A theory on games with self-blind randomization

Junxing Wang, IIIS, Tsinghua University, China
Weiyi Chen, IIIS, Tsinghua University, China
Pingzhong Tang, IIIS, Tsinghua University, China

This paper puts forward a theory on a new type of games where players have an option to play a randomized strategy without observing the realized action. We propose two types of definitions that captures this idea. For each definition, we give a general method for equilibrium analysis and give an example where using this type of strategy increases one’s utility.

Categories and Subject Descriptors: I.2.11 [Distributed Artificial Intelligence]: Multiagent systems; J.4 [Social and Behavioral Sciences]: Economics

General Terms: Economics, Theory

Additional Key Words and Phrases: game theory, games in extensive form, self-blind randomization

1. INTRODUCTION

One of the most central concepts in game theory is mixed strategy. Following Nash [Nash 1951], a mixed strategy of a game, is defined to be a distribution over a set of pure strategies of the game. To carry out a mixed strategy in the original (pure-strategy) game, a player samples from the distribution specified by the mixed strategy, observes the realized pure strategy and then plays it accordingly.

An important assumption, yet, to our best knowledge, always taken for granted, is that the sampling result is perfectly observable by the play taker. In this paper, we relax of this assumption, by allowing the sampling result to be unobservable by the play taker. We coin the resulting strategy self-blind mixed strategy. To implement such a strategy, a player simply turns to a device which carries out an equiprobable mixed-strategy on the player’s behalf, without informing the player actual pure strategy played.

Depends on the extend to which such strategy can inform others, we have two possible definitions of self-blindness. In the first definition, the device that acts on behalf of the play taker announce to others the actual action realized but not the underlying distribution from which the action is drawn. So, the effects such self-blindness have on others are exactly the same as an equiprobable mixed-strategy: others only form “belief” about the distribution, without knowing the actual one based on which the device operates. In the second definition, the device announces to others not only the realized action but also the distribution from which it makes the draw. As a result, the distribution is now a piece of “knowledge” to others: as long as the strategy is taken, others know for sure the play taker is committed to this distribution. In this paper, we restrict attentions on mindless device: the device, if it reports anything at all, always reports truthfully.

For the first definition, unfortunately, self-blindness does not make any difference in one-shot normal-form game. A player, regardless of being able to observe the action played, cannot utilize this information later to change the outcome of the game. In fact, in any normal-form game, playing a self-blind randomized strategy is outcome-equivalent to playing an equiprobable mixed strategy. For the second definition, self-blindness turns a normal form game into a stackelberg game where one player has the commitment power. The problem, aka. “computation of the optimal strategy to commit to”, is thoroughly studied by Conitzer and Sandholm [2006].

In contrast, for the first definition, in extensive-form games, self-blindness can make an essential distinction. A player who plays such a strategy can place herself in a
coarser information partition than the one generated by an equiprobable mixed strategy, due to the fact that she has less information on which action she has taken. More importantly, her ignorance of her own action is credible to other players since the randomization is carried out by a device. This ignorance and the consequent trusts gained from other players, enable the player to make a credible commitment/threat that would otherwise be incredible. It is well known that the ability to commit/threat never hurt, often increase the player’s utility (cf. e.g., [Stengel and Zamir 2004]).

For the second definition of self-blindness, in extensive-form game, we find surprisingly that, the current game theory does not suffice to capture the information flow of the game: different (yet both reasonable) interpretations of the same game and same strategy profile can lead to different information partitions. To this end, we put forward an extended theory using the so-called time-stamp technique, which essentially helps fix the flow of the game. When a game has a reasonable time-stamping, all interpretations leads to the same information partition and the game is well-defined. We give a detail procedure for calculating a Nash equilibrium in the new game.

From now on, we thus restrict attentions on self-blindness for extensive form game. We next discuss relevance of self-blind randomization to several existing notions.

2. RELATED WORK

In the previous section, we have already discussed differences between mixed strategy and self-blind mixed strategy. Here, we further discuss several existing notions. We only compare these notions with the first definition (where the device only informs the realized the action without informing anything related to the distribution) because we believe the second definition is so distinct that one can immediately tell the difference between it and other notions (except for the case of normal form game, which coincide with what is studied in Conitzer and Sandholm [2006], as mentioned).

2.1. Extensive games with imperfect recall

One of the appealing features of self-blindness is to change the information partition of an extensive form game. As shown in our previous example, it can create an extensive games with imperfect recall from one with perfect recall [Kuhn et al. 1950]. However, the difference between self-blindness and imperfect recall is essential. Games with imperfect recall is a class of games where player cannot remember information he/she knew in the past. Here, the ability to remember/forget is the rule of the game — the game tree is so fixed that there is nothing the player can do. In contrast, self-blind randomization is a strategy, whose consequence sometimes creates a game with imperfect recall. In other words, in a game with imperfect recall, when a player makes a deviation, he compares his utilities in different outcomes (leaf nodes) of the same game; while in our game with self-blind randomization, when a player makes a deviation, he compares his utilities in two different games.

