Reasoning about others' knowledge

Yingying Cheng Yanjing Wang

4th November 2015

Abstract

The idea is to look at the proof systems of epistemic logic and discover how we reason about others' minds syntactically. This is a joint project by Yingying Cheng and Yanjing Wang.

1 Introduction

1.1 Motivation

- There are no explicit axioms and rules in Epistemic Logic to tell us how to reason about others.
- Theory of mind in cognitive science
- Reasoning about knowledge in law
- Epistemic assumptions and reasoning in Game theory

1.2 Reasoning about others' knowledge (and belief)

Why does it matter?

- As a basic ability: to interact with other people (strategically), e.g., game theory, theory of mind in psychology;
- Based on the assumption about this ability we can use various tools to regulate the information flows, e.g., design of protocols.
- Social use: e.g., to identify responsibility (in criminal law).

1.3 Epistemic logic seems to be a good tool

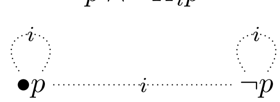
Propositional *modal logics* that reason about knowledge (and belief) [von Wright 1951, Hintikka 1962].

Language: "agent i *knows that* ϕ":

$$\phi ::= \top \mid p \mid \neg\phi \mid (\phi \wedge \phi) \mid K_i\phi (\text{sometimes also } C_G\phi)$$

Model: possibilities with (*equivalence*) relations.
Semantics: you know that ϕ iff ϕ is true in all epistemic alternatives (of the current world).

$$p \wedge \neg K_i p$$



$\mathbb{S}5$ system (strongest epistemic logic)

System $\mathbb{S}5$

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

The logic is powerful in combination with other modalities to handle changes of knowledge, e.g., epistemic temporal logic, dynamic epistemic logic, epistemic alternating-time temporal logic, and so on.

1.4 A simple but important question

In the aforementioned epistemic logics (and their friends), there is **no** single axiom or rule in the standard axiomatizations which can tell us explicitly how to reason about others' knowledge (based on your own knowledge).

> How do we reason about others' knowledge using existing epistemic logics syntactically?

This small project is to find out the answer and reveal the subtleties we took for granted.

1.5 In the rest of the talk

We will use muddy children as a running example and try to show:
(1) How to reason about others' knowledge semantically?
(2) How to reason about others' knowledge syntactically?
(2a) How to reason about others' knowledge using epistemic logic?
(2b) How to reason about others' knowledge using dynamic epistemic logic?
(3) Do we really need common knowledge in the examples where common knowledge is taken for granted? If not, how many levels on knowledge do we need?
(4) How does the syntactic approach help us to understand the initial model better?
(6) How do people reason about others' knowledge in (serious) reality?

2 Muddy Children puzzle

Informal description

> A group of n children meet their father after playing in the mud. Their father notices that $k > 0$ of the children have mud on their foreheads. Each child sees everybody else's foreheads, but not his own. The father says:"some of you are muddy," then says:"Do any of you know that you have mud on your forehead? If you do, raise your hand now." No one raises his hand. The father repeats the question, and again no one moves. After exactly k repetitions, all children with muddy foreheads raise their hands simultaneously.

First, consider a simple version of the above muddy children puzzle, suppose there are only 2 children in total saying Child 1 and Child 2 who are both muddy. Then the father will ask the question twice and both children raise their hands in the second round. After this, we will consider the case of extension to n children.

3 Semantic reasoning

Suppose you are one of the two children. After the first announcement, you see another child muddy, so you think perhaps he's the only muddy one. But you note that this child did not raise his hand, and you realize you are also muddy. So you raise your hand in the next round, and so does the other muddy child.

In a model-theoretical solution, we depict all possible situations. We determine a situation by stating for each child if it is muddy or not muddy.

In 2 children case, there are four situations: $(m_1, m_2) = \{(0,0), (0,1), (1,0), (1,1)\}$

$$\begin{tikzcd}
(1,1) \arrow[r, dash, "1"] \arrow[d, dash, "2"] & (0,1) \arrow[d, dash, "2"] \\
(1,0) \arrow[r, dash, "1"] & (0,0)
\end{tikzcd}$$

After the father publicly announces $m_1 \vee m_2$, node(0,0) is no longer possible:

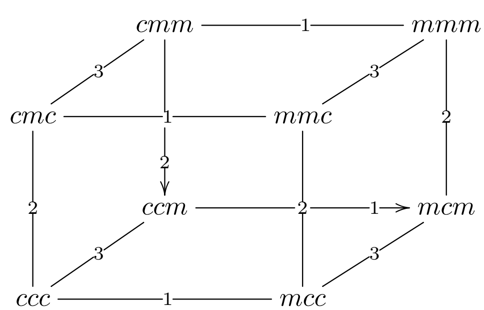
$$\begin{tikzcd}
(1,1) \arrow[r, dash, "1"] \arrow[d, dash, "2"] & (0,1) \\
(1,0)
\end{tikzcd}$$

Both children realize that in (1,0) child 1 would know whether he is muddy (no other distinguishable worlds), and in (0,1), Child 2 would know. Therefore, after both children answer "No" to the question of whether they know what is on their foreheads, worlds (1,0) and (0,1) are no longer possible, and each child eliminate them from the set of possible worlds. The only remaining logical possibility here is

$$(1,1)\bullet$$

Now both children know that their foreheads are muddy.

When $n = k = 3$: (c represents "clean"or "not muddy", m represents "muddy")



We can also get the only node (mmm) finally by eliminating nodes with the reasoning going on.

With the number coming bigger, the semantic method may become more complex, so we have to find another way-syntactic method.

4 Syntactic reasoning within Epistemic Logic

First, We describe a childs knowledge on the general background. These knowledge describe the environment, and are constant for all stages of the process. We fix a Child 1 and list some basic knowledge of Child 1(In the following formulas, m_1 is short for "Child 1 is muddy" and r_1 is short for "Child 1 raises his hand").

(Announcement by the teacher) Child 1 has knowledge of the announcement at two epistemic levels:

$(a1)K_1(m_1 \vee m_2)$;

$(a2)K_1K_2(m_1 \vee m_2)$.

