Deliberating between Backward and Forward Induction Reasoning: First Steps

Aleks Knoks
University of Maryland
aknoks@umd.edu

Eric Pacuit
University of Maryland
epacuit@umd.edu

ABSTRACT

Backward and forward induction can be viewed as two styles of reasoning in dynamic games. Since each prescribes taking a different attitude towards the past moves of the other player(s), the strategies they identify as rational are sometimes incompatible. Our goal is to study players who are able to deliberate between backward and forward induction, as well as conditions under which one is superior to the other. This extended abstract is our first step towards this goal. We present an extension of Stalnaker's game models [34, 35], in which the players can make "trembling hand" mistakes. This means that when a player observes an unexpected move, she has to figure out whether it is a result of a deliberate choice or a mistake, thereby committing herself to one of the two styles of reasoning.

1. INTRODUCTION AND MOTIVATION

We begin with a motivating example. Consider the game G_1 depicted in Figure 1. There are two players: Ann (A) and Bob (B). Ann moves first (node h_0) and can either choose to go out (O), immediately ending the game, or stay in the game (I). If she chooses to stay in, node h_1 is reached. At h_1, Ann and Bob move simultaneously (Ann's available actions are u and v while Bob's are a, b and c). The structure of this game is similar to the extensively studied *Battle of the Sexes with an Outside Option* (see, for instance, [7, 14, 37]).

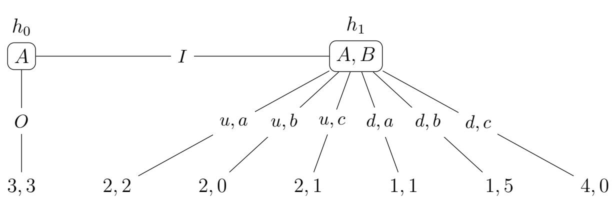


Figure 1: The game G_1

Suppose that Bob initially believes that Ann is going to choose O. To his amazement, however, Ann stays in the game. Now Bob has to figure out why Ann decided to play I, what she will choose at node h_1, and, most importantly, what is his own best response. There are two plausible "lines of reasoning" for Bob. The first goes as follows: Ann chose I at h_0 because she is hoping for a payoff greater than 3. Thus, Ann must be hoping that the game will terminate in the rightmost node. Hence, the rational choice for Bob is b, which—assuming his conjecture about Ann is correct—would result in a payoff of 5. Alternatively, Bob can avoid speculating about the reasons behind Ann's move and focus on trying to figure out what is the rational thing for her to do at h_1. Although Ann might hope initially that the game will end in the rightmost node, she must realize that Bob will never choose c since it is strictly dominated by a. So, guaranteeing a payoff of 2, u is Ann's rational choice at h_1. Clearly, if Bob convinces himself that Ann is choosing u, playing a is his best response. The two lines of reasoning, thus, lead to different recommendations for Bob. The first is an example of the so-called "forward induction reasoning" requiring that the players think critically about the observed past choices of their opponent(s) and find plausible explanations for these choices [7, 24, 28, 35, 37]. The second can be called "backward induction reasoning" requiring the players to only reason about their opponents' future behavior and not about their past moves [2, 12, 28, 30, 35].

There are many characterizations of both forward and backward induction reasoning in the game theory literature (cf. [7, 28, 35]). These formal renderings match the informal explanation given above, recommending that Bob plays a and b, respectively. However, the formal models do not solve what we take to be Bob's real challenge, namely, deciding which of these two lines of reasoning is more plausible in the present case.[1] Notice that a wrong choice leads to an unwelcome consequence. Suppose that Bob interprets Ann's choice of I as an attempt to get a higher payoff, but it turns out that she did it for some other reason—e.g. she was careless—and that, at h_1, she decides to play u. In this case, Bob ends up with 0. Now, suppose that Bob disregards Ann's previous move, as backward induction suggests he should, but it turns out that Ann is hoping to get 4. In this case, Bob's payoff is 1 instead of 5.

[1]We do not mean to suggest that Bob is explicitly applying backward or forward induction himself. Rather, a theorist can identify his reasoning as an instance of one or the other. In our models, the players' "choice" of reasoning style will be traced back to their prior beliefs about how likely it is that their opponents may make a mistake.

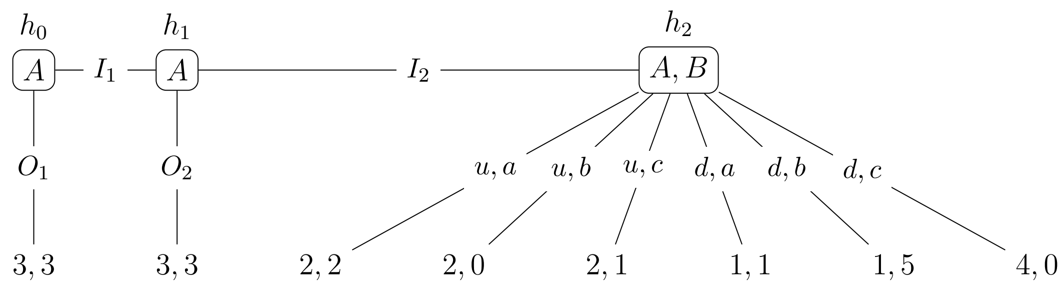


Figure 2: The game G_2

These considerations bring us to the following general question: How can a player deliberate between backward and forward induction in cases in which both seem plausible (at least *prima facie*) while dictating incompatible choices? Admittedly, in the above situation, Bob seems to be faced with a particularly difficult choice, since his information does not seem to sway the scales in favor of one or the other style of reasoning. But suppose that he is in a situation in which the players are prone to making *mistakes* relatively frequently—we elaborate on this notion below; for now simply think of the so-called "trembling hand mistakes" [32]: Ann chooses O, but plays I instead. In this case, backward induction reasoning should be preferred. Or let's say that Ann and Bob are playing a different game, G_2, in which mistakes are still possible, but Ann has two opportunities to go out before reaching the node at which the players move simultaneously—see Figure 2. Intuitively, if Bob observes Ann play I_1, it is reasonable for him to interpret her move as a mistake. Suppose, however, that Bob subsequently observes that Ann plays I_2. Now, the interpretation of her previous choices that is suggested by forward induction seems more plausible (we could of course modify the game further by adding more opportunities for Ann to exit the game). What this all suggests is that additional information about the *context of the game*—e.g., how probable it is that players make mistakes—can help the player settle on a style of reasoning.

Our ultimate goal is to study players that are able to deliberate between forward and backward induction, as well as conditions which make each style of reasoning superior to the other. The above considerations suggest that to do so we need to endow standard game theoretic agents with (relatively) rich beliefs. We use a model introduced by Stalnaker [33, 34, 35, 36] to describe the players' beliefs. In addition to having beliefs about their opponents' strategies and the game—standard for Epistemic Game Theory [16, 27, 29]—our players can also interpret observed behavior as either the result of deliberate action or as a mistake.

