Epistemic Logic and Game Theory

Tai-Wei Hu

Kellogg School of Management,
Northwestern University

Rationality in economics

Dimensions in perfect rationality

- Knowledge of preference/environment
  - knowledge of potential *needs*
  - knowledge of potential *goods*
  - knowledge of the causal relation b/t goods and satisfaction of needs
- Logical ability
  - perfect ability to conduct logical inferences
  - perfect ability to make contingent plans
  - free from logical inconsistency

Logic and Economics

Bounded rationality

- Economics of information/knowledge
  - incomplete information about taste or goods
  - information processing
  - incentive structures
- Complexity and epistemic logic
  - imperfect ability of logical inferences
  - imperfect ability to contemplate all contingent plans

Classical Logic

Formal model of logical inference

- precise meaning of *true* thoughts
- *theory* of theories

Logical inferences and ‘theorems’ as objects of study

- formalize the notion of ‘valid argument’
- formalize the notion of ‘proofs’

Language in CL

Primitive symbols

- Propositional variables $PV = \{p_0, p_1, ..., p_k, ...\}$
- Logical connectives: $\neg$, $\Rightarrow$
- Belief operators: $B_1, B_2, ..., B_n$
- Parentheses: (,)

Formulas

- (F1) $p \in PV$ is a formula
- (F2) if A and B are formulas, so are $(\neg A)$, $(A \Rightarrow B)$, and $B_i(A)$
- (F3) every formula is obtained by a finite number of applications of (F1) and (F2)
- a formula is nonepistemic if it contains no $B_1, ..., B_n$

The set of formulas is denoted $\mathcal{P}$ and set of nonepistemic formulas is denoted $\mathcal{P}^n$

Classical Semantics

A model is a function $\kappa : PV \to \{\top, \bot\}$

- V_κ extends κ to $\mathcal{P}^n$
  - for $p \in PV$, $V_\kappa(p) = \top$ if and only if $\kappa(p) = \top$
  - $V_\kappa(\neg A) = \top$ if and only if $V_\kappa(A) = \bot$
  - $V_\kappa(A \Rightarrow B) = \top$ if and only if $V_\kappa(A) = \bot$ or $V_\kappa(B) = \top$
- κ is a model for a set Γ of formulas if for all $A \in \Gamma$, $V_\kappa(A) = \top$
- $\Gamma \models A$ if and only if for every model κ of Γ, $V_\kappa(A) = \top$

A formula A is *valid*, denoted $\models A$, if and only if $V_\kappa(A) = \top$ for every model κ

Axioms and inference for CL

Axiom schemata and inference rule

- Axioms
  - (L1) $A \Rightarrow (B \Rightarrow A)$
  - (L2) $(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$
  - (L3) $((\neg A) \Rightarrow (\neg B)) \Rightarrow (((\neg A) \Rightarrow B) \Rightarrow A)$
- Inference Rule: from $(A \Rightarrow B)$ and A infers B

Abbreviations

- $A \vee B$ stands for $\neg A \Rightarrow B$
- $A \wedge B$ stands for $\neg(\neg A \vee \neg B)$
- $A \equiv B$ stands for $(A \Rightarrow B) \wedge (B \Rightarrow A)$

Proofs in CL

A *proof* of A from a set of formulas Γ is a finite tree such that

- each node is associated with a formula in $\mathcal{P}^n$
- a leaf is either an axiom of a formula in Γ
- adjoining nodes together form an instance of the inference rule
- A is associated with the root

If there is a proof for A from Γ, we say that A is *provable* from Γ, denoted by $\Gamma \vdash A$

- A is a theorem if there is a proof for A
- theorems and proofs as objects of study

Completeness and soundness

We say that a set of formulas Γ is *inconsistent* if $\Gamma \vdash (C \wedge \neg C)$ for some C

Theorem (Completeness and soundness for CL)

Let Γ be a set of formulas in $\mathcal{P}^n$ and A be a formula.
(1) $\Gamma \vdash A$ if and only if $\Gamma \models A$.
(2) There is a model κ for Γ if and only if Γ is consistent.

Remarks.

- Assertions (1) and (2) are equivalent
- The 'only if' part is called *soundness*, and the 'if' part is called *completeness*.
- Equivalence between provability and validity
- Implies that propositional CL is *decidable*

Epistemic Logics

Epistemic axioms and inference rule

- K: $B_i(A \Rightarrow C) \Rightarrow (B_i(A) \Rightarrow B_i(C))$
- D: $\neg B_i(\neg A \wedge A)$
- T: $B_i(A) \Rightarrow A$
- 4: $B_i(A) \Rightarrow B_i(B_i(A))$
- 5: $\neg B_i(A) \Rightarrow B_i(\neg B_i(A))$
- Necessity: from A infers $B_i(A)$

Various epistemic logics:

- K^n : $CL + K + Nec$
- KD^n : $K^n + D$; KT^n : $K^n + T$
- $KD4^n$: $KD^n + 4$; $S4^n$: $K4^n + T$
- $KD45^n$: $KD4^n + 5$; $S5^n$: $S4^n + 5$

Interpretation and evaluation of epistemic axioms

Basic principles for beliefs

- (G) PLi believes A iff i has an argument for A from basic beliefs
- (G1) i has reasoning ability described by CL
- (G2) i has introspection ability on his own ability described by (G1) and (G2)
- (G3) when thinking about other's beliefs, i assumes (G1-G3) for other players

Correspondence between basic principles and axioms

- (G1) corresponds to knowledge of logical axioms (L1-L3) and (K)
- (G2) corresponds to $KD4$ for single players
- (G1-G3) corresponds to $KD4^n$

Kripke Semantics

A *Kripke frame* is a list $\mathcal{K} = (W, R_1, ..., R_n)$:

- W is the set of possible worlds
- R_i is a binary relation on W, interpreted as the *accessibility relation*.

