DEDUCTIVE REASONING IN EXTENSIVE GAMES

GEIR B. ASHEIM AND MARTIN DUFWENBERG

Abstract. We justify the application to extensive games of the concept of 'fully permissible sets', which corresponds to choice sets when there is common certain belief of the event that each player prefer one strategy to another if and only if the former weakly dominates the latter on the set of all opponent strategies or on the union of the choice sets that are deemed possible for the opponent. The extensive games considered illustrate how our concept yields support to forward induction, without necessarily promoting backward induction. *JEL* Classification Number: C72.

1. Introduction

What happens if players reason deductively in a strategic situation? A classic answer is provided by Bernheim [11] and Pearce [23]. In their modeling of strategic form games, common belief or knowledge of rational choice implies that precisely the strategies surviving iterated strong dominance may be used (provided that players are allowed to hold correlated conjectures concerning the choices of their opponents). Such strategies are called *(strategic form) rationalizable*.

However, since strategic form games suppress information about the sequential structure of a strategic situation, this result is often of limited use. Consider an example: Two persons, 1 and 2, are sitting in front of a button which if pushed sets off a nearby bomb. One after the other, they decide whether or not to push the button. Person 1 moves first, and 2 gets to move only if 1 does not push the button. With obvious motivation for the payoffs, this situation can be modeled by the strategic form game of Fig. 1.

Strategic form rationalizability permits anything to happen. If 1 believes with probability 1 that 2 will choose *push*, then *PUSH* is a utility maximizing choice for 1. Player 1 is justified in this belief in the sense that if 2 believes with probability 1 that 1 will *PUSH*, then to push is indeed a utility maximizing choice for 2. However, this is at best sensible only if information about the sequential structure of the situation is not considered. Clearly, if 1 does not *PUSH*, 2 should not *push* either. 1 should figure this out, and hence never *PUSH* in the first place!

Consideration of this kind led Bernheim [11] and Pearce [23] to propose rationalizability concepts also for extensive games in which the sequential

Date: 29 February 2000.

Acknowledgements: During our work with this paper and the companion paper "Admissibility and common belief", we have received helpful comments from many scholars, including Pierpaolo Battigalli and Larry Samuelson.

	push	*don't*
PUSH	0, 0	0, 0
DON'T	0, 0	1, 1

FIGURE 1. G_1

structure of strategic situations is made explicit. However, these concepts turn out to be difficult to justify using epistemic models; see, however, Battigalli & Siniscalchi [8] for an epistemic characterization of Pearce's concept of 'extensive form rationalizability' (EFR). We have in Asheim & Dufwenberg [2] (AD) proposed a model for deductive reasoning which can be applied to many strategic situations. In the present paper we argue that it is appropriate for analyzing extensive games and we apply the model to several such games. Before going into details, we now provide two more motivating examples.

In Γ_2 of Fig. 2, player 1 chooses D at his last node if he acts in accordance with his preferences. However, it is less clear what happens at the preceding node where 2 moves. Some models that formalize common knowledge or belief of rationality imply behavior which is in line with backward induction. A notable example is the model by Aumann [4], in which common knowledge of rationality implies that the backward induction solution is played in any generic extensive game with perfect information. In the case of Γ_2, Aumann's analysis supports the intuition that 2 chooses f, "figuring out" that 1 chooses D at the last node. One may, however, question how compelling this is: If in Γ_2 player 2 is asked to move, she knows that 1 is not playing according to backward induction. Indeed, she must understand that 1 is *not* choosing a strategy that is maximal given his preferences. Why should 2 believe that 1's behavior will follow backward induction at subsequent nodes? Objections of this kind lead Ben-Porath [9] to propose an alternative model which captures a very different intuition: Each player has an initial belief about the behavior of others. If this belief is contradicted by the play (a "surprise" occurs) he may subsequently entertain any belief consistent with the path of play. In Γ_2, Ben-Porath's model allows player 2 to make any choice.

Ben-Porath's [9] model is in general more permissive than Aumann's [4] model. We believe it is a very important contribution to the literature. If one finds the aforementioned critique of models that are "backward induction supportive" convincing, then Ben-Porath's analysis is a natural next step. However, in this paper we argue that Ben-Porath's approach is *too* permissive. This is because he does not impose certain reasonable constraints on how players reason about the likelihood of opponent choices. The argument can be illustrated by Γ_3 of Fig. 3. What preference should 2 have over her strategies in this game? By extrapolating Ben-Porath's assumption to games with imperfect information (Ben-Porath [9] considers only games with

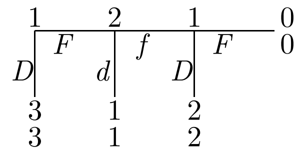


FIGURE 2. Γ_2

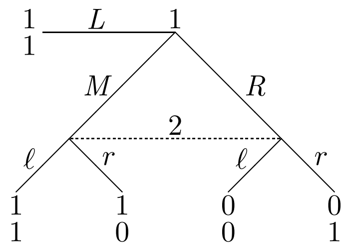


FIGURE 3. Γ_3

perfect information) one could argue as follows:[1] If 2 initially believes with probability one that 1 will choose L, then being asked to play is a surprise to 2. Hence, 2 may entertain any belief consistent with the path of play, and hence, 2 may prefer r to ℓ.

We find this conclusion implausible since L and M are the maximal strategies for 1 independently of his belief concerning 2's choice. Realizing that $\{L, R\}$ is 1's *choice set* (i.e. 1's set of maximal strategies), it would seem that 2 should deem each of L and M much more likely than the remaining non-maximal strategy R. Consequently, conditional on 2 being asked to play (i.e. conditional on 1 having played M or R), she should deem M much more likely than R. This would guarantee that 2 prefers ℓ to r. What we have here is an argument that 2 should deem any opponent strategy that is a rational choice much more likely than any strategy not having this property.

In AD we show that similar concerns may arise in the context of strategic form games. We handle the issue by requiring that a player should ...

1. ... deem any opponent strategy that is a rational choice infinitely more likely (in the sense of Blume, Brandenburger & Dekel [15], Def. 5.1) than any opponent strategy not having this property. This is equivalent to saying that a player should prefer one strategy to another if the former weakly dominates the latter on the set of rational choices for the opponent. Such admissibility of a player's preferences — which we in AD refer to as 'full belief of opponent rationality' — is a key

[1]The reader may wonder whether we need to go to a game with imperfect information to illustrate that Ben-Porath's [9] approach is too permissive. In Sect. 4 we show that the same point can be made concerning a game (Γ_4) with perfect information. However, it is a slightly more complicated to analyze Γ_4 than to analyze Γ_3, so we prefer to use Γ_3 here.

ingredient in the analyses of weak dominance by Samuelson [28] and Börgers & Samuelson [16], and is essentially satisfied by procedures, like EFR (cf. Pearce [23] and Battigalli [6, 7]) and 'iterated elimination of (all) weakly dominated strategies' (IEWDS), that promote forward induction.

2. ... prefer one strategy to another if the former weakly dominates the latter. Such admissibility of a player's preferences — which can be referred to as 'caution' since it means that all opponent strategies are taken into account — has been defended by e.g. Luce & Raiffa ([21], Ch. 13) and is implicit in any procedure that starts out by eliminating all weakly dominated strategies. In an extensive game, 'caution' ensures that each player takes into account the possibility of reaching any information set.

