Dynamic unawareness and rationalizable behavior ☆

Aviad Heifetz [a],[*], Martin Meier [b], Burkhard C. Schipper [c]

[a] *The Economics and Management Department, The Open University of Israel, Israel*
[b] *Institut für Höhere Studien, Wien, Austria*
[c] *Department of Economics, University of California, Davis, United States*

ARTICLE INFO

ABSTRACT

We define generalized extensive-form games which allow for asymmetric awareness of actions. We extend Pearce's (1984) notion of extensive-form (correlated) rationalizability to this setting, explore its properties, and prove existence.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

In real-life dynamic interactions, unawareness of players regarding the relevant actions available to them is at least as prevalent as uncertainty regarding other players' strategies, payoffs or moves of nature. Players frequently become aware of actions they (or other players) could have taken in retrospect, when they can only re-evaluate the past actions chosen by partners or rivals who *were* aware of those actions from the start, and hence re-assess their likely future behavior. Yet, while uncertainty can be captured within the standard framework of extensive-form games with imperfect information, unawareness and mutual uncertainty regarding awareness require an extension of this framework. Such an extension is the first task of the current paper.

At first, one may wonder why the standard framework would not suffice. After all, if a player is unaware of an action which is actually available to her, then for all practical purposes she cannot choose it. Why wouldn't it be enough simply to truncate from the tree all the paths starting with such an action? The reason is that the strategic implications of
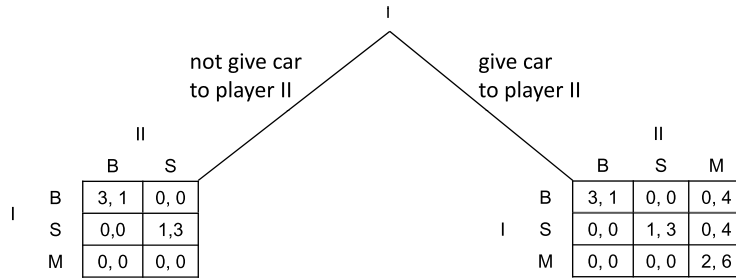
Fig. 1. Modified battle-of-the-sexes game without unawareness.

unawareness of an action are distinct from the unavailability of the same action. To see this, consider the following standard "battle-of-the-sexes" game (where Bach and Stravinsky concerts are the two available choices for each player)

		II	
		B	S
I	B	3, 1	0, 0
	S	0, 0	1, 3

augmented by a dominant Mozart concert for player II:

		II		
		B	S	M
I	B	3, 1	0, 0	0, 4
	S	0, 0	1, 3	0, 4
	M	0, 0	0, 0	2, 6

The new game is dominance solvable, and (M, M) is the unique Nash equilibrium.

Suppose that the Mozart concert is in a distant town, and II can go there only if player I gives him her car in the first place (see Fig. 1): Here, if player I doesn't give the car to player II, player II may conclude by forward induction that player I would go to the Bach concert with the hope of getting the payoff 3 (because by giving the car to II, player I could have achieved the payoff 2). The best reply of player II is to follow suit and attend the Bach concert as well. Hence, in the unique extensive-form rationalizable outcome, player I is not to give the car to player II and to go to the Bach concert.[1]

But what if, instead, the Mozart concert is in town but player II is initially unaware of the Mozart concert, while player I can enable player II to go to the concert simply by telling him about it? If player II remains unaware of the Mozart concert, then neither does he conceive that player I could have told him about the Mozart concert, and in particular he cannot carry out any forward-induction calculation. For him, the game is a standard battle-of-the-sexes game, where both actions of player I are rationalizable. This strategic situation is depicted in Fig. 2.

The strategic situation is not a standard extensive-form game (more on this in Section 2.6 below). If player I chooses not to tell player II about the Mozart concert, then player II's information set (depicted in blue) consists of a node in a simpler game — namely the one-shot battle-of-the-sexes with no preceding move by player I.

This is a simple example of the general novel framework that we define in Section 2 for dynamic interaction with possibly mutual unawareness of actions, generalizing standard extensive-form games. The framework will not only allow modeling of situations in which one player is certain that another player is unaware of portions of the game tree, as in the above example, but also of situations in which a player is uncertain regarding the way another player views the game tree, as well as situations in which the player is uncertain regarding the uncertainties of the other player about yet other players' views of the game tree, and so forth.

In fact, this framework allows not just for unawareness but also for other forms of misconception about the structure of the game that are especially relevant for modeling awareness of unawareness. Section 2.5 specifies further properties obtaining in generalized extensive-form games where the only source of players 'misconception' is unawareness and mutual unawareness of available actions and paths in the game. Since we focus on this type of unawareness, most of the examples in the paper satisfy the further properties specified in Section 2.5. Nevertheless, modeling *awareness of unawareness* does require the general framework in Section 2, as explained at its end.

---

[1] For a discussion of forward induction in battle-of-the-sexes games, see van Damme (1989).
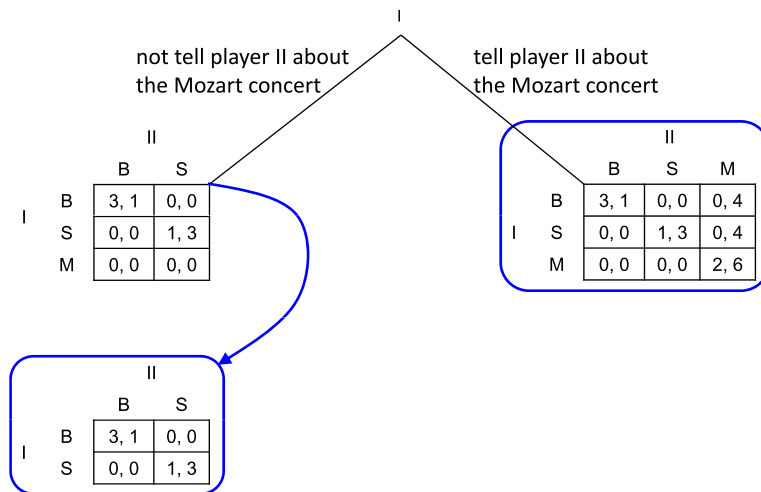
Fig. 2. Modified battle-of-the-sexes with unawareness. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

In this new framework, for each information set of a player her strategy specifies — from the point of view of the modeler — what the player would do if and when that information set of hers is ever reached. In this sense, a player does not necessarily 'own' her full strategy at the beginning of the game, because she might not be initially aware of all of her information sets. That's why a sensible generalization of Pearce's (1984) notion of extensive-form rationalizability is non-trivial. In Section 3 we put forward a modified definition, prove existence, and show the sense in which it coincides with extensive-form rationalizability in standard extensive-form games.

We focus here on a rationalizability solution concept rather than on some notion of equilibrium. While an equilibrium is ideally interpreted as a rest-point of some dynamic learning or adaptation process, or alternatively as a pre-meditated agreement or expectation, we find it difficult to carry over such interpretations to a setting in which every increase of awareness is by definition a shock or a surprise. Once a player's view of the game itself is challenged in the course of play, it is hard to justify the idea that a convention or an agreement for the continuation of the game are readily available.

We chose to focus on extensive-form rationalizability because it embodies forward-induction reasoning. If an opponent makes a player aware of some relevant aspect of reality, it is implausible to dismiss the increased level of awareness as an unintended consequence of the opponent's behavior. Rather, the player should try to rationalize the opponent's choice, re-interpret the opponent's past behavior, and try to infer from it the opponent's future moves. Extensive-form rationalizability indeed captures a 'best rationalization principle' (Battigalli, 1996).

With rationalizability, generalized games are necessary for properly modeling unawareness; trying to model unawareness by having the unaware player assigning probability zero to the contingency of which she is unaware might give rise to a completely different rationalizable behavior, which does not square with unawareness in the proper sense of the word. To see this consider the following example.

A Decision Maker (DM) has to choose between two policies, a_0 and a_1. Before choosing she gets a recommendation from an expert via a narrow communication channel, through which the expert can recommend either "0" or "1". The expert makes the recommendation after observing the state of nature, which may be either γ_0 or γ_1, and which the DM does not see. The interests of the expert and the DM are completely aligned: They each bear a cost of 1 if a_1 is implemented when the state of nature is γ_0 or vice versa. The expert furthermore bears a cost of 10 from "lying", i.e., from recommending "0" when the state of nature is γ_1 or recommending "1" when the state of nature is γ_0.

Assume the DM is aware only of the state γ_0 and unaware of γ_1. The dynamic interaction is hence modeled by the generalized game in Fig. 3. In this generalized game the only extensive-form rationalizable strategy of the DM is to always implement the policy a_0: she does not conceive of a contingency that would make the policy a_1 superior to a_0 even if she hears from the expert the recommendation "1"; in such a case she regrettably concludes that the expert behaved in an irrational way and bore the cost of "lying".

However, if we were to model the DM alternatively as being aware of γ_1 but assigning probability zero to it, the strategic interaction would be modeled by the standard game in Fig. 4.

In this game the unique extensive-form rationalizable strategy of the DM is to choose a_0 upon hearing "0" from the expert, but to implement a_1 upon hearing the recommendation "1". Indeed, extensive-form rationalizability requires the DM to base her choice on a system of beliefs about the expert's strategies with which at every information set of hers she maintains a belief that best rationalizes the choices of the expert which could have led to that information set. In particular, upon hearing the recommendation "1" from the expert, the only way for the DM to rationalize it is to assume that the state of nature is *nevertheless* γ_1, where recommending "1" is strictly dominant for the expert; and in γ_1 the optimal choice for the DM is a_1.
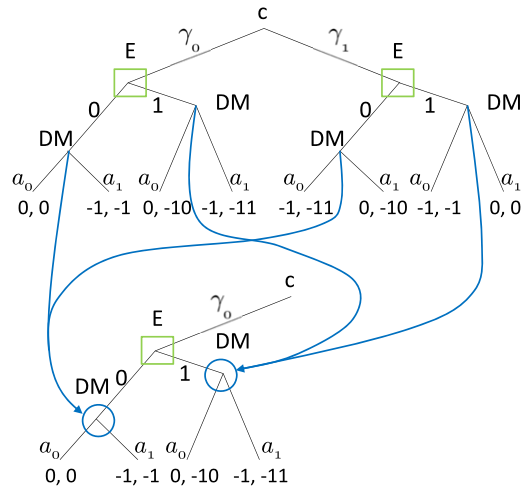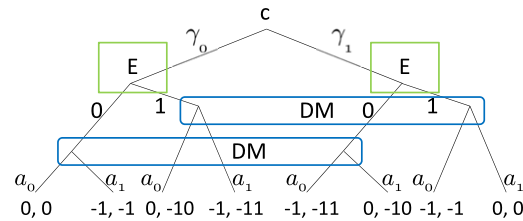
Fig. 3. A game with unawareness.

