博弈论教程

作者： M. J. Osborne (Canada)
A. Rubinstein (USA)
译者：魏玉根

出版社：中国社会科学出版社

出版日期：2000 年 4 月第 1 版 页数：322

前　言

本书提供了博弈论的一些主要思想。它是作为一个学期的研究生课程的教材来设计的，该课程约需 42 学时的面授。

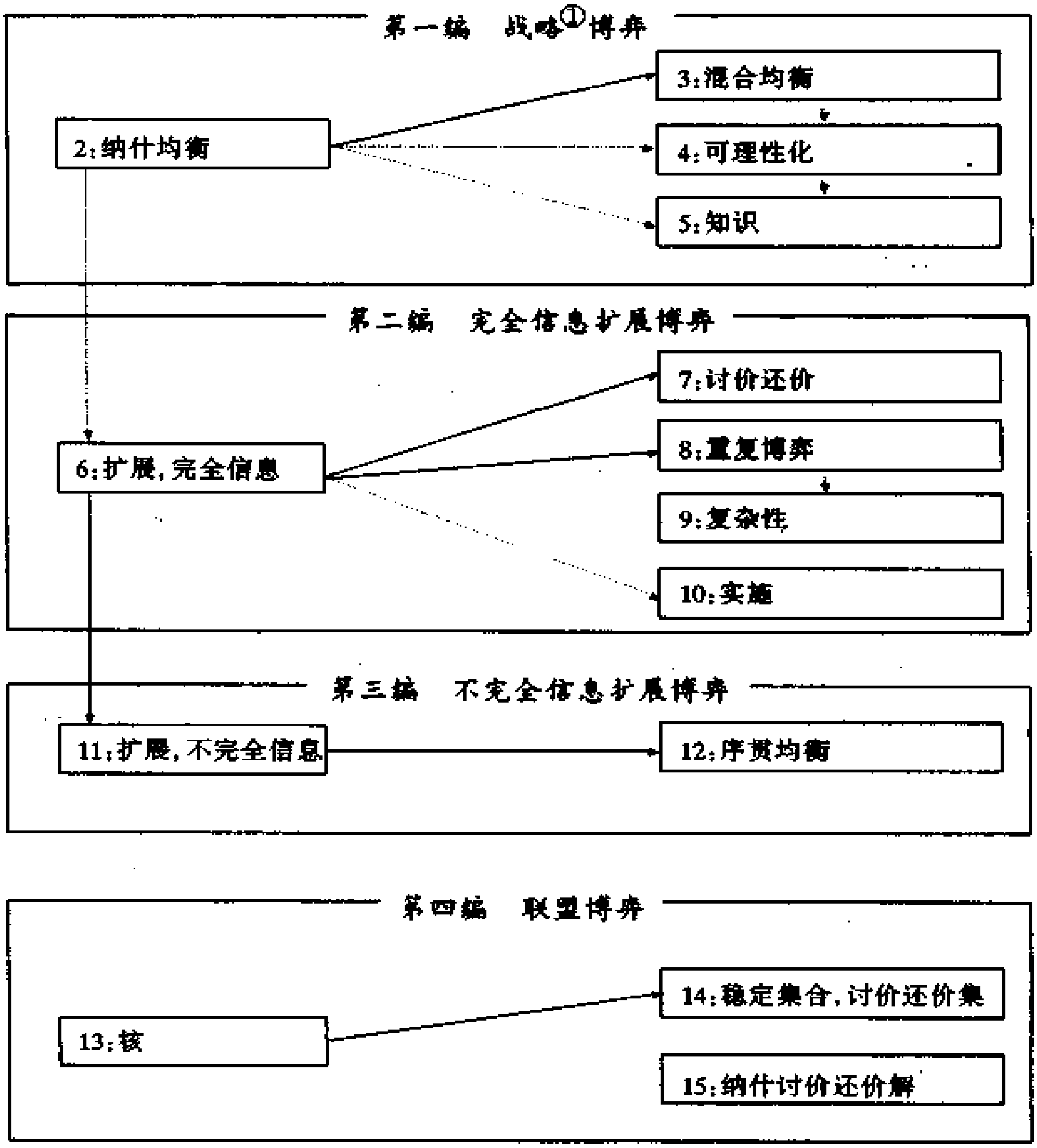
本书所涵盖的问题是那些我们个人认为在这样一个学期的课程中所应包括的内容。我们并不想提供一本关于博弈论的全面参考书，读者也无须将我们未涵盖的主题视为不重要。因此，我们的选择不可避免地反映了自身的偏好和兴趣。（若我们现在动手写这样一本书，则我们可能会增加两章，一章是关于实验博弈论的，另一章是关于学习和演进的。）

书中侧重理论的依据和主要概念的解释。我们的方式是给出精确定义及结论的完整证明，有时为了尽可能容易地达到这些目的而不惜牺牲一般性及限制材料的范围。

我们尽了极大努力去认证所有的概念、结论、例子和练习（参看每章后的“注解”）。我们对可能出现的失误深表歉意，并希望读者能指出这些错误以引起我们的注意。

本书结构

本书包括四部分，在每一部分我们都研究一组相关模型。第 2 页的图表明了各章间的关系。如果作为一个基本教程来看待的话，实际上只需要包括第 2、3、6、11、12、13 章。



各章间的主要关系。各章标题所在的方框大小与各章的篇幅成比例。与两个方框相连的实线箭头表示后一章依赖前一章;虚线箭头表示仅是后一章的主要思想应用于前一章。一个基本教程可包括重线方框中的六章。

① 原书作者认为 strategy 在本书宜翻译为"策略"而非"战略"。译者的这一译法仅供参考。——译者

练习

很多练习都是富有挑战性的；我们经常用练习去陈述次要结论。教师可以给出附加的、较直接的问题，调整练习的水平（通过暗示）以适应于学生的具体需要。练习的答案在本书提供的互联网网址上（参见第5页后）。

作者间的分歧

由不同作者合写的书并非就只能反映一个统一的观点。在某些方面，如同下列的注解，我们会简要地讨论我们之间存在的分歧。

关于人称的注解

我们对如何处理英语第三人称单数有分歧。

AR 认为，我们应使用一个“中性”代词并赞成用“他”，这样能使人明白这是既指男人又指女人。频繁地提醒他/她问题只会转移读者在主要问题上的注意力。语言对形成思想是很重要的，但在叙述材料中过分突出则是无益的，就如同在某些圈子里一样。

MJO 认为没有“中性”的语言。特别是有许多例子，不管是从实际还是从语言运用的分析来看，“他”并不总是被感觉为包括男性和女性的。引用《美国遗产词典》（第三版，第831页）：“这样他并不是一个中性的；而是指这样一个男人，即他被当作由他祖先所涉及的群落成员的代表。传统的用法并非一个简单的语言习惯；它也表明了一个特别的思维方式。”进一步说，用“他”去指并非特定性别的单个人甚至不是自然而然的产物，而是作为一种制度被强加的。这个制度来源于18和19世纪，当时“他们”作为一个单数代词被广泛使用而困扰着语法学者。因为在语法学者看来，男人比女人更

重要，所以他们决定“他”应该被使用。“他”用于指全称个人，这样既有性别主义者态度方面的根源又助长了这些态度。对这个问题没有完美的答案，特别像在这样一本书中有这么多涉及全称个人之处。“他们”作为一个单数代词有很多好处，尽管它的运用会导致混乱（和来自编辑们的抱怨）。我的偏好是对所有个人都用“她”。显然这样使用不是中性主义者的，但在几个世纪“他”占据主导之后，近几十年来“她”的使用不是中性主义思维方式的做法。如果这样的使用使这些读者的注意力偏离于本书所讨论的内容，并且导致他们去思考语言运用中的性别主义（这个问题确实是至少与序贯均衡的细节一样重要的），那么将会获得社会福利的提高。（不管这本书能否被称为“学术材料”，我看没有理由说它的读者与任何其他材料的读者相比应该受到不同对待。）

总而言之，我们俩都强烈地感受到了这个问题，我们俩都认为我们已达成的妥协是不令人满意的。当涉及到具体某个人时，我们有时使用“他”，有时使用“她”。例如在两人博弈中，我们将参与人 1 作为女性，而将参与人 2 作为男性。对于全称个人，我们使用“他”。

致谢

这本书是我们将所教课程与很多朋友及同事讨论后的产物。第 1、8、9 章中的某些材料来自于 AR 的一本关于有限理性模型的书的部分摘要内容。

MJO 我有幸在斯坦福大学向 Rober Aumann、Sergiu Hart、Mordecai Rurz 及 Rober Wilson 学习博弈论。我非常高兴能有机会表达对他们的感谢。多年与 Jean-Pierre Benoît、Haruo Imai、Vijay Krishna 和 Carolyn Pitchik 的讨论增进了我对很多主题的了解。我是在对新西兰坎特伯雷大学经济系的访问中完成本书工作的；对于该系各位同仁的殷勤好客我深表谢意。我也非常感激加拿大社会科学与人类研究委员会和加拿大自然科学与工程研究委员会，在过去六年里，她们对我的博弈论研究给予了经济上的帮助。

AR 我曾在伦敦经济学院（1987 年和 1988 年）、希伯莱大学（1989 年）、特拉维夫大学（1990 年）及普林斯顿大学（1992 年）的教程里使用过本书部分内容。伦敦经济学院、普林斯顿大学和特拉维夫大学的殷勤好客与合作是非常值得怀念的。特别值得感谢的是我的朋友 Asher Wolinsky，我们之

间有过无数次富于启发性的谈话。本书的部分工作得到了美—以双边科学基金的支持(批准号 1011-341)。

我们要向 Pierpaolo Battigalli、Takako、Fujiwara、Wulong Gu、Abhinay Muthoo、Michele、Piccione 和 Doron Sonsino 致以谢意。他们对本书的提纲做了详细的评论,这些评论使我们真正地提高了教材的精确性与可读性。我们还要感谢 Dilip Abreu, Jean-Pierre Benoît, Larry Blume, In-koo Cho, Eddie Dekel, Faruk Gul, Vijay Krishna, Bart Lipman, Bentely MacLeod, Sylvain Sorin, Ran Spiegler 和 Arthur Sweetman 等对本书提出的建议和改进办法。最后,我们要感谢 Wulong Gu 和 Arthur Sweetman,在完成本书的过程中,他们给予了我们特别的帮助:Wulong 从事练习方面的工作,修改了我们的答案并提供了很多他自己的东西;Arthur 建立了本书的索引。

在技术方面我们感谢 Ed Sznyter 劝服一直倔强的 TEX 去执行我们的编号结构。

同 MIT 出版社 Terry Vaughn 打交道是件很愉快的事情;在该项目的初期阶段,他的鼓励对于促成我们完成本书是很重要的。

马丁 J·奥斯本
阿里尔·鲁宾斯坦

MARTIN J. OSBORNE
osborneomcaster. ca
Department of Economics, McMaster University
Hamilton, Canada, L8S 4M4

ARIEL RUBINSTEIN
rarieloccsg. tau. ac. il
Department of Economics, Tel Aviv University
Tel Aviv, Israel, 69978
Department of Economics, Princeton University
Princeton, NJ 08540, USA

我们为本书保留了一个网址。该地址联结由 MIT 出版社提供

http://mitpress. mit. edu/book-home. tcl? isbn = 026250401

我们地址的 URL 现在是

http://www. chass. utoronto. ca/~osborne/cgt

目　　录

第二编 完全信息扩展博弈

第三编 不完全信息扩展博弈

第 1 章 1

绪　论

1.1　博弈论

博弈论是一个分析工具包，它被设计用来帮助我们理解所观察到的决策主体相互作用时的现象。这种理论隐含的基本假设是：决策主体追求确定的外部目标（他们是理性的）并且考虑他们自身的知识或其他决策主体行为的期望（他们**推理具有战略性**）。

博弈论模型是对各种现实生活状况的高度抽象概括。这种抽象性使得它们可被用来研究范围很广的现象。例如，纳什均衡理论（见第 2 章）被用来研究寡头垄断和政治竞争，混和战略均衡（见第 3 章）被用来解释蜂舌长度和花粉管长度的分布状况；重复博弈论（见第 8 章）被用来阐述诸如威胁和承诺等社会现象；核理论（见第 13 章）则揭示了这样一种意义，即在包括多个代理人的经济里，某种价格系统下的交易结果是稳定的。

在博弈论中，纯理论与应用理论间的界限是模糊的，某些纯理论方面的发展是由应用方面的问题引起的。不过，我们相信这个界线依然存在。尽管我们希望这本书能引起那些应用者的兴趣，但本书似乎仅停留在“纯”理论的领域。将抽象模型应用于现实生活的艺术，该是另一部书的主题。

博弈论使用数学来正式地表达它的思想。然而，我们所讨论的博弈论
思想却不是“生”来就带数学味的。从原理上讲，不使用数学，也完全可以写 2
一本与该书内容基本相同的著作。数学形式使得精确地定义概念变得比较容易，可以验证思想的一致性，还可以探求假设的内涵。结果我们采取一种正式的风格，精确地表达定义和结论，并将概念的由来和解释贯穿其间。

数学模型的使用带来了独立的数学兴趣。但在本书中我们将博弈论视为社会科学而非数学的一个分支，其目的是去理解相互作用决策主体的行为。因此，本书并不侧重于数学方面的兴趣，从我们的观点看，数学结论只有被直觉认证后才是有趣的。

1.2 博弈和解

博弈是对战略相互作用的描述，它包括对参与人所能采取的行动的约束和参与人的兴趣，但不强调参与人实际采取的行动。解是对结果的系统描述，这种结果可能产生于一组博弈。博弈论给出各种博弈的合理解并且考察它们的性质。

我们研究四类博弈论模型，它们如本书的四部分标题所示：战略博弈（第一编），具有或不具有完全信息的扩展博弈（第二、三编）及联合博弈（第四编）。下面我们解释这样分类的一些依据。

1.2.1 非合作与合作博弈

在所有的博弈论模型中，基本的实体是参与人(player)。参与人可被解释为单个或一组作某项决策的人群。一旦定义了参与人集合，我们便可区分两类模型：一类是以单个参与人的可能行动集合为基本元素（见第一、二、三编）；另一类是以参与人群的可能联合行动集合为基本元素（见第四编）。有时第一类模型被称为“非合作型”，第二类模型被称为“合作型”（尽管这些术语并不能说清两类模型间的差异）。

从本书各章所占的篇幅上便可看出，近年来大部分研究集中于非合作
3 博弈，但这并不能代表我们对这两个分支相对重要性的评价。我们尤其不同意某些作者的观点，在他们看来非合作博弈模型比合作模型更“基本”。我们认为两者不相上下。

1.2.2 战略博弈和扩展博弈

在第一编我们讨论战略博弈的概念，在第二、三编讨论扩展博弈的概念。战略博弈是这样一种情形的模型：每个参与人选择且仅选择一次行动

计划，并且所有参与人的决策是同时做出的（也就是说，在选择行动计划时每个参与人并不知道其他参与人的行动计划）。与此相反，扩展博弈模型则强调事件的可能顺序：每个参与人不仅可以在博弈开始时考虑自己的行动计划，并且在他不得不做决策的任何时候，也可以考虑他的行动计划。

1.2.3 完全与不完全信息博弈

我们所作的第二个区分是针对第二编和第三编中的模型的。在第二编的模型中参与人对任何其他人的行动都了解，而在第三编的模型中参与人就可能不太清楚别人的行动。前一模型有比较稳定的基础。后一模型在八十年代才开始兴起，我们对其不做重点介绍不是因为它们不切实际或不够重要，而是因为它们不太成熟。

1.3 博弈论和竞争均衡理论

为了近一步澄清博弈论的本质，现在将它与经济学中的竞争均衡理论作一比较。博弈论要考虑决策主体在做出决策前企图获得其他参与人的行为信息，而竞争理论给出的假定是：每个参与人只对某些环境参数感兴趣（例如价格），即使这些参数是被全体参与人的行为所决定的。

我们通过考虑下面一种情形来说明这两个理论的差异：在该情形中，每个参与人的某种行动（如钓鱼）的水准依赖于污染的程度，反过来污染程度
又依赖于全体参与人的活动。若用竞争理论分析，我们便会去寻找一个与 4
全体参与人行动相一致的污染程度，此时每个参与人都认为这个程度是给定的；若用博弈理论分析，我们则要求每个参与人的行动均为最优，此时每个参与人与其他参与人一起造成的污染预期是给定的。

1.4 理性行为

我们研究模型时，假定每个决策主体都是“理性的”，这种理性是建立在这样的意义之上的，即决策主体知道他的选择内容，对未知的事物形成预

期，具有明确的偏好，并在经过一些最优化过程后审慎选择他的行动。排除不确定性因素后下面的要素便组成了一个理性选择模型。

- 一个行动(action)集合 A，决策主体从 A 里做一个选择；
- 一个上述行为的可能结果(consequence)集合 C。
- 一个结果函数(consequence function) $g: A \to C$，g 使每个行动与一个结果相对应。
- 一个集合 C 上的偏好关系(preference relation) $\succsim$。(一个完全的，可传递性的，自反的和二元的关系。)

5 有时专门给定一个效用函数(utility function) $U: C \to \mathbb{R}$ 来表示决策主体偏好。效用函数确定了一个偏好关系 $\succsim$，$x \succsim y$，当且仅当 $U(x) \geqslant U(y)$。$\succsim$ 满足下列条件：

给定任何一个集合 $B \subseteq A$，A 为某个特别情形下的可能行动集合。一个**理性决策主体**选择一个可能行动 $a^*(a^* \in B)$。若 a^* 对所有 $a \in B$ 满足 $g(a^*) \succsim g(a)$，则 a^* 为最优的。相应的它解决了问题 $\max_{a \in B} U(g(a))$。值得注意的是，使用这个决策模型要假定每个决策主体在从不同集合 B 选择时要使用同一个偏好关系。

在我们所研究的模型中，决策主体往往要在不确定条件下进行决策。参与人可能：

- 不能确定环境的客观参数；
- 对博弈中发生的事件不很清楚；
- 不能确定别的不确定参与人的行动；
- 不能确定别的参与人的推理。

为了对不确定情形下的决策建模，几乎所有的博弈论都使用了 von Neuman 和 Morgenstern(1944)及 Savage(1972)的理论。也就是，如果结果函数是随机的并被决策主体已知(即，对每一个 $a \in A$，结果 $g(a)$ 是集合 C 上的一个不确定事件)(概率分布)，那么决策主体就被认为是为了最大化一个函数期望值(v-N-M 效用)去行动，这个函数给每个结果赋一个值。如果行动与结果间的随机联系未给定，那么决策主体就被认为是好像按他心中的一个(主观的)概率分布去行动，这个分布决定了任何行动的结果。在这种情形下决策主体被认为将这样行动，即他心中有一人"状态空间"(state space) Ω，一个 Ω 上的概率测度，一个函数 $g: A \times \Omega \to C$，和一个效用函数 $u: C \to \mathbb{R}$；他被假定为考虑到概率测度去选择一个行动 a 来最大化期望值 $u(g(a, \omega))$。

我们不去讨论那些隐含在理性决策主体之下的种种假设。然而，需要指出的是，这些假定不断地受到实验心理学家的批评，他们经常提出一些其在应用时所需的严格限制条件。

1.5 稳定状态和推论的解释

对于战略和扩展博弈的解，有两种相互矛盾的解释。稳定状态(steady state)(或被 Binmore(1987/88)称之为演进(evolutive))解释，与经济学中的标准紧密相关。博弈论像其他学科一样处理具有规律性的东西。正如 Carnap(1966.p.3)所言："我们日常所见如同较系统的科学观察一样，揭示了世界上具有一定重复性或规律性的东西……科学规律仅是尽可能精确地对这种具有规律性的现象进行解释。"稳定状态解释将博弈作为这样一种模型，即它被设计用来解释一组相似状况中观察到的规律，每个参与人凭借长期实践获得的知识从而"知道"均衡并测试其行为的最优性。推论(deductive)(或者，被 Binmore 称之为演绎的(eductive))解释则将博弈视为孤立的，如同"一瞬间"的事情，并且试图推断理性强加于结果上的限制。它认为每个参与人都简单地从理性原则出发去推断其他参与人的行为。我们则力图避免这两种解释在博弈论中经常引起的混乱。

1.6 有限理性 6

当我们在现实生活中谈论博弈时，我们经常关心各参与人能力的非对称性。例如，某些参与人可能对状况有更清楚的洞察力或有更强的分析能力。这些现实生活中如此关键的差异却被博弈论现有的形式忽略了。

为了说明这个事实所造成的后果，让我们来考虑象棋游戏。在一场真实的象棋游戏中，参与人在他们有关符合规则的行动知识方面和在分析能力方面都可能有差异。然而，若用现代博弈论对象棋游戏建模，则会认为参与人具有的象棋规则知识是充分的，他们的分析能力也是理想的。我们在第2、6章证明的结果(见本书命题22.2和99.2)表明：对于"理性"参与人来说，象棋是一种平庸博弈。存在一个算法来"解"此博弈。该技术为两个

参与人各确定了一个战略，这样就有一个“均衡”结果，该结果的性质是无论其他参与人采用何种战略，只要参与人遵从他自己的战略，那么他的结果至少与均衡结果一样好。这个结果（最早由 Zermelo（1913））表明象棋游戏是没有什么意思的，因为它仅有一个可能结果。然而尽管有这一结论，象棋仍是一种有趣的游戏。它的均衡结果仍有待计算，现在尚不可能采用 Zermelo 的算法。即使有朝一日可证明白方（the white）有一种取胜战略，而要一个人去实行它也可能是行不通的。由此抽象的象棋模型使我们推导出有关该博弈的一个重要事实，但同时又忽略了一场真实象棋游戏结果的最终决定因素：参与人的“能力”。

对不同参与人的能力及形势洞察力的不对称性建模在将来的研究中将是一个吸引人的挑战，对此“有限理性”模型已经开始发挥作用了。

1.7 术语和标记

我们假定读者对数学结果不很熟悉，但本书从头至尾都使用演绎推理。我们采取的标记和数学定义都是标准的，但为了避免混乱在这里列出一些。

7 我们用$\mathbb{R}$表示一个实数集合，$\mathbb{R}_+$表示非负实数集合。$\mathbb{R}^n$表示 n 维实数向量集合，$\mathbb{R}^n_+$表示 n 维非负实数向量集合。对于 $x\in\mathbb{R}^n$和$y\in\mathbb{R}^n$，我们用 $x\geqslant y$ 表示$x_i\geqslant y_i$，其中 $i=1,\cdots n$，用 $x>y$ 表示$x_i>y_i$，其中 $i=1,\cdots n$。我们说一个函数 $f:\mathbb{R}\to\mathbb{R}$，是**递增**（increasing）的，假如每当 $x>y$ 时$f(x)\geqslant f(y)$，每当 $x>y$ 时$f(x)>f(y)$称为是**非递减**（nondecreasing）的。如果对所有 $x\in\mathbb{R}$，$x'\in\mathbb{R}$及 $\alpha\in[0,1]$，$f(\alpha x+(1-\alpha)x')\geqslant\alpha f(x)+(1-\alpha)f(x')$，则函数 $f:\mathbb{R}\to\mathbb{R}$ 为一个凹函数。给定一个函数 $f:X\to\mathbb{R}$，我们用 $\arg\max_{x\in X}f(x)$表示 f 的最大值集合，对任何 $Y\subseteq X$，用 $f(Y)$表示集合$\{f(x):x\in Y\}$。

从始至终我们用 N 表示参与人集合。我们将某个变量的值的集合（每个参与人都对应一个）作为一个组合（profile），用$(x_i)_{i\in N}$表示。或者，假如量词“$i\in N$”是确定的，则简单记为(x_i)。对任何一个组合 $x=(x_j)_{j\in N}$和任何$i\in N$，我们令 x_{-i}为除 i 以外所有参与人组合。给定列表 $x_{-i}=(x_j)_{j\in N\setminus\{i\}}$和一个元素 x_i，我们用(x_{-i},x_i)表示组合$(x_i)_{i\in N}$。如果对每个 $i\in N$，X_i 是一集合，则我们用 X_{-i}表示集合

$\times_{j\in N\setminus\{i\}}X_j$。

对于集合 A 上的二元关系 $\succsim$，如果对任何 $a\in A$ 和 $b\in A$，有 $a\succsim b$ 或 $b\succsim a$，则称为完备的；若对任一 $a\in A$，有 $a\succsim a$，则称为自反的；若对任何 $a\succsim b$，$b\succsim c$，有 $a\succsim c$，则称为传递的。一个偏好关系（preference. relation）是完备的、自反的、传递的、二元的。若 $a\succsim b$ 但没有 $b\succsim a$，那我们写成 $a\succ b$；若 $a\succsim b$ 且 $b\succsim a$，则我们写成 $a\sim b$。如果 A 中的任何序列 $(a^k)_k$ 和 $(b^k)_k$，对所有 k 有 $a^k\succsim b^k$，若 $(a^k)_k$ 和 $(b^k)_k$ 分别收敛到 a 和 b，仍有 $a\succsim b$，则我们说 A 上的偏好关系 $\succsim$ 是**连续的**。对于 $\mathbb{R}^n$ 上的偏好关系 $\succsim$，若对每一个 $b\in\mathbb{R}^n$，集合 $\{a\in\mathbb{R}^n: a\succsim b\}$ 是凸的，则 $\succsim$ 是**拟凹的**（quasi-concave）．若每个集合 $\{a\in\mathbb{R}^n: a\succsim b\}$ 是**严格凸的**，则 $\succsim$ 是**严格拟凹的**（Strictly quasi-concave）。

令 X 为一集合，我们用 $|X|$ 表示 X 中元素的个数。X 的分割（Partition）是 X 的非连通子集的一个集合。这些子集的和为 X。令 N 是一有限集合并令 $X\subseteq\mathbb{R}^N$ 为一集合，若对所有 $i\in N$，没有任何 $y\in X$ 满足 $y_i>x_i$，则 $x\in X$ 是帕累托有效的（Pareto efficient）．若对所有 $i\in N$，没有 $y\in X$ 满足 $y_i\geqslant x_i$ 及对某些 $i\in N$，有 $y_i>x_i$，则称 $x\in X$ 是**严格帕累托有效的**（Strongly Pareto effcient）。

一个有限（或可数）集合 X 上的**概率测度**（Probability measure）μ 是一个可加函数，它将 X 的每一子集与一非负实数联系起来（即当 B 和 C 不相交时，有 $\mu(B\cup C)=\mu(B)+\mu(C)$ 并满足 $\mu(X)=1$。在某些情形下我们研究的概率测度所对应的集合并不一定是有限的。但如果你并不熟悉这种测度，则只要将注意力集中于有限情形即可。对于概率测度的更一般意义，请参看 Chung(1974, ch.2)。

[注解] 8

von Neumann 和 Morgenstern(1944)是博弈论方面的经典作家。Luce 和 Raiffa(1957)是一本早期的教科书，尽管现在已过时了，但它仍包含了该理论基本概念的上乘讨论。Schelling(1960)提供了该理论的一些主要思想的口头讨论。

很多最近出版的书都包含了本书大部分的内容，程度相近的书有：Shubik(1982), Moulin(1986), Friedman(1990), Kreps(1990a, Part Ⅱ), Fudenberg and Tirole(1991a), Myerson(1991), van Damme(1991), 和 Binmore(1992)。Gibbons(1992)是一本更基本的入门书。

Aumann(1985b)包含了有关博弈论目标和成果的讨论, Aumann(1987b)从历史角度说明博弈论。Binmore(1987/88)针对博弈论稳定状态和推论解释两者间的差异进行了批判性的讨论。Kreps(1990b)则对博弈论中的很多问题进行了反省性的讨论。

关于理性选择理论的说明可参看 Kreps(1988)。

第一编

战略博弈

在这部分我们将研究一种战略相互作用模型，它被称之为战略博弈(strategic game)，或者，用 von Neumann and Morgenstern(1944)的术语，称之为“标准形式博弈”。这种模型为每个参与人指定一个可能行动集合和一个建立在可能行动组合集合上的偏好顺序。

第 2 章讨论纳什均衡，对战略博弈来说这是应用最广的解的概念。第 3 章讨论紧密相关的混合战略均衡和相关均衡解，这里参与人的行动不必是确定的。纳什均衡是一个稳定状态解的概念，此间每个参与人的决策依赖于均衡知识。第 4 章研究可理性化和反复剔除劣行动的概念，此间参与人被假定为不知道均衡。第 5 章介绍一个知识模型，它使我们能正式检验在已定义好的各种解之下的种种假设。

11 # 第2章

纳 什 均 衡

纳什均衡是博弈论中最基本的概念。本章我们将在战略博弈范围和贝叶斯博弈相关范围介绍它。

2.1 战略博弈

2.1.1 定义

战略博弈是一种相互作用决策的模型，这种模型假定每个决策主体选择且仅选择一次行动计划，并且这些选择是同时进行的。该模型包括参与人的有限集合 N，对每个参与人 i 有一个行动集合 A_i 和一个建立在行动组合上的偏好关系。我们称一个行动组合 $a=(a_j)_{j\in N}$为结果(outcome)，并且用 A 表示结果集合 $\times_{j\in N}A_j$。这里要求将每个参与人 i 的偏好定义在 A 而不是A_i 上，这是将战略博弈从决策问题中区分出来的特征所在，即每个参与人不仅要考虑自己的行动，还要考虑其他参与人采取的行动。简言之，我们的定义如下：

▶定义 11.1　一个**战略博弈**包括：

• 有限集合 N(**参与人集合**)

• 对每个参与人 $i\in N$ 有一非空集 A_i(对参与人 i 有效的**行动集合**)

• 对每个参与人 $i\in N$，一个建立在集合 $A=\times_{j\in N}A_j$ 上的偏好关系 $\succsim_i$(参与人 i 的**偏好关系**)。

如果每个参与人 i 的行动集合 A_i 是有限的，则博弈是有限的。

该模型如此抽象使其能应用于较广的情形(situation)范围。一个参与 12
人可以是单独的一个人或任何一个决策主体如政府、董事会、革命运动中的领导集体，甚至可以是一朵花或一只动物。该模型没有限制参与人有效的行动集合，例如，它可以包含几个元素或是一个具有多种状态依存的复杂计划集合。然而，由于我们要求一个参与人只对应一个偏好关系，该模型的使用范围也因此而有限。参与人偏好关系可以只是简单地反映参与人对可能结果的感受，在一个有机体不能有意识活动的情形下它可以是繁殖成功的机会。

该模型如此抽象的事实，有利于它应用于很广的范围，但同时也有缺陷，因为模型的内涵不依赖于情形的任何特有性质。实际上，在这个抽象程度上，对于博弈的结果得不出什么结论。要得到有趣的结论，模型应更加具体化。

在一些情形下，参与人的偏好非常自然地被定义在结果上而不是行动组合本身。例如，当要对寡头垄断建模时，我们可能将公司集合作为参与人集合，将价格集合作为每个公司的行动集合；但我们可能希望对这样的假设建模，即每个公司仅关心她的利润，而不关心产生利润的价格组合。为此，就要引进一个结果集合 C；一个函数 $g: A \to C$，它将结果与行动组合联系起来，以及一个 C 上的偏好关系组合($\succsim_i^*$)。那么在战略博弈中每个参与人的偏好关系 $\succsim_i$ 被定义为：$a \succsim_i b$ 当且仅当 $g(a) \succsim_i^* g(b)$。

有时我们希望对这样一种情形建模，即行动组合结果受外来的一个随机变量的影响，而参与人在行动前并不知道这种影响能否实现。对此我们也能作为战略博弈处理：引进一个结果集合 C，概率空间 Ω，和一个函数 $g: A\times\Omega \to C$。该函数的解释是：当行动组合是 $a\in A$ 及随机变量的实现值是 $\omega\in\Omega$，则 $g(a,\omega)$ 便是结果。行动组合包括 C 上一个不确定事件(lottery)；对每个参与人 i，偏好关系 $\succsim_i^*$ 必须是在所有这些不确定事件集合上的具体化。在战略博弈中参与人 i 的偏好关系定义如下：$a \succsim_i b$，当且仅当根据关系 $\succsim_i^*$，C 上被 $g(a,\cdot)$ 引起的不确定事件至少与被 $g(b,\cdot)$ 引起的不确定事件具有相同的后果。

在一个广泛的范围里，战略博弈中参与人 i 的偏好关系 $\succsim_i$，可以用支 13
付函数(payoff function) $u_i: A \to \mathbb{R}$ (亦称为效用函数)来表示，该函数的意义是，只要 $a \succsim ib$，就有 $u_i(a) \geq u_i(b)$。我们称这一函数值为支付(或效用)。我们经常通过给定一个支付函数来确定一参与人的偏好关系。就此而言我们将博弈表示成 $\langle N,(A_i),(u_i)\rangle$，而不是 $\langle N,(A_i),(\succsim_i)\rangle$。

具有两个参与人的有限战略博弈可由图 13.1 中的表便利地表示。用行表示一个参与人的行动,另一个参与人的行动则用列表示。由行 r 和列 c 形成的方框中的两个数,为行参与人选择 r 和列参与人选择 c 时两人的支付。这样在图 13.1 的博弈中,行参与人的行动集合为 $\{T, B\}$,列参与人的行动集合为 $\{L, R\}$,并且从结果 (T, L) 来看,行参与人的支付为 w_1,列参与人的支付为 w_2。若参与人的名字为"1"和"2",则可以方便地说行参与人为参与人 1,列参与人为参与人 2。

	L	R
T	w_1, w_2	x_1, x_2
B	y_1, y_2	z_1, z_2

图 13.1 两个参与人战略博弈的便利表示,假定每个参与人有两种行动

2.1.2 关于解释的评论

战略博弈的一般解释是:这是一个事件只发生一次的模型。每一个参与人都知道博弈的细节及所有参与人都是"理性的"(见第 1.4 节)事实;并且参与人同时独立地选择他们的行动。在这种解释下每个参与人在选择他的行动时并不知道别的参与人的选择,没有信息(除模型的基本元素外)可以让参与人用来形成对别人行动的预期。

另一个解释,也就是本书大部分内容所采用的解释是:参与人可以通过此种博弈或在过去进行的相似博弈的信息来形成别的参与人的行为预期
14 (参看第 1.5 节)。只要博弈行动间不存在战略联系,则博弈行动系列都可用战略博弈来建模。也就是,一个多次从事博弈的人必须只关心他此刻的支付而忽略他现在的行动对其他参与人将来行动的影响。在这种解释下只要排除相互作用事件间的非暂时战略联系,那么将情形模化成战略博弈便是恰当的。(第 8 章所讨论的重复博弈模型即是处理相互作用战略序列的,在那里确实存在这种非暂时联系。)

当称战略博弈参与人行动为"同时的"时候,我们并不强调这些行动是在时间的同一点上完成。如下的一种情形也能用战略博弈来建模:参与人在终点前处于不同位置,开始参与人的可能行动和支付是公开说明了的(这样它们便是参与人间的共同知识),然后每个参与人通过向中央电脑输入一

条信息来选择一个行动，当所有信息被收到后参与人便知道他们的支付了。不过，战略博弈模型应用范围比这个例子所提供的还要广泛得多。应用战略博弈建模的重要之处在于，参与人独立作决策且所有参与人在做决策前并不知道其他参与人的选择。

2.2 纳什均衡

在博弈论中使用最广的解的概念是纳什均衡。这个概念体现了战略博弈行动的稳定状态，在此状态里每一个参与人都拥有对其他参与人行动的正确预期，并且能理性行动。它并不试图去检查稳定状态达到的过程。

▶定义 14.1 **战略博弈**$\langle N, (A_i), (\succsim_i)\rangle$的纳什均衡(a Nash equilibrium of a strategic game)是一个行动组合 $a^* \in A$，a^* 的性质是：对每一个参与人 $i \in N$。我们有：

$$(a^*_{-i}, a^*_i) \succsim_i (a^*_{-i}, a_i)\text{，对所有 } a_i \in A_i.$$

这样对 a^* 为一纳什均衡而言它必须满足：对其中任何一个参与人 i 而言，当其他每个参与人 j 选择均衡行动a^*_j 时，参与人 i 没有其他行动产生的结果优于他选择 a^*_i 所产生的结果。简言之，若给定其他参与人的行动，则参与人没有积极性选择别的行动。 15

下面对定义的重新表述有时是有用的。对任一 $a_{-i} \in A_{-i}$ 定义 $B_i(a_{-i})$为参与人 i 在给定a_{-i}下最佳行动集合：

$$B_i(a_{-i}) = \{a_i \in A_i : (a_{-i}, a_i) \succsim_i (a_{-i}, a'_i)\text{，对所有 } a'_i \in A_i\} \quad (15.1)$$

我们称集值函数 B_i 为参与人 i 的**最优反应函数**(best-response function)。纳什均衡是一个行动组合 a^*，其满足：

$$a^*_i \in B_i(a^*_{-i})\text{；对所有 } i \in N. \quad (15.2)$$

这个定义的替代形式向我们提供了(不一定有效)一个寻找纳什均衡的方法：先计算每个参与人的最优反应函数，再寻找一个行动组合 a^* 满足：对全部 $i \in N$，有 $a^*_i \in B_i(a^*_{-i})$。若函数 B_i 是单值的。则第二步就可解具有$|N|$个未知数$(a^*_i)_{i \in N}$的$|N|$个方程。

2.3 举例

下列的经典博弈代表了一些战略情形。这些博弈非常简单:在每个博弈里仅有两个参与人并且每个参与人仅有两个可能行动。但是,每个博弈都具有一类战略相互作用的本质特征,这些本质特征在更复杂情形下也是经常可见的。

◇例 15.3 巴克与斯特拉温斯基(BoS)(Bach or Stravinsky)? 两个人希望一起去参加音乐会,他们要么经过 Bach,要么经过 Stravinsky。他们主要关心的是一起出去,但一个人喜欢 Bach,另一个人喜欢 Stravinsky。通过支付函数表示个人的偏好,我们有图 16.1 所示的博弈。

这种博弈经常被称为“性别战”,要了解其所包含的标准理论,请参看 Luce and Raiffa(1957, pp.90—91)。为了与这个术语系统相一致,我们称该博弈为“BoS”。

BoS 对这样一种情形建模:参与人希望协调他们的行动,但他们间又有利益冲突。该博弈有两个纳什均衡:(Bach, Bach)和(Stravinsky, Stravinsky)。也就是说,有两个稳定状态:一个是两者都选择 Bach,一个是两者都
16 选择Stravinsky。

	Bach	Stravinsky
Bach	2,1	0,0
Stravinsky	0,0	1,2

图 16.1 BoS 博弈(例 15.3)

◇例 16.1 合作博弈(A cordination game)如在 BoS 中一样,两个人希望一起出去,但在此例中他们就更想去的音乐会达成协议。描述这种情形的博弈如图 16.2 所示。