Formally, a player's preferences over his own strategies, which depend both on his payoff function and on his beliefs about the strategy choice of his opponent, leads to a *choice set* (i.e. a set of maximal strategies). A player's preferences are said to be *fully admissibly consistent* with the preferences of his opponent if one strategy is preferred to another if and only if the former weakly dominates the latter

- on the union of the choice sets that are deemed possible for the opponent (i.e. 'full belief of opponent rationality'), or
- on the set of all opponent strategies (i.e. 'caution').

A subset of strategies is a *fully permissible set* if and only if it can be a choice set when there is common certain belief of full admissible consistency, where an event is 'certainly believed' if the complement is deemed impossible (or more precisely: is Savage-null). Hence, the analysis yields a solution concept that determines a collection of strategy subsets – a family of choice sets – for each player.

The formal definition in AD of the concept of fully permissible sets is given in terms of an elimination procedure — iterative elimination of choice set under full admissible consistency (IECFA) — that iteratively eliminates strategy subsets that cannot be choice sets under full admissible consistency. Subsequently in AD we provide an epistemic characterization as indicated in the previous paragraph. In this paper we will apply the elimination procedure only (which is introduced in Sect. 2), although we will at some places interpret results in a way consistent with the underlying epistemic foundation.

We now propose that some interesting implications of deductive reasoning in any given extensive game can be derived by applying the elimination procedure of AD to the *pure strategy reduced strategic form* (see Mailath, Samuelson & Swinkels [22]) of that extensive game. We present two formal results that serve to justify this application: First, we address the problem of time consistency which is pertinent when applying strategic form analysis to extensive games with an explicit sequential structure. In our case a special problem is that AD's common certain belief of full admissible consistency need not ensure that preferences are complete. To provide a convincing argument that the analysis of AD can be used to analyze deductive reasoning in extensive games, one has to show that any strategy that is maximal at the outset is still maximal when the preferences have been updated upon reaching any information set that the choice of this strategy does not preclude. Drawing on results due to Mailath et al. [22], we prove that this is so. Second, we show that to apply the framework of AD it is sufficient to consider the pure strategy reduced strategic form. Taken together, these results justify investigating the consequences of the concept of fully permissible sets in any extensive games via that game's pure strategy reduced strategic form.

The paper is organized as follows. Section 2 introduces the elimination procedure that determines the fully permissible sets (IECFA) and summarizes some results of AD. Section 3 contains formal results that justify our claim that IECFA is applicable to extensive games. In Sect. 4 we analyze several extensive games via this approach. We first return to the examples of the introduction, then consider a game which allows us to compare our results to those of Ben-Porath [9], and finally analyze some games that relate to issues of backward and forward induction. The conclusion of AD — that our approach yields support to forward induction — is reinforced, and we attempt to shed light on the "backward induction paradox" discussed by many authors. Section 5 concludes.

2. Concepts

Below we make a self-contained presentation of the concept of fully permissible sets through a definition based on the IECFA procedure (cf. Def. 1). Readers that are interested in the underlying epistemic foundation must, however, consult AD. As the purpose here is to apply this concept to extensive games, we start by introducing such games. We refer to standard texts for the general formalism of extensive games and state only those basic and derived notions that will be needed.

2.1. **Extensive Games.** A finite extensive game Γ (without nature) includes a set of players $N \in \{1,2\}$ (we assume 2 players for convenience), a set of terminal nodes Z, and, for each player i, a vNM utility function $\upsilon_i : Z \to \mathbb{R}$ that assigns payoff to any outcome. For any player i, there is a finite collection of information sets H_i, with a finite set of actions $A_i(h)$ being associated with each $h \in H_i$. A pure strategy for player i is a function s_i that to any $h \in H_i$ assigns an action in $A_i(h)$. Let S_i denote player i's finite set of pure strategies, and let $S = S_1 \times S_2$. Write p_i, r_i and s_i ($\in S_i$) for pure strategies and x_i and y_i ($\in \Delta(S_i)$) for mixed strategies. Define $u_i : S \to \mathbb{R}$ by $u_i(s) = \upsilon_i(z)$, where z is the terminal node reached when $s = (s_1, s_2)$ is used, and refer to $G = (S_i, u_i)_{i \in N}$ as the strategic form of the extensive game Γ. Since u_i is a vNM utility function, we may extend u_i to mixed strategies: $u_i(x_i, s_j) = \sum_{s_i \in S_i} x_i(s_i) u_i(s_i, s_j)$. For any $h \in \bigcup_{i \in N} H_i$, let S^h denote the set of strategy vectors for which h is reached. As Γ is a 2-player game with perfect recall, S^h is rectangular: $S^h = S_1^h \times S_2^h$.

2.2. **Fully Permissible Sets.** Say that x_i *weakly dominates* y_i *on* Q_j ($\subseteq S_j$) if, $\forall s_j \in Q_j$, $u_i(x_i, s_j) \geq u_i(y_i, s_j)$, with strict inequality for some $s_j \in Q_j$. Say that player i's preferences over his own strategies are *admissible on* Q_j ($\neq \emptyset$) if x_i is preferred to y_i whenever x_i weakly dominates y_i on Q_j. Player i's *choice set* is the set of pure strategies that are maximal w.r.t. i's preferences over his own strategies: s_i ($\in S_i$) is in i's choice set if and only if there is no x_i ($\in \Delta(S_i)$) such that x_i is preferred to s_i. For the class of preferences considered in the present paper, i's choice set is non-empty and supports any maximal mixed strategy (cf. subsect. 3.5 of AD).

Let the set Q_j be interpreted as the set of strategies that player i deems to be the set of rational choices for his opponent. Assume that player i's preferences over his own strategies are characterized by the property of being admissible on both Q_j and S_j: x_i is preferred to y_i if and only if x_i weakly dominates y_i on Q_j or S_j. Player i's choice set, $C_i(Q_j)$, is then equal to $S_i \backslash D_i(Q_j)$, where, for any ($\emptyset \neq$) $Q_j \subseteq S_j$,

$$D_i(Q_j) := \{s_i \in S_i | \exists x_i \in \Delta(S_i) \text{ s.t. } x_i \text{ weakly dom. } s_i \text{ on } Q_j \text{ or } S_j\}.$$

Let $\Sigma = \Sigma_1 \times \Sigma_2$, where $\Sigma_i := 2^{S_i} \backslash \{\emptyset\}$ denotes the collection of non-empty subsets of S_i. Write π_i, ρ_i, and σ_i ($\in \Sigma_i$) for subsets of pure strategies. For any ($\emptyset \neq$) $\Xi = \Xi_1 \times \Xi_2 \subseteq \Sigma$, write $\alpha(\Xi) := \alpha_1(\Xi_2) \times \alpha_2(\Xi_1)$, where

$$\alpha_i(\Xi_j) := \{\pi_i \in \Sigma_i | \exists (\emptyset \neq) \Psi_j \subseteq \Xi_j \text{ s.t. } \pi_i = C_i(\cup_{\sigma_j \in \Psi_j} \sigma_j)\}.$$

Hence, $\alpha_i(\Xi_j)$ is the collection of strategy subsets that can be choice sets for player i if i's preferences are characterized by the property of being admissible *both* on the union of the strategy subsets in a non-empty subcollection of Ξ_j *and* on the union of all opponent strategies.

