arXiv:2601.06405v1 [econ.TH] 10 Jan 2026

活化但遭截断的决策树中的有限理性与主观评价[1]

Peter J. Hammond： p.j.hammond@warwick.ac.uk

CRETA（经济理论与应用研究中心），华威大学经济学系，考文垂 CV4 7AL，英国。2026年1月9日，根据 ratEnlivenArxiv.tex 排版

摘要：在规范性模型中，通常假设决策者具有贝叶斯理性，因此会在一个完整且设定正确的决策模型中最大化主观期望效用。继 Hammond（2007）对熊彼特（1911、1934）企业家精神概念，以及 Shackle（1953）潜在惊奇概念的讨论之后，我们考察随时间发展、但其增长过程无法以完整细节准确建模的活化决策树。与错误设定模型、未预见的偶发事件或无意识相比，活化决策树所涉及的限制更为严苛；而这几种情形通常都借助普遍状态空间来建模，该空间足够庞大，可以涵盖行动者可能考虑的任何决策模型。我们以荷马关于奥德修斯与塞壬的经典故事为基础，探讨一个启发性例子。尽管我们的新框架超越了风险或不确定性的标准概念，但对于可能因有限理性而被截断的有限决策树，仍然可以采用一种扩展和精细化的贝叶斯理性形式：在发生截断的终端节点上附加实值主观评价，而非结果。此外，这些主观评价也构成了例如近期国际象棋软件所使用的蒙特卡洛树搜索算法的基础。它们也可能有助于为颇具争议的预防原则提供理性依据。 [199 words]

关键词：有界贝叶斯理性、结果主义决策理论、熊彼特式企业家精神、Shackle 的潜在惊奇、截断决策树、活化决策树、延续子树的主观评价、蒙特卡洛树搜索、预防原则。

JEL分类：D11、D63、D81、D91。

[1]本文是 CRETA工作论文第89号 的大幅修订版，原题为“活化决策树中的主观评价与贝叶斯理性”，信息量略显不足。精简版预计将发表于《理论与决策》。

序言

> 理论总是灰色的，亲爱的朋友；而生命的金树永远长青。[2]——歌德《浮士德》第一部中的梅菲斯特。[3]

> .....他说，要完成 [这首] 诗，他离不开那栋房子，因为地窖里有一个阿列夫。他解释说，阿列夫是空间中的一个点，包含了所有其他点。
>
> 阿列夫的直径大概不过一英寸，但整个空间都在那里，真实而完整。每一件事物（比如镜子的表面）都是无限多个事物，因为我清清楚楚地从宇宙的每个角度看到了它。我从每一个位置和角度看着阿列夫；在阿列夫中，我看见了地球，在地球中看见了阿列夫，在阿列夫中又看见了地球；我看见了自己的脸和自己的内脏；我看见了你的脸；我感到头晕目眩，哭了起来，因为我的眼睛看见了那个隐秘而臆想的物体，它的名字为众人所共知，却无人得见——那不可想象的宇宙。我感到无限的惊奇，无限的怜悯。.....因为康托尔的集合论，[4] [阿列夫，也就是 ℵ] 超限数的符号，其中任何一部分都与整体一样大。
>
> 走在街上，沿着宪法站内的楼梯往下走，乘坐地铁时，每一张脸对我来说都似曾相识。我害怕世上再也没有任何事能让我感到惊讶；我害怕自己再也无法摆脱所见的一切。幸好，熬过几个不眠之夜后，遗忘再次降临于我。
>
> ——摘自豪尔赫·路易斯·博尔赫斯《阿列夫》（1945年），由 NormanThomas di Giovanni 与作者合作翻译。

[2]一种可能的译法是：“亲爱的朋友，理论总是灰色的；而生命的金树常青。”[3]这篇论文的主题构成了我于2007年初退休前，在斯坦福大学举办的最后一次研讨会的内容。一两天前，肯尼斯·阿罗给我留了电话留言，问我是否受到歌德这句引文的启发。尽管我只能回答说没有，但我仍觉得，这本该是我灵感的源泉。[4]“Mengenlehre”在德语中是“集合论”。

> 沙克尔选择理论的核心思想是：人们考虑采取某项决策可能带来的结果时，只会关注他们（a）能够想象到，并且（b）在某种程度上认为有可能发生的那些结果。这些结果未必涵盖实际发生的情况；实际发生的可能是他们甚至从未想象过的事件，因此会令他们完全意外。
>
> —— Earl 和 Littleboy（2014年，第162页）。

1 背景与提纲

1.1 为贝叶斯理性辩护

在决策理论中，贝叶斯理性是一种假设，认为行动者所做的选择，其结果（通常是兼具风险和不确定结果的彩票）能够使定义在指定的非空后果域 Y 上的伯努利效用函数 $Y \ni y \mapsto u(y) \in \mathbb{R}$ 的期望值最大化。对于源自 Anscombe 和 Aumann（1963）所称“轮盘彩票”的风险结果，根据定义，非空有限样本空间 Ω 中的每个彩票结果 ω 都有一个“客观”或假设的概率 $\pi(\omega) \in [0,1]$ 。对于源自 Anscombe 和 Aumann（1963）所称“马彩票”的不确定结果，贝叶斯理性要求非空有限状态空间 S 中的每个彩票结果或状态 s 都有一个“主观”或个人概率 $p(s) \in [0,1]$ 。一般的“Anscombe–Aumann”彩票会针对具有主观概率的马彩票的每个可能结果 $s \in S$ ，指定一个适当的轮盘彩票；该轮盘彩票包含定义在非空后果域 Y 上的结果 y ，且具有客观概率 $\lambda_s(y)$。因此，贝叶斯理性规定，恰当的期望效用最大化目标应为双重求和 $\sum_{s\in S} p(s) \sum_{y\in Y} \lambda_s(y) u(y)$，其中涉及主观概率与客观概率的乘积。

以 Hammond（1988a, b；1998a, b；1999）提出的“结果主义”假设为基础，既有研究为决策树中的贝叶斯理性行为提供了规范性论证。该假设要求，任何有限决策树（包括任何延续决策树）中，依照既定行为所产生的所有可能的安斯康姆–奥曼后果彩票，

都能解释为定义在相关定义域上的某个适当选择函数的取值；该定义域由后果彩票的非空有限可行集构成。

Hammond (2022) 对贝叶斯理性的最新论证，以相关的“前理性”概念取代了后果论。该概念适用于一组针对不同事件 E的条件弱基础偏好关系 $\succsim_E$。事实上，前理性偏好要求存在一条满足后果论假设的行为规则，并且与之相关的选择规则从每一对彩票中选出一个弱偏好的结果彩票。

无论采用后果论还是前理性来论证贝叶斯理性，都需要额外满足一个众所周知的连续性公理。该公理适用于每个“Marschak (1950) 三角形”上的偏好。对于任意给定的三个两两不无差异的轮盘彩票 $\{\lambda, \mu, \nu\}$，该三角形定义为三种彩票的所有概率混合 $q_\lambda\lambda + q_\mu\mu + q_\nu\mu$构成的集合 $\Delta(\{\lambda, \mu, \nu\})$，其中三个概率权重 q_λ，q_μ，q_ν 均为非负数，且 $q_\lambda + q_\mu + q_\nu = 1$。经过适当重新标记，可以假设严格偏好关系 $\succ$ 满足 $\lambda \succ \mu$、$\lambda \succ \nu$和 $\mu \succ \nu$。那么，给定相应的弱偏好关系 $\succsim$，连续性公理要求两个集合

$$\begin{aligned} & \{\alpha \in [0,1] \mid \alpha\lambda + (1-\alpha)\nu \succsim \mu\} \\ \text{and} \quad & \{\alpha \in [0,1] \mid \mu \succsim \alpha\lambda + (1-\alpha)\nu\} \end{aligned}$$

相对于单位区间 $[0,1] \subset \mathbb{R}$ 应当是闭集。放弃连续性将允许在彩票上存在某种非贝叶斯理性的字典序偏好关系。

一些早期论证贝叶斯理性的论文也援引了动态一致性假设。该假设要求，树 T 中后续决策节点上的预期或计划行为 n 应与实际行为相符。然而，在决策者于由剔除树 T 中不弱优于 n的所有节点而形成的延续子树 $T_{\geq n}$ 的初始节点 n作出选择之前，树 T 的任何决策节点 n 上的实际行为都尚未确定。因此，将 n视为 $T_{\geq n}$的初始节点后，无须参照此前关于如何在 n 行动的任何意图或计划，就能明确 n 上的实际行为。因此，只考虑 T中每个决策节点 n 上的实际行为，就完全排除了任何可能的动态不一致性：一方面是节点 n 上的实际行为，另一方面是此前关于

在节点 n应选择什么的计划或意图。由此，动态一致性便是构造性地得到满足的。

1.2 有限理性？还是有限建模？

> “所有模型都是错的，但有些模型有用。”——
> George Box（1919–2013）

人类的聪明才智促使我们中的一些人创造谜题和其他决策问题，以供彼此消遣或启发。许多孩子和一些成年人都能从拼图游戏中获得乐趣，或通过学习如何在井字棋（也称为 tic-tac-toe）中不输来获得满足感。还有人尝试填字游戏、数独或魔方。几代学生在数学课程中学习，课程要求他们通过解答难度逐渐提高的习题，或理解这些习题的解，来掌握知识。在这些例子中，挑战都是为一个定义明确的决策问题找到完美的解。

然而，典型的决策问题并不像谜题或数学练习。事实上，它们往往极具挑战性，以至于不可能实现完全的贝叶斯理性。当然，正是这一认识促使 Simon 在1955年和1957年提出有限理性与满意化的概念。然而，满意化似乎很难得到合理解释，除非将其视为一种妥协的结果：人们权衡更深入地搜索贝叶斯最优决策所带来的收益与额外搜索成本后，便会作出这种妥协。因此，满意化似乎更适用于选择要分析的决策模型，而非在给定的待分析模型中选择要作出的决策。基于这一原因，有限模型似乎是一个更令人满意的基础概念。[5]而对于决策树而言，一种显而易见的有限模型形式是截断树，它通过剪去一个或多个完整的延续子树而形成。

因此，为了使贝叶斯理性对行为至多只有有限理性的行动者仍具有相关性，就需要调整贝叶斯理性的期望效用最大化模型，使其也适用于截断决策树。第7节将讨论这一主题。

[5]我不应声称这一想法是原创的。例如，请参阅 Earl 和 Littleboy（2014年）第8.2节中关于 Simon、Shackle 和国际象棋棋局的讨论。

1.3 有界模型中的决策程序

然而，截断并不是贝叶斯理性在决策树中的实际相关性受到严重限制的唯一原因。事实上，截断还助长了另一个原因。毕竟，当局限于一棵固定的决策树时，贝叶斯理性会为行动者规定一个由同样出色的最优决策策略构成的无差异类。就像标准形式博弈中的玩家策略一样，在树的每个决策节点上，最优策略都会指定一个由同样出色、可在该节点采取的着法构成的无差异类。此外，贝叶斯理性规定，行动者应当遵循先前策略选择的延续方案；如果不这样做，就应选择一个期望收益相同的替代方案。

事实上，被迫使用有界决策模型的行动者最终面对的并非单一决策策略的选择，而是我们所称的决策程序的选择。该程序始于行动者面临一个初始决策节点，此时会选择一个以截断决策树形式呈现的初始有界模型进行分析。随后，行动者会在初始决策节点上找出并实际作出一项决策；鉴于初始有界模型，这项决策符合有界贝叶斯理性。

这一初始决策带来了这样的可能性：之后可能发生未曾预见的事件，从而暴露初始模型的某些局限。正如第3节所述的奥德修斯与塞壬的例子所表明的，如果行动者的有界模型排除了重要的可能性，那么贝叶斯理性可能会导致行动者作出糟糕的决策，并造成灾难性后果。

或者，在原决策树 T的任何后续决策节点 n处，如果所有实际发生的事件都已被预见，它们就会将可能性的范围缩小到一棵延续决策树 $T_{\geq n}$；这是到达节点 n后仍与决策相关的树 T 部分。这种范围缩小通常会释放资源，使正在考虑在 n作出何种决策的行动者有机会分析一棵比 $T_{\geq n}$更丰富的修订后的有界决策树；而 $T_{\geq n}$是行动者在节点 n作出决策时所分析的相关延续子树。当然，这意味着行动者隐含地承认，由于资源有限，自己未能充分建模先前那项决策的后果。

无论哪种情况，行动者都可以，而且很可能应该在做出下一项决策之前使用丰富的决策模型。构建相关延续子树的丰富化模型，然后在其中最早的决策节点上做出贝叶斯决策，这一过程就是行动者决策程序的第二步

。当然，除非决策树已经到达末端，否则决策程序将继续进行第三项决策、第四项决策，依此类推。最终将形成一个多步骤决策程序，在以下两者之间循环：

1. 每作出一次决策后，便根据该决策将先前的决策树缩减为其延续子树；2. 丰富该延续子树，从而得到一个新的决策模型；3. 在这个丰富后的决策树的初始决策节点处作出贝叶斯理性决策。

当然，对于足够简化的模型，这三个步骤中的第一步和第三步完全是例行操作；本文讨论的主题是第二步。

1.4 改进有界决策模型

短视地遵循第1.3节所述的决策程序，可能会造成不必要的劣势。这是因为，即使行动者在任何决策节点使用的有界决策模型，也可能包含一个更优的备选决策，因为该决策会努力预判那些无法完全预测的变化。毕竟，如果由于计算方面或其他实际建模限制而无法避免截断，那么它必然会排除一些可能性，而这些可能性原则上本应纳入行动者原先的决策树模型。

Lal（1972年）的研究提出了一个经济学案例。有一个行动者或机构正在考虑一项地方公共项目，例如加深一口村井。为了判断这是否值得，机构可以按照 Lal 的做法，采用著名的 Little–Mirrlees 成本效益分析方法。[6] 这不仅要考虑加深水井的钻探成本，以及为了从更深处抽水而可能需要安装功率更大的水泵，还必须纳入供水量增加后，对村庄内水、农产品和劳动力等相关商品的当地价格可能产生的一般均衡效应。

[6]参见 Little and Mirrlees（1969、1974年、1991年）。

然而，完成初步成本效益分析后，可能会发现，从加深后的水井抽取更多的水，或许会降低的不仅是这个村庄的地下水位，也可能包括邻近村庄的地下水位。这取决于为加深后的水井供水的隐蔽含水层的范围。接下来要考虑的另一项决策，可能是是否开展地质地下水调查，以估算邻近村庄地下水被抽走的数量，以及所有受影响村庄会受到怎样的一般均衡效应。[7] 最终，起初看似简单的决策问题，很可能会因为一些起初未曾预料到的因素而变得复杂得多。

第二个例子与国际象棋有关。假设存在一位拥有无限理性的国际象棋棋手，或一种国际象棋算法。根据定义，这位棋手能够始终进行评估任何决策树所需的全部计算，并得出完美结果，因此总能下出完美棋局。然而在实践中，任何人类棋手，甚至最强大的国际象棋计算机算法，都只能完整分析所有可能棋局后续变化中的一小部分。因此，任何可行的分析通常都必须在棋局结果尚未确定之前就截断。从这个意义上说，玩家几乎总是必须采用有界决策模型。

即便如此，经过初步分析后，高水平棋手通常不仅能够判断哪些可能的后续变化最有可能出现，还能判断哪些变化至关重要，也就是最有可能改变棋局结果的变化。随后，玩家可以在选择下一着之前，决定在丰富化的棋局模型中进一步分析哪些可能出现和／或至关重要的后续变化。

这类考量表明，在树中做出第一个决策之前，理性主体就可能有充分理由改进用于分析该决策的任何有界模型。更显而易见的是，在之后的任何决策节点上，改进此前用于分析各决策节点决策的决策模型，都可能带来益处。

[7]另一个决策是，是否要进行初步的成本效益分析，正如 Little 和 Mirrlees（1991年）在附录中所讨论的那样。

1.5 生动但截断的决策树

因此，为了允许决策树发生变化——甚至是不可预测的变化——本文提出了一种包含“活化”决策树的框架。一个完全短视的行动者遵循“桥到跟前再过”的老话——Savage（1963年，第16页）尤其讨论过这一点——会表现得仿佛这种活化完全无关紧要。这会导致行动者从一个模型贸然跳到另一个模型，始终表现出傲慢自大。

当然，许多未来决策及其不确定后果都无法被详细建模。尽管如此，即便只有一点点阅历的行动者也应该明白，对于任何一项决策，重要的是当前对其终极 事后 价值的预期——从事后回顾时看，这一价值究竟会是多少。沿着 Koopmans（1964年）和 Kreps（1990年、1992年）为讨论灵活性偏好而提出的思路，行动者应尽可能可靠地确定这些预期估值，利用任何被认为相关的有限证据，同时也考虑行动者能够拨出多少有限资源用于决策分析。另请参见 Dekel等 （2001年、2005年）及许多后续研究，它们在应对未预见偶发事件的决策背景下探讨了相关思路；尽管这些事件的可能性，显然已被一位仿佛无所不知且自负的决策分析者预见到。国际象棋的例子在本文更完整的工作论文版本中有更全面的讨论，它说明了假定存在这样一位无所不知的决策分析者是多么不合理。[8]

我们强调，本文与此前关于未预见的偶发事件或无意识的研究存在差异，因为本文不依赖于Karni和Vierø（2017年，第304页）所定义的那类“扩展的可设想状态空间”的存在。相反，我们允许相关状态空间因我们称为“活化”的动态过程而完全不可预测地扩展。具体而言，尽管决策主体可能意识到自身无意识的可能性，但他们甚至无法构建一个实用模型；按照惯例，该模型应建立在涵盖一切可设想可能

[8]另请参阅有关决策和博弈中的无意识的相关研究，例如 Schipper（2014a、b）、Halpern和Rˆego（2014）、Grant 等人 （2015a、b）的研究；尤其是 Vierø（2009年、2021年）以及 Karni和Vierø（2013年、2015年、2017年）关于意识增长和逆贝叶斯主义的相关研究。其他有关无意识的重要研究还包括 Schipper（2014a）编辑的《数学社会科学》特刊中发表的论文，以及 Vierø（2021年）引用的研究。

状态的完备空间之上。这种对未预见的偶发事件或无意识的既有概念的扩展，最初由Hammond（2007年）以非正式方式提出，部分受到熊彼特（1911年、1934年）企业家精神概念以及Shackle（1953年）潜在惊奇概念的启发。[9] 活化决策树的概念，部分可以通过第3节讨论的奥德修斯与塞壬的经典例子来阐释。