This paper is structured as follows. Section 2 describes the formal framework that allows us to represent the two lines of reasoning discussed above: extensive games with simultaneous moves (Section 2.1), our extension of Stalnaker's model (Section 2.2), two notions of rationality (Section 2.3), as well as an illustrative example (Section 2.4). In Section 3, we argue that our model offers an illuminating perspective on epistemic characterizations of backward induction and is a conservative extension of Stalnaker's model. Finally, Sections 4 and 5 discuss related work and outline a number of directions for future research.

2. FRAMEWORK

2.1 Extensive games with simultaneous moves

The examples from the introduction are **extensive games with simultaneous moves**.[2] Following [26, Section 6.3.2], we describe them as structures $\langle N, Act, H, \tau, \{u_i\}_{i\in N}\rangle$, where:

- N is a finite set of players.
- Act is the set of actions available to the players. To simplify notation, we assume that Act is partitioned into sets of actions for each player. For player $i \in N$, let $Act_i \subseteq Act$ denote player i's actions.
- H is a set of finite sequences of finite sequences of elements of Act. Elements $h \in H$ are called **histories**. We assume H satisfies the following constraints:
  - $\epsilon \in H$, where ϵ denotes the empty history.
  - If $h \in H$ and $h' \preceq h$, then $h' \in H$, where $h' \preceq h$ means that h' is a initial segment of h. Formally, we write $h' \preceq h$ provided $h = h'u$ where u is a sequence of sequences from Act, and $h'u$ denotes the concatenation of h' with u.
  - Each $h \in H$ is finite. That is, we restrict attention to *finite horizon* games.[3] We write $len(h)$ for the **length** of h (i.e., the number of elements in h).

  A history $h \in H$ is called a **terminal history** if there is no $h' \in H$ such that $h' \neq \epsilon$ and $hh' \in H$. Let $Z \subseteq H$ denote the set of terminal histories. Let $V = H - Z$ be the set of non-terminal histories. Each non-terminal history is associated with a simultaneous decision problem for a set of players. For this reason, we sometimes call elements of V **decision nodes**. For $h \in V$, let $A(h)$ be the possible extensions of h:

  $$A(h) = \{\vec{a} \mid h\vec{a} \in H \text{ and } \vec{a} \text{ is a sequence of actions}\}.$$

- τ is a turn function $\tau : V \to \wp(N)$ assigning a set of players[4] to each non-terminal history $h \in V$. For each $i \in N$, let $V_i = \{h \in V \mid i \in \tau(v)\}$ be the set of

[2]We assume that the reader is familiar with the basics of game theory. The formal definitions are included here to fix notation.

[3]This is a standard restriction in the literature on epistemic characterization of backward induction.

[4]Throughout this article, we assume that for all $h \in V$, $\tau(h) \neq \emptyset$. If we drop this assumption, then histories in which $\tau(h) = \emptyset$ should be interpreted as a move by nature.

non-terminal histories where player i moves. Similarly, we define the set of actions available to i at a decision node $h \in V_i$. For each $h \in V$ and $i \in \tau(h)$, let $A_i(h)$ be the set of actions available to i at h:

$$A_i(h) = \{a \in Act_i \mid \text{there is an } \vec{a} \in A(h) \text{ containing } a\}$$

If $i \notin \tau(h)$, then let $A_i(h) = \emptyset$. We impose an additional constraint to ensure that each decision node is associated with a *strategic game*:[5]

  – For each $h \in V$, $A(h) = \Pi_{i \in \tau(h)} A_i(v)$.

- For each $i \in N$, $u_i : Z \to \mathbb{R}$ is a utility function.

A **strategy** for player i assigns an action to each of i's decision nodes. Formally, a strategy for player i is a function $s_i : V_i \to Act$ where for all $h \in V_i$, $s_i(h) \in A_i(h)$. Let S_i be the set of all strategies for player i. As usual, a **strategy profile** is a sequence of strategies, one for each player (i.e., an element of $\Pi_{i \in N} S_i$). Given a strategy profile $\mathbf{s}$, let $\mathbf{s}_i$ be player i's component of $\mathbf{s}$ and $\mathbf{s}_{-i}$ the sequence of strategies from $\mathbf{s}$ for all players except i (i.e., $\mathbf{s}_{-i} \in \Pi_{j \neq i} S_j$). Each profile of strategies $\mathbf{s}$ generates a terminal history $\rho_{\mathbf{s}} \in Z$. We say that a non-terminal history is **reached** by a strategy profile provided h is an initial segment of $\rho_{\mathbf{s}}$.

A strategy for player i represents her *conditional plan* for the game. It prescribes a choice for player i at all of i's decision nodes, including those that are *ruled-out* by the strategy itself. Suppose that $h \in V_i$. An action $a \in Act_i(h)$ **rules out** a decision node $h' \in V_i$ provided $h \preceq h'$, but $h\vec{a} \not\preceq h'$ for any $\vec{a} \in A(h)$ containing a. In addition, we say that a rules out action a' provided $a' \in A_i(h')$ for some decision node $h' \in V_i$ that is ruled-out by a.[6] For example, in Figure 1, A's action O at h_0 rules out the actions u and d (because O rules out h_1).

2.2 Game models

A **game model** describes the players' beliefs during a play of the game. As discussed in the introduction, we are interested in representing players that allow for the possibility that one or more of their opponents made a *mistake*. This means that we must include states in which the moves of player i (i.e., the observed *behavior* of player i) does not match i's *choices*. To make this precise, each state in the game model will be associated with both strategies for the players *and* sequences of actions representing the observed behavior of the players.

The players' *behavior* in a game is represented by a sequence of actions. Recall that histories $h \in H$ are sequences of sequences of actions (one action for each player whose turn it is to move). For each $h \in H$, let $beh_i(h)$ be the sequence of i's actions in h. Formally, beh_i is defined by induction on the length of histories: $beh_i(\epsilon) = \epsilon$ (at the initial node, none of the players have made a choice), and

$$beh_i(h\vec{a}) = \begin{cases} beh_i(h)a & \text{if } i \in \tau(h) \text{ and } \vec{a} \text{ contains } a \\ beh_i(h) & i \notin \tau(h) \end{cases}$$

If X is a set, then X^* is the set of all finite strings of X. An i-history is a sequence of actions such that $\alpha = beh_i(h)$ for some $h \in H$. Given an i-history α and a decision node $h \in V_i$, let α_h be the component of α describing the action chosen at h. If α does not specify a move at h (either because the previous moves in α rule out h or α is not a maximal history), then α_h is undefined. For instance, in Figure 1, there are four A-histories (ϵ, O, Iu, and Id) and four B-histories (ϵ, a, b, and c). We use Ou to denote the strategy s_A in which $s_A(h_0) = O$ and $s_A(h_1) = u$ (similarly, for Od). Furthermore, we have $Iu_{h_0} = I$, $Iu_{h_1} = u$, and O_{h_1} is undefined.