(Observational ability) Child 1knows the following observational abilities:

$(b1)K_1(m_1 \leftrightarrow K_2m_1)$;

$(b2)K_1(m_2 \leftrightarrow K_1m_2)$;

$(b3)K_1(\neg m_1 \leftrightarrow K_2\neg m_1)$;

$(b4)K_1(\neg m_2 \leftrightarrow K_1\neg m_2)$;

$(b5)K_1K_2(m_1 \leftrightarrow K_2m_1)$;

$(b6) K_1K_2(m_2 \leftrightarrow K_1m_2)$;
$(b7) K_1K_2(\neg m_1 \leftrightarrow K_2\neg m_1)$;
$(b8) K_1K_2(\neg m_2 \leftrightarrow K_1\neg m_2)$;
$(b9) K_1(K_2m_1 \vee K_2\neg m_1)$.
(Action rule) Child 1 has knowledge of the action rules at two epistemic levels:
$(c1) K_1(K_2m_2 \to r_2)$;
$(c2) K_1(K_1m_1 \to r_1)$;
$(c3) K_1K_2(K_2m_2 \to r_2)$;
$(c4) K_1K_2(K_1m_1 \to r_1)$.
After the observation of another child's inaction in the first round, both children add new knowledge into their knowledge base:
$(d1) K_1(\neg r_2)$;
$(d2) K_1K_2(\neg r_1)$.

Now we shall prove K_1m_1 in the $\mathbb{S}5$ epistemic logic system syntactically.
(1)$K_1(K_2m_2 \to r_2)$ c1
(2)$(K_2m_2 \to r_2) \to (\neg r_2 \to \neg K_2m_2)$ TAUT
(3) $K_1((K_2m_2 \to r_2) \to (\neg r_2 \to \neg K_2m_2))$ (2) GENK
(4)$K_1((K_2m_2 \to r_2) \to (\neg r_2 \to \neg K_2m_2)) \to (K_1(K_2m_2 \to r_2) \to K_1(\neg r_2 \to \neg K_2m_2))$ (3)DISTK
(5) $K_1(K_2m_2 \to r_2) \to K_1(\neg r_2 \to \neg K_2m_2)$ (3)(4)MP
(6) $K_1(\neg r_2 \to \neg K_2m_2)$ (1)(5)MP
(7)$K_1(\neg r_2 \to \neg K_2m_2) \to (K_1\neg r_2 \to K_1\neg K_2m_2)$ (6)DISTK
(8)$K_1\neg r_2 \to K_1\neg K_2m_2$ (6)(7)MP
(9)$K_1\neg r_2$ d1
(10)$K_1\neg K_2m_2$ (8)(9)MP
(11)$K_1K_2(m_1 \vee m_2)$ a2
(12))$m_1 \vee m_2 \to (\neg m_1 \to m_2)$ TAUT
(13)$K_1K_2(m_1 \vee m_2 \to (\neg m_1 \to m_2))$ (12)NECK
(14)$K_1K_2(m_1 \vee m_2 \to (\neg m_1 \to m_2)) \to (K_1K_2(m_1 \vee m_2) \to K_1K_2(\neg m_1 \to m_2))$ DISTK
(15)$K_1K_2(m_1 \vee m_2) \to K_1K_2(\neg m_1 \to m_2)$ (13)(14)MP
(16)$K_1K_2(\neg m_1 \to m_2)$ (11)(15)MP
(17)$K_2(\neg m_1 \to m_2) \to (K_2\neg m_1 \to K_2m_2)$ DISTK
(18)$K_1(K_2(\neg m_1 \to m_2) \to (K_2\neg m_1 \to K_2m_2))$ (17)GENK
(19)$K_1(K_2(\neg m_1 \to m_2) \to (K_2\neg m_1 \to K_2m_2)) \to (K_1K_2(\neg m_1 \to m_2) \to K_1(K_2\neg m_1 \to K_2m_2))$ DISTK
(20)$K_1K_2(\neg m_1 \to m_2) \to K_1(K_2\neg m_1 \to K_2m_2)$ (18)(19)MP
(21)$K_1(K_2\neg m_1 \to K_2m_2)$ (16)(20)MP
(22)$(K_2\neg m_1 \to K_2m_2) \to (\neg K_2m_2 \to \neg K_2\neg m_1)$ TAUT
(23)$K_1(K_2\neg m_1 \to K_2m_2) \to (\neg K_2m_2 \to \neg K_2\neg m_1))$ (22)GENK
(24)$K_1(K_2\neg m_1 \to K_2m_2) \to (\neg K_2m_2 \to \neg K_2\neg m_1)) \to (K_1(K_2\neg m_1 \to K_2m_2) \to K_1(\neg K_2m_2 \to \neg K_2\neg m_1))$ DISTK
(25)$K_1(K_2\neg m_1 \to K_2m_2) \to K_1(\neg K_2m_2 \to \neg K_2\neg m_1)$ (23)(24)MP
(26)$K_1(\neg K_2m_2 \to \neg K_2\neg m_1)$ (21)(25)MP
(27)$K_1(\neg K_2m_2 \to \neg K_2\neg m_1) \to (K_1\neg K_2m_2 \to K_1\neg K_2\neg m_1)$ DISTK
(28)$K_1\neg K_2m_2 \to K_1\neg K_2\neg m_1$ (26)(27)MP
(29)$K_1\neg K_2\neg m_1$ (10)(28)MP
(30)$K_1(\neg m_1 \to K_2\neg m_1)$ b3
(31)$(\neg m_1 \to K_2\neg m_1) \to (\neg K_2\neg m_1 \to m_1)$ T
(32)$K_1((\neg m_1 \to K_2\neg m_1) \to (\neg K_2\neg m_1 \to m_1))$ (31)GENK
(33)$K_1((\neg m_1 \to K_2\neg m_1) \to (\neg K_2\neg m_1 \to m_1)) \to (K_1(\neg m_1 \to K_2\neg m_1) \to K_1(\neg K_2\neg m_1 \to m_1))$ DISTK
(34)$K_1(\neg m_1 \to K_2\neg m_1) \to K_1(\neg K_2\neg m_1 \to m_1)$ (32)(33)MP
(35)$K_1(\neg K_2\neg m_1 \to m_1)$ (30)(34)MP
(36)$K_1(\neg K_2\neg m_1 \to m_1) \to (K_1\neg K_2\neg m_1 \to K_1m_1)$ DISTK
(37)$K_1\neg K_2\neg m_1 \to K_1m_1$ (35)(36)MP
(38)K_1m_1 (29)(37) MP
(39)$K_1(m_1 \leftrightarrow K_2m_1)$ b1
(40)$K_1(m_1 \to K_2m_1) \to (K_1m_1 \to K_1K_2m_1)$ DISTK
(41)$K_1m_1 \to K_1K_2m_1$ (39)(40)MP
(42)$K_1K_2m_1$ (38)(41)MP
We can also prove K_1m_1 after we get $K_1K_2m_1$.
The front 29 steps remain, from the eleventh step, we prove in another way:
(29)$K_1\neg K_2\neg m_1$

