Backward Induction and Subgame Perfection. The justification of a "folk algorithm."

By Marek M. Kaminski#

Abstract: I introduce axiomatically infinite sequential games that extend von Neumann and Kuhn's classic axiomatic frameworks. Within this setup, I define a modified backward induction procedure that is applicable to all games. A strategy profile that survives backward pruning is called a backward induction equilibrium (BIE). The main result compares the sets of BIE and subgame perfect equilibria (SPE). Remarkably, and similarly to finite games of perfect information, BIE and SPE coincide both for pure strategies and for a large class of behavioral strategies. This result justifies the "folk algorithm" of finding SPEs in games with backward induction.

Keywords: game theory, subgame perfect equilibrium, backward induction, refinement, perfect information.

JEL classification: C73

* Acknowledgements: Helpful comments of Michael McBride, Drew Fudenberg, Marc Kilgour, Tom Schelling, Harrie de Swart, Alexei Zakharov and the participants of the IMBS seminar at the University of California, Irvine, are gratefully acknowledged. The usual disclaimers apply.

Department of Political Science and Institute for Mathematical Behavioral Sciences, University of California, Irvine. Email: marek.kaminski@uci.edu.

1 Introduction

The origins of backward induction are murky. Zermelo (1913, 2001) analyzed winning in chess, but his method of analysis was based on a different principle (Schwalbe and Walker 2001). Reasoning based on backward induction was implicit in the Stackelberg's (1934) construction of his alternative to Cournot equilibrium. As a general procedure for solving two-person zero-sum games of perfect information, backward induction appeared in the von Neumann and Morgenstern's founding book (1944: 117). It was used to prove a precursor of Kuhn's Theorem for chess and similar games. The von Neumann's exceedingly complex formulation was later clarified and elevated to the high theoretical status by Kuhn's work (1953, especially Corollary 1), Schelling's (1960) ideas of incredible threats and Selten's (1965) introduction of subgame perfection. It suffered drawbacks when the chain-store paradox, centipede and other games questioned its universal appeal (Selten 1978; Rosenthal 1981). Its profile was further lowered with new refinements. Perfect equilibrium (Selten 1975), sequential equilibrium (Kreps and Wilson 1982) or procedures such as forward induction (Kohlberg and Mertens 1986) offered solutions oftentimes better comforting our intuition. The arguments against backward induction began to multiply and sealed the doubts about its universal validity (Basu 1988, 1990; Bonanno 1988; Binmore 1987, 1988; Reny 1986; Fudenberg, Kreps and Levine 1988; Bicchieri 1989; Pettit and Sugden 1989).

The present paper does not intend to redeem backward induction. Its goal is to extend the procedure to as many games as possible and investigate its relation to subgame perfection. In its standard formulation, backward induction applies only to finite games of perfect information. There, every backward induction equilibrium (BIE), i.e., a strategy profile that survives backward pruning, is also a subgame perfect equilibrium (SPE), and all SPEs result from backward pruning. Yet, game theorists consider it common knowledge that other games can be solved backwards as well, and they routinely apply the procedure to such games. Backward reasoning is implicit in refining Stackelberg equilibrium from other Nash equilibria (NE). Schelling analyzed backward the NE in the iterated Prisoner's Dilemma as early as in 1950s (personal communication, 3/7/2008). Fine textbooks such as Fudenberg and Tirole (1991:72) and Myerson (1991:192) make explicit claims (but without proofs!) that backward induction can be applied to a wider class of games. Backward reasoning can be found in many extensive-form models (parametrized families of extensive-form games) employed by political economy. It is present in the popular argument that voters must "vote sincerely" in the last stage of a majority voting game with a binary agenda.

What results from this alleged abuse? Fudenberg and Tirole (1991: 94) declare "This is the logic of subgame perfection: Replace any "proper subgame" of the tree with one of its Nash-equilibrium payoffs, and perform backward induction on the reduced tree." With caveats, their prescription captures the essence of the algorithm that is presented below. However, if one wanted to find a formal justification of this algorithm in the literature, one couldn't. Such a justification is necessary since, in fact, the definitions of SPE and strategies that survive backward induction are based on different principles and defined differently.

Let's examine this *petite difference*. SPE is a strategy profile that is NE in all subgames. Its intuitive justification focuses on subgames, i.e., parts of a larger game that constitute smaller games. It demands that the players' interaction structure is "rational" in all subgames, i.e., that they result in a NE. In backward induction, a different set of games is considered. The procedure starts at the end of a game and moves backward according to an imagined timeline. Similarly to SPE, the first game (or a set of games) are subgames but at some point a new game appears that is not a subgame of the original game and that has no counterpart in the definition of a SPE. Such a game (and possibly other games

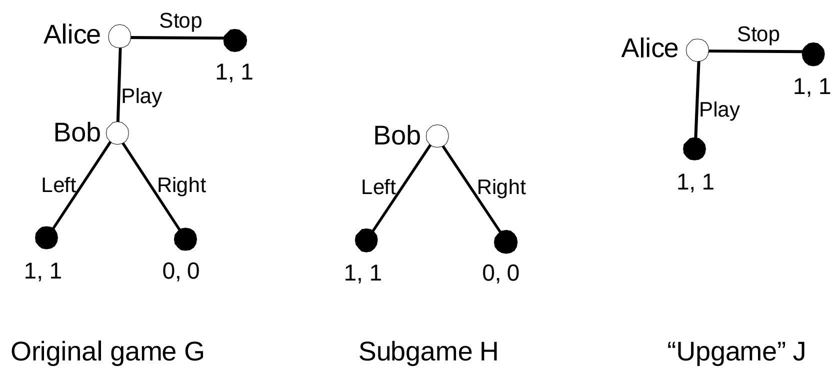

Figure 1:

that follow) is created by substituting a subgame (or subgames) with a NE payoff vector in those subgames.

To make the distinction between SPE and BIE crystal-clear, let's consider a simple game in Figure 1. SPE requires that Bob and Alice play NE strategies in both the original game G and its subgame H. BIE requires that the players play NE strategies in G's subgame H and also the "upgame" J. The upgame is not a subgame of G: it was created by pruning H and assigning the NE payoff to the root of H.

Backward induction seems as such an intuitive procedure for finding SPEs that one can easily ignore the subtle difference in definitions, especially when one begins to struggle with the painful chore of defining BIE formally. Nevertheless, one cannot say a priori that the sets of strategy profiles defined in both cases are identical. In fact, and quite surprisingly, BIE and SPE may differ for mixed strategies in a game of imperfect recall.

Despite the lack of firm foundation, it seems that backward induction acquired in the discipline the status of a "folk algorithm." Game theorists use it but nobody examined its validity. It clearly works – in the sense of producing SPEs – but we do not know exactly for what games it works and what exactly it produces. The formal link between the folk algorithm and subgame perfection is missing.

I investigate the “folk puzzle of backward induction” axiomatically. The axiomatic framework that I introduce below encompasses more games than those of von Neumann and Morgenstern (1944) and Kuhn (1953). Namely, action sets may have any cardinality and the length of a game may be infinite. Despite its slightly higher complexity, the new framework allows for constructive analysis.

It turns out that, after suitable modification, backward induction indeed applies to practically all interesting games and a large class of strategies. Nevertheless, the Fudenberg-Tirole’s prescription has to be modified in a number of subtle ways. First, backward pruning can be applied not only to pure strategies but also to a large class of behavioral and mixed strategies. Second, as it was already mentioned, there are games for which backward induction brings different solutions than subgame perfection. This may happen for mixed strategies in a game of imperfect recall. Third, we must replace every subgame with a SPE and not NE payoff vector. (Obviously, when a subgame has no other subgames than itself, SPE and NE coincide.) Finally, we can replace entire subsets of subgames simultaneously, not only single subgames. Such simultaneous replacement is essential for solving games that are not finite. Remarkably, and in agreement with the case of finite games of perfect information, the conclusion is that for pure strategies and a large class of behavioral strategies the sets of BIE and SPE coincide for all games. This is the main result. It legitimizes informal methods of backward pruning of a game and concatenating resulting partial strategy profiles used by game theorists to solve games .

The generalized algorithm uses the agenda (the tree consisting of the roots of all subgames) instead of the game tree. For games of perfect information, the agenda coincides to the game tree with terminal nodes subtracted. A step in such an algorithm can be informally compared to the classic backward induction as follows:

(1) Prune any subset of disjoint subgames instead of a single decision node whose all followers are terminal nodes;

(2) Substitute all selected subgames with the SPE payoffs instead of the payoffs for best moves;

(3) Concatenate all partial strategy profiles obtained in the previous step; if at any point you will get an empty set, there is no SPE in the game.

The procedure can be applied to pure strategies in all games, to a large family of behavioral strategies in all games, or a large family of mixed strategies in games of perfect recall. Finding all SPEs requires following certain rules of concatenating and discarding partial strategy profiles that are described in Section 5.

The next section introduces axiomatically sequential games with potentially large sets of actions and infinite numbers of moves. While a more accurate name would be "potentially infinite" games, I call such games "infinite" for simplicity of terminology. Then, basic facts linking payoffs to strategies in infinite games are established. Section 3 investigates the decomposition of games into subgames and upgames for pure and behavioral strategies with finite support and crossing. Section 4 explains informally that while mixed strategies cannot be decomposed in a similar fashion, backward induction can be modified to incorporate them as well. Section 5 includes the main result that is a corollary of the results from Section 3. The generalized backward induction is described formally and certain applications are discussed. Section 6 concludes with open questions and speculates on further lines of research.

