

标题 title

作者 *author*

2024 年 9 月 16 日

前言

目录

前言	i
第一部分 AI 的逻辑	1
第一章 合情推理	2
§1.1 命题逻辑的演绎推理	3
§1.2 合情推理的数学模型	11
§1.2.1 合情推理的基本假设，似然	13
§1.2.2 似然与概率	18
§1.2.3 先验与基率谬误	20
§1.3 合情推理的归纳强论证	22
§1.3.1 归纳强论证	22
§1.3.2 有效论证和归纳强论证的比较	27
§1.4 先验模型的存在性	32
§1.5 章末注记	34
§1.6 习题	35

第二章 Markov 链与模型	36
§2.1 Markov 链	37
§2.2 Markov 奖励过程 (MRP)	48
§2.3 Markov 决策过程 (MDP)	54
§2.4 隐 Markov 模型 (HMM)	63
§2.4.1 评估问题	66
§2.4.2 解释问题	68
§2.5 扩散模型	71
§2.5.1 采样逆向过程	75
§2.5.2 训练逆向过程	76
§2.6 章末注记	80
§2.7 习题	80
 第二部分 信息与数据	 81
第三章 熵与 Kullback-Leibler 散度	82
§3.1 熵	83
§3.1.1 概念的导出	83
§3.1.2 概念与性质	88
§3.2 Kullback-Leibler 散度	98
§3.2.1 定义	98
§3.2.2 两个关于信息的不等式	100
§3.3 编码理论	101
§3.3.1 熵与编码	101
§3.3.2 K-L 散度、交叉熵与编码	105

§3.4 在机器学习中的应用：语言生成模型	108
§3.5 附录：Shannon 定理的证明	110
§3.6 习题	113
§3.7 章末注记	115
第四章 高维几何，Johnson-Lindenstrauss 引理	117
§4.1 高维几何	119
§4.1.1 高维球体	119
§4.1.2 Stein 悖论	123
§4.1.3 为什么我们要正则化？远有潜龙，勿用	130
§4.2 集中不等式	131
§4.3 J-L 引理的陈述与证明	138
§4.4 J-L 引理的应用	143
§4.5 附录：Stein 悖论的证明	146
§4.6 习题	147
§4.7 章末注记	147
第五章 差分隐私	148
§5.1 数据隐私问题	149
§5.2 差分隐私的定义与性质	153
§5.3 差分隐私的应用	161
§5.3.1 随机反应算法	161
§5.3.2 全局灵敏度与 Laplace 机制	164
§5.3.3 DP 版本 Llyod 算法	168
§5.4 习题	171
§5.5 章末注记	171

第三部分 决策与优化 172

第六章 凸分析 173

§6.1 决策与优化的基本原理	174
§6.1.1 统计决策理论	174
§6.1.2 优化问题	177
§6.1.3 例子: 网格搜索算法	184
§6.2 凸函数	188
§6.3 凸集	194
§6.3.1 基本定义和性质	195
§6.3.2 分离超平面定理	199
§6.4 习题	202
§6.5 章末注记	202

第七章 对偶理论 203

§7.1 约束的几何意义	206
§7.2 条件极值与 Lagrange 乘子法	214
§7.3 Karush–Kuhn–Tucker 条件	218
§7.4 Lagrange 对偶	223
§7.4.1 原始规划与对偶规划	223
§7.4.2 对偶的几何意义	228
§7.4.3 弱对偶定理	230
§7.4.4 Slater 条件, 强对偶定理	231
§7.5 应用: 支持向量机 (SVM)	236
§7.6 习题	239
§7.7 章末注记	239

第八章 不动点理论	240
§8.1 Banach 不动点定理	241
§8.2 Brouwer 不动点定理	252
§8.3 习题	259
§8.4 章末注记	259

第四部分 逻辑与博弈 260

第九章 逻辑与博弈	261
§9.1 博弈的基本语言：以井字棋为例	263
§9.2 输赢博弈	265
§9.2.1 博弈的不同维度	265
§9.2.2 Zermelo 定理与 AlphaGo Zero	267
§9.3 正则形式博弈	274
§9.3.1 定义	275
§9.3.2 理性与均衡	277
§9.3.3 生成对抗网络	280
§9.3.4 混合策略	283
§9.4 随机博弈（Markov 博弈）	289
§9.5 习题	300
§9.6 章末注记	300

第五部分 认知逻辑 301

第十章 共同知识，Bayes 博弈，Aumann 知识算子	302
§10.1 “泥泞的孩童”谜题	305

§10.2 不完全信息博弈 (Bayes 博弈)	309
§10.3 电子邮件博弈	320
§10.4 Aumann 知识算子	325
§10.5 习题	325
§10.6 章末注记	326
第十一章 模态逻辑, 知识的逻辑	321
§11.1 知识逻辑的语言	321
§11.2 Kripke 语义与框架语义	321
§11.3 模态可定义性	321
§11.4 知识逻辑的基本模型与性质	321
§11.5 对不一致达成一致	321
第六部分 附录: 预备知识	322
附录 A 线性代数基础	323
§A.1 线性空间	323
§A.2 线性映射	330
§A.3 矩阵	336
§A.4 双线性型与二次型	345
§A.5 带内积的线性空间	351
§A.6 行列式	360
§A.7 算子范数与谱理论	364
附录 B 微分学基础	373
§B.1 点集拓扑	373

§B.1.1 度量空间, 范数	373
§B.1.2 开集与闭集	378
§B.1.3 紧致性, 收敛性, 完备性	382
§B.1.4 连续映射	386
§B.1.5 与实数序有关的性质	391
§B.2 一元函数的微分学	394
§B.2.1 导数与微分的定义	395
§B.2.2 微分学基本定理	400
§B.3 多元函数的微分学	402
§B.3.1 微分、偏导数与导数的定义	402
§B.3.2 微分学基本定理	412
§B.3.3 隐函数定理	415
附录 C 概率论基础	421
§C.1 从朴素概率论到公理化概率论	421
§C.1.1 Kolmogorov 概率论	421
§C.1.2 条件概率, 独立性	427
§C.2 随机变量, 分布函数	433
§C.2.1 基本定义	433
§C.2.2 离散型随机变量	439
§C.2.3 连续型随机变量	440
§C.2.4 随机向量, 条件分布, 独立性	445
§C.2.5 随机变量 (向量) 的函数	452
§C.3 随机变量的数字特征, 条件数学期望	456
§C.3.1 数学期望, Lebesgue 积分	456
§C.3.2 数学期望的性质	463

