Cognitively Stable Generalized Nash Equilibrium in Static Games with Unawareness

Yasuo Sasaki

School of Knowledge Science, Japan Advanced Institute of Science and Technology

Oct. 15, 2015
Risk: an agent knows all the contingencies relevant to the decision and is able to assign probabilities to them.

Ambiguity: an agent knows all the contingencies but has difficulties to evaluate them probabilistically.

Unawareness: an agent cannot even conceive all the contingencies.
Standard game theory assumes common knowledge (even in dealing with games with incomplete information). This implies that the agents are aware of all the game components such as agents, actions and states of the nature.

Recently models and solution concepts of games with unawareness have been developed to deal with asymmetric awareness. (Heifetz et al., 2013; Halpern and Rêgo, 2014)
What do we want to model?

- An agent may be unaware of something, she believes another agent may be unaware of something, and so on.
- Her behavior, and thus the equilibrium of the game, depends on such a hierarchy of perception.
Static games with unawareness

**Definition**

\[ \Gamma^U = (\mathcal{G}, \mathcal{F}) \] is a *static game with unawareness*, where:

- \( \mathcal{G} = \{ G^k \}_{k=0,...,n} \) is a set of static games, where \( G^0 \) is the objective game while \( G^1, ..., G^n \) are subjective games. For every \( k \), \( G^k = (N, A^k, u^k) \), where:
  - \( N \) is a finite set of the agents (common in every \( G^k \in \mathcal{G} \)).
  - \( A^k = \times_{i \in N} A^k_i \), where \( A^k_i \) is a finite set of \( i \)'s actions.
  - \( u^k = (u^k_i)_{i \in N} \), where \( u^k_i : A^k \to \mathbb{R} \) is \( i \)'s utility function.

- \( \mathcal{F} = (f_i)_{i \in N} \) is a collection of *awareness correspondences* for each agent. For any \( i \in N \), \( f_i : \mathcal{G} \to \mathcal{G} \). \( f_i(G^k) = G^l \) is interpreted as, “At \( G^k \), \( i \) believes the true game is \( G^l \).”

The model is a simplification of Halpern and Rêgo’s (2014) dynamic model.
C1-C4 are assumed so that the model works and makes sense:

- **C1**: For any $G^k, G^l \in \mathcal{G}$, if $G^k \succeq G^l$, then $A^k_i \supseteq A^l_i$ for any $i \in N$. ($G^k \succeq G^l$ means $G^l$ is reachable from $G^k$ with some awareness correspondence(s).)

- **C2**: For any $a \in A^0$, in any $G^k \in \mathcal{G}$, if $a \in A^k$, then $u^k_i(a) = u^0_i(a)$ for any $i \in N$.

- **C3**: For any $G^k \in \mathcal{G}$ and $i \in N$, if $f_i(G^k) = G^l$, then $f_i(G^l) = G^l$.

- **C4**: For any $i \in N$ and $G^k, G^l \in \mathcal{G}$, if $\mathcal{V}_i(G^k) \simeq \mathcal{V}_i(G^l)$, then $G^k = G^l$. ($\mathcal{V}_i(G^k)$ is $i$’s perception hierarchy at $G^k$. $\mathcal{V}_i(G^k) \simeq \mathcal{V}_i(G^l)$ means $\mathcal{V}_i(G^k)$ is equivalent to $\mathcal{V}_i(G^l)$.)
A two-agent game with unawareness (Feinberg, 2012):

- The objective (true) game is (a).
- Alice perceives (a), while she believes that Bob is unaware of her action, $a_3$, and therefore views (b). She also believes Bob believes (b) is common knowledge.
- Bob actually is aware of $a_3$ and views (a). Moreover he knows such a belief of Alice.
Formulation of the example

\[ G = \{ G^0, G^1, G^2 \}, \] where \( G^0 \) and \( G^1 \) are the same as (a), while \( G^2 \) is the same as (b).
Formulation of the example

Awareness correspondences can generate the agents’ perception hierarchies:

- $f_A(G^0) = G^1$: Alice believes the true game is $G^1$.
- $f_B \circ f_A(G^0) = G^2$: Alice believes Bob believes the true game is $G^2$.
- $f_A \circ f_B \circ f_A(G^0) = G^2$: Alice believes Bob believes Alice believes the true game is $G^2$.
- $\cdots f_B \circ f_A \circ f_B \circ f_A(G^0) = G^2$: Alice believes Bob believes $G^2$ is common knowledge.
Local actions and generalized action profile

In a static game with unawareness $\Gamma^U = (\mathcal{G}, \mathcal{F})$:

**Definition**

For every $i \in N$ and $G^k \in \mathcal{G}_i$, $\sigma_{(i,k)} \in \Delta A^k_i$ is called $i$’s *local action* in $G^k$, where:

- $\mathcal{G}_i = \{G^l \in \mathcal{G} \mid$ for some $G^k \in \mathcal{G}$, $f_i(G^k) = G^l\}$. (the set of static games that $i$ views as the true game somewhere in the model)
- $\Delta A^k_i$ is the set of $i$’s mixed actions in $G^k$.

