模态逻辑 II

逻辑导论

王彦晶
北大哲学系

2023 年 5 月 22 日

- 回顾

- 举例: 知识逻辑 (Epistemic Logic)
    - 知识的公理: 证明系统的作用
    - 知识的动态: 语义的作用
    - 群体知识: 多学科的视角

- 总结

回顾

基本模态逻辑的语言和语义 (盒子与钻石的故事)

基本的模态逻辑语言 ($p \in \mathrm{P}$):

$$\varphi \quad ::= \quad p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi$$

模型 $\mathcal{M} = \langle W, R, V\rangle$ (非空可能世界集, 可能世界之间的关系, 每个可能世界上原子命题的赋值), 语义定义在点模型 $(\mathcal{M}, w)$ 上

$$\boxed{\begin{array}{rcl} \mathcal{M}, w \vDash p & \Leftrightarrow & p \in V(w) \\ \mathcal{M}, w \vDash \neg\varphi & \Leftrightarrow & \mathcal{M}, w \nvDash \varphi \\ \mathcal{M}, w \vDash \varphi \wedge \psi & \Leftrightarrow & \mathcal{M}, w \vDash \varphi \text{ 且} \mathcal{M}, w \vDash \psi \\ \mathcal{M}, w \vDash \Box\varphi & \Leftrightarrow & \text{对所有的} v\text{, 如果} wRv \text{ 则} \mathcal{M}, v \vDash \varphi \\ \mathcal{M}, w \vDash \Diamond\varphi & \Leftrightarrow & \mathcal{M}, w \vDash \neg\Box\neg\varphi \Leftrightarrow \text{存在} v\text{, } wRv \text{ 且} \mathcal{M}, v \vDash \varphi \end{array}}$$

框架是没有赋值V 的 “模型” $\mathcal{F} = \langle W, R\rangle$. φ 在框架 $\mathcal{F}$ 上有效 $(\mathcal{F} \vDash \varphi)$ iff 对所有赋值V, 所有世界$u \in W$, $\langle \mathcal{F}, V\rangle, u \vDash \varphi$.

给定一个框架的类 $\mathbb{C}$, $\Gamma \vDash_{\mathbb{C}} \varphi$ iff 对所有基于$\mathbb{C}$里框架的点模型 $\mathcal{M}, w$: $\mathcal{M}, w \vDash \Gamma \implies \mathcal{M}, w \vDash \varphi$ (保真).

数学性质与公式通过框架有效性的对应

实际上，我们通常会考虑具有特定性质的模型和框架. 我们说 φ 对应于框架性质 X, 如果对任意框架 $\mathcal{F}$:

φ 在框架 $\mathcal{F}$ 上有效 $\iff$ $\mathcal{F}$ 有性质 X.

公式	框架性质
$\Box p \to p$	自反性: $\forall x Rxx$
$\neg\Box\bot$	持续性: $\forall x \exists y Rxy$
$p \to \Box\Diamond p$	对称性: $\forall x \forall y (Rxy \to Ryx)$
$\Box p \to \Box\Box p$	传递性: $\forall x \forall y \forall z (Rxy \land Ryz \to Rxz)$
$\Diamond p \to \Diamond\Diamond p$	稠密性: $\forall x \forall y \forall z (Rxy \to \exists z (Rxz \land Rzy))$
$\Diamond\Box p \to \Box\Diamond p$	合流性: $\forall x \forall y \forall z (Rxy \land Rxz \to \exists t (Ryt \land Rzt))$
$\Box(\Box p \to p) \to \Box p$	传递且没有无穷下降链

这些对应是连接哲学命题与数学性质的桥梁。

正规模态逻辑的极小系统 K: 基本公理 (模式) 和规则

- 命题逻辑的公理
- K 公理 $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$.
- 必然化规则 从 $\vdash \varphi$ 得到 $\vdash \Box\varphi$.
- 分离规则 从 $\varphi \to \psi$ 及 φ 得到 ψ.

注意，在做带前提的推演 ($\Gamma \vdash \varphi$) 的时候，必然化规则只能用在可证的公式上，所以从前提 p 得到 $\Box p$ 是不行的。直观上讲 $\Gamma \vdash \varphi$ 说的是: 假设 Γ 里的东西 (都是真的)，能推出 φ.

可靠性与强完全性:

$$\Gamma \vdash_{\mathsf{K}} \varphi \iff \Gamma \vDash_{\mathcal{C}_{all}} \varphi$$

$\mathcal{C}_{all}$ 是所有框架的框架类 (经常可以省略).

不同的模态逻辑证明系统

基于不同的公理有不同的证明系统:

- T : K+T 公理模式相对于自反的框架类可靠完全.
- B : K+B 公理模式相对于对称的框架类可靠完全.
- D : K+D 公理模式相对于持续的框架类可靠完全.
- S4 : K+T+4 相对于自反传递(或偏序) 的框架类可靠完全
- S5 : K+T+5=K+T+B+4 相对于等价关系的框架类可靠完全
- ...

可以像数学理论一样写哲学理论的公理，然后看他们能推出什么样的结果，可以比较不同理论的强度和相对的一致性。

一些常用的正规模态逻辑的关系

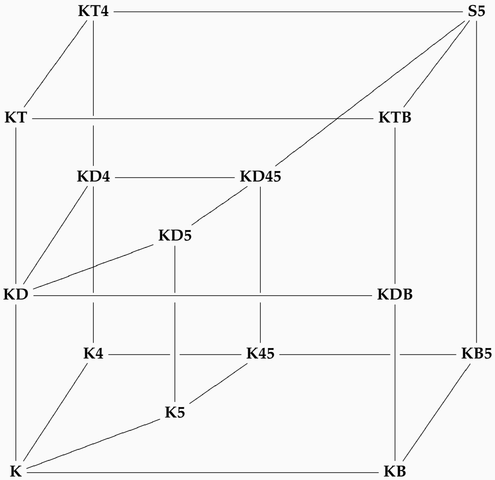


image source: Yang Liu

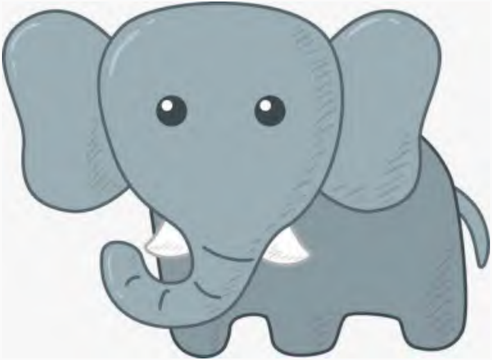
模态语言的特点

- 元语言概念拿到对象语言里精确的说、叠置组合的说
- 语言可以自然的表达各种哲学概念的性质和关系
- 语义直观可应用性广泛
- 公理系统可以像数学理论一样刻画哲学理论
- 有很多良好的元性质 (比如可判定性)
- 可以在哲学与数学及计算机之间架起桥梁

说的还是太抽象, 让我们看一个例子 (其他很多种类的模态逻辑也非常有意思.)

举例: 知识逻辑 (Epistemic Logic)

举例: 知识逻辑 (Epistemic Logic), 也叫认知逻辑



Epistemic Logic

让我们先不从一个谜题开始.