We can now define the concept of a fully permissible set.

Definition 1. Consider the sequence defined by $\Xi(0) = \Sigma$ and, $\forall g \geq 1$, $\Xi(g) = \alpha(\Xi(g-1))$. A non-empty strategy set π_i is said to be a *fully permissible set* for i if $\pi_i \in \bigcap_{g=0}^{\infty} \Xi_i(g)$.

Let $\Pi = \Pi_1 \times \Pi_2$ denote the *collection* of vectors of fully permissible sets. Since the game is finite, $\Xi(g)$ converges to Π in a finite number of iterations. IECFA is the procedure that in round g eliminates sets in $\Xi(g-1) \backslash \Xi(g)$ as possible choice sets. A choice set of player i survives elimination round g if it is a choice set w.r.t. preferences that are characterized by the property of being admissible *both* on the union of some (or all) of opponent choice sets that have survived the procedure up till round $g-1$ *and* on the set of all opponent strategies. A fully permissible set is a choice set which will survive in this way for any g. It follows from the analysis of AD that strategy subsets that this algorithm has not eliminated by round g can be interpreted as choice sets that are compatible with $g-1$ order of mutual certain belief of full admissible consistency.

The algorithm of Def. 1 – IECFA – is an elimination procedure, and in this regard it is reminiscent of procedures that iteratively eliminates dominated strategies. However, IECFA does *not* eliminate strategies. Rather, it eliminates *sets* of strategies that cannot be choice sets under full admissible consistency. It is therefore that IECFA starts with each player's collection of all non-empty strategy subsets, and then iteratively eliminates subsets in this collection. It is important that the appropriate interpretation of IECFA in terms of surviving choice sets be borne in mind.

We reproduce from AD the following proposition, which characterizes the strategy subsets that survive IECFA and thus are fully permissible.

Proposition 1. *(i)* $\forall i \in N$, $\Pi_i \neq \emptyset$. *(ii)* $\Pi = \alpha(\Pi)$. *(iii)* $\forall i \in N$, $\pi_i \in \Pi_i$ *if and only if there exists* $\Xi = \Xi_1 \times \Xi_2$ *with* $\pi_i \in \Xi_i$ *such that* $\Xi \subseteq \alpha(\Xi)$.

Prop. 1(i) establishes existence, but not uniqueness, of each player's fully permissible set(s). Games with multiple strict Nash equilibria illustrate the possibility of such multiplicity; by Prop. 1(iii), any strict Nash equilibrium corresponds to a vector of fully permissible sets. Other (quite different) examples of games with multiple fully permissible sets are provided in Sect. 4 by Γ_5 and Γ_7 as well as a 3-period prisoners' dilemma game. Prop. 1(ii) means that Π is a fixed point in terms of a collection of vectors of strategy sets. By Prop. 1(iii) it is the largest such fixed point.

3. Justifying Extensive Form Application

The concept of fully permissible sets, presented in Sect. 2 of the present paper and epistemically characterized in AD, is designed to analyze the implications of deductive reasoning in strategic form games. In this paper, we propose that this concept can be fruitfully applied for analyzing any extensive game through its strategic form. In fact, we propose that it is legitimate to confine attention to the game's pure strategy reduced strategic form (cf. Def. 2), which is computationally more convenient. In this section we prove two results which, taken together, justify our approach.

3.1. **Dynamic Consistency.** Proposition 2 addresses the dynamic consistency problem inherent in applying AD's strategic form theory to an extensive games with an explicit sequential structure. Consider any strategy that is maximal given preferences that are charcterized by the property of being admissible on both Q_j — the set of strategies that player i deems to be the set of rational choices for his opponent — and S_j — the set of all opponent strategies. Hence, the strategy is maximal at the outset of a corresponding extensive game. We prove that this strategy is still maximal when the preferences have been updated upon reaching any information set that the choice of this strategy does not preclude.

Assume that player i's preferences over his own strategies are given by: x_i is preferred to y_i if and only if x_i weakly dominates y_i on Q_j or S_j. Let, for any $h \in H_i$, $Q_j^h := Q_j \cap S_j^h$ denote the set of strategies in Q_j that are consistent with the information set h being reached. If x_i, y_i $(\in \Delta(S_i^h))$, then i's preferences conditional on the information set $h \in H_i$ being reached is given by: x_i is preferred to y_i if and only if x_i weakly dominates y_i on Q_j^h or S_j^h (where it follows from the definition that weak dominance on Q_j^h is not possible if $Q_j^h = \emptyset$). Furthermore, i's choice set conditional on $h \in H_i$, $C_i^h(Q_j)$, is equal to $S_i^h \backslash D_i^h(Q_j)$, where, for any $(\emptyset \neq)$ $Q_j \subseteq S_j$,

$$D_i^h(Q_j) := \{s_i \in S_i^h | \exists x_i \in \Delta(S_i^h) \text{ s.t. } x_i \text{ weakly dom. } s_i \text{ on } Q_j^h \text{ or } S_j^h\}.$$

By the following proposition, if s_i is maximal at the outset of an extensive game, then it is also maximal at later information sets for i that s_i does not preclude.

Proposition 2. *Let $(\emptyset \neq)$ $Q_j \subseteq S_j$. If $s_i \in C_i(Q_j)$, then $s_i \in C_i^h(Q_j)$ for any $h \in H_i$ with $S_i^h \ni s_i$.*

Proof. Suppose that $s_i \in S_i^h \backslash C_i^h(Q_j) = D_i^h(Q_j)$. Then there exists $x_i \in \Delta(S_i^h)$ such that x_i weakly dominates s_i on Q_j^h or S_j^h. By Mailath, Samuelson & Swinkels ([22], Defs. 2 and 3 and the if-part of Theorem 1), S^h is a *strategic independence* for i. Hence, x_i can be chosen such that $u_i(x_i, s_j) = u_i(s_i, s_j)$ for all $s_j \in S_j \backslash S_j^h$. This implies that x_i weakly dominates s_i on Q_j or S_j, implying that $s_i \in D_i(Q_j) = S_i \backslash C_i(Q_j)$. □

By the assumption of 'caution', each player i takes into account the possibility of reaching any information set $h \in H_i$.

3.2. **Reduced Strategic Form.** It follows from Prop. 3 that it is in fact sufficient to consider the pure strategy reduced strategic form when deriving the fully permissible sets of the game. The following definition is needed.

Definition 2. Let r_i, $s_i \in S_i$. Then r_i and s_i are *equivalent* if, for each $k \in N$, $u_k(r_i, s_j) = u_k(s_i, s_j)$ for all $s_j \in S_j$. The *pure strategy reduced strategic form* (PRSF) of G is obtained by letting, for each i, each class of equivalent pure strategies be represented by exactly one pure strategy.

Since the maximality of one of two equivalent strategies implies that the other is maximal as well, the following observation holds: If r_i and s_i are equivalent and π_i is a fully permissible set for i, then $r_i \in \pi_i$ if and only if $s_i \in \pi_i$. To see this formally, note that if $r_i \in \pi_i$ for some fully permissible set π_i, then, by Prop. 1(ii), there exists $(\emptyset \neq)\Psi_j \subseteq \Pi_j$ such that $r_i \in \pi_i = C_i(\cup_{\sigma_j \in \Psi_j} \sigma_j)$. Since r_i and s_i are equivalent, $s_i \in C_i(\cup_{\sigma_j \in \Psi_j} \sigma_j) = \pi_i$. This observation explains why the following proposition can be established.