Fig. 4. A standard game.
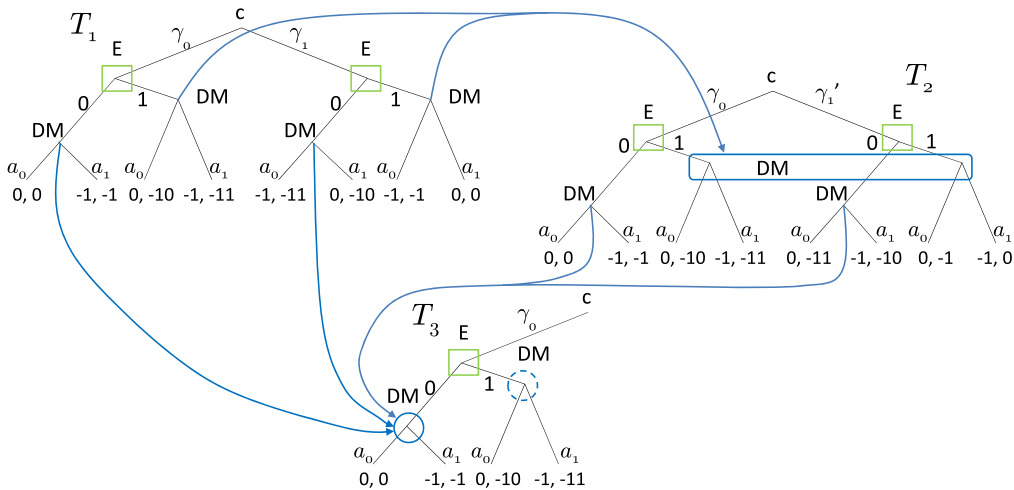
Fig. 5. A game with unawareness and misconceptions.

Conceptually, upon hearing the surprising recommendation "1" both choices of the DM have their internal logic. The former gives priority to "only γ_0 is conceivable", the latter to the rationality of the expert. But in the latter case, if initially the DM is *genuinely* unaware of γ_1, there is no reason why the DM would conceive *precisely* of γ_1 and not of some alternative description γ_1' of nature that would also rationalize the expert's recommendation "1"; some such conceptualizations γ_1' need not necessarily induce the DM to adopt the expert's recommendation. Generalized games lend themselves also to modeling such misconceptions that may arise upon a surprise, as demonstrated in Fig. 5. Here, the DM's rationalizable strategy is to choose a_0 also upon hearing the (surprising) recommendation "1", because the DM believes this recommendation was strictly dominant for the expert but that her interest and those of the expert are actually opposed. Such a subjective, erroneous conception of the strategic interaction by the DM is represented by the tree on the right-hand side of Fig. 5,

where payoffs at some of the terminal nodes are different than those in the objective, modeler's game which is on the upper left side of the figure.

1.1. Related literature

Our framework for dynamic interaction under unawareness seems to be significantly simpler than the one proposed by Halpern and Rêgo (2006) and Rêgo and Halpern (2012), in which they investigated the notions of Nash and sequential equilibrium, respectively. The simplification obtained in our framework is due to the fact that our initial building block is a tree representing *physical* moves, with information sets defined *only* in the subtrees which represent subjective views of the game (and subjective views thereof, etc.); in contrast, Halpern and Rêgo (2006) had information sets defined already in their basic tree. As a result, Halpern and Rêgo (2006) had to postulate additional conditions relating the information sets in subtrees to those of the basic tree.

Our framework is also more parsimonious than the one proposed by Feinberg (2009). Feinberg (2009) defined unawareness with analogous properties both for static and dynamic games, by explicit *unbounded sequences* of mutual "views" of the game. In his dynamic setting, a view is identified with a decision node. This means that even a standard extensive-form game has to be described in Feinberg (2009) by infinitely many copies of the same game tree, specifying explicitly how each player views the game in each of her decision nodes; how each player views, in each of her decision node, the way in which each other player views the game in each of their decision nodes; etc. In our setting such an infinite replication is avoided, since our definition of information sets encapsulates, by its design, all these mutual points of view.[2]

Li (2006) considered a model for dynamic unawareness with perfect information, in which at each decision node a player may have a subjective view of the game tree. Her model is more restricted than ours, since it requires there to be one particular default path of which all players are commonly aware, and since it does not allow for imperfect information.

Ozbay (2007) studied sender–receiver games, in which an 'announcer' can make an unaware decision maker aware of more states of nature before the decision maker takes an action. Such games can also be naturally formulated as a particular instance of our framework. For these games Ozbay studied an equilibrium notion incorporating forward-induction reasoning. Filiz-Ozbay (forthcoming) studied a related setting in which the aware announcer is a risk neutral insurer, while the decision maker is a risk averse insuree. At equilibrium, the insurer does not always reveal all relevant contingencies to the insuree.

Schipper and Woo (2012) use our framework to study political awareness of voters in modern electoral campaigning. Ideologically motivated candidates with a wealth of information about individual voters and sophisticated campaign strategies are faced by voters who lack awareness of some political issues and are uncertain about the exact political positions of candidates. They show that competition among candidates and microtargeting voters with issues and information yield prudent rationalizable election outcomes as if voters have full awareness of issues and complete information of candidates' political positions. They also show political competition is necessary for the result to obtain if voters are initially unaware of some political issues while it is not necessary under just standard uncertainty.[3]

Our aim is to provide a general framework for modeling misperceptions about the availability of actions in dynamic strategic situations. Different kinds of perception biases among players in games have been a popular topic in the recent literature on behavioral game theory. For instance, in static games Eyster and Rabin (2005) analyze players with correct conjectures about opponents' actions but misperceptions about how those opponents' actions are correlated with the opponents' information. In multi-stage games with moves of nature, Jehiel (2005) studies players that bundle nodes at which other players choose into "analogy classes", correctly anticipate the average behavior for each analogy class, and thus may have misperceptions about how others' behavior is related to others' information. Recently there has been a renaissance of non-equilibrium iterative solution concepts in behavioral game theory like level-*k* thinking and related models (e.g. Stahl and Wilson, 1995; Camerer et al., 2004; Crawford and Iriberri, 2007). Note that our iterative solution concept, extensive-form rationalizability, does not only provide behavioral predictions in the limit but also at every finite level of rationalization.

2. Generalized extensive-form games

To define a generalized extensive-form game Γ, consider first, as a building block, a finite game with perfect information and simultaneous moves[4] with a set of players I, a set of decision nodes N_0, active players I_n at node n with finite action sets A_n^i of player $i \in I_n$ (for $n \in N_0$), chance nodes C_0, and terminal nodes Z_0 with a payoff vector $(p_i^z)_{i \in I} \in \mathbb{R}^I$ for the players for every $z \in Z_0$. The nodes $\bar{N}_0 = N_0 \cup C_0 \cup Z_0$ constitute a tree, i.e., they are partially ordered by a precedence relation $\lessdot$ with which $(\bar{N}_0, \lessdot)$ is an arborescence (that is, the predecessors of each node in $\bar{N}_0$ are totally ordered by $\lessdot$), for each decision node $n \in N_0$ there is a bijection ψ_n between the action profiles $\prod_{i \in I_n} A_n^i$ at n and n's immediate successors, and there is a unique node in $\bar{N}_0$ with no predecessors — the root of the tree.

---

[2] Another important difference is that Feinberg (2009) does not define perfect recall, and this might hamper the extension of known solution concepts such as sequential equilibrium or extensive-form rationalizability that rely on perfect recall. Extensive-form rationalizability is the focal solution concept that we extend, define, and analyze in our paper, and to this effect we extend the definition of perfect recall to our setting.

[3] Currently we are unaware of further papers focusing directly and explicitly on dynamic games with unawareness. The literature on unawareness in general is growing fast — see e.g. http://www.econ.ucdavis.edu/faculty/schipper/unaw.htm.

[4] In allowing for simultaneous moves, we follow here Osborne and Rubinstein (1994) and Dubey and Kaneko (1984).

2.1. Partially ordered set of trees

Consider now a family **T** of subtrees of $\bar{N}_0$. A subtree is defined by a subset of nodes $\bar{N}'_0 \subseteq \bar{N}_0$ for which $(\bar{N}'_0, \prec)$ is also a tree (i.e., an arborescence in which a unique node has no predecessors). For two subtrees $T', T'' \in \mathbf{T}$ we write

$$T' \preccurlyeq T''$$

to signify that the nodes of T' constitute a subset of the nodes of T''.

One of the trees $T_1 \in \mathbf{T}$ is meant to represent the modeler's view of the paths of play that are *objectively* feasible.[5] Each other tree $T \in \mathbf{T}$ represents the feasible paths of play as *subjectively* viewed by some player at some node in T_1; or as the frame of mind attributed to the player at some node of T_1 by another player (or even by the same player at a later stage of the game, after her awareness regarding the feasible paths has evolved), whose own frame of mind regarding the feasible paths is represented by yet another $T' \in \mathbf{T}$; and so forth.

Denote by N_i^T the set of nodes in which player $i \in I$ is active in the tree $T \in \mathbf{T}$, and by $N_i = \bigcup_{T \in \mathbf{T}} N_i^T$.

