# Modeling Unimportant Decisions

Pawin Siriprapanukul
Australian National University
e-mail: u4161271@anu.edu.au

October 2006

**Abstract**

We propose an incomplete information game, in which rational players depend on finite belief hierarchies in determining their optimal actions. This is done by introducing costs when players climb up their belief hierarchies. With bounded payoffs, players climb up to a finite order of belief, and just depend on these beliefs in making their decisions. The model is consistent with an experimental literature which shows that people play games with finite belief hierarchies, while it can avoid the problem of misspecification of high order beliefs, which may limit predictive power of some theories.

## 1 Introduction

A recent experimental literature, see for example Stahl and Wilson [12], and Nagel [8], shows that in playing games people depend on *finite* belief hierarchies in determining their actions. This is used as a reason to explain the poor prediction of game theory in many experimental settings. Sakovics [11] uses this literature to support his bounded rationality model. According to Sakovics [11], each player forms a finite belief hierarchy, and the optimal decision, then, is made according to this hierarchy.

Some drawbacks occur to models similar to Sakovics' [11]. Firstly, the model has to assume a common finite order of beliefs that every player would form up to. Secondly, without any reason every player thinks that his opponents form beliefs one order shorter than he does. At last, the model faces a big inconsistency that for the decision on his action each player chooses it optimally, while for the decision to stop forming beliefs he does it without reason.

Here, we propose an incomplete information model, in which rational players choose their optimal actions, relying on just finite belief hierarchies. This is done by introducing costs of using high order beliefs in making optimal decisions. These costs occur because players have to consider additional possible states of the world, while they are more uncertain about beliefs that they use. Hence, they are reluctant to use high order beliefs. With increasing costs and bounded payoffs, it is optimal for players to stop using high order beliefs at a finite order. Then, they use these finite belief hierarchies to determine their optimal actions.

The model has some advantages over bounded rationality models, which have similar features to Sakovics' [11]. There is no inconsistency occurs, because players make every decision optimally. The model allows each player to stop processing his high order beliefs at different order. Also, in equilibrium each player has consistent belief over the stopping points of his opponents. Section 2 provides specification of this model. Section 3 characterizes equilibria in this model. Section 4 gives some simple examples.

Section 5 discusses some advantages of this model over standard Bayesian games, apart from the reason that it is consistent with experimental studies. With finite belief hierarchies, we can get rid off the problem of infinite number of types, resulted from infinite belief hierarchies in general Bayesian games. We can also avoid the extreme assumption of common prior or common knowledge of infinite belief hierarchies implicitly assumed in Bayesian games with finite types of each player. At last, we can avoid the problem related to the impact of high order beliefs, which requires theorists to specify correctly the whole infinite belief hierarchies. Section 6 concludes the paper.

# 2 Specification of the Model

## 2.1 Type space and type notations

In an incomplete information model, some payoff-relevant information is missing. Since the missing information can affect players' payoff, it is in each player's interest to form his belief over this missing information. We call it first-order belief. Since the outcome of the model is also affected by the decisions of other players in the model, each wants to form beliefs on what others will do. Since each player knows that others choose their decisions based on their first-order beliefs, each wants to form a belief over these first-order beliefs. This is his second-order belief. Then, each player knows that others also base their decisions on the second-order beliefs, he wants to form belief over these second-order beliefs. The same argument continues for other high-order beliefs, resulting in an infinite belief hierarchy.

In this paper, we will restrict our attention to a world with only two players[1]. We start by constructing a type space for our model. We have some missing information, and a common space $S$ contains *all* possible values of this missing information. Assume that it contains only finite different objects. This means that each player knows that there can be only finite possible alternative values for this missing information. We denote the space of probability measure on the $\sigma$-field of a metric space $Z$ as $\Delta Z$. Then, define recursively:

[1]The expansion of the model to the case of more than 2 players is rather straightforward.

$$
\begin{aligned}
X_0 &= S \\
X_1 &= X_0 \times \Delta X_0 \\
X_2 &= X_1 \times \Delta X_1 \\
&\vdots \\
X_l &= X_{l-1} \times \Delta X_{l-1} \\
&\vdots
\end{aligned} \tag{1}
$$

Let $\delta_l^i$ denote the probability measure that player $i$ assigns on the set $X_{l-1}$, for $l \geq 1$. That is $\delta_l^i \in \Delta X_{l-1}$. Hence, $\delta_1^i$, which is the belief that $i$ assigns over the space $S$, is the first-order belief of player $i$. Next, $\delta_2^i$ is a way to model second-order belief. We can think of $X_1$ as the set that contains all possible combinations between the missing information, which is the $X_0$ part of the set, and the possible first-order beliefs of player $j$, which is represented by the $\Delta X_0$ part of the set. Marginal probability of $\delta_2^i$ over the set $\Delta X_0$ tells us the belief that player $i$ assigns over the possible values of the first-order belief of player $j$. This way of modeling second-order belief is more general and, hence, widely applied, since it can deal with the case of dependent beliefs over different orders. $\delta_l^i$ is then a way to model the $l$-th-order belief. We assume that each player $i$ assigns positive probability to only finite possible cases in each order of belief, and this is a common knowledge in the model. This means that both players know that there are only finite possible values for each $\delta_l^i$, for $l \geq 1$.

A type $t^i$ of a player $i$ is an infinite belief hierarchy $t^i = (\delta_1^i, \delta_2^i, \delta_3^i, ...) \in \times_{l=0}^{\infty} \Delta X_l$. Actually, $t^i$ is exactly the same belief hierarchy as in the standard Bayesian models, when each order of belief is assumed to be finite. We assume that there is no duplicating belief hierarchy. Hence, a belief hierarchy $t^i$ will have at least one value of $\delta_l^i$, for $l \geq 1$, different from ones of other hierarchies. With different belief structures, a player may choose different decisions. This is the reason that in the standard Bayesian models we have to find a decision for each type $t^i$ of a player $i$ to define the equilibrium outcomes.

To make the model closed, we impose the conditions of "coherency" and "common knowledge of coherency", as defined in Brandenburger and Dekel [2], into our belief system. Let the set $T^{i\prime}$ collects all relevant $t^i$'s. Brandenburger and Dekel [2] show that with the "coherency" and "common knowledge of coherency" conditions, there exists a homeomorphism $g : T^{i\prime} \rightarrow \Delta(S \times T^{j\prime})$. We can think of $g$ as a function associates each $t^i$ with a unique probability measure on the space $(S \times T^{j\prime})$. That is $g(t^i)$ is a joint probability measure over the set of all possible values of the missing information, $S$, and the infinite belief hierarchies of player $j$, $T^{j\prime}$. In other words, a type $t^i$ of player $i$ forms a unique belief on the possible values of missing information and the infinite belief hierarchies of player $j$. Then the decision of type $t^i$ of player $i$ is based on this belief. The space $\Omega = (S \times T^{i\prime} \times T^{j\prime})$ is the universal state space (or universal BL-space in Mertens and Zamir [5]).

It is common in standard Bayesian models that, if $(s, t^i, t^j) \in (S \times T^{i\prime} \times T^{j\prime})$ is a true state, then:

$$(s, t^j) \in \text{supp}[g(t^i)], \;\; i \neq j, \tag{2}$$

where the operator "supp[.]" refers to the support of the probability measure in the bracket. Condition (2) says that a type $t^i$ of player $i$ must give positive probability to the event that $(s, t^j)$ occurs. In other words, each player never excludes the true state from the set of states he considers as possible. To make sure that the condition (2) is met, we will work on a subspace of the previous universal state space, which is:

$$(S \times T^i \times T^j) = \left\{(s, t^i, t^j) \in (S \times T^{i\prime} \times T^{j\prime}) | (s, t^j) \in \text{supp}[g(t^i)], \forall t^i, t^j, \text{ and } i \neq j\right\} \tag{3}$$

This is a common knowledge in the model. Then, from now on the space $(S \times T^i \times T^j)$ in (3) is our state space. A state of the world from this set is denoted by $(s, t^i, t^j) \in (S \times T^i \times T^j)$.

