Epistemic Logic and Inductive Game Theory

by M. Kaneko, 2012

Introductions to two projects and a bridge between them.

1. Prediction/Decision Making in a Game

- Infinite Regress arising from prediction/decision making
- Completion of the Regress argument
- **The Source for Common Knowledge**

2. False Common Beliefs/Knowledge

- Konnyaku Mondo（蒟蒻問答）- - Japanese Comic Story
- At the visual (superficial) level, they have common knowledge;
- At a mental (deeper) level, they mutually misunderstood what they communicated about.

3. Inductive Game Theory

- Experiential Sources for beliefs about the game structure
- Accumulation of experiences ➔ construction of an individual view
- Discriminatory behavior ➔ interpretations including prejudices.

Target Interactions

Inner Mental Worlds ⟺ Outer Experiential Worlds

Research Strategies:

- Logical Clarity - -Explication of Complicated Situation
- Separations of Different Aspects - - Elimination of Irrelevant Aspects
- Gathering Examples having Salient Features
- Integration

Prediction/Decision Making in a Game

1 Prediction/Decision Making is made from the ***ex ante*** perspective.

2 **Logical Inferences** are an **engine** for decision making - - proof theory.

3. An **Infinite regress** arises from **Prediction/Decision Making** in a game with multiple players.

Prisoner's Dilemma

Table 2.1; PD

	s_{21}	s_{22}
s_{11}	$(5,5)$	$(1,6)$
s_{12}	$(6,1)$	$(3,3)^{so}$

Battle of the Sexes:

Interactive Dissonance

Table 2.2; BS

	s_{21}	s_{22}
s_{11}	$(2,1)^{sb}$	$(0,0)$
s_{12}	$(0,0)$	$(1,2)^{sb}$

Decision Criteria from the *Ex Ante* Perspective

$N1^o$: P1 should chooses his best strategy against all of his predictions about P2's choice based on N2;

$N2^o$: P2 should chooses his best strategy against all of his predictions about P1's choice based on N1.

- $N1^o$ is not complete and needs to refer to $N2^o$;
- $N2^o$ is symmetric to $N1^o$.

- PL1 makes a decision based on these criteria.
- The problem is how PL1 infers a decision based on those criteria.

- **Key: Distinction between decisions and predictions.**

Game Theory Language: No Distinction

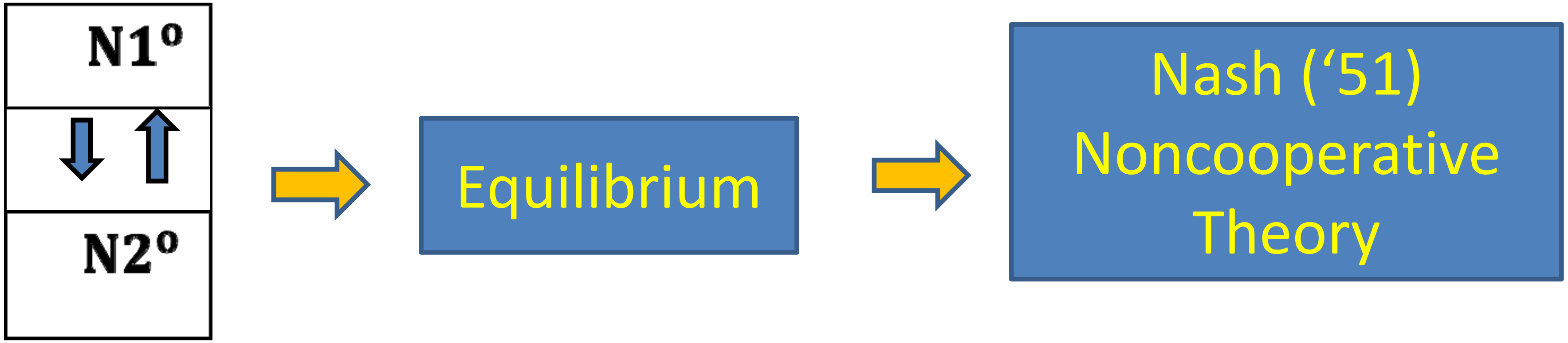


Table 2.1; PD

	s_{21}	s_{22}
s_{11}	$(5, 5)$	$(1, 6)$
s_{12}	$(6, 1)$	$(3, 3)^{so}$

Table 2.2; BS

	s_{21}	s_{22}
s_{11}	$(2, 1)^{sb}$	$(0, 0)$
s_{12}	$(0, 0)$	$(1, 2)^{sb}$

PD: Solvable Game
BS: Unsolvable game

Hu, T., and M. Kaneko (2012), Critical Comparisons between the Nash Noncooperative Theory and Rationalizability, SSM.DP.No.1287, University of Tsukuba.

