Valuation Equilibria

Philippe Jehiel and Dov Samet

May 2003

Abstract

We introduce a new solution concept for games in extensive form with perfect information: the valuation equilibrium. The moves of each player are partitioned into similarity classes. A valuation of the player is a real valued function on the set of her similarity classes. At each node 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 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. An aspiration-based approach is also proposed, in which the similarity partitions are determined endogenously. The corresponding equilibrium is called the aspiration-based valuation equilibrium (ASVE). While the Subgame Perfect Nash Equilibrium is always an ASVE, there are other ASVE in general. But, in zero-sum two-player games without chance moves every player must get her value in any ASVE.

Key words: Game theory, bounded rationality, valuation, similarity, aspiration.

JEL numbers: C72, D81.

1 Introduction

Learning the performance of every possible move in a game in extensive form may be quite difficult if the game has many decision nodes and many moves at every decision node. An alternative that we study here is the partitioning of all possible moves of a player into sets of moves, referred to as similarity classes. Performance is attributed by the player to the similarity classes rather than the individual moves. The performance of the similarity classes of a player is expressed by a *valuation*, that is, a function which assigns a numerical value to each of the similarity classes.

We introduce two new solution concepts for games in extensive form with perfect information, based on similarity classes and their valuation: *valuation equilibrium* and *sequential valuation equilibrium.* A Valuation Equilibrium is a profile of behavioral strategies such that for some system of valuations two conditions are satisfied.

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

The consistency requirement imposes constraints only on the valuations of those 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 applies also to unreached similarity classes. Very much like the sequential equilibrium (Kreps and Wilson 1982) the stronger notion of consistency requires that the valuations of unreached similarity classes must be consistent with small perturbations of the strategy profile.

In the main part of the paper we take the sequential valuation equilibrium as our starting point and we analyze the properties of these equilibria for various similarity partitions.

We provide in subsection 2.4 a motivation for the sequential valuation equilibrium by a learning model. We introduce a simple learning process in which the game is played repeatedly, and the players update their valuations after each round according to the outcome of the round. Whenever this learning process converges, players asymptotically behave according to a sequential valuation equilibrium.[1]

[1] In Jehiel and Samet (2000) we prove the convergence of similar learning processes for maximal similarity classes.

It should be noted that in most of this paper the similarity partition of moves is given exogenously. These partitions should be thought of as being part of the description of the environment much the same as information partitions are part of the description of a game with imperfect information.[2] However, in Section 5 we study similarity classes that are determined endogenously according to a principle based on the aspiration idea.

We start, in Section 3, by making a few preliminary observations. We first show that a sequential valuation equilibrium (SVE) always exists, for any given similarity partitions. For maximal similarity partitions (i.e. no two moves belong to the same similarity class), an SVE coincides with a Subgame Perfect Nash Equilibrium. Moreover, Sequential Equilibria of games with incomplete information and perfect recall can always be represented as SVE by suitable choices of similarity partitions. But, in general an SVE need not even be a Nash Equilibrium.

Illustrations of the solution concept are introduced in Section 4. We first consider decision problems, and then move on to multi-player games. 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, thus illustrating a sharp contrast with standard notions of equilibrium. We next provide a two-player example, in which one of the players is better off in equilibrium whenever he has a coarser similarity partition. We also provide an example of complete information game in extensive form in which the SVE approach forces the players to randomize at each of their decision nodes. The example also serves to contrast the SVE approach with other approaches to the grouping of moves, in particular that of imperfect recall (Piccione and Rubinstein 1997), and that of the analogy-based expectation approach (Jehiel 2001).

In the last section of the paper, the similarity partitions are endogenized. We assume that players categorize moves 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 equilibria as *aspiration-based sequential valuation equilibrium* (ASVE). While the Subgame Perfect Nash Equilibrium is always an ASVE, other strategy profiles may be ASVE. But, in zero-sum two-player games without chance moves a player must get her value in any ASVE.