(30')$K_1(K_2m_1 \vee K_2\neg m_1)$ b9

(31')$(K_2m_1 \vee K_2\neg m_1) \rightarrow (\neg K_2m_1 \rightarrow K_2\neg m_1)$ T

(32')$K_1((K_2m_1 \vee K_2\neg m_1) \rightarrow (\neg K_2m_1 \rightarrow K_2\neg m_1))$ (31')GENK

(33')$K_1((K_2m_1 \vee K_2\neg m_1) \rightarrow (\neg K_2m_1 \rightarrow K_2\neg m_1)) \rightarrow (K_1(K_2m_1 \vee K_2\neg m_1) \rightarrow K_1(\neg K_2m_1 \rightarrow K_2\neg m_1))$ DISTK

(34')$K_1(K_2m_1 \vee K_2\neg m_1) \rightarrow K_1(\neg K_2m_1 \rightarrow K_2\neg m_1)$ (32')(33')MP

(35')$K_1(\neg K_2\neg m_1 \rightarrow K_2m_1)$ (30')(34')MP

(36')$K_1(\neg K_2\neg m_1 \rightarrow K_2m_1) \rightarrow (K_1\neg K_2\neg m_1 \rightarrow K_1K_2m_1)$ DISTK

(37')$K_1\neg K_2\neg m_1 \rightarrow K_1K_2m_1$ (35')(36')MP

(38')$K_1K_2m_1$ (29)(37') MP

(39')$K_1(m_1 \leftrightarrow K_2m_1)$ b1

(40')$K_1(K_2m_1 \rightarrow m_1) \rightarrow (K_1K_2m_1 \rightarrow K_1m_1)$ DISTK

(41')$K_1K_2m_1 \rightarrow K_1m_1$ (39')(40')MP

(42')K_1m_1 (38')(41')MP

The (41')$K_1K_2m_1 \rightarrow K_1m_1$ in the sequence above is within the following principle:

In accepting that another knows a certain fact one is thereby effectively claiming that fact as part of one's own knowledge. And so, to know that another person knows some specific fact one must know this fact oneself. We thus have the following principle: $K_xK_y\phi \rightarrow K_x\phi$

4.1 Common Knowledge

In the resolution above, is it necessary to demand the background knowledge to be common knowledge? When we check the syntactic proof, we can see that we only use at most 2 levels of knowledge (i.e. $K_1K_2\ldots$). So maybe we shall say "No" to the question.

Common knowledge is a special kind of knowledge for a group of agents. There is common knowledge of p in a group of agents G when all the agents in G know p, they all know that they know p, they all know that they all know that they know p, and so on ad infinitum.

We can define an operator E_G with the intended meaning of "everyone in group G knows" by defining it with the axiom

$$E_G\phi \leftrightarrow \bigwedge_{i\in G} K_i\phi$$

By abbreviating the expression $E_GE_G^{n-1}\phi$ with $E_G^n\phi$ and defining $E_G^0\phi = \phi$, we could then define common knowledge with the axiom

$$C\phi \leftrightarrow \bigwedge_{i=0}^{\infty} E^i\phi$$

The axiom is not a well-formed formula. To overcome this difficulty, a fixed-point definition of common knowledge can be given: $C_G\phi = \phi \wedge E_G(C_G\phi)$.

We use the following example to say something about common knowledge:

> Imagine two generals who are planning a coordinated attack on a city. The generals are on two hills on opposite sides of the city, each with their armies, and they know they can only succeed in capturing the city if the two armies attack at the same time. But the valley that separates the two hills is in enemy's hands, and any messenger that is sent from one army base to the other runs a severe risk of getting captured. The generals have agreed on a joint attack, but they still have to settle the time.
>
> So the generals start sending messengers. General 1 sends a soldier with the message We will attack tomorrow at dawn. Call this message p. Suppose his messenger gets across to general 2 at the other side of the valley. Then K_2p holds, but general 1 does not know this because he is uncertain about the transfer of his message. Now general 2 sends a messenger back to assure 1 that he has received his message. Suppose this messenger also gets across without being captured, then K_1K_2p holds. But general 2 does not know this, for he is uncertain about the success of transfer: $\neg K_2K_1K_2p$. General 1 now sends a second messenger. If this one also safely delivers his message we have $K_2K_1K_2p$. But general 1 does not know this, and soon, and soon. In this way, theyll continue sending messages infinitely (and certainly not attack tomorrow at dawn).

Clearly, this procedure will never establish common knowledge between the two generals. They share the knowledge of p but that is surely not enough for them to be convinced that they will both attack at dawn. In case of real common knowledge every formula of the infinite set

$\{K_1p, K_2p, K_1K_2p, K_2K_1p, K_1K_2K_1p, \ldots\}$ holds.

