Self-Admissible Sets[1]

Adam Brandenburger[2]   Amanda Friedenberg[3]

October 2004
September 2009

[1]Some of the results here were presented in two working papers, titled "When Does Common Assumption of Rationality Yield a Nash Equilibrium?" (July 2001) and "Common Assumption of Rationality in Games" (January 2002). We owe much to joint work with H. Jerome Keisler. We are also indebted to Pierpaolo Battigalli, Drew Fudenberg, John Nachbar, Martin Osborne, Marciano Siniscalchi, Gus Stuart, the associate editor, and the referees for important comments. Geir Asheim, Konrad Grabiszewski, Elon Kohlberg, Alex Peysakhovich, and participants at presentations of this work provided valuable input. Brandenburger thanks Harvard Business School and the Stern School of Business for their support. Friedenberg thanks the Olin Business School and the W.P. Carey School of Business for their support.

[2]Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012; abranden@stern.nyu.edu; www.stern.nyu.edu/∼abranden

[3]W.P. Carey School of Business, Arizona State University, Tempe, AZ 85287, amanda.friedenberg@asu.edu, www.public.asu.edu/∼afrieden/

Abstract

Best-response sets (Pearce [29, 1984]) characterize the epistemic condition of "rationality and common belief of rationality." When rationality incorporates a weak-dominance (admissibility) requirement, the self-admissible set (SAS) concept (Brandenburger-Friedenberg-Keisler [18, 2008]) characterizes "rationality and common assumption of rationality." We analyze the behavior of SAS's in some games of interest–Centipede, the Finitely Repeated Prisoner's Dilemma, and Chain Store. We then establish some general properties of SAS's, including a characterization in perfect-information games.

Keywords: Admissibility, Weak Dominance, Self-Admissible Sets, Iterated Admissibility, Epistemic Game Theory, Perfect-Information Games

JEL Codes: C72, D80

1 Introduction

Consider the condition that Ann is rational, she thinks Bob is rational, she thinks he thinks she is rational, and so on. In this case, what strategies will Ann choose? This is a basic question in the epistemic approach to game theory. It has been asked when "rationality" means: (i) ordinary (subjective expected utility) maximization on the matrix, (ii) admissibility (avoidance of weakly dominated strategies) on the matrix, and (iii) maximization at each information set in the tree.

There is a very intuitive answer for (i): Ann will choose an **iteratively undominated** (**IU**) strategy–i.e., a strategy which survives the iterated elimination of strongly dominated strategies. The idea goes back to Bernheim [10] and Pearce [29] (though they made an additional independence assumption). There is also a deeper answer, pioneered by Pearce. He introduced the concept of a **best-response set** (**BRS**): This is a subset $Q^a \times Q^b \subseteq S^a \times S^b$ (where S^a and S^b are Ann's and Bob's strategy sets) such that each of Ann's strategies in Q^a is undominated with respect to Q^b, and likewise with Ann and Bob interchanged.

Here is a sketch of how a formal epistemic analysis leads to a BRS. The first step is to add types (for Ann and Bob) to the description of the game. A particular type for Ann describes what she thinks about which strategy Bob chooses, what she thinks Bob thinks about which strategy she chooses, and so on. Likewise with Bob's types. With these ingredients, we can identify those strategy-type pairs for Ann which are rational, believe (i.e., assign probability 1 to the event) that Bob is rational, and so on. This is the set labelled RCBR ("rationality and common belief of rationality") in the left-hand panel in Figure 1.1. Applying a similar analysis for Bob leads to the right-hand panel.

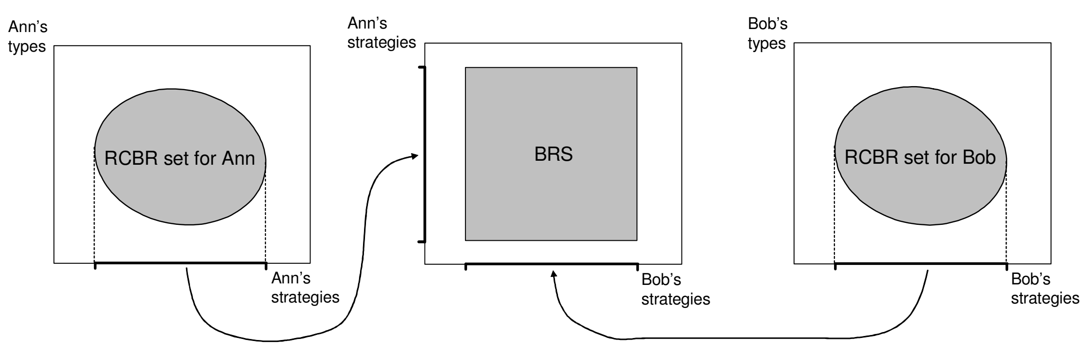

Figure 1.1

The next step is to see what this analysis implies in terms of strategies that can be played. Formally, we project the RCBR sets into their respective strategy sets. This gives the middle panel in Figure 1.1. Can we say what the projections are? Yes, they constitute a BRS. This is the deeper answer to question (i) above.

Two observations. First, a given game can have several BRS's. A particular epistemic analysis will yield one such BRS–which one will depend on what type spaces we happen to begin with.

Second, the greater precision in saying we get a BRS–rather than just saying that the strategies are IU–may seem a small matter. It is well known that every BRS is contained in the IU set, which is itself the largest BRS.

But, the greater precision becomes crucial in the case of admissibility–i.e., when we address question (ii) above. The obvious guess for the answer is that Ann will choose an **iteratively admissible** (**IA**) strategy–i.e., a strategy which survives the iterated elimination of inadmissible strategies. But this is wrong!

Here are the steps in order to see why. Brandenburger-Friedenberg-Keisler [18] (henceforth, BFK) formulate an analysis like that in Figure 1.1. Now, Ann is rational in the sense of admissibility, she assumes Bob is rational in the same sense, and so on. (Here, we say "assumes" rather than "believes." See Sections 3 and 8b.) Likewise for Bob. The middle panel in Figure 1.1 now becomes a concept called a **self-admissible set** (**SAS**) in [18]: This is a subset $Q^a \times Q^b \subseteq S^a \times S^b$ such that each of Ann's strategies in Q^a is admissible with respect to both S^b and Q^b, and, in addition, Q^a satisfies a maximality condition. Likewise with Ann and Bob interchanged. (Details will be given later.)

Note the analogy to BRS. (Obviously, if one of Ann's strategies in Q^a is not strongly dominated with respect to Q^b, it is also undominated with respect to S^b. With admissibility, the two conditions must be explicitly stated. The maximality condition is not a point of difference–again, details later.)

But–and this is the key point–unlike with BRS's and the IU set, an SAS need not be contained in the IA set. Indeed, while the IA set of a game constitutes one SAS of that game, there can be other SAS's which are even disjoint from it.

This is why a separate analysis of SAS's is in order. Much is known about the behavior of IA in various games, but not about SAS. This paper aims to fill in the picture. We preview our investigation in Section 2.

		Bob *L*	Bob *C*	Bob *R*
Ann	*U*	1, 1	1, 1	0, 0
Ann	*M*	1, 1	0, 0	1, 0
Ann	*D*	0, 0	0, 1	0, 0

Figure 1.2

Before that, a little more on the foundation of the concept. In particular, why is the SAS-IA relationship different from the BRS-IU relationship? The reason is a basic non-monotonicity in admissibility: The component-by-component union of two SAS's need not be an SAS. Consider the game in Figure 1.2.[1] There are five SAS's: $\{(U, L)\}$, $\{(U, C)\}$, $\{U\} \times \{L, C\}$, $\{(M, L)\}$, and $\{U, M\} \times$

[1] We thank a referee for this example. It is similar to examples in Asheim-Dufwenberg [2].

$\{L\}$. But $\{U, M\} \times \{L, C\}$ is not an SAS. Kohlberg-Mertens [24, p.1017] gave the 'philosophical' explanation (albeit in a different context): Under admissibility, adding new possibilities can change previously good strategies into bad ones. While M (resp. C) is admissible with respect to $\{L\}$ (resp. $\{U\}$), it becomes inadmissible once C (resp. M) is added. This appears to be the fundamental reason for the greater complexity of SAS vs. BRS theory. We expand on this point in Section 8b.

Finally, let us return to the third version of the question we posed at the outset–concerning epistemic analysis of rationality in the tree. This is not the subject of this paper, but there is a parallel. Two concepts arise in this case: extensive-form rationalizability (Pearce [29, 1984], Battigalli-Siniscalchi [8]) and extensive-form best-response sets (Battigalli-Friedenberg [7]). They are the analogs to IU-BRS and IA-SAS. But, unlike IU-BRS and again like IA-SAS, there is a non-monotonicity: Extensive-form best-response sets aren't necessarily contained in the extensive-form rationalizable set.

2 Preview

In Section 3 we give the formal definition of an SAS and review epistemic foundations of the concept. Section 4 studies how SAS behaves in some of the most commonly studied games–Centipede, the Finitely Repeated Prisoner's Dilemma, and Chain Store. In Sections 5-7 we develop some general properties, first in the strategic form and second in the extensive form–including a characterization in perfect-information games. Section 8 contains some conceptual discussion and also covers related work.

Here, we make two preliminary comments on the philosophy underlying our approach. First, SAS is a strategic-form concept, yet the bulk of the paper investigates its behavior in extensive-form games. Why? The easy answer is that we are following the Kohlberg-Mertens [24, 1986, Sections 2.4-2.5] invariance doctrine: Our analysis should depend only on the strategic form, even if our primary interest is in behavior in game trees.

Recall that Dalkey [20], Thompson [35], and Elmes-Reny [22] showed that two trees have the same reduced strategic form–i.e., the strategic form after elimination of duplicate rows and columns–if and only if they differ by a sequence of four elementary transformations.[2] Kohlberg-Mertens added a fifth convex-combination transformation. If all five transformations are viewed as "irrelevant for correct decision making" ([24, p.1011]), then we get a requirement of invariance to the fully reduced strategic form–i.e., invariance to the strategic form after elimination of convex combinations. We show that SAS satisfies this full invariance requirement (Proposition 5.2).

In fact, there is a deeper level to the invariance issue. Epistemics are a tool to formalize reasoning in games. If the idea of invariance is that reasoning should not change across equivalent games, then shouldn't invariance be stated at the level of the epistemics?[3] The answer, surely, is yes. We would

[2] The Elmes-Reny transformations differ from the Dalkey-Thompson ones by preserving perfect recall throughout.
[3] We are grateful to a referee for raising this point.

like to have a full-blown principle of "epistemic invariance" and to be able to investigate whether various solution concepts–SAS included–satisfy it. See Section 8f for further discussion.

Our second preliminary comment concerns the broader question of what we hope to learn by investigation of SAS–or, indeed, some other epistemically derived concept. Our view is that a primary purpose of epistemic game theory is to formalize intuitive notions which one has about strategic situations–notions such as that of a "best (or rational?) course of action," or that of the importance of "thinking about what the other player is thinking," and so on. The goal is that, by carrying out such formalization and by investigating properties of the resulting concepts, we improve our understanding of the underlying strategic situations. In this paper we take as given that the SAS concept is an embodiment of certain intuitions (following BFK [18]). Our goal here is the next step–to uncover properties of the concept.

3 Self-Admissible Sets

We now give a formal definition of the SAS concept and sketch its epistemic basis. To begin, fix a two-player strategic-form game $\langle S^a, S^b; \pi^a, \pi^b \rangle$, where S^a, S^b are the (finite) strategy sets and π^a, π^b are payoff functions for Ann and Bob, respectively. (We focus on two players, but our analysis extends readily to n players.) The definitions to come all have counterparts with a and b reversed.

Given a finite set X, let $\mathcal{M}(X)$ denote the set of all probability measures on X. We extend π^a to $\mathcal{M}(S^a) \times \mathcal{M}(S^b)$ in the usual way, i.e. $\pi^a(\sigma^a, \sigma^b) = \sum_{s^a \in S^a} \sum_{s^b \in S^b} \sigma^a(s^a)\sigma^b(s^b)\pi^a(s^a, s^b)$. Throughout, we adopt the convention that in a product $X \times Y$, if $X = \emptyset$ then $Y = \emptyset$ (and vice versa).