知识的推理有时是复杂的: 泥孩谜题 (Muddy Children)

泥孩谜题

三个小孩 a, b, c 在门口玩，他们的脸上有可能会不小心弄上了泥巴，不过他们只能看到别人脸上有没有泥巴，而不知道自己脸上有没有泥巴。他们的老爸 (一位逻辑学家) 出门看到孩子们玩的脏兮兮的很生气, 他说: “你们中间有人把泥巴弄到脸上了!”

- 接着他命令: “知道自己脸上有泥巴的给我站出来!”
- a 站出来了, b 和 c 没有动

请问都谁脸上有泥巴? (假设虎父无犬子)

泥孩谜题

三个小孩 a, b, c 在门口玩，a 和 b 的脸上不小心弄上了泥巴，不过他们只能看到别人脸上有没有泥巴，而不知道自己脸上有没有泥巴。他们的老爸 (一位逻辑学家) 出门看到孩子们玩的脏兮兮的很生气, 他说: “你们中间有人把泥巴弄到脸上了!”

- 接着他命令: “知道自己脸上有泥巴的给我站出来!”
- 没人动。
- 老爸继续命令道: “知道自己脸上有泥巴的给我站出来!”
- 这时候 a 和 b 都站出来了, c 没有动。

这是为什么? “高阶” 的知识推理很重要.

知识的推理有时会很“烧脑”：泥孩谜题

三个小孩 a, b, c 在门口玩，他们的脸上都弄上了泥巴，不过他们只能看到别人脸上有泥巴，而不知道自己脸上有没有泥巴。他们的老爸 (一位逻辑学家) 出门看到孩子们玩的脏兮兮的很生气, 他说: “你们中间有人把泥巴弄到脸上了!”

- 接着他命令: “知道自己脸上有泥巴的给我站出来!”
- 没人站出来。他重复道: “知道自己脸上有泥巴的站出来!”
- 还是没人站出来，他最后又大声说了一遍: “现在知道自己脸上有泥巴的给我站出来!”

突然，三个小孩都齐刷刷的站出来了。这是为什么?

更一般的, 如果一共n个小孩里面有k个脸上有泥巴呢?

人脑容易烧坏, 需要靠谱的推理工具.

濠梁之辩

庄子与惠子游于濠梁之上。庄子曰:“鲦鱼出游从容，是鱼之乐也。”惠子曰:“子非鱼，安知鱼之乐?”庄子曰:“子非我，安知我不知鱼之乐?”惠子曰:“我非子，固不知子矣，子固非鱼也，子之不知鱼之乐，全矣。”庄子曰:“请循其本。子曰汝安知鱼乐云者，既已知吾知之而问我，我知之濠上也。”

–《庄子 · 秋水》

想不清楚怎么办?

- “精炼我们的推理的唯一方式是使它们同数学一样切实，这样我们能一眼就找出我们的错误，并且在人们有争议的时候，我们可以简单的说：让我们来算算吧。而无须进一步的忙乱，就能看出谁是正确的”

–莱布尼兹

知识逻辑 (Epistemic Logic)：形式语言、模型、语义

推理知识 (与信念) 的模态逻辑 [von Wright 1951, Hintikka 1962].

- 形式语言: 给定一个主体集 I, 模态词加主体的角标 $\mathcal{K}_i$. “主体 i 知道 (*knows that*) φ” ($\mathcal{K}_i\varphi$), 可以说各种复杂的句子, 例如 $\mathcal{K}_i\varphi \wedge \mathcal{K}_i\neg\mathcal{K}_j\varphi \wedge \mathcal{K}_i\mathcal{K}_j(\mathcal{K}_i\varphi \vee \mathcal{K}_i\neg\varphi)$.
- 模型：$\langle W, \{R_i \mid i \in \mathrm{I}\}, V\rangle$ 可能世界集, 可能世界间的“不可区分关系”(分不清哪个是真实世界), 以及赋值函数. 不可区分关系可以认为是个等价关系 (自反传递对称).
- 语义: 在世界 w 上 i知道φ iff 在所有从 w 出发 i 不可区分的世界上 φ 都为真 iff φ 为假的可能性都被 i 排除了.

例如, 假设纽约确实在下雨 (p), 但 1 不知道, 不过 1 知道 2 知道纽约是否在下雨 (因为 1 知道 2 住在纽约)...

$$
\begin{tikzcd}
\underline{w : p} \arrow[loop above, "1,2"] \arrow[r, leftrightarrow, "1"] & v : \neg p \arrow[loop above, "1,2"]
\end{tikzcd}
$$

知识逻辑 (Epistemic Logic)：形式语言、模型、语义

推理知识 (与信念) 的模态逻辑 [von Wright 1951, Hintikka 1962].

- 形式语言: 给定一个主体集 I, 模态词加主体的角标 $\mathcal{K}_i$. “主体 i 知道 (*knows that*) φ” ($\mathcal{K}_i\varphi$), 可以说各种复杂的句子, 例如 $\mathcal{K}_i\varphi \wedge \mathcal{K}_i\neg\mathcal{K}_j\varphi \wedge \mathcal{K}_i\mathcal{K}_j(\mathcal{K}_i\varphi \vee \mathcal{K}_i\neg\varphi)$.
- 模型：$\langle W, \{R_i \mid i \in \mathbf{I}\}, V\rangle$ 可能世界集, 可能世界间的“不可区分关系”(分不清哪个是真实世界), 以及赋值函数. 不可区分关系可以认为是个等价关系 (自反传递对称).
- 语义: 在世界 w 上 i知道φ iff 在所有从 w 出发 i 不可区分的世界上 φ 都为真 iff φ 为假的可能性都被 i 排除了.

例如, 假设纽约确实在下雨 (p), 但 1 不知道, 不过 1 知道 2 知道纽约是否在下雨 (因为 1 知道 2 住在纽约)...

$$\underline{w : p} \overset{1}{\text{———————}} v : \neg p$$

$$w \vDash p \wedge \neg\mathcal{K}_1 p \wedge \mathcal{K}_2 p \wedge \mathcal{K}_1(\mathcal{K}_2 p \vee \mathcal{K}_2\neg p) \wedge \mathcal{K}_2\neg\mathcal{K}_1 p \wedge \mathcal{K}_1\mathcal{K}_2\neg\mathcal{K}_1 p.$$

泥孩谜题

三个小孩在门口玩，他们的脸上都弄上了泥巴，不过他们只能看到别人脸上有泥巴，而不知道自己脸上有没有泥巴。他们的老爸(一位逻辑学家) 出门看到孩子们玩的脏兮兮的很生气, 他说: “你们中间有人把泥巴弄到脸上了!”

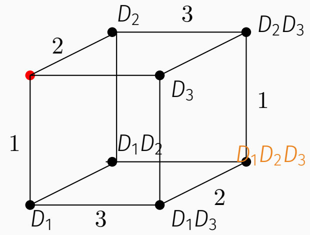
- 接着他命令: “知道自己脸上有泥巴的给我站出来!”
- 没人站出来。他重复道: “知道自己脸上有泥巴的站出来!”
- 还是没人站出来，他最后又大声说了一遍: “现在知道自己脸上有泥巴的给我站出来!”