We require three properties:

1. All the terminal nodes in each tree $T \in \mathbf{T}$ are in Z_0.
2. For every tree $T \in \mathbf{T}$, every node $n \in T$, and every active player $i \in I_n$ there exists a non-empty subset of actions $A_n^{i,T} \subseteq A_n^i$ such that ψ_n maps the action profiles $A_n^T = \prod_{i \in I_n} A_n^{i,T}$ bijectively onto n's successors in T.
3. If for two decision nodes $n, n' \in N_i^T$ (i.e., $i \in I_n \cap I_{n'}$) it is the case that $A_n^i \cap A_{n'}^i \neq \emptyset$, then $A_n^i = A_{n'}^i$.[6]

Property 1 is needed to ensure that each terminal node of each tree $T \in \mathbf{T}$ is associated with well-defined payoffs to the players. Property 2 means that at every node $n \in T$ the actions available to each active player $i \in I_n$ are independent of the actions the other active players choose at n (and hence $A_n^T = \prod_{i \in I_n} A_n^{i,T}$ is a product set — as in any standard dynamic game with simultaneous moves[7]). Property 3 means that i's active nodes N_i^T are partitioned into equivalence classes, such that the actions available to player i are identical within each equivalence class and disjoint in distinct equivalence classes. It will be needed for the definition of information sets which follows shortly.[8]

Within the family **T** of subtrees of $\bar{N}_0$, some nodes n appear in several trees $T \in \mathbf{T}$. In what follows, we will need to designate explicitly these different appearances of such nodes n as distinct entities. To this effect, in each tree $T \in \mathbf{T}$ label by n_T the copy in T of the node $n \in \bar{N}_0$ whenever the copy of n is part of the tree T, with the caveat that if the move $a_n \in A_n^T$ leads from n to n', then a_n leads also from the copy n_T to the copy n'_T. Denote by N the union of all decision nodes in all trees $T \in \mathbf{T}$, by C the union of all chance nodes, by Z the union of terminal nodes, and by $\bar{N} = N \cup C \cup Z$ (copies n_T of a given node n in different subtrees T are distinct from one another, so that $\bar{N}$ is a disjoint union of sets of nodes).

In what follows, when referring to a node in $\bar{N}$ we will typically avoid the subscript T when no confusion may arise. For a node $n \in \bar{N}$ we denote by T_n the tree containing n.

2.2. Information sets

In standard extensive-form games, an information set $\pi_i(n)$ of a player i is both (1) the set of nodes that the player considers as possible at n, and (2) the set of nodes in which the player has the same state of mind as in the nodes which she considers as possible at n.

In generalized games the two notions need not coincide: at a node n of the tree $T_n \in \mathbf{T}$, the player may conceive the feasible paths to be described by a different tree $T' \in \mathbf{T}$, and in particular she may conceive nodes $\pi_i(n)$ which are currently possible in her mind to be a subset of T' rather than of T_n; in such a case n will not be contained in $\pi_i(n)$. The information set $\pi_i(n)$ thus generalizes (1) above; the set of nodes (2) at which the player conceives $\pi_i(n)$ to be possible may include additional nodes which belong to trees outside the tree T' containing $\pi_i(n)$.[9]

Formally, for each decision node $n \in N$, define for each active player $i \in I_n$ a non-empty information set $\pi_i(n)$ with the following properties:

---

[5] In generalized extensive-form games modeling unawareness (see Section 2.5 below), T_1 will coincide with $\bar{N}_0$. In more general applications including delusion (like in the game of Fig. 5 above) or awareness of unawareness (see Section 2.6 below) $\bar{N}_0$ may include additional nodes not in T_1. In such a case $\bar{N}_0$ need not be one of the trees in **T**.

[6] Sometimes the modeler may want to impose an additional property: If in a subtree $T'' \in \mathbf{T}$ the probabilities of reaching $\bar{n}_1, \ldots, \bar{n}_k \in \bar{N}$ from the chance node $c \in C$ are $p_c^{\bar{n}_1} > 0, \ldots, p_c^{\bar{n}_k} > 0$ but some of these nodes do not appear in a subtree $T' \preccurlyeq T''$, then the probabilities of reaching the remaining nodes emanating from c are renormalized in so as to sum to 1 in T'. We do not impose this property here since it may be natural in some contexts but unnatural in others.

[7] See for instance, Osborne and Rubinstein (1994, p. 102, Section 6.3.2).

[8] The idea will be that in a given tree T, each action will correspond only to one view the player can have regarding the way the dynamic interaction has evolved that far, and will hence be available at (all the nodes of) a unique information set.

[9] For example, in the game of Fig. 3, for each node n in the upper tree in which the DM moves, the DM's information set $\pi_i(n)$ is a (singleton) subset of the nodes of the lower tree.

I0 Confinement: $\pi_i(n) \subseteq T$ for some tree T.

I1 No-delusion given the awareness level: If $\pi_i(n) \subseteq T_n$, then $n \in \pi_i(n)$.

I2 Introspection: If $n' \in \pi_i(n)$, then $\pi_i(n') = \pi_i(n)$.

I3 No divining of currently unimaginable paths, no expectation to forget currently conceivable paths: If $n' \in \pi_i(n) \subseteq T'$ (where $T' \in \mathbf{T}$ is a tree) and there is a path $n', \ldots, n'' \in T'$ such that $i \in I_{n'} \cap I_{n''}$, then $\pi_i(n'') \subseteq T'$.

I4 No imaginary actions: If $n' \in \pi_i(n)$, then $A_{n'}^i \subseteq A_n^i$.

I5 Distinct action names in disjoint information sets: For a subtree T, if $n, n' \in T$ and $A_n^i = A_{n'}^i$, then $\pi_i(n') = \pi_i(n)$.

I6 Perfect recall: Suppose that player i is active in two distinct nodes n_1 and n_k, and there is a path $n_1, n_2, \ldots, n_k$ such that at n_1 player i takes the action a_i. If $n' \in \pi_i(n_k)$, then there exist a node $n'_1 \neq n'$ and a path $n'_1, n'_2, \ldots, n'_\ell = n'$ such that $\pi_i(n'_1) = \pi_i(n_1)$ and at n'_1 player i takes the action a_i.

The following figures (Fig. 6) illustrate properties I0 to I6.

Properties I1, I2, I4, and I5 are standard for extensive-form games, and properties I0 and I6 generalize other standard properties of extensive-form games to our generalized setting. The essential new property is I3. At each information set of a player, property I3 confines the player's *anticipation* of her future view of the game to the view she currently holds (even if, as a matter of fact, this anticipation is about to be shattered as the game evolves).

We denote by H_i the set of i's information sets in all trees. For an information set $h_i \in H_i$, we denote by T_{h_i} the tree containing h_i. For two information sets h_i, h'_i in a given tree T, we say that h_i precedes h'_i (or that h'_i succeeds h_i) if for every $n' \in h'_i$ there is a path $n, \ldots, n'$ in T such that $n \in h_i$. We denote $h_i \rightsquigarrow h'_i$.

Remark 1. The following property is implied by I2 and I4: If $n', n'' \in h_i$ where $h_i = \pi_i(n)$ is an information set, then $A_{n'}^i = A_{n''}^i$.

Proof. If $n', n'' \in h_i$ where $h_i = \pi_i(n)$ is some information set, then by introspection I2 we must have $\pi_i(n') = \pi_i(n'') = \pi_i(n)$. Hence by I4 $A_{n'}^i \subseteq A_{n''}^i$ and $A_{n''}^i \subseteq A_{n'}^i$. □

If $n \in h_i$ we write also A_{h_i} for A_n^i.

Remark 2. Properties I0, I1, I2, and I6 imply no absent-mindedness: No information set h_i contains two distinct nodes n, n' on some path in some tree.

Proof. Suppose by contradiction that there exists an information set h_i with a node $n \in h_i$ such that some other node in h_i precedes n in the tree T_n. Denote by n' the first node on the path from the root to n that is also in h_i. By I1 we have $n' \in \pi_i(n') = h_i = \pi_i(n)$, and by perfect recall I6 there exists a path $n'' = n'_1, \ldots, n'_\ell = n'$, such that at n'' player i had the same state of mind as in n', i.e., $\pi_i(n'') = \pi_i(n')$. By I1, we have $n'' \in \pi_i(n'') = \pi_i(n') = h_i$ and n'' is a predecessor of n', a contradiction. □

The perfect-recall property I6 and Remark 2 guarantee that with the precedence relation $\rightsquigarrow$ player i's information sets H_i form an arborescence: For every information set $h'_i \in H_i$, the information sets preceding it $\{h_i \in H_i : h_i \rightsquigarrow h'_i\}$ are totally ordered by $\rightsquigarrow$.

For trees $T, T' \in \mathbf{T}$ we denote $T \rightarrowtail T'$ whenever for some node $n \in T$ and some player $i \in I_n$ it is the case that $\pi_i(n) \subseteq T'$. Denote by $\hookrightarrow$ the transitive closure of $\rightarrowtail$. That is, $T \hookrightarrow T''$ if and only if there is a sequence of trees $T, T', \ldots, T'' \in \mathbf{T}$ satisfying $T \rightarrowtail T' \rightarrowtail \cdots \rightarrowtail T''$.

2.3. Generalized games

A generalized extensive-form game Γ consists of a partially ordered set $\mathbf{T}$ of subtrees of a tree $\bar{N}_0$ satisfying properties 1–3 above, along with information sets $\pi_i(n)$ for every $n \in T$, $T \in \mathbf{T}$ and $i \in I_n$, satisfying properties I0–I6 above.

For every tree $T \in \mathbf{T}$, the T-*partial game* is the partially ordered set of trees including T and all trees T' in Γ satisfying $T \hookrightarrow T'$, with information sets as defined in Γ. A T-partial game is a generalized game, i.e., it satisfies all properties 1–3 and I0–I6.

We denote by H_i^T the set of i's information sets in the T-partial game.

For example, for the generalized game Γ in Fig. 5, the tree $\bar{N}_0$ appears below in Fig. 7. $\bar{N}_0$ starts by nature choosing between γ_0 (following which a_0 is the optimal action for both players), γ_1 (following which a_1 is optimal for both) or γ'_1 (following which a_0 is optimal for the Decision Maker (DM) but suboptimal for the Expert (E)).

However, in the generalized game Γ of Fig. 5 the tree which represents the physical paths of the game is not $\bar{N}_0$ but rather T_1, which appears in the upper left part of Fig. 5. Moreover, $\bar{N}_0$ is none of the trees in $\mathbf{T} = \{T_1, T_2, T_3\}$ of Γ, and hence $\bar{N}_0$ does not appear in Fig. 5.

