What is an Axiom for Backward Induction?*

Adam Brandenburger† Amanda Friedenberg‡

September 2009

Abstract

Which solution concepts satisfy backward induction (BI)? We define a property–we call it Difference–which relates the behavior of a solution concept on a whole tree to its behavior on any truncation of the tree–obtained by pruning from a subtree all moves disallowed by the solution concept in question. Difference (together with some background properties) characterizes the BI algorithm in perfect-information (PI) trees. We propose it as the definition of BI in general (non-PI) trees as well. Our main finding is a *non-monotonicity* in BI: A solution concept $\mathcal{S}$ can satisfy BI while another solution concept $\mathcal{R}$, though a refinement of $\mathcal{S}$, may not. We argue that this has an important implication for the program of refining Nash equilibrium.

1 Introduction

The idea of backward induction (BI) has a long history in game theory, going back to von Neumann and Morgenstern (18, 1944). (See Schwalbe and Walker (15, 2001).) It is a staple of game-theoretic applications, and a standard criterion which solution concepts are expected to satisfy. Yet, we will suggest that even today there are some surprises and puzzles concerning BI.

To start, we need a definition of BI. Even this has not been firmly established. Kohlberg and Mertens (5, 1986, p.1006) wrote:

> In games of perfect information, the meaning of this requirement is clear (Zermelo (19, 1912)). But in games of imperfect information the meaning is ambiguous as best.

By this they mean that, in perfect-information (PI) game trees, the idea of BI is clearly implemented by the BI algorithm. But, the idea of BI should apply beyond these trees—so, there is the question

*We thank Hari Govindan and Bill Zame for very helpful conversations on the topic of this paper. Selvin Akkus, Heski Bar-Isaac, Pierpaolo Battigalli, Willemien Kets, John Nachbar, Alex Peysakhovich, Ariel Ropek, Marciano Siniscalchi, Harborne Stuart, and seminar participants at UCLA provided valuable input. Support from the Stern School of Business, the W.P. Carey School of Business, and the Departments of Economics at UCLA is gratefully acknowledged. wibi-11-10-09

†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

of how it applies, absent being able to refer directly to the algorithm. A number of proposals have been put forth in the literature. However, we will argue that these fail to capture BI at either the formal or the intuitive level.

Here, we propose what we call the Difference Property as the definition of BI. We describe this property in Section 3. For now, simply think of it as a way of building up the solution of a tree from the solution of each of its subtrees—just as the BI algorithm does in PI trees. We will argue that Difference does capture the general idea of BI at both the formal and the intuitive levels.

What do we learn from offering a general definition of BI? We show the following:

> **Main Result:** *There exists solution concepts $\mathcal{S}$ and $\mathcal{R}$, such that $\mathcal{S}$ satisfies BI and $\mathcal{R}$ fails BI, even though $\mathcal{R}$ is a refinement of $\mathcal{S}$.*

This says that there is a basic non-monotonicity in whether or not a solution concept satisfies BI. A solution concept may fail BI, even if it is a refinement of a solution concept which satisfies BI. Moreover, we show that this is the case for "standard" solution concepts—in the proof, we take $\mathcal{S}$ to be sequential equilibrium and $\mathcal{R}$ to be proper equilibrium. (In particular, then, both solution concepts are nonempty.)

We believe this finding has an important implication for the program of refining Nash equilibrium. Since Kohlberg and Mertens (5, 1986), the program has proceeded by searching for a solution concept which satisfies one or other list of desirable axioms—e.g., invariance, admissibility, small worlds, etc. (See (5, 1986, p.1020) for the original list of axioms. There have been many subsequent lists.) Indeed, the Holy Grail for this program would be an agreed-upon list of axioms which succeeds in identifying precisely one solution concept. Refinement theorists would consider such a solution concept to be the 'ultimate' meaning of rationality in the game context.

In practice, the refinements program has adopted a less than purely axiomatic approach. Some of the axioms which have been proposed are truly axioms–i.e., properties demanded of a solution concept. (Our Difference property is of this type.) But, other axioms are not truly axioms. While they talk about how a solution concept should behave, they do so in terms of some other solution concepts, and are therefore circular in nature.

