Are Admissibility and Backward Induction Consistent?*

Adam Brandenburger† Amanda Friedenberg‡

January 4, 2011

Abstract

We revisit a fundamental question in the axiomatic approach to non-cooperative games–viz., the consistency of admissibility (A) and backward induction (BI). The literature has concluded that A and BI are consistent. However, we argue that, to reach this conclusion, the literature has implicitly assumed that BI satisfies a monotonicity property: If a solution concept satisfies BI, then a refinement of the solution concept also satisfies BI. We provide a formalization of BI in terms of fundamentals of the game and, from this, conclude that BI is a non-monotonic property. In fact, we show that A and BI are inconsistent on the domain of all games. It appears to be an open question whether they are consistent on a subdomain of games on which there is a well-defined BI outcome.

1 Introduction

Admissibility (i.e., the avoidance of weakly dominated strategies) and backward induction are basic properties to demand of a solution concept. Indeed, they are viewed as fundamental requirements in the axiomatic approach to non-cooperative game theory. (See, e.g., Kohlberg and Mertens (9, 1986) and, for a recent survey, Govindan and Wilson (6, 2008).) As such, a basic question arises: Are admissibility (A) and backward induction (BI) consistent?

The question of the consistency of A and BI is old. Indeed, the common view is that the question is closed—that A and BI are, in fact, consistent. Two results point to this conclusion:

- Van Damme (19, 1984, Proposition 1, p.9) observed that any quasi-perfect equilibrium is a sequential equilibrium.

*We thank Hari Govindan and Bill Zame for very helpful conversations. Selvin Akkus, Heski Bar-Isaac, Pierpaolo Battigalli, Willemien Kets, Priscilla Man, Andy McLennan, John Nachbar, Alex Peysakhovich, Carlos Pimienta, Ariel Ropek, Marciano Siniscalchi, Harborne Stuart, and seminar participants at UCLA and the World Congress of the Econometric Society provided valuable input. Support from the Stern School of Business, the W.P. Carey School of Business, and the Department of Economics at UCLA is gratefully acknowledged. abc-01-04-11

†Address: Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012, adam.brandenburger@stern.nyu.edu, www.stern.nyu.edu/∼abranden

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

- Kohlberg and Mertens (9, 1986, Proposition 0, p.1009) showed that any proper equilibrium induces a sequential equilibrium.

Recall, both quasi-perfect and proper equilibrium satisfy A. Moreover, the argument goes, both satisfy BI since both induce a sequential equilibrium. The consistency of A and BI follows.

But, do quasi-perfect equilibrium and proper equilibrium satisfy BI? Let us review the argument for "yes." Recall, in perfect-information games (satisfying a no-ties condition), any quasi-perfect equilibrium—and, therefore, any proper equilibrium—yields the BI outcome. In general games (i.e., both perfect- and imperfect-information games), sequential equilibrium is commonly thought to embody BI. As refinements of sequential equilibrium, quasi-perfection and properness are presumed to inherit this property and, therefore, to satisfy BI in general.

Notice two key steps in this argument. The first is that there is a 'direct' definition of BI, and sequential equilibrium satisfies the definition. The second is that if a solution concept satisfies BI, then a refinement of the concept also satisfies BI. Does each of these steps hold up to scrutiny?

Take the first step. What do we mean by a direct definition of BI? We mean a test on a solution concept so that a solution concept satisfies BI if and only if it passes the test. Our position—which is somewhat controversial—is that it is precisely such a test that has been missing from the literature.

The early literature attempted to provide such a test. (See Section 3.) But, it fell short of providing a conclusive definition. Instead, the modern literature turned to, what we would call, an 'indirect' definition. By this we mean a test for whether a given solution concept satisfies BI in terms of other solution concepts (often sequential equilibrium, but also other solution concepts), as opposed to a test based on the solution concept itself. For example, a solution concept is often said to satisfy BI if each component of the solution contains a sequential-equilibrium outcome. (See, e.g., the survey by Govindan and Wilson (6, 2008).) From the point of view of axiomatics, such an indirect definition seems to us to be rather unconventional, precisely because it refers to other solution concepts. More commonly, in axiomatics, when we ask whether a particular member of some family of objects satisfies a certain property, we do not look at other objects in the family.[1]

What, then, is our BI test? It is a requirement that the solution of the whole tree be induced by the solution on each of the subtrees—a property we will call difference (D). We will argue that D captures the essential idea underlying BI. We show that sequential equilibrium does satisfy D and so, in this sense, it satisfies BI.

Now, the second step: If a solution concept satisfies BI, then a refinement of the concept also satisfies BI. Given our definition of BI, this is false. We show that while sequential equilibrium indeed satisfies BI, quasi-perfection and properness do not. There is a basic non-monotonicity in whether or not a solution concept satisfies BI. (This is the case even if we restrict attention to a class of games satisfying a 'no-ties' condition.)

[1] An analogy may help: To test whether a map f is continuous, we need look no further than f itself (and the relevant topologies). In particular, we do not need to look at other functions g in whatever family of functions we have in mind.

In light of this non-monotonicity, the question of the consistency of A and BI appears to be open. Indeed, the question becomes: Are A and D consistent? The question can be asked on at least two domains. First, are A and D consistent on the domain of all (perfect-recall) trees? Second, we can restrict the question to trees satisfying a no-ties condition, so that there is a well-defined BI outcome in perfect-information games satisfying the condition. We propose what appears to be a minimal such no-ties condition and ask: Are A and D consistent on the domain of all such (perfect-recall) trees?

Begin with the first question, i.e., the consistency of A and D on the domain of all trees. An example in Kohlberg and Mertens (9, 1986, Figure 5, p.1013) already shows that the answer is no. It follows that A and BI are inconsistent (on all trees). But, note carefully that Kohlberg and Mertens did not reach this conclusion themselves, precisely because they did not have a direct definition of BI. Indeed, their paper suggests the opposite—that A and BI are consistent.

Next, the consistency of A and D—and, therefore, of A and BI—on the domain of trees satisfying our no-ties condition. This question appears to be open. In our view, answering the question is fundamental to the axiomatic approach to non-cooperative game theory.

To review: This paper takes as given that A and BI are desirable properties for a solution concept to satisfy, and asks about the consistency of these properties. A key step in doing so is formalizing BI. We lay out our formalization, viz. D, in Sections 3-6. For now, we want to emphasize the philosophy behind our approach. We think it is important to give a direct definition of BI, i.e., a test on a solution concept that does not refer to other solution concepts. In our view, the indirect definitions in the literature cannot be viewed as primitive. Of course, this is not to say that an indirect definition is necessarily problematic. Perhaps, there is a theorem that says that a particular indirect approach is equivalent to a direct approach. But, it would seem that, for this, we would still need a direct definition of BI. If we take D as the direct definition of BI, then the non-monotonicity of D shows that the indirect approaches advocated to date do not work this way.

Without doubt, this viewpoint is controversial. Indeed, one of the leading experts in refinements wrote to us:

> "I have to disagree with your comments ... concerning the reduced-form nature of the axiomatic approach.... What you are objecting to is then including an axiom of type (b) [i.e., of the existing form]. I'm afraid if we go down that path, we are constraining ourselves severely and an axiomatic approach may never be possible."[2]

In effect, the plea is for a more pragmatic approach to defining axioms—one that sidesteps potential inconsistency problems. We are not ready to accede to this approach. First, as we said above, we do not know whether or not a consistency problem arises on the family of games that seems to us to be of most interest. Second, suppose inconsistency is, in fact, found. We see no a priori reason to think this would be a dead-end. It is easy to call to mind inconsistency results in various fields (e.g., the inconsistency of naive set theory) that have spurred important subsequent developments.

[2] Hari Govindan: Personal communication.