2 Preliminaries

Sequential (extensive form) games were introduced with a set-theoretic axiomatization by von Neumann and Morgenstern (1944: 73-76). They were conceived and presented in the spirit of early 20th century rigorous decoupling of syntax and semantics, as embodied in the works of Hilbert, Tarski and later the Bourbaki team. In order to prevent a reader from forming any geometric or other intuition, von Neumann announced proudly that "We have even avoided giving names to the mathematical concepts [...] in order to establish no correlation with any meaning which the verbal associations of names may suggest" (1944: 74). Then, he dismissed his own idea of a game tree since "even relatively simple games lead to complicated and confusing diagrams, and so the usual advantages of graphical representation do not obtain." (1944: 77) Despite von Neumann's efforts to turn a sequential game into a highly abstract and incomprehensible for a non-mathematician object, Kuhn's (1953) formulation made games easier on our intuition. He simplified von Neumann's formalism and built the axioms into definitions and assumptions about the tree, players and information. Kuhn also generalized the von Neumann's unnecessarily narrow definition.

The axiomatic setup of this paper goes beyond finite games in an attempt to cover axiomatically most, if not all, games of interest to the discipline. Such games are called for simplicity of terminology *infinite games.* The framework attempts to maintain the compatibility with the pragmatic Kuhn's exposition, and also draws from fine modern presentations of Myerson (1991) and Selten (1975). The axioms are divided into two subsets. The axioms for an infinite tree are listed explicitly; the game axioms are combined with the description of game components and specify how various objects are attached to the tree. In order to establish some intuitive associations, the axioms received in parentheses their names that succinctly describe their content.

The opening paragraphs include a laborious re-establishment of basic and intuitive notions and results.

Rooted tree: Let (T, Υ, τ) be such that T a set of at least two points, Υ is a binary relation over T, and $\tau \in T$. For $y \in T, y \neq \tau$, a *path* to y of length k is any finite set $e_y = \{x_i\}_{i=1}^{k} \subset T$ such that $x_1 = \tau, x_k = y$ and for all $i = 1, ...k-1$

$(x_i, x_{i+1}) \in \Upsilon$. For $y = \tau$,the path to τ is $\{\tau\}$.

Υ is called a *rooted tree* Υ with its *root* τ and the set of *nodes* T if the following axioms AT1-AT4 are satisfied:

AT1 (*domain*) For every $x \in T$, there is $y \in T$ such that $(x, y) \in \Upsilon$;

AT2 (*partial anti-reflexivity*): For every $x \in T$, $(x, x) \in \Upsilon$ iff $x = \tau$;

AT3 (*symmetry*): For all $x, y \in T$ $(x, y) \in \Upsilon$ iff $(y, x) \in \Upsilon$;

AT4 (*unique path*): For every $x \in T$, there is exactly one *path* e_x to x.

AT1-AT3 play a technical role in conceptualizing the tree and allow to reconstruct from Υ the root and the set of nodes. AT4 introduces the defining property of a tree.

The following definitions and related notation are used hereafter (the definitions are slightly redundant in order to maintain the compatibility with other formalisms):

1. Binary relations between two different nodes x, y:

predecessor: $y \in PR(x)$ *precedes* x or *is in the path* to x iff $e_y \subset e_x$;

successor: $y \in SU(x)$ *follows* x iff $e_x \subset e_y$;

immediate predecessor: $y = IP(x)$ *immediately precedes* x iff $e_x - e_y = \{x\}$ (by AT1-4 that for every $x \neq \tau$ there is exactly one immediate predecessor);

immediate successor: $y \in IS(x)$ *immediately follows* x iff $e_y - e_x = \{y\}$;

immediate predecessor in $T_i \subset T$*:* For $x, y \in T_i$, y *immediately precedes* x in T_i iff y precedes x and $(e_x - e_y) \cap T_i = \{x\}$; we write $y = IP_i(x)$ (by AT1-4, for every $x \neq \tau$ there is at most one immediate predecessor);

2. Single nodes and subsets of nodes:

endnode: a node that is not followed by any other node;

the set of all endnodes: T_E;

decision node: a node that is not an endnode;

the set of all decision nodes: $T_D = T - T_E$;

branch: any node except for the root τ;[1]

alternative (originating) at a node x: any immediate successor of x;

terminal path: a path to an endnode or an infinite path;

3. Set of subsets of nodes:

T_t : the set of all terminal paths.

Game: An *n-player sequential game* is a septuple $G =< \Upsilon, N^0, \{T_i\}_{i \in N^0}, I, A, h, P >$ that includes a rooted game tree Υ and the following objects: players with their assigned decision nodes and probability distributions for random moves, the pattern of information, the identification of moves and the probability distributions over random (or pseudorandom) moves, and the payoff functions. The conditions imposed on the components of G and certain useful derived concepts are defined below.

1. *Game tree*: Υ is a rooted tree with the set of nodes T, set of decision nodes T_D, set of endnodes T_E, and the root τ;

2. *Players:* For a positive integer n, $N^0 = \{0, 1, ..., n\}$ consists of *players* $N = \{1, ..., n\}$ and a *random* or *pseudorandom mechanism* labeled with 0;

3. *Player partition:* $\{T_i\}_{i \in N^0}$ is a partition of T_D into (possibly empty) subsets T_i and a (possibly empty) subset T_0 for the random mechanism. The following assumptions are made about T_0:

(i) no path includes an infinite number of nodes from T_0;

(ii) for every $x \in T_0$, the number of alternatives at x is finite and equal at least two.

4. *Information:* $I = \cup_{i=0}^{n} I_i$ is such that every $I_i = \{I_i^k\}_{k \in K_i}$ is a refinement of i's set T_i. We assume that:

(i) all elements of I_0 are singletons;

[1] In a standard definition, a branch is any element $(x, y) \in \Upsilon$. For simplicity of forthcoming notation, it is identified here with the node in the pair that is farther from the root.

(ii) for all $i \in N$, every element of I_i includes only nodes with equal numbers of alternatives and does not include two nodes that are in the same path.

For every $i \in N^0$, a set $I_i^k \in I_i$ is called i's *information set.* An alternative y *originates* at an information set I_i^k if y originates at a node $x \in I_i^k$.

5. *Moves (actions)*: $A = \{A_i^k\}_{i \in N, k \in K_i}$ is a collection of partitions, one for every information set I_i^k of every player i, of all alternatives originating at I_i^k such that for any node $x \in I_i^k$, every member of A_i^k includes exactly one alternative that originates at x. Elements $a \in A_i^k$ are called the *moves (*or *actions)* of i at I_i^k. For any I_0^k, the moves at I_0^k are singletons including branches originating at I_0^k. Since by definition every branch y belongs to precisely one move, for $y \in T - \{\tau\}$ the move a s.t. $y \in a$ is denoted by $A_i^k(y)$;

6. *Random moves:* h is a function that assigns to every information set of the random mechanism $I_0^k = \{x\}$ a probability distribution $\{h_\theta^k\}$ over the alternatives at x with all probabilities being positive . If $T_0 = \varnothing$, h is not defined;

7. *Payoffs (associated with terminal paths):* The *payoff function* $P = (P_1, ..., P_n) : N \times T_t \to \mathbb{R}$ assigns to every terminal path $e \in T_t$ a *payoff vector* at e equal to $P(e) = (P_1(e), ..., P_n(e))$. The component $P_i(e)$ is called the *payoff* of player i at e. Function P_i is called the *payoff function* of player i.

The most extensively studied subset of games defined above are finite games.

Finite game: G is finite if the set of its nodes T is finite.

Subgame: For any game $G =< \Upsilon, N^0, \{T_i\}_{i \in N^0}, I, A, h, P >$, a subgame of G is any game $G' =< \Upsilon', N^0, \{T_i'\}_{i \in N^0}, I', A', h', P' >$ such that

(i) Υ' is a subtree of Υ, i.e., for some $\tau' \in T$, $T' = T \cap \{x \in T : x = \tau'$ or $x \in SU(\tau')\}$ and $\Upsilon' = (\Upsilon \cap [T' \times T']) \cup \{(\tau', \tau')\}$;

(ii) if $x_1, x_2 \in I_i^k$ for some I_i^k in G, then either $\{x_1, x_2\} \subset T'$ or $\{x_1, x_2\} \cap T' = \varnothing$;

(iii) $N' = N$ and $\{T_i'\}_{i \in N^0}, I', A', h', P'$ are restrictions of $\{T_i\}_{i \in N^0}, I, A, h, P$ to T', respectively.

The demonstration that restrictions in (iii) define a game is straightforward. It is also clear that the relation of "being a subgame" is transitive.

Infinite games include a vast majority of interesting sequential games analyzed in the literature. The assumed constraints demand that the numbers of players, random information sets at every path, and random moves at every random information set are finite. Both infinite paths and infinite numbers of moves at players' information sets are allowed.

A player i may be a dummy in a game, i.e., T_i may be empty. Such definition allows to treat subgames as games. Since $|T| \geq 2$, the root of Υ is a decision node and there must be at least one player or random mechanism in the game. Without loss of generality, one can safely assume that there are no dummies in the initial game G for which all results are formulated.

The concepts that follow are derived from the model's primitives. Strategies are defined in order to optimize the introduction of fundamental for this paper ideas of strategy concatenation and decomposition. The adjective "behavioral" is optional since behavioral strategies are our departure point for defining other types of strategies.

Behavioral actions: A *behavioral action of finite support* (in short, a *behavioral action*) α_i^k of player i at his information set I_i^k is a discrete probability distribution over A_i^k.

Strategies (rough behavioral): A *rough behavioral strategy* β_i of player i is any (possibly empty) set of i's behavioral actions that includes exactly one action per information set of i. A *partial rough behavioral strategy* ω_i is any subset of a rough strategy. A partial rough strategy that includes exactly those actions in β_i that are defined for information sets of i in a subgame H of G, is denoted as β_i^H and is called β_i reduced to H.

For any rough strategy β_i, let's denote the probability assigned by β_i at I_i^k to a move a_i by $\beta_i(I_i^k)(a_i)$. A path e is called *relevant* for β_i if β_i chooses every alternative in e that originates at some information set of i with a positive probability, i.e., if for every node $y \in e$ such that $y \in IS(x)$ for some $x \in T_i$, $\beta_i(I_i^k)(A_i^k(y)) > 0$. Finally, a path e *crosses* I_i^k if $e \cap I_i^k \neq \varnothing$.