The leading such 'quasi-axiom' is about BI. In some papers (e.g., Kohlberg and Mertens (5, 1986), Hillas (3, 1990), Govindan and Wilson (2, 2009)), the BI requirement is that a solution concept contain a sequential equilibrium. Why does this circularity matter? Take as given that sequential equilibrium satisfies BI. (Indeed, we say it does, since it satisfies Difference.) But, now add other axioms from some list, just as the literature does. This is effectively asking that a good solution concept contain a refinement of sequential equilibrium, where the refinement is defined implicitly by these other axioms. Because of our non-monotonicity result, there is no guarantee that such a refinement will satisfy BI. That is, in the presence of other axioms, the BI quasi-axiom may not capture the meaning of BI. (Of course, this is if our definition of BI is granted.)

Some papers state a BI quasi-axiom which takes the form of requiring that a solution concept contain a proper equilibrium (e.g., Mertens (9, 1989), (10, 1992)). In this case, the issue is even more immediate: Proper equilibrium fails Difference and therefore, under our definition, does not satisfy BI even in its own right.

Our conclusion is that BI must be stated as a non-circular axiom. Then, to determine whether or not a certain solution concept satisfies BI, one must go back to 'first principles' and give a direct proof that it satisfies the axiom, or give a counterexample. Of course, we think the appropriate BI axiom is Difference.

Kohlberg and Mertens (5, 1986, p.1036) themselves expressed reservations about the use of what we are calling quasi-axioms:

> Our feeling, however, is that the source of the difficulty is in the use of a concept like sequential equilibrium. While sequentiality, invariance, dummy properties, etc., are reasonable properties against which a proposed solution concept may be checked, they cannot serve as a definition or an axiom ....

We go further. We believe our paper shows that a full-fledged axiomatic approach to refinements is not only desirable, but essential.

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 $\mu \in \mathcal{M}(X)$.

First, the formalities of a game tree: We consider finite extensive-form games of perfect recall (Kuhn (7, 1950), (8, 1953)) with the exception that we allow a non-terminal node to have only one outgoing branch (rather than two). We denote a typical such game by Γ, and let N be the set of non-terminal nodes and Z be the set of terminal nodes. 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.

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 $z, z' \in Z$ are **outcome equivalent** if $\Pi(z) = \Pi(z')$. (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.

A strategy profile $\sigma \in \Sigma$ induces a distribution over outcomes, i.e., 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 σ' **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 σ' are strategy profiles in two (possibly different) I-player games. Likewise, given subsets of strategy profiles $Q \subseteq \Sigma$ and $Q' \subseteq \Sigma$ (of two, possibly different, I-player games), say that Q **induces the same outcomes as** Q' if, for each $\sigma' \in Q'$, there is some $\sigma \in Q$ such that σ and σ' are outcome equivalent. Call Q and Q' **outcome equivalent** if Q induces the same outcomes as Q', and 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, 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 (5, 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.

Say 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

What is backward induction? Here is the intuitive idea: 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 don't change our original analysis. Let us now formalize this idea.

First, we 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 (8, 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. Recall, the idea was that we don't change our original 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 $\bar{Q} \in \mathcal{S}(\Gamma_{\mathcal{S},Q^{\Delta}})$, such that $\bar{Q}$ induces the same outcomes as Q.*

Loosely, the Difference property 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 illustrates the definition.

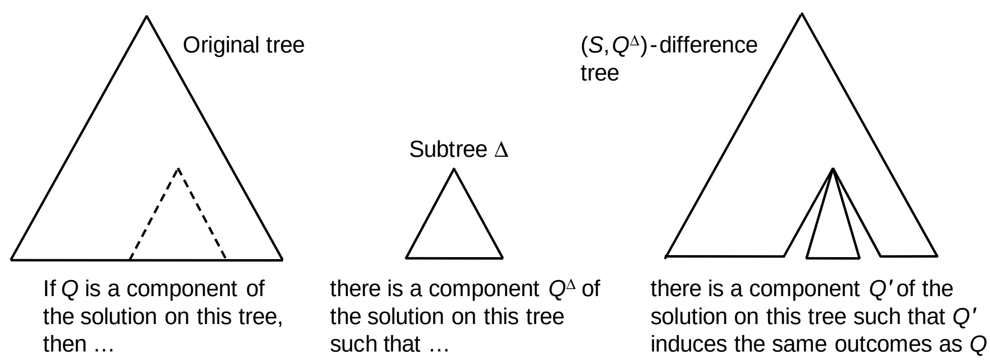

Figure 3.1

How does the definition of Difference capture the idea of BI? The way BI is usually explained is that future play in the game is used to pin down current play. The BI algorithm implements this idea in PI trees. Our Difference property works the same way–in general trees. Solutions on subtrees yield difference trees, which are used to pin down the solution on the overall tree. Difference formalizes this idea as: Each (distribution on) outcome(s) allowed by the solution on the overall tree must also be allowed by the solution of some difference tree.

