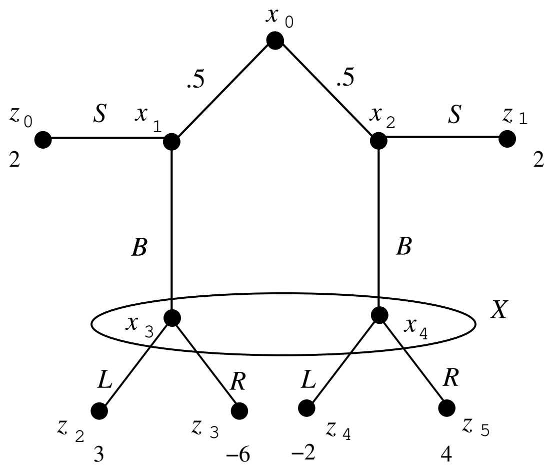

Figure 3: A game of imperfect recall.

It is not hard to show that the strategy that maximizes expected utility in this example chooses move S at node x_1, move B at node x_2, and move R at the information set X consisting of x_3 and x_4. Call this strategy σ. Let σ' be the strategy of choosing move B at x_1, move S at x_2, and move L at X. Piccione and Rubinstein argue that if node x_1 is reached, the player should reconsider, and decide to switch from σ to σ'. While this indeed leads to a better payoff, the resulting strategy (i.e., starting with σ and switching to σ' at x_2, if x_2 is reached) is not a legal strategy in the original game; the player moves left at x_3 and right at x_4, although the two nodes are in the same information set.

The question of what the player's local state is now becomes critical. We did not include the player's behavior strategy in his local state when we modeled extensive-form, but we could have done so with no change. Suppose that we do so in this game of imperfect recall. First note that if we include the player's strategy in his local state, then in the system $\mathcal{R}^\Gamma$ corresponding to the game, if h is the history ending in z_5, at the point $(r^h, 2)$, the player *knows* that he is at x_4, despite the information set. This is an instance of a more general phenomenon: in a game of imperfect recall, the strategy that a player is using gives him information about which node in an information set he is at. This cannot happen in a game of perfect recall.

Suppose for ease of exposition that we include a player's strategy in his local state. What happens if a player switches strategy. How does his local state change then? If the local state includes the new strategy (whether or not it includes the original strategy), then in the set of runs that arises if the player sticks to σ at x_2, but switches from σ to σ' at x_1, when player reaches x_3, he knows that he is at x_3, and if he reaches x_4, he knows that he is at x_4. The "information set" is not correctly representing the player's knowledge at all![7]

The key point here is that the runs-and-systems framework forces a modeler to consider questions like whether a player is able to keep track of his changes of strategy; moreover, the answers must be reflected in the choice of local state. There has been some recent work on defining notions like sequential equilibrium in games of imperfect recall (see, e.g., Halpern

7. This point was already made by Halpern (1997).

& Pass, 2011b; Kline, 2005; Marple & Shoham, 2012). We believe that the program $\mathbf{EQ}^{\Gamma}$ and, more generally, the use of the runs-and-systems framework can provide some insight into this problem.

Interesting new issues arise when we add computation to the picture. Equilibrium notions that take computation into account have been considered by Halpern and Pass (2011a). It seems that the notion of computational Nash equilibrium defined by Halpern and Pass, where players choose a Turing machine to play for them, can be captured using $\mathbf{EQ}^{\Gamma}$ as well. But that is only because, roughly speaking, there is no charge for the computation of which Turing machine is a best response to the Turing machines chosen by the other players. If we were to impose such a cost, then we might need more "computational" or "algorithmic" notions of knowledge (such notions are discussed by Fagin et al., 1995, ch. 10).

We conclude with two other directions for further research. First, although we have focused on using kb programs to characterize solution concepts here, the idea that an agent's actions depend on her knowledge and beliefs seems like a very natural way to characterize strategies in games, and "meta-strategies" for classes of games. Sayings such as "look before you leap" and "trust, but verify" are really shorthand for knowledge-based programs. We believe that useful insights into how agents play games can be gained by thinking at the knowledge level in this way. Indeed, it is not only preconditions on actions that depend on knowledge and belief; an agent's utility can also depend on her beliefs. This is the key insight in *psychological games* (Geanakoplos, Pearce, & Stacchetti, 1989). It would be interesting to extend knowledge-based programs to knowledge-based utilities.