Proposition 3. *Let $\tilde{G} = (\tilde{S}_k, \tilde{u}_k)_{k \in N}$ be a strategic form game where r_i, $s_i \in \tilde{S}_i$ are two equivalent strategies for i. Consider $G = (S_k, u_k)_{k \in N}$, where $S_i = \tilde{S}_i \backslash \{r_i\}$ and $S_j = \tilde{S}_j$ for $j \neq i$, and where, for all $k \in N$, u_k is the restriction of $\tilde{u}_k$ to $S = S_1 \times S_2$. Let, for each $k \in N$, Π_k $(\tilde{\Pi}_k)$ denote the collection of fully permissible sets for k in G $(\tilde{G})$. Then Π_i is obtained from $\tilde{\Pi}_i$ by removing r_i from any $\tilde{\pi}_i \in \tilde{\Pi}_i$ with $s_i \in \tilde{\pi}_i$, while, for $j \neq i$, $\Pi_j = \tilde{\Pi}_j$.*

Proof. By Prop 1(iii) it suffices to show that

1. If $\tilde{\Xi} \subseteq \alpha(\tilde{\Xi})$ for $\tilde{G}$, then $\Xi \subseteq \alpha(\Xi)$ for G, where Ξ_i is obtained from $\tilde{\Xi}_i$ by removing r_i from any $\tilde{\pi}_i \in \tilde{\Xi}_i$ with $s_i \in \tilde{\pi}_i$, while, for $j \neq i$, $\Xi_j = \tilde{\Xi}_j$.
2. If $\Xi \subseteq \alpha(\Xi)$ for G, then $\tilde{\Xi} \subseteq \alpha(\tilde{\Xi})$ for $\tilde{G}$, where $\tilde{\Xi}_i$ is obtained from Ξ_i by adding r_i to any $\pi_i \in \Xi_i$ with $s_i \in \pi_i$, while, for $j \neq i$, $\tilde{\Xi}_j = \Xi_j$.

Part 1. Assume $\tilde{\Xi} \subseteq \alpha(\tilde{\Xi})$. By the observation preceding Prop. 3, if $\tilde{\pi}_i \in \tilde{\Pi}_i$, then $r_i \in \tilde{\pi}_i$ if and only if $s_i \in \tilde{\pi}_i$. Pick any $k \in N$ and any $\tilde{\pi}_k \in \tilde{\Pi}_k$. Let ℓ denote k's opponent. By the definition of $\alpha_k(\cdot)$, there exists $(\emptyset \neq)\tilde{\Psi}_\ell \subseteq \tilde{\Pi}_\ell$ such that $\tilde{\pi}_k = C_k(\cup_{\tilde{\sigma}_\ell \in \tilde{\Psi}_\ell} \tilde{\sigma}_\ell)$. Construct Ψ_i by removing r_i from any $\tilde{\sigma}_i \in \tilde{\Psi}_i$ with $s_i \in \tilde{\sigma}_i$ and replace $\tilde{S}_i$ by S_i, while, for $j \neq i$, $\Psi_j = \tilde{\Psi}_j$ and $S_j = \tilde{S}_j$. Then it follows from the definition of $C_k(\cdot)$ that $C_k(\cup_{\sigma_\ell \in \Psi_\ell} \sigma_\ell) = \tilde{\pi}_k \backslash \{r_k\}$ if $k = i$ and $C_k(\cup_{\sigma_\ell \in \Psi_\ell} \sigma_\ell) = \tilde{\pi}_k$ if $k \neq i$. Since, $\forall k \in N$, $(\emptyset \neq)\Psi_k \subseteq \Xi_k$, we have that $\Xi \subseteq \alpha(\Xi)$. *Part 2* is shown similarly. □

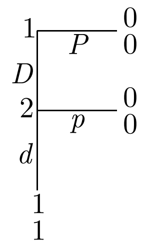
Proposition 3 means that the PRSF is sufficient for analyzing common certain belief of full admissible consistency, which is the epistemic foundation for the concept of fully permissible sets. Consequently, in the strategic form of an extensive game, it is unnecessary to specify actions at information sets that a strategy precludes from being reached. Hence, instead of fully specified strategies, it is sufficient to consider (what Rubinstein [27] calls) *plans of action*. For a generic extensive game, the set of plans of action is identical to the strategy set in the PRSF.

4. Applications

In this section we apply the concept of fully permissible sets to several extensive games. We first return to the examples of the introduction, and then consider a game which allows us to compare our results to those of Ben-Porath [9]. Furthermore, we analyze some games that relate to issues of backward and forward induction, before finally presenting an analysis of a 3-period prisoners' dilemma game.

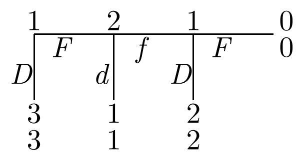
Other support for forward induction through the concept of EFR and the procedure of IEWDS precludes outcomes in conflict with backward induction (cf. e.g. Battigalli [7]). In constrast, the following examples show how our concept promotes forward induction (in a game like Γ_6 of Fig. 9), while not insisting on the backward induction outcome in games (like Γ_5 of Fig. 8 and the 3-period prisoners' dilemma) where earlier constributions, like Basu [5], Reny [25] and others, have argued on theoretical grounds that this is particularly problematic. Still, it should be noticed that the backward induction outcome *is* obtained in Γ_4 of Fig. 7, and that our concept has considerable bite in the 3-period prisoners' dilemma game.

Motivated by Props. 2 and 3, we analyze each extensive game via its PRSF (cf. Def. 2), given in conjunction to the extensive game itself. In each example, each plan of action that appears in the underlying extensive game corresponds to a distinct strategy in the PRSF.



	p	d
P	0, 0	0, 0
D	0, 0	1, 1

FIGURE 4. Γ_1 and its PRSF.



	d	f
D	3, 3	3, 3
FD	1, 1	2, 2
FF	1, 1	0, 0

FIGURE 5. Γ_2 and its PRSF.

4.1. **The Examples of the Introduction.** The extensive game Γ_1 of Fig. 4 is a model of the strategic situation that we used to motivate G_1. Applying the algorithm of Def. 1 – IECFA – to the PRSF of Γ_1 yields:

$$\Xi(0) = \Sigma_1 \times \Sigma_2$$
$$\Pi = \Xi(1) = \{\{D\}\} \times \{\{d\}\}$$

$(\{D\}, \{d\})$ is the unique vector of fully permissible sets in Γ_1.

We now move to Γ_2 which we analyze via its PRSF (cf. Fig. 5). Our algorithm IECFA applied to the PRSF of Γ_2 yields:

$$\Xi(0) = \Sigma_1 \times \Sigma_2$$
$$\Xi(1) = \{\{D\}\} \times \Sigma_2$$
$$\Pi = \Xi(2) = \{\{D\}\} \times \{\{d, f\}\}$$

We interpret this result as follows: Irrespective of what strategies that 1 deems as rational choices for 2, D is the only strategy that is maximal for 1. Player 2 considers each of the strategies FD and FF infinitely less likely than D. However, conditional on 2's node being reached, i.e. conditional on 1 not choosing D, 2 does not have any assessment of likelihood concerning which non-maximal strategy FF or FD that 1 has chosen. Hence each of d and f is maximal for 2.