It is convenient to have a type notation which also tells us the finite belief hierarchy equipped with it. Denote $t^i_{m^i}$ as a "type $t^i \in T^i$ with beliefs just up to his $m^i$-th order". That is $t^i_{m^i} = (\delta^i_1, \delta^i_2, ..., \delta^i_{m^i}) = \times_{l=0}^{m^i-1} \text{proj}_{\Delta X_l} t^i$, where the operator "$\text{proj}_Z z$" means projection of an element $z$ on the set $Z$. Let the set $T^i_{m^i}$ collect all relevant values of $t^i_{m^i}$. Then we define the set of $t^i_{m^{i\prime}}$ associated with $t^i_{m^i}$, when $m^{i\prime} \geq m^i$ as:

$$\Upsilon^i_{m^{i\prime}}(t^i_{m^i}) = \left\{ t^i_{m^{i\prime}} \in T^i_{m^{i\prime}} \,\middle|\, \forall l \leq m^i - 1,\ \text{proj}_{\Delta X_l} t^i_{m^{i\prime}} = \text{proj}_{\Delta X_l} t^i_{m^i} \right\}. \tag{4}$$

The set $\Upsilon^i_{m^{i\prime}}(t^i_{m^i})$ in (4) collects all possible types $t^i_{m^{i\prime}}$'s that share the same first $m^i$ components with $t^i_{m^i}$.

Define $\overline{m}^i$ as the highest order of belief that all types $t^i$'s of a player $i$ would form up to. It turns out that each type $t^i_{\overline{m}^i} \in T^i_{\overline{m}^i}$ is important in defining equilibrium in our model[2]. Before proceeding further, define $(t^i_{\overline{m}^i}, m^i) \in T^i_{\overline{m}^i} \times \mathbb{N}$ as a "type $t^i_{\overline{m}^i}$ who forms up to his $m^i$ -th-order belief". Note that $(t^i_{\overline{m}^i}, m^i)$ has exactly the same belief structure as $t^i_{m^i}$, if $t^i_{\overline{m}^i} \in \Upsilon^i_{\overline{m}^i}(t^i_{m^i})$. Hence, this $(t^i_{\overline{m}^i}, m^i)$ would make the same decisions as its associated $t^i_{m^i}$. Note also that $\Upsilon^i_{m^{i\prime}}(t^i_{\overline{m}^i}, m^i) = \Upsilon^i_{m^{i\prime}}(t^i_{m^i})$.

## 2.2 Equilibrium conditions and other specifications

Let $A^i$ be a set of all possible actions, $a^i$, for player $i$. Define $\phi^i : T^i_{\overline{m}^i} \times \mathbb{N} \rightarrow A^i$ as a strategy function of player $i$. Let $\Phi^i$ collect all possible strategy functions $\phi^i$. Then define $\mu^i : T^i_{\overline{m}^i} \rightarrow \mathbb{N}$ as a belief formation function. $\mu^i(t^i_{\overline{m}^i})$ tells us the highest order of belief that a type $t^i_{\overline{m}^i}$ would process up to. The set $M^i$ collects all possible functions $\mu^i$. We can define equilibrium decisions of each player $i$, in the spirit of Bayesian-Nash

[2] We will discuss about this in proposition 1 in section 3 below.

equilibrium, as a pair $(\phi^{i*}, \mu^{i*})$, in which:

- Each type $t^i_{\overline{m}^i}$ has a correct belief on the pair $(\phi^{j*}, \mu^{j*})$.
- $\phi^{i*}(t^i_{\overline{m}^i}, m^i)$ is the optimal action of a type $t^i_{\overline{m}^i}$, $\forall t^i_{\overline{m}^i} \in T^i_{\overline{m}^i}$ and $\forall m^i \leq \overline{m}^i$, given $(\phi^{j*}, \mu^{j*})$.
- $\mu^{i*}(t^i_{\overline{m}^i})$ is the optimal belief formation of a type $t^i_{\overline{m}^i}$, for all $t^i_{\overline{m}^i} \in T^i_{\overline{m}^i}$, given $\phi^{i*}$ and $(\phi^{j*}, \mu^{j*})$.
- $\overline{m}^i$ is the maximum value of $\mu^{i*}(t^i_{\overline{m}^i})$, $\forall i$, when comparing across all $t^i_{\overline{m}^i} \in T^i_{\overline{m}^i}$.

The first condition above is an equilibrium condition. The optimal conditions for the second and the third statements need some discussions, and, hence, will be specified below. The equilibrium outcome of the model is:

$$\left(\phi^{i*}(t^i_{\overline{m}^i}, \mu^{i*}(t^i_{\overline{m}^i})), \phi^{j*}(t^j_{\overline{m}^j}, \mu^{j*}(t^j_{\overline{m}^j}))\right), \forall t^i_{\overline{m}^i}, t^j_{\overline{m}^j} \tag{5}$$

We can look at this equilibrium outcome as optimal decisions of every type $(t^i_{\overline{m}^i}, \mu^{i*}(t^i_{\overline{m}^i}))$. Each $t^i_{\overline{m}^i}$ has his optimal decision in processing his beliefs $\mu^{i*}(t^i_{\overline{m}^i})$. Then, he decides on his optimal action according to $\phi^{i*}(t^i_{\overline{m}^i}, \mu^{i*}(t^i_{\overline{m}^i}))$.

Now we look at optimal condition for $\phi^{i*}(t^i_{\overline{m}^i}, m^i)$. Given the value of $\overline{m}^j$, a type $(t^i_{\overline{m}^i}, m^i)$ knows that the relevant beliefs for him are just ones up to his $(\overline{m}^j+1)$-th-order belief. This is because he knows that each relevant type $t^j_{\overline{m}^j}$ would form his beliefs up to at most the $\overline{m}^j$-th order before making his decision, and the $(\overline{m}^j+1)$-th-order belief of $(t^i_{\overline{m}^i}, m^i)$ can capture his belief over $t^j_{\overline{m}^j}$'s $\overline{m}^j$-th-order belief already. Hence, it is enough to base his own decision with his $(\overline{m}^j+1)$-th-order belief. If $m^i \geq \overline{m}^j+1$, a type $(t^i_{\overline{m}^i}, m^i)$ already has his unique subjective probability distribution over $X_{\overline{m}^j} = S \times T^j_{\overline{m}^j}$, which is his $\delta^i_{\overline{m}^j+1} = \text{proj}_{\Delta X_{\overline{m}^j}} t^i_{m^i}$. However, for a type $(t^i_{\overline{m}^i}, m^i)$ with $m^i < \overline{m}^j+1$, he does not. For such a type, he must form an alternative guess of probability distribution over $S \times T^j_{\overline{m}^j}$.

There is a common set of beliefs that all players rely on when they do not want to process their own subjective beliefs. We can think of this belief as a fairly reliable and easy to process one. One alternative candidate is the uniform probability distribution. However, here we calculate the altermative belief from the structure of the type space of the model. The case of players using uniform probability distribution as their alternative belief can also be formed by appropriate specification of the type space[3]. Denote $d^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)$ to be the alternative guess of a type $(t^i_{\overline{m}^i}, m^i)$ for probability distribution over $S \times T^j_{m^i}$, when he does not want to process his own $(\overline{m}^j + 1)$-th-order belief. It can be expressed as:

$$d^i_{m^i+1}(t^i_{\overline{m}^i}, m^i) = \sum_{t^i_{m^i+1} \in \Upsilon^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)} \frac{1}{|\Upsilon^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)|} \cdot \text{proj}_{\Delta X_{m^i}} t^i_{m^i+1}, \tag{6}$$

where $|Z|$ denotes the number of elements in the set $Z$. The formula in (6) allows us to form the alternative belief of the next consecutive order. However, in many cases players want to form beliefs of the higher order. Apply the same formula to each type $t^i_{\overline{m}^i+1} \in \Upsilon^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)$, we can define recursively:

$$d^i_{m^i+l}(t^i_{\overline{m}^i}, m^i) = \sum_{t^i_{m^i+1} \in \Upsilon^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)} \frac{1}{|\Upsilon^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)|} \cdot d^i_{m^i+l}(t^i_{m^i+1}). \tag{7}$$

Hence, using (7) with (6), for example, the value of alternative belief $d^i_{m^i+2}(t^i_{\overline{m}^i}, m^i)$ is $\sum_{t^i_{m^i+1} \in \Upsilon^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)} \frac{1}{|\Upsilon^i_{m^i+1}(t^i_{\overline{m}^i}, m^i)|} \left( \sum_{t^i_{m^i+2} \in \Upsilon^i_{m^i+2}(t^i_{m^i+1})} \frac{1}{|\Upsilon^i_{m^i+2}(t^i_{\overline{m}^i+1})|} \cdot \text{proj}_{\Delta X_{m^i+1}} t^i_{m^i+2} \right)$. For a type $(t^i_{\overline{m}^i}, m^i)$ with $m^i < \overline{m}^j + 1$, his alternative guess of probability distribution over $S \times T^j_{\overline{m}^j}$, then, is $d^i_{\overline{m}^j+1}(t^i_{\overline{m}^i}, m^i)$.