A player history may be a *partial* description of what that player does in the game. This happens when the i-history α does not specify a choice for i at a decision node h not ruled out by α. Of course, if an i-history α specifies an action for player i at a decision node $h \in V_i$, then α specifies an action for i at each h' such that $h' \preceq h$ and $h' \in V_i$. We are interested in sets of player histories that represent possible plays of the game. A set of player histories $\{\alpha_i\}_{i \in N}$ is **coherent** if there is a history $h \in H$ such that for all $i \in N$, $beh_i(h) = \alpha_i$. Note that a set of i-histories may be coherent, yet not completely describe a path trough the game. For instance, $\{I, c\}$ is a coherent set of player histories in the game pictured in Figure 1: There are two histories $h = (I)(u, c)$ and $h' = (I)(d, c)$ such that $beh_B(h) = beh_B(h') = c$ and $beh_A(h) = beh_A(h') = I$. However, there is a unique history representing the play of the game associated with a coherent set of player strategies. The play of the game generated by a coherent set of i-histories $\{\alpha_i\}_{i \in N}$ is the longest history h such that $h \preceq h'$ for each h' such that for all $i \in N$, $beh_i(h') = \alpha_i$. The play of the game associated with the coherent set $\{I, c\}$ in the game in Figure 1 is $h = (I)$. The play of the game associated with a coherent set of player histories may be empty and need not be maximal. For example, the following table lists the coherent sets of strategies and the corresponding play of the game for the game pictured in Figure 1.

Coherent sets player strategies	Play of the game
$\{\epsilon, \epsilon\}$	ϵ
$\{O, \epsilon\}$	(O)
$\{I, a\}$, $\{I, b\}$, $\{I, c\}$, $\{I, \epsilon\}$	(I)
$\{Iu, \epsilon\}$, $\{Id, \epsilon\}$	(I)
$\{Iu, a\}$	$(I)(u, a)$
$\{Iu, b\}$	$(I)(u, b)$
$\{Iu, c\}$	$(I)(u, c)$
$\{Id, a\}$	$(I)(d, a)$
$\{Id, b\}$	$(I)(d, b)$
$\{Id, c\}$	$(I)(d, c)$

The sets $\{O, a\}$, $\{O, b\}$ and $\{O, c\}$ are not coherent.

[5] There is a hidden notational difficulty here. Since different players move at different decision nodes, the indices of the sequences of actions change from decision node to decision node. Formally, we represent a sequence $\vec{a}$ at decision node h as a function $\vec{a} : \tau(h) \to \cup_{i \in \tau(h)} A_i(h)$ where for each $i \in \tau(h)$, $\vec{a}(h) \in A_i(h)$. We write $\vec{a}_i$ to denote the action $a \in A_i(h)$ such that $\vec{a}(i) = a$ and say $\vec{a}$ contains a. This implicitly assumes that $i \in \tau(h)$ (otherwise $\vec{a}_i$ is not well-defined). Alternatively, we could assume that all players move at every decision node and introduce notation to distinguish "active" players from "passive" players. The passive players at a decision node would only have a single action available for their choice. We follow the first approach in this article.

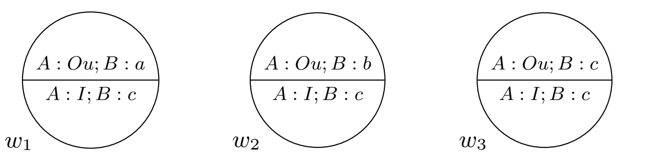
[6] We are implicitly assuming that all the action labels are unique. This assumption can be dropped, although it does simplify the notation.

Suppose that W is a nonempty set, elements of which are called **states**. Each player i will be associated with two functions β_i and σ_i subject to the following constraints:

1. For each $i \in N$, $\beta_i(w)$ is a (possibly empty) i-history and $\sigma_i(w)$ is a strategy for player i.
2. The i-histories $\{\beta_i(w)\}_{i\in N}$ are **coherent**.

We say that a player made a **mistake at a history** $h \in V_i$ in the world w provided her behavior is different than what is prescribed by her chosen strategy $\sigma_i(w)$ at h. Formally, i made a mistake at $h \in V_i$ provided $\beta_i(w)_h \neq \sigma_i(w)(h)$ (if $\beta_i(w)_h$ is defined).

Example. Recall the game in Figure 1 and consider three states w_1, w_2 and w_3. Suppose that $\sigma_A(w_1) = \sigma_A(w_2) = \sigma_A(w_3) = Ou$ (recall that Ou is the strategy in which A chooses O at h_0 and u at h_1, but it is not an A-history) and $\beta_A(w_1) = \beta_A(w_2) = \beta_A(w_3) = I$. Thus, A made a mistake at h_0. The strategies for player B are $\sigma_B(w_1) = a$, $\sigma_B(w_2) = b$, $\sigma_B(w_3) = c$ (again, these are the strategies in which B chooses, respectively, a, b, and c at h_1), and $\beta_B(w_1) = \beta_B(w_2) = \beta_B(w_3) = c$. These states are pictured as follows:



The strategies $\sigma_A(w)$ and $\sigma_B(w)$ are displayed in the top half of the circles and the histories $\beta_A(w)$ and $\beta_B(w)$ in the bottom half (where $w \in \{w_1, w_2, w_3\}$). If these states describe A's beliefs (i.e., they are the set of doxastic possibilities for A), then A is certain that B will play c, but is uncertain about exactly *why* B is playing c. It might be because B made a mistake (as in states w_1 and w_2) or because B simply followed through on his plan to play c. Furthermore, A arrived at these beliefs under the supposition that she (contrary to her chosen strategy) selected I at h_0.

The players' beliefs and belief revision policies are represented in the standard way (cf. [4, 10, 36]). Each player $i \in N$ is associated with a **prior probability** on the set of states, $P_i \in \Delta(W)$,[7] and a **plausibility ordering** $\succeq_i \subseteq W \times W$ satisfying the following constraints: for each $i \in N$ and for each $w \in W$, $P_i(w) > 0$ (i.e., P_i is a full support probability measure); $\succeq_i$ is a locally connected (for all w, v, x, if $w \succeq_i x$ and $w \succeq_i v$, then $v \succeq_i x$ or $x \succeq_i v$) partial order (reflexive and transitive relation) on W. The plausibility ordering $\succeq_i$ represents player i's belief revision policy. For each $i \in N$ and states $w, v \in W$, let $w \approx_i v$ iff $w \succeq_i v$ or $v \succeq_i w$. Since, $\succeq_i$ is a locally complete partial order, $\approx_i$ is an equivalence relation. For $w \in W$, let $[w]_i = \{v \mid w \approx_i v\}$ denote the equivalence class of w for $\approx_i$, called i's *information cell.* The intended interpretation is that $w \approx_i v$ means that w and v are *subjectively indistinguishable* to player i (i's beliefs, knowledge, and conditional beliefs are the same in both states).[8] The players' *full beliefs* at a state w are defined as usual: For each $w \in W$, let $\max_{\succeq_i}([w]_i) = \{x \mid \text{there is no } y \in [w]_i \text{ such that } y \succ_i x\}$, where $y \succ_i x$ means that $y \succeq_i x$ but $x \not\succeq_i y$.