[2] From the viewpoint of learning, this amounts to assuming that players do not change their similarity relations during the learning process.

2 Valuation Equilibria

2.1 Games and strategies

Consider a finite game G with perfect information and a finite set of players I. The game is described by a tree (Z, N, r, A), where Z and N are the (finite) sets of terminal and non-terminal nodes, correspondingly, r is the root of the tree, and A the set of arcs. Elements of A are ordered pairs (n, m), where m is the immediate successor of n.

For each $i \in I$, the function $f_i : Z \longrightarrow R$ is i's payoff. The set N_i for $i \in I$ is the set of nodes in which it is i's turn to play. The sets N_i are disjoint. The *moves* of player i at node $n \in N_i$ are the nodes in $M_i(n) = \{m \mid (n, m) \in A\}$. Denote $M_i = \bigcup_{n \in N_i} M_i(n)$.

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 $M_i(n)$.

The nodes in $N \setminus \cup_{i \in I} N_i$ belong to Nature, which has a fixed strategy.

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

2.2 Similarity and valuation

Player i has a relation of *similarity* on M_i, her set of moves. We assume that it is an equivalence relationship and denote by Λ_i the partition of M_i into similarity classes. For $m \in M_i$, $\lambda(m)$ 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, if i chooses in each of her nodes, with probability 1, only those actions that belong to similarity classes with maximal valuation.

We say that the valuation v_i is *consistent* with the profile σ if for each $\lambda \in \Lambda_i$, with $P^\sigma(Z(\lambda)) > 0$, $v_i(\lambda) = \sum_{z \in Z(\lambda)} P^\sigma(z) f_i(z) / P^\sigma(Z(\lambda))$. That is, if the valuation of a similarity class λ is i's expected payoff given that (at least) one of the nodes in λ was reached. We note that if σ is completely mixed (i.e. $\sigma_i(n)(m) > 0$ for all i and $(n, m) \in A, n \in N_i$ - we write $\sigma > 0$) then there exists a unique valuation v which is consistent with σ.

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 for the "wrong" reason. Player i may avoid all actions in a certain similarity class because it has a low valuation. This low valuation, in turn, is arbitrarily small, and bears no relation to the payoffs at terminal nodes that are reached from the class. Still, consistency is maintained because the class is never reached.

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 valuation v_i is *sequentially consistent* with the strategy profile σ, if there exists a sequence of completely mixed 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.

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.

2.4 Learning processes that lead to valuation equilibria

We illustrate reinforcement learning processes, the limit points of which must be valuation, or sequential valuation equilibria. In Jehiel and Samet (2000) such a process has been studied for the case that the similarity relation is maximal (i.e., no two distinct nodes are similar). The processes here are variants of the learning model in Jehiel and Samet (2000).

Consider the infinitely repeated game of G. We denote by h an infinite history in the repeated game, and by h_t the finite history of the first t rounds in h. Assume that at the beginning of the repeated game each player has some initial valuation, and after each round she revises her valuation. After a given finite history h_t of the repeated game, the valuation of a move m of player i is her average payoff in those rounds in h_t in which she made a move similar to m. If no such move was made, then m has the initial valuation.

Assume further that the strategies of the players in the repeated game satisfy the following condition. For each finite history h_t and player i there exists $\varepsilon = \varepsilon(h_t, i)$, such that player i's strategy in G, after h_t, $\sigma_i^{h_t}$, is ε-optimal for her valuation $v_i^{h_t}$. That is, for any $m' \in M_i(n)$, if $v_i^{h_t}(m') < \max v_i^{h_t}(m) - \varepsilon$, then m' is played with probability 0. Moreover, we assume that for each i, $\varepsilon(h_t, i) \to 0$ when $t \to \infty$.

Consider now the event E that $\sigma_i^{h_t}$ converges to σ. That is, E is the set of infinite histories h such that the strategy profiles along h converge to σ. For each set of end nodes $Z(\lambda)$ that has a positive probability under σ, v^{h_t} must converge. Let v be valuation which is the limit of v^{h_t} on sets $Z(\lambda)$ with positive probability, and defined arbitrarily small on all other such sets.