2.2. Markov decision processes and stochastic games

It is also natural to relate self-blindness to Markov decision process (MDP) and its multi-agent generalization, namely, stochastic game [Shapley 1953]. In a stochastic game, there are several states where the game progresses from one state to others stochastically, determined by an joint action profile (In MDP, the transition is determined by a single action). States after transition might be partially observable, leading to partially observable stochastic game (POSG) [Emery-montemerlo et al. 2004] (partially observable MDP (POMDP) respectively). It might seem that our game can be modeled as a POSG, where station transition after a self-blind randomization is unobservable. However, there are at least three differences between POSG and our game. First, there is an intermediate payoff for each player for each state, while we only de-
fine payoff when the game terminates (i.e., the absorbing states). Second, in POSG, the actions are often carried out simultaneously among all players while we do not have this constraint. Third, the state transition probability, as well as the observability of states, are part of the rules of the game, which is pre-specified and cannot be controlled by a single player, while in a game with self-blind randomization, one can flexibly change the transition probability by simply changing the probability of randomization. Similarly for observability of states. The first and second differences are not essential: the intermediate payoff can be set as 0 for internal nodes and simultaneously move can be simulated by adding a dummy move for each inactive player for each sequential move. The third can be regarded as an essential difference, which resembles the difference mentioned in the previous subsection.

2.3. Stackelberg model

The Stackelberg model is an two-round extensive-form game where a leader first commits to a strategy, followed by a follower who responds optimally. A common usage of this model is that, for some simultaneous-move game, if one player (leader) can commit to a strategy, he/she can turn the original game into a Stackelberg game and increase his/her utility. We have already discussed its relevance to self-blindness in the introduction: it is exactly the consequence of applying the second definition of self-blindness to a normal-form game.

2.4. More information may hurt

This is a well-known observation, but often stated in the context of games of incomplete information (namely, Bayesian game) where a person holding less/no information on the other player’s type can sometimes end up with higher utility (cf. eg. Page 28, [Osborne and Rubinstein 1994] and Page 281, [Osborne 2001]). Here, we provide another justification of this claim, but from a different angle: a player can now manipulate his information partition in a credible way and increase his utility. So, here, by “less information”, we mean a coarser information partition.

3. PRELIMINARY

We adapt notations from Osborne and Rubinstein [1994], where a game in extensive-form is defined to be a tuple \( \langle N, H, P, f_c, (I_i)_{i \in N}, (u_i)_{i \in N} \rangle \), where

— \( N \) is a finite set of players.
— \( H \) is a set of history sequences (finite or infinite) inductively defined as follows.
  — The empty sequence \( \varnothing \) is a member of \( H \).
  — If \( (a^k)_{k=1, \ldots, K} \in H \), then for any \( L < K \), \( (a^k)_{k=1, \ldots, L} \in H \).
  — If an infinite sequence \( (a^k)_{k=1, \ldots} \) satisfies that \( (a^k)_{k=1, \ldots, L} \in H \) for every positive integer \( L \), then \( (a^k)_{k=1, \ldots} \in H \).
— A history \( (a^k)_{k=1, \ldots, K} \in H \) is terminal if it is infinite or if there is no \( a^{K+1} \) such that \( (a^k)_{k=1, \ldots, K+1} \in H \). The set of actions available after the nonterminal history \( h \) is denoted \( A(h) = \{ a : (h, a) \in H \} \), As usual, the set of terminal histories is denoted \( Z \).
— \( P \) is a player function that assigns to each nonterminal history (each member of \( H \setminus Z \)) a member of \( N \cup \{ c \} \). \( P(h) \) is the player who takes an action after the history \( h \). If \( P(h) = c \) then chance determines the action taken after the history \( h \).
— \( f_c \) is a function that associates with every history \( h \) for which \( P(h) = c \) a probability measure \( f_c(\cdot|h) \) on \( A(h) \), where each such probability measures is independent of every other such measures. \( f_c(a|h) \) is the probability that \( a \) occurs after the history \( h \).

I_i is a partition of \( \{ h \in H : P(h) = i \} \) that satisfies the property that \( A(h) = A(h') \) whenever \( h \) and \( h' \) are in the same member of the partition, for each player \( i \in N \). For \( I_i \in I_i \), we denote by \( A(I_i) \) the set \( A(h) \) and by \( P(I_i) \) the player \( P(h) \) for any \( h \in I_i \). (\( I_i \) is the information partition of player \( i \); a set \( I_i \in I_i \) is an information set of player \( i \).)

For each \( i \in N \), \( u_i \) is the utility function of player \( i \) over \( Z \). As usual, \( u_i \) can be extended to a lottery over \( Z \) using expectations.

4. THE FIRST TYPE SELF-BLINDNESS

In this section, we study the relative easier definition of self-blindness. As mentioned, the device that acts on the player’s behalf announces the realized action to all the other players except the player herself. The device keep as a secret the distribution from which the action is sampled.