[3]See, for example, example 1 in section 4 below.

Our players are resource-bounded rational and reluctant to use his own subjective high-order beliefs. By resource-bounded rationality, we mean that in making a decision it is costly for a type $(t^i_{\overline{m}^i}, m^i)$ to analyze each possible state of the world. He is also reluctant to use high-order beliefs because it involves a high content of error. Hence, in processing up to his $m^i$-th-order belief, where $m^i \in \mathbb{N}$, he has to pay the utility cost of $C^i(m^i)$. Then, a type $(t^i_{\overline{m}^i}, m^i)$ has two alternatives, which are first to use $d^i_{\overline{m}^j+1}(t^i_{\overline{m}^i}, m^i)$ for free, or second to process his own subjective belief up to the $m^{i\prime}$-th-order, which is of course higher than the $m^i$-th-order, and pay the additional utility cost $(C^i(m^{i\prime}) - C^i(m^i))$. When the type $(t^i_{\overline{m}^i}, m^i)$ processes an additional order of his belief, he has to encounter additional possible states of the world, and he thinks that it involves higher errors. Hence, we set that for any $m^{i\prime} > m^i$, $C^i(m^{i\prime}) > C^i(m^i)$. This cost function is defined as a part of the state $s \in S$, which means that both players know the exact value of $C^i(m^i)$, $\forall i$.

The players' preferences satisfy standard axioms of subjective expected utility maximizers. Hence, their preferences can be represented by expected utility functions. Given functions $\phi^j(t^j_{\overline{m}^j}, m^j)$ and $\mu^j(t^j_{\overline{m}^j})$ of the player $j$, we specify utility function of a type $(t^i_{\overline{m}^i}, m^i)$, when he plays action $a^i$ as:

$$U^i(a^i, (t^i_{\overline{m}^i}, m^i); \phi^j(t^j_{\overline{m}^j}, \mu^j(t^j_{\overline{m}^j}))) = \sum_{s \times t^j_{\overline{m}^j} \in S \times T^j_{\overline{m}^j}} \left[u^i\left(a^i; \phi^j(t^j_{\overline{m}^j}, \mu^j(t^j_{\overline{m}^j})), s\right) - C^i(m^i)\right] \cdot p(s \times t^j_{\overline{m}^j}), \tag{8}$$

where $u^i(.)$ is a Bernoulli payoff function of player $i$, the probability distribution function $p$ is equal to $\delta^i_{\overline{m}^j+1} = \text{proj}_{\Delta X_{\overline{m}^j}} t^i_{m^i}$, for a type $(t^i_{\overline{m}^i}, m^i)$ with $m^i \geq \overline{m}^j + 1$, and it is equal to $d^i_{\overline{m}^j+1}(t^i_{\overline{m}^i}, m^i)$ for a type $(t^i_{\overline{m}^i}, m^i)$ with $m^i < \overline{m}^j + 1$.

Define the best response correspondence $BR^i : T^i_{\overline{m}^i} \times \mathbb{N} \times \Phi^j \times M^j \rightarrow A^i$ for a type $(t^i_{\overline{m}^i}, m^i)$, $\forall m^i \leq \overline{m}^i$ as:

$$BR^i((t^i_{\overline{m}^i}, m^i); \phi^j(t^j_{\overline{m}^j}, \mu^j(t^j_{\overline{m}^j}))) = \arg\max_{a^i \in A^i} U^i(a^i, (t^i_{\overline{m}^i}, m^i); \phi^j(t^j_{\overline{m}^j}, \mu^j(t^j_{\overline{m}^j}))). \quad (9)$$

Then, optimal decision $\phi^{i*}(t^i_{\overline{m}^i}, m^i)$ for a type $(t^i_{\overline{m}^i}, m^i)$, given $\phi^j(t^j_{\overline{m}^j}, m^j)$ and $\mu^j(t^j_{\overline{m}^j})$, is:

$$\phi^{i*}(t^i_{\overline{m}^i}, m^i) \in BR^i(t^i_{\overline{m}^i}, m^i; \phi^j(t^j_{\overline{m}^j}, \mu^j(t^j_{\overline{m}^j}))), \forall t^i_{\overline{m}^i} \text{ and } \forall m^i \leq \overline{m}^i. \quad (10)$$

Note that for a given pair of $(\phi^j, \mu^j)$, $\phi^{i*}(t^i_{\overline{m}^i}, m^i)$ may not be equal to $\phi^{i*}(t^i_{\overline{m}^i}, m^{i\prime})$, if $m^i \neq m^{i\prime}$.

A type $t^i_{\overline{m}^i}$ processes his high-order beliefs up until the point that the additional benefit from processing his belief is less than or equal to zero. This involves comparing the value of expected utility $U^i(\phi^{i*}(t^i_{\overline{m}^i}, m^i), (t^i_{\overline{m}^i}, m^i); \phi^{j*}(t^j_{\overline{m}^j}, \mu^{j*}(t^j_{\overline{m}^j})))$ at different levels of $m^i$. Consider the case that currently a type $t^i_{\overline{m}^i}$ forms his beliefs up to the $m^i$-th order and he is considering whether to process his $m^{i\prime}$-th-order belief, in which $m^{i\prime} > m^i$. He knows his $m^{i\prime}$-th-order belief, but processing it involves positive additional cost of $(C^i(m^{i\prime}) - C^i(m^i))$. If processing this high order belief does not alter his decision, there is no need for him to process it. This is because there is no additional benefit, while he has to pay additional cost. Also, he will not process his $m^{i\prime}$-th-order belief when the benefit of processing his additional beliefs is less than the additional cost. Hence, he processes his $m^{i\prime}$-th-order belief, only when (i) it alters his decision, *and* (ii) the expected *loss* from *not* processing it is more than the additional cost.

Formally, given $\phi^{i*}$ and $(\phi^{j*}, \mu^{j*})$, a type $t^i_{\overline{m}^i}$ processes his high order belief up to the optimal order $\mu^{i*}(t^i_{\overline{m}^i})$. For convenient, here we write $\phi^{i*}(m^i)$ to represent $\phi^{i*}(t^i_{\overline{m}^i}, m^i)$, and $U^i(\phi^{i*}(m^i), (t^i_{\overline{m}^i}, m^{i\prime}); \phi^{j*}, \mu^{j*})$ for $U^i(\phi^{i*}(t^i_{\overline{m}^i}, m^i), (t^i_{\overline{m}^i}, m^{i\prime}); \phi^{j*}(t^j_{\overline{m}^j}, \mu^{j*}(t^j_{\overline{m}^j})))$. Note that $m^{i\prime}$ may not be equal to $m^i$. The optimal value $\mu^{i*}(t^i_{\overline{m}^i})$ will make, $\forall m^i \leq \mu^{i*}(t^i_{\overline{m}^i})$:

$$U^i(\phi^{i*}(\mu^{i*}(t^i_{\overline{m}^i})), (t^i_{\overline{m}^i}, \mu^{i*}(t^i_{\overline{m}^i})); \phi^{j*}, \mu^{j*}) \geq U^i(\phi^{i*}(m^i), (t^i_{\overline{m}^i}, \mu^{i*}(t^i_{\overline{m}^i})); \phi^{j*}, \mu^{j*}), \quad (11)$$

and, $\forall m^{i\prime} > \mu^{i*}(t^i_{\overline{m}^i})$,

$$U^i(\phi^{i*}(\mu^{i*}(t^i_{\overline{m}^i})), (t^i_{\overline{m}^i}, m^{i\prime}); \phi^{j*}, \mu^{j*}) \geq U^i(\phi^{i*}(m^{i\prime}), (t^i_{\overline{m}^i}, m^{i\prime}); \phi^{j*}, \mu^{j*}). \quad (12)$$

For the case of $m^i \leq \mu^{i*}(t^i_{\overline{m}^i})$, if the type $t^i_{\overline{m}^i}$ does not process his $\mu^{i*}(t^i_{\overline{m}^i})$-th-order belief, he chooses $\phi^{i*}(m^i)$, while if he processes it, he chooses $\phi^{i*}(\mu^{i*}(t^i_{\overline{m}^i}))$. Using utility function from (8) (which considers processing cost already), the condition in (11) says that it is worth for a type $t^i_{\overline{m}^i}$ to process up to his $\mu^{i*}(t^i_{\overline{m}^i})$-order belief, since the expected loss from not processing it is more than the additional cost. For the case of $m^{i\prime} > \mu^{i*}(t^i_{\overline{m}^i})$, if the type $t^i_{\overline{m}^i}$ process his $m^{i\prime}$-order belief, he chooses $\phi^{i*}(m^{i\prime})$. Then, the condition (12) says that he does not want to process it.

