

# 通用智能框架下的纽康姆难题

李熙

**摘要:** 面对纽康姆难题, 主流的决策理论出现了分歧, 采用期望效用最大化原则和占优原则分别会导向两种不同的选择, 这对决策理论构成了一个挑战。通过合适的概率计算, 占优原则可以看作一种极端的期望效用最大化原则, 所以决策理论的基础——追求效用最大化并不与占优原则冲突, 问题是——如何计算期望效用。纽康姆难题背后的关键也是如何通过概率把握因果, 其核心是归纳预测问题。通用人工智能领域的通用归纳、通用智能模型刻画的恰恰是因果预测问题, 本文认为, 为了解决纽康姆难题发展各种新奇的决策理论是不必要的, 而借助通用智能模型 AIXI 探讨纽康姆难题中涉及到的因果性问题。

**关键词:** 纽康姆难题; 证据决策论; 因果决策论; 通用归纳

**中图分类号:** B81

**文献标识码:** A

对于一个理性的主体来说, 如果已知每个行为可能带来的后果, 他会选择能够带来最大期望效用的行为。但如何精确地刻画并计算期望效用却是一个难题, 尤其在涉及不确定情境下的决策时, 对于如何用概率刻画可能后果的不确定性, 并不是一件平凡的事。证据决策论与因果决策论是两种比较主流的决策理论, 在分析纽康姆难题时, 通常认为, 证据决策论采用期望效用最大化进行决策, 因果决策论采用“占优原则”进行决策, 事实上, 通过合适的概率计算, “占优原则”可以看作特殊的期望效用最大化, 也就是说, 两种决策论都通过期望效用最大化进行决策, 但二者对于“期望”的计算方式不一样。比如在纽康姆难题中, 证据决策论 ([4]) 与因果决策论 ([5]) 导出了两种完全不同的行为方式。

纽康姆难题最初由物理学家纽康姆提出、由哲学家诺齐克 (R. Nozick) 发表 ([8]), 它指如下这种谜题:

一个女巫向你展示了两个箱子  $a$  和  $x$ , 其中  $a$  透明  $x$  不透明。透明箱子  $a$  中有很少的钱 ( $L$  元), 不透明箱子  $x$  中可能有很多钱 ( $B$  元,  $B \gg L$ ), 也可能什么都没有, 你可以两个箱子都拿走也可以只拿不透明的箱子  $x$ 。

---

收稿日期: 2017-05-18

作者信息: 李熙      中南大学哲学系  
xieshenlixi@163.com

基金项目: 国家社科基金项目“通用人工智能的哲学基础研究”(17CZX020)。

致谢: 感谢评审专家提出的修改意见和建议, 使得本文的可读性得到了提高。

女巫声称她可以提前预测你的行为，如果她预测到你会两个箱子都拿，她就让  $x$  空着，如果她预测到你只拿  $x$ ，她就会放很多钱（ $B$  元）在里面。你们已经玩了好多回合了，女巫每次都猜对。下一次你该怎么选？<sup>1</sup>

		女巫	
你	选择 \ 预测	$x^-$	$x^+$
	$a^-$	0	$B$
	$a^+$	$L$	$L + B$

这个难题涉及到因果决定论与自由意志这一宏大哲学主题，所以倍受哲学家关注。除了因果决策论和证据决策论，为了解决这个难题，甚至产生了几个专门的决策理论。比如连贯决策论 ([1])，无时间决策论 ([14])，无更新决策论 ([10])，量子决策论 ([3]) 等等。从结果来看，这些决策理论要么支持选择一个箱子，要么支持选择两个箱子。但还有另外一种消解难题的方式，那就是彻底否定这个游戏本身的合理性，认为这不是一个严格定义的博弈框架，不同的解读可以导出不同的博弈，比如 [13]。

一般认为，证据决策论采用经典的条件概率，因果决策论采用反事实条件概率，其他形形色色的决策理论也基本是在设计各种“条件概率”。采用“消解”方案的人一般认为，经典的条件概率理论已足够，对游戏的不同解读方式对应不同的条件概率计算方式。本文也认为，发展一套专门的反事实条件概率或其他形式的条件概率的理论是不必要的，本文将从胡特尔 (M. Hutter) 的通用智能模型 AIXI<sup>2</sup> 入手，认为在纽康姆难题中，关键在于如何用概率论刻画未知的“因果性”，从算法概率的角度看，关键在于，环境类/假设空间如何设定。