Formally, an extensive-form game with (first type) self-blindness is defined below. All the definitions are the same except the definition of Information partition. To define Information partition, we introduce the concept of equivalence relation as follows.

An equivalence relation \( \approx_i \) on histories over \( H \) (the equivalent relation of player \( i \)). We say \( h_1 \approx_i h_2 \) where \( h_1 = (a^k_1)_{k=1, \ldots, L_1}, h_2 = (a^k_2)_{k=1, \ldots, L_2} \in H \), if it satisfies one of the following three properties.

- (1) The realization of \( h_1 \) and \( h_2 \) are in the same information set of \( G \). Denote \( h|k \) the length \( k \) prefix of history \( h \).
  (a) \( \forall k < L, A(h_1|k) = A(h_2|k) \)
  (b) \( \forall k < L, a^{k+1}_1 \neq a^{k+1}_2, P(h_1|k) = P(h_2|k) = i \)

- (2) \( h_1 \) and \( h_2 \) are in the information set caused by self-blindness.
  (a) \( \exists I, h_1, h_2 \in I \)
  (b) \( \forall k \leq L_1, (h_1|k) =_i (h_2|k) \)

- (3) Transitivity.
  \( \exists h_3 \in H, h_1 \approx_i h_3, h_2 \approx_i h_3 \)

An information partition \( I'_i \) of \( \{ h \in H : P(h) = i \} \) such that each member of the partition \( I'_i \in I'_i \) is a set of equivalent histories, say \( I'_i(h) = \{ h'|h' \approx_{P(h)} h \} \).

As one can see, the resulting game is still a game in extensive form with imperfect information. The standard equilibrium computation techniques applies (such as convert it to an equivalent normal form game). We simply give an example to illustrate the benefit from applying self-blindness.
For example, for the game shown in the Fig above. Denote the histories on level 3 by \{h_1, h_2, h_3, h_4\}. In game without self-blind randomization, the Nash Equilibrium returned by Backward Induction is RRL where the utility is (2, 2). If Player 1 plays the first type self-blind randomization on level 1, Player 1's information set would be \{(h_1, h_3),\{h_2, h_4\}\}.

Consider Player 1's action in the information set \{h_1, h_3\}, undoubtedly he would play L because utilities of playing L are larger than playing R wherever he is (at history \(h_1\), the utility of playing is larger than R; and at history \(h_3\), playing L is better than playing R). But in the information set \{h_2, h_4\}, he cannot decide because the utility relation between playing L and playing R is unknown (at history \(h_3\), playing L yields -1, which is smaller than playing R at 0; but at history \(h_4\), playing L is better than playing R at -2).

Consider Player 2’s action on level 2, undoubtedly he would play R on the left because utilities of playing R are not smaller than playing L whichever Player 1 plays on level 3 (if Player 1 plays L at history \(h_2\), playing R better than L; and if Player 1 plays R at history \(h_4\), playing R equals L). But on the right, Player 2 cannot decide because the utility relation between playing L and R is unknown.

With these analyses, we can convert the tree into a normal form game as follows, where rows are Player 1’s actions and columns are Player 2’s actions.

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>LL</td>
<td>(-1, 1)</td>
<td>(-1, 1)</td>
</tr>
<tr>
<td>LR</td>
<td>(0, -2)</td>
<td>(0, -2)</td>
</tr>
<tr>
<td>RL</td>
<td>(7, -1)</td>
<td>(2, 2)</td>
</tr>
<tr>
<td>RR</td>
<td>(7, -1)</td>
<td>(0, -2)</td>
</tr>
</tbody>
</table>

There are two Nash Equilibriums in the normal game form above, which are (RR, L) and (RL, R) where the utilities are (7, -1) and (2, 2) respectively.

5. THE SECOND TYPE SELF-BLINDNESS

In this section, we study the relative more complex definition of self-blindness. In this type, the device that acts on the player’s behalf announces both the realized action to all the other players except the player herself and the distribution from which the action is sampled.

5.1. Some observations

A seemingly straightforward way to define the concept of self-blind randomization is as follows.

— Think of it as a lottery \(\vec{p}\) over the set \(A(h)\) of available actions of the player
— Turn to a device that generates an action \(a \in A(h)\) according to the probability distribution \(\vec{p}\) on the player’s behalf
— The device informs everyone else except the play taker of the actual action \(a\) played. In other words, everyone else is clear with the history path after the action except the play taker.
— As a result, this particular action might create new information sets when the player needs to make decision later. Consider the following example, Take a look at Fig 5.4, if Bob uses self-blind randomization \(b \in A^*(h)\) on level 1, denote the histories in level 3 as \{h_1, h_2, h_3, h_4\}, Bob’s information set becomes \{(h_1, h_3),\{h_2, h_4\}\}.
In general, if the original game comes with an information set partition $I_0$, and let the information set generated by a self-blind randomization $b$ be $I_b$. It is straightforward to define the result information partition as the finest common coarsening partition of $I_0$ and $I_b$ (the finest partition $I_f$ such that any set in $I_0$ or $I_b$ is a subset of some set in $I_f$). In other words, the result information partition captures the player’s ignorance caused by the original structure of the game as well as an earlier self-blind randomization. If there are several such self-blind randomization, the resulting information partition can be defined similarly.

