Non-Equivalence between All and Canonical Elaborations

Satoru Takahashi

June 2018

Introduction

A seemingly small departure from complete information may have a large impact on strategic behavior.

- the e-mail game (Rubinstein, 1989),
- global games (Carlsson and van Damme, 1993),
- the structure theorem (Weinstein and Yildiz, 2007),
- ...

But not "anything goes."

Kajii and Morris (1997) introduce the notion of **robustness to all elaborations**.

- a strict Nash equilibrium may not be robust;
- some equilibrium is shown to be robust.

Elaborations

A complete-information game, denoted by $\mathbf{g}$, consists of

- a finite set of players, N;
- a finite set of actions, A_i;
- payoffs $g_i \colon A \to \mathbb{R}$.

An **elaboration** of $\mathbf{g}$, denoted by $(\mathbf{u}, P)$, is an incomplete-information game consisting of

- the same sets of players and actions as $\mathbf{g}$;
- a countable set of types, T_i;
- a common prior $P \in \Delta(T)$;
- type-dependent payoffs $u_i \colon A \times T \to \mathbb{R}$.

I say that $(\mathbf{u}, P)$ is an **ε-elaboration of g** if

$$P(\{t \in T \mid u_i(\cdot, t_i, t'_{-i}) = g_i \ \forall i \forall t'_{-i}\}) \geq 1 - \varepsilon.$$

Robustness to All Elaborations

An action profile a^* is **robust to all elaborations in g** if for any $\delta > 0$, there exists $\varepsilon > 0$ such that every ε-elaboration has a Bayesian Nash equilibrium that plays a^* with probability at least $1 - \delta$.

Kajii and Morris (1997) show that

- a game may have no robust equilibrium;
- sufficient conditions for robustness:
  - a unique correlated equilibrium;
  - a **p**-dominant equilibrium with $\sum_i p_i < 1$;
- a necessary condition for robustness: no other equilibrium is strictly **p**-dominant with $\sum_i p_i \leq 1$.

There is no known generic game with multiple robust equilibria.

A 2 × 2 Coordination Game

Two investors decide to invest on a project (I) or not (N):

$$\mathbf{g} = \begin{array}{c|c|c|} & I & N \\ \hline I & 1,1 & -2,0 \\ \hline N & 0,-2 & 0,0 \\ \hline \end{array} \quad .$$

Both (I, I) and (N, N) are strict Nash equilibria:

- (I, I) is not robust;
- (N, N) is robust.

A 2×2 Coordination Game: Non-Robustness of (I, I)

$$\mathbf{g} = \begin{array}{c|c|c|} & I & N \\ \hline I & 1,1 & -2,0 \\ \hline N & 0,-2 & 0,0 \\ \hline \end{array} \quad .$$

Consider the following "e-mail game" elaboration $(\mathbf{u}, P)$:

- $T_1 = T_2 = \mathbb{N}$;
- $P(k, k+1) = P(k+1, k) = \varepsilon(1-\varepsilon)^k/2$;
- type 0 has N as a dominant action; other types have the same payoff as g_i.

$(\mathbf{u}, P)$ is an ε-elaboration of $\mathbf{g}$, and "always N" is a unique Bayesian Nash equilibrium.

Hence other action profiles are not robust.

A 2×2 Coordination Game: Robustness of (N, N)

	I	N
I	$1, 1$	$-2, 0$
N	$0, -2$	$0, 0$

$\mathbf{g} =$ (the table above).

Kajii and Morris (1997) show that any **p**-dominant equilibrium with $\sum_i p_i < 1$ is robust to all elaborations.

Since (N, N) is a $(1/3, 1/3)$-dominant equilibrium, (N, N) is robust.

Canonical Elaborations

An elaboration $(\mathbf{u}, P)$ of $\mathbf{g}$ is **canonical** if every type is either

- a normal type: knows that his own payoff is the same as g_i, or
- a commitment type: has some action as a dominant action.

An action profile a^* is **robust to canonical elaborations in g** if for any $\delta > 0$, there exists $\varepsilon > 0$ such that every canonical ε-elaboration has a Bayesian Nash equilibrium that plays a^* with probability at least $1 - \delta$.