§C.3.3 随机变量的内积空间	467
§C.3.4 特征函数	470
§C.3.5 条件数学期望	472
§C.4 多元正态分布 (Gauss 向量)	479

第一部分

AI 的逻辑

第二部分

信息与数据

第三部分

决策与优化

第四部分

逻辑与博弈

第五部分

认知逻辑

第十章 共同知识，Bayes 博弈，Aumann 知识算子

真实世界是一个巨大的游戏，参与者之间的信息不对称是一个普遍存在的现象，这些不对称往往形成了“丛林法则”。“丛林法则”从来不是一个书面的规则。刚来美国旅游的人可能会非常担心自己被抢劫；然而，“盗亦有道”，绝大部分时候，抢劫犯只会抢二十美元，以便吃一顿饭。如果坏了规矩，反而会被惩罚。

我们可以考虑一个看似非常疯狂的想法：既然有丛林法则，何不把“一次抢劫最多二十美元”写入法律，这会有什么区别呢？抛开道德和法律的问题，这样做似乎不无道理。

1982 年，经济学家 Alvin E. Roth 和 J. Keith Murnighan 对讨价还价这一经典的市场博弈现象进行了同样的实验。实验中，有两个玩家，他们可能会收到特定价值的奖品，每个玩家的奖品价值可以不同。然而玩家并不总是能获得奖品，他有一个概率获得这一奖品。他们需要在规定的时间内就他们各自获得奖品的概率进行讨价还价。

具体来说，玩家其实在商讨如何分配一张 100% 概率的“彩票券”，

这个券决定了每个玩家赢得奖品的概率。例如，甲如果获得 40% 的彩票券，就有 40% 的概率赢得奖品，60% 的概率一无所获；而乙则完全相反，有 60% 的概率赢得奖品，40% 的概率一无所获。

然而，如果在规定时间内未能达成协议，那么所有玩家都将一无所获。因此，只有在双方就彩票券的分配达成了协议，并且该玩家在随后的抽奖中中奖时，玩家才能获得相应的奖品。否则，他将得不到任何奖励。我们将这种每个玩家只有两种可能金钱收益的游戏称为“二元彩票游戏”。

在实际的实验中，玩家甲的奖品价值 20 美元，玩家乙的奖品价值为 5 美元，谈判时间限制为 12 分钟。在实验开始前，实验人员会公开告知甲乙完整的游戏规则，包括奖品价值和谈判时间限制。实验人员还会公开告知甲乙关于他们私有信息的情况，例如，

“在游戏开始后，你们奖品的价值会告诉对方。”

“在游戏开始后，甲的奖品价值会告诉乙，但乙的不会被告诉甲。”

实验人员也可以选择告知这些信息。

实验开始后，实验人员会告知甲乙他们各自的奖品价值，并且选择性地告知甲（或乙）乙（或甲）的奖品价值，这些信息要和实验开始前的信息保持一致。然后，甲乙开始商讨如何分配彩票券。如果在规定时间内达成了协议，那么他们将按照协议分配彩票券。如果未能达成协议，那么他们将一无所获。

上面的实验设置中，在游戏开始前是否公开宣告双方的信息结构，和上面那个疯狂的想法是一样的。然而，在讨价还价的情景下，我们似乎不会觉得这有什么区别。然而，实验结果却显示，在信息不对称的情况下，双方信息结构的公开宣告会对博弈结果产生显著的影响。

谈判过程展现出了明显的策略性行为。例如，甲（20 美元的玩家）通常不会提及自己奖品的价值；但如果游戏开始前宣布过“乙（5 美元的玩家）不清楚甲的奖品价值”，20 美元的玩家往往会虚报自己的奖品。一个典型的例子是：“我知道你的奖品是 5 美元。我的只有 2 美元。所以我应该得到超过 50% 的份额。”另一方面，当乙知道甲的奖品价值时，往往会透露信息。这两种策略通常都不被对方相信。

更重要的是，实验结果显示，当实验人员在游戏开始前宣布双方的信息结构时，玩家变得更加没有策略性，博弈的结果很可能没有达到 Nash 均衡：当只有 20 美元的玩家知道两个奖品时，其总体平均收益（包括达成协议和未达成协议的情况）为 34.9，显著低于 5 美元玩家的相应收益 53.6。

另一方面，当双方的信息结构在游戏开始前不被公开宣告时，玩家表现出更多的策略性。在这种情况下，因为他们不能确定对手是否不知道自己的奖品，所以玩家无法像前一种情况那样自由地谎报自己的奖品价值。然而，因为玩家都不知道对方是否指导双方奖品的价值，所以，如果一个玩家知道两个奖品的价值，他完全可以假装自己只知道自己的奖品价值。由于更复杂的策略性行为，博弈的结果往往达到 Nash 均衡！

上面的故事告诉我们，信息结构（也就是我们知道什么）是否被公开宣告，对于博弈的结果有着重要的影响。实际上，这一问题是一个关于共同知识的问题。

在本章，我们将更系统地讨论共同知识，并定义 *Bayes* 博弈，它是我们研究博弈论中信息结构的基本语言。最后，我们将介绍 *Aumann* 知识算子，它是从 *Bayes* 博弈中起源的，一个关于“知识”的数学模型。

§10.1 “泥泞的孩童”谜题

我们先从一个经典的谜题开始。有 n 个孩子在玩泥巴，他们互相泼泥巴。母亲告诉孩子们，如果他们脸上沾上了泥巴，会受到严厉的惩罚。孩子们不能看到自己的脸，但是可以看到其他所有人的脸。所有孩子都希望保持自己的脸干净，但是弄脏别人的脸。