As one can imagine, the original information set of the game is arbitrarily given. Consequently, the resulting information partition may group several histories with drastically different forms in the same information set. Take a look at the following example,

In this example, the Game Tree is shown as the picture above. Denote the histories on level 4 as $\{h_1, h_2, h_3, h_4\}$. Player 1’s information set is $\{\{h_1, h_2\}, \{h_3, h_4\}\}$, which is given originally. The game flow is as follows.

— Time 1: Player 1 plays L or R;
— Time 2: Player 2 plays L or R if the history is $h = (R)$
— Time 3: Player 2 plays L or R if the history is $h = (L)$ or $h = (R, L)$

Player 1’s strategy is playing self-blind randomization on level 1. Player 2’s strategy is playing pure L on level 2, self-blind randomization at node 1, level 3, and self-blind randomization at node 2, level 3.
Consider Player 1's information at node 3, level 4. His knowledge is the self-blind probabilities published at former time and the information set he is located. Note that Player 1 has no information about specific actions in histories since he is in the information set \(\{h_1, h_3\}\); otherwise Player 1 is able to distinguish himself in the information set \(\{h_1, h_3\}\) in game without self-blind randomizations, which is a contradiction. In other words, Player 1’s information at node 3, level 4 is self-blind probability at Time 1, self-blind probability at Time 3 and the information set \(\{h_1, h_3\}\).

Therefore, Player 1 cannot distinguish himself between \(h_1\) and \(h_3\) because he is not sure whether the self-blind probability at Time 3 is played by Node 1, Layer 3 or Node 2, Layer 3.

In the second example, the Game Tree is shown above. Denote the first two histories on level 3 as \(h_1, h_2\) and the histories on level 4 as \(h_3, h_4\). Player 1’s information set is \(\{\{h_1, h_2\}, \{h_3, h_4\}\}\), which is given originally. The game flow is as follows.

— Time 1: Player 1 plays L or R;
— Time 2: Player 2 plays L or R
— Time 3: Player 2 plays L or R if the history realization is \(h = (R, L)\)

Player 1’s strategy is playing self-blind randomization on level 1;
Player 2’s strategy is playing self-blind randomization at Node 1, level 2, pure L at Node 2, level 2 and self-blind randomization at Node 3, level 3;

Consider Player 1’s information at Node 3, level 4. It is self-blind probability at Time 1, self-blind probability at Time 3 and the information set \(\{h_1, h_3\}\). Therefore, Player 1 is able to distinguish himself between \(h_1\) and \(h_3\) because there is only one Player 2’s node at Time 3. If he receives Player 2’s self-blind probability at Time 3, the history is \(h_3\); otherwise he receives Player 2’s self-blind probability at Time 2, the history is \(h_1\).

Compare the two examples above. We happen to find that their tree structures and players’ strategies are the same, but the result information partitions are different because of the concept of time, which is “Time” in the example. Since the finest common coarsening can only return unique information partition, it cannot be right in this case. Consequently, it is necessary for us to introduce time stamp in self-blind game.

5.2. Formal definition

An extensive-form game with (second type) self-blindness is defined on a standard extensive-form game \(G = (N, H, P, f_c, (I_i)_{i \in N}, (u_i)_{i \in N})\) as follows,
A time-stamping over $H$ is a strict total order (a total order with strict part)\(^1\) over $H$. Formally, we have one and only one of the relations in $(>_{i}, <_{i}, =_{i})$ satisfied between any two histories in $H$. Equivalently, there exists a function $f$ that maps a real value to histories in $H$ satisfying $h_{1} >_{i} h_{2} \Leftrightarrow f(h_{1}) > f(h_{2})$. We call $f$ a possible time-stamp over $H$.

A time-stamping is valid if the following two properties are satisfied.

Any history is greater than its prefix. That is, For $h_{L} = (a^{k})_{k=1,2,\ldots,L} \in H$ and for any $L' < L$, $h_{L'} = (a^{k})_{k=1,2,\ldots,L'}$, we have $h_{L'} <_{i} h_{L}$.

Indifferent between two histories implies both the active players and available action sets are the same at these two histories. For $h_{1}, h_{2} \in H$ so that $h_{1} =_{i} h_{2}$, $A(h_{1}) = A(h_{2})$ and $P(h_{1}) = P(h_{2})$.

A self-blindability function $\tau$ that maps each nonterminal histories in the set $H \setminus Z$ to boolean value. It specifies at which history self-blindness is allowed. It has the following two properties:

- $P(h) = c, \tau(h) = 0$. That is, for chance move, self-blindness is forbidden.
- For any two histories with the same time-stamps $h_{1} =_{i} h_{2} \in H$, $\tau(h_{1}) = \tau(h_{2})$.