Finite crossing in subgames: For every $i \in N$, every subgame H of G, a rough strategy β_i, and every path e^H in H that is relevant for β_i reduced to H, β_i^H, e^H crosses only a finite number of information sets from I_i such that $\beta_i(I_i^k)$ is non-degenerate.

Behavioral strategy: For any $i \in N$, a behavioral strategy of i, or simply a strategy of i, is any rough strategy of i that it satisfies *finite crossing in subgames.*

Comment: Finite support and finite crossing guarantee that, in all subgames, the payoffs (as they will be defined later) for behavioral strategies can be derived from a discrete probability distribution with finite support. The class of strategies that satisfy these conditions is pretty large since it includes all behavioral strategies in a finite game and all pure strategies in any game. Relaxing these conditions would introduce complications of measure-theoretic nature along lines examined by Aumann (1953). Certainly, a more general treatment of behavioral strategies would simply consider all rough strategies but whether the results would survive is at this moment unknown.

All behavioral strategies of i form i's *behavioral strategy space* B_i. Elements of $B = \times_{i=1}^n B_i$, *behavioral strategy profiles,* are denoted by $\boldsymbol{\beta}$.

I use a set-theoretic interpretation of strategies that will greatly simplify the definitions of strategy decomposition and concatenation, as well as the treatment of partial strategy profiles. There is a simple isomorphism between strategies or strategy profiles defined in a set theoretic and standard way. Thus, every strategy profile $\boldsymbol{\beta}$ is interpreted as a union of players' strategies (which are obviously disjoint); the Cartesian product $\times_{i=1}^n B_i$ is interpreted as taking all possible unions of individual strategies, one per player; the notation for a strategy profile $(\beta_i)_{i=1}^n$ represents an alternative notation for $\cup_{i=1}^n \beta_i$. An example of notational difficulty that is avoided is the interpretation of $\times_{i=1}^n B_i$ when at least one strategy set is empty. Another example is provided by the definition that comes next.

A strategy profile $\boldsymbol{\beta}$ with the strategy of player i removed, i.e., $\boldsymbol{\beta}-\beta_i$, is denoted by $\boldsymbol{\beta}_{-i} \in \times_{N-\{i\}} B_i$; $(\boldsymbol{\beta}_{-i}, \gamma_i)$ denotes $\boldsymbol{\beta}$ with β_i substituted with γ_i, i.e., $\boldsymbol{\beta}-\beta_i \cup \gamma_i$.

When such a distinction is necessary, the payoff functions, strategies, strategy profiles, etc. in games or subgames G and H will be given identifying superscripts P^G, P^H, etc.

The most important step towards building the framework for infinite games is expressing payoffs in terms of strategies.

Recall that the probability assigned by β_i at I_i^k to a move a_i was denoted by $\beta_i(I_i^k)(a_i)$. For any $x, y \in T$, such that $x \in IS(y)$, $y \in I_i^k$, and $x \in a_i$, the probability of *the move to* x is defined as $p_\beta^m(x) = \beta_i(I_i^k)(a_i)$. By convention, $p_\beta^m(\tau) = 1$ for all $\boldsymbol{\beta}$. A path e is *included* in $\boldsymbol{\beta}$ if $p_\beta^m(x) > 0$ for all $x \in e$ and is denoted $e \subset \boldsymbol{\beta}$. The set of all terminal paths included in $\boldsymbol{\beta}$ is denoted by T_β. The probability of *playing* e *under* $\boldsymbol{\beta}$, $p_\beta(e)$, is defined as:

$p_\beta(e) = \prod_{x \in e} p_\beta^m(x)$

Thus, $p_\beta(e)$ is the product of the probabilities assigned by $\boldsymbol{\beta}$ to all alternatives in e.[2]. The probability of *reaching* a node y, $p_\beta(y)$ is defined as the probability of playing e_y under $\boldsymbol{\beta}$:

$p_\beta(y) = p_\beta(e_y) = \prod_{x \in e_y} p_\beta^m(x)$.

The assumptions of finite support and finite crossing are used below to establish the fundamental fact that p_β defines a probability distribution over a finite subset of all terminal paths:

[2] By definition of the game and behavioral strategy, for a path of infinite length, only a finite number of alternatives may be assigned probabilities different than zero or one. The multiplication over an infinite series of numbers is assumed to be commutative and associative, with an infinite product that has at least one zero equal to zero and an infinite product of ones equal to one.

Lemma 1 *For every game G, every subgame H of G, and behavioral strategy profile $\boldsymbol{\beta}$:*

(a) The set of all terminal paths in H included in $\boldsymbol{\beta}^H$, T_{β^H}, is nonempty and finite;

(b) For every $e^H \in T_t^H$, $p_\beta^H(e^H) > 0$ iff $e^H \in T_{\beta^H}$;

(c) $\Sigma_{e \in T_{\beta^H}} p_{\beta^H}(e) = 1$.

Proof. Lemma 1 will be proved for game G. The assumptions of finite support and crossing apply to subgames as well as they apply to G, which allows to repeat all steps for every subgame of G.

Ad. (a): The construction of a terminal path $e \in T_\beta$ is by induction. Both when $\tau \in T_0$ and when $\tau \in T_i$ for $i = 1, ..., n$, there is a node $v \in SU(\tau)$ such that $p_\beta^m(v) > 0$. Let's choose v as the second (after τ) node of the path.

Let's assume now that a path e of length l reached a node x. If $x \in T_E$, e is the desired path. If $x \in T_D$, let I_i^k be such that $x \in I_i^k$. Both when $I_i^k \subset T_0$ and when $I_i^k \subset T_i$ for $i = 1, ..., n$, there is a node $y \in SU(x)$ such that $p_\beta^m(y) > 0$. Path $e_y \subset \boldsymbol{\beta}$ and its length is $l + 1$. The construction either ends at some endnode or proceeds indefinitely, producing some infinite path. In both cases, the resulting path e is terminal. By the definition of the game and the assumption of finite crossing, only a finite number of factors in the product $\prod_{z \in e} p_\beta^m(z)$ is in the interval $(0, 1)$; by construction, the remaining factors are equal to 1. Thus, $p_\beta(e) > 0$ and $e \in T_\beta$.

The finiteness of T_β is established by induction over the number of non-degenerate information sets that are crossed by a path. By definitions of game and behavioral strategy, such a number is finite for every path and every probability distribution at such a set is finite.

Let's assume that T_β is infinite. An infinite path will be constructed such that it crosses an infinite number of information sets with non-degenerate probability distributions. If such a distribution at the information set $\{\tau\}$ is non-degenerate, the construction starts with τ. Otherwise, let y be the first $SU(\tau)$ such that for I_i^k, $y \in I_i^k$, either $i = 0$ or $\beta(I_i^k)(y)$ is non-degenerate. Thus, for at least one $x \in T$, such that $x \in SU(y)$ and such that $p_x = p_x^m > 0$, there is an infinite number of paths in T_β that include x . Let's assume now that a path $e_z \subset T_\beta$ that crosses l non-degenerate information sets was found such that $|\{e \subset T_\beta \text{ s.t. } z \in e\}| = \infty$. Either the probability distribution at z is non-degenerate, in which case let $w = z$, or we can find w, the first $SU(z)$ with such a property. We can choose $v \in SU(w)$ such that $p_v > 0$, $|\{e \subset T_\beta \text{ s.t. } v \in e\}| = \infty$ and $p_\beta(v) = p_\beta(e_v) > 0$. The path e_v crosses $l + 1$ non-degenerate information sets. The path $\cup e_v$ crosses an infinite number of non-degenerate information sets, which is a contradiction.

Ad. (b): If $e \notin T_\beta$, then by definition of T_β, for some $x \in e$ $p_\beta(x) = 0$, which implies $p_\beta(e) = 0$.

If $e \in T_\beta$, then, by definition of game and behavioral strategy, only for a finite number $x \in e$ it may happen that $0 < p_\beta(x) < 1$; for all other $y \in e$, $p_\beta(y) = 1$. This implies $p_\beta(e) > 0$.

Ad. (c) (outline): Since by (a) T_β is finite and non-empty, the argument goes by induction over the sum of probabilities assigned to all paths of the same length. ■

Lemma 1 allows to define payoffs for behavioral strategies.

Payoffs (for behavioral strategy profiles): For every behavioral strategy profile $\boldsymbol{\beta} \in B$, $P_i(\boldsymbol{\beta}) = \Sigma_{e \in T_\beta} P_i(e) \times p_\beta(e)$ for $i = 1, ..., n$.

Since no confusion is anticipated, in the spirit of conserving letters, the original letter P that denotes payoffs assigned to terminal paths is recycled here.

Finally, pure strategies are defined as a special case of behavioral strategies.

Pure strategies: β_i such that $\beta_i(I_i^k)$ is always degenerate is called a *pure strategy* and denoted by π_i. For pure strategies, letter π is used in place of β and Π in place of B.

Decomposition of strategies: The definitions offered below introduce certain partial strategies or strategy profiles G and $\boldsymbol{\beta}$:

$\boldsymbol{\beta}^H$ is $\boldsymbol{\beta}$ *reduced* to H if $\boldsymbol{\beta}^H = \cup_{i=1}^n \beta_i^H$;

β_i^{-H} is a *complement* of β_i with respect to H if $\beta_i^{-H} = \beta_i - \beta_i^H$;

$\boldsymbol{\beta}^{-H} = \cup_{i=1}^n \beta_i^{-H}$ is a *complement* of $\boldsymbol{\beta}$ with respect to H;

B_i^{-H} is the set of all β_i^{-H} for all $\beta_i \in B_i$;

$B^{-H} = \times_{i=1}^n B_i^{-H}$.

Let $\delta_i^H : B_i \to B_i^H \times B_i^{-H}$ denote the *decomposition function for player i* that assigns to $\boldsymbol{\beta}_i$ its reduced strategy $\boldsymbol{\beta}_i^H$ and its complement $\boldsymbol{\beta}_i^{-H}$. The *decomposition function* $\delta^H : B \to B^H \times B^{-H}$ is defined as $(\delta_i^H)_{i=1}^n$. The following simple but useful result holds for every game G and its subgame H:

Lemma 2 *(a) for every i, δ_i^H is 1-1 and onto;*

(b) δ^H is 1-1 and onto.