像 BoS 一样,该博弈有两个均衡:(Mozart, Mozart)和(Mahler, Mahler)。与 BoS 不同,参与人有共同兴趣达到这其中的一个均衡,即(Mozart, Mozart),然而,纳什概念不排除稳定状态(Mahler, Mahler),当然该状态结果较差。

	Mozart	Mahler
Mozart	2,1	0,0
Mahler	0,0	1,2

图 16.2 合作博弈(例 16.1)

◇例 16.2 囚徒困境(The Prisoner's Dilemma)

两名犯罪嫌疑人被分别关在两个牢房里。如果他们都坦白,则每人将判3年徒刑;如果只有一人坦白,则坦白者被释放并做为对另一个人的证人,另一个人将判4年徒刑;如果两人都不坦白,则他们两人都从轻发落只判刑一年。选择一个便利的支付来表示偏好,我们有图17.1所示的博弈。

	不坦白	坦白
不坦白	3,3	0,4
坦白	4,0	1,1

图 17.1 囚徒困境(例 16.2)

这是一种通过合作可获利的博弈——对参与人来说最好的结果是两人都不坦白——但每个参与人都愿成为无罪释放者。不管其中一个怎么做,另一个都偏好于坦白而不是不坦白,这时博弈有惟一纳什均衡(坦白、坦白)。

◇例 16.3 鹰—鸽(Hawk-Dove)

两动物为某一食物而争斗。每只动物都能像鸽或像鹰那样行动。对每只动物来说最好的结果是它像鸽一样;最坏的结果是两个都是像鹰一样。 17
每只动物都偏向于若对手像鸽一样则它自己像鹰一样或若对于像鹰一样则它自己像鸽一样。具有这种情形的博弈如图17.2所示。该博弈有两个纳什均衡,(鸽,鹰)。和(鹰,鸽),其对应于参与人产生的两种不同约定。

◇例 17.1 猜谜博弈(Matching Pennies)

两人中的每个人都可选择"头"或"尾",若两者选择不一样,则第一个人

	鸽	鹰
鸽	3,3	1,4
鹰	4,1	0,0

图 17.2 鹰—鸽(例 16.3)

给第二个人一元钱;若两者选择一样,则第二个人给第一个人一元钱。每个人仅关心他所得的钱的数量。对这种情形建模的博弈如图 17.3 所示。这种博弈,参与人的兴趣是完全相反的,被称为“严格竞争的”。猜谜博弈没有纳什均衡。

	头	尾
头	1,−1	−1, 1
尾	−1, 1	1,−1

图 17.3 猜谜博弈(例 17.1)

18 战略博弈概念包含的情形比上述 5 个例子所述的要复杂得多。下面是三类被广泛研究的博弈:拍卖、时间博弈和位置博弈。

◇例 18.1 拍卖(An auction)

某物将被移交给集合$\{1,\cdots,n\}$中的某一个参与人以换取一定的支付。参与人 i 对某物的估值是 v_i,且 $v_1>v_2>\cdots>v_n>0$。移交某物的机制是现场(密封价格)拍卖:参与人同时提供价格(非负数),该物移交给所有出价最高者中指数最低的参与人,以获得一定效用。

在一场一阶(first price)拍卖中胜者的支付就是他的出价。

?练习 18.2 试将一阶拍卖系统地表示成战略博弈形式并分析它的纳什均衡。特别地,试证明在所有均衡中参与人 1 获得该物。

在一场二阶(second price)拍卖中胜者的支付是这些未胜的参与人所提供的价格中最高的(这样若只有一个参与人提出最高价则支付的价格是第二高价)。

⑦ 练习 18.3 试证明在一场二阶拍卖中任何一个参与人的出价 v_i 都是次优(weakly dominant)行动不管其他参与人的行动，当参与人 i 出价为 v_i 时，他的支付至少与当他出其他任何价时的支付一样高。试证明胜者不是参与人 1 仍有("非效率的")均衡。

◇例 18.4 消耗战(A war of attrition)

两个参与人争夺同一个物体。物体对参与人 i 的价值为 $v_i>0$，时间设为连续变量，从 0 开始直到无穷。每个参与人选择向另一个参与人交出物体的时间；若一个参与人在时间 t 移交，则另一个人在此时同时得到。若两个参与人同时交出，则物体在他们间平分，参与人 i 得到支付 $v_i/2$。时间是有价值的：每一单位时间每个参与人失去一个单位支付，直到第一次移交发生。

⑦ 练习 18.5 将这一情形系统表示成战略博弈形式并证明在所有纳什均衡中某一个参与人会立即移交。

◇例 18.6 置位博弈(A location game)

n 个人中的每个人都面临是否做政治候选人的选择，如果选择了是，则还要选位置。有一个市民的群体，每一市民都有一个喜欢的位置，喜欢的位 19
置分布由 $[0,1]$ 上的一个密度函数 f 给定，f 满足：对所有 $x\in[0,1]$，有 $f(x)>0$。对候选人来说，如果他距市民喜欢的位置较其他候选人更近，则他可获得该市民的选票；如果 k 个候选人选择了同一位置，则每个候选人将得到该位置所能吸引到的选票 $1/k$。竞选的胜者为获选票最多者。每个参与人的偏好为：自己是惟一获胜者优于与他人并列第一，并列第一优于不参加竞选；不参加竞选优于虽参加却失败了。

⑦ 练习 19.1 将这种情形系统地表达成一个战略博弈，试找出 $n=2$ 时的纳什均衡集合，并证明当 $n=3$ 时没有纳什均衡。

2.4 纳什均衡的存在性

就像猜谜博弈(见图 17.3)所表明的那样，不是每个战略博弈都有纳什均衡。博弈的纳什均衡集合非空的条件已被广泛地考察过。我们现在给出

此类结论中最简单的一种。(尽管该结论的数学水平较之于本书其他部分要高,但它并不依赖于细节。)

存在性结论有两个目的。首先,如果有一个符合结论假设的博弈,那我们为寻求均衡付出的努力就会有成功的希望。其次,更重要的是,均衡存在性表明博弈是与一个稳定状态解相一致的。进一步说,一类博弈的均衡存在性使得我们可以去研究均衡的性质(例如通过使用"比较静态"技巧)而不用去具体地找到均衡,也不用冒研究空集的风险。

为了说明一个博弈具有纳什均衡就不得不去说明行动组合 a^*,其满足对所有 $i \in N$, $a_i^* \in B_i(a_{-i}^*)$(参看(15.2))。通过 $B(a) = \times_{i \in N} B_i(a_{-i})$ 来定义集值函数 $B: A \to A$。那么(15.2)可以简单地写成向量形式 $a^* \in B(a^*)$。不动点定理给出了 B 上的条件,在此条件下确定存在一个值 a^* 满足 $a^* \in B(a^*)$。我们所使用的不动点定理如下所述(这要归功于 Kakutani(1941))。

■引理 20.1(Kakutani 的不动点定理)。

20 令 X 是 $\mathbb{R}^n$ 的一个紧凸子集,令 $f: X \to X$ 是一集值函数且满足:

- 对所有 $x \in X$,集合 $f(x)$ 非空且凸;
- f 的图形是闭的(亦即:对所有序列 $\{x_n\}$ 和 $\{y_n\}$,对所有 n 有 $y_n \in f(x_n)$, $x_n \to x$, $y_n \to y$,我们有 $y \in f(x)$)。

则存在 $x^* \in X$ 满足 $x^* \in f(x^*)$。

? 练习 20.2　试证明下面的 4 个条件都是 Kakutani 定理必须的:(ⅰ) X 是紧的;(ⅱ) X 是凸的;(ⅲ)对每个 $x \in X$, $f(x)$ 是凸的;(iv) f 有一个闭图形。

如果对每个 $a^* \in A$,集合 $\{a_i \in A_i : (a_{-i}^*, a_i) \succsim_i a^*\}$ 是凸的,则可定义 A 上的偏好关系 $\succsim_i$ 为 A_i 上拟凹的(quasi-concave)。

■命题 20.3　战略博弈 $\langle N, (A_i), (\succsim_i) \rangle$ 若对所有 $i \in N$ 满足:

- 参与人 i 的行动集合 A_i 是欧氏空间的一个非空紧凸子集;

并且偏好关系 $\succsim_i$ 是

- 连续的
- 在 A_i 上拟凹的。

则,该博弈有一个纳什均衡。

证明:通过 $B(a) = \times_{i \in N} B_i(a_{-i})$来定义 $B: A \to A$(这里 B_i 是参与人 i 的最优反应函数,定义见(15.1))。对任一 $i \in N$,由于 $\succsim_i$ 是连续的且 A_i 是紧的,则 $B_i(a_{-i})$ 非空,又由于 $\succsim_i$ 在 A_i 上是拟凹的,所以 $B_i(a_{-i})$ 又是凸的;因为每个 $\succsim_i$ 是连续的,所以 B 有闭图。这样应用 Kakutani 定理 B 有一个不动点;正如我们所注意的那样任何不动点都是博弈的纳什均衡。 □

请注意:这个结论保证了满足一定条件的战略博弈至少有一个纳什均衡;正如我们所见,一个博弈可以有不止一个均衡。(我们并没有讨论结论在什么条件下仅有一个纳什均衡。)同时也应注意到命题 20.3 不适用于任何其中有某个参与人有有限多个行动的博弈,因为这种博弈违背了每个参与人行动集合是凸的条件。

?练习 20.4 对称博弈(Symmetric games)。

考虑一个满足命题 20.3 条件的两人战略博弈。令 $N = \{1, 2\}$ 并且假定博弈是对称的。$A_1 = A_2$ 并且 $(a_1, a_2) \succsim_1 (b_1, b_2)$ 当且仅当 $(a_2, a_1) \succsim_2 (b_2, b_1)$,对所有 $a \in A, b \in A$。利用 Kakutani 定理去证明有一个行动 a_1^* 21
$\in A_1$ 使得(a_1^*, a_1^*)是博弈的一个纳什均衡。(这一均衡称为对称均衡)试给出一个仅有非对称解的有限对称博弈的例子。

2.5 严格竞争博弈

对于一个任意的战略博弈纳什均衡集合,我们并没有太多的话可说,仅在一些确定的博弈类型里我们能对均衡的质的特性作些说明。一种确定类型的博弈是这样的,在这种博弈中有两个参与人,他们的偏好是正好相反的。为讨论方便我们在这部分假定参与人的名字为“1”和“2”(即:$N = \{1,2\}$)。

▶定义 21.1 战略博弈$\langle \{1, 2\}, (A_i), (\succsim_i) \rangle$是**严格竞争的**(strictly competitive),如果:对任一 $a \in A$ 和 $b \in A$ 我们有 $a \succsim_1 b$,当且仅当 $b \succsim_2 a$。

一个严格竞争博弈有时也叫做零和(zerosum)博弈,因为如果用支付函数 u_1 表示参与人 1 的偏好关系 $\succsim_1$,则用 u_2 表示参与人 2 的偏好关系,且 $u_1+u_2=0$。

对于参与人 i,如果他在假设不管自己采取什么行动,参与人 j 都将选择使他损失尽可能大的行动基础上选择对他自己最好的行动,则我们说参与人 i 最大最小化(maxminimizes)。我们现在要说明的是,对于一个具有纳什均衡的严格竞争博弈来说,一对行动是纳什均衡当且仅当每个参与人的行动都是最大最小化者。这个结论很有意义,因为它提供了一种个人决策与纳什均衡概念背后的理论之间的联系。在推导这一结论的过程中我们也要证明一个具有纳什均衡的严格竞争博弈的所有均衡产生同一支付这一强结论。在非严格竞争博弈中纳什均衡的这个特性是很少能得到满足的。

▶定义 21.2 令$\langle\{1,2\}(A_i),(u_i)\rangle$是一个严格竞争战略博弈。行动 $x^*\in A_1$ 是**参与人 1 的最大最小化行动**如果有:

$$\min_{y\in A_2}u_1(x^*,y)\geqslant\min_{y\in A_2}u_1(x,y),\qquad \text{对所有 } x\in A_1.$$

同样,行动 $y^*\in A_2$ 是**参与人 2 的最大最小化运行**如果有:

$$\min_{x\in A_1}u_2(x,y^*)\geqslant\min_{x\in A_1}u_2(x,y),\qquad \text{对所有 } y\in A_2.$$

22 简言之,对参与人 i 来说最大最小化行动就是最大最小化参与人 i 能保证获得的支付的行动。对参与人 1 来说最大最小化行动解决了问题 $\max_x\min_y u_1(x,y)$,相应对参与人 2 来说最大最小化行动解决了问题 $\max_y\min_x u_2(x,y)$。

在下文里为了方便起见假定参与人 1 的偏好关系用支付函数 u_1 表示,并且不失一般性,$u_2=-u_1$,下面的结论表明,参与人 2 支付的最大最小化值等同于参与人 1 支付的最小最大化值。

■引理 22.1 令$\langle\{1,2\},(A_i),(u_i)\rangle$为一严格竞争战略博弈。那么有

$$\max_{y\in A_2}\min_{x\in A_1}u_2(x,y)=-\min_{y\in A_2}\max_{x\in A_1}u_1(x,y).$$

进一步而言,$y\in A_2$ 解决了问题 $\max_{y\in A_2}\min_{x\in A_1}u_2(x,y)$当且仅当它解决了问题 $\min_{y\in A_2}\max_{x\in A_1}u_1(x,y)$。

证明:对任何函数 f 我们有 $\min_z(-f(z))=-\max_z f(z)$和 $\arg\min_z(-f(z))=\arg\max_z f(z)$。从而对每一 $y\in A_2$ 我们有 $-\min_{x\in A_1}u_2(x,y)$

$=\max_{x\in A_1}(-u_2(x,y))=\max_{x\in A_1}u_1(x,y)$。这样 $\max_{y\in A_2}\min_{x\in A_1}u_2(x,y)=-\min_{y\in A_2}[-\min_{x\in A_1}(u_2(x,y))]=-\min_{y\in A_2}\max_{x\in A_1},u_1(x,y)$;并且有:$y\in A_2$ 是问题 $\max_{y\in A_2}\min_{x\in A_1}u_2(x,y)$的解,当且仅当它是问题 $\min_{y\in A_2}\max_{x\in A_1}u_1(x,y)$的解。□

下面的结果给出了严格竞争博弈纳什均衡与最大最小化行动集合之间的一种联系。

■命题 22.2　令 $G=\langle\{1,2\},(A_i),(u_i)\rangle$是一严格竞争战略博弈。

(a)如果(x^*,y^*)是 G 的一个纳什均衡,则对参与人 1 来说 x^*是最大最小化行动,对参与人 2 来说 y^*是最大最小化行动;

(b)如果(x^*,y^*)是 G 的一个纳什均衡,则 $\max_x\min_y u_1(x,y)=\min_y\max_x u_1(x,y)=u_1(x^*,y^*)$,并且有 G 的所有纳什均衡产生同一支付;

(c)如果 $\max_x\min_y u_1(x,y)=\min_y\max_x u_1(x,y)$(特别地,如果 G 有一个纳什均衡(参看 b)),x^*是参与人 1 的最大最小化行动,y^*是参与人 2 的最大最小化行动,那么(x^*,y^*)是 G 的一个纳什均衡。

证明:我们先证明(a)和(b),令(x^*,y^*)是 G 的一个纳什均衡,那么对所有$y\in A_2$有 $u_2(x^*,y^*)\geqslant u_2(x^*,y)$。或者,既然 $u_2=-u_1$,则对所 23
有$y\in A_2$有 $u_1(x^*,y^*)\leqslant u_1(x^*,y)$,因此,$u_1(x^*,y^*)=\min_y u_1(x^*,y)\leqslant\max_x\min_y u_1(x,y)$。同样地,对所有 $x\in A_1$,$u_1(x^*,y^*)\geqslant u_1(x,y^*)$从而对所有 $x\in A_1$,$u_1(x^*,y^*)\geqslant\min_y u_1(x,y)$,于是 $u_1(x^*,y^*)\geqslant\max_x\min_y u_1(x,y)$。这样 $u_1(x^*,y^*)=\max_x\min_y u_1(x,y)$并且 x^*是参与人 1 的最大最小化行动。

对参与人 2 可类似证明,y^*是参与人 2 的最大最小行动,并且 $u_2(x^*,y^*)=\max_y\min_x u_2(x,y)$,于是 $u_1(x^*,y^*)=\min_y\max_x u_1(x,y)$。

为了证明(c),令 $v^*=\max_x\min_y u_1(x,y)=\min_y\max_x u_1(x,y)$。应用引理 22.1 我们有 $\max_y\min_x u_2(x,y)=-v^*$。因为 x^*是参与人 1 的最大最小化行动,我们有:对所有$y\in A_2$,$u_1(x^*,y)\geqslant v^*$;因为 y^*是参与人 2 的最大最小化行动,我们有:对所有 $x\in A_1$,$u_2(x,y^*)\geqslant -v^*$。在上述两个不等式中令 $y=y^*$及 $x=x^*$,我们得到 $u_1(x^*,y^*)\geqslant v^*$,并且利用$u_2=-u_1$ 这一事实,我们得到(x^*,y^*)是 G 的一个纳什均衡的

结论。 □

由(c)可注意到一个纳什均衡可通过求解问题 $\max_x\min_y u_1(x,y)$ 找到，当计算一个博弈的纳什均衡，特别是当参与人随机行动时(参看例 36.1)，上述事实是非常有用的。

从(a)和(c)也可注意到一个严格竞争博弈的纳什均衡是可互换的(interchangeable)：如果(x,y)和(x',y')是均衡解，则(x,y')和(x',y)也是均衡解。

(b)表明对任何有纳什均衡解的严格竞争博弈，$\max_x\min_y u_1(x,y)=\min_y\max_x u_1(x,y)$。请注意不等式 $\max_x\min_y u_1(x,y)\leqslant\min_y\max_x u_1(x,y)$具有更多一般性质：对所有 y，给定任一 x' 我们有 $u_1(x',y)\leqslant\max_x u_1(x,y)$，于是 $\min_y u_1(x',y)\leqslant\min_y\max_x u_1(x,y)$，(若最大、最小不能被有效定义，则可用上、下确界分别来代替最大、最小)这样在任何博弈中(无论它是否为严格竞争的)参与人 1 都能保证自己获得的支付就是参与人 2 能牵制她的最多数量。"博弈有一纳什均衡"这一假设在建立反向不等式中是必要的。为了明白这个论断，可以考虑猜谜博弈(图 17.3)，在那里 $\max_x\min_y u_1(x,y)=-1<\min_y\max_x u_1(x,y)=1$。

如果 $\max_x\min_y u_1(x,y)=\min_y\max_x u_1(x,y)$那我们就说这个支付，即参与人 1 的均衡支付，是这个博弈的值(value)。从命题 22.2 可知如果 v^* 是严格竞争博弈的值，那么参与人 1 的任何均衡战略都可保证她的支付至少与她的均衡支付 v^* 一样多，同时参与人2 的任何均衡战略都能保证他的支付至少与他的均衡支付 $-v^*$ 一样多，于是参与人 2 任一这样的战略都能保证参与人 1 的支付至多与她的均衡支付一样多。在一个非严格竞争博弈中参与人的均衡战略一般不具有这些性质(例如，可考虑
24 BoS(图 16.1))。

☑练习 24.1 令 G 为一个具有纳什均衡的严格竞争博弈。

a. 试证明：如果在 G 中增加参与人 1 的一些支付来形成一个新的严格竞争博弈 G'，则 G'不存在这样的均衡，即在该均衡里参与人 1 的状况比在 G 的均衡里恶化了。(注意，G'可能根本没有均衡。)

b. 试证明：若参与人 1 被禁止使用 G 中的某一行动从而形成一个新的博弈 G'，则 G'不存在这样的均衡，即在该均衡里参与人 1 的支付高于在均衡 G 里的支付。

c. 试举例证明：对一个非严格竞争博弈来说 a、b 中的两个性质都不是必要的。

2.6 贝叶斯博弈:不完全信息战略博弈

2.6.1 定义

我们经常希望对一部分人并不确知另一部分人的特征的情形建模:与战略博弈紧密相关的贝叶斯博弈模型便是为此而设计的。

作为一种战略博弈,贝叶斯博弈有两个基本元素:参与人集合 N 和行动集合组合(A_i),为了对参与人相互间不确定性建模,我们引进一个可能"自然状态"集合 Ω,其中的每个元素都是对所有参与人相关特征的描述。为方便起见我们假定 Ω 是有限的。通过给定 Ω 上的一个概率测度 p_i,每个参与人 i 便有一个关于自然状态的先验概率(prior belief)。在任何一个给定的博弈行动中,某一自然状态 $\omega\in\Omega$ 是实现了的。我们通过引进**信号函数**(signal functions)组合(τ_i)来对参与人关于自然状态的信息建模。当自然状态为 ω,在参与人1选择他的行动前,用 $\tau_i(\omega)$表示他所观察到的信号。令 T_i 为 τ_i所有可能值的集合,我们称 T_i 为参与人 i 的类型(*type*)集合。我们假定对所有 $t_i\in T_i$, $p_i(\tau_i^{-1})(t_i)>0$(参与人1给 T_i 中的每一个元素赋正先验概率)。如果参与人 i 收到信号 $t_i\in T_i$,则他推断状态在集合 $\tau_i^{-1}(t_i)$中;他关于已实现状态的**后验概率**(posterior belief))赋予每一状态 $\omega\in\Omega$,以概率 $p_i(\omega)/p_i(\tau_i^{-1}(t_i))$,此时要求 $\omega\in\tau_i^{-1}(t_i)$,否则,赋零概率(亦即,$\omega$ 的概率是以其在 $\tau_i^{-1}(t_i)$上为条件的)。例如,对所有$\omega\in\Omega$,若 $\tau_i(\omega)=\omega$ 则 25
参与人拥有关于自然状态的全部信息。或者,如果 $\Omega=\times_{i\in N}T_i$ 并且对每一参与人 i 概率测度 p_i 是 Ω 上的乘积测度,同时 $\tau_i(\omega)=\omega_i$,则参与人的信号是独立的且参与人 i 从他的信号里不能获得任何关于其他参与人的信息。

像在战略博弈里一样,每个参与人关心他的行动组合,同时他也可能关心自然状态。现在即使他知道在任何一个自然状态下其他每一个参与人所采取的行动,在给定他所采取的行动条件下,他也可能不能确定将要实现的二元组(a,ω)。因为他只有关于自然状态的不完全信息。因此我们在模型中引进一个有关 $A\times\Omega$ 上不确定事件的偏好关系组合($\succsim_i$)(这里,同前 $A=\times_{j\in N}A_j$)。简言之,我们作以下定义。

▶定义 25.1 一个**贝叶斯博弈**(Bayesian game)包括:

- 有限集合 N(参与人集合)
- 有限集合 Ω(状态集合)

并且对每个参与人 $i \in N$ 有

- 集合 A_i(对参与人 i 有效的行动集合)
- 有限集合 T_i(可能被参与人 i 所观察到的信号集合)和一个函数 τ_i $\Omega \rightarrow T_i$(参与人 i 的信号函数)
- Ω 上的一个概率测度 p_i(参与人 i 的先验概率)满足对所有 $t_i \in T_i$ 有 $p_i(\tau_i^{-1}(t_i))>0$
- 一个关于 $A \times \Omega$ 之上概率测度集合的偏好关系 $\succsim_i$(参与人 i 的偏好关系),这里 $A = \times_{j \in N} A_j$。

请注意这个定义允许所有参与人有不同的先验概率。这些概率可能是相关的,一般来说它们是同分布的,并与一个"客观的"测度相一致。这个模型被经常运用于自然状态是参与人偏好参数组合的情形(例如,他们对一个物体估值的组合)。不过,这个模型是很一般的,在2.6.3部分我们将讨论它被应用于每个参与人不确定别人所知的情形。

同时也应注意到有时描述贝叶斯博弈并不涉及暗含的状态空间 Ω,而
26 是描述成"简化形式",即与参与人信息相关的基本元素是可能类型集合的组合。

现在我们回头来谈贝叶斯博弈均衡的定义。在任何给定的博弈行动里每个参与人都知道他的类型,并且他不需要考虑在其他类型情形下他将做什么。因此,一个参与人可能会认为应对每个孤立的自然状态定义均衡。然而,在任一给定的状态里,一个希望决定他最优行动的参与人可能需要考虑别的参与人在别的状态里将如何做,因为他可能对状态不完全了解。进一步说,这一考虑的帮助信息可能依赖于参与人自己在别的状态里将选择的行动,因为别的参与人也可能对状态不完全了解。

这样我们就可将一个贝叶斯博弈 $\langle N, \Omega, (A_i), (T_i), (\tau_i), (p_i), (\succsim_i)\rangle$ 的纳什均衡定义成一个战略博弈 G^* 的纳什均衡,G^* 中对每个 $i \in N$ 和每个可能信号 $t_i \in T_i$ 都有个参与人,这个参与人称为 (i, t_i)。("具有类型 t_i 的参与人 i")。每个这种参与人 (i, t_i) 的行动集合为 A_i;这样在 G^* 中行动组合的集合为 $\times_{j \in N}(\times_{t_j \in T_j} A_j)$。每个参与人 (i, t_i) 的偏好如下定义。参与人 i 的后验概率,与 G^* 中的一个行动组合 a^* 一起,产生关于 $A \times \Omega$ 的一个不确定事件 $L_i(a^*, t_i)$:由 $L_i(a^*, t_i)$ 赋给 $((a^*(j, \tau_j$

$(\omega)))_{j\in N},\omega)$的概率,是参与人在收到信号 t_i 时状态为 ω 的后验概率,$(a^*(j,\ \tau_j(\omega))$是参与人$(j,\ \tau_j(\omega))$在组合 a^*中的行动。在 G^*中参与人(i,t_i)偏好行动组合 a^*优于行动组合 b^*,当且仅当参与人 i 在贝叶斯博弈中对不确定事件$L_i(a^*,t_i)$的偏好优于不确定事件 $L_i(b^*,t_i)$。由此,我们有下述内容。

▶定义 26.1 **一个贝叶斯博弈$\langle N,\Omega,(A_i),(T_i),(\tau_i),(p_i),(\succsim_i)\rangle$的纳什均衡是一个定义如下的战略博弈纳什均衡:**

- 参与人集合是所有二元组(i,t_i)的集合,其中 $i\in N,t_i\in T_i$。
- 每个参与人(i,t_i)的行动集合是 A_i。
- 每个参与人(i,t_i)的偏好顺序$\succsim^*_{(i,t_i)}$定义如下:

$$a^*\succsim^*_{(i,t_i)}b^* \text{ 当且仅当 } L_i(a^*,t_i)\succsim_i L_i(b^*,t_i),$$

这里 $L_i(a^*,t_i)$是关于 $A\times\Omega$ 的不确定事件,若 $\omega\in\tau_i^{-1}(t_i)$,则其将概率 $p_i(\omega)/p_i(\tau_i^{-1}(t_i))$赋给$((a^*(j,\ \tau_j(\omega)))_{j\in N},\omega)$,否则,赋零值。

也就是说,在一个贝叶斯博弈纳什均衡中,每个参与人在给定他所收到的信号及他所持的从信号推断的有关状态及别的参与人行动的信念条件下选择对他有效的最优行动。注意为了确定一个行动组合是否为贝叶斯博弈的纳什均衡,我们仅需知道贝叶斯博弈中的每个参与人是如何比较关于 $A\times\Omega$ 的不确定事件的,这里 Ω 上的概率分布是一样的:参与人从来不必比较概率分布相异的不确定事件。这样从纳什均衡的观点看,在贝叶斯博弈中参与人偏好的具体化包含了冗余信息。(这种冗余在战略博弈中类似存在:为了定义一个战略博弈的纳什均衡我们仅需知道任何一个参与人 i 是如何将一个结果(a_{-i},a_i)与另一个结果(a_{-i},b_i)相比较的。) 27

2.6.2 例

◇例 27.1 二阶拍卖(Second-price auction)考虑例 18.1 所述的二阶封价拍卖的一个变形,在例 18.1 中每个参与人知道他的估价 v_i,但不能确定别人的估价。作为特例,假定可能估价集合是有限集合 V 及每个参与人都相信任何一个其他参与人独立作出的估价都是从 V 上的同一分布出发的。则我们能将此情形建成贝叶斯博弈模型,其中:

- 参与人集合 N 是 $\{1,\cdots,n\}$
- 状态集合 Ω 是 V^n(估价的组合集合)
- 参与人 i 的行动集合 A_i 是 $\mathbb{R}_+$
- i 能收到的信号集合 T_i 是 V
- i 的信号函数 τ_i 定义为 $\tau_i(v_1,\cdots,v_n)=v_i$
- 先验概率 p_i 为:对 V 上的某个概率分布 π, $p_i(v_1,\cdots,v_n)=\prod_{j=1}^{n}\pi(v_j)$
- 参与人 i 的偏好关系由某一随机变量的期望所表示。在状态$(v_1,\cdots,v_n)$下,如果参与人 i 具有最低指数,对所有 $j\in N$,有 $a_i\geqslant a_j$,则该随机变量期望值为 $v_i-\max_{j\in N\setminus\{i\}}a_j$,否则,随机变量期望值为0。

这个博弈有一个纳什均衡 a^*,在此均衡中对所有 $i\in N$ 及 $v_i\in V=T_i$ 有 $a^*(i,v_i)=v_i$(每个参与人提出他的估价)。实际上(如练习 18.3)每个类型中每个参与人提出的估价都是弱占优行动。

?练习 27.2　两个参与人希望通过 Bach 或 Stravinsky 一起去参加一
28 场音乐会。如在 BoS 中一样,他们主要关心的是一起出去,但他们两个人都不知道另一个人是偏好 Bach,还是 Stravinsky。每个参与人的偏好由他的支付的期望值表示,对纯结果来说支付类似于图 16.1 所示。试将上述情形建成贝叶斯模型并找到对所有可能概率的纳什均衡。并特别证明存在这样的均衡:在此均衡里存在两个参与人不去同一音乐会的正概率。

?练习 28.1　**交换博弈**(An exchange game)每个参与人收到一张票,票上有一个在区间[0,1]的某个有限子集合 S 中的数字。参与人票上的数字是他可能得到的价格。两个价格具有独立同分布函数 F。每个参与人被单独、同时询问他是否愿把自己的价格与另一人的价格交换。若两个人都同意则交换成功,否则每个参与人只接受自己的价格。每个参与人的目标是使自己的期望支付最大化。试将此情形建成贝叶斯博弈模型并证明在任何纳什均衡中每个参与人愿交换的最高价是最小的可能价。

?练习 28.2　通过构造一个具有如下特征的两人贝叶斯博弈来说明更多的信息可能损害某个参与人;参与人 1 完全知道信息而参与人 2 并非如此;该博弈只有惟一纳什均衡,在此均衡中参与人 2 的支付高于他在任一个他知道参与人 1 类型的相应博弈的惟一均衡中的支付。

2.6.3 关于贝叶斯博弈模型的评价

参与人都不确知相互特征的情形可被建成贝叶斯博弈模型，在此模型里参与人的不确定性通过某个“状态”集合上的一个概率测度来表示，这些思想要归功于 Harsanyi(1967/68)。Harsanyi 假定每个参与人的先验概率都是一样的，他认为参与人知识的所有差异都应从赋予每个参与人信息的客观机制里导出，而不是来源于参与人初始信念的差异。在第 5.3 节我们指出了共同先验概率的假设对参与人后验概率间的关系具有较强的涵义。(例如，在一对参与人接收到他们的信号后下述这些内容不可能是他们间的“共同知识”，即参与人 1 相信自然状态的概率在某一给定集合下是 α 而参与人 2 相信此概率为 $\beta\neq\alpha$。尽管参与人 1 相信概 29
率为 α 且参与人 2 相信概率为 β 是可能的，但他们中的任何一人都并不确知另一位的信念。)

贝叶斯博弈不仅可用于对如练习 27.1 所述的每个参与人不确知别的参与人支付的情形建模，还可应用于每个参与人并不确知其他参与人知识(knowledge)的情形。

例如，考虑这样一个贝叶斯博弈：参与人集合为 $N=\{1,2\}$，状态集合为 $\Omega=\{\omega_1,\omega_2,\omega_3\}$，每个参与人的先验概率对每一状态赋概率1/3，信号函数定义为 $\tau_1(\omega_1)=\tau_1(\omega_2)=t'_1$，$\tau_1(\omega_3)=t''_1$，和 $\tau_2(\omega_1)=t'_2$，$\tau_2(\omega_2)=\tau_2(\omega_3)=t''_2$，参与人 1 的偏好对 $j=1,2$ 和行动组合 $b.c$ 满足 $(b,\omega_j)\succ_1(c,\omega_j)$ 及 $(c,\omega_3)\succ_1(b,\omega_3)$。而参与人 2 对所有 (a,ω) 均有同等偏好。状态 ω_1 下在这种博弈中参与人 2 知道参与人 1 对 b 的偏好优于 c，而状态 2 下他不知道参与人是更偏好 b 还是 c。因为状态 1 下参与人 1 不知道状态是 w_1 还是 w_2。此情形下她不知道(ⅰ)参与人 2 是否知道她对 b 的偏好优于 c，或是(ⅱ)是否参与人 2 不确知她是更偏好 b 还是更偏好 c。

参与人不确知相互知识的情形是否都能用贝叶斯博弈来建模呢？假定参与人的支付仅依赖于一个参数 $\theta\in\Theta$。有 X_i 表示每个参与人 i 的可能信念集合。那么任一参与人 j 的信念是一个 $\Theta\times X_{-j}$ 上的概率分布。也就是，任何参与人的信念集合都会涉及其他所有参与人的信念集合才能确定。这样对我们所提问题的回答就很有意义并且等同于这样的问题：是否我们能找到这样一个集合组合 $\{X_j\}_{j\in N}$，该集合满足对所有 $i\in N$。集合 X_i 同构于 $\Theta\times X_{-i}$ 上的概率分布集合，如果能这样，我们就可令 $\Omega=\Theta\times(\times_{i\in N}X_i)$ 为状态空间并用贝叶斯博弈模型去处理任何参与人不仅不确知每个其

他人的支付，而且还不确知每个其他人信念的情形。对这一问题的积极回答由 Mertens and Zamir(1985)给出。我们省略了证明。

[注解]

在 Borel(1921)和 von Neumann(1928)的文章中首次提出抽象战略博
30 弈的概念。Nash(1950a)在此类博弈的背景下形成了纳什均衡概念；其中所包含的基本思想至少可追溯到 Cournot(1838)。命题 20.3 证明的思想首创于 Nash(1950a, 1951)和 Glicksberg(1952)，尽管他们所证结论有细微差异。正如所述，此结论类似于 Nikaidô and Isoda(1955)中的定理 3.1。最大最小化思想至少在 18 世纪初就已存在(参看 Kuhn(1968))。命题 22.2 的主要思想应归于 von Neumann(1928)，严格竞争博弈理论由 von Neumann 和 Morgenstern(1944)发展。贝叶斯博弈由 Harsanyi(1967/68)定义和研究。

"囚徒困境"似乎最早在 Raiffa(1951)和 Flood(1952，与 Dresher 合作的论文)的未出版的文献中提出，该博弈的标准解释应归于 Tucker(参看 Raiffa(1992.p.173))。BoS 应归于 Luce 和 Raiffa(1957)。"鹰一兔"又被称作"斗鸡"。拍卖(例 18.1 和 27.1)最早由 Vickrey(1961)正式研究。例 18.4 中的消耗战应归于 Maynard Smith(1974)，例 18.6 中的置位博弈应归于 Hotelling(1929)，例 28.1 中的博弈应归于 Brams、Kilgour 和 Davis(1993)。

第3章 31

混合、相关及演进均衡

本章我们研究两个均衡概念：混合战略纳什均衡和相关均衡。在这些均衡中参与人的行动都是不确定的。我们还将简单考察一个纳什均衡的变形，它用来对一个演进过程建模。

3.1 混合战略的纳什均衡

3.1.1 定义

混合战略纳什均衡概念被设计用来对某种博弈的稳定状态建模，在此种博弈中参与人的选择并不确定而是受概率规则调节。我们首先讨论正式定义，然后再讨论它们的解释。

在前面章节中我们将一个战略博弈定义为一个三元组$\langle N,(A_i),(\succsim_i)\rangle$，这里每个参与人$i$的偏好关系$\succsim_i$被定义在行动组合集合$A=\times_{i\in N}A_i$上（见定义11.1）。在本章我们允许参与人的选择是非确定的，这样就需要给模型的基本元素增加一个每个参与人关于在A上不确定事件的偏好关系的详细说明。遵从于现代博弈理论的习惯，我们假定每个参与人i的偏好关系满足冯·诺依曼和摩根斯坦恩的假设，这样偏好关系就可用某一函数$u_i:A\to\mathbb{R}$的期望值来表示。因此本章中我们关于战略相互作用的模型是一个三维组$\langle N,(A_i),(u_i)\rangle$，它不同于我们原来定义的战略博弈，在这里对每一$i\in N$，$u_i:A\to\mathbb{R}$是一个期望值，代表了参与人$i$关于$A$上的不确定事件集合的偏好关系的函数。不过，我们均将这类模型简称为**战略博弈**。

令$G=\langle N,(A_i),(u_i)\rangle$是一个这样的战略博弈。我们用$\Delta(A_i)$来表示 32

A_i 上的概率分布集合并且称 $\Delta(A_i)$ 的一个元素为参与人 i 的一个**混合战略**(mixed strategy)。我们假定参与人的混合战略是独立随机化的。为明确起见，我们又称 A_i 的一个元素为一个**纯战略**(pure strategy)。对任一有限集 X 和 $\delta \in \Delta(X)$ 我们用 $\delta(x)$ 表示 δ 赋给 $x \in X$ 的概率并且将 δ 的支集(support)定义为由 $x \in X$ 构成的集合，其中 $\delta(x) > 0$。混合战略的一个组合 $(\alpha_j)_{j \in N}$ 产生了集合 A 上的一个概率分布。例如，如果每个 A_j 是有限的，则在给定随机化独立性的条件下行动组合 $a = (a_j)_{j \in N}$ 的概率便是 $\Pi_{j \in N}\alpha_j(a_j)$，这样参与人 i 对 $(\alpha_j)_{j \in N}$ 的估值是 $\sum_{a \in A}(\Pi_{j \in N}\alpha_j(a_j))u_i(a)$。

