(命题) 模态逻辑进阶

哲学数学计算机中的逻辑课程 (2016 年秋)

王彦晶
北大哲学系

2016 年 11 月 24 日

回顾

举例: 知识逻辑

回顾

基本模态逻辑的语言和语义

基本的模态逻辑语言 ($p \in \mathsf{PV}$):

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

语义定义在点模型 (pointed (Kripke) model)$\mathcal{M}, w = \langle W, R, V, w \rangle$ 上:

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

框架是没有赋值 V 的模型.

$\Gamma \vDash \varphi$ iff 对所有点模型 $\mathcal{M}, w \vDash \Gamma \implies \mathcal{M}, w \vDash \varphi$.

给定一个框架的类 $\mathbb{C}$, $\Gamma \vDash_{\mathbb{C}} \varphi$ iff 对所有基于$\mathbb{C}$里框架的点克里普克模型 $\mathcal{M}, w \vDash \Gamma \implies \mathcal{M}, w \vDash \varphi$.

基本系统 K 和其他常见公理

K 公理 $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$.

NEC 从 φ 得到 $\Box\varphi$ (必然化规则).

TAUT 命题逻辑的公理.

MP 从 $\varphi \to \psi$ 及 φ 得到 ψ.

可靠性和完全性:

$$\Gamma \vdash_{\mathsf{K}} \varphi \iff \Gamma \vDash \varphi$$

其他公理:

T $\Box p \to p$ 必然的都是真的

B $p \to \Box\Diamond p$ 来源于布劳威尔 (Brouwer)

D $\Box p \to \Diamond p$ 来源于道义逻辑 (Ought implies can)

4 $\Box p \to \Box\Box p$ 永远意味着永远永远

5 $\neg\Box p \to \Box\neg\Box p$ 你知道你不知道

$$\Gamma \vdash_{\mathsf{K}+?} \varphi \iff \Gamma \vDash_{\mathbb{C}?} \varphi$$

框架及其性质

我们说公理 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:

- T $\Box p \to p$ 自反性 (Reflexivity): 任意 x, xRx
- B $p \to \Box\Diamond p$ 对称性 (Symmetry): 任意 x, y, 若 xRy 则 yRx
- D $\Box p \to \Diamond p$ 持续性 (Seriality): 任意 x, 存在 y 使得 xRy
- 4 $\Box p \to \Box\Box p$ 传递性 (Transitivity) 任意 xyz 若 $xRy\&yRz$ 则 xRz
- 5 $\neg\Box p \to \Box\neg\Box p$ 欧几里得性 (Euclidean property) 任意 xyz 若 xRy 且 xRz 则 yRz 来源于《几何原本》: *things which equal the same thing also equal one another.*

语形的公理和语义的性质都要合理!

举例: 知识逻辑

举例: 知识逻辑 (EPISTEMIC LOGIC) 通常也叫认知逻辑 (马希文)

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

- 语言: “主体 i 知道 (*knows that*) φ” ($\mathcal{K}_i\varphi$).
- 模型里对**每个** i 有个关系 R_i.
- 关系要满足一些性质, 例如最强的要求: 自反传递对称 (等价关系, 直观上也可理解为 “不可区分关系”: 分不清 w 和 v).
- 在世界 w 上 i 知道 φ 当且仅当在所有从 w 出发 i 不可区分的世界上 φ 都为真.

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...

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

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

逻辑工具主要用来处理人脑想不清楚的事情.

经济学中的知识模型

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 (Common Knowledge). 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 是子集 (“公式”), $\mathcal{K}_i(E) = \{s \mid I_i(s) \subseteq E\}$. 例如 $\mathcal{K}_1(\{w\}) = \emptyset$, $\mathcal{K}_2(\{w\}) = \{w\}$. 这个模型和上一页的克里普克模型等价 (把 {w} 看成 p).

Aumann 用它证明了经济学里另一个著名的不可能定理: 如果两个人有相同的先验知识, 则他们不可能对有分歧的后验知识 (经过各自的实验获取新信息) 形成公共知识. 简而言之, 如果出发点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 agree to disagree!

公共知识 (COMMON KNOWLEDGE)

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

令命题 p 为 “应该靠右行驶”. 假设在一条比较窄的路上, 迎面快速开来了一辆车, 什么样的知识状态能让你觉得往右躲是安全的?