For each action set $\Sigma$, define the set of self-blind actions $\rho(\Sigma)$ to be the set of all lotteries over $\Sigma$. A member $\rho(\Sigma)$ is a self-blind action that assigns $\rho(a)$ to each action $a \in \Sigma$. Further define $\rho^{+}(\Sigma) = \rho(\Sigma) \times \Sigma$, a member of $\rho^{+}(\Sigma)$ is a pair that denotes a self-blind action and its realization. These are the two things the device announces to others. Define $\rho^{*}(\Sigma) = \Sigma \cup \rho^{+}(\Sigma)$, i.e. the set of all possible actions including self-blindness.

A realization function $r$ defined as follows: $r(a) = a$ for all $a \in \Sigma$ and $r((p,a)) = a$ where $(p,a) \in \rho^{+}(\Sigma)$.

Denote $H^{*}$ the set of history sequences (now contains self-blindness and its realization). A function $r$ assigns each new history a original history. The set $A^{*}(h)$ of actions available at nonterminal history. The three terms are defined inductively and jointly as follows.

Basic Step:

- The empty sequence $\phi \in H^{*}$.
- $r(\phi) = \phi$.

Induction Step:

- For any $h \in H^{*}$, if $\tau(r(h)) = 1$, then $A^{*}(h) = \rho^{*}(A(r(h)))$; if $\tau(r(h)) = 0$, $A^{*}(h) = A(r(h))$.
- For any $h \in H^{*}$ such that $r(h) \in H \setminus Z$, we have $(h,b) \in H^{*}$, where $b \in A^{*}(h)$.
- Denote $h[i] = (b^{k})_{k=1,\ldots,i} \in H^{*}$ the prefix of a history $h = (b^{k})_{k=1,\ldots,L} \in H^{*}$. We have $r(h[i]) = (r(h)(L - 1)), r(b^{k})$.
- A history $(b^{k})_{k=1,\ldots,K} \in H^{*}$ is terminal if it is infinite or if there is no $b^{K+1}$ such that $(b^{k})_{k=1,\ldots,K+1} \in H^{*}$. The set of terminal histories is denoted $Z^{*}$.
- At any $h \in H^{*} \setminus Z^{*}$, with $r$ defined, we define the following notation: $A(h) = A(r(h))$, player function $P(h) = P(r(h))$, the self-blindability function $\tau(h) = \tau(r(h))$ and time-stamping $h_{1} >_{i} h_{2} \Leftrightarrow r(h_{1}) > r(h_{2})$ for all $h_{1}, h_{2} \in H^{*}$.

An equivalence relation $\cong_{i}$ on histories over $H^{*}$ (the equivalent relation of player $i$). We say $h_{1} \cong_{i} h_{2}$ where $h_{1} = (b^{k})_{k=1,\ldots,L_{1}}, h_{2} = (b^{k})_{k=1,\ldots,L_{2}} \in H^{*}$, if it satisfies one of the following three properties.

- (1) The realization of $h_{1}$ and $h_{2}$ are in the same information set of $G$.
- (a) $\exists I \in I_{i}, r(h_{1}), r(h_{2}) \in I$

---

\(^{1}\)A binary relation that is antisymmetric, transitive and total and a strict part, defined as asymmetric sub-relation

(b) For $i = 1, 2$:
\[ \forall k_1 < L_i, b_{i+1}^{k_1} = (p, a_1), p \in \rho(A(h_i|k_1.)) \]
\[ \exists k_2 < L_{i-1}, (h_i|k_1.) = (h_{i-1}|k_2.), b_{i+2}^{k_2} = (p, a_2), a_2 \in A(h_{i-1}|k_2.) \]

(c) $h_1 = h_2$

— (2) $h_1$ and $h_2$ are in the information set caused by self-blindness.

(a) $L_1 = L_2$

(b) $\forall k \leq L_1, (h_1|k.) = (h_2|k.)$

(c) $\exists k < L_3 - i, (h_1|k.) = t(h_3 - i|k_2.), b_{k+1}^{i+2} = (p, a_2), a_2 \in A(h_3 - i|k_2.)$

— Transitivity.

$\exists h_3 \in H^*, h_1 \cong_i h_3, h_2 \cong_i h_3$

— A information partition $I^*_i$ of $\{h \in H^* : P(r(h)) = i\}$ such that each member of the partition $I^*_i \in I^*_i$ is a set of equivalent histories, say $I^*_i(h) = \{h'|h' \cong_i h\}$

5.3. Strategy and Nash equilibrium

— Strategy Definition

The set of pure strategies of player $i \in N$ in a self-blind game is $\Sigma_{i,pure} = \{\sigma_{i,pure} = (a_I)_{I \in I^*_i} | a_I \in A^*(I)\}$

A mixed strategy $\sigma_{i,mixed}$ of player $i \in N$ in a self-blind game is a probability measure over the set of player $i$’s strategies. The set of mixed strategies is $\Sigma_{i,mixed} = \rho(\Sigma_{i,pure})$.