## 1 通用智能模型 AIXI 简介

**记号.** 字母表  $\mathcal{X}$  上的有穷序列记为  $\mathcal{X}^*$ ，无穷序列记为  $\mathcal{X}^\infty$ ，记  $\mathcal{X}^\# := \mathcal{X}^* \cup \mathcal{X}^\infty$ ，空序列记为  $\epsilon$ 。有穷序列  $x \in \mathcal{X}^*$  的长度记为  $|x|$ ， $\mathcal{X}^t := \{x \in \mathcal{X}^* : |x| = t\}$ ， $\mathcal{X}^{\leq t} := \bigcup_{i=0}^t \mathcal{X}^i$ 。长度为  $n$  的序列  $x_1 x_2 \dots x_n$  简记为  $x_{1:n}$ ，类似的， $x_{t:n} := x_t x_{t+1} \dots x_{n-1} x_n$ ，

<sup>1</sup>在原始的纽康姆难题 ([8]) 设定中，你看到女巫跟很多人玩，她几乎从来没输过。现在轮到你了，你该如何选择？本文对此做了微小的改动。

之所以称其为纽康姆“难题”，是因为，根据主流的决策理论，比如证据决策理论，你最好只拿一个箱子；而根据因果决策理论，两个箱子都拿是一种“占优”策略。这在一定程度上非常反直观，所以虽然不构成一个悖论，却是一个难题。虽然本文对游戏设定略有修改，但这种反直观的情形依然存在。关于纽康姆难题更细致的哲学分析可以参看张建军 ([15])。

<sup>2</sup>关于 AIXI 的详细知识请参考 [2] 和 [6]，其中涉及到的关于柯尔莫哥洛夫复杂性的知识请参考 [7]，博弈论的知识请参考 [9]。

$x_{<n} := x_1 \dots x_{n-1}$ ,  $ax_{1:n} := a_1x_1a_2x_2 \dots a_nx_n$ ,  $ax_{<j} := a_1x_1a_2x_2 \dots a_{j-1}x_{j-1}$ , 两个序列  $x$  和  $y$  的连接为  $xy$ .  $x*$  表示序列  $x$  后连接任意长的序列.  $x$  是  $y$  的前缀定义为  $x \preceq y \Leftrightarrow \exists z(y = xz)$ , 真前缀  $x \prec y \Leftrightarrow x \preceq y \wedge x \neq y$ . 如果没有非空序列  $x, y \in A$  使得  $xy \in A$ , 则称集合  $A \subset \mathcal{X}^*$  为无前缀集.

字母表  $\mathcal{X}$  上的全序  $<$  可以诱导出有穷字符串  $\mathcal{X}^*$  上的准字典序,

$$x < y \Leftrightarrow [|x| < |y| \vee (|x| = |y| \wedge \exists z \in \mathcal{X}^*, a \in \mathcal{X}, b \in \mathcal{X} (x = za^* \wedge y = zb^* \wedge a < b))]$$

艾佛森括号

$$[\![\phi]\!] := \begin{cases} 1 & \text{如果 } \phi \text{ 成立} \\ 0 & \text{否则} \end{cases}$$

柯尔莫哥洛夫复杂性函数  $K(x|y) := \min_p \{|p| : U(\langle p, y \rangle) = x\}$ ,  $K(x) := K(x|\epsilon)$ , 其中  $U$  是通用前缀图灵机. 对于那些不是字符串的对象  $o$ , 定义  $K(o) := K(\langle o \rangle)$ , 其中  $\langle o \rangle \in \mathcal{X}^*$  是对  $o$  的某种标准编码.

记  $\Delta(X)$  为  $X$  上的概率分布的集合.