Later in this section, we will come back to discuss a subtlety in the definition of Difference–viz., why we require that the solution on the whole tree be included in the solution on the difference tree, and not vice versa. In Sections 5b-c we will explain why some other possible definitions of BI do not work. But, first, we record the formal connection between Difference and BI. We show that Difference–plus some background properties–characterizes the BI algorithm in PI trees satisfying a no-ties requirement.

First, this requirement:

Definition 3.1 *A tree Γ satisfies the **Single Payoff Condition (SPC)** if, for all $z, z' \in Z$, the following holds: If i moves at the last common predecessor of z and z', then $\Pi_i(z) = \Pi_i(z')$ implies $\Pi(z) = \Pi(z')$.*

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 PI game satisfying SPC, there is a unique BI outcome. Moreover, SPC appears to be a minimal requirement for this purpose.

We will also have two background properties:

(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 $\mathcal{S}(\Gamma)$.*

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 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.*

We can now state the formal connection between Difference and BI.

Theorem 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.

Proof. 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). ■

Back to the definition of (D). Why not require instead the reverse inclusion–i.e., that the solution on a difference tree be contained in the solution on the whole tree?

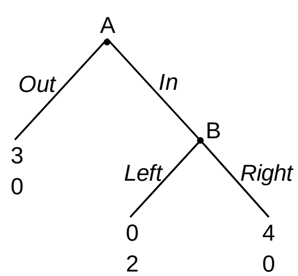

Figure 3.2

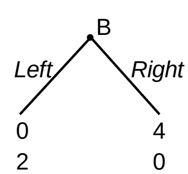

Figure 3.3

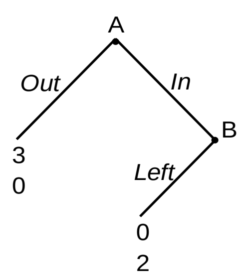

Figure 3.4

The reason is simple. Consider the tree Γ in Figure 3.2, the tree Δ in Figure 3.3 (which is the subtree of Γ that begins after Ann chooses In), and the tree $\bar{\Gamma}$ in Figure 3.4. Let $\mathcal{G} = \{\Gamma, \Delta, \bar{\Gamma}\}$. If $\mathcal{S}$ satisfies (E) and (R), then $\mathcal{S}(\Delta) = \{\{Left\}\}$. So, $\bar{\Gamma}$ is the $(\mathcal{S}, \{Left\})$-difference tree. Again, (E) and (R) imply that $\mathcal{S}(\bar{\Gamma}) = \{\{(Out, Left)\}\}$. By (D), Ann must then play Out in any $Q \in \mathcal{S}(\Gamma)$, yielding the BI outcome (as required by Theorem 3.1). Now change the definition of Difference to require instead the reverse inclusion. Consider a solution concept $\mathcal{R}$ on $\mathcal{G}$ given by $\mathcal{R}(\Delta) = \{\{Left\}\}$, $\mathcal{R}(\bar{\Gamma}) = \{\{(Out, Left)\}\}$, and $\mathcal{R}(\Gamma) = \{\{Out, In\} \times \{Left\}\}$. This satisfies (E), (R), and the reverse version of Difference on the domain $\mathcal{G}$. But, $\mathcal{R}$ is not outcome equivalent to BI on this family of trees.

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. Indeed, (D) uses solutions on difference trees to pin down the solution on the whole tree. If we change Difference to require instead the reverse inclusion, then we see that solutions on difference trees do not pin down the solution on the whole tree.

In Sections 5b-c , we will review some other proposed definitions of BI–these also fail to implement the idea that solutions on parts of the tree pin down the solution on the overall tree. We will also point out another subtlety in the definition of Difference.

4 Main Theorem

We state and prove our main result.

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.