Definition 3.1 *Fix $X \times Y \subseteq S^a \times S^b$ and some $s^a \in X$. Say s^a is* ***weakly dominated with respect to*** *$X \times Y$ if there exists $\sigma^a \in \mathcal{M}(S^a)$, with $\sigma^a(X) = 1$, such that $\pi^a(\sigma^a, s^b) \geq \pi^a(s^a, s^b)$ for every $s^b \in Y$, and $\pi^a(\sigma^a, s^b) > \pi^a(s^a, s^b)$ for some $s^b \in Y$. Otherwise, say s^a is* ***admissible with respect to*** *$X \times Y$. If s^a is admissible with respect to $S^a \times S^b$, simply say that s^a is* ***admissible****.*

Definition 3.2 *Fix $X \subseteq S^a$. Say $s^a \in X$ is* ***optimal under*** *$\mu^b \in \mathcal{M}(S^b)$* ***given*** *$X$ if $\pi^a(s^a, \mu^b) \geq \pi^a(r^a, \mu^b)$ for all $r^a \in X$. If $s^a \in S^a$ is optimal under μ^b given S^a, simply say that s^a is* ***optimal under*** *μ^b.*

Remark 3.1 *A strategy s^a is admissible with respect to $X \times Y$ if and only if there exists $\mu^b \in \mathcal{M}(S^b)$, with $\operatorname{Supp} \mu^b = Y$, such that s^a is optimal under μ^b given X.*

Next, we need:

Definition 3.3 *Fix some $s^a \in S^a$ and suppose there is $\varphi^a \in \mathcal{M}(S^a)$ with $\pi^a(\varphi^a, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S^b$. Then if $r^a \in \operatorname{Supp} \varphi^a$, say r^a* ***supports*** *s^a (****via*** *φ^a). Write $\operatorname{su}(s^a)$ for the set of $r^a \in S^a$ that support s^a.*

In words, this says that su(s^a) consists of those strategies for Ann that are part of some convex combination equivalent (for her) to s^a. With this, we can give the definition of a self-admissible set.

Definition 3.4 *Fix $Q^a \times Q^b \subseteq S^a \times S^b$. The set $Q^a \times Q^b$ is a **self-admissible set (SAS)** if:*

(i) each $s^a \in Q^a$ is admissible with respect to $S^a \times S^b$;

(ii) each $s^a \in Q^a$ is admissible with respect to $S^a \times Q^b$;

(iii) for any $s^a \in Q^a$, if $r^a \in \mathrm{su}(s^a)$ then $r^a \in Q^a$;

and likewise for each $s^b \in Q^b$.

Definition 3.4 brings out the analogy to best-response sets (BRS's). To repeat from earlier, a strong-dominance version of condition (i) is implied by a strong-dominance version of condition (ii). For weak dominance, we need to stipulate the additional condition. Condition (iii) could be added to the definition of a BRS. It is without loss of generality in the following sense: Any BRS $Q^a \times Q^b$ not satisfying condition (iii) is contained in a larger BRS that does satisfies condition (iii).[4] (This is a consequence of Lemma A1 below.) By contrast, a set $Q^a \times Q^b$ satisfying only conditions (i)-(ii) of Definition 3.4 may not be contained in any SAS. (See BFK [18, Section 2.3] for an example.)

Let us briefly review how an epistemic analysis of admissibility leads to the SAS concept. The analysis starts from a basic challenge for admissibility in games, identified by Samuelson [31]. On the one hand, admissibility requires that, if Ann is rational, she should not rule out any of Bob's strategies (per Remark 3.1). On the other hand, if Ann thinks Bob is rational, then she should rule out the possibility that Bob plays an inadmissible strategy. There appears to be a conflict between the requirements that: (i) Ann is rational, and (ii) Ann thinks that Bob is rational. An epistemic analysis of admissibility must face this tension.

BFK [18] resolve the tension by asking that Ann consider it infinitely less likely–but not impossible–that Bob is irrational vs. rational. Ann is equipped with a lexicographic probability system (LPS) on Bob's strategies and types–i.e., on $S^b \times T^b$. This is a sequence of measures $(\mu_1, \ldots, \mu_n)$, where μ_1 represents Ann's primary hypothesis, μ_2 represents her secondary hypothesis, and so on. (See Blume-Brandenburger-Dekel [11].) Now, Ann can consider one strategy-type pair (s^b, t^b) for Bob infinitely more likely than another pair (r^b, u^b)—e.g., if μ_1 assigns probability 1 to (s^b, t^b), while μ_2 assigns positive probability to (r^b, u^b). More generally, say Ann **assumes** an event $E \subseteq S^b \times T^b$ if, under her LPS, all of E is infinitely more likely than all of not-E.

With these ingredients, the epistemic condition of rationality and common assumption of rationality (**RCAR**) is expressible. (A strategy-type pair (s^a, t^a) is **rational** if t^a is associated with a full-support LPS and s^a lexicographically maximizes the sequence of expected payoffs Ann gets under the LPS associated with t^a.) Theorem 8.1 in BFK says that RCAR is characterized by the

[4] In private communication, David Pearce told us that he was aware of the maximality condition, but, given this property, did not include it in his definition. In fact, if the definition of BRS is derived epistemically, per Figure 1.1, maximality would be automatically incorporated.

SAS concept. That is, fixing a game and a type structure (analogous to the one in Figure 1.1), the strategies that are played under RCAR constitute an SAS of the game. Conversely, fixing a game and an SAS of the game, there is a type structure so that the strategies played under RCAR are those in the SAS. In Section 8c, we will discuss other epistemic analyses of admissibility.

As we proceed to investigate properties of SAS's, we will compare what we find with properties of the IA set. (We take IA to mean simultaneous maximal deletion.) This is a natural comparison since the two concepts are related at the epistemic level. (Again, see BFK [18].) Here is the formal definition:

Definition 3.5 *Set $S^i_0 = S^i$ for $i = a, b$, and define inductively*

$$S^i_{m+1} = \{s^i \in S^i_m : s^i \text{ is admissible with respect to } S^a_m \times S^b_m\}.$$

*A strategy $s^i \in S^i_m$ is called m-**admissible**. A strategy $s^i \in \bigcap_{m=0}^{\infty} S^i_m$ is called **iteratively admissible (IA)**.*

Since the game is finite, there is an M such that $S^i_M = \bigcap_{m=0}^{\infty} S^i_m$ for $i = a, b$. Moreover, each S^i_M is nonempty and therefore the IA set is nonempty.

4 Applications

We now begin our investigation of properties of SAS's. In this section, we look at what the SAS concept gives in three canonical examples–Centipede (Rosenthal [30]), the Finitely Repeated Prisoner's Dilemma, and Chain Store (Selten [32]). This will give some pointers to general properties to investigate in the following sections.

Example 4.1 (Centipede) *Consider n-legged Centipede, as in Figure 4.1. If $Q^a \times Q^b$ is an SAS of Centipede, and $(s^a, s^b) \in Q^a \times Q^b$, then s^a is Ann's strategy of playing Out at the first node.*[5]

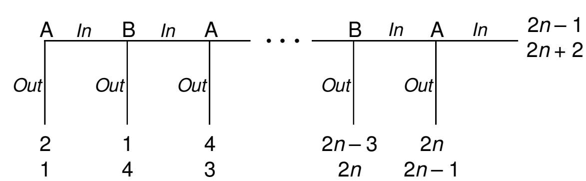

Figure 4.1

Indeed, suppose, to the contrary, that there is an $(s^a, s^b) \in Q^a \times Q^b$ where s^a involves Ann's playing In at the first node. In particular, pick a profile (s^a, s^b) which yields the longest path of play (before Ann or Bob plays Out). (The fact that some player does play Out on this path follows from

[5] We consider the reduced strategic form of the game. This will suffice, given Proposition 5.2 to come.

admissibility–i.e., condition (i) of the definition of an SAS.) Let h be the node on this path at which Out is played. Suppose Bob moves at h. (A similar argument applies if Ann moves at h.) Then, by condition (ii) of the definition of an SAS, and Remark 3.1, Ann's strategy s^a must be optimal under a measure that assigns: (i) probability 1 to Bob's playing Out at node h or earlier; and (ii) positive probability to Bob's playing In until node h and Out at h. Now consider the strategy r^a for Ann that plays In until node h' (where h' is the immediate predecessor of h) and plays Out at h'. Then r^a does strictly better than s^a under any such measure–a contradiction.

The IA set for Centipede is easily verified to be $\{(Out, Out)\}$, and $\{(Out, Out)\}$ is also an SAS of the game. Thus, in particular, the set of SAS's is nonempty.

This analysis of Centipede seems very intuitive. It starts at the root of the tree and works forwards to reach a contradiction: If it is Bob who ends the game (by playing *Out* at node h), then Ann should have ended the game earlier.[6] (Of course, our goal is not to defend *Out* as the only rational choice. Indeed, the analysis says that RCAR–a much stronger condition–yields this conclusion.) The next example also involves a forward-looking argument.[7]

Example 4.2 (Finitely Repeated Prisoner's Dilemma) *Consider the Prisoner's Dilemma, as in Figure 4.2. Fix an SAS of the game played T times (for some integer T). Any strategy profile in the SAS yields the Defect-Defect path throughout.*

A \ B	*C*	*D*
C	*c*, *c*	*e*, *d*
D	*d*, *e*	0, 0

$d > c > 0 > e$

Figure 4.2

The proof will use a projection property of SAS's established later (Proposition 6.2). We argue by induction on the number of rounds. For $T = 1$, the result is immediate from the fact that any strategy in an SAS is admissible (a fortiori, not strongly dominated). Now, assume the result for T, and fix an SAS $Q^a \times Q^b$ of the $(T+1)$-fold game. Suppose $s^a \in Q^a$ involves Ann's playing C on the first round. Then, for any $s^b \in Q^b$, Ann gets a first-round payoff of c if s^b involves Bob's playing C on the first round, and e if s^b involves Bob's playing D on the first round. These are also Ann's total payoffs from the game when (s^a, s^b) is played, since the induction hypothesis, together with the fact that SAS's induce SAS's on subtrees (see Proposition 6.2 for a precise statement), implies that

[6] The same proof applies to Nash equilibrium. If (σ^a, σ^b) is an equilibrium of Centipede, then σ^a puts probability 1 on Ann's playing *Out* at the first node. Just apply our argument to $\mathrm{Supp}\,\sigma^a \times \mathrm{Supp}\,\sigma^b$.

[7] Other papers–albeit with different epistemics–employ forward-looking arguments. For Centipede, see Aumann [4]. For FRPD, see Stuart [34].

the profile (s^a, s^b) *must yield the Defect-Defect path on rounds* $2, \ldots, T+1$. *Suppose instead that Ann chooses the "Defect always" strategy. Then she gets a first-round payoff of* d *if* s^b *involves Bob's playing* C *on the first round, and* 0 *if* s^b *involves Bob's playing* D *on the first round. On subsequent rounds Ann gets at least* 0. *But then the "Defect always" strategy does strictly better than* s^a *against every* $s^b \in Q^b$, *contradicting the definition of an SAS.*

The IA set for FRPD consists of a unique strategy pair, where each player chooses Defect regardless of history.

Example 4.3 (Chain Store) *Consider the version of Chain Store in Figure 4.3. By admissibility, the unique SAS is* $\{(In, Cede)\}$.

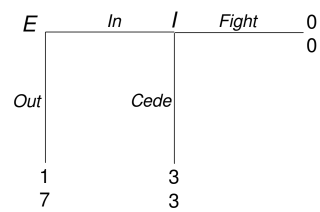

Figure 4.3

Now consider twice-repeated Chain Store. The tree is given in Figure 4.4 and the (reduced) strategic form in Figure 4.5.