Equivalence?

By the definitions,

robust to all elaborations $\Rightarrow$ robust to canonical elaborations.

Does the converse hold? I do not know the answer.

- Whenever the literature establishes the non-robustness of some equilibrium in some game, it always uses canonical elaborations.
- Ui (2001) shows that if $\mathbf{g}$ is a potential game with a unique potential maximizer a^*, then a^* is robust to canonical elaborations. His proof relies on canonicality.
- Morris and Ui (2005) show that if $\mathbf{g}$ has a monotone potential and either $\mathbf{g}$ or the monotone potential is supermodular, then the potential maximizer is robust to all elaborations.
- Pram (2018) shows the equivalence when correlated equilibrium is used as a solution concept. His proof relies on the convexity of correlated equilibria.

The Result

I establish the non-equivalence between all and canonical elaborations by means of a counterexample, but based on set-valued notions of robustness.

A closed set $\mathcal{E} \subseteq \Delta(A)$ is **robust to all** (resp., **canonical**) **elaborations in g** if for any $\delta > 0$, there exists $\varepsilon > 0$ such that every (resp., canonical) ε-elaboration has a Bayesian Nash equilibrium that induces an action distribution in the δ-neighborhood of $\mathcal{E}$.

The Counterexample

Balkenborg and Vermeulen (2016) introduce a class of **minimal diversity games**.

With three players and two actions,

- $A_1 = A_2 = A_3 = \{0, 1\}$;
-
$$g_1(a) = g_2(a) = g_3(a) = \begin{cases} 0 & \text{if } a = (0,0,0) \text{ or } (1,1,1), \\ 1 & \text{otherwise.} \end{cases}$$

Let $\mathcal{E} = \Delta(A \setminus \{(0,0,0),(1,1,1)\})$.

Proposition 1. $\mathcal{E}$ is robust to canonical elaborations in $\mathbf{g}$.

Proposition 2. $\mathcal{E}$ is not robust to all elaborations in $\mathbf{g}$.

Proof of Proposition 1

The proof is essentially the same as that of Ui (2001).

Fix any canonical ε-elaboration $(\mathbf{u}, P)$. Consider

$$\max_{\sigma} \sum_{t \in T} P(t) g_1(\sigma(t)),$$

where max is taken over all strategy profiles σ where all commitment types play their dominant actions.

Since players have common payoffs, any maximizer σ^* is a Bayesian Nash equilibrium of $(\mathbf{u}, P)$.

Proof of Proposition 1, Continued

Let $\bar{\sigma}$ be the (possibly non-equilibrium) strategy profile where all normal types play $(0,0,1)$ and all commitment types play their dominant actions. Then

$$\begin{aligned}
\min_{\mu \in \mathcal{E}} \left\| \sum_t P(t)\sigma^*(t) - \mu \right\|_1 &= 2 \sum_t P(t)\sigma^*(t)(\{(0,0,0),(1,1,1)\}) \\
&= 2\left(1 - \sum_{t \in T} P(t) g_1(\sigma^*(t))\right) \\
&\leq 2\left(1 - \sum_{t \in T} P(t) g_1(\bar{\sigma}(t))\right) \leq 2\varepsilon.
\end{aligned}$$

Proof of Proposition 2: h

I use the following game $\mathbf{h} = (h_\alpha, h_\beta, *)$ among players α, β, and γ as a building block of my construction of elaborations:

$$\mathbf{h} = \begin{array}{c} \\ 0 \\ 1 \\ \\ \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{|c|c|} \hline 1, 0, * & 0, 1, * \\ \hline 0, 2, * & 2, 0, * \\ \hline \end{array} \\ 0 \end{array} \qquad \begin{array}{c} \\ 0 \\ 1 \\ \\ \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{|c|c|} \hline 2, 0, * & 0, 2, * \\ \hline 0, 1, * & 1, 0, * \\ \hline \end{array} \\ 1 \end{array} .$$

I denote by $\mathbf{h}^x$ the induced two-player game between players α and β given player γ's mixed action $x \in [0, 1]$ (x denotes the probability of action 1):

$$\mathbf{h}^x = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{|c|c|} \hline 1 + x, 0 & 0, 1 + x \\ \hline 0, 2 - x & 2 - x, 0 \\ \hline \end{array} \end{array} .$$

Game $\mathbf{h}^x$ has a unique equilibrium $((1 + x)/3, (1 + x)/3)$.