Since $\sigma_i^{h_t}$ is ε-optimal for $v_i^{h_t}$, and since ε converges to 0 when $t \to \infty$, it follows that the limiting strategy σ_i is optimal for the limiting valuation v_i.

Consider now a set of end nodes in $Z(\lambda)$ that has a positive probability under σ. By the stability theorem (see Loève (1963)) the frequency of the end nodes in $Z(\lambda)$ along histories in E, converges to the conditional probability of these nodes according to σ. This shows that v is consistent with σ. Thus, the limit σ is a valuation equilibrium.

Suppose now that for each history h_t, $\sigma^{h_t} > 0$. Let E be the event that for each subtree of G the conditional probability of σ^{h_t} converges. In particular σ^{h_t} converges to a strategy profile σ. Moreover, the valuations v^{h_t} also converge on each set $Z(\lambda)$. As before, σ is optimal for v. Consider the valuation u^{h_t} which is the unique valuation consistent with σ^{h_t}. Then, using again the stability theorem we can show that u^{h_t} converges to v, which shows that σ is a sequential valuation equilibrium.

3 General Properties

3.1 Existence

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

Proposition 1 *For each game G there exists at least one sequential valuation equilibrium.*

Proof. The strategy of proof is the same as that for the existence of sequential equilibria (Kreps-Wilson 1982). Consider the set Σ^ε of strategy profiles σ^ε that satisfy $\sigma^\varepsilon > \varepsilon$. 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, $\hat{\sigma}^\varepsilon_i$ 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^ε_i is ε-optimal for the valuation v^ε_i which is the unique valuation i consistent with σ^ε.

By compactness, there are 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 it is a sequential valuation equilibrium. ∎

3.2 The trivial similarity relations

For the two trivial similarity relations, the largest and the smallest, the characterization of VE's and SVE's is rather simple.

Proposition 2 *Suppose that for each player i all the nodes in M_i are similar. Then every strategy profile is a SVE.*

Proposition 3 *Suppose that for each player i no two different nodes in M_i are similar. Then a strategy profile is SVE iff it is a Subgame Perfect Nash Equilibrium.*

For completeness, we also include the characterization of VE's, when the similarity relation is maximal, which is simple for the generic case:[3]

Proposition 4 *Suppose that for each player i no two different nodes in M_i are similar, and for every two terminal nodes $z \neq z'$, $f_i(z) \neq f_i(z')$. Then a strategy profile σ is VE iff the probability P^σ it induces over the terminal nodes assigns probability 1 to one of these nodes.*

3.3 Games with imperfect information

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 *information sets* of player i, and Λ_i the

[3] When $f_i(z) = f_i(z')$ a mixed distribution over z and z' can be sustained provided player i gets a chance to choose between the subgames containing z and z' at some decision node. In the generic case considered in the proposition it can be shown by contradiction that no randomization can occur on the equilibrium path (take the last reached node such that the behavior at that node is random).

partition of i's moves, M_i, into *actions*. Thus for any action $\lambda = \{m_1, \ldots, m_k\} \in \Lambda_i$ all the nodes in λ have different immediate predecessors, and the set of these immediate predecessors $\{n_1, \ldots, n_k\}$ is an information set in Υ_i.