2 Formulation

We fix the following notation throughout. Given sets $X_1, \ldots, X_I$, write $X = \times_{i=1}^I X_i$ and $X_{-i} = \times_{j \neq i} X_j$. Likewise, given maps $f_i : X_i \to Y_i$, $i = 1, \ldots, I$, write $f : X \to Y$ for the product map, i.e., $f(x_1, \ldots, x_I) = (f_1(x_1), \ldots, f_I(x_I))$. Define product maps $f_{-i} : X_{-i} \to Y_{-i}$ analogously. If X is either a finite or a closed subset of $\mathbb{R}^n$, let $\mathcal{M}(X)$ be the set of Borel probability measures on X. Write $\operatorname{Supp} \mu$ for the support of μ.

We consider finite extensive-forms of perfect recall. We take the definition of Kuhn (11, 1950)-(12, 1953), with the exception that we allow a non-terminal node to have only one outgoing branch (rather than two). Let N (resp. Z) be the set of non-terminal (resp. terminal) nodes of such an extensive-form. The players are labelled $i = 1, \ldots, I$. Write H_i for the family of information sets for player i and $H = \bigcup_{i=1}^I H_i$ for the family of all information sets. (Recall, under the Kuhn definition of a tree, an information set is a subset of N.) Write $M_i[h]$ for the set of moves m available to i at $h \in H_i$. (Recall, under the Kuhn definition of a tree, a move is a subset of N.) A pure strategy s_i for player i maps each $h \in H_i$ to some $m_i \in M_i[h]$. Write S_i for the set of pure strategies for player i, and $\Sigma_i = \mathcal{M}(S_i)$ (with typical element σ_i) for the set of mixed strategies. The map $\zeta : S \to Z$ takes each pure-strategy profile into the terminal node it reaches.

A extensive-form game, written Γ, is an extensive-form plus extensive-form payoff functions for each player. Let $\Pi_i : Z \to \mathbb{R}$ be the payoff function for player i. The outcome map $\Pi : Z \to \mathbb{R}^I$ is given by $\Pi(z) = (\Pi_1(z), \ldots, \Pi_I(z))$. Terminal nodes $z, \widetilde{z} \in Z$ are **outcome equivalent** if $\Pi(z) = \Pi(\widetilde{z})$. (Note that Π need not be injective.) Write $\pi_i : S \to \mathbb{R}$ for player i's strategic-form payoff function, i.e., $\pi_i = \Pi_i \circ \zeta$. Extend π_i to $\Sigma_i \times \Sigma_{-i}$ in the usual way. At times, we will consider payoff functions that satisfy a no-ties condition:

Definition 2.1 *A game tree Γ satisfies the **Single-Payoff Condition (SPC)** if, for all $z, \widetilde{z} \in Z$, the following holds: If i moves at the last common predecessor of z and $\widetilde{z}$, then $\Pi_i(z) = \Pi_i(\widetilde{z})$ implies $\Pi(z) = \Pi(\widetilde{z})$.*

In words, a game satisfies SPC if, whenever player i is indifferent between two terminal nodes over which he is decisive, those two terminal nodes are outcome equivalent. It is clear that in a perfect-information (PI) tree satisfying SPC, there is a unique BI outcome. Moreover, SPC appears to be a minimal requirement for this purpose.

A strategy profile $\sigma \in \Sigma$ induces a distribution over outcomes, viz., the measure in $\mathcal{M}(\mathbb{R}^I)$ given by the image measure of σ under $\Pi \circ \zeta$. In particular, the probability of outcome $x \in \mathbb{R}^I$ is $\sigma((\Pi \circ \zeta)^{-1}(x))$. Call strategy profiles σ and $\widetilde{\sigma}$ **outcome equivalent** if they induce the same distribution on outcomes. Note, we can (and do) define this notion of outcome equivalence, even when σ and $\widetilde{\sigma}$ are strategy profiles in two (possibly different) I-player games. Likewise, given subsets of strategy profiles $Q \subseteq \Sigma$ and $\widetilde{Q} \subseteq \Sigma$ (of two, possibly different, I-player games), say that Q **induces the same outcomes as** $\widetilde{Q}$ if, for each $\widetilde{\sigma} \in \widetilde{Q}$, there is some $\sigma \in Q$ such that σ and $\widetilde{\sigma}$

are outcome equivalent. Call Q and $\widetilde{Q}$ **outcome equivalent** if Q induces the same outcomes as $\widetilde{Q}$, and $\widetilde{Q}$ induces the same outcomes as Q.

Say $\sigma_i \in \Sigma_i$ (resp. $\sigma_{-i} \in \Sigma_{-i}$) **allows** an information set h if there is some s_i with $\sigma_i(s_i) > 0$ (resp. s_{-i} with $\sigma_{-i}(s_{-i}) > 0$) such that s_i (resp. s_{-i}) allows h. Say $\sigma_i \in \Sigma_i$ (resp. $\sigma_{-i} \in \Sigma_{-i}$) **reaches** an information set h if, for each s_i with $\sigma_i(s_i) > 0$ (resp. s_{-i} with $\sigma_{-i}(s_{-i}) > 0$), s_i (resp. s_{-i}) allows h. Write $\Sigma_i(h)$ (resp. $\Sigma_{-i}(h)$) for the set of strategies σ_i (resp. σ_{-i}) that reach h. (Note carefully that we abuse notation here, since $\Sigma_{-i}(h)$ need not be a product set.)

Say a strategy profile $\sigma \in \Sigma$ **allows a move** m if $m \in M_i[h]$, where h is allowed by σ, and m is played with strictly positive probability under σ. Given a subset of strategy profiles $Q \subseteq \Sigma$, say Q **allows a move** m if there is some $\sigma \in \Sigma$ which allows m.

A **solution concept** $\mathcal{S}$ associates with each game tree Γ a family of subsets of strategy profiles for Γ. Formally, a solution concept $\mathcal{S}$ (on a family of games $\mathcal{G}$) maps each tree (in $\mathcal{G}$) to a family of subsets of strategy profiles for the tree, i.e. $\mathcal{S}(\Gamma) \subseteq 2^{\Sigma}$. The family $\mathcal{S}(\Gamma)$ is called the **solution** of Γ. Each element of $\mathcal{S}(\Gamma)$, i.e., each subset of mixed-strategy profiles $Q \in \mathcal{S}(\Gamma)$, is called a **component** of the solution. Some familiar examples: For Nash equilibrium, we could take the solution of a game to consist of multiple components, where each component is a singleton and consists of a particular Nash equilibrium. Or, following Kohlberg and Mertens (9, 1986) and their successors, we could take each component to consist of a connected set of Nash equilibria. For iterated (strong or weak) dominance, we could take the solution to consist of a single component—viz., all the iterated undominated profiles.

A solution concept $\mathcal{R}$ is a **refinement** of $\mathcal{S}$ if, for each game Γ and every $R \in \mathcal{R}(\Gamma)$, there is a $Q \in \mathcal{S}(\Gamma)$ so that Q induces the same outcomes as R.

We note that we have given our definitions in terms of mixed strategies. Of course, some solution concepts (e.g., sequential equilibrium) are defined using behavioral strategies. When needed, we will understand all the preceding definitions to be in terms of behavioral strategies, and use the notation β_i for a behavioral strategy for player i.

3 Backward Induction

The intuitive idea behind BI is clear: Fix a tree Γ and a subtree Δ of Γ. Now discard Δ, leaving behind only the solution on this subtree—leaving behind the 'ghost' of the subtree, if you like. Then, we do not change our original analysis.

While the idea is intuitively clear, it is less clear how to formalize it. There have been a number of attempts at such a definition, especially in the earlier literature. (See, especially, Kohlberg and Mertens (9, 1986, pp.1012-1013) and Hillas and Kohlberg (7, 2002).) Here, we argue that the correct definition is given by, what we call, the Difference property. We also argue that alternative definitions proposed by the literature do not succeed in capturing BI. Moreover, the Difference definition does not suffer from the same drawbacks as these alternative definitions.