Recall some definitions from Kreps and Wilson (6, 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. $\bar{H}$) 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 $\eta : \bar{H} \to H$ with $\bar{h} \subseteq \eta(\bar{h})$. Write $\bar{M}_i[\bar{h}_i]$ for the moves available to i at $\bar{h}_i$ in the difference tree, and note that, for each $\bar{h}_i$, there is an injective mapping $\xi[\bar{h}_i] : \bar{M}_i[\bar{h}_i] \to M_i[\eta(\bar{h}_i)]$ so that $\bar{m}_i \subseteq \xi[\bar{h}_i](\bar{m}_i)$. If $\bar{s}_i$ is a pure strategy for i in the difference tree, we write $[\bar{s}_i]$ for the set of pure strategies for i in Γ which coincide with $\bar{s}_i$ in the difference tree.

Proposition 4.1 *The solution concept $\mathcal{S}_{SE}$ satisfies (D).*

Proof. Fix a tree Γ and some $\beta = (\beta_1, \ldots, \beta_I)$ with $\{\beta\} \in \mathcal{S}_{\mathrm{SE}}(\Gamma)$. 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., $\{\beta^{\Delta}\} \in \mathcal{S}_{\mathrm{SE}}(\Delta)$.

Construct the difference tree $\Gamma_{\mathrm{SE},\{\beta^{\Delta}\}}$ 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 $\beta_i(\eta(\bar{h}_i))(\xi[\bar{h}_i](\bar{M}_i[\bar{h}_i])) = 1$. Moreover, if $\eta(\bar{h})$ is in Δ, $\eta(\bar{h})$ is reached with strictly positive probability under β^{Δ}. So, in this case, $\mu(\eta(\bar{h}))(\bar{h}) = \mu^{\Delta}(\eta(\bar{h}))(\bar{h}) = 1$. Indeed, this is true more generally, i.e., for each $\eta(\bar{h})$ (whether or not it is in Δ) $\mu(\eta(\bar{h}))(\bar{h}) = 1$. We use these facts repeatedly below.

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

It is immediate from the construction that each $\bar{\beta}_i(\bar{h}_i)$ is a best reply under $\bar{\mu}$. So, it suffices to show that $(\bar{\beta}, \bar{\mu})$ 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, $\beta_i^k(\eta(\bar{h}_i))(\xi[\bar{h}_i](\bar{M}_i[\bar{h}_i])) > 0$ and $\mu^k(\eta(\bar{h}_i))(\bar{h}_i) > 0$ for all $\bar{h}_i$. Define $(\bar{\beta}^k, \bar{\mu}^k)$ as follows: For each $\bar{h}_i$ and each $\bar{m}_i \in \bar{M}_i\left[\bar{h}_i\right]$, set $\bar{\beta}_i^k(\bar{h}_i)(\bar{m}_i) = \beta_i^k(\eta(\bar{h}_i))(\xi[\bar{h}_i](\bar{m}_i)|\xi[\bar{h}_i](\bar{M}_i[\bar{h}_i]))$. Likewise, for each $\bar{h}_i$ and each $n \in \bar{h}_i$, set $\bar{\mu}^k(\bar{h}_i)(n) = \mu^k(\eta(\bar{h}_i))(n|\bar{h}_i)$. Note, by construction $\bar{\beta}^k$ is completely mixed and $\bar{\mu}^k$ is derived from $\bar{\beta}^k$ by conditioning. Moreover, using the fact that each $\beta_i^k(\eta(\bar{h}))(\xi[\bar{h}_i](\bar{M}_i[\bar{h}_i])) \to 1, \mu^k(\eta(\bar{h}_i))(\bar{h}_i) \to 1$, it follows that $(\bar{\beta}_i^k, \bar{\mu}_i^k) \to (\bar{\beta}_i, \bar{\mu}_i)$ as required. ■

Next, recall the following definitions from Myerson (11, 1978). A profile of completely mixed strategies $\sigma^{\varepsilon} = (\sigma_1^{\varepsilon}, \ldots, \sigma_I^{\varepsilon})$ is an **ε-proper equilibrium** of Γ if, whenever $\pi_i\left(s_i, \sigma_{-i}^{\varepsilon}\right) < \pi_i\left(r_i, \sigma_{-i}^{\varepsilon}\right)$, then $\sigma_i^{\varepsilon}(s_i) \leq \varepsilon \sigma_i^{\varepsilon}(r_i)$. 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

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

Proposition 4.2 *The solution concept $\mathcal{S}_{PE}$ fails (D).*

Proof. Consider the game Γ given in Figure 4.1. 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-left, Left-right*$\}$, $\frac{2}{3}\varepsilon :$ *Right-left*, $\frac{1}{3}\varepsilon :$ *Right-right*) and Bob uses (unnormalized) weights (ε : { *Left-Out-left, Left-Out-right*}, ε^3 : { *Right-Out-left, Right-Out-right*}, $\frac{3}{5}$: *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.

