Theoretical Economics 2 (2007), 163–185 1555-7561/20070163

Valuation equilibrium

PHILIPPE JEHIEL
Paris School of Economics and Department of Economics, University College London

DOV SAMET
Graduate School of Business Administration, Tel Aviv University

We introduce a new solution concept for games in extensive form with perfect information, *valuation equilibrium*, which is based on a partition of each player's moves into *similarity classes*. A *valuation* of a player is a real-valued function on the set of her similarity classes. In this equilibrium each player's strategy is *optimal* in the sense that at each of her nodes, a player chooses a move that belongs to a class with maximum valuation. The valuation of each player is *consistent* with the strategy profile in the sense that the valuation of a similarity class is the player's expected payoff, given that the path (induced by the strategy profile) intersects the similarity class. The solution concept is applied to decision problems and multi-player extensive form games. It is contrasted with existing solution concepts. The valuation approach is next applied to stopping games, in which non-terminal moves form a single similarity class, and we note that the behaviors obtained echo some biases observed experimentally. Finally, we tentatively suggest a way of endogenizing the similarity partitions in which moves are categorized according to how well they perform relative to the expected equilibrium value, interpreted as the aspiration level.

KEYWORDS. Game theory, bounded rationality, valuation, similarity, aspiration.
JEL CLASSIFICATION. C72, D81.

1. INTRODUCTION

"Buy low, sell high" is obviously an oversimplified rule for investors. It disregards many aspects of the market that could be taken into account when one makes investment decisions. Still, this popular wisdom captures an essential aspect of decision making in complex situations where a decision maker finds it impossible to evaluate the future consequences of her move precisely. This simple rule groups together many such moves. Indeed, it suggests that one should disregard all aspects of the market other than

Philippe Jehiel: jehiel@enpc.fr
Dov Samet: DovS@tauex.tau.ac.il
We thank the editor and three anonymous referees as well as seminar participants at Paris (Roy seminar), CORE, Prague, Gerzensee (ESSET 2003), the Conference on Bounded Rationality and Learning, Urbana Champaign 2003, Northwestern, the Second World Congress of Game Theory (Marseilles 2004), UCL, and the Center for Rationality, Jerusalem.

the price. In all states of the market in which the price is high, the investor evaluates all "buy" moves as superior to "sell" moves, and the opposite is true in all states of the market in which the price is low.

More generally, the difficulty of evaluating different moves is a feature of most complex games. The strategic form of such games may be so big that it cannot be considered by real players. In such cases, the task of choosing the right move at each node is too hard as one has a limited understanding of the future consequences associated with each possible move. Instead, players can group several moves together, at different decision nodes, consider them similar, and evaluate the whole group rather than each move individually. Where does the similarity of moves come from? It may depend on some conventional wisdom, either derived from a "narrative" behind the game or possibly gleaned from the experience of previous players, which contributes to our understanding of what could possibly be more desirable. Thus, in the investment game we know enough economic theory (or at least believe that we do) to tell that the price level should be an important ingredient in making our investment decision. When we think of a complex game like chess, the similarity of moves can be deduced from the structure of the game (the board and the rules of the game, which can hardly be seen in the game tree), which helps us to compare certain moves in different configurations of the board (for example, by assessing the strength of a position by the profile of pieces rather than the board position itself). Having this picture in mind, our premise is that the grouping of moves into similarity classes is given to players externally, and that it is not a matter of choice by individual players.

Our analysis is only a first step in the study of the grouping of moves. A complete study should endogenize the formation of this grouping. Here we study mainly the implications of the grouping of moves (assumed to be exogenously given) on the equilibrium analysis. We illustrate how the solutions obtained differ from other equilibrium approaches and explore whether interesting phenomena can be explained by the approach.

We consider games in extensive form with perfect information and assume that each player has a partition of her nodes into similarity classes. A *valuation* of a player assigns a numerical value to each of her similarity classes.

We introduce two solution concepts for extensive games based on similarity classes and their valuation: *valuation equilibrium* and *sequential valuation equilibrium*.[1] A valuation equilibrium (VE) is a profile of behavioral strategies for which players have valuations that satisfy two conditions.

- Each player's strategy is *optimal* for her valuation. By this we mean that at each node where she plays she chooses one of the moves that belongs to a class with maximum valuation.

[1] Steiner and Stewart (2006) study a learning model applied to coordination games in which the payoffs obtained with the same strategy are aggregated over nearby states. They show how such a learning process may select among several equilibria of the underlying coordination games using the techniques of global games. While such a learning process is related in spirit to our approach (in their model as in ours the notion of similarity is exogenously given), their limiting outcome does not however correspond to a valuation equilibrium because their model of similarity is not in terms of partition of moves.

- Each player's valuation is *consistent* with the strategy profile. That is, the valuation attached to a player's similarity class is the expected payoff of the player given that the path (induced by the strategy profile) intersects this class.

We think of the consistency requirement as resulting from a learning process in which similarity classes are kept fixed (they are externally given) and players keep updating the valuations assigned to similarity classes along the learning process.[2] Observe that the consistency requirement imposes constraints only on the valuations of similarity classes that are reached with positive probability in equilibrium. Our second and main solution concept, the sequential valuation equilibrium, imposes a stronger notion of consistency that also applies to unreached similarity classes. Very much like sequential equilibrium (Kreps and Wilson 1982), sequential consistency requires that the valuations of unreached similarity classes be consistent with small perturbations of the strategy profile.

In Section 2 we formally define the concepts. In Section 3, we discuss the motivation for our approach and review some simple ideas in chess (like the values assigned to pieces) in light of our solution concept.

Sections 4 and 5 make a number of observations regarding valuation equilibrium and its link to other approaches. We first show that in finite environments a sequential valuation equilibrium (SVE) always exists, for any given similarity partitions. We also note that for maximal similarity partitions (i.e., when each move forms a similarity class), an SVE coincides with a subgame perfect Nash equilibrium.

In Section 5 valuation equilibria are related to, and contrasted with, other solutions. The examples in this section are deliberately simple and serve to illustrate a number of theoretical insights. We first consider decision problems. In sharp contrast with standard notions of equilibrium, we provide a one-agent decision problem involving chance moves such that in equilibrium the agent makes the worst possible decision at every decision node. In another one-agent setup, the decision maker must make a binary decision in either problem a or problem b as selected by nature. For a similarity grouping involving three classes (one class contains one move in each problem a and b, and the other classes are singletons), we find that there are two strict SVEs,[3] thus showing that it is not possible to interpret SVE as a standard solution of a different game possibly with different final payoffs and different information structures but the same move structure. We next move on to multi-player games. We first observe that any sequential equilibrium of games with incomplete information and perfect recall can be represented as an SVE by natural choices of similarity partitions. Thus, the valuation approach covers the usual information treatment while allowing for more flexibility (as can be inferred from the one-agent decision problem described above).[4] We also contrast the valuation approach with the imperfect recall approach (Piccione and Rubinstein 1997).

[2]From the more general perspective in which similarity classes can also be adjusted along the learning process, our approach implicitly requires that similarity classes are adjusted much more slowly than the valuations attached to them. The case in which the adjustments of the similarity classes and of their corresponding valuations take place at the same pace may require another solution concept to be considered.

[3]An SVE is strict if for every player only one strategy is optimal for the valuation.

[4]The valuation approach allows for grouping that would not even make sense in the incomplete

In Section 6 we apply the valuation approach to stopping decision problems and games where we consider the grouping of all non-terminal moves into one similarity class while terminal moves are singleton similarity classes. We observe that a decision maker facing several stopping decision problems of length T must either stop immediately in all but one of these problems or go on till the very end, i.e., period T (with positive probability) in at least one of these problems. This holds irrespective of the payoffs chosen, which illustrates a systematic timing bias implied by the valuation approach in such decision problems. Such biases are further explored in more structured situations in which a positive correlation is assumed between the payoffs obtained with an immediate stop and the payoffs obtained when the decision maker goes ahead (in this case we assume that $T=2$). In stopping games, we observe that the grouping of non-terminal moves may allow players to sustain threats that would not otherwise be credible.