突然，这三个小孩都齐刷刷的站出来了。这是为什么? (假设虎父无犬子)

要刻画三个小孩最开始的知识状态应该怎么做?

三个泥孩的初始知识模型

- 每个点是一个可能的情况
- D_i 表示 i 脸上有泥巴 ($i = 1, 2, 3$)
- 两个点之间有 i 的连线代表 i 搞不清 ta 到底在哪个情况. (省略了箭头和自反关系)



在右下角的世界上, $\neg\mathcal{K}_1 D_1 \wedge \mathcal{K}_2 D_1 \wedge \mathcal{K}_1(\mathcal{K}_2 D_1 \vee \mathcal{K}_2 \neg D_1)$ 为真.

在每个世界上都有 $\neg(\mathcal{K}_i D_i \vee \mathcal{K}_i \neg D_i) \wedge (\mathcal{K}_i D_j \vee \mathcal{K}_i \neg D_j)$ $(i \neq j)$.

之后我们会谈怎么处理知识的更新以处理泥孩谜题.

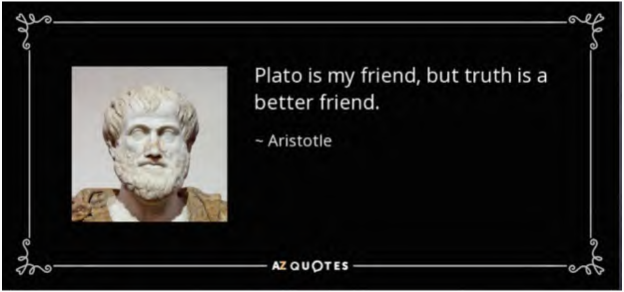
理想化的知识推理系统 S5 (对自反传递对称的框架类可靠完全)

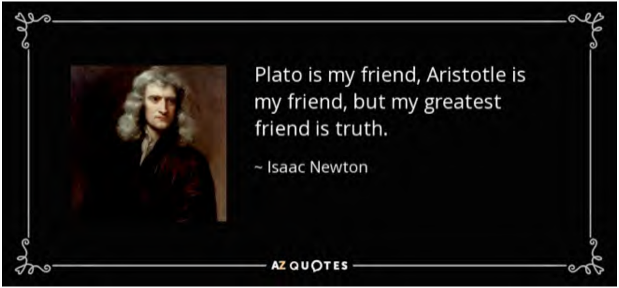
公理模式		推理规则	
TAUT	命题重言式	MP	$\dfrac{\varphi, \varphi \to \psi}{\psi}$
K	$\mathcal{K}_i(\varphi \to \psi) \to (\mathcal{K}_i\varphi \to \mathcal{K}_i\psi)$	NEC	$\dfrac{\varphi}{\mathcal{K}_i\varphi}$
T	$\mathcal{K}_i\varphi \to \varphi$		
4	$\mathcal{K}_i\varphi \to \mathcal{K}_i\mathcal{K}_i\varphi$		
5	$\neg\mathcal{K}_i\varphi \to \mathcal{K}_i\neg\mathcal{K}_i\varphi$		

- 4, 5: 正负自省公理 (Introspection axioms).
- T: 理想的知识应该是真的.

  “知之为知之，不知为不知，是知也。” – 《论语》

少了条T 公理? 知 = 箭头 + 口.





真理很难获得, 但求真更重要.

信念逻辑的常用系统 KD45 (对持续传递欧性的框架类完全)

语言中用 $\mathcal{B}_i\varphi$ 表示 i 相信 φ.

公理模式		推理规则	
TAUT	命题重言式	MP	$\dfrac{\varphi, \varphi \to \psi}{\psi}$
DISTK	$\mathcal{B}_i(\varphi \to \psi) \to (\mathcal{B}_i\varphi \to \mathcal{B}_i\psi)$	NEC	$\dfrac{\varphi}{\mathcal{B}_i\varphi}$
D	$\neg\mathcal{B}_i\bot$		
4	$\mathcal{B}_i\varphi \to \mathcal{B}_i\mathcal{B}_i\varphi$		
5	$\neg\mathcal{B}_i\varphi \to \mathcal{B}_i\neg\mathcal{B}_i\varphi$		

信念系统放弃了 T 公理, 但是要求信念至少是一致的: 你不能相信矛盾.

有了可靠性完全性，其实也不用记形式的推理规则，只要语义上保有效就行。

使用推理系统讨论一个哲学问题的简单例子

问题: 是不是所有真命题都是可知的? 假设 $p \wedge \neg\mathcal{K}_i p$ 是真的, i 能知道这个命题么? $\mathcal{K}_i(p \wedge \neg\mathcal{K}_i p)$ 在知识逻辑中是一个矛盾式.

1	$\mathcal{K}_i(p \wedge \neg\mathcal{K}_i p)$	假设
2	$\vdash (p \wedge \neg\mathcal{K}_i p) \to p$	TAUT
3	$\vdash (p \wedge \neg\mathcal{K}_i p) \to \neg\mathcal{K}_i p$	TAUT
4	$\vdash \mathcal{K}_i((p \wedge \neg\mathcal{K}_i p) \to p)$	NEC(2)
5	$\vdash \mathcal{K}_i((p \wedge \neg\mathcal{K}_i p) \to \neg\mathcal{K}_i p)$	NEC(3)
6	$\mathcal{K}_i p$	MP(公理K, 4, 1)
7	$\mathcal{K}_i\neg\mathcal{K}_i p$	MP(公理K, 5, 1)
8	$\neg\mathcal{K}_i p$	公理T

6和8矛盾! 事实上, 上面只用到了 KT 的推理能力. 可知性谜题: Fitch Paradox (什么叫可知? 被谁知道? 知道的时候还为真么?)
相关的 Moore paradox: 一个人说 “天在下雨但我不知道天在下雨” 会很奇怪. 当 i 断言 φ 的时候会有预设 $\mathcal{K}_i\varphi$.

弱化推理系统

没有 T 公理呢?

是不是所有真命题都可信?

$\mathcal{B}_i(p \wedge \neg \mathcal{B}_i p)$ 在信念逻辑 KD45 系统中也是一个矛盾式.

1	$\mathcal{B}_i(p \wedge \neg \mathcal{B}_i p)$	
2	$\vdash (p \wedge \neg \mathcal{B}_i p) \to p$	TAUT
3	$\vdash (p \wedge \neg \mathcal{B}_i p) \to \neg \mathcal{B}_i p$	TAUT
4	$\vdash \mathcal{B}_i((p \wedge \neg \mathcal{B}_i p) \to p)$	NEC(2)
5	$\vdash \mathcal{B}_i((p \wedge \neg \mathcal{B}_i p) \to \neg \mathcal{B}_i p)$	NEC(3)
6	$\mathcal{B}_i p$	MP(公理K, 4, 1)
7	$\mathcal{B}_i \neg \mathcal{B}_i p$	MP(公理K, 5, 1)
8	$\mathcal{B}_i \mathcal{B}_i p$	MP(公理4, 6)
9	$\mathcal{B}_i(\mathcal{B}_i p \wedge \neg \mathcal{B}_i p)$	正规模态逻辑由 (7,8) 可证

9 与 D 公理矛盾! 只用到了系统 KD4 的推理能力.