Then, we can specify the exact conditions for an equilibrium of the model as a pair $(\phi^{i*}, \mu^{i*})$, $\forall i$, where:

- Each type $t^i_{\overline{m}^i}$ has a correct belief on the pair $(\phi^{j*}, \mu^{j*})$.
- $\phi^{i*}(t^i_{\overline{m}^i}, m^i)$ satisfies condition (10), given $(\phi^{j*}, \mu^{j*})$.
- $\mu^{i*}(t^i_{\overline{m}^i})$ satisfies conditions (11) and (12), given $\phi^{i*}$ and $(\phi^{j*}, \mu^{j*})$.
- $\overline{m}^i$ is the maximum value of $\mu^{i*}(t^i_{\overline{m}^i})$, $\forall i$, when comparing across all $t^i_{\overline{m}^i} \in T^i_{\overline{m}^i}$.

# 3 Some Characterization of Equilibria of the Model

In this section, we want to discuss some characteristics of the equilibrium in our model. These results will lead to a practical way to figure out an equilibrium of the model.

The below proposition 1 tells us the reason why our analyses complete with the decisions of each of $t^i_{\overline{m}^i}$ and $t^j_{\overline{m}^j}$. There is no need to look at the decisions of each $t^i$, which is the same as $t^i_\infty$ in our notation.

**Proposition 1** *Let $\overline{m}^i$, and $\overline{m}^j$ be given. For each $m^{i,1} \geq \overline{m}^i$, let $t^i_{m^{i,1}} \in \Upsilon^i_{m^{i,1}}(t^i_{\overline{m}^i})$. Then $\forall m^{i\prime} \leq \overline{m}^i$, $\phi^{i*}(t^i_{m^{i,1}}, m^{i\prime}) = \phi^{i*}(t^i_{\overline{m}^i}, m^{i\prime})$. Moreover, $\mu^{i*}(t^i_{m^{i,1}}) = \mu^{i*}(t^i_{\overline{m}^i})$.*

**Proof.** From our definitions, $(t^i_{m^{i,1}}, m^{i\prime}) = (t^i_{\overline{m}^i}, m^{i\prime}) = t^i_{m^{i\prime}}$. Hence, given $\overline{m}^j$:

$$U^i(a^i, (t^i_{m^{i,1}}, m^{i\prime}); \phi^{j*}, \mu^{j*}) = U^i(a^i, (t^i_{\overline{m}^i}, m^{i\prime}); \phi^{j*}, \mu^{j*}), \text{ for any given } \phi^{j*},\ \mu^{j*}.$$

The problem that each has to solve is the same, which makes $\phi^{i*}(t^i_{m^{i,1}}, m^{i\prime}) = \phi^{i*}(t^i_{\overline{m}^i}, m^{i\prime})$. Then, if $t^i_{\overline{m}^i}$ decides to stop forming high-order belief at the $\mu^{i*}(t^i_{\overline{m}^i})$-th order, which is lower than the $\overline{m}^i$-th order by definition, this means, $\forall m^{i\prime\prime} \leq \mu^{i*}(t^i_{\overline{m}^i})$:

$$U^i(\phi^{i*}(\mu^{i*}(t^i_{\overline{m}^i})), (t^i_{\overline{m}^i}, \mu^{i*}(t^i_{\overline{m}^i})); \phi^{j*}, \mu^{j*}) \geq U^i(\phi^{i*}(m^{i\prime\prime}), (t^i_{\overline{m}^i}, \mu^{i*}(t^i_{\overline{m}^i})); \phi^{j*}, \mu^{j*}), \quad (13)$$

and, $\forall m^{i\prime\prime\prime} > \mu^{i*}(t^i_{\overline{m}^i})$:

$$U^i(\phi^{i*}(\mu^{i*}(t^i_{\overline{m}^i})), (t^i_{\overline{m}^i}, m^{i\prime\prime\prime}); \phi^{j*}, \mu^{j*}) \geq U^i(\phi^{i*}(m^{i\prime\prime\prime}), (t^i_{\overline{m}^i}, m^{i\prime\prime\prime}); \phi^{j*}, \mu^{j*}). \quad (14)$$

The condition in (13) and the fact that $\phi^{i*}(t^i_{m^{i,1}}, m^{i\prime}) = \phi^{i*}(t^i_{\overline{m}^i}, m^{i\prime})$, $\forall m^{i\prime} \leq \overline{m}^i$, makes, $\forall m^{i\prime\prime} \leq \mu^{i*}(t^i_{\overline{m}^i})$:

$$U^i(\phi^{i*}(\mu^{i*}(t^i_{\overline{m}^i})), (t^i_{m^{i,1}}, \mu^{i*}(t^i_{\overline{m}^i})); \phi^{j*}, \mu^{j*}) \geq U^i(\phi^{i*}(m^{i\prime\prime}), (t^i_{m^{i,1}}, \mu^{i*}(t^i_{\overline{m}^i})); \phi^{j*}, \mu^{j*}). \quad (15)$$

Then, for $m^{i\prime\prime\prime} > \mu^{i*}(t^i_{\overline{m}^i})$, we need that:

$$U^i(\phi^{i*}(\mu^{i*}(t^i_{\overline{m}^i})), (t^i_{m^{i,1}}, m^{i\prime\prime\prime}); \phi^{j*}, \mu^{j*}) \geq U^i(\phi^{i*}(m^{i\prime\prime\prime}), (t^i_{m^{i,1}}, m^{i\prime\prime\prime}); \phi^{j*}, \mu^{j*}). \quad (16)$$

The condition in (16) holds because, firstly, if $\overline{m}^i \geq m^{i'''} > \mu^{i*}(t^i_{\overline{m}^i})$, (14) and the fact that $\phi^{i*}(t^i_{m^{i,1}}, m^{i\prime}) = \phi^{i*}(t^i_{\overline{m}^i}, m^{i\prime})$, $\forall m^{i\prime} \leq \overline{m}^i$, make (16) hold, and, secondly, if $m^{i'''} > \overline{m}^i$, (16) holds by the definition of $\overline{m}^i$.

Then, from (15) and (16), $t^i_{m^{i,1}}$ also stops forming his high-order belief at the $\mu^{i*}(t^i_{\overline{m}^i})$-th order. That is $\mu^{i*}(t^i_{m^{i,1}}) = \mu^{i*}(t^i_{\overline{m}^i})$. ■

From proposition 1, we can conclude that for any type $t^i \in \Upsilon^i_\infty(t^i_{\overline{m}^i})$, he behaves in the same way as his associated $t^i_{\overline{m}^i}$. Hence, there is no need to find optimal action for each $t^i$.

Next, our main objective is to show corollary 3, which says that given the value of $\overline{m}^j$, the value of $\overline{m}^i$ must be between $\overline{m}^j - 1$ and $\overline{m}^j + 1$. This is because there is no need for a type $t^i_{\overline{m}^i}$ to go further than his $\overline{m}^j + 1$-th-order belief, which makes $\overline{m}^i \leq \overline{m}^j + 1$. It is also true that a type $t^j_{\overline{m}^j}$ will not go further than his $\overline{m}^i + 1$-th-order belief, which makes $\overline{m}^j - 1 \leq \overline{m}^i$. Proposition 2 tells the reason for this.

