

# 模态逻辑

杨睿之

复旦大学哲学学院

2024 年秋季

# 前情提要

- $\kappa$ -饱和模型、基于可数不完全超滤的超幂
- 迂回引理
- Van Benthem 刻画定理

# 认知逻辑

考虑一个人 (agent) 的 **认知逻辑** (epistemic logic)。中我们用  $K\varphi$  表示这个人 “知道  $\varphi$ ”。即其语言为

$$\phi ::= p \mid \perp \mid \neg\phi \mid \phi \vee \psi \mid K\phi$$

## 例

- $K(\varphi \rightarrow \psi) \rightarrow K\varphi \rightarrow K\psi$
- $K\varphi \rightarrow \varphi$
- $K\varphi \rightarrow KK\varphi$

# 认知逻辑

## 认知逻辑的语义

- $K$  一般被看作是  $\Box$  算子, 即在一个关系模型中  $(W, R, V), w \Vdash K\varphi$ , 当且仅当对任意  $v \in W, R_{wv}$  蕴涵  $v \Vdash \varphi$
- 我们可以用  $\langle K \rangle$  表示  $K$  的  $\Diamond$  对偶

# 认知逻辑

不可区分性解释 (indistinguishability interpretation)

- 在认知逻辑框架  $(W, R)$  中,  $W$  中元素被理解为可能的 **状态**,  $R_{wv}$  被理解为处在状态  $w$  的 agent 在认知上无法区分状态  $w$  和状态  $v$
- $w \models K\varphi$  (agent 知道  $\varphi$ ), 当且仅当在任何 agent **无法与  $w$  区分的状态  $w'$  中**  $\varphi$  都成立

# 认知逻辑

## 例 (批改论文的教师)

- $p$ : 论文优秀,  $q$ : 论文由 LLM 生成
- $w_1 = \{p, q\}$ ,  $w_2 = \{p, \neg q\}$ ,  $w_3 = \{\neg p, q\}$ ,  $w_4 = \{\neg p, \neg q\}$
- 假设该教师在面对一篇不优秀的论文时可以识别出该论文的问题是由于学生不熟练还是 LLM 的缺陷, 从而判断该论文是否为学生独立完成; 而面对一篇优秀的论文是则难以判断。基于不可区分性解释, 对应的框架  $(W, R)$  是?

# 认知逻辑

一般认为，不可区分性解释下，一个认知逻辑框架  $(W, R)$  中的可及关系是

- 自反的
- 对称的
- 传递的

也即是一个等价关系

# 认知逻辑

对应地，下列公式在  $R$  是等价关系的框架中有效

- $T : K\varphi \rightarrow \varphi$
- $B : \varphi \rightarrow K\langle K \rangle\varphi$
- $4 : K\varphi \rightarrow KK\varphi$
- $5 : \langle K \rangle\varphi \rightarrow K\langle K \rangle\varphi$



# 认知逻辑

回忆:  $K$  公式  $K(\varphi \rightarrow \psi) \rightarrow K\varphi \rightarrow \psi$  在任何关系框架中有效, 我们定义模态逻辑

- $K = \{K\}$
- $S4 = \{K, T, 4\}$
- $S5 = \{K, T, 5\}$

# 认知逻辑

## 事实

- $S5 = \{K, T, 5\} = \{K, T, B, 4\}$
- 5 公理逻辑等价于:  $\neg K\varphi \rightarrow K\neg K\varphi$

# 认知逻辑

- 不可区分性解释下关于可及关系是等价关系的假设或许有问题?
- 一般认可认知逻辑是  $S4$  的

# 知道与相信

JTB: 知识是被确证的真信念

- 在语言中分别用模态词  $K$ 、 $B$  表示 “知道” 与 “相信”

$$\phi ::= p \mid \perp \mid \neg\phi \mid \phi \vee \psi \mid K\phi \mid B\phi$$

- 对应的框架中分别用  $R_K$ 、 $R_B$  表示对应的可及关系
- 语义上,  $K$ 、 $B$  分别是对应  $R_K$ 、 $R_B$  的 “ $\Box$  算子”
- 如果认为  $K\phi \rightarrow B\phi$  是有效的, 那么  $R_B \subset R_K$

# 知道与相信

- 一般认为知识是真的, 即  $K\varphi \rightarrow \varphi$ , 信念未必
- 一般认为信念至少是一致的, 即  $\neg B\perp$