- 仅仅我知道 p ($\mathcal{K}_i p$) 行么?
- 我和他 (j) 都知道 p ($\mathcal{K}_i p \wedge \mathcal{K}_j p$) 行么?
- 两个人都知道两个人都知道 p 行么?

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”?

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

公共知识的语义定义

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:

- $\mathcal{E}_G\varphi := \bigwedge_{i\in G}\mathcal{K}_i\varphi$
- $\mathcal{C}_G\varphi := \mathcal{E}_G\varphi \wedge \mathcal{E}_G\mathcal{E}_G\varphi \wedge \mathcal{E}_G\mathcal{E}_G\mathcal{E}_G\varphi....$

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

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 i 的关系组成的路径走到哪里 φ 都成立.

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

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

公共知识的要求很高!

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:

$$
\begin{array}{ccccccc}
\underline{(2,3)} & \overset{B}{\text{——}} & (4,3) & \overset{A}{\text{——}} & (4,5) & \overset{B}{\text{——}} & (6,5) \cdots\cdots\cdots \\
| A & & & & & & \\
(2,1) & \overset{B}{\text{——}} & (0,1) & & & &
\end{array}
$$

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

公共知识的要求很高!

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...

- 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)!

博弈论的解概念也有类似问题!

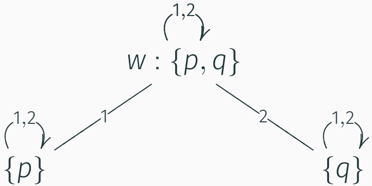
其他的群体知识: 分布式知识 (DISTRIBUTED KNOWLEDGE)

直观: 把我们知道的都放一起得到的知识. 语义上我能排除可能性 w 你能排除可能性 v, 我们合起来就能排除更多的可能性, 也就知道更多了. 用 $\mathcal{D}_G\varphi$ 表示 φ 是群体 G 的分布式知识

$$\boxed{\mathcal{M}, w \vDash \mathcal{D}_G\varphi \quad \Leftrightarrow \quad \text{所有 } w \text{ 如果} wR_{DG}v, \text{则} \mathcal{M}, v \vDash \varphi}$$

这里 $R_{DG} = \bigcap_{i \in G} R_i$.

在下面的模型上 $w \vDash \mathcal{K}_1 p \wedge \neg\mathcal{K}_2 q \wedge \mathcal{K}_2 q \wedge \neg\mathcal{K}_2 p \wedge \mathcal{D}_{\{1,2\}}(p \wedge q)$:



最强的 s5 系统 (对自反传递对称的框架类完全)

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

如何理解这几条公理:

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

知之为知之，不知为不知，是知也

公共知识和分布式知识的公理: 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.)

处理 D_G 需要在 S5 系统基础上增加:

- NEC 和 K 的 $\mathcal{D}_G$ 版本
- $\mathsf{D}_{\{i\}}\varphi \leftrightarrow \mathcal{K}_i\varphi$
- $\mathsf{D}_G\varphi \to \mathsf{D}_{G'}\varphi$ 如果 $G \subseteq G'$.

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 $B_i\varphi$ 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	$\mathcal{B}_i(\varphi \to \psi) \to (\mathcal{B}_i\varphi \to \mathcal{B}_i\psi)$	NECK	$\dfrac{\varphi}{\mathcal{B}_i\varphi}$
D	$\mathcal{B}_i\varphi \to \neg\mathcal{B}_i\neg\varphi$		
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, 但是要求信念至少是一致的, 你不能相信矛盾: $\mathcal{B}_ip \to \neg\mathcal{B}_i\neg p$ 等价于 $\neg\mathcal{B}_i\bot$ (有效性的意义下).

推理的一个例子

是不是所有真命题都可知?

假设 $p \wedge \neg K_i p$ 是真的, 但似乎不能说我知道我不知道原来今天是感恩节.

$\mathcal{K}_i(p \wedge \neg \mathcal{K}_i p)$ 在知识逻辑 S5 系统中是一个矛盾式.