尽管如此，假定存在一棵完全活化、且不再需要进一步活化的决策树，会引发与假定存在普遍扩展可设想状态空间相同类型的概念问题。为避免这一问题，我们承认递归活化决策树具有相关性，其生长过程永远无法在单一的普遍模型中得到完整描述。尽管如此，我们仍主张，即使是完全活化的树，也可以且应当化简为一棵具有随机结果的树；这些随机结果不是后果彩票，而是附加于截断决策树终端节点上的主观评价。

1.6 论文概要

第1.1节简要回顾的早期研究表明，在补充连续性条件后，结果论或前理性为在任何给定决策树中的贝叶斯理性行为提供了规范性依据 T。本文的任务是说明如何修改贝叶斯理性公设，使之成为一种决策程序，允许行动者适应一棵既因活化而不断演变、又不可避免地受到截断的决策树；这种截断是避免面对过于复杂、无法进行全面分析的问题所必需的。

第 2 节开启论文的后续部分，简要回顾无界理性与有限理性之间的一些区别。其中包括一些著名的例子，例如 Simon 的程序理性概念，以及 Manzini 和 Mariotti（2007年）提出的“理性候选名单”方法。

接下来，第3节将重新探讨荷马史诗中著名的奥德修斯与塞壬的故事。以 Strotz（1956）、Pollak（1968）、Ham- mond（1976）和 Elster（1979）为代表的既有研究，通常将其视为偏好变化的典型例子，借此说明天真选择与精明选择之间的区别，以及承诺装置可能具有的价值。与此相反，本文将这个荷马史诗中的例子视为一棵神话决策树

[9] 关于这两个概念之间联系的最新讨论，参见 Metcalfe 等（2021年）。

女巫喀耳刻（或 Circe）将这棵树赋予了生命：奥德修斯在听取她提出的明智建议时，历经了多个阶段。[10]

接下来的第 4 节和第 5 节聚焦于另外两个更具体的例子。第一个例子涉及一位消费者，他作为投资者，为了在满足线性预算约束的条件下最大化两期二次效用函数，而选择一个金融资产组合。使这位消费者的决策问题更鲜活，可能只会影响参数值，但也可能使新商品变得相关；这些商品甚至可能尚未被发明出来。

第 5 节讨论这两个具体例子中的第二个，即国际象棋，无论是由计算机还是人类来下。当然，作为一种双人博弈，国际象棋表面上超出了本文主要讨论的单人决策问题范畴。尽管如此，分析一盘过去下过或正在进行的国际象棋棋局的人，仍面临着设法找出双方最优策略的任务。[11] 无论如何，国际象棋相当鲜明地体现了决策者的有界模型的重要性，无论该模型用于单人决策问题还是双人博弈。

第 5 节简要解释了蒙特卡洛树模拟如何使计算机算法能够评估这样的局面：在这些局面中，所有可能的未来着法都已在可能范围内经过详细分析。本节还简要介绍了一个特别不幸的人类棋步案例，该着法当时被称为“世纪大失误”。这被视为一位特别杰出的玩家未能修正其对棋局可能如何发展的有界模型。

第6节概述了我们需要的关键概念，用以描述在既未截断也未赋予生机的经典有限决策树中的贝叶斯理性行为。后续分析中使用的一项关键工具，是对任意决策树 T 的评价$v(T)$。其定义为：在 T 的每个决策节点上都进行最优决策后，所能产生的任一结果彩票的标准化期望效用。依照动态规划中的最优性原理，可以通过反向递归计算这一评价，从 T 的终端节点开始；每个终端节点都对应一个已指定的结果，通常以彩票的形式呈现。

[10]在荷马希腊语中，她的名字拼写为 $K\acute{\iota}\rho\kappa\eta$。[11]第 2.1 节的脚注 12 描述了可以进行完美分析的棋局。

正如第1.4节中简要讨论的，任何行动者若要考虑一个其后果难以全面分析的决策，就必须求助于有界模型。第7节介绍了一种特殊但尤为重要的有界决策模型，其中许多延续子树会被截断。由此得到一棵“截断”决策树，其结果起初仅为截断操作未移除的树的终端节点所指定。于是，对于截断决策树中那些最终可能接续“截断”节点、而非原完整树终端节点的节点，第6节所用的反向递归评价方法便会失效。第7.5节提出通过填补这一缺口来解决问题：为每个截断节点赋予一项“主观评价”，即一个实值估计，用以表示如果能够对从该节点开始的缺失延续子树进行完整分析，其评价本应是多少。

第8节最终转向更正式地描述决策树如何实现活化。我们考察一个被分解为一系列步骤的过程。其中每一步都是该过程中前一棵树的一种特殊的“最小”活化，要么发生在一个指定的终端节点上，要么沿着一条指定的边发生。整个过程从给定的“基础”决策树 T 开始，最终得到一棵活化决策树 T^{+}。

接下来，第9节说明如何将贝叶斯理性决策程序的概念扩展到有限理性主体可能采用的那类活化但截断的决策树。此项扩展的一个关键部分指出，正如第3节讨论的奥德修斯与塞壬的例子所示，活化可能涉及扩展结果域和世界状态的定义域，因而也会扩展活化后的效用函数和主观概率分布的定义域。第9节最后将我们的贝叶斯理性模型中效用与主观概率的任意性（该模型包含活化的主观评价），与 Anscombe and Aumann（1963）主观概率模型中相应概念的任意性进行比较。

倒数第二节第10节从第10.1节开始，讨论预防原则及其与期权价值的关系。接下来，第10.2节简要讨论Ullmann-Margalit（2006）、Paul（2014，2015a、b、c）以及其他哲学家近期的研究，他们提出了“变革性体验”这一概念。第10.3节简要讨论如何将本文的结果扩展到 Hammond and Troccoli Moretti（2025）中的框架；在该框架下，决策树可以包含非终端的定时后果节点。

其中包括“菜单后果”节点，其结果取决于延续子树，而该子树的初始节点就是该结果节点。最后，第10.4节讨论Karni和Vierø（2013，2015，2017）关于“逆贝叶斯主义”概念的近期研究。

第11节的简短总结着重描述一种有界贝叶斯理性决策程序，该程序可应用于生动化但截断的决策树。

2 超越无界理性

> 在一粒沙中看见一个世界，在一朵野花中看见一
> 片天堂，将无穷握于掌中，将永恒浓缩于一小时。——
> 摘自威廉·布莱克的《天真预兆》

2.1 无界理性与程序理性

Simon（1955年、1957年）提出的著名概念“有限理性”，或许最好通过其反面来定义。无限理性的决策主体会基于所有可能选择的行为及其所有潜在结果的完美模型，做出完美决策。其结果可能是那种颇令人不安的假想完整模型，豪尔赫·路易斯·博尔赫斯在短篇小说《阿列夫》中对它作了精妙的描绘；本文序言中引用了其中一些相关摘录。

在完美模型中，无论结果是确定的（明确的），还是用 Anscombe和Aumann（1963）的术语来说：(i) 有风险的，其假设的“客观”概率如轮盘彩票中的概率；(ii) 不确定的，其个人或“主观”概率如马彩票中的概率，无界理性的定义始终不变。这种无界理性会使国际象棋或围棋这类复杂而引人入胜的棋局，变得不比儿童玩的井字棋有趣；井字棋也称为“tic-tac-toe”。[12] 而且，不会存在

[12]请注意，在国际象棋中，目前可从 http: //tb7.chessok.com/ 在线获取的“罗蒙诺索夫残局库”，描述了从任何合法局面开始的完美对弈，前提是棋盘上双方剩余棋子的总数不超过七枚，其中包括双方的王。当然，通常的国际象棋棋局开始时，双方各有

所谓“意外后果法则”；每一种可能的结果都应当经过计算，从某种意义上说，这就使它成了有意为之，即便它可能只是某个风险决策带来的不幸结果。

除了有限理性之外，Simon还提出了一个重要的相关概念：“程序理性”。这一概念承认，决策程序本身可能是理性的，即使它们导致的决策从无界理性的角度来看并不理性。Simon强调了志向水平等概念，以及满意化。后者似乎是指找到一个达到志向水平的决策，也就是做出一个看起来足够好的决策，而不是最优决策。然而，最优（甚至只是灵活的）搜索意味着，如果很快且轻易地达到志向水平，那么这一水平就要求过低，因此应该提高。

相比之下，本文提出的规范性框架认为，满意化行为不应发生在既定决策模型之内，而应体现在选择模型中纳入多少细节、排除多少细节。最终，行为应相对于所选的任何有界模型达到最优。

2.2 理性候选名单与其他有界模型

一种有界模型涉及Manzini和Mariotti（2007年）提出的“理性候选名单”，他们用这一概念讨论了“序贯可理性化”，针对的是行为只有确定性结果的决策问题。他们的观点是，面对一个大得不便处理的可行选项集，行动者可以在初始阶段先列出一个相对较小的子集，以便之后认真考虑。此外，这份候选清单应足够小，使得从中找出最优决策成为一个可处理的决策问题。因此，任何候选清单都可以被视为可行集的有界模型。此外，当人们认识到观察和/或计算可能需要付出成本时，受Sims（2003年、2011年）以及Hansen和Sargent（2007年）启发的“理性忽视”研究，会探讨何种有界决策模型可能是最优的。

候选清单的选择可以被认为是通过某种有限理性搜索程序，以相当任意、甚至随机的方式产生的。当然，有些选项进入候选清单的可能性可能远高于其他选项。此外，如果将候选清单的构成视为随机的，那么不同的随机

两名玩家各自在棋盘上有16个棋子。

变量（用于指示每个选项是否属于候选清单）很可能彼此相关。

然而，一旦在第一阶段确定了候选清单，就完全可以合理地假设，行动者确实会在随后的第二阶段从候选清单中选择一个最优选项。也就是说，从候选清单中进行选择，符合 Simon（1955年、1957年）所说的“实质理性”。

候选清单可以看作程序理性的一种特定形式，它包含一个两阶段程序。然而，这里要指出的重点是，无论候选清单具体如何，它都代表了对完整决策问题的有限 模型。事实上，计算机无法完美下国际象棋等局限，适用于所有困难的决策问题，包括生活中出现的大多数问题，而非经济学家和其他决策科学家通常分析并应用的那些过度简化的模型。因此，我们用来为决策制定提供依据的任何模型，都应足够灵活，以便在可能发生变化时灵活调整；否则，任何实际模型都不得不忽略这些变化。

假设我们确实尽可能认真地看待 Simon 的“程序理性”理念。具体而言，探讨以下假设对决策主体的影响，想必颇有意义：

1. 他们的时间、注意力和计算资源都极其有限，因此除了简化模型之外，别无选择；2. 而且事实上，他们会将自己局限于足够简单的有界模型，从而确实能够找到在其有界模型范围内最优的决策。

然而，一旦认识到行动者用于决策的模型是有界的，就必须同时认识到，事件最终可能迫使人们考虑一个扩展或“活化”的模型，其中包含未建模变化。

3 再论奥德修斯与塞壬

3.1 一位天真水手的模型

作为活化决策树的第一个“经典”例子，我们重新审视荷马史诗神话中奥德修斯与塞壬的故事。根据这则史诗神话，

na¨ıve 水手的最短航线若经过塞壬岛附近，他们或许曾使用过一种有界模型来表示自己的决策树，如图1所示。[13] 具体而言，这些 na¨ıve 水手的行动仿佛表明，他们认为自己面临的选择是：

- 要么 靠近 塞壬岛并抵达目的地较早，走直达路线；
- 要么 避开 塞壬岛，较晚抵达，绕行路线。

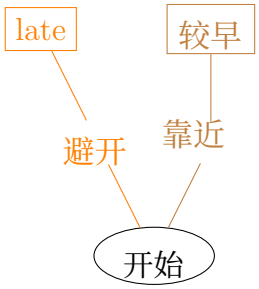

图1：Na¨ıve Sailor

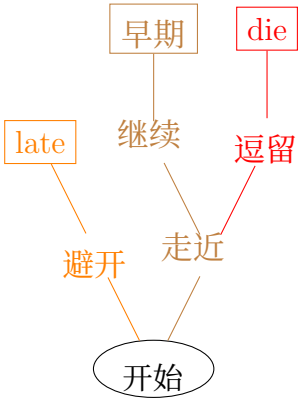

图2：老练的水手

3.2 老练水手的模型

然而，据荷马所述，奥德修斯有女巫喀耳刻担任顾问，她拥有超凡的见识。她警告奥德修斯，塞壬的歌声能诱使毫无戒备的水手走向死亡，而且塞壬岛上的草地散落着水手的骸骨。因此，如果有任何¨ıve 水手选择靠近图1所示决策树中的路径，进入塞壬的听力范围，他们就会发现自己面对的其实是标记为靠近的决策节点，该节点出现在增活后扩展的决策

[13]请注意，在图1–4中，每个方框都表示一个终端节点，其结果由框内的词语描述。除这些终端节点外，树中的所有其他节点都是决策节点。除标记为“开始”的初始决策节点外，其他每个决策节点都用一个词或短语标记，极简地描述导致该节点的最后一个行动。

树（如图2所示）中。在增活树的这个节点上，他们表面上面临的选择是：

- 要么 继续听到塞壬的歌声后回家，
- 要么 逗留，被她们的歌声迷住，最终死去 在她们的岛上，永远无法抵达原本的目的地。

当然，这一新增特征在于，听到塞壬的歌声后，此前从未有水手拥有足够的意志力逃离这座岛屿。这正是 Hammond（1976）所称的偏好变化“潜在成瘾者”例子的本质特征。面对图2所示的决策树，一位理解塞壬歌声说服力的老练水手会避开她们的岛屿，待在听不到歌声的地方，即使这意味着要在大幅延迟之后才能抵达原本的目的地。

3.3 喀耳刻关于奥德修斯的第一个激活模型

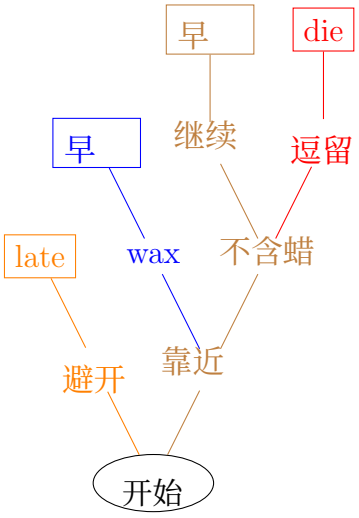

图3：Kirke 的第一模型

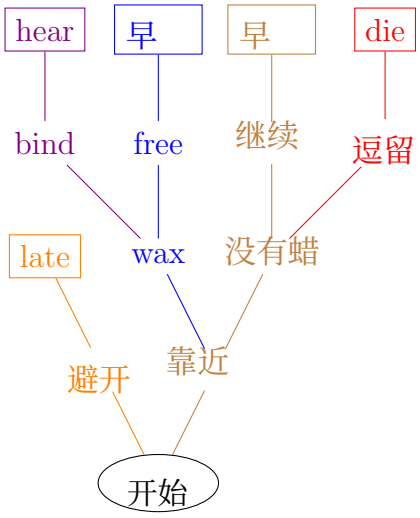

图4：喀耳刻的最终模型

不过，喀耳刻的建议并不只是警告。关于避开塞壬岛、绕远路返回伊萨卡、只是稍晚抵达的平淡故事，算不上令人难忘的史诗。相反，喀耳刻指出，只要采取

用蜡堵住所有水手耳朵这一预防措施，就能安全地驶过塞壬岛。因此，在决定驶向塞壬岛之后、尚未进入可听见歌声的范围之前，图3中的节点 靠近 处的选择将是：

- 要么 用蜡 塞住全体船员的耳朵（包括奥德修斯自己的耳朵），这样他们就都听不到塞壬的歌声；
- 要么不用 蜡，就像所有更早的那些天真的水手一样，他们的骸骨如今散落在塞壬草地上。

3.4 喀耳刻为奥德修斯构建的最终鲜活模型

然而，荷马为我们讲述的史诗版本要有趣得多。荷马笔下的喀耳刻为奥德修斯提出了更好的行动方案，使奥德修斯至少能够听到塞壬的歌声，同时又能保住性命。事实上，喀耳刻建议，除了给全体船员的耳朵涂上蜡之外，还应把他本人牢牢绑在桅杆上。此外，还应严格嘱咐船员：如果他们看到奥德修斯请求放开他，不仅要置之不理，还要把绳子绑得更紧。因此，喀耳刻为奥德修斯构建的最终模型中还多了一个标记为 蜡 的决策节点，如图4所示，其选项为：

- 要么将奥德修斯绑在桅杆上，并让他的耳朵不封蜡，以便他能够听到塞壬的歌声，
- 要么让奥德修斯自由，但像其他船员一样用蜡封住耳朵。

我们强调，这一额外决策可以在驶向塞壬岛之后作出，但必须在接近到足以听见她们歌声之前作出。

3.5 迈向充满活力的决策树

更早之前那些天真的水手，尸骨散落在塞壬草地上，他们的模型如图1所示。听过塞壬的歌声并得知她们存在之后，他们或许才意识到，更恰当的模型本应如图2所示，但为时已晚。奥德修斯（及其船员）则

足够幸运，得到了一个实用得多的模型，一开始甚至超越图2，直接达到图3，最终则达到图4。

图1–4中的每棵决策树单独来看都毫无生气。然而，四棵树合在一起，却讲述了一个关于学习的宏大故事。但这并非通常所说的在固定样本空间内不断加深对世界状态了解的统计模型。相反，随着丰富化模型纳入越来越多的可能性，可能性集合也在不断扩大。我引入“增活树”这一术语时，忍不住将其与一棵鲜活生长的树作类比；我也借此暗示它与熊彼特（ 1911，1934）关于创新的著作，以及 Shackle（1953）关于“潜在惊奇”的论述有着密切的相似之处——进一步讨论请参见 Hammond（2007）。

4 线性—二次型投资组合问题

4.1 两期投资组合问题

我们的两个扩展示例中的第一个，涉及一位消费者，其两期伯努利效用函数取二次型

$$u(\mathbf{x}_1, \mathbf{x}_2) = -\tfrac{1}{2}(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 (\mathbf{x}_1 - \mathbf{a}_1) - \tfrac{1}{2}(\mathbf{x}_2 - \mathbf{a}_2)^\top \mathbf{Q}_2 (\mathbf{x}_2 - \mathbf{a}_2) \quad (1)$$

这里$\mathbf{x}_1$ 和$\mathbf{x}_2$ 表示两期的有限维消费向量，两者的维度可能不同；而$\mathbf{a}_1$ 和$\mathbf{a}_2$则是相应的参数向量。此外，假设$\mathbf{Q}_1$, $\mathbf{Q}_2$ 是维度适当的对称正定方阵。

假设消费者每期都面临一个预算约束，共有两个预算约束，可表示为

$$\mathbf{p}_1^\top \mathbf{x}_1 + \mathbf{q}^\top \mathbf{b} = m_1 \quad \text{and} \quad \mathbf{p}_2^\top \mathbf{x}_2 = m_2 + \mathbf{r}^\top \mathbf{b} \quad (2)$$