Defining Backward Induction

To formalize the idea of BI, we first need to specify what it means to delete a subtree, leaving behind only the solution on the subtree. The relevant concept goes back to Kuhn (12, 1953, p.208); we will call it a **difference tree**. A difference tree is defined relative to a solution concept $\mathcal{S}$. Begin with a tree Γ and a subtree Δ of Γ. Fix a nonempty component of $\mathcal{S}(\Delta)$, which we will denote Q^{Δ}. The $(\mathcal{S}, Q^{\Delta})$-difference tree is obtained by deleting from the original tree Γ any move not allowed by Q^{Δ}. It is readily verified that each $(\mathcal{S}, Q^{\Delta})$-difference tree is a well-defined game tree. (This uses the fact that we required Q^{Δ} to be nonempty.) Write $\Gamma_{\mathcal{S},Q^{\Delta}}$ for the $(\mathcal{S}, Q^{\Delta})$-difference tree. Note, the difference tree depends on a solution concept, subtree, and particular component of the solution on the subtree.

Now, the Difference property (D). To repeat, the idea is that we do not change our analysis when we replace a tree with a difference tree. We can now state this precisely:

(D) *A solution concept $\mathcal{S}$ satisfies* ***Difference*** *(on $\mathcal{G}$) if for each tree Γ (in $\mathcal{G}$) and each subtree Δ of Γ the following holds: If $Q \in \mathcal{S}(\Gamma)$, there is a nonempty component $Q^{\Delta} \in \mathcal{S}(\Delta)$ and a component $\widetilde{Q} \in \mathcal{S}(\Gamma_{\mathcal{S},Q^{\Delta}})$, such that $\widetilde{Q}$ induces the same outcomes as Q.*

Difference is a modification of property (BI3) in Kohlberg and Mertens (9, 1986, p.1012). (We discuss the relationship in the Appendix.) Figure 3.1 illustrates the definition. Loosely, it says that the solution on the whole tree should be included in the solution on what is left after replacing a subtree with what the solution allows on the subtree.

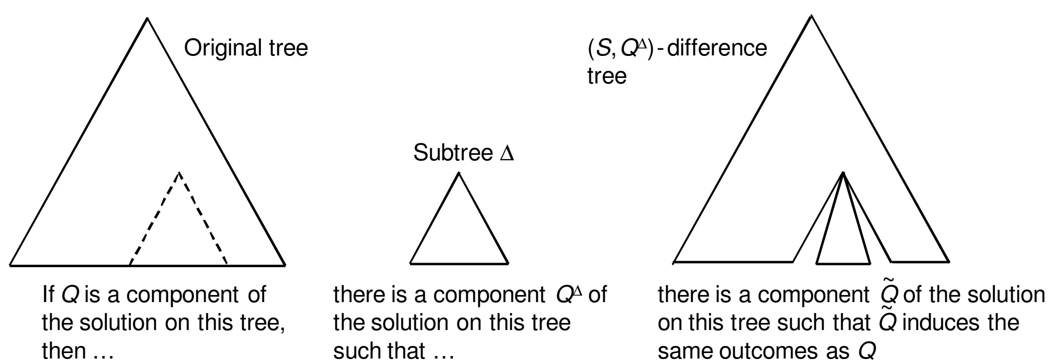

Figure 3.1

The Difference property works in much the same way as the BI algorithm. The algorithm works by using what it prescribes on future subtrees to pin down behavior on the current subtree. The Difference property applies this same principle to general trees. Solutions on subtrees yield difference trees, which are used to pin down the solution on the overall tree. Formally, each (distribution on) outcome(s) allowed by the solution on the overall tree must also be allowed by the solution on some difference tree.

Note, by itself, D is missing two essential components, which are built into the BI algorithm. First, it does not impose optimizing behavior. Second, it does not impose the requirement that the solution contain a nonempty component. Thus, to capture fully the idea of BI, we must impose D in the presence of two additional axioms: Rationality and Existence. We now define these axioms.

Say a strategy σ_i is optimal under $\sigma_{-i} \in \mathcal{M}(\Sigma_{-i})$, among strategies in $Q_i \subseteq \Sigma_i$, if $\sigma_i \in Q_i$ and $\pi_i(\sigma_i, \sigma_{-i}) \geq \pi_i(\rho_i, \sigma_{-i})$ for each $\rho_i \in Q_i$. (Writing $\sigma_{-i} \in \mathcal{M}(\Sigma_{-i})$ is a slight notational abuse.) A strategy σ_i is **(extensive-form) rational** if, for each information set $h \in H_i$ allowed by σ_i, there is some $\sigma_{-i} \in \mathcal{M}(\Sigma_{-i})$, with $\sigma_{-i}(\Sigma_{-i}(h)) = 1$, under which $\sigma_i(\cdot|\Sigma_i(h))$ is optimal among strategies in $\Sigma_i(h)$.

(R) *A solution concept $\mathcal{S}$ satisfies **Rationality** (on $\mathcal{G}$) if, for each tree Γ (in $\mathcal{G}$) and each component $Q \in \mathcal{S}(\Gamma)$, any profile $\sigma \in Q$ consists of rational strategies.*

Note, carefully, R does not require that players have 'correct' beliefs about the strategies played. (Take $\mathcal{S}$ to be the solution that maps each game into a single component, viz. $\mathcal{S}(\Gamma) = \{Q_i \times Q_{-i}\}$, where Q_i contains all the rational strategies for i in the game. Note, a strategy $s_i \in Q_i$ need not optimal under any $\sigma_{-i} \in \mathcal{M}(\Sigma_{-i})$, with $\sigma_{-i}(Q_{-i}) = 1$.) Thus, imposing R does not confine the solution concept to be an equilibrium refinement.

Next:

(E) *A solution concept $\mathcal{S}$ satisfies **Existence** (on $\mathcal{G}$) if, for each game Γ (in $\mathcal{G}$), there is a nonempty component of $\mathcal{S}(\Gamma)$.*

Taken together, D, R, and E are our definition of BI.

Alternative Proposals

BI is the idea that the solution of the whole tree should be pinned down by the solution on each of the subtrees. The literature has attempted to formalize this idea via various properties that relate the solution of the whole tree to the solution of its parts. Here, we look at two prominent such definitions: Reverse Difference and Projection. We show that neither truly captures BI. (The Appendix discusses other properties in the literature.)

Let us review the two definitions:

- *A solution concept $\mathcal{S}$ satisfies **Reverse Difference** (on $\mathcal{G}$) if for each tree Γ (in $\mathcal{G}$) and each subtree Δ of Γ the following holds: If $Q \in \mathcal{S}(\Gamma)$, there is a nonempty component $Q^{\Delta} \in \mathcal{S}(\Delta)$ and a component $\widetilde{Q} \in \mathcal{S}(\Gamma_{\mathcal{S},Q^{\Delta}})$, such that Q induces the same outcomes as $\widetilde{Q}$.*

- *A solution concept $\mathcal{S}$ satisfies **Projection** (on $\mathcal{G}$) if, for each tree Γ (in $\mathcal{G}$) the following holds: For each subtree Δ and component $Q \in \mathcal{S}(\Gamma)$, there is a component $Q^{\Delta} \in \mathcal{S}(\Delta)$ such that for each $\sigma \in Q$, the restriction of σ to the subtree Δ is contained in Q^{Δ}.*

Reverse Difference asks that the solution on the difference tree be pinned down by the solution on the whole tree. Projection asks that the solution on the whole tree be pinned down by the solution on the subtree (as opposed to the solution on the difference tree). Both properties are aimed at relating the solution on the whole tree to the solution on parts of the tree (i.e., on difference trees or subtrees). However, both are insufficient for the solution on the parts of the tree to pin down the solution on the whole tree.