Distinction with Belief Operators

B_1

$N1^o$		$B_2B_1(N1^o)$		$\cdots$
$\downarrow$	$\nearrow$	$\downarrow$	$\nearrow$	$\downarrow$
$B_2(N2^o)$		$B_2B_1B_2(N2^o)$		$\cdots$

Infinite Regress

Two approaches infinite regresses:

1. **Infinitary Logic Approach GL_ω : KD^n-type**
2. **Fixed Point Logic Approach**

Hu T.-W., and M. Kaneko, Infinite Regresses arising from Prediction/Decision Making in Games, in Preparation.

Logical Axioms and Inference Rules for $\mathbf{GL}_\omega$

L1: $A \supset (B \supset A)$;

L2: $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$;

L3: $(\neg A \supset \neg B) \supset ((\neg A \supset B) \supset A)$;

L4: $\wedge\Phi \supset A$, where $A \in \Phi$;

L5: $A \supset \vee\Phi$, where $A \in \Phi$;

$$\frac{A \supset C \quad A}{C}\ \text{(MP)} \qquad \frac{\{A \supset B : B \in \Phi\}}{A \supset \wedge\Phi}\ (\wedge\text{-rule}) \qquad \frac{\{B \supset A : A \in \Phi\}}{\vee\Phi \supset A}\ (\vee\text{-rule}).$$

K: $\mathrm{B}_i(A \supset C) \supset (\mathrm{B}_i(A) \supset \mathrm{B}_i(C))$;

D: $\neg\mathrm{B}_i(\neg A \wedge A)$;

$\wedge$-Barcan: $\wedge\mathrm{B}_i(\Phi) \supset \mathrm{B}_i(\wedge\Phi)$; $\qquad$ $$\frac{A}{\mathrm{B}_i(A)}\ \text{(Necessitation)}.$$

Infinite Regress: Let $D_i(p_1, ..., p_n) = \mathrm{B}_i(\wedge_{j \neq i} p_j), i \in N$, where $p_1, ..., p_n$ are propositional variables. Let $A = (A_1, ..., A_n)$ be an n-tuple of formulae in $\mathcal{P}_\omega$. The sequence $\langle \mathrm{Ir}_i^0(A), \mathrm{Ir}_i^1(A), ...\rangle, i \in N$ are generated

$$\mathrm{Ir}_i^0(A) = \mathrm{B}_i(A_i) \text{ for } i \in N; \text{ and } \mathrm{Ir}_i^{\nu+1}(A) = \mathrm{B}_i(\wedge_{j \neq i} \mathrm{Ir}_j^{\nu}(A)) \text{ for } i \in N. \quad (2.1)$$

$\mathrm{Ir}_1^0(A) = \mathrm{B}_1(A_1)$		$\mathrm{Ir}_1^1(A) = \mathrm{B}_1(\mathrm{Ir}_2^0(A)) = \mathrm{B}_1\mathrm{B}_2(A_2)$		$\cdots$
	$\nearrow$		$\nearrow$	
$\mathrm{Ir}_2^0(A) = \mathrm{B}_2(A_2)$		$\mathrm{Ir}_2^1(A) = \mathrm{B}_2(\mathrm{Ir}_1^0(A)) = \mathrm{B}_2\mathrm{B}_1(A_1)$		$\cdots$

- We define $\mathrm{Ir}_i(A) := \wedge\{\mathrm{Ir}_i^{\nu}(A) : \nu \geq 0\}$

- It is a subjective notion **in the mind of a single player**.
- Collapse Theorem: When we add Axiom T: $\mathrm{B_i}(A) \supset A$, the infinite regress collapses into the common knowledge, i.e., in $\mathbf{GL}_{\boldsymbol{\omega}}$**(T)**, $\vdash \mathbf{Ir}_{\boldsymbol{i}}(\boldsymbol{A_1}, \dots, \boldsymbol{A_n}) \equiv \mathbf{C}(\wedge_{\boldsymbol{i}} \boldsymbol{A_i})$.