## 事实

在任意框架  $(W, R)$  中,  $\neg B\perp$  是有效的, 当且仅当

**D**:  $Bp \rightarrow \langle B \rangle p$  是有效的

# 知道与相信

## $K$ 与 $B$ 的联系

- **KB1** :  $K\varphi \rightarrow B\varphi$
- **KB2** :  $B\varphi \rightarrow KB\varphi$
- **KB3** :  $B\varphi \rightarrow BK\varphi$

# 知道与相信

## 事实

- $\vdash_{K_K, K_B, 5_K, D_B, KB1} BK\varphi \rightarrow K\varphi$
- $\vdash_{K_K, K_B, 5_K, D_B, KB1, KB3} K\varphi \leftrightarrow B\varphi$

# 多主体认知

- 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？”



# 多主体认知

- 多主体认知逻辑 (multi-agent epistemic logic) 的语言

$$\phi ::= p \mid \perp \mid \neg\phi \mid \phi \vee \psi \mid K_a\phi, (a \in A)$$

其中  $A$  是一个 agent 集

- 多主体认知逻辑的语义：对每个  $a \in A$  有可及关系  $R_a$

# 多主体认知

- 在单一主体认知逻辑的框架中，假设需要考虑的基本事态（命题变元）个数为  $n$ ，那么所需考虑的状态的个数不超过  $2^n$
- 在多主体认知逻辑中，可能需要更多

## 例

假设只考虑一个基本事态  $p$ ：

$$K_a p \wedge K_b p \wedge \neg K_b K_a p$$

# 群体认知

在多主体认知逻辑的基础上，我们可以考虑群体认知。令

$G \subset A$  代表一个群体

- 我们用  $D_G\varphi$  表示  $\varphi$  是群体  $G$  的 **分布式知识** (distributed knowledge), 即如果群体  $G$  中的成员充分交流后能掌握  $\varphi$
- 我们用  $E_G\varphi$  表示群体  $G$  中的每个主体都知道  $\varphi$
- 我们用  $C_G\varphi$  表示  $\varphi$  是群体  $G$  中的 **公共知识** (common knowledge), 即群体中的每个主体都知道  $\varphi$  并且也知道别的主体知道  $\varphi$ .....

# 群体认知

## 定义

- 定义  $R_G^D = \bigcap_{a \in G} R_a$ ,  $\mathfrak{M}, w \Vdash D_G \varphi$ , 当且仅当对任意  $v$ , 若  $R_G^D wv$ , 则  $v \Vdash \varphi$
- 如果  $G$  是有穷的, 可以定义  $E_G \varphi := \bigwedge_{a \in G} K_a \varphi$
- 如果允许无穷语句, 可以定义  $C_G \varphi := \bigwedge_{k \in \mathbb{N}} E_G^k \varphi$
- 定义  $R_G^C = (\bigcup_{a \in G} R_a)^*$ , 其中  $(\cdot)^*$  表示取自反传递闭包

# 动态认知逻辑

**动态认知逻辑** (dynamic epistemic logic) 引入了行动算子, 对应于模型转变, 可以用于刻画一些多主体的认知博弈。

# 动态认知逻辑

## 例 (丙的生日)

甲和乙问丙的生日是几号。丙公开列出了下列候选

- 5.15、5.16、5.19、6.17、6.18、7.14、7.16、8.14、8.15、8.17

接着丙告诉了甲生日是几月，告诉了乙几号。然后丙问二位，我的生日是几号。甲先说：我不知道你的生日是几号，但我知道乙也不知道；接着乙说：我之前是不知道，但我现在知道了；然后甲说：好吧，我现在知道了。请问丙的生日是？

# 动态认知逻辑

## 例 (Muddy Children)

三个小孩顽皮归来，每个人额头上都可能沾着泥（自己感觉不到）。老师让三个小孩围成一圈，这样他们可以看到另两位的脸。老师说：你们中至少有一个额头上沾着泥。如果你知道你自己额头上是否沾着泥，请上前一步。话毕，没有小孩上前。老师重复道：如果你知道你自己额头上是否沾着泥，请上前一步。这次有小孩上前一步，但不是所有小孩。老师第三次重复那句话，这次所有小孩都上前了一步。