自省公理的讨论: 负自省公理 5

$$5 : \neg\mathcal{K}_i p \to \mathcal{K}_i\neg\mathcal{K}_i p$$

Lenzen (1978) 的例子:

- 假设 $\neg p \land \mathcal{B}_i\mathcal{K}_i p$: i 错误的相信他知道 p.
- 则 $\neg\mathcal{K}_i p$ (由 T 公理的逆否形式 $\neg p \to \neg\mathcal{K}_i p$)
- 所以 $\mathcal{K}_i\neg\mathcal{K}_i p$ (公理 5及 MP)
- 则有 $\mathcal{B}_i\neg\mathcal{K}_i p$ (假设知识至少是信念)
- 最后有 $\mathcal{B}_i\bot$, 又与 D 公理矛盾了!

似乎在这些前提下，我们没法表达关于知识的错误信念! 那么哪个假设有问题? Lenzen 认为是公理 5.

转换种形式: $\neg\mathcal{K}_i\neg\mathcal{K}_i p \to \mathcal{K}_i p$ 似乎更有问题: 如果你认为可能知道则你知道.

Lenzen 倾向于弱化公理 5: $\neg\mathcal{K}_i\neg\mathcal{K}_i p \to \mathcal{K}_i\neg\mathcal{K}_i\neg p$ (公理 4.2)

混合知识与信念时的奇怪现象 (Stalnaker 06)

假设 KD 的信念和没有负自省的 S4 知识, 以及非常自然的一些交互公理:

- $\mathcal{B}_i(p \to q) \to (\mathcal{B}_i p \to \mathcal{B}_i q)$
- $\neg \mathcal{B}_i \bot$
- $\mathcal{K}_i(p \to q) \to (\mathcal{K}_i p \to \mathcal{K}_i q)$
- $\mathcal{K}_i p \to \mathcal{K}_i \mathcal{K}_i p$
- $\mathcal{K}_i p \to p$
- $\mathcal{K}_i p \to \mathcal{B}_i p$
- $\mathcal{B}_i p \to \mathcal{K}_i \mathcal{B}_i p$
- $\neg \mathcal{B}_i p \to \mathcal{K}_i \neg \mathcal{B}_i p$
- $\mathcal{B}_i p \to \mathcal{B}_i \mathcal{K}_i p$ (强信念)

我们可以证明这样的信念可以用知识定义: $\mathcal{B}_i p \leftrightarrow \neg \mathcal{K}_i \neg \mathcal{K}_i p$! 进一步还可以推 .2 公理 $\neg \mathcal{K}_i \neg \mathcal{K}_i p \to \mathcal{K}_i \neg \mathcal{K}_i \neg p$.

自省公理的讨论: 正自省公理 4 也叫 KK 原则 ($\mathcal{K}\varphi \to \mathcal{K}\mathcal{K}\varphi$)

Williamson (1992) 关于非精确知识的讨论 (inexact knowledge):

令 p_k 表示一棵树是 k 厘米高:

1. 假设对任意 k: $\mathcal{K}(p_{k+1} \to \neg\mathcal{K}\neg p_k)$
2. 有 $\mathcal{K}(\mathcal{K}\neg p_k \to \neg p_{k+1})$
3. 假设 $\mathcal{K}\neg p_0$
4. 由 (3) 和 公理 4: $\mathcal{K}\mathcal{K}\neg p_0$
5. 由 (2) 和 NEC 及 K 公理我们有 $\mathcal{K}\mathcal{K}\neg p_0 \to \mathcal{K}\neg p_1$
6. 由 (4) 和 (5): $\mathcal{K}\neg p_1$
7. 重复上述对任意 k 推出 $\mathcal{K}\neg p_k$
8. 假设 p_{666} 为真, 则与 $\mathcal{K}\neg p_{666}$ 矛盾了 (T 公理)

Williamson 也攻击对应的框架性质: 传递性在非精确知识这里不对: A,B 两个颜色你无法区分, B,C 你也无法区分, C,D 你也无法区分, 但是可能你能区分 A 和 D.

使用逻辑进行论证的一种常见方式

要攻击一个断言 a, 可以先假设它 (无条件的) 成立, 然后:

- 考虑一些情形 (case) 下比较自然的前提 b, c
- 在逻辑语言中形式化地表达 a, b, c 为 A, B, C
- 借助 (某种广为接受的) 逻辑在以 BC 为前提, 把 A 当公理的情况下推出一个反直观或者矛盾的 D.
- 如果接受逻辑推理是可靠的, 似乎只能怀疑前提假设有问题
- a, b, c 中的 b, c 都很自然，只能怀疑 a 了...

"排除一切不可能的，剩下的即使看上去再不可能, 那也是真相。"

– 《福尔摩斯》

不过很多时候问题出在从自然语言的 a, b, c 到形式语言的 A, B, C 的翻译上，或者看上去 "自然" 的假设也可能有问题, 或者不应该用那个逻辑做推理... 很多与非经典逻辑有关系.

祖师爷 Hintikka (1962) 对正负自省的看法

赞成正自省 (KK 或者公理 5):

- 如果 $\{\mathcal{K}\varphi, \neg\mathcal{K}\neg\psi\}$ 是一致的
- 则直觉上 $\{\mathcal{K}\varphi, \psi\}$ 也是一致的
- 把上面的 ψ 换成 $\neg\mathcal{K}\varphi$ 就得到了正自省公理 (下面不一致则上面不一致. $\mathcal{K}\varphi, \neg\mathcal{K}\neg\psi$ 不一致也就意味着 $\mathcal{K}\varphi \to \neg\neg\mathcal{K}\neg\psi$ 即 $\mathcal{K}\varphi \to \mathcal{K}\mathcal{K}\varphi$).

反对负自省:

- 假设负自省公理 5 和 T (换种形式 $p \to \neg\mathcal{K}\neg p$), 我们能得到 B : $p \to \mathcal{K}\neg\mathcal{K}\neg p$.
- 这就会排除完全错误信念的可能 ($\neg\mathcal{K}\neg\mathcal{K}\neg p$ 信念有些时候可以理解为我想我知道 $\neg\mathcal{K}\neg\mathcal{K}$).

Jaakko Hintikka: Jan. 12, 1929 - Aug. 12, 2015

一阶逻辑的分配范式, Hintikka 集, 模态关系语义, 知识逻辑, 博弈语义, 独立友好 (IF) 逻辑... 逻辑学, 哲学, 数学, 语言学...

开创了知识逻辑的领域: 将知识和主体联系起来, 给出直观的语义, 讨论了一系列哲学问题, 使用一阶模态讨论了其他种类的从物 (*de re*) 的知识...

逻辑全知性 (Logical omniscience)

很多哲学家认为如下的封闭性规则可能太强了:

- 从 $\vdash \varphi$ 得到 $\mathcal{K}\varphi$ (必然化规则)
- 从 $\vdash \varphi \to \psi$ 得到 $\mathcal{K}\varphi \to \mathcal{K}\psi$
- 从 $\vdash \varphi \leftrightarrow \psi$ 得到 $\mathcal{K}\varphi \leftrightarrow \mathcal{K}\psi$

甚至 K 公理也被一些哲学家看成一种封闭原则:
$\mathcal{K}(\varphi \to \psi) \land \mathcal{K}\varphi \to \mathcal{K}\psi$ 也被攻击, 如 Dretske (1970).