**Proposition 2** *Let $\overline{m}^j$ be given. For any $m^i$, where $\overline{m}^i \geq m^i \geq \overline{m}^j + 1$,*

$$\phi^{i*}(t^i_{\overline{m}^i}, m^i) = \phi^{i*}(t^i_{\overline{m}^i}, m^i + 1)$$

**Proof.** Let $a^{i*} = \phi^{i*}(t^i_{\overline{m}^i}, m^i) \in \arg\max_{a^i \in A^i} U^i(a^i, (t^i_{\overline{m}^i}, m^i); \phi^{j*}, \mu^{j*})$. Then, $\forall a^i \in A^i$ we have:

$$\begin{aligned} & \sum_{(s, t^j_{\overline{m}^j}) \in S \times T^j_{\overline{m}^j}} \left[ u\left(a^{i*}; \phi^{j*}(t^j_{\overline{m}^j}, \mu^{j*}(t^j_{\overline{m}^j})), s\right) - C^i(m^i) \right] \cdot p(s, t^j_{\overline{m}^j}) \\ \geq & \sum_{(s, t^j_{\overline{m}^j}) \in S \times T^j_{\overline{m}^j}} \left[ u\left(a^{i}; \phi^{j*}(t^j_{\overline{m}^j}, \mu^{j*}(t^j_{\overline{m}^j})), s\right) - C^i(m^i) \right] \cdot p(s, t^j_{\overline{m}^j}), \end{aligned}$$

which makes:

$$\sum_{(s,t^j_{\overline{m}^j})\in S\times T^j_{\overline{m}^j}} \left[u\left(a^{i*};\phi^{j*}(t^j_{\overline{m}^j},\mu^{j*}(t^j_{\overline{m}^j})),s\right)\right]\cdot\delta^i_{\overline{m}^j+1}(s,t^j_{\overline{m}^j}) \tag{17}$$
$$\geq \sum_{(s,t^j_{\overline{m}^j})\in S\times T^j_{\overline{m}^j}} \left[u\left(a^{i};\phi^{j*}(t^j_{\overline{m}^j},\mu^{j*}(t^j_{\overline{m}^j})),s\right)\right]\cdot\delta^i_{\overline{m}^j+1}(s,t^j_{\overline{m}^j}).$$

When $m^i \geq \overline{m}^j + 1$, we have $p$ to be equal to $\delta^i_{\overline{m}^j+1} = \text{proj}_{\Delta X_{\overline{m}^j}} t^i_{m^i}$. This is also be the case for a type $(t^i_{\overline{m}^i}, m^i+1)$. With constant $A^i$ set, minus $C^i(m^i+1)$ both sides of (17), we have that $\forall a^i \in A^i$:

$$\sum_{(s,t^j_{\overline{m}^j})\in S\times T^j_{\overline{m}^j}} \left[u\left(a^{i*};\phi^{j*}(t^j_{\overline{m}^j},\mu^{j*}(t^j_{\overline{m}^j})),s\right)-C^i(m^i+1)\right]\cdot\delta^i_{\overline{m}^j+1}(s,t^j_{\overline{m}^j})$$
$$\geq \sum_{(s,t^j_{\overline{m}^j})\in S\times T^j_{\overline{m}^j}} \left[u\left(a^{i};\phi^{j*}(t^j_{\overline{m}^j},\mu^{j*}(t^j_{\overline{m}^j})),s\right)-C^i(m^i+1)\right]\cdot\delta^i_{\overline{m}^j+1}(s,t^j_{\overline{m}^j})$$

We have $a^{i*}$ is also an $\phi^{i*}(t^i_{\overline{m}^i}, m^i+1)$. The relationship in the opposite direction is also true (actually, both are the same problem). Hence, given $\phi^{j*}$and $\mu^{j*}$, $\phi^{i*}(t^i_{\overline{m}^i}, m^i) = \phi^{i*}(t^i_{\overline{m}^i}, m^i+1)$. ■

**Corollary 3** *It must be the case that* $\overline{m}^j - 1 \leq \overline{m}^i \leq \overline{m}^j + 1$, $\forall i$.

**Proof.** Firstly, at $m^i = \overline{m}^j + 1$, under the condition in proposition 2, with $C^i(m^i) > 0$ and $C^i(m^{i\prime}) > C^i(m^i)$, for $m^{i\prime} > m^i$, we have $\forall m^{i\prime} > m^i$:

$$U^i(\phi^{i*}(m^i),(t^i_{\overline{m}^i},m^{i\prime});\phi^{j*},\mu^{j*}) \geq U^i(\phi^{i*}(m^{i\prime}),(t^i_{\overline{m}^i},m^{i\prime});\phi^{j*},\mu^{j*}).$$

This is because we have $\phi^{i*}(m^i) = \phi^{i*}(m^{i\prime})$, while $C^i(m^{i\prime}) > C^i(m^i)$. Hence, each type $t^i_{\overline{m}^i}$ does not want to go up to the next order belief. Then, $\mu^{i*}(t^i_{\overline{m}^i})$ for each $t^i_{\overline{m}^i}$ is less than or equal to $\overline{m}^j + 1$, which makes $\overline{m}^i \leq \overline{m}^j + 1$.

Since the above must be true for both $i$ and $j$, $\overline{m}^j - 1 \leq \overline{m}^i \leq \overline{m}^j + 1$. ■

With corollary 3, our problem can be reduced a lot. In deciding which order of beliefs to stop processing, a type $t^i_{\overline{m}^i}$ needs to compare his current order of beliefs to ones up to his $\overline{m}^j + 1$-th order, which must be less than or equal to his $\overline{m}^i + 2$-th order according to corollary 3, only. There is no need for him to consider ones higher than the $\overline{m}^j + 1$-th order.

The below proposition 4 says that without processing cost $C^i(m^i)$, a type $t^i$ wants to climb up all his belief hierarchy.

**Proposition 4** *If $C^i(m^i) = 0 \ \forall i$ and $\overline{m}^j \geq m^i \geq 0$, $U^i(\phi^{i*}(m^i), (t^i_{\overline{m}^i}, m^i + 1); \phi^{j*}, \mu^{j*})$ is always less than or equal to $U^i(\phi^{i*}(m^i + 1), (t^i_{\overline{m}^i}, m^i + 1); \phi^{j*}, \mu^{j*})$. This means that the type $t^i$ would want to climb up all his belief hierarchy.*

**Proof.** Given $\overline{m}^j \geq m^i \geq 0$, functions $\phi^{j*}, \mu^{j*}$, and $C^i(m^i) = 0$, let $\phi^{i*}(t^i_{\overline{m}^i}, m^i) = a^{i*}_{m^i} \in A^i$. Then, from the definition of $\phi^{i*}(t^i_{\overline{m}^i}, m^i + 1)$, we have that $\forall a^i \in A^i$:

$$\begin{aligned} & \sum_{(s, t^j_{\overline{m}^j}) \in S \times T^j_{\overline{m}^j}} \left[ u\left(\phi^{i*}(t^i_{\overline{m}^i}, m^i + 1); \phi^{j*}(t^j_{\overline{m}^j}, \mu^{j*}(t^j_{\overline{m}^j})), s\right)\right] \cdot p(s, t^j_{\overline{m}^j}) \\ \geq & \sum_{(s, t^j_{\overline{m}^j}) \in S \times T^j_{\overline{m}^j}} \left[ u\left(a^i; \phi^{j*}(t^j_{\overline{m}^j}, \mu^{j*}(t^j_{\overline{m}^j})), s\right)\right] \cdot p(s, t^j_{\overline{m}^j}), \end{aligned}$$

or

$$U^i(\phi^{i*}(m^i + 1), (t^i_{\overline{m}^i}, m^i + 1); \phi^{j*}, \mu^{j*}) \geq U^i(a^i, (t^i_{\overline{m}^i}, m^i + 1); \phi^{j*}, \mu^{j*}).$$

Since $a^{i*}_{m^i} \in A^i$, we also have:

$$U^i(\phi^{i*}(m^i + 1), (t^i_{\overline{m}^i}, m^i + 1); \phi^{j*}, \mu^{j*}) \geq U^i(a^{i*}_{m^i}, (t^i_{\overline{m}^i}, m^i + 1); \phi^{j*}, \mu^{j*}).$$

The inequality holds for every realization of $(t^i_{\overline{m}^i}, m^i + 1)$. At any level of $m^i$, if there is no processing cost and a type $(t^i_{\overline{m}^i}, m^i)$ thinks that $\overline{m}^j \geq m^i$, he wants to climb up an additional order of belief. This is also true for any type $(t^j_{\overline{m}^j}, m^j)$ of player $j$. Hence, both players know this and want to climb up all their infinite hierarchies of beliefs. ■

From proposition 4, if there is no processing costs for each player, our model turns back to be a standard Bayesian game.