Proposition 5 *Consider a game with imperfect information and perfect recall defined on (Z, N, r, A) with payoff functions $(f_i)_{i \in I}$ and action 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 (Z, N, r, A) with payoff functions $(f_i)_{i \in I}$ and similarity relations $(\Lambda_i)_{i \in I}$.*

Proof. For each player i we define a valuation v_i as follows. For an action $\lambda \in (\Lambda_i)$, $v_i(\lambda)$ is i's expected payoff when she chooses action λ, conditional on being at information set in which λ is available. This expected payoff is computed using the probability of the nodes in the information set (given by μ) and the 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 strategy profile σ is the limit of strategy profiles $\sigma^k > 0$. For each σ^k we define the valuation v^k as above (where the probability of each node in the information set is given by σ^k). The expected value of an action λ of player i is exactly $\sum_{z \in Z(\lambda)} P^{\sigma}(z) f_i(z) / P^{\sigma}(Z(\lambda))$, and therefore v^k is consistent with σ^k. Since μ is the limit of the the conditional probabilities of σ^k on information sets, it follows that $v^k \to v$. ∎

4 Illustrations

We now provide some illustrations of how the SVE concept works. We start with decision problems and then move to multi-player setups.

4.1 Decision problems

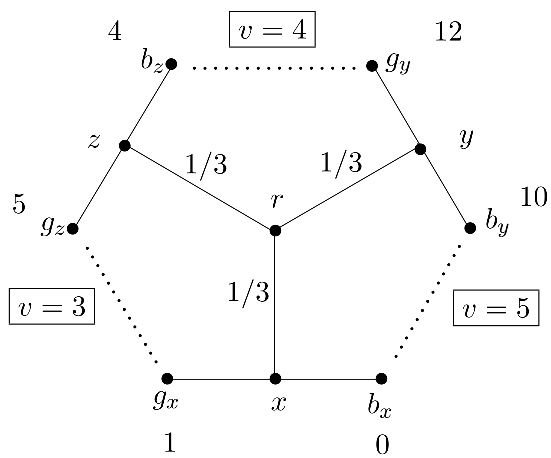
Obviously, in a decision problem the agent cannot benefit from the grouping of moves into similarity classes. It can only prevent him from making optimal decisions in all circumstances. The following example illustrates a more dramatic case in which due to similarity grouping, making the worst decision is a valuation equilibrium, while making the best one is not. It 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 action or a bad one, where the payoff is higher in the first. The payoffs are written next to these nodes. The three doted 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 at each of the nodes x, y and z the bad action 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 a SVE. Indeed, for each k let σ^k be the strategy in which the good action in each node has probability $1/k$ and the bad one $1 - 1/k$. 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 action in 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 u (consider node z).

Figure 1: A valuation equilibrium in a decision problem



In Example 1 the role of nature is crucial. In a decision problem *without* moves of nature, similarity grouping cannot be so detrimental as can easily be shown:

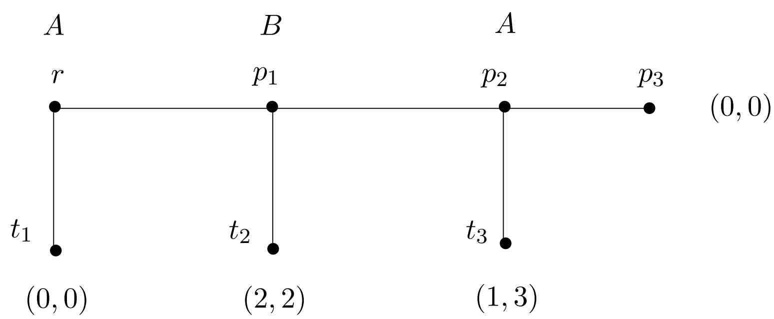
Proposition 6 *In a decision problem (i.e., a game with one player) without moves of nature, any strategy σ that guarantees the maximal payoff is a sequential valuation equilibrium.*

4.2 Multi-player setups

In decision problems, we have seen that the grouping of moves may hurt the agent. Clearly, this may also happen in multi-agent setups. But, a new observation is that grouping may sometimes help the player when there are other players around. The following example illustrates the claim.

Example 2. In the game in Figure 2 there are two players A and B. Each player in her turn can chose between two actions "take" or "pass". Nodes are labelled as in the figure and payoffs appear next to the final leaves.