此时，孩子的父亲出现了，于是，孩子们停止泼泥巴。孩子们互相不说话。父亲看到了 k ($k \geq 1$) 个人脸上有泥巴，于是宣布：“你们至少有一个人脸上沾了泥巴。”之后，父亲会公开地问若干轮如下问题：“你们知道自己脸上有泥巴了吗？”孩子们回答“知道”或者“不知道”。

假设孩子们观察力敏锐、聪慧且诚实，并且每一轮他们都同时回答。接下来会发生什么？乍一看，似乎大家每次都会回答“不知道”，因为他们不能从这句话知道自己脸上有泥巴。但是，这个谜题远比想象的要复杂。读者可以先不看后文，先自己试一试。

假设有 k 个孩子脸上有泥巴。这个问题的谜底是：在前 $k - 1$ 轮中，所有孩子都会说“不知道”，在第 k 轮中，所有脸上有泥巴的孩子都会说“知道”。

这一结论的论证来源于对较小 k 的归纳总结：

- 当 $k = 1$ 时，脸上沾满泥巴的孩子看到其他人都没有泥巴。既然他知道至少有一个孩子的脸上有泥巴，他就能推出那个人肯定是他自己。
- 现在假设 $k = 2$ ，脸上沾满泥巴的孩子是 a 和 b 。一开始，因为他们分别看到了对方的脸上有泥巴，所以他们每个人都回答“不知道”。但是，当 b 回答“不知道”时， a 可以代入 b 的角色，意识到 b 看到了 a 脸上有泥巴。否则， b 在第一轮中就会知道泥巴在 b 的脸上，

并回答“知道”。因此， a 在第二轮回答“知道”。 b 也会通过同样的推理得出相同的结论。

- 现在假设 $k = 3$ ，脸上沾满泥巴的孩子分别是 a ， b 和 c 。孩子 a 的论证如下。假设我没有泥巴落在脸上。根据 $k = 2$ 的情况， b 和 c 在第二轮都会回答“是”。他们没有这样做，我意识到假设是错误的，我的脸上也有泥巴。因此在第三轮我会回答“知道”。 b 和 c 的论证也是类似的。

容易看出， $k = 3$ 的论证具有一般性，对一般的 k 也成立。

注。“泥泞的孩童”还有其他流行的陈述方式，比如“蓝眼睛红眼睛”。一个岛上有 100 个人，其中有 5 个红眼睛，95 个蓝眼睛。这个岛有三个奇怪的宗教规则。

1. 他们不能照镜子，不能看自己眼睛的颜色。
2. 他们不能告诉别人对方的眼睛是什么颜色。
3. 一旦有人知道了自己的眼睛是红色，他就必须在当天夜里自杀。

岛民不知道具体有几个红眼睛。

某天，有个旅行者到了这个岛上。由于不知道这里的规矩，所以他在和全岛人一起狂欢的时候，一不留神说了一句话：“你们这里有红眼睛的人。”假设这个岛上的人足够聪明，每个人都可以做出缜密的逻辑推理。请问这个岛上将会发生什么？

谜题就解到这里。然而，一个好的谜题，知道答案之后一定会带来更多的谜题。为什么谜题的答案是这样的呢？比如，如果 $k > 1$ ，那么所有人本来就都知道 p ：“至少有一个人脸上有泥巴”。那么父亲说这句话的意义是什么？

我们可以设想，如果父亲没有说 p ，会发生什么？容易发现，无论

父亲问多少轮，所有孩子都只会回答“不知道”！（见习题[\[hy: 出一下\]](#)）。因此，这里会产生极其反直觉的事实：即便大家都知道 p ，父亲说不说 p ，会导致完全不同的结果。为什么会这样？

我们重新审视这一问题的过程。

- 假设 $k = 2$ ，脸上沾满泥巴的孩子是 a 和 b 。在父亲宣布 p 之前， a 和 b 都知道 p 。然而，他们并不知道对方知道 p 。 a 可能会有两种想法：

- 我的脸上有泥巴，所以 b 知道 p 。
- 我的脸上没有泥巴， b 是唯一一个有泥巴的， b 看不到其他人脸上有泥巴，所以 b 不知道 p 。

当父亲宣布 p 之后， a 知道了 b 知道 p 。当第一轮 b 回答“不知道”之后， a 可以用“ b 知道 p ”这一事实排除第二种情况，从而推出自己脸上有泥巴。

- 假设 $k = 3$ ，脸上沾满泥巴的孩子是 a 、 b 和 c 。在父亲宣布 p 之前， a 、 b 和 c 不仅知道 p ，而且知道彼此知道 p 。比如说，以 a 的视角看， b 能看到 c 脸上有泥巴，所以 a 知道 b 知道 p 。

但是， a 不知道 b 知道 c 知道 p ，因为此时有两种情况：

- a 脸上有泥巴， b 能看到 c 和 a 脸上有泥巴，所以 b 知道 c 能看到 a 脸上有泥巴，从而知道 c 知道 p 。
- a 脸上没有泥巴， b 能看到 c 脸上有泥巴， a 脸上没有泥巴，但因为 b 不知道自己脸上有没有泥巴，所以 c 不一定知道 p ， b 不知道 c 知道 p 。

更一般地, a, b, c 都不知道所有人知道所有人知道 p ! 然而, 当父亲宣布 p 之后, a, b, c 都知道了所有人知道所有人知道 p .

用 $E^m p$ 表示所有人知道所有人知道……所有人知道 (m 次) p . 在一般情况下, 父亲没有宣布 p 之前, $E^k p$ 并不成立. 父亲宣布了 p 之后, 对任意 $m \geq 1$, $E^m p$ 都成立! 因此, 父亲宣布 p 带来了共同知识. 有了共同知识, 这一谜题就可以按照我们所讨论的方式进行下去.

我们曾经假设过所有人“观察力敏锐、聪慧且诚实”. 然而, 这一假设并不足够. 上面的论证其实暗含了, 所有人都知道所有人“观察力敏锐、聪慧且诚实”, 所有人都知道所有人都知道所有人“观察力敏锐、聪慧且诚实”, ……换言之, 我们需要假设“所有人观察力敏锐、聪慧且诚实”是共同知识.