Turn now to the pure reduced strategic form of Γ_3, which is illustrated in Fig. 6. Applying IECFA to the PRSF of Γ_3 yields:

$$\Xi(0) = \Sigma_1 \times \Sigma_2$$
$$\Xi(1) = \{\{L, M\}\} \times \Sigma_2$$
$$\Pi = \Xi(2) = \{\{L, M\}\} \times \{\{\ell\}\}$$

Interpretation: Irrespective of what strategies that 1 deems as rational choices for 2, L and M are the maximal strategies for 1. Player 2 deems each of L and M infinitely more likely than R. Conditional on her information set being reached, 2 considers it infinitely more likely that 1 is using M rather than R. Then only ℓ can be a maximal strategy for 2.

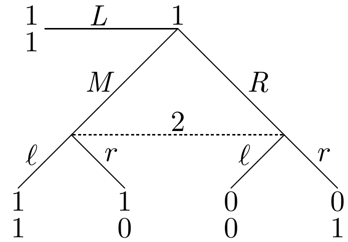
4.2. **Comparison to Ben-Porath** [9]**.** Ben-Porath [9] models "initial common certainty of rationality" (initial CCR) in extensive games with perfect information. He proves that in generic games (with no payoff ties at terminal nodes for any player) the outcomes consistent with that assumption coincide with those that survive the Dekel-Fudenberg [17] procedure (where one round of elimination of all weakly dominated strategies is followed by iterated elimination of strongly dominated strategies). The concept of fully permissible sets generally refines the Dekel-Fudenberg procedure (see AD, Prop. 2). Game Γ_4 of Fig. 7 shows that the refinement may be strict even for generic extensive games with perfect information.

Γ_4, which was introduced by Reny ([24], Fig. 1) and has appeared in many contributions, is a generic game. The set of profiles surviving the Dekel-Fudenberg procedure is $\{D, FF\} \times \{d, f\}$, and hence these profiles are consistent with initial CCR. We refer to Ben-Porath [9] for formal details, and here give only the rough intuition for why the strategies D and d are possible: D is 1's unique best strategy if he believes with probability one that 2 plays d. Player 1 is justified in this belief in the sense that d is 2's best strategy if she initially believes with probability one that 1 will choose D, and if called upon to play revises this belief so as to believe with high enough probability that 1 is using FD. Since only initial beliefs must be supported by strategies consistent with rationality, such belief revision is acceptable.

This is at odds with the implications of common certain belief of full admissible consistency. Applying IECFA to the PRSF of Γ_4 yields:

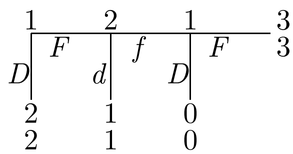
$$
\begin{aligned}
\Xi(0) &= \Sigma_1 \times \Sigma_2 \\
\Xi(1) &= \{\{D\}, \{FF\}, \{D, FF\}\} \times \Sigma_2 \\
\Xi(2) &= \{\{D\}, \{FF\}, \{D, FF\}\} \times \{\{f\}, \{d, f\}\} \\
\Xi(3) &= \{\{FF\}, \{D, FF\}\} \times \{\{f\}, \{d, f\}\} \\
\Xi(4) &= \{\{FF\}, \{D, FF\}\} \times \{\{f\}\} \\
\Pi = \Xi(5) &= \{\{FF\}\} \times \{\{f\}\}
\end{aligned}
$$

Interpretation: $\Xi(1)$: FD cannot be a maximal strategy for 1 since it is a dominated strategy. $\Xi(2)$: Player 2 certainly believes that only $\{D\}$, $\{FF\}$ and $\{D, FF\}$ are candidates for 1's choice set. This excludes $\{d\}$ as 2's choice set, since $\{d\}$ is 2's choice set only if 2 deems $\{FD\}$ or $\{D, FD\}$ possible. $\Xi(3)$: 1 certainly believes that only $\{f\}$ and $\{f, d\}$ are candidates for 2's choice set. This excludes $\{D\}$ as 1's choice set, since $\{D\}$ is 1's choice set only if 1 deems $\{d\}$ possible. $\Xi(4)$: Player 2 certainly believes that only



	ℓ	r
L	1, 1	1, 1
M	1, 1	1, 0
R	0, 0	0, 1

FIGURE 6. Γ_3 and its PRSF.



	d	f
D	2, 2	2, 2
FD	1, 1	0, 0
FF	1, 1	3, 3

FIGURE 7. Γ_4 and its PRSF.

$\{FF\}$ and $\{D, FF\}$ are candidates for 1's choice set. This implies that 2's choice set is $\{f\}$. $\Xi(5)$: 1 certainly believes that 2's choice set is $\{f\}$, and hence $\{FF\}$ is 1's choice set. No further elimination of choice sets is possible, so $\{FF\}$ and $\{f\}$ are the respective players' unique fully permissible sets.

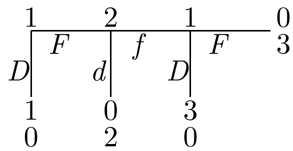
4.3. **Backward Induction.** Does deductive reasoning in extensive games imply backward induction? During the last couple of decades, many authors have debated various aspects of this issue.[2] The background is the following paradoxical aspect of backward induction: Why should a player believe that an opponent's future play will satisfy backward induction if the opponent's previous play is incompatible with backward induction? We now discuss what our approach has to say.

Reny [25] studies the "Take-it-Or-Leave-it" game with k stages (TOL(k)) (a version of Rosenthal's [26] centipede game), where at ℓth stage of the game, the total pot is ℓ dollars. If ℓ is odd (even), player 1 (2) may take ℓ dollars and end the game, or leave it, in which case the pot increases with one dollar. Should the game continue until the kth stage and the player whose turn it is decides to leave the k dollars, it is given to the other player. We analyze TOL(3) in detail.

Applying our algorithm IECFA to the PRSF of Γ_5 yields:

$$\begin{aligned} \Xi(0) &= \Sigma_1 \times \Sigma_2 \\ \Xi(1) &= \{\{D\}, \{FD\}, \{D, FD\}\} \times \Sigma_2 \\ \Xi(2) &= \{\{D\}, \{FD\}, \{D, FD\}\} \times \{\{d\}, \{d, f\}\} \\ \Pi = \Xi(3) &= \{\{D\}, \{D, FD\}\} \times \{\{d\}, \{d, f\}\} \end{aligned}$$

[2]These papers include Aumann [4], Basu [5], Ben-Porath [9], Bicchieri [12], Binmore [13, 14], Gul [18], and Reny [25].



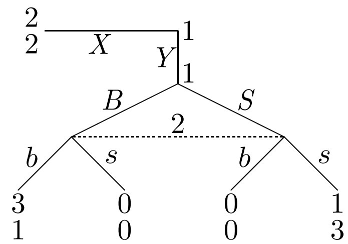
	d	f
D	1, 0	1, 0
FD	0, 2	3, 0
FF	0, 2	0, 3

FIGURE 8. Γ_5 (= TOL(3)) and its PRSF.