Another visual example:

An illustration is provided by the puzzle of consecutive numbers, where two agents each is given a natural number, with the public rule that the numbers are consecutive. A situation in which agent a has a 2 and b a 3 can be depicted by the following Kripke model:

$$\begin{tikzcd} (0,1) \arrow[loop above, "a,b"] \arrow[r, leftrightarrow, "b"] & (2,1) \arrow[loop above, "a,b"] \arrow[r, leftrightarrow, "a"] & \underline{(2,3)} \arrow[loop above, "a,b"] \arrow[r, leftrightarrow, "b"] & (4,3) \arrow[loop above, "a,b"] \arrow[r, leftrightarrow, "a"] & (4,5) \arrow[loop above, "a,b"] & \dots \end{tikzcd}$$

In that situation, the standard epistemic semantics predicts that it is not common knowledge between a and b that their numbers are less than 100000, or even less than any positive number however large it may be, despite the fact that each of a and b knows that the numbers are less than 5. Informally, this corresponds to the fact that a considers it possible that b considers it possible that a has a 4, and so on and so forth.

MaCarthy's early work on epistemic logic with "any fool knows"

MaCarthy added a special person constant called 'any fool' and denoted it by 0. He says:" It is convenient to introduce any fool because whatever he knows, everyone knows that everyone else knows. 'Any fool' is especially useful when an event occurs in front of all the knowers, and we need sentences like, 'S_1knows that S_2 knows that S_3 knows etc.'" In MaCarthy's Formal Systems, $S * p$ stands for "Person S knows proposition p". e.g. $0 * (S * p \rightarrow p)$ means "Any fool knows that what a person knows is true".

4.2 Extension to n children

Suppose n children are all muddy (i.e.$k = n$), so after the father ask the same question n times, all children will raise their hands. Similar to the 2 children case, We describe a child's knowledge on the general background.

(Announcement by the teacher) :

$K_1K_2 \dots K_k(m_1 \vee \dots \vee m_n), \ \ \forall k \leqslant n$ (a)

(Observational Abilities):

$K_1K_2 \dots K_k(\neg(m_1 \vee \dots \vee m_{s'}) \rightarrow K_s\neg(m_1 \vee \dots \vee m_{s'})), \ \ \forall k \leqslant n$ and $n \geqslant s > s'$ (b)

(Action Rules):

$K_1K_2 \dots K_k(K_sm_s \rightarrow r_s), \ \ \forall k, s \leqslant n$ (c)

We call the set that exactly contains (a),(b) and (c) Γ_0. Then we define $\Gamma_k := \{K_1 \dots K_{k-1}(K_km_k \rightarrow r_k)(c1), K_1 \dots K_{k-1}(\neg(m_1 \vee \dots \vee m_{k-1}) \rightarrow K_k(\neg(m_1 \vee \dots \vee m_{k-1}))(b1), K_1 \dots K_k(m_1 \vee \dots \vee m_k)(a1), \forall n \geqslant k \geqslant 2\}$. It is easy to see that $\Gamma_k \subset \Gamma_0$

As nobody answer to the fathers question at first time, so we add new knowledge to the knowledge base of Child 1. $K_1 \dots K_k\neg r_s \ \ \forall k, s \leqslant n$. Call it A. It is easy to see that $K_1K_2 \dots K_{k-1}\neg r_k \subset$ A.

Proposition 1 $\Gamma_n, K_1 \dots K_{n-1}\neg r_n \vdash_{S5} K_1 \dots K_{n-1}(m_1 \vee \dots \vee m_{n-1}) \ \ \forall n \geqslant 2$

PROOF $n = 2$, we are to prove $\Gamma_2, K_1\neg r_2 \vdash K_1m_1$, this is done in the 2 children case.

Suppose the proposition holds when $n = k - 1$, we now prove when $n = k$, the proposition also holds. That is, $\Gamma_k, K_1 \dots K_{k-1}\neg r_k \vdash_{S5} K_1 \dots K_{k-1}(m_1 \vee \dots \vee m_{k-1})$.

$(1)K_1 \dots K_{k-1}\neg r_k$ Pre

$(2)K_1 \dots K_{k-1}(K_km_k \rightarrow r_k)$ (c1)

$(3)(K_km_k \rightarrow r_k) \rightarrow (\neg r_k \rightarrow \neg K_km_k)$ T

$(4)K_1 \dots K_{k-1}((K_km_k \rightarrow r_k) \rightarrow (\neg r_k \rightarrow \neg K_km_k))$ (3)GENK

$(5)K_1 \dots K_{k-1}((K_km_k \rightarrow r_k) \rightarrow (\neg r_k \rightarrow \neg K_km_k)) \rightarrow (K_1 \dots K_{k-1}(K_km_k \rightarrow r_k) \rightarrow K_1 \dots K_{k-1}(\neg r_k \rightarrow \neg K_km_k))$ DISTK

$(6)K_1 \dots K_{k-1}(K_km_k \rightarrow r_k) \rightarrow K_1 \dots K_{k-1}(\neg r_k \rightarrow \neg K_km_k)$ (4)(5)MP

$(7)K_1 \dots K_{k-1}(\neg r_k \rightarrow \neg K_km_k)$ (2)(6)MP

$(8)K_1 \dots K_{k-1}(\neg r_k \rightarrow \neg K_km_k) \rightarrow (K_1 \dots K_{k-1}\neg r_k \rightarrow K_1 \dots K_{k-1}\neg K_km_k)$ DISTK

$(9)K_1 \dots K_{k-1}\neg r_k \rightarrow K_1 \dots K_{k-1}\neg K_km_k$ (7)(8) MP

$(10)K_1 \dots K_{k-1}\neg K_km_k$ (1)(9)MP

$(11)K_1 \dots K_k(m_1 \vee \dots \vee m_k)$ (a1)

$(12)K_1 \dots K_k(\neg(m_1 \vee \dots \vee m_{k-1}) \rightarrow m_k)$

$(13)K_1 \dots K_{k-1}(K_k\neg(m_1 \vee \dots \vee m_{k-1}) \rightarrow K_km_k)$

$(14)K_1 \dots K_{k-1}(\neg K_km_k \rightarrow \neg K_k\neg(m_1 \vee \dots \vee m_{k-1}))$