其中$\mathbf{b}$ 表示第1期期末净资产持有量构成的有限维投资组合向量，$\mathbf{q}$ 为第1期的资产价格向量，而$\mathbf{r}$ 为总回报向量。当然，$\mathbf{p}_1$ 和$\mathbf{p}_2$ 分别表示各期的商品价格向量，且均假定严格为正； $m_1, m_2 \in \mathbb{R}$ 则为外生财富转移。我们允许$\mathbf{a}_2$, $\mathbf{r}$ 和 m_2 都具有不确定性，但将$\mathbf{p}_2$ 视为确定的，正如希克斯（1946年）在其暂时均衡理论中使用未来价格的点预期时所做的那样。[14] 为简单起见，我们

[14]关于一个有些类似的观点，参见迈尔森（1983年）。

还假设对称矩阵$\mathbf{Q}_2$ 在第1期已知。最后，我们假设随机总收益向量$\mathbf{r}$与随机变量$\mathbf{a}_2$ 和 m_2均随机独立。

4.2 第二期最优解

假设到第2期开始时，参数向量 $\mathbf{a}_2$, 、总回报向量 $\mathbf{r}$和非劳动收入 m_2 都已确定，投资组合向量 $\mathbf{b}$ 则是由消费者在第1期自行选择而预先确定的。因此，消费者的第二期优化问题与其在第1期选择何种 $\mathbf{x}_1$ 无关，可简化为

$$\max_{\mathbf{x}_2} \left\{-\tfrac{1}{2}(\mathbf{x}_2-\mathbf{a}_2)^\top \mathbf{Q}_2(\mathbf{x}_2-\mathbf{a}_2)\right\} \quad \text{subject to} \quad \mathbf{p}_2^\top \mathbf{x}_2 = m_2 + \mathbf{r}^\top \mathbf{b} \tag{3}$$

为求解这个约束最大化问题，引入拉格朗日函数 n

$$\mathcal{L}_{\lambda_2}(\mathbf{x}_2) = -\tfrac{1}{2}(\mathbf{x}_2-\mathbf{a}_2)^\top \mathbf{Q}_2(\mathbf{x}_2-\mathbf{a}_2) - \lambda_2(\mathbf{p}_2^\top \mathbf{x}_2 - m_2 - \mathbf{r}^\top \mathbf{b}) \tag{4}$$

那么，$\mathcal{L}_{\lambda_2}(\mathbf{x}_2)$作为$\mathbf{x}_2$的函数是凹函数。因此，在任何满足一阶条件的点$\mathbf{x}_2$ 处，它都达到最大值

$$\mathbf{0} = \mathcal{L}'_{\lambda_2}(\mathbf{x}_2) = -(\mathbf{x}_2-\mathbf{a}_2)^\top \mathbf{Q}_2 - \lambda_2 \mathbf{p}_2^\top \tag{5}$$

由于假设$\mathbf{Q}_2$ 正定，因而可逆，所以这个一阶条件显然等价于

$$(\mathbf{x}_2-\mathbf{a}_2)^\top = -\lambda_2 \mathbf{p}_2^\top \mathbf{Q}_2^{-1} \tag{6}$$

或者，转置并整理后，可得

$$\mathbf{x}_2 = \mathbf{a}_2 - \lambda_2 \mathbf{Q}_2^{-1}\mathbf{p}_2 \tag{7}$$

将其代入方程（3）中的预算方程，得到

$$\mathbf{p}_2^\top \mathbf{x}_2 = \mathbf{p}_2^\top(\mathbf{a}_2 - \lambda_2 \mathbf{Q}_2^{-1}\mathbf{p}_2) = m_2 + \mathbf{r}^\top \mathbf{b} \tag{8}$$

这意味着

$$\lambda_2 = \frac{\mathbf{p}_2^\top \mathbf{a}_2 - m_2 - \mathbf{r}^\top \mathbf{b}}{\mathbf{p}_2^\top \mathbf{Q}_2^{-1}\mathbf{p}_2} \tag{9}$$

注意，解 λ_2 存在，因为$\mathbf{p}_2 \neq \mathbf{0}$ 且$\mathbf{Q}_2$ 是正定的。最后，我们可以将（9）与（7）结合起来，确定最优需求向量，即

$$\mathbf{x}_2^* = \mathbf{a}_2 - \frac{\mathbf{p}_2^\top \mathbf{a}_2 - m_2 - \mathbf{r}^\top \mathbf{b}}{\mathbf{p}_2^\top \mathbf{Q}_2^{-1} \mathbf{p}_2} \mathbf{Q}_2^{-1} \mathbf{p}_2 \tag{10}$$

当然，要使这个解在经济上合理，我们应要求 $\lambda_2 \geq 0$，或者等价地，要求$\mathbf{p}_2^\top \mathbf{a}_2 \geq m_2 + \mathbf{r}^\top \mathbf{b}$。由于这个不等式涉及第一期选择的资产向量 $\mathbf{b}$ ，我们将在推导出消费者的第一期最优决策后，再回过头来讨论这个问题。

请注意，这个解意味着事后，在$\mathbf{a}_2$, $\mathbf{r}$和 m_2 都已成为已知量，且$\mathbf{x}_2^*$ 已被最优选择之后，方程（7）和（9）表明，消费者最大化后的第二期效用为

$$\begin{aligned} -\tfrac{1}{2}(\mathbf{x}_2^* - \mathbf{a}_2)^\top \mathbf{Q}_2 (\mathbf{x}_2^* - \mathbf{a}_2) &= -\tfrac{1}{2}\lambda_2^2 \mathbf{p}_2^\top \mathbf{Q}_2^{-1} \mathbf{Q}_2 \mathbf{Q}_2^{-1} \mathbf{p}_2 \\ &= -\frac{(\mathbf{p}_2^\top \mathbf{a}_2 - m_2 - \mathbf{r}^\top \mathbf{b})^2}{2\mathbf{p}_2^\top \mathbf{Q}_2^{-1} \mathbf{p}_2} \end{aligned} \tag{11}$$

4.3 第一期期望效用

回到第一期，我们假设$\mathbf{p}_2$ 和$\mathbf{Q}_2$ 都预先已知。因此，利用（11）后，跨期伯努利效用函数（1）的事前期望值可以表示为以下函数

$$v(\mathbf{x}_1, \mathbf{b}) = -\tfrac{1}{2}(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 (\mathbf{x}_1 - \mathbf{a}_1) - \frac{\mathbb{E}(\mathbf{p}_2^\top \mathbf{a}_2 - m_2 - \mathbf{r}^\top \mathbf{b})^2}{2\mathbf{p}_2^\top \mathbf{Q}_2^{-1} \mathbf{p}_2} \tag{12}$$

关于第一期选择变量$\mathbf{x}_1$ 和$\mathbf{b}$。式（12）右侧第二项中分数的分子可以展开为

$$\begin{aligned} &\mathbb{E}(\mathbf{p}_2^\top \mathbf{a}_2 - m_2 - \mathbf{r}^\top \mathbf{b})^2 \\ &\qquad = \mathbb{E}(\mathbf{p}_2^\top \mathbf{a}_2 - m_2)^2 - 2\mathbb{E}[(\mathbf{p}_2^\top \mathbf{a}_2 - m_2)(\mathbf{r}^\top \mathbf{b})] + \mathbb{E}(\mathbf{r}^\top \mathbf{b})^2 \end{aligned} \tag{13}$$

令 $\bar{}\mathbf{a}_2 := \mathbb{E}\mathbf{a}_2$、 $\bar{m}_2 := \mathbb{E}m_2$ 以及 $\bar{}\mathbf{r} := \mathbb{E}\mathbf{r}$ 分别表示各自均值，且假定这些均值均存在。我们关于$\mathbf{r}$与$\mathbf{a}_2$ 和 m_2 相互独立的假设意味着，式（13）右侧的中间项可化简为

$$\mathbb{E}[(\mathbf{p}_2^\top \mathbf{a}_2 - m_2)(\mathbf{r}^\top \mathbf{b})] = (\mathbf{p}_2^\top \bar{\mathbf{a}}_2 - \bar{m}_2)(\bar{\mathbf{r}}^\top \mathbf{b}) \tag{14}$$

至于式（13）右侧的最后一项，注意到

$$(\mathbf{r}^\top \mathbf{b})^2 = (\mathbf{b}^\top \mathbf{r})\,(\mathbf{r}^\top \mathbf{b}) = \mathbf{b}^\top (\mathbf{r}\,\mathbf{r}^\top)\mathbf{b} \quad \text{and so} \quad \mathbb{E}(\mathbf{r}^\top \mathbf{b})^2 = \mathbf{b}^\top \mathbf{R}\mathbf{b} \tag{15}$$

其中，$\mathbf{R}$ 表示收益二阶矩的对称方阵 $\mathbb{E}[\mathbf{r}\mathbf{r}^\top]$ ，我们也假设该矩阵存在。根据假设，矩阵$\mathbf{R}$ 是正定的，因为任何投资组合$\mathbf{b} \neq \mathbf{0}$ 的收益二阶矩 $\mathbb{E}(\mathbf{r}^\top \mathbf{b})^2$ 始终为正。

将式（14）和（15）代入式（13），得

$$\begin{aligned} \mathbb{E}(\mathbf{p}_2^\top \mathbf{a}_2 - m_2 - \mathbf{r}^\top \mathbf{b})^2 &= \mathbb{E}(\mathbf{p}_2^\top \mathbf{a}_2 - m_2)^2 - 2(\mathbf{p}_2^\top \bar{\mathbf{a}}_2 - \bar{m}_2)\bar{\mathbf{r}}^\top \mathbf{b} + \mathbf{b}^\top \mathbf{R}\mathbf{b} \\ &= c + (\mathbf{b}^* - \mathbf{b})^\top \mathbf{R}(\mathbf{b}^* - \mathbf{b}) \end{aligned} \tag{16}$$

其中$\mathbf{b}^{*\top}\mathbf{R} = (\mathbf{p}_2^\top \bar{\mathbf{a}}_2 - \bar{m}_2)\bar{\mathbf{r}}^\top$，这意味着$\mathbf{b}^* = \mathbf{R}^{-1}\bar{\mathbf{r}}(\mathbf{p}_2^\top \bar{\mathbf{a}}_2 - \bar{m}_2)$，并且

$$c = \mathbb{E}(\mathbf{p}_2^\top \mathbf{a}_2 - m_2)^2 - \mathbf{b}^{*\top}\mathbf{R}\mathbf{b}^* = \mathbb{E}(\mathbf{p}_2^\top \mathbf{a}_2 - m_2)^2 - (\mathbf{p}_2^\top \bar{\mathbf{a}}_2 - \bar{m}_2)^2 \bar{\mathbf{r}}^\top \mathbf{R}^{-1}\bar{\mathbf{r}} \tag{17}$$

因此，最后，在忽略一个无关紧要的加性常数后，消费者第一期的最大化目标可写成如下二次型

$$v(\mathbf{x}_1, \mathbf{b}) = -\tfrac{1}{2}(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 (\mathbf{x}_1 - \mathbf{a}_1) - \tfrac{1}{2}(\mathbf{b}^* - \mathbf{b})^\top \mathbf{S}(\mathbf{b}^* - \mathbf{b}) \tag{18}$$

其中 $\mathbf{S} := \mathbf{R}/\mathbf{p}_2^\top \mathbf{Q}_2^{-1}\mathbf{p}_2$。

4.4 第一期优化问题

因此，消费者第一期的优化问题是：在预算约束 $\mathbf{p}_1^\top \mathbf{x}_1 + \mathbf{q}^\top \mathbf{b} = m_1$下，对函数（18）关于$\mathbf{x}_1$和$\mathbf{b}$进行最大化。我们通过引入拉格朗日函数来求解这个约束最大化问题

$$\begin{aligned} \mathcal{L}_{\lambda_1}(\mathbf{x}_1, \mathbf{b}) = -\tfrac{1}{2}(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 (\mathbf{x}_1 - \mathbf{a}_1) - \tfrac{1}{2}(\mathbf{b}^* - \mathbf{b})^\top \mathbf{S}(\mathbf{b}^* - \mathbf{b}) \\ - \lambda_1(\mathbf{p}_1^\top \mathbf{x}_1 + \mathbf{q}^\top \mathbf{b} - m_1) \end{aligned} \tag{19}$$

其作为（$\mathbf{x}_1$,$\mathbf{b}$）的函数是凹函数，因此当两个一阶条件

$$\begin{aligned} \mathbf{0} &= \mathcal{L}'_{\lambda_1, \mathbf{x}_1} &&= -(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 - \lambda_1 \mathbf{p}_1^\top \\ \text{and} \quad \mathbf{0} &= \mathcal{L}'_{\lambda_1, \mathbf{b}} &&= (\mathbf{b}^* - \mathbf{b})^\top \mathbf{S} - \lambda_1 \mathbf{q}^\top \end{aligned} \tag{20}$$

均得到满足时，关于（$\mathbf{Q}_1$ 和$\mathbf{S}$ ）的函数达到最大值。由于$\mathbf{Q}_1$ 和$\mathbf{S}$ 均为正定矩阵，因而可逆，这些一阶条件等价于

$$(\mathbf{x}_1 - \mathbf{a}_1)^\top = -\lambda_1 \mathbf{p}_1^\top \mathbf{Q}_1^{-1} \quad \text{and} \quad (\mathbf{b}^* - \mathbf{b})^\top = \lambda_1 \mathbf{q}^\top \mathbf{S}^{-1} \tag{21}$$

或者，转置并整理后，得到

$$\mathbf{x}_1 = \mathbf{a}_1 - \lambda_1 \mathbf{Q}_1^{-1} \mathbf{p}_1 \quad \text{and} \quad \mathbf{b} = \mathbf{b}^* - \lambda_1 \mathbf{S}^{-1} \mathbf{q} \tag{22}$$

将这些代入预算方程，得到

$$\mathbf{p}_1^\top \mathbf{x}_1 + \mathbf{q}^\top \mathbf{b} = \mathbf{p}_1^\top (\mathbf{a}_1 - \lambda_1 \mathbf{Q}_1^{-1} \mathbf{p}_1) + \mathbf{q}^\top (\mathbf{b}^* - \lambda_1 \mathbf{S}^{-1} \mathbf{q}) = m_1 \tag{23}$$

这意味着

$$\lambda_1 = \frac{\mathbf{p}_1^\top \mathbf{a}_1 + \mathbf{q}^\top \mathbf{b}^* - m_1}{\mathbf{p}_1^\top \mathbf{Q}_1^{-1} \mathbf{p}_1 + \mathbf{q}^\top \mathbf{S}^{-1} \mathbf{q}} \tag{24}$$

注意，这一定义良好，因为$\mathbf{p}_1 \neq \mathbf{0}$ 和$\mathbf{q} \neq \mathbf{0}$，而且两个对称矩阵$\mathbf{Q}_1$ 和$\mathbf{S}$ 均为正定矩阵，因此可逆，且其逆矩阵也为正定矩阵。[15] 最后，我们可以利用（22）和（24）确定最优商品需求向量和资产需求向量，结果为

$$\mathbf{x}_1^* = \mathbf{a}_1 - \frac{\mathbf{p}_1^\top \mathbf{a}_1 + \mathbf{q}^\top \mathbf{b}^* - m_1}{\mathbf{p}_1^\top \mathbf{Q}_1^{-1} \mathbf{p}_1 + \mathbf{q}^\top \mathbf{S}^{-1} \mathbf{q}} \mathbf{Q}_1^{-1} \mathbf{p}_1 \tag{25}$$

$$\text{and} \quad \mathbf{b} = \mathbf{b}^* - \frac{\mathbf{p}_1^\top \mathbf{a}_1 + \mathbf{q}^\top \mathbf{b}^* - m_1}{\mathbf{p}_1^\top \mathbf{Q}_1^{-1} \mathbf{p}_1 + \mathbf{q}^\top \mathbf{S}^{-1} \mathbf{q}} \mathbf{S}^{-1} \mathbf{q} \tag{26}$$

当然，要使这个解在经济上合理，我们应要求 $\lambda_1 \geq 0$，或者等价地，要求

$$\mathbf{p}_1^\top \mathbf{a}_1 + \mathbf{q}^\top \mathbf{b}^* = \mathbf{p}_1^\top \mathbf{a}_1 + \mathbf{q}^\top \mathbf{R}^{-1} \bar{\mathbf{r}} \, (\mathbf{p}_2^\top \bar{\mathbf{a}}_2 - \bar{m}_2) \geq m_1 \tag{27}$$

此外，为了使我们先前求得的第二期解在经济上合理，还应要求 $\mathbf{p}_2^\top \mathbf{a}_2 \geq m_2 + \mathbf{r}^\top \mathbf{b}$。因为

[15]矩阵理论中的一个标准结果是，任意对称 $n \times n$ 矩阵$\mathbf{A}$ 都可以对角化，即存在一个正交矩阵$\mathbf{E}$ （意味着$\mathbf{E}^{-1} = \mathbf{E}^\top$）和一个对角矩阵$\mathbf{D}$，使得$\mathbf{E}\mathbf{A}\mathbf{E}^\top = \mathbf{D}$，因此$\mathbf{A} = \mathbf{E}^\top \mathbf{D}\mathbf{E}$。由此容易看出，以下四个条件在逻辑上等价：（i）$\mathbf{A}$ 是正定矩阵；（ii）$\mathbf{D}$ 的所有对角元素均为正；（iii）$\mathbf{D}^{-1}$ 存在，且其所有对角元素均为正；（iv）$\mathbf{A}^{-1} = \mathbf{E}^\top \mathbf{D}^{-1} \mathbf{E}$ 存在且为正定矩阵。

两个随机变量 $\mathbf{p}_2^\top \mathbf{a}_2 - m_2$ 和 $\mathbf{r}^\top \mathbf{b}$ 相互独立，因此这一要求意味着，存在一个实数 α，使得在给定最优选择$\mathbf{b}$的情况下，有

$$\mathbf{p}_2^\top \mathbf{a}_2 - m_2 \geq \alpha \geq \mathbf{r}^\top \mathbf{b} \tag{28}$$

对于随机对 $(\mathbf{p}_2^\top \mathbf{a}_2 - m_2, \mathbf{r}^\top \mathbf{b}) \in \mathbb{R}^2$ 几乎所有可能的取值。

4.5 一个活化的决策问题

为了使这个线性—二次决策模型更加生动，我们考虑这样一种可能性：在第1期已经选定数对 $(\mathbf{x}_1, \mathbf{b})$ 之后，发生不可预见的变化。一般而言，可能会出现一个新的第二期目标函数

$$-\tfrac{1}{2}(\mathbf{x}_2^+ - \mathbf{a}_2^+)^\top \mathbf{Q}_2^+ (\mathbf{x}_2^+ - \mathbf{a}_2^+) \tag{29}$$