Finally, although we have talked about kb programs as "procedural", in fact, there is no procedure given for the calculation of the relevant knowledge, which really amounts to a best-response computation. In a non-probabilistic setting, there are conditions under which a kb program can be implemented by a unique standard program (i.e., one without tests for knowledge) as shown by Fagin et al. (1995, Section 7.2). Such results do not carry over to probabilistic systems (since they give no indication of how to compute the relevant probabilities). Nevertheless, given beliefs, kb programs can be viewed as defining how an agent should act. When computing an equilibrium, the beliefs are typically determined by the strategy profile. That is, we do not start with beliefs and then determine how to act. Rather, in most of the solutions concepts we have considered here, we have a fixed point: the beliefs determine the strategies (each player's strategy is a best response to her beliefs), and the strategies determine the beliefs. However, we believe that, in other applications of kb programs, it may well be possible to view a kb program as providing a procedural specification. We leave this topic to further research.

Acknowledgments

We thank the reviewers for their perceptive comments, which led to many improvements in the paper. Some material in this paper appeared in preliminary form in (Halpern & Moses, 2007). Joe Halpern's work on this paper was supported in part by NSF grants IIS-0534064, IIS-0812045, IIS-0911036, and CCF-1214844, by AFOSR grants FA9550-08-1-0438 and FA9550-09-1-0266, by the DoD Multidisciplinary University Research Initiative (MURI) program administered by AFOSR under grant N00014-01-1-0795, and by ARO grant W911NF-09-1-0281. Yoram Moses is the Israel Pollak academic chair at the Technion; his work was supported in part by the Israel Science Foundation under grant ISF 1520/11.

References

Aghassi, M., & Bertsimas, D. (2006). Robust game theory. *Mathematical Programming, Series B*, *107*(1–2), 231–273.

Aumann, R. J. (1974). Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, *1*, 67–96.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, *55*, 1–18.

Aumann, R. J., & Brandenburger, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica*, *63*(5), 1161–1180.

Benthem, J. van (2007). Rational dynamics and epistemic logic in games. *International Game Theory Review*, *9*(1), 13–45. (Reprinted with corrections in *International Game Theory Review*, **9**:2, 377-409.).

Benthem, J. van (2010). *Modal Logic for Open Minds*. Center for Study of Logic and Information-Lecture Notes.

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica*, *52*(4), 1007–1028.

Bonanno, G. (2002). Modal logic and game theory: two alternative approaches. *Risk Decision and Policy*, *7*, 309–324.

Brafman, R. I., Latombe, J.-C., Moses, Y., & Shoham, Y. (1997). Applications of a logic of knowledge to motion planning under uncertainty. *Journal of the ACM*, *44*(5), 633–668.

Brandenburger, A., & Dekel, E. (1987). Rationalizability and correlated equilibria. *Econometrica*, *55*, 1391–1402.

Bruin, B. de (2010). *Explaining Games: The Epistemic Programme in Game Theory*, Vol. 346. Synthese Library.

Chu, F., & Halpern, J. Y. (2003). Great expectations. Part I: On the customizability of generalized expected utility. In *Proc. Eighteenth International Joint Conference on Artificial Intelligence (IJCAI '03)*, pp. 291–296.

Dwork, C., & Moses, Y. (1990). Knowledge and common knowledge in a Byzantine environment: crash failures. *Information and Computation*, *88*(2), 156–186.

Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning About Knowledge*. MIT Press, Cambridge, Mass. A slightly revised paperback version was published in 2003.

Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1997). Knowledge-based programs. *Distributed Computing*, *10*(4), 199–225.

Geanakoplos, J., Pearce, D., & Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, *1*(1), 60–80.

Hadzilacos, V. (1987). A knowledge-theoretic analysis of atomic commitment protocols. In *Proc. 6th ACM Symposium on Principles of Database Systems*, pp. 129–134.

Halpern, J. Y. (1997). On ambiguities in the interpretation of game trees. *Games and Economic Behavior*, *20*, 66–96.

Halpern, J. Y. (2009). A nonstandard characterization of sequential equilibrium, perfect equilibrium, and proper equilibrium. *International Journal of Game Theory*, *38*(1), 37–50.

Halpern, J. Y. (2013). A nonstandard characterization of sequential equilibrium, perfect equilibrium, and proper equilibrium: Erratum. Unpublished manuscript.