Application to Prediction/Decision Making in a Game

$$\mathbf{B}_i \begin{cases} \mathbf{D0}_i^o : \wedge_{t \in S}[\mathrm{I}_i(t_i) \supset \langle \wedge_{j \neq i} \mathrm{B}_j(\mathrm{I}_j(t_j)) \supset \mathrm{best}(t_i; t_{-i}) \rangle]; \\ \mathbf{D1}_i^o : \wedge_{t_i \in S_i}[\mathrm{I}_i(t_i) \supset \wedge_{j \neq i} \mathrm{B}_j \mathrm{B}_i(\mathrm{I}_i(t_i))]. \\ \mathbf{D2}_i^o : \wedge_{t_i \in S_i}[\mathrm{I}_i(t_i) \supset \vee_{t_{-i} \in S_{-i}} \wedge_{j \neq i} \mathrm{B}_j(\mathrm{I}_j(t_j))]. \end{cases}$$

← Ni^o

Let $\mathrm{D0}_i := \mathrm{B}_i(\mathrm{D0}_i^o)$, $\mathrm{D1}_i := \mathrm{B}_i(\mathrm{D1}_i^o)$, and $\mathrm{D2}_i := \mathrm{B}_i(\mathrm{D2}_i^o)$.

We consider the infinite regresses:

$\mathrm{Ir}_i(\mathrm{D0}) := \mathrm{Ir}_i[\mathrm{D0}_1, ..., \mathrm{D0}_n]$, $\mathrm{Ir}_i(\mathrm{D1})$, and $\mathrm{Ir}_i(\mathrm{D2})$.

We characterize each $\mathrm{I_i}(s_i)$ by these infinite regresses.

- Best(…) is expressed by preferences (primitives);
- each $\mathrm{I_i}(s_i)$ is unknown.

$\mathrm{Ir}_i(\mathrm{D012})$: −Infinite Regress of Axioms D0, D1, D2:

WF_i−Choice of the Weakest Formulae satisfying D0, D1,

$\mathrm{Ir}_i(\mathrm{B}_1(g_1), \ldots, \mathrm{B}_n(\mathrm{g}_n))$ − Beliefs of the game structure.

We denote $\mathrm{Ir}_i[\mathrm{D012}]$,$\mathrm{Ir}_i[\mathrm{B}_1(g_1)$,$\mathrm{B}_2(g_2)]$,$\mathrm{WF}_i$ by Γ_i.

Theorem: When *g* satisfies Interchangeability, Γ_i determines $\mathbf{I}_i(s_i)$.

Theorem 1 (Playable): Let g be the Prisoner's Dilemma. Then, for $i = 1, 2,$

$$\Gamma_i \vdash \neg \mathrm{I}_i(\mathrm{s}_{i1}) \text{ and } \Gamma_i \vdash \mathrm{I}_i(\mathrm{s}_{i2}).$$

Interactive Dissonance

Theorem 2 (Incompleteness): Let g be the Battle of the Sexes. Then, for $i = 1, 2,$

$$\Gamma_i \nvdash \neg \mathrm{I}_i(\mathrm{s}_{i1}) \text{ and } \Gamma_i \nvdash \mathrm{I}_i(\mathrm{s}_{i2}).$$

Infinite Regress

- **Logical Completion of an interpersonal regress argument rather than reality.**
- **Subjective concept representing what is occurring in the mind of a single person.**
- **When the infinite regress happens to be objectively true, it is the common knowledge - - one source for CK.**