其中向量 $\mathbf{x}_2^+$、$\mathbf{a}_2^+$ 的维度以及相应正定方阵 $\mathbf{Q}_2^+$ 的维度可能有所增加，这或许是因为出现了新商品。当然，第二期预算约束也必须改变；我们将其写作

$$\mathbf{p}_2^{+\top}\, \mathbf{x}_2^+ \leq \mathbf{r}^\top\, \mathbf{b} + m_2 \tag{30}$$

资产向量 $\mathbf{b}$ 与之前相同，因为它已由消费者在第1期的决策确定。三元组 $(\mathbf{a}_2^+, m_2, \mathbf{r})$ 的联合分布也可能改变；事实上，如果 $\mathbf{a}_2^+$ 的维度大于 $\mathbf{a}_2$ 的维度，就必然会改变。

如果能够提前得知这些变化，那么在第1期，消费者面临的将是最大化一个修订后的二次目标函数，而不是由式（18）定义的二次评价函数 $v(\mathbf{x}_1, \mathbf{b})$

$$v^+(\mathbf{x}_1, \mathbf{b}) = -\tfrac{1}{2}(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 (\mathbf{x}_1 - \mathbf{a}_1) - \tfrac{1}{2}(\mathbf{b}^{+*} - \mathbf{b})^\top \mathbf{S}^+ (\mathbf{b}^{+*} - \mathbf{b}) \tag{31}$$

具有相同的选择变量 $\mathbf{x}_1$ 和$\mathbf{b}$，并受与式（2）相同的第一期预算约束 $\mathbf{p}_1^\top \mathbf{x}_1 + \mathbf{q}^\top \mathbf{b} = m_1$ 限制。然而，发生变化的是出现在式（31）最后一项中的向量参数 $\mathbf{b}^{+*}$ 和矩阵参数 $\mathbf{S}^+$，它们的变化如今完全无法预测。活化要求我们认识到，这些参数本身也必须被视为不确定的。根据定义，一位仍然坚信某个二次模型依然适用的贝叶斯理性消费者，会对参数$\mathbf{b}^*$，$\mathbf{S}$这一对无法预测的参数持有某种主观概率信念，而这对参数

刻画了二次评价函数参数族中的每一个成员

$$\widetilde{v}(\mathbf{x}_1, \mathbf{b}; \widetilde{\mathbf{b}}^*, \widetilde{\mathbf{S}}) \equiv -\tfrac{1}{2}(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 (\mathbf{x}_1 - \mathbf{a}_1) - \tfrac{1}{2}(\widetilde{\mathbf{b}}^* - \mathbf{b})^\top \widetilde{\mathbf{S}} (\widetilde{\mathbf{b}}^* - \mathbf{b}) \quad (32)$$

按照主观期望效用最大化的理性要求，第1期的$\widehat{\ }$ $\widetilde{\ }$ $\widetilde{\ }$最优策略须最大化函数（32）相对于关于参数对$\widetilde{\ }$ $\widetilde{\ }$(**b***，**S**)的概率信念的期望值 $\mathbb{E}[\widetilde{v}$(**x**$_1$, **b**;**b***，**S**)]。不过，忽略一个无关紧要的加性常数后，这一期望可表示为以下便于处理的形式

$$\widehat{\mathbb{E}}[\widetilde{v}(\mathbf{x}_1, \mathbf{b}; \widetilde{\mathbf{b}}^*, \widetilde{\mathbf{S}})] \equiv -\tfrac{1}{2}(\mathbf{x}_1 - \mathbf{a}_1)^\top \mathbf{Q}_1 (\mathbf{x}_1 - \mathbf{a}_1) - \tfrac{1}{2}(\widehat{\mathbf{b}}^* - \mathbf{b})^\top \widehat{\mathbf{S}} (\widehat{\mathbf{b}}^* - \mathbf{b}) \quad (33)$$

$\widehat{\ }$ $\widehat{\ }$ $\widetilde{\ }$这涉及随机矩阵 **S** 的适当主观期望值 $\mathbf{S} := \mathbb{E}[\mathbf{S}]$ 。注意，矩阵 **S** 是正定的，因此可逆，因为它是随机正定矩阵 **S** 的期望值。[16] 这$\widehat{\ }$使向量 $\mathbf{b}^*$ 能够被唯一选定，从而满足一阶$\widehat{\ }$ $\widehat{\ }$ $\widehat{\ }$ $\widetilde{\ }$ $\widetilde{\ }$条件 $\mathbf{S}\mathbf{b}^* = \mathbb{E}[\mathbf{S}\mathbf{b}^*]$。这意味着存在唯一的最优向量$\widehat{\ }$$\mathbf{b}^*$ ，其形式为

$$\widehat{\mathbf{b}}^* = \widehat{\mathbf{S}}^{-1}\, \widehat{\mathbb{E}}[\widetilde{\mathbf{S}}\, \widetilde{\mathbf{b}}^*] = (\widehat{\mathbb{E}}[\widetilde{\mathbf{S}}])^{-1}\, \widehat{\mathbb{E}}[\widetilde{\mathbf{S}}\, \widetilde{\mathbf{b}}^*] \quad (34)$$

$\widehat{\ }$ $\widehat{\ }$ $\widetilde{\ }$当然，方程（34）通常意味着$\mathbf{b}^* \neq \mathbb{E}[\mathbf{b}^*]$ ； 不过，例如，当随机参数对（$\mathbf{b}^*$,**S**）不相关时，这显然是一个特殊情形，因此 $\mathbb{E}[\mathbf{S}\mathbf{b}^*] = \mathbb{E}[$**S**$]\, \mathbb{E}[\mathbf{b}^*]$。

5 计算机国际象棋

5.1 简化国际象棋

我们的两个扩展示例中的第二个，涉及一位国际象棋棋手面对一个已知局面 n_0 时必须选择着法所面临的决策问题。要具体说明这个局面，需要指出轮到谁走棋，以及棋盘上的64个方格中每一格（如果有的话）由什么棋子占据。[17] 那么，令 $N_1 := N_{+1}(n_0)$ 表示由所有可通过在局面 n_0 下走出合法着法而到达的局面组成的集合。

[16]要说明这一点，请注意，如果随机对称矩阵 $\widetilde{\ }$**S** 几乎必然是正定的，那么对所有**u** $\in \mathbb{R}^n \setminus \{\mathbf{0}\}$ 都有$\mathbf{u}^\top(\widehat{\mathbb{E}}\widetilde{\ }\mathbf{S})\mathbf{u} = \widehat{\mathbb{E}}(\mathbf{u}^\top \widetilde{\ }\mathbf{S}\mathbf{u}) > 0$。[17]实际上，即使是在简化版国际象棋中——没有用于限制每位玩家总思考时间的棋钟，也没有超出.....的和棋规则

回顾一下，在国际象棋棋局中，只有当一方的王受到对手的棋子攻击时，该王才算被将军；也就是说，若没有其他着法介入，该棋子便能吃掉王。玩家的着法只有在不会使己方王处于被将军状态时才是合法的。如果轮到某位玩家走棋，而该玩家没有合法着法，那么：（i）若该玩家的王正被将军，则该玩家已被将死并输掉棋局；（ii）若该玩家的王未被将军，则形成逼和，棋局为和棋。

根据 Zermelo（1913）的著名结果，以及 von Neumann（1928）对完全信息双人“零和”博弈中的最大最小或极小极大策略所作的开创性分析，在白方和黑方玩家都采取最佳策略的情况下，存在一个客观的结果函数

$$N_1 \ni n_1 \mapsto r^+(n_1) \in \{W, D, L\} \tag{35}$$

该函数将每个可能的局面 $n_1 \in N_1$ 映射为棋局的一个确定结果$r^+(n_1)$ $\in \{W, D, L\}$；对于即将行棋的玩家而言，结果要么是胜局（W），要么是和棋（D），要么是败局（L）。可以使用计分规则将这一结果转换为收益，例如白方胜记 1 分，-1 黑方胜记 1 分，而和棋记 0 分。于是，给定一个从局面 n开始的国际象棋延续子博弈，该子博弈中双方玩家采取最佳策略所产生的结果，将由一个客观的评估函数给出

$$N_1 \ni n_1 \mapsto v(n_1) \in \{1, 0, -1\} \tag{36}$$

对于在 n_0 轮到着法的玩家，当且仅当以下条件成立时，从 n_0 到 n_1 的着法是最优的：

1. 若在 n_0 轮到白方着法，则 n_1 使评估函数 $v(n_1)$ 最大化；

2. 若轮到黑方在 n_0行棋，n_1 会使评估函数 $v(n_1)$最小化。

逼和、三次重复或长将——国际象棋规则规定：(i) 如果王或相关的车此前曾移动过，则不允许王车易位；(ii) 兵可以吃过路兵，但只能在即将被吃掉的兵从初始位置前进两格后立即进行。因此，许多国际象棋局面的完整描述需要显著更多的信息。

然而，式（36）中的客观标准化估值函数只能针对少数相对简单的局面计算，这些局面满足：

- 或者可以证明，一方能在少数几步之内通过将死强制取胜；否则，只要他们愿意，便可通过以下任一方式达成和棋：（i）逼和；（ii）局面三次重复；（iii）长将；

- 或者，棋盘上的棋子不超过7枚（包括双方的王）。在这种情况下，第2.1节脚注7中引用的罗蒙诺索夫“残局数据库”软件，将指出双方玩家均采取极大极小策略时棋局的结果。

因此，玩家在 n_0选择要走哪一步，从而决定棋盘上的下一局面 $n_1 \in N_1$ 时，通常不得不对收益函数形成主观信念。通过向前推演几步，可以为这些信念提供指引。但除非能够确定无疑地算出一种方法，强制进入一个评价结果明确的简单局面，否则最终仍须对前方几步后的许多局面赋予此类评价。由此，便构建出一个将国际象棋局面映射为主观期望收益的主观评估函数。用于此目的的国际象棋计算机程序采用的算法性能优异，甚至超越人类，但仍不可避免地存在不足。目前，一些最有效的软件采用基于蒙特卡洛树搜索（MCTS）的算法，第7.4节将进一步讨论这一算法——Browne等人（2012年）对其作了概述，该综述在计算机科学文献中被广泛引用。将MCTS应用于国际象棋时，为了评价给定局面 n，它会考虑许多从局面 n开始的模拟后续子博弈，并在选择每一步后续着法的过程中加入少量、经过精心控制的随机性。随后，任一局面 n的最终评价，便是所有从局面 n开始的模拟棋局的平均得分。

5.2 真实国际象棋

真实国际象棋复杂得多。首先，即将走棋的玩家可以证明下一步棋既可以使同一局面第三次出现，也可以使双方都已走满至少50步，期间没有棋子被吃或兵被移动，从而要求和棋。此外，棋局通常以以下方式之一结束：(i) 一方

处于劣势的玩家选择认输；或(ii)双方都认为自己获胜的机会不足，于是同意和棋。

最后，比赛还设有由国际象棋计时钟监控的时限；更准确地说，是一对相连的棋钟，每位玩家各有一个，用来显示该玩家在下一个时限控制开始前还剩多少总时间。任一玩家走完一步后，都可以按下一个操纵杆，同时停止自己的棋钟并启动对手的棋钟。这些额外因素使得对任何国际象棋局面 n 的描述都复杂得多，因为描述中必须包括例如每位玩家在超时判负前还可以使用多少时间。

5.3 有界模型中的人类失误

国际象棋专家通过在每个局面中只关注少数几个看似合理的着法来运用自己的棋艺。给定任意一个合法国际象棋局面n_0，考虑集合 $N_1 := N_{+1}(n_0)$，其中包含从局面 n_0 出发，通过合法着法 n_1 后可能到达的所有局面 n_1 。国际象棋专家能够辨别出，集合 N_1中的许多局面 n_1 虽然允许到达，却差得太远，不值得投入太多关注，甚至完全不值得考虑。当然，国际象棋专家也非常擅长评估他们可能想到要走到的任何局面 n_1 的价值。从这个意义上说，他们拥有良好的有界模型。

然而，毕竟只是人类，即使是最优秀的玩家，对不同局面的模型和评估有时也可能存在严重缺陷。2006年，时任世界冠军的弗拉基米尔·克рамник就犯下了“世纪大失误”：他漏看了一步将死，导致立即败局。这次失误发生在他与计算机程序 Deep Fritz 进行的六局比赛中的第二局。[18]

在这盘棋中，Deep Fritz执白方棋子。图5所示局面出现之前，它走出的最后一着是第34着，记作34. Ne6×f8。这个记谱表示，白方的马从e6格吃掉了黑方位于f8格的棋子，实际上是一辆车。作为回应，Kramnik（执黑方）犯下了严重失误，走出后的一着34 . . . Qa7–e3，如图5中的箭头所示，从而形成了该图所示的局面。随后，计算机程序Deep

[18]参见 https://en.chessbase.com/post/how-could-kramnik-overlook-the-mate-。请注意，网址末尾的连字符是网址的重要组成部分，但末尾不应加句号。

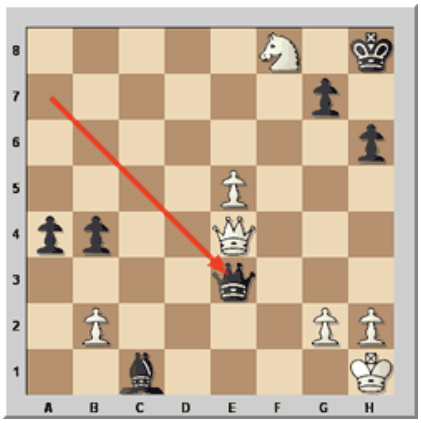

图5：Deep Fritz 对 Kramnik，第2局

Fritz立即指出，它的下一着——后从e4格走到h7格的35. Qe4–h7——将通过将死白方而立即赢得胜局。

恐怕有成千上万，甚至数百万名棋力远不及 Kramnik 的国际象棋棋手，只要让他们看到第34.....Qa7–e3着法之前的局面，就一定会注意到白方威胁着下出第35.Qe4–h7将死。那么，Kramnik 是怎么漏看这一点的呢?

一个很好的线索，来自 Kramnik 本人在棋局结束几分钟后举行的新闻发布会上所说的一番话:

> “其实，问题不只在最后一着.....我把这条变化算了很多很多遍，一遍遍地复查。我在下第29.....Qa7时就已经算过这条变化，此后每走一着，我都会反复重新计算，结果最后还是漏看了一着将死。其实这是第一次发生在我身上，我实在找不到任何解释。我当时并不觉得累，我想整盘棋我都算得很好.....这实在太奇怪了，我解释不了。”

一种解读是，Kramnik执黑时，早在选择第29着时，就已经计划走后来证明是灾难性的一着：34 . . . Qa7–e3，如图5所示。毕竟，如果图5中的黑王位于g8格而不是h8格，那么黑方就可以用35. . . . K×f8应对35. Qe4–h7将军。在

然而，在实际棋局中，即使到了最后一刻、他即将走出34 . . . Qa7–e3时，他仍未像理应做的那样修正自己的棋局心理模型，没有意识到这一着会让对手一步将死。最终，Kramnik仿佛过于执着于早先走34 . . . Qa7–e3的计划，以至于没有察觉自己走出这着致命棋步后可能立即被将死。

最后，尽管Kramnik声称自己无法解释这一失误，但事实上，上面的引述确实提供了部分解释。它没有解释的是：为什么他会如此执着于自己先前那个有缺陷的模型。

6 决策树中的贝叶斯理性

6.1 轮盘彩票

沿用Anscombe和Aumann（1963）的术语，对于任意非空集合 Z，令 $\Delta(Z)$表示所有轮盘彩票或简单概率测度的集合。这些彩票是形如函数 $Z \ni z \mapsto \lambda(z) \in [0,1]$ 的对象，具有有限支撑$\operatorname{supp}\lambda \subseteq Z$，满足

$$\lambda(z) > 0 \iff z \in \operatorname{supp}\lambda \quad \text{and} \quad \sum\nolimits_{z\in Z} \lambda(z) = \sum\nolimits_{z\in \operatorname{supp}\lambda} \lambda(z) = 1 \tag{37}$$

于是，对于任意 $z \in Z$，令 $Z \ni z' \mapsto \delta_z(z') \in \Delta(Z)$表示满足 $\delta_z(z) = 1$的唯一退化彩票。此外，每当 Z 为有限集时，令 $\Delta^0(Z)$表示满足$\operatorname{supp}\lambda = Z$的完全支撑彩票 $\lambda \in \Delta(Z)$的集合，或者等价地，对于所有 $z \in Z$，$\lambda(z) > 0$。最后，给定任意λ，$\mu \in \Delta(Z)$以及任意标量 $\alpha \in [0,1]$，令 $\nu := \alpha\lambda + (1-\alpha)\mu \in \Delta(Z)$表示彩票混合 $Z \ni z \mapsto \nu(z) \in [0,1]$，它对所有 $z \in Z$满足

$$\nu(z) = [\alpha\lambda + (1-\alpha)\mu](z) = \alpha\lambda(z) + (1-\alpha)\mu(z) \tag{38}$$

6.2 Anscombe–Aumann（AA）结果彩票

贝叶斯理性假设，或等价地说，主观期望效用最大化假设，适用于存在一个非空的状态空间 S，其中包含可能的世界状态 s，并且主观概率映射 $S \ni s \mapsto p(s) \in [0,1]$ 在该空间上有定义，其中 $\sum_{s\in S} p(s) = 1$。再次沿用Anscombe and Aumann（1963）的做法，我们假设 S 是有限的。此外，根据他们在第1.1节中介绍的术语，任何随机过程

若用于确定一个不确定的世界状态，则称为“马彩票”。

贝叶斯理性关注的是对安斯康姆–奥曼彩票结果的偏好。根据定义，这类彩票既可能涉及由轮盘彩票带来的风险，也可能涉及由赛马彩票带来的不确定性。令 Y 表示一个非空的后果域，再令 $\Delta(Y)$表示 Y上的轮盘彩票定义域。

接下来，给定状态有限集 S s 以及后果域 Y，对每个状态 $s \in S$，令 Y_s 为 Y的一个副本。[19] 再令

$$L^S(Y) := \prod_{s \in S} \Delta(Y_s) = \{\langle \lambda_s \rangle_{s \in S} \mid \forall s \in S : \lambda_s \in \Delta(Y_s)\} \tag{39}$$

表示安斯康姆–奥曼彩票（或*AA* 彩票）的空间，其形式为列表 $\langle \lambda_s \rangle_{s \in S}$ 或映射 $S \ni s \mapsto \lambda_s \in \Delta(Y)$。每个这样的映射都指定了一个组合：首先是决定状态 $s \in S$的赛马彩票，其次是决定结果 $y \in Y$的状态依赖轮盘彩票 λ_s 。