In T_1 if the Expert, after learning nature's move, announces '0', the DM is unaware of the fact that nature could have chosen γ_1. The DM's frame of mind is hence represented by the tree T_3 at the bottom of Fig. 5. According to the DM's
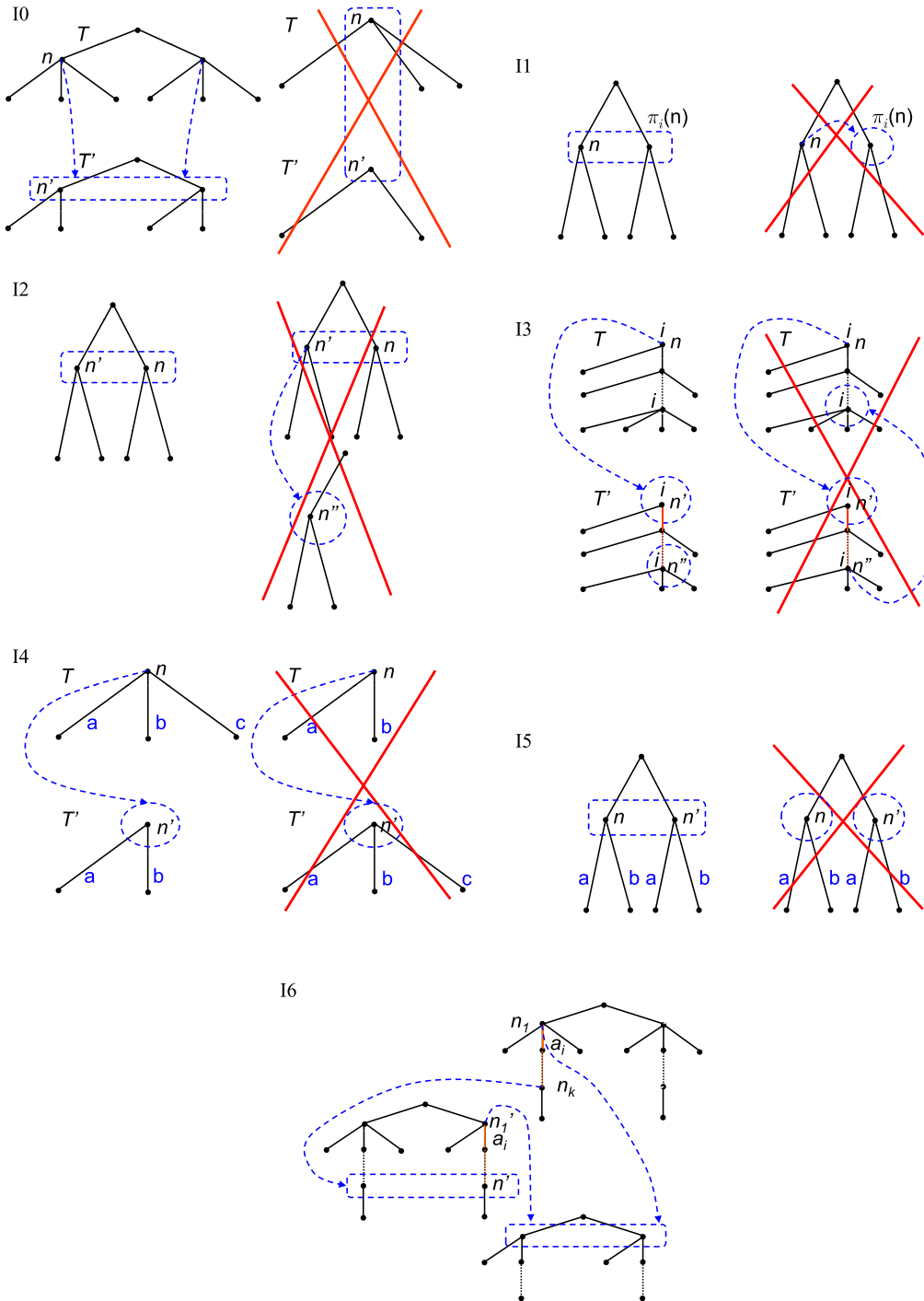
Fig. 6. Properties I0 to I6.

conception there she does not miss any information, and her information set is a singleton. This single node is also the unique node considered as possible by the DM at T_1 after hearing the message '0' (and hence the two arrows going from T_1 downwards to that singleton in T_3). Moreover, under this frame of mind T_3, by which only γ_0 is feasible, the DM conceives that she would have a singleton information set in T_3 even if E were to announce '1'.

The truth of the matter, as portrayed in the tree of physical paths T_1, is different. If E announces '1', the DM gets to believe that nature could have chosen not between γ_0 and γ_1 (as was actually the case) but rather between γ_0 and γ_1'. This is portrayed by the fact that after hearing '1' in T_1, the nodes that the DM considers as possible are in T_2, the tree in the
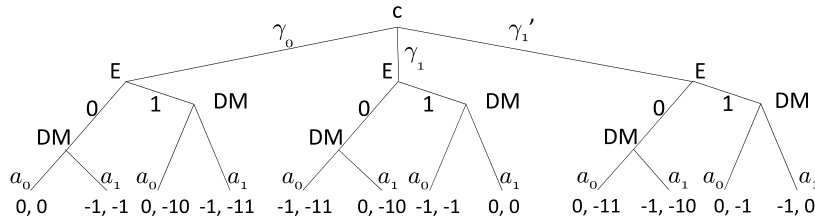
Fig. 7. The tree $\bar{N}_0$ for the generalized game in Fig. 5.

upper right part of Fig. 5. The information set of the DM in T_2 after hearing '1' contains two nodes, corresponding to the two possible choices of nature that the DM considers as possible there, γ_0 and γ_1'.

Moreover, in this state of mind the DM is surprised, in the sense that she realizes that had she heard '0' from the Expert, she would not suspect that nature could choose anything beyond γ_0. This is reflected by the fact that in T_2, the information set of the DM had she heard '0' is contained in T_3.

The T_1-partial game is the entire game Γ with the set of trees $\{T_1, T_2, T_3\}$. The T_2-partial game consists only of the trees $\{T_2, T_3\}$. The T_3-partial game is a standard extensive-form game with the unique tree T_3.

2.4. Strategies

A (pure) strategy

$$s_i \in S_i := \prod_{h_i \in H_i} A_{h_i}$$

for player i specifies an action of player i at each of her information sets $h_i \in H_i$.[10] Denote by

$$S = \prod_{j \in I} S_j$$

the set of strategy profiles in the generalized extensive-form game.

If $s_i = (a_{h_i})_{h_i \in H_i} \in S_i$, we denote by

$$s_i(h_i) = a_{h_i}$$

the player's action at the information set h_i.

With the strategy s_i, at node $n \in N_i$ define the player's action at n to be $s_i(\pi_i(n))$. Thus, the strategy s_i specifies what player i does at each of her active nodes $n \in N_i$, both in case $n \in \pi_i(n)$ and in case $\pi_i(n)$ is a subset of nodes of a tree which is distinct from the tree T_n to which n belongs.

In a generalized game Γ only the tree $T_1 \in \mathbf{T}$ represents the physical paths in the game; every other tree in $\mathbf{T}$ represents the subjective view of the feasible paths in the mind of a player, or the view of the feasible paths that a player believes that another player may have in mind, etc. Moreover, as the actual game in T_1 evolves, a player may become aware of paths of which she was unaware earlier, and the way she views the game may alter as well.

Thus, in generalized extensive-form games, a strategy cannot be conceived as an ex ante plan of action. Formally, a strategy s_i of player i is a list of answers to the questions "what would the player do if h_i were the set of nodes she considered as possible?", for $h_i \in H_i$. This list of answers should be interpreted as follows. For every *given* frame of mind $T \in \mathbf{T}$ that player i may entertain about the feasible paths (a frame of mind which i actually has at some node in the actual game tree T_1, or is *attributed* to i by player j at some node of T_1 [at which j's frame of mind may be yet in another tree $T' \in \mathbf{T}$], etc.),

1. for every information set $\pi_i(n) \subseteq T$ the action $s_i(\pi_i(n))$ should be interpreted as the action that player i actually takes at n under the strategy s_i, if and when n is reached; and
2. for every information set $\pi_i(n'') \subseteq T'' \neq T$ the action $s_i(\pi_i(n''))$ should be interpreted as the action that player i would have taken at n'' if her frame of mind were T'' rather than T. This means that when player j considers as possible that the node n'' can be reached, j *believes* that under the strategy s_i player i would take the action $s_i(\pi_i(n''))$ at n'', if and when n'' were reached.

---

[10] The set of *behavioral* strategies is

$$\prod_{h_i \in H_i} \Delta(A_{h_i}).$$

For a strategy $s_i \in S_i$ and a tree $T \in \mathbf{T}$, we denote by s_i^T the strategy in the T-partial game induced by s_i (i.e., $s_i^T(h_i) = s_i(h_i)$ for every information set h_i of player i in the T-partial game). If $R_i \subseteq S_i$ is some set of strategies of player i, denote by R_i^T the set of strategies induced by R_i in the T-partial game. The set of i's strategies in the T-partial game is thus denoted by S_i^T. Denote by $S^T = \prod_{j \in I} S_j^T$ the set of strategy profiles in the T-partial game.

We say that a strategy profile $s = (s_j)_{j \in I} \in S$ *reaches a node* $n \in T$ if the players' actions $s_j^T(\pi_j(n'))_{j \in I_{n'}}$ and nature's moves in the nodes $n' \in T$ lead to n with a positive probability. Notice that by property I4 ("no imaginary actions"), $s_j^T(\pi_j(n'))_{j \in I}$ is indeed well defined: even if $\pi_j(n') \notin T$ for some $n' \in T$, the action profile $s_j^T(\pi_j(n'))_{j \in I_{n'}}$ is an action profile which is actually available in T to the active players $j \in I_{n'}$ at n'.

We say that a strategy profile $s \in S$ *reaches* the information set $h_i \in H_i$ if s reaches some node $n \in h_i$.

We say that the strategy $s_i \in S_i$ *reaches* the information set h_i if there is a strategy profile $s_{-i} \in S_{-i}$ of the other players such that the strategy profile (s_i, s_{-i}) reaches h_i. Otherwise, we say that the information set h_i is *excluded* by the strategy s_i.

Similarly, we say that the strategy profile $s_{-i} \in S_{-i}$ *reaches* the information set h_i if there exists a strategy $s_i \in S_i$ such that the strategy profile (s_i, s_{-i}) reaches h_i.

As is the case also in standard games, for every given node, a given strategy profile of the players induces a distribution over terminal nodes in each tree, and hence an expected payoff for each player in the tree.

To exemplify the above definitions, consider again the game of Fig. 5. The Expert (E) has two information sets in T_1, two information sets in T_2, and one information set in T_3. The following therefore describes a strategy s_E of the Expert: '0' after γ_0 in T_1; '1' after γ_1 in T_1; '0' after γ_0 in T_2; '1' after γ_1' in T_2; '0' after γ_0 in T_3.