Halpern, J. Y., & Moses, Y. (2007). Characterizing solution concepts in games using knowledge-based programs. In *Proc. Twentieth International Joint Conference on Artificial Intelligence (IJCAI '07)*, pp. 1300–1307.

Halpern, J. Y., & Moses, Y. (2010). Characterizing solution concepts in games using common knowledge of rationality. Unpublished manuscript.

Halpern, J. Y., Moses, Y., & Waarts, O. (2001). A characterization of eventual Byzantine agreement. *SIAM Journal on Computing*, *31*(3), 838–865.

Halpern, J. Y., & Pass, R. (2011a). Algorithmic rationality: Game theory with costly computation.. Available at www.cs.cornell.edu/home/halpern/papers/algrationality.pdf; to appear, *Journal of Economic Theory.* A preliminary version with the title "Game theory with costly computation" appears in *Proc. First Symposium on Innovations in Computer Science*, 2010.

Halpern, J. Y., & Pass, R. (2011b). Sequential equilibrium in games of imperfect recall. Unpublished manuscript; available at www.cs.cornell.edu/home/halpern/papers/imperfect.pdf.

Halpern, J. Y., & Rêgo, L. C. (2013). Extensive games with possibly unaware players. *Mathematical Social Sciences.* To appear.

Halpern, J. Y., & Zuck, L. D. (1992). A little knowledge goes a long way: knowledge-based derivations and correctness proofs for a family of protocols. *Journal of the ACM*, *39*(3), 449–478.

Harrenstein, P., Hoek, W. van der, Meyer, J.-J. C., & Witteveen, C. (2002). On modal logic interpretations of games. In *ECAI*, pp. 28–32.

Hyafil, N., & Boutilier, C. (2004). Regret minimizing equilibria and mechanisms for games with strict type uncertainty. In *Proc. Twentieth Conference on Uncertainty in Artificial Intelligence (UAI 2004)*, pp. 268–277.

Kline, J. J. (2005). Imperfect recall and the relationships between solution concepts in extensive games. *Economic Theory*, *25*, 703–710.

Kreps, D. M., & Wilson, R. B. (1982). Sequential equilibria. *Econometrica*, *50*, 863–894.

Lang, J., & Zanuttini, B. (2012). Knowledge-based programs as plans—the complexity of plan verification. In *Proceedings of the 20th European Conference on AI (ECAI 2012)*, pp. 504–504.

Lang, J., & Zanuttini, B. (2013). Knowledge-based programs as plans: succinctness and the complexity of plan existence. In *Theoretical Aspects of Rationality and Knowledge: Proc. Fourteenth Conference (TARK 2013)*, pp. 138–147.

Lorini, E., & Schwarzentruber, F. (2010). A modal logic of epistemic games. *Games*, *1*(4), 478–526.

Marple, A., & Shoham, Y. (2012). Equilibria in finite games with imperfect recall. Unpublished manuscript.

Mazer, M. S. (1990). A link between knowledge and communication in faulty distributed systems. In *Theoretical Aspects of Reasoning about Knowledge: Proc. Third Conference*, pp. 289–304.

Mazer, M. S., & Lochovsky, F. H. (1990). Analyzing distributed commitment by reasoning about knowledge. Tech. rep. CRL 90/10, DEC-CRL.

Moses, Y., & Kislev, O. (1993). Knowledge-oriented programming. In *Proc. 12th ACM Symposium on Principles of Distributed Computing*, pp. 261–270.

Moses, Y., & Tuttle, M. R. (1988). Programming simultaneous actions using common knowledge. *Algorithmica*, *3*, 121–169.

Neiger, G., & Bazzi, R. (1992). Using knowledge to optimally achieve coordination in distributed systems. In *Theoretical Aspects of Reasoning about Knowledge: Proc. Fourth Conference*, pp. 43–59.

Neiger, G., & Toueg, S. (1993). Simulating real-time clocks and common knowledge in distributed systems. *Journal of the ACM*, *40*(2), 334–367.

Osborne, M. J., & Rubinstein, A. (1994). *A Course in Game Theory.* MIT Press, Cambridge, Mass.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, *52*(4), 1029–1050.

Piccione, M., & Rubinstein, A. (1997). On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior*, *20*(1), 3–24.

Reiter, R. (2001). On knowledge-based programming with sensing in the situation calculus. *ACM Transactions on Computational Logic*, *2*(4), 433–457.

Selten, R. (1965). Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für Gesamte Staatswissenschaft*, *121*, 301–324 and 667–689.

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, *4*, 25–55.