To verify this claim, consider the game trees Γ in Figure 3.2, Δ in Figure 3.3 (which is the subtree of Γ that begins after Ann chooses In), and $\widetilde{\Gamma}$ in Figure 3.4. Suppose $\mathcal{S}$ satisfies E and R. Then $\mathcal{S}(\Delta) = \{\{Left\}\}$, so $\widetilde{\Gamma}$ is the $(\mathcal{S}, \{Left\})$-difference tree. Again, E and R imply that $\mathcal{S}(\widetilde{\Gamma}) = \{\{(Out, Left)\}\}$. If $\mathcal{S}(\Gamma) = \{\{Out, In\} \times \{Left\}\}$, then $\mathcal{S}$ satisfies both Reverse Difference, Projection, Rationality, and Existence on the domain $\{\Gamma, \Delta, \widetilde{\Gamma}\}$, despite the fact that $\mathcal{S}$ is not outcome equivalent to BI on this family of trees. Note, in this case, $\mathcal{S}$ fails D, since $(In, Left) \in \mathcal{S}(\Gamma)$ but $(In, Left) \notin \mathcal{S}(\widetilde{\Gamma})$. Of course, if $\mathcal{S}$ instead satisfies D, then $\mathcal{S}(\Gamma) = \{\{(Out, Left)\}\}$.

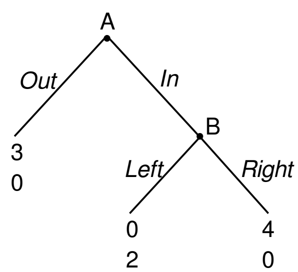

Figure 3.2

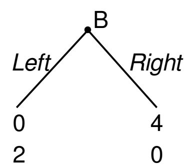

Figure 3.3

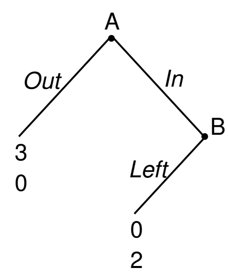

Figure 3.4

The problem here is clear: The idea of BI is that solutions on parts of the tree should be used to pin down the solution on the whole tree. If the notion of a "part" is a difference tree (as in Reverse Difference), but we instead require the reverse inclusion to D, then the solutions on difference trees do not pin down the solution on the whole tree. Likewise, if the notion of a "part" is a subtree (as in Projection), then, again, the solutions on the subtrees do not pin down the solution on the whole tree. To see the latter, return to the example. While the solution $\mathcal{S}(\Delta)$ is used to pin down Bob's behavior in Γ, it cannot be used to pin down Ann's behavior in Γ, because she has no move in Δ. Under D, even if Ann has no move in Δ, we can use a component Q^{Δ} of the solution on Δ to pin down Ann's behavior in Γ, since Ann does have a move in the associated difference tree.

Background Check

We have seen that a solution concept can satisfy Reverse Difference, Projection, Rationality, and Existence on a domain of PI trees, and yet the solution may fail to deliver the (unique) BI outcome. As such, these are incorrect implementations of the idea that the solution on the whole tree can be pinned down by the solution on the parts.

Now, we show that the definition of Difference does not suffer from the same flaw:

Proposition 3.1 *Fix a solution concept $\mathcal{S}$.*

(i) If $\mathcal{S}$ satisfies E, R, *and* D *on the domain of PI trees satisfying SPC, then each component of $\mathcal{S}$ is outcome equivalent to the* BI *algorithm on these trees.*

(ii) If each component of $\mathcal{S}$ is outcome equivalent to the BI *algorithm on every PI tree satisfying SPC, then $\mathcal{S}$ satisfies* E, R, *and* D *when restricting the domain of the solution concept to these trees.*

Proposition 3.1 says that E, R, and D characterize BI on the domain of trees satisfying SPC. Part (ii) is standard. But, part (i) is key: In part (ii), we could replace Difference with Reverse Difference or Projection. But, as we have seen, we cannot do the same for part (i).

Proof of Proposition 3.1. Part (i): The proof is by induction on the length of the tree. For a tree of length 1, the result is immediate from E, R, and the fact that the game satisfies SPC. So, suppose the statement holds for any tree of length l or less.

Fix a tree of length $l+1$, where i moves first and write Δ^k, $k = 1, \ldots, K$, for the immediate subtrees. For each such subtree Δ^k, fix a component Q^k of $\mathcal{S}(\Delta^k)$. Using the induction hypothesis, $Q^k \neq \emptyset$ and any $(\sigma_1^k, \ldots, \sigma_I^k) \in Q^k$ gives the unique BI outcome on that subtree.

Consider the tree obtained by deleting from each immediate subtree Δ^k any move not allowed by Q^k. Call this tree $\Gamma_{\mathcal{S}}$. Then, in $\Gamma_{\mathcal{S}}$, each of i's choices $k = 1, \ldots, K$ leads to a unique outcome in the associated subtree. Of course, these outcomes do not depend on the particular choices of Q^k.

Using SPC, all rational strategies for i (in $\Gamma_{\mathcal{S}}$) are outcome equivalent. By E, there is a nonempty component of $\mathcal{S}(\Gamma_{\mathcal{S}})$. By R, any such component must be outcome equivalent to BI in $\Gamma_{\mathcal{S}}$, and, therefore, outcome equivalent to BI in Γ.

Now, successively apply D to each subtree, so that any outcome allowed by any component of the solution on the overall tree must be allowed by the component on $\Gamma_{\mathcal{S}}$. (This uses the fact that there is a unique outcome in this difference tree and this outcome does not depend on the initial choice of solutions Q^k.) It follows that any outcome allowed by any component of the solution on the overall tree must be the BI outcome in that tree. By E, the solution must have some nonempty component, establishing part (i).

Part (ii): Fix a solution concept $\mathcal{S}$, as in the premise. It is immediate that $\mathcal{S}$ satisfies E and R. We show D. Fix a tree Γ satisfying SPC, so that there is a unique BI outcome. Fix also a subtree Δ and a component $Q^\Delta \in \mathcal{S}(\Delta)$. Consider the $(\mathcal{S}, Q^\Delta)$-difference tree $\Gamma_{\mathcal{S}, Q^\Delta}$. It, too, is a PI tree satisfying SPC, and so has a unique BI outcome. But this must coincide with the BI outcome in Γ, since deleting (from the subtree) any move precluded by Q^Δ does not delete the BI outcome in Γ. This establishes D. ■

4 Non-Monotonicity of Backward Induction

Now that we have a formal definition of BI, we can substantiate our claim that while sequential equilibrium satisfies BI, neither proper equilibrium nor quasi-perfect equilibrium does. This implies that the question of the consistency of A and BI is not answered by the results in van Damme (19, 1984, Proposition 1, p.9) and Kohlberg and Mertens (9, 1986, Proposition 0, p.1009). We believe that this non-monotonicity of the BI property is an issue that has been missed to date.

Begin with proper equilibrium and quasi-perfect equilibrium. Recall the definitions (Myerson (15, 1978), van Damme (19, 1984)). A profile of completely mixed strategies $\sigma^{\varepsilon} = (\sigma_1^{\varepsilon}, \ldots, \sigma_I^{\varepsilon})$ is an ε-**proper equilibrium** of Γ if, for each i, $\pi_i\left(s_i, \sigma_{-i}^{\varepsilon}\right) < \pi_i\left(r_i, \sigma_{-i}^{\varepsilon}\right)$ implies $\sigma_i^{\varepsilon}(s_i) \leq \varepsilon\sigma_i^{\varepsilon}(r_i)$. A profile σ is a **proper equilibrium** of Γ if there is a sequence of ε-proper equilibria σ^{ε} of Γ with $\lim_{\varepsilon\to 0}\sigma^{\varepsilon} = \sigma$. We define the proper equilibrium solution concept $\mathcal{S}_{\mathrm{PE}}$ by

$$\mathcal{S}_{\mathrm{PE}}(\Gamma) = \{\{\sigma\} : \sigma \text{ is a proper equilibrium of } \Gamma\}.$$

A profile of behavioral strategies $\beta = (\beta_1, \ldots, \beta_I)$ is a **quasi-perfect equilibrium** if there exists a sequence of completely mixed behavioral strategy profiles, viz. $\beta^k = (\beta_1^k, \ldots, \beta_I^k)$, so that $\beta^k \to \beta$ and, for each i, every $\beta_i(h_i)$ is optimal under μ (among strategies in $\Sigma_i(h_i)$). We define the quasi-perfect equilibrium solution concept $\mathcal{S}_{\mathrm{QPE}}$ by

$$\mathcal{S}_{\mathrm{QPE}}(\Gamma) = \{\{\beta\} : \beta \text{ is a quasi-perfect equilibrium of } \Gamma\}.$$

Proposition 4.1 *The solution concepts $\mathcal{S}_{PE}$ and $\mathcal{S}_{QPE}$ fail* D.