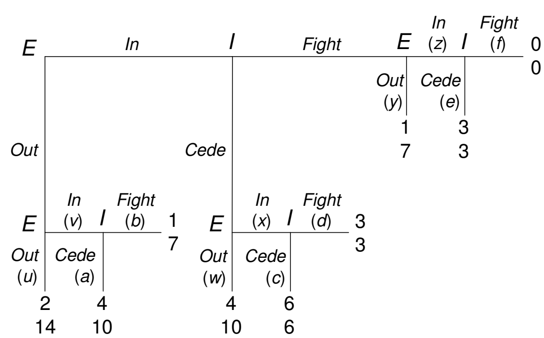

Figure 4.4

For the entrant E, *all strategies are admissible. For the incumbent* I, *only strategies ac and ae are admissible. From this, we can see that if* $Q^E \times Q^I$ *is an SAS either: (i)* $Q^I = \{ac\}$ *and* $Q^E = \{xy, xz\}$ *or* $\{xz\}$*; (ii)* $Q^I = \{ae\}$ *and* $Q^E = \{v\}$*; (iii)* $Q^I = \{ac, ae\}$ *and* $Q^E = \{v\}$. *Two paths of play are possible under the SAS concept. In one,* E *enters in both periods and* I *cedes in both periods. In the other,* E *stays out in the first period (because there is sufficient chance the* *incumbent would fight) and enters only in the second period, at which point I cedes.*

	ac	ad	bc	bd	ae	af	be	bf
u	2, 14	2, 14	2, 14	2, 14	2, 14	2, 14	2, 14	2, 14
v	4, 10	4, 10	1, 7	1, 7	4, 10	4, 10	1, 7	1, 7
wy	4, 10	4, 10	4, 10	4, 10	1, 7	1, 7	1, 7	1, 7
wz	4, 10	4, 10	4, 10	4, 10	3, 3	0, 0	3, 3	0, 0
xy	6, 6	3, 3	6, 6	3, 3	1, 7	1, 7	1, 7	1, 7
xz	6, 6	3, 3	6, 6	3, 3	3, 3	0, 0	3, 3	0, 0

Figure 4.5

The IA set for twice-repeated Chain Store is the singleton $\{(xz, ac)\}$*, corresponding to enter regardless for I and cede regardless for E.*

What do these three examples tell us about the behavior of SAS's?

First, we see that, in each case, the IA set is one SAS of the game. We will show this is true in general. A consequence is that any game possesses at least one SAS.

We also see that there are SAS's distinct from the IA set. For example, $\{Out\} \times \{Out, In\}$ is an SAS of 3-legged Centipede, while the IA set is the singleton $\{(Out, Out)\}$. The difference between SAS and IA is more stark in twice-repeated Chain Store. In this game, an SAS may even give a different outcome from the IA set. There are SAS's in which the entrant stays out in the first period–not so in the IA set.

In particular, in Centipede, SAS yields the backward-induction (BI) outcome, while, in twice-repeated Chain Store, SAS allows non-BI outcomes. This is different from IA, and prompts the questions: Why does SAS yield the BI outcome in some games but not others? When does SAS yield the BI outcome? Can we characterize SAS's in perfect-information games?

In the following sections we will give answers to these and other questions about the SAS concept.

5 Strategic-Form Properties of SAS's

Here we record two basic strategic-form properties: existence and invariance. The proofs are relegated to the Appendix.

Proposition 5.1 *Any finite game possesses a nonempty SAS. In particular, the IA set is an SAS.*

Next, invariance. Referring back to our discussion in Section 2, we see that, since SAS is defined on the strategic form, we have only to show invariance with respect to the addition of strategies that are convex combinations (for all players) of existing strategies. (Of course, this covers addition of duplicate strategies.) To establish that this is true, consider games $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$ and $\overline{G} = \left\langle S^a \cup \{q^a\}, S^b, \overline{\pi}^a, \overline{\pi}^b \right\rangle$, where $q^a \notin S^a$, and:

(i) $\overline{\pi}^a | S^a \times S^b = \pi^a$ and $\overline{\pi}^b | S^b \times S^a = \pi^b$ (where "|" denotes the restriction);

(ii) there is a $\varphi^a \in \mathcal{M}(S^a)$ such that $\overline{\pi}^a(q^a, s^b) = \pi^a(\varphi^a, s^b)$ and $\overline{\pi}^b(s^b, q^a) = \pi^b(s^b, \varphi^a)$ for each $s^b \in S^b$.

Proposition 5.2

(a) Let $\overline{Q}^a \times \overline{Q}^b$ be an SAS of $\overline{G}$. Then $(\overline{Q}^a \backslash \{q^a\}) \times \overline{Q}^b$ is an SAS of G.

(b) Let $Q^a \times Q^b$ be an SAS of G. If q^a does not support any strategy in Q^a, then $Q^a \times Q^b$ is an SAS of $\overline{G}$. Otherwise, $(Q^a \cup \{q^a\}) \times Q^b$ is an SAS of $\overline{G}$.

Proposition 5.1 states that the IA set is one SAS. So, Proposition 5.2 says that the IA set remains an SAS after the addition or deletion of convex combinations. Indeed, we can go further–the IA set is also invariant to the fully reduced strategic form. (See Proposition A1 in the Appendix.)

6 Extensive-Form Properties of SAS's

We will consider extensive-form games with perfect recall (Kuhn [25], [26]). Let S^a, S^b be the strategy sets associated with an extensive form Γ.[8] Write H^a (resp. H^b) for the information sets at which Ann (resp. Bob) moves, and $S^a(h)$ (resp. $S^b(h)$) for the subset of S^a (resp. S^b) that allows information set h. Let Z be the set of terminal nodes. Let $\zeta : S^a \times S^b \to Z$ be the map from strategy profiles to terminal nodes. Extensive-form payoff functions are maps $\Pi^a : Z \to \mathbb{R}$ and $\Pi^b : Z \to \mathbb{R}$. The strategic form induced by Γ is then $G = \langle S^a, S^b, \pi^a, \pi^b \rangle$, where $\pi^a = \Pi^a \circ \zeta$ and $\pi^b = \Pi^b \circ \zeta$. (Note the abuse of notation: ζ maps $S^b \times S^a$ to Z in the definition of π^b.)

A very basic requirement of a solution concept defined on the matrix is that it should induce optimal behavior in the tree.

Definition 6.1 *A strategy $s^a \in S^a$ is **(extensive-form) rational** if, for each information set $h \in H^a$ allowed by s^a, there is some $\mu^b \in \mathcal{M}\left(S^b\right)$, with $\mu^b\left(S^b(h)\right) = 1$, under which s^a is optimal among all strategies in $S^a(h)$.*

It is a standard argument that a strategy which is admissible in the matrix is extensive-form rational in every associated tree with perfect recall. (Use the full-support measure in Remark 3.1 to build a measure μ^b at each information set h.) So, we certainly get:

[8] In light of Proposition 5.2, we can (and often do) conflate strategies with plans of action. No confusion should result.

Proposition 6.1 *Fix an extensive-form game* Γ *with induced strategic form* G. *If* $Q^a \times Q^b$ *is an SAS of* G, *then any* $s^a \in Q^a$ *(resp.* $s^b \in Q^b$*) is extensive-form rational in* Γ.

More subtle extensive-form properties involve relating what a solution concept gives on the whole tree to what it gives on parts of the tree. One such property (introduced by Kohlberg-Mertens [24, p.1012]) is projection: If a strategy profile lies in a solution for the whole tree, then it should also lie in a solution for any part of the tree that it reaches.

SAS's satisfy projection: Any SAS of game G that allows subtree Δ induces an SAS on the strategic form of Δ. We will need some notation to prove this. Fix an extensive form Γ and associated strategic form G. Let Δ be a proper subtree of Γ, with strategic form D. Let S^a_Δ (resp. S^b_Δ) be the subset of S^a (resp. S^b) consisting of those strategies that allow subtree Δ. Note that, up to duplication of strategies, we identify S^a_Δ with the set of Ann's strategies on subtree Δ. (Likewise for Bob.) Since SAS's are invariant to the addition or deletion of duplicate strategies (Proposition 5.2), we can indeed make this identification. Now the formal statement:

Proposition 6.2 *Let* $Q^a \times Q^b$ *be an SAS of* G, *and suppose that* $(Q^a \cap S^a_\Delta) \times (Q^b \cap S^b_\Delta) \neq \emptyset$. *Then* $(Q^a \cap S^a_\Delta) \times (Q^b \cap S^b_\Delta)$ *is an SAS of* D, *up to the addition of strategies that are duplicates on* Δ.

Proof. Each $s^a \in Q^a \cap S^a_\Delta$ is optimal under some $\mu^b \in \mathcal{M}\left(S^b\right)$ with $\operatorname{Supp}\mu^b = S^b$ (condition (i) of the definition of an SAS). We have $\mu^b(S^b_\Delta) > 0$, so $\mu^b(\cdot|S^b_\Delta)$ is well defined with $\operatorname{Supp}\mu^b(\cdot|S^b_\Delta) = S^b_\Delta$. Suppose s^a is not optimal under $\mu^b(\cdot|S^b_\Delta)$, i.e., there is some $r^a \in S^a_\Delta$ with

$$\sum_{s^b \in S^b_\Delta} \pi^a(r^a, s^b)\mu^b(s^b|S^b_\Delta) > \sum_{s^b \in S^b_\Delta} \pi^a(s^a, s^b)\mu^b(s^b|S^b_\Delta).$$

Define a new strategy $q^a \in S^a_\Delta$ that agrees with r^a at each information set (for a) in Δ, and agrees with s^a at information sets not in Δ. Then, using the fact that $\mu^b(S^b_\Delta) > 0$,

$$\sum_{s^b \in S^b} \pi^a(q^a, s^b)\mu^b\left(s^b\right) > \sum_{s^b \in S^b} \pi^a(s^a, s^b)\mu^b\left(s^b\right),$$

a contradiction. Thus s^a is admissible with respect to $S^a_\Delta \times S^b_\Delta$.

The argument for condition (ii) is very similar. Each $s^a \in Q^a \cap S^a_\Delta$ is optimal under some $\nu^b \in \mathcal{M}\left(S^b\right)$ with $\operatorname{Supp}\nu^b = Q^b$. Since $Q^b \cap S^b_\Delta \neq \emptyset$, we have $\nu^b(Q^b \cap S^b_\Delta) > 0$, so that $\nu^b(\cdot|Q^b \cap S^b_\Delta)$ is well defined with $\operatorname{Supp}\nu^b(\cdot|Q^b \cap S^b_\Delta) = Q^b \cap S^b_\Delta$. Suppose there is some $r^a \in S^a_\Delta$ with

$$\sum_{s^b \in S^b_\Delta} \pi^a(r^a, s^b)\nu^b(s^b|Q^b \cap S^b_\Delta) > \sum_{s^b \in S^b_\Delta} \pi^a(s^a, s^b)\nu^b(s^b|Q^b \cap S^b_\Delta).$$

Define a new strategy $q^a \in S^a_\Delta$ that agrees with r^a at each information set (for a) in Δ, and agrees with s^a at information sets not in Δ. Then

$$\sum_{s^b \in S^b} \pi^a(q^a, s^b)\nu^b\left(s^b\right) > \sum_{s^b \in S^b} \pi^a(s^a, s^b)\nu^b\left(s^b\right),$$

a contradiction.

Finally, fix $s^a \in Q^a \cap S^a_\Delta$, and suppose there is $\varphi^a \in \mathcal{M}(S^a_\Delta)$ such that $\pi^a(\varphi^a, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S^b_\Delta$. Given each $r^a \in \operatorname{Supp}\varphi^a$, let $f(r^a)$ be the strategy that agrees with r^a at each information set (of a) in Δ, and agrees with s^a at all other information sets. Define a measure $\rho^a \in \mathcal{M}(S^a)$ that where $\rho^a(f(r^a)) = \varphi^a(r^a)$, for $r^a \in \operatorname{Supp}\varphi^a$. Then $\pi^a(\rho^a, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S^b$. Thus condition (iii) applied to $Q^a \times Q^b$ implies that each $q^a \in \operatorname{Supp}\rho^a$ is in Q^a. But each $r^a \in \operatorname{Supp}\varphi^a$ agrees with $f(r^a) \in \operatorname{Supp}\rho^a$ on subtree Δ, by construction. So, by our identification of strategies, we have that condition (iii) is satisfied. ■

By Proposition 5.1, the IA set is one SAS, so, by the projection property, the projection of the IA set into a reached subtree constitutes an SAS of the subtree. Notice this does not say that IA projects to the IA set of the subtree–only to some SAS of the subtree. The game in Figure 6.1 is an example.[9] The IA set is $\{(In\text{-}D, L)\}$. But the IA set of the subtree following In consists of the entire set $\{U, D\} \times \{L, R\}$–different from the projection $\{(D, L)\}$. Of course, the projection $\{(D, L)\}$ does form an SAS, as it must. Is there a game where the projection of the IA set of the whole tree is even disjoint from the IA set on a reached subtree? We don't know and leave this as an open question.