Proof. Ad (a) (outline): The proof follows directly from the following premises: (i) every β_i can be uniquely partitioned into two partial strategies defined for H and all information sets in G but not in H, with two different strategies producing different partitions; (ii) every pair of partial strategies may result from such a partition; (iii) the union of such partial strategies produces the initial strategy β_i.

Ad (b): it is a simple consequence of (a). ■

Lemma 2 allows to define the function of *concatenation* of strategies that is an inverse of decomposition: For every subgame H of G, and every pair of partial strategy profiles $\boldsymbol{\beta}^H \in B^H$ and $\boldsymbol{\beta}^{-H} \in B^{-H}$, $\sigma^H(\boldsymbol{\beta}^H, \boldsymbol{\beta}^{-H}) = \boldsymbol{\beta}^H \cup \boldsymbol{\beta}^{-H} = \delta^H(\boldsymbol{\beta})$. Moreover, $\sigma^H = (\sigma_i^H)_{i=1}^n$, where every σ_i^H is an inverse of a respective δ_i^H. It is clear that both σ^H and all its all components σ_i^H are 1-1 and onto.

The final two definitions of this section introduce two familiar equilibrium concepts due to Nash (1951) and Selten (1965). For any game G and a strategy profile $\boldsymbol{\beta} \in B$, the equilibrium conditions for $\boldsymbol{\beta}$ are stated as follows:

Nash equilibrium (NE): For every $i \in N$, $\beta_i \in ArgMax_{t_i \in B_i} P_i(\boldsymbol{\beta}_{-i}, t_i)$;

Subgame perfect equilibrium (SPE): For every subgame H of G, $\boldsymbol{\beta}^H$ is a NE in H.

Analogous definitions hold when all considered strategies are pure.

3 Decomposition of pure and behavioral strategies

Throughout this section, G is any game, H any subgame of G, τ the root of G, and ϕ the root of H.

The letters s_i, $\mathbf{s}$, S_i, S, etc. are used to denote strategies, strategy profiles, strategy spaces, joint strategy spaces, etc. that are simultaneously either pure or behavioral, in order to simultaneously process both cases. Lemma 2 guarantees that the operations δ and σ are well-defined and bring unique outcomes within the same family of strategies. Moreover, the definition of finite support of every strategy guarantees that the outcomes of δ and σ have finite support.

Note that, with a few additional definitions, all results of this paper can be strengthened to any type of strategies that are closed under concatenation and decomposition. Pure and behavioral strategies clearly have this property. Further potentially interesting strategies could include stationary strategies.

Profiles $\mathbf{s}$ or $\mathbf{s}^G$ denote any strategy profiles in G and $\mathbf{s}^H$, $\mathbf{s}^{-H}$ (or $\mathbf{s}^{G-H}$) their decomposition with respect to H.

The sets $T_{s(G-H)}$ and $T_{s(H)}$ denote all terminal paths from T_s that do not include the root of H, ϕ, or include ϕ, respectively:

$T_{s(G-H)} = \{e \in T_s \text{ s. t. } \phi \notin e\}$;

$T_{s(H)} = \{e \in T_s \text{ s. t. } \phi \in e\}$.

Lemma 3 says that the payoff in any game G from any profile s is the sum of the payoffs from all terminal paths that do not include ϕ and the payoff of s

reduced to H multiplied by the probability of reaching ϕ.

Lemma 3 $P^G(s) = p_{s^G}(\phi)P^H(s^H) + \Sigma_{e\in T_{s(G-H)}} p_{s^G}(e)P^G(e)$.

Proof. A terminal path in s either includes ϕ or does not include it. By Lemma 1, the total number of paths in s is finite, hence $P^G(s)$ can be represented as the sum of payoffs

$$P^G(\mathbf{s}) = \Sigma_{e\in T_{s(H)}} p_{s^G}(e)P^G(e) + \Sigma_{e\in T_{s(G-H)}} p_{s^G}(e)P^G(e) \tag{1}$$

We need to show that the first term in equation 1 is equal to the first term in the right-hand side of the equation in the lemma. If $p_{s^G}(\phi) = 0$, then we have from the definition of $p_{s^G}(e)$:

$p_{s^G}(e) = 0$ for all $e \in T_{s(H)}$

Since the summation is over a finite set, this means that

$\Sigma_{e\in T_{s(H)}} p_{s^G}(e)P^G(e) = 0 = p_{s^G}(\phi)P^H(s^H) = 0$.

Let's assume now that $p_{s^G}(\phi) > 0$.

First, notice that

(i) every terminal path $e \in T_{s(H)}$ defines a terminal path of s^H in H, and that all terminal paths of s^H in H can be obtained this way.

Moreover, for every $e \in T_{s(H)}$ and the corresponding $e^H \in T^H_{s^H}$, we have:

(ii) $P^G(e) = P^H(e^H)$ and

(iii) $p_{s^G}(e) = \prod_{y\in e} p^m_s(y) = \prod_{y\in e^H} p^m_s(y) \prod_{y\in e, y\notin e^H} p^m_s(y) =$

$= \prod_{y\in e^H} p^m_{s^H}(y) \times p_{s^G}(\phi) = p_{s^G}(\phi) \times p_{s^H}(e^H)$

We can use (i)-(iii) to make suitable substitutions:

$\Sigma_{e\in T_{s(H)}} p_{s^G}(e)P^G(e) = \Sigma_{e^H\in T^H_{s^H}} p_{s^G}(\phi) \times p_{s^H}(e^H) \times P^H(e^H) =$

$= p_{s^G}(\phi)P^H(s^H)$. ■

Upgame: For any sequential game $G =< \Upsilon, N^0, \{T_i\}_{i\in N^0}, I, A, h, P >$, a game $G' =<, \Upsilon', N^{0\prime}, \{T'_i\}_{i\in N^0}, I', A', h', P' >$ is an *upgame* of G (with respect to a subgame H of G) if (a) Υ' is a subtree of Υ such that ϕ, the root of H in G, and all nodes that follow ϕ are substituted with a terminal node ϕ in G' and a payoff vector $P^F(\phi)$ that is of the same dimension as payoffs in G. (b) The players are unchanged and $\{T_i'\}_{i\in N^0}, I', A', h', P'$ are restrictions of f, I, A, h, P to Υ', respectively (with ϕ excluded from restriction). The demonstration that such restrictions define a game is straightforward. It is clear that, in a similar fashion, we can substitute any set of disjoint subgames of G. Every game resulting from such an operation will be called an upgame.

Prune: If F is an upgame of G such that for every root ϕ of a subgame H, $P^F(\phi)$ is a SPE payoff vector in H, F is called a *prune* of G. If H is a subgame with no proper subgames, F is called a *close prune*.

An upgame is obtained when we substitute a subgame, or a set of subgames, with an arbitrary payoff vector. It becomes a prune when the payoff vector results from a SPE in the removed subgames. A close prune is when the removed subgame is the smallest possible, i.e., it has no proper subgames. It is useful to notice a few simple facts. While upgames exist for every game and its every subgame, if a subgame has no SPE, then no prune exists for such a subgame and no SPE exists for the game. Also, a game with a subgame with multiple SPEs resulting in different payoff vectors has multiple prunes. Finally, for close pruning, the condition of SPE in the definition of a prune is equivalent to NE.

By definition of an upgame F corresponding to a subgame H, any complement strategy s^{-H} is also a strategy in F and vice versa.

Let F be any upgame of G, let H be the removed subgame with its root ϕ, $\mathbf{s}$ any strategy profile in G, and let e_ϕ denote the terminal path in F that ends with ϕ.

Lemma 4 $P^G(s) = P^F(s^{-H}) + p_{s^G}(\phi)[P^H(s^H) - P^F(e_\phi)]$.

Proof. By Lemma 3, it is sufficient to show that

$$P^F(\mathbf{s}^{-H}) - p_{s^G}(\phi)P^F(e_\phi) = \Sigma_{e\in T_{s(G-H)}} p_{s^G}(e)P^G(e) \qquad (2)$$

This follows from the fact that s^{-H} is identical with s for all information sets outside of H, and that all terminal paths in F, except for e_ϕ, have the same probabilities and payoffs assigned as in G. For e_ϕ, only the probabilities of reaching ϕ are equal, i.e., $p_{s^F}(\phi) = p_{s^G}(\phi)$. Thus, the left-hand side of equation 2 can be rewritten as follows:

$\Sigma_{e \in T_F} p_{s^F}(e) P^F(e) - p_{s^G}(\phi) P^F(e_\phi) = [\Sigma_{e \in T_{s(G-H)}} p_{s^G}(e) P^G(e) + p_{s^G}(\phi) P^F(e_\phi)] - p_{s^G}(\phi) P^F(e_\phi) = \Sigma_{e \in T_{s(G-H)}} p_{s^G}(e) P^G(e)$ ■

The next result characterizes the critical aspect of pruning a game. Since concatenation and reduction of strategies will be applied to subgames of subgames, we need additional notation:

s^{HJ} is a strategy profile s reduced to a subgame H, and then further reduced to J, a subgame of H;

s^{H-J} is a complement of s^J in H.

A similar notation is applied to individual strategies and payoff profiles.

For any game G and any of its subgames H, and any behavioral strategy profile s^G in G, let F be the upgame of G obtained by substituting H with $P^H(s^H)$.

Theorem 1 *(decomposition) For any game G, any subgame H, and any behavioral strategy profile s^G in G, let F be the upgame of G obtained by substituting H with $P^H(s^H)$. The following conditions are equivalent:*

(a) s^G is a SPE for G;

(b) s^H is a SPE for H and s^{-H} is a SPE for F.

Proof. If $H = G$, the proof is immediate. Thus, we can assume that H is a proper subgame of G.

$(a) \rightarrow (b)$: decomposing a SPE strategy profile must result in a pair of SPE strategy profiles.