现在我们从 G 派生出另一个战略博弈，称之为 G 的"混合扩展"，在该博弈里每个参与人 i 的行动集合是他在 G 中混合战略集合 $\Delta(A_i)$。

▶定义 32.1　战略博弈 $\langle N, (A_i), (u_i)\rangle$ 的**混合扩展**(mixed extension)是这样一个战略博弈 $\langle N, (\Delta(A_i)), (U_i)\rangle$：$\Delta(A_i)$ 是 A_i 上的概率分布集合，并且 $U_i: \times_{j \in N}\Delta(A_j) \to \mathbb{R}$ 将 U_i 下关于 A 上由 α 引起的不确定事件的期望值赋给每个 $\alpha \in \times_{j \in N}\Delta(A_j)$（这样若 A 有限，则 $U_i(\alpha) = \sum_{a \in A}(\Pi_{j \in N}\alpha_j(a_j))u_i(a)$。

请注意每个函数 U_i 都是多重线性的。也就是，对任意混合战略组合 α、参与人 i 的任何混合战略 β_i 和 Y_i 以及任一 $\lambda \in [0,1]$ 我们有 $U_i(\alpha_{-i}, \lambda\beta_i + (1-\lambda)Y_i) = \lambda U_i(\alpha_{-i}, \beta_i) + (1-\lambda)U_i(\alpha_{-i}, Y_i)$。同时要注意在每一 A_i 为有限情形下，对任一混合战略组合 α 我们有

$$U_i(\alpha) = \sum_{a_i \in A_i} \alpha_i(a_i) U_i(\alpha_{-j}, e(a_i)) \tag{32.2}$$

这里 $e(a_i)$ 是参与人 i 的退化混合战略，它使 $a_i \in A_i$ 的概率为 1。

我们现在来定义本章要研究的主要均衡概念。

▶定义 32.3　**一个战略博弈的混合战略纳什均衡**(a mixed strategy Nash equilibrium of a strategic game)是该博弈混合扩展的纳什均衡。

假定 $\alpha^* \in \times_{j \in N}\Delta(A_j)$ 是 $G = \langle N, (A_i), (u_i)\rangle$ 的一个混合战略纳什均衡，该均衡中每个参与人 i 的混合战略 α_i^* 在它将概率 1 赋给一个元素(如 A_i 中的 a_i^*)的意义下是退化的。那么，因为 A_i 可用 $\Delta(A_i)$ 的一个子集来
33 确定，所以行动组合 a^* 是 G 的一个纳什均衡。反之，假定 a^* 是 G 的一个纳什均衡，那么由 U_i 在 α_i 中的线性可知，没有 A_i 中的行动之上的概率分布使参与人 i 获得的支付高于由 $e(a_i^*)$ 产生的支付，这样组合 $(e(a_i^*))$ 是

G 的一个混合战略。

我们刚才已经讨论了一个战略博弈的纳什均衡集是它的混合战略纳什均衡集合的一个子集。在第2章我们看到存在纳什均衡集为空的博弈。不过，如下所述，任意每个参与人只有有限多个行动的博弈至少有一个混合战略纳什均衡。

■命题 33.1 任何有限战略博弈都有一个混合战略纳什均衡。

证明：令 $G=\langle N,(A_i),(u_i)\rangle$ 是一战略博弈，并且对每一参与人 i 令 m_i 为集合 A_i 元素的个数。那我们就能用向量$(p_1,\cdots p_{m_i})$集合来确定参与人 i 的混合战略集合 $\Delta(A_i)$，其中对所有 k，$p_k\geqslant 0$ 且 $\sum_{k=1}^{m_i}p_k=1$（p_k 为参与人 i 使用第 k 个纯战略的概率）。这个向量集合是非空、凸和紧的。既然在概率上期望支付是线性的，那么每个参与人在 G 的混合扩展中的支付函数对于他自己的战略是拟凹和连续的。这样 G 的混合扩展满足命题 20.3 所有的条件。 □

对此证明最关键的假设是每个参与人 i 的行动集合是有限的。克里克斯伯格(Glicksberg, 1952)证明了若一个博弈中每个行动集合是欧氏空间的凸紧子集且每个支付函数是连续的，则它有混合战略纳什均衡。（如果每个参与人的支付函数在他的行动中也是拟凹的，那么命题 20.3 表明这样的博弈有纯战略纳什均衡）。

下面的结论给出了纳什均衡的一个非常重要的性质，它在计算均衡时很有用。

■引理 33.2 令 $G=\langle N,(A_i),(u_i)\rangle$ 为一有限战略博弈。那么 $\alpha^*\in\times_{i\in N}\Delta(A_i)$ 为 G 的一混合战略纳什均衡，当且仅当对每个参与人 $i\in N$，α_i^* 的支集上的每个纯战略是对 α_{-i}^* 的最优反应。

证明：先假定在 α_i^* 的支集中有一个行动 a_i，它不是 α_{-i}^* 的最优反应。接着由 U_i 在 α_i 中的线性性质（参看(32.2)）参与人 i 可通过从 a_i 向一个最优反应行动转移概率来增加他的支付；因此 α_i^* 不是对 α_{-i}^* 的最优反应。

接着假定有一混合战略 α_i' 提供的期望支付高于 α_i^* 对 α_{-i}^* 的反应支付。接着再由 U_i 的线性至少有一 α_i' 支集上的行动一定能比 α_i^* 支集上的行动提供更高的支付，因此并不是 α_i^* 支集上所有行动都是 α_{-i}^* 的最优反应。 34

应。 □

从而有：在任一个参与人的均衡混合战略支集中的每个行动都使该参与人产生同样支付。

如果某个参与人的行动集合不是有限的，那么结论就需要修改。在此情形下，α^* 是 G 的一个混合战略纳什均衡当且仅当（ⅰ）对于每个参与人 i 在给定 α_{-i}^* 的条件下没有 A_i 中的任何行动产生对参与人 i 的支付超过他的均衡支付，并且（ⅱ）在给定 α_{-i}^* 条件下，产生小于他的均衡支付的行动集合有 α_i^*——零测度。

请注意参与人的偏好可用期望支付函数来表示这一假设对这些混合战略均衡的特征具有重要意义。对于别的一些不确定性下的决策理论，这些结论不一定都有。

3.1.2 例

下面这个例子说明一个人如何能找到有限博弈的混合战略纳什均衡。

◇例 34.1 （BoS）考虑博弈 BoS，见图 35.1 上部分。在第 2 章我们将此表中参与人 i 的支付解释为代表了参与人 i 对（纯）结果集合的偏好。这里，给定我们在混合战略均衡中的兴趣，我们将支付解释为 v-N-M 效用。

正如我们以前所说明的那样，该博弈有两个（纯）纳什均衡，(B,B) 和 (S,S)，这里 $B = Bach$ 及 $S = Stravinsky$。假定 (α_1,α_2) 是一个混合战略纳
35 什均衡。若 $\alpha_1(B)$ 为 0 或 1，我们就获得两个纯纳什均衡，若 $0<\alpha_1(B)<1$，那么在给定 α_2 条件下，由引理 33.2 参与人 1 的行动 B 和 S 必定产生同一支付，这样我们一定有：$2\alpha_2(B)=\alpha_2(S)$，这样有 $\alpha_2(B)=\frac{1}{3}$。因为 $0<\alpha_2(B)<1$ 则可有同样结论，那参与人 2 的行动 B 和 S 一定产生同一支付，因此 $\alpha_1(B)=2\alpha_1(S)$，或 $\alpha_1(B)=\frac{2}{3}$。这样该博弈的惟一非退化混合战略纳什均衡是 $((\frac{2}{3},\frac{1}{3}),(\frac{1}{3},\frac{2}{3}))$。

在此博弈的混合扩展中建立参与人最优反应函数是富有启发性的。如果 $0\leqslant\alpha_2(B)<\frac{1}{3}$，那么参与人 1 的惟一最优反应 α_1 有 $\alpha_1(B)=0$，如果 $\frac{1}{3}<\alpha_2(B)\leqslant 1$，那么她的惟一最优反应有 $\alpha_1(B)=1$；如果 $\alpha_2=\frac{1}{3}$，那么，如我们前面所见，她所有混合战略都是最优的。对参与人 2 作一个类似的计算我们得到图 35.1 下半部所示的函数。

	Bach	Stravinsky
Bach	2,1	0,0
Stravinsky	0,0	1,2

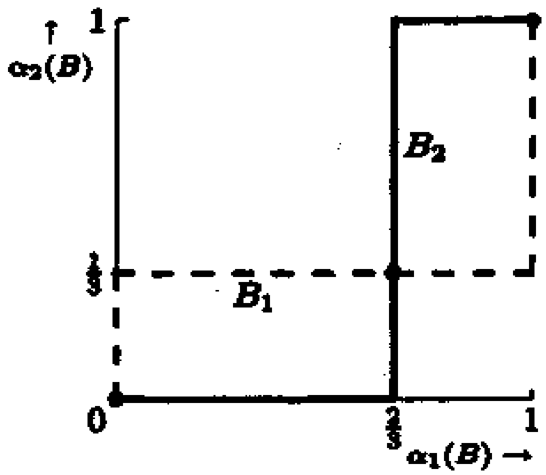


图 35.1 战略博弈 BoS(上半部)和该博弈混合扩展中的参与人最优反应函数(下半部)。参与人 1 的最优反应函数由虚线表示；参与人 2 的则由实线表 1 示。小圆盘表示两个纯战略纳什均衡及混合战略纳什均衡

?练习 35.1 猜均值(Guess the average) n 个人中的每一个人都说出集合 $\{1,\cdots,K\}$ 中的一个数。美元奖金在他们中说出的数字离这些数字均值的 $\frac{2}{3}$ 最近的人之间平分。试证明该博弈有惟一一个混合战略纳什均衡，其中每个参与人的战略是纯的。

?练习 35.2 **投资竞赛**(An investment race)两个投资者处在一场奖金为 1 元的竞赛中。每个投资者可花费区间 $[0,1]$ 中的任何一个数量。胜者是花费最多的投资者，在平局的情况下每个投资者得到奖金 0.5 元。试将此情形系统地表示成一个战略博弈并且找到它的混合战略纳什均衡。(注意参与人的支付函数是不连续的，这样 Glicksberg 的结论就不能用，不过该博弈有一混合战略纳什均衡。) 36

在第 2.5 节我们定义并研究了严格竞争博弈。我们说明了(命题 22.2)在任一个具有纳什均衡的严格竞争战略博弈中均衡集合是与最大最小化行动二元组集合相一致的。这一事实可被用来寻找混合扩展是严格竞

争的博弈的混合战略纳什均衡集合。(请注意这样的事实:一个博弈是严格竞争的并不意味着它的混合扩展是严格竞争的。为说明这点,可以考虑有三个可能结果果 a^1, a^2 和 a^3 的博弈。这样我们可有 $a^1 \succ_1 a^2 \succ_1 a^3$ 和 $a^3 \succ_2 a^2 \succ_2 a^1$,因此博弈是严格竞争的。但是两个参与人可能都对 a^2 的偏好优于对 a^1 和 a^3 以同样概率发生的不确定事件的偏好,所以它的混合扩展不是严格竞争的。)

?练习 36.1 猜对(Guessing right)参与人 1 和 2 每人选择集合 $\{1, \cdots, K\}$ 中的一个元素。若两个参与人选择同一个数则参与人 2 付给参与人 1 一元钱;否则不需付钱。每个参与人最大最小化他的期望货币支付。试找出这个(严格竞争的)博弈的混合战略纳什均衡。

?练习 36.2 空袭(Air Strike)A 军有一架飞机,用它可摧毁三个可能目标中的任一个。B 军有一门防空高炮,它能放在三个目标中的任一个里面。目标 k 的值为 v_k,且有 $v_1 > v_2 > v_3 > 0$。只要目标不设防且 A 去打击它,则该目标就可被摧毁。A 军希望最大化损失的期望值而 B 军希望最小化它。试将此情形系统地表示成一个(严格竞争的)战略博弈并找到它的混合战略纳什均衡。

?练习 36.3 试证明如下数学结论,我们将在练习 64.2 中应用它。对任何 $\mathbb{R}^k$ 上的两个紧凸子集 X 和 Y,存在 $x^* \in X$ 和 $y^* \in Y$ 满足对所有 $x \in X$ 和 $y \in Y$ 有 $x^* \cdot y \leqslant x^* \cdot y^* \leqslant x \cdot y^*$。[证明这个结论你既可以求助于在战略博弈中纳什均衡存在性(命题 20.3),也可以通过以下的基本证明(它可避免 Kakutani 不动点定理的盲目使用。)令 (x^k), (y^k) 各为 X 和 Y 上的稠密序列,并且对每个正整数 n 考虑每个参与人有 n 个行动和参与人 i 的支付函数由 $u_1(i, j) = x^i \cdot y^j$ 给定的严格竞争博弈,使用命题 33.1 及
37 22.2]。

3.2 关于混合战略纳什均衡的解释

在这部分我们讨论一些关于混合战略均衡的解释。一些观点我们并不赞成,仅代表我们中某一位观点的段落由该作者姓名中首个大写字母引导。

3.2.1　混合战略作为选择的目标

从朴素的观点看，通过参与人给他们的行为引进随机性，一个混合战略便产生了一个深思熟虑的决策：一个选择混合战略的参与人使他自己致力于一个随机方法，该方法就是随机地选择他的行动集合的元素。在所有参与人这样做后，方法就被实行了，同时一个行动组合也就实现了。这样每个参与人 i 选择了 $\Delta(A_i)$ 中的一个元素，其方式类似于我们在第 2 章所讨论的战略博弈中选择 A_i 中的一个元素。

当然有很多参与人将随机性引进他们行动的情形。例如，参与人在扑克牌游戏中随机地"虚张声势"，政府随机地稽核纳税人，一些商店随机地提供折扣等。

AR 在战略博弈中混合战略均衡这一概念并没有抓住参与人将随机性引进他们行动的动机。一般地参与人深思熟虑地去随机化是为了影响别的参与人的行动。例如考虑孩子式的猜迷博弈（例 17.1），在其中参与人选择是出单数手指还是出双数手指。这个博弈被经典地用来说明混合战略均衡这一概念的动机，但是随机化是关于博弈中参与人深思熟虑战略的奇怪描述。参与人的行动是对别人选择的猜测的反应；猜是种极精细而非随机的心理活动。另外，可考虑另一个说明混合战略均衡动机的例子，即征税当局与纳税人间的关系。当局的目的是阻止纳税人逃税，成本方面的考虑导致他们只能随机地稽核。他们希望纳税人知道他们的战略并且他们对是否稽核的战略并非无差别，就如同在一个混合战略均衡中所要求的那样。该情形应建成一个这样的博弈模型：当局首先选择稽核的概率，接着纳税人被告知这一概率而去采取行动。在这样一个模型里可能随机化集合是纯战略集合。

MJO 将参与人的均衡混合战略解释为一个深思熟虑的选择，其主要问 38
题是在一个混合战略均衡中，每个参与人对所有支集为她的均衡战略的子集的混合战略偏好一致：给定别的参与人的均衡行为，她的均衡战略只是众多使她获得相同预期支付战略中的一个。不过，这个问题不仅仅限于混合战略均衡。例如，它困扰着很多序贯博弈（包括所有重复博弈）中的均衡，在这些均衡里参与人对她的均衡战略和对很多非均衡战略偏好一致。进一步来说，在一些博弈里可能还有别的选择均衡混合战略的理由。例如，在严格竞争博弈中，我们已经看到一个均衡混合战略可以严格地最大化参与人能保证获得的支付。（例如在猜谜博弈便是如此）。最后，Harsanyi（1973）的

巧妙证明多少摆脱了均衡混合战略的这一特征。

MJO 看起来猜谜混合战略均衡似乎提供了一个关于与随机挑选的对手重复进行博弈的参与人的稳定状态行为的极好描述。此情形下，在任何一次交锋中参与人都没有猜测对手行动的方法，这样对她来说采用"最大化她能保证获得的支付"这一战略便是合理的。若两个参与人重复相互影响则猜测的心理可能会提供对他们行为的洞察力，尽管即使此情形下战略的混合战略均衡可以提供对他们行为的极好描述。税收稽核情形同样能被模化成参与人的选择是同时进行的战略博弈。此博弈中由当局选择的均衡稽核概率与当局首先行动的博弈是一样的，给定纳税人的行为，当局对稽核与否的偏好是一样的。

3.2.2 混合战略纳什均衡作为稳定状态

39 在第 2 章我们将纳什均衡解释为这样一种情形的稳定状态：参与人重复行动并且忽略任何存在于行动间的战略联系。类似地我们能将一个混合战略纳什均衡解释为一个随机稳定状态。参与人拥有过去行动被采用频率的信息（如"参与人作为参与人 1 的角色在 80％的时间采取行动 a_1，20％的时间采用行动 b_1"），每个参与人使用这些频率信息去形成他的关于别的参与人未来行动的信念，因此可以系统表达他的行动。在均衡中这些频率随时间保持不变并且在这样的意义下是稳定的：给定稳定状态信念，由参与人用正概率选择的任何行动是最优的。

混合战略均衡预示着博弈的结果是随机的，所以对博弈的单个行动来说它的预测比纯战略均衡的预测精确性差。但正如我们在第 1.5 节所讨论那样，理论的作用是用来解释规则的；混合战略均衡的概念抓住了随机规则。

此解释的一个变形建立在将一个 n 个参与人的博弈解释为一个 n 个大总体相互作用模型的基础上。在 n 个参与人被逐个从每个总体随机抽出来后，博弈的每个事件就发生了。在参与人 i 的均衡混合战略中的概率被解释为 A_i 中的元素在第 i 个总体中被使用的稳定状态概率。在这个解释下博弈就是模型的一种简化形式；其中总体被明确描述。

隐含于稳定状态解释之下的一个解释是：没有任何参与人看出存在于别的参与人行动间的或是存在于别的参与人行动与他自己行动间的任何联系。放弃这一假设将形成相关均衡概念，我们将在第 3.3 节讨论它。

3.2.3　混合战略作为扩展博弈的纯战略

在选择他的行动前，一个参与人可能收到他的行动或许要依赖的随机个人信息，从别的参与人的观点看，这些信息并不重要。参与人可能不会有意识地选择利用这种存在于他的行动与他得到的个人信息间的联系，这种联系也许仅是使他的行动被其他的参与人或外部观察者看来似乎是“随机的”。在把参与人的行为模化为随机的过程中，混合纳什均衡抓住了有关参与人视作不相关因素方面的行为依存关系。换言之，参与人可能知道决定他对手行为的外部因素，但是又可能发现要确定这种关系是不可能的或是成本极大的。（出于同样的理由，我们将掷币结果模化成随机的，而不将它 40
描述成以下因素相互作用的结果：它的起始位置和速度，风速及其他因素等）。简言之，从这个方式来看混合战略纳什均衡是对系统稳定状态的描述，这个系统反映了博弈初始描述所缺失的元素。

为了更具体形象，让我们来考虑博弈 BoS（例 34.1）。如我们原来所知，此博弈有一个混合战略纳什均衡$((\frac{2}{3},\frac{1}{3}),(\frac{1}{3},\frac{2}{3}))$。现在假设每个参与人有三种由他所不知的因素确定的“心情”。每个参与人在每个心情里呆1/3时间，并与别的参与人心情无关；他的心情不影响他的支付。假设：无论在心情 1 还是在心情 2，参与人 1 选择 Bach，当在心情 3 时选择 Stravinsky；在心情 1 参与人 2 选择 Bach，在心情 2 或心情 3 下选择 Stravinsky。将此情形视为每个参与人三种类型对应于他的可能心情的贝叶斯博弈，则此行为确定了一个恰与原始 BoS 博弈的混合纳什均衡相对应的纯战略纳什均衡。请注意这个混合战略均衡的解释不依赖于每个参与人具有三个相同可能性的且独立的心情，我们仅需要参与人个人信息满足它们能产生恰当的随机变量。不过，对这种信息结构存在的要求限制了解释。

AR 关于该解释有三个批评意见。首先，参与人深思熟虑的行动依赖于不影响他支付的因素是很难接受的。人们通常为他们的选择给出理由；在任何特定情形下，希望应用混合战略均衡概念的建模者(modder)都应该指出支付不相关的理由及解释参与人个人信息与他的选择间所需要的依存关系。

MJO 混合战略均衡中每个人对她的均衡战略支集上的所有行动具有相同偏好，所以说选择的行动依赖于被建模者认为“不相关的”因素不是没道理的。当被问到他们为什么从每个元素都一样具有吸引力的集合中选定

一个行动，人们经常给出诸如“我不知道——我只是觉得喜欢它”之类的答案。

AR 其次，此种解释下由均衡所预测的行为是非常脆弱的。如果一个经理的行为是由他所吃早餐的类型决定的，那么模型之外的种种因素，例如他的饮食习惯或鸡蛋价格的变化，都可能改变他用于选择他行动的概率，这样就会导致关于别的参与人概率的变化及造成不稳定性。

MJO 对于随机事件的每个结构都有一个导致同一均衡的行动模式。
41 例如，如果在鸡蛋价格上涨前有一个均衡，在此均衡里经理在早餐吃鸡蛋及早上 7:30 前起床的日子里提供折扣；那么在价格上涨之后可能有这样的均衡，即经理当她吃鸡蛋和上午 8:00 以前起床时才提供折扣。在价格改变之后她原有的行为模式对别的参与人的战略就不再是最优反应；系统是否在一个稳定的方式下调节到一个新均衡依赖于调节过程。一个混合战略纳什均衡在下列的意义下是脆弱的：参与人没有正面激励去坚持他们均衡行为模式（因为均衡战略并不是惟一最优的）；除此之外，在此解释下的均衡不比在任何别的解释下更脆弱。（再一次指出，这是一个由 Harsanyi 模型所阐述的问题，将在下节讨论。）

AR 最后，为了用这种方式去解释一个特定问题的均衡，就需要去指出“真实生活”以外的变量，参与人将他们的行为建立在这些变量之上。例如，为了解释价格竞争模型中的混合战略纳什均衡，人们就应该既要具体指出模型中没有的，但又要作为公司制定价格政策基础的各因素，又要指出信息结构足够丰富使它能跨越所有混合战略纳什均衡集合。这些应用混合战略均衡概念的人很少这样做。

MJO 世上的参与人能见到众多她的行动可依赖的随机变量：她早上起床的时间，她所处的心情，她的报纸被送来的时间，……这些随机变量的结构是如此丰富以至于不必要在每个理论应用中将它们写出来。为了更好地将一个更大博弈中的混合战略解释为纯战略就应该抓住这样的思想，即参与人选择的行动可能依赖于模型之外的因素。

3.2.4 混合战略作为不确定化博弈中的纯战略

对由 Harsanyi(1973)提出的混合战略均衡，我们现在提出一个理论根
42 据。若在一个博弈中参与人的偏好服从于一个小随机变化，则该博弈视为一种经常发生的情形。（这样，如同在前面部分讨论的那样，随机因素被引进了，但这里它们是与支付不相关的。）在情形的每次发生中，每个参与人知

道他自己的偏好而不知其他参与人的偏好。一个混合战略均衡就是对参与人随时间选择他们行动频率的一个概括。

令 $G=\langle N,(A_i),(u_i)\rangle$ 为一有限战略博弈，$\epsilon=(\epsilon_i(a))_{i\in N,a\in A}$ 为区间$[-1,1]$中随机变量的一个族，这里 $\epsilon_i=(\epsilon_i(a))_{a\in A}$ 有一个连续可微的密度函数和一个绝对连续的分布函数，并且随机向量 $(\epsilon_i)_{i\in N}$ 是独立的。考虑一组不确定化博弈，其中每个参与人 i 对结果 a 的支付服从小随机变化 $\epsilon_i(a)$，每个参与人 i 知道 ϵ_i 的实现值(realization)为 $(\epsilon_i(a))_{a\in A}$，但不知道别的参与人随机变量的实现情况。这样，可考虑贝叶斯博弈 $G(\epsilon)$，其中自然状态集合为 ϵ 所有可能实现值的集合，每个参与人的(共同)先验概率是由 ϵ 确定的概率分布，参与人 i 的信号函数仅告诉他 $(\epsilon(a))_{a\in A}$ 的实现情况及参与人在结果为 a 和状态为 ϵ 情况下他的支付是 $u_i(a)+\epsilon_i(a)$。(请注意每个参与人有无限多类型。)

Harsanyi 的主要结论(1973，定理2和7)是：对几乎任一博弈 G 和任一满足上述条件的随机变量族 ϵ^*，几乎任一 G 的混合战略纳什均衡都是与某个极限相联系的混合战略组合，而该极限为贝叶斯博弈 $G(\gamma\epsilon^*)$ 的纯战略均衡序列，随不确定性的大小，γ 逐渐消失的极限，当然 $G(\gamma\epsilon^*)$ 应满足对于 $G(\gamma\epsilon^*)$ 中的每个博弈而言由每个类型的参与人所选择的行动都是严格最优的。进一步来说，任何这种收敛序列的极限都与 G 的一个混合战略均衡相联系(Harsanyi(1973，定理5))。也就是，当支付中的随机变化小的时候，博弈 G 的几乎任一混合均衡都接近于相联系的贝叶斯博弈的纯均衡，反之亦然。我们称具有这种性质的 G 的混合战略均衡在 ϵ^* 下是**可逼近的**(approachable)(由于与这些结论相关的数学的复杂性，我们没引进证明。)

☑ 练习 42.1 考虑两个参与人的博弈，其中每个参与人 i 有两个纯战略 a_i 和 b_i。对 $i=1,2$，令 δ_i 为独立的随机变量，δ_i 具有$[-1,1]$上的同一分布，并且对 $i=1,2$ 和 $a\in A$，令随机变量 $\epsilon_i(A)$ 具有性质：对 $x=a_2,b_2$，有 $\epsilon_1(a_1,x)-\epsilon_1(b_1,x)=\delta_1$ 和对 $x=a_1,b_1$，有 $\epsilon_2(x,a_2)-\epsilon_2(x,b_2)=\delta_2$。

a. 试证明在 ϵ 下 BoS(例15.3)的所有均衡都是可能接近的。 43

b. 对于 $u_i(a_1,a_2)=1(i=1,2)$ 和所有其他支付为 0 的博弈，试证明只有纯战略纳什均衡 (a_1,a_2) 在 ϵ 下是可逼近的。

c. 对于 $u_i(a)=0(i=1,2$；所有的 $a\in A$ 的博弈，试证明仅有混合战略纳什均衡 α 在 ϵ 下是可逼近的。其中：对 $i=1,2$，$\alpha_i(a_i)=\alpha_i(b_i)=\frac{1}{2}$。

(别的均衡在别的不确定性下是可逼近的。)

这样 Harsanyi 对混合战略均衡的理论根据是：即使没有参与人根据需要的频率去使用他的纯战略，支付函数中的随机变化也会导致每个参与人根据正确的频率去选择他的纯战略。别的参与人的均衡行动使得自己支付函数的每个实现值选择惟一最优纯战略的参与人根据他的均衡混合战略所需的频率来选择他的行动。

MJO Harsanyi 的结论是对下列断言的极好反应：因为参与人对所有具有相同支集的战略无偏好差别，所以她没有理由去选择她的均衡混合战略。在上面我们已论证了对一些博弈，(包括严格竞争博弈)这个批评是中肯的。因为参与人有别的理由去选择他们的均衡混合战略。Harsanyi 的结论表明，在几年中任一个博弈批评的力量都是有限的，因为几乎任一混合战略纳什均衡都接近于博弈任一不确定化的一个严格纯战略均衡，其中参与人支付服从小随机变化。

3.2.5 混合战略作为概率

在另一种解释下(这种解释在第 5.4 节我们将仔细分析)，一个混合战略纳什均衡是一个概率(beliefs)组合 β，在 β 中 β_i 是所有其他参与人关于参与人 i 的行动的**共同**(common)信念。β_i 的性质是，对每个参与人 i，在给定 β_{-i} 条件下，每个 β_i 支集上的行动都是最优的。在此解释下，每个参与人都选择一个单一行动而非一个混合战略。均衡是参与人概率的稳定状态，而不是他们行动的稳定状态。这些信念需满足两个性质：在所有参与人中它们是共同的，并且与每个参与人是一个期望效用最大化者的假设相一致。

若欲从此思想出发，我们将形成如下的均衡概念：

44 ▶定义 44.1 一个有限战略博弈的**混合战略纳什均衡**是混合战略组合 α^*，α^* 满足：对每个参与人 i，α_i^* 支集上的每个行动都是对 α_{-i}^* 的最优反应。

引理 33.2 表明了该定义等同于我们以前的定义(32.3)，并且这也保证了该思想的确是对混合战略均衡的一个解释。

不过，请注意，当我们用这种方式来解释混合战略均衡时，均衡的预测内容是少的：它仅预测了每个参与人使用一个对均衡概率反应最优的行动。这些最优反应集合包括了在参与人均衡混合战略支集上的任何行动，甚至

也可能包括在战略的支集之外的行动。

3.3　相关均衡

在第 3.2.3 节我们讨论了将一个混合战略纳什均衡解释成一个稳定状态，在此状态中每个参与人的行动依赖于他从"自然"收到的一个信号。在此解释中信号是私有的，也是独立的。

若信号不是私有的和独立的又会怎样呢？例如，假设在 BoS（见图 35.1）中两个参与人都观察到一个随机变量，该随机变量以概率 $\frac{1}{2}$ 取 x 值或 y 值。这样就有一个新均衡，在均衡中若实现值为 x，则两个参与人都选 Bach，若实现值为 y，则选 Stravinsky。给定每个参与人的信息，他的行动是最优的：如果实现值为 x，那么他知道别的参与人选择 Bach，所以对他来说选 Bach 是最优的，若实现值为 y 则有对称结果。

此例中参与人观察到同一随机变量。更一般地，他们的信息可能并非完全相关。例如假设有一随机变量可取三个值 x，y 和 z，参与人 1 仅知道实现值要么为 x 要么为 $\{y,z\}$ 中的一个，而参与人 2 仅知道它要么为 $\{x,y\}$ 中的一个，要么为 z。也就是，参与人 1 的**信息分割**是 $\{\{x\},\{y,z\}\}$，参与人 2 的是 $\{\{x,y\},\{z\}\}$。在这些假设下，参与人 1 的一个战略包括两个行动：一个是当她知道实现值为 x 所采取的；另一个是当她知道实现值为 $\{y,z\}$ 中的一个所采取的。同理，参与人 2 的一个战略包括两个行动：一个对 $\{x,y\}$，另一个对 z。在下列条件下，一个参与人的战略是最优的，即给定其他参与人的战略，对于他得到的全部信息，他选择任一个不同于由他的战略所决定的行动，他都不能做得更好。为了说明一个参与人在选择最优 45
行动时如何利用他的信息，假定 y 和 z 的概率为 η 和 ζ 并且参与人 2 的战略是：若他知实现值在 $\{x,y\}$ 中，则他选择行动 a_2，若他知道实现值是 z 则采取行动 b_2，那么如果参与人 1 知道 y 或 z 中的某个已发生了，则在给定参与人 2 以概率 $\eta/(\eta+\zeta)$（以 $\{y,z\}$ 为条件的 y 的概率）选择 a_2 及以概率 $\zeta/(\eta+\zeta)$ 选择 b_2 的条件下参与人 1 选择最优行动。

这些例子使我们有下列均衡的概念。

▶**定义 45.1**　**一个战略博弈** $\langle N,(A_i),(u_i)\rangle$ **的相关均衡包括**

• 有限概率空间(Ω,π)（Ω是状态集合，π是Ω上的一个概率测度）

• 对每个参与人$i\in N$，Ω的一个分割$\mathcal{P}_i$（参与人i的**信息分割**(information partition)）

• 对每个$i\in N$，函数$\sigma_i:\Omega\to A_i$，满足对某个$P_i\in\mathcal{P}_i$，只要$\omega\in P_i$，$\omega'\in P_i$，则有$\sigma_i(\omega)=\sigma_i(\omega')$（$\sigma_i$是参与人$i$的战略）

这样对每个$i\in N$及每个函数$\tau_i:\Omega\to A_i$，该函数满足对某一$P_i\in\mathcal{P}_i$。（即对参与人i的每个战略）只要$\omega\in P_i$及$\omega'\in P_i$，则有$\tau_i(\omega)=\tau_i(\omega')$，我们有：

$$\sum_{\omega\in\Omega}\pi(\omega)u_i(\sigma_{-i}(\omega),\sigma(\omega))\geqslant\sum_{\omega\in\Omega}\pi(\omega)u_i(\sigma_{-i}(\omega),\ \tau_i(\omega)).\tag{45.2}$$

请注意概率空间和信息分割不是外来的而是均衡的组成部分。同时也要注意(45.2)等同于这样的必要条件：给定别的参与人的战略和参与人i关于ω的知识，对以正概率发生的每个状态ω行动$\sigma_i(\omega)$是最优的。（这个等价依赖于参与人的偏好服从期望效用理论的假设。）

我们先说明相关均衡集合包括混合战略纳什均衡集合。

■命题 45.3　一个有限战略博弈$\langle N,(A_i),(u_i)\rangle$的每一个混合战略纳什均衡$\alpha$都有一个相关均衡$\langle(\Omega,\pi),(\mathcal{P}_i),(\sigma_i)\rangle$，其中对每个参与人$i\in N$由$\sigma_i$导致的$A_i$上的分布是$\alpha_i$。

46 证明：令$\Omega=A(=\times_{j\in N}A_j)$并且由$\pi(a)=\Pi_{j\in N}\alpha_j(a_j)$定义$\pi$。对每个$i\in N$和$b_i\in A_i$令$P_i(b_i)=\{a\in A:a_i=b_i\}$且令$\mathcal{P}_i$包含$|A_i|$个集合$P_i(b_i)$。对每一$a\in A$由$\sigma_i(a)=a_i$定义$\sigma_i$。那么$\langle(\Omega,\pi),(\mathcal{P}_i),(\sigma_i)\rangle$是一相关均衡，因为(45.2)对每一战略$\tau_i$都满足：左边是参与人$i$在混合战略纳什均衡$\alpha$中的支付，右边是当他使用以概率$\alpha_i(a_i)$选择行动$\tau_i(a)$的混合战略和别的参与人$j$使用混合战略$\alpha_j$时他的支付。进一步说，由$\sigma_i$导致的$A_i$上的分布是$\alpha_i$。 □

下面的例子是对本部分开篇例子的正式表示。

◇例 46.1　BoS 中（例 34.1）三个混合战略纳什均衡支付组合是(2, 1)，(1,2)和$(\frac{2}{3},\frac{2}{3})$。并且某个相关均衡产生支付组合$(\frac{3}{2},\frac{3}{2})$：令$\Omega=$

$\{x,y\}$，$\pi(x)=\pi(y)=\frac{1}{2}$，$\mathcal{P}_1=\mathcal{P}_2=\{\{x\},\{y\}\}$，$\sigma_i(x)=$ Bach 且 $\sigma_i(y)=$ Stravinsky($i=1,2$)。这个均衡的一个解释是参与人观察一个当面掷币的结果，它决定了他们所博弈的均衡是两个纯战略纳什均衡中的哪一个。

这个例子表明了下面的结论：

■命题 46.2 令 $G=\langle N,(A_i),(u_i)\rangle$ 为一战略博弈。G 的相关均衡支付的任一凸组合仍是 G 的一个相关均衡支付。

证明：令 $u^1,\cdots,u^K$ 为相关均衡支付组合且令 $(\lambda^1,\cdots,\lambda^K)\in\mathbb{R}^K_+$ 对所有 k 有 $\lambda^k\geqslant 0$ 且 $\sum_{k=1}^K\lambda^k=1$。对 k 的每个值令 $\langle(\Omega^k,\pi^k),(\mathcal{P}_i^k),(\sigma_i^k)\rangle$ 为一产生支付组合 u^k 的相关均衡；不失一般性假设集合 Ω^k 是不相交的。则下述内容便定义了一个支付组合为 $\sum_{k=1}^K\lambda^K\mu^k$ 的相关均衡：令 $\Omega=U_k\Omega^k$，且对任一 $\omega\in\Omega$ 由 $\pi(\omega)=\lambda^k\pi^k(\omega)$ 定义 π，这里 k 满足 $\omega\in\Omega^k$。对任一 $i\in N$，令 $\mathcal{P}_i=U_k\mathcal{P}_i^k$ 且由 $\sigma_i(\omega)=\sigma_i^k(\omega)$ 定义 σ_i，这里 k 使得 $\omega\in\Omega^k$。 □

我们可将本证明中构造的相关均衡解释如下：首先一个公共随机方法决定 K 个相关均衡中哪一个将被形成，接着与第 k 个相关均衡相对应的随机变量被实现。

◇例 46.3 考虑图 47.1 左边的博弈：纳什均衡支付组合是(2,7)和(7,2)(纯的)和$(4\frac{2}{3},4\frac{2}{3})$(混合的)。下列的相关均衡产生一个位于这三个组合凸包之外的支付组合。令 $\Omega=\{x,y,z\}$ 和 $\pi(x)=\pi(y)=\pi(z)=$ 47
$\frac{1}{3}$，令参与人 1 的分割为 $\{\{x\},\{y,z\}\}$ 和参与人 2 的为 $\{\{x,y\},\{z\}\}$。将战略定义如下：$\sigma_1(x)=B$ 及 $\sigma_1(y)=\sigma_1(z)=T$；$\sigma_2(x)=\sigma_2(y)=L$ 及 $\sigma_2(z)=R$。(选择与状态间相关关系如图 47.1 右边所示。)在给定参与人 2 的行动下参与人 1 的行动是最优的：在状态 x，参与人 1 知道若参与人 2 采用 L 则对她来说采用 B 是最优的，在状态 y 和 z，若她给参与人 2 采用 L 和 R 赋同样概率，则对她来说采用 T 是最优的，对称地，给定参与人 1 的行动，参与人 2 的行动也是最优的，因此我们有一个相关均衡，支付组合是(5,5)。

这个例子中我们能用结果的集合确定状态的集合，表明了下列结论。

■命题 47.1 令 $G=\langle N,(A_i),(u_j)\rangle$ 为一有限战略博弈。在 G 的一

个相关均衡中所能获得的结果上的每一个概率分布都能在这样的一个相关均衡中获得，即状态集合为 A 并且对每个 $i\in N$ 参与人 i 的信息分割包括(对某个行动 $b_i\in A_i$)所有形式为$\{a\in A:a_i=b_i\}$的集合。

证明：令$\langle(\Omega,\pi),(\mathcal{P}_i),(\sigma_i)\rangle$为 G 的一相关均衡。那么，$\langle(\Omega',\pi'),(\mathcal{P}_i'),(\sigma_i')\rangle$也是一个相关均衡，这里 $\Omega'=A$。对每个 $a\in A$ 有$\pi'(a)=\pi(\{\omega\in\Omega:\sigma(\omega)=a\})$，$\mathcal{P}_i'$ 包括类型$\{a\in A:a_i=b_i\}$(对某个 $b_i\in A_i$)的集合，并且 σ_i'由$\sigma_i'(a)=a_i$ 所确定。 □