Interpretation: $\Xi(1)$: FF cannot be a maximal strategy for 1 since it is a dominated strategy. $\Xi(2)$: Player 2 certainly believes that only $\{D\}$, $\{FD\}$ and $\{D, FD\}$ are candidates for 1's choice set. This excludes $\{f\}$ as 2's choice set since $\{f\}$ is 2's choice set only if 2 deems $\{FF\}$ or $\{FD, FF\}$ possible. $\Xi(3)$: 1 certainly believes that only $\{d\}$ and $\{d, f\}$ are candidates for 2's choice set, implying that $\{FD\}$ cannot be 1's choice set. No further elimination of choice sets is possible and the collection of vectors of fully permissible sets is as specified.

Note that backward induction is *not* implied. To illustrate why, we focus on player 2 and explain why $\{d, f\}$ may be a choice set for her. Player 2 certainly believes that 1's choice set is $\{D\}$ or $\{D, FD\}$. This leaves room for two basic cases. First, suppose 2 deems $\{D, FD\}$ possible. Then $\{d\}$ must be her choice set, since she must consider it infinitely more likely that 1 uses *FD* than that he uses *FF*. Second, and more interestingly, suppose 2 does not deem $\{D, FD\}$ possible. Then conditional on 2's node being reached 2 certainly believes that 1 is not choosing a maximal strategy. As player 2 is assumed not to assess the relative likelihood of strategies that are not maximal, $\{d, f\}$ is her choice set in this case. Note that even in the case when 2 deems $\{D\}$ to be the only possible choice set for 1, she still considers it possible that 1 may choose one of his non-maximal strategies *FD* and *FF* (cf. the property of 'caution'), although each of these strategies is in this case deemed infinitely less likely than the unique maximal strategy D.

We now compare our results to the very different findings of Aumann [4]. In his analysis of common knowledge of rational choice in perfect information games it is crucial to specify full strategies (rather than plans of actions). In Aumann's model, common knowledge of rational choice implies in TOL(3) that all strategies for 1 but *DD* (where he takes the 1 dollar at his first node and takes the 3 dollars at his last node) are impossible. Hence, it is impossible for 1 to play *FD* or *FF* and thereby ask 2 to play. However, in the counterfactual event that 2 is asked to play, she optimizes as if player 1 at his last node follows his only possible strategy *DD*, implying that it is impossible for 2 to choose f (see Aumann's Sects. 4b, 5b, and 5c). Thus, in Aumann's analysis, if there is common knowledge of rational choice, then each player chooses the backwards induction strategy. By contrast, in our analysis player 2 being asked to play is seen to be incompatible with 1 playing *DD* or *DF*. For the determination of 2's preference over her strategies it is the relative likelihood of *FD* versus *FF* that is important to her. As seen



	b	s
X	2, 2	2, 2
B	3, 1	0, 0
S	0, 0	1, 3

FIGURE 9. Γ_6 (= BoSwOO) and its PRSF.

above, this assessment depends on whether she deems $\{D, FD\}$ as a possible candidate for 1's choice set.

4.4. **Forward Induction.** In AD we analyze the "Battle-of-the-Sexes-with-an-Outside-Option" (BoSwOO) game (introduced by Kreps & Wilson [20] who credit Elon Kohlberg) and the "Burning money" game (van Damme ([29], Fig. 5), Ben-Porath & Dekel ([10], Fig. 1.2)) in the strategic form and show how the concept of fully permissible sets yields forward induction outcomes. We refer the reader to AD for a detailed discussion. Here we briefly consider an extensive form version of the BoSwOO game (cf. Fig. 9), and then analyze a modification due to Dekel & Fudenberg [17].

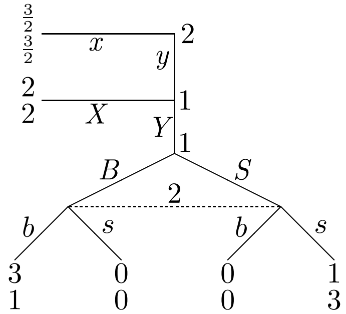
Applying IECFA to the PRSF of Γ_6 yields:

$$\begin{aligned} \Xi(0) &= \Sigma_1 \times \Sigma_2 \\ \Xi(1) &= \{\{X\}, \{B\}, \{X, B\}\} \times \Sigma_2 \\ \Xi(2) &= \{\{X\}, \{B\}, \{X, B\}\} \times \{\{b\}, \{b, s\}\} \\ \Xi(3) &= \{\{B\}, \{X, B\}\} \times \{\{b\}, \{b, s\}\} \\ \Xi(4) &= \{\{B\}, \{X, B\}\} \times \{\{b\}\} \\ \Pi = \Xi(5) &= \{\{B\}\} \times \{\{b\}\} \end{aligned}$$

The profile (B, b) corresponds to the usual forward induction outcome. Props. 2 and 3 together is our justification for claiming that common certain belief of full admissible consistency captures forward induction in the same way in any extensive game underlying the PRSF of Γ_6.[3]

Turn now to a game introduced by Dekel & Fudenberg ([17], Fig 7.1), which is discussed also by Hammond [19], and which is reproduced here as Γ_7 of Fig. 10. It is a modification of Γ_6 which introduces an "extra

[3]Analogous remarks apply to the Burning money game, but with a twist. The player who cannot burn money will have two strategies in her unique fully permissible set; at the information set reached if money is burnt a maximal strategy may prescribe any action. By contrast IEWDS permits only one specific action. This has been taken as troublesome as it may seem to suggest that burning is viewed as a "signal of a rational player's intentions", despite burning in the end being an action a rational player would never use. Our solution of the Burning money game is robust to this critique. See AD for more discussion of and details about the Burning money game.



	x	b	s
X	$\frac{3}{2}, \frac{3}{2}$	2, 2	2, 2
B	$\frac{3}{2}, \frac{3}{2}$	3, 1	0, 0
S	$\frac{3}{2}, \frac{3}{2}$	0, 0	1, 3

FIGURE 10. Γ_7 and its PRSF.

outside option" for player 2. In this game there may seem to be a tension between forward and backward induction: For player 2 not to choose x may seem to suggest that 2 "signals" that she seeks a payoff of at least $\frac{3}{2}$, in contrast to the payoff of 1 that she gets when the subgame structured like Γ_6 is considered in isolation (as seen in the analysis of Γ_6). However, this intuition is not quite supported by the concept of fully permissible sets. Applying our algorithm IECFA to the PRSF of Γ_7 yields:

$$\begin{aligned}
\Xi(0) &= \Sigma_1 \times \Sigma_2 \\
\Xi(1) &= \{\{X\}, \{B\}, \{X, B\}\} \times \{\{x\}, \{s\}, \{x, b\}, \{x, s\}, \{b, s\}, \{x, b, s\}\} \\
\Xi(2) &= \{\{X\}, \{B\}, \{X, B\}\} \times \{\{x\}, \{x, b\}, \{b, s\}\} \\
\Xi(3) &= \{\{B\}, \{X, B\}\} \times \{\{x\}, \{x, b\}, \{b, s\}\} \\
\Pi = \Xi(4) &= \{\{B\}, \{X, B\}\} \times \{\{x\}, \{x, b\}\}
\end{aligned}$$

The only possibility for X being a maximal strategy for player 1 is that he deems $\{x\}$ as the only possible candidate for 2's choice set, in which case 1's choice set is $\{X, B\}$. Else $\{B\}$ is 1's choice set. Furthermore, 2 can have a choice set different from $\{x\}$ only if she deems $\{X, B\}$ as a possible candidate for 1's choice set. Intuitively this means that if 2's choice set differs from $\{x\}$ (i.e. equals $\{x, b\}$), then she deems it possible that 1 considers it impossible that b is a maximal strategy for 2. Since it is only under such circumstances that b is a maximal element for 2, perhaps this strategy is better thought of in terms of "strategic manipulation" than in terms of "forward induction". Note that the concept of fully permissible sets has more bite than the Dekel-Fudenberg procedure; in addition to the strategies appearing in fully permissible sets also s survives the Dekel-Fudenberg procedure.