Since H is a subgame of G and s^G is SPE in G, s^H must be a SPE for H. We need to prove that s^F is a SPE for F. Let's assume that this is not the case and that we can find J, a subgame of F, such that s^J is not a Nash equilibrium. Two cases are possible:

Case 1: J does not include ϕ. In such a case, J is disjoint with H. Thus, J is also a subgame of G and s^J is identical with s^{FJ}. However, (a) implies that s^J is a Nash equilibrium, which also must be the case for s^{FJ}.

Case 2: J includes ϕ. Similarly to case 1, we can identify s^J with s^{FJ}. Let's assume that there is a player i whose strategy t_i^J gives him a higher payoff than $P_i(s^J)$, i.e.,

$$P_i^J(\mathbf{s}_{-i}^J, t_i^J) > P_i^J(\mathbf{s}^J) \tag{3}$$

We will construct K, a subgame of G, such that player i can change his strategy in the SPE s^G and receive a higher payoff. The subgame K is simply J with the node ϕ developed into H.

Let K be defined as J with ϕ substituted with subgame H. By construction, K is a subgame of G and s^K is a Nash equilibrium in K. Let's define a new strategy t_i^K of player i in subgame K as $(s^{K-J}, t_i^J) = \sigma(s^{K-J}, t_i^J)$. We will now apply Lemma 4 to K, J, strategy profiles (s^{K-J}, t_i^J) and s^K, and to player i:

$$P_i^K(\mathbf{s}^{K-J}, t_i^J) = P_i^{K-J}(\mathbf{s}^{K-J}) + p_{s^K}(\phi)[P_i^J(\mathbf{s}_{-i}^J, t_i^J) - P_i^{K-J}(e_\phi)] \tag{4}$$

and

$$P_i^K(\mathbf{s}^K) = P_i^{K-J}(\mathbf{s}^{K-J}) + p_{s^K}(\phi)[P_i^J(\mathbf{s}^J) - P_i^{K-J}(e_\phi)] \tag{5}$$

where e_ϕ is a path to ϕ in K. Subtracting the respective sides of equation 5 from equation 4, we have:

$P_i^K(s^{K-J}, t_i^J) - P_i^K(s^K) =$

$= P_i^{K-J}(s^{K-J}) + p_{s^K}(\phi)[P_i^J(s_{-i}^J, t_i^J) - P_i^{K-J}(e_\phi)] - P_i^{K-J}(s^{K-J}) + p_{s^K}(\phi)[P_i^J(s^J) - P_i^{K-J}(e_\phi)] =$

$= p_{s^K}(\phi)P_i^J(s_{-i}^J, t_i^J) - p_{s^K}(\phi)P_i^J(s^J) =$

$= p_{s^K}(\phi)[P_i^J(s_{-i}^J, t_i^J) - P_i^J(s^J)] >$

$> 0.$

This inequality means that s restricted to subgame K is not a Nash equilibrium and contradicts our assumption (a).

$(b) \rightarrow (a)$: every strategy profile resulting from the concatenation of SPE strategy profiles is SPE.

Let s^H be SPE in H, the subgame of G, s^F be SPE in F, the prune of G resulting from pruning H, and s be the concatenated strategy profile (s^H, s^F). We need to show that s is SPE.

Let's assume that this is not the case. Thus, we can find a subgame K of G such that s^K is not a Nash equilibrium. This implies that some player i could improve his payoff in K against s_{-i}^K by playing some strategy t_i^K, i.e.,

$$P_i^K(\mathbf{s}_{-i}^K, t_i^K) > P_i^K(\mathbf{s}^K) \tag{6}$$

Three cases are possible.

Case 1: K is also a subgame of F that does not include ϕ. By subgame perfection of F and contrary to 6, s^K must be a Nash equilibrium.

Case 2: K is also a subgame of H. By subgame perfection of H and contrary to 6, s^K must be a Nash equilibrium.

Case 3: K includes the node ϕ and at least one more node from F. In such a case, K must include all nodes that follow ϕ and H must be a subgame of K. Let J be the upgame resulting from substituting H in K with $P^H(s^H)$.

We will construct a pair of strategy profiles in H and in J that will contradict our assumption that s^H and s^F are SPE.

Let $(s_{-i}^H, t_i^H), (s_{-i}^J, t_i^J) = \rho^K(s_{-i}^K, t_i^K)$, i.e., let these profiles be the result of decomposition of profile (s_{-i}^K, t_i^K) in K with respect to its subgame H. By Lemma 4 applied to game K, its subgame H, profiles $\mathbf{s}$ and (s_{-i}^K, t_i^K), and player

i, and noting that we can substitute indices $-H$ with J, we have:

$P_i^K(s) = P_i^J(s^J) + p_{s^K}(\phi)[P_i^H(s^H) - P_i^J(e_\phi)]$

$P_i^K(s_{-i}^K, t_i^K) = P_i^J(s_{-i}^J, t_i^J) + p_{s^K}(\phi)[P_i^H(s_{-i}^H, t_i^H) - P_i^J(e_\phi)]$

where e_ϕ is the path to ϕ in K.

The assumption 6 implies that

$P_i^J(s_{-i}^J, t_i^J) + p_{s^K}(\phi)[P_i^H(s_{-i}^H, t_i^H) - P_i^J(e_\phi)] > P_i^J(s^J) + p_{s^K}(\phi)[P_i^H(s^H) - P_i^J(e_\phi)]$

and, after simplification, that $P_i^J(s_{-i}^J, t_i^J) + p_{s^K}(\phi)P_i^H(s_{-i}^H, t_i^H) > P_i^J(s^J) + p_{s^K}(\phi)P_i^H(s^H)$

This inequality means that at least one of the following inequalities must hold:

$$P_i^J(\mathbf{s}_{-i}^J, t_i^J) > P_i^J(\mathbf{s}^J) \tag{7}$$

$$P_i^H(\mathbf{s}_{-i}^H, t_i^H) > P_i^H(\mathbf{s}^H \tag{8}$$

However, inequality 7 cannot hold due to the fact that s^F is SPE, J is a subgame of F,and s^J is a reduction of s^F to J. The inequality 8 cannot hold due to the fact that s^H is SPE in H. ■

The Decomposition Theorem says that every SPE can be obtained by a concatenation of two SPE subgame-upgame profiles, and that every concatenation of two SPE profiles produces a SPE. The next result extends this finding to a simultaneous pruning of certain subsets of subgames.

Agenda: Consider the graph Υ^A that includes the roots of all subgames of G, has the same root as G, and whose branches are defined as follows: for all $\phi, \psi \in T_D$, (ϕ, ψ) is a branch in Υ^A if ϕ and ψ are roots of some subgames of G, ψ follows ϕ in G, and there is no root of another subgame χ such that χ follows ϕ and ψ follows χ. It is straightforward that such a graph is a game tree. By an obvious association with voting models, it is called the *agenda* of Υ, and the set of all nodes in the agenda is denoted with T_A.

Subgame level: For a subgame H of game G with a root ϕ, the *level of* H is the total number of nodes that are in the path to ϕ in the agenda of G (including both τ and ϕ). Alternatively, the level of H is the number of different subgames (including G and H) of which H is a subgame. It is straightforward that the level of any subgame is a positive integer.

Lemma 5 *For any game G, any positive integer k, and any two different subgames H, J of G of level k, the sets of nodes of H and J are disjoint.*

Proof. Let's assume first that neither the root ϕ of H follows ψ, the root of J, nor ψ follows ϕ. If there is a third node χ that belongs to both subgames, then by definition of subgame both ϕ and ψ must be in the path to χ. Since neither ϕ is in the path to ψ nor ψ in the path to ϕ, we could find at least two different paths to χ: one through ϕ and one through ψ, what is inconsistent with the definition of tree.

Let's assume now that ϕ follows ψ or vice versa. But this would imply that the path to ϕ in the subgame subtree is longer than to ψ (or vice versa), and the subgames cannot be at the same level. ■

Now, the procedure used for obtaining an upgame can be applied to simultaneous pruning of different subgames of the same level.

Lemma 5 implies, after simple calculations, that we can substitute any set of subgames of the same level with payoffs of the appropriate dimension, and obtain an upgame of F. Let $\{H_\theta\}_{\theta\in\Theta}$ be a set of subgames of the same level.

$P^{H_\theta}(s^{H_\theta})$ is the payoff from $\mathbf{s}$ in H_θ;

$\{H_\theta\}_{\theta\in\Theta}$-upgame of G is an upgame of G resulting from substituting subgames $\{H_\theta\}_{\theta\in\Theta}$ with payoff vectors in G;

$(s, \{H_\theta\}_{\theta\in\Theta})$-upgame of G is an upgame of G resulting from substituting subgames $\{H_\theta\}_{\theta\in\Theta}$ with respective payoffs from $\{P^{H_\theta}(s^{H_\theta})\}_{\theta\in\Theta}$.

For any game G, any positive integer k, any subset of subgames $\{H_\theta\}_{\theta\in\Theta}$ of G of level k, and any profile s in G, let F be the $(s,\{H_\theta\}_{\theta\in\Theta})$-upgame of G.

Theorem 2 *(simultaneous decomposition) The following conditions are equivalent:*

(a) s is a SPE for G;

(b) s^F is a SPE for F and for every $H_\theta \in \{H_\theta\}_{\theta\in\Theta}$, s^{H_θ} is a SPE for H_θ.

Proof (outline). When $\{H_\theta\}_{\theta\in\Theta}$ is finite, the proof in both directions follows from the Decomposition Theorem by an inductive argument over the number of removed subgames.

Let's assume that $\{H_\theta\}_{\theta\in\Theta}$ is infinite. By the assumption of a finite support of involved strategies, this case can be reduced to the finite case as follows:

To complete the proof for $(b) \to (a)$, it is sufficient to notice that, by Lemma 1, both s and any alternative profile (s_{-i}, t_i), include a finite number of terminal paths. Thus, only a finite number of subgames from $\{H_\theta\}_{\theta\in\Theta}$ have a nonempty intersection with at least one terminal path that is included either in s or (s_{-i}, t_i). Thus, if any change of strategy for a player i in a subgame J of G could bring i a higher payoff, this could be done for pruning only a finite number of subgames. The application of this observation allows to reduce the infinite case to the case of a finite decomposition.