Definition 1. For each $w \in W$ and $i \in N$, player i's **(partial) beliefs** at state w are given by the probability measure $P_{i,w} \in \Delta(W)$ defined as follows: For each $E \subseteq W$,

$$P_{i,w}(E) = P_i(E \mid \max_{\succeq_i}([w]_i))$$

The players' partial beliefs $P_{i,w}$ represent their beliefs about the possible choices, behaviors and beliefs of their opponents at state w.[9] The belief revision policy describes how the players revise their beliefs given *any* evidence $F \subseteq W$:

$$P_{i,w}(E \mid F) = P_i(E \mid \max_{\succeq_i}(F \cap [w]_i)).$$

Note that this conditional probability is well-defined for any set F such that $F \cap [w]_i \neq \emptyset$. In particular, there may be a set F such that $P_{i,w}(F) = 0$, yet $P_{i,w}(\cdot \mid F)$ is well-defined. This is a very general model of belief revision for the players, since it describes how each player revises her beliefs given *any* evidence consistent with her current information (i.e., any F such that $[w]_i \cap F \neq \emptyset$). However, we are primarily interested in how the players revise their beliefs given the actions that they observe in the game.[10] Each state $w \in W$ is associated with a history $h \in H$ as follows. Let h_w be the history corresponding to the play of game associated with $\{\beta_i(w)\}_{i\in N}$ (see the discussion above). Note that h_w need not be a maximal history, so h_w is the behavior that is observed at state w. For any $h \in H$, let $[h] = \{w \mid \beta_i(w) = beh_i(h) \text{ for all } i \in N\}$ be the event that the players behaved according to history h. Then, $P_{i,w}(E \mid [h_w])$ is i's probability of E given her most plausible explanation of the actions she observed at state w. Thus, the belief revision policy describes how the players' beliefs change during a play of the game.[11]

Putting everything together, a **game model** for a game G is a tuple $\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i\in N}, \{\succeq_i\}_{i\in N}, \{P_i\}_{i\in N}\rangle$. In addition, we impose the following two constraints:

- For all $w \in W$ and $i \in N$, if $v \in [w]_i$, then $\sigma_i(w) = \sigma_i(v)$. That is, players *know* their own strategy.[12]
- For all $w \in W$ and $i \in N$, for each initial segment $h' \subseteq h_w$ (including the empty history), there is a $w' \in [w]_i$ such that $h_w = h'$.

The last constraint ensures that if a sequence of choices in the game is consistent with a player's information, then all of

[7]For a set X, let $\Delta(X)$ be the set of all probability measures on X. In this paper, we assume that the set of states is finite, so we can assume that P is defined on all subsets of W.

[8]That is, the equivalence classes of $\approx_i$ are the different "types" for player i.

[9]That is, beliefs about the possible types of their opponents.

[10]Our models of games are closely related to Bayesian extensive games with observable actions [26, Section 12.3]. However, there are important methodological and conceptual differences between Bayesian games and epistemic models of games (see [27, Section 1.4]). For this reason, we postpone a complete comparison between our game models and Bayesian extensive games with observable actions to the full version of the paper.

[11]Thus, our models are related to the *type spaces* based on conditional probability systems from [7, Section 2.2].

[12]Each player can be associated with a standard knowledge operator where for all E, $K_i(E) = \{w \mid [w]_i \subseteq E\}$.

its initial segments must be consistent with the player's information. This is a consequence of assuming that the structure of the game is (commonly) known to all the players and that players cannot think it is possible to observe a history without observing the sequence of choices that generated the history. Compare the above constraint with the stronger assumption that for all $w \in W$, for all $h \in H$, there is a $w' \in [w]_i$ such that $h_{w'} = h$. This ensures that it is consistent with the players' information that every possible history in the game could be realized. Of course, *ex ante*, the players do not rule out any histories.[13] However, our models represent the players' *ex iterim* beliefs. In such models, it may be consistent with a player's information (which includes her chosen strategy) that some history of the game will not be played (cf. the discussion of richness conditions on the model in Section 5).

2.3 Rationality

A player chooses rationally provided her strategy choice at a state maximizes the players subjective expected utility with respect to her beliefs about the past and expected moves of her opponents. We do not assess the rationality of the players' moves themselves. Thus, a player may choose rationally at a state, though she may not carry out her plan because she made a mistake.

Suppose that G is an extensive game with simultaneous moves and $\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i \in N}, \{\succeq_i\}_{i\in N}, \{P_i\}_{i\in N}\rangle$ is a game model for G. For each $w \in W$, the strategy realized at w by player i is $s_i(w) : V_i \to Act_i$ defined as follows:

$$s_i(w)(h) = \begin{cases} \beta_i(w)_h & \text{if } \beta_i(w)_h \text{ is defined} \\ \sigma_i(w)(h) & \text{otherwise} \end{cases}$$

Then, $\mathbf{s}(w) = (s_1(w), \dots, s_n(w))$ is a profile of strategies, and let $Out(\mathbf{s})$ be the (unique) terminal history generated by $\mathbf{s}$.

Definition 2. For any strategy $s_i \in S_i$ for player i, the **expected utility** of s_i at state w is:

$$EU_{i,w}(s_i) = \sum_{w' \in W} P_{i,w}(\{w'\} \mid [h_w])u_i(Out(s_i, \mathbf{s}_{-i}(w))).$$

A player chooses optimally at state w provided her current strategy maximizes her subjective expected utility at w, given the actions that she observed. Let $S_i(w) \subseteq S_i$ be the set of strategies for player i that conform to player i's moves in state w. That is, $s_i \in S_i(w)$ implies that for all $h \in V_i$, if $\beta_i(w)_h$ is defined, then $s_i(h) = \beta_i(w)_h$. Then,

$$\begin{array}{rcl} Opt_i & = & \{w \mid \sigma_i(w) \text{ maximizes expected utility} \\ & & \text{with respect to } P_{i,w} \text{ and } S_i(w)\}. \end{array}$$

If $w \in Opt_i$, then player i is adopting the best possible strategy given i's observations at w. Rationality is more demanding. There are two versions of rationality. The first requires that a player is rational at a state w provided her strategy at w is optimal given her beliefs at w *and* was optimal at all previous decision nodes given her beliefs at the moment of decision. We say that a state $w' \in [w]_i$ is an **earlier choice state** provided $\beta_i(w')$ is an initial segment of $\beta_i(w)$.

[13] Assuming that all the players are *aware* (in the sense of [22, 23]) of the structure of the game.

Definition 3. Player i is **rational-1** at state w provided $w' \in Opt_i$ for *all* earlier choice states w'. Let Rat_i^1 be the set of all states w such that i is rational-1 in w.