$(15)K_1 \dots K_{k-1}(\neg K_km_k \rightarrow \neg K_k\neg(m_1 \vee \dots \vee m_{k-1})) \rightarrow (K_1 \dots K_{k-1}\neg K_km_k \rightarrow K_1 \dots K_{k-1}\neg K_k\neg(m_1 \vee \dots \vee m_{k-1}))$ DISTK

$(16)K_1 \dots K_{k-1}\neg K_km_k \rightarrow K_1 \dots K_{k-1}\neg K_k\neg(m_1 \vee \dots \vee m_{k-1})$ (14)(15)MP

$(17)K_1 \dots K_{k-1}\neg K_k\neg(m_1 \vee \dots \vee m_{k-1})$ (10)(16)MP

$(18)K_1 \dots K_{k-1}(\neg(m_1 \vee \dots \vee m_{k-1}) \rightarrow K_k(m_1 \vee \dots \vee m_{k-1}))$ (b1)

$(19)K_1 \dots K_{k-1}(\neg K_k(m_1 \vee \dots \vee m_{k-1}) \rightarrow (m_1 \vee \dots \vee m_{k-1}))$

(20)$K_1 \dots K_{k-1}(\neg K_k(m_1 \vee \cdots \vee m_{k-1}) \rightarrow (m_1 \vee \cdots \vee m_{k-1})) \rightarrow (K_1 \dots K_{k-1}\neg K_k(m_1 \vee \cdots \vee m_{k-1}) \rightarrow K_1 \dots K_{k-1}(m_1 \vee \cdots \vee m_{k-1}))$ (19)DISTK

(21)$K_1 \dots K_{k-1}\neg K_k(m_1 \vee \cdots \vee m_{k-1}) \rightarrow K_1 \dots K_{k-1}(m_1 \vee \cdots \vee m_{k-1})$ (19)(20)MP

(22)$K_1 \dots K_{k-1}(m_1 \vee \cdots \vee m_{k-1})$ (17)(21)MP $\mathcal{K}$

From the proposition, we can get that once a game round is over, the knowledge of Child 1 transfers from $K_1 \dots K_k(m_1 \vee \cdots \vee m_k)$ to $K_1 \dots K_{k-1}(m_1 \vee \cdots \vee m_{k-1})$, thus, after $n-1$ round, Child 1 knows that $K_1K_2(m_1 \vee m_2)$, therefore Child 1 will raise his hand after the nth question asked by the father.

5 Syntactic reasoning within Public Announcement Logic

If we put the Muddy Children puzzle within the Public Announcement Logic, we can find a new resolution.

The language of Public Announcement Logic is usually presented as follows:

$$\phi := \top \mid p \mid \neg\phi \mid \phi \wedge \phi \mid \Box_i\phi \mid [\phi]\phi$$

where $p \in P$. As usual, we define $\bot, \phi \vee \psi, \phi \rightarrow \psi$ and $< \psi > \phi$ as the abbreviations of $\neg\top, \neg(\neg\phi \wedge \neg\psi), \neg\phi \vee \psi$ and $\neg[\psi]\neg\phi$ respectively. The original reading of $\Box_i\phi$ is that "Agent i knows that ϕ" and $[\psi]\phi$ expresses "After announcing ψ publicly, ϕ holds." While $< \psi > \phi$ says that ψ can be truthfully announced publicly, and after its announcement ϕ holds.

The language of PAL is interpreted on Kripke models. A Kripke model over a nonempty set P of basic propositions is a triple $(S, \rightarrow, V)$ where S is a nonempty set of possible worlds, $\rightarrow \subseteq S \times S$ is a binary relation over S and $V : P \rightarrow 2^S$ is a valuation function assigning each basic proposition letter a set of worlds where it is true. Given a Kripke model $M = (S, \rightarrow, V)$ over P , the truth value of PAL formulas at a state s in M is defined as follows:

$M, s \models \top \Leftrightarrow$ always
$M, s \models p \Leftrightarrow s \in V(p)$
$M, s \models \neg\phi \Leftrightarrow M, s \not\models \phi$
$M, s \models \phi \wedge \psi \Leftrightarrow M, s \models \phi$ and $M, s \models \psi$
$M, s \models \Box_i\phi \Leftrightarrow \forall t \rhd_i s : M, t \models \psi$
$M, s \models [\psi]\phi \Leftrightarrow M, s \models \psi$ implies $M\|_\psi, s \models \phi$

where $(\forall t \rhd s : \dots)$ denotes for all $t : s \rightarrow t$ implies $\dots$, and $M|_\psi = (S', \rightarrow', V')$ such that: $S' = \{s | M, s \models \psi\}, \rightarrow' = \rightarrow |_{S' \times S'}$ and $V'(p) = V(p) \cap S'$. According to this semantics, an announcement action $[\psi]$ is interpreted as a model transformer which deletes the worlds in the model that do not satisfy ψ.

5.1 Axiom systems of Public Announcement Logic

Standard reduction-based proof system for PAL ($\mathbb{S}5$ omitted).

Axiom Schemas	
TAUT	all the instances of tautologies
DISTK	$\Box_i(\phi \rightarrow \psi) \rightarrow (\Box_i\phi \rightarrow \Box_i\psi)$
UATOM	$[\psi]p \leftrightarrow (\psi \rightarrow p)$
UNEG	$[\psi]\neg\phi \leftrightarrow (\psi \rightarrow \neg[\psi]\phi)$
UCON	$[\psi](\phi \wedge \chi) \leftrightarrow ([\psi]\phi \wedge [\psi]\chi)$
UK	$[\psi]\Box_i\phi \leftrightarrow (\psi \rightarrow \Box_i(\psi \rightarrow [\psi]\phi))$
UCOMP	$[\psi][\chi]\phi \leftrightarrow [\psi \wedge [\psi]\chi]\phi$
Rules	
NECK	$\dfrac{\phi}{\Box_i\phi}$
MP	$\dfrac{\phi, \phi \rightarrow \psi}{\psi}$

No uniform substitution!