这个结论让我们在计算相关均衡支付时把注意力集中到状态集合是结果集合的均衡。不过要注意这一均衡可能无自然的解释。

	L	R
T	6,6	2,7
B	7,2	0,0

	L	R
T	y	z
B	x	–

图 47.1 相关均衡的一个例子。左边为一个战略博弈。右边的表格作为一个博弈相关均衡的状态的函数给出了参与人的选择

48 在相关均衡的定义中我们假定了所有参与人具有关于状态发生概率的共同信念。如果有一个参与人持不同信念的随机变量，则附加的均衡支付组合是可能的。假定参与人 1 确信在某些比赛里队 T_1 将击败队 T_2，而参与人 2 确信队 T_2 将赢！那么就有 BoS(例 34.1)的一个均衡：若 T_1 赢则结果为(Bach, Bach)，若 T_2 赢则结果为(Stravinsky, Stravinsky)，它给每个参与人一个期望支付 2！(第 5.3 节我们将说明：如果两个参与人有同样的先验概率则他们的信念如我们假设那样相异就不可能是参与人间的共同知识了。)

? 练习 48.1 考虑由图 48.1 给定支付的三人博弈。(参与人 1 选择两行中的某行，参与人 2 选择两列中的某列，参与人 3 选择三表中的某表。)

a. 试证明纯战略均衡支付是(1,0,0)，(0,1,0)，和(0,0,0)。

b. 试证明有一个这样的相关均衡：参与人 3 选择 B，参与人 1 和 2 以同样概率选择(T,L)和(B,R)。

c. 解释如下涵义：参与人 3 宁愿不去拥有参与人 1 与参与人 2 用于协调他们行动的信息。

	L	R
T	0,0,3	0,0,0
B	1,0,0	0,0,0

A

	L	R
T	2,2,2	0,0,0
B	0,0,0	2,2,2

B

	L	R
T	0,0,0	0,0,0
B	0,1,0	0,0,3

C

图 48.1　一个三人博弈。参与人 1 选择两行中的某一行，参与人 2 选择两列中的某一列，参与人 3 选择三个表中的某一表

3.4　演进均衡

本部分我们讨论隐含在被称为演进均衡(evolutionary equilibrium)的一种纳什均衡概念变形之中的基本思想。这个概念被设计用来对参与人的行动是由演进的力量所决定的情形建模。我们将讨论限于一种简单情形，在其中有机体(动物、人类、植物，……)单个总体的成员们两两相互作用。在 49
每个竞争中每个有机体从集合 B 中选取一个行动。有机体不是有意识地选择行动，而是他们要么从他们的先辈那里继承行为模式，要么由变异赋给他们。我们假设有一函数 u 测度每个有机体的生存能力：如果一个有机体面对它的潜在对手总体中的行动分布 β 它选择行动 a，那么它的生存能力由 β 下的期望 $u(a,b)$ 来测度。这个描述对应于一个两人对称战略博弈 $\langle\{1,2\},(B,B),(u_i)\rangle$，这里 $u_1(a,b)=u(a,b)$ 且 $u_2(a,b)=u(b,a)$。

演进均衡的一个备择解是 B 中的一个行动。均衡的概念是被设计用来抓住一个稳定状态，在其中所有有机体采取这个行动并且没有变异体侵入总体。更精确地说，该思想是对每个可能行动 $b\in B$ 进化过程有时会将总体的一小部分转变成采取行动 b 的变异体。在均衡中任何这样的变异一定获得一个比均衡行动的期望支付低的期望支付，因此它会消失。现在，如果总体的 $\epsilon>0$ 部分包含采取行动 b 的变异而所有其他有机体采取行动 b^*，则一个变异体的平均支付为 $(1-\epsilon)u(b,b^*)+\epsilon u(b,b)$(因为它以概率 $1-\epsilon$ 碰到一个非变异体而以概率 ϵ 碰到另一个变异体)而非变异体的平均支付是 $(1-\epsilon)u(b^*,b^*)+\epsilon u(b^*,b)$。因此对于 b^* 为一演进均衡我们要求：对所有 ϵ 充分小的值

$$(1-\epsilon)u(b,b^*)+\epsilon u(b,b)<(1-\epsilon)u(b^*,b^*)+\epsilon u(b^*,b)$$

此不等式是成立的当且仅当对每一 $b \neq b^*$，要么 $u(b,b^*)<u(b^*,b^*)$ 要么 $u(b,b^*)=u(b^*,b^*)$，且 $u(b,b)<u(b^*,b)$。由此我们可将演进均衡定义如下：

▶定义 49.1　令 $G=\langle\{1,2\},(B,B),(u_i)\rangle$ 为一对称战略博弈，这里对某一函数 u 有 $u_1(a,b)=u_2(b,a)=u(a,b)$。G 的一个**演进稳定战略**(evolutionarily stable strategy)(ESS)是一个行动 $b^* \in B$，其满足 (b^*,b^*) 是 G 的一个纳什均衡且对于 b^* 的每个最优反应 $b \in B$ 且 $b \neq b^*$ 有 $u(b,b)<u(b^*,b)$。

在下列例子中，如在很多经典文献中一样，集合 B 被看作是某个有限行动集合上的混合战略集合。

◇例 49.2　鹰—鸽(Hawk－Dove)每时每刻一个总体中的动物都在成对为一个具有价值 1 的猎物而争斗。每只动物能像鸽子(D)或像鹰(H)一
50 样行动。在竞争中如果两只动物都是鸽式的，则它们平分猎物的价值；如果都是鹰式的，则猎物的价值减少 c 且在它们间平分；如果有一个是鹰式的，另一个是鸽式的，则鹰式的得 1，鸽式的为 0。该博弈用图 50.1 表示，(若 $c>1$，则它有如图 17.2 的同样结构)，令 B 是 $\{D,H\}$ 上所有混合战略纳什均衡，在其中每个参与者使用战略 $(1-1/c, 1/c)$；该战略是惟一的 ESS。(特别地，在此情形中排除了鹰的总体不是进化稳定的。)若 $c<1$，博弈有惟一混合战略纳什均衡，在其中每个参与者使用纯战略 H；此战略是惟一的 ESS。

	D	H
D	$\frac{1}{2},\frac{1}{2}$	$0,1$
H	$1,0$	$\frac{1}{2}(1-c),\frac{1}{2}(1-c)$

图 50.1　一个鹰—鸽博弈

从定义 49.1 立即可知如果(b^*,b^*)是一对称纳什均衡并且仅有 b^* 是对 b^* 的最优反应(即，(b^*,b^*)是严格均衡)那么 b^* 是一个 ESS。一个非严格均衡战略可能不是一个 ESS：考虑两个参与人的对称博弈，其中每

γ,γ	$1,-1$	$-1,1$
$-1,1$	γ,γ	$1,-1$
$1,-1$	$-1,1$	γ,γ

图 50.2 无 ESS 的博弈。每个纯战略给变异体产生的支付高于惟一对称均衡混合战略

个参与人有两个行动且对所有$(a,b)\in B\times B$ 有 $u(a,b)=1$。一个更有趣的非严格均衡战略不是 ESS 的例子，可考虑图 50.2 中的博弈，其中 B 包括在一个包含三个元素的集合上的所有混合战略且 $0\leqslant\gamma\leqslant1$。该博弈有惟一对称混合战略纳什均衡，在均衡中每个参与人的混合战略是$(\frac{1}{3},\frac{1}{3},\frac{1}{3})$；每个参与人的期望支付是 $\gamma/3$。使用三个纯战略中的任何一个的变异体，当它碰到一个非变异体时获得 $\gamma/3$ 的期望支付，而当它碰到另一个变异体时获得更高的期望支付 γ。因此均衡混合战略不是一个 ESS(从其可知不是每个有纳什均衡的博弈都有一个 ESS)。 51

? 练习 51.1 试证明：在每两个参与人的对称战略博弈中，每个参与人有两个纯战略且对四个战略组合支付是不一样的，则该博弈有一个是 ESS 的混合战略。

[注解]

混合战略的现代系统表达应归于 Borel(1921；1924, pp.204—221；1927)。尽管这些思想至少可追溯到 18 世纪早期(参看 Guilbaud (1961)和 kuhn (1968))。Borel 创立了某些特别严格竞争博弈混合战略纳什均衡存在的理论，von Neumann(1928)证明了对所有严格竞争博弈都存在一个均衡。我们证明的存在性结论(命题 33.1)(它包含了所有有限战略博弈)应归于 Nash (1950 a, 1951)。相关均衡的概念应归于 Aumann(1974)，他的论文同样也是第 3.3 节中另一内容的基础。进化稳定战略的思想应归于 Maynard Smith 和 Price(参看 Maynard Smith (1972))和 Maynard Smith and

Price (1973);也可参看 Maynard Smith (1974,1982)。

第3.2.2节提及的总体模型应归于 Rosenthal(1979)。在第3.2.3节讨论的将混合战略解释为扩展博弈中的纯战略思想和第3.2.4节的内容归于 Harsanyi(1973)。在第3.2.5节所给的关于混合战略纳什均衡的解释在 Aumann(1987a)中被讨论。在第3.2节给定的某些关于混合战略纳什均衡的批评摘录于 Rubinstein(1991)。第3.3节的例子应归于 Aumann(1974)。

我们对命题33.1的证明应归于 Nash(1950a),它求助了命题20.3而命题20.3的证明使用了 Kakutani 的不动点定理。Nash(1951)提供了命题33.1的另一个证明,该证明使用了更基本的 Brouwer 不动点定理,它应用了点值函数。

52 练习35.1中的博弈来源于 Moulin(1986,p.72)。练习36.3来源于 Arrow, Barankin, and Blackwell(1953)。

关于参与人的偏好不满足由效用函数所表示的必要假设时的混合战略纳什均衡的讨论参看 Crawford(1990)。我们在第3.4节所讨论的 ESS 概念已在很多方面被扩展了,请参看 Damme(1991, Chapter 9)。

我们并没阐述是否存在动态调整过程导致均衡这一问题。一个被称为**虚拟参与**(fictitious play)的这种过程,由 Brown(1951)一提出,最近又被重新考虑了。在这个过程中每个参与人经常选择一个对别的参与人过去行动统计频率的最优反应。Robinson(1951)说明了在任一严格竞争博弈中该过程收敛于一混合战略纳什均衡;Shapley(1964, Section 5)说明了在非严格竞争的博弈中它不必要是这样的。最近的研究关注于明确拥有进化及学习力的模型;对于这方面工作的介绍可参看 Battigalli, Gilli and Molinari (1992)。

第 4 章 53

可理性化和反复剔除劣行动[①]

本章我们考察在给定与下列观点相一致的信念条件下要求参与人的选择是最优的结果：每个参与人是理性的；每个参与人认为其他参与人是理性的，每个参与人认为其他参与人认为每个参与人是理性的，等等。

4.1 可理性化

在第 2、3 章我们讨论了战略博弈解的概念，在这些博弈中每个参与人的选择在给定他关于别的参与人行动的信念条件下都是最优的，且这种信念还必须是正确的。也就是，我们假定每个参与人知道别的参与人的均衡行动。如果参与人重复进入博弈所模化的情形，则他们能从他们所观察到的稳定状态行动中获得这一知识，不过如果博弈是瞬间事件，在其中所有参与人同时选择他们的行动，则每个参与人如何知道别的参与人的均衡行动是不清楚的；由于这个原因博弈论者们发展了解的概念，它不需引入那一假设。

本章我们研究一些这样的解的概念，即参与人关于每个其他人的行动的信念虽未被假设为正确的，但受理性思考的约束，每个参与人都相信由每个其他参与人所采取的行动都是对某一信念的反应。进一步说，每个参与人都假设其他参与人都用此公式推理，所以认为其他参与人都相信每个参与人的行动是对某一信念的最优反应，等等。

我们将要研究的解的概念弱于纳什均衡。实际上，在很多博弈中它们 54

① 原书作者认为本书中 dominated 应译为“被控的”，而不宜译为“劣的”。译者的这一译法仅供参考。——译者

不排除任何行动的采用。不过我们发现该方法的有趣之处在于，它利用了对于参与人知识假设的逻辑内涵，参与人的知识比那些前面章节所讨论的要少。

固定一战略博弈$\langle N,(A_i),(u_i)\rangle$(其中$u_i$的期望代表了参与人$i$关于$A=\times_{j\in N}A_j$(对每一$i\in N$)上不确定事件的偏好)。在本章中为了发展这一思想就不必假定每个参与人的行动集合A_i是有限的，尽管为了简单我们在某些讨论中采用这个假设。参与人i的一个信念(belief)(关于其他参与人行动的)是$A_{-i}(=\times_{j\in N\setminus\{i\}}A_j)$上的一个概率测度。注意这个定义允许参与人相信别的参与人的行动是相关的：该信念不必是每个行动集合A_j(对$j\in N\setminus\{i\}$)上的独立概率测度的乘积。同以前一样，在给定的信念下如果没有别的行动使参与人i获得更高支付，则参与人i的一个行动$a_i\in A_i$就是对该信念的一个最优反应(best response)。我们经常使用短语"参与人i认为某个别的参与人j是理性的"，这意味着：参与人i认为参与人j选择的任何行动都是对参与人j关于非j参与人行动信念的最优反应。

如果参与人i认为每个其他参与人j是理性的，则他必能将他的关于别的参与人行动的信念μ_i理性化如下：被信念μ_i赋给正概率的每个其他参与人j的任一行动都必定是对参与人j的信念的一个最优反应。如果参与人i进一步认为每个其他参与人j都认为每个参与人$h\neq j$(包括参与人i)是理性的，那么他(参与人i)必定对参与人j关于参与人h的信念的观念有看法。如果参与人i的推理有无限深度，我们就有下列定义：

▶定义 54.1　行动$a_i\in A_i$在战略博弈$\langle N,(A_i),(u_i)\rangle$中是**可理性化的**(rationalizable)。如果存在：

- 一个集合族$((X_j^t)_{j\in N})_{t=1}^{\infty}$，其中$X_j^t\subseteq A_j$(对所有$j$和$t$)，
- 参与人i的一个信念μ_i^1它的支集是X_{-i}^1的一个子集，并且
- 对每一$j\in N$，每一$t\geqslant 1$和每一$a_j\in X_j^t$，参与人j的一个信念$\mu_j^{t+1}(a_j)$它的支集是X_{-j}^{t+1}的一个子集。

这样

- a_i是对参与人i集合μ_i^1的最优反应
- $X_i^1=\varnothing$和对每一个$j\in N\setminus\{i\}$集合X_j^1是所有$a'_j\in A_j$的集合；这样在μ_i^1的支集上有某一a_{-i}满足$a_j=a'_j$
- 对每个参与人$j\in N$和每个$t\geqslant 1$，每个行动$a_j\in X_j^t$是对参与人j的

信念 $\mu_j^{t+1}(a_j)$ 的一个最优反应

• 对每一 $t \geq 2$ 和每一 $j \in N$ 集合 X_j^t 是所有 $a'_j \in A_j$ 的集合，这样有某 55
个参与人 $k \in N \setminus \{j\}$，某一行动 $a_k \in X_k^{t-1}$ 和在 $\mu_k^t(a_k)$ 支集上的某个 a_{-k} 满足 $a'_j = a_j$。

注意在形式上本定义第 2 部分的第二、四条件是多余的，我们引进它们是为了使定义更紧密地与我们所给的动机相对应。同时也要注意我们引进族 $((X_j^t)_{j \in N})_{t=1}^{\infty}$ 中的集合 X_i^1，只是为了简化概念，即使它被要求是空的，若 $|N| \geq 3$ 则 X_i^1 是惟一这种多余集合，而若 $|N| = 2$ 就有很多了（对任一奇数 t 的 X_i^t 和对 $j \neq i$，任一偶数 t 的 X_j^t）。

集合 X_j^1（对 $j \in N \setminus \{i\}$）被解释为：参与人 j 的由参与人 i 关于非 i 参与人行动的信念（它证明 i 选择 a_i 是有道理的）赋予正概率的行动组合。对任一 $j \in N$，X_j^2 的解释是：它是参与人 j 所有行动 a_j 的集合，这样就至少存在某个参与人 $k \neq j$ 的一个行动 $a_k \in X_k^1$，它由赋给 a_j 正概率的信念 $\mu_k^2(a_k)$ 证明为是有道理的。

为了说明该定义的意义，假设有三个参与人，他们中的每个人有两个可能行动 A 和 B。假设参与人 1 的行动 A 是可理性化的且参与人 1 在理性化中所使用的信念 μ_1^1 赋给参与人 2 和 3 的选择 (A, A) 或 (B, B) 正概率，则 $X_2^1 = X_3^1 = \{A, B\}$。参与人 2 的信念 $\mu_2^2(A)$ 和 $\mu_2^2(B)$ 证明他选择 A 和 B 影响参与人 1 和 3 的行动是有道理的；同样参与人 3 的信念 $\mu_3^2(A)$ 和 $\mu_3^2(B)$ 影响参与人 1 和 2。这四个信念不一定都导致关于参与人 1 的同一信念，且不一定都赋正概率给行动 A。集合 X_1^2 包括参与人 1 所有由 $\mu_2^2(A)$ 或 $\mu_3^2(A)$ 或 $\mu_2^2(B)$ 或 $\mu_3^2(A)$ 赋给正概率的行动。

这个可理性化定义等价于下列定义。

▶定义 55.1　在战略博弈 $\langle N, (A_i), (u_i) \rangle$ 中，一个行动 $a_i \in A_i$ 是**可理性化的**，如果对每个 $j \in N$ 有一集合 $Z_j \subseteq A_j$ 满足：

• $a_i \in Z_i$

• 每个行动 $a_j \in Z_j$ 是对参与人 j 的信念 $\mu_j(a_j)$（它的支集是 Z_{-j} 的一个子集）的最优反应。

注意若 $(Z_j)_{j \in N}$ 和 $(Z'_j)_{j \in N}$ 满足该定义则 $(Z_j \cup Z'_j)_{j \in N}$ 也满足该定义，所以可理性化的行动组合集合是最大集合 $\times_{j \in N} Z_j$，对于它 $(Z_j)_{j \in N}$ 满足该定义。 56

■引理　定义 54.1 与 55.1 是等价的

证明：根据定义 54.1 若 $a_i \in A_i$ 是可理性化的则可定义 $Z_i = \{a_i\} \cup (\cup_{t=1}^{\infty} X_i^t)$ 和 $Z_j = (\cup_{t=1}^{\infty} X_j^t)$（对每一 $j \in N \setminus \{i\}$）。根据定义 55.1，若它是可理性化的，则定义 $\mu_i^1 = \mu_i(a_i)$ 和 $\mu_j^t(a_j) = \mu_j(a_j)$（对每一 $j \in N$ 和每一整数 $t \geq 2$）。那么在定义 54.1 第二和第四部分所定义的集合 X_j^t 是 Z_j 的子集且满足在第一和第三部分中的条件。□

从定义 55.1 明显可知：任一有限博弈中参与人在某一混合战略纳什均衡中以正概率所采取的任何行动都是可理性化的（将 Z_j 作为参与人 j 的混合战略的支集）。下列结果表明对在某个相关均衡中以正概率所采取的行动同样是正确的。

■引理 56.2　参与人在有限战略博弈的相关均衡中以正概率所采取的每一行动都是可理性化的。

证明：用 $\langle N, (A_i), (u_i) \rangle$ 表示战略博弈；选择一个相关均衡，并且对每一个参与人 $i \in N$ 令 Z_i 为参与人 i 在均衡中以正概率采取的行动集合。那么在参与人 i 选择 a_i 的条件下，任一 $a_i \in Z_i$ 是对由非 i 参与人的战略所产生的 A_{-i} 的分布的一个最优反应。这个分布的支集是 Z_{-i} 的一个子集且由定义 55.1 可知 a_i 是可理性化的。□

在囚徒困境中（例 16.2）只有纳什均衡行动“坦白”是可理性化的。在第 2.3 节别的博弈中每一个参与人的两个行动都是可理性化的，因为在每个案例中两个行动都在某一混合战略纳什均衡中以正概率被采用。这样可理性化未将任何约束置于那些博弈的结果之上。对于很多别的博弈可理性化所给的约束都是弱的。不过，在一些博弈中可理性化提供的答案是明确的，就如下列练习所展示的那样。

?练习 56.3　找出图 57.1 中两人博弈的每个参与人可理性化行动的集合。

?练习 56.4　库诺特双头垄断（Cournot duopoly）考虑战略博弈 $\langle \{1, 2\}, (A_i), (u_i) \rangle$，其中对 $i = 1, 2$ 有 $A_i = [0, 1]$ 和 $u_i(a_1, a_2) = a_i(1 - a_1 - a_2)$。试证明每个参与人惟一的可理性化行动是他惟一的纳什均衡行动。

?练习 56.5　猜均值（Guess the average）在练习 35.1 的博弈中试证

	b_1	b_2	b_3	b_4
a_1	0,7	2,5	7,0	0,1
a_2	5,2	3,3	5,2	0,1
a_3	7,0	2,5	0,7	0,1
a_4	0,0	0,−2	0,0	10,−1

图 57.1　练习 56.3 中的两人博弈

明每个参与人的均衡行动是他惟一的可理性化行动。 57

?练习 57.1　假使两个参与人在单位区间选择位置 a_1 和 a_2；每个参与人都希望尽可能与另一个参与人靠近，每个参与人的支付是 $-|a_1-a_2|$。试证明每个参与人的每个行动都是可理性化的，而纳什均衡集合是 $\{(a_1,a_2):a_1=a_2\}$。现在假设每个参与人都知道他与对手的距离。通过增加这样的条件来修改定义 55.1，即理性化含有行动 a_i 和距离 d 的二元组 (a_i,d) 的信念的子集是 $\{a_j-d,a_j+d\}$ 的一个子集。试证明没有 $d>0$ 使得存在一个行动 a_i 满足 (a_i,d) 在此意义下是可理性化的，而 $(a_i,0)$ 对每个 a_i 都是可理性化的。

注意在定义 54.1 及 55.1 中我们将参与人 i 的信念作为 A_{-i} 上的一个概率分布，它允许每个参与人相信他的对手们间的行动是相关的。在很多经典文献中，参与人不允许有这样的信念：它被假设为每个参与人的信念是独立概率分布的乘积，每个别的参与人都有一个（这一限制在两人博弈中明显是不重要的）。这个假设是与在混合战略均衡概念之后的动因（motivation）相一致的。我们关于可理性化的定义要求在所有理性化的水平上参与人是理性的，关于可理性化的另一定义则还要求在所有理性化水平上假设信念是独立的。

如同图 58.1 中的博弈所表示的那样，两个定义有不同的内涵。在此博弈中有三个参与人，参与人 1 选择两行中的某行，参与人 2 选择两列中的某列，参与人 3 选择四个表中的某一表。所有三个参与人获得同一支付，支付由方框中的数字所给出。我们说在定义 54.1 和 55.1 的意义下参与人 3 的行动 M_2 是可理性化的，在此意义下一个参与人可能相信其对手们间的行动是相关的，但如果参与人局限于信念是独立概率分布的乘积，则他对手的 58

行动就不是可理性化的。为明白这点，请注意参与人 1 的行动 U 是对赋概率 1 给(L, M_2)的信念的最优反应和行动 D 是对赋概率 1 给(R, M_2)的信念的最优反应；同理，参与人 2 的两个行动都是对赋正概率给 U、D 和 M_2 的信念的最优反应。进一步说，参与人 3 的行动 M_2 是对这样信念的最优反应：参与人 1 和 2 以相同概率选取(U, L)和(D, R)。这样 M_2 在我们已定义的意义下是可理性化的(在定义 55.1 中取 $Z_1 = \{U, D\}$，$Z_2 = \{L, R\}$ 和 $Z_3 = \{M_2\}$)。不过，它对任一(独立的)混合战略二元组不是最优反应并且因此在修改的定义下不是可理性化的，在那定义中每个参与人的信念局限为独立概率分布的一个乘积。[为了使 M_2 成为一个最优反应，我们需要 $4pq + 4(1-p)(1-q) \geqslant \max\{8pq, 8(1-p)(1-q), 3\}$，这里$(p, 1-p)$和$(q, 1-q)$分别为参与人 1 和 2 的混合战略，对 p 和 q 的任何值不等式都不满足。]

	M_1: L	M_1: R	M_2: L	M_2: R	M_3: L	M_3: R	M_4: L	M_4: R
U	8	0	4	0	0	0	3	3
D	0	0	0	4	0	8	3	3

图 58.1 一个三人战略博弈。参与人 1 选择两行中的某行，参与人 2 选择两列中的某列，参与人 3 选择四表中的某表。所有三个参与人获得同一支付，支付由方框中的数字所给出

4.2 反复剔除强劣行动

同可理性化概念一样，我们现在所研究的解的概念也是从单个(single)参与人的角度来看一个博弈。每个参与人基于这样的计算来采取行动，即不需要有关别的参与人采取行动的知识。为了确定解，我们从剔除某个参与人确定的应该不采取的行动开始。在一个复杂的博弈中，这种假设是特别有价值的：参与人为了寻找简化他们所面对的情形的方法，将采用这一技巧。我们假设参与人不考虑别的参与人无论作什么都不是最优反应的行
59 动。一个知道别的参与人是理性的参与人能假定他们会从考虑中排除这些行动。现在考虑博弈 G' 是通过从初始博弈 G 中剔除所有这些行动所得到

的。再一次，在 G' 中知道别的参与人是理性的参与人，无论别的参与人干什么他都不会选择一个不是最优反应的行动。进一步说，在 G' 中一个知道别的参与人认为他是有理性的参与人，会确信他们也不会选择永非最优反应的行动。按此方法连续论证就意味着 G 的结果一定能经受住无穷次这样的剔除。现在我们系统表达这种思想并要说明它等价于可理性化概念。

4.2.1 永非最优反应

▶定义 59.1 在一战略博弈中参与人 i 的行动是**永非最优反应**(never-best response)如果它对参与人 i 的任何信念都不是最优反应。

显然任何永非最优反应都不是可理性化的。如果参与人 i 的一个行动 a_i 是一永非最优反应则对参与人 i 的每一信念存在某一行动(它可能依赖于信念)对参与人 i 来说比 a_i 更好。现在我们要证明在一有限博弈中如果 a_i 是一永非最优反应则无论参与人 i 持何信念都存在一个混合战略对参与人 i 来说比 a_i 更好。这一性质被精确地定义为如下所述。

▶定义 59.2 在战略博弈 $\langle N,(A_i),(u_i)\rangle$ 中参与人 i 的行动 $a_i \in A_i$ 是**强劣的**(strictly dominated)如果存在一个参与人 i 的混合战略 α_i 满足对所有 $a_{-i} \in A_{-i}$ 有 $U_i(a_{-i},\alpha_i) > u_i(a_{-i},a_i)$，这里 $U_i(a_{-i},\alpha_i)$ 是当参与人 i 使用混合战略 α_i 和其他参与人的行动向量是 a_{-i} 时参与人 i 的支付。

实际上我们表明了在每个参与人的行动集合都是有限的博弈中一个行动是永非最优反应 当且仅当它是强劣的。这样在这些博弈中强劣概念有着决策理论的基础，它并不涉及混合战略。可以说一个人即使会拒绝混合战略成为选择目标的思想，他仍能会认为一个参与人不会采用强劣行动。

■引理 60.1 在有限战略博弈中参与人的行动是永非最优反应当且 60
仅当它是强劣的。

证明：令战略博弈为 $G=\langle N,(A_i),(u_i)\rangle$ 令 $a_i^* \in A_i$。考虑辅助的严格竞争博弈 G'(参看定义 21.1)，在 G' 中参与人 1 的行动集合是 $A_i \setminus \{a_i^*\}$，参与人 2 的行动集合是 A_{-i}，参与人 1 的偏好在给定 $v_1(a_i,a_{-i}) = u_i(a_{-i},a_i) - u_i(a_{-i},a_i^*)$ 条件下由支付函数 v_1 所代表。(注意 v_1 的自

变量(a_i, a_{-i})是 G'中的行动二元组而自变量(a_{-i}, a_i)和(a_{-i}, a_i^*)是 G 中的行动组合。)对任何给定的 G'中混合战略组合(m_1, m_2)我们用 $v_1(m_1, m_2)$表示参与人 1 的期望支付。

行动 a_i^* 是G 中的永非最优反应当且仅当对G'中参与人 2 的任一混合战略都有参与人 1 的一个行动产生正支付;那即是,当且仅当 $\min_{m_2}\max_{a_i} v_1(a_i, m_2) > 0$。这同样是当且仅当 $\min_{m_2}\max_{m_1} v_1(m_1, m_2) > 0$(由 m_1 中 v_1 的线性)。

现在,由命题 33.1 博弈 G'有一混合战略纳什均衡,这样由命题 22.2 的(b)部分应用于 G'的混合扩展我们有 $\min_{m_2}\max_{m_1} v_1(m_1, m_2) > 0$ 当且仅当 $\max_{m_1}\min_{m_2} v_1(m_1, m_2) > 0$;那也是,当且仅当存在 G'中参与人 i 的一混合战略 m_1^* 满足对所有 m_2 有 $v_1(m_1^*, m_2) > 0$(即对所有 A_{-i}上的信念)。因为 m_1^* 是 $A_i \setminus \{a_i^*\}$上的一概率测度,所以它是 G 中参与人 1 的一个混合战略;条件 $v_1(m_1^*, m_2) > 0$(对所有 m_2)等价于 $U_i(a_{-i}, m_1^*) - U_i(a_{-i}, a_i^*) > 0$(对所有 $a_{-i} \in A_{-i}$),这等价于 a_i^* 是强劣的。 □

注意在此证明中的论证依赖于我们的假设:参与人对于不确定事件的偏好满足 von Neuman 和 Morgenstern 假设,如果偏好不满足这些假设,则为一永非最优反应与强劣的性质一般是不等价的。

4.2.2 反复剔除强劣行动

我们现在正式定义本部分开始时我们所描述的过程。

▶定义 60.2 有限战略博弈$\langle N, (A_i), (u_i)\rangle$的结果集合 $X \subseteq A$ **是反复剔除强劣行动剩下的**(survives iterated elimination of strictly dominated actions)。如果 $X = \times_{j\in N} X_j$ 且有一集合族$((X_j^t)_{j\in N})_{t=0}^{T}$对每个 $j \in N$ 满足下列条件:

61 • $X_j^0 = A_j$ 且 $X_j^T = X_j$

• $X_j^{t+1} \subseteq X_j^t$,对每一个 $t = 0, \cdots, T-1$

• 对 $t = 0, \cdots, T-1$ 参与人 j 在 $X_j^t \setminus X_j^{t+1}$ 中的每个行动在博弈$\langle N, (X_i^t), (u_i^t)\rangle$中都是强劣的,这里 u_i^t 对每一个 $i \in N$ 都局限于$\times_{j\in N} X_j^t$ 的函数 u_i。

• X_j^T 中没有行动在博弈 $\langle N,(X_i^T),(u_i^T)\rangle$ 中是强劣的。

◇例 61.1　图 61.1 的博弈中，行动 B 对于 T 和 M 以同等概率 $\frac{1}{2}$ 被采用的混合战略是劣的。在 B 从博弈中剔除后，对于 R，L 是劣的；在 L 被剔除后，对于 M，T 又是劣的。这样 (M,R) 是反复剔除强劣行动后剩下的惟一结果。

	L	R
T	3,0	0,1
M	0,0	3,1
B	1,1	1,0

图 61.1　两人战略博弈。参与人 1 惟一的可理性化行动是 M，参与人 2 惟一的可理性化行动是 R

我们现在要说明反复剔除劣行动剩下的结果集合是存在的，且是可理性化的行动组合集合

■命题 61.2　若 $X=\times_{j\in N}X_j$ 是在一个有限战略博弈 $\langle N,(A_i),(u_i)\rangle$ 中反复剔除劣行动剩下的，则对每一个 $j\in N$，X_j 是参与人 j 的可理性化的行动集合。

证明：假设 $a_i\in A_i$ 是可理性化的，且令 $(Z_j)_{j\in N}$ 为在定义 55.1 中支撑 a_i 的集合组合。对任何 t 值我们有 $Z_j\subseteq X_j^t$，因为 Z_j 中的每个行动都是对关于 Z_{-j} 的某一信念的最优反应，因此在博弈 $\langle N,(X_i^t),(u_i^t)\rangle$（由引理 60.1）中不是强劣的。所以 $a_i\in X_i$。

现在我们要证明对每一个 $j\in N$，X_j 的每一元素都是可理性化的。由可理性化定义：X_j 中的任何行动在这样的博弈中都不是强劣的，即每个参与人 i 的行动集合是 X_i。这样由引理 60.1，X_j 中的每一个行动对某个 X_{-j} 上的信念在 X_j 的元素中都是一个最优反应。我们需要证明的是 X_j 中的每个行动在集合 A_j 的所有元素中对 X_{-j} 上的某个信念的最优反应。如果 $a_j\in X_j$ 在 A_j 的所有元素中不是一个最优反应那么有一个 t 值使得 a_j 对 X_{-j} 上的信念 μ_j 在 X_j^t 的所有元素中是一个最优反应，但在 X_j^{t-1} 的元素中不 62
是一个最优反应。这样就有一个行动 $b_j\in X_j^{t-1}\setminus X_j^t$，它对 μ_j 在 X_j^{t-1} 的元素

中是一个最优反应,这就与 b_j 在该过程的第 t 阶段被剔除的事实相矛盾。□

请注意定义 60.2 中的过程并不要求所有强劣战略在任一阶段被全部剔除。因此该结论表明剔除的顺序和速度不影响剩下的结果集合。

如果我们将可理性化的定义修改为要求参与人相信他们对手的行动是独立的,则引理 60.1 和反复剔除强劣行动的概念与可理性化的等价都将不成立。为说明之,考虑图 58.1 中的博弈。行动 M_2 对参与人 3 的关于参与人 1 和 2 以同等概率采取(U,L)和(D,R)的信念是一个最优反应因而不是强劣的。不过,如我们以前所知,在修改的定义下(其中每个参与人的信念局限为独立信念的一个乘积。)它对(独立的)混合战略的任一二元组不是一最优反应因而不是可理性化的。

4.3 反复剔除弱劣行动

我们说一个参与人的行动是弱劣的,如果无论别的参与人做什么这个参与人有另一行动至少与该行动一样好,且对至少某个别的参与人的行动向量来说(比该行动)更好。

▶定义 62.1　在战略博弈$\langle N,(A_i),(u_i)\rangle$中参与人 i 的行动$a_i\in A_i$是**弱劣的**(weakly dominated),如果有一个参与人 i 的混合战略α_i 满足对所有 $a_{-i}\in A_{-i}$有 $U_i(a_{-i},\alpha_i)\geqslant u_i(a_{-i},a_i)$和对某一 $a_{-i}\in A_{-i}$有 $U_i(a_{-i},\alpha_i)>u_i(a_{-i},a_i)$,这里 $U_i(a_{-i},\alpha_i)$是参与人 i 使用混合战略 α_i 和别的参与人行动向量是a_{-i}时参与人 i 的支付。

由引理 60.1,一个弱劣而非强劣的行动是对某一信念的最优反应。这一事实使得反对使用一个弱劣行动的观点弱于反对使用一个强劣行动的观点。然而因为使用弱劣行动并无好处,所以在简化复杂博弈的过程中剔除
63 这些行动就是很自然的了。

弱劣概念导致了一个类似于反复剔除强劣行动的过程(定义 60.2)。不过,这个过程并不怎么吸引人,因为反复剔除弱劣行动剩下的行动集合可能依赖于行动被剔除的顺序,如图 63.1 中的两人博弈所示。我们先剔除 T (对于 M 弱劣)接着剔除 L(对 R 弱劣)的序列导致了参与人 2 选择 R 和支付组合是(2,1)的结果。另一方面,我们先剔除 B(对 M 弱劣),接着剔

除 R(对 L 弱劣)的序列导致了参与人 2 选择 L 和支付组合是(1,1)的结果。我们将在第 6.6 节进一步讨论反复剔除弱劣行动的过程。

	L	R
T	1,1	0,0
M	1,1	2,1
B	0,0	2,1

图 63.1　两人博弈,其中反复剔除弱劣行动剩下的行动集合依赖于行动被剔除的顺序

?练习 63.1　考虑练习 18.6 中博弈的一个变形,在其中有两个参与人,市民喜欢的位置的分布是同一的,每个参与人局限于选择一个形如 l/m 的位置,这里 m 是偶数,$l\in\{0,\cdots,m\}$。试证明反复剔除弱劣行动剩下的惟一结果是这样的,即两个参与人都选择位置 1/2。

?练习 63.2　占优可解性(Dominance Solvability)一个战略博弈是占优可解的,如果所有参与人对由在每个阶段每个参与人的所有弱劣行动都被剔除的反复剔除过程剩下的所有结果都是无偏好的,试给出一个占优可解的战略博弈例子,但该博弈不是所有参与人对反复剔除弱劣行动剩下的所有结果都无偏好的情形(并不是所有弱劣行动在每一阶段都可能被剔除的过程)。

?练习 64.1　两个参与人中的每人都报一个最多为 100 的非负整数。若 $a_1+a_2\leqslant 100$,这里 a_i 是由参与人 i 所报的数,那么每个参与人 i 收到支付 a_i。若 $a_1+a_2>100$ 且 $a_i<a_j$,那么参与人 i 收到 a_i 且参与人 j 收到 $100-a_i$;若 $a_1+a_2>100$ 且 $a_i=a_j$,则每个参与人收到 50。试说明该博弈是占优可解的(参看前面的练习)并找到剩下的结果集合。 64

引理 60.1 表明在一个有限博弈中一个非强劣行动对某个信念是最优反应。下面的练习对一个非 弱劣行动(或混合战略)强化了这一结论。

?练习 64.2　试证明在有限战略博弈中一个参与人的任一非弱劣混合战略对赋正概率给别的参与人的每一个行动向量的信念都是最优反应。〔提

示:令$\langle N,(A_i),(u_i)\rangle$为该博弈且令 U 为形如$(u_1(a^1_{-1}, m_1),\cdots,u_1(a^k_{-1}, m_1))$的所有向量集合,这里 m_1 属于参与人 1 的混合战略;$\{a^1_{-1},\cdots,a^k_{-1}\}$是除参与人 1 之外其他参与人的所有行动向量集合。令 $u^*\in U$ 对应于参与人 1 的一个非弱劣混合战略。你需要证明存在一个正向量 p^* 满足对所有 $u\in U$ 有 $p^*\cdot u^*\geqslant p^*\cdot u$。这样做为了不失一般性,令 $u^*=0$,并且对任一 $\epsilon>0$ 令 $P(\epsilon)=\{p\in\mathbb{R}^k: p_i\geqslant\epsilon$,对所有 i 和$\sum_{i=1}^k p_i=1\}$。对集合 $P(\epsilon)$和 U 使用练习 36.3 的结论且令$\epsilon\to 0$;同时要应用 U 为有限个向量的凸包这一事实。]

