Modal Logic of Permission and Normality

关于允许和常态的模态逻辑

董惠敏

浙江大学哲学系

复旦大学哲学学院 · Fudan Logic Seminar

2020 年 10 月 16 日，线上讲座

内容简介

- 允许：作为道义充分条件；一个简单的哲学想法
- 允许：中文“可以”的语言示例；博弈解的逻辑刻画
- 一个关于可废止允许的模态逻辑
- 一个动态更新模型

在谈论“什么是好的”和“什么是坏的”时，可能涉及如下方面的概念：

1. 规范/道义 (normative/deontic)：义务，允许，禁止；
2. 规则：规则、命令；
3. 实践：主体，活动；
4. 法律：主张，权利，豁免；
5. 策略：偏好、收益、效用；
6. 道德：价值；
7. ...

道义逻辑的简单历史

- 希腊语“déon”：“什么是受到约束的”、“什么是适当的”。
- 对道义的研究是一种“关于道德的科学（The science of morality)”(Benthem 1993)。
- 在莱布尼茨 1671 年的著作《*Elementa iuris naturalis*》中就伦理学与道义逻辑进行研究，尤其是他提出模态语义的基本思想去刻画义务、允许、禁止等道义概念。
  - 义务 (the obligatory, *debitum*)：
    一个好人所必须做的（what is necessary for a good man to do)。
  - 标准道义逻辑 (standard deontic logic) 继承了他的模态思想，通过义务这个模态来定义允许、禁止。
  - 允许是不被禁止的、没有义务不去做，存在一种实施它的好的方式。
- 动态逻辑、计算机程序语言 (Digum et al. 1996)：
  - 允许：所做的不会导致坏结果，即充分好 (what is sufficient for being good)。

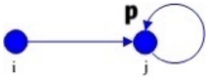
允许：作为道义充分条件

- 模态语言 $\{\neg, \wedge, \rightarrow, A, P, O\}$；
- 持续框架 $F = \langle W, R\rangle$ 为道义框架；
- wRu：“u 是 w 可通达/可想象/可看到的理想世界”；
- 持续模型 $M = \langle W, R, ||\cdot||\rangle$ 作为道义模型；（$M \in F$）
- 模型可满足： $M, w \vDash \varphi$
- 模型有效： $M \vDash \varphi$ iff $\forall w \in W.\ M, w \vDash \varphi$
- 框架有效： $F \vDash \varphi$ iff $\forall M \in F.\ M \vDash \varphi$

$P\varphi$： φ 是允许的

$$\overbrace{||\varphi|| \subseteq R[w]} \subseteq ||\varphi|| \qquad ||\varphi|| \subseteq \underbrace{R[w] \subseteq ||\varphi||}$$

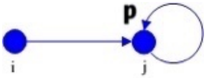
$O\varphi$： φ 是必须的



$P\varphi$：φ 是允许的

$$||\varphi|| \subseteq R[w] \subseteq ||\varphi||$$

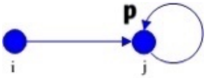
$O\varphi$：φ 是必须的



例子：令 $||p|| = \{j\}$ 且 $||q|| = \emptyset$。

1. $\neg(Op \wedge O\neg p)$ 在可能世界 i, j 上为真吗？

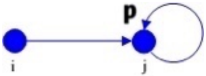
$$\overbrace{\|\varphi\| \subseteq \rlap{$\underbrace{}_{O\varphi:\ \varphi\ \text{是必须的}}$}R[w]}^{P\varphi:\ \varphi\ \text{是允许的}} \subseteq \|\varphi\|$$



例子：令 $\|p\| = \{j\}$ 且 $\|q\| = \emptyset$。

1. $\neg(Op \land O\neg p)$ 在可能世界 i, j 上为真吗？是，且框架有效。
2. $Op \to p$ 在可能世界 i, j 上为真吗？

$$\overbrace{||\varphi|| \subseteq \mathrlap{\underbrace{}_{O\varphi:\ \varphi\text{ 是必须的}}}R[w]}^{P\varphi:\ \varphi\text{ 是允许的}} \subseteq ||\varphi||$$