What about the DM? She has no information sets contained in T_1 (!), one information set in T_2 (with two nodes in it) and two singleton information sets in T_3. The following therefore describes a strategy s_{DM} of the DM: play a_0 in every information set.

The profile of strategies (s_E, s_{DM}) described above induces the following paths in T_1:

1. If nature chooses γ_0, the path is $(\gamma_0, '0', a_0)$. Notice that after γ_0, '0' in T_1, we read the choice of the DM by following the arrow that leads to her information set down in T_3, and check her choice with the strategy s_E there.
2. If nature chooses γ_1, the path is $(\gamma_1, '1', a_0)$. After γ_1, '1' in T_1, we read the choice of the DM by following the arrow that leads to her information set, which this time is in T_2.

Observe that the Expert is never actually (in T_1) deluded to think that the strategic interaction is described by the T_2-partial game, nor is he ever actually (in T_1) unaware so as to think that the strategic interaction is described by T_3. Thus, the moves of s_E in T_2 (in particular after nature chooses γ_1', which can never actually happen in reality, i.e., in T_1) describe the DM's belief about the Expert's choice that led to her information set if she believes that the Expert is using the strategy $s_E^{T_2}$ — the restriction of the strategy s_E to the T_2-partial game. Under any solution concept we would need indeed to analyze what the DM believes at her information set in T_2 about the Expert's past actions that have led her to that information set, and those past moves are determined by a strategy of the expert in the T_2-partial game. This is the reason that we need to define a strategy at all the information sets of each player, including those in which he will never actually move: the latter parts of the strategy become the object of contemplation and analysis of the other player (or players) when they are deluded or unaware of parts of the actual game.

2.5. Unawareness

Generalized games can describe many types of games with subjective reasoning. In a generalized game, a player cannot imagine that she can take an action which is physically unavailable to her (property I4), but at a given information set $\pi_i(n)$ she can nevertheless imagine that in a succeeding information set she will have an action which is actually nowhere available in the tree T_n as in the example of Fig. 8. Furthermore, she can imagine that along the path of play another player will forget the history of play, i.e., that at a later information set this other player will imagine he is playing in a game tree which is completely unrelated to the game tree he imagined at an earlier stage along the path.

Since our main motivation is to analyze games with unawareness rather than games with arbitrary kinds of subjective reasoning, it is worthwhile spelling out additional properties of generalized games in which the only reason for players' misconception of the game is (differing) unawareness of available actions. In *extensive-form games with unawareness* the set of trees $\mathbf{T}$ forms a join semi-lattice under the inclusion partial order relation $\preccurlyeq$. The maximal tree in this join semi-lattice is the modeler's objective description of feasible paths of play.

The following additional properties parallel properties of static unawareness structures in Heifetz et al. (2006).[11]

U0 Confined awareness: If $n \in T$ and $i \in I_n$, then $\pi_i(n) \subseteq T'$ with $T' \preccurlyeq T$.
U1 Generalized reflexivity: If $T' \preccurlyeq T$, $n \in T$, $\pi_i(n) \subseteq T'$ and T' contains a copy $n_{T'}$ of n, then $n_{T'} \in \pi_i(n)$.
U2 Introspection: If $n' \in \pi_i(n)$, then $\pi_i(n') = \pi_i(n)$. (I.e., property I2.)

---

[11] The number of each property corresponds to the respective property in Heifetz et al. (2006).

U3 Subtrees preserve awareness: If $n \in T'$, $n \in \pi_i(n)$, $T \preccurlyeq T'$, and T contains a copy n_T of n, then $n_T \in \pi_i(n_T)$.

U4 Subtrees preserve ignorance: If $T \preccurlyeq T' \preccurlyeq T''$, $n \in T''$, $\pi_i(n) \subseteq T$ and T' contains the copy $n_{T'}$ of n, then $\pi_i(n_{T'}) = \pi_i(n)$.

U5 Subtrees preserve knowledge: If $T \preccurlyeq T' \preccurlyeq T''$, $n \in T''$, $\pi_i(n) \subseteq T'$ and T contains the copy n_T of n, then $\pi_i(n_T)$ consists of the copies that exist in T of the nodes of $\pi_i(n)$.

The following remark is analogous to Remark 3 in Heifetz et al. (2006).

Remark 3. U5 implies U3.

Proof. If $n \in T'$, $n \in \pi_i(n)$, $T \preccurlyeq T'$, and T contains a copy n_T of n, then by U5 $\pi_i(n_T)$ must consist of the copies that exist in T of the nodes of $\pi_i(n)$. Since by assumption $n \in \pi_i(n)$ and the copy n_T exists in T, we must have $n_T \in \pi_i(n_T)$. □

Remark 4. U0 implies I0. U1 implies I1.

Remark 5. U0 is equivalent to I0 and $T \rightarrowtail T'$ implies $T' \preccurlyeq T$.

Proof. I0 and $T \rightarrowtail T'$ implies $T' \preccurlyeq T$ are equivalent to if there exist $n \in T$ and $i \in I_n$ such that $\pi_i(n) \subseteq T'$ then $T' \preccurlyeq T$. □

All these properties are static properties in the sense that they relate nodes in one tree with copies of those nodes in another tree. One may wonder about dynamic properties of unawareness. The following property states that a player cannot become unaware during the play.

DA Awareness may only increase along a path: If there is a path $n, \ldots, n'$ in some subtree T'' such that player i is active in n and n', and $\pi_i(n) \subseteq T$ while $\pi_i(n') \subseteq T'$, then $T' \succcurlyeq T$.

Perfect recall and confined awareness imply that awareness may only increase along a path.

Remark 6. U0 and I6 imply DA.

Proof. Suppose $n, \ldots, n'$ is path in T'' such that $i \in I_n \cap I_{n'}$, $\pi_i(n) \subseteq T$ while $\pi_i(n') \subseteq T'$. Then by the non-emptyness of $\pi_i(n')$ there exists $n'' \in \pi_i(n')$. For this n'' there is by I6 a path $n'_1, \ldots, n'_\ell = n''$ in T' such that $\pi_i(n'_1) = \pi_i(n)$. By U0, we have $T \preccurlyeq T'$. □

2.6. Awareness of unawareness

In some strategic situations a player may be aware of her unawareness in the sense that she is suspicious that *something* is amiss without being able to conceptualize this 'something'. Such a suspicion may affect her payoff evaluations for actions that she knows are available to her. More importantly, she may take actions to investigate her suspicion if such actions are physically available.

To model awareness of unawareness some of the trees may include *imaginary actions* as placeholders for actions that a player may be unaware of and terminal nodes/evaluations of payoffs that reflect her awareness of unawareness. (The approach of modeling awareness of unawareness by "imaginary moves" was proposed by Halpern and Rêgo, 2006.)

Consider the example in Fig. 8. There are two players with a common interest. The payoffs at the terminal nodes in each tree correspond to the payoffs to each player. Player 1 is a professor of player 2, a doctoral student. The doctoral student has to choose a topic for his dissertation. He can conceive of two topics, topic a and topic b. This is modeled in the lowest tree, T_3. Before he selects his topic, the professor can intervene or keep silent. If she intervenes, she could simply suggest a topic to the student. Or, she could point the student to a book and suggest to him to read it in anticipation of finding a suitable topic. Before reading it, the student does not anticipate of what he will become aware upon reading it. Yet, he suspects that he will become aware of *something*. This is modeled in the upper left tree, T_2. Moreover, he realizes that he benefits more from reading the book than if the professor were to simply tell him this *something*. The payoffs upon reading are increased by 1. When the student reads the book, he becomes aware of two additional topics, one of them, D', is strictly worse than any previously conceived topic and one topic, C', is much more promising than any previously conceived topic. Moreover, he realizes that reading the book was actually more beneficial to him than anticipated before as all payoffs upon reading are increased by 2 in upper right tree, T_1. Note that before the student decides whether or not to read the book, he is not modeled as anticipating to be in the upper right tree T_1, because he cannot conceptualize the nature of dissertation topics he discovers upon reading the book.

As before, a strategy of a player ascribes an available action at each of his information sets in all trees. The strict dominant action of player 1 in both of her information sets is to suggest reading the book. (Recall that players have a common interest; the payoffs at terminal nodes are each player's payoffs.) Moreover, since the student realizes that reading
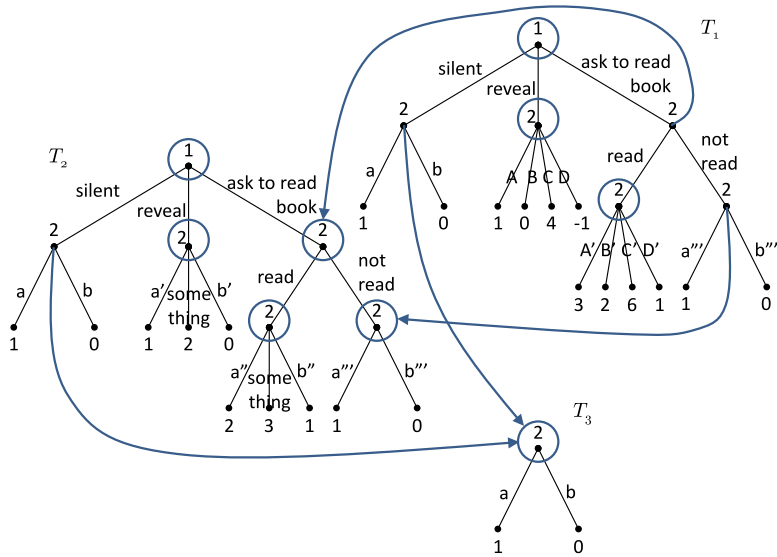
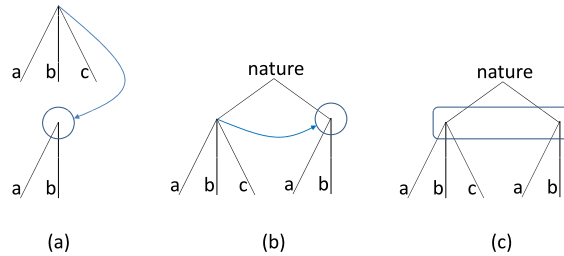**Fig. 8.** Game form with awareness of unawareness.

Fig. 9. Transformations of (a) into "standard" game forms (b) and (c).

is beneficial, the unique extensive-form rationalizable outcome corresponds to the terminal node in T_1 that is reached by the path in which the professor suggests reading the book, the student indeed reads, and after reading chooses topic C'.