A player may be rational-1 even if she does not correctly implement her strategy. The second version of rationality requires that a player's strategy is optimal even when the player learns that her beliefs are mistaken. That is, the strategy is optimal and remains optimal after any belief revision.

Definition 4. Player i is **rational-2** at state w provided $w' \in Opt_i$ for *all* states $w' \in [w]_i$. I.e., $[w]_i \subseteq Opt_i$. Let Rat_i^2 be the set of all states w such that i is rational-2 in w.

Of course, $Rat_i^1 \subseteq Rat_i^2$ (if a player is rational-1, then the player is rational-2). However, in general, the converse is not true (this is illustrated by an example in the next section).

2.4 Example

Figure 3 depicts models of the games from Figures 1 and 2. These models represent the players' initial beliefs and dispositions to change their beliefs that we discussed in the introduction. The model on the left, $\mathcal{M}_1$, represents one play of the game in Figure 1, and the model on the right, $\mathcal{M}_2$, represents a play of the game in Figure 2. We draw an arrow from state v to state w when $w \succeq_i v$. The solid arrows represent Bob's plausibility ordering $\succeq_B$ and the dashed arrows represent Ann's plausibility ordering $\succeq_A$ (we only represents Bob's beliefs in $\mathcal{M}_2$). To keep down the clutter in the pictures, we assume that the remaining arrows can be inferred by transitivity and reflexivity. The strategies $\sigma_A(w)$ and $\sigma_B(w)$ are displayed in the top half of the circles and the histories $\beta_A(w)$ and $\beta_B(w)$ in the bottom half (empty histories are left blank). We think of the players strategy choices and moves as discrete random variables. Thus, $[Choose_i^h = a] = \{w \mid \sigma_i(w)(h) = a\}$ is the event that player i chooses action a at decision node h. Similarly, $[Move_i^h = a] = \{w \mid \beta_i(w)_h = a\}$ is the event that player i played a at history h. The (common) prior probabilities are displayed next to the states.

Suppose that w_4 is the actual world in model $\mathcal{M}_1$. Thus, Ann chose the strategy Ou, but made a mistake and played I followed by u (as originally planned). Bob chose strategy a which he correctly implemented when given the chance to move. His (overall) most plausible worlds are w_1 and w_2. This means that he is certain that Ann plays O at h_0 (i.e., $P_{B,w_4}([Choose_A^{h_0} = O]) = 1$). Moreover, he (initially) thinks that Ann's strategies Ou and Od are equally likely (i.e., $P_{B,w_4}([Choose_A^{h_1} = u]) = P_{B,w_4}([Choose_A^{h_1} = d]) = 0.5$). If Ann surprises Bob by playing I, he is disposed to interpret this as a mistake on her part, rather than as revealing that she is following a different strategy (i.e., $\max_{\succeq_B}([w_4]_B \cap [Move_A^{h_0} = I]) = \{w_3\}$, $\beta_A(w_3) = I$ while $\sigma_A(w_3) = Ou$). Furthermore, after observing Ann play I, Bob is certain that her next move will be u: $P_{B,w_4}([Choose_A^{h_1} = u] \mid [Move_A^{h_0} = I]) = 1$. This model also illustrates what it means for a player to be rational-1. Note that Ann made a mistake in w_4, yet she is still rational-1 ($w_4 \in Rat_A^1$). Both w_1 and w_3 are earlier choice states for Ann (as is w_4), and she chooses optimally in all these states: $Opt_A = \{w_1, w_2, w_3\}$.

The model $\mathcal{M}_2$ in Figure 3 represents Bob's beliefs in the game from Figure 2 in which Ann has two opportunities to exit the game. Suppose that w_6 is the actual world. No

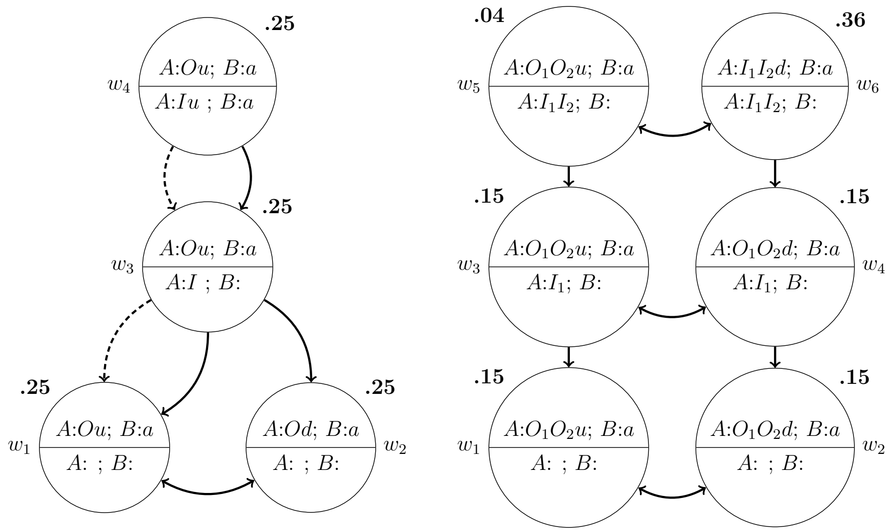


Figure 3: The models $\mathcal{M}_1$ and $\mathcal{M}_2$

mistakes are made with Ann playing $I_1 I_2 c$ and Bob playing u. Initially, Bob believes that Ann is going to choose O_1 ($\max_{\succeq_B}([w_6]_B) = \{w_1, w_2\}$ with $\sigma_A(w_1)(h_0) = \sigma_A(w_2)(h_0) = O_1$). On the condition that Ann actually plays I_1, he is disposed to interpret her move as a mistake, predicting that she is going to go out at the next opportunity ($\max_{\succeq_B}([w_6]_B \cap [Move_A^{h_1} = I_1]) = \{w_3, w_4\}$, $\sigma_A(w_3)(h_1) = \sigma_A(w_4)(h_1) = O_2$. If Ann surprises Bob the second time by playing I_2, he is disposed to conclude that it is very likely that Ann actually chose to play I_1 and I_2 ($P_{B,w_6}([Choose_A^{h_0} = I_1] \cap [Choose_A^{h_1} = I_2] \mid [Move_A^{h_0} = I_1] \cap [Move_A^{h_1} = I_2]) = .9$). Intuitively, if Ann surprises Bob, he is disposed to reason in the backward induction style (ignoring her mistake), but if she surprises him a second time, Bob switches to forward induction and conjectures that (it is highly probable that) Ann is going to play d. The model $\mathcal{M}_2$ also illustrates the difference between rationality-1 and rationality-2. In w_2, Bob is rational-1, since $\{w_1, w_2\} \subseteq Opt_B$, but he is not rational-2, since $w_6 \in [w_2]_B$, but $w_6 \notin Opt_B$. That is, the strategy that Bob chooses at w_2 is not optimal *with respect to the beliefs he would have* after revising his initial beliefs with the information that Ann plays I_1 and I_2.