Finally, in Section 7 we briefly suggest a way to endogenize the similarity partitions based on the idea of an aspiration level. Moves are categorized according to whether they deliver less, more, or the same level of payoff as a benchmark payoff, referred to as the aspiration level, which is assumed to be the equilibrium payoff. We refer to such an equilibrium as an *aspiration-based sequential valuation equilibrium* (ASVE). After briefly providing a learning motivation for the aspiration approach, we observe that the subgame perfect Nash equilibrium is always an ASVE, but other strategy profiles may be ASVEs as well. Still, in zero-sum two-player games without chance moves, a player must get her value in any ASVE. This provides an interesting class of games in which the aspiration grouping delivers good outcomes.

2. Valuation equilibrium

2.1 *Games and strategies*

Consider a finite extensive game with perfect information. It is specified by (1) a finite set of players I, (2) a tree (Z,N,r,A), where Z and N are the (finite) sets of terminal and non-terminal nodes, respectively, $r\in N$ is the root of the tree, and A the set of arcs, (3) a non-intersecting collection of subsets $(N_i)_{i\in I}$ of N where N_i is the set of nodes at which it is i's turn to play, and (4) a collection $(f_i)_{i\in I}$ of functions where $f_i : Z \longrightarrow R$ is i's payoff function defined over the set of terminal nodes Z.

Elements of A are ordered pairs (n,m), where $m\in Z\cup N$ is the immediate successor of $n\in N$. The *moves* of player i at node $n\in N_i$ are the nodes in $M_i(n)=\{m\mid (n,m)\in A\}$. The set of moves of player i is denoted $\overline{M}_i=\bigcup_{n\in N_i} M_i(n)$.

A (behavioral) strategy for player i is a function σ_i defined on N_i such that for each $n\in N_i$, $\sigma_i(n)$ is a probability distribution on $M_i(n)$, where $\sigma_i(n)(m)$ should be interpreted as the probability that move $m\in M_i(n)$ is selected at n according to $\sigma_i(n)$.

The nodes in $N\setminus\cup_{i\in I}N_i$ belong to nature, which has a fixed strategy. We assume without loss of generality that at each of its nodes n, nature assigns a positive probability to each of the moves at n.

information treatment. Two moves at two different nodes may be members of the same similarity class even though the number of moves available at the nodes differ.

For a strategy profile $\sigma = (\sigma_i)_{i \in I}$, we let P^σ be the probability over Z induced by σ and nature's strategy. That is, for each $z \in Z$, $P^\sigma(z)$ is the probability that z is reached when σ is played.

2.2 *Similarity and valuation*

Player i has a relation of *similarity* on $\overline{M}_i$, her set of moves. We assume that it is an equivalence relationship and denote by Λ_i the partition of $\overline{M}_i$ into similarity classes. For $m \in \overline{M}_i$, $\lambda(m)$ denotes the similarity class in Λ_i that contains m. For each similarity class $\lambda \in \Lambda_i$, we let $Z(\lambda)$ be the set of all terminal nodes that are descendants of some node in λ.

A *valuation* for player i is a function $v_i : \Lambda_i \longrightarrow R$.

2.3 *Equilibria*

We say that the strategy σ_i is *optimal for the valuation* v_i if for each $n \in N_i$ and $m \in M_i(n)$, $\sigma_i(n)(m) = 0$ whenever $m \notin \arg\max_{m' \in M_i(n)} v_i(\lambda(m'))$. That is, at each of her nodes, with probability 1 player i chooses only moves that belong to similarity classes with maximal valuation.

In equilibrium we require the valuations to be *consistent* with the strategy profile used by the players. Formally, we say that the valuation v_i is *consistent* with the profile σ if for each $\lambda \in \Lambda_i$ such that $P^\sigma(Z(\lambda)) > 0$, we have

$$v_i(\lambda) = E^\sigma(f_i \mid Z(\lambda)), \tag{1}$$

or equivalently,

$$v_i(\lambda) = \sum_{z \in Z(\lambda)} P^\sigma(z) f_i(z) / P^\sigma(Z(\lambda)).$$

We think of the consistency requirement as resulting from a learning process in which each player i keeps track only of the average value of picking a move in λ, for every $\lambda \in \Lambda_i$. More precisely, assume that strategies have settled down at σ, and that at least one $m \in \lambda$ is played with positive probability according to σ (i.e., $P^\sigma(Z(\lambda)) > 0$). By keeping track of the average payoff she obtained whenever she played a move in the various similarity classes in Λ_i, player i will (eventually) value λ according to (1), as this valuation corresponds to the expected payoff obtained by player i given that (at least) one of the moves in λ was played (and that the strategy profile is σ). She will next pick a strategy that is optimal for her valuation, which gives rise to the following solution concept.

Definition 1. A strategy profile $\sigma = (\sigma_i)_{i \in I}$ is a *valuation equilibrium* (VE) if there exists a valuation profile $v = (v_i)_{i \in I}$ such that for each i,

- σ_i is optimal for v_i
- v_i is consistent with σ.

Note that being consistent with σ does not impose any restriction on the valuation of similarity classes that are not reached under σ. Thus, it is possible that a strategy profile is supported by a valuation even though the valuations of unreached classes bear no relation to the true payoffs of the game. Specifically, a player may avoid all moves in a certain similarity class because it has a low valuation. This low valuation, in turn, may be arbitrarily small, and bear no relation to the payoffs at terminal nodes that are reached from the class. Still, consistency is maintained because the class is never reached.[5]

To avoid such equilibria we refine the notion of VE in a way that parallels the notion of sequential equilibrium. We require that the valuation v reflects possible payoffs at nodes that are not reached, much the same as beliefs in sequential equilibrium reflect possible beliefs at nodes that are not reached.

We say that a strategy σ is *positive on similarity classes* if $P^\sigma(Z(\lambda)) > 0$ for each i and each $\lambda \in \Lambda_i$. We say that σ is *positive* (or *completely mixed*) if $P^\sigma(z) > 0$ for each terminal node z. Clearly, if σ is positive it is positive on similarity classes. The following claim is obvious.

CLAIM 1. *If σ is positive on similarity classes then there exists a unique valuation v that is consistent with σ.*

We say that a valuation v_i is *sequentially consistent* with the strategy profile σ if there exists a sequence of strategy profiles $(\sigma^k)_{k=1}^\infty$ that are positive on similarity classes and such that σ^k converges to σ and v_i^k converges to v_i, where v_i^k is the unique valuation consistent with σ^k.

DEFINITION 2. A strategy profile σ is a *sequential valuation equilibrium* (SVE) if there exists a valuation profile $v = (v_i)_{i\in I}$ such that for each i,

- σ_i is optimal for v_i
- v_i is sequentially consistent with σ.

It is easy to see that sequential consistency implies consistency, and thus an SVE is also a VE.

We could possibly strengthen the notion of sequential consistency by requiring that the strategies $(\sigma^k)_{k=1}^\infty$ are not only positive on similarity classes but also completely mixed. But as the following claim shows these requirements are equivalent.

CLAIM 2. *A valuation v_i is sequentially consistent with the strategy profile σ if and only if there exists a sequence of positive strategy profiles $(\sigma^k)_{k=1}^\infty$ such that σ^k converges to σ and v_i^k converges to v_i, where v_i^k is the unique valuation consistent with σ^k.*

PROOF. Obviously if the condition in the claim is satisfied, then in particular the strategies $(\sigma^k)_{k=1}^\infty$ are positive on similarity classes, and therefore v_i is sequentially consistent with σ.

[5]This is similar in spirit to the theme developed in the notion of self-confirming equilibrium (Fudenberg and Levine 1993).