在强化学习框架中, 主体与环境不断交互, 在第  $k$  个回合, 主体向环境采取动作  $a_k \in \mathcal{A}$ , 环境反馈给主体一个观察  $x_k \in \mathcal{X}$ , 主体感知到  $x_k$  的同时体验到效用  $u(ax_{1:k})$ , 然后进入第  $k+1$  个回合, 如此循环往复. 形成的交互历史是  $(\mathcal{A} \times \mathcal{X})^*$  的一个元素. 一个策略是一个函数  $\pi: (\mathcal{A} \times \mathcal{X})^* \rightarrow \Delta(\mathcal{A})$ , 它映射一个历史  $ax_{<t}$  到动作  $a$  上的概率分布  $\pi(a|ax_{<t})$ , 环境是函数  $\nu: (\mathcal{A} \times \mathcal{X})^* \times \mathcal{A} \rightarrow \Delta(\mathcal{X})$ , 它映射一个历史  $ax_{<t}$  和一个动作  $a_t$  到观察  $x$  上的概率分布  $\nu(x|ax_{<t}a_t)$ . 一个环境也可以视为一系列条件测度的序列, 对于任意动作序列  $a_{<t}$ , 反馈回  $\nu(x_{<t}|a_{<t}) := \prod_{i=1}^{t-1} \nu(x_i|ax_{<i}a_i)$ , 而且, 环境是时序的, 未来的动作不会影响过去的观察. 环境与主体共同构成一个交互历史上的概率分布:

$$\pi_\nu(ax_{<t}) := \prod_{i=1}^{t-1} \pi(a_i|ax_{<i})\nu(x_i|ax_{<i}a_i)$$

一般默认用  $\mu$  作为真实环境, 如果真实环境  $\mu$  已知, 那么一个理性的主体会采取某个能使其获得最大的期望效用的“最优”策略

$$\pi_\mu^* := \operatorname{argmax}_\pi V_\mu^\pi$$

其中策略  $\mu$  在环境  $\mu$  中的期望效用为,

$$V_\mu^\pi(ax_{<t}) := \frac{1}{\Gamma_t} \mathbb{E}_\mu^\pi \left[ \sum_{i=t}^\infty \gamma_i u(ax_{1:i}) \mid ax_{<t} \right]$$

其中  $u: (\mathcal{A} \times \mathcal{X})^* \rightarrow [0, 1]$  是效用函数,  $\gamma$  是贴现函数, 满足  $\sum_{t=1}^{\infty} \gamma_t < \infty$ ,  $\Gamma_t := \sum_{i=t}^{\infty} \gamma_i$  为贴现归一化因子。

假设  $\mathcal{M} := \{\nu_1, \nu_2 \dots\}$  是下半可计算的环境集。其中对每个环境  $\nu \in \mathcal{M}$ , 主体都有一个相信其为真实环境的先验信念  $w(\nu)$ , 即  $w \in \Delta(\mathcal{M})$ , 那么对所有可能环境的贝叶斯混合就是

$$\xi := \sum_{\nu \in \mathcal{M}} w(\nu) \nu$$

后验信念用  $w(\nu|ax_{<t}) := w(\nu) \frac{\nu(ax_{<t})}{\xi(ax_{<t})}$  更新。当环境未知时, Hutter ([2]) 的通用智能主体 AIXI 会采用以所罗门诺夫 (Solomonoff, [11, 12]) 先验  $w(\nu) := 2^{-K(\nu)}$  为先验信念的贝叶斯混合追逐期望效用  $V_{\xi}^{\pi}$  最大化。因为

$$\xi(x_{1:m}|a_{1:m}) \triangleq M(x_{1:m}|a_{1:m}) := \sum_{q: U(q, a_{1:m}) = x_{1:m}} 2^{-|q|}$$

所以, 可以不区分确定性环境还是不确定性环境, 而追求最大化  $V_M^{\pi}$ 。AIXI 可以看作在各种可能的环境中以算法概率追求期望效用最大化的主体。

## 2 AIXI 与纽康姆难题

纽康姆难题的原始版本中, 主体对于女巫胜率的估计是题目直接给定或基于频率的。虽然纽康姆难题涉及因果性的话题, 但对于概率本身的理论基础却不作讨论。事实上, 纽康姆难题的关键是因果性, 或主体做决策时眼中的主观因果性, 对这个因果性的刻画恰恰应该被囊括在主体对于女巫胜率的主观信念之中, 体现为某种主观概率。而在 AIXI 的框架中, AIXI 是以所罗门诺夫先验概率作为评估各种可能性的初始信念, 用贝叶斯混合或与其等价的算法概率为基础进行决策, 因果性体现为可计算函数或可计算概率分布。下面对比证据决策论和因果决策论分析 AIXI 在纽康姆难题上的表现。