Figure 2: A sequential valuation equilibrium in a two player game



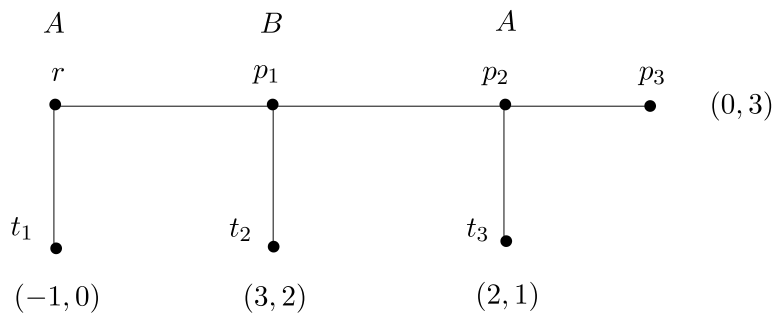
In the Subgame Perfect Nash Equilibrium of this game players A and B pass in the first two moves, and A takes in the third and her payoff is 1.

Assume now that player A bundles the moves p_1 and p_3 into a single similarity class denoted $Pass$, while all other similarity classes are singletons. Consider the following strategy profile σ: player A passes at nodes r and p_2, and B takes at node p_1. To see that σ is a SVE, consider the valuations v_A and v_B that assigns to each of the moves t_i the payoff of the appropriate player at this node, while $v_A(Pass) = 2$ and $v_B(\{p_2\}) = 0$. It is readily verified that σ_i, for $i = A, B$, is optimal for v_i. In this SVE player A's payoff is 2, which is more than what he gets in the SPNE.

The next example serves to illustrate that SVE may force randomization in circumstances in which it would not arise in standard notions of equilibrium.

Example 3. The game in Figure 3 differs from that in Figure 2 only in the payoffs. We assume that the similarity relation is the same as in the previous example. Here there is a unique SVE, and it employs mixed strategies. At the root r player A chooses p_1, while at p_2 she chooses each of t_3 and p_3 with probability 1/2. Player B chooses each of t_2 and p_2 with probability 1/2. The valuations that make this strategy profile a SVE are as follows. The valuation of each of the moves t_i is the appropriate payoff at this node. For p_3, $v_B(\{p_3\}) = 2$, which is the expected payoff of this move for B. To find $v_A(Pass)$, we note that $Z(Pass)$, the set of terminal nodes reached from $Pass$ is $\{t_2, t_3, p_3\}$. The conditional probability of these nodes are 1/2, 1/4 and 1/4 respectively. Thus $v_A(Pass) = \frac{1}{2} \cdot 3 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 0 = 2$.

Figure 3: An SVE may require mixed strategies when Nash does not



4.3 Other approaches to move bundling

We use Example 3 for comparison between valuation equilibria and some alternative ways of bundling moves.

Imperfect Recall (Piccione-Rubinstein 1997): Suppose that player A does not recall whether he is at node r or at node p_2 with a common action space $\{take, pass\}$.[4] The corresponding equilibrium (multi-self approach) is that i) Player A mixes $\frac{1}{4}take + \frac{3}{4}pass$ in his unique memory set $\{r, p_2\}$ and ii) Player B passes at node p_1. In this case, player A behaves in the same way at his two nodes. The outcome is never t_2. The other three outcomes t_1, t_3, p_3 all occur with positive probability.

Other similarity classes and Imperfect Recall: It is also instructive to analyze the same game with the valuation equilibrium approach whenever player A uses two similarity classes, $Pass = \{p_1, p_3\}$ and $Take = \{t_1, t_3\}$, and player B treats his two moves t_2 and p_2 separately. In this case, there are several SVE's. For example, one SVE is that: Player A mixes $\frac{1}{5}t_1 + \frac{4}{5}p_1$ at node r and $\frac{1}{2}t_3 + \frac{1}{2}p_3$ at node $t = 3$; Player B Passes at node $t = 2$. The valuations that support this equilibrium are $v_A(Take) = v_A(Pass) = 1$, and $v_B(\{t_2\}) = v_B(\{p_2\}) = 2$. Thus, in this case, even though player A bundles $\{p_1, p_3\}$ on the one hand and $\{t_1, t_3\}$ on the other, the situation is very different from the one analyzed above in which player A does not distinguish between nodes $t = 1$ and 3. Here, in contrast to imperfect recall, SVE does not require that player A behaves similarly at r and p_2. Moreover, with our definition of consistency, the equilibrium arising with the imperfect recall approach is *not* an SVE with our assumed similarity partitioning. It would be if the consistency criterion were defined by weighting final payoffs according to the number of times the corresponding path intersects