Conversely, suppose that v_i is sequentially consistent with σ, and let $(\sigma^k)_{k=1}^{\infty}$ be a sequence of strategies that are positive on similarity classes such that $\sigma^k \to \sigma$ and $v_i^k \to v_i$, where v_i^k is the unique valuation consistent with σ^k. Let ν be a positive strategy profile. For each k let γ_k be the minimum of $P^{\sigma^k}(Z(\lambda))$ over all λ ($\gamma_k > 0$ because σ^k is positive on similarity classes and there are finitely many similarity classes). Define $\hat{\sigma}^k = (1-\gamma_k 2^{-k})\sigma^k + \gamma_k 2^{-k}\nu$. Then $\hat{\sigma}^k \to \sigma$ and $\hat{v}_i^k \to v_i$, while $\hat{\sigma}^k$ is positive for each k. □

3. Motivation and interpretation

The valuation approach is closely related to the familiar notion of evaluating board positions in chess, checkers, and many other games by a few simple criteria such as the profile of pieces on each side or the position around the center (see Samuel 1959 for an early investigation of the game of checkers). Adopting the valuation approach, the set of moves leading to positions with the same features can be viewed as one similarity class. The optimality condition means that players always choose moves that lead to board positions with the highest valuations (as determined by the criteria). The consistency condition endogenizes the valuation assigned to clusters of moves that form a similarity class. This is in line with the popular view in chess that a queen is worth twice as much as rook, which we interpret as saying that the chance of winning is roughly the same on average over all board positions where the queen is replaced by two rooks (whenever applicable).[6]

But, the valuation approach is in our view broadly applicable to many interactions other than chess or checkers. When we are told that it pays to be tough, we attach a single valuation to a big cluster of moves (as there are many different contexts where one can be tough). Since being tough may have very different consequences in different contexts (in bargaining being tough may be good when the other party has no outside option but not otherwise), attaching a single valuation to being tough introduces a form of bounded rationality that the valuation approach is designed to capture. Similarly, when financial advisers suggest selling when the market price is high and buying when it is low they do not make their advice contingent on whether the market is in a bubble or not, even though this information may obviously affect the true assessment of selling or buying at the current market price. Again, attaching a single valuation to selling when the price is high or when it is low induces an oversimplification that the valuation approach is designed to capture.

More generally, the valuation approach is aimed at modeling the interaction of players in complex environments in which it is too hard to assess the strength of every single move separately. While most (if not all) of the applications considered below have a small number of nodes (designed to illustrate theoretical properties), one must bear in mind that they should be considered as complex situations from the viewpoint of the players. For example, in the stopping decision problems considered in Section 6, the

[6]The idea of a criterion obtained by linearly adding the values of pieces introduces an extra element (of additive separability) that is not present in the valuation approach. Yet, the spirit is clearly related.

potential duration of the interaction should be thought of as large from the viewpoint of the players.

4. General properties

4.1 *Existence*

Since each SVE is also a VE it is enough to prove the existence of an SVE.

Proposition 1. *Each game has at least one sequential valuation equilibrium.*

Proof. The strategy of proof is the same as that for the existence of sequential equilibria (Kreps and Wilson 1982). Consider the set Σ^ε of strategy profiles σ^ε with $\sigma_i^\varepsilon(n)(m) > \varepsilon$ for all $n \in N_i$ and $m \in M_i(n)$. For any strategy profile $\sigma^\varepsilon \in \Sigma^\varepsilon$ there exists a unique valuation $v(\sigma^\varepsilon)$ such that for each i, $v_i(\sigma^\varepsilon)$ is consistent with σ^ε. By the equations that define valuations, $v(\sigma^\varepsilon)$ depends continuously on σ^ε in Σ^ε.

We say that player i's strategy σ_i is *ε-optimal for the valuation v_i*, if for each $n \in N_i$ and $m \in M_i(n)$, $\sigma_i(n)(m) = \varepsilon$ whenever $m \notin \arg\max_{m \in M_i(n)} v_i(\lambda(m))$. Consider the correspondence that associates with each $\sigma^\varepsilon \in \Sigma^\varepsilon$ the set of all strategy profiles $\hat\sigma^\varepsilon \in \Sigma^\varepsilon$ such that for each i, $\hat\sigma_i^\varepsilon$ is ε-optimal for the valuation $v_i(\sigma^\varepsilon)$. It is easy to see that this correspondence is upper hemicontinuous with non-empty closed convex values. It follows by Kakutani's fixed-point theorem that there exists σ^ε such that for each i, σ_i^ε is ε-optimal for the valuation v_i^ε, which is the unique valuation of i consistent with σ^ε.

By compactness, there exist σ and v and a subsequence of σ^{ε_k} with $\varepsilon_k \to 0$ such that both $\sigma^{\varepsilon_k} \to \sigma$ and $v(\sigma^{\varepsilon_k}) \to v$. By continuity, σ is optimal for v and hence is a sequential valuation equilibrium. □

4.2 *The trivial similarity relations*

For the two trivial similarity relations, the largest and the smallest, the characterization of VEs and SVEs is simple enough.

Proposition 2. *(a) If for each player i all moves in $\overline{M}_i$ are similar, then every strategy profile is an SVE. (b) If for each player i no two different moves in $\overline{M}_i$ are similar, then a strategy profile is an SVE if and only if it is a subgame perfect Nash equilibrium.*

5. Link to other solution concepts

In this section we illustrate the working of the SVE concept in a variety of settings. We first consider decision problems that illustrate that the SVE concept is new and cannot be interpreted in general as a standard equilibrium even by varying the payoff and information structures of the players. We next discuss the link of the approach to sequential equilibrium in games with incomplete information, and to imperfect recall.

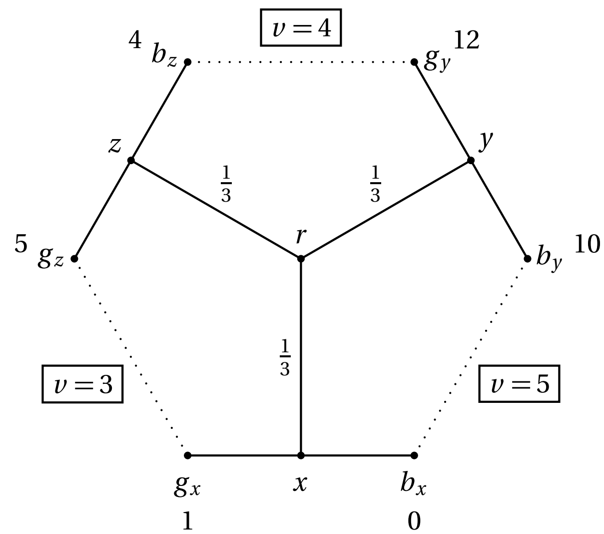

FIGURE 1. A valuation equilibrium in a decision problem

5.1 *Decision problems*

Obviously, in a decision problem the grouping of moves into similarity classes cannot benefit the agent. It can only prevent him from making optimal decisions in all circumstances. The following example illustrates a more dramatic case in which due to the similarity grouping, making the worst decision is a valuation equilibrium, while making the best one is not. This is somewhat surprising in light of the optimality requirement in valuation equilibrium.

EXAMPLE 1. The decision tree is depicted by the solid lines in Figure 1. At the root r, nature chooses one of three nodes x, y, and z with equal probability. At each of these nodes, the decision maker can choose between a good move or a bad one, where the payoff is higher in the first. The payoffs are written next to these nodes. The three dotted lines connect similar nodes. Thus, the set M is partitioned into the similarity classes $\{g_x, g_z\}$, $\{b_x, b_y\}$, and $\{g_y, b_z\}$.

The strategy σ that selects the bad move at each of the nodes x, y, and z is a VE. To see this, consider the valuation v given in the figure. Obviously, it is consistent with σ, and σ is optimal for v. Moreover, σ is also an SVE. Indeed, for each k let σ^k be the strategy for which the good move at each node has probability $1/k$ and the bad one probability $1-1/k$. The unique valuation that is consistent with σ^k is given by $v^k(\{g_x, g_z\}) = 3$, $v^k(\{b_x, b_y\}) = 5$, and $v^k(\{g_y, b_z\}) = 4(1-1/k) + 12(1/k)$. Obviously, $\sigma^k \to \sigma$, and for small enough k, σ is optimal for v^k.

Note, however, that the strategy τ that selects the good move at each node is not a valuation equilibrium. Indeed, for a valuation u to be consistent with τ, it must satisfy $u(\{g_x, g_z\}) = 3$ and $u(\{g_y, b_z\}) = 12$. But τ is not optimal for such a valuation u (consider node z). ◊

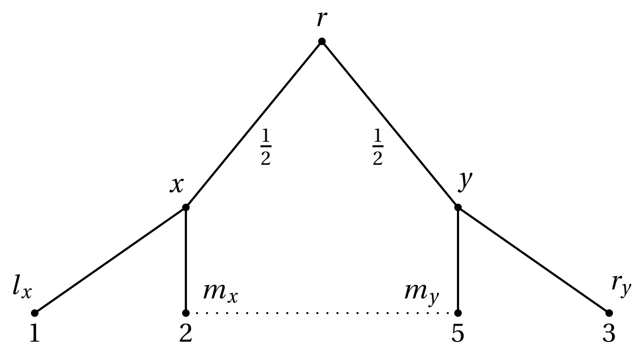

FIGURE 2. Two strict SVEs

COMMENT. In Example 1 the role of nature is crucial. In a decision problem (i.e., a game with one player) without moves of nature, any strategy σ that guarantees the maximal payoff is a sequential valuation equilibrium.