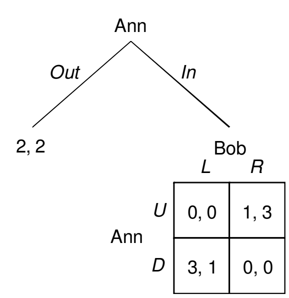

Figure 6.1

It is also important to note that a solution concept may satisfy the projection property, but, nonetheless, may not yield the BI outcome in perfect-information (PI) games. SAS is a case in point. Refer back to the twice-repeated Chain Store game (Example 4.3 in Section 4). Pure-strategy Nash equilibrium is another example: Let (s^a, s^b) be a pure-strategy profile for Γ that reaches subtree Δ. Then, if the restrictions of s^a and s^b to Δ fail to constitute a Nash equilibrium of Δ, the profile (s^a, s^b) must fail to be a Nash equilibrium of Γ.

We can go further: A solution concept may satisfy the (extensive-form) rationality and projection properties, but may not give the BI outcome. Again, SAS is a case in point. So is pure-strategy

[9] We thank a referee for the example.

Nash equilibrium in extensive-form rational strategies.

If not BI, what does SAS yield in general PI games? The parallel between SAS and Nash equilibrium gives a clue. So do the analyses of the Centipede and Chain Store games (Examples 4.1 and 4.3). In any SAS of Centipede, Ann plays *Out* immediately. Footnote 6 pointed out that the same is true of any Nash equilibrium of Centipede (and that the proof is the same). In the once-repeated Chain Store game, there is a Nash equilibrium in which the entrant stays out and the incumbent fights. But, this involves an inadmissible strategy for the incumbent. In the unique admissible equilibrium, the entrant enters and incumbent cedes–the same as in the unique SAS. In twice-repeated Chain Store, there is an SAS in which the entrant stays out in the first period. This is not possible under BI, but is possible if we look at admissible Nash equilibria.

The conjecture, then, is that SAS yields an admissible Nash equilibrium in PI games. This is almost correct. The next section gives an exact statement and proof.

7 Perfect-Information Games

We now come to a characterization of SAS in PI games. We impose a no-ties condition on payoffs. For this, we need some definitions: An **outcome** is a payoff vector $\Pi(z) = (\Pi^a(z), \Pi^b(z))$. Two terminal nodes z and z' are **outcome equivalent** if $\Pi(z) = \Pi(z')$. We also say that two strategy profiles (s^a, s^b) and (r^a, r^b) are outcome equivalent if $\zeta(s^a, s^b)$ and $\zeta(r^a, r^b)$ are outcome equivalent.

Definition 7.1 *A tree Γ satisfies the **Single Payoff Condition (SPC)** if for all z, z' in Z, if Ann (resp. Bob) moves at the last common predecessor of z and z', then $\Pi^a(z) = \Pi^a(z')$ implies $\Pi^b(z) = \Pi^b(z')$ (resp. $\Pi^b(z) = \Pi^b(z')$ implies $\Pi^a(z) = \Pi^a(z')$).*

In words, Ann is indifferent between two terminal nodes over which she is decisive, only if those two terminal nodes are outcome equivalent. SPC appears to be a minimal no-ties condition if BI is to select a unique outcome in the tree. A generic tree satisfies SPC. Non-generic trees can also satisfy SPC and many games of interest are non-generic. (Zero-sum games are one example. The twice-repeated Chain Store game, as given in Figure 4.4, satisfies SPC. For additional examples, see the discussions of non-genericity in Mertens [28, p.582] and Marx-Swinkels [27, pp.224-225].) A tree that satisfies "no relevant ties" (Battigalli [6, p.48]) satisfies SPC, but the converse need not hold. In a PI tree, SPC is equivalent to the "transfer of decisionmaker indifference" condition (Marx-Swinkels [27]).[10]

Now the characterization result:

Proposition 7.1 *Fix a PI tree Γ satisfying SPC, and let G be the strategic form of Γ.*

(a) Fix an SAS $Q^a \times Q^b$. Then there is a pure Nash equilibrium of G such that each profile $(s^a, s^b) \in Q^a \times Q^b$ is outcome equivalent to this equilibrium.

[10] This latter is a condition on the strategic form. In the proof to come, we need to make use of properties of the extensive form–hence our use of SPC.

(b) *Fix an admissible pure Nash equilibrium* (s^a, s^b) *of* G. *Then there is an SAS of* G *that contains* (s^a, s^b).

Before the formal argument, we give a sketch of the proof. For part (a), the argument is forward looking, by induction on the length of the tree. Suppose the result is true for a tree of length l. Here is why it is then true for a tree of length $(l+1)$. Refer to Figure 7.1.

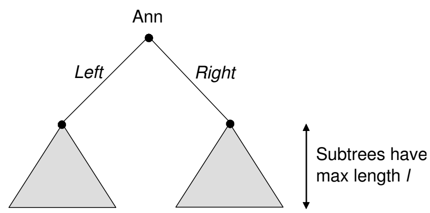

Figure 7.1

Fix an SAS $Q^a \times Q^b$ and suppose there is a strategy in Q^a where Ann plays $Left$. Then, the projection of $Q^a \times Q^b$ into the $Left$ subtree–which we will denote $Q^a_L \times Q^b_L$–forms an SAS of this subtree. By the induction hypothesis, we can find a Nash equilibrium (s^a_L, s^b_L) of the $Left$ subtree, so that each profile in $Q^a_L \times Q^b_L$ is outcome equivalent to this Nash equilibrium. We want to show that (s^a_L, s^b_L) can be made into a Nash equilibrium of the whole tree. To do this, we give Ann the strategy that selects $Left$ and then follows the choices prescribed by s^a_L. Denote this strategy by s^a. On the $Left$ subtree, Bob will follow the choices prescribed by s^b_L. When Ann plays s^a, Bob then has no incentive to deviate. When Bob plays s^b_L, Ann has no incentive to deviate to strategies that lead to the $Left$ subtree. It remains to specify choices for Bob on the $Right$ subtree, so that Ann also has no incentive to deviate to the $Right$ subtree. This step is achieved via the Minimax Theorem for PI games.

This sketch omits some important details. One is the role of SPC, which is, in fact, necessary for both parts (a) and (b) of Proposition 7.1.

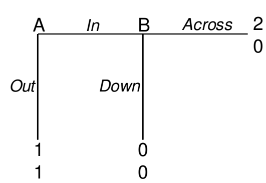

Figure 7.2

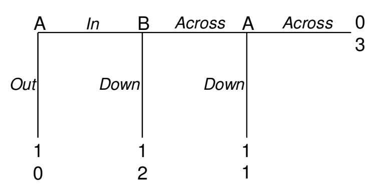

Figure 7.3

Figures 7.2 and 7.3 are two trees that fail SPC. In Figure 7.2, we see that part (a) is false without SPC.[11] Here, the set $\{Out, In\} \times \{Down, Across\}$ is an SAS, but $(In, Down)$ is not outcome equivalent to a Nash equilibrium. In Figure 7.3, we see that part (b) is false without SPC. Here, $(Out, Across)$ is an admissible Nash equilibrium (is even outcome equivalent to a BI

[11] Drew Fudenberg kindly provided this example.

strategy profile), but is not contained in any SAS. In particular, the unique SAS (and so the IA set) is $\{Out, In\text{-}Down\} \times \{Down\}$. (To see this, fix an SAS $Q^a \times Q^b$. For Ann, the strategy *In-Across* is inadmissible, so cannot be part of an SAS. Using this and condition (iii) of an SAS, $Q^a = \{Out, In\text{-}Down\}$. For Bob, only *Down* is admissible with respect to $\{Out, In\text{-}Down\}$.)

Let us also comment on the gap between parts (a) and (b) of Proposition 7.1. Part (a) involves a strategy profile which is admissible (condition (i) of an SAS) and outcome equivalent to a Nash equilibrium. Part (b) starts with an admissible Nash profile and says that there is an SAS containing this profile. Can the gap between the two directions be closed?

Part (b) cannot be improved to read: *Fix an admissible pure profile* (s^a, s^b) *of* G *that is outcome equivalent to a Nash equilibrium of* G*. Then* (s^a, s^b) *is contained in some SAS.* The tree in Figure 7.4 satisfies SPC, and $(Out, Across)$ is an admissible profile that is outcome equivalent to the Nash equilibrium $(Out, Down)$. But the unique SAS is $\{(In\text{-}Across, Across)\}$. Suppose we try instead to improve part (a) to read: *Fix an SAS* $Q^a \times Q^b$*. Then there is an admissible pure Nash equilibrium of* G *such that each profile* $(s^a, s^b) \in Q^a \times Q^b$ *is outcome equivalent to this equilibrium.* We don't know if this stronger statement is true.

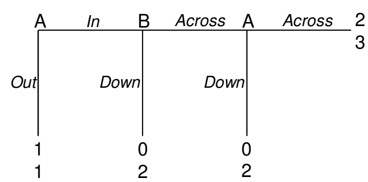

Figure 7.4

Proof of Proposition 7.1(a). By induction on the length of the tree.

Consider a tree of length 1, and assume that Ann moves at the initial node. (We then define π^a and π^b on S^a alone.) If $s^a \in Q^a$, then $\pi^a(s^a) \geq \pi^a(r^a)$ for all $r^a \in S^a$. Thus (s^a) is an admissible Nash equilibrium. If $r^a \in Q^a$, then we must have $\pi^a(r^a) = \pi^a(s^a)$. Since Ann moves at the last common predecessor of $\zeta(r^a)$ and $\zeta(s^a)$, SPC implies $\pi^b(r^a) = \pi^b(s^a)$, establishing that (r^a) is outcome equivalent to (s^a).

Now suppose the result is true for any tree of length l or less, and consider a tree of length $l+1$. We can again assume that Ann moves at the initial node. Denote the subtrees that follow Ann's initial move by $\Delta_1, \ldots, \Delta_K$. Also, let S^a_k denote the strategies for Ann that allow subtree Δ_k. (Up to duplication, this subset can be identified with Ann's strategies for Δ_K.) Let S^b_k be the corresponding set for Bob. We have $S^b = \times_{k=1}^{K} S^b_k$.[12] (Thus, a strategy for Bob on the whole tree is some $s^b = (s^b_1, \ldots, s^b_K)$, where each s^b_k specifies Bob's choices at the information sets in subtree Δ_k.)

By the projection property (Proposition 6.2), for each k such that $Q^a \cap S^a_k \neq \emptyset$, we have that $(Q^a \cap S^a_k) \times \mathrm{proj}_{S^b_k} Q^b$ is an SAS of (the strategic form of) Δ_k. By the induction hypothesis, for

[12] Do notice, our approach here differs slightly from that in the proof of Proposition 6.2.

each such k, each profile in $(Q^a \cap S_k^a) \times \mathrm{proj}_{S_k^b} Q^b$ is outcome equivalent.

Suppose there are strategies $s^a, r^a \in Q^a$ that reach two distinct subtrees, say Δ_1 and Δ_2 respectively. By condition (ii) of an SAS, there is a $\sigma^b \in \mathcal{M}\left(S^b\right)$, with $\mathrm{Supp}\,\sigma^b = Q^b$, such that s^a is optimal under σ^b. Since s^a reaches Δ_1 and r^a reaches Δ_2, this implies

$$\sum_{s_1^b \in \mathrm{proj}_{S_1^b} Q^b} \mathrm{marg}_{S_1^b}\, \sigma^b(s_1^b)\pi^a(s^a, s_1^b) \geq \sum_{s_2^b \in \mathrm{proj}_{S_2^b} Q^b} \mathrm{marg}_{S_2^b}\, \sigma^b(s_2^b)\pi^a(r^a, s_2^b).$$

(We use π^a for the induced payoff functions on $S_1^a \times S_1^b$ and $S_2^a \times S_2^b$; no confusion should result.)

The induction hypothesis gives that $\pi^a(s^a, s_1^b)$ is constant for all $s_1^b \in \mathrm{proj}_{S_1^b} Q^b$, and, likewise, that $\pi^a(r^a, s_2^b)$ is constant for all $s_2^b \in \mathrm{proj}_{S_2^b} Q^b$. Thus we have $\pi^a(s^a, s_1^b) \geq \pi^a(r^a, s_2^b)$ for all such s_1^b and s_2^b. But, symmetrically, we can apply condition (ii) of an SAS to r^a, to get the opposite inequality. Thus $\pi^a(s^a, s_1^b) = \pi^a(r^a, s_2^b)$ for all $s_1^b \in \mathrm{proj}_{S_1^b} Q^b$ and $s_2^b \in \mathrm{proj}_{S_2^b} Q^b$. Also, since Ann moves at the last common predecessor of $\zeta(s^a, s_1^b)$ and $\zeta(r^a, s_2^b)$, SPC implies that $\pi^b(s_1^b, s^a) = \pi^b(s_2^b, r^a)$.