[4]"take" corresponds to t_1 and t_3; "pass" corresponds to p_1 and p_3.

the similarity classes.[5]

Analogy-based Expectation (Jehiel 2001): Suppose that player B only has an expectation about the average pass/take behavior of player A all over the game (i.e., player B bundles player A's two nodes r and p_2 into a single analogy class, see Jehiel 2001). The only analogy-based expectation equilibrium is such that: i) Player A *passes* at node r and *takes* at node p_2, ii) Player B *passes* at node p_1 (with the belief that player A mixes $\frac{1}{2}take + \frac{1}{2}pass$ at his two nodes r and p_2). In this case, the outcome is t_3. The behavior of player A at nodes r and p_2 is not the same and there is no mixing.

5 Aspiration-based equilibria

We now investigate similarity relations on moves that are based on the performance of these moves relative to the equilibrium payoff. We refer to the idea of aspiration because the classification of a move in these similarity relations depends only on whether the move performs better, worse or similarly to the benchmark equilibrium payoff.

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 will be 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

[5] If player A were to mix $\frac{1}{4}take + \frac{3}{4}pass$ at his two nodes and player B were to *pass* at node p_1 (as in the imperfect recall approach), the corresponding valuations for player A would be:

$$v_A(Take) = \frac{(1/4)(-1) + (3/4)(1/4)(2)}{1/4 + (3/4)(1/4)} = \frac{2}{7}$$

and

$$v_A(Pass) = \frac{(3/4)(1/4)(2)}{3/4} = \frac{1}{2}$$

Thus, $v_A(Pass) > v_A(Take)$ and player A should Pass rather than mix at his two nodes. (Hence, player A's strategy is not a best-response to his valuations.)

By contrast, if the consistency criterion incorporates the number of times pathes intersect the similarity classes, then while $v_A(Take)$ is unchanged, $v_A(Pass)$ should be replaced by

$$\frac{(3/4)(1/4)(1)(2)}{3/4(1/4 + (2)(3/4))} = \frac{2}{7}$$

and with this alternative definition, the (multi-self) imperfect recall equilibrium would be an SVE.

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 in the game G 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.*

It follows from Proposition 8 below that for each game G there exists an ASVE. To study the properties of ASVE 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 $M = \{m^1, \ldots, m^k\}$ be a maximal set of points in $\lambda_i^+(\sigma)$, such that each point in M does not have a descendant in $\lambda_i^+(\sigma)$. Then $Z(\lambda_i^+(\sigma)) = \cup_{j=1}^k Z(m^j)$, and the latter set is a disjoint union. Choose $\varepsilon > 0$ such that $u_i(m^j,\sigma) > u_i(\sigma)+\varepsilon$ for $j = 1, \ldots k$. For a strategy profile ν which is close enough to σ, also $u_i(m^j,\nu) > u_i(\nu)+\varepsilon$ for $j = 1, \ldots k$. 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,

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

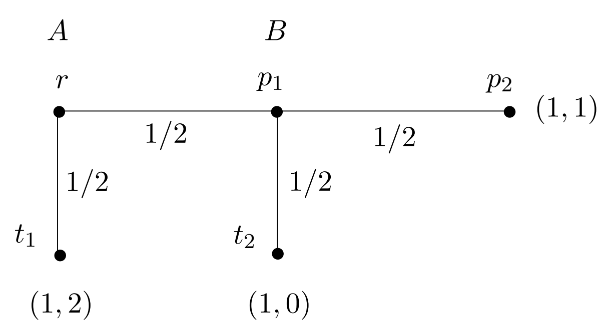
By the sequential consistency of v_i with σ it follows that $v_i(\lambda_i^+(\sigma)) \geq u_i(\sigma) + \varepsilon > u_i(\sigma)$.