6.3 从二元集合中选择与基础偏好

令 $\mathcal{F}_{\setminus\emptyset}(L^S(Y))$ 表示 AA-彩票域 $L^S(Y)$ 的非空有限子集族。在该彩票域上的一个 选择函数 是一个映射

$$\mathcal{F}_{\setminus\emptyset}(L^S(Y)) \ni F \mapsto C(F) \in \mathcal{F}_{\setminus\emptyset}(L^S(Y)) \tag{40}$$

对于每个非空的 可行集 $F \in \mathcal{F}_{\setminus\emptyset}(L^S(Y))$，确定一个非空的 选择集 $C(F) \in \mathcal{F}_{\setminus\emptyset}(L^S(Y))$，满足 $C(F) \subseteq F$。

对于任意定义在 $\mathcal{F}_{\setminus\emptyset}(L^S(Y))$ 上的选择函数 $F \mapsto C(F)$，当 F 是包含 # $F = 2$ 个元素的二元集合 时，其取值确定一个严格偏好关系

[19]这里讨论的是<style id='2'>状态无关的后果域情形；为简化记号，我们作此假设。当 Y_s 取决于 s时，就会出现更一般的<style id='4'>状态依赖的后果域情形。在这种情况下，令 $Y^\cup := \cup_{s \in S} Y_s$ 表示所有后果 y 的<style id='10'>并域，其中 y 在某个状态 $s \in S$ 下是可行的。于是，在可行的状态–后果对的定义域 $D := \{(s, y) \in S \times Y^\cup \mid y \in Y_s\}$ 上，可能存在一个<style id='14'>状态依赖效用函数 $D \ni (s, y) \mapsto u(s, y) \to \mathbb{R}$。这类状态依赖效用函数已在 Dr`eze（1962年）、Karni（1985年）、Schervish 等（1990年）、Dr`eze 和 Rustichini（2004年）以及 Seidenfeld 等（2010年）的研究中得到探讨。参见 Hammond（1998b年、1999年、2022年），其中给出了一种统一处理，推广了后果域为状态无关时的特殊情形，并且即使在后果域为状态依赖的一般情形下，也推导出了状态无关效用函数。

$\succ_C$、严格不偏好关系 $\prec_C$ 和无差异关系 $\sim_C$。这三种二元关系的定义如下：对于每一对 $\lambda^S,\ \mu^S \in L^S(Y)$，均有

$$\lambda^S \left\{ \begin{array}{c} \succ_C \\ \sim_C \\ \prec_C \end{array} \right\} \mu^S \quad \text{according as} \quad C(\{\lambda^S, \mu^S\}) = \left\{ \begin{array}{c} \{\lambda^S\} \\ \{\lambda^S, \mu^S\} \\ \{\mu^S\} \end{array} \right\} \tag{41}$$

由（40）所规定的选择函数 $F \mapsto C(F)$ 背后，存在一个与之对应的唯一二元弱偏好关系 $\succsim_C$，定义在 $L^S(Y)$ 上，称为基础关系。对于每一对 $\lambda^S, \mu^S \in L^S(Y)$，该基础关系满足

$$\lambda^S \succsim_C \mu^S \iff \lambda^S \in C(\{\lambda^S, \mu^S\}) \iff \lambda^S \succ_C \mu^S \text{ or } \lambda^S \sim_C \mu^S \tag{42}$$

最后，我们提及定义在 $L^S(Y)$ 上的相应弱不偏好关系 $\precsim_C$，使得

$$\lambda^S \precsim_C \mu^S \iff \mu^S \in C(\{\lambda^S, \mu^S\}) \iff \lambda^S \prec_C \mu^S \text{ or } \lambda^S \sim_C \mu^S \tag{43}$$

显然，弱偏好关系 $\succsim_C$ 和弱不偏好关系 $\precsim_C$ 都是完全的，也就是说，对于每一对 $\lambda^S, \mu^S \in L^S(Y)$，以下两项都必须成立：

1. $\lambda^S \succsim_C \mu^S$、 $\mu^S \succsim_C \lambda^S$ 或二者同时成立；
2. $\lambda^S \precsim_C \mu^S$、 $\mu^S \precsim_C \lambda^S$ 或二者同时成立。

6.4 贝叶斯理性与期望效用

设 $F \mapsto C(F)$ 是任意一个同时满足（40）和对所有 $F \in \mathcal{F}_{\setminus\emptyset}(L^S(Y))$ 均有 $C(F) \subseteq F$ 的选择函数，该选择函数对应于由（41）定义在 AA 彩票 λ^S 的空间 $L^S(Y)$ 上的基础偏好关系 $\succsim_C$。那么，映射$L^S(Y) \ni \lambda^S \mapsto U^S(\lambda^S) \in \mathbb{R}$ 是一个效用函数，当且仅当它表示 定义在 $L^S(Y)$ 上的偏好关系 $\succsim_C$，也就是说，对所有 $\lambda^S, \mu^S \in L^S(Y)$，均有

$$\lambda^S \succsim_C \mu^S \iff U^S(\lambda^S) \geq U^S(\mu^S) \tag{44}$$

以下定义排除了主观概率为零的情况，这与决策理论中已成为标准的做法有所不同。正如 Hammond (1998b) 第 5.4 节以及 Hammond (2022年) 所述，这一限制旨在避免在允许概率为零时，决策树的延续子树中出现的困难。

定义 1。 令 S 表示世界中不确定状态的固定非空有限集。

1. 内点主观概率质量函数是一个映射 $S \ni s \mapsto \mathbb{P}(s) \in (0,1]$，满足 $\sum_{s\in S} \mathbb{P}(s) = 1$。*2.* 定义在定义域 $\mathcal{F}_{\setminus\emptyset}(L^S(Y))$ 上的选择函数 $F \mapsto C(F)$，以及与之相关、定义在空间 $L^S(Y)$ 上的基础偏好关系 $\succsim_C$，当且仅当存在一个内点主观概率质量函数 $S \ni s \mapsto \mathbb{P}(s) \in (0,1]$，以及一个伯努利效用函数 $Y \ni y \mapsto u(y) \to \mathbb{R}$，使得 $\succsim_C$ 由冯·诺依曼主观期望效用函数表示（即满足式（44））时，二者都是贝叶斯理性的。该函数对所有 AA 彩票 $\lambda^S = \langle \lambda_s \rangle_{s\in S} \in L^S(Y)$ 的定义由双重求和给出

$$U^S(\lambda^S) = \sum\nolimits_{s\in S} \mathbb{P}(s) \sum\nolimits_{y\in Y} \lambda_s(y) u(y) \tag{45}$$

6.5 规范化效用

回顾第6.1节中的说明，对于每个 $y \in Y$，我们使用 δ_y 表示 $\Delta(Y)$ 中满足 $\delta_y(\{y\}) = 1$ 的唯一 <style id='3'> 退化概率测度。为避免琐碎情形，我们假设定义域 Y 中至少存在三个结果y、y^0和y，使得对于相应的三个退化彩票 δ_y、δ_{y^0}和δ_y，基础严格偏好关系 $\succ$ 满足严格偏好性质 $\delta_y \succ \delta_{y^0} \succ \delta_{y^0}$。

考虑任意伯努利效用函数 $Y \ni y \mapsto u(y) \to \mathbb{R}$，以及定义在 $\Delta(Y)$ 上、与之相关的偏好关系 $\succsim$，二者共同满足

$$\lambda \succsim \mu \iff U(\lambda^S) \geq U(\mu^S) \tag{46}$$

对于定义在 $\Delta(Y)$ 上的冯·诺依曼客观期望效用函数，其定义为

$$U(\lambda) = \sum\nolimits_{y\in Y} \lambda(y) u(y) \tag{47}$$

如 Hammond（1998a年，第 2.3 节）所述，假设贝叶斯理性，效用差之比 $u(y^0) - u(y)\ u(y) - u(y)$ 用经济学家的术语来说，等于无差异曲线上概率变动的恒定边际替代率：（i）概率从结果 y 变为 y^0；（ii）从

结果 y 变为 $\underline{y}$。因此我们称两个伯努利效用函数$y \mapsto u(y)$和 $y \mapsto \tilde{u}(y)$等价，当且仅当对于 Y 中满足 $\delta_y \succ \delta_{y^0} \succ \delta_y$ 的每组三个结果 $\underline{y}$、 y^0、 $\overline{y}$，相应的效用差之比满足 $u(\underline{y}) < u(y^0) < u(\overline{y})$。

$$\frac{u(y^0) - u(\underline{y})}{u(\overline{y}) - u(\underline{y})} = \frac{\tilde{u}(y^0) - \tilde{u}(\underline{y})}{\tilde{u}(\overline{y}) - \tilde{u}(\underline{y})} \tag{48}$$

但是，对于满足 $u(\underline{y}) < u(y^0) < u(\overline{y})$ 的每一组三元组 $\underline{y}$、 y^0和 $\overline{y}$，(48) 都成立，当且仅当存在一个加性常数 $\alpha \in \mathbb{R}$ 和一个正的乘性常数 $\rho \in \mathbb{R}$，使得对于所有 $y \in Y$，均有

$$\tilde{u}(y) = \alpha + \rho u(y) \tag{49}$$

现在，给定任意一对满足 $\bar{u} > \underline{u}$ 的实数 $\underline{u}$和 $\bar{u}$，以及任意伯努利效用函数 $Y \ni y \mapsto u(y) \to \mathbb{R}$，都存在两个唯一的常数 α 和 $\rho > 0$，使得由 (49) 定义的变换后效用函数是一个规范化的效用函数，并满足 $\tilde{u}(\underline{y}) = \underline{u}$ 以及 $\tilde{u}(\overline{y}) = \bar{u}$。事实上，我们所需的两个常数为

$$\rho = \frac{\bar{u} - \underline{u}}{u(\overline{y}) - u(\underline{y})} \quad \text{and then} \quad \alpha = \bar{u} - \rho u(\overline{y}) = \underline{u} - \rho u(\underline{y}) \tag{50}$$

从现在起，用 u 表示唯一的伯努利效用函数 $Y \ni y \mapsto u(y) \to \mathbb{R}$，其由 (47) 定义的期望值满足 (46)，并且已规范化为满足

$$u(\underline{y}) = \underline{u} \quad \text{and} \quad u(\overline{y}) = \bar{u} \tag{51}$$

6.6 有限决策树及其延续

在数学术语中，有限有向图(N, E)由以下部分组成：

1. 一个非空有限集 N，由顶点或节点 n组成；

2. 一个指定子集 $E \subseteq N \times N$，由有向边组成，这些边是有序对 $e = (n, n')$，也可记作 $n \to n'$，其中 $n \neq n'$。

n_1、 n_2、.....、 n_ℓ这一 ℓ 节点序列在 N中是长度为 $\ell \in \mathbb{N}$ 的路径，当且仅当对于$k = 1$、2、.....、 $\ell - 1$， $n_k \to n_{k+1}$ 是 E 中的一条有向边。

图（N、E）是有根有向树，当且仅当存在唯一的初始节点 $n_0 \in N$（或称根节点、种子节点或入口点），使得对于图中的每个其他节点 $n \in N \backslash \{n_0\}$，都存在唯一的路径（$n_0$、$n_1$、$n_2$、.....、$n$），该路径从节点 n_0 开始并在节点 n结束。

定义 2。令 $T = (N, E)$ 表示任意一棵有限有根有向树。

1. 给定任意一对节点 $n, n' \in N$，当且仅当存在一条由 N 中长度为 ℓ 的节点组成、连接 $n_1 = n$ 与 $n_\ell = n'$ 的路径 $(n_1, n_2, \ldots, n_\ell)$ 时，称 $n' \in N$（最终）先于 n。令 $N_{>n}$ 表示所有先于 n 的节点 $n' \in N$ 的集合。

2. 对每个 $n \in N$，将初始节点为 n 的延续子树 $T_{\geq n} = (N_{\geq n}, E_{\geq n})$定义为满足以下条件的唯一一棵树：

- $N_{\geq n} = \{n\} \cup N_{>n}$；
- $E_{\geq n}$ 是将 E 中的边限制在 $E \cap (N_{\geq n} \times N_{\geq n})$ 上的结果。

3. 当且仅当有序对 (n, n') 或 $n \to n'$ 是 T 的一条有向边时，称任意其他节点 $n' \in N$ 直接后继于 n。

4. 将 n 的所有直接后继节点集合记为

$$N_{\geq n}^{+1} := \{n' \in N \mid (n, n') \in E\} \tag{52}$$

回顾第1.1节中 Anscombe 和 Aumann（1963）提出的轮盘彩票与赛马彩票之间的区别。以下定义以 Hammond（1988b年；1998a年、b；2022年）的定义为基础。

定义 3。给定非空后果域 Y 以及表示可能世界状态的非空有限集 S，将 $\mathcal{T}^S(Y)$ 定义为所有有限决策树的集合。这些决策树由一棵有根有向树 $T = (N, E)$ 和一个事件对应 $N \ni n \mapsto S_{\geq n} \in 2^S \setminus \{\emptyset\}$ 构成；该事件对应指定了每个延续子树 $T_{\geq n}$中可能的状态集合 $S_{\geq n}$。此外，决策树中的节点集合 N 被划分为四个两两不相交的子集：

1. N^d 决策节点 n 的集合，在这些节点处：（*i*）对于由(52)定义的直接后继集合中的某个节点 n'，行动者必须选择有向边 $n \to n'$；（*ii*）$S_{\geq n'} = S_{\geq n}$ 对于所有 $n' \in N_{\geq n}^{+1}$；

2. N^c 机会节点 n 的集合，在这些节点处：（*i*）从 n 出发的有向边 $n \to n'$ 由指定的轮盘彩票$N^{+1}_{\geq n} \ni n' \mapsto \pi(n'|n) \in (0,1]$；[20]（*ii*） $S_{\geq n'} = S_{\geq n}$ 对于所有 $n' \in N^{+1}_{\geq n}$；

3. 事件节点 N^e 的集合 n，其中从 n 发出的有向边 $n \to n'$ 由一个马彩票 确定，该彩票的结果将事件 $S_{\geq n}$ 划分为非空且两两不相交的子事件集合 $\{S_{\geq n'} \mid n' \in N^{+1}_{\geq n}\}$；对于每个 $n' \in N^{+1}_{\geq n}$，这些子事件都由可在子树 $T_{\geq n'}$中出现的状态$s \in S$ 组成；

4. 非空的终端节点 集合 N^t n，在这些节点处 $N^{+1}_{\geq n} = \emptyset$，因此没有边从其发出；每个节点都映射到一个安斯科姆–奥曼后果彩票 $\gamma(n) = \langle \gamma_s \rangle_{s \in S_{\geq n}} \in L^{S_{\geq n}}(Y)$，其结果 y属于指定的后果域 Y。

称决策树 $T \in \mathcal{T}^S$（Y）为:

- 确定性的当且仅当；
- 风险 仅在 $N^c \neq \emptyset$但 $N^e = \emptyset$；[21]
- 一棵 萨维奇树 仅在 $N^e \neq \emptyset$但 $N^c = \emptyset$;
- 一棵 安斯科姆–奥曼树 仅在 $N^c \neq \emptyset$且 $N^e \neq \emptyset$。

6.7 延续子树的评价

考虑正统的“未赋予生动性”的决策模型，它由任意有限决策树表示，该决策树属于树的定义域 T，其非空有限状态空间为 $\mathcal{T}^S$，安斯科姆–奥曼彩票结果属于定义域 Y。那么，第 1.1 节所述的贝叶斯理性论证意味着，和动态规划中通常的做法一样，可以在每棵决策树中使用一个向后递归的程序来计算评价 $v(T_{\geq n})$，即主观期望延续

[20]关于为什么如果存在一个机会节点 $n \in N^c$，以及一个节点 $n' \in N^{+1}_{\geq n}$，在该节点处 $\pi(n'|n) = 0$，那么第1.1节所述的贝叶斯理性的结果主义和等价的前理性论证都意味着所有结果彩票必然无差异，参见 Hammond（1988b）。[21]Raiffa（1968）关注的是具有金钱后果 的风险决策树，其中的收益以美元计量。

到达延续子树 $T_{\geq n}$的初始节点 $n \in N$ 的价值。本节余下部分将说明这一向后递归过程的细节。

对于任意有限决策树 T，反向递归从任意终端节点 $n \in N^t$ 开始。如第 6.6 节所述，到达树 T 的任意终端节点 $n \in N^t$ 时指定的结果，是 AA 结果彩票$\gamma(n) = \langle \gamma_s \rangle_{s \in S_{\geq n}} \in L^{S_{\geq n}}(Y)$。于是，子树 $T_{\geq n}$ 的评价 $v(T_{\geq n})$——该子树仅包含终端节点 n——就是 AA 结果彩票 $\gamma(n)$ 的期望效用，该彩票由公式 (45) 指定为以下双重求和

$$v(T_{\geq n}) = U^{S_{\geq n}}(\gamma(n)) = \sum\nolimits_{s \in S_{\geq n}} \mathbb{P}(s) \sum\nolimits_{y \in Y} \gamma_s(y)\, u(y) \tag{53}$$

对于任意非终端节点 $n \in N \setminus N^t$，从初始节点 n 开始的延续子树 $T_{\geq n}$ 的评价 $v(T_{\geq n})$，取决于集合$\{v(T_{\geq n'}) \mid n' \in N^{+1}_{\geq n}\}$，该集合包含所有从紧接在 n 之后的节点 $n' \in N^{+1}_{\geq n}$ 开始的延续子树的评价。要确定 $v(T_{\geq n})$ 的正确公式，需分别考虑三种情况，具体取决于 n 是机会节点、事件节点还是决策节点。

第一种情况是，当 n 为机会节点，且其直接后继节点$n' \in N^{+1}_{\geq n}$ 分别以给定的正概率 $\pi(n'|n)$发生时，相关递归显然为

$$v(T_{\geq n}) = \sum\nolimits_{n' \in N^{+1}_{\geq n}} \pi(n'|n)\, v(T_{\geq n'}) \tag{54}$$

第二种情况是，当 n 为事件节点，且其每个直接后继节点 $n' \in N^{+1}_{\geq n}$ 都确定了相关的单元格时。在这种情况下，式（54）中出现的、将事件 $S_{\geq n}$ 划分为两两不相交集合的划分 $\{S_{\geq n'} \mid n' \in N^{+1}_{\geq n}\}$的客观给定概率 $\pi(n'|n)$，需要替换为主观条件概率 $p(n'|n)$；这些概率是根据不同状态 $s \in S_{\geq n}$对应的相关主观概率 $\mathbb{P}(s)$推导得出的。由于我们要求对所有 $s \in S$，$\mathbb{P}(s) > 0$，因此这些条件概率$p(n'|n)$都有定义，并可计算如下：

$$p(n'|n) = \sum\nolimits_{s \in S_{\geq n'}} \mathbb{P}(s) \;/\; \sum\nolimits_{s \in S_{\geq n}} \mathbb{P}(s) \tag{55}$$