如果没有这样的假设, 上面的论证都将不成立. 例如, 还是只有两个孩子 a, b 脸上有泥巴. 假如 a 不知道 b 是诚实的, 即便 b 回答了“不知道”, a 也无法从 b 的回答中得到任何额外的知识!

除了假设“所有人观察力敏锐、聪慧且诚实”是共同知识, 我们还需要假设以下陈述是共同知识:

- 每个人都能看到所有除自己外的人.
- 每个人都听到了父亲说的话.
- 父亲是诚实的.
- 每个人都在每一轮进行了充分的推理.
- ……

任何假设的破坏都会导致之前的讨论失效. 那么, 为什么父亲宣布 p 就可以让 p 变成共同知识呢?

所有人都听到父亲说 p 并不能产生共同知识. 假如父亲只是对每一个孩子单独宣布 p . 所有人并不知道所有人都知道 p , 因而仅仅可以做到 $E p$, 没有共同知识.

那么, 所有人都知道所有人听到父亲说 p 会如何呢? 进一步假设每个孩子给每一个孩子都安装了窃听器, 每个人都能够偷听每个人与父亲的谈话内容. 所有人并不知道所有人都知道所有人都知道 p , 因而仅仅有 $E^2 p$.

所以关键在于, 父亲宣布 p 的过程是公开的, 每个人都可以仔细观察别人有没有听到父亲说 p , 也可以观察到别人有没有观察到别人有没有听到父亲说 p , 等等. 此时对每一个 m 都有 $E^m p$.

“泥泞的孩童”谜题足以表明, 关于“知道”的讨论远比想象的复杂. 关于“知道”和知识的研究在哲学中划归为知识论. 我们将介绍处理知识的两种数学模型:

- 一种源自 Aumann, Harsanyi 和 Rubinstein 等人, 以 Bayes 概率论为基础, 是偏经济学的学术风格;
- 另一种源自 Kripke, Hintikka 和 Halpern 等人, 以模态逻辑为基础, 是偏计算机科学和哲学的学术风格。

在这一章, 我们主要讨论 Bayes 概率论的方法.

§10.2 不完全信息博弈 (Bayes 博弈)

接下来, 我们介绍讨论“知识”的博弈论语言. 我们先从正则形式博弈和 Nash 均衡开始说起. 正则形式博弈隐含了一个重要的假设: 所有玩家对整个世界有一致、完全的共同知识. Nash 均衡建立在这一假设

之上：每个玩家可以在博弈结束后，根据其他玩家的策略，确定自己的最优反应。

然而，现实世界中，玩家对世界的认识是有限的，不能获得完全的信息。比如，我们可能不知道对手的收益函数，然而这在现实中极其普遍。

另一个问题是，Nash 均衡只有假设混合策略的情况下才能保证存在。我们在现实中并不真的在选择混合策略：所有的交易其实都是“一锤子买卖”，绝对不可能有人说“我今天以 0.5 的概率花一块钱买你的苹果，0.5 的概率花 5 块钱买你的苹果”。英语也有一句谚语：

“Decision makers do not flip coins in the real world.”

相比之下，纯策略 Nash 均衡更加符合实际。然而，即便是纯策略 Nash 均衡也可能是不合理的状态。考虑如下的二人博弈：

$$\begin{pmatrix} 1,1 & 0,0 \\ 0,0 & 0,0 \end{pmatrix}.$$

显然，两个人玩家都选择第二策略就达到了纯策略 Nash 均衡。

然而，当行玩家对列玩家的选择有任意小的不确定性时，他都更倾向于选择第一个策略。因此，我们给出的这个纯策略 Nash 均衡实际上描述了一种不太可能出现的状态。

因此，一种 Nash 均衡的修正概念被提出：颤抖的手完美化，它指的是， s 是一个纯策略 Nash 均衡，并且当对手玩家的策略有任何微小不确定性的时候， s 中的策略依然是最优反应。

“颤抖的手”给了我们一个例子，说明对对手的不确定性会影响玩家的决策。因此，进一步的问题是，如何量化对对手的不确定性？Harsanyi

给出了一个现在已经是“标准答案”的解决方案：引入玩家的“类型”（可能世界）以及其他玩家的对此的先验的信念。他的采用了 Bayes 解释的概率论，信念在数学上被建模为对可能世界的概率分布。

我们先给出不完全信息博弈的定义：

定义 10.1 (不完全信息博弈, Bayes 博弈) 一个不完全信息博弈 (Bayes 博弈) 包含了以下组成部分：

- 玩家集合： I .
- 行动空间： $A = (A_i)_{i \in I}$, A_i 表示玩家 A_i 的所有可能行动.
- 类型空间： $\Theta = (\Theta_i)_{i \in I}$, Θ_i 表示玩家 i 的所有可能类型.
- 收益函数： $u_i : A \times \Theta \rightarrow \mathbb{R}$, 当所有人的行动和类型都确定的时候, 玩家 i 能拿到的收益.

所有玩家的行动 $a = (a_i)_{i \in I}$ 形成了一个行动组合, 所有玩家的类型 $\theta = (\theta_i)_{i \in I}$ 形成了一个类型组合. □

$P_i \in \Delta(\Theta_i)$ 是玩家 i 类型的概率分布. 比较不直观的一点是, P_i 表示了其他玩家对玩家 i 类型的信念. 因此, Bayes 博弈其实做了简化：

- 不论哪个玩家, 对特定玩家 i 的信念是一致的.
- 所有 P_i 是相互独立的, 因此玩家的类型之间不会有相互的关联.

因此, 玩家 i 在博弈中的全部不确定性都来自其他玩家的类型, 他对此的信念是

$$P_{-i} = \prod_{j \neq i} P_j.$$

最后, 我们假设玩家 i 知道自己的类型.