Objectivity vs. Subjectivity

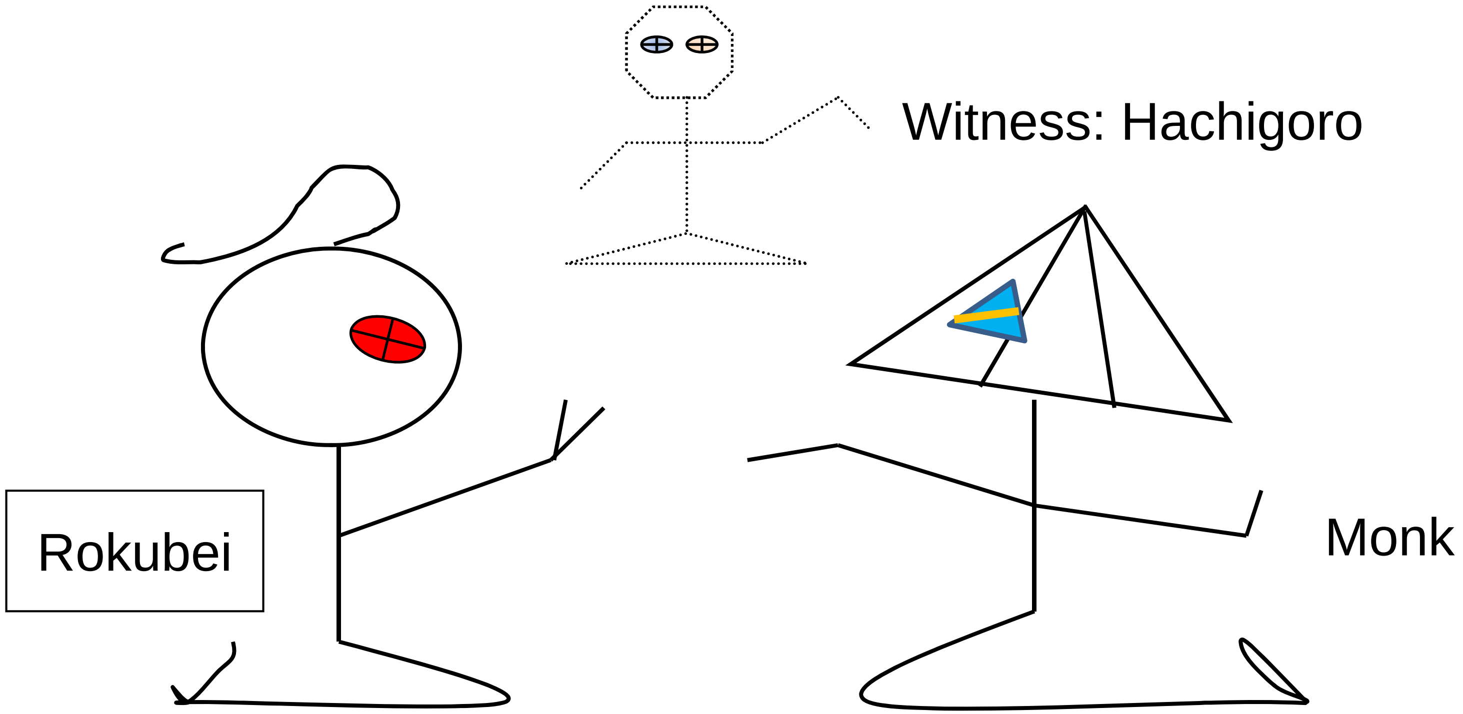
- **Each person's mind is separated from the other's.**
  - How different are the minds of different persons?
  - Visual Levels vs. Mental Levels

One Story called Konnyaku Mondo is suggestive.

Konnyaku Mondo (蒟蒻問答)

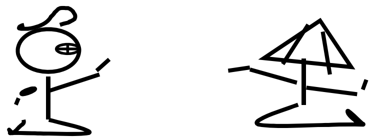
Kaneko '04, Act 2, *Game Theory and Mutual Misunderstanding*.