[注解]

可理性化概念源于 Bernheim(1984)和 Pearce(1984)(他们两个都限制参与人去相信他们对手的行动是独立的)。(Spohn(1982)讨论了这一思想,但对其没有系统表达。对反复剔除劣战略的过程看法首先由 Gale(1953)和 Luce and Raiffa(1957;pp. 108—109, 173)详细研究;我们所给的系统表述应归于 Moulin(1979)。引理 60.1 归于 Pearce(1984);它与 van Damme(1983)的引理 3.2.1 紧密相关。命题 61.2 应归于 Pearce(1984;p.1035)。

练习 56.4 中的结论应归于 Gabay and Moulin (1980), Bernheim(1984)和 Moulin(1984)。练习 56.5 摘于 Moulin(1986, p, 72)。练习 57.1 摘于 Rubinstein and Wolinsky(1994)。练习 63.2 中占优可解性概念应归于 Moulin(1979),它与 Luce and Raiffa(1957, p.109)的"在完全弱意义下的可
65 解性"的概念紧密相关。练习 64.1 归功于Brams and Taylor(1994),练习 64.2 应归于 Arrow, Barankin and Blackwell(1953)。

对于一类可理性化给出明确答案的博弈可参看 Vives(1990)和 Milgrom and Roberts(1990)。

第 5 章

知识与均衡

在本章，我们描述一个知识模型并用它来系统表达一个事件是“共同知识”的思想，用它去探究人们“彼此同意保留不同意见”是否可能，并用它去正式说明关于参与人隐含于纳什均衡概念和可理性化之后的知识的假设。

5.1 一个知识模型

战略博弈是对参与人之间的相互作用建模。因此我们不反对参与人对有关外来参数的知识感兴趣，对有关别的参与人的知识感兴趣。我们首先对单一决策主体的知识模型作一简单介绍。

该模型的基础是一个状态(state)集合 Ω。状态的概念在经典文献中有两种解释。在一个极端，状态被视为决策主体在确定的决策问题背景下感觉为相关的不确定性的描述。这种解释在标准的不确定性经济模型中被采用。在另一极端，状态被视为对世界的完整描述；不仅包括决策主体的信息和信念，还包括他的行动。

5.1.1 信息函数

确定决策主体对有关状态的知识范围的一个方法是确定一个信息函数 P，它将每一状态 $\omega\in\Omega$ 与 Ω 的一个非空子集 $P(\omega)$ 相联系。其解释是当状态是 ω 时决策主体仅知道状态在集合 $P(\omega)$ 中。即，他认为实际状态可 68
能是 $P(\omega)$ 中的任一状态而非 $P(\omega)$ 之外的任一状态。

▶定义 68.1 对于状态集合 Ω 的一个**信息函数**(information function)

是这样一个函数 P，它将每一状态 $\omega\in\Omega$ 与 Ω 的一个非空子集 $P(\omega)$联系起来。

当我们使用一个信息函数去对一个决策主体的知识建模时我们经常设定包含状态集合和信息函数的二元组 $\langle\Omega,P\rangle$ 满足下列两个条件：

P1　对每一 $\omega\in\Omega$，$\omega\in P(\omega)$。

P2　若 $\omega'\in P(\omega)$，则 $P(\omega')=P(\omega)$。

P1 是说决策主体从来不会将真实状态从他认为可能的状态集合中排除出去：他永远不能确认该状态是与真实状态不同的。P2 是说决策主体利用状态与他的信息的一致性或不一致性来作关于状态的推断。假设，与 P2 相反，$\omega'\in P(\omega)$且有一个状态 $\omega''\in P(\omega')$，$\omega''\notin P(\omega)$，那么如果状态是 ω，决策主体会认为既然 ω''与他的信息不一致，则真实状态不会是 ω'。同理，如果有一状态 $\omega''\in P(\omega)$，$\omega''\notin P(\omega')$，那么当状态是 ω 时决策主体就会认为既然 ω''与他的信息一致则真实状态不会是 ω'。

下列条件等价于 P1 和 P2。

▶定义68.2　对于状态集合 Ω，信息函数 P 是**分割的**(partitional)如果有 Ω 的一个分割使得对任一 $\omega\in\Omega$，集合 $P(\omega)$是包含 ω 的分割的元素。

■引理 68.3　一个信息函数是分割的当且仅当它满足 P1 和 P2。

证明：若 P 是分割的则它显然满足 P1 和 P2。现在假设 P 满足 P1 和 P2。若 $P(\omega)$和 $P(\omega')$相交且 $\omega''\in P(\omega)\cap P(\omega')$，则由 P2 我们有 $P(\omega)=P(\omega')=P(\omega'')$；由 P1 我们有 $\bigcup_{\omega\in\Omega}P(\omega)=\Omega$。这样 P 是分割的。　□

给定该结论，满足 P1 和 P2 的一个信息函数可被它导出的信息分割所确定。

◇例 68.4　令 $\Omega=[0,1)$且假设决策主体仅观察到一个数的十进制展
69 开的前 4 个数字。那么对每一个 $\omega\in\Omega$，集合 $P(\omega)$是所有这种状态 $\omega'\in\Omega$ 的集合，即 ω'的前 4 个数字同 ω 的前 4 个数字是一样的。这个信息函数是分割的。

☒练习 69.1　令 Q 为一个问题集合，问题的答案要么是“Yes”，要么

是“No”。一个状态是对 Q 中所有问题答案的一览表。假设信息函数 P 有下列性质：对某一状态 ω_1 集合 $P(\omega_1)$ 包括所有对前两个问题的答案与 ω_1 一样的状态，对另一状态 ω_2 集合 $P(\omega_2)$ 包括所有对前 3 个问题的答案与 ω_2 一样的状态。P 必须是分割的吗？

?练习 69.2　一个决策主体被告知一个整数 Ω 但仅记得该数是 $n-1, n, n+1$ 中的某个。用一信息函数对决策主体的知识建模并确定该函数是否是分割的。

5.1.2　知识函数

我们称一个状态集合（Ω 的一子集）为一**事件**（event）。给定我们关于信息函数的解释，一个 $P(\omega)\subseteq E$ 的决策主体在状态 ω 中，知道事件 E 中的某一状态已经发生了。在此情形下我们说在状态 ω 中决策主体**知道** E。给定 P 我们现在定义决策主体的**知识函数**（knowledge function）K 为：

$$K(E)=\{\omega\in\Omega: P(\omega)\subseteq E\} \tag{69.3}$$

对任一事件 E 集合 $K(E)$ 是所有状态（在其中决策主体知道 E）的集合。从任一信息函数派生出的知识函数 K 都满足下面三个性质。

K1　$K(\Omega)=\Omega$

这是说在所有状态中决策主体都知道 Ω 中的某一状态已经发生。

K2　若 $E\subseteq F$ 则 $K(E)\subseteq K(F)$

这是说若无论 E 什么时候发生 F 都发生且决策主体知道 E，则他知道 F：若 E 包含 F 则 E 的知识包含 F 的知识。

K3　$K(E)\cap K(F)=K(E\cap F)$

该性质的解释是若决策主体既知道 E 又知道 F，则他知道 $E\cap F$。 70

若 P 满足 P1 则相关的知识函数 K 满足下列附加性质。

K4　（知识公理）　$K(E)\subseteq E$

这是说无论何时决策主体知道 E，那么确实有 E 中的某一元素是真实状态：决策主体不知道任何错误的事情。该公理是这样从 P1 中导出的：若 $\omega\in K(E)$，则 $P(\omega)\subseteq E$，所以由 P1 我们有 $\omega\in E$。

若 P 是分割的（即：满足 P1 和 P2）那么 $K(E)$ 是作为 E 的子集的分割的所有元素的并集。（若 E 不含分割的任一元素，则 $K(E)$ 是空的。）在此情形下知识函数 K 满足下列两个附加性质。

K5　（透明度公理）　$K(E)\subseteq K(K(E))$

给定将 $K(E)$ 解释为决策主体知道 E 的事件，我们将 $K(K(E))$ 解释为决策主体知道他知道 E 的事件。正如我们上面所提到的，若 P 满足 P1 和 P2，那么集合 $K(E)$ 是由 P 导致的分割的元素的并集；由此评述从 K5 便可导出若 F 是分割的元素的一个并集，则 $K(F)=F$。

K6 （智慧公理） $\Omega \setminus K(E) \subseteq K(\Omega \setminus K(E))$

该公理的解释是决策主体知道什么是他所不知的。若他不知道 E 则他知道他不知道 E。因为 P 是分割的，所以 $K(E)$ 是由 P 导致的分割的元素的一个并集，这样 $\Omega \setminus K(E)$ 也是这样一个并集，K6 由此成立。

注意给定 K 满足 K4，则 K5 和 K6 中的性质在实际上同等拥有。

我们已经将一个信息函数作为原生的并且从它派生了一个知识函数。反过来我们可先对集合 Ω 定义一个**知识函数**为函数 K，它将 Ω 的一个子集与每一事件 $E \subseteq \Omega$ 联系起来。然后我们可从它派生出一个信息函数 P 如下：对每一状态 ω 令

$$P(\omega)=\bigcap\{E \subseteq \Omega : K(E) \ni \omega\} \qquad (70.1)$$

（如果没有事件 E 满足 $\omega \in K(E)$，则我们将交集视作 Ω。）

71 ?练习 71.1

a. 给定一信息函数 P，令 K 为由(69.3)所定义的知识函数，且令 P' 为由(70.1)中的派生出的信息函数。试证明 $P'=P$。

b. 给定满足 K1，K2 和 K3 的知识函数 K，令 P 为由(70.1)定义的信息函数，且令 K' 为由(69.3)中的 P 派生出的知识函数。试证明 $K'=K$。

?练习 71.2 使用我们已描述的架构，我们可将一个个人决策问题系统表达如下：令 A 为一行动集合，Ω 为一状态集合，P 为一分割信息函数，π 为一个 Ω 上的概率测度，$u: A \times \Omega \to \mathbb{R}$ 为一期望值代表了个人关于 A 上不确定事件的偏好的函数。个人的问题是去选择一个函数 $a: \Omega \to A$（称为一行动(act)，其满足：只要 $\omega \in P(\omega)$ 和 $\omega' \in P(\omega)$，则 $a(\omega)=a(\omega')$，去解决问题 $\max_a E_\pi u(a(\omega), \omega)$（这里 E 为期望算符）。若对所有 $\omega \in \Omega$，$P(\omega) \subseteq P'(\omega)$（即若由 P' 导致的分割的每一个元素都是由 P 导致的分割的元素的一个并集），则定义分割信息函数 P' 是比信息函数 P 更粗糙的(coarser)。试说明若 P' 比 P 更粗糙，则在信息函数 P' 下的最优行动不优于在信息函数 P 下的最优行动。试将此结论与练习 28.2 中的结论作比较。

5.1.3 一个阐述性的例子：帽子之谜(The Puzzle of the Hats)

下面这个在20世纪前半叶的一段时间曾“风靡欧洲”的谜(Littlewood(1953,p.3),阐述了我们已定义的概念。几个“完全理性”的人围绕一张桌子而坐,他们每人戴一顶颜色或白或黑的帽子。每个人都能看到别的 $n-1$ 个人的帽子,但看不到自己的帽子。一个旁观者宣布:“你们中的每位都戴着顶颜色或白或黑的帽子,这些帽子中至少有一顶是白的,我将开始慢慢数数。每次数数后你们都有机会举一只手。不过你只能在你知道你帽子颜色的情况下才能这样做。”第一次在什么时候有人会举手?

下面我们用我们已介绍的正式模型来回答该问题。开始,在旁观者宣布后,状态集合是帽子颜色的所有结构 $c=(c_1,\cdots,c_n)$ 的集合,这里每一个 72
c_i 要么是 W 要么是 B 且至少有一个 c_i 是 W。这 2^n-1 个状态组成集合。

$$\Omega=\{c\in\{B,W\}^n:|\{i:c_i=W\}|\geqslant 1\}$$

任何个人 i 的初始信息函数 P_i^1 给定如下:在任一状态 c 中集合 $P_i^1(c)$ 包括所有与 i 的观察相一致的状态,因而包含至多两个状态,它们仅随 i 的帽子颜色不同而不同。精确地说,如果 c 是这样一种状态,在其中一个不是 i 的个人戴了顶白帽,那么 $P_i^1(c)=\{(c_{-i},W),(c_{-i},B)\}$,如果 c 是这样一种状态,在其中别的所有个人都戴黑帽,那么 $P_i^1(c)=\{c\}$(因为不是所有人都戴黑帽)。

对于拥有信息函数 P_i 的个人 i“知道他的帽子颜色”意味着什么呢?这意味着要么他知道事件 $\{c:c_i=W\}$ 发生,要么他知道事件 $\{c:c_i=B\}$ 发生。因此事件“i 知道他的帽子颜色”是

$$E_i=\{c:P_i(c)\subseteq\{c:c_i=B\}\text{或 }P_i(c)\subseteq\{c:c_i=W\}\}$$

仅在状态 c(在其中恰有一个人 i 对于他有 $c_i=W$)中对某一 j 有 $P_j^1(c)\subseteq E_j$,且在此情形下 $P_i^1(c)\subseteq E_i$,所以 i 举手。

现在令 $F^1=\{c:|\{i:c_i=W\}|=1\}$,它是在第一阶段有某个人举手的状态集合。若在第一阶段没有人举手,那么所有个人获得状态不在 F^1 中这一附加信息,因而对所有 i 和对所有 $c\notin F^1$ 我们有 $P_i^2(c)=P_i^1(c)\setminus F^1$。也就是,在任一这种状态中每个人都推知至少有两个人戴白帽。我们有 $P_i^2(c)=P_i^1(c)=\{(c_{-i},W),(c_{-i},B)\}$,除非对恰有一个人 $j\neq i$, $c_j=W$,在此情形下 $P_i^2(c_{-i},W)=\{(c_{-i},W)\}$(且 $P_j^2(c_{-j},W)=\{(c_{-j},W)\}$.换言

之，在任一状态中为精确起见，设有两个人 j 和 h，满足 $c_j = W$ 和 $c_h = W$，则我们有 $P_j^2(c)\subseteq E_j$ 和 $P_h^2(c)\subseteq E_h$，因此 h 和 j 在第二阶段都举手。现在令 $F^2=\{c:|\{i:c_i=W\}|=2\}$，在该状态集合中过程在第二阶段结束。在旁观者 2 后仍无人举手（$c\notin F^1\cup F^2$），则在此状态中所有个人都推断至少有三顶帽子是白的，过程以 $P_i^3(c)=P_i^2(c)\setminus F^2$ 继续。很容易可以看出若 k 顶帽子都是白的，则直到旁观者数 k 才有人举手，在这时戴白帽的 k 个人举手。

73 5.2 共同知识

如果在某状态中每个人都知道某一事件则我们说在这个状态中该事件是“共有的知识”。我们说一个事件是“共同知识”，如果它不仅是共有知识，而且每个人都知道所有别的参与人都知道它，每个人都知道所有其他人都知道所有人都知道它，如此等等。为了简单，限于两人情形的共同知识概念系统表达成如下定义。

▶定义 73.1　令 K_1 和 K_2 为个人 1 和 2 对状态集合 Ω 的知识函数。一个事件 $E\subseteq\Omega$ 是**状态 $\omega\in\Omega$ 中 1 和 2 间的共同知识**（common knowledge between 1 and 2 in the state），如果 ω 是无穷序列 $K_1(E)$, $K_2(E)$, $K_1(K_2(E))$, $K_2(K_1(E))$, ……中每个集合的一个元素。

共同知识的另一个定义（在命题 74.2 中我们要说明它是等价的）是用个人信息函数来叙述的。

▶定义 73.2　令 P_1 和 P_2 为个人 1 和 2 对状态集合 Ω 的信息函数。一个事件 $F\subseteq\Omega$ **是在 1 和 2 间自明的**（self-evident between 1 and 2），如果对所有 $\omega\in F$，我们有 $P_i(\omega)\subseteq F$（对 $i=1,2$）。一个事件 $E\subseteq\Omega$ **是在状态 $\omega\in\Omega$ 中 1 和 2 间的共同知识**（common knowledge between 1 and 2 in the state $\omega\in\Omega$）如果有一个自明的事件 F 满足 $\omega\in F\subseteq E$。

简言之，一个事件在两个人间是自明的，如果不管它什么时候发生，两个人都知道它发生了（即不管它什么时候发生它都是两人间的共有知识），并且是状态 ω 中的共同知识，如果有一包含 ω 的自明事件，它的发生蕴含了 E。

◇例 73.3　令 $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}$，令 P_1 和 P_2 为个人 1 和 2 的分割信息函数，并令 K_1 和 K_1 为相关的知识函数。令由信息函数导致的分割为

$$\mathcal{P}_1=\{\{\omega,\omega_2\},\{\omega_3,\omega_4,\omega_5\},\{\omega_6\}\}$$
$$\mathcal{P}_2=\{\{\omega_1\},\{\omega_2,\omega_3,\omega_4\},\{\omega_5\},\{\omega_6\}\}$$

事件 $E=\{\omega_1,\omega_2,\omega_3,\omega_4\}$ 不包含任何 1 和 2 间的自明事件，因此没有任何状态使 E 在第二种定义(73.2)的意义下是 1 和 2 间的共同知识。事件 E 在第一种定义(73.1)的意义下，在任何状态都不是共同知识，因为 $K_1(K_2$ 74
$(K_1(E_1)))=\phi$，如下列计算所展示的那样：

$$K_1(E)=\{\omega,\omega_2\},\quad K_2(E)=E,$$
$$K_2(K_1(E))=\{\omega_1\},\quad K_1(K_2(E))=\{\omega_1,\omega_2\},$$
$$K_1(K_2(K_1(E)))=\phi,\quad K_2(K_1(K_2(E)))=\{\omega_1\}.$$

事件 $F=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5\}$ 在 1 和 2 间是自明的，因此在第二种定义的意义下，在 F 的任一状态中都是 1 和 2 间的共同知识。因为 $K_1(F)=K_2(F)=F$，事件 F 在第一种定义的意义下在 F 的任一状态中同样是 1 和 2 间的共同知识。

在证明共同知识的两个定义是等价的之前，我们先建立下列结论。

■引理 74.1　令 P_1 和 P_2 为个人 1 和 2 对状态集合 Ω 的分割信息函数，令 K_1 和 K_2 为相关的知识函数，且令 E 为一事件。那么下列的三个条件是等价的。

a. $K_i(E)=E$，对 $i=1,2$。

b. E 在 1 和 2 间自明的。

c. E 是由 $P_i(i=1,2)$ 导致的分割的元素的一个并集。

证明：假设(a)成立，则对每个 $\omega\in E$ 我们有 $P_i(\omega)\subseteq E$(对 $i=1,2$)因此(b)满足；假设(b)成立，则对 $i=1,2$，$E=\bigcup_{\omega\in E}P_i(\omega)$，因此 E 是两个分割的元素的一个并集，所以(c)满足。最后由(c)立即可知(a)成立。

现在我们来证明定义 73.1 和 73.2 是等价的。

■命题 74.2　令 Ω 为一有限状态集合，令 P_1 和 P_2 为个人 1 和 2 的分割信息函数，且令 K_1 和 K_2 是相关的知识函数。那么根据定义 73.1 一个事件 $E\subseteq\Omega$ 在状态 $\omega\in\Omega$ 中是 1 和 2 间的共同知识，当且仅当它根据定义 73.2 在状态 ω 中是 1 和 2 间的共同知识。

证明：假定事件 E 根据定义 73.1 是状态 ω 中 1 和 2 间的共同知识。对每一个 $i\in\{1,2\}$ 和 $j\neq i$ 我们有 $E\supseteq K_i(E)\supseteq K_j(K_i(E))\supseteq K_i(K_j(K_i(E)))\supseteq\cdots$ 且 ω 是所有这些集合的一个元素，所以它们不是空的。这样因为 Ω 是有限的，所以有一集合 $F_i=K_i(K_j(K_i\cdots K_i(E)\cdots))$ 满足 $K_j(F_i)=F_i$；因为 P_i 是分割的，K_i 满足 K4 和 K5，所以我们也有 $K_i(F_i)=F_i$。这样
75 由引理 74.1 事件 F_i 在 1 和 2 间是自明的，所以 E 根据定义 73.2 是 ω 中的共同知识。

现在假定根据定义 73.2，$E\subseteq\Omega$ 在状态 ω 中是 1 和 2 间的共同知识。那么存在一自明事件 F 有 $\omega\in F\subseteq E$。由引理 74.1，每个形为 $K_i(K_j(K_i\cdots K_i(F)\cdots))$ 的集合都与 F 重叠。从 K2 可知 ω 是所有形如 $K_i(K_j(K_i\cdots K_i(E)\cdots))$ 集合的一个元素，因此 E 根据定义 73.1 是 ω 中的共同知识。 □

5.3 人们能彼此同意保留不同意见吗？

下面是一个能用我们已经描述的框架来叙述的有趣问题。在两个具有相同先验概率的个人间，个人 1 赋概率 η_1 给某一事件，而个人 2 赋概率 $\eta_2\neq\eta_1$ 给同一事件，能成为共同知识吗？答案似乎是肯定的：当他们拥有不同信息时，在这种方式下个人可能“彼此同意保留不同意见”。不过，现在我们要证明的是如果个人的信息函数是分割的则答案是否定的。

贝叶斯博弈（第 2.6 节）是一个使该结论很有趣的背景。在经典文献中常作的假设是在这个博弈中所有参与人有同一先验概率。结论所包含的是在这一假设下参与人赋不同后验概率给同一事件在他们间不可能是共同知识。因此如果我们想对概率方面的差异是共同知识这样的情形建模，我们必须假定参与人的先验概率是不同的。

令 ρ 为状态集合 Ω 上的一概率测度，其被解释为个人的共同先验概率，且令 P_1 和 P_2 为个人的信息函数。如果 E 是一事件且 $\rho(E\mid P_i(\omega))=\eta_i$（这里 $\rho(E\mid P_i(\omega)$ 是 E 的以 $P_i(\omega)$ 为条件的概率），则给定状态 ω 中他的信息，个人 i 赋概率 η_i 给事件 E。因此事件“个人 i 赋概率 η_i 给 E”是 $\{\omega\in\Omega:\rho(E\mid P_i(\omega))=\eta_i\}$。

■命题 75.1　假定状态集合 Ω 是有限的且个人 1 和 2 有相同的先验

概率。如果每个人的信息函数是分割的且它们在个人1赋概率 η_1 给某一事件 E 和个人2赋概率 η_2 给 E 的某一状态 $\omega^* \in \Omega$ 中是1和2间的共同知识，那么 $\eta_1 = \eta_2$。 76

证明：若假设满足有一个自明的事件 $F \ni \omega^*$，它是 $\{\omega \in \Omega: \rho(E \mid P_1(\omega)) = \eta_1\}$ 和 $\{\omega \in \Omega: \rho(E \mid P_2(\omega)) = \eta_2\}$ 支集的一个子集，且因此为这两个集合的一个子集，这里 ρ 是共同先验概率。由引理74.1，对每个人 i 事件 F 是 i 的信息分割的元素的一个并集。因为 Ω 是有限的，所以每个并集中集合的个数也是有限的；令 $F = \cup_k A_k = \cup_k B_k$。对任何不相交集 C 和 D，满足 $\rho(E \mid C) = \eta_i$ 和 $\rho(E \mid D) = \eta_i$，我们有 $\rho(E \mid C \cup D) = \eta_i$。所以，既然对每一个 k 我们有 $\rho(E \mid A_k) = \eta_1$，从而 $\rho(E \mid F) = \eta_1$；同理 $\rho(E \mid F) = \eta_2$。故 $\eta_1 = \eta_2$。 □

?练习76.1 试证明如果具有分割信息函数的两个人有相同先验概率，那么他们赋不同概率于某一事件可能是他们间的共同知识。不过，试证明由个人1所赋的概率超过由个人2所赋的概率不可能是共同知识。

?练习76.2 试证明如果具有分割信息函数的两个人有相同先验概率，那么个人1相信某一不确定事件的期望超过某个数 η 而个人2相信这个期望少于 η 在他们俩间不可能是共同知识。试举例说明该结论依赖于个人信息函数是分割的这一假设。

5.4 知识和解的概念

在前面各章中我们讨论了纳什均衡和可理性化的概念。当寻求这些概念的动因时我们往往不正式地求助于关于参与人所知的假设。本部分我们将应用上面描述的模型去正式考察那些隐含于解的概念之后的关于参与人知识的假设。

从始至终我们将注意力集中于给定的战略博弈 $G = \langle N, (A_i), (\succsim_i) \rangle$ 上（见定义11.1）。

令 Ω 为一状态集合，其中每一个状态都是对与博弈相关的环境的一个描述：也就是对每个参与人的知识、行动和信念的一个描述。正式地，每一

个状态 $\omega\in\Omega$ 包括对每一个参与人 i 的下列详细规定：

• $P_i(\omega)\subseteq\Omega$，它描述了参与人 i 在状态 ω 中的知识(这里 P_i 是一分割信息函数)。

• $a_i(\omega)\in A_i$，由参与人 i 在状态 ω 中选择的行动。

77 • $\mu_i(\omega)$，$A_{-i}=\times_{j\in N\setminus\{i\}}A_j$ 上的一概率测度，是参与人 i 在状态 ω 中关于别的参与人行动的信念。(注意这允许参与人去相信别的参与人的行动是相关的。)

请注意状态的概念，因为它包含了对每个参与人的知识、行动和信念的一个详细规定，它可以是自反的：如果在状态 ω_1 中某个参与人不知道状态是 ω_1 还是 ω_2 那么关于 ω_1 的描述就指它自己。

在这一状态集合的定义中我们暗含的假设是博弈为 G 在所有参与人中是共同知识。因此我们忽略了下列这些所列举的可能：某一个参与人不知道他自己的行动集合或别的参与人的行动集合，或者有某一个参与人 i 不知道参与人 j 是否知道参与人 i 的偏好。这一假设强于我们对某些结论所需要的假设。为正式表示参与人关于博弈知识的弱假设我们需要拓展状态集合的定义，要求每一个状态包含一个关于所进行的博弈的详细规定。

我们现在分离出一个状态的某些性质，这些性质包含着在此状态中行动是与各种解的概念相一致的。我们的第一个结论是如果在某一状态中每个参与人是理性的，他们知道别的参与人的行动且有一个与他的知识相一致的信念，那么在此状态中所选择的行动组合是博弈的一个纳什均衡。

■命题 77.1 假设在状态 $\omega\in\Omega$ 中每个参与人 $i\in N$

a. 知道别的参与人的行动：$P_i(\omega)\subseteq\{\omega'\in\Omega: a_{-i}(\omega')=a_{-i}(\omega)\}$；

b. 有一个与他的知识相一致的信念：$\mu_i(\omega)$ 的支集是 $\{a_{-i}(\omega')\in A_{-i}:\omega'\in P_i(\omega)\}$ 的一个子集；

c. 是理性的：$a_i(\omega)$ 是参与人 i 对 $\mu_i(\omega)$ 的一个最优反应。

那么：$(a_i(\omega))_{i\in N}$ 是 G 的一纳什均衡。

证明：由(c)行动 $a_i(\omega)$ 是参与人 i 对自己信念的一个最优反应，由(b)它赋概率 1 给集合 $\{a_{-i}(\omega')\in A_{-i}:\omega'\in P_i(\omega)\}$；由(a)这集合是 $\{a_{-i}(\omega)\}$。 □

假设每个参与人都知道所有其他参与人的行动是很强的。我们现在要说明的是在两人博弈中我们可用下列假设代替它：若我们加强(c)去要求不

仅每个参与人是理性的且每个参与人知道别的参与人是理性的则每个参与人知道别的参与人的信念。既然该结论涉及混合战略我们现在令考虑下的战略博弈为 $G=\langle N,(A_i),(u_i)\rangle$，这里对每一 $i\in N$ 函数 u_i 的期望值代表了参与人 i 关于 A 上不确定事件的偏好。 78

■命题 78.1 假设 $|N|=2$ 且在状态 $\omega\in\Omega$ 中每个参与人 $i\in N$。

a. 知道别的参与人的信念：对 $j\neq i$，$P_i(\omega)\subseteq\{\omega'\in\Omega:\mu_j(\omega')=\mu_j(\omega)\}$；

b. 有一个与他的知识相一致的信念：$\mu_i(\omega)$ 的支集是 $\{a_j(\omega')\in A_j:\omega'\in P_i(\omega)\}$ 的一个子集，这里 $j\neq i$。

c. 知道别人是理性的：对任一 $\omega'\in P_i(\omega)$ 行动 $a_j(\omega')$ 是参与人 j 对 $\mu_j(\omega')(j\neq i)$ 的一个最优反应。

那么：混合战略组合 $(\alpha_1,\alpha_2)=(\mu_2(\omega),\mu_1(\omega))$ 是 G 的一个混合战略纳什均衡。

证明：令 a_i^* 为参与人的一个行动，它在 $\alpha_i=\mu_j(\omega)$ 的支集中。由(b)存在一个状态 $\omega'\in P_j(\omega)$ 有 $a_i(\omega')=a_i^*$。由(c)行动 a_i^* 是参与人 i 对 $\mu_i(\omega')$ 的最优反应，由(a)它等价于 $\mu_i(\omega)$。 □

注意两个命题都不需要参与人从 Ω 上的某一共同先验概率去派生出他们的信念。特别地，注意在(b)中我们仅要求每个参与人的信念与他的知识相一致。同时也应注意博弈是共同知识这一假设在两个结论中都可弱化：在命题 77.1 中假设每个参与人知道他自己的行动集合和偏好就够了，在命题 78.1 中假设博弈是共有知识就够了。

下列的例子表明当有两个以上的参与人时命题 78.1 无类似结果。考虑图 79.1 上部的博弈。(注意参与人 3 的支付常为 0。)令状态集合为 $\Omega=\{\alpha,\beta,\gamma,\delta,\epsilon,\xi\}$，且令参与人的行动函数和信息函数由该图下部分给出；假定参与人的信念来自于同一先验概率，它在表的第一行被给出。

考虑状态 δ。我们断言命题的三个条件都是满足的。条件(b)满足是因为每个参与人在 δ 中的信念决定于同一先验，它还可证明在此状态中每个参与人知道别的参与人的信念且知道别的参与人是理性的。考虑参与人 1。她知道状态要么是 δ 要么是 ϵ，所以她知道参与人 2 的信息是 $\{\gamma,\delta\}$ 或 $\{\epsilon,\xi\}$。在两种情形下，参与人 2 都相信以概率 $\frac{2}{3}$ 由参与人 1 和 3 所选择

的行动二元组为(D,A)和以概率$\frac{1}{3}$选择的是(D,B)。给定这一信念，参与
79 人 2 选择的行动 L 是最优的。所以参与人 1 知道参与人 2 是理性的。

	L	R
U	2,3,0	2,0,0
D	0,3,0	0,0,0

A

	L	R
U	0,0,0	0,2,0
D	3,0,0	3,2,0

B

状　态	α	β	γ	δ	ϵ	ξ
概率×63	32	16	8	4	2	1
1的行动	U	D	D	D	D	D
2的行动	L	L	L	L	L	L
3的行动	A	B	A	B	A	B
1的行动	$\{\alpha\}$	$\{\beta$	$\gamma\}$	$\{\delta$	$\epsilon\}$	$\{\xi\}$
2的行动	$\{\alpha$	$\beta\}$	$\{\gamma$	$\delta\}$	$\{\epsilon$	$\xi\}$
3的行动	$\{\alpha\}$	$\{\beta\}$	$\{\gamma\}$	$\{\delta\}$	$\{\epsilon\}$	$\{\xi\}$

图 79.1　上半部是一个三人博弈，在其中参与人 1 选择两行中的某行；参与人 2 选择两列中的某列；参与人 3 选择两表中的某表。下半部是该博弈中参与人的行动函数、信息函数和信念。

同理参与人 2 知道参与人 1 的信息是$\{\beta,\gamma\}$或$\{\delta,\epsilon\}$。在两种情形下参与人 1 相信以概率$\frac{2}{3}$参与人 2 和 3 选择(L,B)且以概率$\frac{1}{3}$他们选择(L,A)。给定这一信念，她选择的行动 D 是最优的，所以参与人 2 知道参与人 1 是理性的。

参与人 3 知道参与人 1 的信息是$\{\delta,\epsilon\}$和参与人 2 的信息是$\{\gamma,\delta\}$。因此，类似上述证明，参与人 3 知道参与人 1 和 2 是理性的。

在 γ,δ 和ϵ 这三种状态中，参与人 3 的信念是参与人 1 和 2 的行动二元组为(D,L)，因此在状态 δ 中参与人 1 和 2 知道参与人 3 的信念。因为她的支付经常为 0 所以 1 和 2 知道她是理性的。

不过在状态 δ 中这些信念并不能确定一纳什均衡。实际上，各个参与人关于别的参与人行动的信念都不是一致的：参与人 1 相信参与人 3 以概
80 率$\frac{1}{3}$选择 A 而参与人 2 相信参与人 3 以概率$\frac{2}{3}$选择 A。任何一个这种信念与行动 D 和 L 都不能形成博弈的一个混合战略纳什均衡。

使这个例子说明问题的原因在于状态 δ 中参与人1不知道参与人2知道她的信念：参与人1认为状态可能是ϵ，在其中参与人2不知道参与人1是相信参与人3采用 β 还是参与人3以概率$\frac{2}{3}$选 B 和以概率$\frac{1}{3}$选 A。

Aumann 和 Brandenburger(1995)证明了如果所有参与人分享一个共同先验概率，且在某一状态中理性是共有知识及参与人的信念是共同知识，那么即使有多于两个的参与人在那状态中的信念也能形成一个混合战略纳什均衡。关键点是如果参与人1和2的关于参与人3行动的信念是共同知识且如果所有参与人分享同一先验概率，那么信念一定是一样的(通过一个像命题75.1的证明中那样的论证)。

下面的结论正式化了第4章中的论证，使得可理性化的概念建立在一个弱于纳什均衡的关于参与人知识的假设上，它仅要求"所有参与人是理性的"，在所有参与人之间是共同知识。(这个结论并非要依赖有两个参与人的假设，尽管在这儿论述是比较简单的。)

■命题80.1 假定$|N|=2$，在状态 $\omega\in\Omega$ 中，每个参与人的信念与他的知识相一致是参与人之间的共同知识，并且每个参与人是理性的。也就是，假设有一自明事件 $F\ni\omega$ 使得对每一$\omega'\in F$ 和每一$i\in N$。有

a. $\mu_i(\omega')$的支集是$\{a_j(\omega'')\in A_j:\omega''\in P_i(\omega')\}$的一个子集，$j\neq i$；

b. 行动 $a_i(\omega')$是参与人 i 对$\mu_i(\omega')$的一个最优反应。

那么，对每一 $i\in N$，行动 $a_i(\omega)$在 G 中是可理性化的。

证明：对每一 $i\in N$。令 $Z_i=\{a_i(\omega')\in A_i:\omega'\in F\}$。由(b)我们知道对任一个 $\omega'\in F$ 行动$a_i(\omega')$是对 $\mu_i(\omega')$的一个最优反应，由(a)它的支集是$\{a_j(\omega'')\in A_j:\omega''\in P_i(\omega')\}$的一个子集。因为 F 是自明的，我们有 $P_i(\omega')\subseteq F$，因而$\{a_j(\omega'')\in A_j:\omega''\in P_i(\omega')\}\subseteq Z_j$。故(使用定义55.1)$a_i(\omega)$是可理性化的。 □

本部分三个结论都从关于他们在那状态中的知识的假设出发导出了参与人在某一特定状态中的行动或信念的内涵。下面练习中的结论基于一个 81
不同类型的假设——即在每一状态中参与人的理性都是共同知识。若该假设满足且所有参与人的信念派生于同一先验，则参与人的行动分布在 Ω 上是一个相关均衡。

?练习81.1 假设对所有 $\omega\in\Omega$ 所有参与人是理性的。(因此在每一个状态中他们的理性是共同知识，因为在每一个状态中任何在所有状态中

是真实的事实都是共同知识)试证明如果在每一个状态中每个参与人的信念派生于 Ω 上的一共同先验概率 ρ,其满足对所有 $i \in N$ 和所有 $\omega \in \Omega$, $\rho(P_i(\omega))>0$ 及对每一 $i \in N$ 和每一 $\omega' \in P_i(\omega)$, $a_i(\omega') = a_i(\omega)$,那么 $\langle(\Omega,\rho),(\mathcal{P}_i),(a_i)\rangle$(这里 $\mathcal{P}_i$ 是由 P_i 导致的分割)是 G 的一个相关均衡。(证明很简单,主要任务是理解结论的内容。)

	A	B
A	M,M	1,−L
B	−L,1	0,0

G_a (概率 $1-p$)

	A	B
A	0,0	1,−L
B	−L,1	M,M

G_b (概率 p)

图 81.1 电子邮件博弈的支博弈。参数满足 $L>M>1$ 和 $p<\frac{1}{2}$

5.5 电子邮件博弈

本部分我们研究一种阐述本章所导入的概念的博弈。两个参与人都不得不选取行动 A 或 B 中的某一个。以概率 $p<\frac{1}{2}$ 参与人所涉及的博弈是 G_b;以概率 $1-p$ 涉及的是 G_a。在 G_a 和 G_b 中对于参与人来说选择同一行动是互利的,但是最优的行动依赖于博弈:在 G_a 中结果 (A,A) 是最优的,而在 G_b 中结果 (B,B) 是最优的。支付由图 81.1 所示,那里 $L>M>1$。注意即使一个参与人确定博弈为 G_b,对他来说选择 B 也是有风险的,除非他充分相信他的伙伴也打算选 B。

开始只有参与人 1 知道哪一个是真实博弈。先假定参与人 2 不能获得
82 这一信息。那么我们能将此情形模化为一个贝叶斯博弈(定义 25.1),在其中有两个状态 a 和 b 及由信号函数导致的信息结构:对于参与人 1 为 $\{\{a\},\{b\}\}$,对参与人 2 为 $\{\{a,b\}\}$。这个博弈有惟一纳什均衡:两个参与人都经常选择 A;每个参与人的期望支付是 $(1-p)M$。