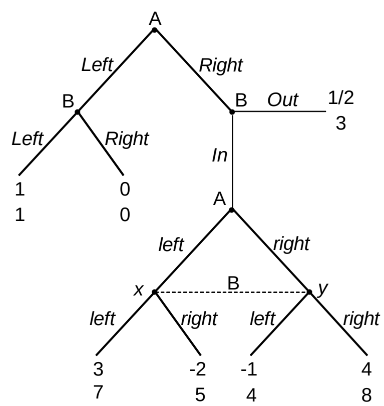

Figure 4.1

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-left, Out-right*}, *In-left, In-right*)), there are three proper 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-left* and *Out-right* are strongly dominated in the subtree, and so can't be part of a proper equilibrium.)

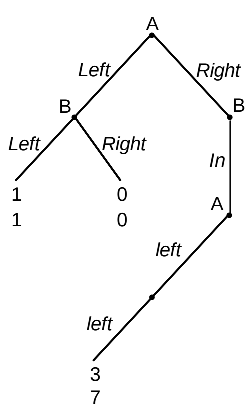

Figure 4.2

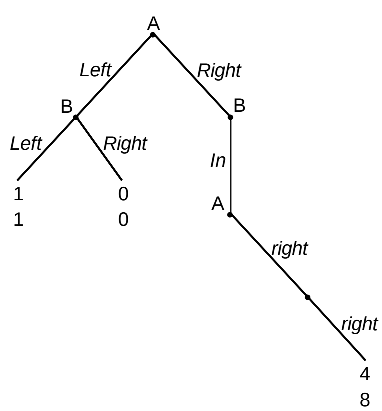

Figure 4.3

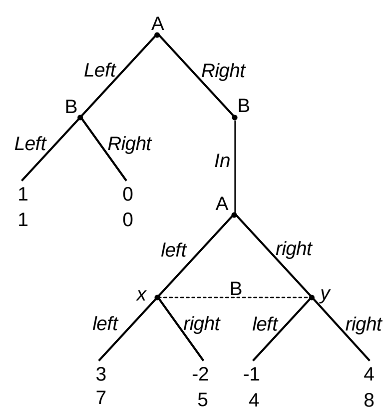

Figure 4.4

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

Now we can prove Theorem 4.1.

Proof of Theorem 4.1. By Propositions 1 and 3 in Kreps and Wilson (6, 1982), $\mathcal{S}_{\text{SE}}$ satisfies (E) and (R). By our Proposition 4.1, $\mathcal{S}_{\text{SE}}$ also satisfies (D). By Myerson (11, 1978, p.79), $\mathcal{S}_{\text{PE}}$ 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 $\sigma^{\varepsilon}_{-i}$. (See, e.g., Lemma 2.3.2 in van Damme (17, 1991).) Since $\sigma^{\varepsilon}_{-i}$ has full support, it follows by a standard argument that σ_i is then (extensive-form) rational. So, $\mathcal{S}_{\text{PE}}$ satisfies (R). By our Proposition 4.2, $\mathcal{S}_{\text{PE}}$ fails (D). Finally, by Theorem 1 in van Damme (16, 1984), $\mathcal{S}_{\text{PE}}$ is a refinement of $\mathcal{S}_{\text{SE}}$. ■

It is instructive to compare the behavior of sequential equilibrium with that of proper equilibrium in the game of Figure 4.1. Much as with proper 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 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 the Difference property is satisfied (as required by Theorem 4.1).

Under properness, the situation is different. The strategies *Left-left* and *Left-right* for Ann are undominated in the original game of Figure 4.1. In fact, *Left* is played in a proper equilibrium. It is supported by a belief that assigns ε less weight to the event that Bob plays *Right-In* vs. *Left-Out*. But, in the difference tree *Out* is eliminated for Bob and so Ann cannot consider the event the Bob plays *Right-In* "ε less likely" than the event that Bob plays *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 equilibrium of these trees.