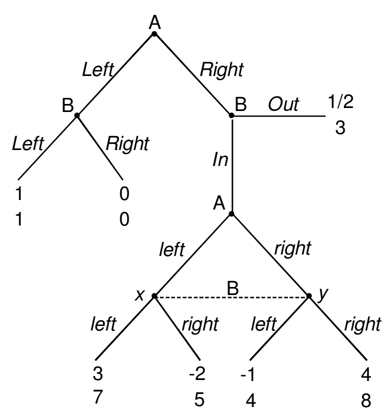

Figure 4.1

Proof. Consider the game Γ given in Figure 4.1. We first show that there is a proper equilibrium where Ann plays $Left$ (at the initial node) with probability one. To see this, note that there is an ε-proper equilibrium where Ann uses (unnormalized) weights (1 : *Left*, $\frac{2}{3}\varepsilon$: *Right-left*, $\frac{1}{3}\varepsilon$: *Right-right*) and Bob uses (unnormalized) weights (ε : *Left-Out*, ε^3 : *Right-Out*, $\frac{3}{5}$: *Left-In-left*, $\frac{3}{5}\varepsilon^2$: *Right-In-left*,

$\frac{2}{5}$: *Left-In-right*, $\frac{2}{5}\varepsilon^2$: *Right-In-right*). So, the outcome $(1,1)$ is allowed under properness. It follows that the outcome $(1,1)$ is also allowed under quasi-perfection. (See van Damme (19, 1984, Theorem 1, p.9).)

Now take Δ to be the subtree beginning at the node where Bob can choose *Out*. Writing Ann's (resp. Bob's) strategies for Δ in the order (*left, right*) (resp. (*Out*, *In-left*, *In-right*)), there are three subgame-perfect equilibria of the subtree: $((1,0),(0,1,0))$, $((0,1),(0,0,1))$, and $((\frac{2}{3},\frac{1}{3}),(0,\frac{3}{5},\frac{2}{5}))$. (Note that $Out\text{-}left$ and *Out-right* are strongly dominated in the subtree, and so cannot be part of a subgame-perfect equilibrium.) Each of these subgame-perfect equilibria are both proper and quasi-perfect.

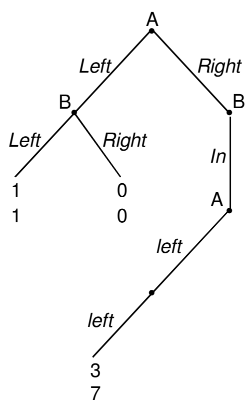

Figure 4.2

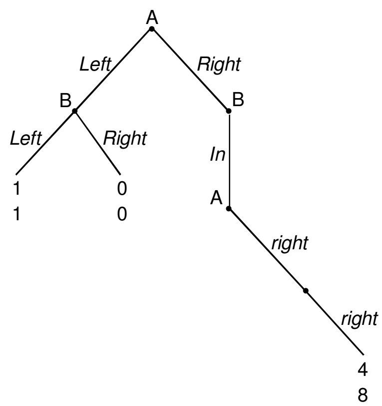

Figure 4.3

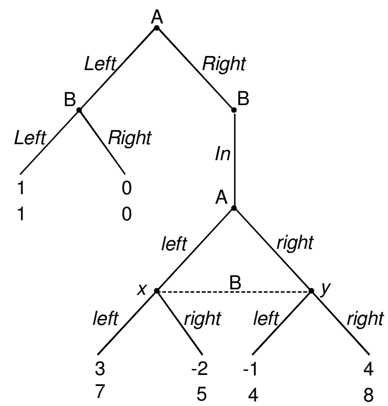

Figure 4.4

Thus, there are three possible difference trees for both properness and quasi-perfection. They are given as Figures 4.2-4.4. In each of these trees, the strategies $Left\text{-}left$ and $Left\text{-}right$ (for Ann) are weakly dominated. (In Figure 4.4, they are weakly dominated by a $\frac{1}{2}$:$\frac{1}{2}$ mixture of $Right\text{-}left$:*Right-right*.) Therefore, the outcome $(1,1)$ cannot arise in a proper or quasi-perfect equilibrium. This contradicts D. ■

Remark 4.1 *We could define the proper-equilibrium solution concept to consist of a single component containing all the proper equilibria of a game. We could do likewise for the quasi-perfect-equilibrium solution concept. The proof of Proposition 4.1 shows that these versions of the solution concepts fail* D*, too.*

Having established that quasi-perfection and properness fail D, we can now see that there is a basic non-monotonicity in whether or not a solution concept satisfies D.

Theorem 4.1 *There exists a solution concept $\mathcal{S}$ and a refinement $\mathcal{R}$ of $\mathcal{S}$, such that $\mathcal{S}$ satisfies* E*,* R*, and* D*, while $\mathcal{R}$ satisfies* E *and* R *but fails* D*.*

In the proof of the theorem, we will take $\mathcal{S}$ to be sequential equilibrium and $\mathcal{R}$ to be proper equilibrium. (We could equally take $\mathcal{R}$ to be quasi-perfect equilibrium.)

Recall the definition of sequential equilibrium (Kreps and Wilson (10, 1982)). A pair (β, μ) is an **assessment** if β is a profile of behavioral strategies and μ is a system of beliefs. (That is: $\mu : H \to \mathcal{M}(N)$ with each $\mu(h)(h) = 1$.) The assessment is **consistent** if there is a sequence $(\beta^k, \mu^k) \to (\beta, \mu)$ where each β^k is a profile of completely mixed behavioral strategies. (That is: For each i and $h_i \in H_i$, $\operatorname{Supp} \beta_i^k(h_i) = M_i[h_i]$ and each μ^k is derived from β^k by conditioning.) An assessment (β, μ) is a **sequential equilibrium** if it is consistent and, for each i, every $\beta_i(h_i)$ is optimal under μ (among strategies in $\Sigma_i(h_i)$). We define the sequential equilibrium solution concept $\mathcal{S}_{\text{SE}}$ by

$$\mathcal{S}_{\text{SE}}(\Gamma) = \{\{\beta\} : \text{there is a system of beliefs } \mu \text{ s.t. } (\beta, \mu) \text{ is a sequential equilibrium of } \Gamma\}.$$

For the connection to D, we need some more notation. Fix a solution concept $\mathcal{S}$, a tree Γ, a subtree Δ of Γ, and consider a difference tree $\Gamma_{\mathcal{S},Q^\Delta}$. We write $\bar{H}_i$ (resp. $\widetilde{H}$) for the family of i's (resp. the family of all) information sets in this difference tree. Write H for the family of information sets in Γ, and note that there is an injective mapping $\eta : \widetilde{H} \to H$ with $\widetilde{h} \subseteq \eta(\widetilde{h})$. Write $\widetilde{M}_i[\widetilde{h}_i]$ for the moves available to i at $\widetilde{h}_i$ in the difference tree, and note that, for each $\widetilde{h}_i$, there is an injective mapping $\xi[\widetilde{h}_i] : \widetilde{M}_i[\widetilde{h}_i] \to M_i[\eta(\widetilde{h}_i)]$ so that $\widetilde{m}_i \subseteq \xi[\widetilde{h}_i](\widetilde{m}_i)$. If $\widetilde{s}_i$ is a pure strategy for i in the difference tree, we write $[\widetilde{s}_i]$ for the set of pure strategies for i in Γ which coincide with $\widetilde{s}_i$ in the difference tree.

Proposition 4.2 *The solution concept $\mathcal{S}_{SE}$ satisfies* D.