- **Other's Mind and False Beliefs**:
- CK as shared information, but
- **Common Misunderstanding** in interpretations of exchanged gestures



Konnyaku - - Devil's Tongue Jelly Product
Konnyaku Mondo – a Japanese comic story in the Rakugo-style

Konnyaku Mondo 1



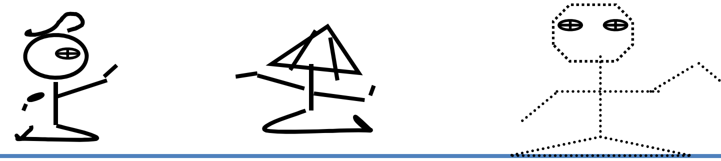
There was a temple where no monks were living any longer. A devil's tongue jelly maker, named Rokubei, lived next door. He moved into the temple and started pretending to be a monk.

One day, a traveling Zen Buddhist monk passed by and challenged Rokubei to a debate on Buddhism. Rokubei had no knowledge on Buddhism and was not able to have a debate. He tried to refuse, but he couldn't escape and finally agreed.

The Buddhist dialogue started but Rokubei didn't know how to perform and he kept silent. The Buddhist monk tried to communicate to Rokubei in many ways.

After some time, Rokubei started responding with gestures to the body movements the monk made.

Konnyaku Mondo 2



The monk took this as a style of dialogue and tried to answer in gestures, too. They exchanged gestures, and after some time, the monk told Rokubei, “your thoughts are profound and mine are of no comparison. I am very sorry to have bothered you”.

After saying this, he left the temple.

Hachigoro, a neighbor of Rokubei, witnessed the whole debate, and followed the monk to ask what had happened.

The monk answered, "I’m not trained enough in Buddhist thoughts to compete with that master. Please convey to him my earnest apology for having left so abruptly". Almost as quickly as the words had left his lips, he ran away.

Hachigoro returned to the temple and asked Rokubei if he knew anything about Buddhism thoughts.

Rokubei answered, ‘No, I have no idea about Buddhism, but the guy is, in fact, a beggar and he talked badly about my jelly products. That’s why I gave him a lesson’.

Interpretations and Implications

- **CK between R and M: At the visual level**
  - Exchanged gestures
  - The fact that R has beaten M.

- **Mutual Misunderstanding: At a mental level**
  - Interpretations of the gestures exchanged
  - The reasons for how R has succeeded in beating M.

Very Basic Question: How does a player know the game structure?

Either EL or GT starts with given beliefs/knowledge.

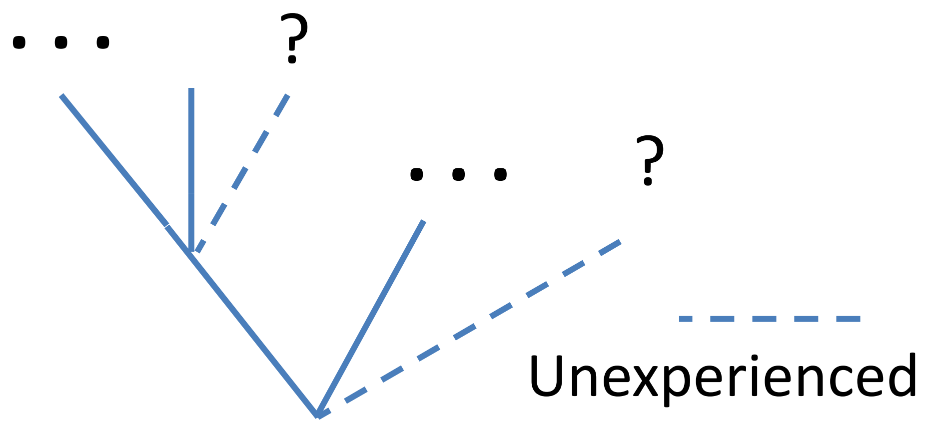
Inductive Game Theory

- no (little) a priori knowledge
- experiences – trials and errors
- accumulation of partial experiences
- partial understanding of the situation
- revision of behavior/beliefs