Proof of Proposition 2: $\tilde{\mathbf{h}}$

I also construct another game $\tilde{\mathbf{h}} = (\tilde{h}_\alpha, \tilde{h}_\beta, *)$ by relabeling player β's action 0 as action 1, and action 1 as action 0:

$$\tilde{\mathbf{h}} = \begin{array}{c|c|c|} & 0 & 1 \\ \hline 0 & 0,1,* & 1,0,* \\ \hline 1 & 2,0,* & 0,2,* \\ \hline \multicolumn{3}{c}{0} \end{array} \qquad \begin{array}{c|c|c|} & 0 & 1 \\ \hline 0 & 0,2,* & 2,0,* \\ \hline 1 & 1,0,* & 0,1,* \\ \hline \multicolumn{3}{c}{1} \end{array} \quad .$$

I denote by $\tilde{\mathbf{h}}^x$ the induced two-player game given player γ's mixed action $x \in [0,1]$:

$$\tilde{\mathbf{h}}^x = \begin{array}{c|c|c|} & 0 & 1 \\ \hline 0 & 0, 1+x & 1+x, 0 \\ \hline 1 & 2-x, 0 & 0, 2-x \\ \hline \end{array} \quad .$$

Game $\tilde{\mathbf{h}}^x$ has a unique equilibrium $((1+x)/3, (2-x)/3)$.

Proof of Proposition 2: Construction of $(\mathbf{u}, P)$

- $T_i = \{t_i^*, t_i^{i+1,f}, t_i^{i+2,f}, t_i^{i,s}, t_i^{i+1,s}\}$,
- $$P(t) = \begin{cases} 1-\varepsilon & \text{if } t = (t_1^*, t_2^*, t_3^*), \\ \varepsilon/9 & \text{if } t = (t_i^*, t_{i+1}^{i,f}, t_{i+2}^{i,f}), (t_i^{i,s}, t_{i+1}^{i,f}, t_{i+2}^{i,s}), \\ & (t_i^{i,s}, t_{i+1}^*, t_{i+2}^{i,s}) \text{ with some } i \in N, \\ 0 & \text{otherwise.} \end{cases}$$

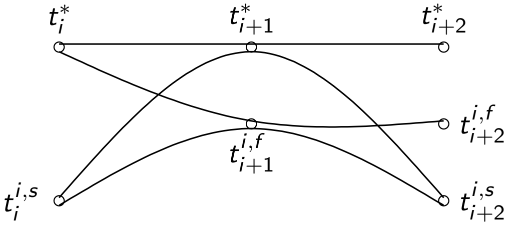
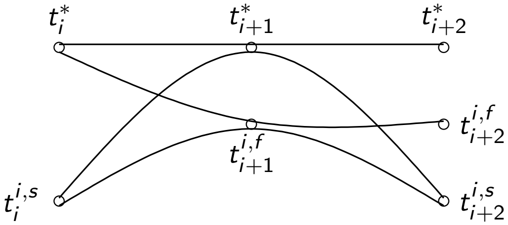


Figure: Interactions among t_i^*, t_{i+1}^*, t_{i+2}^*, $t_{i+1}^{i,f}$, $t_{i+2}^{i,f}$, $t_i^{i,s}$, and $t_{i+2}^{i,s}$.