因此，当 n 是一个事件节点，其主观概率 $p(n'|n)$ 由式（55）给出，而不是具有假设概率或客观概率的机会节点时，

$\pi(n'|n)$，先前的公式（54）变为

$$v(T_{\geq n}) = \sum\nolimits_{n' \in N_{\geq n}^{+1}} p(n'|n)\, v(T_{\geq n'}) \tag{56}$$

在 n 为决策节点的第三种也是最后一种情形中，我们应用随机动态规划的标准最优性原理。这要求任何当前最优决策 $n^* \in N_{\geq n}^{+1}$ 都必须是实现最高可能期望值的第一步，而这一期望值取决于针对所有后续决策所制定的适当计划。考虑如下归纳假设：对于每个节点 $n' \in N_{\geq n}^{+1}$，值 $v(T_{\geq n'})$ 是行动者在 $T_{\geq n'}$ 的每个决策节点上都选择最优决策时所能获得的最大评价。当 n' 是终端节点时，这显然成立，因为在节点 n' 无须作出决策。如果该假设对每个节点 $n' \in N_{\geq n}^{+1}$ 都成立，那么节点 n 处的任何最优决策都必须是沿着一条边 $n \to n^*$ 移动到一个紧接着的节点 n^*，使其相对于 n' 的评价 $v(T_{\geq n'})$ 最大，其中须满足 $n' \in N_{\geq n}^{+1}$，且集合 $N_{\geq n}^{+1}$ 是有限的。换言之，必须满足

$$n^* \in \arg\max\nolimits_{n' \in N_{\geq n}^{+1}} v(T_{\geq n'}) \tag{57}$$

因此，当 n 是决策节点时，适当的递推式为

$$v(T_{\geq n}) = v(T_{\geq n^*}) = \max\nolimits_{n' \in N_{\geq n}^{+1}} v(T_{\geq n'}) \tag{58}$$

因此，方程 (53)、(54)、(56) 和 (58) 合在一起，确实能在四种不同情形下，通过向后递推确定 $v(T_{\geq n})$。

7 截断决策树

> “限制未来想象创造的第二个因素是：所考虑的日期越遥远，不确定性就越不受可能性因素的约束。”——Shackle（1969，第 224 页）

7.1 作为有限建模的有限理性

与进入决策树 $T \in \mathcal{T}^S(Y)$ 的任意延续子树 $T_{\geq n}$ 相关联的条件期望活化评价为 $v(T_{\geq n})$。原则上，这可以通过遵循反向递归的程序来计算

该程序已在第6.7节中阐述。然而，随着至少一些延续子树 $T_{\geq n}$ 变得更加复杂，通过反向递归执行计算所需的全部步骤显然会变得更具挑战性，甚至实际上难以实现。

如第1.2节所述，计算最优决策能力的这些实际限制，可能会迫使决策者转而采用某种¯有界模型，其形式为有界决策树 T。现在，只要有界树的每个终端节点都是原始树 $T \in \mathcal{T}^S$ 中的终端节点（Y），就有 $\bar{T} \in \mathcal{T}^S$（$Y$）。于是，可以将第6节中使用的动态¯规划论证应用于树 T，而非 T。

本节聚焦于更具挑战性的情况：原始决策树被截断，部分延续子树及其终端节点和对应结果被整体移除，由此形成有界模型。在国际象棋棋局中，玩家最多向前考虑 k 步时（其中 $k \in \mathbb{N}$ 较小），便会得到这种截断树。

7.2 割与截断节点

下面的定义阐明了这样一个观点：决策树的任何截断都发生在截断节点处。该节点具有这样的性质：所有严格后继节点都会被移除，因此截断节点便成为终端节点。

定义4。令 T 为定义域 $\mathcal{T}^S(Y)$ 中的任意决策树。如果存在一个由一个或多个割或非终端截断节点 $x \in N \setminus N^t$ 组成的集合 X，使得决策树 $\hat{} \hat{} \hat{} T = (N, E)$ 是 T 的截断，当且仅当：

1. 存在一组两两不相交的非初始节点删除集 $\{N_{>x} \mid x \in X\}$ $N_{>x} = N_{\geq x} \setminus \{x\}$ $n \in N \setminus \{n_0\}$，每个删除集都关联着一个延续子树 $T_{\geq x}$，其初始节点位于截断点 $x \in X$；$\hat{}$*2.* 截断后的剩余节点集 N 为 $N \setminus \cup_{x \in X} N_{>x}$；$\hat{}$ $\hat{}$*3.* 截断后的剩余有向边集 E 是原有向边集 $E \subset N \times N$ $n \to n'$ 在 T 中的限制，即 $(N \times \hat{} \hat{} \hat{} N) \cap E$ 至 $N \times N$；*4.* 每个截断点 $x \in X$ 都是截断树 T 的终端节点 $x \in \hat{} N^t$。

7.3 截断树中的主观评价

给定有限决策树 $T=(N, E)$，为了使其截断树 $\hat{T}=(\hat{N}, \hat{E})$ 成为一棵决策树，必须为 $\hat{T}$的每个终端节点 $n \in \hat{N}^t$指定一个结果 $\hat{\gamma}(n)$。若 $n \in \hat{N}^t \cap N^t$，即 T 的终端节点 n 恰好也是未截断树 T的终端节点，则 T 中的 n 的结果 $\hat{\gamma}(n)$应与其在 T中的结果相同，这意味着 $\hat{\gamma}(n)=\gamma(n)$。

如果节点 n 是截断点 $x \in \hat{N}^t \setminus N^t$，然而，节点 $n=x$处尚未指定任何结果 $\hat{\gamma}(n)=\hat{\gamma}(x)$；该节点已成为截断树 T $\hat{}$的终端节点，但并非原始未截断树 T中的终端节点。在无界理性的完整树 T中，这一菜单后果的适当理想评价 $v(T_{>x})$，以及节点 $n=x \in \hat{N}^t \setminus N^t$的评价 $v(T_{\geq n})$，都是按照第6.7节所述规则通过反向递归计算出的期望效用水平。然而，在截断树 T中，行动者能做的最好选择是采用一种主观评价 $\hat{\gamma}(n) \in \mathbb{R}$。这应当是对期望效用 $v(T_{\geq n})$的估计：如果行动者能够在完整的延续子树 $T_{\geq n}$中进行计算，且不受任何截断，那么通过反向递归本会得到这一期望效用。[22]

在本文余下部分，为了允许截断树，我们区分以下两类：

1. “结果”终端节点 n，每个节点都附有一个指定的 Anscombe–Aumann 彩票结果 $\gamma(n) \in L^{S_{\geq n}}(Y)$；2. “截断”节点 $x \in X$，每个节点都附有一个主观评价 $\hat{\gamma}(x) \in \mathbb{R}$ 。

7.4 蒙特卡洛树搜索与情景规划

早在电脑游戏成为流行消遣之前，数学家就将博弈视为决策模型。然而，由于博弈树搜索中的一些基本组成部分含义模糊，人们对决策的普遍理解一直受到阻碍。

[22]或许可以将附加到每个截断节点上的实值主观评价，替换为一种主观菜单后果；这种后果将在即将发表的 Hammond and Troccoli Moretti (2026) 一文中讨论——另见第10.3节。关于这一想法的进一步探讨，我们留待后续研究。

> 尤其是静态评估函数，也就是根据可直接检测的特征来判定某个节点优劣的方法，始终未能得到充分定义。期望结果模型认为，赋予某个节点的恰当值，应是从该节点开始随机进行棋局时，棋局结果的期望值。
>
> ——摘自 Bruce Abramson 的（1987年）博士论文，最终以 Abramson（1991年）的形式出版。

接下来，我们将讨论一种强大的计算机算法。这种算法在估计至少某些特定类型的截断博弈的规范性有用延续价值方面，已被证明极其有效。事实上，利用上文引文中所述的 Abramson 关于蒙特卡洛树搜索（MCTS）的关键思想，促成了整整一代计算机程序的诞生，这些程序：

1. 在国际象棋中，催生了 Stockfish 软件引擎；在足够长的一系列棋局中，它可以轻松击败任何人类玩家；
2. 但在围棋中，却无法击败最优秀的人类玩家。

在决策树中，任何相关版本的 MCTS 程序显然都需要评估每棵延续决策树：通过对一组“蒙特卡洛”模拟进行抽样，在这些模拟中，行动者在该延续的每个决策节点上都进行适当的随机化选择，从而估计规范化的期望效用。

这种评估无法被完全建模的延续子树的方法，是 Deep Fritz 的成功以及随后用于计算机国际象棋的 Stockfish 开源引擎的基础。然而，最终这一代算法被 AlphaZero 取代。AlphaZero 将 MCTS 和强化学习，与一种人工神经网络结合起来；这种网络已成为如今所谓“人工智能”的关键组成部分。事实证明，AlphaZero 这类算法在国际象棋方面的表现远胜 Stockfish 这一代程序，同时也终于能够击败围棋领域最优秀的人类玩家。不过，Silver等人（2018年）的重要论文指出，MCTS 仍然是 AlphaZero 的组成部分。[23]

一种适用范围广得多、并已在一些商业应用中证明有效的方法，涉及“情景规划”。这与

[23]另请参阅 https://en.wikipedia.org/wiki/AlphaZero，了解更多最新信息。

第2.2节所讨论的理性候选名单方法有些相似。它要求详细考虑一系列有限的有界不完整情景，这些情景经过筛选，旨在适当关注那些可能对最终政策选择产生重大影响的可能情况。这与一种可能很常见的观点形成对比：重点应放在最有可能发生的情况上。[24]

7.5 每个截断节点处的主观评价

为了给第7.3节所讨论的这类主观估计提供依据，我们扩展有限决策树的定义域 $\mathcal{T}^S(Y)$，其中结果属于原始定义域 Y，以便为每个截断点 $x \in X$指定一个主观评价 $\hat{\gamma}(x) \in \mathbb{R}$。经过这一修改，第6节中阐述的贝叶斯理性分析便可应用于扩展定义域 $\mathcal{T}^S(Y \cup \mathbb{R})$，其中决策树的结果属于扩展定义域 $Y \cup \mathbb{R}$。这一扩展分析表明，贝叶斯理性意味着最大化一个唯一的扩展归一化伯努利效用函数 $\hat{u}$ 在扩展定义域 $Y \cup \mathbb{R}$上定义。这个扩展效用函数 $\hat{u}$ 结合了先前构造的函数$Y \ni y \mapsto u(y) \in \mathbb{R}$ 的一个副本，以及一个补充函数；该补充函数为每个截断点 $x \in X$ 赋予效用值 $\hat{\gamma}(x) := \hat{u}(T_{>x}) = \hat{u}(T_{\geq x}) \in \mathbb{R}$，从而也为可建模树的有界定义域 $\mathcal{T}^S(Y \cup \mathbb{R})$ 中每个与之对应的延续子树 $T_{>x}$ 赋予效用值。

在每个截断节点上附加实数值的主观评价后，第6.7节对普通决策树中贝叶斯理性行为的描述便可扩展到典型的截断树 T T。此外，第1.1节中引用的、曾用于论证贝叶斯理性的结果，显然也适用于每个截断节点都赋有实数值主观评价的截断决策树。

[24]关于荷兰皇家壳牌集团前首席经济学家的观点及其情景规划实践经验，请参见 Jefferson（2012）。关于 Shackle（1953，1969）提出的潜在惊奇这一概念与壳牌所采用的情景规划之间可能存在的联系，另请参见 Jefferson（2014）和 Derbyshire（2017）。

8 赋活决策树

8.1 赋活审议过程

回顾第6节对有限决策树中贝叶斯理性行为的描述。在第7节中，这一描述被扩展到截断决策树，其中主观评价附加在任意截断节点上。纳入截断决策树，意味着可以采用有限理性的多步决策程序，其思路与第1.3节非正式讨论的内容一致。还要回顾的是，每个此类程序都会在几个步骤之间循环：作出决策后，先构建相关的延续决策树；随后转向一个新的决策模型，该模型可能会丰富这棵子树；最后，在这棵新树的初始决策节点上作出决策。当然，这一决策会让行动者进入新一轮循环。还要回顾第1.5节的启示：考虑活化决策树可能会给行动者带来潜在益处，因为这类树允许探索预期未来模型修订可能产生的影响。

因此，现在是时候将活化正式引入第6节和第7节讨论过的决策树（包括截断树）了。活化和截断都将被视为审议过程的一部分；在相邻的决策节点之间，该过程会修改行动者的有界决策模型。正如第7.5节所讨论的，这种模型的一般形式是截断决策树 $\hat{T} \in \mathcal{T}^S(Y \cup \mathbb{R})$。正如第9.1节将要讨论的，活化可能会影响后果域 Y 和不确定的世界状态的定义域 S。因此，从现在起，我们用 Y 和 S 分别表示适用于当前决策节点的当前定义域。

事实上，在行动者已承诺作出一项即将执行的决策之前，审议过程通常会立即暂停。随后，除非行动者的决策树模型中已无可行之步，否则暂停只是暂时的；描述行动者在任意截断决策树 T 中的决策程序的赋活审议过程，会在暂停后继续，直至行动者再次即将在之后的决策节点 $n' \in N$ 采取一步。之后的节点 n' 是一个修订后的有界决策模型的初始节点，其形式为延续子树 $\hat{T}_{\geq n'}$。

任何审慎决策过程——无论是否采用情境化方式——一个显而易见的关键环节，都包括剪除决策树模型中行动者知道永远不会到达的部分，因为这些部分已被行动者此前的决策排除，或者

行动者已了解到一个或多个彩票的结果后被排除。这个剪枝过程通常会释放行动者的一些建模资源。正如第1.3节所述，这些资源随后可以重新用于使先前有界模型中剩余的部分情境化，这通常会让行动者受益，

8.2 激活不完整决策模型

我们首先区分模型化激活与非模型化激活。请再回想第3节讨论过的荷马史诗神话《奥德修斯与塞壬》。那些天真的水手一旦进入塞壬歌声的听力范围，便遭遇了出乎意料、因而未被建模的激活事件；他们的骸骨散落在塞壬岛的草地上。那时，想要逃脱悲惨的命运已经为时已晚；水手们缺乏扬帆离岛的意志力。然而，在喀耳刻详尽而出色的建议下，奥德修斯得以构建一个更准确的塞壬力量激活模型，并思考自己如何既能听到塞壬的歌声，又能与船员一同活下来，继续航行。

我们在此讨论的活化现象介于模型化激活与非模型化激活这两种情况之间。一方面，对于完全没有建模的激活，我们无从置喙；天真的水手葬身塞壬岛便是如此。事实上，若没有至少一个关于激活的最低限度模型，行动者便缺乏能够容纳激活可能性的逻辑框架。另一方面，如果激活的可能性已得到全面且恰当的建模，那么这只会改变相关的决策树，也就无需再讨论激活。

因此，我们聚焦于一种中间情形：行动者意识到活化的可能性，甚至对其进行建模，但建模并不完整。在一种再典型不过的情形中，当一个能够充分处理任何潜在活化的理想决策树变得复杂到难以实际应用时，无论采用何种建模程序，最终都会得到这样一个不完整模型。因此，原本理想的决策树不得不截断，而且截断的方式当然无法得到令人满意的建模。这就留下了我们感兴趣的可能性：决策者愿意承认修改决策模型的可能性，因而会考虑如何活化原始决策树。

因此，我们在此探讨行动者的决策模型通常如何在一段时间内通过一系列调整经历激活过程，

正如奥德修斯受到喀耳刻指点时，向他揭示的那些逐步演变的模型所经历的那样。不过，在那个例子中，所有活化都发生在奥德修斯与喀耳刻相遇之后、他和船员再次扬帆启航之前。事实上，给定一棵一般的决策树，行动者可能会因为对模型所建议的决策之结果不满意而修正当前模型。在这方面，或许值得重申我们在第1.2节中所写的内容：我们强调，与 Simon 的满意化概念不同，后者是在给定的决策模型中寻找一个令人满意的决策；而本文所倡导的方法，则是寻找一个更优但仍有界的决策模型，其中包含一个能够带来更好结果的可行决策。

不过，更一般地说，当行动者做出决策树中所建模的实际决策时，活化过程会受到中断，甚至遭到扰乱。此后进行的任何修订都应针对行动者经过改进的新模型。这就是在行动者后续做出的任何步以及已经观察到的任何彩票结果的情况下，原延续决策树所剩下的部分。或者，如果行动者正在进行国际象棋之类的棋局，活化过程也可能因观察到对手的一步棋而受到中断和扰乱。

8.3 两种最小活化

考虑一个行动者，其决策程序一开始面对给定的有限截断决策树 T，我们称之为基树，其节点集合为 N，边集合为 E。我们假设 $\hat{T}$ 属于第6.6节定义3中规定的决策树定义域 $\mathcal{T}^S(Y \cup \mathbb{R})$，其中 Y 是后果域，S 是不确定的世界状态的定义域。我们所考虑的活化过程从这棵基树开始。随后，它将一系列连续活化组合起来；所有这些活化都是最小的，其含义是尽可能不影响树 T。

首先要注意，如果在树 T 的初始节点 $\hat{n}_0$ 之前发生了任何活化，那么行动者就有时间修正此前的任何决策，使其适用于增活树。因此，树 T 本可以、也应该已经重新定义，以纳入这次活化。当然，重新定义后得到的仍是未增活树。

第一种最小增活可称为*终端节点增活*，它发生在单个终端节点处。增活会将该

终端节点（无论它是否为截断节点）替换为一个决策节点、机会节点或事件节点，而该节点是一个新的有限延续子树的初始节点。

第二种最小增活，我们可以称之为边活化，要复杂得多。它以截断树 $T = (N, E)$ 的一条有向边 $n \to m$ 为基础，如图6所示； T的所有其他边，包括 n 和 m 这两个分别位于$n \to m$起点和终点的节点，都保持不变。当然，边 $n \to m$的初始节点 n 就是延续子树 $T_{\geq n}$ 的初始节点，而该延续子树的初始节点为 n。到达边 $n \to m$的终端节点 m 时，活化之前的行动者面临的延续决策树为 $T_{\geq m}$ ，其初始节点为 m。