$(a) \to (b)$ can be reduced to a finite case in a similar way as $(b) \to (a)$. ■

4 Decomposition of mixed strategies (informal discussion)

The present section contains an informal discussion of how the results for behavioral strategies can be applied to mixed strategies. A formal treatment of the subject is complex enough to deserve a separate paper.

While it is easy to show that a mixed strategy profile defines a unique pair of mixed strategy profiles in a subgame and an upgame, different profiles may produce identical pairs of partial profiles. In other words, the decomposition of mixed profiles is not necessarily 1-1, as demonstrated by the following simple example:

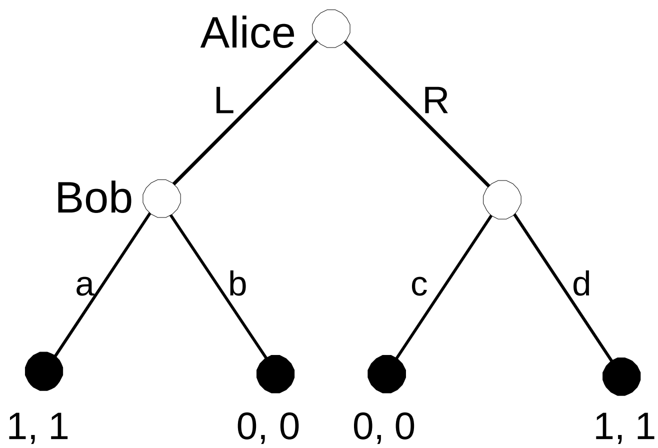

Figure 2: Pure Coordination with perfect information

Bob has four pure strategies: ac, ad, bc, and bd. Every mixed strategy can be represented as a vector of four numbers that represent a probability distribution over ac, ad, bc, and bd, respectively (in fact, one could use just three numbers but parsimony is not the objective in this example). Consider a family of Bob's mixed strategies that are parametrized by a number $r \in [0, \frac{1}{2}]$: $\mu_r = (r, (\frac{1}{2} - r), (\frac{1}{2} - r), r)$. Different values of r define different mixed strategies. Thus, we have a continuum of mixed strategies. It is easy to see that all those strategies decompose into the same pair of partial strategies in subgames $(\frac{1}{2}, \frac{1}{2})$, i.e., such that in every information set, any of Bob's two moves is played with probability $\frac{1}{2}$. Assuming any one-to-one definition of concatenation for mixed strategies, we can concatenate this pair into exactly one mixed strategy.

It is pretty clear that the problem is present in all sufficiently complex games, i.e., in games such that at least one player has at least two information sets that include at least two actions each. It may also appear for equilibrium strategies.

Since concatenation may fail to produce certain mixed profiles, the nice algorithmic flavour of backward induction is lost. However, in games of perfect recall, any mixed strategy is equivalent with respect to payoffs that it generates to some behavioral strategy (this is the essence of Kuhn's (1953) Theorem 4). One can speculate that whenever in a game every mixed strategy is payoff equivalent to some behavioral strategy in all subgames of this game, one can check whether such a mixed strategy is SPE by examining the corresponding behavioral strategy. A further conjecture would be that the condition of perfect recall is sufficient to guarantee payoff equivalence in all games. Thus, backward induction could be extended in the following way: a mixed strategy profile is BIE if the corresponding behavioral profile is BIE.

When recall is imperfect, a player could have a mixed strategy that generates a payoff higher than any of his behavioral strategies (see Kuhn 1953:63-65 for a discussion of an example). Consequently, in the presence of imperfect recall, a mixed profile obtained from the concatenation of SPE partial behavioral profiles is not necessarily SPE within the realm of all mixed strategies.

5 Backward induction for pure, behavioral, and mixed strategies

Theorem 2 allows to define a general procedure of backward induction for any game and pure or behavioral strategies of finite support and finite crossing. Let's fix the game G.

Pruning sequence: The *sequence of pruning* $\{\Theta_j\}_{j=1}^{l}$ is a partition of T_A, the set of agenda nodes, where n is a positive integer, such that for all $j = 1, ..., l$, if $\chi \in \Theta_k$ follows $\psi \in \Theta_m$, then $k \leq m$.

The pruning sequence denotes the order of removing subgames, with Θ_j

denoting the roots of subgames removed in step j. The condition imposed on $\{\Theta_j\}_{j=1}^{l}$ asserts that a subgame H of a subgame J is removed before, or simultaneously with, J. In words, in every step we remove a set of complete subgames.

As an illustration, consider Pure Coordination with perfect information from Figure 2. The agenda includes three nodes: A1, B1 and B2 (the letter represents the player and the number denotes the player's information set, from left to right). There are 6 possible pruning sequences (parentheses are omitted for singleton sets): B1, B2, A1; B2, B1, A1; {B1, B2}, A1; B1, {B2, A1}; B2, {B1, A1}; {B1, B2, A1}. According to the first two sequences, single subgames are pruned; according to the next sequences, both proper subgames are pruned; in the next two sequences, one proper subgame is pruned first and then the entire game is pruned; finally, the entire game is pruned in one step. The condition imposed on the pruning sequence guarantees that A1 is pruned in the last step, possibly with other nodes.

Consider a pruning sequence $\{\Theta_j\}_{j=1}^{l}$ and any set of games $\{G_j\}_{j=1}^{l}$ such that $G_1 = G$ and, for $l > 1$ and for $j = 1, ..., n-1$, G_{j+1} is some upgame of G_j resulting from the substitution of the subgames $\{H_\theta\}_{\theta\in\Theta_j}$ with some payoff vectors. $\{G_j\}_{j=1}^{l}$ is called a *pruning set* for G according to $\{\Theta_j\}_{j=1}^{l}$. For a strategy profile $\mathbf{s} \in S$, $\{G_j\}_{j=1}^{l}$ is a specific pruning set defined recursively as follows: $G_1 = G$ and, for $l > 1$ and for $j = 1, ..., n-1$, G_{j+1} is a ($\mathbf{s}$,$\{H_\theta\}_{\theta\in\Theta_j}$)-upgame of G_j, where $\{H_\theta\}_{\theta\in\Theta_j}$ are subgames of G_j with roots in Θ_j. In words, the procedure begins with the entire game. In every step of pruning, a new tentative game G_{j+1} is created by substituting a subset of subgames from the previous tentative game G_j selected according to $\{\Theta_j\}_{j=1}^{l}$ with some payoff vectors. When pruning is conducted according to a specific profile $\mathbf{s}$, the payoffs in G_{j+1} are generated by strategy profile $\mathbf{s}$ working in the removed subgames in G_j.

A backward induction equilibrium preserves subgame perfection for at least one sequence of pruning:

Backward induction equilibrium: A strategy profile $\mathbf{s}$ is a backward induction equilibrium (BIE) according to a pruning sequence $\{\Theta_j\}_{j=1}^{l}$ if either (a) $l = 1$ and $\mathbf{s}$ is SPE for G or (b) $l > 1$ and (a) the reduction of $\mathbf{s}$ to G_l is SPE and (b) for all $j = 1, ..., l-1$, the reduction of $\mathbf{s}$ to every $H_\theta \in \{H_\theta\}_{\theta\in\Theta_j}$ is SPE. It is a BIE if it is BIE according to at least one pruning sequence.

In words, a strategy profile $\mathbf{s}$ is BIE if we can prune a game using $\mathbf{s}$ in such a way that at every stage $\mathbf{s}$ is SPE in removed subgames and $\mathbf{s}$ is also SPE in the final game resulting from pruning. As we will soon see, the exact sequence of pruning does not matter for the property of being BIE.

For a fixed game G and a set of strategy profiles (behavioral or pure) S, let's denote the subset of all BIE's with S^{BIE} and the subset of all SPE's by S^{SPE}. Let's examine the relationship between S^{BIE} and S^{SPE}:

By our definition of BIE as resulting from *any* sequence of pruning, if $\mathbf{s}$ is SPE, then assuming $l = 1$ implies that it is also BIE.

Conversely, if $\mathbf{s}$ is a BIE, then we can find a pruning sequence $\{\Theta_j\}_{j=1}^{l}$ that satisfies the conditions from the definition of BIE. Excluding the trivial case of $l = 1$, Theorem 2 applied $l-1$ times guarantees that $\mathbf{s}$ is SPE. The argument goes as follows: First, by definition of BIE, $\mathbf{s}$ reduced to G_l must be SPE and, second, by Theorem 2, if $\mathbf{s}$ reduced to G_{j+1} and $\{H_\theta\}_{\theta\in\Theta_j}$ is in all cases SPE, then $\mathbf{s}$ reduced to G_j must be SPE. Those facts imply by induction that $\mathbf{s}$ reduced to G_1 (i.e., $\mathbf{s}$ itself) is SPE.

The relationship between subgame perfection and backward induction can now be stated formally. It is straightforward:

Corollary 1 *For any game G, $B^{SPE} = B^{BIE}$ and $\Pi^{SPE} = \Pi^{BIE}$.*

In fact, the Corollary remains valid not only behavioral or pure strategies but for all families of sets of behavioral strategies closed under concatenation and decomposition. A simple consequence of the Corollary (in combination with Theorem 2) is that if $\mathbf{s}$ is BIE with one pruning sequence, then it must be BIE with any pruning sequence. The reasoning is simple: If $\mathbf{s}$ is BIE under one sequence, then it is SPE by Corollary 1, then it is (by Theorem 2 applied consecutively) SPE in all upgames and pruned subgames according to the second sequence.