根据证据决策论, 行为  $a$  带来的期望效用为

$$V(a) = \sum_{x \in \mathcal{X}} u(ax) P[x|a]$$

而根据因果决策论, 行为  $a$  带来的期望效用为

$$V(a) = \sum_{x \in \mathcal{X}} u(ax) P[a > x]$$

二者的差别仅在于, 证据决策论采用条件概率  $P[x|a]$ , 而因果决策论采用反事实条件概率  $P[a > x]$ 。在纽康姆难题中, 条件概率与反事实条件概率并不相等。但不管条件概率还是反事实条件概率, 背后起作用的都是因果性。

如果你相信女巫的说辞,那么你就会相信女巫是通过某种你不知道的策略  $\hat{\pi}$  进行预测  $\hat{a}_t := \operatorname{argmax}_{a_t} \hat{\pi}(a_t|h_{<t})$ , 然后据此决定放不放钱  $\mu(x_t|h_{<t}\hat{a}_t) \in \{0,1\}$ ,  $\hat{x}_t := \iota x_t(\mu(x_t|h_{<t}\hat{a}_t) = 1)$ 。

根据 AIXI 的框架,你的博弈环境可以看作

$$P[\hat{x}_t|h_{<t}] = \sum_{a_t} P[a_t\hat{x}_t|h_{<t}] = P[\hat{a}_t\hat{x}_t|h_{<t}] + 0 = \mu(\hat{x}_t|h_{<t}\hat{a}_t)\hat{\pi}(\hat{a}_t|h_{<t}) = \hat{\pi}(\hat{a}_t|h_{<t})$$

$$P[x_t^\pm|h_{<t}] = \hat{\pi}(a_t^\mp|h_{<t}) = 1 - \hat{\pi}(a_t^\pm|h_{<t})$$

你在博弈环境  $P$  中的期望效用为

$$V^P(h_{<t}) = \lim_{m \rightarrow \infty} \sum_{h_{t:m}} \left( \sum_{i=t}^m \gamma_i u(h_{1:i}) \right) P[h_{t:m}|h_{<t}]$$

在纽康姆博弈中,主体把每一局都当做最后一局来玩,不考虑未来的情况,所以  $\gamma_i = \llbracket i = t \rrbracket$ 。虽然不考虑未来可能继续的博弈,但计算概率时,整个的过去历史在起作用,主体所能依赖的也只有过去历史。

因为无法判断  $a$  (选择箱子) 和  $x$  (放钱) 的先后顺序,所以根据时间顺序谁先谁后有二种分解方式,一种我们称它为“证据/因果”的分解方式,另一种称它为“自由意志”的分解方式。<sup>3</sup>

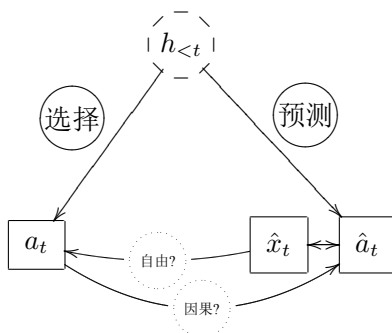
$$\begin{aligned} V^P(h_{<t}) &= \sum_{a_t x_t} u(h_{1:t}) P[a_t x_t|h_{<t}] \\ &= \sum_{a_t x_t} u(h_{1:t}) P[x_t|h_{<t} a_t] P[a_t|h_{<t}] && \text{(证据/因果)} \\ &= \sum_{a_t x_t} u(h_{1:t}) P[a_t|h_{<t} x_t] P[x_t|h_{<t}] && \text{(自由意志)} \end{aligned}$$

按“自由意志”的分解方式,

$$\begin{aligned} V^P(h_{<t} a_t^-) &= 0 \times P[a_t^-|h_{<t} x_t^-] \times P[x_t^-|h_{<t}] + B \times P[a_t^-|h_{<t} x_t^+] \times P[x_t^+|h_{<t}] \\ V^P(h_{<t} a_t^+) &= L \times P[a_t^+|h_{<t} x_t^-] \times P[x_t^-|h_{<t}] + (L+B) \times P[a_t^+|h_{<t} x_t^+] \times P[x_t^+|h_{<t}] \end{aligned}$$

自认为拥有自由意志的人 would 认为自己的行为完全独立于女巫的行为  $P[a_t|h_{<t} x_t] = \pi(a_t|h_{<t})$ , 自己的行为由自己决定  $\pi(a_t|h_{<t})$ , 所以, 根据  $P[x_t^\pm|h_{<t}] = \hat{\pi}(a_t^\mp|h_{<t})$