# 4 An Example

We provide an example in this section. We write $\{Z_1, z_1; Z_2, z_2; ...; Z_n, z_n\}$ to denote a probability distribution, in which event $Z_l$ has probability $z_l$ to occur, $l = 1, ..., n$. The probability distribution $\delta^i_{m^i,n}$ is the $n$-th possible $m^i$-th order belief of player $i$, which is equal to $\text{proj}_{X_{m^i-1}} t^i_{m^i,n}$.

**Example 1**: Figure 1 below shows the gross Bernoulli payoffs of player 1 and 2, respectively. The only uncertainty here is the payoffs of player 1 when he chooses action (row) $C$. We can call the left table as the $L$ state, and the right the $H$ state. Player 1 knows exactly in which state he is. Player 2 does not know that. We construct the type space for each player in this example after the payoffs table:

Figure 1: Bernoulli payoffs matrices of Example 1

| | A | B | C |
|---|---|---|---|
| A | 40,40 | 28,14 | 24,36 |
| B | 14,28 | 36,36 | 32,30 |
| C | 18,24 | 22,32 | 32,40 |

$L$

| | A | B | C |
|---|---|---|---|
| A | 40,40 | 28,14 | 24,36 |
| B | 14,28 | 36,36 | 32,30 |
| C | 32,24 | 33,32 | 34,40 |

$H$

- Alternative beliefs of player 1:
  - 1st-order beliefs:

    $\delta^1_{1,1} = \{S_1, 1\}$, and $\delta^1_{1,2} = \{S_2, 1\}$.

- 2nd-order beliefs:

  $\delta^1_{2,1} = \{(S_1, t^2_{1,1}), \frac{3}{4}; (S_1, t^2_{1,2}), \frac{1}{4}\}$, $\delta^1_{2,2} = \{(S_1, t^2_{1,1}), \frac{1}{4}; (S_1, t^2_{1,2}), \frac{3}{4}\}$,

  $\delta^1_{2,3} = \{(S_2, t^2_{1,1}), \frac{3}{4}; (S_2, t^2_{1,2}), \frac{1}{4}\}$, and $\delta^1_{2,3} = \{(S_2, t^2_{1,1}), \frac{1}{4}; (S_2, t^2_{1,2}), \frac{3}{4}\}$.

- 3rd-order beliefs:

  $\delta^1_{3,1} = \{(S_1, t^2_{1,1}, t^2_{2,1}), \frac{3}{4}; (S_1, t^2_{1,2}, t^2_{2,2}), \frac{1}{4}\}$,

  $\delta^1_{3,2} = \{(S_1, t^2_{1,1}, t^2_{2,1}), \frac{1}{4}; (S_1, t^2_{1,2}, t^2_{2,2}), \frac{3}{4}\}$,

  $\delta^1_{3,3} = \{(S_2, t^2_{1,1}, t^2_{2,1}), \frac{3}{4}; (S_2, t^2_{1,2}, t^2_{2,2}), \frac{1}{4}\}$,

  and $\delta^1_{3,4} = \{(S_2, t^2_{1,1}, t^2_{2,1}), \frac{1}{4}; (S_2, t^2_{1,2}, t^2_{2,2}), \frac{3}{4}\}$.

- Higher-order beliefs:

  Any coherent beliefs.

• Type space of player 1:

  - $t^1_{1,1} = \{\delta^1_{1,1}\}$, and $t^1_{1,2} = \{\delta^1_{1,2}\}$.
  - $t^1_{2,1} = \{\delta^1_{1,1}, \delta^1_{2,1}\}$, $t^1_{2,2} = \{\delta^1_{1,1}, \delta^1_{2,2}\}$, $t^1_{2,3} = \{\delta^1_{1,2}, \delta^1_{2,3}\}$, and $t^1_{2,4} = \{\delta^1_{1,2}, \delta^1_{2,4}\}$.
  - $t^1_{3,1} = \{\delta^1_{1,1}, \delta^1_{2,1}, \delta^1_{3,1}\}$, $t^1_{3,2} = \{\delta^1_{1,1}, \delta^1_{2,2}, \delta^1_{3,2}\}$, $t^1_{3,3} = \{\delta^1_{1,2}, \delta^1_{2,3}, \delta^1_{3,3}\}$, and $t^1_{3,4} = \{\delta^1_{1,2}, \delta^1_{2,4}, \delta^1_{3,4}\}$.
  - Types with higher-order beliefs can be ones with any coherent higher-order beliefs.

• Alternative beliefs of player 2:

  - 1st-order beliefs:

    $\delta^2_{1,1} = \{S_1, \frac{3}{4}; S_2, \frac{1}{4}\}$, and $\delta^2_{1,2} = \{S_1, \frac{1}{4}; S_2, \frac{3}{4}\}$.

  - 2nd-order beliefs:

    $\delta^2_{2,1} = \{(S_1, t^1_{1,1}), \frac{3}{4}; (S_2, t^1_{1,2}), \frac{1}{4}\}$, and $\delta^2_{2,2} = \{(S_1, t^1_{1,1}), \frac{1}{4}; (S_1, t^1_{1,2}), \frac{3}{4}\}$.

  - 3rd-order beliefs:

    $\delta^2_{3,1} = \{(S_1, t^1_{1,1}, t^1_{2,1}), \frac{9}{16}; (S_1, t^1_{1,1}, t^1_{2,2}), \frac{3}{16}; (S_2, t^1_{1,2}, t^1_{2,3}), \frac{3}{16}; (S_2, t^1_{1,2}, t^1_{2,4}), \frac{1}{16}\}$,

$\delta^2_{3,2} = \{(S_1, t^1_{1,1}, t^1_{2,1}), \frac{3}{16}; (S_1, t^1_{1,1}, t^1_{2,2}), \frac{9}{16}; (S_2, t^1_{1,2}, t^1_{2,3}), \frac{1}{16}; (S_2, t^1_{1,2}, t^1_{2,4}), \frac{3}{16}\}$,

$\delta^2_{3,3} = \{(S_1, t^1_{1,1}, t^1_{2,1}), \frac{3}{16}; (S_1, t^1_{1,1}, t^1_{2,2}), \frac{1}{16}; (S_2, t^1_{1,2}, t^1_{2,3}), \frac{9}{16}; (S_2, t^1_{1,2}, t^1_{2,4}), \frac{3}{16}\}$,

and $\delta^2_{3,4} = \{(S_1, t^1_{1,1}, t^1_{2,1}), \frac{1}{16}; (S_1, t^1_{1,1}, t^1_{2,2}), \frac{3}{16}; (S_2, t^1_{1,2}, t^1_{2,3}), \frac{3}{16}; (S_2, t^1_{1,2}, t^1_{2,4}), \frac{9}{16}\}$.

- 4th-order beliefs:

  $\delta^2_{4,1} = \{(S_1, t^1_{1,1}, t^1_{2,1}, t^1_{3,1}), \frac{9}{16}; (S_1, t^1_{1,1}, t^1_{2,2}, t^1_{3,2}), \frac{3}{16}; (S_2, t^1_{1,2}, t^1_{2,3}, t^1_{3,3}), \frac{3}{16};$
  $(S_2, t^1_{1,2}, t^1_{2,4}, t^1_{3,4}), \frac{1}{16}\}$,

  $\delta^2_{4,2} = \{(S_1, t^1_{1,1}, t^1_{2,1}, t^1_{3,1}), \frac{3}{16}; (S_1, t^1_{1,1}, t^1_{2,2}, t^1_{3,2}), \frac{9}{16}; (S_2, t^1_{1,2}, t^1_{2,3}, t^1_{3,3}), \frac{1}{16};$
  $(S_2, t^1_{1,2}, t^1_{2,4}, t^1_{3,4}), \frac{3}{16}\}$,

  $\delta^2_{4,3} = \{(S_1, t^1_{1,1}, t^1_{2,1}, t^1_{3,1}), \frac{3}{16}; (S_1, t^1_{1,1}, t^1_{2,2}, t^1_{3,2}), \frac{1}{16}; (S_2, t^1_{1,2}, t^1_{2,3}, t^1_{3,3}), \frac{9}{16};$
  $(S_2, t^1_{1,2}, t^1_{2,4}, t^1_{3,4}), \frac{3}{16}\}$,

  and $\delta^2_{4,4} = \{(S_1, t^1_{1,1}, t^1_{2,1}, t^1_{3,1}), \frac{1}{16}; (S_1, t^1_{1,1}, t^1_{2,2}, t^1_{3,2}), \frac{3}{16}; (S_2, t^1_{1,2}, t^1_{2,3}, t^1_{3,3}), \frac{3}{16};$
  $(S_2, t^1_{1,2}, t^1_{2,4}, t^1_{3,4}), \frac{9}{16}\}$.