The generalized extensive-form game in Fig. 8 does not satisfy all properties of unawareness of Section 2.5. For instance, at the node after player 1 asked player 2 to read the book in tree T_1, player 2's information set is in tree T_2. But T_2 is not a subtree of T_1. Thus, property U0 is violated. Modeling awareness of unawareness requires our more general framework of generalized extensive-form games.

2.7. The connection to standard extensive-form games

Harsanyi (1967) showed how to transform games with asymmetric information into games with imperfect information about a move of nature. Can a similar idea be used to transform any generalized extensive-form game into a standard extensive-form game? Given a generalized extensive-form game Γ with a partially ordered set of trees **T**, one could define the transformation of Γ to be the extensive-form game with an initial move of nature, in which nature chooses one of the trees in **T**.

Notice, however, that the resulting structure would not be a standard extensive-form game. To see this, notice that every standard extensive-form game has the following property (E): the equivalence class of nodes in which a player considers as possible a given possibility set of nodes is *identical* with that possibility set; this set is called an information set of the player, and in all of its nodes the player has the same set of available actions. In contrast, in the transformation considered above for games with misperceptions, this equivalence class may be a *strict super-set* of the possibility set. For example, when the generalized game in Fig. 9(a) is transformed so as to have an initial move of nature, the possibility set for the (unique) player is the right node, while the equivalence class contains both the right and left node.

Thus, if after adding the initial move of nature the information sets are defined to be synonymous with the possibility sets, the resulting game would be non-standard, because for some information set there may be additional nodes *outside it* in which the player considers it as possible (as in Fig. 9(b), where in the left node the player considers only the right node as possible). If, in contrast, we choose the alternative definition, by which an information set is the equivalence class in which a player has a particular set of nodes that she considers as possible, the resulting game would again be non-standard,

this time because the actions available to the player in the nodes of a given information set might not be identical across these nodes (as in Fig. 9(c), where in the left node the player has more available actions than in the right node, even though both are within the same information set).[12]

There is also another aspect that prevents the above transformation from yielding a standard extensive-form game. In a standard extensive-form game each player has a full-support prior on the moves of nature.[13] Using Bayes rule, the player therefore has a well-defined belief about nature at each stage of the game. In contrast, in the above transformation each player ascribes probability 1 only to one of the initial moves of nature; moreover, along the path of play the player may switch completely the move of nature in which she confides even if nothing in the path of play itself imposed such a switch. Such a switch corresponds to a node in the generalized game in which the player is **defined** as *becoming aware* of new aspects of the dynamic interaction; such an increase of awareness may occur even when the physical path of play *per se* did not imply a surprise, and may have also been compatible with the player's previous conception of the game. Thus, if we do add an initial move of nature to connect the trees of the generalized game, the player's (evolving) belief about nature cannot be encapsulated within an initial probabilistic belief about nature, and must be represented explicitly by a *belief system* as part of the *definition* of the game.

Adding an initial move of nature has a further conceptual drawback. In classical extensive-form games the implicit assumption is that the players understand the entire structure of the dynamic interaction as embodied in the game tree.[14] Assigning probability zero to some move of nature is still compatible with realizing what could have happened if this zero-probability move were nevertheless to materialize. This is conceptually distinct from being completely *unaware* of a subset of paths in the game, and it is the latter concept that we want to model here. Moreover, as we have seen in the example of the introduction (Figs. 3 and 4), it may lead to behavioral predictions different from unawareness.

Thus, standard extensive-form games are neither technically fit (without further generalization) for modeling behavior under dynamic misperceptions and unawareness, nor do they convey the appropriate conceptual apparatus for modeling such interactions, hence the need for our definition of generalized games.

3. Extensive-form rationalizability

Pearce (1984) defined extensive-form (correlated) rationalizable strategies by a procedure of an iterative elimination of strategies (see also Battigalli, 1997). The idea behind the definition involves a notion of forward induction. In generic perfect-information games, extensive-form rationalizable strategy profiles yield the backward induction outcome, though they need not be subgame-perfect equilibrium strategies (Reny, 1992; Battigalli, 1997; Chen and Micali, 2011; Perea, 2012). In what follows we extend this definition to generalized extensive-form games.

A *belief system* of player i

$$b_i = \left(b_i(h_i)\right)_{h_i \in H_i} \in \prod_{h_i \in H_i} \Delta\left(S_{-i}^{T_{h_i}}\right)$$

is a profile of beliefs — a belief $b_i(h_i) \in \Delta(S_{-i}^{T_{h_i}})$ about the other players' strategies in the T_{h_i}-partial game, for each information set $h_i \in H_i$, with the following properties:

- $b_i(h_i)$ reaches h_i, i.e., $b_i(h_i)$ assigns probability 1 to the set of strategy profiles of the other players that reach h_i.
- If h_i precedes h_i' ($h_i \rightsquigarrow h_i'$) then $b_i(h_i')$ is derived from $b_i(h_i)$ by Bayes rule whenever possible.

Denote by B_i the set of player i's belief systems.

For a belief system $b_i \in B_i$, a strategy $s_i \in S_i$ and an information set $h_i \in H_i$, define player i's expected payoff *at* h_i to be the expected payoff for player i in T_{h_i} given $b_i(h_i)$, the actions prescribed by s_i at h_i and its successors, assuming that h_i has been reached.

We say that with the belief system b_i and the strategy s_i player i is *sequentially rational* at the information set $h_i \in H_i$ if either s_i doesn't reach h_i or there exists no strategy s_i' which is distinct from s_i only at h_i and/or at some of h_i's successors in T_{h_i} and yields player i a higher expected payoff in the T_{h_i}-partial game given the belief $b_i(h_i)$ on the other players' strategies $S_{-i}^{T_{h_i}}$.

---

[12] In this example of a game with a single player who is unaware of her action c, one could obviously describe the game simply as a single-person decision problem between a and b. This would not be possible, however, in more complex games like the one in Fig. 2. There, one cannot do away with any of the nodes in the upper tree or in the lower tree; if these two trees are joined by a preceding move of nature, then when player 1 doesn't tell player 2 about the Mozart concert, player 2's information set becomes non-standard.

[13] Moreover, in the classical definition of an extensive-form game the priors of the different players about nature are actually identical, i.e., the players have a common prior about nature.

[14] For instance, Myerson (1991, p. 4) puts forward explicitly the tenet that game theory deals with *intelligent* players, where "a player in the game is *intelligent* if he knows everything that we know about the game and he can make any inference about the situation that we can make".

We now turn to define extensive-form rationalizability in generalized extensive-form games. The following definition generalizes Battigalli's (1997) definition of (correlated) extensive-form strategies, which he proved to be equivalent to that of Pearce (1984).

Definition 1 *(Extensive-form rationalizable strategies).* Define, inductively, the following sequence of belief systems and strategies of player i:

$$B_i^1 = B_i,$$

$$R_i^1 = \left\{ s_i \in S_i: \begin{array}{l} \text{there exists a belief system } b_i \in B_i^1 \text{ with which for every} \\ \text{information set } h_i \in H_i \text{ player } i \text{ is sequentially rational at } h_i \end{array} \right\},$$

$$\vdots$$

$$B_i^k = \left\{ b_i \in B_i^{k-1}: \begin{array}{l} \text{for every information set } h_i, \text{ if there exists some profile of} \\ \text{the other players' strategies } s_{-i} \in R_{-i}^{k-1} = \prod_{j \neq i} R_j^{k-1} \text{ such} \\ \text{that } s_{-i} \text{ reaches } h_i, \text{ then } b_i(h_i) \text{ assigns probability 1 to } R_{-i}^{k-1, T_{h_i}} \end{array} \right\},$$

$$R_i^k = \left\{ s_i \in S_i: \begin{array}{l} \text{there exists a belief system } b_i \in B_i^k \text{ with which for every} \\ \text{information set } h_i \in H_i \text{ player } i \text{ is sequentially rational at } h_i \end{array} \right\}.$$

The set of player i's **extensive-form rationalizable strategies** is

$$R_i^\infty = \bigcap_{k=1}^\infty R_i^k.$$

The definition captures rationality and common strong belief in rationality (Battigalli and Siniscalchi, 2002): At each information set, a rationalizable strategy should be optimal vis-a-vis some belief over the opponents strategy; if the information set is reached by some tuple of *optimal* opponents' strategies (vis-a-vis some beliefs of theirs), then the player's belief is further required to be concentrated on such tuples; if, furthermore, the information set is reached by some tuple of the opponents' strategies which are optimal vis-a-vis a belief system of theirs concentrated on optimal strategies of *their* opponents, the player's belief should be concentrated on *those* tuples; and so forth.

In other words, along each feasible path of play, in the first information set an active player believes that all her opponents will behave rationally, will believe that *their* opponents will behave rationally, etc. If at some information set in the game all the opponents' strategy profiles which could lead to that information set fail this ideal condition, the player seeks a *best rationalization* (Battigalli, 1996) which could have led to that information set.

For example, if player i has a unique opponent j, who has only two strategies that lead to an information set of i — s_j' which is strictly dominated for j, and s_j which is optimal for j but only under a belief of j that i is (or was, or will be) irrational, then at that information set i is required to believe that in the sequel j will continue to employ s_j (because s_j embodies a better rationalization of j's past behavior than does s_j'). Forward-induction reasoning then implies that from that information set onwards, i's rationalizable strategy should be optimal vis-a-vis s_j, unless a further information set h_i' is reached which is compatible only with s_j'; at h_i' player i has no choice but to revert to the belief that j is irrational, and react accordingly.

The definition of this solution concept for generalized extensive-form games highlights the need to define the notion of a strategy as we did, by the actions taken not only at the tree T_1 which represents the physical paths of the game, but also at all the other trees $T \in \mathbf{T}$. True, to track the physical paths compatible with profiles of extensive-form rationalizable strategies it is enough to look at their restrictions to T_1. However, at each given node $n \in T_1$ in which player i is active, the set of nodes $\pi_i(n)$ that she considers as possible is a subset of her subjective view of the feasible paths $T_{\pi_i(n)}$, and at that point she can only contemplate her strategy in terms of the $T_{\pi_i(n)}$-partial game. Furthermore, in order to rank the opponents' strategies according to their rationality, player i has to weigh them in the terms the opponents conceive the game, i.e., in the T-partial games which represent *their* subjective view of the strategic interaction *within* the $T_{\pi_i(n)}$-partial game (which may be different than the actual subjective views the opponents have on the game at various nodes of T_1); and so forth.