**Definition**

Let $\sigma_i$ be a combination of agent $i$’s local actions in all the games in $\mathcal{G}_i$ and denote the set of such combinations by $\Sigma_i$. Then let us denote $\Sigma = \times_{i \in N} \Sigma_i$ and call $\sigma \in \Sigma$ a *generalized action profile*. 
Example

- $G_A = \{G^1, G^2\}$ and $G_B = \{G^0, G^2\}$.
- $\Sigma = (\Delta A^1_A \times \Delta A^2_A) \times (\Delta A^0_B \times \Delta A^2_B)$.

A local action (except for the one used in the objective game) should be interpreted as a belief about the agent’s choice rather than her actual choice.

Ex.) $\sigma_{(A,1)}$ is Alice’s choice in $G^1$ (which is used in the objective game), as well as Bob’s belief about Alice’s choice.
Generalized Nash equilibrium (GNE)

Definition (Halpern and Rêgo, 2014)

In a static game with unawareness $\Gamma^U = (G, F)$, $\sigma^* \in \Sigma$ is a generalized Nash equilibrium if and only if for every $i \in N$, $G^k \in G_i$ and $\sigma_{(i,k)} \in \Delta A^k_i$, $Eu^k_i(\sigma^*) \geq Eu^k_i(\sigma_{(i,k)}, \sigma^*_{-(i,k)})$, where:

- $Eu^k_i(\sigma)$ is $i$’s expected utility in $G^k \in G_i$ when $\sigma \in \Sigma$ is used.
- $\sigma_{-(i,k)}$ is all the local actions in $\sigma \in \Sigma$ other than $\sigma_{(i,k)}$.

GNE is such a generalized action profile $\sigma^*$ that, for every $i$ and $G^k$, if $i$ believes $G^k$ is the true game, then, in $G^k$, her local action is a best response to the others’ choices in $\sigma^*$.

In a GNE, the combination of each agent’s choice used in the objective game is called its objective outcome.
Example: Two (pure-action) GNEs of the game

- \( \sigma_1^* : (\sigma(A,1), \sigma(A,2)) = (a_2, a_2) \) and \( (\sigma(B,0), \sigma(B,2)) = (b_1, b_1) \)
- \( \sigma_2^* : (\sigma(A,1), \sigma(A,2)) = (a_3, a_1) \) and \( (\sigma(B,0), \sigma(B,2)) = (b_3, b_2) \)
An observation

See $\sigma^*_2: (\sigma_{(A,1)}, \sigma_{(A,2)}) = (a_3, a_1)$ and $(\sigma_{(B,0)}, \sigma_{(B,2)}) = (b_3, b_2)$. Once this equilibrium is played, some cognitive problems arise:

- Alice had expected that Bob would choose $b_2$. But he has chosen $b_3$ (which is weakly dominated in $G^2$).
- Alice believed Bob was unaware of $a_3$, so she notices using $a_3$ will give him new knowledge about the game.
- Bob had expected that Alice had expected that he would take $b_2$, so he notices that using $b_3$ may be a surprise to her.

\[
\begin{align*}
G^0 & \xrightarrow{f_A} G^1 & \xrightarrow{f_B} G^2
\end{align*}
\]
An observation

When $\sigma_1^*$, $(\sigma(A,1), \sigma(A,2)) = (a_2, a_2)$ and $(\sigma(B,0), \sigma(B,2)) = (b_1, b_1)$, is played, such problems do not arise because:

- Alice had expected that Bob would choose $b_1$, and indeed he chooses $b_1$;
- Alice considers Bob had expected that Alice would choose $a_2$, and indeed she chooses $a_2$;
- Alice considers Bob considers Alice had expected that Bob would choose $b_1$, and indeed he chooses $b_1$;
- And so on. (The same thing goes for Bob’s view.)

