Logical Dynamics I

Yanjing Wang

Department of Philosophy, Peking University

Sept. 15, 2014

A Card Game

Three cards J, Q, K are randomly distributed by D to three people A, B, C. Now agent A asks agent B do you have...? Then agent B answers truthfully. Now try to answer the following questions (please keep the answers to yourself!):

1. Who knows the actual distribution of the cards now?
2. Does C think that B knows the actual distribution?
3. Does C think that A knows the actual distribution?
4. What is the actual distribution?

A Card Game

1. Who knows the actual distribution of the cards now?
2. Does *C* think that *B* knows the actual distribution?
3. Does *C* think that *A* knows the actual distribution?
4. What is the actual distribution?

A	**B**	**C**	**D**
C	A,B	C	A,B
No	Yes	No	Yes
No	Yes	No	Yes
?			?

For me (*A*), I do not know the distribution but I know that *C* now knows and *B* doesn't.

Now let's see the actual distribution of cards.

A Card Game

Teachers often ask questions while they do know the answers, but sometimes they indeed do not know the answer... You just never know :)

We can simplify the situation by letting B announce: I do not have...

1. Who knows the actual distribution of the cards now?
2. Does C think that B knows the actual distribution?
3. Does C think that A knows the actual distribution?
4. Does C know whether B knows whether A knows the distribution?
5. Does A know whether B knows whether C knows the distribution?

A Card Game

- The example involves: multiple agents, knowledge, belief, communication, protocol, observation, strategy, intention...
- To reason about knowledge and belief based on interaction requires “higher-order” reasoning power (Theory of Mind).
- We need a reliable way to reason correctly when facing complicated situations.
- There must be some rules governing the valid reasoning in such scenarios, which we are going to discover.

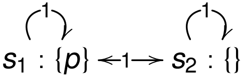
Let's try to get a little bit more formal...

We may represent the situation in a graph...The question itself does not carry information. The answer is an announcement "B does not have the card ..." What should happen afterwards?

1. Who knows the actual distribution of the cards now?
2. Does B know whether A knows the actual distribution?
3. Does A know whether C knows the distribution?
4. Does C know whether A knows the actual distribution?
5. Does D know whether C knows the actual distribution?
6. Does C know whether B knows whether A knows the distribution?
7. Does A know whether B knows whether C knows the distribution?

The idea: Graph+Changes (extended mind)

Kripke models (Relational models)



Kripke models $\mathcal{M} = (S, P, I, \{\rightarrow_i\}_{i \in I}, V)$ where:

- S is a non-empty set of "(epistemically) possible worlds".
- P is a non-empty set of proposition symbols, e.g.,$\{p, q, \ldots\}$.
- I is a non-empty set of agent names.
- $\rightarrow_i \subseteq S \times S$ is a binary relation.
- $V : S \rightarrow 2^P$, e.g., $V(s_1) = \{p\}$, $V(s_1) = \{\}$.

An *epistemic (S5) model* is a Kripke model where $\rightarrow_i$ is an equivalence relation for each $i \in I$ (often write $\sim_i$).

Let's get a little bit more formal...

Consider the following *epistemic* language (Hintikka 62, von Wright 1951):

$$\phi \quad ::= \quad \top \mid p \mid \neg\phi \mid (\phi \wedge \phi) \mid K_i\phi$$

where $p \in P$ and $i \in I$, with the following semantics on Kripke models $\mathcal{M} = (S, P, I, \{\rightarrow_i\}_{i \in I}, V)$:

$$\begin{array}{rcl} \mathcal{M}, s \models \top & \Leftrightarrow & \text{always} \\ \mathcal{M}, s \models p & \Leftrightarrow & p \in V(s) \\ \mathcal{M}, s \models \neg\phi & \Leftrightarrow & \mathcal{M}, s \not\models \phi \\ \mathcal{M}, s \models \phi \wedge \psi & \Leftrightarrow & \mathcal{M}, s \models \phi \text{ and } \mathcal{M}, s \models \psi \\ \mathcal{M}, s \models K_i\psi & \Leftrightarrow & \forall t : s \rightarrow_i t \implies \mathcal{M}, t \models \psi \end{array}$$

As usual, we define $\bot$, $\phi \vee \psi$, $\phi \rightarrow \psi$ and $\hat{K}_i\phi$ as the abbreviations of $\neg\top$,$\neg(\neg\phi \wedge \neg\psi)$, $\neg\phi \vee \psi$ and $\neg K_i \neg\phi$.

Basic system **S5=KT45=KT5=KT4B**

Axiom Schemas	
TAUT	all the instances of tautologies
DISTK(K)	$K_i(p \to q) \to (K_i p \to K_i q)$
T	$K_i p \to p$
4	$K_i p \to K_i K_i p$
5	$\neg K_i p \to K_i \neg K_i p$
Rules	
NECK	$\dfrac{\phi}{K_i \phi}$
SUB	$\dfrac{\phi(p)}{\phi(\psi)}$
MP	$\dfrac{\phi, \phi \to \psi}{\psi}$

A frame makes T valid iff it is reflexive.
A frame makes 4 valid iff it is transitive.
A frame makes 5 valid iff it is Euclidean.
S5 is sound and strongly complete w.r.t. epistemic frames.

Basic system **KD45** for doxastic logic

Axiom Schemas	
TAUT	all the instances of tautologies
DISTK(K)	$B_i(p \to q) \to (B_i p \to B_i q)$
D	$\neg B_i \bot$
4	$B_i p \to B_i B_i p$
5	$\neg B_i p \to B_i \neg B_i p$
Rules	
NECK	$\dfrac{\phi}{B_i \phi}$
SUB	$\dfrac{\phi(p)}{\phi(\psi)}$
MP	$\dfrac{\phi, \phi \to \psi}{\psi}$

A frame makes D valid iff it is serial.

KD45 is sound and strongly complete w.r.t. transitive, Euclidean and serial frames.

You can show that there are "unbelievable" true facts.

Further developments and challenges

Group notions:

$$\mathcal{M}, s \models C_G\psi \quad \Leftrightarrow \quad \forall t : s \to_G^* t \implies \mathcal{M}, t \models \psi$$
$$\mathcal{M}, s \models D_G\psi \quad \Leftrightarrow \quad \forall t : s \to_i t \text{ for each } i \in G \implies \mathcal{M}, t \models \psi$$

Some Challenges:

- Logical omniscience
- What are the “relevant/compatible alternatives”?

Epistemology ⇒ Logic (syntactic) ⇒ Logic (semantics) ⇒ Computer science/Economics ⇒ Logic (dynamic) ⇒ Epistemology.

In CS: Distributed System, Security verification, Multi-agent systems. In Economics: Game theory, Decision Theory

Gerbrandy, J. and Groeneveld, W. (1997).
Reasoning about information change.
Journal of Logic, Language and Information, 6(2):147–169.

Groenendijk, J. and Stokhof, M. (1991).
Dynamic predicate logic.
Linguistics and Philosophy, 14(1):39 – 100.

Hintikka, J. (1962).
Knowledge and Belief: An Introduction to the Logic of the Two Notions.
Cornell University Press.

Plaza, J. A. (1989).
Logics of public communications.
In Emrich, M. L., Pfeifer, M. S., Hadzikadic, M., and Ras, Z. W., editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216.

Veltman, F. (1996).
Defaults in update semantics.
Journal of Philosophical Logic, 25(3):221–261.

Von Wright, G., H. (1951).
An Essay on Modal Logic.
North-Holland Publishing Company.