现在假定参与人 1 能与参与人 2 以某种方式交流从而使得博弈成为他们间的共同知识。在此情形下每个参与人的信息结构为 $\{\{a\},\{b\}\}$,(退化的)贝叶斯博弈有一纳什均衡;在其中每个参与人在状态 a 选择 A,在状态

b 选择 B；每个参与人的支付是 M。

在本部分我们所研究的情形中，参与人能交流，但是对他们公开的手段不允许博弈成为共同知识。特别是，参与人限于在下列协议下通过电脑来交流。若博弈是 G_b，那么参与人 1 的电脑会自动发送一条信息给参与人 2 的电脑，若博弈是 G_a，那么没有信息发送。若电脑收到信息则它会自动地发送一条确认信息；不仅对初始信息，对确认信息及确认信息的确认信息等等也会发送一条确认信息。协议是被设计用来发送确认信息的，因为技术有这样的性质：对任何给定的信息存在一个小概率 $\epsilon > 0$ 使之不能达到它的目的地。若信息没到达则交流结束。在交流阶段的末尾每个参与人的屏幕显示他的机器已发送信息的次数。

为了讨论这种情形下参与人的知识，我们需要确定一个状态集合和参与人的信息函数。确定状态集合为 $\Omega = \{(Q_1, Q_2): Q_1 = Q_2$ 或 $Q_1 = Q_2 + 1\}$。在状态 (q, q) 中，参与人 1 的电脑发送 q 个信息，它们都到达参与人 2 的电脑，且由参与人 2 的电脑发送的第 q 个信息丢失了。在状态 $(q+1, q)$ 中，参与人 1 的电脑发送 $q+1$ 个信息，除了最后一个其余的全都到达参与人 2 的电脑。参与人 1 的信息函数定义为：若 $q \geqslant 1$ 为 $P_1(q, q) = \{(q, q), (q, q-1)\}$，否则为 $P_1(0,0) = \{(0,0)\}$；参与人 2 的信息函数定义为：对所有 q，$P_2(q, q) = \{(q, q), (q+1, q)\}$。因 $G(Q_1, Q_2)$ 表示在状态 (Q_1, Q_2) 中进行的博弈；即 $G(0,0) = G_a$，否则 $G(Q_1, Q_2) = G_b$。参与人 1 知道在所有状态中的博弈。参与人 2 知道在除了 $(0,0)$ 和 $(1,0)$ 的其他所有状态中的博弈。在状态 $(1,0)$ 和 $(1,1)$ 的每个状态中，参与人 1 知道博弈是 G_b，但不知道参与人 2 知道它。同理在状态 $(1,1)$ 和 $(2,1)$ 的每一状态中参与人 83
2 知道博弈是 G_b 但不知道参与人 1 是否知道参与人 2 知道博弈是 G_b。如此等等。在任一状态 (q, q) 或 $(q+1, q)$ 中，q 的值越大，这种类型的叙述"参与人 i 知道参与人 j 知道参与人 i 知道……博弈为 G_b"就越准确，但没有使博弈为 G_b 是共同知识的状态。

若 ϵ 是很小的，则以一大概率每一个参与人在他的屏幕上看到一个很大的数。当参与人 1 在她的屏幕上看到"1"，她不能确定参与人 2 是否知道博弈是 G_b，因此可能会迟疑地选 B。但如果在她屏幕上的数字为 17 的话，那么博弈是 G_b 就"几乎"是共同知识了，因此很可能她将遵从更想要的博弈 G_b 的均衡 (B, B)。她的决策将依赖于她的关于若他屏幕上的数字是 16 或 17 则参与人 2 将会做什么的信念。反之，参与人 2 的决策依赖于他的关于若她屏幕上的数字为 16 则参与人 1 将做什么的信念，如此等等。为了研究这些需考虑的事实，我们现在定义如下的贝叶斯博弈，它被称为电子邮件

博弈(electronic mail game)。

- 状态集合为 $\Omega=\{(Q_1,Q_2):Q_1=Q_2$ 或 $Q_1=Q_2+1\}$。
- 每个参与人 i 的信号函数 τ_i定义为 $\tau_i(Q_1,Q_2)=Q_i$。
- 每个参与人在 Ω 上的信念是一样的,它派生于技术(以ϵ 为特征)和以概率 $1-p$ 博弈为 G_a 的假设:$p_i(0,0)=1-p$,对任一非负整数 q 有 $p_i(q+1,q)=p\epsilon(1-\epsilon)^{2q}$和 $p_i(q+1,q-1)=p\epsilon(1-\epsilon)^{2q+1}$。
- 在每一状态(Q_1,Q_2)中支付由博弈 $G(Q_1,Q_2)$确定。

■命题 83.1 电子邮件博弈有惟一纳什均衡,在其中两个参与人经常选择 A。

证明:在状态(0,0)中对于参与人 1 行动 A 是强占优的,所以在任一纳什均衡中当参与人 1 收到信号 0 时便会选择 A。如果参与人 2 未收到信息(即他的信号为 0),那么他知道要么参与人 1 没发送信息(概率为 $1-p$ 的一个事件),要么参与人 1 发送的信息没到达(概率为 $p\epsilon$的一个事件)。如果参与人 2 选择 A,那么既然参与人 1 在状态(0,0)中选择了 A,则无论参与人 1 在状态(1,0)中选择什么,参与人 2 的期望支付至少为$(1-p)M/[(1-p)+p\epsilon]$,如果参与人 2 选择 B 则他的支付至多为$[-L(1-p)+p\epsilon M]/[(1-p)+p\epsilon]$。因此对参与人 2 来说当他的信号为 0 时选择 A 是惟一最优的。

假定现在我们已经证明了对所有满足 $Q_1+Q_2<2q$ 的(Q_1,Q_2)参与
84 人 1 和 2 在任一均衡中都会选择 A。考虑参与人1 当她发出 q 个信息时的决策。在此情形下参与人 1 不能确定是 $Q_2=q$ 还是$Q_2=q-1$。假定她没收到她的第 q 个信号的确认信息,她赋给 $Q_2=q-1$ 的概率是 $z=\epsilon/[\epsilon+(1-\epsilon)\epsilon]>\frac{1}{2}$。因此她相信比起参与人 2 收到信息,她的最后信息没有到达更可能发生(这是论证中的关键点)。如果她选择 B 那么她的期望支付至多是$z(-L)+(1-z)M$(因为在推导的假设下,她知道若 $Q_2=q-1$ 那么参与人 2 选择 A);如果她选择 A 那么她的支付至少为 0。给定 $L>M$ 和 $z>\frac{1}{2}$,她的最优行动因而是 A。由类似证明,如果参与人 1 和 2 在任一均衡中对所有满足 $Q_1+Q_2<2q+1$ 的(Q_1,Q_2)都选择 A,那么参与人 2 当他的信号为 q 时选择A。故每一个参与人根据每个可能的信号选择 A。 □

因此即使两个参与人都知道博弈是 G_b 且即使网络中的噪声(概率为

ϵ)任意小,参与人也好像他们并无信息那样行动并选 A,就如同在没有电子邮件系统时那样做。

如果在你的屏幕上数字为 17,你将如何做?很难想像当 L 稍微超过 M 和 ϵ 是小的时候,一个看到屏幕上的数值为 17 的参与人不会选择 B。我们的直觉与博弈理论分析间的冲突使得均衡是自相矛盾的。在这方面有一长串博弈的例子(像有限重复囚徒困境(参看命题 155.1),连锁店博弈(参看第 6.5.1 节)和蜈蚣博弈(第 6.5.2 节),在其中我们的直觉与分析间的矛盾源于这样的事实,即数学推导并不是人类推理过程的一部分。

[注解]

在第 5.1 节所描述的基本知识模型形成于 1950 年和 1960 年,Hintikka (1962)是首创者。共同知识的概念应归于 Lewis(1969)和 Aumann(1976)。Lewis 给了一个非正式的定义(并且讨论了对第 5.1 和第 5.2 节的哲学背景),Aumann 给了一正式定义且证明了命题 75.1。第 5.4 节基于 Aumann and Brandenburger(1995)和 Brandenburger(1992)。(Spohn(1982)包含了一 85
个堪称命题 78.1 先驱者的结论)。第 5.5 节的电子邮件博弈由 Rubinstein (1989)所研究,在精神上它近乎于由电脑科学家所研究的"协作攻击问题"(例如可参看 Halpern(1986))。

第 5.1.3 节中帽子之谜的来源不详;参看 Littlewood(1953, p.3)。练习 76.2 基于 Milgrom 和 Stokey(1982),练习 81.1 基于 Aumann(1987a)。

关于相互作用知识模型(在其中参与人的信息函数不是分割的)的讨论参看 Bacharach(1985)和 Samet(1990)。对文献的综述参看 Binmore and Brandenburger(1990)和 Geanakoplos(1992, 1994)。

第二编

完全信息扩展博弈

扩展博弈(extensive game)具体描述参与人在战略情形中所遇到的决策问题的序列结构。该模型允许我们研究这样的解,即每个参与人不仅可以在博弈开始时考虑他的行动计划,且在任何一个不得不做决策的时点上,他都可以考虑他的行动计划。与此相反,战略博弈模型将我们限于这样的解,即每个参与人选择且仅选择一次他的行动计划,这个计划可包纳无限的变数,但战略博弈模型在博弈中的某些事件已知后则不允许参与人去重新考虑他的行动计划。

扩展博弈的一般模型不要求每个参与人在做决策时都知道以前所发生的所有事件。我们将在第三编研究这种模型。在本编我们考察一个简单模型,每个参与人在博弈中的每个时点上都完全地知道所有参与人以前的行动。在第 6 章我们描述基本模型,在随后的三章我们研究两类有趣的具有完全信息的扩展博弈:轮流出价讨价还价博弈(第 7 章)和重复博弈(第 8、第 9 章)。在第 10 章我们提供一些关于实施理论的主要结论(同时使用战略和扩展博弈模型)。

89 # 第 6 章

完全信息扩展博弈

本章将研究具有完全信息的扩展博弈模型。我们将论证在这一模型中纳什均衡解的概念是不令人满意的，因为它忽略了决策问题的序列结构。我们还将定义子博弈精炼均衡这一概念，在其中参与人被要求随计划进行去重估他的计划。在本章末尾我们将把这一解的概念与反复剔除弱劣行动的解的概念作比较。

6.1 完全信息扩展博弈

6.1.1 定义

扩展博弈是对参与人在战略情形中所遇到的决策问题的序列结构的详细描述。如果每个参与人在做决策时都完全知道以前所发生的所有事件，则在这一博弈中存在完全信息。为了简便我们将从这样的博弈开始，即没有两个参与人同时作决策，且所有相关行动都是参与人所为（无随机因素干扰）。（我们将在 6.3 部分除掉这两个限制。）

▶定义 89.1 **完全信息扩展博弈**（extensive game with perfect information）有下列组成部分：

• 一个集合 N（参与人的集合）。

90 • 一个序列集合 H（有限或无限的）满足下列三条性质：

• 空序列 ϕ 是 H 的一个元素。

• 如果 $(a^k)_{k=1,\cdots,K}^{\infty} \in H$（这里 K 可能是无限的）且 $L<K$，那么 $(a^k)_{k=1,\cdots,L}^{\infty} \in H$。

• 如果一个无限序列$(a^k)_{k=1,\cdots}$对每个正整数 L 满足$(a^k)_{k=1,\cdots,L}\in H$ 那么$(a^k)_{k=1,\cdots}\in H$。

(H 的每个元素都是一段**历史**(history);一段历史的每一组成部分都是一个参与人采取的**行动**。)一段历史$(a^k)_{k=1,\cdots,K}\in H$ 是**终点**(terminal)如果它是无限的或如果没有 a^{K+1}使得$(a^k)_{k=1,\cdots,K+1}\in H$。终点历史集合由 Z 表示。

• 一个函数 P,它赋给每个非终点历史($H\setminus Z$ 的每一元素)一个 N 的元素。(P 是**参与人函数**(player function),$P(h)$是历史 h 后采取行动的参与人。)

• 对每个参与人 $i\in N$ 有一个Z 上的偏好关系$\succsim_i$(参与人 i 的**偏好关系**)。

有时为了方便在确定一个扩展博弈的结构时不须确定参与人的偏好,我们称组成部分满足定义中前三个条件的一个三元组$\langle N,H,P\rangle$为一**完全信息扩展博弈形式**(extensive game form with perfect information)。

如果可能的历史集合 H 是有限的,则博弈是有限的。如果最长历史的长度是有限的,则博弈有一有限边界(finite horizon)。令 h 为一段长度为k的历史,我们用(h,a)表示长度为 $k+1$ 且包含由 a 紧随的h 的历史。

本章从头至尾将一个完全信息扩展博弈称为"扩展博弈"。我们解释这一博弈如下。在任一非终点历史 h 之后,参与人 $P(h)$从下列集合选择一个行动

$$A(h)=\{a:(h,a)\in H\}.$$

空历史是博弈的起始点;我们有时称它为*初始历史*(initial history)。在这点上参与人 $P(\varnothing)$选择 $A(\varnothing)$的一个元素。对这集合中每个可能的选择 a^0 参与人 $P(a^0)$随后选择集合 $A(a^0)$的一个元素;这个选择决定下一个参与人如何行动,如此等等。一段历史在其之后没有选择要做,则它是终点。注意一段历史可能是一无限行动序列。将一段历史定义为一个序列(而不是作为一个更复杂的数学对象,像一序列串)包含了这样的假设,即在任一无限长历史之后没有行动可被采用,所以每一段这种历史都是终点。同在战略博弈情形中一样,我们经常通过给定代表偏好的支付函数来确定参与人关于终点历史的偏好。

◇例 91.1 两人使用下列过程去分配两个想要的独立的不可分割的 91
物体。他们中的某一个人提出一种分配方式,另一个人可能接受也可能拒

绝。如果拒绝，两人都得不到任何东西。每个人仅关心所得的物体数量。

对个人困境进行建模的一个扩展博弈是$\langle N, H, P, (\succsim_i)\rangle$，其中，

- $N=\{1,2\}$；
- H 包括 10 段历史：$\varnothing$，$(2,0)$，$(1,1)$，$(0,2)$，$((2,0),y)$，$((2,0),n)$，$((1,1),y)$，$((1,1),n)$，$((0,2),y)$，$((0,2),n)$；
- $P(\phi)=1$ 和 $P(h)=2$，(对任一非终点历史 $h\neq\phi$)；
- $((2,0),y)\succ_1((1,1),y)\succ_1((0,2),y)\sim_1((2,0),n)\sim_1((1,1),n)\sim_1((0,2),n)$且$((0,2),y)\succ_2((1,1),y)\succ_2((2,0),y)\sim_2((0,2),n)\sim_2((1,1),n)\sim_2((2,0),n)$。

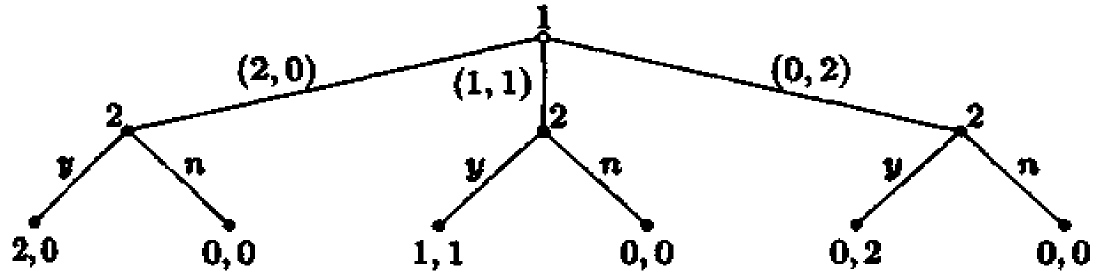


图 91.1　一个扩展博弈，它模化了例 91.1 中在两人间分配两个独立的不可分割的物体的过程

该博弈一个便利的表示见图 91.1。图表顶部的小圈代表初始历史 $\varnothing$（博弈的起始点）。这个圈上的 1 指 $P(\varnothing)=1$(参与人 1 首先行动)。从圈发出的三个线段对应于 $A(\varnothing)$的三个元素(在初始历史参与人 1 可能的行动)；在三个线段旁的标记是行动的名称，$(k,2-k)$是 k 个物体给参与人 1 和剩下的 $2-k$ 个给参与人 2 的提议。每条线段都指向一个小圆饼，在它旁边的标记 2 指参与人 2 在任一长度为 1 的历史后采取一行动。在从圆饼发出的线段旁的标记是参与人 2 的行动名，y 意指“接受”，n 意指“拒绝”。端点历史之下的数字是代表了参与人偏好的支付。(在每一个二元组中第一个数代表了参与人 1 的支付，第二个数字代表了参与人 2 的支付。)

92 图 91.1 表示了扩展博弈的第一个定义，在此定义中基本的成分是树(无圈的连通图)。在这种形式中每个结对应一段历史且任一对相连的结对应于一个行动；行动名称不是定义的一个部分。这个定义更常用，不过我们发现定义 89.1 将参与人的行动作为基本要素，它更自然。

6.1.2　战略

在扩展博弈中参与人的一个战略是对每段历史(在它之后轮着他去行

动)确定由参与人选择的行动。

▶**定义 92.1**　在一个完全信息扩展博弈$\langle N, H, P, (\succsim_i)\rangle$中参与人 $i \in N$ **的一个战略**(strategy of player i)是一个函数,它对每个非终点历史 $h \in H \setminus Z$(满足 $P(h)=i$)赋给 $A(h)$中的一个行动。

注意在博弈$\langle N, H, P, (\succsim_i)\rangle$中参与人的战略仅依赖于博弈形式$\langle N, H, P\rangle$。

为了阐述战略的概念让我们来考虑图 91.1 中的博弈。参与人 1 仅在初始历史 $\varnothing$ 之后采取一个行动,所以我们对于他的每一个战略可以用这段历史后所能采取的三个可能行动(2,0),(1,1)和(0,2)中的一个来确定。参与人 2 在三段历史(2,0),(1,1),(0,2)中的每个之后采取一个行动并且在每一情形下他有两个可能行动。因此我们可用一个三元组 $a_2b_2c_2$ 来确定他的每一个战略,这里 a_2、b_2 和 c_2 是在历史(2,0),(1,1)和(0,2)之后采取的行动。参与人 2 的战略 $a_2b_2c_2$ 的解释是它是一个权变计划:若参与人 1 选择(2,0)则参与人 2 选择 a_2;若参与人 1 选择(1,1)则参与人 2 选择 b_2;若参与人 1 选择(0,2)则参与人 2 将选 c_2。

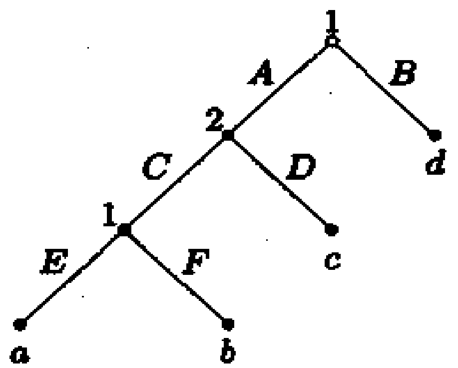


图 93.1　参与人 1 与参与人 2 前后都行动向扩展博弈

图 93.1 中的博弈阐述了重要的一点:一个战略对每段历史(在此历史之后轮着参与人采取行动)确定由参与人选择的行动,即使在战略被领悟的情况下对永远到达不了的历史也一样。在此博弈中参与人 1 有四个战略 AE, AF, BE, BF。那即是,她的战略在历史(A, C)后确定一个行动,即使该战略已确定在博弈开始时她选择 B。在此意义下战略不同于我们自然所认为的战略是行动计划;在第 6.4 节我们还要谈及这点。正如我们马上要看到的那样,为了某些目的我们可以将 BE 和 BF 视为同一战略;不过,在 93
别的情形下将它们区分开是很重要的。

对扩展博弈$\langle N, H, P, (\succsim_i)\rangle$中的每个战略组合$s=(s_i)_{i\in N}$我们定义$s$的**结果**(outcome)$O(s)$为当每个参与人$i\in N$遵从$s_i$的规程时由$s$所导致的终点历史。也就是$O(s)$是(可能无限的)历史$(a^1,\cdots,a^K)\in Z$使得对$0\leqslant k<K$我们有$s_{P(a^1,\cdots,a^k)}(a^1,\cdots,a^k)=a^{k+1}$。

同在战略博弈中一样,我们可定义一个混合战略为(纯)战略集合上的一个概率分布。在完全信息扩展博弈中考虑这种战略只要稍加点东西。因此当参与人采取行动时他们并不完全知道信息的扩展博弈;在这类博弈中混合战略概念有更多意义。

6.1.3 纳什均衡

对于扩展博弈我们定义的第一个解的概念忽略了博弈的序列结构;它把战略当作在采取行动前做一次且仅做一次的选择。

▶定义 93.1 **完全信息扩展博弈$\langle N, H, P, (\succsim_i)\rangle$的纳什均衡**是一战略组合$s^*$使得对每个参与人$i\in N$我们有

$O(s^*_{-i}, s^*_i)\succsim_i O(s^*_{-i}, s_i)$,对参与人$i$的每个战略$s_i$.

相应地,我们能定义扩展博弈Γ的纳什均衡为定义如下的由Γ派生的战略博弈纳什均衡。

94 ▶定义 94.1 **完全信息扩展博弈$\Gamma=\langle N, H, P, (\succsim_i)\rangle$的战略形式**(strategic form of the extensive game with perfect information)是战略博弈$\langle N, (S_i), (\succsim'_i)\rangle$,在其中对每个参与人$i\in N$有

- S_i是Γ中参与人i的战略集合。
- $\succsim'_i$由$s\succsim'_i s'$确定当且仅当对每一个$s\in\times_{i\in N}S_i$和$s'\in\times_{i\in N}S_i$有$O(s)\succsim_i O(s')$。

?练习 94.2 令G为一个两人战略博弈$\langle\{1,2\}, (A_i), (\succsim_i)\rangle$,其中每个参与人有两个行动:对$i=1,2$有$A_i=\{a'_i, a''_i\}$。试说明$G$是完全信息扩展博弈的战略形式当且仅当要么对某一$a_1\in A_1$,我们对$i=1,2$有$(a_1, a'_2)\sim_i(a_1, a''_2)$,或者对某一$a_2\in A_2$我们对$i=1,2$有$(a'_1, a_2)\sim_i(a''_2, a_2)$。

如果纳什均衡是我们为扩展博弈所定义的惟一解,那么我们可以定义

范围更小的一个战略：我们可以要求一个战略仅在与它在博弈中较早时点上所确定的行动并不一致的历史之后，去确定一个参与人的行动。之所以这样，是因为战略组合 s 的结果 $O(s)$ 在与 s_i 不一致的权变之后并不受任一参与人 i 的战略 s_i 所确定的行动的影响。准确地说，我们能定义参与人 i 的一个简化战略(reduced strategy)为一函数 f_i，它的定义域是 $\{h \in H : P(h) = i\}$ 的一个子集且有下列性质：(ⅰ)它将 f_i 定义域中的每段历史 h 与 $A(h)$ 中的一行动联系起来；(ⅱ)满足 $P(h) = i$ 的一段历史 h 在 f_i 的定义域里当且仅当在 h 中参与人 i 的所有行动都由 f_i 决定(也就是，若 $h = (a^k)$ 和 $h' = (a^k)_{k=1,\cdots,L}$ 是满足 $P(h') = i$ 的 h 的一个子集，那么 $f_i(h') = a^{L+1}$)。参与人 i 的每一个简化战略都对应于参与人 i 的一个战略集合；对于别的参与人的任一战略向量，在这个集合中的每一个战略都产生同一结果(也就是，集合中的战略是结果等价的(outcome-equivalent)。扩展博弈的纳什均衡集合对应于战略博弈的纳什均衡，在其中每个参与人的行动集合是他的简化战略集合。(一个战略的完美定义对于子博弈精炼均衡是必要的，我们将在下一部分定义它。)

作为参与人在扩展博弈中简化战略集合的一个例子，可以考虑图 93.1 中的博弈。参与人 1 有三个简化战略：一个由 $f_i(\varnothing) = B$(具有定义域 $\{\varnothing\}$)确定，一个由 $f_i(\varnothing) = A$ 和 $f_i(A, C) = E$(具有定义域 $\{\varnothing, (A, C)\}$)确定，一个由 $f_i(\varnothing) = A$ 和 $f_i(A, C) = F$(具有定义域 $\{\varnothing, (A, C)\}$ 确定。

对于一些博弈，在下列意义下某个参与人的一些简化战略是等价的：不 95
管别的参与人的战略，对所有参与人它们产生同样的支付(尽管不是同一结果)。也就是从参与人支付的观点看，对一些博弈在战略的定义中冗余成分，其超过了由简化战略的概念所得到的。例如，在图 93.1 的博弈中若 $a = b$，则参与人 1 的两个简化战略(其中她在博弈开始时选择 A)从支付的观点看是等价的。为了抓住这一深层次的冗余及由简化战略概念所得到的冗余，我们可定义下列战略形式的变形。

▶定义 95.1　令 $\Gamma = \langle N, H, P, (\succsim_i)\rangle$ 为一完全信息扩展博弈且令 $\langle N, (Si), (\succsim'_i)\rangle$ 为它的战略形式。对任一 $i \in N$ 定义参与人 i 的战略 $s_i \in S_i$ 和 $s'_i \in S_i$ 为等价的如果对每一 $s_{-i} \in S_{-i}$ 我们有：对所有 $j \in N$ 有 $(s_{-i}, s_i) \sim'_j (s_{-i}, s'_i)$。$\Gamma$ **的简化战略形式**(reduced strategic form)是战略博弈 $\langle N, (S'_i), (\succsim''_i)\rangle$，其中对每一 $i \in N$，每个集合 S'_i 包含 S_i 中每个等价战略的一个元素，并且 $\succsim''_i$ 是由 $\succsim'_i$ 导致的关于 $\times_{j \in N} S'_j$ 的偏好次序。

(注意这个定义指定了在简化战略形式中行动的名称;这些行动的每个选择都确定了一个不同的简化战略形式。不过,在任何传统的博弈理论分析中行动的名称没有多大的意义,所以我们称为博弈简化战略形式。)

图93.1中博弈的战略形式和简化战略形式在图96.1中表示出来了。如果 $a=b$,那么参与人1的战略 AE 和 AF 是等价的,所以参与人1在博弈的简化战略形式中只有两个行动。

下个例子阐明了纳什均衡的概念并揭示了该均衡可能具有的一个不受欢迎的特征。

◇例95.2 图96.2 中的博弈有两个纳什均衡:(A,R)和(B,L),支付组合为(2,1)和(1,2)。战略组合(B,L)是一纳什均衡,因为给定在历史 A 后参与人2选择 L,则对参与人1来说在博弈开始时选择 B 是最优的(若她选 A,则给定参与人2的选择她获得0而非1);给定参与人1的选择 B 则对参与人2来说选择 L 是最优的(因为他的选择不会使结果变化。)

	C	D
AE	a	c
AF	b	c
BE	d	d
BF	d	d

	C	D
AE	a	c
AF	b	c
B	d	d

图96.1 战略形式(左)和简化战略形式(右)

我们将一个非终点历史解释为一个时点(在此点上一个参与人可能会重估他的行动计划)导致了一个关于此博弈中纳什均衡(B,L)缺乏说明力
96 的讨论。如果历史 A 将发生,那么参与人似乎会选 R 而非 L,因为这样做他能获得一个更高的支付。均衡(B,L)由若参与人1选择 A,则参与人2选择 L 的"威胁"所维持。这个威胁并不可信,因为参与人2没有办法让自己去做这一选择。因此参与人1会确信若她选择 A,则参与人2将选择 R;因为她对结果(A,R)的偏好优于纳什均衡结果(B,L),所以她愿去偏离均衡且选择 A。在下一部分我们将定义一个抓住了这些可能发生事件

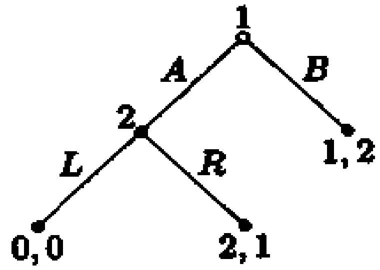


图 96.2 一个有两个参与人的扩展博弈的例子

的均衡的概念。

◇例 96.1 图 91.1 中博弈的纳什均衡是((2,0), *yyy*),((2,0), *yyn*),((2,0) *yny*,),((2,0), *ynn*),((1,1), *nyy*),((1,1), *nyn*),((0,2), *nny*),((2,0), *nny*)和((2,0), *nnn*)。前四个导致分配(2,0),随后两个导致分配(1,1),最后两个各导致分配(0,2)和(0,0)。除了((2,0), *yyy*)和((1,1), *nyy*)所有这些均衡都涉及参与人 2 在某段历史之后缺乏说服力的一个行动(因为他拒绝了至少给他一个物体的提议);像例 95.2 中的均衡(B, L) 97
一样,它们都被我们现在定义的均衡概念排除了。

6.2 子博弈精炼均衡

以前一部分结束时的讨论为动因,我们现在定义子博弈精炼均衡的概念。我们先定义子博弈的概念。

▶定义 97.1 **完全信息扩展博弈 $\Gamma=\langle N,H,P,(\succsim_i)\rangle$ 继承历史 h 的子博弈**(the subgame of the extensive game with perfect information that follows the history)是扩展博弈 $\Gamma(h)=\langle N,H|_h,P|_h,(\succsim_i|_h)\rangle$,这里 $H|_h$ 是满足$(h,h')\in H$ 的行动序列h'的集合,$P|_h$ 由对每一$h'\in H|_h$ 有$P|_h(h')=P(h,h')$确定,$\succsim_i|_h$ 由 $h'\succsim_i|_h h''$确定当且仅当$(h,h')\succsim_i(h,h'')$。

我们现在定义的均衡概念要求在每段历史之后给定别的参与人的战略,且每个参与人的战略所规定的行动是最优的。给定扩展博弈 Γ 中参与人 i 的战略 s_i 和一段历史 h,用 $s_i|_h$ 表示在子博弈 $\Gamma(h)$中由 s_i 所导致的战

略(即对每一 $h' \in H|_h$ 有 $s_i|_h(h') = s_i(h, h')$;用 O_h 表示 $\Gamma(h)$的结果函数。

▶定义 97.2 **完全信息扩展博弈 $\Gamma = \langle N, H, P, (\succsim_i)\rangle$ 的一个子博弈精炼均衡**(a subgame perfect equilibrium of an extensive game with perfect information)是一战略组合 s^* 使得对每个参与人 $i \in N$ 和每段满足 $P(h) = i$ 的非终点历史 $h \in H \setminus Z$ 我们有对子博弈 $\Gamma(h)$中参与人 i 的每个战略 s_i,

$$O_h(s^*_{-i}|_h, s^*_i|_h) \succsim_i|_h O_h(s^*_{-i}|_h, s_i).$$

等价地,我们可定义子博弈精炼均衡为 Γ 中的战略组合 s^* 满足对任一段历史 h 战略组合 $s^*|_h$ 是子博弈 $\Gamma(h)$的一个纳什均衡。

子博弈精炼均衡的概念排除了这样的纳什均衡,即参与人的威胁并不可信,例如图 96.2 的博弈中惟一的子博弈精炼均衡是(A, R),图 91.1 中的博弈中仅有的子博弈精炼均衡是$((2,0), yyy)$和$((1,1)nyy)$。

◇例 97.3 斯塔克伯格博弈(Stackelberg games)。斯塔克伯格博弈是两人完全信息扩展博弈,在其中一个"领导者"从集合 A_1 中选择一个行动,
98 另一"追随者"在知道领导者的选择后从集合 A_2 中选择一个行动。经常用于经济学中的这类博弈的解是子博弈精炼均衡解(尽管这个术语不常用)。斯塔克伯格博弈的一些(非全部)子博弈精炼均衡对应下列最大化问题的解

$$\max_{(a_1, a_2) \in A_1 \times A_2} u_1(a_1, a_2),\text{以 } a_2 \in \arg\max_{a'_2 \in A_2} u_2(a_1, a'_2)\text{为条件},$$

这里 u_i 是代表参与人 i 的偏好的支付函数。如果每个参与人 i 的行动集合 A_i 是紧的,且支付函数 u_i 是连续的,则该最大化问题有解。

☑练习 98.1 试举一例:一个斯塔克伯格博弈的一子博弈精炼均衡并不对应于上述最大化问题的一个解。

为了验证战略组合 s^* 是子博弈精炼均衡,定义 97.2 要求我们对每个参与人 i 和每个子博弈去检验不存在导致参与人 i 所偏好的结果的战略。下列的结论说明了在一有限边界博弈中对每个参与人 i 和每一子博弈我们可将注意力限于这些可选择的战略——仅在一段历史之后在由这些战略所确定的行动中它们不同于 s^*_i。特别地,一个战略组合是一子博弈精炼均衡,当且仅当对每个子博弈首先采取行动的参与人通过仅改变他的初始行动不能获得一个更好的结果。对于一扩展博弈 Γ 用 $l(\Gamma)$表示在 Γ 中最长历史的长度;我们称 $l(\Gamma)$为 Γ 的长度。

■引理 98.2 （一次偏离性质）(The one deviation property) 令 $\Gamma=\langle N,H,P,(\succsim_i|_h)\rangle$ 为一有限边界完全信息扩展博弈。战略组合 s^* 是 Γ 的子博弈精炼均衡当且仅当对每一个参与人 $i\in N$ 和每段满足 $P(h)=i$ 的历史 $h\in H$ 我们有：对子博弈 $\Gamma(h)$ 中参与人 i 的每个战略 s_i

$$O_h(s_{-i}^*|_h,s_i^*|_h)\succsim_i|_h O_h(s_{-i}^*|_h,s_i)$$

其中 s_i 在 $\Gamma(h)$ 的初始历史之后仅在由它所规定的行动中不同于 $s_i^*|_h$。

证明：如果 s^* 是 Γ 的一个子博弈精炼均衡，则它满足条件。现在假定 s^* 不是子博弈精炼均衡；假定参与人 i 在子博弈 $\Gamma(h')$ 中能有利地偏离（即采取别的行动可获得更好的结果）。那么在 $\Gamma(h')$ 中存在参与人 i 的一个有利的偏离战略，满足对具有多个阶段且不大于 $\Gamma(h')$ 长度的历史 h，使得 $s_i(h)\neq(s_i^*|_{h'})(h)$；因为 Γ 有有限边界，则这个数是有限的。从 $\Gamma(h')$ 中参 99
与人 i 的所有有利的偏离中选取一战略 s_i，满足使 $s_i(h)\neq(s_i^*|_{h'})(h)$ 的 h 的长度是最小的。令 h^* 为 $\Gamma(h')$ 的满足 $s_i(h)\neq(s_i^*|_{h'})(h)$ 最长历史 h。那么 $\Gamma(h^*)$ 的初始历史是 $\Gamma(h^*)$ 中惟一这样的历史，即在该点由 s_i 所确定的行动不同于由 $s_i^*|_{h'}$ 所确定的行动。进一步说，$s_i|_{h^*}$ 是 $\Gamma(h^*)$ 中的一个有利偏离，因为不是那样的话，在 $\Gamma(h^*)$ 中就有一个有利偏离在比 s_i 更短的历史后不同于 $s_i^*|_{h'}$。因此 $s_i|_{h^*}$ 是 $\Gamma(h^*)$ 中的一个有利偏离，在 $\Gamma(h^*)$ 的初始历史之后仅在由它所规定的行动中不同于 $s_i^*|h^*$。 □

☑练习 99.1 试举一个不具有一次偏离性质的有限边界博弈的例子。

我们现在要证明每个有限完全信息扩展博弈都有一个子博弈精炼均衡。我们的证明是构造式的：对博弈中每一个最长的非终点历史，我们选择一个对于轮到要行动的参与人来说是最优的行动，并且用一个这样的终点历史来代替那些历史，即在这个终点历史中，支付组合是最优行动被选择后所产生的；然后我们重复这个过程，一直回到博弈的开始。（下面的结论是著名的库恩(Kuhn)定理）

■命题 99.2 每一个有限完全信息扩展博弈都有一个子博弈精炼均衡。

证明：令 $\Gamma=\langle N,H,P,(\succsim_i)\rangle$ 为一有限完全信息扩展博弈。我们通过对 $l(\Gamma(h))$ 的归纳来构造 Γ 的一个子博弈精炼均衡；同时我们定义一个将每段历史 $h\in H$ 与一个终点历史相联系起来的函数 R，并且要证明这个历史是子博弈 $\Gamma(h)$ 的子博弈精炼均衡结果。

如果 $l(\Gamma(h))=0$（即 h 是 Γ 的终点历史）定义 $R(h)=h$。现在假定

$R(h)$的定义域是这样的：对某个 $k\geqslant 0$，所有满足 $l(\Gamma(h))\leqslant k$ 的 $h\in H$。令 h^* 为满足 $l(\Gamma(h^*))=k+1$ 的一段历史，且令 $P(h^*)=i$。因为 $l(\Gamma(h^*))=k+1$，我们对所有 $a\in A(h^*)$ 有 $l(\Gamma(h^*,a))\leqslant k$。定义 $s_i(h^*)$ 为 $R(h^*,a)$在 $a\in A(h^*)$上的一个 $\succsim_i$ 最大化者，并且定义 $R(h^*)=R(h^*,s_i(h^*))$。通过归纳我们现在已定义了 Γ 中的一个战略组合 s；由引理 98.2 这个战略组合是 Γ 的一个子博弈精炼均衡。 □

该证明中所用的方法常被称为逆向归纳法(backwards induction)。除了作为证明该命题的一个技术手段，这个方法也是计算有限博弈子博弈精
100 炼均衡集合的算法。子博弈精炼均衡概念的部分奇妙之处为派生于该算法描述了这样的方法，即只要边界是相对短的，则对参与人分析这种博弈来说似乎是自然的。

配以第 2.5 节中关于严格竞争博弈的结论，我们能从上述结论中得到的一个推论是：象棋博弈中的每个参与人都有一个确保其均衡支付的战略(该结论最早由 Zermelo(1913)证明)。因为象棋有有限多可能的历史(一旦一个位置被重复三次则博弈宣布为平局)，命题 99.2 表明了它有一个子博弈精炼均衡，且因而也有一个纳什均衡；因为它是严格竞争的，命题 22.2 表明了均衡支付是惟一的，且一个参与人的任一纳什均衡战略都能保证这个参与人得到他的均衡支付。要么白方有一种保证它能赢的战略，要么黑方有一种保证博弈的结果对它来说或赢或平局的战略。

? 练习 100.1 试证明 Kuhn 定理(命题 99.2)中的条件“博弈是有限的”不能被条件“它有一有限边界”或条件“在任一段历史每个参与人有有限多可能的行动”所代替。