Proposition 1 *For any game G, any strategy profile $\mathbf{s}$, and any two sequences of pruning $\{\Theta_j\}_{j=1}^{l}$ and $\{\Omega_j\}_{j=1}^{m}$, $\mathbf{s}$ is BIE according to $\{\Theta_j\}_{j=1}^{l}$ iff $\mathbf{s}$ is BIE according to $\{\Omega_j\}_{j=1}^{m}$.*

All sequences of pruning return precisely the same strategy profiles. The only differentiating factor is the convenience of using one sequence over another. The following algorithm describes finding all SPE's with a given non-trivial (i.e., for $l > 2$) pruning sequence:

1. *Sequence of pruning:* Set a pruning sequence $\{\Theta_j\}_{j=1}^{l}$. Denote all partial strategy profiles obtained in step j with S_j^{SPE}. Thus, $S^{SPE} = S_l^{SPE}$. Set the initial set of partial strategy profiles S_1^{SPE} defined as the set of all partial strategy profiles in G that are SPE for all subgames of G with roots from Θ_1;

2. *Concatenation:* For step j, $1 \leq j \leq l$, the procedure generated S_j^{SPE}.

If $j = l$, stop. Otherwise,

If $S_j^{SPE} = \varnothing$, then set $S^{SPE} = \varnothing$.

If $S_j^{SPE} \neq \varnothing$, then for every $\mathbf{s}_j^{\varkappa} \in S_j^{SPE}$ perform the following procedure. Let's denote the upgame corresponding to $\mathbf{s}_j^{\varkappa}$ by G_j and the corresponding subgames with $\{H_\theta\}_{\theta \in \Theta_j}$.

For every s^{θ_j}, a partial strategy profile for all subgames $\{H_\theta\}_{\theta \in \Theta_j}$ exactly one must hold:

(i) s^{θ_j} is a SPE for all $\{H_\theta\}_{\theta \in \Theta_j}$. In such a case, include $\mathbf{s}_j^{\varkappa} \cup s^{\theta_j}$ in S_{j+1}^{SPE};

(ii) s^{θ_j} is not a SPE for at least one of $\{H_\theta\}_{\theta \in \Theta_j}$. In such a case, discard $\mathbf{s}_j^{\varkappa} \cup s_\theta^*$.

Below, a few special (partially overlapping) cases and applications of Corollary 1 are discussed. Before starting the discussion, a rather obvious limitation of the generalized backward induction procedure should be acknowledged. If a game has no proper subgames (the game's agenda is a singleton), then the method offers no computational or other benefits since the only pruning sequence is the trivial one. In general, the usefulness of the method depends on the structure of the agenda.

1. Pure strategies in finite games of perfect information.

In this simplest classic case, the agenda is identical with the game tree minus the endnodes. The game tree is pruned one subgame at a time and every subgame includes only a singleton decision node. Looking for a NE or SPE in every such a subgame is equivalent to finding the best move (or the best moves) of a player. The existence of a BIE for finite games is the thesis of Kuhn's Corollary 1 (Kuhn, 1953: 61). If payoffs at some stage are identical, one may obtain many SPEs.

2. Pure strategies in general games of perfect information.

Among well-known examples of non-finite games of perfect information there are games of fair division (Steinhaus 1948, Brams and Taylor 1996) or Romer-Rosenthal (1978) agenda setter model. The algorithm for such games closely resembles simple backward induction.

Close pruning is a particularly simple version of backward induction. Below, it is applied to the following Romer-Rosenthal's agenda setter model.

Example 1 *Romer-Rosenthal agenda setter model.*

Two players, the Agenda setter A and the Legislator L, have Euclidean preferences in the issue space [0,3] and the ideal points $a = 0$ and $l = 2$, respectively. The status quo is $q = 3$. First, A proposes a policy $x \in [0, 3)$. Next, L chooses the final law from $\{x, q\}$.

The agenda-setter story defines a unique game G such that the pure strategy spaces and payoffs can be represented as follows:

$\Pi_A = [0,3)$; $\Pi_L = \{X : X \subset [0,3)\}$

$$P_A = \begin{cases} -x & \text{if } x \in X \\ -3 & \text{if } x \notin X \end{cases}; \; P_L = \begin{cases} -|x-2| & \text{if } x \in X \\ -1 & \text{if } x \notin X \end{cases}$$

Set X represents all policy points that L would accept.

There are two levels in the game that correspond to the periods. We are interested in pure strategies only. The adopted sequence of pruning removes all subgames of the same level at a time.

Step 1: At level two, there is a continuum of subgames that are parametrized by the issue space $[0,3)$. The subgame at x offers two options to L: when x is proposed, to accept it or to reject it (which implies the acceptance of q). The best actions for L for a subgame following x can be described as follows:

If $x < 1$, reject x;

If $x > 1$, accept x;

If $x = 1$, reject or accept x.

Applying simultaneous pruning to level 2 brings our first set of partial SPEs that includes two partial SPE profiles which are two strategies of L:

$S_1^{SPE} = \{X_1, X_2\}$, where $X_1 = [1,3)$ and $X_2 = (1,3)$, i.e., "accept every offer not smaller than 1", and "accept every offer greater than 1."

The two partial strategy profiles produce two upgames G_1 and G_2 with the unique active player A and his strategy space [0,3), where his payoffs are defined as follows:

$$G_1 : P_L(x) = \begin{cases} -3 & \text{if } x < 1 \\ -x & \text{if } x \geq 1 \end{cases}$$

$$G_2 : P_L(x) = \begin{cases} -3 & \text{if } x \leq 1 \\ -x & \text{if } x > 1 \end{cases}$$

Step 2: We have to consider all partial strategy profiles from S_1^{SPE} that were obtained in Step 1, i.e., X_1 and X_2.

Case 1 (X_1): The unique best action in G_1 is 1. The partial strategy profile in the removed subgames corresponding to G_1 is X_1. When this profile is concatenated with L's choice, the resulting SPE in the entire game is $(1, [1, 3))$;

Case 2 (X_2): There is no best action for G_2. The corresponding strategy profile in the removed subgames X_2 is discarded.

Solution: There is a unique SPE in G equal to $(1, [1, 3))$.

3. Behavioral and pure strategies in finite games.

In all finite games, we can prune subgames one close prune at a time and look for Nash equilibria in behavioral or pure strategies in a fashion very similar to looking for pure strategy equilibria in finite games of perfect information. An especially interesting case is a when the game is finitely repeated.

Let's assume first that a game G has precisely one equilibrium in either pure or behavioral strategies. Let G^k be G repeated k times, $k \geq 2$. It is straightforward that, when close pruning is applied, there is precisely one possible SPE in every removed subgame of G^k, i.e., the one corresponding to the SPE in G. When all subgames of the same level are pruned, the resulting game is G^{k-1} plus a payoff adjustment for all players equal to the equilibrium payoff in G. This reasoning implies the following result ($\mathbf{s}$ denotes again either pure or behavioral strategy):

Corollary 2 *For any finite game G that has a unique SPE $\mathbf{s}$ and for any integer $k \geq 2$, G^k has exactly one SPE that is equal to the repeated concatenation of $\mathbf{s}$.*

A simple consequence of Corollary 2 are well-known facts that the finitely repeated games such as the Prisoner's Dilemma or Matching Pennies have precisely one SPE.

When the underlying one-shot game has many equilibria, the task of calculating the total number of equilibria in a finitely repeated game may become quite complex since different equilibria may contribute backward different payoffs vectors. Nevertheless, our procedure simplifies the calculations greatly.

Example 2 *Twice repeated Pure Coordination.*

In one-shot Pure Coordination, two players simultaneously choose one of their two strategies (denoted with the same labels for both players, i.e., L and R). If their choices coincide, they get the payoff of one each. Otherwise, they both receive zero (the game is the one from Figure 2 with simultaneous moves assumed).

There are two levels in the twice-repeated Pure Coordination. Our sequence of pruning again coincides with the levels.

Step 1: There are four subgames at level two with three NE in each subgame. This produces 81 NE strategy profiles; sixteen of them are in pure strategies. The partial strategy profile in every subgame can be parametrized in the following simple way: LL and RR denote the NE in which the players coordinate on (L, L) or (R, R), respectively, and $\frac{1}{2}\frac{1}{2}$ denotes the remaining NE in completely mixed strategies. For instance, the label $LLLLRR\frac{1}{2}\frac{1}{2}$ would denote the following partial strategy profile: both players play L in the first and second subgame, they both play R in the third subgame, and they both play $\frac{1}{2}$ in the fourth subgame. Thus, the sets of partial SPEs in pure and behavioral strategies are defined as follows, respectively:

$\Pi_1^{SPE} = \{yzvw \text{ such that } y, z, v, w \in \{LL, RR\}\}$;

$B_1^{SPE} = \{yzvw \text{ such that } y, z, v, w \in \{LL, RR, \frac{1}{2}\frac{1}{2}\}\}$.

Step 2: At this point, it makes sense to separate the cases of pure and behavioral strategies.

Pure strategies: The characterization of all SPE in pure strategies is easy. Every partial profile from Π_1^{SPE} obtained in Step 1 adds exactly 1 to the payoff in the upgame, or the first iteration of the game. Thus, there are two SPEs in the upgame, LL and RR, per every partial strategy profile obtained in Step 1. Consequently, the total number of SPEs in pure strategies is 32 and the set of SPE in pure profiles can be defined as follows:

$\Pi^{SPE} = \{xyzvw \text{ such that } x, y, z, v, w \in \{LL, RR\}\}$.

Behavioral strategies: counting the total number of SPEs is also easy because all prunes resulting from removing subgames have similar structures and each of them has exactly three NEs. In every prune, players receive the payoffs of $1\frac{1}{2}$ or 2 for coordinating their strategies and 0 or $\frac{1}{2}$ for discoordination. Thus, every prune is either Pure Coordination or a variant of Asymmetric Coordination. In all cases, there are exactly two NE in pure strategies and one NE in completely mixed strategies. Thus, every one of 81 partial strategy profiles from Step 1 can be preceded by three NE in the upgame. The total number of SPE in behavioral strategies is 243.