Proof. Fix a tree Γ and some $\beta = (\beta_1, \ldots, \beta_I)$ with $\{\beta\} \in \mathcal{S}_{\text{SE}}(\Gamma)$. Then, there exists some system of beliefs $\mu : H \to \mathcal{M}(N)$ such that (β, μ) is a sequential equilibrium. Fix a subtree Δ. For each information set h_i of Δ, set $\beta_i^\Delta(h_i) = \beta_i(h_i)$ and $\mu^\Delta(h_i) = \mu(h_i)$. It is immediate that $(\beta^\Delta, \mu^\Delta)$ is a sequential equilibrium of Δ, i.e., $\{\beta^\Delta\} \in \mathcal{S}_{\text{SE}}(\Delta)$.

Construct the difference tree $\Gamma_{\text{SE},\{\beta^\Delta\}}$ by deleting from Γ any path (in Δ) that is played with zero probability under β^Δ. This amounts to deleting from Γ any path which is in Δ and which is played with zero probability under β. So, certainly, each $\beta_i(\eta(\widetilde{h}_i))(\xi[\widetilde{h}_i](\widetilde{M}_i[\widetilde{h}_i])) = 1$. Moreover, if $\eta(\widetilde{h})$ is in Δ, then $\eta(\widetilde{h})$ is reached with strictly positive probability under β^Δ. So, in this case, $\mu(\eta(\widetilde{h}))(\widetilde{h}) = \mu^\Delta(\eta(\widetilde{h}))(\widetilde{h}) = 1$. Indeed, this is true more generally, i.e., for each $\eta(\widetilde{h})$ (whether or not it is in Δ) $\mu(\eta(\widetilde{h}))(\widetilde{h}) = 1$. We use these facts repeatedly below.

Now, we define an assessment $(\widetilde{\beta}, \widetilde{\mu})$ of the difference tree $\Gamma_{\text{SE},\{\beta^\Delta\}}$. Choose $\widetilde{\beta} = (\widetilde{\beta}_1, \ldots, \widetilde{\beta}_I)$ so that each $\widetilde{\beta}_i(\widetilde{h}_i)$ satisfies $\widetilde{\beta}_i(\widetilde{h}_i)(\widetilde{m}_i) = \beta_i(\eta_i(\widetilde{h}_i))(\xi[\widetilde{h}_i](\widetilde{m}_i))$, for all $\widetilde{m}_i \in \widetilde{M}_i[\widetilde{h}_i]$. (Recall that each $\beta_i(\eta(\widetilde{h}_i))(\xi[\widetilde{h}_i](\widetilde{M}_i[\widetilde{h}_i])) = 1$, so this is well defined.) Likewise, choose $\widetilde{\mu}$ so that each $\widetilde{\mu}(\widetilde{h})(n) = \mu(\eta(\widetilde{h}))(n)$, for each node in $\widetilde{h}$. (Recall that each $\mu(\eta(\widetilde{h}))(\widetilde{h}) = 1$, so this is well-defined.) We will show that $(\widetilde{\beta}, \widetilde{\mu})$ is a sequential equilibrium of the difference tree, so that $\{\widetilde{\beta}\} \in \mathcal{S}(\Gamma_{\text{SE},\{\beta^\Delta\}})$. Since, by construction, any outcome allowed by β is allowed by $\widetilde{\beta}$, this will establish the result.

It is immediate from the construction that each $\widetilde{\beta}_i(\widetilde{h}_i)$ is a best reply under $\widetilde{\mu}$. So, it suffices to show that $(\widetilde{\beta}, \widetilde{\mu})$ is consistent.

Since (β, μ) is consistent, there is some $(\beta^k, \mu^k) \to (\beta, \mu)$ where each β^k is completely mixed and each μ^k is derived from β^k by conditioning. As such, $\beta_i^k(\eta(\widetilde{h}_i))(\xi[\widetilde{h}_i](\bar{M}_i[\widetilde{h}_i])) > 0$ and $\mu^k(\eta(\widetilde{h}_i))(\widetilde{h}_i) > 0$ for all $\widetilde{h}_i$. Define $(\widetilde{\beta}^k, \widetilde{\mu}^k)$ as follows: For each $\widetilde{h}_i$ and each $\bar{m}_i \in \bar{M}_i\left[\widetilde{h}_i\right]$, set $\widetilde{\beta}_i^k(\widetilde{h}_i)(\bar{m}_i) = \beta_i^k(\eta(\widetilde{h}_i))(\xi[\widetilde{h}_i](\bar{m}_i)|\xi[\widetilde{h}_i](\bar{M}_i[\widetilde{h}_i]))$. Likewise, for each $\widetilde{h}_i$ and each $n \in \widetilde{h}_i$, set $\widetilde{\mu}^k(\widetilde{h}_i)(n) = \mu^k(\eta(\widetilde{h}_i))(n|\widetilde{h}_i)$. Note, by construction $\widetilde{\beta}^k$ is completely mixed and $\widetilde{\mu}^k$ is derived from $\widetilde{\beta}^k$ by conditioning. Moreover, using the fact that each $\beta_i^k(\eta(\widetilde{h}))(\xi[\widetilde{h}_i](\bar{M}_i[\widetilde{h}_i])) \to 1, \mu^k(\eta(\widetilde{h}_i))(\widetilde{h}_i) \to 1$, it follows that $(\widetilde{\beta}_i^k, \widetilde{\mu}_i^k) \to (\widetilde{\beta}_i, \widetilde{\mu}_i)$ as required. ■

Remark 4.2 *We could define the sequential-equilibrium solution concept to consist of a single component containing all the sequential equilibria of a game. A proof like the one just given shows that this version of the solution concept satisfies* D*, too.*

Now we can prove Theorem 4.1.

Proof of Theorem 4.1. By Propositions 1 and 3 in Kreps and Wilson (10, 1982, p.876), $\mathcal{S}_{\mathrm{SE}}$ satisfies E and R. By Proposition 4.2 here, $\mathcal{S}_{\mathrm{SE}}$ also satisfies D. By Myerson (15, 1978, p.79), $\mathcal{S}_{\mathrm{PE}}$ satisfies E. For R, start with a proper equilibrium σ and an associated sequence of ε-proper equilibria σ^ε. There is an ε such that σ_i is optimal under σ^ε_{-i}. (See, e.g., Lemma 2.3.2 in van Damme (20, 1987, p.29).) Since σ^ε_{-i} has full support, it follows by a standard argument that σ_i is then (extensive-form) rational. So, $\mathcal{S}_{\mathrm{PE}}$ satisfies R. By Proposition 4.1 here, $\mathcal{S}_{\mathrm{PE}}$ fails D. Finally, by Theorem 1 in van Damme (19, 1984, p.9), $\mathcal{S}_{\mathrm{PE}}$ is a refinement of $\mathcal{S}_{\mathrm{SE}}$. ■

It is instructive to compare the behavior of sequential equilibrium with that of proper (or quasi-perfect) equilibrium in the game of Figure 4.1. Much as with proper (or quasi-perfect) equilibrium, there is a sequential equilibrium where: (i) Ann puts weight 1 on *Left*; and (ii) Bob puts weight 1 on *Left*, weight 1 on *In*, and weights $\frac{3}{5}$:$\frac{2}{5}$ on *left* vs. *right*. This is supported by an assessment for Bob that puts weights $\frac{2}{3}$:$\frac{1}{3}$ on node x vs. node y. Likewise, corresponding to the three proper (or quasi-perfect) equilibria of the subtree, there are three sequential equilibria. In particular, Figure 4.4 is again a difference tree under sequential equilibrium. The distinction is that there is a sequential equilibrium of this third difference tree in which Ann plays *Left*. (The details are the same as for the sequential equilibrium of the original tree.) So, this time, D is satisfied (as required by Theorem 4.1).

Under properness or quasi-perfection, the situation is different. The strategies *Left-left* and *Left-right* for Ann are admissible in the original game of Figure 4.1. In fact, *Left* is played in a proper (and, therefore, quasi-perfect) equilibrium. It is supported by a mixed strategy for Bob that puts ε-times less weight on *Right-In* vs. *Left-Out*. But, in the difference tree, *Out* is eliminated for Bob, and so there cannot be a mixed strategy for Bob that puts ε-times less weight on *Right-In* vs. $Left$-Out. As a result, $Left$ is weakly dominated in each of the difference trees of Figures 4.2-4.4 and so cannot be part of a proper (or quasi-perfect) equilibrium of these trees.