- Higher-order beliefs:

  Any coherent beliefs.

- Type space of player 2:

  - $t^2_{1,1} = \{\delta^2_{1,1}\}$, and $t^2_{1,2} = \{\delta^2_{1,2}\}$.
  - $t^2_{2,1} = \{\delta^2_{1,1}, \delta^2_{2,1}\}$, and $t^2_{2,2} = \{\delta^2_{1,1}, \delta^2_{2,2}\}$.
  - $t^2_{3,1} = \{\delta^2_{1,1}, \delta^2_{2,1}, \delta^2_{3,1}\}$, $t^2_{3,2} = \{\delta^2_{1,1}, \delta^2_{2,1}, \delta^2_{3,2}\}$, $t^2_{3,3} = \{\delta^2_{1,2}, \delta^2_{2,2}, \delta^2_{3,3}\}$, and $t^2_{3,4} = \{\delta^2_{1,2}, \delta^2_{2,2}, \delta^2_{3,4}\}$.
  - Types with higher-order beliefs can be ones with any coherent higher-order beliefs.

According to the type spaces above, players are allowed to form only two alternative beliefs over the space of uncertainty. The first belief is that there is high possibility that the first event will occur, which is represented by the probability distribution

$\{Z_1, \frac{3}{4}; Z_2, \frac{1}{4}\}$. The second alternative is that there is low possibility that the first event will occur, which is represented by the probability distribution $\{Z_1, \frac{1}{4}; Z_2, \frac{3}{4}\}$. Hence, for example, player 2 with his first-order belief, can only be either $t^2_{1,1}$, or $t^2_{1,2}$, and the type $t^1_{1,1}$ of player 1 can become only either $t^1_{2,1}$, or $t^1_{2,2}$.

There are some orders of beliefs that players need not to form additional beliefs. These are ones that he can infer from his knowledge. For example, player 1 needs not to form his first-order belief, since he knows exactly his own payoffs. Player 2 needs not to form second-order beliefs, since he knows that player 1 knows his own payoffs, and, hence, if it is in state $L$, it must be the type $t^1_{1,1}$ that he faces, and if it is in state $H$, it must be the type $t^1_{1,2}$. We set up processing costs accordingly, as follow:

$$\begin{aligned} C^1(1) &= 0, C^1(2) = 1, C^1(3) = 1. \\ C^2(1) &= 1, C^2(2) = 1, C^2(3) = 3. \end{aligned}$$

We also assume that when there is no cost, players will process the next order of belief.

At this point the description of the example is complete. It can be checked that there is an equilibrium in this example in which:

- $\overline{m}^1 = 3$, and $\overline{m}^2 = 2$.
- Optimal actions for all types $(t^1_{\overline{m}^1}, m^1)$ of player 1 are:
  - $\phi^{1*}(t^1_{3,1}, 0) = \phi^{1*}(t^1_{3,2}, 0) = \phi^{1*}(t^1_{3,3}, 0) = \phi^{1*}(t^1_{3,4}, 0) = A$.
  - $\phi^{1*}(t^1_{3,1}, 1) = \phi^{1*}(t^1_{3,2}, 1) = A$, and $\phi^{1*}(t^1_{3,3}, 1) = \phi^{1*}(t^1_{3,4}, 1) = C$.
  - $\phi^{1*}(t^1_{3,1}, 2) = A$, $\phi^{1*}(t^1_{3,2}, 2) = C$, $\phi^{1*}(t^1_{3,3}, 2) = A$, and $\phi^{1*}(t^1_{3,4}, 2) = C$.
  - $\phi^{1*}(t^1_{3,1}, 3) = A$, $\phi^{1*}(t^1_{3,2}, 3) = C$, $\phi^{1*}(t^1_{3,3}, 3) = A$, and $\phi^{1*}(t^1_{3,4}, 3) = C$.
- Optimal actions for all types $(t^2_{\overline{m}^2}, m^2)$ of player 2 are:

- $\phi^{2*}(t^2_{2,1}, 0) = \phi^{2*}(t^2_{2,2}, 0) = C$.
- $\phi^{2*}(t^2_{2,1}, 1) = A$, and $\phi^{2*}(t^2_{2,2}, 1) = C$.
- $\phi^{2*}(t^2_{2,1}, 2) = A$, and $\phi^{2*}(t^2_{2,2}, 2) = C$.
- $\phi^{2*}(t^2_{3,1}) = A$, $\phi^{2*}(t^2_{3,2}) = A$, $\phi^{2*}(t^2_{3,3}) = A$, and $\phi^{2*}(t^1_{3,4}) = C$.
- $\phi^{2*}(t^2_{4,1}) = A$, $\phi^{2*}(t^2_{4,2}) = A$, $\phi^{2*}(t^2_{4,3}) = A$, and $\phi^{2*}(t^1_{4,4}) = C$.

- Optimal beliefs processing point for all types $t^1_{\overline{m}^1}$ of player 1 are:
  - $\mu^{1*}(t^1_{3,1}) = \mu^{1*}(t^1_{3,2}) = 1$, $\mu^{1*}(t^1_{3,3}) = 3$, and $\mu^{1*}(t^1_{2,4}) = 1$.
- Optimal beliefs processing point for all types $t^2_{\overline{m}^2}$ of player 2 are:
  - $\mu^{2*}(t^2_{2,1}) = 2$, and $\mu^{2*}(t^2_{2,2}) = 0$.
- The optimal outcome:
  - $\phi^{1*}(t^1_{3,1}, \mu^{1*}(t^1_{3,1})) = \phi^{1*}(t^1_{3,2}, \mu^{1*}(t^1_{3,2})) = A$, $\phi^{1*}(t^1_{3,3}, \mu^{1*}(t^1_{3,3})) = A$, and $\phi^{1*}(t^1_{3,4}, \mu^{1*}(t^1_{3,4})) = A$.
  - $\phi^{2*}(t^2_{2,1}, \mu^{2*}(t^2_{2,1})) = A$, and $\phi^{2*}(t^2_{2,2}, \mu^{2*}(t^2_{2,2})) = C$.

For player 1, if he knows that $\overline{m}^2 = 2$, his relevant beliefs are ones up to his 3rd order, according to corollary 3. Look at $(t^1_{3,1}, 3)$, given $\overline{m}^2$, $\phi^{2*}$, and $\mu^{2*}$, he believes that he is facing type $t^2_{2,1}$, whose optimal action $\phi^{2*}(t^2_{2,1}, \mu^{2*}(t^2_{2,1}))$ is $A$, with probability $\frac{3}{4}$, and facing type $t^2_{2,2}$, whose optimal action $\phi^{2*}(t^2_{2,2}, \mu^{2*}(t^2_{2,2}))$ is $C$, with probability $\frac{1}{4}$. He know exactly that he is in state $L$. With these knowledge and belief, his optimal action is $A$. We can find optimal action for each $(t^1_{3,l}, 3)$, $l = 2, 3, 4$, in the same way.

Then, for $(t^1_{3,l}, 2)$, given $\overline{m}^2$, $\phi^{2*}$, and $\mu^{2*}$, he uses alternative belief $d^1_3(t^1_{3,l}, 2)$, which is the same as $\delta^1_{3,l}$, for $l = 1, ..., 4$. Optimal action for $(t^1_{3,l}, 2)$, then, is the same as $(t^1_{3,l}, 3)$. For $(t^1_{3,1}, 1)$ (which is the same as $(t^1_{3,2}, 1)$), he uses alternative belief $d^1_3(t^1_{3,1}, 1)$, which can be calculated to be $\{(S_2, t^2_{1,1}, t^2_{2,1}), \frac{1}{2}; (S_2, t^2_{1,2}, t^2_{2,2}), \frac{1}{2}\}$. Then, his optimal action is $A$. We do the same way for $(t^1_{3,3}, 1)$ (which is the same as $(t^1_{3,4}, 1)$). At last, for $(t^1_{3,1}, 0)$ (which is the same as $(t^1_{3,l}, 1)$, for $l = 2, 3, 4$), he uses alternative belief $d^1_3(t^1_{3,1}, 0)$, which can be calculated to be $\{(S_1, t^2_{1,1}, t^2_{2,1}), \frac{1}{4}; (S_1, t^2_{1,2}, t^2_{2,2}), \frac{1}{4}; (S_2, t^2_{1,1}, t^2_{2,1}), \frac{1}{4}; (S_2, t^2_{1,2}, t^2_{2,2}), \frac{1}{4}\}$. His optimal action is $A$.