Wang & Cao system of PAL

System PAN

Axiom Schemas		Rules	
TAUT	all the instances of tautologies	MP	$\frac{\phi, \phi \to \psi}{\psi}$
DISTK	$\Box(\phi \to \chi) \to (\Box\phi \to \Box\chi)$	NECK	$\frac{\phi}{\Box\phi}$
DISTU	$[\psi](\phi \to \chi) \to ([\psi]\phi \to [\psi]\chi)$	NECU	$\frac{\phi}{[\psi]\phi}$
INV	$(p \to [\psi]p) \wedge (\neg p \to [\psi]\neg p)$		
DET(optional)	$\langle\psi\rangle\phi \to [\psi]\phi$		
PRE	$\psi \leftrightarrow \langle\psi\rangle\top$		
NM	$\Diamond\langle\psi\rangle\phi \to [\psi]\Diamond\phi$		
PR	$\langle\psi\rangle\Diamond\phi \to \Diamond\langle\psi\rangle\phi$		

where $p \in \mathbf{P} \cup \{\top\}$

Then how about the relation between public announcement and knowledge? I know ϕ as long as I heard a public announcement ϕ, is it true?

Certainly not! There are many counterexamples about this: $[\phi \wedge \neg\Box\phi]\Box(\phi \wedge \neg\Box\phi)$ doesn't hold. $\phi \wedge \neg\Box\phi$ is itself consistent, but $\Box(\phi \wedge \neg\Box\phi)$ is equivalent to $(\Box\phi \wedge \Box\neg\Box\phi)$, $\Box\neg\Box\phi \to \neg\Box\phi$ in $\mathbb{S}5$ system, therefore $(\Box\phi \wedge \neg\Box\phi)$, contradicts!

But we can prove that when ϕ is a proposition letter, i.e. p, I know p as long as I heard the public announcement p.

5.2 Syntactic proof of $[p]\Box p$

Proposition 2 *$[p]\Box p$ is a theorem in PA and PAN.*

First in the reduction-based proof system:

(1)$p \to p$ TAUT
(2)$(p \to p) \to [p]p$ UATOM
(3)$[p]p$ (1)(2)MP
(4)$[p]p \to (p \to [p]p)$ TAUT
(5)$p \to [p]p$ (3)(4)MP
(6)$\Box(p \to [p]p)$ (5)NECK
(7)$\Box(p \to [p]p) \to (p \to \Box(p \to [p]p))$ TAUT
(8)$p \to \Box(p \to [p]p)$ (6)(7)MP
(9)$p \to \Box(p \to [p]p) \to [p]\Box p$ UK
(10)$[p]\Box p$ (8)(9)MP

Then in Wang &Cao system of PAL

(1)$(p \to [p]p) \wedge (\neg p \to [p]\neg p)$ INV
(2)$(p \to [p]p) \wedge (\neg p \to [p]\neg p) \to (p \to [p]p)$ TAUT
(3)$p \to [p]p$ (1)(2)MP
(4)$p \to p$ TAUT
(5)$p \to (p \wedge [p]p)$ (3)(4)TAUT
(6)$\Box(p \to (p \wedge [p]p))$ (5)NECK
(7)$(p \wedge [p]p) \leftrightarrow < p > p$ PFUNC
(8)$\Box(p \to < p > p)$
(9)$\Box(p \to < p > p) \to \Box(\neg(p \wedge [p]\neg p))$ TAUT
(10)$\Box\neg(p \wedge [p]\neg p)$ (8)(9)MP
(11)$(p \wedge [p]\neg p) \leftrightarrow < p > \neg p$ PFUNC
(12)$\Box\neg < p > \neg p$
(13)$\Box\neg < p > \neg p \to \neg\Diamond < p > \neg p$ TAUT
(14)$\neg\Diamond < p > \neg p$ (13)(14)MP
(15)$< p > \Diamond\neg p \to \Diamond < p > \neg p$ PR
(16)$\neg\Diamond < p > \neg p \to \neg < p > \Diamond\neg p$ (16)TAUT,MP
(17)$\neg < p > \Diamond\neg p$ (15)(17)MP

(18)$\neg < p > \Diamond\neg p \to [p]\Box p$ TAUT
(19)$[p]\Box p$ (18)(19)MP

Next we prove $(\phi \to [\psi]\phi) \wedge (\neg\phi \to [\psi]\neg\phi)$ when ϕ is a boolean formula by induction on the length of the ϕ.
PROOF
$\phi := p$.
$\phi := \neg\chi$, we need to prove $(\neg\chi \to [\psi]\neg\chi) \wedge (\neg\neg\chi \to [\psi]\neg\neg\chi)$, by inductive hypothesis, we have $(\chi \to [\psi]\chi) \wedge (\neg\chi \to [\psi]\neg\chi)$, while $\neg\neg\chi \leftrightarrow \chi$.
$\phi := \chi_1 \wedge \chi_2$, we need to prove $(\chi_1 \wedge \chi_2 \to [\psi](\chi_1 \wedge \chi_2)) \wedge (\neg(\chi_1 \wedge \chi_2) \to [\psi]\neg(\chi_1 \wedge \chi_2))$, that is $(\chi_1 \wedge \chi_2) \to ([\psi]\chi_1 \wedge [\psi]\chi_2)$ and $(\neg\chi_1 \vee \neg\chi_2) \to ([\psi]\neg\chi_1 \vee [\psi]\neg\chi_2)$. By inductive hypothesis, we have $\chi_1 \to [\psi]\chi_1$,$\chi_2 \to [\psi]\chi_2$, $\neg\chi_1 \to [\psi]\neg\chi_1$and $\neg\chi_2 \to [\psi]\neg\chi_2$. According to propositional tautology, we have $(\chi_1 \wedge \chi_2) \to ([\psi]\chi_1 \wedge [\psi]\chi_2)$, $\neg\chi_1 \to ([\psi]\neg\chi_1 \vee [\psi]\neg\chi_2)$ and $\neg\chi_2 \to ([\psi]\neg\chi_1 \vee [\psi]\neg\chi_2)$, then we have $(\neg\chi_1 \vee \neg\chi_2) \to ([\psi]\neg\chi_1 \vee [\psi]\neg\chi_2)$. Done. $\mathcal{K}$