比如: 我知道如果我是钵中之脑(p) 则我没有手 ($\neg q$), 这似乎可以形式化为 $\mathcal{K}(p \to \neg q)$, 在正规模态逻辑里可以等价为 $\mathcal{K}(q \to \neg p)$, 进而由 K 公理有 $\mathcal{K}q \to \mathcal{K}\neg p$. 然而我知道我有手 $\mathcal{K}q$, 但面对怀疑论者的质疑, 似乎作为一个谨慎的哲学家我并不能说我知道我不是钵中之脑 ($\neg\mathcal{K}\neg p$).

不过两个 “我知道” 是在不同意义下说的 (相对不同的考查范围)!

逻辑全知性

不仅哲学家, 数学家和计算机科学家也不喜欢知识推理中的逻辑全知性, 在每个具体的数学理论下，所有定理都等值于 $\top$, 但知道 $\top$ 为真，并不能让我知道那些定理为真，特别在涉及加密的时候.

解决的办法 (调整语义破坏相应公理或者规则的有效性):

- 重新解释知识算子: 隐含知识 (implicit knowledge) 等
- 知识就是公式集
- 引入有逻辑矛盾的不可能世界
- 引入觉知 (awareness): 知道 = 觉知 + 隐含知识
- 算法知识 (Algorithmic knowledge): 知道 = 算法给结果
- 更一般性的语义如邻域语义
- 加上时间的概念: 推理要花时间

技术处理要得到哲学辩护才可以, 不是只找个语义让特定的公理或者规则不再成立而已。虽然哲学上有很多争议，但现实当中最有用的还是S5的系统和对应的模型。

再回到泥孩谜题

三个小孩在门口玩，他们的脸上都弄上了泥巴，不过他们只能看到别人脸上有泥巴，而不知道自己脸上有没有泥巴。他们的老爸(一位逻辑学家) 出门看到孩子们玩的脏兮兮的很生气, 他说: “你们中间有人把泥巴弄到脸上了!”

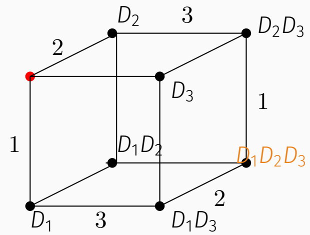
- 接着他命令: “知道自己脸上有泥巴的给我站出来!”
- 没人站出来。他重复道: “知道自己脸上有泥巴的站出来!”
- 还是没人站出来，他最后又大声说了一遍: “现在知道自己脸上有泥巴的给我站出来!”

突然，这三个小孩都齐刷刷的站出来了。这是为什么? (假设虎父无犬子)

更一般的, 如果一共 n 个小孩里面有 k 个脸上有泥巴呢?

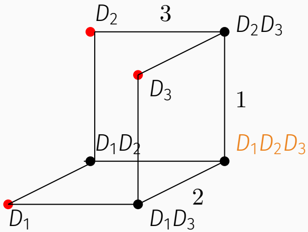
我们之前有了处理静态知识建模和推理的办法, 怎么严格处理知识的动态变化呢?

当有三个脏小孩的情况



“至少有一个人脸上有泥巴!” $\psi = D_1 \vee D_2 \vee D_3$

当有三个脏小孩的情况

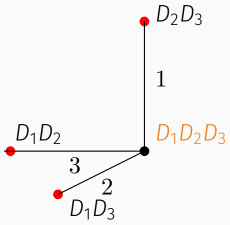


知道自己脸上有泥巴的给我站出来!

没人 (敢) 往前迈步. 其实这就在说: 我们都不知道!

$\chi = \neg \mathcal{K}_1 D_1 \wedge \neg \mathcal{K}_2 D_2 \wedge \neg \mathcal{K}_3 D_3$

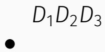
当有三个脏小孩的情况



现在知道自己脸上有泥巴的给我站出来!

还是没人站出来.

$\chi = \neg\mathcal{K}_1 D_1 \wedge \neg\mathcal{K}_2 D_2 \wedge \neg\mathcal{K}_3 D_3$



现在所有小孩都知道他们脑袋上有泥巴了:
$D_1D_2D_3 \vDash K_1D_1 \wedge \mathcal{K}_2D_2 \wedge K_3D_3$, 而且所有人都知道所有人都知道了.

看似无用的反复宣告无知反而得到了知识!

如何设计逻辑语言与语义描述新信息对知识的变化呢?

受语言哲学的启发:

- (Stalnaker) 1978 On Assertion：断言的内容（命题）依赖于语境，断言改变语境
- (Groenendijk & Stokhof, 1991), (Veltman 1996)：句子的意义在于它改变语境的潜力
- (Gerbrandy & Groeneveld, 1997)，(Plaza, 1989)：信息交流活动的意义在于它对人知识状态的改变
- (Baltag et al., 1998) 从公开的动作到半公开甚至私密的事件

公开宣告逻辑 (Public Announcement Logic) by Plaza (1989)

公开宣告逻辑的语言 ($p \in \mathbf{P}, i \in \mathbf{I}$):

$$\varphi \quad ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathcal{K}_i\varphi \mid [!\varphi]\varphi$$

语义也是通过 (S5) 克里普克模型给出 $\mathcal{M} = (W, \{R_i\}_{i \in I}, V)$:

$$\boxed{\mathcal{M}, w \vDash [!\psi]\varphi \quad \Leftrightarrow \quad \text{若}\mathcal{M}, w \vDash \psi \text{ 则}\mathcal{M}|_{\psi}, w \vDash \varphi}$$

其中 $\mathcal{M}|_{\psi} = (W', \{R'_i \mid i \in \mathbf{I}\}, V')$ 使得: $W' = \{w \mid \mathcal{M}, w \vDash \psi\}$, $R'_i = R_i|_{W' \times W'}$ 且 $V'(p) = V(p) \cap W'$.其实就是做了一个子模型.

动作的意义在于改变! 如下模型有: $\mathcal{M}, w_1 \vDash \neg\mathcal{K}_1 p \wedge [!p]\mathcal{K}_1 p$.

$$\begin{tikzcd}
w_1 : p \arrow[loop above, "1"] \arrow[r, leftrightarrow, "1"] & w_2 : \neg p \arrow[loop above, "1"] & !p \quad \Longrightarrow & w_1 : p \arrow[loop above, "1"]
\end{tikzcd}$$

刚才 3 个泥孩的谜题可形式化为 ($\mathcal{M}$ 是初始模型):

$\mathcal{M}, (D_1 D_2 D_3) \vDash [!\psi][!\chi][!\chi](\mathcal{K}_1 D_1 \wedge \mathcal{K}_2 D_2 \wedge \mathcal{K}_3 D_3)$.