The next example illustrates the possibility of multiple equilibria in a one-agent decision problem such that in each equilibrium there is a unique strategy that is optimal for the valuation (incentives are strict). This simple example thereby illustrates that it is not possible to interpret the set of SVEs as the set of sequential equilibria of a game with modified payoff and information structures, since such modifications are incapable of producing two strict Nash equilibria in a one-agent decision problem.[7] Thus, SVE is a new solution concept that cannot be reduced to existing ones.[8]

EXAMPLE 2. The decision tree is depicted by the solid lines in Figure 2. At the root r, nature chooses each of the nodes x and y with equal probability. At node x, the decision maker can choose between nodes m_x and l_x, and at y between m_y and r_y. The dotted line connects similar nodes. Thus, the nodes are grouped according to their names. That is, the set of moves is partitioned into three similarity classes, $\mathit{Left}=\{l_x\}$, $\mathit{Middle}=\{m_x, m_y\}$, and $\mathit{Right}=\{r_y\}$.

The strategy that selects the move in *Middle* at each of the nodes x and y is an SVE. The corresponding valuation is $v(\mathit{Left})=1$, $v(\mathit{Middle})=3.5$, $v(\mathit{Right})=3$, and the strategy is clearly optimal for this valuation.

The strategy that selects the moves in *Middle* at x and *Right* at y is also an SVE. The corresponding valuation is $v(\mathit{Left})=1$, $v(\mathit{Middle})=2$, $v(\mathit{Right})=3$, and the strategy is clearly optimal for the valuation.

There is no other pure strategy SVE. ◊

[7]If the decision maker does not observe whether she is at x or y and faces the choice between m and an move other than m (in both x and y), then the only equilibrium is that the decision maker chooses m.

[8]The psychological game approach of Geanakoplos et al. (1989), in which agents directly care about the belief of others about their strategies, also allows for multiple equilibria in decision problems. Here there is no dependence of the utility on the belief, and the multiplicity comes from the dependence of the valuation on the strategy.

5.2 *Games with imperfect information*

As noted above, in Example 2, the solution of an SVE cannot be reduced to that of a sequential equilibrium in some associated game. We now show how any sequential equilibrium of a game with imperfect information can be interpreted as an SVE of a game with perfect information and a similarity relation that is the partition of moves into actions in the former game.

Formally, consider an imperfect information game defined on the tree (Z,N,r,A) with payoff function f_i. Let Υ_i be the partition of i's nodes, N_i, into the *information sets* I_i of player i. For each $I_i \in \Upsilon_i$ let $L(I_i)$ be the set of labels of arcs (or actions) that start at nodes in I_i. The set of successors (i.e., moves) of nodes in I_i can be partitioned by the labels of the arcs in $L(I_i)$ that lead to them. By partitioning the successors of the nodes in each information set we obtain a partition Λ_i of all of i's moves.

PROPOSITION 3. *Consider an extensive game with imperfect information and perfect recall defined on (Z,N,r,A) with players I, payoff functions $(f_i)_{i\in I}$, and move partitions $(\Lambda_i)_{i\in I}$. Let an assessment (σ,μ) of this game (where μ denotes a belief system) be a sequential equilibrium. Then σ is a sequential valuation equilibrium of the game defined over (Z,N,r,A) with payoff functions $(f_i)_{i\in I}$ and similarity relations $(\Lambda_i)_{i\in I}$.*

PROOF. Let (σ^k) be the sequence of strategy profiles in the definition of the sequential equilibrium (σ,μ), such that $\sigma^k > 0$ for each k and $\sigma^k \to \sigma$. For each σ^k, define the valuation v^k by (1). Because of the positivity of σ^k, v^k is defined for each λ and is consistent with σ^k.

Let v be the limit of (v^k). Since μ is the limit of the conditional probabilities of σ^k on the information sets, it follows that for each player i and each similarity class $\lambda \in \Lambda_i$, $v_i(\lambda)$ is i's expected payoff conditional on being at one of the nodes in λ. This expected payoff is computed using the probability (given by μ) of the nodes in the information set that lead to λ, and the probability of reaching each of the terminal nodes (given by σ). By the very definition of sequential rationality, σ_i is optimal for v_i. □

The converse of this proposition does not hold. That is, an SVE σ of the game with the similarity classes $(\Lambda_i)_{i\in I}$ need not be the strategy profile σ of a sequential equilibrium (σ,μ) of the original game with incomplete information. This is so because σ may not even be a strategy in the original game. Indeed, suppose that in a given information set $\{n_1,\dots,n_k\} \in \Upsilon_i$, there are two moves that maximize i's expected payoff. In that case, the strategy σ may assign one move at some nodes and the other move at other nodes.

5.3 *Imperfect recall*

The bundling of a player's moves into similarity classes resembles the bundling of a player's decision nodes into information sets in games with imperfect memory. However, the two bundlings are very different. Information sets reflect the inability of a player to distinguish between different *pasts*, while similarity classes reflect her inability to distinguish between different *futures*, and in particular the payoffs that result from

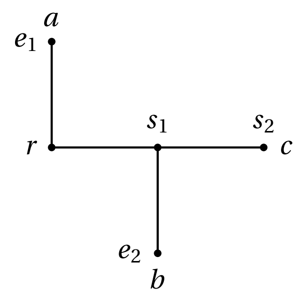

FIGURE 3. The absent-minded driver game

these futures. Thus valuation equilibrium can differ very much from imperfect recall in its modeling and predictions. To illustrate this, we consider the absent-minded driver game analyzed by Piccione and Rubinstein (1997).

At the intersection r, the driver may either exit to e_1 (turn left) and get a payoff of a or go straight ahead to the next intersection s_1. At s_1 he may either exit to e_2 (turn right) and get a payoff of b or go straight ahead to s_2 and get a payoff of c. It is assumed that b is greater than a and c so that ideally the driver wants to continue to the next intersection and exit there.

The imperfect recall approach assumes that, being unable to distinguish between the two decision nodes, the driver mixes the decision of exiting at both of them and the decision of going straight at each of them.

There is no imposition in the valuation approach that if the moves of going straight are bundled into a single similarity class the moves of exiting must be bundled too. For example, suppose the driver distinguishes between the turn left and the turn right moves but bundles together the go straight moves. His similarity partition is thus $\mathit{straight} = \{s_1, s_2\}$, $\mathit{left} = \{e_1\}$ and $\mathit{right} = \{e_2\}$. The only SVE is that the driver goes straight first and exits at the second intersection node, resulting in the maximal payoff of b. (This is, of course, very different from the outcome of the imperfect recall approach, in which the probability of exit should be the same at the two decision nodes.)

To see this we first check that exiting at the second intersection is an SVE. The corresponding valuations are $v(\mathit{straight}) = b$, $v(\mathit{left}) = a$, $v(\mathit{right}) = b$ and the assumed strategy is optimal given the valuation. No other equilibrium can arise, as sequential consistency implies that $v(\mathit{left}) = a$, $v(\mathit{right}) = b$ and it is readily verified that one cannot sustain an equilibrium with $v(\mathit{straight}) < b$ (since optimality requires that the driver turns right at the second decision node and sequential consistency in turn requires that $v(\mathit{straight}) = b$, leading to a contradiction). Thus, one should have $v(\mathit{straight}) = b$, implying that the driver goes straight first and turns right at the second intersection (if he were to mix at the second intersection this would induce a valuation for *straight* strictly less than b).

Our setup permits a bundling similar to that of imperfect recall. But, as we now observe, the consistency requirement in the valuation approach does not give rise to either of the solutions proposed in Piccione and Rubinstein (1997). To see this, assume now that the driver bundles the exit moves on the one hand ($exit = \{e_1, e_2\}$) and the straight moves on the other ($straight = \{s_1, s_2\}$).