例子：令 $||p|| = \{j\}$ 且 $||q|| = \emptyset$。

1. $\neg(Op \wedge O\neg p)$ 在可能世界 i, j 上为真吗？是，且框架有效。
2. $Op \to p$ 在可能世界 i, j 上为真吗？仅在 j 上为真。
3. $P(p \vee q) \to Pp \wedge Pq$ 在可能世界 i, j 上为真吗？是；框架有效。
4. $Pp \to P(p \wedge q)$ 在可能世界 i, j 上为真吗？是；框架有效。

$P(p \vee q) \to Pp \wedge Pq$ 在任意道义模型 M 下是有效的。

证明.

任意给定一个点 $w \in M$，要证 $M, w \vDash P(p \vee q) \to Pp \wedge Pq$。要注意，$||p \vee q|| = ||p|| \cup ||q||$ 成立。

$$
\begin{aligned}
& M, w \vDash P(p \vee q) \\
\text{iff } & ||p \vee q|| \subseteq R[w] \\
\text{iff } & ||p|| \cup ||q|| \subseteq R[w] \\
\text{iff } & ||p|| \subseteq R[w] \,\&\, ||q|| \subseteq R[w] \\
\text{iff } & M, w \vDash P\varphi \,\&\, M, w \vDash P\psi \\
\text{iff } & M, w \vDash P\varphi \wedge P\psi
\end{aligned}
$$

通过上面的证明也可以知道，$Pp \wedge Pq \to P(p \vee q)$ 也是有效的。 □

“可以”：中文的语言示例

在中文里，“可以”这个情态动词，至少有三种义项：

1. 表示主客观条件下所容许的行为（ablility/capacity）：

   你可以上班了。

2. 表示情理上的许可：

   你可以讲你想讲的。

3. 表示有价值的行为：

   你可以读读这本书。

我们通常使用“可以”作为表达允许这个道义概念的模态词。

“可以”：两种连词连用

- 在复句中，“可以”这个情态动词与不同连词结合，形成的句群可以表达不同的关联关系。
- 在这里，主要关注于递进关系和转折关系。

“并且”：表递进关系

后一分句关涉的内容比前一分句的更推进一层：

- 你可以声明保留不同的意见，并且可以把自己的意见向上级提出。

当声明保留意见是充分好的，那么，一般地/常规来说，进一步的行为向上级提意见也是充分好的。

“但是”：表转折关系

但并非所有的递进行为都是充分好的。

1. 你可以向上级提出不同意见，但不可以擅自作出决定。(伦理上)
2. 你可以爱你的孩子，但你不可以溺爱他们。(伦理上)
3. 你可以上学，但你不可以摘下心率监测仪。(策略上)

后一分句的递进行为并非也是（伦理上、策略上）充分好。

如何挑出那些充分好的递进行为？

可采用反事实假设 (Lewis, Counterfacturals, 1973)，最常见的（most normal）递进行为通常来说是足够充分的。

逻辑与博弈

- 近年来，不同的形式方案被提出用以刻画博弈的逻辑结构 (Tamminga 2013; van Benthem 2014)；
- 优点：
  - 对博弈相关概念进行逻辑表征（可表达性）；
  - 构建相关的可计算博弈模型；
  - 以可满足性、模型检测等逻辑技术，检查博弈概念的一致性；
  - ...

允许：作为博弈解

其中一种是道义方案 (Roy et al. 2014)：

通过允许和义务去表征与博弈解，分别是博弈解的充分、必要条件。

在“囚徒困境”中，假设博弈解为纳什均衡，那么：

- $R[w]$：在当前状态 w 下的博弈解集，$\{(\mathrm{Defect}, \mathrm{Defect})\}$；
- 可检测 $P(\mathrm{Defect}, \mathrm{Defect})$、$O(\mathrm{Defect}, \mathrm{Defect})$ 是否为真。

		Player A	
		Coordinate	Defect
Player B	Coordinate	3, 3	0, 5
	Defect	5, 0	1, 1*

Number left (right) of comma refers to A’s (B’s) preference ordering (0 = worst outcome; 2 = best outcome). * indicates the Nash equilibrium.