4.5. **Prisoners' Dilemma.** As a final application, consider a 3-period prisoners' dilemma game with each player's set of actions being {*cooperate*, *defect*} in each stage. The payoffs of the stage game are given as follows, using Aumann's [3] description: Each player decides whether he will receive 1 (*defect*) or the other will receive 3 (*cooperate*). There is no discounting. Hence, the action *defect* is strongly dominant in the stage game, but still,

	s_2^{NT}	s_2^{NV}	s_2^{NE}	s_2^{RT}	s_2^{RV}	s_2^{RE}
s_1^{NT}	7, 7	4, 8	4, 8	5, 5	2, 6	2, 6
s_1^{NV}	8, 4	5, 5	5, 5	4, 8	1, 9	1, 9
s_1^{NE}	8, 4	5, 5	5, 5	5, 5	2, 6	2, 6
s_1^{RT}	5, 5	8, 4	5, 5	3, 3	6, 2	3, 3
s_1^{RV}	6, 2	9, 1	6, 2	2, 6	5, 5	2, 6
s_1^{RE}	6, 2	9, 1	6, 2	3, 3	6, 2	3, 3

FIGURE 11. Reduced form of the 3-period PD game.

each player is willing to *cooperate* in one stage if this induces the other player to *cooperate* instead of *defect* in the next stage. It follows from Prop. 3 that we need only consider (what Rubinstein [27] calls) plans of action.

There are 6 plans of actions for each player that survive the Dekel-Fudenberg procedure. In any of these, a player always defects in the 3rd stage, and does not always cooperate in the 2nd stage. The 6 plans of actions of each player i are denoted s_i^{NT}, s_i^{NV}, s_i^{NE}, s_i^{RT}, s_i^{RV} and s_i^{RE}, where N denotes that i is *nice* in the sense of *cooperating* in the 1st stage, where R denotes that i is *rude* in the sense of *defecting* in the 1st stage, where T denotes that i plays *tit-for-tat* in the sense of *cooperating* in the 2nd stage if and only $j \neq i$ has *cooperated* in the 1st stage, where V denotes that i plays *inverse tit-for-tat* in the sense of *defecting* in the 2nd stage if and only if $j \neq i$ has *cooperated* in the 1st stage, and where E denotes that i is *exploitive* in the sense of *defecting* in the 2nd stage independently of what $j \neq i$ has played in the 1st stage. The strategic form after elimination of all other plans of actions is given in Fig. 11. Note that none of these plans of actions are weakly dominated in the full strategic form.

Prop. 2 of AD implies that any fully permissible set is a subset of the set of strategies surviving the Dekel-Fudenberg procedure. Hence, only subsets of $\{s_i^{NT}, s_i^{NV}, s_i^{NE}, s_i^{RT}, s_i^{RV}, s_i^{RE}\}$ can be i's choice set under common certain belief of full admissible consistency. Furthermore, under common certain belief of full admissible consistency, we have for each player i that

- any choice set that contains s_i^{NT} must also contain s_i^{NE}, since s_i^{NT} is a maximal strategy only if s_i^{NE} is a maximal strategy,
- any choice set that contains s_i^{NV} must also contain s_i^{NE}, since s_i^{NV} is a maximal strategy only if s_i^{NE} is a maximal strategy,
- any choice set that contains s_i^{RT} must also contain s_i^{RE}, since s_i^{RT} is a maximal strategy only if s_i^{RE} is a maximal strategy,
- any choice set that contains s_i^{RV} must also contain s_i^{RE}, since s_i^{RV} is a maximal strategy only if s_i^{RE} is a maximal strategy,

Given that the choice set of the opponent satisfies these conditions, this implies that

- if s_i^{NE} is included in i's choice set, only the following sets are candidates for i's choice set: $\{s_i^{NT}, s_i^{NE}, s_i^{RT}, s_i^{RE}\}$, $\{s_i^{NV}, s_i^{NE}, s_i^{RV}, s_i^{RE}\}$, or $\{s_i^{NE}, s_i^{RE}\}$. The reason is that s_i^{NE} is a maximal strategy only if i considers it possible that j's choice set contains s_j^{NT} (and hence, s_j^{NE}) or s_j^{RT} (and hence, s_j^{RE}).
- if s_i^{RE}, but not s_i^{NE}, is included in i's choice set, only the following sets are candidates for i's choice set: $\{s_i^{RT}, s_i^{RE}\}$, $\{s_i^{RV}, s_i^{RE}\}$, or $\{s_i^{RE}\}$. The reason is that s_i^{RE} is a maximal strategy only if i considers it possible that j's choice set contains s_j^{NV}, s_j^{NE}, s_j^{RV}, or s_j^{RE}.

This in turn implies that

- i's choice set does not contain s_i^{NV} or s_i^{RV} since any candidate for j's choice set contains s_j^{RE}, implying that s_i^{NE} is preferred to s_i^{NV} and s_i^{RE} is preferred to s_i^{RV}.

Hence, the only candidates for i's choice set under common certain belief of full admissible consistency are $\{s_i^{NT}, s_i^{NE}, s_i^{RT}, s_i^{RE}\}$, $\{s_i^{NE}, s_i^{RE}\}$, $\{s_i^{RT}, s_i^{RE}\}$, and $\{s_i^{RE}\}$. Moreover, it follows from Prop. 1(iii) that all these sets are indeed fully permissible since

- $\{s_i^{NT}, s_i^{NE}, s_i^{RT}, s_i^{RE}\}$ is i's choice set if he deems $\{s_j^{RT}, s_j^{RE}\}$, but not $\{s_j^{NE}, s_j^{RE}\}$ and $\{s_j^{NT}, s_j^{NE}, s_j^{RT}, s_j^{RE}\}$, as possible candidates for j's choice set,
- $\{s_i^{NE}, s_i^{RE}\}$ is i's choice set if he deems $\{s_j^{NT}, s_j^{NE}, s_j^{RT}, s_j^{RE}\}$ as a possible candidate for j's choice set,
- $\{s_i^{RT}, s_i^{RE}\}$ is i's choice set if he deems $\{s_j^{RE}\}$ as the only possible candidate for j's choice set,
- $\{s_i^{RE}\}$ is i's choice set if he deems $\{s_j^{NE}, s_j^{RE}\}$, but not $\{s_j^{RT}, s_j^{RE}\}$ and $\{s_j^{NT}, s_j^{NE}, s_j^{RT}, s_j^{RE}\}$, as possible candidates for j's choice set.