A first observation is that our solution concept does not force the driver to exit with the same probability at nodes r and s_1. (This is similar to the observation made above for games with incomplete information.) However, in order to highlight another (more interesting) difference from the imperfect recall approach, we restrict ourselves to equilibria in which the behavioral strategies are the same at r and s_1. We assume that $c < (a+b)/2$ so that always going straight is not an equilibrium of the imperfect recall kind (whatever the approach considered in Piccione–Rubinstein).

Let α be the probability that the driver goes straight at his two decision nodes r and s_1. Under the assumed similarity partition, the valuations consistent with such a strategy should be $v(exit) = (a+\alpha b)/(1+\alpha)$ and $v(straight) = (1-\alpha)b + \alpha c$. For such a strategy to be an equilibrium, we need $v(exit) = v(straight)$ or[9]

$$\alpha^{VE} = \frac{-(b-c) + \sqrt{(b-c)^2 + 4(b-a)(b-c)}}{2(b-c)}.$$

This probability of going straight does not correspond to that arising from either the modified multi-self approach[10] (or the ex ante optimal approach), which yields $\alpha^* = (b-a)/(2(b-c))$, or the multi-self approach[11] proposed by Piccione–Rubinstein, which yields $\alpha^{PR} = (b-a)(3(b-c))$.

Note that $\alpha^{VE} > \alpha^*$ (whereas for some parameter values, say $a = 0$, $b = 4$, $c = 1$, we have $\alpha^* > \alpha^{PR}$). Interestingly, had we considered a notion of consistency so that $v_i(\lambda) = \sum_{z \in Z(\lambda)} P^\sigma(z) h(z,\lambda) f_i(z) / \left[\sum_{z \in Z(\lambda)} P^\sigma(z) h((z,\lambda)\right]$ where $h(z,\lambda)$ is the number of times the path leading to z intersects the similarity class λ, we would have obtained the probability α^* of going straight with the valuation approach. We obtain $\alpha^{VE} > \alpha^*$ because our notion of consistency makes the valuation of *straight* higher for a given probability of going straight (our notion of consistency attaches more weight to b than to c and $b > c$).

REMARK. Whether our notion of consistency or the one just described is preferable should be the subject of further investigation. It is immaterial for most of our discussion here. From a psychological viewpoint, our notion corresponds to a situation in which after each round of the learning stage resulting say in the final node z the player

[9] Observe that α lies in $(0,1)$ whenever $c < (a+b)/2$.

[10] This is obtained as a solution to

$$\frac{a+\alpha^* b}{1+\alpha^*} = \frac{1}{1+\alpha^*}[\alpha^* c + (1-\alpha^*)b] + \frac{\alpha^*}{1+\alpha^*}c.$$

[11] This is obtained as a solution to

$$\alpha^{PR} = \arg\max_\alpha \{\lambda[(1-\alpha)a + \alpha(1-\alpha)b + \alpha^2 c] + (1-\lambda)[(1-\alpha)b + \alpha c]\}$$

where $\lambda = 1/(1+\alpha^{PR})$.

only remembers that he played a move in λ at least once (as opposed to the number of times he chose a move in λ) and the payoff attached to it.

6. Application to stopping games

In this section we study games in which each player in turn can choose either to continue to play—choose a "go-ahead" move—or stop the game. We are interested in the effect of bundling all go-ahead moves into one similarity class. The motivation behind such a grouping is that it may be hard to assess the strength of moves that do not lead to an immediate end, thus forcing some grouping for such moves. We take the extreme view that all such moves are bundled into one similarity class. We start with decision problems and then move on to two-person stopping games.

6.1 *Multiple timing decision problems*

To start with, nature chooses among finitely many timing decision problems indexed by $k=1,\dots,K$. The probability that nature selects problem k is denoted p_k and is assumed to be strictly positive.[12] Each k-decision problem has the following structure. Non-terminal nodes are indexed by dates $t=1,\dots,T$. We write (k,t) for a non-terminal node at date t and decision problem k. At each (k,t), $t=1,\dots,T-1$ the decision maker may either stop or go ahead. Stopping at (k,t) leads to the terminal node $\mathit{stop}(k,t)$ with payoff a_k^t. Go ahead at (k,t) leads to the node $(k,t+1)$. At (k,T), the decision maker must choose between l_k, yielding payoff b_k, and r_k, yielding payoff c_k. We let $a_k^T=\max(b_k,c_k)$. We assume generic values of a_k^t such that $a_k^t \neq a_{k'}^{t'}$ whenever $(k,t)\neq(k',t')$. A multiple timing decision problem is depicted in Figure 4 for the case $K=2$ and $T=3$.

When the decision maker is fully rational, he should go till date $t_k=\arg\max_t a_k^t$ in problem k and stop if $t_k<T$ or choose whatever is best among l_k or r_k if $t_k=T$.

Assume now that the decision maker bundles all non-terminal moves into one similarity class called *go* while all other moves are singleton similarity classes.[13] That is, $go=\{(k,t+1)\mid k=1,\dots,K \text{ and } t=1,\dots,T-1\}$ and $\{go,\{\mathit{stop}(k,t)\}_{k,t},\{l_k\}_k,\{r_k\}_k\}$ is the similarity partition.

Our main observation follows.

Proposition 4. *Consider an SVE in the above setting. Either the decision maker stops immediately with probability 1 (i.e., he chooses $\mathit{stop}(k,1)$ with probability 1) in at least $K-1$ decision problems or there is a number k^* such that the probability that node (k^*,T) is reached is positive.*

[12] In line with our learning narrative (see Section 2.3), one may alternatively view our decision maker as facing each of the decision problems k in sequence where the frequency of occurrence of problem k coincides with p_k

[13] This idea of grouping is a bit reminiscent of the model of limited foresight (Jehiel 1995) in that the decision maker has a coarse forecast about the effect of going on. Yet, the forecast is about one own's future play in Jehiel (1995) whereas it is about the average value of going on over several k-problems here. In Jehiel's (1995) model, the decision maker never goes till node (k,T) whenever $a_k^T<a_k^{T-1}$.

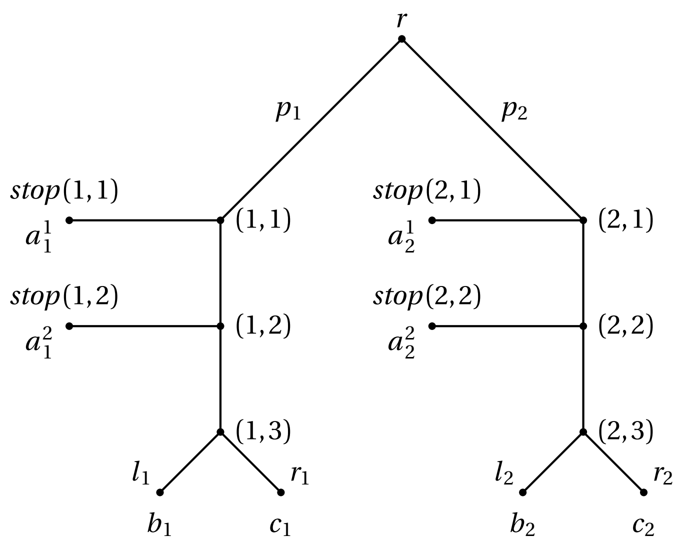

FIGURE 4. A multiple timing decision problem for $K=2$ and $T=3$

PROOF. Let $\overline{K}$ be the set of $k \in \{1,\dots,K\}$ such that the decision maker goes ahead at $(k,1)$ with positive probability and assume that $\overline{K}>1$. Since for $k \notin \overline{K}$ the player does not go ahead, it follows by consistency that $v(go)$ is computed only by the payoffs for games $k \in \overline{K}$.