<sup>3</sup>根据我们的分解方式,一种情况下可以讨论女巫的行为多大程度上影响主体选择箱子的动作,这跟主体的自由意志相关,所以我们称这种分解方式为“自由意志”分解。另一种情况下可以讨论选择箱子的动作是否可能影响女巫的行为,而这种情况下又可以分别讨论证据决策者和因果决策者的不同选择,所以我们称这种分解方式为“证据/因果”分解。



有,

$$V^P(h_{<t}a_t^-) = B \times P[x_t^+|h_{<t}] = B \times \hat{\pi}(a_t^-|h_{<t})$$

$$V^P(h_{<t}a_t^+) = B \times P[x_t^+|h_{<t}] + L = B \times \hat{\pi}(a_t^-|h_{<t}) + L$$

所以  $V^P(h_{<t}a_t^+) > V^P(h_{<t}a_t^-)$ 。

也就是说, 不管女巫放没放钱, 她的行为的结果都是已经确定无法更改的了, 所以自己两个箱子都拿可以多赚  $L$ , 是一种“占优”策略。

$$\operatorname{argmax}_{a_t} V^P(h_{<t}a_t) = \operatorname{argmax}_{a_t} \pi[a_t|h_{<t}] = a_t^+$$

而相信自己不完全拥有自由意志、自己的行为部分受女巫的控制的人 would 认为  $P[a_t^+|h_{<t}x_t^+] \neq P[a_t^+|h_{<t}x_t^-]$ , 这时, 选择一个箱子还是两个箱子就取决于女巫的预测行为和自己拥有多少自由意志。比如, 假如你认为自己没有任何自由意志, 这意味着你的行为将完全被女巫决定  $P[a_t^\pm|h_{<t}x_t^\mp] = 1$ , 那么,

$$V^P(h_{<t}a_t^-) = B \times P[x_t^+|h_{<t}] = B \times \hat{\pi}(a_t^-|h_{<t})$$

$$V^P(h_{<t}a_t^+) = L \times P[x_t^-|h_{<t}] = L \times \hat{\pi}(a_t^+|h_{<t})$$

所以, 只要你相信“女巫可能会以微小的概率 ( $\geq \frac{L}{L+B}$ ) 预测你拿不透明的箱子”, 你就应该只拿不透明的箱子,

$$\operatorname{argmax}_{a_t} V^P(h_{<t}a_t) = a_t^- \iff \hat{\pi}(a_t^-|h_{<t}) \geq \frac{L}{L+B}$$

对于 AIXI 来说, 因为女巫的策略  $\hat{\pi}$  未知, 所以要用算法概率

$$M(x_t|h_{<t}a_t) = \frac{P[\{q \in \mathcal{M} : q \rightsquigarrow h_{<t} \ \& \ q(h_{<t}a_t) = x_t\}]}{P[\{q \in \mathcal{M} : q \rightsquigarrow h_{<t}\}]} = \frac{\sum_{q \rightsquigarrow h_{1:t}} P[q]}{\sum_{q \rightsquigarrow h_{<t}} P[q]}$$

替换  $P$  进行估测, 其中  $q \rightsquigarrow h$  指程序  $q$  输出历史  $h$ 。虽然放钱  $x_t$  发生在选择箱子  $a_t$  之前, 但 AIXI 在做出选择之前完全看不到箱子里是否放了钱  $x_t$ , 所以他只会根据他已有的经验  $h_{<t}$  做决策。在“类型”确定的情况下 AIXI 的博弈可看作可观察行为的完美信息博弈, 所以 AIXI 是一个完全的自由意志主义者, 他不确定的是——自己的行为是否会影响女巫的预测, 所以他根本不会按照“自由意志”的方式进行分解, 而会选择按照“证据/因果”的方式建模。因此, 对于 AIXI 来说, 首先, 是否要考虑女巫的说辞、还是只考虑她的行为, 如果考虑说辞该如何建立模型, 如果完全相信女巫, 相信放钱  $x_t$  确实发生在选择箱子  $a_t$  之前, 就会按照“自由意志”的方式分解, 又因为存在一些没有自由意志的可能世界, 所以 AIXI 会是一个不完全的自由意志主义者, 也有可能拿一个箱子。

