Epistemic Logic

III Some (philosophical) problems of the basic framework

Yanjing Wang
Department of Philosophy, Peking University

Sept. 27th 2023

Recap: basic systems

Introspections

Logical omniscience

External vs. Internal

Group notions

Recap: basic systems

Epistemic Language (EL) and semantics

The Epistemic Language:

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

It is interpreted on Kripke models $\mathcal{M} = (S, \{\rightarrow_i\}_{i \in I}, V)$ where $\rightarrow_i$ has certain properties (such as being an equivalence relation).

$$\boxed{\mathcal{M}, s \vDash K_i\phi \quad \Leftrightarrow \quad \forall t : s \rightarrow_i t \implies \mathcal{M}, t \vDash \phi}$$

$S5$ system (strongest epistemic logic)

System $S5$

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

Note that $S5 = KT5 = KTB4 = S4B$ modulo theorems.

$S5$ is sound and strongly complete for modal logic over frames with equivalence relations.

Hintikka thinks $S4$ is acceptable as an epistemic logic.

$KD45$ system (strongest doxastic logic)

System $KD45$

Axioms		Rules	
TAUT	all the instances of tautologies	MP	$\dfrac{\phi, \phi \to \psi}{\psi}$
DISTK	$B_i(p \to q) \to (B_i p \to B_i q)$	NECK	$\dfrac{\phi}{B_i \phi}$
D	$\neg B_i \bot$ (or $B_i p \to \neg B_i \neg p$)	SUB	$\dfrac{\phi}{\phi[p/\psi]}$
4	$B_i p \to B_i B_i p$		
5	$\neg B_i p \to B_i \neg B_i p$		

$KD45$ is sound and strongly complete for modal logic over frames that are transitive, euclidean and serial.

Moore sentence and Church-Fitch (knowability) paradox

$$\vdash_{KT} K_i(p \wedge \neg K_i p) \rightarrow \bot$$

$$\vdash_{KD4} B_i(p \wedge \neg B_i p) \rightarrow \bot$$

Verificationist's theory of truth requires $\phi \rightarrow \Diamond K\phi$ for any ϕ

Under the condition that $\neg\Diamond\bot$:

requiring $\phi \rightarrow \Diamond K\phi$ for all ϕ will result in $\phi \rightarrow K\phi$. Why?

What if we revise the formalization of verificationist's theory?

Moore sentences will play yet an important role in the dynamic setting.

Compare with another type of sentences: *p* and might not *p*. It is different from the Moore sentence: "suppose *p* and might not *p*" is still strange.

More on Moore sentences

About the lack of evidence linking the government of Iraq with the supply of weapons of mass destruction to terrorist groups, Rumsfeld stated (February 12, 2002):

> *Reports that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns;that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know. And if one looks throughout the history of our country and other free countries, it is the latter category that tends to be the difficult ones.*

$I\phi := \neg K\phi \wedge \neg K\neg\phi$. Second-order ignorance ($II\phi$) and Rumsfeld ignorance ($I\phi \wedge \neg KI\phi$) cannot be known! (Fine 18).

(S5) knowledge = true (KD45) belief?

What do KD45 models look like?

KD45 agents thought they are S5 agents!

KD45 belief plus the following definition:

$$K_i\phi \equiv_{df} \phi \wedge B_i\phi$$

Does it give us the S5 axioms of K_i? Negative introspection does not work!

True KD45 belief is S4.4! (Stalnaker, Lenzen, Aucher, Yang). You can never strengthen the belief to obtain S5 (Yang)!

.4 Axiom is $p \to (\neg\Box p \to \Box\neg\Box p)$. Frame property:
for all $w, v, u \in W$, if wRv, wRu and $w \neq u$, then vRu.

A basic system combing knowledge and belief (Stalnaker 06)

Consider the system of KD belief and S4 knowledge plus some interaction axioms:

- $B(p \to q) \to (Bp \to Bq)$
- $\neg B\bot$
- $K(p \to q) \to (Kp \to Kq)$
- $Kp \to KKp$
- $Kp \to p$
- $Kp \to Bp$
- $Bp \to KBp$
- $\neg Bp \to K\neg Bp$
- $Bp \to BKp$ (strong belief)

B is actually $KD45$ and we can show that $Bp \leftrightarrow \neg K\neg Kp$. Then we have Geach theorem (.2) $\neg K\neg Kp \to K\neg K\neg p$ from D for B.