A behavioral strategy of player $i \in N$ is a collection $\{\beta_I\}_{I \in I^*_i}$ of independent probability measures, where $\{\beta_I\}_{I \in I^*_i}$ is a probability measure over $A^*(I)$. The set of behavioral strategies is $\Sigma_{i,behavioral} = \{\beta_{i,behavioral} = (\beta_I)_{I \in I^*_i} | \beta_I \in \rho(A^*(I))\}$.

— Nash Equilibrium Definition

Nash Equilibrium in Strategies of a self-blind game is a profile $\sigma^*$ of strategies with the property that for every player $i \in N$ we have

\[ u(\sigma^*_i, \sigma^*_i) > u(\sigma^*_i, \sigma_i) \]

for every strategy $\sigma_i$ of player $i$.

— Nash Equilibrium Computation on 3 layer binary tree

---

The traditional history set is $H = \{ \varphi, a, b, ac, ad, bc, bd, ace, acf, ade, adf, bce, bcf, bde, bdf \}$. We denote the histories on level 3 by $(h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8)$. Then the player function $P(h) = 1$ when history $h \in \{ \varphi, ac, ad, bc, bd \}$ and $P(h) = 2$ when history $h \in \{ a, b \}$.

The information partition is $\pi = \{ \{ h_1 \}, \{ h_2 \}, \{ h_3 \}, \{ h_4 \}, \{ h_5 \}, \{ h_6 \}, \{ h_7 \}, \{ h_8 \} \}$. Time stamp on each node is the number of history actions. Let $\tau(\varphi) = true$ and $\tau(h)_{h \neq \varphi} = false$. The information sets caused by self-blind randomization for player 1 is $\{ \{ ac, bc \}, \{ ad, bd \} \} \times \{ p_\varphi | p_\varphi \in [0, 1] \} \in I_1$.

---

**Analysis**

Denote $pb + (1-p)a$ by $p \begin{array}{c} a \\ b \end{array}$.

Assume Player 1’s strategy is playing self-blind randomization $(p_\varphi, 1-p_\varphi)$ on $h = \varphi$, playing mixed strategy $(p_1, 1-p_1)$ when $h \in \{ ac, bc \}$ and playing mixed strategy $(p_2, 1-p_2)$ when $h \in \{ ad, bd \}$; Player 2’s strategy is playing mixed strategy $(q_1, 1-q_1)$ when $h = a$ and playing $(q_2, 1-q_2)$ when $h = b$.

At first we fix $p_\varphi$, Player 1’s mixed strategy $(p_1, p_2)$ and analyze mixed Player 2’s strategy $(q_1, q_2)$. To maximize its utility, Player 2 would set $(q_1, q_2)$ by

$$
(q_1^*, q_2^*) = \max_{(q_1, q_2)} \left\{ p_{q_1} \cdot q_1 \left| \begin{array}{l} p_2 \cdot u_2(h_4) \\ u_2(h_3) \\ u_2(h_2) \\ u_2(h_1) \end{array} \right| q_2 \left| \begin{array}{l} p_2 \cdot (1-p_{q_2}) \cdot q_2 \\ u_2(h_7) \\ u_2(h_6) \\ u_2(h_5) \end{array} \right| 0 \right\}
$$

Then

$$
\Delta_1(p_1, p_2) = p_1 \left| \begin{array}{l} u_2(h_2) \\ u_2(h_1) \end{array} \right| - p_2 \left| \begin{array}{l} u_2(h_4) \\ u_2(h_3) \end{array} \right| \text{ and } \Delta_2(p_1, p_2) = p_1 \left| \begin{array}{l} u_2(h_6) \\ u_2(h_5) \end{array} \right| - p_2 \left| \begin{array}{l} u_2(h_8) \\ u_2(h_7) \end{array} \right|
$$

Denote $\Delta_i(p_1, p_2) = \begin{cases} 1, \Delta_i(p_1, p_2) > 0 \\ 0, \Delta_i(p_1, p_2) < 0 \end{cases}$

for $i = 1, 2$. We only consider general cases $\Delta_i(p_1, p_2) \neq 0$ here. For the specific case $\Delta_i(p_1, p_2) = 0$, it can be reduced to general cases. Then Player 2’s mixed strategy probability pair satisfies $(q_1, q_2) \in Q = \{ (0, 0), (0, 1), (1, 0), (1, 1) \}$. Let $P(q_1, q_2) = \{ (p_1, p_2) | q_i = f_i(p_1, p_2), i = 1, 2 \}$ be the set of Player 1’s mixed strategy probability pairs that satisfies the value of Player 2’s mixed strategy probability pair is $(q_1, q_2)$. Since $(q_1, q_2) \in Q = \{ (0, 0), (0, 1), (1, 0), (1, 1) \}$, there are only four different $P(q_1, q_2)$, which are $P_{(0,0)}, P_{(0,1)}, P_{(1,0)}, P_{(1,1)}$.

Now we fix $p_\varphi$, Player 2’s strategy $(q_1, q_2)$ and analyze Player 1’s strategy $(p_1, p_2)$.