3. STALNAKER AND AUMANN

It is easy to see that our game model is a conservative extension of Stalnaker's [34, 35, 36]. Since we extend his model by allowing states in which the players' moves differ from what is prescribed by their chosen strategy, the players can *know* each other's strategies and still be uncertain about the way the game is going to end. In spite of this, however, our models can accommodate the standard epistemic characterizations of backward induction found in the literature, and, in particular, Aumann's classic characterization.

Aumann proved that if there is common knowledge that all of the players are rational, then the backward induction path will be realized [2]. Our models are richer than Aumann's: We describe the players' beliefs and belief revision policies in addition to the players' knowledge. Recently, Samet extended Aumann's result to doxastic models, which are much closer to the models we use. He proved that if there is *common belief*[14] that all the players' strategies are *doxastically substantively rational*, then the backward induction path is realized [31]. Since we allow for mistakes, we will have models in which there is common belief that the players choose optimally, but the backward induction path does not obtain. There is another difference between our models and Aumann's and Samet's. The behavior functions can be viewed as a temporal parameter. That is, our model includes states that describe the players' beliefs at different moments during the play of the game (cf. [5, 12]). In general, the players' beliefs may change even if the game unfolds according to their chosen strategies. We can recover Samet's characterization of backward induction with an additional constraint:

For all $i \in N$ and $w, w' \in W$, if $w' \in [w]_i$, then for each $w'' \in \max_{\succeq_i}([w]_i \cap [h_{w'}])$ there is a $w''' \in \max_{\succeq_i}([w]_i)$ such that $\sigma_i(w'') = \sigma_i(w''')$ for all $i \in N$.

This constraint says that the players cannot learn anything about their opponents' strategies that they did not already know at the beginning of the game.

Proposition 1. *Suppose that G is an extensive game (without simultaneous moves) in "general position" (see Appendix A) and $\mathcal{M}_G$ is a model for G satisfying the above constraint and that every possible mistake is considered: for all $w \in W$, every possible mistake that i can make given i's strategy at w is realized by the behavior at some state $w' \in [w]_i$. Suppose that $w \in W$ is a state in which the histories $(\beta_1(w), \ldots, \beta_n(w))$ generate a maximal path through the game. If there are no mistakes in w and common belief at w that all the players are rational-1, then the path that is generated by the histories is the backward induction path.*

[14] We assume that the reader is familiar with the formal definition of common belief. See Appendix A for the formal definition in our framework.

There are many other epistemic characterizations of backward induction.[15] What is more relevant for our purposes is Stalnaker's criticism of Aumann's epistemic characterization of backward induction [35, Section 5]. The problem lies with Aumann's notion of rationality which is captured and refined by our rationality-1. A player is rational-1 provided her strategy is optimal (given a sequence of moves) and was optimal at all previous choices *with respect to her beliefs at the moment of choice*. Stalnaker argues that this notion of rationality is much too strong.[16] His idea is that a strategy for player i is optimal provided i would choose optimally at node v (according to his strategy) given i's beliefs *under the hypothesis that node v is reached*. To formalize this idea, Stalnaker introduces the notion of *perfect rationality*: "In cases where two or more [strategies] are [optimal], the agent should consider, in choosing between them, how he should act if he learned he was in error about something." [34, pg. 148]. It is not hard to see that our definition of rationality-2 is equivalent to Stalnaker's definition of perfect rationality. Thus, our models can accommodate both Stalnaker's and Aumann's analysis of backward induction. Of course, this, by itself, is not new (cf. [5] and [19]). However, our analysis also opens the door for further refinements of the notions of rationality they use.

For instance, perfect rationality (or rationality-2) requires that a player's strategy is *robustly optimal*. That is, it is optimal even after the player learns that her beliefs are mistaken. Variants of rationality-2 can be defined by fixing the set of *evidence* that may induce a change of belief for a player. For instance, we can require that a player's strategy must be robustly optimal with respect to evidence about her opponents' moves (cf. [7, Section 2.2]), strategy choices, beliefs, or even evidence about the player's own moves. A complete analysis of the different options will be left for the full version of the paper.

4. RELATED WORK

Our model allows for states in which the players' moves differ from what is prescribed by their chosen strategy. This general idea (i.e., trembling hand mistakes) was used by Selten and others to characterize *refinements* of the Nash equilibrium (cf. [18, 32]). Within the equilibrium refinement program, Bicchieri's work on forward and backward induction [9] comes closest to ours. In [9], the players respond to (hypothetical) surprising moves in an extensive game (that may be the result of a trembling-hand mistake) by revising their beliefs à la AGM [1]. Our models differ in both important technical details and the underlying motivations. Most importantly, we downplay the role that the Nash equilibrium (and its refinements) plays in the analysis of rational behavior in game situations (this is in line with much of the epistemic game theory literature, cf. [13]).

More recently, Cubitt and Sugden develop a model in which a player's behavior may, in principle, differ from her (rational) choice [15]. They include a postulate stating that the players' behavior must all conform to the same principles of rational choice. Among other things, they are interested in highlighting the role that this assumption plays in the players' *reasoning* about what to do in a game situation (cf. also Bacharach's discussion of the *transparency of reason* in [3, Section 4.2]). There are some intriguing connections between our work and theirs, but a complete discussion will be left for the full version of this paper.

5. CONCLUDING REMARKS

We have imposed only two minimal constraints on our models: every information cell must include the player's beliefs at all previous choice points, and the players "know" their own strategy choice. The literature on forward induction, and, more generally, belief revision in games [5, 11, 35], contains other natural constraints that we may want to impose. One belief revision policy that has been extensively discussed in relation to forward induction reasoning is the so-called **rationalizability principle** [8]: "A player should always try to interpret her information about the behavior of her opponents assuming that they are not implementing 'irrational' strategies." (cf. [6]). In order to represent this belief revision policy, Stalnaker includes a "richness" condition on his models [35, pg. 35, footnote 5] ensuring that the players have the conditional beliefs needed to rationalize any observed behavior.[17] With such a richness condition, we can formally prove Stalnaker's characterization of the belief revision policy in which the players apply the rationalizability principle at most once.[18]

Another direction for future research is to compare our approach to belief revision with non-standard probabilities, lexicographic probability systems, and conditional probability systems [21, 25]. Once the relationship between these different models is understood, we can connect our work with Battgalli and Siniscalchi's characterizations of common *strong belief* of rationality [7] and Halpern's recent epistemic characterizations of trembling-hand equilibria using non-standard probabilities [20].

Finally, note that the games in Figures 1 and 2 have the same reduced normal form. However, our analysis in this paper suggests that there are strategically relevant differences between the two games (cf. [24]). In particular, the players may be able to *learn* about their opponents' strategies during a play of the game. This suggests possible connections with models of learning in extensive games [17].

6. REFERENCES

[1] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510 – 530, 1985.