Ex post, the non-monotonicity in BI which we identify in this paper is, perhaps, not that surprising. At least, it may not be that surprising once one has a direct definition of BI, as we provide. Here is the essence of the argument:

- Start with a solution concept that satisfies BI. (In our case, this is sequential equilibrium.)
- Next consider a stronger solution concept. (In our case, this is proper equilibrium.)
- The stronger solution concept may prune more moves in forming a particular difference tree. (In our case, this is the move *Out* for Bob.)
- From elementary game theory, we know that when we prune a move for one player in a game, we can change previously good strategies for other players into bad strategies. (In our case, these are the strategies $Left$-$left$ and $Left$-*right* for Ann.)
- Suppose such a previously good strategy is played under the stronger solution concept on the overall tree. Then, this solution concept will fail D—and, therefore, BI. (In our case, Ann's playing $Left$ is indeed part of a proper equilibrium of the overall tree.)

The argument is actually very elementary. Propositions 4.1 and 4.2 serve simply to convert the in-principle argument into a specific instance of interest.

We note in passing that there is another (potential) source of non-monotonicity. Figure 4.4 was a difference tree for both the solution concept $\mathcal{S}$ and the refinement $\mathcal{R}$. However, a refinement might rule out a difference tree altogether. This, too, could lead to a failure of D.

5 Are Admissibility and Difference Consistent?

Now we return to our main question: Are admissibility and BI consistent? Recall, a strategy $\sigma_i \in \mathcal{M}(S_i)$ is **admissible** if there is no strategy $\rho_i \in \mathcal{M}(S_i)$ which weakly dominates it. We define:

(A) *A solution concept satisfies **Admissibility** if it contains only admissible strategies.*

A standard argument gives that A implies R. Thus, the question of whether A and BI are consistent amounts to: Does there exist a solution concept that satisfies D, A, and E?

The answer depends on the family of extensive-form games to which we apply the solution concept. Let us begin with two extreme cases:

Q1 Does there exist a solution concept that satisfies D, A, and E on the domain of all finite (perfect-recall) games?

Q2 Does there exist a solution concept that satisfies D, A, and E on the domain of all "generic" finite (perfect-recall) games?

Answers to these questions can already be found in the literature. Let us review.

The answer to Q1 is no. This can already be seen from Figure 5 in Kohlberg and Mertens (9, 1986, p.1013). Figure 5.4 provides a modification of the figure. (The modification is to allow us to talk about outcomes rather than strategies.) Consider the subtree following Ann's play of $Left$. By E and A, the solution on the subtree requires Bob to play $Left$. Now, refer to the difference tree in Figure 5.5 and note that, by E and A, Ann must play $Left$ in this tree. So, by E and D, Ann must play $Left$ in the original tree. This yields the $(2, 2)$ outcome. But a similar argument applies to the subgame following Ann's play of *Right*. This yields the $(2, 3)$ outcome—a contradiction.

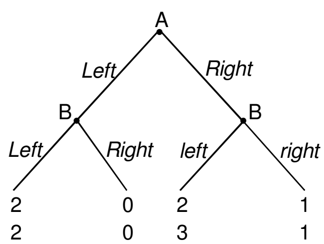

Figure 5.4

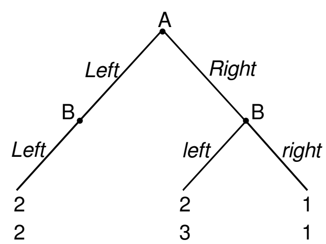

Figure 5.5

The answer to Q2 is yes. This uses a recent result by Pimienta and Shen (17, 2010). Fix an extensive form, written G, and let $\mathrm{PF}(G)$ be the set of payoff functions for the extensive form G. Say $X \subseteq \mathrm{PF}(G)$ is **generic relative to** G if its complement is a lower-dimensional semi-algebraic set. (See Blume and Zame (3, 1994) and the references there.) Theorem 1 in Pimienta and Shen (17, 2010) shows that, for a given extensive form G, the set of associated payoffs so that sequential equilibrium and quasi-perfect equilibrium coincide is generic relative to G. Recall, sequential equilibrium satisfies D and E (Proposition 4.2) and quasi-perfect equilibrium satisfies A. So, this implies that, for a given extensive form G, the set of payoffs so that the associated games satisfy D, A, and E is generic relative to G.

But, arguably, both Q1 and Q2 are too extreme. Q1 asks the question on a domain that includes certain 'pathological' extensive-form games, such as those in Figures 5.4-5.5. Indeed, we can already see the non-monotonicity of D by looking at these games.[3] (There is a sequential and proper equilibrium of Figure 5.4 where Ann plays *Right* and Bob plays $Left$-$left$ with probability one. This induces a sequential equilibrium of the game in Figure 5.5, but not a proper equilibrium of that game.) The reason for the non-monotonicity is that Ann is indifferent between two terminal nodes over which she is decisive, despite the fact that Bob is not. Put differently, these two terminal nodes correspond to distinct outcomes, despite the fact that one player is indifferent between them.

[3]We thank Priscilla Man and Andy McLennan for this point.

On the other hand, Q2 asks the question on a domain that excludes many 'non-pathological' extensive-form games—in fact, that excludes many extensive-form games of applied interest. In particular, the question is asked on a strict subdomain of the domain of trees where any two terminal nodes must correspond to different outcomes. Many games of interest fail this property—e.g., voting games, auctions, Bertrand competition, chess, etc. (See the discussions in Mertens (14, 1989) and Marx and Swinkels (13, 1997).) In the particular case of voting and auction games, both BI and A are often important aspects of the analysis. (Examples include Besley and Coate (2, 1997), Caillaud and Mezzetti (4, 2004), Gerardi and Yariv (5, 2007), and Hörner and Jamison (8, 2008), among many others.)

In sum, Q1 does not rule out games where ties are pathological, but Q2 rules out games where ties arise naturally from the application. This leads us to the question:

Q3 Does there exist a solution concept that satisfies D, A, and E on the domain of all finite (perfect-recall) games satisfying SPC?

We do not know the answer and leave the question as open.

6 Conclusion

We have taken as given that A and BI are desirable properties for a solution concept to satisfy. This is the classical view. Of course, one may question this view and ask if either A or BI is a desirable property. In this paper, we refrain from this question and, instead, show that there are puzzles even in the classical world.

To reach this conclusion, we formulate the BI property as E, R, and D. Property D formalizes the basic idea that the solution on the whole tree should be pinned down by the solution on its parts. Proposition 3.1 showed that E, R, and D characterize BI on PI games satisfying SPC. Of course, we cannot rule out the possibility that some another formalization of the basic idea might deliver an analog to our Proposition 3.1.

In the existing literature, direct definitions of BI appear to fall into one of two categories: those that fail to characterize BI (i.e., do not satisfy both directions of Proposition 3.1) and those that do characterize BI. We reviewed definitions—Reverse Difference and Projection—that fall into the first category, in Section 3. The Appendix discusses definitions that fall into the second category. Those definitions can be viewed as (perhaps, minor) modifications of D. The key is that the main messages of our paper hold with those definitions, too. In particular, sequential equilibrium satisfies those definitions, but proper equilibrium does not. So, those definitions are also non-monotonic. The implications for the question of the consistency of A and BI are the same: The answer to Q1 is no, the answer to Q2 is yes, and the answer to Q3 appears to be open.

One might ask for additional requirements on BI—above and beyond E, R, and D. For example, we could require that, to satisfy BI, a solution concept should satisfy E, R, D, and also Projection.

(We looked for a minimal set of properties that deliver Proposition 3.1. For this purpose, we can drop Projection—but not D—from this list.) Again, the main messages of our paper hold with such additions, provided that sequential equilibrium satisfies the additional requirement(s). Take Projection. In the proof of Proposition 4.2, we pointed out that sequential equilibrium satisfies Projection. So, sequential equilibrium satisfies E, R, D, and Projection, but proper does not. The non-monotonicity still arises. The implications for the question of the consistency of A and BI would be the same: The answer to Q1 is no, the answer to Q2 is yes, and the answer to Q3 appears to be open.