Introspections

Hintikka's arguments

On positive introspection:

- If $\{K\phi, \neg K\neg\psi\}$ is consistent
- Then $\{K\phi, \psi\}$ is consistent
- Substitute ψ above with $\neg K\phi$ we will have the KK principle (Axiom 4).

On negative introspection:

- Assuming 5, we have axiom $B : p \to K\hat{K}p$.
- Hintikka thinks it is not always possible to know the possibility of a fact.

Negative introspection

$$\neg K_i p \to K_i \neg K_i p$$

Lenzen (1978)'s example:

- Suppose $B_i K_i p$ but $\neg p$: i falsely believes that he knows p.
- then $\neg K_i p$ ($\mathtt{T}$ axiom)
- therefore $K_i \neg K_i p$ ($\mathtt{5}$ axiom)
- thus $B_i \neg K_i p$ (knowledge implies belief)
- conclude $B_i \bot$ (under normal axioms for B_i)

We cannot express false belief of knowledge!

The equivalent form: $\neg K_i \neg K_i p \to K_i p$ sounds more problematic: if you "think" it is possible to know then you know it.

Lenzen prefers S4.2: $\neg K_i \neg K_i p \to K_i \neg K_i \neg p$

Positive introspection

Williamson (1992, 2000) on inexact knowledge (assuming axiom 4 and T) where p_k means a tree is k cm tall (simplified version):

1. assuming (for all reasonable k): $K(p_{k+1} \to \neg K \neg p_k)$
2. assuming $K \neg p_0$
3. by (2) and axiom 4: $KK \neg p_0$
4. by (1) and `NEC` and `DIST` we have $KK \neg p_0 \to K \neg p_1$
5. by (3) and (4): $K \neg p_1$
6. repeat to derive $K \neg p_k$
7. Suppose p_{666} is true then it contradicts to $K \neg p_{666}$ (under `T` axiom)

(1) is fishy: From (1) we can derive $p_{k+1} \to \neg K \neg p_k$. Moreover, by T axiom $Kp_{k+1} \to p_{k+1}$. Hence, we can infer $Kp_{k+1} \to \neg K \neg p_k$. What exactly is assumed by (1)?

Transitivity

Transitivity is not that reasonable especially in the setting of inexact knowledge, e.g., a spectrum of redness.

Can we have positive introspection without transitivity? Yes...

Logical omniscience

Logical omniscience

Problematic closure rules:

- From $\vdash \phi$ infer $K\phi$
- From $\vdash \phi \to \psi$ infer $K\phi \to K\psi$
- From $\vdash \phi \leftrightarrow \psi$ infer $K\phi \leftrightarrow K\psi$

'Deductive closure principle': $K(\phi \to \psi) \land K\phi \to K\psi$ is a different thing (which can also be challenged though, e.g., examples by Dretske (1970)).

The problem is more about the complexity of the logical reasoning!

Logical omniscience

Ideas of the ‘solutions’:

- Reinterpretation: implicit knowledge; of an ideal agent
- Syntactic approach: a set of formulas
- Impossible worlds: inconsistent alternatives
- Awareness: K=awareness+implicit knowledge
- Algorithmic knowledge: K=answer by algorithm
- Neighbourhood semantics: still problematic (From $\vdash \phi \leftrightarrow \psi$ infer $K\phi \leftrightarrow K\psi$)
- Timed knowledge: reasoning takes time
- Hardness to verify the knowledge
- but, can we do better? I think so...

External vs. Internal

Modeller or modelled

- Help to model other's reasoning
- Help us to reason by ourselves

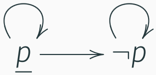
However, can the agents modelled 'see' the model (without the real world)?

It seems OK for S5 models, what about KD45?

For the KD45 agent the model is actually:

Is it good enough to have generated submodel based on the states which are reachable from the real world?

Now consider the following S4 model (reflexive and transitive):



If the agent can see the model (without the real world) himself, then he can reason as follows (suggested by Yang Liu):

I am not considering $\neg p$ as the only possibility, thus I must be on a p world, then I should know p! What is wrong?