[2] R. Aumann. Backward induction and common knowledge of rationality. *Games and Economic Behavior*, 8(1):6 – 19, 1995.

[3] M. Bacharach. A theory of rational decision in games. *Erkenntnis*, 27(1):17 – 55, 1987.

[15] It is beyond the scope of this article to survey all of the different approaches. See [29, Section 8.11] and [27] for a discussion and pointers to the literature.

[16] We will not repeat Stalnaker's argument here. The gist of it is that it is important not to conflate "action a would be optimal if node v were reached" and "if node v *is* reached, then action a is optimal".

[17] This is related to Battigalli and Siniscalchi's use of *complete* (conditional) type spaces in their characterization of extensive-form rationalizability and related concepts. See [7, Section 6.2.3] for a discussion.

[18] Stalnaker states and hints at a proof of such a theorem in [35, pg. 51, footnote 22].

[4] A. Baltag and S. Smets. Conditioanl doxastic models: A qualitative approach to dynamic belief revision. In *Electornic notes in theoretical computer science*, volume 165, pages 5 – 21. Springer, 2006.
[5] A. Baltag, S. Smets, and J. Zvesper. Keep 'hoping' for rationality: a solution to the backwards induction paradox. *Synthese*, 169:301–333, 2009.
[6] P. Battigalli. On rationalizability in extensive games. *Journal of Economic Theory*, 74:40 – 61, 1997.
[7] P. Battigalli and M. Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106(2):356 – 391, 2002.
[8] D. Bernheim. Rationalizable strategic behavior. *Econometrica*, 52:1007 – 1028, 1984.
[9] C. Bicchieri. *Rationality and Coordination (Cambridge Studies in Probability, Induction, and Decision Theory)*. Cambridge: Cambridge University Press, 1993.
[10] O. Board. Dynamic interactive epistemology. *Games and Economic Behavior*, 49(1):49–80, 2004.
[11] G. Bonanno. AGM belief revision in dynamic games. In K. R. Apt, editor, *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK XIII)*, pages 37 – 45, 2011.
[12] G. Bonanno. A doxastic behavioral characterization of generalized backward induction. *Games and Economic Behavior*, 88:221 – 241, 2014.
[13] A. Brandenburger. Introduction. In *The Language of Game Theory: Putting Epistemics into the Mathematics of Games*. World Scientific Series in Economic Theory, 2014.
[14] R. Cooper, D. DeJong, R. Forsythe, and T. Ross. Forward induction in the battle-of-the-sexes games. *The American Economic Review*, 83(5):1303 – 1316, 1993.
[15] R. P. Cubitt and R. Sugden. Common reasoning in games: A Lewisian analysis of common knowledge of rationality. *Economics and Philosophy*, 30:285 – 329, 2014.
[16] E. Dekel and M. Siniscalchi. Epistemic game theory. In H. P. Young and S. Zamir, editors, *Handbook of Game Theory, Volume 4*. Elsevier, 2014.
[17] D. Fudenberg and D. M. Kreps. Learning in extensive-form games I. self-confirming equilibria. *Games and Economic Behavior*, 8:20 – 55, 1995.
[18] S. Govindan and R. Wilson. Nash equilibrium, refinements of. In S. N. Durlauf and L. E. Blume, editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke, 2008.
[19] J. Halpern. Substantive rationality and backward induction. *Games and Economic Behavior*, 37(2):425 – 435, 2001.
[20] J. Halpern. A nonstandard characterization of sequential equilibrium, perfect equilibrium, and proper equilibrium. *International Journal of Game Theory*, 3838(1):37–50, 2009.
[21] J. Halpern. Lexicographic probability, conditional probability, and nonstandard probability. *Games and Economic Behavior*, 68(1):155 – 179, 2010.
[22] J. Halpern and L. Rêgo. Extensive games with possibly unaware players. *Mathematical Social Sciences*, 70:42 – 58, 2014.
[23] A. Heifetz, M. Meier, and B. Schipper. Interactive unawareness. *Journal of Economic Theory*, 130(1):78 – 94, 2006.
[24] E. Kohllberg and J. Francois Mertens. On the strategic stability of equilibria. *Econometrica*, 54(5):1003 – 1037, 1986.
[25] D. Lehman and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55(1):1 – 60, 1992.
[26] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.
[27] E. Pacuit and O. Roy. Epistemic foundations of games. In E. N. Zalta, editor, *Stanford Encyclopedia of Philosophy*, 2015.
[28] A. Perea. Backward induction versus forward induction reasoning. *Games*, 1(3):168–188, 2010.
[29] A. Perea. *Epistemic Game Theory: Reasoning and Choice*. Cambridge UP, 2012.
[30] A. Perea. Belief in the opponents' future rationality. *Games and Economic Behavior*, 83:231 – 254, 2014.
[31] D. Samet. Common belief of rationality in games of perfect information. *Games and Economic Behavior*, 79:192 – 200, 2013.
[32] R. Selten. A reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25 – 55, 1975.
[33] R. Stalnaker. On the evaluation of solution concepts. *Theory and Decision*, 37(1):49 – 73, 1994.
[34] R. Stalnaker. Knowledge, belief and counterfactual reasoning in games. *Economics and Philosophy*, 12(02):133 – 163, 1996.
[35] R. Stalnaker. Belief revision in games: forward and backward induction. *Mathematical Social Sciences*, 36(1):31 – 56, 1998.
[36] R. Stalnaker. Extensive and strategic forms: Games and models for games. *Research in Economics*, 53(3):293 – 319, 1999.
[37] E. van Damme. Stable equilibria and forward induction. *Journal of Economic Theory*, 48:476 – 496, 1989.

APPENDIX

A. PROOF OF PROPOSITION 1

In this appendix, we restrict attention to extensive games $G = \langle N, Act, H, \tau, \{u_i\}_{i\in N}\rangle$ without simultaneous moves. So, for all decision nodes $v \in V$, $|\tau(v)| = 1$. In this case, we can view histories as sequences of actions rather than sequences of sequences of actions. Furthermore, following [2], we assume that the payoff for each of the players is different at different terminal nodes (the game is in "general position"). This implies that the result of applying the backward induction algorithm[19] is uniquely defined.

[19] The terminal nodes are labeled with the payoffs for each player. For each non-terminal history h with $\tau(h) = \{i\}$, label h with the maximum of all the labels of the successors of h. This labeling is then used to identify the so-called *backward induction path*.

Belief operators: Suppose that

$$\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i\in N}, \{\succeq_i\}_{i\in N}, \{P_i\}_{i\in N}\rangle$$

is a game model. For each event $E \subseteq W$, we say that player i believes E, $B_i(E)$, provided E is implied by i's full beliefs. That is, $B_i(E) = \{w \mid \max_{\succeq_i}([w]_i) \subseteq E\}$.