While play in accordance with strategies surviving the Dekel-Fudenberg procedure does not provide any prediction other than both players *defecting* in the 3rd stage, the concept of fully permissible sets has more bite. In particular, a player *cooperates* in the 2nd stage only if the opponent has *cooperated* in the 1st stage. This implies that only the following paths can be realized if players choose strategies in fully permissible sets:

((*cooperate*, *cooperate*), (*cooperate*, *cooperate*), (*defect*, *defect*))

((*cooperate*, *cooperate*), (*cooperate*, *defect*), (*defect*, *defect*)) and v.v.

((*cooperate*, *defect*), (*defect*, *cooperate*), (*defect*, *defect*)) and v.v.

((*cooperate*, *cooperate*), (*defect*, *defect*), (*defect*, *defect*))

((*cooperate*, *defect*), (*defect*, *defect*), (*defect*, *defect*)) and v.v.

((*defect*, *defect*), (*defect*, *defect*), (*defect*, *defect*)).

That the path ((*cooperate*, *defect*), (*cooperate*, *defect*), (*defect*, *defect*)) or v.v. cannot be realized if players choose strategies in fully permissible sets can be interpreted as an indication that the present analysis seems to produce some element of reciprocity in the 3-period prisoners' dilemma game.

5. Concluding Remark

In this paper we apply to extensive games AD's concept of fully permissible sets, and explore its implications in several examples. AD characterize this concept as choice sets under common certain belief of full admissible consistency. Full admissible consistency entails that one strategy is preferred to another if and only the former weakly dominates the latter on the union of the choice sets that are deemed possible for the opponent (this property is called 'full belief of opponent rationality'), or on the set of all opponent strategies (this corresponds to 'caution'). Hence, full admissible consistency is associated with certain properties of *preferences*. In closing the paper we would like to emphasize that the full belief of opponent rationality relates to the strategy choices of the opponents *in the whole game*. It does not relate to choices among the remaining available strategies *at each and every information set*.

To illustrate this point, look back at Γ_2. Conditional on 2's node being reached it is clear that 1 cannot be choosing a strategy that is maximal given his preferences. Conditional on 2's node being reached, our modeling then imposes no constraint on 2's assessment of likelihood concerning which non-maximal strategy FF or FD that 1 has chosen. Note here how crucially the analysis presumes that 2 assesses the likelihood of different strategies as chosen by player 1 *in the whole game*.

It is possible to imagine a distinct modeling approach which relates to choices among the remaining available strategies *at different information sets*. In Γ_2 this would amount to the following: Conditional on 2's node being reached she realizes that 1 cannot be choosing a strategy which is maximal given his preferences. However, 2 considers it infinitely more likely that 1 at his last node chooses a strategy that is maximal among his remaining available strategies given his conditional preferences at that node. In the introduction we argued (with Ben-Porath [9]) that it is not intuitively clear that this is reasonable, a view which permeates the working hypotheses on which the current work in grounded. Yet, the alternative approach is logically conceivable, and research on this basis may be illuminating and worthwhile. However, we leave for other contributions to go in this alternative direction.[4]

References

[1] Asheim, G.B.: On the epistemic foundation for backward induction. Memorandum No. 30, Dept. of Economics, University of Oslo (1999).

[2] Asheim, G.B., Dufwenberg, M.: Admissibility and common belief. University of Oslo (2000).

[4]As a contribution in this other direction, Asheim [1] provides an epistemic foundation for backward induction in generic perfect information games by imposing as a consistency principle that each player believes in each subgame that his opponent chooses rationally in the subgame.

[3] Aumann, R.: Game theory. In "The New Palgrave: A Dictionary of Economics" (Eatwell, Milgate, Newman Eds.), Macmillan Press, London and Basingstoke (1987).
[4] Aumann, R.: Backward induction and common knowledge of rationality. Games Econ. Beh. **8** (1995) 6–19.
[5] Basu, K.: On the non-existence of a rationality definition for extensive games. Int. J. of Game Theory **19** (1990), 33–44.
[6] Battigalli, P.: Strategic rationality orderings and the best rationalization principle. Games Econ. Beh. **13** (1996) 178–200.
[7] Battigalli, P.: On rationalizability in extensive games. J. Econ. Theory **74** (1997) 40–61.
[8] Battigalli, P., Siniscalchi, M.: Interactive beliefs and forward induction. Princeton University (2000).
[9] Ben-Porath, E.: Rationality, Nash equilibrium, and backwards induction in perfect information games. Rev. Econ. Stud. **64** (1997) 23–46.
[10] Ben-Porath, E., Dekel, E.: Coordination and the potential for self-sacrifice. J. Econ. Theory **57** (1992) 36–51.
[11] Bernheim, D.: Rationalizable strategic behavior. Econometrica **52** (1984) 1007–1028.
[12] Bicchieri, C.: Self-refuting theories of strategic interaction: A paradox of common knowledge. Erkenntnis **30** (1989) 69–85.
[13] Binmore, K.: Modelling rational players I. Econ. and Phil. **3** (1987) 179–214.
[14] Binmore, K.: Backward induction and rationality. DP 95–10, University College London (1995).
[15] Blume, L., Brandenburger, A., Dekel, E.: Lexicographic probabilities and choice under uncertainty. Econometrica **59** (1991) 61–79.
[16] Börgers, T., Samuelson, L.: 'Cautious' utility maximization and iterated weak dominance, Int. J. Game Theory **21** (1992) 13–25.
[17] Dekel, E., Fudenberg, D.: Rational behavior with payoff uncertainty. J. Econ. Theory **52** (1990) 243–267.
[18] Gul, F.: Rationality and coherent theories of strategic behavior. J. Econ. Theory **70** (1997) 1–31.
[19] Hammond, P.: Aspects of rationalizable behavior. In "Frontiers of Game Theory" (Binmore, Kirman, Tani Eds.), MIT Press, Cambridge, MA (1993).
[20] Kreps, D., Wilson, R.: Sequential equilibria. Econometrica **50** (1993) 863–894.
[21] Luce, D., Raiffa, H.: "Games and Decisions", Wiley, New York, NY (1957).
[22] Mailath, G., Samuelson, L., Swinkels, J.: Extensive form reasoning in normal form games. Econometrica **61** (1993) 273–302.
[23] Pearce, D.G.: Rationalizable strategic behavior and the problem of perfection. Econometrica **52** (1984) 1029–1050.
[24] Reny, P.J.: Backward induction, normal form perfection and explicable equilibria. Econometrica **60** (1992) 627–649.
[25] Reny, P.J.: Common belief and the theory of games with perfect information. J. Econ. Theory **59** (1993) 257–274.
[26] Rosenthal, R.: Games of perfect information, predatory pricing and the chain-store paradox. J. Econ. Theory **25** (1981) 92–100.
[27] Rubinstein, A.: Comments on the Interpretation of Game Theory. Econometrica **59** (1991) 909–924.
[28] Samuelson, L.: Dominated strategies and common knowledge. Games Econ. Beh. **4** (1992) 284–313.
[29] van Damme, E.: Stable equilibria and forward induction. J. Econ. Theory **48** (1989) 476–496.

Department of Economics, University of Oslo, P.O. Box 1095 Blindern, N-0317 Oslo, Norway; Department of Economics, Stockholm University, SE-10691 Stockholm, Sweden

E-mail address: g.b.asheim@econ.uio.no, md@ne.su.se