Suppose now by contradiction that there is no k^* as claimed by the proposition. Then for each $k \in \overline{K}$, the decision maker reaches with probability 1 the set of nodes $\{stop(t,k)\}$ such that $t<T$. Reaching such a node, $stop(t,k)$, with positive probability, rather than going ahead with probability 1 at (t,k), means that $a_t^k \geq v(go)$. Therefore, $v(go)$ is a weighted average of payoffs $a_k^t \geq v(go)$ with $k \in \overline{K}$ and $t<T$. As $|\overline{K}|>1$, there are at least two such payoffs, and by the genericity assumption all of them are different. Thus, this weighted average must be strictly greater than $v(go)$, which is impossible. □

To illustrate Proposition 4 consider the following example. Let $k=1,2$ and $T=3$ with $p_1=p_2=0.5$ and $a_1^1=a_2^1=-1$, $a_1^2=1$, $a_2^2=1.1$, $b_1=b_2=-2$, and $c_1=c_2=-3$.[14] In this setting, there is a unique SVE. It is such that $v(go)=1$, the decision maker goes on till node $(1,3)$ with probability $\frac{1}{30}$ in problem 1 and goes to $stop(2,2)$ with probability 1 in problem 2. In problem 1, the decision maker goes till the end with positive probability because otherwise if he were to stop with probability 1 earlier, the high payoff obtained by going on in problem $k=2$ would lead the decision maker not to prefer stopping before the last node $(1,3)$ in problem 1. Observe that even though the payoffs b_1 and c_1 are quite low, the expected payoff obtained by the decision maker is 1 under the SVE.

Applied to multiple investment decision problems played in sequence (see footnote 12), Proposition 4 means that a decision maker who invests in more than one

[14]These payoffs are non-generic in the sense defined above. Yet, the same conclusion as in Proposition 4 applies to this specification as well.

enterprise and follows the valuation approach will keep investing till the very end (with positive probability) in at least one of his enterprises, no matter what payoffs are attached to long-term investments. This obviously implies some suboptimality if keeping investing till the end is a bad option in all enterprises, but the magnitude of the loss need not necessarily be large as illustrated in the previous example.

We now consider a modified version of the above timing decision problem, but instead of a finite number of problems we consider a density of problems indexed by $a \in [0,1]$ and assume that each problem is drawn according to a uniform distribution on $[0,1]$. We also simplify by assuming that $T=2$ in each problem and we assume that a is the payoff obtained by the decision maker in problem a if he stops immediately, i.e., if he reaches node $stop(a,1)$. We also let $h(a)=\max(b_a,c_a)$ be the payoff obtained in problem a if the decision maker reaches nodes $(a,2)$. The similarity structure is the same as before, with $go=\{(a,2) \mid a \in [0,1]\}$ and all other moves equal to singleton similarity classes.

The structure of the SVE is very simple. Assume $v(go)=x \in (0,1)$. The decision maker will choose $stop(a,1)$ in problem a when $a > x$ and he will choose $(a,2)$ when $x > a$ (what he does when $x=a$ is irrelevant for the fixed-point calculation). Assuming that $x \neq 0,1$, consistency then boils down to

$$v(go)=\frac{\int_0^x h(a)\,da}{x}$$

or

$$x^2=\int_0^x h(a)\,da \tag{2}$$

We wish to compare the resulting decision rule with the efficient decision rule when h is a smooth increasing function that satisfies $h'(a)>1$ for all $a \in [0,1]$. This means that the difference of payoffs obtained when the decision maker goes ahead in two situations a and b is magnified as compared with the difference of payoffs obtained after an immediate stop $|h(a)-h(b)|>|a-b|$. We assume also that $h(0)<0$, and $\int_0^1 h(a)\,da>1$ (which implies $h(1)>1$ because $h'>0$), which ensures the existence of x^V satisfying (2), and which also guarantees that there is a unique $x^{FB} \in (0,1)$ satisfying $x^{FB}=h(x^{FB})$. This x^{FB} characterizes the first-best decision rule: at the optimum, the decision maker should go when $a > x^{FB}$ (i.e., when $h(a)>a$), and he should stop when $a < x^{FB}$ (i.e., when $a > h(a)$).

Since $\int_0^x h(a)\,da/x < h(x)$ (because $h'>0$), we have $x^V < h(x^V)$, which implies that $x^V > x^{FB}$ (because $h'>1$). Thus, under the valuation approach, the decision maker goes ahead when $a < x^V$ and stops when $a > x^V$, which leads to an erroneous stopping decision i) when $a < x^{FB}$—the decision maker chooses to go when he should stop and ii) when $a > x^V$—the decision maker stops when he should go. For intermediate values of $a \in (x^{FB}, x^V)$, the optimal decision is made.

Interpreting our decision problem in terms of investment decisions, the valuation approach highlights two types of biases: sometimes returns are being taken too quickly

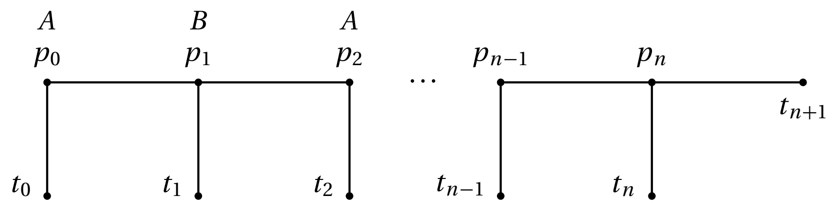

FIGURE 5. A finite horizon stopping game

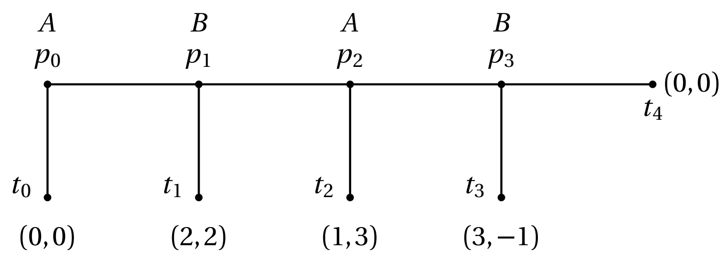

FIGURE 6. A sequential valuation equilibrium in a two-player game

in favorable markets (where it would pay staying longer) and sometimes positions are being kept too long in unfavorable markets (where it would pay liquidating earlier).

6.2 *Finite horizon stopping games*

We now consider two-player stopping games, and comment on some notable differences between the valuation approach and the analogy-based expectation approach (Jehiel 2005). Two players A and B move alternately. At all turns p_k except the last, a player decides whether to go ahead (play p_{k+1}) or to stop the game (play t_k). The player at the last turn p_n chooses between down (t_n) or straight (t_{n+1}). Figure 5 illustrates such a game.

We are interested in the effect of the grouping of non-terminal nodes. The following example shows that such a similarity grouping may sometimes allow a player to sustain threats that would otherwise not be credible, thereby making the player better off.

EXAMPLE 3. Consider the stopping game in Figure 6. In the subgame perfect Nash equilibrium of this game players A and B go ahead in the first two moves, and A stops in the third with a payoff of 1.

Assume now that player A bundles the non-terminal moves p_1 and p_3 into a single similarity class denoted *go*, while all other similarity classes are singletons.[15] Consider the following strategy profile σ: player A goes ahead at nodes p_0 and p_2; player B stops

[15]Observe that the two moves p_1 and p_3 are such that the average payoff obtained by player A over the terminal nodes in the corresponding subgames are the same, which may provide another rationale for the grouping.

at node p_1 and chooses t_4 at node p_3. To see that σ is an SVE, consider the valuations v_A and v_B that assign to each of the moves t_i the payoff of the player at this node, while $v_A(go) = 2$ and $v_B(\{p_2\}) = 0$. It is readily verified that σ_i, for $i = A, B$, is optimal for v_i. At this SVE player A's payoff is 2, which is more than he gets in the SPNE.

While the grouping of non-terminal moves may explain some non-standard behavior, as illustrated above, we note that in Rosenthal's (1982) centipede game such a grouping gives rise to the same outcome as in the standard case: players stop as soon as they can.[16]

To see this, consider the following payoff functions f_A and f_B in the stopping game depicted in Figure 5. If player X makes the choice at p_k for $0 \le k < n$, then for all $k \ge j \ge 1$, $f_X(t_{k-j}) < f_X(t_k) > f_X(t_{k+1})$. For such payoffs, the backward induction strategies stop the game at each node. Moreover, in any Nash equilibrium or correlated equilibrium player A stops at p_0.

Consider the grouping of all non-terminal moves while all terminal moves are treated separately. Thus, player A bundles $go_A = \{p_1, p_3, \ldots\}$ and player B bundles $go_B = \{p_2, p_4, \ldots\}$, while all other similarity classes are the singletons $\{t_i\}$.