While listing all pure strategy SPEs was trivial, enumerating all behavioral SPEs is more laborious but still manageable. The task can be simplified by dividing all 81 partial strategies from B_1^{SPE} into equivalence classes. Let's group them in the following way:

Sixteen partial strategies $xyzv$ where $x, y, z, v \in \{LL, RR\}$;

Four partial strategies $x\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}y$ where $x, y \in \{LL, RR\}$;

Four partial strategies $\frac{1}{2}\frac{1}{2}xy\frac{1}{2}\frac{1}{2}$ where $x, y \in \{LL, RR\}$;

One partial strategy $\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$.

In all 25 cases, there are three identical NEs in the upgame: LL, RR, and $\frac{1}{2}\frac{1}{2}$. I leave to the reader the fascinating chore of calculating the remaining NEs. A hint: the following sets of partial strategy profiles will have identical or symmetric NE in the upgame:

Sixteen cases $xyw\frac{1}{2}\frac{1}{2}$ or $\frac{1}{2}\frac{1}{2}xyw$ where $x, y, w \in \{LL, RR\}$;

Sixteen cases $x\frac{1}{2}\frac{1}{2}yw$ or $xy\frac{1}{2}\frac{1}{2}w$ where $x, y, w \in \{LL, RR\}$;

Four cases $x\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$ or $\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}x$ where $x \in \{LL, RR\}$;

Four cases $\frac{1}{2}\frac{1}{2}x\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$ or $\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}x\frac{1}{2}\frac{1}{2}$ where $x \in \{LL, RR\}$;

Sixteen cases $xy\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$ or $x\frac{1}{2}\frac{1}{2}y\frac{1}{2}\frac{1}{2}$ or $\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}xy$ or $\frac{1}{2}\frac{1}{2}x\frac{1}{2}\frac{1}{2}y$ where $x, y \in \{LL, RR\}$.

A note on mixed strategies: Characterizing mixed strategy SPE is most laborious since an infinite number of mixed strategies may share the same subgame behavioral representation. Section 4 offers the following informal suggestion for finding all mixed strategy equilibria: for every SPE in behavioral strategies $\boldsymbol{\beta}$, calculate all mixed strategy profiles $\boldsymbol{\mu}$ that are payoff-equivalent to $\boldsymbol{\beta}$ in all subgames. As an example, consider a mixed strategy of player 1 μ_p that is equal, for any $p \in [0, \frac{1}{2}]$ to:

$\mu_p = (pLLRRL, pLLLLL, (1-p)LLLRL, (1-p)LLRLL)$

Strategy μ_p plays the pure strategies $LLRRL$ and $LLLLL$ with probability p and the pure strategies $LLLRL$ and $LLRLL$ with probability $1-p$. Strategies μ_p are payoff equivalent in all subgames for all $p \in [0, \frac{1}{2}]$, and their shared behavioral representation is $\beta = LL\frac{1}{2}\frac{1}{2}L$. In words, every mixed strategy μ_p plays L initially and in all subgames that follow coordination; it plays $\frac{1}{2}$ in all subgames that follow discoordination.

It is easy to show that every mixed strategy that is payoff equivalent in all subgames to β must be one of the μ_ps. Thus, for the unique SPE in behavioral strategies that involves β, $(LL\frac{1}{2}\frac{1}{2}L; LL\frac{1}{2}\frac{1}{2}L)$, we obtain the corresponding two-parameter family of strategy profiles that are payoff-equivalent in subgames $M_{p,r} = \{(\mu_{1,p}\ \mu_{2,r})$ s. t. $\mu_{1,p}$ and $\mu_{2,r}$ satisfy our Condition for μ_p for $p \in [0, \frac{1}{2}]$ and $r \in [0, \frac{1}{2}]$, respectively$\}$. The family $M_{p,r}$ completely characterizes all mixed strategy profiles that pairwisely satisfy the requirement of payoff-equivalence in all subgames to a behavioral strategy profile $\beta\beta$.

4. Parametrized families of games.

A variant of the procedure may be applied when the parametrization of a family of games is sufficiently regular. In Example 1, one can set any status quo q and any ideal points l and a. The most interesting case is when $a < l < q$ and $l - a > q - l$ (or when both inequalities are reversed). Finding the unique SPE for the family of games parametrized with a, q and l is almost as simple as in Example 1.

6 Conclusion

An obvious open question is whether the results for behavioral strategies of finite support and crossing can be generalized to all "rough" behavioral strategies. Attacking this question would demand leaving the comfortable world of finite probability distributions and using measure theory in the spirit of Aumann's (1953) pioneering contribution. The framework presented in this paper goes around measure-theoretic difficulties with suitable constraints imposed on a subset of behavioral strategies. But even with those constraints suspended, I couldn't find an example that would violate the equivalence of BIE and SPE. Thus, a natural conjecture is that it holds universally.

Two more general problems deserve a comment.

First, a general axiomatic framework applied in the present paper encompasses more games than the axiomatic approaches of von Neumann and Kuhn. When game theory was born, considering finite games only seemed natural, and non-finite games appeared in the literature infrequently. Today we routinely go beyond the limitations of finite games either with a continuum of strategies that represent quantity, price or position in the issue space, or with infinite repetition of a game. I believe that contemporary game theory deserves sound axiomatic foundations that would include infinite games, and that would lead towards a more unified and complete discipline. Concepts that were axiomatically analyzed for finite games, such as Kreps and Wilson's (1982) sequential equilibrium, seem to be obvious targets for axiomatic investigation. The present paper demonstrates that a general framework of infinite games is manageable enough to allow us to obtain interesting new results or extensions of well-known ones.

Second, another line of inquiry would ask whether backward induction can be modified to go beyond subgame perfection. An immediate ad-hoc modification would consider only those SPE's that exclude partial equilibria with weakly dominated strategies. Perhaps, after a suitable modification of the main principle, backward induction-like reasoning could also produce solutions such as perfect equilibrium, sequential equilibrium, or other refinements. Proving that this is not the case would be an interesting finding as well.

Further refinements of backward induction could produce computational benefits similar to those obtained for subgame perfection. Backward solving is equivalent to hierarchical concatenation of solutions. Thus, solving a game with backward reasoning is equivalent to collecting together those independent solutions and connects the global solution with stage-wise decision-making. Moreover, for games with proper subgames, applying backward induction could simplify finding solutions and obtaining results the same way it does so for SPE.

References

AUMANN, R. J. (1964): "Mixed and Behavior Strategies in Infinite Extensive Games," in *Advances in Game Theory, 2nd Ed.*, ed. by R. J. Aumann, M. Dresher, L. S. Shapley, and A. W. Tucker. Princeton: Princeton University Press.

BASU, K. (1988): "Strategic Irrationality in Extensive Games," *Mathematical Social Sciences*, 15, 247-260.

— (1990): "On the Non-Existence of a Rationality Definition for Extensive Games," *International Journal of Game Theory*, 19, 33-44.

BICCHIERI, C. (1989): "Self-Refuting Theories of Strategic Interaction: A Paradox of Common Knowledge," *Erkenntnis*, 30, 69-85.

BINMORE, K. (1987): "Modeling Rational Players, Part 1," *Economics and Philosophy*, 3, 179-214.

— (1988): "Modeling Rational Players, Part 2," *Economics and Philosophy,* 4, 9-55.

BONANNO, G. (1991): "The Logic of Rational Play in Games of Perfect Information," *Economics and Philosophy*.

BRAMS, S., and A. TAYLOR (1996): *Fair Division. From Cake-Cutting to Dispute Resolution*. Cambridge: Cambridge University Press.

FUDENBERG, D., D. M. KREPS, and D. K. LEVINE (1988): "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44, 354-380.

FUDENBERG, D., and J. TIROLE (1991): *Game Theory*. Cambridge, MA: The MIT Press.

KOHLBERG, E., and J.-F. MERTENS (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 1003-37.

KREPS, D. M., and R. WILSON (1982): "Sequential Equilibria," *Econometrica*, 50, 863-894.

KUHN, H. W. (1953): "Extensive Games and the Problem of Information.," in *Contributions to the Theory of Games 1*, ed. by H. W. Kuhn, and A. W. Tucker. Princeton: Princeton University Press, 193-216.

MYERSON, R. (1991): *Game Theory. Analysis of Conflict*. Cambridge, MA: Harvard University Press.

NASH, J. F. (1951): "Non-Cooperative Games," *Annals of Mathematics*, 54, 155-62.

PETTIT, P., and R. SUGDEN (1989): "The Backward Induction Paradox," *Economics and Philosophy*, 14, 95-125.

RENY, P. J. (1988): "Rationality, Common Knowledge and the Theory of Games," Princeton: Princeton University.

ROMER, T., and H. ROSENTHAL (1978): "Political Resource Allocation, Controlled Agendas, and the Status Quo," *Public Choice*, 33, 27-44.

ROSENTHAL, R. W. (1981): "Games of Perfect Information, Predatory Pricing, and the Chain-Store Paradox," *Journal of Economic Theory*, 25, 92-100.

SCHELLING, T. C. (1960): *The Strategy of Conflict*. Cambridge, Ma: Harvard University Press.

SCHWALBE, U., and P. WALKER (2001): "Zermelo and the Early History of Game Theory," *Games and Economic Behavior*, 34, 123-137.

SELTEN, R. (1965): "Spieltheoretische Behandlung Eines Oligopolmodels Mit Nachfrageträgheit," *Zeitschrift für die gesampte Staatswissenschaft*, 301-324, 667-689.
— (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 25-55.
— (1978): "The Chain Store Paradox," *Theory and Decision*, 9, 127-159.
STACKELBERG, H. F. V. (1934): *Marktform Und Gleichgewicht (Market Structure and Equilibrium)*. Vienna.
STEINHAUS, H. (1948): "The Problem of Fair Division," *Econometrica*.
VON NEUMANN, J., and O. MORGENSTERN (1944): *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.
ZERMELO, E. (1913): "Über Eine Anwendung Der Mengenlehre Auf Die Theorie Des Schachspiels," Cambridge: Cambridge University Press, 501-504.