At the p world: $K(\neg p \to K\neg p)$ is true thus $K\neg K\neg p \to Kp$. Note that $\neg K\neg p$ is also true at p but we don't have $K\neg K\neg p$ since it is S4 and you are not allowed to use negative introspection!

By having the model, you are implicitly using negative introspection: if I don't know $\neg p$ then I know I don't know it. The meta reasoning violates the assumption (the model is your brain).

Group notions

Common knowledge

David Lewis' *Convention* (1969): $p :=$ every driver must drive on the right.
What kind of knowledge is enough to let people feel safe in driving on the right? Is 'everybody knows that p' enough (Ep)? What about everybody knows that everybody knows that p (EEp)?

No, since agent i considers possible that agent j considers possible that agent i does not know ($\neg K_i K_j K_i p$) and thus agent j may drive on the left. You may argue by induction that $E^k p$ is not enough.

$$C\phi := E\phi \land EE\phi \land EEE\phi....$$

Alternating bit protocol

There are two processors 'Sender S' and a 'Receiver R'. The goal is for S to read a tape $X = x_0, x_1, \ldots$, and to send all the inputs it has read to R over a communication channel. R in turn writes down everything it receives on an output tape Y.Unfortunately the channel is not trustworthy, i.e., there is no guarantee that all messages arrive. On the other hand, some messages will not get lost, or more precisely: if you repeat sending a certain message long enough, it will eventually arrive. This property is called fairness. Now the question is whether one can write a protocol (or a program) that satisfies the following two constraints, provided that fairness holds:

- safety: at any moment, Y is a prefix of X;
- liveness: every x_i will eventually be written as y_i on Y.

$$
\begin{array}{|lcl|}
\hline
\mathcal{M}, s \models E_G\phi & \Leftrightarrow & \text{for all } t \text{ such that } s \overset{EG}{\to} t, \mathcal{M}, t \models \phi \\
\mathcal{M}, s \models C_G\phi & \Leftrightarrow & \text{for all } t \text{ such that } s \overset{CG}{\to} t, \mathcal{M}, t \models \phi \\
\hline
\end{array}
$$

where $\overset{EG}{\to} = \bigcup_{i \in G} \overset{i}{\to}$ and $\overset{CG}{\to} = (\overset{EG}{\to})^*$ (the reflexive transitive closure of $\overset{EG}{\to}$).

Axiomatization:

- S5 system for K plus `NEC` and `DIST` for C_G and
- Fixed-Point Axiom: $C_G\phi \leftrightarrow (\phi \wedge E_G C_G \phi)$
- Induction Axiom: $(\phi \wedge C_G(\phi \to E_G\phi)) \to C_G\phi$ or
- Induction *Rule*: $\vdash \psi \to E_G(\psi \wedge \phi)$ infer $\vdash \psi \to C_G\phi$

Common knowledge

Questions:

- Can we easily have full common knowledge? e.g., consecutive numbers.
- Do we really need the full power of common knowledge in many cases? tricky case: muddy children.
- What if I don't know who are there in the group? Common knowledge w.r.t. "group agent" without the explicit set of agents.

Distributed knowledge

Intuition: what we know if we put all of our knowledge together.

$$\boxed{\mathcal{M}, s \models D_G\phi \quad \Leftrightarrow \quad \text{for all } t \text{ such that } s \stackrel{DG}{\to} t, \mathcal{M}, t \models \phi}$$

where $\stackrel{DG}{\to} = \bigcap_{i \in G} \stackrel{i}{\to}$.

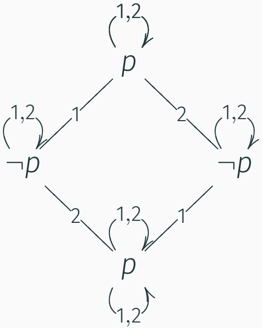
Finite axiomatization (a wise man's knowledge):

- S5 axioms for D_G plus
- $K_i\phi \to D_G\phi$ (when $i \in G$)

Distributed knowledge

Questions:

- It is not invariant under bisimulation;
- It is not the case: $\{\psi \mid \mathcal{M}, s \vDash K_i\psi \text{ for some } i \in G\} \vDash \phi$ iff $\mathcal{M}, s \vDash D_i\phi$ for all ϕ.



is bisimilar to

Other problems

- Modeling vs. model hecking
- Justification
- Agency and identity
- Understanding
- Other notions of knowing: correlation, procedure