按照“证据/因果”的分解方式, 可区分证据决策与因果决策两种决策方式,

$$\begin{aligned} V^P(h_{<t}a_t^-) &= 0 \times P[x_t^-|h_{<t}a_t^-] \times P[a_t^-|h_{<t}] + B \times P[x_t^+|h_{<t}a_t^-] \times P[a_t^-|h_{<t}] \\ V^P(h_{<t}a_t^+) &= L \times P[x_t^-|h_{<t}a_t^+] \times P[a_t^+|h_{<t}] + (L+B) \times P[x_t^+|h_{<t}a_t^+] \times P[a_t^+|h_{<t}] \end{aligned}$$

对于因果决策者来说, 他相信自己的选择行为发生在女巫的预测行为之后, 所以对女巫的预测没有影响, 所以,

$$P[x_t^+|h_{<t}a_t^-] = P[x_t^+|h_{<t}a_t^+] = P[x_t^+|h_{<t}] = \hat{\pi}(a_t^-|h_{<t})$$

所以,

$$\begin{aligned} V^P(h_{<t}a_t^-) &= B \times P[x_t^+|h_{<t}a_t^-] = B \times \hat{\pi}(a_t^-|h_{<t}) \\ V^P(h_{<t}a_t^+) &= B \times P[x_t^+|h_{<t}a_t^+] + L = B \times \hat{\pi}(a_t^-|h_{<t}) + L \end{aligned}$$

所以, 虽然对放钱  $x$  和选择箱子  $a$  的“时序”的理解与自由意志的情形完全相反, 但结果跟完全相信自由意志的情况一样, 都认为拿两个箱子是“占优”策略, 可以多赚  $L$ , 必须两个箱子都拿,

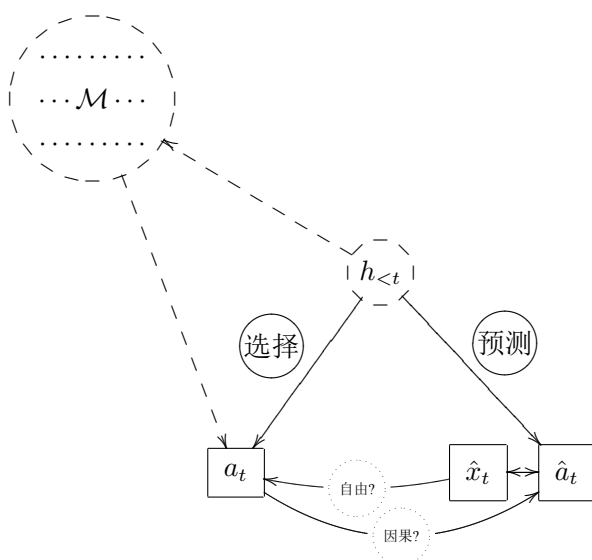
$$\operatorname{argmax}_{a_t} V^P(h_{<t}a_t) = \operatorname{argmax}_{a_t} \pi(a_t|h_{<t}) = a_t^+$$

而对于证据决策者来说, 他不确信自己的行为是否影响女巫的预测, 所以会认为,

$$P[x_t^+|h_{<t}a_t^-] \neq P[x_t^+|h_{<t}a_t^+]$$

所以, 是否只拿一个箱子取决于二者相差多少,

$$\begin{aligned} \operatorname{argmax}_{a_t} V^P(h_{<t}) = a_t^- &\iff B \times P[x_t^+|h_{<t}a_t^-] \geq B \times P[x_t^+|h_{<t}a_t^+] + L \\ &\iff P[x_t^+|h_{<t}a_t^-] - P[x_t^+|h_{<t}a_t^+] \geq \frac{L}{B} \end{aligned}$$



因此，因果决策与证据决策的区别就在于对条件概率  $P[x|ha]$  的赋予方式不同；而对于 AIXI 来说，他会用算法概率  $M$  来逼近  $P(x_t|h_{<t}a_t)$ 。这意味着 AIXI 通过综合分析女巫所有可能的（可计算的）预测策略来猜测  $P$ 。也就是说，通过虚拟所有可能世界的运行来逼近现实世界。所以，如果没有充分的证据，AIXI 不会贸然相信  $M(x_t|h_{<t}a_t^+) = M(x_t|h_{<t}a_t^-)$ 。所以，AIXI 是一个证据决策者。但是，AIXI 该如何对待女巫的说辞？如果把女巫的话作为先验信息考虑进来？由于 AIXI 的环境类/假设空间 ( $\mathcal{M}$ ) 不包含未来影响过去的情况，所以 AIXI 不相信“时空穿梭”。这意味着，AIXI 对于诸如“女巫看到我的决策然后穿越回过去再决定放不放钱”之类的假设的信念度为 0。然而，虽然女巫声称选择行为发生在她的预测和放钱行为之后，但是，看到  $x$  的结果是在自己的选择做出之后，所以，AIXI 可以怀疑存在某些假设，根据这些假设，选择行为仍然可以影响女巫的决策（比如，假设“女巫可以扫描你大脑神经元的运作”，而且假设“女巫可以扫描你大脑神经元的运作”这个假设本身是可计算的），那么，这会使得  $M(x_t^+|h_{<t}a_t^-) \gg M(x_t^+|h_{<t}a_t^+)$ 。因此，AIXI 有很大的可能会选择一个箱子。