We have now shown that every profile in $Q^a \times Q^b$ is outcome equivalent.

The final step is to construct a Nash equilibrium to which (s^a, s_1^b) is outcome equivalent. By the induction hypothesis, there is a Nash equilibrium of Δ_1 to which (s^a, s_1^b) is outcome equivalent. Denote it as (q_1^a, q_1^b). We show that for each $k = 2, \ldots, K$, there is a strategy $q_k^b \in S_k^b$ such that $\pi^a(q_1^a, q_1^b) \geq \pi^a(r^a, q_k^b)$ for all $r^a \in S_k^a$. The profile $(q_1^a, (q_1^b, \ldots, q_K^b))$ will then be the desired equilibrium.

Since s^a is optimal under σ^b, as above, we have for each $k = 2, \ldots, K$,

$$\pi^a(q_1^a, q_1^b) = \pi^a(s^a, s_1^b) \geq \sum_{s_k^b \in \mathrm{proj}_{S_k^b} Q^b} \mathrm{marg}_{S_k^b}\, \sigma^b(s_k^b)\pi^a(r^a, s_k^b)$$

for all r^a that allow Δ_k. Letting $(\overline{r}_k^a, \overline{r}_k^b) \in \arg\max_{S_k^a} \min_{S_k^b} \pi^a(\cdot, \cdot)$, we have in particular

$$\pi^a(q_1^a, q_1^b) \geq \textstyle\sum_{s_k^b \in \mathrm{proj}_{S_k^b} Q^b} \mathrm{marg}_{S_k^b}\, \sigma^b(s_k^b)\pi^a(\overline{r}_k^a, s_k^b).$$

But $\pi^a(\overline{r}_k^a, r_k^b) \geq \pi^a(\overline{r}_k^a, \overline{r}_k^b)$ for any $r_k^b \in S_k^b$, by definition. So

$$\pi^a(q_1^a, q_1^b) \geq \sum_{s_k^b \in \mathrm{proj}_{S_k^b} Q^b} \mathrm{marg}_{S_k^b}\, \sigma^b(s_k^b)\pi^a(\overline{r}_k^a, \overline{r}_k^b) = \pi^a(\overline{r}_k^a, \overline{r}_k^b).$$

Set $(\underline{r}_k^a, \underline{r}_k^b) \in \arg\min_{S_k^b} \max_{S_k^a} \pi^a(\cdot, \cdot)$. By the Minimax Theorem for PI games (see, e.g., Ben Porath [9]), $\pi^a(\underline{r}_k^a, \underline{r}_k^b) = \pi^a(\overline{r}_k^a, \overline{r}_k^b)$. It follows that $\pi^a(q_1^a, q_1^b) \geq \pi^a(\overline{r}_k^a, \overline{r}_k^b) = \pi^a(\underline{r}_k^a, \underline{r}_k^b)$. But $\pi^a(\underline{r}_k^a, \underline{r}_k^b) \geq \pi^a(r_k^a, \underline{r}_k^b)$ for any $r_k^a \in S_k^a$, by definition. So $\pi^a(q_1^a, q_1^b) \geq \pi^a(r_k^a, \underline{r}_k^b)$. Setting $q_k^b = \underline{r}_k^b$ gives the desired profile. ∎

For the proof of Proposition 7.1(b), we will need a preliminary lemma, which is proved in the Appendix.

Lemma 7.1 *Fix a PI tree. If s^a is admissible, then $\pi^a(s^a, s^b) = \pi^a(r^a, s^b)$ for each $r^a \in \mathrm{su}(s^a)$ and $s^b \in S^b$.*

Proof of Proposition 7.1(b). Fix a PI tree Γ satisfying SPC. Let G be the associated strategic form. We show that if (s^a, s^b) is an admissible Nash equilibrium of G, then $\mathrm{su}(s^a) \times \mathrm{su}(s^b)$ is an SAS of G.

Each $r^a \in \mathrm{su}(s^a)$ is admissible, since s^a is admissible (Corollary A1). So condition (i) of an SAS is satisfied.

Next, condition (ii). It suffices to show that, for each $q^a \in S^a$ and $r^b \in \mathrm{su}(s^b)$, $\pi^a(q^a, s^b) = \pi^a(q^a, r^b)$. If so, s^a must be admissible with respect to $\mathrm{su}(s^b)$. To see this last statement: Fix $r^b \in \mathrm{su}(s^b)$ and notice that if the claim holds then

$$\pi^a\left(s^a, r^b\right) = \pi^a\left(s^a, s^b\right) \geq \pi^a\left(q^a, s^b\right) = \pi^a\left(q^a, r^b\right),$$

where the two equalities follow from the claim and the inequality follows from the fact that (s^a, s^b) is a Nash equilibrium. With this, s^a is optimal under any measure $\sigma^b \in \mathcal{M}\left(S^b\right)$ with $\mathrm{Supp}\,\sigma^b = \mathrm{su}(s^b)$, and so s^a is admissible with respect to $\mathrm{su}(s^b)$.

Fix $q^a \in S^a$ and $r^b \in \mathrm{su}(s^b)$. By Lemma 7.1 applied to s^b (which is admissible), we have $\pi^b\left(s^b, q^a\right) = \pi^b\left(r^b, q^a\right)$. If $\left(q^a, s^b\right)$ and $\left(q^a, r^b\right)$ reach the same terminal node, then certainly $\pi^a\left(q^a, s^b\right) = \pi^a\left(q^a, r^b\right)$. If not, Bob moves at the last common predecessor of $\zeta\left(q^a, s^b\right)$ and $\zeta\left(q^a, r^b\right)$, so that SPC establishes the desired result.

Finally, condition (iii) follows immediately from Lemma A3. ■

To sum up: In a PI tree satisfying SPC, each SAS is outcome equivalent to some Nash equilibrium, and each admissible Nash equilibrium is contained in some SAS. In particular, then, an SAS need not yield the (unique) BI outcome in such a game. However, a converse does hold.

Proposition 7.2 *Fix a PI tree satisfying SPC. There is an SAS that is outcome equivalent to the BI outcome.*

Proof. Proposition A2 in the Appendix shows that in a PI tree satisfying SPC, there is an admissible BI strategy profile. Of course, this profile is a Nash equilibrium, so, by Proposition 7.1(b), there is an SAS that contains it. ■

Note that this is only a partial converse. A BI strategy profile need not be admissible, and, therefore, need not be contained in any SAS. (In particular, then, a BI profile need not be contained in the IA set.)

8 Discussion

We conclude with some comments on conceptual matters and related literature.

a. Player-Specific Type Structures In Figure 1.1, we presented the basic 'architecture' of an epistemic analysis. The starting point is to add types to the description of the game. With this, we can analyze the conditions of RCBR or RCAR. The projection of the RCBR (resp. RCAR) set into the strategy sets forms a BRS (resp. SAS) of the game. As for which BRS or SAS obtains, this depends on which particular types we add.

How should we think about the choice of one vs. another type structure? In any particular structure, certain beliefs, beliefs about beliefs, ..., will be present (i.e., will be induced by a type) and others won't be. So, there is an important implicit assumption behind the choice of a structure. This is that it is "transparent" to the players that the beliefs in the type structure–and only those beliefs–are possible. (See Battigalli-Friedenberg [7, Appendix A] for a formal treatment of this point, for the RCBR case.) Why would there be such a "transparent" restriction on beliefs? The idea is that there is a "context" (BFK [18, Section 2.8]) to the strategic situation (e.g., history, conventions, etc.) and this "context" causes the players to rule out certain beliefs.

Notice what is involved: Ann and Bob think the same way about which beliefs are–and are not–possible. This is a substantive (if again implicit) assumption. While the assumption is standard in epistemic game theory, it is clearly important to investigate the implications of dropping it.[13]

Presumably, the new ingredient would be the concept of a "player-specific type structure," where we specify a (potentially different) type structure for each player. Now, a type t^a for Ann, in Ann's type structure, could consider possible a certain type t^b for Bob, even though, in Bob's type structure, there is no type for Bob with the same hierarchy of beliefs as t^b. Of course, a special case will be when the player-specific type structures coincide. Call such a (common) type structure a "player-independent type structure."

The BRS concept characterizes RCBR across all player-independent type structures. If we now characterize RCBR across all player-specific type structures, we will get all the BRS's and also some new sets. It can be shown these new sets need not be BRS's. We conjecture that they will be contained in the IU set. If so, the extra precision we get in identifying these sets above and beyond BRS's may seem a small matter—much as the extra precision in identifying BRS's above and beyond the IU set may seem a small matter.

Contrast with the case of admissibility. In the Introduction, we pointed to a basic non-monotonicity in admissibility, and explained that, precisely because of this non-monotonicity, SAS's needed to be understood separately from the IA set. For the same reason, it seems that a full characterization of RCAR across all type structures, including the player-specific type structures, will be needed. This is left for future work.

b. Non-Monotonicity Continued We began with the observation that there is a basic non-

[13] We are grateful to a referee for very fruitful exchanges on this issue.

monotonicity in admissibility: Adding new possibilities can change previously good strategies into bad ones. Refer back to Figure 1.2. There, we said that when we add the strategy C to the set $\{U, M\} \times \{L\}$, we introduce a new possibility and, so, the strategy M may now turn into a bad strategy. But this argument was incomplete. After all, the strategy C is already in the matrix. So, presumably, Ann should have already considered this possibility.

The answer brings us back to Samuelson's [31] basic tension between admissibility and strategic reasoning (mentioned in Section 3). On the one hand, admissibility requires that Ann includes all of Bob's strategies. On the other hand, strategic reasoning requires that Ann excludes Bob's irrational strategies. Back to Figure 1.2, and consider a solution concept with $Q^a \times Q^b = \{U, M\} \times \{L\}$. Yes, Ann should include C, since she should include all possibilities. But, she should also exclude C, since it is inconsistent with the solution. So, if C is now added to Q^b, it does 'make a difference.' Ann should no longer exclude C, which turns the previously good strategy M into a bad one. We conclude that non-monotonicity of the solution concept is part and parcel of any epistemic analysis that addresses the inclusion-exclusion problem.

How does this verbal argument play out at a more formal level? Let us specialize to the case of an LPS $(\mu_1, \ldots, \mu_n)$. Assumption lies strictly between the concepts of "belief at the 1st level" and "belief." (See BFK [18, Proposition 5.1].) In particular, Ann believes at the 1st level (resp. believes) the event $E =$ "Bob is rational" if $\mu_1(E) = 1$ (resp. $\mu_i(E) = 1$ for all i). Back again to Figure 1.2. We can construct an LPS $(\mu_1, \ldots, \mu_n)$ which has full-support and is such that Ann believes at the 1st level the event $\{L, C\}$.[14] But, we also want M to be optimal under this LPS—and, for this, it must be that the irrational R is considered at least as likely as C. That is, if $\mu_i(C) > 0$ then there is some $j \leq i$ with $\mu_j(R) > 0$. Of course, by inclusion, there is some i with $\mu_i(C) > 0$. Thus, by asking for inclusion, we forgo exclusion. Under belief, we get exclusion, but lose inclusion. Ann can believe the event "Bob is rational" only if each measure μ_i assigns probability 0 to one of Bob's strategies–viz., the irrational R.

Assumption gives both inclusion and exclusion. The 'cost' is non-monotonicity: An LPS may assume $\{L\}$ but not $\{L, C\}$. In particular, then, the strategy M can be optimal for Ann if she assumes L. Yet M cannot be optimal if she assumes $\{L, C\}$. Thus, the non-monotonicity of assumption yields a non-monotonicity of the solution concept: $Q^a \times Q^b$ is an SAS but $Q^a \times (Q^b \cup \{s^b\})$ is not.

As far as we can see, any resolution of the inclusion-exclusion problem must have this feature.

c. Relationship to Other Solution Concepts The literature has proposed a number of weak-dominance analogs to the Pearce [29] BRS concept. To the best of our knowledge, no paper has provided foundations which fully resolve the inclusion-exclusion problem. Each gives up on one or other criterion.