Case 1: $(p_1, p_2) \in P_{(1,1)}$

Player 1’s expected utility is

$$
u_{(1,1)}(p_\varphi, p_1) = p_\varphi \times p_1 \left| \begin{array}{l} u_1(h_2) \\ u_1(h_1) \end{array} \right| + (1-p_\varphi) \times p_1 \left| \begin{array}{l} u_1(h_6) \\ u_1(h_5) \end{array} \right|
$$

To maximize its utility, Player 1 would set

$$
p_1^* = \max_{p_1} \left\{ \nu_{(1,1)}(p_\varphi, p_1) \right\} = \max_{p_1} \left\{ p_1 \left| \begin{array}{l} p_\varphi u_1(h_2) + (1-p_\varphi)u_1(h_6) \\ p_\varphi u_1(h_1) + (1-p_\varphi)u_1(h_5) \end{array} \right| 1 \right\}.
$$

Denote $\Delta_{(1,1)}(p_\varphi) = p_\varphi u_1(h_1) + (1 - p_\varphi)u_1(h_5) - p_\varphi u_1(h_2) - (1 - p_\varphi)u_1(h_6)$, then

$$p_1 = g_{(1,1)}(p_\varphi) = \begin{cases} 
1, & \Delta_{(1,1)}(p_\varphi) > 0 \\
0, & \Delta_{(1,1)}(p_\varphi) < 0 \\
\forall \alpha \in [0, 1], & \Delta_{(1,1)}(p_\varphi) = 0 
\end{cases}$$

. If there exists $p_2 \in [0, 1]$ such that $(p_1, p_2) \in \mathcal{P}_{(1,1)}$, then the result of this case is $o_{(1,1)}(p_\varphi) = u_{(1,1)}(p_\varphi, p_1)$ where $p_1 = g_{(1,1)}(p_\varphi)$, otherwise there is no Nash Equilibrium in this case, let $o_{(1,1)}(p_\varphi) = -\infty$.

Case 2: $(p_1, p_2) \in \mathcal{P}_{(0,0)}$

This case is similar to Case 1, in the same way we can get the result: If there exists $p_1 \in [0, 1]$ such that $(p_1, p_2) \in \mathcal{P}_{(0,0)}$, then the result of this case is $o_{(0,0)}(p_\varphi) = u_{(0,0)}(p_\varphi, p_2)$ where $p_2 = g_{(1,1)}(p_\varphi)$, otherwise there is no Nash Equilibrium in this case, let $o_{(0,0)}(p_\varphi) = -\infty$.

Case 3: $(p_1, p_2) \in \mathcal{P}_{(1,0)}$

Player 1’s expected utility is

$$u_{(1,0)}(p_\varphi, p_1, p_2) = p_\varphi \times p_1 \left| \frac{u_1(h_2)}{u_1(h_1)} + (1 - p_\varphi) \times p_2 \right| \frac{u_1(h_8)}{u_1(h_7)}$$

. To maximize its expected utility, Player 1 would set $(p_1^*, p_2^*) = g_{(1,0)}(p_\varphi) = \max_{(p_1, p_2)} \left\{ \max_{p_1} \left\{ \frac{u_1(h_2)}{u_1(h_1)} \right\}, \max_{p_2} \left\{ \frac{u_1(h_8)}{u_1(h_7)} \right\} \right\}$. Similarly, if there exists $(p_1, p_2) = g_{(1,0)}(p_\varphi) \in \mathcal{P}_{(1,0)}$ then the result is $o_{(1,0)}(p_\varphi) = u_{(1,0)}(p_\varphi, p_1, p_2)$, otherwise there is no Nash Equilibrium in this case.

Case 4: $(p_1, p_2) \in \mathcal{P}_{(0,1)}$

This case is similar to Case 3, in the same way we can get the result: if there exists $(p_1, p_2) = g_{(0,1)}(p_\varphi) \in \mathcal{P}_{(0,1)}$ then the result is $o_{(0,1)}(p_\varphi) = u_{(0,1)}(p_\varphi, p_1, p_2)$, otherwise there is no Nash Equilibrium in this case.

All Strategies with $p_\varphi$ that satisfies $o_{(q_1, q_2)}(p_\varphi) \neq -\infty$ are Nash Equilibrium.

5.4. Example

Take a look at the picture above. All the settings are the same in Nash Equilibrium Computation with utility value given in the picture. In the game without self-blind randomization, the Nash Equilibrium returned by Backward Induction is RRL, where the utility is $(2, 2)$.

In game with self-blind randomization, Player 1’s strategy is playing self-blind randomization on level 1, mixed strategies in information sets $\{LL, RL\}$ and $\{LR, RR\}$, and Player 2’s strategy is playing mixed strategies when $h \in \{L, R\}$.