### 3 小结与讨论

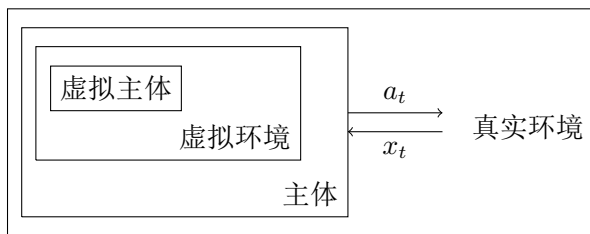
纽康姆难题是一个有哲学韵味的决策难题。面对纽康姆难题，主流的决策理论出现了分歧，采用期望效用最大化原则和占优原则分别会导向两种不同的选择。证据决策论者会拿一个箱子，因果决策论者会拿两个箱子。拿一个箱子的一般是



宿命决定论者，拿两个箱子的一般是自由意志论者。一般来说，证据决策论采用经典的条件概率，因果决策论采用反事实条件概率。本文认为，解决纽康姆难题的关键在于，如何用概率方法刻画其中隐藏的因果性。如果用经典概率论可以刻画出其中的因果性，那么发展一套专门的反事实条件概率或其他形式的条件概率的理论是不必要的。本文通过对联合概率进行的“证据/因果”和“自由意志”两种不同的分解方式，说明不同的分解方式组合，会产生不同的期望效用。占优原则可以看作一种极端的期望效用最大化原则。所以，决策理论的基础——追求效用最大化并不与占优原则冲突。问题是——如何计算期望效用。本文主要从胡特尔的通用智能模型 AIXI 出发，从 AIXI 与女巫博弈的视角分析纽康姆悖论。AIXI 采用算法概率计算期望效用，因为算法概率可以用来处理序列预测问题，而序列预测在一定程度上是对因果性的反映。而从算法概率的角度看，计算概率的关键在于，环境类/假设空间如何设定。借助通用智能模型 AIXI，我们可以清楚地看出，基于什么样的环境类/假设空间，主体会采取什么样的选择。当环境类/假设空间足够大到包含类似于“自己的行为可能会影响女巫的决策”的可能假设时，主体都会倾向于选择一个箱子。

另外，如果 AIXI 的源代码和内存可以被修改，而且它自己也知道可以被修改（即，假设空间中包含自己的源代码和内存可以被修改的环境），那么，这就类似于主体不具有完全的自由意志的情形，也会倾向于只选择一个箱子。

如果 AIXI 意识到自己是完全内嵌于真实环境之中的，自己对真实环境的理解只是一种虚拟建模，主体完全是环境的一部分，环境可以修改主体的任何部分，环境执行主体的代码，主体所能做的顶多是在模拟环境的时候能够同时模拟主体在模拟环境中的表现。这种完全内嵌于环境的主体跟内存和源代码可以被修改的主体类似，都会考虑到自己不完全具有自由意志，都可能倾向于选择一个箱子。



虽然 AIXI 的“算法概率”对于估测“女巫”的预测行为有着重要作用，但是，在 AIXI 的框架中，尽管信息是不完全的，但时序关系是严格确定的，所以“主体”与“环境”的博弈可以看作“可观察行为的贝叶斯扩展博弈”。但在纽康姆难题中，对于“女巫”的预测行为“主体”无法观察到，所以无法确定自己的“选择”行为是否会影响“女巫”的“预测”、或“女巫”的“预测”是否会影响自己的