Samet's game model: Samet's game model is a tuple $\langle W, \{\Pi_i, t_i\}_{i\in N}, \mathbf{s}\rangle$, where W is a non-empty set of states, for each $i \in N$, Π_i is a partition on W, $t_i : W \to \Delta(W)$ is a type function assigning a probability measure to each state, and $\mathbf{s} : W \to S$ (where $S = \Pi_i S_i$) assigns a strategy to each state. Let $[\mathbf{s}_i(w) = s_i]$ be the set of states w such that $\mathbf{s}_i(w) = s_i$. The knowledge and belief operators are defined as usual: for all $E \subseteq W$, $K_i(E) = \{w \mid \Pi_i(w) \subseteq E\}$ and $B_i(E) = \{w \mid t_i(w)(E) = 1\}$. Samet includes the following constraints:

- For all $w \in W$, if $v \in \Pi_i(w)$, then $t_i(w) = t_i(v)$
- For all $w \in W$, $t_i(w)(\Pi_i(w)) = 1$
- For all $w \in W$, if $v \in \Pi_i(w)$, then $[\mathbf{s}_i(w) = s_i] \subseteq B_i([\mathbf{s}_i(w) = s_i])$.

Rationality in Samet's model: Building on the notation introduced in Section 2, for a strategy profile $\mathbf{s}$, let $Out_h(\mathbf{s})$ be the (unique) terminal history that is reached if the players follow their strategies in $\mathbf{s}$ starting at h. Then, for a state $w \in W$ and strategy $s_i \in S_i$, $Out_v(s_i, \mathbf{s}_{-i}(w))$ is the terminal node that is reached, starting at h, if player i follows the strategy s_i and the other players follow the strategies associated with state w. Then, let

$$[Out_h(s_i, \mathbf{s}_{-i}) >_i Out_h(\mathbf{s})] \quad = \{w \mid u_i(Out_h(s_i, \mathbf{s}_{-i}(w))) > u_i(Out_h(\mathbf{s}(w)))\}.$$

Player i is said to be **doxastically substantively rational** at all states when:

$$R_i^{ds} = \bigcap_{h\in V_i} \bigcap_{s_i\in S_i} \neg B_i([Out_h(s_i, \mathbf{s}_{-i}) >_i Out_h(\mathbf{s})])$$

Let $R^{ds} = \bigcap_{i\in N} R_i^{ds}$.

Common belief: Given belief operators $B_i : \wp(W) \to \wp(W)$ for each player $i \in N$ (defined in Samet's game model or our game model), we define a common belief operator $CB : \wp(W) \to \wp(W)$ in the usual way. First, define *everyone believes*: For all $E \subseteq W$, $B(E) = \bigcap_{i\in N} B_i(E)$. Then define the nth power of B, B^n, as follows: for all $E \subseteq W$, $B^1(E) = B(E)$ and for $n > 1$, $B^n(E) = B(B^{n-1}(E))$. Finally, **common belief** of an event E is $CB(E) = \bigcap_{n\geq 1} B^n(E)$

Samet's Theorem 3 states that, in any of his models, $CB(R^{ds}) \subseteq I$, where I is the set of states in which the backward induction path is played.

Suppose that $\mathcal{M}_G = \langle W, \{(\beta_i, \sigma_i)\}_{i\in N}, \{\succeq_i\}_{i\in N}, \{P_i\}_{i\in N}\rangle$ is a game model for our game G. The **forgetful projection of** $\mathcal{M}_G$, denoted $\mathcal{M}_G^\circ$, is the tuple $\langle W, \{\Pi_i, t_i\}_{i\in N}, \mathbf{s}\rangle$, where for each $w \in W$, let $\Pi_i(w) = [w]_i$, $t_i(w) = P_{i,w}$, and $\mathbf{s}(w) = (\sigma_1(w), \ldots, \sigma_n(w))$. It is not hard to see that $\mathcal{M}_G^\circ$ satisfies the constraints imposed by Samet. For instance, we have, for all $w' \in \Pi_i(w)$, $t_i(w) = t_i(w')$, since if $w' \approx_i w$, then $\max_{\succeq_i}([w]_i) = \max_{\succeq_i}([w']_i)$.

We first state and prove a simple Lemma that will be used to relate Samet's notion of doxastic substantive rationality with our rationality-1.

LEMMA 1. *Suppose that the game G and game model $\mathcal{M}_G$ and state in w satisfy the assumption of Proposition 1. Then, for all players $i \in N$, for all $w' \in W$, if $w \in Rat_i^1$, then for all $v \in V_i$, there is some $w' \in \max_{\succeq_i}([w]_i)$ such that $u_i(Out_h(\mathbf{s}(w'')) > u_i(Out_h(s_i; \mathbf{s}_{-i}(w')))$.*

PROOF. First of all, it is easy to see that if a strategy s_i is optimal for player i at state w, then for all strategies $t_i \neq s_i$, there must be at least one state $w' \in \max_{\succeq_i}([w]_i \cap [Move_i^w = \beta_i(w))$ such that $u_i(Out(s_i; \mathbf{s}_{-i}(w))) > u_i(Out(t_i; \mathbf{s}_{-i}(w)))$.

Suppose that $w \in Rat_i^1$. Then, for all $h \in V_i$, if $\beta_i(w)_h$ is defined (i.e., h on the path generated by the behavior of the players in state w), then for all $s_i \in S_i(w)$, there is at least one state $w' \in \max_{\succeq_i}([w]_i \cap [h_w])$ such that $u_i(Out_h(\mathbf{s}(w))) > u_i(Out_h(s_i; \mathbf{s}_{-i}(w)))$. Note that we can move from $Out(\cdot)$ to $Out_h(\cdot)$ since we restrict attention to strategy profiles that conform to the behavior of the players at w. By the constraint stated before Proposition 1, this implies that there is a $w'' \in \max_{\succeq_i}([w]_i)$ such that $u_i(Out_h(\mathbf{s}(w''))) > u_i(Out_h(s_i; \mathbf{s}_{-i}(w'')))$. This, together with the assumption that all mistakes are realized by some state in i's information cell, ensures that, for every decision node $h \in V_i$, there is some $w' \in \max_{\succeq_i}([w]_i)$ such that

$$u_i(Out_h(\mathbf{s}(w'))) > u_i(Out_h(s_i; \mathbf{s}_{-i}(w'))).$$

This completes the proof of the Lemma. $\square$

The proof of the proposition follows immediately:

PROOF OF PROPOSITION 1. Suppose that $w \in W$ and $(\beta_1(w), \ldots, \beta_n(w))$ generate a maximal path through the game. If $w \in CB(\bigcap_j Rat_j^1)$ in $\mathcal{M}_G$, then Lemma 1 implies that $w \in CB(R^{ds})$ in the forgetful projection $\mathcal{M}_G^\circ$. Since $\mathcal{M}_G^\circ$ is a Samet model of a game, Samet's Theorem 3 implies that $w \in I$. Since no mistakes are made in w, this implies that $(\beta_1(w), \ldots, \beta_n(w))$ is the backward induction path. $\square$