Moore 句的幽灵: 不是所有被宣告的都会变成大家的知识

公开宣告逻辑也有可靠完全的推理系统.但是, 注意 $[!\varphi]\mathcal{K}_i\varphi$ 不是有效的! 事实上连 $[!\varphi]\varphi$ 都不是有效的! 有的真的公式被宣告之后会变成假的!

$$\begin{tikzcd} w_1 : \{p\} \arrow[loop above, "1"] \arrow[r, leftrightarrow, "1"] & w_2 : \{\} \arrow[loop above, "1"] \end{tikzcd} \qquad !(p \wedge \neg\mathcal{K}_i p) \implies \qquad \begin{tikzcd} w_1 : \{p\} \arrow[loop above, "1"] \end{tikzcd}$$

$w_1 \vDash (p \wedge \neg\mathcal{K}_1 p) \wedge [p \wedge \neg\mathcal{K}_1 p]\mathcal{K}_1 p$.

有哲学涵义的技术问题：哪些公式是在公开宣告下保持的?

甚至有时候假话说出来之后会使得它自己变成真的 (“True lies”).

类似的知识谜题都可以被形式化并通过计算机自动化的解决

- Cheryl’s Birthday
- Unfaithful wives/husbands
- Sum and Product
- Russian cards
- Puzzle of three hats
- 100 prisoner and a light bulb
- The hardest logic puzzle ever
- …

McCarthy 推广了很多 puzzle，告诉了很多聪明人 (包括北大的马希文老师)，推动了知识逻辑的发展。

Cheryl’s Birthday (新加坡奥赛题)

艾伯特和柏纳刚认识谢丽尔，想要知道谢丽尔的生日，谢丽尔列出了十个可能的日期：

5 月 15 日、5 月 16 日、5 月 19 日、6 月 17 日、6 月 18 日、7 月 14 日、7 月 16 日、8 月 14 日、8 月 15 日、8 月 17 日

接着谢丽尔分别告诉艾伯特及柏纳她生日的月及日，以下是艾伯特和柏纳的对话。

艾伯特：我不知道谢丽尔的生日是哪一天，但我知道你也不知道

柏纳：一开始我不知道谢丽尔的生日，但现在我知道了

艾伯特：那我也知道谢丽尔的生日了

请问谢丽尔的生日是那一天?

Russian Cards (俄罗斯奥赛题, van Ditmarsch 2003)

从 1, 2, 3, 4, 5, 6, 7 中 B 和 C 分别被保密的发了三张牌, 而 E 拿走剩下的牌. 问题: B 和 C 有没有可能通过公开的 (真实的) 宣告使得互相都知道对方的牌是什么, 但同时 E 还是不知道任何一张不在手里的牌的归属?

比如: 假设 B 拿的是 123, C 那的是 456, E 拿的是 7. B 说 "我或者 C 拿的是 123", 然后 C 说 "我或者 B 拿的是 456", 这样行么?

不行! 因为宣告者还得保证宣告是真的, 这也就意味着 B 知道: B 拿的是 123 或者 C 拿的是 123 ($[!\mathcal{K}_B(123_B \vee 123_C)]$) . 可是 B 怎么可能知道 C 拿的是什么牌? 所以 B 肯定手里拿的是 123 (有跳步).

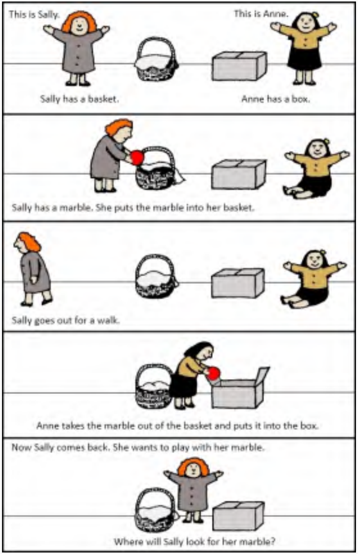
同时, 我们还需要保证拿到任何三张牌都可以有办法让对家知道. 甚至即使 E 知道我们的宣告方式, 他还是不能知道我们手里的牌! 这是可以做到的.

"纸牌密码学"(Card Cryptography)

对谜题感兴趣的

谜题的书: 一百囚徒与一个灯泡 [One Hundred Prisoners and a Light Bulb] [荷] 汉斯·范·狄马斯，[荷] 巴塔德·寇易著，[印度] 易兰车兹彦绘，马明辉译，科学出版社

Theory of mind (心理理论) 理解别人的错误信念



Baron-Cohen, Leslie and Frith (1985)

I. Dumontheil et al. (2010)
"把大球往上拿一格!" 到底拿哪个要看是谁说的

Table of Contents

从一个主体到多个主体

我们至今谈论的主要是一个主体的知识, 或者这个主体关于其他主体的知识, 有没有有意思的群体知识?

公共知识 (common knowledge)

- 哲学
- 计算机
- 经济学
- 语言学

公共知识 (Common knowledge)

David Lewis (D. 刘) 的 *Convention* (1969) 一书中提出.

令 p 为"X 国靠右行驶". 假设你到一个陌生的 X 国租了一辆车在一本地人很少的景区玩, 在一条比较窄的路上, 迎面快速开来了一辆车, 什么样的关于 p 的知识状态能让你觉得往右躲是安全的?

- 仅仅知道 p ($\mathcal{K}_i p$) 够么?
- 我和对面司机 (j) 都知道 p ($\mathcal{K}_i p \wedge \mathcal{K}_j p$) 够么?
- 两个人都知道两个人都知道 p 够么?

也不行, 我能想象也许你以为我以为是靠左走的 ($\neg\mathcal{K}_i\mathcal{K}_j\mathcal{K}_i p$)...

大刘的 "黑暗丛林法则"?

公共知识直观上讲就是: 所有人都知道所有人都知道所有人都知道..... 的知识.

McCarthy (1979) 也有类似的思想: any fool knows

两个将军的问题: 计算机通信中的公共知识

两个将军 A, B 分别在两个山头上, 他们相对两个山头中间谷底的敌人发起联合进攻, 但是他们没有无线通讯设备, 只能让侦察兵通过中间的谷地去送信, 可是侦查兵有可能会被敌人抓住.

不幸的是, 两个将军必须同时发起进攻才能赢得胜利, 否则就会全军覆没. 那么怎么能保证两个将军能同时发起进攻呢? 假设 A 将军给 B 将军成功送信 "明晨九点发起进攻", B 将军通过信使回复 "收到", 但是信使可能被敌人截获, A 将军即使收到 B 将军的回复, 还得给 B 将军再发一个回执, 可是 B 将军收到后也得再给 A 将军一个回执... 回来回去已经过了第二天 9 点....

- A: 如果我知道 B 一定会进攻的话, 我就也进攻
- B: 如果我知道 A 一定会进攻的话, 我就也进攻

可以证明,"A,B 都决定进攻" 需要是公共知识才行, 可这在信息可能丢失的情况下是不可能的 (J. Halpern and Y. Moses 1990)!

公共知识是可以精确定义和精确推理的!