After we have $\phi \to [\psi]\phi \wedge \neg\phi \to [\psi]\neg\phi$ holds when ϕ is a boolean formula, apply the same method as in the proof above, we will get the following proposition:

Proposition 3 *If ϕ is a boolean formula, $[\phi]\Box\phi$ is a theorem of the PAN system.*

5.3 Formal proof for 2-muddy children within $\mathbb{S}5$ PAN

We list Child 1s knowledge on the general background in the public announcement language:
(a1)$[m_1 \vee m_2]\Box_1(m_1 \vee m_2)$
(a2)$[m_1 \vee m_2]\Box_1\Box_2(m_1 \vee m_2)$
(a3)$\Box_1(m_1 \leftrightarrow \Box_2 m_1)$
(a4)$\Box_1(\neg m_1 \leftrightarrow \Box_2\neg m_1)$
(a5)$\Box_1(\Box_2 m_1 \vee \Box_2\neg m_1)$
(a6)$\Box_1 m_2$
No body raised his hand, that is to say each child publicly announced they don't know whether they are muddy, that is: $\neg\Box_1 m_1 \wedge \neg\Box_2 m_2$
We have to prove that $[m_1 \vee m_2][\neg\Box_1 m_1 \wedge \neg\Box_2 m_2](\Box_1 m_1 \wedge \Box_2 m_2)$ We only prove one part: $[m_1 \vee m_2][\neg\Box_1 m_1 \wedge \neg\Box_2 m_2]\Box_1 m_1$:

Hint: If we have proved that $[\phi](\Diamond_1\chi_1 \wedge \Box_1(\chi_1 \vee \chi_2))$, then after the announcement of $(\neg\chi_1)$ we will get $[\phi][\neg\chi_1]\chi_2$.

5.4 Return to the initial question

After all the syntactic proof, we can sum up the alternative axiomatizations of EL and PAL with axioms and rules which explicitly state how we reason about other's knowledge:
In EL, we add a necessitation rule for multi agents:

$$\frac{\phi}{\Box_1 \ldots \Box_n\phi}$$

In PAL:

$$[\phi]\Box_1 \ldots \Box_n\phi \text{ when } \phi \text{ is a boolean formula.}$$

We prove the adding axiom in PAL as following:
(1)$(\phi \to [\phi]\phi)$
(2)$\phi \to (\phi \wedge [\phi]\phi)$
(3)$\Box_1 \ldots \Box_n(\phi \to (\phi \wedge [\phi]\phi))$
(4)$\Box_1 \ldots \Box_n(\phi \to (< \phi > \phi))$
(5)$\Box_1 \ldots \Box_n\neg(\phi \wedge ([\phi]\neg\phi))$
(6)$\Box_1 \ldots \Box_n\neg < \phi > \neg\phi$
(7)$\Box_1 \ldots \Box_n[\phi]\phi$
(8)$\Box[\psi]\phi \to [\psi]\Box\phi$ **PR**
(9)$\Box_1 \ldots \Box_{n-1}(\Box_n[\phi]\phi \to [\phi]\Box_n\phi)$
(10)$\Box_1 \ldots \Box_{n-1}[\phi]\Box_n\phi$
(11)...
(12)$[\phi]\Box_1 \ldots \Box_n\phi$

As for the necessitation rule, there is a substantial amount of literature where the apparent problem of the failure of the necessitation rule is discussed. why I should know every valid formula? But in our proof above, we do apply the neessitation rule and it seems that no other method is available. So we have to say that necessitation rule plays an important role when reasoning about others' knowledge.

If we only care about the first-level knowledge in terms of $K_i\phi$, we always need higher order knowledge, in the muddy children puzzle, we need to prove $[m_1 \vee m_2][\neg\Box_1 m_1 \wedge \neg\Box_2 m_2](\Box_1 m_1 \wedge \Box_2 m_2)$, If we apply the reduction method, we can see what it is in EL:

$[m_1 \vee m_2][\neg\Box_1 m_1 \wedge \neg\Box_2 m_2]\Box_1 m_1$(Let $a = m_1 \vee m_2$ and $b = \neg\Box_1 m_1 \wedge \neg\Box_2 m_2$)

$\Leftrightarrow [a \wedge [a]b]\Box_1 m_1$

$\Leftrightarrow (a \wedge [a]b \to \Box_1(a \wedge [a]b \to [a \wedge [a]b]m_1)$

$\Leftrightarrow (a \wedge [a]b \to \Box_1(a \wedge [a]b \to (a \wedge [a]b \to m_1))$

$\Leftrightarrow (a \wedge [a]b \to \Box_1(a \wedge [a]b \to m_1)$

$\Leftrightarrow a \wedge [a]\neg\Box_1 m_1 \wedge [a]\neg\Box_2 m_2 \to \Box_1(a \wedge [a]\neg\Box_1 m_1 \wedge [a]\neg\Box_2 m_2 \to m_1)$

$\Leftrightarrow a \wedge (a \to \neg[a]\Box_1 m_1) \wedge (a \to \neg[a]\Box_2 m_2) \to \Box_1(a \wedge (a \to \neg[a]\Box_1 m_1) \wedge (a \to \neg[a]\Box_2 m_2) \to m_1)$

$\Leftrightarrow a \wedge (a \to \neg(a \to \Box_1(a \to [a]m_1))) \wedge (a \to \neg(a \to \Box_2(a \to [a]m_2))) \to \Box_1(a \wedge (a \to \neg(a \to \Box_1(a \to [a]m_1))) \wedge (a \to \neg(a \to \Box_2(a \to [a]m_2))) \to m_1)$

$\Leftrightarrow a \wedge (a \to \neg(a \to \Box_1(a \to (a \to m_1)))) \wedge (a \to \neg(a \to \Box_2(a \to (a \to m_2)))) \to \Box_1(a \wedge (a \to \neg(a \to \Box_1(a \to (a \to m_1)))) \wedge (a \to \neg(a \to \Box_2(a \to (a \to m_2)))) \to m_1)$

$\Leftrightarrow a \wedge (a \to \neg(a \to \Box_1(a \to m_1))) \wedge (a \to \neg(a \to \Box_2(a \to m_2))) \to \Box_1(a \wedge (a \to \neg(a \to \Box_1(a \to m_1))) \wedge (a \to \neg(a \to \Box_2(a \to m_2))) \to m_1)$

$\Leftrightarrow a \wedge (a \to a \wedge \Diamond_1(a \wedge \neg m_1)) \wedge (a \to a \wedge \Diamond_2(a \wedge \neg m_2)) \to \Box_1(a \wedge (a \to a \wedge \Diamond_1(a \wedge \neg m_1)) \wedge (a \to a \wedge \Diamond_2(a \wedge \neg m_2)) \to m_1)$

There is no announcement operator in the last formula, we can see that the second-order knowledge($\Box_1 \ldots \Diamond_1 \ldots$) is included.