注意 Kuhn 定理并未谈及惟一性。实际上图 91.1 中的博弈有两个子博弈精炼均衡((2,0), yyy)和((1,1), nyy)，按照每个参与人偏好的说法它们是不等价的。不过，显然一个博弈若在其中没有参与人对任何两个结果的偏好不一样则它有一个惟一的子博弈精炼均衡。进一步地说，只要任一参与人是无差别的，则所有参与人对任何两个结果偏好都一样，于是即使可能有多于一个的子博弈精炼均衡，所有参与人对所有子博弈精炼均衡偏好都一样。这个结论在下列练习中被表示出来。

? 练习 100.2 如果只要对某一 $i\in N$ 有 $z\sim_i z'$ 则对所有 $j\in N$ 有 $z\sim_j z'$，那么说一个有限完全信息扩展博弈满足非无差别条件(no indiffer-

ence condition)。这里 z 和 z' 是终点历史。试通过应用对子博弈长度的归纳来证明每个参与人对满足该条件博弈的所有子博弈均衡结果是无差别的。同时证明如果 s 和 s' 是子博弈精炼均衡，那么 s'' 也是，这里对每个参与 101
人 i 战略 s''_i 等价于 s_i 或 s'_i（即博弈的均衡是可互换的）。

?练习 101.1 试证明扩展博弈 Γ 的一个子博弈精炼均衡也是由下列方法从 Γ 得到的博弈的一个子博弈精炼均衡：删除在均衡中不能到达的子博弈并将已删除的子博弈中均衡的结果赋给这样产生的终点历史。

?练习 101.2 令 s 为完全信息扩展博弈 Γ 中的一个战略组合；假设 $P(h)=i$，$s_i(h)=a$ 和 $a'\in A(h)$ 满足 $a'\neq a$。考虑由 Γ 对某一个行动序列 h' 删除所有形如 (h,a',h') 的历史所得到的博弈 Γ'，且令 s' 为 Γ' 中由 s 所导致的战略组合。试证明如果 s 是 Γ 的一子博弈精炼均衡，那么 s' 是 Γ' 的一子博弈精炼均衡。

?练习 101.3 军队 1 和 2 正为争夺一个岛而战斗，该岛开始是由军队 2 的一个营所占领的。军队 1 有 K 个营，军队 2 有 L 个。只要该岛由某支军队占领，另一支军队就可发动一场进攻。战斗的结果是占领的营和攻击中的一个营被消灭；攻击军队胜，且只要它有剩下的营，则只要一个营来占领该岛。每支军队的指挥官对最大化生存的营的数量感兴趣，且将对岛的占领视为值一个营。（如果在一场战斗后两支军队都无剩余的营，那么每个指挥官的支付为 0。）将此情形作为一个扩展博弈来分析，并利用子博弈精炼均衡概念来预测胜者为 K 和 L 的一个函数。

6.3 博弈定义的两种扩展

像在定义 89.1 中一样，完全信息扩展博弈模型很容易从两个方面扩展。

6.3.1 外部不确定性

首先我们扩展该模型使其涵盖有某一外部不确定性。一个完全信息和机会行动扩展博弈（extensive game with perfect information and chance moves)是一五元组$\langle N, H, P, f_c, (\succsim_i)\rangle$，这里同前，$N$ 是一个有限参与人集合，H 是历史集合，且

102 • P 是一个从 H 中的非终点历史到 $N\cup\{c\}$ 上的函数。（若 $P(h)=c$，则机会确定历史 h 后所采取的行动。）

• 对每个满足 $P(h)=c$ 的 $h\in H$，$f_c(\cdot\mid h)$是 $A(h)$上的一个概率测度；每个这种概率测度都被假定为与其他任一个这种概率测度独立。（$f_c(a\mid h)$是历史 h 后 a 发生的概率。）

• 对每个参与人 $i\in N$，$\succsim_i$ 是关于终点历史集合上的不确定事件的偏好关系。

对每个参与人 $i\in N$ 战略的定义如前。一个战略组合的结果是关于终点历史的一个概率分布。子博弈精炼均衡的定义也同前(参看定义 97.2)。

[?] 练习 102.1　试证明一个完全信息和机会行动扩展博弈同时满足一次偏离性质(引理 98.2)和 Kuhn 定理(命题 99.2)。

6.3.2　同时行动

假定全部参与人在一段历史之后同时行动，并且他们中的每一位在做选择时都完全地知道以往所有的事件，我们可将完全信息扩展博弈的定义修改如下。一个完全信息和同时行动扩展博弈(extensive game with perfect information and simultaneous moves)是一四元组$\langle N, H, P, (\succsim_i)\rangle$，这里 N，H 和对每个参与人 i 的 $\succsim_i$ 都与定义 89.1 中一样，P 是一个对每个非终点历史赋一个参与人集合的函数，且 H 和 P 联合满足下列条件：对每段非终点历史 h，有一集合族$\{A_i(h)\}_{i\in P(h)}$，对于它 $A(h)=\{a:(h,a)\in H\}=\times_{i\in P(h)}A_i(h)$。

这种博弈中的一段历史是一个向量序列；每个向量 a^k 的组成部分是在历史$(a^l)_{l=1}^{k-1}$之后由轮到行动的参与人所选择的行动。在其中每个参与人 $i\in P(h)$可在历史 h 之后做选择的行动集合是 $A_i(h)$；其解释是在 $P(h)$中参与人的选择是同时做出的。

在这种博弈中参与人 $i\in N$ 的一个战略是将 $A_i(h)$中的一个行动赋给每个满足 $i\in P(h)$的非终点历史 h 的一个函数。除了“$P(h)=i$”由“$i\in P(h)$”代替外，子博弈精炼均衡的定义同在定义 97.2 中的一样。

⑦练习 103.1 假设三个参与人用下述过程分享一块蛋糕。参与人 1 首先提出一种分法，参与人 2 和 3 同时回答为“yes”或“no”。如果参与人 2 和 3 都说：“yes”，那么分割完成；否则任何参与人一无所得。每个参与人的偏好是宁多勿少。试将此情形系统表达成一个同时行动扩展博弈，并找出它的子博弈精炼均衡。 103

⑦练习 103.2 考虑下述两人博弈。开始参与人 1 可选择“停”或“继续”。如果她选择停则博弈以支付(1,1)结束；如果她选择继续，则所有参与人同时报出非负整数且每个参与人的支付是这些数字的结果。试将此情形系统表达成一个同时行动扩展博弈并找出它的子博弈精炼均衡。

⑦练 103.3 试证明对于同时行动扩展博弈满足一次偏离性质(引理 98.2)而不满足 Kuhn 定理(命题 99.2)。

6.4 关于战略的解释

正如我们已经注意到的，战略的定义(92.1)并不对应于一个行动计划，因为它要求参与人在这样的历史之后确定他的行动，即若他执行他的计划，则这段历史不可能发生。如我们以前所见，在图 104.1 的博弈中，参与人 1 的战略既要确定在博弈开始又要确定在历史(A,C)之后她所采取的行动，即使在博弈开始时她所采取的行动是 B。

对若战略被实行则不可能发生的历史相对应的某个参与人战略的组成部分，可作如下解释：它们是别的参与人关于那个参与人在他不遵从他计划的事件中将要做什么的信念。例如，在图 104.1 的博弈中，参与人 1 在历史(A,C)之后的行动，可被认为是参与人 2 关于参与人 1 在该历史之后将作的选择的信念，参与人 2 需要拥有这个信念去理性地选择一个行动。如果参与人 1 计划去选择 A，那么参与人 2 的信念就与参与人 1 计划好的在历史(A,C)之后的行动相一致。不过如果参与人 1 计划选择 B，那么这样一个信念不可能派生于参与人 1 的行动计划。在此情形下，参与人 1 的战略仍然提供这样一个信念。注意：参与人 2 关于参与人 1 的信念即使参与人 104
1 计划去选择 B，也是与博弈的分析相关的，因为为了理性化 B 的选择参与人 1 需要形成一个关于参与人 2 在历史 A 之后的计划的信念。

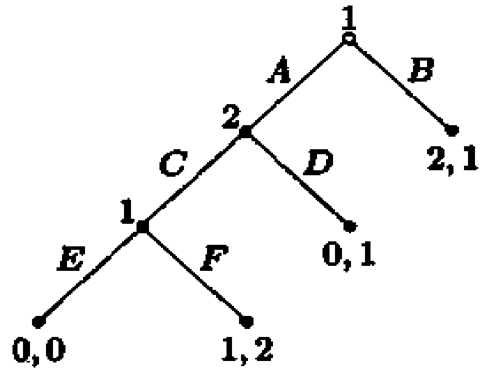


图 104.1 一个扩展博弈，在其中参与人 1 既在参与人 2 之前又在参与人 2 之后行动

这个解释有很多内涵。第一，谈论“战略的选择”变得有疑问了，因为一个参与人不会选择别的参与人的信念。第二，在任何一个具有两个以上参与人的博弈的均衡中有一隐含假设：除了任一给定的参与人 i 之外的所有参与人拥有同一个关于参与人 i 的行动的信念，这不存在于他遵从其行动计划的情形，而存在于在他偏离此计划的情形下。第三，如果有人想对战略施加限制条件，那他就得小心，因为他那时所做的假设不仅与参与人的行动有关，也与这些行动计划被违背时他们考虑每个其他人意图的信念有关。

这个对于战略的解释也减少了子博弈精炼均衡概念的吸引力。请再一次考虑图 104.1 中的博弈。在博弈的结构里对于参与人 2 没有办法去理性化参与人 1 对 A 的选择（因为参与人 1 对历史 B 的偏好优于当她选择 A 时所能产生的每段历史）。因此如果她观察到参与人 1 选择 A，参与人 2 必定放弃关于博弈的一个基本假设：她必须相信要么参与人 1 不是理性的，参与人 1 感觉到该博弈不同于图 104.1 的博弈；要么参与人 1 因为“错误”而选择 A（尽管这些错误在博弈的具体规定中是想象不到的）。可是子博弈精炼均衡概念要求不管他观察到什么历史，参与人 2 继续维持他的这些初始假设：参与人 1 是理性的，知道该博弈且不会犯错误。

105 6.5 两个值得注意的有限边界博弈

本节我们通过考察两个著名的博弈来展示子博弈精炼均衡概念的一些优点和弱点。为了便于描述这种博弈，我们引进一个离散的，从 1 期开始的

时间变量,这个变量对扩展博弈的正规模型不是一个附加物;它仅是用来简化博弈的描述和准确描述其结构的一个设计。

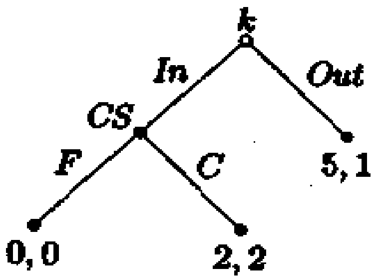


图 105.1　连锁店博弈中在城市 k 参与人选择的结构。每个二元组中的第一个数为连锁店的支付,第二个数为参与人 k 的支付

6.5.1　连锁店博弈

某连锁店(参与人 CS)在 K 个城市中有分店,城市标号为 $1,\cdots,K$。在每个城市 k 有惟一一个潜在竞争者,即参与人 k。在每个时期潜在竞争者中的其中一个决定是否与参与人 CS 竞争;在时期 k 轮着参与人 k 这样做。如果参与人 k 决定去竞争,那么连锁店可以抵制(F)也可合作(C)。连锁店在参与人 $k+1$ 做决定前对参与人 k 的决定作出反应。因此在时期 k 可能的结果集合是 $Q=\{Out,(In,C),(In,F)\}$。若在任一给定城市连锁店遇到挑战,则它宁愿合作而不抵制,不过连锁店在没有竞争者进入的情况下获得最大支付。每个潜在竞争者呆在外面比进入且遭到抵制更好,但是当它进入且连锁店是合作的时候获得最高支付。在单个时期参与人的选择和他们的考虑的结构在图 105.1 中被简化。

再加两个假设便完成了博弈的描述。第一,在博弈的每个时点上所有参与人都知道以前所选择的所有行动。这允许我们将该情形模化为一个完全信息扩展博弈,在其中历史集合是 $(\cup_{k=0}^{K}Q^k)\cup(\cup_{k=0}^{K-1}(Q^k\times\{In\}))$,这 106
里 Q^k 是 Q 的 k 个元素的所有序列的集合;参与人函数给定为:若 $h\in Q^k$ 则 $P(h)=k+1$ 且若 $h\in Q^k\times\{In\}$ 则对 $k=0,\cdots,K-1$ 有 $P(h)=CS$。第二,博弈中连锁店的支付是在 K 个城市它的支付的和。

该博弈有很多纳什均衡:每一个任何时点上的结果是 Out 或 (In,C) 的终点历史是一个纳什均衡结果。(在任一参与人 k 选择 Out 的均衡中连锁店的战略确定为:如果参与人 k 进入,它将抵制。)

相反,该博弈有惟一子博弈精炼均衡,在此均衡中每个挑战者选择 In 和连锁店经常选择 C。(在城市 K 不管历史连锁店必须选择 C,所以

在城市 $K-1$ 它必须同样做；继续这个过程人们可明白连锁店必须经常选择 C。

当 K 值较小时，非子博弈精炼的纳什均衡在直观上不太吸引人，而子博弈精炼均衡都是吸引人的。不过，当 K 较大时子博弈精炼均衡就失去吸收力了。在这个均衡中，连锁店的战略决定了不管它以前的行动如何，它都与每个进入者合作。给定我们关于战略的解释（参看前一部分），这就意味着：挑战者仍相信连锁店会与它合作——尽管他已经观察到连锁店会同很多进入者抗争。虽然连锁店的惟一子博弈精炼均衡战略确定了它会与每个进入者合作，但是对一个已经观察到连锁店重复抗争的竞争者来说，认为它的进入将会遇到一个侵略性的反应似乎更合理，特别是如果有很多城市有待竞争。如果一个挑战者进入，那么选择合作是连锁店眼前的利益，但直觉意味着为了阻止将来的进入，对侵略性行动建立一个威信是连锁店的长远利益。在第 12.3.2 节我们研究一种试图抓住该思想的连锁店不确化博弈，在那里挑战者不完全知道连锁店的动机。

6.5.2 蜈蚣博弈

两个参与人卷入了一个他们轮流有机会去终止的过程中。每个人对当他在任何时期 t 去终止该过程的结果的偏好优于当另一个参与人在 $t+1$ 期时这样做。不过，较好的仍然是过程在任一时期都不被终止所能产生的
107 任一结果。在 T 时之后（这里 T 是偶数）过程结束。对 $T=6$ 的情形由图 107.1 所示。（名字“蜈蚣”来自于图表的形状。）

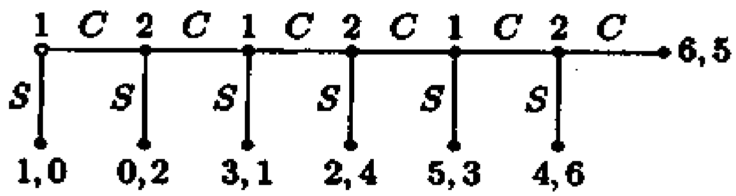


图 107.1 6 个时期的蜈蚣博弈

正式地，博弈中的历史集合包括所有长度为 $t(0\leqslant t\leqslant T)$ 的序列 $C(t)=(C,\cdots,C)$ 和所有包含由一个 S 跟随的 $t-1(1\leqslant t\leqslant T)$ 个重复的 C 的序列 $S(t)=(C,\cdots,C,S)$。参与人函数定义为：若 t 为偶数且 $t\leqslant T-2$，则 $P(C(t))=1$；若 t 为奇数，则 $P(C(t))=2$。对 $t\leqslant T-3$ 参与人 $P(C(t))$ 的偏好顺序是：$S(t+3)$ 优于 $S(t+1)$，$S(t)$ 优于 $S(t+2)$，参与人 1 的偏

好顺序是：$C(T)$优于 $S(T-1)$，$S(T-1)$优于 $S(T)$，参与人 2 的偏好顺序是：$S(T)$优于 $C(T)$。

该博弈有惟一子博弈精炼均衡，在此均衡里每个参与人在每一时期都选择 S。这个均衡的结果同每个纳什均衡的结果一样。为说明之，首先注意不存在结果为 $C(T)$的均衡。现在假定有一以参与人 i 在 t 期选择 S 而结束的纳什均衡(即在历史 $C(t-1)$之后)。如果 $t\geqslant 2$，那么参与人 1 可通过在 $t-1$ 期选择 S 来增加他的支付。因此在任一均衡中，参与人 1 在第 1 期选择 S。为使这对参与人 1 最优，参与人 2 必须在 2 期选择 S。纳什均衡的概念对参与人在较后时期的选择未加任何约束：参与人 1 在 1 期选择 S 和参与人 2 在 2 期选择 S 的任一战略二元组都是一个纳什均衡。(不过要注意该博弈的简化形式有惟一纳什均衡。)

在这个博弈的惟一子博弈精炼均衡中，每个参与人都相信另一个参与人将在下一个机会终止博弈，哪怕在一段参与人过去已选择了多次继续的历史之后。如同在连锁店博弈的精炼均衡中一样，这样一种信念在直观上不太吸引人；除非 T 很小，否则参与人 1 在博弈的开始立即选择 S 是不太可能的。蜈蚣博弈中的直觉与连锁店博弈中的直觉在任一长的历史之后，两个参与人都重复违背了深嵌于子博弈精炼均衡概念之中的理性规程。在一段参与人和他的对手过去已选择了多次的历史之后， 108
参与人形成关于他的对手在下期的行动的信念所依赖的基础就更不清楚了。

? 练习 108.1 对任一 $\epsilon>0$，定义战略博弈的一个 ϵ—均衡为一个行动组合，它满足这样的性质，即任何参与人都没有另一个可选择的行动，使他增加的支付超过 ϵ。试证明对任一正整数 k 和任一 $\epsilon>0$，有一边界 T 足够长使得修正为所有支付被 T 除的蜈蚣博弈战略形式有一个 ϵ—均衡，在其中第一个终止博弈的参与人在时期 k 这样做。

6.6 反复剔除弱劣战略

6.6.1 与子博弈精炼均衡的关系

在第 4.3 节我们明确了一个战略博弈的反复剔除弱劣行动的过程，并

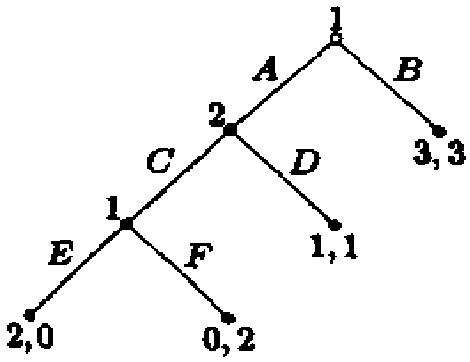
且论证了它比反复剔除强劣行动的过程缺少吸引力(尽管一个弱劣行动是对某一信念的最优反应),对于参与人来说它仍是用来简化博弈的自然方法。在 Kuhn 定理(命题 99.2)的证明中,我们明确了有限完全信息扩展博弈的逆向归纳过程,并且说明了它产生博弈的子博弈精炼均衡集合。

这两个过程是相关的。令 Γ 为一有限完全信息扩展博弈,在其中没有参与人对任何两个终点历史是无差别的。那么 Γ 有惟一的子博弈精炼均衡,现在定义一个在 Γ 的战略形式 G 中剔除弱劣行动的序列(Γ 中弱劣战略)满足:在过程结束时仍保留的所有 G 的行动组合产生 Γ 的惟一子博弈精炼均衡结果。

令 h 为一满足 $P(h)=i$ 和 $l(\Gamma(h))=1$ 的 Γ 的一段历史且令 $a_i^*\in A(h)$为对历史 h 由逆向归纳过程所选择的惟一行动。逆向归纳剔除了参与人 i 的在历史 h 之后选择不同于 a_i^* 的每个战略。一个不同于 a_i^* 的行动的参与人 i 的每个战略。在所有这些战略中,这些与 h 相一致的(即这些只
109 要 h' 是 h 的一段子历史($P(h')=i$ 就选择紧随 h' 的 h 的组成部分的)行动是 G 中的弱劣行动。在我们所定义的剔除序列中,所有这些弱劣行动在此阶段都从 G 中被剔除了。对每段满足 $l(\Gamma(h))=1$ 的历史 h 完成这种剔除后,我们转向满足 $l(\Gamma(h))=2$ 的历史 h 并且完成一个类似的剔除,用此方法我们一直回到博弈的开始。在此过程结束时仍保留的参与人 i 的每个战略选择由逆向归纳法在任何与参与人 i 的子博弈精炼均衡战略相一致的历史之后所选择的行动。因此特别地子博弈精炼均衡保留且每一个保留的战略组合产生惟一的子博弈精炼均衡结果。

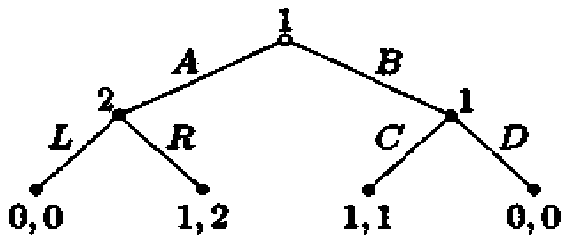
不过要注意别的剔除次序可能会去掉所有子博弈精炼均衡。例如在图 109.1 博弈中,惟一的子博弈精炼均衡是(BE,D),但如果在战略形式中弱劣行动 AE 被剔除则 D 在保留的博弈中是弱劣的;如果 AF 在 D 后被剔除,那么剩下的两个行动组合(BE,C)和(BF,C)都不是扩展博弈的子博弈精炼均衡。

也要注意若某一个参与人对两段终点历史是无差别的,那么可能(ⅰ)有一个剔除子博弈精炼均衡结果的剔除次序和(ⅱ)没有剔除次序使得所有剩下的战略组合产生子博弈精炼均衡结果,图 110.1 的博弈表明(ⅰ):参与人 1 的战略 AC, AD 和 BD 对 BC 是弱劣的,在它们都被剔除后剩下的行动二元组都不能产生子博弈精炼均衡结果(A,R)。若支付$(1,2)$由$(2,0)$所替代,那么调整后的博弈表明(ⅱ):结果(A,L)甚至不是一个纳什均衡结果,但对任一剔除顺序都能保留。



	C	D
AE	2,0	1,1
AF	0,2	1,1
BE	3,3	3,3
BF	3,3	3,3

图 109.1 扩展博弈(左)和它的战略形式(右)。存在一个剔除战略形式中弱劣行动的次序,它剔除扩展博弈的惟一子博弈精炼均衡



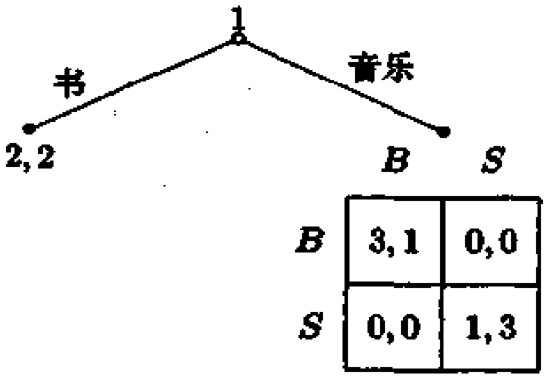
	L	R
AC	0,0	1,2
AD	0,0	1,2
BC	1,1	1,1
BD	0,0	0,0

图 110.1 一个扩展博弈(左)和它的战略形式(右)。存在一个剔除战略形式中弱劣行动的剔除次序,它剔除扩展博弈的一个子博弈精炼均衡结果

6.6.2 顺向归纳法 110

我们现在提供两个例子,它们表明反复剔除弱劣行动战略抓住了一些在扩展博弈中参与人推理的有趣特征。

◇例 110.1 (具有外部选择的 *BoS*)(*BoS with an outside option*)
考虑图 111.1 所示的完全信息和同时行动扩展博弈。在该博弈中参与人 1 先决定是呆在家里看书还是去听音乐会。若她决定看书则博弈结束;若她决定去听音乐会,则她与参与人 2 进行 BoS 博弈(例 15.3)。(在历史



	B	S
书	2,2	2,2
B	3,1	0,0
S	0,0	1,3

图 111.1 具有外部选择的 BoS(左边:完全信息和同时行动扩展博弈)和它的简化战略形式(右)

音乐会之后参与人同时选择行动)每个参与人偏好有人陪他去听他喜欢的作曲家的音乐会。

在这个博弈的简化战略形式中,对于参与人 1 来说 S 严格劣于*书*;若它被剔除则对参与人 2 来说 S 对 B 弱劣,最后,对参与人 1 来说*书*对B 是严格劣的。剩下的结果是(B,B)。

这个剔除序列与下列关于扩展博弈的论证相一致。如果参与人 2 不得不做一个决策,则他知道参与人 1 没有选择*书*。对参与人 1 来说只有她计
111 划选 B,这个选择才有意义。因此参与人 2 应该也选 B。这一论证的逻辑在经典文献中被称为"顺向归纳"(forward induction)。

在下一个例子中反复剔除弱劣战略导致了一个更引人注目的结论。

例 111.1(焚钱)(Burning money)考虑在图 112.1 上部的博弈。两个人打算用图左边表格所示的货币支付来进行 BoS 博弈。在这样做之前,参与人 1 可放弃一美元(采取行动 D)或避免这样做(采取行动 0);她的行动被参与人 2 观察到。两个参与人都是风险中立者。(注意:参与人初始行动之后的两个子博弈在战略上都是同一的。)

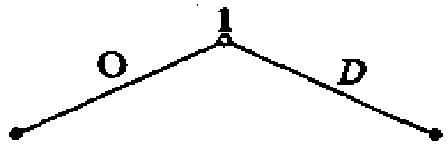
博弈的简化战略形式由图 112.1 的下部所示。弱劣行动如下可被反复剔除:

1. 对于参与人 1 来说 DS 对 OB 弱劣
2. 对于参与人 2 来说 SS 对 SB 弱劣
3. 对于参与人 2 来说 BS 对 BB 弱劣
4. 对于参与人 1 来说 OS 对 DB 强劣
5. 对于参与人 2 来说 SB 对 BB 弱劣
6. 对于参与人 1 来说 DB 对 OB 强劣

剩下的惟一战略二元组为(OB,BB);参与人 1 能放弃一美元的事实包含了

在反复剔除弱劣行动下，结果是参与人 1 所偏好的。

一个符合这个剔除序列的直观论证如下。参与人 1 必须预测到如果她选择 O，那么她将获得一个至少为 3/4 的期望支付，因为对每个关于参与人 2 行动的信念她有一至少使她获得这个期望支付的行动。因此如果参与人 2 观察到参与人 1 选了 D，则他必定期望参与人 1 将随后选择 B（因为选择 S 不可能使参与人 1 获得超过 3/4 的支付）。给定这点，如果参与人 1 选 D 则参与人 2 应该选 B；参与人 1 知道这点，所以若她选 D 则她希望获得 2 的支付。但现在参与人 2 仅能通过相信参与人 1 会选 B 来理性化由参与人 1 所选择的 O（因为 S 只能给参与人 1 不多于 1 的支付），所以在观察到 O 之后，参与人 2 的最优行动是 B。这就使得 O 对参与人 1 是最优行动。 112



	B	S
B	3,1	0,0
S	0,0	1,3

	B	S
B	2,1	−1,0
S	−1,0	0,3

	BB	BS	SB	SS
OB	3,1	3,1	0,0	0,0
OS	0,0	0,0	1,3	1,3
DB	2,1	−1,0	2,1	−1,0
DS	−1,0	0,3	−1,0	0,3

图 112.1 一个完全信息和同时行动的扩展博弈，其中参与人 1 在进行博弈 BoS 前可选择毁掉一美元。扩展博弈在上部分给出，简化战略形式在下部分

为了阐明在第 1.5 节中我们所作的“稳定状态”与“推断”这两个博弈理论方法间的区别，现在我们讨论这两个例子。从稳定状态解释的观点看两个例子有同样的证明：在结果为方 （例 110.1）或（O，（S，S））（例 111.1）的均衡中参与人 2 的信念在下列意义下都是不合理的：若参与人 1 偏离（到音 113

乐或 D),那么参与人 2 能得到的惟一合理的结论是参与人 1 打算随后采取 B,其意味着参与人 2 应该选 B 使得偏离对参与人 1 有利。从推断解释的观点看两个博弈有差别,至少在例 2 中证明更复杂这点上如此。在第 1 例中参与人不得不研究参与人 2 将如何解释她采取的行动(Concert)。在第 2 例中参与人 1 关于参与人 2 对她行动 0 的意图的解释推理,涉及到她的关于参与人 2 将如何理性化她没采用的行动(D)的信念。

第 2 个例子提出了一个这样的问题,即如何具体化一个给定情形下的博弈。我们已进行的证明很明显建立在这样假设的基础上,即图 112.1 的博弈将情形反应为与参与人理解的一样。特别地,它们假设参与人理解处理一美元的可能性是与 BoS 的进行相关的。我们认为这是一个不令人信服的假设:任何有理性的人不会考虑处理一美元的可能性与参加哪场音乐会的选择相关。因此我们认为模化该情形的博弈应该简单排除处理的合理性。[AR 认为即使(包括了参与人 1 可能焚钱的行动在内的)博弈被一裁判员清楚地提供给全部参与人也应该如此,因为在一个参与人战略地分析情形前,他会"编辑"情形的描述并剔除"不相关"因素。]我们将"处理一美元的可能性是不相关的"这一断言建立在什么原则基础上呢?答案是不很清楚的;一些思想如下:(a)处理不影响 BoS 中参与人的支付。(b)如果处理对于参与人 1 的理性有教益的话,那么一个合理的结论应该是毁坏一美元的参与人不过是不理性的。(相反,例如花钱在广告上可能预示着有用的信息。)(c)博弈的两部分间的不相似使得参与人 2 不太可能在第一阶段试图从参与人 1 的行动来推断第二阶段她将如何行动。

对本部分证明的一个解释是,每个参与人通过能解释他将来意图的信息来完成他的行动。因此为了进一步考察证明,通过增加具有这种明显意图的行动来扩张博弈似乎是自然的。不过,如下例所示,如果这样做,我们将面临困难。

114 假定 BoS 将被迫行动且参与人 1 在 BoS 开始前能发出一个信息(任一信号串)给参与人 2。再假定每个参与人只关心 BoS 的结果,而不关心任何被发送信息的内容或他所采取的行动与信息间的相互关系。这个扩展博弈有子博弈精炼均衡结果,其中(B,B)和(S,S)都在 BoS 中被采用;特别地,有一个参与人 2 完全忽略参与人 1 的信息的均衡。之所以这样是因为没有强迫参与人 2 去将由参与人 1 发送的信息解释为有意思的,即使信息是"我要采取 B"。信息可被发送的事实并不影响结果,因为行动的名称在纳什均衡概念中不起任何作用。一个合理的结论似乎是,如果我们希望对参与人间的交流建模,那么对扩展博弈模型的调整是必须的。

?练习 114.1 考察图 112.1 上部博弈的一个变形，在其中参与人 1 先有放弃一美元的选择，然后参与人 2 在观察到参与人 1 的行动后也被允许放弃一美元，最后参与人 1 和 2 进行 BoS。试找出反复剔除弱劣行动后剩下的结果集合并将其与在图 112.1 的博弈中这样做所得的结果相比较。

?练习 114.2 考虑一个博弈，它与图 112.1 上部的博弈不同之处仅仅在于在参与人 1 作了放弃一美元的行动后由参与人进行的博弈如图 114.1 所示。试找出反复剔除弱劣行动后剩下的结果集合。

	A	*B*
A	2,2	0,0
B	0,0	1,1

图 114.1 与练习 114.2 相关的博弈

[注解]

扩展博弈的概念首创于 von Neumann and Morgenstern(1944)；Kuhn(1953)提供了我们描述的系统表达。子博弈精炼均衡的概念应归功于 Selten(1965)。

一次偏离性质(引理 98.2)与动态规划的一个原理紧密相关。命题 99.2 115
应归功于 Kuhn(1953)。同时将行动博弈视为完全信息博弈的思想应归功于 Dubey and Kaneko(1984)。我们关于第 6.4节中战略解释的一些讨论以 Rubinstein(1991)为基础。在第6.5.1节研究的蜈蚣博弈应归功于 Rosenthal (1981)。由这些博弈所引起的一些问题由 Reny(1993)研究。[参看第12.3.2节对连锁店博弈的一具体变形应归功于 Kreps and Wilson(1982a)和 Milgrom and Roberts(1982)]。Moulin(1986)给出了有关反复剔除弱劣行动的过程和子博弈精炼均衡解的结论。图109.1中的博弈(经调整)摘于 Reny (1992)。顺向归纳的思想(与例110.1中的博弈一起)应归功于 Kohlberg；它在 Kohlberg and Mertens(1986)中被讨论。例111.1中的博弈应归功于 van Damme(1989)，也可参看 Ben-Porath and Dekel(1992)和 Osborne(1990)。(关于在此博弈中所引起的问题的更多讨论参看 Rubinstein(1991))

练习 103.2 基于 Kreps 的一个思想；练习 108.1 应归功于 Radner (1980)(也可参看 Radner(1986))。

117 # 第 7 章

讨价还价博弈

人们经常要在一种对最优结果缺乏完全一致看法的情形下集体选择一个结果。这里我们研究一个反应了这种情形并基于完全信息扩展博弈的模型。

7.1 讨价还价和博弈论

博弈论可用来处理人们的利益相冲突的情形。相关的人可能试图通过自愿地采取一个对他们中的所有人都有利的行动进程来解决这种冲突。如果对所有的人来说有不止一个比分歧更值得采纳的行动进程,且对追求哪个行动进程又有冲突,那么某种关于如何解决冲突的协议就是必要的。协议过程可以通过使用博弈论的工具来建模;本章中的模型是这种分析的一个例子。

因为利益冲突的存在是博弈论情形的中心环节,讨价还价理论就不仅仅是博弈论的一个应用;讨价还价模型自它诞生以来就处于该主题的中心且已引起了广泛注意。大部分早期工作都是使用由 John Nash 创导的公理方法,他的工作在第 15 章我们要讨论。本章中我们应用完全信息扩展博弈模型来研究讨价还价的一些特征,特别是参与人的不耐心和风险厌恶对结果的影响。

118 ## 7.2 轮流出价讨价还价模型

考虑这样一种情形,即两个讨价还价的人都有机会就集合 X 中的一个

结果达成协议并且都领会如果他们达不成协议则结果将会是某一固定事件 D。例如，集合 X 可以是一块想要的蛋糕的可行分法集合，D 可以是两人都得不到蛋糕的事件。为了将这一情形模化为一个扩展博弈，我们得确定在谈判中参与人应遵循的过程。

我们研究的过程是一种参与人轮流出价的过程。通过引进值为非负整数的"时间"变量，我们能便利地对它进行描述。博弈的第一个行动发生在时期0，这时参与人1出价（X 的一个元素），参与人2要么接受要么拒绝。接受则终止博弈而拒绝则导致时期1，此时参与人2出价，参与人1要么必须接受或拒绝。再次，接受则终止博弈，拒绝则导致时期2，此时再一次轮到参与人1出价。博弈按这种方法继续：只要没有出价被接受，在每个偶数时期参与人1出价而参与人2必须接受或拒绝；在每个奇数时期参与人2出价而参与人1必须接受或拒绝。对谈判轮数没有限制：博弈有一个无限边界。（关于在将一种情形模化为博弈时在有限与无限边界间选择的讨论参看第8.2节）某个出价被拒绝的事实对后来可能的出价无任何限制。特别地，一个拒绝出价 x 的参与人可能在其后提出一个对他来说比 x 更糟的出价。如果没有任何出价被接受则结果是分歧事件 D。

我们现在给一个关于该情形作为完全信息扩展博弈（见定义89.1）的正式描述。参与人集合是 $N=\{1,2\}$。令可能协议集合 X 为欧氏空间的一个紧的且连通的子集。令 T 为非负整数集合。历史集合 H 是所有为下列类型之一的序列的集合，这里 $t\in T$，对所有 s 有 $x^s\in X$，A 意即"接受"，R 意即"拒绝"。

Ⅰ. ϕ（初始历史），或 $(x^0,R,x^1,R,\cdots,x^t,R)$

Ⅱ. $(x^0,R,x^1,R,\cdots,x^t)$

Ⅲ. $(x^0,R,x^1,R,\cdots,x^t,A)$

Ⅳ. $(x^0,R,x^1,R,\cdots)$

从这个关于历史的描述可知：轮到行动的参与人在一段类型Ⅰ的历史 119
之后选择 X 的一个元素，且在一段类型Ⅱ的历史之后选择 $\{A,R\}$ 中的一个元素。类型Ⅲ和Ⅳ的历史是终点；类型Ⅲ的那些为有限的，而类型Ⅳ的那些为无限的。参与人函数被确定如下：如果 h 属于类型Ⅰ或类型Ⅱ且 t 是奇数，或 h 为空，则 $P(h)=1$；如果 h 属于类型Ⅰ或类型Ⅱ且 t 是偶数，则 $P(h)=2$。

为了完成对博弈的描述我们需要确定参与人在终点历史上的偏好。我们假定每个参与人仅关心协议是否达成以及协议的时间和内容，而不关心在达成协议前出价的路径。精确地说，终点历史集合被分割如下：对每个 x

$\in X$ 和 $t \in T$ 所有满足 $x^t = x$ 的类型Ⅲ的历史集合是分割的一个元素，用 (x, t) 表示；所有类型Ⅳ的历史集合是分割的一个元素，用 D 表示。对每个参与人 i 在历史上的偏好关系来自于这个分割的一些元素组成的集合 $(X \times T) \cup \{D\}$ 上的一个偏好关系 $\succsim_i$（那就是，每个参与人 i 对存在于分割的同一元素中的两段历史是无差别的）。我们假定每个参与人 i 的偏好关系 $\succsim_i$ 满足下列条件：

- 任何协议都比未达成协议好：对所有 $(x, t) \in X \times T$，有 $(x, t) \succsim_i D$。
- 时间是有价值的：对每个时期 $t \in T$ 和每个协议 $x \in X$，有 $(x, t) \succsim_i (x, t+1)$，若 $(x, 0) \succ_i D$ 则有严格偏好。
- 偏好是平稳的：$(x, t) \succsim_i (y, t+1)$ 当且仅当 $(x, 0) \succsim_i (y, 1)$，$(x, t) \succsim_i (y, t)$ 当且仅当 $(x, 0) \succsim_i (y, 0)$。
- 偏好是连续的：若对所有 n 有 $x_n \in X$ 和 $y_n \in X$，$\{x_n\}$ 收敛于 $x \in X$，$\{y_n\}$ 收敛于 $y \in X$，且对所有 n 有 $(x_n, t) \succsim_i (y_n, s)$ 那么 $(x, t) \succsim_i (y, s)$。