更进一步，在一般情况下，玩家 i 对这个世界的信念应该包含：

- 博弈中其他玩家都有谁，
- 自己的可能行动，
- 自己的类型，
- 自己的收益函数，
-

在一个真实的博弈中，以上信息都会是不确定的：对手可以藏在暗处，我们会不知道自己本来拥有的选择，我们可以不了解自己的性格，我们甚至也不知道自己究竟在追求什么。尽管有很多的不确定性最终都可以归结为“类型”（见后面正文的若干例子和习题[\[lby: 出一下\]](#)），Bayes 博弈依然是一个高度理想化的模型。

然而，Bayes 博弈的成功，正如 Robert Aumann 所说，不在于多么贴合实践，而是给了一种系统的方法，让我们可以建模不确定性和信念，从而理解这些概念在博弈中的作用：

“一个典型的例子是 Roth 和 Murnighan (1982) 在完全信息和不完全信息讨价还价方面的实验工作¹……他们将这此结果与早期 Fouraker 和 Siegel (1960) 的实验进行了比较。

“Fouraker 和 Siegel 进行了类似的实验，但由于缺乏 Harsanyi 的模型，只能将不完全信息的情况描述为双方都没有被告知对方的收益。

¹也就是本章开头讲的故事。

“然而，Roth 和 Murnighan 则从类型的角度详细阐述了不完全信息，并明确考虑了博弈的共同知识方面。”

接下来，我们看一个 Bayes 博弈的例子。

例 10.1 (“工作还是偷懒” 博弈) 在这个博弈中，有两个玩家，他们共同完成一个项目。两个玩家的行动都是“工作”(W) 或“偷懒”(S)。行玩家的类型集合是单点集，列玩家的类型是“勤奋”(D) 或“懒惰”(L)。收益矩阵为

	$\theta_2 = D,$			$\theta_2 = L,$	
	W	S		W	S
W	3, 3	-1, 0	W	1, 1	-1, 2
S	2, 1	0, 0	S	2, -1	0, 0

换言之，我们其实不确定的只有玩家 2 是喜欢偷懒还是努力工作，但他具体是什么人又会影响双方的收益，从而影响双方的决策。□

至此，我们已经建立了博弈的语言，下一任务就是定义一个玩家的策略。和第 9 章的博弈极为不同，Bayes 博弈中的玩家要面对不确定性。这给我们定义策略带来了一定的困难。为此，我们先要讨论清楚，究竟什么是“不确定性”。

设想如下的情景：你选了一门课，期末要考试。于是，你认真学习，并且把往年题 c 都找出来，认真做完，老师也划定了考纲和难度：你对这个考试胸有成竹。真正上场考试的时候，你的心理有一个对难度和考点的预期，因此，你在考试时候面对的是具有明确风险的不确定性。

然而，当你考的不是期末考，而是英语四级考试的时候，你可能就会变得非常随性：挂了还可以重新考，所以考前根本没有学习，甚至连题型都不知道有什么。这个时候，你甚至连四级总分是多少都不知道，

所以，你甚至连考试结果的预期都没有。这个时候，你面对的是模糊的不确定性。

还有一种情景，你现在不是考四级，而是考 GRE，这是上机考试。GRE 的每道题的难度都不一样，题目是随机出现的，难度分布极其不均匀，并且他会根据你答题情况来自动调整题目的难度。此时，你面对的是不稳定的不确定性。

上面讲到了三种不确定性。那么，Bayes 博弈是他们中的是哪一种呢？显然，在 Bayes 博弈中，玩家依然只能做一次行动，所以不是不稳定的不确定性。此外，Bayes 博弈中，所有玩家的类型集是固定的，甚至连类型的似然也是确定的，因此，这是具有明确风险的不确定性。

接下来，我们可以定义玩家的策略和理性了。注意，玩家知道自己的类型，但不知道其他人的类型，所以，玩家只能根据自己的类型来决定自己的行为。因此，我们应该定义玩家的策略如下：

定义 10.2 (策略) 玩家 i 的策略是一个映射

$$s_i : \Theta_i \rightarrow A_i,$$

其中 Θ_i 是玩家 i 的类型集合， A_i 是玩家 i 的行动集合， $s_i(\theta_i)$ 表示玩家 i 在类型 θ_i 下的行动。 □

注. 自然，我们也可以定义混合策略，此时， s_i 是一个 Θ_i 到 $\Delta(A_i)$ 的映射。不过，在 Bayes 博弈中，混合策略会让情况（不论概念上还是计算上）变得复杂，所以我们避免混合策略的讨论。

当面对具有明确风险（似然）不确定性的时候，我们可以遵循 von Neumann-Morgenstern 的期望效用理论，定义玩家的理性。首先，我们定义玩家的事中期望收益：

定义 10.3 (事中期望收益) 玩家 i 在类型 θ_i 下, 采取策略 s_i , 对手的策略是 s_{-i} 时, i 的事中期望收益为

$$\tilde{u}_i(s_i, \theta_i, s_{-i}) = \mathbb{E}_{\theta_{-i} \sim P_{-i}}[u_i(s_i(\theta_i), s_{-i}(\theta_{-i}), \theta_i, \theta_{-i})]. \quad \square$$

于是, 按照期望理论, 玩家的事中理性就是在知道对手的策略之后, 会最大化自己的事中期望收益。

据此, 我们就可以定义最优反应:

定义 10.4 (事中最优反应) 考虑玩家 i 的策略 s_i , 对手的策略是 s_{-i} , 如果对任意的其他策略 s'_i 和任意类型 θ_i , 都有

$$\tilde{u}_i(s_i, \theta_i, s_{-i}) \geq \tilde{u}_i(s'_i, \theta_i, s_{-i}),$$

那么 s_i 是玩家 i 的事中最优反应. \square

用最优反应, 我们可以很容易定义 Bayes Nash 均衡:

定义 10.5 (Bayes Nash 均衡, BNE) 考虑策略组合 $s = (s_i)_{i \in I}$, 如果对任意 i, θ_i, a_i 都有

$$\tilde{u}_i(s(\theta_i), \theta_i, s_{-i}) \geq \tilde{u}_i(a_i, \theta_i, s_{-i}),$$

那么 s 是一个 **Bayes Nash 均衡 (BNE)**. \square

接下里, 我们来看一个 BNE 的例子。

例 10.2 (猜硬币游戏) 考虑经典的猜硬币游戏 (见例 9.3), 这是一个正

则形式博弈，收益矩阵为

	H	T
H	$1, -1$	$-1, 1$
T	$-1, 1$	$1, -1$

如果两个人都出 H 的时候收益有独立微小的扰动，我们就得到了一个 Bayes 博弈：

	H	T
H	$1 + \epsilon\theta_1, -1 + \epsilon\theta_2$	$-1, 1$
T	$-1, 1$	$1, -1$

其中 $\theta_i \sim \mathcal{U}[-1, 1]$ 且相互独立，这就是玩家 i 的类型。

那么，这个博弈的 BNE 会是什么呢？我们可以猜一个。考虑策略：
 $s_i : [-1, 1] \rightarrow \{H, T\}$ 满足

$$s_i(\theta_i) = \begin{cases} H, & \theta_i \in [0, 1], \\ T, & \theta_i \in [-1, 0). \end{cases}$$

我们来验证，这的确是一个 BNE.