This means that profiles of extensive-form rationalizable strategies have a different significance in their different domains. In T_1 they define paths which could actually be realized; for $n \in T_1$ for which $T_{\pi_i(n)} \neq T_1$, in $T_{\pi_i(n)}$ these profiles define paths conceived as feasible by player i when the actual node at T_1 is n; for $n' \in T_{\pi_i(n)}$ for which $T_{\pi_j(n')} \neq T_{\pi_i(n)}$, in $T_{\pi_j(n')}$ these profiles define paths that at node $n \in T_1$ player i conceives player j to conceive as possible if and when n' is reached in i's subjective view of the game $T_{\pi_i(n)}$; etc.

Although extensive-form rationalizability is defined as a reduction procedure on the set of beliefs, it implies a reduction of strategy sets.

Remark 7. $R_i^k \subseteq R_i^{k-1}$ for every $k > 1$.

Proof. Consider $s_i \in R_i^k$. By definition, s_i is sequentially rational at each of player i's information sets given some belief system $b_i \in B_i^k$. Since $B_i^k \subseteq B_i^{k-1}$, s_i is also be sequentially rational at each of player i's information sets given a belief system in B_i^{k-1}, namely given b_i. Hence $s_i \in R_i^{k-1}$. □

Proposition 1. *The set of rationalizable strategies is non-empty.*

The proof is in Appendix A.

It may be instructive to compare explicitly the extensive-form rationalizability strategies in our battle-of-the-sexes example from the introduction (Figs. 1 and 2).[15]

Remark 8. In the Bach–Stravinsky–Mozart example with unavailability of actions from the introduction (Fig. 1) there is a unique extensive-form rationalizable outcome while in the Bach–Stravinsky–Mozart example with unawareness (Fig. 2), there are several extensive-form rationalizable outcomes.

The proof is contained in Appendix A.

When we compare these examples, then the main difference arises from the lack of forward induction of player II under unawareness. In the Bach–Stravinsky–Mozart example with unawareness (Fig. 2), player II cannot forward-induce anything from the action "don't tell" taken by player I since former is unaware of this action. Yet, in the Bach–Stravinsky–Mozart example with the unavailability of an action (Fig. 1) player II can forward-induce from the action "don't give the car" player I's intention to go to the Bach concert. In other words, awareness of an available action (providing the car for going to the Mozart concert) and certainty that it hasn't been taken has stronger strategic implications than unawareness of the very same action.

In the Bach–Stravinsky–Mozart example with unawareness (Fig. 2), the rationalizable outcome is not unique. This is in contrast to the example with unavailability of actions instead, where there is a rationalizable outcome. However, there exist also games where with unavailability of actions there are more rationalizable outcomes than with unawareness of the same actions. Such an example is presented in Meier and Schipper (2012).

3.1. The one-deviation principle

In dynamic decision problems, the one-deviation principle reflects a property of dynamic optimization, namely that a sequence of "locally optimal" decisions is also "globally optimal". Since in games with unawareness, both the relevant optimization horizon and the available set of actions may change during the play, it is not a priori immediate whether the one-deviation principle still holds for generalized extensive-form games. In this section we show that the one-deviation principle applies to generalized extensive-form games as well.

We say that with the belief system b_i and the strategy s_i player i is *rational* at the information set $h_i \in H_i$ if there exists no action $a'_{h_i} \in A_{h_i}$ such that only replacing the action $s_i(h_i)$ by a'_{h_i} results in a new strategy which yields player i a higher expected payoff at h_i given the belief $b_i(h_i)$ on the other players' strategies $S_{-i}^{T_{h_i}}$.

The *one-deviation principle* states that requiring rationality (which is a local notion of optimality) at all of i's information sets yields also sequential rationality at all of i's information sets:

Proposition 2. *If with the belief system b_i and the strategy s_i player i is rational at all of her information sets h_i, then she is also sequentially rational at all of her information sets.*

The proof is in Appendix A. The argument is the same as in the proof of the one-deviation property for standard extensive-form games with perfect recall, exploiting the fact that if h_i precedes h'_i then both these information sets of player i are subsets of the same tree $T \in \mathbf{T}$ in the generalized game.

Proposition 2 implies that if, in Definition 1 of extensive-form rationalizability, we were to replace "sequentially rational" by "rational", a smaller set of strategies could emerge. This is only due to the fact, however, that sequential rationality does not require optimization at information sets which were excluded by the strategy itself at an earlier information set. Thus, in terms of plans of action (equivalence classes of strategies which coincide at all information sets that none of them excludes), using "rationality" instead of "sequential rationality" would not affect the set of extensive-form rationalizable

---

[15] We thank an anonymous referee for this suggestion.

outcomes. Using "rationality" rather than "sequential rationality" seems, nevertheless, conceptually advantageous, because it rules out strategies with which a player behaves sub-optimally at some information sets — even if reaching them is counterfactual given the player's own earlier choices.[16]

4. Conclusion

We proposed a general framework for analyzing dynamic strategic interaction with asymmetric unawareness. We extended Pearce's (1984) notion of extensive-form (correlated) rationalizability to this setting, explored its properties, and proved existence. In a companion paper, Heifetz et al. (2011) we introduce a refinement of extensive-form rationalizability, called prudent rationalizability, and show that it rules out implausible outcomes in examples due to Pearce (1984) and Ozbay (2007). We apply it to a model of verifiable communication of Milgrom and Roberts (1986) and show that prudent rationalizability implies full unraveling of information in their model, but if the receiver is *unaware* of a dimension, then full unraveling does not need to occur. Thus, this is yet another example in which unawareness has strategic implications which are genuinely different than those implied by asymmetric information. In another companion paper, Meier and Schipper (2012), we define the normal-form game associated to a generalized extensive-form game with unawareness. We use it to characterize extensive-form rationalizability (resp. prudent rationalizability) in generalized extensive-form games by iterated elimination of conditional strictly (resp. weakly) dominated strategies in the associated generalized normal-form game. Moreover, we extend iterated admissibility to such associated generalized normal-form games.

Appendix A. Proofs

A.1. Proof of Proposition 1

We proceed by induction.

B_i^1 is non-empty. Indeed, to construct a belief system b_i, for each information set h_i with no predecessors (according to the precedence relation $\rightsquigarrow$) in the arborescence of information sets H_i, assign to player i a full-support belief $b_i(h_i)$ on the other players' strategies $S_{-i}^{T_{h_i}}$ that reach h_i. The full-support guarantees that Bayes rule is applicable for deriving the beliefs of player i in all her remaining information sets.

Suppose, by induction, we have already shown that B_i^k is non-empty. We have to show that R_i^k is non-empty. For a typical belief system $b_i \in B_i^k$ we have to construct a strategy $s_i \in R_i^k$, i.e., a strategy with which player i is sequentially rational at each of her information sets H_i given the belief system b_i. Since H_i is an arborescence, it is standard to construct such a strategy s_i by backward induction on H_i.

To complete the induction step, observe that B_i^{k+1} is non-empty, because by definition it singles out a non-empty subset of B_i^k.

Now, since player i's set of strategies S_i is finite and by Remark 7, $R_i^{k+1} \subseteq R_i^k$ for every $k \geqslant 1$, for some ℓ we eventually get $R_i^\ell = R_i^{\ell+1}$ for all $i \in I$ and hence $B_i^{\ell+1} = B_i^{\ell+2}$ for all $i \in I$. Inductively,

$$\emptyset \neq R_i^\ell = R_i^{\ell+1} = R_i^{\ell+2} = \cdots$$

and therefore

$$R_i^\infty = \bigcap_{k=1}^\infty R_i^k = R_i^\ell \neq \emptyset$$

as required.

A.2. Proof of Proposition 2

Suppose, by way of contradiction, that with the belief system b_i and the strategy s_i player i is rational at all of her information sets in all the trees $\hat{T} \in \mathbf{T}$, but there exist a tree $T \in \mathbf{T}$ and an information set h_i in T such that player i is not sequentially rational at h. This means that s_i^T reaches h_i in T and that there exists a strategy s_i' which is distinct from s_i only at h_i and/or at information sets succeeding h_i in T, and such that s_i' yields player i a higher expected payoff with her belief $b_i(h_i)$ over her opponents' strategy profiles S_{-i}^T.

For every information set h_i' in T, denote by $s_{h_i'}'$ the strategy of player i which coincides with s_i' at h_i' and its successors in T, and with s_i at the remaining information sets of player i.

The contradiction assumption implies that there exists at least one information set h_i' with the property

($*$)   with the belief $b_i(h_i')$ the strategy $s_{h_i'}'$ yields player i a higher expected payoff at h_i' than does s_i,

namely the information set $h_i' = h_i$.

Among the information sets h_i' satisfying ($*$), there exists at least one $\tilde{h}_i'$ which is *furthest* in the tree T, i.e., with the property that $\tilde{h}_i'$ satisfies ($*$) but no successor of $\tilde{h}_i'$ in T does. Such a furthest information set exists since the tree T is finite, and since by the perfect-recall property I6 the information sets *within* each given tree T form an arborescence (i.e., for every information set h_i in T, the information sets in T preceding h_i are linearly ordered).

For such a furthest $\tilde{h}_i'$, let $\tilde{s}_i'$ be the strategy which coincides with $s_{\tilde{h}_i'}'$ at $\tilde{h}_i'$ and with s_i elsewhere. In other words, we get $\tilde{s}_i'$ from $s_{\tilde{h}_i'}'$ by restoring s_i's choices at all the successors of $\tilde{h}_i'$. Since $\tilde{h}_i'$ is furthest with the property ($*$), with the belief $b_i(\tilde{h}_i')$ the expected payoff that $\tilde{s}_i'$ yields player i at $\tilde{h}_i'$ is greater than or equal to the expected payoff that $s_{\tilde{h}_i'}'$ yields her,[17] while — due to ($*$) — $s_{\tilde{h}_i'}'$ yields player i a higher expected payoff at $\tilde{h}_i'$ than does s_i.

Together this implies that even though $\tilde{s}_i'$ is distinct from s_i only at $\tilde{h}_i'$, nevertheless with the belief $b_i(\tilde{h}_i')$ it yields player i a higher expected payoff at $\tilde{h}_i'$ than does s_i. This contradicts the assumption that with the belief system b_i and the strategy s_i player i is rational at all of her information sets.