In terms of inclusion: Samuelson [31] provided foundations for the consistent pairs concept. (See also Börgers-Samuelson [13].) A consistent pair may contain an inadmissible strategy, and

[14] This is a semi-formal discussion only. To be precise, we would need to include the types, too.

so inclusion is not satisfied.[15] Asheim [1], Asheim-Dufwenberg [2], and Asheim-Perea [3] take an interesting different approach: They require that Ann consider every strategy for Bob possible, but not every type for Bob. So, epistemically speaking, they have partial rather than full inclusion.

In terms of exclusion: Dekel-Fudenberg [21] introduced the $S^{\infty}W$ concept of one round of deletion of inadmissible strategies followed by iterated deletion of strongly dominated strategies. There is an obvious BRS-like version of the definition (BFK [18, Section 11B]). The epistemic foundations (Brandenburger [14], Börgers [12], and [18, Section 11B]) rest on the idea of belief at the 1st level. As mentioned above, this concept fails exclusion.

SAS also differs from each of these solution concepts in terms of the strategies that can be played. (See BFK [19, Section S.3] for details.)

d. Related PI Results Proposition 7.1 resembles an earlier result on PI trees, due to Ben Porath [9, Theorem 2]. Ben Porath defines an extensive-form analog to belief at the 1st level, and gives epistemic conditions which yield a solution concept we will call $S^{\infty}CD$. This concept first eliminates all conditionally dominated strategies (Shimoji-Watson [33])–i.e., all strategies which are not extensive-form rational. It then iteratively eliminates strategies which are strongly dominated in the matrix (equivalently, strongly dominated at the root of the tree). It is an extensive-form analog to the $S^{\infty}W$ concept, and, indeed, in generic PI trees the two concepts coincide.

The $S^{\infty}CD$ concept need not yield a Nash outcome in PI trees satisfying SPC. For example, in Centipede (Figure 4.1), the only elimination via $S^{\infty}CD$ is the strategy In at every node for Ann.

Ben Porath then adds an assumption which does result in Nash outcomes in (generic) PI games. This is a 'grain of truth' condition on his epistemic analysis, which says that each player assigns positive probability to the actual state–and, therefore, to the actual strategies. See [9, p.38]. Presumably, the admissibility requirement (recall Remark 3.1) plays a similar role in our analysis.

Battigalli-Friedenberg [7] study the epistemic condition of "rationality and common strong belief of rationality" (RCSBR), due to Battigalli-Siniscalchi [8]. RCSBR does not impose an admissibility requirement. For instance, in simultaneous games, RCSBR is characterized by BRS.[16] Nevertheless, there is a point of connection with this paper. Battigalli-Friedenberg show that in PI games satisfying no-relevant ties (refer back to the discussion after Definition 7.1 for a definition), RCSBR is characterized by Nash outcomes in extensive-form rational strategies.

e. n-Player Games We have treated two-player games, but the analysis extends to n-player games.

> A set $\times_{i=1}^{n} Q^i \subseteq \times_{i=1}^{n} S^i$ is an (n-**player**) **SAS** if, for each player i: (i) each $s^i \in Q^i$ is admissible with respect to $\times_{j=1}^{n} S^j$; (ii) each $s^i \in Q^i$ is admissible with respect to $S^i \times \times_{j \neq i} Q^j$; (iii) for any $s^i \in Q^i$, if $r^i \in \text{su}(s^i)$ then $r^i \in Q^i$.

[15] The modified consistent pairs concept (Ewerhart [23]) has a flavor of solving the inclusion-exclusion problem. We do not know of epistemic foundations for this concept.

[16] Here, the definition of a BRS incorporates a maximality condition; see [7]. See also Footnote 4.

Of course, conditions (i) and (ii) are equivalent to: (i′) each $s^i \in Q^i$ is optimal under some $\mu^{-i} \in \mathcal{M}(\times_{j\neq i}S^j)$, with $\operatorname{Supp}\mu^{-i} = \times_{j\neq i}S^j$, given S^i; (ii′) each $s^i \in Q^i$ is optimal under some $\nu^{-i} \in \mathcal{M}(\times_{j\neq i}S^j)$, with $\operatorname{Supp}\nu^{-i} = \times_{j\neq i}Q^j$, given S^i. Under this definition, all the results in this paper hold for the n-player case, including Proposition 7.1. (We now use the n-player Minimax Theorem for PI games: $\min_{\times_{j\neq i}S^j} \max_{S^i} \pi^i(\cdot,\cdot) = \max_{S^i} \min_{\times_{j\neq i}S^j} \pi^i(\cdot,\cdot)$.)

f. Invariance Continued SAS satisfies invariance (Proposition 5.2). In Section 2, we also mentioned the idea of a deeper notion of invariance–that epistemic reasoning should not change across equivalent games. Does SAS satisfy such "epistemic invariance," too?

To answer, we need a formalization of this notion. Here is one possibility. Fix a tree Γ and an associated type structure, where a type t^a for Ann is associated with a conditional probability system (CPS) on $S^b \times T^b$, and the family of conditioning events (at least) includes all events in $S^b \times T^b$ that correspond to an information set in some tree with the same fully reduced strategic form as Γ. Now bring in SAS's and their epistemic underpinnings–viz., lexicographic probability systems (LPS's) à la BFK [18]. A full-support LPS on $S^b \times T^b$ naturally induces a CPS on $S^b \times T^b$, where the family of conditioning events consists of all non-empty open sets in $S^b \times T^b$ (BFK [17], [19]). This family includes the events that correspond to the information sets. So, SAS can be said to pass one test of epistemic invariance. Still, this is only a sketch, and a full-blown formulation of epistemic invariance is warranted.

g. Relationship to the Stability Literature SAS is a solution concept derived from the epistemic program. In this paper, the focus was on SAS as a solution concept in its own right. We have covered several general properties of a solution concept–existence, invariance, extensive-form rationality, and projection. Of course, there are other properties one can consider. Starting with Kohlberg-Mertens [24], the stability literature has developed a long list of potentially desirable properties of solution concepts. One is the difference property (Kohlberg-Mertens [24, Section 2.6]). Roughly speaking, this requires that for any tree and subtree, an outcome allowed by the solution concept on the original tree is also allowed by the solution concept on the difference tree–i.e., on the tree obtained by pruning the subtree according to the solution concept. Elsewhere ([16, Theorem 3.1]), we show that if a solution concept satisfies existence, extensive-form rationality, and difference on the domain of PI trees satisfying SPC, then this concept must yield the BI outcome in these trees. It follows that SAS fails the difference property. Still, we do not view this as a flaw in the SAS concept, since we do not insist on BI. Mertens himself has expressed a similar view in the context of equilibrium analysis: "I had (and still have) some instinctive liking for the bruto Nash equilibrium, or a modest refinement like admissible equilibria" ([28, pp.582-583]).

It would certainly be interesting to conduct a comprehensive examination of which properties various epistemic solution concepts do or do not satisfy. Speaking as epistemic game theorists, we believe that this would help determine which properties are desirable and which ones are not. Of course, this exercise would also serve as a kind of 'audit' of different epistemic concepts. Our investigation of SAS in this paper is only a first step in this direction.

Appendix

We begin with the proof of Proposition 5.1.

Lemma A1 *Fix a strategy* s^a *and some* $\varphi^a \in \mathcal{M}(S^a)$ *such that* $\pi^a(\varphi^a, s^b) = \pi^a(s^a, s^b)$ *for all* $s^b \in S^b$. *Then* s^a *is optimal under* $\mu^b \in \mathcal{M}(S^b)$ *if and only if all* $r^a \in \operatorname{Supp}\varphi^a$ *are optimal under* μ^b.

Proof. Routine. ■

Corollary A1 *Fix a strategy* s^a *that is admissible given* $Q^a \times Q^b$. *Then, for each* $r^a \in \operatorname{su}(s^a)$ *with* $r^a \in Q^a$, r^a *is admissible given* $Q^a \times Q^b$.

The next lemma is Lemma F1 in BFK [18], but we give a statement and proof here for ease of reference.

Lemma A2 *If* $s^a \in S^a_m$ *then there is a* $\mu^b \in \mathcal{M}(S^b)$, *with* $\operatorname{Supp}\mu^b = S^b_{m-1}$, *such that* $\pi^a(s^a, \mu^b) \geq \pi^a(r^a, \mu^b)$, *for each* $r^a \in S^a$.

Proof. By Remark 3.1, there is a $\mu^b \in \mathcal{M}(S^b)$, with $\operatorname{Supp}\mu = S^b_{m-1}$, such that $\pi^a(s^a, \mu^b) \geq \pi^a(r^a, \mu^b)$ for all $r^a \in S^a_{m-1}$. Suppose there is an $r^a \in S^a \backslash S^a_{m-1}$ with

$$\pi^a(r^a, \mu^b) > \pi^a(s^a, \mu^b). \tag{A.1}$$

Then $r^a \in S^a_l \backslash S^a_{l+1}$, for some $l < m-1$. Choose r^a (and l) so that there does not exist $q^a \in S^a_{l+1}$ with $\pi^a(q^a, \mu^b) > \pi^a(s^a, \mu^b)$.

Fix some $\nu^b \in \mathcal{M}(S^b)$, with $\operatorname{Supp}\nu^b = S^b_l$, and define a sequence of measures $\mu^b_n \in \mathcal{M}(S^b)$, for each $n \in \mathbb{N}$, by $\mu^b_n = (1 - \frac{1}{n})\mu^b + \frac{1}{n}\nu^b$. Note that $\operatorname{Supp}\mu^b_n = S^b_l$ for each n. Using $r^a \notin S^a_{l+1}$, and Remark 3.1 applied to the $(l+1)$-admissible strategies, it follows that for each n there is a $q^a \in S^a_l$ with

$$\pi^a(q^a, \mu^b_n) > \pi^a(r^a, \mu^b_n). \tag{A.2}$$

We can assume that $q^a \in S^a_{l+1}$. (Choose $q^a \in S^a_l$ to maximize the left-hand side of equation (A.2) among all strategies in S^a_l.) Also, since S^a_{l+1} is finite, there is a $q^a \in S^a_{l+1}$ such that equation (A.2) holds for infinitely many n. Letting $n \to \infty$ yields

$$\pi^a(q^a, \mu^b) \geq \pi^a(r^a, \mu^b). \tag{A.3}$$

From equations (A.1) and (A.3) we get $\pi^a(q^a, \mu^b) > \pi^a(s^a, \mu^b)$, contradicting our choice of r^a. ■

Proof of Proposition 5.1. We show that the IA set is an SAS. Fix $s^a \in S^a_M$. Certainly, $s^a \in S^a_1$, and so s^a is admissible with respect to S^b, establishing Condition (i). Since $S^i_M = S^i_{M+1}$, we know that s^a is admissible with respect to $S^a_M \times S^b_M$. Using Lemma A2 and Remark 3.1, it follows that

s^a is admissible with respect to $S^a \times S^b_M$, establishing Condition (ii). For Condition (iii), we show, by induction on m, that if $r^a \in \mathrm{su}(s^a)$, then $r^a \in S^a_m$. The result is immediate for $m = 0$, so assume $r^a \in S^a_m$. Then, using the fact that $s^a \in S^a_{m+1}$ and Corollary A1, $r^a \in S^a_{m+1}$. ■

To prove Proposition 5.2, we need two lemmas.

Lemma A3 *If $q^a \in \mathrm{su}(r^a)$ and $r^a \in \mathrm{su}(s^a)$, then $q^a \in \mathrm{su}(s^a)$.*

Proof. Immediate. ■

Lemma A4 *Fix $s^a \in S^a$ and $\varphi^a \in \mathcal{M}(S^a)$ with $\pi^b(s^b, \varphi^a) = \pi^b(s^b, s^a)$ for all $s^b \in S^b$. Fix also $X \subseteq S^a$ with $\mathrm{Supp}\,\varphi^a \subseteq X$ and some $Y \subseteq S^b$. Then s^b is admissible with respect to $X \times Y$ if and only if it is admissible with respect to $(X \cup \{s^a\}) \times Y$.*

Proof. We can obviously assume that $s^a \notin X$. Now, if s^b is admissible with respect to $X \times Y$, there is a $\mu^a \in \mathcal{M}(S^a)$, with $\mathrm{Supp}\,\mu^a = X$, such that s^b is optimal under μ^a given Y. Define $\nu^a \in \mathcal{M}(S^a)$ by

$$\nu^a(r^a) = \begin{cases} \varepsilon & \text{if } r^a = s^a, \\ \mu^a(r^a) - \varepsilon\varphi^a(r^a) & \text{otherwise,} \end{cases}$$

where $\varepsilon > 0$ is chosen small enough that every $\mu^a(r^a) - \varepsilon\varphi^a(r^a) > 0$. (This is possible since $\varphi^a(r^a) > 0$ implies $\mu^a(r^a) > 0$.) Then $\mathrm{Supp}\,\nu^a = X \cup \{s^a\}$, and $\pi^b(r^b, \nu^a) = \pi^b(r^b, \mu^a)$ for all $r^b \in S^b$. Thus s^b is admissible with respect to $(X \cup \{s^a\}) \times Y$.