1	$\mathcal{K}_i(p \wedge \neg \mathcal{K}_i p)$	
2	$(p \wedge \neg \mathcal{K}_i p) \rightarrow p$	TAUT
3	$(p \wedge \neg \mathcal{K}_i p) \rightarrow \neg \mathcal{K}_i p$	TAUT
4	$\mathcal{K}_i((p \wedge \neg \mathcal{K}_i p) \rightarrow p)$	NECK(2)
5	$\mathcal{K}_i((p \wedge \neg \mathcal{K}_i p) \rightarrow \neg \mathcal{K}_i p)$	NECK(3)
6	$\mathcal{K}_i p$	MP(公理K, 4, 1)
7	$\mathcal{K}_i \neg K_i p$	MP(公理K, 5, 1)
8	$\neg \mathcal{K}_i p$	公理T

其实上面只用到了 KT 的推理能力.

推理的一个例子

没有 T 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	$(p \wedge \neg \mathcal{B}_i p) \to p$	TAUT
3	$(p \wedge \neg \mathcal{B}_i p) \to \neg \mathcal{B}_i p$	TAUT
4	$\mathcal{B}_i((p \wedge \neg \mathcal{B}_i p) \to p)$	NECK(2)
5	$\mathcal{B}_i((p \wedge \neg \mathcal{B}_i p) \to \neg \mathcal{B}_i p)$	NECK(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	$\mathcal{B}_i(\mathcal{B}_i p \wedge \neg \mathcal{B}_i p)$	

只用到了系统 KD 的推理能力.

自省公理的讨论: 负自省公理 5

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

Lenzen (1978) 的例子:

- 假设 $\neg p \wedge \mathcal{B}_i\mathcal{K}_i p$: i 错误的相信他知道 p.
- 则 $\neg\mathcal{K}_i p$ (T 公理逆否形式 $\neg p \rightarrow \neg\mathcal{K}_i p$)
- 所以 $\mathcal{K}_i\neg\mathcal{K}_i p$ (公理 5及 MP)
- 则有 $\mathcal{B}_i\neg\mathcal{K}_i p$ (知识至少是信念)
- 最后有 $\mathcal{B}_i\bot$

我们没法表达关于知识的错误信念!

转换种形式: $\neg\mathcal{K}_i\neg\mathcal{K}_i p \rightarrow \mathcal{K}_i p$ 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: $\neg\mathcal{K}_i\neg\mathcal{K}_i p \rightarrow \mathcal{K}_i\neg\mathcal{K}_i\neg p$(公理 4.2)

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

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.

赞成正自省 (KK):

- 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 $\mathcal{K}\varphi \to \neg\neg\mathcal{K}\neg\psi$.).

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 $\neg\mathcal{K}\neg\mathcal{K}$).

如下的封闭性规则可能太强了:

- 从 $\vdash \varphi$ 得到 $\mathcal{K}\varphi$
- 从 $\vdash \varphi \to \psi$ 得到 $\mathcal{K}\varphi \to \mathcal{K}\psi$
- 从 $\vdash \varphi \leftrightarrow \psi$ 得到 $\mathcal{K}\varphi \leftrightarrow \mathcal{K}\psi$

K 公理也可看成一种封闭原则: $\mathcal{K}(\varphi \to \psi) \wedge \mathcal{K}\varphi \to \mathcal{K}\psi$ 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 ($\neg q$) $\mathcal{K}(p \to \neg q)$, 而且我知道我有手 $\mathcal{K}q$, 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 ($\neg\mathcal{K}\neg p$). 其实两个 "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, 计算机科学家也不喜欢逻辑全知性, 特别在涉及加密通信的时候.

解决的办法:

- 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

计算机中的应用

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: 如通信协议和程序
- 基于知识的程序 (knowledge-based program)
- 基于不确定性的自动规划
- 非完美信息博弈的求解
- 多主体系统
- 信念修正
- ...

知识逻辑领域的 TARK 会议系列

Theoretical Aspects of Rationality and Knowledge

- 从 1986 开始每两年一届, J. Halpern 发起
- 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相关的问题.
- 86 年的会议群星闪耀: Aumann, Hintikka, McCarthy, Stalnaker, Halpern, Vardi, Moses, Fagin, Dwok, Smullyan, Moore, Asher, Kamp, Levesque, Immerman, Plotkin, Ladner, Fischer....
- `http://www.tark.org/`