There is only one SVE: players stop at all the nodes. The proof is by backward induction. Obviously this is true at p_n in which the player has no other choice. Suppose that we have shown this for all the nodes $p_n, \ldots, p_{k+1}$, and consider node p_k. Obviously, $v_X(t_k) = f_X(t_k)$. Suppose to the contrary that player X plays p_{k+1} with some positive probability. Then, either there is a positive probability that the game reaches node t_{k+1}, where by the induction hypothesis it ends, or the game never reaches p_k. In either case $v_X(go_A)$ is a convex combination of the payoffs $f_X(t_{k+1}), f_X(t_k), \ldots, f_X(t_0)$ where the weight of $f_X(t_k)$ is less than 1. Thus, $v_X(go_A) < v_X(t_k)$, and player X must choose t_k with probability 1, which is a contradiction. ◊

7. An aspiration approach to similarity classes

It is beyond the scope of this paper to propose a general theory that accounts for the emergence of similarity partitions, as many factors (outside the interaction itself) may have an influence. However, in this section we suggest a narrative that may be of relevance to the endogenizing of the similarity partitions in contexts in which players have no preconceived view about how to bundle moves.

Specifically, we look at a situation in which moves are partitioned based on their performance relative to the equilibrium payoff, thus implying an additional link between the strategy profile and the similarity partitions. We refer to the idea of aspiration level because the classification of a move in these similarity relations depends only on whether the move performs better than, worse than, or similarly to the benchmark equilibrium payoff. After formally defining the idea we suggest a learning narrative to motivate it.

[16]By contrast, the analogy-based expectation equilibrium approach, which studies an alternative form of grouping based on the idea that players have a coarse understanding of the reaction function of their opponents, explains why players need not stop immediately in the centipede game (see Jehiel 2005).

Formally, for a strategy profile σ and a node $n \in N \cup Z$, we denote by $u_i(n,\sigma)$ the expected payoff of player i in the subgame G^n with root n, with the strategy σ^n induced on G^n by σ. That is, denoting by $Z(n)$ the terminal nodes of G^n,

$$u_i(n,\sigma) = \sum_{z \in Z(n)} P^{\sigma^n}(z) f_i(z).$$

We denote the expected payoff of player i in the game $u_i(r,\sigma)$ by $u_i(\sigma)$. This expected payoff is interpreted as the aspiration level of player i induced by σ.

Given a strategy profile σ, we define for each player i the aspiration-based similarity partition $\Lambda_i(\sigma) = \{\lambda_i^+(\sigma), \lambda_i^0(\sigma), \lambda_i^-(\sigma)\}$ by

$$\begin{aligned}
\lambda_i^+(\sigma) &= \{m \in M_i \mid u_i(m,\sigma) > u_i(\sigma)\} \\
\lambda_i^0(\sigma) &= \{m \in M_i \mid u_i(m,\sigma) = u_i(\sigma)\} \\
\lambda_i^-(\sigma) &= \{m \in M_i \mid u_i(m,\sigma) < u_i(\sigma)\}.
\end{aligned}$$

Note that one or two of these three sets may be empty.

Definition 3. A strategy profile σ is an *aspiration-based sequential valuation equilibrium* (ASVE) if σ is a sequential valuation equilibrium with respect to the aspiration-based similarity partitions $\Lambda(\sigma) = (\Lambda_i(\sigma))_{i \in I}$ induced by it.

7.1 *Learning to play an ASVE*

The concept of an ASVE has a simple interpretation in terms of learning. We sketch here a learning process the (asymptotic) properties of which should be the subject of future research. Suppose that players repeatedly play game G. Each player i starts with an arbitrary grouping of moves into three similarity classes λ_i^+, λ_i^0, λ_i^- some of which may be empty, and updates them after each history. After history h_t the classes are $\lambda_i^+(h_t)$, $\lambda_i^0(h_t)$, $\lambda_i^-(h_t)$. Given history h_t, player i chooses with probability $1 - \varepsilon(h_t, i)$ at node $n \in N_i$ a move $m \in M_i(n)$ that belongs to λ_i^+ if this set is not empty. Otherwise she selects a move in λ_i^0 if it is not empty, or else a move in λ_i^-. She chooses with probability $\varepsilon(h_t, i) > 0$ any move in $M_i(n)$. We assume that $\varepsilon(h_t, i)$ goes to 0 when t tends to ∞. At each stage, players observe their payoffs and update their aspiration levels by taking the average payoff obtained over all previous stages. After a given move has been played a sufficient number of times (the number of times should be increasing with t), the average payoff resulting from the move is compared with the aspiration level. If it is sufficiently above the aspiration level the move is assigned to λ^+; if it is sufficiently below the aspiration level, it is assigned to λ^-; otherwise it is assigned to λ^0. If the strategies converge along such a learning process, they must converge to an ASVE.

Compared to our initial learning motivation for the SVE, the above narrative assumes that the similarity partition varies along the learning process (as moves may be reassigned to different similarity partitions at different rounds of the learning process). However, the change in the similarity partition is slow compared to the change in the valuations (this can be seen from our assumption that the valuations of moves are compared to the aspiration level only after the move has been played a sufficient number of times), which ensures that given a similarity relation the game has enough time to converge to a SVE of the corresponding grouping.

7.2 *Analysis*

We show that a subgame perfect Nash equilibrium is always an ASVE, thereby proving a constructive argument for why an ASVE always exists. To establish this it is useful to note that sequential consistency with σ of a valuation v_i on $\Lambda_i(\sigma)$ implies that v_i reflects the objective differences of utility in the three elements of the partition.

LEMMA 1. *Suppose that a valuation v_i on the aspiration-based similarity partition $\Lambda_i(\sigma)$ is sequentially consistent with σ. Then,*

- *if $\lambda_i^+(\sigma)\neq\emptyset$, then $v_i(\lambda_i^+(\sigma))>u_i(\sigma)$*
- *if $\lambda_i^0(\sigma)\neq\emptyset$, then $v_i(\lambda_i^0(\sigma))=u_i(\sigma)$*
- *if $\lambda_i^-(\sigma)\neq\emptyset$, then $v_i(\lambda_i^-(\sigma))<u_i(\sigma)$.*

PROOF. To see the first inequality, let $M=\{m^1,\dots,m^k\}$ be a maximal set of points in $\lambda_i^+(\sigma)$ such that each point in M is not a descendant of any other point in $\lambda_i^+(\sigma)$, and let $Z(m^j)$ be the set of terminal nodes in the subgame starting at m_j. We have $Z(\lambda_i^+(\sigma))=\cup_{j=1}^k Z(m^j)$, where the latter set is a disjoint union. Choose $\varepsilon>0$ such that $u_i(m^j,\sigma)>u_i(\sigma)+\varepsilon$ for $j=1,\dots,k$. For a strategy profile ν which is close enough to σ, $u_i(m^j,\nu)>u_i(\nu)+\varepsilon$ for $j=1,\dots,k$. Let ν be such a completely mixed strategy profile and let v_i' be i's valuation for ν. Note that for a descendant z of m^j, $P^{\nu^{m^j}}(z)=P^\nu(z)/P^\nu(Z(m^j))$. Thus,

$$
\begin{aligned}
v_i'(\lambda_i^+(\sigma)) &= \sum_{z\in Z(\lambda_i^+(\sigma))} P^\nu(z)f_i(z)/P^\nu(Z(\lambda_i^+(\sigma)))\\
&=\sum_{j=1}^k\Big[\sum_{z\in Z(m^j)} P^\nu(z)f_i(z)/P^\nu(Z(m^j))\Big]P^\nu(Z(m^j))/P^\nu(Z(\lambda_i^+(\sigma)))\\
&=\sum_{j=1}^k u_i(m^j,\nu)P^\nu(Z(m^j))/P^\nu(Z(\lambda_i^+(\sigma)))>u_i(\nu)+\varepsilon.
\end{aligned}
$$

By the sequential consistency of v_i with σ it follows that $v_i(\lambda_i^+(\sigma))\geq u_i(\sigma)+\varepsilon>u_i(\sigma)$.