Yasuo Sasaki  
IUKM 2015
In general, GNE can be categorized into two classes:

- Every agent’s expectation about the others’ choices, every agent’s expectation about the others’ expectations about the others’ choices, and so on are the same as the actions used in the objective outcome. ($\sigma_1^*$)
  $\Rightarrow$ Each agent has no reason to change her behavior as well as her perception hierarchy: This class of GNE can be motivated as an equilibrium convention. (*cognitively stable*)

- This is not the case. ($\sigma_2^*$)
  $\Rightarrow$ Some cognitive problems may arise. An agent may update her expectation about the others’ choices or even revise some part of her perception hierarchy, and thus change her behavior next time: This class of GNE cannot be motivated as an equilibrium convention. (*cognitively unstable*)
Cognitively stable GNE

**Definition**

In a static game with unawareness $\Gamma^U = (\mathcal{G}, \mathcal{F})$, let $\sigma^* \in \Sigma$ be a GNE. Then it is **cognitively stable** if and only if for every $i \in N$ and $G^k, G^l \in \mathcal{G}_i$, $\sigma^*_{(i,k)} \equiv \sigma^*_{(i,l)}$, where:

- $\sigma^*_{(i,k)} \equiv \sigma^*_{(i,l)}$ means $\sigma^*_{(i,k)}$ and $\sigma^*_{(i,l)}$ are equivalent in the sense that they have common supports and moreover probabilities on them are all same.

Cognitively stable GNE is a class of GNE in which, for every agent $i$, the local actions in every game in $\mathcal{G}_i$ are all equivalent.
\(\sigma^*_1\) is cognitively stable, while \(\sigma^*_2\) is not.

- \(\sigma^*_1\): \((\sigma(A, 1), \sigma(A, 2)) = (a_2, a_2)\) and \((\sigma(B, 0), \sigma(B, 2)) = (b_1, b_1)\)
- \(\sigma^*_2\): \((\sigma(A, 1), \sigma(A, 2)) = (a_3, a_1)\) and \((\sigma(B, 0), \sigma(B, 2)) = (b_3, b_2)\)

Only \(\sigma^*_1\) can be motivated as an equilibrium convention.
Let us examine how unawareness can or cannot change the set of possible equilibrium conventions of the game.

**Definition**

In a static game with unawareness $\Gamma^U$, let $\Delta A^0 = \times_{i \in N} \Delta A_i^0$, and define:

- $E \subseteq \Delta A^0$: the set of Nash equilibria in $G^0$.
- $E_g \subseteq \Delta A^0$: the set of objective outcomes of GNEs.
- $E_c \subseteq \Delta A^0$: the set of objective outcomes of cognitively stable GNEs.

$\delta \in E$ ($\delta \in E_c$) is interpreted as a candidate of an equilibrium convention in the absence (presence, resp.) of unawareness.
Properties (general)

- \( E \neq \emptyset \) (Nash, 1950) and \( E_g \neq \emptyset \) (Halpern and Rêgo, 2014), but \( E_c \) may be empty. By definition, \( E_c \subseteq E_g \).

- In some game with unawareness, there exists \( \delta \in E_c \) such that \( \delta \notin E \): Unawareness can make a new equilibrium convention.

- In the previous example, \((a_2, b_1) \in E, (a_2, b_1), (a_3, b_3) \in E_g\) and \((a_2, b_1) \in E_c\).
Proposition 1

In a static game with unawareness $\Gamma^U = (G, F)$ in which, for every $i \in N$, $A_i^k = A_i^0$ when $f_i(G^0) = G^k$, $E_c \subseteq E$.

When every agent is aware of all of her own actions (i.e. her actions in $G^0$), always $E_c \subseteq E$. (a sufficient condition that unawareness cannot make a new equilibrium convention)
Proposition 2

In a static game with unawareness $\Gamma^U = (\mathcal{G}, \mathcal{F})$, $E_c \supseteq E'$, where:

- $E' = \{\delta \in E | \text{for every } i \in N \text{ and } G^k \in \mathcal{G}_i, \text{supp}(\delta) \subseteq A^k \text{ in } G^k\}$, where $\text{supp}(\delta)$ is the set of combinations of pure actions that can be played with positive probability under $\delta$.

When $G^0$ has a Nash equilibrium such that the existence of the outcome is common knowledge, it is in $E_c$: $\delta \in E'$ is always a candidate of an equilibrium convention under unawareness.
(Summary)
I have defined cognitive stability of GNE in static games with unawareness and shown some properties.

(Future works)
The analysis is just the first step and we need to characterize the concept in a more rigorous way.

- Cognitive stability, belief hierarchy and equilibrium convention.
- Extension to more general frameworks of games with unawareness.
Thank you for your attention.