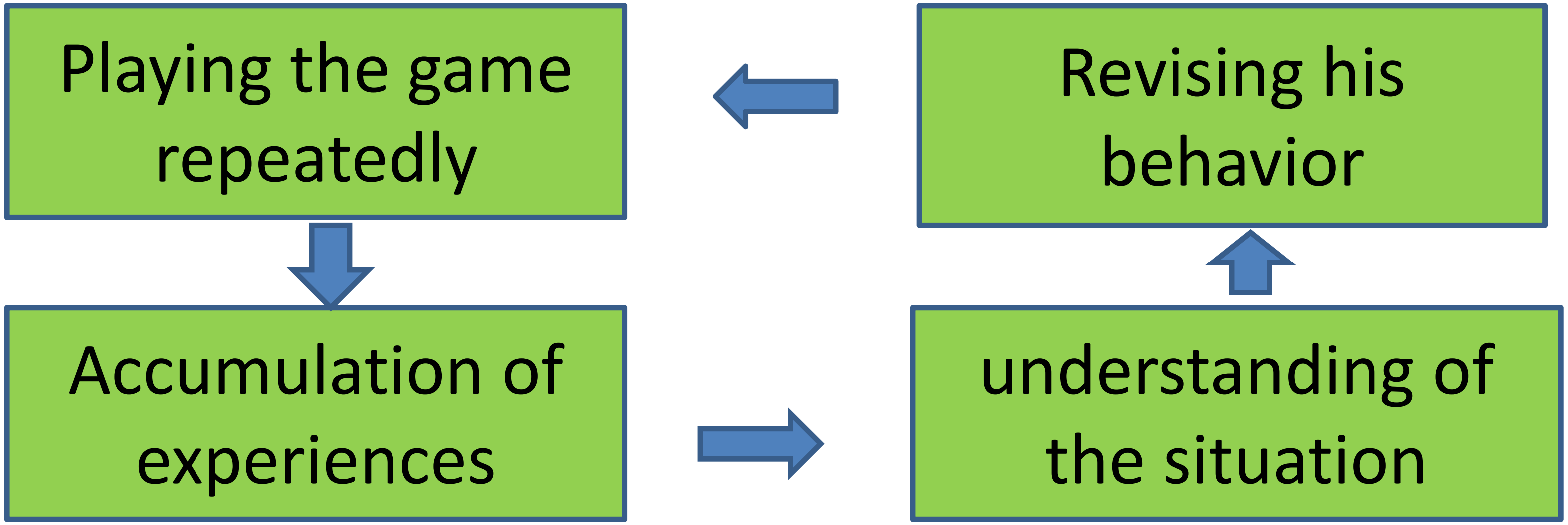
M. Kaneko, J. J. Kline, Inductive Game Theory: A Basic Scenario, J. Math. Economics 44 (2008) 1332-1362. Corrigendum (the same journal, 46 (2010) 620-622

	a	b
a	?, ?	1, ?
b	6, ?	3, ?

The other person's Positions:
Role-switching



Incorporations of Partialities and Restrictions



Theoretical Works:

- Inductive Constructions of Individual Views form Experiences
- New Definition of an Extensive Game
- The mind of another person

Computer Simulations

Experimental Study.

Japanese Example: the Buraku (Village) People

- The Buraku people consist of, **officially,** 1% of the total population (125millions).
- They have been discriminated against by the other people in various occasions, e.g., marriage, job-search etc.
- They have indistinguishable looking from other people, but those people are distinguished for their origins.

Typical Example of Discrimination

Suppose:

- A boy (non-Buraku), **B**, and a girl (Buraku), **G**, have falling in love, and decided to get married.

- B brought G to see the parents in order to obtain permission for their marriage. The parents liked G and were pleased very much.
- But it was customary for a rich family to hire a personal investigator to look into the family background. Then, he found G is from Buraku.
- The parents started objecting their marriage; they may give some reasons for it.

- Discrimination exists → prejudices are induced after the fact - - Buraku people are contaminated, or simply, PA dislike them.

Kaneko, M., and A. Matsui, (1999), Inductive game theory: discrimination and prejudices, *Journal of Public Economic Theory* 1, 101-137.

Conclusions

- EL approach provides a lot of important concepts:
  - logical Inferences to prediction/decision making
  - Objectivity vs subjectivity
  - A lot of false elements
- Konnyaku Mondo suggests a lot about interpersonal subjective thinking and reality
- IGT provides experiential foundations of game theory/social sciences

- Further Developments
  - Many notions of bounded cognitive and epistemic abilities.
  - Unified approach of the EL and IGT.