固定列玩家的策略 s_2 ，我们来计算行玩家的最优反应。对于行玩家来说，不论他的类型是什么，他面对的是一个 50% 的概率选择 H 和 T 的对手，因此，如果选择 H ，行玩家的期望收益是

$$\frac{1}{2} \cdot (1 + \epsilon\theta_1) + \frac{1}{2} \cdot (-1) = \frac{\epsilon\theta_1}{2}.$$

如果选择 T ，行玩家的期望收益是

$$\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (1) = 0.$$

因此，如果 $\theta_1 > 0$ ，行玩家应该选择 H ，如果 $\theta_1 < 0$ ，行玩家应该选择 T ，这是一个最优反应，恰好是我们猜测的策略。

对于列玩家的最优反应，我们可以做类似的计算，也可以得到， s_2 是列玩家的最优反应。因此， (s_1, s_2) 是一个 BNE。□

以上例子有极其特殊的含义：策略 (s_1, s_2) 导致的结果实际上是，每个玩家计算最优反应的时候，面对的对手其实仿佛是一个混合策略玩家，他以等概率选择 H 和 T 。注意，当 $\epsilon \rightarrow 0$ ，这个博弈收益矩阵回到了原始博弈。因此，Bayes 博弈里行动概率分布其实可以被视作原始博弈的混合策略。

猜硬币游戏的例子其实说明，正则形式博弈的混合策略 Nash 均衡被理解为：当不确定性趋于消失时候，BNE 形成的行动概率分布。这不是偶然的，实际上，所有的正则形式博弈的混合策略均衡都可以用一系列 Bayes 博弈的 BNE 纯化。

下面我们把猜硬币游戏的过程一般化。考虑一个正则形式博弈 (I, A, u) ，其中 I 是玩家集合， A 是行动空间， u 是收益函数。我们可以定义一个 Bayes 博弈，被称为扰动博弈：

- 给定一个扰动参数 $\epsilon > 0$ ，定义类型组合为 $\theta = (\theta_i)_{i \in I}$ ， $\theta_i \in [-1, 1]$ ，然后，将收益扰动为：

$$u'_i(s, \theta) = u_i(s) + \epsilon \theta_i.$$

- 假设 $\theta_i \sim F_i$ ，相互独立， F_i 是具有连续可微密度的分布。

定义一个从 Bayes 博弈策略到正则形式博弈混合策略的映射 ϕ ，给定一个 Bayes 博弈的策略组合 s ， $\phi(s)$ 是一个混合策略，满足

$$\phi(s)_i(a_i) = \Pr_{\theta_i \sim F_i} [s_i(\theta_i) = a_i].$$

换言之，在原本的 Bayes 博弈中，概率定义在了玩家的类型上，这个映射的作用是将这个概率转化为了玩家的行动上。

接下来，我们可以正式给出纯化定理的表述：

定理 10.1 (Harsanyi 纯化定理) 给定玩家集 I 和行动空间 A 。对于一般的收益函数 u 和连续分布族 $\{F_i\}_{i \in I}$ ，对任意完全信息正则形式博弈 (I, A, u) 的混合策略 Nash 均衡 σ ，存在一列扰动博弈纯策略 BNE s_ϵ ，当扰动参数 $\epsilon \rightarrow 0$ ， $\phi(s_\epsilon) \rightarrow \sigma$ 。

这一定理的证明比较长，并且需要用到较为复杂的数学技巧。由于证明本身与本章的讨论无关，所以这里略去。

这一定理给了 Nash 均衡（也即混合策略）一种新的解释：混合策略定义的 Nash 均衡可以被看作不确定性趋于消失的时候的（纯策略）BNE。

尽管我们说，“Decision makers do not flip coins in the real world.”然而，如果玩家对自己的收益有微小的不确定性，他的行为就会仿佛在抛硬币。这就是混合策略的似然解释。

注意，在 Bayes 博弈中，玩家对自己的类型是确定的，所以玩家在决策时不应该对自己的收益有微小的不确定，因而上面这一解释并不完全正确。然而，我们可以重新定义理性的概念，它会产生等价的 BNE 定义，但是玩家此时要面对自己类型的不确定性。

为了说明这一理性的概念，我们先看一个例子。假如你是一个顶尖斗地主玩家，知道自己拿到身份、拿到手牌之后要如何行动，这是你一贯的策略。当然，在游戏开始前，你其实并不知道自己的身份和手牌，这是你对自己的不确定性。不过，这并不影响你评估自己的胜率：你只需要知道其他玩家的水平，你就可以大概估计一个总体的胜率。

上面的例子里，玩家其实在博弈开始前就已经选好了一个类型到行动的策略，但是他要面临自己类型的不确定性。此时，我们计算的胜率其实是事前期望收益：

定义 10.6 (事前期望收益) 给定玩家集 I 、行动空间 A 和类型上的联合分布 P ，如果玩家 i 的策略是 s_i ，对手的策略是 s_{-i} ，那么 i 的事前期望收益为

$$\hat{u}_i(s_i, s_{-i}) = \mathbb{E}_{\theta \sim P}[u_i(s_i(\theta_i), s_{-i}(\theta_{-i}), \theta_i, \theta_{-i})]. \quad \square$$