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

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

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

Figure 4: An ASVE which is not an equilibrium



An ASVE is not necessarily an equilibrium, as demonstrated in the game in Figure 4. Consider the strategy profile σ where player A plays t_1 and p_1 with probability 1/2 each, and player B plays t_2 and p_2 with probability 1/2 each. Obviously, σ is not an equilibrium. However, player's B expected payoff is 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 requirement for ASVE are satisfied for σ.

Even though the above ASVE is not an equilibrium, in ASVE player B obtains no less than her individually rational payoff (which is 1). This is no coincidence.

Proposition 7 *Suppose that G is a game without moves of nature. Let ρ_i be the individual rational payoff of player i in the game G. If σ is an* ASVE, *then for each i, i's 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 individual 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. This 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 at n, with probability 1, nodes in $\lambda_i^+(\sigma)$. 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 $u_i(r, \sigma) > u_i(\sigma)$. ∎

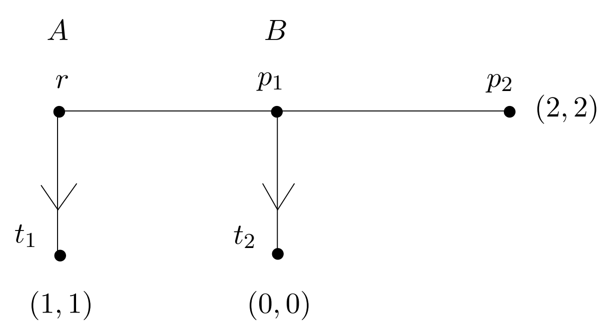
In particular for two-person zero-sum game we conclude:

Corollary 1 *If G is a two-person zero-sum game without moves of nature, and σ is an* ASVE, *then players' equilibrium payoffs correspond to the value of the game.*

The case of a single decision maker also follows immediately from Proposition 7.

Corollary 2 *If G is a decision problem without moves of nature, then an* ASVE *is an optimal decision.*

Figure 5: An equilibrium which is not an ASVE



We have seen that an ASVE is not necessarily an equilibrium. Conversely, an equilibrium of G is not necessarily an ASVE. To see this consider the game tree in Figure 5. The strategy profile $\sigma = (t_1, t_2)$ is a (non-perfect) Nash equilibrium. But $\lambda_2^-(\sigma) = \{t_1\}$ and $\lambda_2^+(\sigma) = \{p_2\}$ and the consistent valuations of these classes must be 0 and 2 correspondingly. Since σ_2 is not optimal for v_2, it follows that σ is not an ASVE.

The situation is different for Subgame Perfect Nash Equilibria.

Proposition 8 *A Subgame Perfect Nash Equilibrium of G 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 $\Lambda_i(\sigma)$ which is sequentially consistent with σ. At each node $n \in N_i$, σ_i selects with probability 1 nodes $m \in M_i$ which 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. ∎

Remark: Observe that unlike Proposition 7 and Corollaries 1 and 2, Proposition 8 holds also when there are moves of nature. This is because it relies only on Lemma 1, which holds for such cases.

Finally, we consider the following behavioral assumption:

Behavioral Assumption (BA): Whenever two successors belong to the same similarity class, they are chosen with equal probability (i.e. $\sigma_i(n)$ assigns the same probability to m, m' in $M_i(n)$ whenever $\lambda(m) = \lambda(m')$).