Let us review the proof. Begin with proper equilibrium. In the original tree, there is a proper equilibrium where Ann plays *Left*. This is supported by a belief that assigns ε less weight to the event {(Right-In-left),(Right-In-right) vs. the event {(Left-Out-left),(Left-Out-right). More colloquially, *Left* for Ann is supported by a belief that assings infinitely more weight to the event that Bob plays *Left-Out* vs. the event that Bob plays *Right-In*. But, there can be no such belief in the difference tree, since *Out* is dominated in the subtree Δ.

prob(Out,L)=$\varepsilon \geq \varepsilon^2 = \Pr\left(\{l - R, r - R\}\right)$

, in Δ,

It is instructive to compare the behavior of sequential equilibrium with that of proper equilibrium in the game of Figure 4.1. Much as with proper 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 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 the Difference property is satisfied (as required by Theorem 4.1).

Under properness, the situation is different. The strategies $Left$-$left$ and $Left$-$right$ for Ann are undominated in the original game of Figure 4.1–in fact, $Left$ is played in a proper equilibrium. But they are weakly dominated in the difference tree of Figure 4.4–and so, cannot be part of a proper equilibrium. This happens because, in the course of forming the difference tree, a move for Bob has been eliminated. After this elimination, a previously undominated strategy for Ann becomes dominated.

We can now see that, at least in 'hindsight,' the non-monotonicity in BI which we identify in this paper is not at all surprising. Here are the key steps:

- Start with a solution concept which satisfies BI. (In our example, this is sequential equilibrium.)
- Next consider a stronger solution concept. (In our example, this is proper equilibrium.)
- The stronger solution concept may prune more moves in forming a particular difference tree. (In our example, 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 example, 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 Difference–hence BI. (In our example, Ann's playing $Left$ is indeed part of a proper equilibrium of the overall tree.)

The point is actually a very elementary one. Of course, we need our Propositions 4.1 and 4.2 to convert the in-principle argument into a specific instance of interest.

We note in passing that there is another (potential) source of a failure of Difference. In our example, Figure 4.4 was a difference tree for both the solution concept $\mathcal{S}$ and the refinement $\mathcal{R}$. However, a refinement might also rule out a difference tree altogether–again leading to a failure of Difference.

5 Discussion

We have proposed a definition of BI, namely Difference, shown that it exhibits a basic non-monotonicity, and described the implication we think this has for the refinements program. In this section, we discuss some other possible definitions of BI, the question of whether some non-equilibrium solution concepts satisfy Difference, and some open issues.

a. History The idea of relating the solution on the whole tree to solutions on subtrees has a long history in game theory. As already mentioned, the idea of a difference tree goes back to Kuhn (8, 1953, p.204), who also showed that subgame perfect equilibrium satisfies a difference-like property (8, 1953, p.208). Kohlberg and Mertens (5, 1986, pp.1012-1013) proposed a difference-like property as one of several possible definitions of BI. They stated–but did not prove–that sequential equilibrium satisfies their difference property. See also Pimienta (13, 2009).

Next, we review some other attempts in the literature, and, continuing the discussion at the end of Section 3, point to an another subtlety in the definition of Difference.

b. Projection Several papers have put forward a Projection property as the definition of BI. (See, e.g., Kohlberg and Mertens (5, 1986, p.1012) and Hillas and Kohlberg (4, 2002, Section 10).) This is the property that "a solution of the game induces a solution in any subgame" (5, 1986, p.1012).

At first sight, Projection seems to fit the idea of BI. It looks like it uses the solutions on the parts of the tree to pin down the solution on the overall tree—this time, using the solutions on the subtrees rather than on the difference trees to do so. But, it turns out that solutions on subtrees may be insufficient for this purpose. As such, Projection may fail to deliver the BI outcome.

To see why, let us try to formalize the Projection property. Start with a game Γ and a subtree Δ. Given a pure strategy s_i in Γ, write s_i^Δ for the restriction of s_i to Δ–i.e., for the restriction of s_i to the information sets in Δ. Given a mixed strategy σ_i in Γ, define a mixed strategy σ_i^Δ in Δ by

$$\sigma_i^\Delta(s_i^\Delta) = \sum_{\{s_i : s_i^\Delta \text{ is the restriction of } s_i\}} \sigma_i(s_i).$$

Call σ_i^Δ the restriction of σ_i to Δ.

(P) *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 $(\sigma_1, \ldots, \sigma_I) \in Q$ which reaches Δ, the restriction of $(\sigma_1, \ldots, \sigma_I)$ to the subtree Δ is contained in Q^Δ.*

That is, (P) attempts to use the behavior on a reached subtree to pin down behavior on the overall tree.