对一个主体集 G, 对于 φ 的普遍知识 (general knowledge) 和公共知识 φ 分别定义为:

- $\mathcal{E}_G\varphi := \bigwedge_{i\in G} \mathcal{K}_i\varphi$
- $\mathcal{C}_G\varphi := \mathcal{E}_G\varphi \wedge \mathcal{E}_G\mathcal{E}_G\varphi \wedge \mathcal{E}_G\mathcal{E}_G\mathcal{E}_G\varphi \ldots.$ 看上去很难把握

$$\boxed{\begin{array}{lcl} \mathcal{M}, w \vDash \mathcal{E}_G\varphi & \Leftrightarrow & \text{对所有 } v \text{ 如果} wR_{EG}v, \text{ 则} \mathcal{M}, v \vDash \varphi \\ \mathcal{M}, w \vDash \mathcal{C}_G\varphi & \Leftrightarrow & \text{对所有 } v \text{ 如果} wR_{CG}v, \text{ 则} \mathcal{M}, v \vDash \varphi \end{array}}$$

这里 $R_{EG} = \bigcup_{i\in G} R_i$, $R_{CG} = (R_{EG})^*$ (R_{EG} 的自反传递闭包).

直观上, $\mathcal{C}_G\varphi$ 在 w 上为真当且仅当从 w 出发通过 G 里的各种 i 的关系组成的路径走到哪里 φ 都成立.

$$\begin{tikzcd} \underline{w : \{p\}} \arrow[loop above, "i,j"] \arrow[r, leftrightarrow, "i"] & v : \{p\} \arrow[loop above, "i,j"] \arrow[r, leftrightarrow, "j"] & u : \{\} \arrow[loop above, "i,j"] \end{tikzcd}$$

$w \vDash \mathcal{E}_{\{i,j\}}p \wedge \neg\mathcal{C}_{\{i,j\}}p$

假设 C 秘密的分别给了 A 和 B 两数字 2 和 3, 他只告诉他们俩这两数字是连着的自然数 (但没说谁拿到的大), A 和 B 只能看到自己手里的数字. 令 p 为命题 “这两个数字的和小于一千万”, 请问 p 是 A 和 B 的公共知识么?

模型的样子:

$$\begin{tikzcd}
\underline{(2,3)} \arrow[r, dash, "B"] \arrow[d, dash, "A"] & (4,3) \arrow[r, dash, "A"] & (4,5) \arrow[r, dash, "B"] & (6,5) \cdots\cdots\cdots \\
(2,1) \arrow[r, dash, "B"] & (0,1) & &
\end{tikzcd}$$

我们会有: $(2,3) \vDash \neg \mathcal{K}_B \mathcal{K}_A \mathcal{K}_B\, (x + y \leq 10)$

公共知识的公理: 可以刻画 C_G 这样似乎无穷的算子

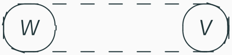
处理 $\mathcal{C}_G$ 需要在 S5 系统基础上增加 (把 $\mathcal{E}_G$ 看成简写):

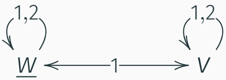
- NEC 和 K 的 $\mathcal{C}_G$ 版本
- 不动点公理: $\mathcal{C}_G\varphi \leftrightarrow (\varphi \wedge \mathcal{E}_G\mathcal{C}_G\varphi)$ ($\mathcal{C}_G\varphi$ 是一个 “不动点”.)
- 归纳公理: $(\varphi \wedge \mathcal{C}_G(\varphi \to \mathcal{E}_G\varphi)) \to \mathcal{C}_G\varphi$ (如果 φ 在当前点成立而且不管走到哪如果 φ 成立则下一步怎么走 φ 都成立, 则 φ 不管走到哪都成立.)

这样的逻辑也是可判定的, 可以做自动化的推理.

群体知识: 经济学中的模型

与 S5 的克里普克模型等价的模型被经济学家 Robert Aumann (1976) 独立的发现, 在经济学里叫做 Aumann's structure, 本质上是状态集上有划分 (Partition) 还带概率分布, 而划分可以导出等价关系. Aumann 用这种结构来定义和处理公共知识. 划分 I_i 代表不同主体得到的信息 (虚线 I_1, 实线 I_2), 事件e 是状态的子集 (相当于我们的 "公式"), $\mathcal{K}_i(e) = \{s \mid I_i(s) \subseteq e\}$ (i 知道 e 的状态的集合), $\mathcal{E}_G(e) = \bigcap_{i \in G} \mathcal{K}_i(e)$, $\mathcal{C}_G(e) = \bigcap_{n=1}^{\infty} \underbrace{\mathcal{E}_G \ldots \mathcal{E}_G}_{n}(e)$.





Aumann 用它证明了经济学里一个著名的不可能定理: 如果两个人初始的先验信息是公共知识, 则不管怎么根据进一步的私人证据进行充分的更新和交流, 大家都不可能最后形成有分歧的后验知识 (不可能 agree to disagree)!

理论计算机, AI 及博弈论中的应用

墙内开花墙外香: Reasoning about knowledge by Fagin, Halpern, Moses, Vardi (1995).

处理 (不) 确定性, 主体建模等.

- 分布式系统 (一个计算机只 “知道” 一部分整个系统的情况)
- 形式化验证: 特别是安全协议验证
- 基于知识的程序 (knowledge-based program)
- 多主体系统
- 基于不确定性的自动规划
- 知识表示
- 信念修正
- 非完美信息博弈的求解
- ...

知识逻辑领域的 TARK 会议系列

Theoretical Aspects of Rationality and Knowledge

- J. Halpern 发起, 从 1986 开始每两年一届,
- 把计算机科学家, 哲学家, 经济学家, 语言学家聚在一起讨论讨论和形式化的知识与信念相关的问题.
- 86 年第一次会议群星闪耀: Aumann, Hintikka, McCarthy, Stalnaker, Halpern, Vardi, Moses, Fagin, Dwok, Smullyan, Moore, Asher, Kamp, Levesque, Immerman, Plotkin, Ladner, Fischer....
- `http://www.tark.org/`

对哲学讨论的启发

- 传统知识论：知识定义 (e.g., JTB+?)、证成、应对怀疑论.
- 用逻辑工具刻画知识论的不同理论, 比较, 考虑后果. 与传统知识论的连接点: Justification Logic 等.

但更重要的是提出新的问题: 形式化知识论

- 用各种形式化手段研究不确定性、知识更新、信念修正、经验学习、归纳、博弈... 从个体到群体到社交网络.
- 各种手段：逻辑、概率论、博弈论、机器学习
- Hendricks, Kelly, Hájek, Hansson, van Benthem, Parikh, Halpern, Williamson ...
- Readings in Formal Epistemology, Springer 2016

有了知识逻辑的概念可以帮我们每天更能看清这个世界

- 什么是密码？你知我知, 但别人不知
- 邮件发给他密送你是干嘛的？你知他知, 但他不知你知, 而且这些是你我的公共知识.
- 微信群是干嘛的? (试图) 制造公共知识
- “代我问他好” 是干嘛的？让你知道我尊重他.
- 送什么礼物给太太? 我知道她也知道对她有用的
- 广告语的意义? 制造带意义的动作传递知识.
- 《三体》中的黑暗森林法则: 爱好和平的公共知识难以达成.
- 如何建设健康学术环境: 让他知道你知道学术规范
- 狼人杀? 帮助不大... 但有时可以理性利用人类的不理性
- 会经常挑刺, e.g., 微信每次更新：“修复了一些已知问题”
- 写 Intro 的一个目的: 让别人知道你知道
- 各种付费知识分享平台: 让你以为 (相信) 你知道.