With these optimal decisions, conditions (11) and (12) can be used to check for the optimal belief processing point. For $t^1_{3,1}$ with his 1st-order belief, using the free alternative belief his optimal decision is $A$. If he uses his own 3rd-order belief, his optimal decision is still $A$, but he has to pay processing cost for one unit. Hence, he will choose to stay at his 1st-order belief. Contrary to $t^1_{3,3}$ with his 1st-order belief, using the free alternative belief his optimal decision is $C$. If he uses his own 3rd-order belief, his optimal decision is $A$. If he does not process his 3rd-order belief, he chooses $C$ and loss 3.5 units of forgone utility, which is higher than the additional processing cost of 2 units. Hence, he will decide to process his 3rd-order belief, paying for the processing cost, but save his forgone utility.

The optimal decisions and belief processing point of player 2 can be checked in the same way.

# 5 Discussion in Relation to Standard Bayesian Games

## 5.1 An advantage over Bayesian game with finite type space

In an attempt to model a standard Bayesian game with finite type space, a modeller must follow the following steps:[4]

*Step 1*: For decision-maker 1, assume that he must know that there are only 2 possible types for decision-maker 2, namely:

[4]Here, we consider the simplest case that there are only 2 possible types for each of the two decision-makers.

$$t^{2,1} = (\delta_1^{2,1}, \delta_2^{2,1}, \delta_3^{2,1}, ...),$$

and $$t^{2,2} = (\delta_1^{2,2}, \delta_2^{2,2}, \delta_3^{2,2}, ...).$$

Suppose that the $n^{th}$ belief is the first order of belief that $\delta_n^{2,1} \neq \delta_n^{2,2}$. Therefore if $n > 1$, it means that $\delta_l^{2,1} = \delta_l^{2,2}$ for $1 \leq l < n$.

*Step 2*: Construct hierarchy of beliefs for a type $t^{1,i}$ of player 1 with additional conditions that:

1. For $1 \leq l < n$, $\delta_{l+1}^{1,i}(\delta_l^{2,1}) = 1$.

2. For $l \geq n$, supp $\delta_{l+1}^{1,i} \subseteq \Delta X_l = \left\{ \text{proj}_{\Delta X_l} t^{2,1}, \text{proj}_{\Delta X_l} t^{2,2} \right\}$.

*Step 3*: Impose the same conditions to all types of each decision-maker.

These three steps prevent the expansion of type space into infinite space, while allow each player to form his infinite hierarchies of belief. This category of models is very restrictive. It requires that there are limited number of types (hierarchies of beliefs) available for each decision-maker. Moreover, decision-maker $i$ must know all the details of the hierarchy of beliefs of each type of player $j$. The assumption of common prior helps us to avoid exposing with these restrictive assumptions. However, common prior assumption has its own problems. Morris [6] clearly outlines arguments against the assumption.

Our model does not have this problem. We can use any type space without imposing the common knowledge assumption. Even we need finite state space, $S$, and we need players to form finite number of beliefs in each order, our model is more general in that it allows many different beliefs in each order, which will result in the infinite number of types for each player.

## 5.2 An advantage over general Bayesian games

In a general standard Bayesian game, in which a player does not have unique belief in every order, even if we start from a finite set $S$, the universal state space, $(S \times T^i \times T^j)$, is an infinite space. As discussed in section 3 of Brandenburger and Dekel [2], we will have infinite support for belief of a type $t^i$ of player $i$ over $(S \times T^j)$, $g^i(t^i)$. Thus, in making a decision, decision-maker $i$ must consider all the infinite types of $j$. The equilibrium outcome of the model is difficult the calculate, unless we impose some restrictions to make continuous strategy functions.

This category of models is interesting because it is closer to the reality. It is hard to believe that a person can know all the details of belief hierarchies of another person. However, allowing more general beliefs posts a problem of infinite number of types, and consequently infinite number of optimization problems, one for each type.

Our model solves this problem by forcing each player to use just finite belief hierarchy. As long as the state space and the number of different beliefs in each order are finite, we will face with finite number of optimization problems.

## 5.3 Problem with predictive power

Recent literature shows that higher-order uncertainty can affect the outcome of a game (For example, see Rubinstein [9], Weinstein and Yildiz [13], and Feinberg and Skrzypacz [3]). We can see this in example 1 of section 4 above that the type $t^1_{2,3}$ has different optimal action from the type $t^1_{2,4}$, even if the only difference between them is their different 2nd-order beliefs. This finding posts a serious problem to Bayesian games. To create a good model with predictive power, we have to specify correctly the whole infinite belief hierarchy of each type of each player. If not, the incorrect specification can divert the outcome of the game. Weinstein and Yildiz [13] and Morris, Postlewaite, and Shin [7] find some restriction to soothe this problem.

Our model can also reduce the problem, since players here depend on just finite belief hierarchies in making their own decisions. Hence, our model just requires correct specification of beliefs in these finite orders. However, our model is subject to another criticism that is we must specify correctly the processing cost functions of each player in order to allow our model to have predictive power[5]. A way to avoid this criticism is to put the cost functions into the space of uncertainty, $S$. This will expand our type space, but it does not affect other parts of the model.

### 5.4 Augmented Bayesian game with finite type space

If we augment the belief hierarchy of a type $t^i_{\overline{m}^i}$, by firstly for $l \leq \mu^{i*}(t^i_{\overline{m}^i})$, preserve the value of $\delta^i_l = \text{proj}_{\Delta X_{l-1}} t^i_{\overline{m}^i}$, and, secondly for $l > \mu^{i*}(t^i_{\overline{m}^i})$, augment the beliefs to be common knowledge of simple alternative beliefs, we can change our model into a standard Bayesian game with finite type space. Hence, this model can also provide an alternative interpretation of the Bayesian games with appropriate finite type spaces. A Bayesian game with appropirate finite type space can be thought of as a general game, in which players have costs in processing their own beliefs.

## 6 Conclusion

We propose an incomplete information game, in which rational players depend on finite belief hierarchies in determining their optimal actions. This is done by introducing processing costs to players in the model. This model goes along well with recent experimental literature, which shows that people depend on finite belief hierarchies in determining their actions. We discuss some advantages of our model over standard Bayesian games and a variant of bounded rationality models.

We still have some problems in interpreting our processing costs. Also, we realize that the presence of the cost function and its exact form can not be easily proven. This reduces the predictive power of our model, when we want to use it with some real issues.

[5]We thanks Professor Martin Richardson in pointing out this problem to us.

# References


[1] Robert J. Aumann. Rationality and bounded rationality. *Games and Economic Behavior*, 21:2–14, 1997.

[2] Adam Brandenburger and Eddie Dekel. Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, 59:189–198, 1993.

[3] Yossi Feinberg and Andrzej Skrzypacz. Uncertainty about uncertainty and delay in bargaining. *Econometrica*, 73(1):69–91, 2005.

[4] Kin Chung Lo. Equilibrium in beliefs under uncertainty. *Journal of Economic Theory*, 71:443–484, 1996.

[5] Jean-Francois Mertens and Shmuel Zamir. Formulation of bayesian analysis for games with incomplete information. *International Journal of Game Theory*, 14(1):1–29, 1985.

[6] Stephen Morris. The common prior assumption in economic theory. *Economics and Philosophy*, 11:227–253, 1995.

[7] Stephen Morris, Andrew Postlewaite, and Hyun Song Shin. Dept of knowledge and the effect of higher order uncertainty. *Economic Theory*, 6:453–467, 1995.

[8] Rosemarie Nagel. Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326, 1995.

[9] Ariel Rubinstein. The electronic mail game: A game with almost common knowledge. *American Economic Review*, 79:385–391, 1989.

[10] Matthew J. Ryan. What do uncertainty-averse decision-makers believe? *Economic Theory*, 20:47–65, 2002.

[11] Jozsef Sakovics. Games of incomplete information without common knowledge priors. *Theory and Decision*, 50(4):347–366, 2001.

[12] Dale O. Stahl and Paul W. Wilson. On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10:218–254, 1995.

[13] Jonathan Weinstein and Muhamet Yildiz. Impact of higher-order uncertainty. *Mimeo*, 2004.