At first, we fix $(p_1, p_2)$. The delta we define in Nash Equilibrium Computation is $\Delta_1(p_1, p_2) = p_1 \left| \begin{array}{c} -1 \\ -3 - p_1 \end{array} \right| \begin{array}{c} -2 \\ 1 \end{array}$ and $\Delta_2(p_1, p_2) = p_1 \left| \begin{array}{c} -2 \\ -1 - p_1 \end{array} \right| \begin{array}{c} -2 \\ 2 \end{array}$ . Next, we need to compute $P_{(q_1, q_2) \in \mathcal{Q}}$. To satisfy the condition $(q_1, q_2) = (1, 1)$, let $\Delta_1(p_1, p_2) > 0$ and $\Delta_2(p_1, p_2) > 0$, we get $p_2 \in \left[ -\frac{2}{3}, \frac{1}{3} \right)$ and $p_2 < \frac{1}{3} p_1$, which is $P_{(1,1)}$ Analogously, we can get $P_{(0,0)}, P_{(1,0)}, P_{(0,1)}$ as follows.

$$P_{(1,1)} = \{(p_1, p_2) | p_2 < -\frac{2}{3} p_1 + \frac{1}{3}, p_2 < \frac{1}{3} p_1\}$$

$$P_{(0,0)} = \{(p_1, p_2) | p_2 > -\frac{2}{3} p_1 + \frac{1}{3}, p_2 > \frac{1}{3} p_1\}$$

$$P_{(1,0)} = \{(p_1, p_2) | p_2 < -\frac{3}{5} p_1 + \frac{1}{5}, p_2 > \frac{1}{5} p_1\}$$

$$P_{(0,1)} = \{(p_1, p_2) | p_2 > -\frac{3}{5} p_1 + \frac{1}{5}, p_2 < \frac{1}{5} p_1\}$$
We draw these four sets on a rectangular coordinate system where abscissa presents $p_1$ and ordinate presents $p_2$.

Next, we fix $(q_1, q_2)$ and analyze $(p_1, p_2)$.

Case 1: $(p_1, p_2) \in P_{(1,1)}$

Since for any $p_φ$, $\Delta_{(1,1)}(p_φ) = 4p_φ + 7(1 - p_φ) > 0$, Player 1 will always set $p_1 = g_{(1,1)}(p_φ) \equiv 1$. However, for all $p_2 \in [0, 1]$, $(p_1, p_2) = (1, p_2) \notin P_{(1,1)}$, therefore the utility of Player 1 in this case is $o_{(1,1)}(p_φ) = -\infty$.

Case 2: $(p_1, p_2) \in P_{(0,0)}$
In this case, $\Delta_{(0,0)}(p_\varphi) = -p_\varphi + 2(1 - p_\varphi) > 0$ when $p_\varphi < \frac{2}{3}$. Since $(x, 1) \in P_{(0,0)}$ for any $x \in (0, 1]$, $o_{(0,0)}(p_\varphi) = -p_\varphi + 2(1 - p_\varphi) = -3p_\varphi + 2$ when $p_\varphi < \frac{2}{3}$.

Analogously when $p_\varphi > \frac{2}{3}$, $p_2 = g_{(0,0)}(p_\varphi) = 0$. Since $(p_1, 0) \notin P_{(0,0)}$ for any $p_1 \in [0, 1]$, then $o_{(0,0)}(p_\varphi) = -\infty$ when $p_\varphi < \frac{2}{3}$.

Case 3: $(p_1, p_2) \in P_{(1,0)}$

\[
\begin{align*}
\begin{bmatrix}
 u_{(1,0)}(p_\varphi, p_1, p_2) = p_\varphi \times p_1 & 0 & 0 \\
 0 & (1 - p_\varphi) \times p_2 & 0 \\
 0 & 0 & 2 \end{bmatrix}
\end{align*}
\]
\[\text{and } (p_1, p_2) = (1, 1)\]

Since $(1, 1) \notin P_{(1,0)}$, $o_{(1,0)}(p_\varphi) = -\infty$.

Case 4: $(p_1, p_2) \in P_{(0,1)}$

\[p_1, p_2 = g_{(0,1)}(p_\varphi) = (1, 0)\]

Since $(1, 1) \in P_{(0,1)}$, $o_{(1,0)}(p_\varphi) = 0 \cdot p_\varphi + 7(1 - p_\varphi) = 7 - 7p_\varphi$.

When $p_\varphi \leq \frac{2}{3}$, there are two Nash Equilibria for any specific $p_\varphi$. One is good for Player 1 while the other one is bad. For the good Nash Equilibrium, $p_1 = 1, p_2 = 0, q_1 = 0, q_2 = 1$, Player 1’s payoff is $7 - 7p_\varphi$. For the bad one, $p_1 \in [0, 1], p_2 = 1, q_1 = q_2 = 0$, Player 1’s payoff is $-3p_\varphi + 2$.

When $p_\varphi = \frac{2}{3} + \varepsilon$ where $\varepsilon \to 0^+$, there is only one Nash equilibrium where $p_1 = 1, p_2 = 0, q_1 = 0, q_2 = 1$, and Player 1’s payoff is $\frac{4}{3} - \varepsilon$ larger than game without self-blind randomization.

REFERENCES

CONITZER, V. AND SANDHOLM, T. 2006. Computing the optimal strategy to commit to.