The last inequality is similarly proved. To show the equality we choose a subset M of $\lambda_i^0(\sigma)$ as above. For each m^j, $u_i(m^j,\sigma)=u_i(\sigma)$. Let $\varepsilon>0$. Then for a strategy profile ν which is close enough to σ, $|u_i(m^j,\nu)-u_i(\nu)|<\varepsilon$. For a completely mixed ν and its corresponding valuation v_i' we conclude by the above equations that

$$
\begin{aligned}
|v_i'(\lambda_i^0(\sigma))-u_i(\nu)| &= \Big|\sum_{j=1}^k u_i(m^j,\nu)-u_i(\nu)P^\nu(Z(m^j))/P^\nu(Z(\lambda_i^0(\sigma)))\Big|\\
&\leq\sum_{j=1}^k |u_i(m^j,\nu)-u_i(\nu)|P^\nu(Z(m^j))/P^\nu(Z(\lambda_i^0(\sigma)))|<\varepsilon.
\end{aligned}
$$

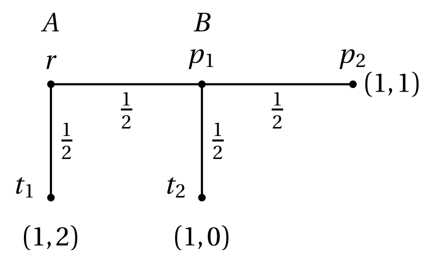

FIGURE 7. An ASVE that is not an equilibrium

Since this is true for any v close enough to σ it follows that $|v_i(\lambda_i^0(\sigma)) - u_i(\sigma)| \le \varepsilon$, and since this is true for any ε it follows that $v_i(\lambda_i^0(\sigma)) = u_i(\sigma)$. □

We can now show the following result.

PROPOSITION 5. *A subgame perfect Nash Equilibrium is an ASVE.*

PROOF. Let σ be a subgame perfect equilibrium of G. Using completely mixed strategy profiles that converge to σ we can define for each player i a valuation v_i on $\Lambda_i(\sigma)$ that is sequentially consistent with σ. At each node $n \in N_i$, σ_i selects with probability 1 nodes $m \in M_i$ that maximize $u_i(m,\sigma)$. By Lemma 1, σ selects with probability 1 nodes with the highest valuation at n. Thus, σ_i is optimal for v_i. □

Even though the subgame perfect Nash equilibrium is always an ASVE, an ASVE is not necessarily an equilibrium, as demonstrated by the game in Figure 7. Consider the strategy profile σ where player A plays t_1 and p_1 with probability $\frac{1}{2}$ each, and player B plays t_2 and p_2 with probability $\frac{1}{2}$ each. Obviously, σ is not an equilibrium. However, player B's expected payoff is $\frac{5}{4}$ and therefore $\lambda_2^-(\sigma) = \{t_2, p_2\}$. Thus, σ_2 is optimal for v_2. It is easy to see that the rest of the requirements for an ASVE are satisfied for σ.

In zero-sum games without moves of nature, the aspiration grouping results in the value of the game no matter what ASVE is considered, thus suggesting an interesting class of games in which the aspiration grouping leads to nice properties.

PROPOSITION 6. *Let σ be an ASVE of a two-person zero-sum game without moves of nature. Then the players' equilibrium payoffs in σ correspond to the value of the game.*

We prove this result as a corollary of the next result.

PROPOSITION 7. *Suppose that G is a game without moves of nature. Let ρ_i be the individually rational payoff of player i in the game G. If σ is an ASVE, then for each i, its expected payoff in G under σ, $u_i(\sigma)$, is at least ρ_i.*

PROOF. Assume to the contrary that $u_i(\sigma) < \rho_i$. We show that for each $n \in N \cup Z$, if i's individually rational payoff in the subgame G^n, $\rho_i(G^n)$, is at least ρ_i, then $u_i(n,\sigma) >$

$u_i(\sigma)$. The proof is by induction on the depth of the subgame. The claim trivially holds for $n \in Z$. Suppose now that $\rho_i(G^n) \geq \rho_i$ and the claim holds for all the subgames of G^n. If $n \in N_j$ for $j \neq i$, then it must be the case that for each $m \in M_j(n)$, $\rho_i(G^m) \geq \rho_i$. Thus by the induction hypothesis, for all $m \in M_i(n)$, $u_i(m,\sigma) > u_i(\sigma)$. Therefore also $u_i(n,\sigma) > u_i(\sigma)$. Suppose now that $n \in N_i$. Then there exists at least one $m \in M_i(n)$ such that $\rho_i(G^m) \geq \rho_i$. By the induction hypothesis, $u_i(m,\sigma) > u_i(\sigma)$. It follows that $m \in \lambda_i^+(\sigma)$. Since the latter set is not empty, and σ_i is optimal for v_i, it follows by Lemma 1 that σ_i selects nodes in $\lambda_i^+(\sigma)$ at n, with probability 1. Hence, by the definition of this set, $u_i(n,\sigma) > u_i(\sigma)$. In particular, since $\rho_i(G^r) = \rho_i$, we derive the contradiction $u_i(r,\sigma) > u_i(\sigma)$. □

REMARK. Another corollary of the above proposition is that in a decision problem without moves of nature, an ASVE is an optimal decision.

8. CONCLUDING REMARKS

We have introduced in this paper a new solution concept in which players know/learn only the average performance of playing over bundles of moves. We have suggested a learning narrative to motivate the consistency requirement imposed on equilibrium valuations. This learning narrative belongs to the family of reinforcement learning models such as those considered in AI in the tradition of Samuel (1959) (see Sutton and Barto 1998 for a recent textbook on this literature). Note that in contrast to how reinforcement learning is modeled in game theory (see Fudenberg and Levine 1998 for an exposition) our underlying reinforcement learning does not consider the reinforcement of strategies (but rather the reinforcement of similarity classes). In Jehiel and Samet (2005) we consider the case where moves rather than strategies are reinforced and we showed the convergence to the subgame perfect Nash equilibrium in extensive form games with complete information. In this paper, we have gone one step further by assuming that moves are bundled together into similarity classes and that reinforcement bears on the similarity classes rather than on the moves separately. The convergence properties of the corresponding learning models should be studied.[17]

It should be stressed that our solution concept, valuation equilibrium, assumes that the similarity classes are exogenously given and do not vary along the learning process (see however Section 7). In some cases though, it may be argued that as players learn they also adjust their way of forming similarity classes, thereby leading to potentially more complex learning dynamics. Clearly, more work is required to analyze such dynamics and their corresponding limit points.

REFERENCES

Fudenberg, Drew and David K. Levine (1993), "Self-confirming equilibrium." *Econometrica*, 61, 523–545. [168]

[17]The belief learning analog of the SVE was introduced by Jehiel (2005), who assumes that players bundle decision nodes of their opponents and form expectations only about the average play of their opponents over the bundle.

Fudenberg, Drew and David K. Levine (1998), *The Theory of Learning in Games*. MIT Press, Cambridge, Massachusetts. [184]

Geanakoplos, John, David Pearce, and Ennio Stacchetti (1989), "Psychological games and sequential rationality." *Games and Economic Behavior*, 1, 60–79. [172]

Jehiel, Philippe (1995), "Limited horizon forecast in repeated alternate games." *Journal of Economic Theory*, 67, 497–519. [176]

Jehiel, Philippe (2005), "Analogy-based expectation equilibrium." *Journal of Economic Theory*, 123, 81–104. [179, 180, 184]

Jehiel, Philippe and Dov Samet (2005), "Learning to play games in extensive form by valuation." *Journal of Economic Theory*, 124, 129–148. [184]

Kreps, David M. and Robert B. Wilson (1982), "Sequential equilibria." *Econometrica*, 50, 863–894. [165, 170]

Piccione, Michele and Ariel Rubinstein (1997), "On the interpretation of decision problems with imperfect recall." *Games and Economic Behavior*, 20, 3–24. [165, 174, 175]

Rosenthal, Robert W. (1982), "Games of perfect information, predatory pricing and the chain-store paradox." *Journal of Economic Theory*, 25, 92–100. [180]

Samuel, Arthur L. (1959), "Some studies in machine learning using the game of checkers." *IBM Journal of Research and Development*, 3, 210–229. [169, 184]

Steiner, Jakub and Colin Stewart (2006), "Learning by similarity in coordination problems." Unpublished paper, Yale University. [164]

Sutton, Richard S. and Andrew G. Barto (1998), *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, Massachusetts. [184]

Submitted 2006-1-13. Final version accepted 2007-5-7. Available online 2007-5-8.