A *Kripke model* is a pair $(\mathcal{K}, \sigma)$ of a frame and an assignment $\sigma : W \times PV \to \{\top, \bot\}$, which can be extended to $W \times \mathcal{P}$ as follows:

- if $p \in PV$, then $(\mathcal{K}, \sigma, w) \models p$ iff $\sigma(w, p) = \top$
- $(\mathcal{K}, \sigma, w) \models \neg A$ iff $(\mathcal{K}, \sigma, w) \not\models A$
- $(\mathcal{K}, \sigma, w) \models A \Rightarrow B$ iff $(\mathcal{K}, \sigma, w) \not\models A$ or $(\mathcal{K}, \sigma, w) \models B$
- $(\mathcal{K}, \sigma, w) \models B_i(A)$ iff $(\mathcal{K}, \sigma, w) \models A$ for all u such that wR_iu

Epistemic axioms and conditions on accessibility

- No condition $\leftrightarrow K$
- Seriality $\leftrightarrow D$
  - for any $w \in W$, there exists some u such that $wR_i u$
- Reflexibility $\leftrightarrow T$
  - for any $w \in W$, $wR_i w$
- Transitivity $\leftrightarrow 4$
  - for any $u, v, w \in W$, $wR_i u$ and $uR_i v$ imply $wR_i v$
- Euclidean $\leftrightarrow 5$
  - for any $u, v, w \in W$, $wR_i u$ and $wR_i v$ imply $uR_i v$

Soundness and completeness

Theorem

$\vdash_{KD4^n} A$ *if and only if* $(\mathcal{K}, \sigma, w) \models A$ *for any Kripke frame* $\mathcal{K}$ *and any assignment* σ *and any* $w \in W$ *such that* R_i *is serial and transitive for all* i.

Remarks.

- the theorem holds for any epistemic logic we listed
- the inference is made by the *outside observer*; however, a parallel version for each player's mind is possible

Decision criterion and predictions

Consider the following criteria for decisions and predictions:

- (N1): player 1 chooses his best strategy against *all* of his predictions about player 2's choice based on (N2)
- (N2): player 2 chooses his best strategy against *all* of his predictions about player 1's choice based on (N1)

Remarks.

- Ideal criterion leads to circular definition
- Common knowledge is involved to obtain a solution for this criterion
- Alternative:
  - play a default strategy
  - dominant strategies
  - best response against dominant strategies
  - play Nash equilibrium strategies

Common Knowledge Logic

Let C be the common knowledge operator

Syntax

- axiom and inference rule
  - (CA) $C(A) \Rightarrow A \wedge B_1(C(A)) \wedge \ldots \wedge B_n(C(A))$
  - (CI) from $D \Rightarrow A \wedge B_1(D) \wedge \ldots \wedge B_n(D)$ infer $D \Rightarrow C(A)$
- if $\vdash D \Rightarrow B_e(A)$ for all $e = (i_1, \ldots, i_m)$, then $\vdash D \Rightarrow C(A)$

Semantics

- $(\mathcal{K}, \sigma, w) \models C(A)$ if and only if $(\mathcal{K}, \sigma, w) \models A$ for all u reachable from w, i.e., for all u such that there is a sequence $w = w_0, w_1, \ldots, w_m = u$ with the property that for all k, $w_k R_j w_{k+1}$ for some j

Soundness and completeness holds in the common knowledge logic

Epistemic conditions for Nash theory

(N1) and (N2) can be formalized as following:

- (Ni1) $I_i(s_i) \Rightarrow (\bigvee_{s_{-i} \in S_{-i}} I_{-i}(s_{-i}))$
- (Ni2) $I_i(s_i) \Rightarrow B_i(I_i(s_i))$
- (Ni3) $I_i(s_i) \Rightarrow \bigwedge_{s_{-i} \in S_{-i}} (I_{-i}(s_{-i}) \Rightarrow \text{Best}_i(s_i; s_{-i}) \wedge B_i(I_{-i}(s_{-i})))$
- $(Ni) = (Ni1) \wedge (Ni2) \wedge (Ni3)$, $i = 1, 2$

Theorem

Let G be a 2-person game with interchangeability in pure strategies.
(1) $C(N1 \wedge N2)$, RN, $C(g) \vdash \wedge_{s_1, s_2}[I_1(s_1) \wedge I_2(s_2) \equiv C(Nash(s_1, s_2))]$.
(2) for $i = 1, 2$,
$C(N1 \wedge N2)$, RN, $C(g) \vdash \wedge_{s_i}[I_i(s_i) \equiv \vee_{t_{-i}} C(Nash(s_i; t_{-i}))]$.

Conclusion

Ex ante decision making in games: the idealized case

- Nash solution is a result of
  - common knowledge of game structure and payoffs
  - common knowledge of criteria for decision and prediction
  - perfect logical abilities
  - unbounded ability in interpersonal inferences
- but Nash solution may *not* exist

Bounded rationality

- Lack of common knowledge
- Complexity of logical inferences
- Complexity of interpersonal inferences