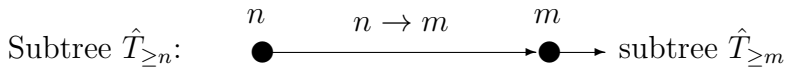

图6：截断树 $T = (N, E)$ 的有向边 $n \to m$

接下来，图7展示了由以下内容构成的最小增活对特定有向边 $n \to m$ 所产生的影响：

1. 首先，在 $n \to m$中插入额外的活化节点 e_m^- ；该节点是一个事件节点，其中，根据第6.2节的定义，马彩票的结果决定 $\{m, e_m^+\}$ 这一对节点中的哪个节点接续于 e_m^-之后；

2. 第二，添加一条额外活化边 $e_m^- \to e_m^+$ ，该边终止于额外的增活后节点 e_m^+ ，其为：

   - 除了 m 本身之外， e_m^-唯一的其他直接后继；
   - 该活化延续子树的初始节点

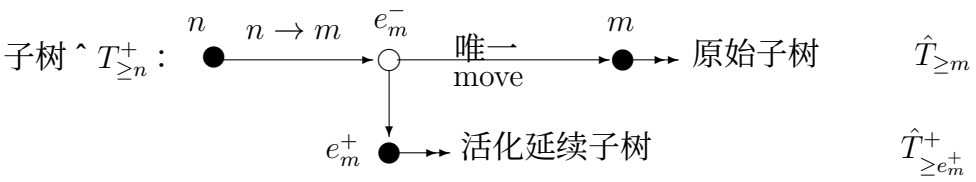

图 7：有向边 $n \to m$ 的一种最小增活

因此，发生在事件节点 e_m^- 处的彩票决定了从原始截断树 T 中的路径 $n \to e_m^- \to m$ 向替代路径$n \to e_m^- \to e_m^+$ 的不确定偏离是否发生。具体而言，如图 7 所示：

1. 首先，如果没有发生偏离，那么在$\hat{}$ $T_{\geq n}$ 的最小增活 $T^{+}_{\geq n}$ 中，相关的直接后继e^{-}_{m} 是从原始未增活树 T中的 $m\hat{}$发出的$\hat{}$原始未增活延续子树 $T_{\geq m}$ 的初始节点 m 。

2. 或者，如果确实发生了偏离，那么 e^{-}_{m} 的相关直接后继就是增活后节点 e^{+}_{m}。这是另一个有限延续子树 $\hat{}$ $T^{+}_{\geq e^{+}_{m}}$ 的初始节点；在将 $\hat{}$ $T_{\geq n}$ 增活为新树 $\hat{}$ $T^{+}_{\geq n}$ 的过程中，该子树会在 e^{+}_{m} 处附加到 $\hat{}T_{\geq n}$ 上。

8.4 增活的三个层次

第8.3节所述的最小增活仅属于第一层中最低的层次。第二层则来自基本增活：其中，T 的每个终端 $\hat{}$节点都可以进行一次终端节点增活，而 T 的每条边 $\hat{}$都可以进行一次边增活。根据决策树的定义，原始决策树以及每次基本增活所产生的新延续决策树，都具有有限集合的终端节点和边。因此，任何基本增活总会产生一棵有限的增活树。

然而，在第3节对荷马史诗中的例子的叙述中，大多数活化都不是基本活化。相反，喀耳刻向奥德修斯描述的决策树经历了多个连续阶段的活化，其中只有第一阶段涉及对基树的基本活化。事实上，在这一阶段之后，所有后续的活化都针对活化决策树中的某些部分，而这些部分仅是在先前的活化过程中才添加进去的。这表明还存在第三层次的递归增活：在激活过程中添加到基树中的树的部分，可以在之后的阶段本身也被活化。再举一个例子，任何曾以高于最基础水平下过国际象棋的人都知道，递归增活会影响玩家对所下棋局不断发展的理解，进而影响其对该棋局中任意局面的评估方式。

请注意，即使是最小增活，也至少会在原始树中增加一个节点。并且，根据定义，任何基本增活都是由有限个最小增活组合而成的。然而，递归增活可以无限进行，从而生成一棵含有无限多个节点的树。不过，本文仅讨论有限决策树。因此，我们的分析仅限于经过有限个阶段后结束的增活过程。

9 活化树中的贝叶斯理性

9.1 活化的结果域和状态域

在第 6 节和第 7 节中，我们考察了一个固定的后果域 Y 以及一个固定的世界状态不确定状态定义域 S 。随后，我们讨论了相应的有限决策树固定定义域 $\mathcal{T}^S(Y \cup \mathbb{R})$ 中的贝叶斯理性和主观评价，其中包括这些决策树的延续与截断 $\hat{}T$。

然而，正如第3节所讨论的，荷马笔下奥德修斯与塞壬的故事应当提醒我们，活化可能会扩展 Y 和 S中的一个或两者。毕竟，想想从图1中第一棵天真水手的树，到喀耳刻的最终模型（即图4中的第四棵树），经历活化的三个阶段后会发生什么。显而易见的不确定状态域扩展了，纳入了塞壬存在的可能性。显而易见的结果域也扩展了，纳入了奥德修斯和全体船员死在塞壬岛上的惨烈可能性，以及另一种远为理想的可能性：奥德修斯听到了塞壬的歌声，而他和船员们都成功脱身。

因此，我们允许活化过程的任一步骤扩展结果域 Y 和状态域 S中的一个或两者。然而，正如我们刚在第8.4节所论证的那样，将注意力限制在有限决策树上，就必然意味着我们的分析仅限于经过有限步后便告结束的活化过程。由此可知，任何获准的过程都必须以极限结果域 Y^+ 和极限状态域 S^+告终。

9.2 具有活化效用函数的理性

回想一下，在第6.5节中使用了两个不同的固定结果 y， $\bar{y} \in Y$，以构造满足式（51）的唯一归一化伯努利效用函数。请注意，活化决策树及其截断的新定义域 $\mathcal{T}^{S^+}(Y^+ \cup \mathbb{R})$ 必须包含对每个结果 $y \in Y^+$ 而言的风险型有限决策树集合 $\mathcal{T}(\Delta(\{y, \bar{y}, y\}))$；这些决策树不含任何事件节点、赛马彩票或不确定状态，且其轮盘结果彩票仅限于三种退化彩票 δ_y、 $\delta_{\bar{y}}$ 和 δ_y 的概率混合。然后，可以重复第6.5节所述的构造过程，以确定每个额外结果 $y^+ \in Y^+ \setminus Y$ 的效用 $u(y^+)$。所得结果是一个

归一化活化伯努利效用函数

$$Y^+ \ni y^+ \mapsto u^+(y^+) \in \mathbb{R} \tag{59}$$

它不仅对原始定义域 Y 中的结果 y 满足（51），也对整个扩展后的定义域 Y^+ 中的任何额外结果 y^+ 满足（51）。此外，由（59）定义的 u^+ 的期望，将在整个彩票域 $\Delta(Y^+)$ 上表示行动者的扩展偏好关系 $\succsim^+$。而且，由于早先归一化（51）中使用的两个特定结果 y，y 既属于 Y^+，也属于 Y，因此我们可以施加显而易见的对应归一化

$$u^+(\underline{y}) = \underline{u} \quad \text{and} \quad u^+(\overline{y}) = \overline{u} \tag{60}$$

条件。由此得到的函数（59）便会将 $Y \ni y \mapsto u(y) \to \mathbb{R}$ 扩展至扩大的定义域 Y^+。同样，通过考虑定义域 $\mathcal{T}^{S^+}(Y^+ \cup \mathbb{R})$ 而不是 $\mathcal{T}^S(Y^+ \cup \mathbb{R})$，便可将为每个状态 $s \in S$ 构造主观概率 $\mathbb{P}(s)$ 的方法，扩展为构造每个状态 $s \in S^+$ 的扩展主观概率 $\mathbb{P}^+(s)$。

通过构造活化后果域 Y^+ 和状态域 S^+，以及构造满足（59）和（60）的、定义在 Y^+ 上的归一化活化伯努利效用函数，并构造每个状态 $s \in S^+$的活化主观概率 $\mathbb{P}^+(s)$，第1.1节为属于定义域$\mathcal{T}^S(Y \cup \mathbb{R})$的决策树 T中的贝叶斯理性行为所提供的论证，显然也适用于属于定义域 $\mathcal{T}^{S^+}(Y^+ \cup \mathbb{R})$的活化决策树 T^+ 中的贝叶斯理性行为。

9.3 贝叶斯理性的局限

一种比单纯的贝叶斯理性在哲学上更为精细的理性概念，或许会要求理性主体在后果域上使用某种“规范上有正当依据”的伯努利效用函数，并对不确定的世界状态赋予某种“规范上有正当依据”的主观概率。[25] 因此，这种更丰富的理性概念将超越单纯的贝叶斯理性。

[25]正如一位审稿人善意指出的那样，本文并不适合开始讨论“规范上正当”这一短语究竟能够且应当意味着什么这一哲学问题。希望大致了解此类讨论应当深入到何种程度的读者，可以参阅Gibbard（1990年、2003年、2012年）、Broome（2013年、2021年）或Bratman（2018年）等权威著作。

活化树中的贝叶斯理性同样存在局限。事实上，完全理性应当要求行动者在终端评价节点上的主观评价具有规范上的正当依据，此外还要求行动者具备归一化伯努利效用函数和主观概率。当然，行动者对相关主观概率和主观评价的估计，完全可以通过蒙特卡洛树搜索或情景规划等程序加以改进，正如第7.4节所讨论的那样。

10 更多示例与扩展

10.1 预防原则与期权价值

欧盟委员会对预防原则给出了官方定义：[26]

> 预防原则是一种风险管理方法：如果某项政策或行动可能对公众或环境造成危害，且相关问题尚未达成科学共识，则不应实施该政策或行动。不过，随着更多科学信息的出现，可以重新审视该政策或行动。
>
> 预防原则这一概念最早见于欧盟委员会于2000年2月通过的一份通报，该通报对这一概念作出界定，并说明了其适用方式。
>
> 只有在存在潜在风险时，才可援引预防原则；不得以此为任意决策辩护。

不过，Sunstein（2003年）的结语指出了该原则可能存在的一些重大局限：[27]

> 我的目标是……解释预防原则为何具有看似令人费解的吸引力，并找出有助于使其付诸实施的策略。在个人层面上，这些策略

[26]https://eur-lex.europa.eu/EN/legal-content/glossary/precautionary-principle.html[27]参见 Sunstein（2003年），第1057–1058页。有关相关批评，参见 Sunstein（2005）。Steele（2006）至少部分回应了其中许多批评。

> 并非毫无道理，尤其是对于信息匮乏的人，或是尽力只关注当前情境某一方面的人而言。但对政府来说，预防原则并不合理，原因很简单：一旦拓宽视野，就会发现该原则根本无法提供任何指引。理性的风险监管体系当然会采取预防措施，但不会采纳预防原则。

例如，正如 Gollier 和 Treich（2003年）所讨论的那样，预防原则背后的决策理论原则似乎源自 Arrow 和 Fisher（1974年）关于环境保护与不可逆投资的研究。在第317页讨论一个两期模型时，Arrow 和 Fisher 写道：

> 鉴于人们能够从经验中学习，投资不足可以在第二期之前得到纠正，而错误的过度投资却无法挽回，其结果会一直持续下去。

他们论文的结语（第319页）如下：[28]

> 本质上，这一点是说，不可逆决策的预期收益应作调整，以反映该决策造成的选项损失。

Arrow 和 Fisher 使用“错误的过度投资”这一说法，清楚表明他们考虑的是这样一种情形：投资决策至多只能是有限理性的。这可能是因为，典型的决策程序很可能会忽略决策可能产生的一些结果，尤其是涉及环境的结果。这些结果即使并非完全不可逆，逆转起来也可能代价极高。因此，有人可能会主张采用一种决策程序，对这样一种不可能性有所考虑：我们永远无法完整地对所有可能相关的根本性结果建模。实现这一点的一种方式，是让正在分析的决策树更加生动。这样一来，只要切实可行，这棵树便能考虑到这种可能性：它所遗漏的至少一些重要结果，或许必须纳入之后修订的分析中。预防原则似乎是实现这一点的众多可能方式之一。

[28]Arrow 和 Fisher 还加了一条脚注，引用了 Hart（1942年）的研究；Arrow 早先曾多次称该研究引发了他对风险与不确定性经济学的兴趣。

10.2 转化性经历

变革性体验这一概念在哲学中的兴起，归功于 Paul（2014，2015a、b、c）；不过，Ullmann-Margalit 在其较早发表于 2006 年、探讨“重大决策”的论文中也讨论过类似观点。重大决策的首要特征（第 158 页）是，它们必须“具有变革性，或‘影响核心’”。正如 Paul（2015b，第 761 页）所写：

> 转化性体验是指……对行动者而言既极为新颖，又会以深刻而根本的方式改变她的经历；例如，成为父母、发现一种新的信仰、移居到一个新的国家，或投身战争……
>
> 一种认知层面上具有变革性的体验，是指能让你学到一些没有这种体验便无法学到的东西的体验。经历这种体验，会让你获得新的能力，去想象、识别并在认知上建模这类可能发生的未来体验。一种个人层面上具有变革性的体验，会以某种深刻且关乎个人根本的方式改变你，例如改变你的核心个人偏好，或改变你理解自身欲望以及自我认定为何种人的方式。因此，变革性体验是指兼具认知层面和个人层面变革性的体验。
>
> 变革性决策的主要问题在于，当我们无法认知获取可能结果的主观价值时，标准决策模型就会失效。

作为两个在哲学上颇具代表性的转化性经历案例，她考察了是否要成为吸血鬼（！）或是否要生孩子的决策。另见 Pettigrew（2015年、2020年）、Barnes（2015年）和 Campbell（2015年）的讨论，以及 Bykvist 和 Stef´ansson（2017年）、Randell（2023年），还有 Sunstein 和 Paul（2024年）等作者后来展开的讨论。

考虑任何一个决定行动者是否会经历变革性体验的决策。本文提出的主张是，可以用一棵活化决策树来建模这一决策，其潜在结果包括体验带来的结果。对于被归类为在认识层面具有变革性、但在个人层面不具有变革性的体验，可以通过允许一种活化来建模这种转变，而这种活化

只影响相关不确定的世界状态的定义域和／或这些状态的概率。

然而，被归类为个人转化性的体验，可能会以某种方式改变行动者的个人身份，而大多数哲学家通常认为，这会给规范性决策理论带来严重的概念挑战。Pettigrew（2020年）是少数显著尝试应对这些挑战的哲学家之一。他的研究值得在别处作更充分的讨论。在此我仅指出，一些经济学家可能希望将个人转化性体验视为只会导致一种以非常宽泛的方式理解的个人特征发生变化；这种个人特征可以在一个特定的个人特征定义域内变动，而该定义域涵盖例如个人身份的一切可能变化。参见 Fleurbaey 和 Hammond（2004年，尤其是第7节）所讨论的思想；他们试图在回应 Broome（1993年、1994年）等批评的同时，超越 Kolm（1972年、1994年、2002年）关于“基本偏好”的研究。[29] 一个值得探索的可能性是，可以利用活化决策树来探究相关个人特征定义域中未预见的变化所带来的影响，包括重大伦理决策的相关结果。

10.3 带时间的结果节点的决策树

在 Hammond and Troccoli Moretti（2026）中，我们考察了有限决策树。这些决策树除了包含决策节点、机会节点、事件节点和终端节点外，还包含带时间的结果节点。假设每个带时间的结果节点都有唯一的直接后继。

给定后果域 Y，对于每个 $m \in \mathbb{N}$，令 Y^m 表示 m 个 Y 的笛卡尔积。接着，令 $\mathbf{Y}(Y)$ 表示所有可能的带时间的结果流构成的定义域；对于某个 $m \in \mathbb{N}$，这些结果流的形式为 $\mathbf{y} = (t_j, y_j)_{j=1}^m \in \mathbb{R}^m \times Y^m$，其中时间序列 $(t_j)_{j=1}^m$ 满足 $t_1 < t_2 < \ldots < t_m$。通过这一构造，任何带有定时后果节点的确定性截断ˆ决策树 T，都可以化简为一个<合适>决策树 $T^* \in \mathcal{T}(\mathbf{Y}(Y))$，其中：

[29] Pivato（2025）是相关文献中一篇非常近期且数学上颇为复杂的贡献。

1. 在每个终端节点 $n \in N^t$ 处（该节点是一个结果节点），将指定的终端结果 $\gamma(n) \in Y$ 替换为唯一一个以节点 n 结束的带时间的结果流 $\mathbf{y}(n) \in \mathbf{Y}(Y)$；2. 随后删除每个定时后果节点，并将紧邻被删除节点之前和之后的两条有向边合并为一条有向边。

除了确定性决策树之外，对于所有包含机会节点和/或事件节点的决策树，也可以进行类似的简化。

在包含定时后果节点的决策树中，到达终端节点 n 所产生的最终结果，并非节点 n处单个结果的一个彩票，而是沿着决策树中通向节点 n的唯一路径逐步累积的结果流和后果彩票所构成的一个彩票。除非到达任何结果节点 n所产生的结果包含一种"菜单后果"，而这种后果取决于从节点 n出发的延续决策树 $T_{\geq n}$ 中有哪些后果彩票可行，否则允许定时后果节点所带来的影响便是常规的。即便如此，第1.1节概述的、能够证明贝叶斯理性的早期结果仍然成立；为了涵盖每一种包含菜单后果的结果流，唯一需要改变的是相关结果流的定义域变得丰富得多。

不过，这些变化可能影响重大。例如，它们使偏好灵活性成为可能，如 Koopmans（1950年、1964年）和 Kreps（1979卷）所述。它们也使人们能够偏好不确定性的消解时机，如 Kreps and Porteus（1978）所述。其他可能性还包括反映对彩票偏好的菜单后果，例如 Machina（1989）以及 Gilboa and Schmeidler（1989）所研究的那类偏好；这类偏好无法用定义域为带时间的结果流的期望效用函数来表示。此外，菜单后果还使跨期消费流的递归偏好成为可能，这类偏好最早由 Epstein and Zin（1989）加以研究。

事实上，对于包含定时后果节点以及决策节点、机会节点和事件节点的树，第6.7节的四个方程（53）、（54）、（56）和（58）中所列规则显然可以扩展。此扩展处理节点 n 为结果节点的情形；在这种情况下，附加于节点 n 的带时间的结果会被置于延续子树 $T_{\geq n}$ 中产生的每个结果流 $\hat{}\mathbf{y}$ 的开头。通过此扩展，

第6节中关于活化决策树的贝叶斯理性的描述仍然成立。

然而，在描述充实但经过截断的决策树中的主观评价贝叶斯理性时，确实会出现一个小小的复杂之处。问题在于，如何截断沿决策树中的一条路径累积起来的、非平凡的结果流和结果彩票。按照第7节的思路，显而易见的解决办法是：为每个截断点或截断节点 x 附上一个实值的综合主观评价，涵盖截断前累积的结果流，同时也对被截去的、未建模的延续决策树作出补偿。两者结合起来，应该能够估算出：如果能够完整执行第7.3节所述的反向递归，将会得到怎样的主观评价。在经济学应用中，原则上，这种方法可以适用于消费者和生产者在动态模型中共同面临的那类决策问题，例如希克斯（1946年）和迈尔森（1983年）等人所研究的问题。