Proof of Proposition 2: Construction of $(\mathbf{u}, P)$



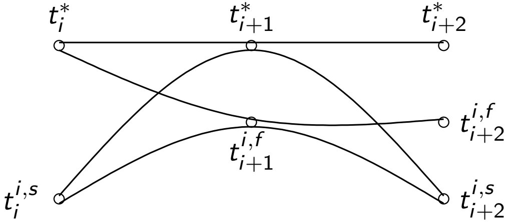
-

$$
\begin{aligned}
u_i(a, t_i^*, t_{i+1}, t_{i+2}) &= g_i(a), \\
u_{i+1}(a, t_i^*, t_{i+1}^{i,f}, t_{i+2}^{i,f}) &= \tilde{h}_\alpha(a_{i+1}, a_{i+2}, a_i), \\
u_{i+2}(a, t_i^*, t_{i+1}^{i,f}, t_{i+2}^{i,f}) &= \tilde{h}_\beta(a_{i+1}, a_{i+2}, a_i), \\
u_{i+1}(a, t_i^{i,s}, t_{i+1}^{i,f}, t_{i+2}^{i,s}) &= 0, \\
u_i(a, t_i^{i,s}, t_{i+1}^*, t_{i+2}^{i,s}) &= u_{i+2}(a, t_i^{i,s}, t_{i+1}^*, t_{i+2}^{i,s}) = 0, \\
u_i(a, t_i^{i,s}, t_{i+1}^{i,f}, t_{i+2}^{i,s}) &= h_\alpha(a_i, a_{i+2}, a_{i+1}), \\
u_{i+2}(a, t_i^{i,s}, t_{i+1}^{i,f}, t_{i+2}^{i,s}) &= h_\beta(a_i, a_{i+2}, a_{i+1}).
\end{aligned}
$$

Proof of Proposition 2: Lemma

Let σ be any Bayesian Nash equilibrium of $(\mathbf{u}, P)$.

Lemma If $\sigma_i(t_i^*) < 1/2$ (resp., $< 1/2$) and $\sigma_i(t_i^*) + \sigma_{i+2}(t_{i+2}^*) \leq 1$ (resp., ≥ 1), then $\sigma_{i+1}(t_{i+1}^*) = 1$ (resp., 0). In particular, if $\sigma_i(t_i^*) = 0$ (resp., 1), then $\sigma_{i+1}(t_{i+1}^*) = 1$ (resp., 0).



Proof of Lemma If $\sigma_i(t_i^*) < 1/2$, then $\sigma_{i+1}(t_{i+1}^{i,f}) < 1/2$ via $\tilde{\mathbf{h}}$.
So $\sigma_i(t_i^{i,s}) = \sigma_{i+2}(t_{i+2}^{i,s}) < 1/2$ via $\mathbf{h}$.
Hence $\sigma_{i+1}(t_{i+1}^*) = 1$. (Notice that $\sigma_i(t_i^{i+1,f}) + \sigma_{i+2}(t_{i+2}^{i+1,f}) = 1$.)

Proof of Proposition 2

Suppose that $\sigma_i(t_i^*) \neq 1/2$ for some $i \in N$.
Without loss of generality, I assume that $i = 1$ maximizes $|\sigma_i(t_i^*) - 1/2|$ and $\sigma_1(t_1^*) < 1/2$.
By the maximality and $\sigma_1(t_1^*) < 1/2$, I have

$$\frac{1}{2} - \sigma_1(t_1^*) = \left|\sigma_1(t_1^*) - \frac{1}{2}\right| \geq \left|\sigma_3(t_3^*) - \frac{1}{2}\right| \geq \sigma_3(t_3^*) - \frac{1}{2},$$

and hence $\sigma_1(t_1^*) + \sigma_3(t_3^*) \leq 1$.
By Lemma, I have $\sigma_2(t_2^*) = 1$.
Applying Lemma iteratively, I have $\sigma_3(t_3^*) = 0$ and hence $\sigma_1(t_1^*) = 1$, a contradiction.
Thus I have $\sigma_i(t_i^*) = 1/2$ for all $i \in N$.
(In fact, $(\mathbf{u}, P)$ has a unique Bayesian Nash equilibrium of "always play 50-50.")

Conclusion

I provide an example to show the non-equivalence between all and canonical elaborations.

Open questions:

- to prove or disprove the equivalence for singleton-valued robustness notions;
- to prove or disprove the equivalence for approximate robustness (Haimanko and Kajii, 2016).