Figure 1: Prisoner’s Dilemma

合作博弈

在合作博弈“高低博弈”中，当博弈解为**纳什均衡**时：

- $R[w]$：在当前状态 w 下的博弈解集，$\{(\mathrm{High}, \mathrm{High}), (\mathrm{Low}, \mathrm{Low})\}$；
- 可检测到 $P(\mathrm{High}, \mathrm{High})$、$P((\mathrm{Low}, \mathrm{Low}))$、$O[(\mathrm{High}, \mathrm{High}) \vee (\mathrm{Low}, \mathrm{Low})]$ 为真；
- 问题：参与者应该怎样选？合作如何可能？(合作背后的理性假设？)

		Player A High	**Player A** Low
Player B	High	2, 2*	0, 0
Player B	Low	0, 0	1, 1*

Number left (right) of comma refers to A's (B's) preference ordering (0 = worst outcome; 2 = best outcome). * indicates the Nash equilibrium.

Figure 2: The Hi-Lo Game

博弈转换

在合作博弈“高低博弈”中，当博弈解为帕累托最优时：

- $R[w]$：在当前状态 w 下的博弈解集，$\{(\mathrm{High}, \mathrm{High})\}$；
- 可检测到 $P(\mathrm{High}, \mathrm{High})$、$\neg P((\mathrm{Low}, \mathrm{Low}))$、$O(\mathrm{High}, \mathrm{High})$ 为真；
- 将最常见的博弈解选出来。

		Player A: High	Player A: Low
Player B	High	2, 2*	0, 0
Player B	Low	0, 0	1, 1*

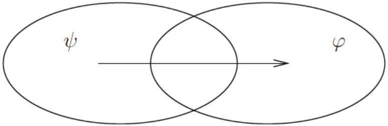
Number left (right) of comma refers to A's (B's) preference ordering (0 = worst outcome; 2 = best outcome). * indicates the Nash equilibrium.

Figure 2: The Hi-Lo Game

允许与常态的道义逻辑

$$\varphi := p \mid \varphi \wedge \varphi \mid \neg\varphi \mid P\varphi \mid O\varphi \mid \varphi \trianglelefteq \varphi$$

- $\varphi \trianglelefteq \psi$：$\psi$ **至少和** φ **类似**。(Boutilier 1994; Halpern 1997)
- **全称模态** $A\varphi := ((\neg\varphi) \trianglelefteq \bot)$；**存在模态** $E\varphi := \neg A \neg\varphi$；
- 道义偏好**框架** $F = \langle W, R, \leq \rangle$：$R$ **持续**；$\leq$ **自返传递**；
- 道义偏好框架类**记为** $\mathcal{F}$
- 道义偏好**模型** $M = \langle F, || \cdot || \rangle$
  - $M, w \models \psi \trianglelefteq \varphi$ iff $\forall u \in W \exists v \in W. (M, u \models \psi \Rightarrow M, v \models \varphi \;\&\; u \leq v)$

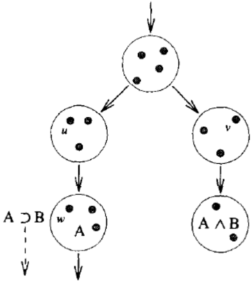


最常见

$$M, w \models \Box(\psi/\varphi)$$
$$\text{iff}$$
$$\forall v \in W\, [M, v \models \varphi \Rightarrow \exists u \geq v\, (M, u \models \varphi\ \&\ \forall s \geq u\, (M, s \models \varphi \rightarrow \psi))].$$

在最常见的 φ-状态下，ψ 总成立；φ 是 ψ 的递进、常见子类。

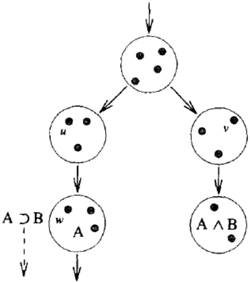
$$\Box(\psi/\varphi) \leftrightarrow [E\varphi \rightarrow (\varphi \wedge \neg\psi) \trianglelefteq (\varphi \wedge \psi)].$$