“选择”。从 AIXI 的角度看，这就要慎重对待“女巫”给出的先验信息。因果决策者相当于把“女巫”的说辞理解为“自己的行为与女巫的行为完全独立”，所以，凡是模型类中自己的行为与女巫的行为不独立的所有假设都提前剔除了。而证据决策者只重视博弈的交互过程本身，不会把“女巫”的说辞作为先验信息从而据此提前对假设空间做那么大的删减。因此，二者的差别源于对“游戏设定”本身的理解不同。如何“合理地”对待游戏设定本身的不确定性已经超出了算法概率所能处理的不确定性的范畴，因为算法概率仅仅以“简单性”为先验，而不会考虑“游戏设定”给出的先验信息。

另外，一个有意思的问题是，假如跟你博弈的不是女巫而是 AIXI，而且你知道跟你博弈的是 AIXI，并且你知道自己是某个可计算的策略，那么，你会选一个箱子还是两个箱子？因为 AIXI 能够逼近任何可计算的策略，而且逼近的误差大小与可计算策略的柯尔莫哥洛夫复杂度相当，根据胡特尔 ([2])，

$$\sum_{k=1}^{\infty} \sum_{x_{1:k}} \mu(x_{<k}|a_{<k}) \left( \mu(x_k|ax_{<k}a_k) - \xi(x_k|ax_{<k}a_k) \right)^2 \leq K(\mu) \ln 2 < \infty$$

这意味着，AIXI 可以近似正确地预测你的选择。所以，这种情况下最好永远只拿一个箱子。

## 参考文献

- [1] W. Eckhardt, 2012, *Paradoxes in Probability Theory*, Dordrecht: Springer.
- [2] M. Hutter, 2005, *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*, Berlin, Heidelberg: Springer.
- [3] N. Ishibashi-Ohmura and T. Takahashi, 2015, “Quantum decision-making in Newcomb’s problem: Effect of reward size”, *NeuroQuantology*, **13**(3): 293–298.
- [4] R. C. Jeffrey, 1990, *The Logic of Decision*, Chicago, USA: University of Chicago Press.
- [5] J. M. Joyce, 1999, *The Foundations of Causal Decision Theory*, Cambridge, United Kingdom: Cambridge University Press.
- [6] S. Legg, 2008, *Machine Super Intelligence*, PhD dissertation, University of Lugano.
- [7] L. Ming and P. Vitányi, 2008, *An Introduction to Kolmogorov Complexity and Its Applications*, New York, USA: Springer.
- [8] R. Nozick, 1969, “Newcomb’s problem and two principles of choice”, *Essays in Honor of Carl G. Hempel*, pp. 114–146, Dordrecht: Springer.
- [9] Y. Shoham and K. Leyton-Brown, 2009, *Multiagent Systems: Algorithmic, Game-theoretic, and Logical Foundations*, Cambridge, United Kingdom: Cambridge University Press.
- [10] N. Soares and B. Fallenstein, 2015, “Toward idealized decision theory”, arXiv:1507.01986, preprint.

- [11] R. Solomonoff, 1978, "Complexity-based induction systems: Comparisons and convergence theorems", *IEEE Transactions on Information Theory*, **24(4)**: 422–432.
- [12] R. J. Solomonoff, 1964, "A formal theory of inductive inference. Part I and Part II", *Information and control*, **7(1,2)**: 1–22, 224–254.
- [13] D. H. Wolpert and G. Benford, 2013, "The lesson of Newcomb's paradox", *Synthese*, **190(9)**: 1637–1646.
- [14] E. Yudkowsky, 2010, "Timeless decision theory", <https://intelligence.org/files/TDT.pdf>.
- [15] 张建军, 逻辑悖论研究引论 (修订本), 2014 年, 北京: 人民出版社。

(责任编辑: 罗心澄)

# Newcomb's Problem under the Frame of Universal Intelligence

Xi Li

## Abstract

Newcomb's problem challenges the classical decision theory because the expected utility principle and dominance principle appear to provide conflicting recommendations for what you should choose in the Newcomb's game. Since the dominance principle can be taken as a special case of the expected utility principle, the foundation of the decision theory — the utility maximization model is safe and sound, the problem is — how to calculate the expected utility properly. The key behind Newcomb's problem is how to grasp the causation with (counterfactual) conditional probability. So the real problem behind the Newcomb's problem is the problem of induction and prediction. The famous universal induction/universal intelligence model in artificial general intelligence precisely characterises causality with algorithmic prediction. Maybe it is not necessary to develop various novel decision theories to cope with Newcomb's problem. We will explore the causality problem involved in Newcomb's problem by means of the universal intelligent model AIXI.