Conversely, if s^b is admissible with respect to $(X \cup \{s^a\}) \times Y$, there is a $\mu^a \in \mathcal{M}(S^a)$, with $\mathrm{Supp}\,\mu^a = X \cup \{s^a\}$, such that s^b is optimal under μ^a given Y. Define $\nu^a \in \mathcal{M}(S^a)$ by $\nu^a(r^a) = \mu^a(r^a) + \mu^a(s^a)\varphi^a(r^a)$ for $r^a \in X$. Then $\mathrm{Supp}\,\nu^a = X$, and $\pi^b(r^b, \nu^a) = \pi^b(r^b, \mu^a)$ for all $r^b \in S^b$. Thus s^b is admissible with respect to $X \times Y$. ■

Proof of Proposition 5.2. Begin with part (a). It is immediate that each $s^a \in \overline{Q}^a \backslash \{q^a\}$ satisfies conditions (i)-(iii) of an SAS. So we will turn to Bob.

Since each $s^b \in \overline{Q}^b$ is admissible with respect to $(S^a \cup \{q^a\}) \times S^b$ (condition (i) of an SAS applied to $\overline{G}$), Lemma A4 implies that each $s^b \in \overline{Q}^b$ is admissible with respect to $S^a \times S^b$. Next, note that $s^b \in \overline{Q}^b$ is admissible with respect to $\overline{Q}^a \times S^b$ (condition (ii) of an SAS applied to $\overline{G}$). It suffices to consider the case when $q^a \in \overline{Q}^a$. Then $\mathrm{Supp}\,\varphi^a \subseteq \overline{Q}^a$ (condition (iii) of an SAS applied to $\overline{G}$). It follows from Lemma A4 that s^b is admissible with respect to $(\overline{Q}^a \backslash \{q^a\}) \times S^b$, establishing condition (ii) of an SAS for Bob.

For condition (iii) of an SAS, suppose r^b supports $s^b \in \overline{Q}^b$, via $\rho^b \in \mathcal{M}(S^b)$, in the game G. We have to show that $r^b \in \overline{Q}^b$. This will follow from condition (iii) applied to $\overline{G}$, provided

$\pi^b(\rho^b, q^a) = \pi^b(s^b, q^a)$. Notice that

$$\pi^b(\rho^b, q^a) = \sum_{u^b \in S^b} \pi^b(u^b, q^a)\,\rho^b(u^b) = \sum_{u^b \in S^b} \sum_{s^a \in S^a} \pi^b(u^b, s^a)\,\varphi^a(s^a)\,\rho^b(u^b) =$$

$$= \sum_{s^a \in S^a} \sum_{u^b \in S^b} \pi^b(u^b, s^a)\,\rho^b(u^b)\,\varphi^a(s^a) = \sum_{s^a \in S^a} \pi^b(s^b, s^a)\,\varphi^a(s^a) = \pi^b(s^b, q^a)$$

as required.

For part (b) of the proposition, first suppose that q^a does not support any strategy in Q^a. Any $s^a \in Q^a$ is admissible with respect to $S^a \times S^b$ (resp. $S^a \times Q^b$) among strategies in S^a. It follows from Lemma A1 that each $s^a \in Q^a$ is also admissible with respect to $(S^a \cup \{q^a\}) \times S^b$ (resp. $(S^a \cup \{q^a\}) \times Q^b$). This establishes conditions (i)-(ii) of Definition 3.4 for $s^a \in Q^a$. Condition (iii) is immediate for this case. Next, because each $s^b \in Q^b$ is admissible with respect to $S^a \times S^b$, it is also admissible with respect to $(S^a \cup \{q^a\}) \times S^b$ by Lemma A4. Condition (ii) is immediate. Finally, notice that if r^b supports s^b in $\overline{G}$, then certainly it does in G, so condition (iii) follows.

Next, suppose q^a supports some $s^a \in Q^a$, and write $\overline{Q}^a = Q^a \cup \{q^a\}$. It is immediate that each $r^a \in Q^a$ satisfies conditions (i)-(ii) of Definition 3.4. Lemma A1 implies that q^a also satisfies (i)-(ii), since s^a does. Condition (iii) is clearly satisfied for any $r^a \in Q^a$, since $q^a \in \overline{Q}^a$. Condition (iii) is also satisfied for q^a. To see this, use Lemma A3 to get that if u^a supports q^a, then u^a also supports s^a (since q^a supports s^a). Applying condition (iii) to G then implies $u^a \in Q^a$.

Next, consider some $s^b \in Q^b$. Conditions (i) and (iii) are as above (i.e., as in the case where q^a does not support any strategy in Q^a). Turn to condition (ii). Using condition (iii) already established for q^a, $\operatorname{Supp}\varphi^a \backslash \{q^a\} \subseteq Q^a$. So, by Lemma A4, any $s^b \in Q^b$ is admissible with respect to $\overline{Q}^a \times S^b$. ■

The IA set is one SAS (Proposition 5.1). We have just proved that SAS's are invariant to the fully reduced strategic form of the game. So, we have: Fix two matrices with the same fully reduced strategic form, and let $S^a_M \times S^b_M$ be the IA set for one of these matrices. Then, $S^a_M \times S^b_M$ induces an SAS of the second matrix. But does it induce the IA set of the other matrix? That is, is the IA set itself invariant? The answer is yes, in the following (restricted) sense.

Let $\overline{S}^i_m$ denote the set of m-admissible strategies for player i (where $i = a, b$) in the game $\overline{G}$.

Proposition A1 *For all m,*

(a) *if $\operatorname{Supp}\varphi^a \subseteq S^a_m$, then $\overline{S}^a_m \times \overline{S}^b_m = (S^a_m \cup \{q^a\}) \times S^b_m$,*

(b) *$\overline{S}^a_m \times \overline{S}^b_m = S^a_m \times S^b_m$ otherwise.*

Proof. The proof is by induction on m. For $m = 0$, the result is immediate. Assume the claim holds for m. We show it then holds for $m + 1$. This is immediate from the induction hypothesis if $\overline{S}^a_m = S^a_m$. So, we will suppose that $\overline{S}^a_m = S^a_m \cup \{q^a\}$.

We first show that if $\operatorname{Supp}\varphi^a \subseteq S^a_{m+1}$, then $\overline{S}^a_{m+1} = S^a_{m+1} \cup \{q^a\}$; and $\overline{S}^a_{m+1} = S^a_{m+1}$ otherwise. Certainly, if $s^a \in \overline{S}^a_{m+1}$ and $s^a \neq q^a$, then s^a is admissible given $S^a_m \times \overline{S}^b_m$. So, by the induction hypothesis, $\overline{S}^a_{m+1} \subseteq S^a_{m+1} \cup \{q^a\}$. Fix $s^a \in S^a_m$ that is admissible given $S^a_m \times S^b_m$. By Lemma A1, s^a is also admissible given $(S^a_m \cup \{q^a\}) \times S^b_m$. Again using Lemma A1, $\operatorname{Supp}\varphi^a \subseteq S^a_{m+1}$ if and only if $q^a \in S^a_{m+1}$. The claim then follows from the induction hypothesis.

Next, we show that $\overline{S}^b_{m+1} = S^b_{m+1}$. Since $\overline{S}^a_m = S^a_m \cup \{q^a\}$, $\operatorname{Supp}\varphi^a \subseteq S^a_m \subseteq \overline{S}^a_m$. (This is the induction hypothesis.) The result now follows from Lemma A4. ■

Proof of Lemma 7.1. Fix an admissible s^a, and also some $\varphi^a \in \mathcal{M}(S^a)$ with $\pi^a(\varphi^a, s^b) = \pi^a(s^a, s^b)$ for all $s^b \in S^b$. Without loss of generality, take $s^a \in \operatorname{Supp}\varphi^a$. Suppose, contra hypothesis, there exists $r^a, q^a \in \operatorname{Supp}\varphi^a$ with $\pi^a(r^a, r^b) > \pi^a(q^a, r^b)$ for some $r^b \in S^b$. Let h_1 be the last common predecessor of the terminal nodes $\zeta(r^a, r^b)$ and $\zeta(q^a, r^b)$ and note that Ann moves at h_1. Of course, there may be many choices of profiles (r^a, r^b) and (q^a, r^b). If so, choose profiles so that there does not exist another with a last common predecessor (strictly) following h_1.

We will first argue that there is some q^b that allows h_1, with $\pi^a(q^a, q^b) > \pi^a(r^a, q^b)$. If not, then $\pi^a(r^a, s^b) \geq \pi^a(q^a, s^b)$ for all s^b that allow h_1, with strict inequality for some s^b that allows h_1 (in particular, for r^b). Note that we can construct a strategy that allows h_1, agrees with r^a at node h_1 and onwards, but otherwise agrees with q^a. As such, q^a must be inadmissible. But this implies that s^a is inadmissible (Corollary A1), a contradiction.

For the remainder of the proof, we will take $\pi^a(r^a, r^b) \neq \pi^a(r^a, q^b)$. If this is not the case, then $\pi^a(q^a, q^b) > \pi^a(r^a, q^b) = \pi^a(r^a, r^b) > \pi^a(q^a, r^b)$, and a corresponding argument can then be made with respect to the pairs (q^a, q^b) and (q^a, r^b).

Let h_2 be the last common predecessor of $\zeta(r^a, r^b)$ and $\zeta(r^a, q^b)$ and note that Bob moves at h_2. Refer to Figure A.1 and note that, since Ann moves at h_1 and Bob moves at h_2, these nodes are distinct. Moreover, since (r^a, r^b) passes through both these nodes, they must be (strictly) ordered. In particular, h_2 must strictly follow h_1, since q^b allows h_1.

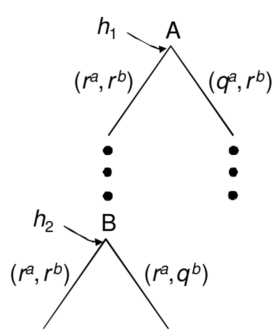

Figure A.1

Now pick strategies $\widehat{r}^b$ and $\widehat{q}^b$, each of which allows h_2. The two strategies agree in all but one respect: At and after h_2, the strategy $\widehat{r}^b$ agrees with r^b and the strategy $\widehat{q}^b$ agrees with q^b.

Fix a strategy $u^a \in \operatorname{Supp} \varphi^a$ that allows h_2. Then, for every strategy s^b that allows h_2, we have $\pi^a\left(u^a, s^b\right) = \pi^a\left(r^a, s^b\right)$. (If not, we contradict our choice of $\left(r^a, r^b\right)$ and $\left(q^a, r^b\right)$.) Given this, we can write

$$\pi^a(s^a, \widehat{r}^b) = \varphi^a\left(S^a\left(h_2\right)\right)\pi^a(r^a, \widehat{r}^b) + c,$$
$$\pi^a(s^a, \widehat{q}^b) = \varphi^a\left(S^a\left(h_2\right)\right)\pi^a(r^a, \widehat{q}^b) + c,$$

where

$$c = \sum_{w^a \notin S^a(h_2)} \varphi^a\left(w^a\right)\pi^a(w^a, \widehat{r}^b) = \sum_{w^a \notin S^a(h_2)} \varphi^a\left(w^a\right)\pi^a(w^a, \widehat{q}^b).$$

(This last equality uses the fact that, by construction, $\zeta(w^a, \widehat{r}^b) = \zeta(w^a, \widehat{q}^b)$ whenever w^a does not allow $S^a\left(h_2\right)$.)