可废止允许

$$M, w \models P\varphi$$
iff
$$\forall v \in W[M, v \models \varphi \Rightarrow \exists u \geq v(M, u \models \varphi \,\&\, \forall s \geq u(s \in ||\varphi|| \Rightarrow s \in R[w]))].$$

最常见的 φ-状态，总是理想的、理性的、合理的。



合理子类

$$\heartsuit(\varphi/\psi) \leftrightarrow (\psi \trianglelefteq \varphi) \wedge (\varphi \trianglelefteq \psi) \wedge \Box(\varphi/\psi)$$

最常见的 ψ-状态也是最常见的 φ-状态；ψ 是 φ 的合理子类。

有效式

- $O\varphi \wedge P\psi \to \Box(\varphi/\psi)$：可允许的是可允许的常见子类。
- $P\varphi \wedge \heartsuit(\varphi/\psi) \to P\psi$：可允许活动的合理子类，也是可允许的。

道义逻辑系统

浙江大学 ZHEJIANG UNIVERSITY

刻画关于允许、义务、常态等概念的道义逻辑系统 **N** 定义如下：

- Tautologies
- The binary modality $\trianglelefteq$ satisfies:

  REF: $(\psi \to \varphi) \to \psi \trianglelefteq \varphi$

  TRAN: $(\varphi \trianglelefteq \chi) \wedge E(\chi \trianglelefteq \psi) \to (\varphi \trianglelefteq \psi)$

  CET: $(\varphi \trianglelefteq \psi) \vee (\psi \trianglelefteq \varphi)$

  $\mathrm{B}_{\trianglelefteq}$: $(\psi \trianglelefteq \varphi) \wedge (\chi \trianglelefteq \varphi) \to (\psi \vee \chi) \trianglelefteq \varphi$

  $\bot_{\trianglelefteq:l}$: $\bot \trianglelefteq \varphi$

  $\bot_{\trianglelefteq:r}$: $\neg(\varphi \trianglelefteq \bot)$

  $\mathrm{RM}_{\trianglelefteq:l}$: $\chi \to \varphi / \varphi \trianglelefteq \psi \to \chi \trianglelefteq \psi$

  $\mathrm{RM}_{\trianglelefteq:r}$: $\varphi \to \psi / \chi \trianglelefteq \varphi \to \chi \trianglelefteq \psi$
- The reduction axioms:

  $\mathrm{Red}_{\Box}$: $\Box(\varphi/\psi) \leftrightarrow [E\varphi \to (\varphi \wedge \neg\psi) \trianglelefteq (\varphi \wedge \psi)]$

  $\mathrm{Red}_{\heartsuit}$: $\heartsuit(\varphi/\psi) \leftrightarrow (\psi \trianglelefteq \varphi) \wedge (\varphi \trianglelefteq \psi) \wedge \Box(\varphi/\psi)$

- The modalities A and E are defined as:

  $A\varphi \leftrightarrow (\neg\varphi) \trianglelefteq \bot$

  $E\phi \leftrightarrow \neg A \neg\phi$
- The modality P satisfies:

  OWP: $O\varphi \wedge P\psi \to \Box(\varphi/\psi)$

  PtF: $P\bot$

  FCP: $P\varphi \wedge \heartsuit(\varphi/\psi) \to P\psi$

The modality O satisfies:

  K_O: $O(\varphi \to \psi) \wedge O\varphi \to O\psi$

  NEC_O: $\varphi / O\varphi$

  OiC: $O\varphi \to E\varphi$

  D_O: $\neg O\bot$

定理 (可靠性)

道义逻辑系统 N 相对于道义偏好框架类 $\mathcal{F}$ 是可靠的，即

$$\vdash_{\mathbf{N}} \varphi \Rightarrow \mathcal{F} \vDash \varphi$$

定理 (完全性)

道义逻辑系统 N 相对于道义偏好框架类 $\mathcal{F}$ 是（强）完全的，即

$$\mathcal{F} \vDash \varphi \Rightarrow \vdash_{\mathbf{N}} \varphi$$

特定语境

语境模型 $C^\Gamma = \langle C(\Gamma), \preceq \rangle$:

- $C(\Gamma) = \{\{\pm p \mid p\text{是出现于}\Gamma\text{的原子命题}\} \mid \text{或} \pm p = p\text{或} \pm p = \neg p\}$;
- $\preceq \subseteq C \times C$: 自返、传递。

动态模型：刻画 Pareto 最优的字典更新

给出道义模型 $M = \langle W, R, \leq, ||\cdot||\rangle$ 和语境模型 $\mathcal{C}^{\Gamma} = \langle C, \preceq\rangle$。

更新模型 $M \otimes \mathcal{C}^{\Gamma} = \langle W^*, R^*, \leq^*, V^*\rangle$ 定义如下：

- $W^* = \{(u, c) \mid M, u \vDash c$，其中$c \in C\}$;
- $(u, c) \leq^* (v, d)$ iff
  - 或 $c \prec d$
  - 或 $(c \sim d$ and $u \leq v)$;
- $(u, c)R^*(v, d)$ iff
  - 或 $c \prec d$
  - 或 $(c \sim d$ 且 $uRv)$，使得 R^* 是持续的：

$$\forall(u, c)\exists(v, d).\text{或}c \prec d\text{或}(c \sim d \,\&\, uRv);$$

- $(u, c) \in V^*(p)$ iff $u \in V(p)$.

为了保证这个更新模型也是一个道义模型，需要在系统中加入公理 $c \to \bigvee_{d \succ c} Ed \vee \bigvee_{d \sim c} \neg O \neg d$。

更新算子的真值

$$M, w \vDash \langle\Gamma\rangle\varphi \text{ iff } \exists(w, c) \in W^{*} \text{ s.t. } M \otimes C^{\Gamma}, (w, c) \vDash \varphi.$$

由于 $(w, c) \in W^{*}$ 在更新模型中，这意味着 $M, w \vDash c$。于是可证明：
$M \otimes C^{\Gamma}, (w, c) \vDash \varphi$ iff $M, w \vDash c \wedge \langle\Gamma\rangle\varphi$.

更新的归约公理

- $[\Gamma](\varphi \trianglelefteq \psi) \leftrightarrow \bigwedge_{c\in C}\bigwedge_{d\in C}[c \to \bigvee_{e\succ d} A(\varphi^d \to E\psi^e) \vee \bigvee_{e\sim d}(\varphi^d \trianglelefteq \psi^e)]$
- $[\Gamma]\Box(\psi/\varphi) \leftrightarrow [(\bigwedge|\varphi| \leftrightarrow \bot) \to$
  $\bigwedge_{c\in C}[c \to \bigwedge_{d\in \max(|\varphi|)} \Box(\bigwedge_{f\sim d}(\varphi^f \to \psi^f)/\bigvee_{b\sim d}\varphi^b)]]$
- $[\Gamma]O\varphi \leftrightarrow \bigwedge_{c\in C}[c \to \bigwedge_{d\succ c} A\varphi^d \wedge \bigwedge_{d\sim c} O\varphi^d]$
- $[\Gamma]P\varphi \leftrightarrow [(\bigwedge|\varphi| \leftrightarrow \bot) \to \bigwedge_{c\in C}[c \to \bigwedge_{d\in \max(|\varphi|)\cap\sim[c]} P\bigvee_{b\sim d}\varphi^b]]$

其中，

- $\varphi^e := e \wedge \langle\Gamma\rangle\varphi$；
- $|\varphi| = \{a \in A \mid\not\vdash \neg(a \wedge \langle\Gamma\rangle\varphi)\}$；
- $\sim[c] = \{a \in A \mid a \sim c\}$。

结语

- 基于反事实的哲学思想，通过“最常见”这个概念去限制作为道义充分条件的允许，可以更好地刻画自然语言、博弈解；
- 这里构建的刻画允许、义务、常态的动态道义逻辑，是可靠和强完全的；
- 可尝试与其他可废止框架进行比较；
- 可尝试对比这里的反事实假设下的形式框架与 Ceteris Paribus “其他条件不变”等与因果推理紧密关联的形式框架。

感谢！