知识与其他模态词的整合

- 知识与时间
- 知识与道义
- 知识与偏好
- 知识与策略
- ...

超越 “知道如是”(knowing that) 的知识逻辑

- 我知道是否这个定理是对的 (knowing whether).
- 我知道这瓶牛奶的价格是什么 (knowing what).
- 我知道如何做蛋糕 (knowing how).
- 我知道为什么他来晚了 (knowing why).
- 我知道是谁没交作业 (knowing who)
- 知道 + 一个 Wh-疑问从句

上面的 “知道” 能换成 “相信” 么? 如果知识至少是信念为什么不能换? 这涉及语言学, 认知科学和哲学. 这些句子表示的知识和一般通过 know-that 表达的知识一样么? 怎么做这样的表达是的推理? 怎么让机器自动化的做?

我们就在推动这方面的研究.

量词与模态词的交互

Hintikka (1962) 使用一阶模态逻辑讨论过 "知道是谁"，比如 "知道玛丽是谁" 可以被形式话为：$\exists x \mathcal{K}(Mary \approx x)$, 也可以理解为知道 "谁是玛丽?" 这个问题的答案.

Hintikka 还使用带量词的知识逻辑来理解问题(questions)，比如，考虑一个问题 Q : "是谁杀了 Bob?"

- 问题 Q 有个预设 提问人知道 Bob 被杀了: $\mathcal{K}\exists x M(x, Bob)$
- Q 的目的是知道谁杀了 Bob: $\exists x \mathcal{K} M(x, Bob)$.
- 一个可能的回答是 Adam 杀了 Bob: $M(Adam, Bob)$.
- 但 Hintikka 认为上面这个回答可能还不足以达到 Q 的目标, 因为提问人还得知道 Adam 是谁才行 $\exists x \mathcal{K}(Adam \approx x)$.

$\mathcal{K}\exists x \varphi$ 与 $\exists x \mathcal{K}\varphi$ 体现了哲学上讲得从言 (*de dicto*) 与从物 (*de re*) 知识的区别.

总结

其他重要的模态逻辑

每个主题都至少可以讲一次...

- 时态逻辑
- 动态逻辑
- 道义逻辑
- 可证性逻辑
- 混合逻辑
- 一阶模态逻辑
- ...

最新的前沿理论研究可以关注模态逻辑领域的双年会议:
Advances in Modal Logic (www.aiml.net).

时态逻辑 (Temporal Logic) 命题逻辑之上加上...

- 线性时间 (Linear time) 时态逻辑 (LTL): $X\varphi$ (Next), $\varphi U \psi$ (Until), $F\varphi := \top U \varphi$ (未来), $G\varphi := \neg F \neg \varphi$(永远)
- 分叉时间 (Branching time) 时态逻辑 (CTL): $EX\varphi$, $EG\varphi$, $\psi EU \varphi$

例如, 管理系统资源的程序需要处理进程 $i \in \{1, 2\}$ 的请求 (req_i) 和特定资源的分配 (own_i), 每个 (无穷) 执行路径要满足

- $G\neg(own_1 \wedge own_2)$ 同一个资源不能同时分给两个进程
- $G(req_i \to F\ own_i)$ 资源请求肯定会在未来某个时刻被满足
- $GF(req_1 \wedge \neg(own_1 \vee own_2)) \to GF\ own_1$ 如果进程 1 无穷多次的在资源闲置时申请, 则它会无穷多次的得到这个资源.

除了使用推理系统...

模型检测 (Model Checking)

把程序或者硬件抽象成模型 $\mathcal{M}, s$，用时态逻辑等逻辑语言的公式写出我们想要验证的性质 φ, 检查 $\mathcal{M}, s$ 是否满足 φ (不成立则指出模型里哪不对)

- 为什么这是个问题? 时态逻辑的语义定义没那么简单.
- 同一个逻辑的模型检测一般比定理证明或可满足性判断计算复杂度更低. “A picture is worth a thousand words”.
- 核心问题: 状态空间爆炸 (state space explosion). E.g., 文史楼里灯开关的所有状态组合可能超过整个宇宙的原子数量.
- 办法: 更有效的表示模型，抽象化，去掉等价的部分，拆分...

时态逻辑公式往往可以转换成自动机, 可满足性变为自动机识别语言判空问题。

07 年图灵奖: 基于时态逻辑的模型检测

颁给 Clarke, Emerson 以及 Sifakis 基于他们对模型检测的开创性贡献.

"Model Checking started as an academic research idea. The continuing research of Clarke, Emerson, and Sifakis as well as others in the international research community over the last 27 years led to the creation of new logics, as well as new algorithms and surprising theoretical results. This in turn has stimulated the creation of many Model Checking tools by both academic and industrial teams, resulting in the widespread industrial use of Model Checking. "

Handbook of Philosophical Logic 2nd.

模态逻辑及其他哲学逻辑在计算机里有大量的应用, 计算机的应用反过来又推动了逻辑学的进步. 其他做过重要模态逻辑贡献的图灵奖得主还包括 McCarthy, Pnueli, Lamport, Scott, Milner …

应用模态逻辑的方式

不光是做做推理.

把看似不好把握的抽象的概念变成可以在哲学上辩护的, 可以精确定义和计算的 working definition, 从 “天上” 拉到 “地上”.

- 从语义的角度解释概念, 并使用这些形式化的概念做事情
- 从语形的角度给出基本公理, 讨论这些假设的推论和与其他假设的一致性

当代前沿的哲学研究很多会借助逻辑技术

学一点逻辑，特别是模态逻辑，至少可以帮助理解这些技术内容.

不过，技术复杂也不等于深刻，技术要与哲学的想法结合好，才能相辅相成，尽量做到清晰且深刻。

There are those who think that clarity, because it is difficult and rare, should be suspect. The rejection of this view has been the deepest impulse in all my philosophical work.

—Bertrand Russell

模态逻辑的一些特点

- 元语言概念拿到对象语言精确说、组合说
- 语言自然可以直接表达各种哲学概念的性质和关系
- 模态公理自然的对应于很多数学性质
- 基于克里普克模型关系模型的语义直观可以应用于各种领域
- 基本的模态逻辑和各种变种都有很好的计算性质
- 公理系统可以像数学理论一样刻画哲学理论
- 语言语义模型的定义可根据需要定义比较自由
- 很多时候从语义出发找语形推理规则
- 可以在哲学与数学及计算机之间架起桥梁
- ....

哲学在我国古书本名为道学。今日哲学者，希腊语斐罗梭斐之译文。其原义为爱智。故哲学家不忌怀疑而忌武断，不妨有所不知，而切不可强不知以为知。愿以孔子之言，与哲学系诸同学共勉之。

——蔡元培《北京大学民国十四年哲学系级友会纪念刊题词》

Take home message

$$\mathcal{K}p \rightarrow \mathcal{K}\mathcal{K}p$$

$$\neg\mathcal{K}p \rightarrow \mathcal{K}\neg\mathcal{K}p$$

$$\mathcal{K}p \rightarrow p$$

$$\mathcal{K}\exists x Px \rightarrow \exists x \mathcal{K} Px$$