Since $\varphi^a(S^a\left(h_2\right)) > 0$, we have $\pi^a(s^a, \widehat{r}^b) \neq \pi^a(s^a, \widehat{q}^b)$. By choice of $\widehat{r}^b$ and $\widehat{q}^b$, it follows that s^a allows h_2. So, given that we chose $s^a \in \operatorname{Supp} \varphi^a$ (again using the facts established in the previous paragraph), we have $\pi^a(s^a, \widehat{r}^b) = \pi^a(r^a, \widehat{r}^b)$ and $\pi^a(s^a, \widehat{q}^b) = \pi^a(r^a, \widehat{q}^b)$. This implies

$$(1 - \varphi^a(S^a(h_2)))\pi^a(s^a, \widehat{r}^b) = c = (1 - \varphi^a(S^a(h_2)))\pi^a(s^a, \widehat{q}^b).$$

Since $\pi^a(s^a, \widehat{r}^b) \neq \pi^a(s^a, \widehat{q}^b)$ this can occur only if $\varphi^a(S^a\left(h_2\right)) = 1$. But, q^a does not allow h_2 (refer to Figure A.1) and $q^a \in \operatorname{Supp} \varphi^a$. So, $\varphi^a(S^a\left(h_2\right)) < 1$, a contradiction. ■

Finally, we prove that in a PI tree satisfying SPC, there is an admissible BI strategy profile. This was used in the proof of Proposition 7.2.

Fix a PI tree Γ and let N^a (resp. N^b) be the set of nodes at which a (resp. b) moves. It will also be useful to fix the following terminology, tailored to a PI tree.

Definition A1 *Say r^a **weakly dominates** s^a **at node** $n \in N^a$ if s^a, r^a allow n and:*

(i) for each s^b that allows n, $\pi^a\left(r^a, s^b\right) \geq \pi^a\left(s^a, s^b\right)$;

(ii) for some s^b that allows n, $\pi^a\left(r^a, s^b\right) > \pi^a\left(s^a, s^b\right)$.

*Say s^a is **weakly dominated at** n if there is an r^a that weakly dominates s^a at n. Say s^a is **admissible at** n if it is not weakly dominated at n.*

Note, if s^a precludes n, then s^a is admissible at n. The following lemma is immediate by finiteness of S^a.

Lemma A5 *If s^a is inadmissible at n, then there is an admissible strategy r^a that weakly dominates s^a at n.*

Lemma A6 *Fix a PI tree. A strategy s^a is admissible if and only if it is admissible at each $n \in N^a$.*

Proof. Suppose s^a is admissible. Then, applying Proposition 3.1 in Brandenburger-Friedenberg [15], s^a is admissible at each $n \in N^a$. For the converse, suppose, s^a is admissible at each node $n \in$

N^a. Then, applying Lemma 4 in Battigalli [6] and Proposition 3.1 in Brandenburger-Friedenberg [15], we get that s^a is admissible. ■

Proposition A2 *Fix a PI tree satisfying SPC. There is an admissible BI strategy profile.*

Proof. Fix a BI profile (s^a, s^b). We will construct a new BI profile, viz. (r^a, s^b), so that r^a is admissible. Once we do this, we can apply the same argument to construct a new BI profile, viz. (r^a, r^b), so that r^b is also admissible. This will complete the proof.

If s^a is admissible, we are done. So, suppose s^a is inadmissible. Then, by Lemma A6, there is a node $n \in N^a$ at which s^a is weakly dominated. Let $n_1, .., n_K$ be a list of all nodes (for a) where: (i) s^a is weakly dominated at each n_k; and (ii) s^a is admissible at each node n that precedes n_k. (Note, the nodes $n_1, .., n_K$ cannot be ordered. Also, each $n_k \in N^a$.) We will use these nodes to construct inductively strategies $f(s^a, n_k)$, $k = 1, \ldots, K$, for a.

Begin with node n_1. By Lemmas A6 and A5, there is an admissible strategy q^a that weakly dominates s^a at n_1. Construct a strategy $f^a(s^a, n_1)$ as follows: Let $f^a(s^a, n_1)$ coincide with q^a at n_1 and each node that follows n_1, provided q^a allows the node. Otherwise, let $f^a(s^a, n_1)$ coincide with s^a. Now suppose $f^a(s^a, n_k)$ is defined. Consider the node n_{k+1} and an admissible strategy q^a (possibly different from earlier q^a's, of course) that weakly dominates s^a at n_{k+1}. (Again, we use Lemmas A6 and A5.) Construct $f^a(s^a, n_{k+1})$ analogously: Let $f^a(s^a, n_{k+1})$ coincide with q^a at n_{k+1} and each node that follows n_{k+1}, provided q^a allows the node. Otherwise, let $f^a(s^a, n_{k+1})$ coincide with $f^a(s^a, n_k)$. Denote by r^a the resulting strategy $f^a(s^a, n_K)$.

Note that r^a is admissible. Indeed, by the above construction and Lemma A6, r^a is admissible at each node $n \in N^a$. Now apply Lemma A6 to get the result. Next, we turn to showing that (r^a, s^b) is a BI profile. Here is the idea. First, we show that, for each node n, both a and b are indifferent between any (r^a_n, s^b_n) and (s^a_n, s^b_n), where we write r^a_n (resp. s^a_n or s^b_n) for a strategy that allows n and thereafter agrees with r^a (resp. s^a or s^b). From this we will conclude that (r^a, s^b) is a BI profile.

Step I: For each node n, $\pi^a(r^a_n, s^b_n) = \pi^a(s^a_n, s^b_n)$ and $\pi^b(s^b_n, r^a_n) = \pi^b(s^b_n, s^a_n)$. To show this, it suffices to consider a node n at which a moves. If r^a coincides with s^a at n and each node that follows n, certainly $\pi^a(r^a_n, s^b_n) = \pi^a(s^a_n, s^b_n)$. If not, then r^a must weakly dominate s^a at some node that (weakly) proceeds n. It follows that $\pi^a(r^a_n, s^b_n) \geq \pi^a(s^a_n, s^b_n)$. Moreover, given that (s^a, s^b) is a BI profile, it follows that $\pi^a(s^a_n, s^b_n) \geq \pi^a(r^a_n, s^b_n)$. So, $\pi^a(r^a_n, s^b_n) = \pi^a(s^a_n, s^b_n)$. Now, note that a moves at the last common predecessor of $\zeta(r^a_n, s^b_n)$ and $\zeta(s^a_n, s^b_n)$. So, by SPC, $\pi^b(s^b_n, r^a_n) = \pi^b(s^b_n, s^a_n)$, as desired.

Step II: For each node $n \in N^a$ (resp. $n \in N^b$), r^a_n (resp. s^b_n) maximizes $\pi^a(\cdot, s^b_n)$ (resp. $\pi^b(s^b_n, \cdot)$) among all strategies that allow n. Notice, this statement is immediate from Step I and the fact that (s^a, s^b) is a BI profile, provided n is a penultimate node (i.e., a "last move" in the tree). Then, assuming the statement holds for all nodes that follow n, the result again follows (for n) from Step I and the fact that (s^a, s^b) is a BI profile.

It is immediate from Step II that (r^a, s^b) is a BI profile. ■

References

[1] Asheim, G.B., "Proper Rationalizability in Lexicographic Beliefs," *International Journal of Game Theory*, 30, 2001, 453-478.

[2] Asheim, G.B., and M. Dufwenberg, "Admissibility and Common Belief," *Games and Economic Behavior*, 42, 2003, 208-234.

[3] Asheim, G.B., and A. Perea, "Sequential and Quasi-Perfect Rationalizability in Extensive Games," *Games and Economic Behavior*, 53, 2005, 15-42.

[4] Aumann, R., "On the Centipede Game," *Games and Economic Behavior*, 23, 1998, 97-105.

[5] Battigalli, P., "Strategic Rationality Orderings and the Best Rationalization Principle," *Games and Economic Behavior*, 13, 1996, 178-200.

[6] Battigalli, P., "On Rationalizability in Extensive Games," *Journal of Economic Theory*, 74, 1997, 40-61.

[7] Battigalli, P., and A. Friedenberg, "Context-Dependent Forward-Induction Reasoning," 2009. Available at www.public.asu.edu/∼afrieden.

[8] Battigalli, P. and M. Siniscalchi, "Strong Belief and Forward-Induction Reasoning," *Journal of Economic Theory*, 106, 2002, 356-391.

[9] Ben Porath, E., "Rationality, Nash Equilibrium and Backwards Induction in Perfect-Information Games," *Review of Economic Studies*, 64, 1997, 23-46.

[10] Bernheim, D., "Rationalizable Strategic Behavior," *Econometrica*, 52, 1984, 1007-1028.

[11] Blume, L., A. Brandenburger, and E. Dekel, "Lexicographic Probabilities and Choice Under Uncertainty," *Econometrica*, 59, 1991, 61-79.

[12] Börgers, T., "Weak Dominance and Approximate Common Knowledge," *Journal of Economic Theory*, 64, 1994, 265-276.

[13] Börgers, T. and L. Samuelson, "Cautious Utility Maximization and Iterated Weak Dominance," *International Journal of Game Theory*, 21, 1992, 13-25.

[14] Brandenburger, A., "Lexicographic Probabilities and Iterated Admissibility," in Dasgupta, P., D. Gale, O. Hart, and E. Maskin (eds.), *Economic Analysis of Markets and Games*, MIT Press, 1992, 282-290.

[15] Brandenburger, A. and A. Friedenberg, "The Relationship Between Rationality on the Matrix and the Tree," 2003. Available at www.stern.nyu.edu/∼abranden.

[16] Brandenburger, A. and A. Friedenberg, "What is Backward Induction?" 2009. Available at www.stern.nyu.edu/∼abranden.

[17] Brandenburger, A, A. Friedenberg, and H.J. Keisler, "Notes on the Relationship Between Strong Belief and Assumption," 2006. Available at www.stern.nyu.edu/∼abranden.

[18] Brandenburger, A, A. Friedenberg, and H.J. Keisler, "Admissibility in Games," *Econometrica*, 76, 2008, 307–352.

[19] Brandenburger, A, A. Friedenberg, and H.J. Keisler, "Supplement to 'Admissibility in Games,'" *Econometrica Supplementary Material*, 76, 2008. Available at http://econometricsociety.org/ecta/Supmat/5602_extensions.pdf.

[20] Dalkey, N., "Equivalence of Information Patterns and Essentially Determinate Games," in H. Kuhn and A. Tucker (eds.), *Contributions to the Theory of Games*, Vol. 2, 1953, Princeton: Princeton University Press, 217-244.

[21] Dekel, E., and D. Fudenberg, "Rational Behavior with Payoff Uncertainty," *Journal of Economic Theory*, 52, 1990, 243-267.

[22] Elmes, S. and P. Reny., "On the Strategic Equivalence of Extensive Form Games," *Journal of Economic Theory*, 62, 1994, 1-23.

[23] Ewerhart, C., "Rationality and the Definition of Consistent Pairs," *International Journal of Game Theory,* 27, 1998, 49–59.

[24] Kohlberg, E., and J-F. Mertens, "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1986, 1003-1038.

[25] Kuhn, H., "Extensive Games," *Proceedings of the National Academy of Sciences*, 36, 1950, 570-576.

[26] Kuhn, H., "Extensive Games and the Problem of Information," in Kuhn, H., and A. Tucker (eds.), *Contributions to the Theory of Games*, Vol. 2, 1953, Princeton University Press, 193-216.

[27] Marx, L. and J. Swinkels, "Order Independence for Iterated Weak Dominance," *Games and Economic Behavior*, 18, 1997, 219-245.

[28] Mertens, J-F., "Stable Equilibria–A Reformulation," *Mathematics of Operations Research*, 14, 1989, 575-625.

[29] Pearce, D., "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1984, 1029-1050.

[30] Rosenthal, R., "Games of Perfect-Information, Predatory Pricing, and the Chain-Store Paradox," *Journal of Economic Theory*, 25, 1981, 92-100.

[31] Samuelson, L., "Dominated Strategies and Common Knowledge," *Games and Economic Behavior*, 4, 1992, 284-313.

[32] Selten, R., "The Chain Store Paradox," *Theory and Decision,* 9, 1978, 127–159.

[33] Shimoji, M., and J. Watson, "Conditional Dominance, Rationalizability, and Game Forms," *Journal of Economic Theory*, 83, 1998, 161-195.

[34] Stuart, H., "Common Belief of Rationality in the Finitely Repeated Prisoners' Dilemma," *Games and Economic Behavior*, 19, 1997, 133-143.

[35] Thompson, F., "Equivalence of Games in Extensive Form," Research Memorandum RM-759, 1952, The RAND Corporation.