References

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

[2] Besley, T., and S. Coate, "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 112, 1997, 85-114.

[3] Blume, L., and W. Zame, "The Algebraic Geometry of Perfect and Sequential Equilibrium," *Econometrica*, 62, 1994, 783-794.

[4] Caillaud, B., and C. Mezzetti, "Equilibrium Reserve Prices in Sequential Ascending Auctions," *Journal of Economic Theory*, 117, 2004, 78-95.

[5] Gerardi, D., and L. Yariv, "Deliberative Voting," *Journal of Economic Theory*, 134, 2007, 317-338.

[6] Govindan, S., and R. Wilson, "Refinements of Nash Equilibrium," in Durlauf, S., and L. Blume (eds.), *The New Palgrave Dictionary of Economics*, 2nd edition, 2008, at http://www.dictionaryofeconomics.com.

[7] Hillas, J., and E. Kohlberg, "Conceptual Foundations of Strategic Equilibrium," in R. Aumann and S. Hart (eds.), *Handbook of Game Theory*, Vol. III, Elsevier, 2002, 1597-1663.

[8] Hörner, J., and J. Jamison, "Sequential Common-Value Auctions with Asymmetrically Informed Bidders," *Review of Economic Studies*, 75, 2008, 475–498.

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

[10] Kreps, D., and R. Wilson, "Sequential Equilibria," *Econometrica*, 50, 1982, 863-894.

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

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

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

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

[15] Myerson, R., "Refinement of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, 1978, 73-80.

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

[17] Pimienta, C., and J. Shen, "On the Equivalence between (Quasi-)Perfect and Sequential Equilibria," 2010, at http://research.economics.unsw.edu.au/cpimienta.

[18] Reny, P., "Backward Induction, Normal Form Perfection and Explicable Equilibria," *Econometrica*, 60, 1992, 627-649.

[19] van Damme, E., "A Relation Between Perfect Equilibria in Extensive Form Games and Proper Equilibria in Normal Form Games," *International Journal of Game Theory*, 13, 1984, 1-13.

[20] van Damme, E., *Stability and Perfection of Nash Equilibria*, Springer-Verlag, 1987.

Appendix

In Section 3, we argued that, taken together, E, R, and D capture BI. We showed that some other proposals do not capture BI. We now return to investigate two alternative formulations of BI, both suggested by property BI3 in Kohlberg and Mertens (9, 1986, p.1012).

a. Strategies vs. Outcomes We formulated our property D in terms of outcomes not strategies. In accordance with this, the statement of Proposition 3.1 also involves outcomes: It says that E, R, and D give the BI outcome—not the BI strategies.

First note, we cannot improve Proposition 3.1 so that it delivers the BI strategies, even if we restrict attention to the family of PI trees satisfying No Relevant Ties (Battigalli (1, 1997, p.48)). (This is a subfamily of the PI trees satisfying SPC.) The solution concept of extensive-form rationalizability (EFR, due to Pearce (16, 1984)) is outcome equivalent to BI on PI trees satisfying NRT (1, 1997, Theorem 4, p.53). So, by Proposition 3.1(ii), extensive-form rationalizability satisfies E, R, and D on this family of trees. But, it need not yield the BI strategies on such trees. See Figure 3 in Reny (18, 1992, p.637) for an example.

In light of this, perhaps we should restate D, so that it is a requirement on strategies and not outcomes. Specifically:

(SD) *A solution concept $\mathcal{S}$ satisfies **Strategy-wise Difference** (on $\mathcal{G}$) if for each tree Γ (in $\mathcal{G}$) and each subtree Δ of Γ the following holds. Let $Q \in \mathcal{S}(\Gamma)$. Then there exists a $Q^{\Delta} \in \mathcal{S}(\Delta)$ and a $P^{\mathcal{S}} \in \mathcal{S}(\Gamma_{\mathcal{S},Q^{\Delta}})$ such that for each $\sigma \in Q$, the restriction of σ to $\Gamma_{\mathcal{S},Q^{\Delta}}$ is contained in $P^{\mathcal{S}}$.*

The proof of Proposition 4.2 shows that sequential equilibrium satisfies SD. Moreover, since quasi-perfect and proper equilibrium fail D, they also fail SD. Therefore, our main results still hold if we replace D with SD.

What if we take the definition of BI to be E, R, and SD? One might think that, in the proof of Proposition 3.1, we can replace D line-by-line with the stronger requirement of SD and reach a stronger conclusion, viz., that we get BI strategy-wise and not just outcome-wise. But this is false.

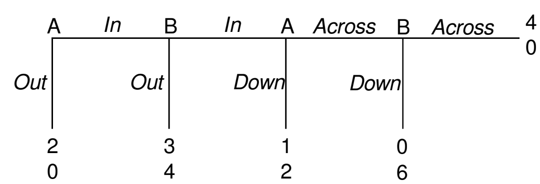

Figure A.1

Consider the game in Figure A.1. Here the BI strategies are $(In\text{-}Down, Out\text{-}Down)$. The proof of Proposition 4.2(i) requires the following analysis: Consider the subtree in Figure A.2. Per the new induction hypothesis, suppose that the solution on this subtree gives the BI strategies. Now consider the associated difference tree in Figure A.3. By E and R, Ann must choose In. From this, E and SD say that, in the original tree, Ann must choose some strategy and this strategy must be consistent with In. But this strategy need not be $In\text{-}Down$; it could be $In\text{-}Across$. Certainly, then, if replace D with SD, our proof will not yield the stronger conclusion. We conjecture that a solution concept can satisfy E, R, and SD, even though it fails to give the BI strategies. (Of course, it must give the BI outcome.)

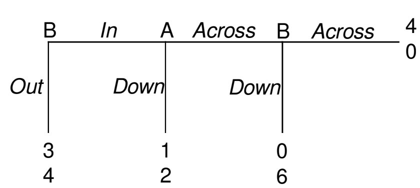

Figure A.2

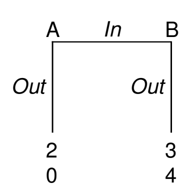

Figure A.3

b. Expected Payoffs vs. Pruning Fix a solution concept $\mathcal{S}$ where each nonempty component is a singleton. In this case, we could formulate D in terms of expected payoffs rather than outcomes:

Given a component $Q \in \mathcal{S}(\Gamma)$, we could ask that there is a nonempty (singleton) component $Q^{\Delta} \in \mathcal{S}(\Delta)$ such that, when we replace Δ with a terminal node whose payoffs are the expected payoffs under Q^{Δ}, there is a component $\widetilde{Q}$ of the solution on the new tree that induces the same outcomes as Q. We can mimic the proofs of Propositions 4.1 and 4.2 to show that sequential equilibrium will satisfy this expected-payoff version of D, but proper (and quasi-perfect) equilibrium will not.

This said, there is no clear way to extend this version of D to the common cases of solution concepts with multi-valued components. Yet, there are prominent examples of solution concepts which have multi-valued components and are outcome equivalent to BI on perfect-information games satisfying a no-ties condition. Two such examples are extensive-form rationalizability (EFR) and iterated admissibility (IA). (By IA, we mean simultaneous maximal deletion.) EFR and IA are outcome equivalent to BI on perfect-information games satisfying NRT (1, 1997, Theorem 4, p.53).

We already noted that EFR satisfies E, R, and D on the family of PI trees satisfying NRT. We do not know if it satisfies D on all SPC trees. IA does not satisfy D on SPC trees. To see this, consider again the game of Figure 4.1. It is easily checked that the IA set allows the outcome $(1, 1)$. Next, consider again the subtree beginning at the node where Bob can choose *Out*. Calculating the IA set here leads to the difference tree in Figure 4.4. But, in this tree, Ann's strategy $Left$ is weakly dominated—so, the outcome $(1, 1)$ is inconsistent with IA in this tree, contradicting D.