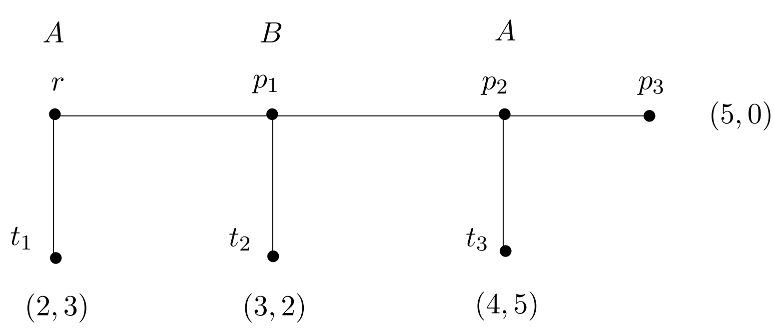
Such a behavioral assumption may reflect the idea that if two successors belong to the same similarity class nothing allows the player to distinguish their relative strength, thus suggesting that the two moves should be chosen with equal probability.

We note that while a subgame perfect equilibrium is always an ASVE, it need not in general satisfy this equal probability assumption with respect to the associated aspiration-based partitioning.

The following example illustrates this claim in the strong sense that the outcome of the unique subgame perfect equilibrium may not be supported as the outcome of an ASVE with such a behavioral requirement.

Example 4. The game in Figure 6 differs from that in Figure 3 only in the payoffs.

Figure 6: Moves in the same class must not have equal probability



The subgame perfect equilibrium results in the outcome t_2. The associated aspiration-based similarity classes are: $\lambda_A^- = \{t_1\}$, $\lambda_A^0 = \{p_1\}$, $\lambda_A^+ = \{t_3, p_3\}$ and $\lambda_B^- = \{p_2\}$, $\lambda_B^0 = \{t_2\}$. By our assumed behavioral requirement, player A should play t_3 and p_3 each with probability half at node p_2. But, this in turn implies that $v_B(p_2) = \frac{5+0}{2} = 2.5 > 2 = f_2(t_2)$. Hence, by the optimality of player B's strategy, player B should play p_2 and not t_2 at node p_1. Hence, the perfect equilibrium outcome cannot be supported as the outcome of an ASVE satisfying (BA).

Remark: In zero-sum two player games without chance moves, an ASVE satisfying (BA) exists, and it leads to the value of the game, as shown in Corollary 1.

6 Concluding remarks

This paper has introduced a new solution concept viewed as the limiting point of a learning process in which players would only try to learn the average performance of playing over bundles of moves. The underlying model (as outlined in subsection 2.4) belongs to the family of reinforcement learning models such as the ones 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 a textbook on this) our underlying reinforcement learning does not consider the reinforcement of strategies (but rather the reinforcement of similarity classes). In Jehiel and Samet (2000) we considered 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 went 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.

Clearly, it remains to analyze further how players bundle moves into similarity classes, but the present paper should be suggestive enough that this line of thought leads to considerations previously unexplored in game theory.

References

[1] Fudenberg, D. and D. Levine (1998): *The Theory of Learning in Games,* The MIT Press.

[2] Jehiel, P. (2001): 'Analogy-based Expectation Equilibrium,' mimeo CERAS and UCL.

[3] Jehiel, P. and D. Samet (2000): 'Learning to Play Games in Extensive Form by Valuation,' mimeo Tel Aviv, CERAS and UCL.

[4] Kreps, D. and R. Wilson (1982): 'Sequential Equilibria,' *Econometrica* **50**: 863-894.

[5] Loève, M. (1963): *Probability Theory*, D. Van Nostrand, Third ed.

[6] Piccione, M. and A. Rubinstein (1997): 'On the Interpretation of Decision Problems with Imperfect Recall,' *Games and Economic Behavior* **20**: 3-24.

[7] Samuel, A. L. (1959): 'Some Studies in Machine Learning Using the Game of Checkers,' *IBM J. Res. and Devel.* **3**: 210-229.

[8] Sutton, R. S. and A. G. Barto (1998): *Reinforcement Learning: An Introduction,* The MIT Press.