6 Muddy Children scenario is a lucky exception

In Artemov's(Russian, 1951) report "Syntactic Epistemic Logic"in the cerebration event for Johan van Benthem, He proposed that Muddy Children scenario is a lucky exception.

Let us review the informal description at the beginning of our report,

"… Each child sees everybody else's foreheads, but not his own.…"

In fact, we presuppose two basic knowledge reasoning ability for each person, that is:

1. Knowing about others:

$\bigwedge\limits_{i \neq j} \{K_i(m_j) \vee K_i(\neg m_j)\}$

2.Not knowing about himself:

$\bigwedge\limits_{i=1,2} \{\neg K_i(m_i) \wedge \neg K_i(\neg m_i)\}$

What is missing in the semantic solution?

The first presupposition was presented in the list of general knowledge in our syntactic process of the previous section. But not the second one, so is it unnecessary to write it into our informal description?

Consider our semantic solution to the 2-Muddy Children puzzle, we adopt the following model as the whole possibility:

$$\begin{tikzcd}
(1,1) \arrow[r, dash, "1"] \arrow[d, dash, "2"] & (0,1) \arrow[d, dash, "2"] \\
(1,0) \arrow[r, dash, "1"] & (0,0)
\end{tikzcd}$$

M_1

Does the model M_1 characterize the scenario of 2-Muddy Children puzzle completely?

We should first check if the model M_1 satisfies the two syntactic presuppositions above. There seems no problem.

How about the other direction? If we omit the second presupposition, what will happen?

A nontrivial example

Let us consider the case when the second presupposition is omitted. That is:

Two children have muddy foreheads and each child sees the other child's forehead. The father announces publicly "some of you are muddy." The father then says: "Do any of you know that you have mud on your forehead? If you do, raise your hand now." No one raises his hand. The father repeats the question, and both children raise their hands simultaneously.

Which model characterizes the scenario best?

The possible epistemic models

The model M_1 is certainly a choice.

$$\begin{tikzcd}
(1,1) \arrow[r, dash, "1"] \arrow[d, dash, "2"] & (0,1) \arrow[d, dash, "2"] \\
(1,0) \arrow[r, dash, "1"] & (0,0)
\end{tikzcd}$$

Problem arises!

How about the following model M_2? Obviously, it is not equivalent to M_1

$$(1,0) \text{ —}2\text{— } (1,1)$$

$$(0,0) \text{ —}2\text{— } (0,1)$$

Model M_2

. . .

What can we learn from this abnormal example?

> If an epistemic scenario is given syntactically, but formalized by an epistemic model, it makes sense to examine its syntactic formalization as well and try to establish their equivalence.
>
> Sergei Artemov

7 'Knowingly' on criminal

There are a lot of "knowingly" in the rules of criminal law and judicial interpretation. Theorists have different opinions on "knowingly", as for "ascertain"or "should have known", so it is such a difficult discrimination that judicial departments try different methods to identify "knowingly".

The following is several examples in criminal law:

Article 138 If a person who is directly responsible **knowingly** fails to adopt measures against dangers in school buildings or in educational. . .

Article 144 Whoever mixes the foods that he produces or sells with toxic or harmful non-food raw materials or **knowingly** sells such foods . . .

Article 172 Whoever **knowingly** holds or uses counter-feit currencies shall, . . .

Article 399 Any judicial officer who, bending the law for selfish ends or twisting the law for a favor, subjects to investigation for criminal responsibility a person he **knows** to be innocent or intentionally protects from investigation for criminal responsibility a person he **knows** to be guilty . . .

. . .

Then how to identify responsibility in criminal law? How to reason that the doer did things knowingly?

There is a term in law called "legal reasoning", which is an evidence law to reason from known basic facts to unknown presumption facts and allow the party propose counterevidence in the process.

Let Φ be the set of basic facts and ψ be presumption fact, ψ usually has the form of $K\chi$. As we can reason from Φ to ψ, there must be some relation between Φ and ψ, we call it 'Permanent Connection', i.e.PC. Then $\bigwedge \Phi \to \psi \in PC$.

Reasoning about knowingly in law usually follow the following pattern:

(1)Solidify the permanent connection between knowingly subjectively and objective basic facts.

(2)Determine the existence of objective basic facts according to the permanent connection.

(3)Excluded the doer's counterevidence.

In logic, to say $\bigwedge \Phi \to \psi$ valid iff ψ holds whenever Φ holds, no tolerance of exception. Is it that rigid in criminal law? How can our logic help with it?

8 Reference

[1] J. Jude Kline. Evaluations of Epistemic Components for resolving the Muddy Children Puzzle. *Economic Theory*,53(1):61-83, 2013.

[2] John McCarthy, Masahiko Sato, Takeshi Hayshi, Shigeru Igarashi:1978, 'On the model theory of knowledge' Unpublished manuscript.

[3] Sergei Artemov. Syntactic Epistemic Logic. On: celebrations for Johan van Benthem, Unpublished presentation.

[4] Yanjing Wang and Qinxiang Cao. On axiomatizations of Public announcement logic. *Synthese*, pages 1-32, 2013.

[5] Yanjing Wang. A note on equivalence of two semantics for epistemic logic of shallow depths. Unpublished notes.