在 Bayes 博弈中，事前理性指的就是，玩家在不知道自己的类型的情况下，最大化自己的事前期望收益。根据事前期望收益，我们可以定义事前最优反应：

定义 10.7 (事前最优反应) 考虑玩家 i 的策略 s_i ，对手的策略是 s_{-i} ，如果对任意的其他策略 s'_i 和任意类型 θ_i ，都有

$$\hat{u}_i(s_i, s_{-i}) \geq \hat{u}_i(s'_i, s_{-i}),$$

那么 s_i 是玩家 i 的事前最优反应。 □

注意，事前理性和事中理性考虑的都是同样的策略，但是玩家面对的不确定性是不同的。然而，事前理性和事中理性其实是等价的：

定理 10.2 给定玩家集 I 行动空间 A ，类型空间 Θ 、收益函数 u 和类型空间的联合分布 P ，考虑策略组合 $s = (s_i)_{i \in I}$ ，对任意玩家 i ， s_i 是 s_{-i} 的事前最优反应当且仅当 s_i 是 s_{-i} 的事中最优反应。

这一定理的证明类似正则形式博弈的无差别原理（定理 9.7）的证明，见习题 [hy: 出一下]。

这一定理其实有一些反直觉：在 Bayes 博弈中，如果玩家面对的仅仅只是具有确定风险的不确定性，那么他是否不确定自己的类型并不会影响他的理性决策。这其实是因为，玩家面对的仅仅是具有确定风险的不确定性，而不是更加复杂的不确定性，因此，无论博弈如何进行，他在开始前就已经可以确定自己的最优策略。

这一定理的直接推论是，我们可以根据事前最优反应来定义 BNE：

$$\hat{u}_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i})$$

对任意 i 和任意策略 s'_i 成立。

当所有的不确定性都消失的时候，我们得到的收益是真实的，被称为事后收益，此时不再有任何的概率，因此博弈退化为了正则形式博弈。事前、事中、事后分别表明了信息的确定（披露）程度。

§10.3 电子邮件博弈

作为一个例子，接下来我们使用 Bayes 博弈来研究知识的性质。这个例子由 Ariel Rubinstein 给出，它说明在二人正则形式博弈中，共同知识对到底实现哪一个 Nash 均衡非常关键。

考虑两个玩家和两个可能的收益矩阵：

	A	B		A	B
A	(0,0)	(-10,1)	A	(8,8)	(-10,1)
B	(1,-10)	(8,8)	B	(1,-10)	(0,0)

在左边的矩阵中, (B, B) 是唯一的 Nash 均衡. 在右边的矩阵中有多个 Nash 均衡: (A, A) 和 (B, B) . (A, A) 给出比 (B, B) 更高的收益, 但行动 A 比 B 更有风险.

左边矩阵是真实矩阵概率是 $p > 1/2$. 玩家 1 知道真实的矩阵, 而玩家 2 不知道. 如果选择了右边矩阵, 玩家 1 会给玩家 2 发送一条消息. 如果玩家 2 收到了消息, 她会回复. 如果玩家 1 收到了回复, 她会发送第二条消息来确认她收到了玩家 2 的回复. 以此类推. 每条消息都以 ϵ 的概率独立等可能丢失.²

以上传信的过程可以用 Bayes 博弈的类型来刻画. 具体来说, 两个玩家的类型集合为 $\Theta_i = \{\theta_i^0, \theta_i^1, \theta_i^2, \dots\}$. θ_i^m 表示玩家 i 发了 m 封邮件. θ_i^m 有直观的含义. 例如, 类型 θ_1^0 表示真实收益矩阵是左边的, 而类型 θ_1^1 表示真实收益矩阵是右边的, 1 发送了一封电子邮件, 但 2 没有收到.

实际上, θ 包含了所有可能的情况:

- (θ_1^0, θ_2^0) : 真实收益矩阵是左边的.
- (θ_1^1, θ_2^0) : 真实收益矩阵是右边的, 1 发送了一封电子邮件, 但 2 没有收到.
- (θ_1^1, θ_2^1) : 真实收益矩阵是右边的, 2 收到了第一封电子邮件, 但 1 没有收到 2 的回复.
-

²注意: 发送电子邮件不是一个行动, 而是一个规则.

我们可以算出来, 当真实矩阵为左边矩阵时, 每个类型出现的概率. 首先, 以概率 p 选择左边的矩阵, 而且没有人发送消息. 因此, (θ_1^0, θ_2^0) 的概率是 p , 其他项概率都是 0. 我们得到如下的概率表:

左	θ_2^0	θ_2^1	θ_2^2	\dots
θ_1^0	p	0	0	\dots
θ_1^1	0	0	0	\dots
θ_1^2	0	0	0	\dots
\vdots	\vdots	\vdots	\vdots	\ddots

同样可以算出来, 当真实矩阵为右边矩阵时, 每个类型出现的概率. 首先, 以概率 $1 - p$ 选择右边的矩阵, 玩家 1 发送一条消息, 它会以概率 ϵ 丢失. 因此, (θ_1^1, θ_2^0) 的概率是 $\epsilon(1 - p)$. 以此类推, 可以得到计算. 我们得到如下的概率表:

右	θ_2^0	θ_2^1	θ_2^2	\dots
θ_1^0	0	0	0	\dots
θ_1^1	$\epsilon(1 - p)$	$\epsilon(1 - \epsilon)(1 - p)$	0	\dots
θ_1^2	0	$\epsilon(1 - \epsilon)^2(1 - p)$	$\epsilon(1 - \epsilon)^3(1 - p)$	\dots
θ_1^3	0	0	$\epsilon(1 - \epsilon)^4(1 - p)$	\dots
\vdots	\vdots	\vdots	\vdots	\ddots

容易看出来, 当真实类型为 θ_t^m 时, 收益矩阵是到第 m 层的共同知识, 即 E^m . 所以对于很大的 m , 收益矩阵是“几乎公共知识”. 所以, 这个模型在研究的问题是: 如果收益矩阵是“几乎公共知识”, 那么 Nash 均衡是什么? 为此, 我们需要求出来这个 Bayes 博弈的 BNE.