However, a solution concept may satisfy (E), (R), and (P) on the family of PI trees satisfying SPC, yet fail to deliver the BI outcome in these trees. Indeed, consider the solution concept of extensive-form rational Nash equilibrium, which we denote $\mathcal{S}_{\text{RNE}}$. (Thus: $\{\sigma\} \in \mathcal{S}_{\text{RNE}}(\Gamma)$ if $\sigma = (\sigma_1, \ldots, \sigma_I)$ is a Nash equilibrium in extensive-form rational strategies for Γ.) It is readily verified that $\mathcal{S}_{\text{RNE}}$ satisfies (E), (R), and (P) (on all trees). But, clearly, $\mathcal{S}_{\text{RNE}}(\Gamma)$ may include a component which is not outcome equivalent to BI (on an SPC tree).

Perhaps, we should modify the definition of Projection, so that we can use all subtrees—not just reached subtrees—to pin down behavior in the overall tree.

(P′) *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 $(\sigma_1, \ldots, \sigma_I) \in Q$, the restriction of $(\sigma_1, \ldots, \sigma_I) \in Q$ to the subtree Δ is contained in Q^{Δ}.*

Do (E), (R), and (P′) give the BI outcome? Still not. Set $\mathcal{G} = \{\Gamma, \Delta\}$, where Γ is again the tree in Figure 3.2 and Δ is again the subtree which begins after Ann's choice of In. Define a solution concept $\mathcal{S}$ on $\mathcal{G}$ by $\mathcal{S}(\Gamma) = \{\{Out, In\} \times \{Left\}\}$ and $\mathcal{S}(\Delta) = \{\{Left\}\}$. Then, $\mathcal{S}$ satisfies (E), (R), and (P′) on $\mathcal{G}$, but fails to deliver the BI outcome on this family of trees.

The problem is that 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. With (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. We conclude that even (P′) does not give a method for using the solution on the parts to pin down the solution on the whole, while (D) does give such a method.

c. Definition of Difference Now, back to the definition of (D). We already explained (in Section 3) why we require the solution on the whole tree to be contained in the solution on a difference tree, and not vice versa. Here, we point out another subtlety.

We formulated our Difference property in terms of outcomes not strategies. In accordance with this, the statement of Theorem 3.1(i) also involves outcomes: It says that (E), (R), and (D) give the BI outcome–not the BI strategies.

We cannot improve Theorem 3.1(i) 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)). (This is a subfamily of the PI trees satisfying SPC.) The solution concept of extensive-form rationalizability (Pearce (12, 1984)) is outcome equivalent to BI on PI trees satisfying NRT (1, 1997, Theorem 4). So, using Theorem 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 (14, 1992) 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* $(s_1, \ldots, s_I) \in Q$*, the restriction of* $(s_1, \ldots, s_I)$ *to* $\Gamma_{\mathcal{S},Q^{\Delta}}$ *is contained in* $P^{\mathcal{S}}$.

One might think that, in the proof of Theorem 3.1(i), 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.

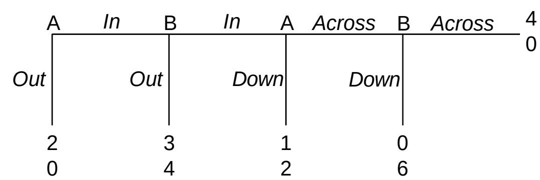

Figure 5.1

Consider the game in Figure 5.1. Here the BI strategies are $(In\text{-}Down, Out\text{-}Down)$. The proof of Theorem 3.1(i) requires the following analysis: Consider the subtree in Figure 5.2. Per the new induction hypothesis, suppose that the solution on this subtree gives the BI strategies. Now consider the associated difference tree in Figure 5.3. 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.)

One more variation. Fix a solution concept where each nonempty component is a singleton. In this case, we could formulate Difference 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 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 Difference, but proper equilibrium won't. As such, the message of this paper would be unchanged. This said, there is no clear way to extend this version of Difference to solution concepts with multi-valued components. Many solution concepts have multi-valued components. The non-equilibrium solution concepts discussed next are good examples.