10.4 逆贝叶斯主义

“逆贝叶斯主义”由 Karni 和 Vierø 在一系列合作论文中加以阐述（2013年、2015年、2017年）。对于本文所讨论的一般有限决策树，逆贝叶斯主义所表达的结果是：在任何活化决策树 T^+ 中，如果以活化未发生这一事件为条件，计算不同结果的概率，那么所得结果应当是原始未活化决策树 T 中相应的概率。更具体地说，原始主观概率 $\mathbb{P}(s)$ 可以通过考察以下条件概率来恢复：以 s 属于未活化状态空间S（而非活化状态空间 S^+）这一事件为条件，计算活化后的主观概率 $\mathbb{P}^+(s)$。

Steele 和 Stef´ansson（2021a、b）对逆贝叶斯主义的一些近期批评，基于 Bradley（2017年，第257页）对该原则的描述：

> ……当我们意识到一些与先前纳入考虑的前景不一致的前景时
> …… [逆贝叶斯主义要求] 我们应将关系态度扩展到新的集合，
> 同时保留所有先前的关系信念。

本文所考察的所有决策树，包括经过活化但遭到截断的决策树，都有一个关键特征：它们遵循 Hammond（1988a、b；1998a、b；1999、2022），所有机会节点或事件节点上附加的概率都为正。此外，截断采取第7节所述的特殊形式：移除延续子树，但保留其初始节点，使之成为截断树的终端节点。因此，Steele 和 Stef´ansson 所提出的、被视为逆贝叶斯主义例外的修正后信念，揭示了这样一种理论局限：只考察由决策树活化及允许的截断所导致的信念修正，而这些决策树上的所有概率都为正。

11卷 活化的有界理性

第1.2节开头引用了统计学家 George Box 那句广为流传的格言，它应当提醒我们：除最琐碎的决策问题外，任何形式模型都不可避免地会受到种种限制。从荷马的《奥德赛》到现代的国际象棋算法，跨越漫长时间的例子表明，对于有一个或多个决策需要作出的行动者而言，任何模型的效用都很可能只是暂时的。本文已开始探讨如何调整贝叶斯理性的概念，使其能够适用于这样一种行动者：他认识到活化带来的根本困难，并认真尝试加以应对，但这种尝试不可避免地并不完整，也不尽完善。

具体而言，本文认为，此类行动者能够采用的最佳决策程序是：（i）在到达任何决策节点之前，尽可能考虑可能发生的活化，并据此构建一个主观的截断概率模型；该模型以对每项可能决策最终事后评价的估计期望值为基础；（ii）在任何决策节点，作出使这一估计期望值最大化的决策；（iii）在每次后续决策之前，将第8节所讨论的适当活化应用于阶段（i）所用模型的相关后续部分，以识别自上次决策以来可能显现的模型缺陷。

致谢

本文所报告的研究最初于2007年4月至2010年3月期间获得玛丽·居里卓越讲席的支持，该讲席由欧盟委员会资助，合同编号为MEXC-CT-2006-041121。谨此感谢肯尼斯·阿罗、肯·宾莫尔和布克哈德·席珀在早期阶段提出的宝贵建议；感谢斯坦·梅特卡夫让我更加欣赏Shackle 的原创思想；也感谢马库斯·米勒和乔安妮·杨参与充满活力的讨论，同时声明他们对我的任何错误或遗漏概不负责。

也衷心感谢以下听众的耐心聆听与鼓励：斯坦福大学商学院/经济系理论研讨会（2007年3月）；伦敦政治经济学院偏好变化会议（2009年5月）；华威大学（2010年11月）；鲁汶天主教大学CORE（2011年9月）；波恩豪斯多夫数学中心（2013年8月）；米兰圣心天主教大学（2013年10月）；昆士兰大学无意识与博弈论会议（2014年2月）；布里斯托大学博弈论及其应用会议（2016年5月）；华威大学微观经济学进行中线上研讨会（2021年3月）；康斯坦茨大学“动态决策理论问题”研讨会（2023年7月）；巴黎索邦经济中心（2023年11月）；华沙经济学院及波兰科学院数学研究所巴拿赫中心举办的第九届赫维茨机制设计理论研讨会（2024年12月）；利物浦大学经济系研讨会（2025年4月）；最后还有2025年6月在曼彻斯特举行的RUD（风险、效用与决策）会议，以及2025年7月在意大利伊斯基亚岛举行的SAET（经济理论促进学会）会议。

近期，在与相关项目的合作者 Agust´ín Troccoli Moretti，以及我在华威大学的同事柯克·瑟吉纳和巴勃罗·贝克的讨论帮助下，本文当前版本得以历经漫长过程完成。最后但同样重要的是，这一经过大幅改进的版本源自较早的一篇 CRETA 工作论文，并在很大程度上得益于两位格外勤勉的匿名审稿人提出的宝贵意见。

参考文献

Abramson, B.（1991）《双人博弈的预期结果模型》（圣马特奥：Morgan Kaufman）。Anscombe, F.J. 和 R.J. Aumann（1963）“主观概率的定义”《数学统计年刊》34：199–205。Arrow, K.J. 和 A.C. Fisher（1974）“环境保护、不确定性与不可逆性”《经济学季刊》88（2）：312–319。Barnes, E.（2015）“当你不想期待时，你能期待什么”《哲学与现象学研究》91（3）：775–786。Bradley, R.（2017）《具有人性面孔的决策理论》（剑桥：剑桥大学出版社）。Bratman, M.E.（2018）《规划、时间与自我治理：实践理性论文集》（牛津：牛津大学出版社）。Broome J.（1993）“偏好的原因并非偏好的对象”《社会选择与福利》10：57–68。Broome, J.（1994）“答 Kolm”《社会选择与福利》11：199–201。Broome, J.（2013）《通过推理实现理性》（牛津：Wiley-Blackwell）。Broome, J.（2021）《规范性、理性与推理：论文选集》（牛津：牛津大学出版社）。Bykvist, K. 和 H.O. Stef´ansson（2017）“认识论转变与理性选择”《经济学与哲学》33(1)：125–138。Campbell, J.（2015）“L.A. Paul 的变革性体验”《哲学与现象学研究》91(3)：787–793。Dekel, E., B.L. Lipman 和 A. Rustichini（2001）“用唯一的主观状态空间表示偏好，”《计量经济学》69：891–934。Dekel, E., B.L. Lipman, A. Rustichini 和 T. Sarver（2007）“用唯一的主观状态空间表示偏好：勘误”《计量经济学》75：591–600。Derbyshire, J.（2017）“潜在惊讶理论作为情景规划的理论基础”《技术预测与社会变迁》124：77–87。

Dr`eze, J.H.（1962）《人类生命的社会效用》法国运筹学评论*23*：*93–118*，*Dreze, J.H.* 与 *A. Rustichini*（*2004*）《状态依赖效用与决策理论》，载于 *S. Barbera*、*P.J. Hammond* 和 *C. Seidl*（编）《效用理论手册》第*2*卷：拓展篇（多德雷赫特：Kluwer Academic）第16章，第839–892页。Earl, P.E. 与 B. Littleboy（2014）《G.L.S. Shackle》（英国贝辛斯托克：帕尔格雷夫·麦克米伦）。Elster, J.（1979）《尤利西斯与海妖：理性与非理性研究》（剑桥：剑桥大学出版社）。Epstein, L.G. 与 S.E. Zin（1989）《替代、风险厌恶与消费及资产收益的时间行为：理论框架》《计量经济学》57（4）：937–969。Fleurbaey, M. 与 P.J. Hammond（2004）《人际可比效用》，载于 S. Barber`a、P.J. Hammond 和 C. Seidl（编）《效用理论手册》第*2*卷：拓展篇（多德雷赫特：Kluwer Academic）第21章，第1181–1285页。Gibbard, A.（1990）《明智的选择，恰当的感受：规范性判断理论》（哈佛大学出版社）。Gibbard, A.（2003）《思考如何生活》（哈佛大学出版社）。Gibbard, A.（2012）《意义与规范性》（牛津大学出版社）。Gilboa, I. 与 D. Schmeidler（1989）《非唯一先验下的最大最小期望效用》《数学经济学杂志》18：141–153。Gollier, C. 与 N. Treich（2003）《科学不确定性下的决策：预防原则的经济学》《风险与不确定性杂志27：77–103。Grant, S. 与 J. Quiggin（2015a）《受意外影响的选择偏好模型》《理论与决策》79（1）：167–180。Grant, S.、J. Kline、P. O'Callaghan 与 J. Quiggin（2015b）《互动无意识的子模型》《理论与决策》79（4）：601–613。Halpern, J.Y. 与 L.C. Rˆego（2014）《可能存在无意识玩家的扩展式博弈》《数学社会科学》70：42–58。

Hammond, P.J.（1976）《偏好的变化与连贯的动态选择》《经济研究评论》43：159–173。Hammond, P.J.（1988a）《后果论与独立性公理》，载于 B.R. Munier（编）《风险、决策与理性》（第三届效用、风险与决策理论基础及应用国际会议论文集）（多德雷赫特：D. Reidel），第503–516页。Hammond, P.J.（1988b）《期望效用的后果主义基础》《理论与决策》25：25–78。Hammond, P.J.（1998a）《客观期望效用：一种后果主义视角》，载于 S. Barber`a、P.J. Hammond 和 C. Seidl（编）《效用理论手册》第*1*卷：原理（多德雷赫特：Kluwer Academic）第5章，第145–211页。Hammond, P.J.（1998b）《主观期望效用》，载于 S. Barber`a、P.J. Ham- mond 和 C. Seidl（编）《效用理论手册》第*1*卷：原理（多德雷赫特：Kluwer Academic）第6章，第213–271页。Hammond, P.J.（1999）《状态依赖后果域上的主观期望状态独立效用》，载于 M.J. Machina 和 B. Munier（编）《决策中的信念、互动与偏好（多德雷赫特：Klu- wer Academic），第7–21页。Hammond, P.J.（2007）《决策、博弈与经济行为建模中的熊彼特式创新》《经济思想史》XV：179–195。Hammond, P.J.（2022）《将前理性视为避免可预见的遗憾后-́果》《经济评论》73：943–976。Hammond, P.J. 与 A. Troccoli Moretti（2026）《带有定时和菜单式结果节点的风险决策树中的前理性》，华威大学 CRETA 工作论文（编写中）。Hansen, L.P. 与 T. Sargent（2007）《稳健性》（普林斯顿：普林斯顿大学出版社）。Hart, A. G.（1942）《风险、不确定性与概率复合的无利可图性》，载于 O. Lange、F. McIntyre 和 T. O. Yntema（编）《数理经济学与计量经济学研究》（芝加哥：芝加哥大学出版社）。

Hicks, J.R.（1946）《价值与资本：经济理论若干基本原理探究》（牛津：克拉伦登出版社）。Jefferson, M.（2012）“《壳牌情景》：20世纪70年代究竟发生了什么，以及对当前世界前景有何启示”《技术预测与社会变革》79（1）：186–197。Jefferson, M.（2014）“《时间的流逝》：Shackle、壳牌与情景”，载于Earl, P.和B. Littleboy（编）《*G.L.S. Shackle*》（贝辛斯托克：帕尔格雷夫·麦克米伦）第198–214页。Karni, E.（1985）《不确定性下的决策：状态依赖偏好的情形》（马萨诸塞州剑桥：哈佛大学出版社）。Karni, E.和M.-L. Vierø（2013）“‘逆贝叶斯主义’：一种基于选择的意识增长理论”《美国经济评论》103：2790–2810页。Karni, E.和M.-L. Vierø（2015）“概率精细化与逆贝叶斯主义”《风险与不确定性杂志》50：189–208页。Karni, E.和M.-L. Vierø（2017）“对无知的觉察：无知情境下决策的一种理论”《经济理论杂志》168：301–328页。Kolm, S.-C.（1972）《正义与 ´公平 ´法国国家科学研究中心。Kolm, S.-C.（1994）“‘基本偏好’的含义”《社会选择与福利》11：193–198页。Kolm, S.-C.（2002）《正义与公平》（麻省理工学院出版社）。Koopmans, T.C.（1950）“影响未来福祉的决策的效用分析”（摘要）《计量经济学》18：174–175页。Koopmans, T.C.（1964）“未来偏好的灵活性”，载于M.W. Shelly和G.L. Bryan（编）《人类判断与最优性》（纽约：约翰·威立），第13章，第243–254页。Kreps, D.M.（1990）《博弈论与经济建模》（牛津：克拉伦登出版社）。Kreps, D.M.（1992）“存在未预见偶发事件时的静态选择”，载于P. Dasgupta、D. Gale、O. Hart和E. Maskin（编）《市场与博弈的经济分析：献给*Frank Hahn*的论文集》（马萨诸塞州剑桥：麻省理工学院出版社），第258–281页。

Kreps, D.M.和E.L. Porteus（1978）“不确定性的时间解析与动态选择理论”《计量经济学》46（1）：185–200页。Lal, D.（1972）《水井与福利：马哈拉施特拉邦小型灌溉的探索性成本—收益研究》（巴黎：经合组织发展中心）。Little, I.M.D.和J.A. Mirrlees（1969）《工业项目分析手册》第二卷（巴黎：经合组织发展中心）。Little, I.M.D.和J.A. Mirrlees（1974）《项目评估与规划》（伦敦：海涅曼出版社）。Little, I.M.D.和J.A. Mirrlees（1991）“《项目评估与规划》二十年后”《世界银行*1990*年发展经济学年度会议论文集》第351–382页。Machina, M.（1989）“不确定条件下选择的动态一致性与非期望效用模型”《经济文献杂志》27：1622–1668页。Manzini, P.和M. Mariotti（2007）“序贯可理性化选择”《美国经济评论》97：1824–1839页。Marschak, J.（1950）“理性行为、不确定前景与可测效用”《计量经济学》18（2）：111–141页。Metcalfe, S.、S. Salles-Filho、L.T. Duarte、A. Bin、A.T. Azevedo和P.H.A. Feitosa（2021）“Shackle关于科学、技术与创新中优先事项设定和决策的方法”《未来》134：102838。Myerson, R.B.（1983）“具有耐用商品和适应性预期的动态微观经济模型”《经济行为与组织》4：309–351页。Paul, L.A.（2014）《变革性体验》（牛津大学出版社）。Paul, L.A.（2015a）“《当你满怀期待时无法期待什么》”《哲学研究》 [在线] 92（2）：1–22页。Paul, L.A.（2015b）“《变革性体验》的′概要《变革性体验》”《哲学与现象学研究》91（3）：760–764页。Paul, L.A.（2015c）“《变革性体验：对Pettigrew、Barnes和Campbell的回应》”《哲学与现象学研究》91（3）：794–813页。

Pettigrew, R.（2015）“变革性体验与决策理论”《哲学与现象学研究》91（3）：766–774。Pettigrew, R.（2020）《为变化中的自我做选择》（牛津大学出版社）。Pivato, M.（2025）“普遍递归偏好结构”，可见 SSRN：https://ssrn.com/abstract=5410051。Pollak, R.A.（1968）“一致性规划”《经济研究评论》35：201–208。Raiffa, H.（1968）《决策分析：不确定性下选择导论》（Addison-Wesley）。Randell, P.（2023）“熟悉的变革性体验”《综合》202：45。Savage, L.J.（1954，1972）《统计学基础》（纽约：约翰·威立；纽约：多佛出版社）。Schipper, B.C.（2014a）“无意识——文献与专题特刊简介”《数学社会科学》70：1–9。Schipper, B.C.（2014b）“基于偏好的无意识”《数学社会科学》70：34–41。Schumpeter, J.A.（1911） [2版。 1926] 《经济发展理论；一项关于 ¨企业家利润、资本、信贷、利息和经济周期的研究》（莱比锡，邓克尔与洪布洛特出版社）。Schumpeter, J.A.（1934） [1961]《经济发展理论：关于利润、资本、信贷、利息与商业周期的探究》由 R. Opie 从 Schumpeter（1911）译出，并由 J.E. Elliott 撰写新导言。Schervish, M.J.、T. Seidenfeld 和 J.B. Kadane（1990）“状态依赖效用”《美国统计学会会刊》85（411）：840–847。Seidenfeld, T.、M.J. Schervish 和 J.B. Kadane（2010）“不确定性下的连贯选择函数”《综合》172：157–176。Shackle, G.L.S.（1953）“《惊奇的逻辑》”《经济学》（新系列）20：112–117。

Shackle, G.L.S.（1969）《人类事务中的决策、秩序与时间》（剑桥：剑桥大学出版社）。Silver, D.等（2018）“一种通过自我对弈掌握国际象棋、日本将棋和围棋的通用强化学习算法”《科学》362（第6419期）：1140–1144。Simon, H.A.（1955）“理性选择的行为模型”《经济学季刊》69：99–118。Simon, H.A.（1957）《人的模型》（纽约：约翰·威立）。Sims, C.A.（2003）“理性忽视的含义”《货币经济学杂志》50（3）：665–690。Sims, C.A.（2011）“理性忽视与货币经济学”《货币经济学手册》，第*3*卷（North-Holland），第155–181页。Steele, K.（2006）“预防原则：公共决策的新方法？”《法律、概率与风险》5：19–31。Steele, K. 和 H.O. Stef´ansson（2021a）《超越不确定性：对未知可能性的推理》（剑桥：剑桥大学出版社）。Steele, K. 和 H.O. Stef´ansson（2021b）“意识增长中的信念修正”《心智》130（520）：1207–1232。Strotz, R.H.（1956）“动态效用最大化中的短视与不一致性”《经济研究评论》23：165–180。Sunstein, C.R.（2003）“超越预防原则”《宾夕法尼亚大学法律评论》151：1003–1058。Sunstein, C.R.（2005）《恐惧法则：超越预防原则》（剑桥大学出版社：剑桥）。Sunstein, C.R. 和 L.A. Paul（2024）“自由、变革性体验、法律与证言”，哈佛公法工作论文第24-13号。可见：http://dx.doi.org/10.2139/ssrn.4929159。Ullmann-Margalit, E.（2006）“重大决策：选择、转变与漂移”《皇家哲学研究所增刊》58：157–172。

Vierø, M.-L.（2009年）“安斯康姆–奥曼竞赛之后究竟发生了什么？——在模糊环境中表示偏好”《经济理论》41, 175–212。

Vierø, M.-L.（2021年）“意识增长的跨期模型”《经济理论杂志》197: 105351。