A.3. Proof of Remark 8

Note first that a strategy for player I in the game of Fig. 1 is a function that prescribes an action at the root of the tree and each matrix whereas in the game of Fig. 2 it is a function that prescribes an action at the root of the tree, the left and right matrices in the upper tree as well as an action in the lower matrix. Consequently, the belief systems of player II differ accordingly in those examples.

To make the differences and similarities between the examples more transparent, we will derive the extensive-form rationalizable strategies for both examples side-by-side.

At the first level, any strategy is rational for player I except all strategies that both prescribe "don't give" at the root of the tree and "Mozart" at the left matrix. Thus,

$$R_I^1 = \left\{ \begin{array}{l} (\text{"don't give"}, B, *), (\text{"don't give"}, S, *), \\ (\text{"give"}, *, *) \end{array} \right\},$$

where the first component of each strategy corresponds to player I's action at the root of the tree, the second component corresponds to his action at the left matrix, and the last component corresponds to his action at the right matrix. We denote by "$*$" any available action of the player at his corresponding information set.

For player II, both the Bach concert and the Stravinsky concert are rational if player I does not give him the car, i.e., at the right matrix. If player I does give him the car, then only the Mozart concert is rational since it is a dominant action conditional on being the right matrix. Thus,

$$R_{II}^1 = \left\{ (B, M), (S, M) \right\},$$

where the first component of a strategy refers to player II's action in the left matrix and the second refers to the right matrix.

At the first level, any strategy is rational for player I except all strategies that both prescribe "don't tell" and going to the Mozart concert at the upper left matrix. Thus,

$$R_I^1 = \left\{ \begin{array}{l} (\text{"don't tell"}, B, *, *), (\text{"don't tell"}, S, *, *), \\ (\text{"tell"}, *, *, *) \end{array} \right\},$$

where the first component of each strategy corresponds to player I's action at the root of the upper tree, the second component corresponds to his action at the left upper matrix, the third component corresponds to his action at the right upper matrix, and the last component refers to his action in the lower matrix. We denote by "$*$" any available action of the player at his corresponding information set.

For player II, both the Bach concert and the Stravinsky concert are rational if he is unaware of the Mozart concert. If he is aware of the Mozart concert, then only this concert is rational since it is a dominant action conditional of being in the right matrix. Thus,

$$R_{II}^1 = \left\{ (M, B), (M, S) \right\},$$

where now the first component of a strategy refers to player II's action in the upper right matrix and the second refers to the lower matrix.

---

[17]   Because the expected payoff that $s_{\tilde{h}_i'}'$ yields with the belief $b_i(\tilde{h}_i')$ is a weighted average of its expected payoff with the beliefs

$$\{ b_i(\tilde{h}_i''): s_{\tilde{h}_i'}' \text{ reaches } \tilde{h}_i'' \text{ and } \tilde{h}_i'' \text{ is a successor of } \tilde{h}_i' \}$$

at the corresponding information sets $\tilde{h}_i''$ — due to the fact that the belief system b_i satisfies Bayes rule whenever possible.

At the second level, player I is certain that player II will go to the Mozart concert when given the car. Thus, any second level rational strategy for player I that gives the car to player II must also prescribe going to the Mozart concert after "give the car". Not giving the car to player II and going to the Stravinsky concert is dominated by giving the car to player II and going to the Mozart concert. Not giving player II the car and going to the Bach concert is rational for player I assuming that she believes with probability at least $\frac{1}{4}$ that player II will go to the Bach concert. Giving the car to player II and going to the Mozart concert is rational for player I if she believes with probability at least $\frac{3}{4}$ that player II would go to the Stravinsky if not given the car. To summarize,

$$R_I^2 = \big\{(\text{"don't give"}, B, *), (\text{"give"}, *, M)\big\}.$$

For player II, $R_{II}^2 = R_{II}^1$ since the deletion of strategies that prescribe M in the left matrix for player I (and reach the left matrix) at the first level does not influence the optimality of any strategy of player II because when player I takes M any of player II's actions yields the same payoff in the left matrix.

At the second level, player I is certain that player II will go to the Mozart concert when told about it. Thus, any second level rational strategy for player I that specifies "tell" at the root of the upper tree must also prescribe going to the Mozart concert in the upper right matrix. Not telling player II about the Mozart concert and going to the Stravinsky concert is dominated by telling player II about the Mozart concert and going to the Mozart concert. Not telling player II about the Mozart concert and going to the Bach concert is rational for player I assuming that she believes with probability at least $\frac{1}{4}$ that (the unaware) player II will go to the Bach concert. Telling player II about the Mozart concert and going to the Mozart concert is rational for player I if she believes with probability at least $\frac{3}{4}$ that (the unaware) player II would go to the Stravinsky concert if not told about the Mozart concert. In the lower tree, both players are unaware of the Mozart concert. Going to Bach is rational for player I if she believes with probability at least $\frac{1}{4}$ that player II goes to Bach as well. Going to the Stravinsky concert is rational for player I if she believes with probability at least $\frac{3}{4}$ that player II goes to Stravinsky. This is just the standard battle-of-sexes game. To summarize,

$$R_I^2 = \big\{(\text{"don't tell"}, B, *, *), (\text{"tell"}, *, M, *)\big\}.$$

For player II, note that his strategy does not prescribe an action in the left matrix. Hence, the deletion of M in the left matrix for player I at the first level has no effect on the optimality of any strategy of player II. Thus $R_{II}^2 = R_{II}^1$.

So far, the arguments are analogous in both examples. A difference arises at the third level for player II. For player I, any second level strategy is also third level rational for player I since in both examples no strategies of player II have been eliminated at the second level.

At the third level, when player II is not given the car, he can forward-induce that player I will go to the Bach concert. This is because any second level rational strategy of player I that reaches the right matrix must prescribe the Bach concert. Consequently, a third level rational strategy of player II must prescribe going to the Bach concert as well when not given the car. Thus,

$$R_{II}^3 = \big\{(B, M)\big\}.$$

At the fourth level, if player I gives the car to player II, then latter will go to the Bach concert. Otherwise, if player I does not give the car to player II, then latter will go to the Mozart concert. Since player I strictly prefers to the Mozart concert together with player II, any fourth level rational strategy of player I must involve her not giving the car to player II. Thus,

$$R_I^4 = \big\{(\text{"don't give"}, B, *)\big\} = R_I^k, \quad \text{for all } k \geqslant 4.$$

Hence, the extensive-form rationalizable strategies are

$$R_I^\infty = \big\{(\text{"don't give"}, B, *)\big\},$$
$$R_{II}^\infty = \big\{(B, M)\big\}.$$

At the third level, when player II is not told about the Mozart concert, he can *not* forward-induce that player I will go to the Bach concert. This is because he is unaware of the Mozart concert and his information set is located in the lower subtree of Fig. 2. At this matrix, both Bach and Stravinsky are second level rational actions for player I. Thus, no strategies can be eliminated

$$R_{II}^3 = R_{II}^1.$$

For both players, no strategies were eliminated at the third level. Thus, no further strategies can be eliminated at any level $k \geqslant 3$.

Hence, the extensive-form rationalizable strategies are

$$R_I^\infty = \big\{(\text{"don't tell"}, B, *, *), (\text{"tell"}, *, M, *)\big\},$$
$$R_{II}^\infty = \big\{(M, B), (M, S)\big\}.$$

References

Battigalli, P., 1996. Strategic rationality orderings and the best rationalizability principle. Games Econ. Behav. 13, 178–200.

Battigalli, P., 1997. On rationalizability in extensive games. J. Econ. Theory 74, 40–61.

Battigalli, P., Siniscalchi, M., 2002. Strong belief and forward induction reasoning. J. Econ. Theory 106, 356–391.

Camerer, C., Ho, T.H., Chong, J.K., 2004. A cognitive hierarchy model of games. Quart. J. Econ. 119, 861–898.

Chen, J., Micali, S., 2011. The robustness of extensive-form rationalizability. Mimeo.

Crawford, V., Iriberri, N., 2007. Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? Econometrica 75, 1721–1770.

Dubey, P., Kaneko, M., 1984. Information patterns and Nash equilibria in extensive games: I. Math. Soc. Sci. 8, 111–139.

Eyster, E., Rabin, M., 2005. Cursed equilibrium. Econometrica 73, 1623–1672.

Feinberg, Y., 2009. Games with unawareness. Mimeo.

Filiz-Ozbay, E., forthcoming. Incorporating unawareness into contract theory. Games Econ. Behav.

Halpern, J., Rêgo, L., 2006. Extensive games with possibly unaware players. In: Proc. Fifth International Joint Conference on Autonomous Agents and Multi-agent Systems, pp. 744–751.

Harsanyi, J., 1967. Games with incomplete information played by "Bayesian" players, Part I. Manage. Sci. 14, 159–182.

Heifetz, A., Meier, M., Schipper, B.C., 2006. Interactive unawareness. J. Econ. Theory 130, 78–94.

Heifetz, A., Meier, M., Schipper, B.C., 2011. Prudent rationalizability in generalized extensive-form games. Mimeo.

Jehiel, P., 2005. Analogy-based expectation equilibrium. J. Econ. Theory 123, 81–104.

Li, J., 2006. Dynamic games with perfect awareness information. Mimeo.

Meier, M., Schipper, B.C., 2012. Conditional dominance in dynamic games with unawareness. Mimeo.

Milgrom, R., Roberts, J., 1986. Relying on the information of interested parties. RAND J. Econ. 17, 18–32.

Myerson, R., 1991. Game Theory: Analysis of Conflict. Harvard University Press.

Osborne, M., Rubinstein, A., 1994. A Course in Game Theory. MIT Press.

Ozbay, E., 2007. Unawareness and strategic announcements in games with uncertainty. In: Samet, D. (Ed.), Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge. Presses Universitaires de Louvain, pp. 231–238.

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

Perea, A., 2012. Epistemic Game Theory: Reasoning and Choice. Cambridge University Press.

Rêgo, L., Halpern, J., 2012. Generalized solution concepts in games with possibly unaware players. Int. J. Game Theory 41, 131–155.

Reny, P., 1992. Backward induction, normal form perfection and explicable equilibria. Econometrica 60, 627–649.

Schipper, B.C., Woo, H.Y., 2012. Political awareness and microtargeting of voters in electoral competition. University of California, Davis.

Stahl, D., Wilson, P.W., 1995. On players' models of other players: Theory and experimental evidence. Games Econ. Behav. 10, 218–254.

van Damme, E., 1989. Stable equilibria and forward induction. J. Econ. Theory 48, 476–496.