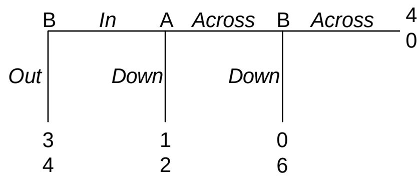

Figure 5.2

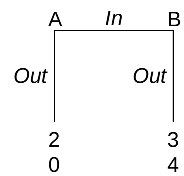

Figure 5.3

d. The Axiomatic Approach As we emphasized in the Introduction, our definition of BI is axiomatic—it is a requirement of a solution concept, but does not make reference to other solution concepts. So, in particular, it does not make reference to equilibrium concepts. Therefore, our approach allows us to ask whether non-equilibrium solution concepts satisfy BI. Two natural candidates are extensive-form rationalizability (EFR) and the iterated elimination of weakly dominated strategies (IA for "iterated admissibility").[1]

We already noted that EFR satisfies (E), (R), and (D) on the family of PI trees satisfying NRT. We don't 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).

e. Open Questions We have seen that (E), (R), and (D) are consistent. In particular, sequential equilibrium satisfies these axioms. But, what about the consistency of (E), (R), (D), and additional axioms? There are two obvious candidates to investigate: admissibility and forward induction.

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. Then:

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

It is a standard argument that (A) implies (R). But, (E), (D), and (A) are inconsistent–at least, if we require that a solution concept satisfy (E), (D), and (A) on all trees. Refer to Figure 5.4. (This is essentially Figure 5 in Kohlberg and Mertens (5, 1986), modified 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

[1] In keeping with the literature, we adopt the following conventions: EFR and IA each have one component, consisting of pure-strategy profiles. We allow a pure strategy to be dominated (conditionally dominated or weakly dominated) by a mixed strategy. We also take IA to be simultaneous maximal deletion.

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.

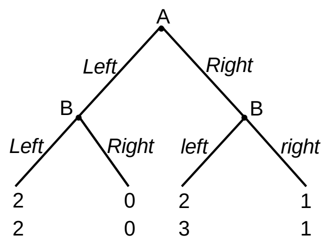

Figure 5.4

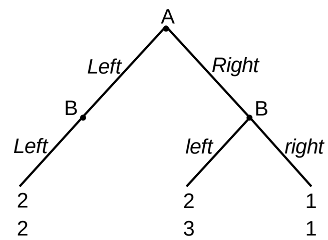

Figure 5.5

But, this is perhaps too harsh a test. The tree in Figure 5.4 doesn't satisfy SPC. This raises the question: Are (E), (D), and (A) consistent on the family of trees satisfying SPC? We don't know.

Now turn to forward induction (FI). There have been a number of proposals in the literature on how to define FI. See, e.g., Govindan and Wilson (2, 2009, Definition 3.5), whose proposal is stated in terms of what they call weak sequential equilibrium. In keeping with the main message here, we would argue that FI should be stated as a full-fledged axiom in its own right–i.e., as an axiom which does not make reference to other solution concepts. But, what is this axiom? What about the consistency of such an axiom with Difference–i.e., with BI? We leave this as an open issue.

References

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

[2] Govindan, H., and R. Wilson, "On Forward Induction," *Econometrica*, 77, 2009, 1-28.

[3] Hillas, J., "On the Definition of the Strategic Stability of Equilibria," *Econometrica*, 58, 1990, 1365-1390.

[4] 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.

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

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

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

[8] 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.

[9] Mertens, J-F., "Stable Equilibria: A Reformulation Part I. Definition and Basic Properties," *Mathematics of Operations Research*, 14, 1989, 575-625.

[10] Mertens, J-F., "The Small Worlds Axiom for Stable Equilibria," *Games and Economic Behavior*, 4, 1992, 553-564.

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

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

[13] Pimienta, C., "Backwards Induction and Sequential Equilibrium," 2009. Available at http://research.economics.unsw.edu.au/cpimienta.

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

[15] Schwalbe, U., and P. Walker, "Zermelo and the Early History of Game Theory," *Games and Economic Behavior,* 34, 2001, 123-137.

[16] 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.

[17] van Damme, E., *Stability and Perfection of Nash Equilibria*, Springer 1991.

[18] von Neumann, J., and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1944 (60th Anniversary Edition, 2004).

[19] Zermelo, E., "Uber eine Anwendung der Mengenlehre auf die Theorie des Schachspiels," 1912, published in *Proceedings of the Fifth International Congress of Mathematicians*, Vol. 2, 1913, Cambridge University Press, 501-514.