我们需要弄清楚对每个类型 θ_i^m ，玩家会做什么。假设玩家 1 的类型为 θ_1^0 。玩家 1 知道 (θ_1^0, θ_2^0) 是真实的类型，所以左边的矩阵被选择。据此推理：玩家 1 选择占优策略 B 。

假设玩家 2 的类型为 θ_2^0 。我们对玩家 1 的所有可能类型分类讨论：

- 如果玩家 1 的类型为 θ_1^0 ，那么左边的矩阵被选择，对于玩家 2 来说，这种情况的概率为

$$\Pr(\theta_1^0 | \theta_2^0) = \frac{p}{p + \epsilon(1 - p)} := \mu_2^0.$$

- 如果玩家 1 的类型为 θ_1^1 ，那么右边的矩阵被选择，对于玩家 2 来说，这种情况的概率为

$$\Pr(\theta_1^1 | \theta_2^0) = 1 - \mu_2^0.$$

现在我们来分析玩家 2 的两种选择： A 和 B 。

- 选择 B 的期望收益至少是 $8\mu_2^0$ 。推理如下：
 - 玩家 1 的类型是 θ_1^0 时，这是左边的矩阵，玩家 1 肯定选择 B ，此时玩家 2 选择 B 的收益是 8。
 - 玩家 1 的类型是 θ_1^1 时，这是右边的矩阵，无论玩家 1 怎么选，此时玩家 2 选择 B 的收益至少是 0。

因此，按照全概率公式计算， B 的期望收益至少是 $8\mu_2^0$ 。

- 选择 A 的期望收益至多是 $-10\mu_2^0 + 8(1 - \mu_2^0)$ 。推理如下：
 - 玩家 1 的类型是 θ_1^0 时，这是左边的矩阵，玩家 1 肯定选择 B ，此时玩家 2 选择 A 的收益是 -10 。

- 玩家 1 的类型是 θ_1^1 时, 这是右边的矩阵, 无论玩家 1 怎么选, 此时玩家 2 选择 A 的收益至多是 8.

因此, 按照全概率公式计算, A 的期望收益至多是 $-10\mu_2^0 + 8(1 - \mu_2^0)$.

注意,

$$\begin{aligned}
 & 8\mu_2^0 - (-10\mu_2^0 + 8(1 - \mu_2^0)) \\
 &= 10\mu_2^0 - 8 \\
 &= \frac{10p - 8(p + \epsilon(1 - p))}{p + \epsilon(1 - p)} \\
 &= \frac{(2 + \epsilon)p - 8\epsilon}{p + \epsilon(1 - p)} \\
 &> \frac{1 - 8\epsilon}{p + \epsilon(1 - p)}.
 \end{aligned}$$

对充分小的 $\epsilon > 0$, 这个值是正的. 因此, B 更好.

假设玩家 1 的类型为 θ_1^1 , 于是, 右边的矩阵被选择. 同样, 对玩家 2 的类型分类:

- 如果玩家 2 的类型为 θ_2^0 , 对于玩家 1 来说, 这种情况的概率为

$$\Pr(\theta_2^0 | \theta_1^1) = \frac{\epsilon(1 - p)}{\epsilon(1 - p) + \epsilon(1 - \epsilon)(1 - p)} := \mu_2^1.$$

- 如果玩家 2 的类型为 θ_2^1 , 对于玩家 1 来说, 这种情况的概率为

$$\Pr(\theta_2^1 | \theta_1^1) = 1 - \mu_2^1.$$

同样可以计算玩家 1 的两种选择的期望收益：

- 选择 B 的期望收益至少为 0. 推理如下：玩家 2 是类型 θ_2^0 时肯定选择 B （上面已经证明），因此最坏的情况是玩家一类型为 θ_2^1 .
- 选择 A 的期望收益至多为 $-10\mu_2^1 + 8(1 - \mu_2^1)$. 推理如下：玩家 2 是类型 θ_2^0 时肯定选择 B （上面已经证明），因此最好的情况是玩家一类型为 θ_2^1 .

综合两方面， B 更好，因为对于所有 ϵ , $\mu_1^1 > 1/2$.

逐步迭代上述过程，我们发现，在唯一的 BNE 中，所有玩家在所有类型下都选择 B .

然而，如果右边的矩阵是共同知识， (A, A) 也是一个真正的 Nash 均衡，然而，因为对于收益矩阵的不确定性，即便收益矩阵是“几乎共同知识”，这个均衡也不会被实现！这个例子说明了，共同知识对于均衡的实现是非常关键的. 我们在习题[thy: 出一下]中会进一步讨论 Nash 均衡与共同知识的关系.

§10.4 Aumann 知识算子

§10.5 习题

关于均衡的进一步思考

- 用 $Nash(x)$ 表示“ x 是 Nash 均衡”，那么 $\exists x C(Nash(x))$ 和 $C(\exists x C(Nash(x)))$ 的含义是否一样？

- 如果不引入不确定性，在完全信息下，实现特定的 Nash 均衡是否还需要共同知识？
- 如果玩家不是逻辑全知的，或者说她的推理、计算能力是有限的，那么 Nash 均衡是否还会达到？是否可接近？

§10.6 章末注记

第六部分

附录：预备知识

参考文献

- [Bre57] Leo Breiman. The Individual Ergodic Theorem of Information Theory. *The Annals of Mathematical Statistics*, 28(3):809–811, 1957.
- [CT12] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2012.
- [Huf52] David A. Huffman. A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the IRE*, 40(9):1098–1101, September 1952.
- [Inf] Information | Etymology, origin and meaning of information by etymonline. <https://www.etymonline.com/word/information>. (accessed 2023-07-10).
- [Jay02] Edwin T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2002.

- [KL51] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- [LLG⁺19] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, October 2019.
- [McM53] Brockway McMillan. The Basic Theorems of Information Theory. *The Annals of Mathematical Statistics*, 24(2):196–219, June 1953.
- [RHW86] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, pages 318–362. MIT Press, Cambridge, MA, USA, January 1986.
- [Rob49] Robert M. Fano. *The Transmission of Information*. March 1949.
- [Sha48] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948.
- [Shi96] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer, New York, NY, 1996.

- [Uff22] Jos Uffink. Boltzmann's Work in Statistical Physics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2022 edition, 2022.
- [李 10] 李贤平. 概率论基础. 高等教育出版社, 2010.