这些假设蕴含了对任何一个 $\delta \in (0,1)$ 有一连续函数 $u_i: X \to \mathbb{R}$ 使得在下列意义下偏好关系 $\succsim_i$ 在 $X \times T$ 上由函数 $\delta^t u_i(x)$ 代表：$(x, t) \succsim_i (y, s)$ 当且仅当 $\delta^t u_i(x) \geqslant \delta^s u_i(y)$。[这来自于 Fishburn 和 Rubinstein，(1982，定理 1 和 2)]注意若 $\delta^t u_i(x)$ 代表了 $\succsim_i$，那么对任一 $\epsilon \in (0,1)$ 函数 $\epsilon^t v_i(x)$ 也代表了 $\succsim_i$，这里 v_i 由 $v_i(x) = (u_i(x))^{(\ln\epsilon)/(\ln\delta)}$ 确定。因此如果 $\delta^t u_i(x)$ 和 $\epsilon^t v_i(x)$ 是两个偏好关系的表示且 $\delta > \epsilon$，那么除非 $v_i = u_i$ 否则我们不能得出第一个偏好关系比第二个更"有耐心的"结论。

120 我们称这样确定的完全扩展博弈 $\langle N, H, P, (\succsim_i) \rangle$ 为**轮流出价讨价还价博弈**(bargaining game of alternating offers) $\langle X, (\succsim_i) \rangle$。

这种博弈的前两期在图 120.1 中被表示出来。（注意 x^0 仅是在博弈开始时对参与人 1 有效的出价中的一个，x^1 仅是在他拒绝 x^0 后对参与人 2 有效的多个出价中的一个。）

下列是一个重要的轮流出价讨价还价的例子。

◇例 120.1（分蛋糕）(Split-the-pie) 可能的协议集合 X 是一块想要的蛋糕的所有分法的集合：

$$X = \{(x_1, x_2): x_i \geqslant 0, \quad i = 1, 2 \text{ 且 } x_1 + x_2 = 1\}.$$

每个参与人 i 在 $(X \times T) \cup \{D\}$ 上的偏好关系 $\succsim_i$ 有下列性质：$(x, t) \succsim_i (y, t)$ 当且仅当 $x_i \geqslant y_i$（蛋糕是想要的）；同时 $D \sim_1 ((0, 1), 0)$，且 $D \sim$

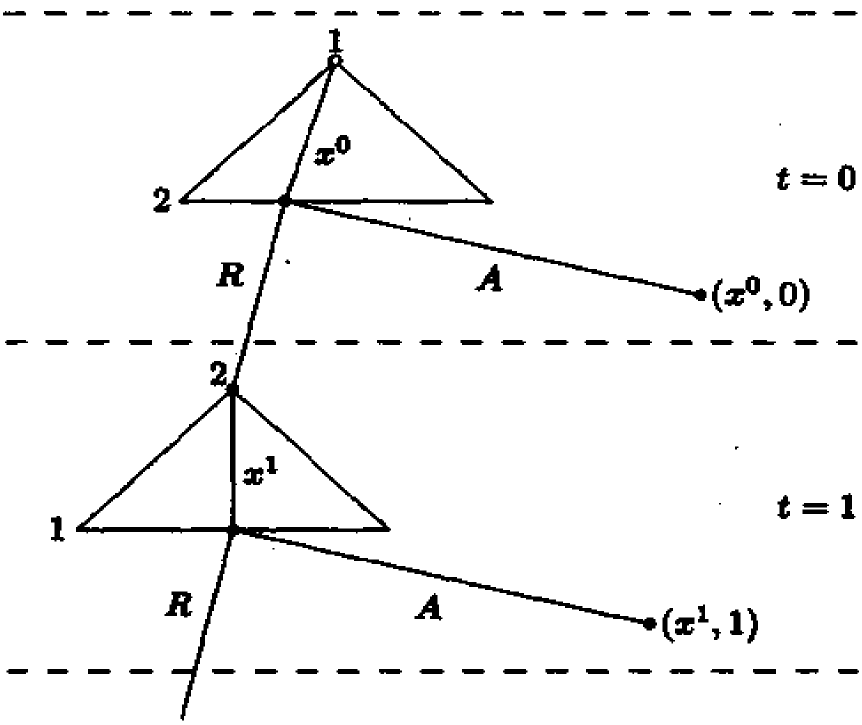


图 120.1 轮流出价讨价还价博弈前两期的表示

$_2((1,0),0)$(在达不成协议的事件中两个参与人一无所获)。因此 $\succsim_i$ 在 $X\times T$ 上可由形如 $\delta_i^t w_i(x_i)$ 的一个函数表示,其中 $0<\delta_i<1$ 且 w_i 是递增的和连续的且 $w_i(0)=0$。

轮流出价讨价还价博弈的纳什均衡集合是很大的。特别地,对任一 $x^*\in X$ 存在一个纳什均衡,在其中参与人立即就 x^* 达成协定(即参与人 1 的均衡战略赋 x^* 给初始历史,参与人 2 的战略赋 A 给历史 x^*)。一个这种均衡是,在其中两个参与人经常接受一个出价 x 当且仅当 $x=x^*$。(换言之,每个参与人 i 能在时期 t 接受一出价 x 当且仅当 $(x,t)\succsim_i(x^*,t)$。)还有,对于参与人偏好的很多具体规定都有协议不能立即达成的纳什均衡。例如,在分蛋糕中对任一协议 x 和时期 t 都有一个纳什均衡,对于它结果是在时期 t 接受 x。一个这种均衡是,在其中直到 $t-1$ 期每个参与人都要求得整块蛋糕且拒绝所有出价,并从时期 t 起出价 x 且仅接受 x。 121

这些纳什均衡阐明了我们在第 6.1.3 节结束时所形成的观点:纳什均衡概念不排除"不可信的威胁"的使用。考虑分蛋糕博弈的纳什均衡,在其中两个参与人经常出价 x^* 且参与人 i 在时期 t 接受一个出价 x 当且仅当 $(x,t)\succsim_i(x^*,t)$。如果 $(x^*,0)\succ_2(x^*,1)$ 则由参与人偏好的连续性有一协议 x,在其中 x_2 略少于 x_2^*,且满足 $(x,0)\succ_1(x^*,0)$ 和 $(x,0)\succ_2(x^*,$

1)。在均衡中参与人 2 的战略决定了在任何时期他拒绝这样一个出价 x：这个“威胁”导致参与人 1 去出价 x^*。给定参与人 1 的战略则参与人 2 的威胁是不可信的：如果参与人 2 实行他的威胁去拒绝 x，则可能发生的最优结果是在下一时期有一个关于 x^* 的协议，这比起参与人 2 在时期 0 通过接受 x 而得到的关于 x 的协议来，是参与人 2 较不喜欢的结果。正如我们在以前章节中所解释的那样，子博弈精炼均衡概念被设计用来分离没有参与人的战略有这种不吸引人的性质的均衡。

7.3 子博弈精炼均衡

7.3.1 特性

我们现在要说明，在某些附加的假设条件下，一个轮流出价讨价还价博弈有一个在本质上是惟一的子博弈精炼均衡，我们将描述它的特征。第一个假设被设计用来避免冗余。

122 A1 设有两个协议 x 和 y 对 $i=1,2$ 使 $(x,0)\sim_i(y,0)$ 成立。

下一个假设简化了分析。

A2 对于 $i=1,2, j\neq i, (b^i,1)\sim_j(b^i,0)\sim_j D$，这里 b^i 是对参与人 i 最优的协议。

为了叙述以下两个假设我们定义协议集合 X 的**帕累托边界**(Pareto frontier)为满足对 $i=1,2$ 没有协议 y 使得 $(y,0)\succ_i(x,0)$ 的协议 x 的集合。我们称帕累托边界的一个元素为一有效的协议。

A3 X 的帕累托边界是严格单调的：如果一个协议 x 是有效的则不存在另一个协议 y 使得对两个参与人 i 都有 $(y,0)\succsim_i(x,0)$。

A4 存在惟一的协议二元组 (x^*,y^*) 满足：$(x^*,1)\sim_1(y^*,0)$，$(y^*,1)\sim_2(x^*,0)$，且 x^* 和 y^* 都是有效的。

这些假设中最重要的是 A4。在分蛋糕博弈中一个它可被满足的充分条件是：每个参与人的偏好关系显示“延迟增加损失”：$x_i-f_i(x)$ 对每个参与人 i 是 x_i 的一个增函数，$f(x)$ 是满足 $(f(x),0)\sim_i(x,1)$ 的协议。假设 A4 可被满足的另一个情形是在其中协议集合 X 的帕累托边界函数是集合 $\{x\in\mathbb{R}^2, x_2=g(x_1)\}$，其中 g 为某个递减凹函数且每个参与人 i 的偏好关系由函数 $\delta_i^t x_i$ $(0<\delta_i<1)$ 表示。

■命题122.1 满足A1至A4的轮流出价讨价还价博弈$\langle X,(\succsim_i)\rangle$有一个子博弈精炼均衡。令$(x^*,y^*)$为惟一有效协议二元组满足

$$(x^*,1)\sim_1(y^*,0)\text{和}(y^*,1)\sim_2(x^*,0). \tag{122.2}$$

在每个子博弈精炼均衡中，参与人经常出价x^*，接受y和任一个满足$(x,0)\succ_1(y^*,0)$的出价x以及拒绝任一个满足$(x,0)\prec_1(y^*,0)$的出价x；参与人2经常出价y^*，接受x^*和任一个满足$(x,0)\succ_2(x^*,0)$的出价x以及拒绝任一个满足$(x,0)\prec_2(x^*,0)$的出价x。

证明：首先我们要证明在命题中定义的战略二元组是一个子博弈精炼均衡。为了这样做我们要应用该博弈有“一次偏离性质”这一事实：战略二元组是一子博弈精炼均衡，当且仅当对每段历史h轮到行动的参与人不可能仅在历史h之后通过改变他的行动来有利地偏离。如引理98.2所示，每个有限边界的博弈具备这一性质的证明留作练习。 123

?练习123.1 试证明每个轮流出价讨价还价博弈都满足一次偏离性质。

我们现在仅需要检查在任一可能的非终点历史之后每个参与人行动的最优性。最优性的状况是一段类型Ⅱ的历史。假设在时期t轮着参与人2去对出价x^t作反应，如果他接受这个出价则结果为(x^t,t)；如果他拒绝则结果是$(y^*,t+1)$。因为他的偏好是平稳的，从(122.2)可知$(x^t,t)\succsim_2(y^*,t+1)$，当且仅当$(x^t,0)\succsim_2(x^*,0)$。因此他的接受规则是最优的。

现在我们转向证明较困难的部分：要证明子博弈精炼均衡在本质上是惟一的。（惟一的不确定性在于每个参与人对这样出价的反应，即他认为不同于别的参与人的均衡出价且别的参与人认为更坏；注意没有这样的出价是有效的。）

给定参与人偏好的平稳性，对$i=1,2$所有以参与人i的出价开始的子博弈都是同一的。令G_i为这样的一个子博弈（G_1是博弈本身）。这样$\delta\in(0,1)$且对$i=1,2$，令$u_i:X\to\mathbb{R}$使得$\delta^t u_i(x)$代表了$X\times T$上的$\succsim_i$。令$M_i(G_i)$为G_i中参与人i的子博弈精炼均衡支付集合的上确界(SPE)：

$$M_i(G_i)=\sup\{\delta^t u_i(x):\text{存在结果为}(x,t)\text{的}G_i\text{的一个SPE}\}。$$

令$m_i(G_i)$为相应的下确界。

步骤1 $M_1(G_1)=m_1(G_1)=u_1(x^*)$和$M_2(G_2)=m_2(G_2)=u_2(y^*)$。[那即是，在$G_1$中的每个SPE中参与人1的支付是$u_1(x^*)$，在$G_2$

的每个 SPE 中参与人 2 的支付是 $u_2(y^*)$]。

证明：通过函数 ϕ 来描述在 X 的帕累托边界上的支付二元组：如果 x 是有效的则 $u_2(x)=\phi(u_1(x))$。由 X 的连通性和偏好关系的连续性则 ϕ 的定义域是一个区间且 ϕ 是连续的；由 A3 它是一对一的和递减的。

我们先证明

$$m_2(G_2)\geqslant\phi(\delta M_1(G_1))。\qquad(123.2)$$

124 如果参与人 1 在 G_2 的第一期拒绝参与人 2 的一个出价，那么她的支付不多于 $\delta M_1(G_1)$。因此在 G_2 的任一 SPE 中她必须接受任一给她多于 $\delta M_1(G_1)$的出价。故在 G_2 的任一 SPE 中参与人 2 的支付不少于 $\phi(\delta M_1(G_1))$。

我们现在证明

$$M_1(G_1)\leqslant\phi^{-1}(\delta m_2(G_2))。\qquad(124.1)$$

在 G_1 的任一 SPE 中参与人 2 的支付不少于 $\delta m_2(G_2)$，因为参与人 2 会经常拒绝参与人 1 开始的出价。因此，参与人 1 在 G_1 的任一 SPE 中能获得的支付不超过 $\phi^{-1}(\delta m_2(G_2))$。

最后我们证明 $M_1(G_1)=u_1(x^*)$。因为存在 G_1 的一个 SPE，在其中即刻协议在 x^* 上达成，所以我们有 $M_1(G_1)\geqslant u_1(x^*)$。现在我们要证明我们关于协议集合和(122.2)解的惟一性的假设包含了 $M_1(G_1)\leqslant u_1(x^*)$。

由 A2 我们有 $\delta u_2(b^1)=u_2(b^1)$，所以 $u_2(b^1)=0$；由 A3 和 ϕ 的定义我们有 $\delta\phi(\delta u_1(b^1))>0=u_2(b^1)=\phi(u_1(b^1))$。因为 ϕ 是递减的，我们能得到结论 $u_1(b^1)>\phi^{-1}(\delta\phi(\delta u_1(b^1)))$。现在，由(123.2)和(124.1)我们有 $M_1(G_1)\leqslant\phi^{-1}(\delta\phi(\delta M_1(G_1)))$。因此由 ϕ 的连续性存在 $U_1\in[M_1(G_1),u_1(b^1)]$满足 $U_1=\phi^{-1}(\delta\phi(\delta U_1))$。如果 $M_1(G_1)>u_1(x^*)$那么 $U_1\neq u_1(x^*)$，所以采用 a^* 和 b^* 作为有效协议(其满足 $u_1(a^*)=U_1$ 和 $u_1(b^*)=\delta u_1(a^*)$)可知(a^*,b^*)是(122.2)的一个不同于(x^*,y^*)的解。这与 A4 矛盾。因此 $M_1(G_1)\leqslant u_1(x^*)$。

同理我们可证明 $m_1(G_1)=u_1(x^*)$，$M_2(G_2)=u_2(y^*)$，和 $m_2(G_2)=u_2(y^*)$，从而完成该步骤的证明。

步骤 2 在 G_1 的每个 SPE 中参与人 2 会立即接受参与人 1 的初始出价是 x^*。

证明：在 G_1 的每个 SPE 中参与人 1 的支付是 $u_1(x^*)$(由步骤 1)且参

与人2的支付至少是$\delta u_2(y^*)=u_2(x^*)$。因为对参与人出价的拒绝导致了子博弈G_2,在其中参与人2的SPE的支付是$u_2(y^*)$。因此由A1和x^*是有效的这一事实,参与人1的开始出价是x^*,它被参与人2接受。

步骤3 在G_1的每一个SPE中参与人2的战略接受任一满足$(x,0)\succ_2(x^*,0)$的出价x且拒绝任一满足$(x,0)\prec_2(x^*,0)$的出价x。

证明:参与人2的一个拒绝导致了G_2,在其中参与人2的支付是$u_2(y^*)$(由步骤1)。因为$u_2(x^*)=\delta u_2(y^*)$,参与人2必须接受任一满足 125
$(x,0)\succ_2(x^*,0)$的出价x且拒绝任一满足$(x,0)\prec_2(x^*,0)$的出价x。(不限制参与人2对出价$x\neq x^*$且满足$(x,0)\sim_2(x^*,0)$的反应。)

最后,将类似于步骤3.4的证明应用于G_2的每个SPE中参与人2的出价和参与人1的接受规则。 □

注意,如果协议集合X的所有元素是有效的(例如在分蛋糕博弈中)那么轮流出价讨价还价博弈有一个惟一的(不仅是在本质上惟一的)子博弈精炼均衡。下面的例子在应用中常被使用。

◇例125.1 考虑每个参与人i的偏好由函数$\delta_i^t x_i$($\delta_i\in(0,1)$)表示的分蛋糕博弈。那么我们有$x^*=(a^*,1-a^*)$和$y^*=(1-b^*,b^*)$。这里a^*和b^*解了二元方程组

$$\begin{cases}1-b^*=\delta_1 a^*\\1-a^*=\delta_2 b^*,\end{cases}$$

所以$a^*=(1-\delta_2)/(1-\delta_1\delta_2)$,$b^*=(1-\delta_1)/(1-\delta_1\delta_2)$。

未被命题122.1所涵盖的一个有趣情形是分蛋糕博弈的一个变形,在其中每个参与人对未达成协议的每个时期招致费用$c_i>0$(且对这些由一个参与人招致的费用总和没有上界)。也就是,如果协议x在时期t达成则参与人i的支付是x_i-c_it。这一情形违背了A2,因为对每个协议x有$(x,0)\succ_i(x,1)$。它同样违背了A4:如果$c_1\neq c_2$,那么没有满足那两个条件的协议二元组,而如果$c_1=c_2$,那么有很多这样的协议二元组。

? 练习125.2

a. 试证明如果$c_1<c_2$那么在上一段中所描述的博弈有惟一子博弈精炼均衡,并且该均衡有同命题122.1中的均衡有同样的结构,此时$x^*=(1,0)$且$y^*=(1-c_1,c_1)$。

b. 试证明如果 $c_1=c_2=c<1$，那么该博弈有很多子博弈精炼均衡结果，若 $c<\frac{1}{3}$，则包括协议被延迟的均衡。

7.3.2 均衡的性质

效率(Efficiency)轮流出价讨价还价博弈的结构允许讨价还价永远继续下去，但在假设 A1 到 A4 下，在所有的子博弈中精炼均衡协议在 X 的帕
126 累托边界协议上能立即达成(所以博弈的结果是有效的)。

战略的平稳性(Stationarity of Strategies)子博弈精炼均衡的战略是平稳的：对任一段在其之后轮到参与人 i 去提出一个协议方案的历史，他都提出同样的协议方案。并且对任一段在其之后轮到他去对一提议作反应的历史，他使用同一准则去选择他的反应。我们没有限制参与人只使用平稳的战略；可是，这些战略作为结论出现。

先动优势(First Mover Advantage)再次考虑例 125.1。如果 $\delta_1=\delta_2=\delta$，那么参与人 1 能得到的蛋糕的数量是 $a^*=1/(1+\delta)>\frac{1}{2}$。博弈中惟一的不对称是参与人 1 先动；她获得大半蛋糕的事实表明首先出价是有优势的。更一般地先动优势有：利用 A3 和 x^* 及 y^* 是有效的事实，在任一满足 A1 到 A4 的轮流出价讨价还价博弈中我们有 $(x^*,0)\succ_1(y^*,0)$。

不耐心的比较静态分析(Comparative Statics of Impatience)参与人偏好的关键特征是它们显示了不耐心。希望一个参与人越不耐心则他在均衡中越不利似乎是合理的。在例 125.1 的博弈中的确是这样，因为 a^* 和 b^* 的值分别在 δ_1 和 δ_2 中是递增的。对任一满足 A1 至 A4 的轮流出价讨价还价博弈我们将这个结论一般化。

定义 $\succsim'_i$ 至少与 $\succsim_i$ 一样不耐心　如果两者导致在 $X\times\{0\}$ 上的同一次序且只要 $(x,1)\sim_i(y,0)$ 就有 $(x,1)\precsim'_i(y,0)$。

■**命题 126.1**　令 $\langle X,(\succsim_i)\rangle$ 和 $\langle X,(\succsim'_i)\rangle$ 为满足 A1 至 A4 的轮流出价讨价还价博弈，假设 $\succsim'_i$ 至少与 $\succsim_1$ 一样不耐心且 $\succsim'_2=\succsim_2$。令 x^* 为在 $\langle X,\langle\succsim_i\rangle\rangle$ 的每一个子博弈精炼均衡中达成的协议，同时令 x' 为在 $\langle X,(\succsim'_i)\rangle$ 的每一个子博弈精炼均衡中达成的协议。那么：$(x^*,0)\succsim_1(x',0)$。

证明：假设结论不成立(所以特别地有 $x^*\neq x'$)。考虑 $X\times X$ 的子

集 S，它包括所有满足 x 和 y 是有效的且 $(y,1)\sim_2(x,0)$ 的二元组 $(x,$
$y)$。令 y' 为满足 $(x',y')\in S$ 的协议。因为 $(x',1)\sim'_1(y',0)$（由
(122.2)）可知 $(x',1)\succsim_1(y',0)$，且由 A4 因而有 $(x',1)\succ_1(y',0)$。
由 A2 我们有 $(b^1,b^1)\in S$，且由 A4 和时间是有价值的这一假设我们有
$(b^1,1)\prec_1(b^1,0)$。因为 X 是紧的和连通的，所以存在在某一路径上的
一个协议 X，该路径在帕累托边界上且连接 x' 和 b^1 使得 $(\hat{x},\hat{y})\in S$ 和 127
$(\hat{x},1)\sim_1(\hat{y},0)$。因为 $(x',0)\succ_1(x^*,0)$ 和 $(b^1,0)\succsim_1(x',0)$ 我们有 $\hat{x}\neq$
x^*，这与 A4 相矛盾。 □

7.4　变形和扩展

7.4.1　过程的重要性

为了将讨价还价模化为扩展博弈，我们需要对参与人所遇到的决策问题的序列结构作一个明确的描述：我们需要确定一个讨价还价过程。在下面练习中，所考虑的轮流出价讨价还价博弈模型的变形表明了讨价还价过程的结构在决定结果的过程中起着重要作用。

? 练习 127.1　假设参与人 1 提出所有出价（而不是轮流出价）。试证明，在 A1 至 A3 下，修改后的博弈有一在本质上是惟一的子博弈精炼均衡，在其中，若不考虑参与人的偏好，则达到的协议是 b^1，它对参与人 1 来说是最优可能协议。

轮流出价讨价还价博弈过程几乎是对称地对待全部参与人。首先行动的参与人比他的对手好这一事实只不过是证明了作为惟一出价者的参与人能享受极端优势。

7.4.2　剔除模型的一个关键特征的变形

轮流出价讨价还价博弈模型的一个关键特征是，一个参与人强迫另一个参与人在现在的一个协议和后来的一个更值得要的协议间作选择的能

力。为了阐明这点，先考虑这样的博弈，在其中每个时期参与人同时出价，在任一给定的时期里仅当出价是相容的协议才达成，参与人都不能强迫别人在现在的一个协议与后来的一个更好的协议间作选择；每个有效协议是一个子博弈精炼均衡结果。

为了进一步阐明这点，考虑协议集合 X 包含了有限个元素的情形，则
128 在其中参与人今天提供的一个略好于反应者明天所期望的协议的能力是有限的。在此情形下子博弈精炼均衡支付的范围依赖于集合 X 的丰富程度。下列练习在一个特别的情形中表明了这点。

?练习 128.1 考虑分蛋糕博弈的一个变形，在其中蛋糕仅分成一个基本可分单元 $\epsilon>0$ 的整数倍，且每个参与人 i 的偏好由函数 $\delta^t x_i$ 表示。用 $\Gamma(\epsilon)$ 来表示该博弈并用 $\Gamma(0)$ 来表示蛋糕可被完全分割的博弈。

a. 试证明如果 δ 足够靠近 1 则对每个协议 $x\in X$ 都存在一个结果为 $(x,0)$ 的 $\Gamma(\epsilon)$ 的子博弈精炼均衡。

b. 试证明如果 δ 足够靠近 1 则对每个结果 $z\in(X\times T)\cup\{D\}$ 存在一个结果为 z 的 $\Gamma(\epsilon)$ 的子博弈精炼均衡（对 $x=(1,0)$ 和 $x=(0,1)$ 利用 a 部分的均衡战略来阻止偏离）。

c. 试用反证法证明：对每个 $\delta\in(0,1)$ 和每个 $\eta>0$ 存在 $\bar{\epsilon}>0$，使得若 $\epsilon<\bar{\epsilon}$，则参与人 $i(i=1,2)$ 在 $\Gamma(\epsilon)$ 的每个子博弈精炼均衡中的支付与在 $\Gamma(0)$ 的惟一子博弈精炼均衡中的支付间的差异少于 η 且协议立即达成。

7.4.3 退出

一类有趣的轮流出讨价还价博弈模型的扩展可通过这样的方式获得，即允许一个或两个参与人在博弈的各时点上“退出”（无须别的参与人同意）而不是继续讨价还价。一种简单的情形是，在其中仅有一个参与人，比如说参与人 2 可退出来，并且当他对一个出价作反应时才能这样做。用 (Out,t) 来表示在时期 t 他这样做的结果，且假定对所有 $t\in T$ 有 $(Out,t)\sim_1 D$。

先假定 $(Out,0)\prec_2(y^*,1)$。这里 y^* 是参与人 2 在标准轮流出价讨价还价博弈中惟一子博弈精炼均衡出价。那么参与人 2 退出的能力对结果没有影响：尽管参与人 2 有一附加“威胁”，但它一无所值，因为他偏好去继续讨价还价且用一期的延迟获得结果 y^*。

现在假定 $(Out,0)\succ_2(y^*,1)$。那么参与人 2 的威胁就不是一无所值

的。此情形中(在 A1 至 A4 下),在任一子博弈精炼均衡中参与人 1 经常提出满足$(\hat{x},0)\sim_2(Out,0)$并会被参与人 2 接受的有效协议 $\hat{x}$,且参与人 2 经常提出满足$(\hat{y},0)\sim_1(\hat{x},1)$并会被参与人 1 接受的有效协议 $\hat{y}$。因此在 129
这情形中参与人 2 运用外部选择的能力使讨价还价的结果对他来说等价于他退出所导致的结果。

?练习 129.1 试证明我们刚刚对分蛋糕博弈所描述的结论,在其中每个参与人 i 在$(X\times T)\cup\{D\}\cup(\{Out\}\times T)$上偏好关系由 u_i 表示,这里对所有 t 有 $u_i(x,t)=\delta^t x_i$, $u_i(D)=0$, $u_1(Out,t)=0$,并且对某一 $b<1$ 和 $\delta\in(0,1)$有 $u_2(Out,t)=\delta^t b$。

该结论有时被称为"外部选择原理",它对于"参与人能运用他们的外部选择的时点"这一假设并不强。例如,如果其中一个参与人可在每个时期结束时退出,而不是恰在他拒绝出价之后,那么博弈有多重子博弈精炼均衡(参看 Shaked(1994)和 Osborne and Rubinstein(1990,第 3.12 节)。

7.4.4 具有破裂风险的一个模型

最后,考虑轮流出价讨价还价博弈模型的一个修正,在其中每个时期结束时都有一个以概率 $\alpha\in(0,1)$结束博弈的机会行动(在第 15.4 节我们将再次考虑这种情形)。假定参与人并不关心协议达成的时间;对每个参与人达成协议的压力不是参与人的不耐心而是谈判将破裂的风险。在模化该情形的(完全信息和机会行动)扩展博弈中有六类历史。其中四类类似于类型Ⅰ至Ⅳ(参看第 7.2 节),在其中 R 的每次发生被(R,C)所取代,这里 C 是讨价还价继续进行(而不是破裂)的机会行动。类型Ⅴ的一段历史形为$(x^0,R,C,x^1,R,C,\cdots,x^t,R)$,在其后轮到机会行动:类型的一段历史是终点且形为$(x^0,R,C,x^1,R,C,\cdots,x^t,R,B)$,这里 B 代表破裂。我们假定所有参与人在所有没有达成协议的终点历史中(即:在所有类型Ⅳ和Ⅵ的历史中)是无差别的。给定机会行动的存在,我们需要确定参与人对终点历史上不确定事件集合的偏好。同前所述,我们假定这些偏好仅依赖于最终达成的协议(不依赖于拒绝的协议的路径)。进一步,我们假定每个参与人 i 的偏好关系由一个效用函数 $u_i:X\cup\{B\}\to\mathbb{R}$ 表示。(因为类型的历史不 130
以正概率发生,不管参与人采用何种战略,参与人的偏好关系都无须非得对 D 排序。)最后,我们假定:对所有 $x\in X$ 有 $u_i(B)=0$, $u_i(x)\geqslant 0$ 且存在惟一的有效协议二元组(x^*,y^*)满足

$$u_1(y^*)=(1-\alpha)u_1(x^*)\text{和 } u_2(x^*)=(1-\alpha)u_2(y^*). \quad (130.1)$$

(假如下列情形中也是这样,参与人正分一块蛋糕,对某一递增凹函数 w_i 有 $u_i(x_1,x_2)=w_i(x_i)$和 $w_i(0)=0$。

⑦练习 130.2 对上段中所描述的轮流出价讨价还价博弈的变形证明类似命题 122.1 的结论。

7.4.5 多于两个的参与人

命题 122.1 并没有扩展到这样一种情形,即如下分蛋糕博弈(例 120.1)的三人变形所示,在其中有多于两个的参与人。可能的协议集合是

$$X=\{(x_1,x_2,x_3): x_i\geqslant 0, \quad i=1,2,3\text{ 且 } x_1+x_2+x_3=1\}$$

且每个参与人 i 的偏好由函数 $u_i(x,t)=\delta^t x_i(0<\delta<1)$表示。讨价还价过程如下。参与人 1 首先出价。由参与人 j 在时期 t 提出的出价 x 最先由参与人 $j+1(\mathrm{mod}3)$考虑,他可接受或拒绝它。如果他接受了,则参与人 $j+2(\mathrm{mod}3)$可接受或拒绝它。如果两个人都接受它,则博弈结束且 x 被实行。否则参与人 $j+1(\mathrm{mod}3)$在 $t+1$ 期提出下一个出价。

令$\frac{1}{2}\leqslant\delta<1$。我们断言:对每个协议 x 存在一个 x 被立即接受的子博弈精炼均衡。这样一种均衡可用四个共同拥有的"状态"x,e^1,e^2 和 e^3 来描述,这里 e^i 是第 i 个单位向量。在状态 y 每个参与人 i 出价 y 且接受出价 z 当且仅当 $z_i\geqslant\delta y_i$。初始状态是 x。状态间的变换仅发生在一个出价被提出后和一个反应发生前。如果在任一状态 y 参与人 i 出价 z,满足 $z_i>y_i$,那么状态变成 e^j,这里 $j\neq i$ 是拥有最低指数且对他 $z_j<\frac{1}{2}$ 的参与人。这样一个参与人 j 的存在以及条件 $\delta\geqslant\frac{1}{2}$ 保证了对他来说拒绝参与人 i 的
131 出价是最优的。促使这个均衡的主要力量是一个参与人会因为拒绝一个偏离出价而受偿:在他拒绝之后,他获得所有的蛋糕。

⑦练习 131.1 试证明如果每个参与人都只限于使用平稳战略(在其中只要他是出价人,则他出同一出价;并且,只要他是第一个或第二个反应人,则他使用同一规则接受出价),那么上面所描述的博弈的惟一子博弈精炼均衡将蛋糕的 $\delta^{k-1}/(1+\delta+\delta^2)$分给参与人 $k,k=1,2,3$。

[注解]

本章中的模型，及命题 122.1 应归于 Rubinstein(1982)。对模型的解释和分析及它的应用参看 Osborne and Rubinstein(1990)。

有效将注意力限于有限边界博弈的两个模型原形在 Ståhl(1972)和 Krelle(1976, pp. 607—632)中可找到。对时间偏好的讨论见 Fishburn and Rubinstein(1982)。命题 122.1 的证明是对 Rubinstein(1982)原始证明的一个改进，且继承了 Shaked and Sutton(1984a)的思想。第 7.4.2 节的材料在 Muthoo(1991)和 van Damme, Selten, and Winter(1990)中被讨论。第 7.4.3 节中一个参与人可退出的模型由 Binmore, Shaked, and Sutton 提出；对于例子可参看 Shaked and Sutton(1984b)和 Binmore(1985)。第 7.4.4 节中的模型在 Binmore, Rubinstein and Wolinsky(1986)中被讨论。第 7.4.5 节中讲座的例子应归功于 Shaked；更详细的可参看 Osborne and Rubinstein (1990, 第 3.13 节)。对于轮流出价讨价还价博弈模型的另一个解释可参看 Rubinstein(1995)。

133 # 第 8 章

重复博弈

重复博弈模型能设计用来考察长期相互作用的关系。它抓住了参与人会考虑自己当前的行动对其它参与人将来行动的影响这一思想，并且试图去解释诸如合作、报复和威胁等现象。

8.1 基本思想

该理论蕴含的基本思想可用两人重复进行囚徒困境博弈（在图 134.1 中复制）的情形来阐述。记得这个博弈有惟一的纳什均衡，在其中每个参与人选 D；进一步来说，对每个参与人来说行动 D 严格优于行动 C，所以结果 (D, D) 的理由是很强的。尽管这样，如果两个参与人“合作”且选 C 则他们都会更好。在重复博弈理论背后的主要思想是：如果每个参与人相信背叛将终止合作从而导致其后的对他来说超过短期所得的损失，那么若博弈被重复进行，则共同想要的结果（在其中 (C, C) 在每个时期发生）是稳定的。

该理论的主要成就是在博弈中分离出了支持共同想要的结果的战略类型。该理论使我们洞察到了个人重复相互作用时的行为结构，这个结构可用“团体规范”来解释。我们描述的结论表明，维持共同想要的结果所需的团体规范涉及到每个参与人“惩罚”任一一个这样的参与人，即他们的行动是不值得要的。当我们要加强“惩罚的威胁应该是可信的”这一深嵌于子博弈精炼均衡概念的要求时，在团体规范要求惩罚者去实行威胁的情形中，团
134 体规范也必须确保惩罚者有动力这样去做。在此情形中，惩罚的本质依赖于参与人如何评估将来的结果。有时惩罚阶段维持一段有限的时间就足够了，在此之后参与人转而追求共同想要的结果；有时团体规范必须给予花费了成本实行惩罚的参与人补偿。

尽管我们将这些关于均衡战略结构的结论视为该理论的主要成就，但是在经典文献中大部分结论却是将注意力集中于能被均衡维持的支付集合，以及给出这个集合所需的几乎包含所有的合理支付组合的条件。这些“无名氏定理”有两面性：一方面，它们表明了如果参与人是短视的，则不能维持的团体想要的结果在参与人有长期目标的情况下可以被维持；另一方面，它们表明重复博弈的均衡结果集合是巨大的，所以均衡概念缺乏预测的能力，“无名氏定理”在本章中是很多形式发展的焦点。不过，我们强调，从我们的观点看来，该理论的主要贡献是发现了支持共同想要的支付组合的有趣的、稳定的团体规范（战略），而不是简单地证明产生这些支付的均衡存在。

8.2 无限次重复博弈与有限次重复博弈

重复博弈模型有两个样式：期界可以是有限的，也可以是无限的。如我们要看到的那样，在两种情形中结论是不一样的，一个不同的极端（这非一般的）情形是在其中成分博弈是囚徒困境。下面我们会看到，在任一这个博弈的有限次重复中，惟一的纳什均衡结果是参与人在每个时期所选(D, D)的结果；另一方面，在无限次重复博弈中子博弈精炼均衡支付组合的集合是很大的。因此在应用重复博弈模型的特定情形中，我们需要确定是有限边 135
界还是无限边界是比较恰当的。

按照我们的观点，一个模型应试图抓住参与人所领悟的现实特征；它不必试图去描述一个旁观者所领悟的现实，尽管很明显在这两种领悟间存在联系。因此具有一个在某种物理意义上是有限（或无限）边界的情形并不必意味着该情形的最优模型有有限（或无限）边界。如果在每个时期后参与人相信博弈还会持续一段时期，则一个具有无限边界的模型是恰当的，而如果参与人明显地觉察到有一个确定的最后时期，则有限边界模型是恰当的。例如，参与人有有限生命的事实，并不意味着人们会经常将他们的战略相互作用模型化为一个有限的重复博弈。如果他们进行博弈如此频繁推移非常缓慢，那么他们可能忽略边界的存在直至它的到来近在眼前，且直到这时他们的战略想法才可能被一个无限边界的博弈较好地抓住。

AR 在一个客观上是有限的情形中，决定我们应该使用有限边界模型还是无限边界模型的一个关键准则是，最后的时期是否明确地进入了参与人的战略考虑。因为这个原因，甚至一些涉及小数量重复的情形作为无限

重复博弈也能被较好地分析。例如，当实验对象被指导去用图 134.1 所示的支付(解释为美元)进行二十次囚徒困境博弈时，我相信他们的推理思路用无限次重复博弈比用有限次重复博弈能被更好地模型化，因为除非非常靠近博弈的结束，他们很可能忽略最后时期的存在。

	C	D
C	3,3	0,4
D	4,0	1,1

图 134.1 囚徒困境

MJO 重复进行囚徒困境博弈有限次的实验主体的行为，是与有限次重复博弈的惟一子博弈精炼均衡不一致的。它可能与无限次重复博弈的某个子博弈精炼均衡相一致的事实是没什么意思的，因为这样一致的结果范围是巨大的。当然无限次重复博弈的子博弈精炼均衡不能洞察到实验人的行
136 动对支付大小和博弈长度的依赖。(证明的一个概览见 Rapoport(1987)。)实验结论明确地指明了在有限次重复的囚徒困境博弈中，子博弈精炼均衡的概念并没有抓住人类行动特征。不过，这个缺陷似乎更与子博弈精炼均衡概念固有的逆向归纳有关，而不是边界的有限性。一个能使我们理解事实的模型可能是有限重复博弈的一个变形；但是作为一个解释工具，此模型本身似乎没什么前途。而且，在成分博弈有多个纳什均衡的背景下，有限次重复博弈均衡较好地符合偶然观察；当边界较远时，人们会合作行动而当边界较近时，人们会随机行动；无限次重复博弈均衡不能使我们洞察这种行动。最后，在人们的贴现因子随时间递减至零的情形中，即使它们永不会变为零(即无固定有限边界被察觉)，均衡结果与有限次重复博弈的结果，与无限次重复博弈的结果相比更有共同之处。

AR 很多现存的经典文献中，一个长期的有限次重复博弈中均衡集合可能不同于相同成分博弈有限次重复的均衡集合这一事实被称为“摄动”。我发现它很有趣：两个模型抓住了生命的一个非常现实的特征，即一个预先确定的有限时期的存在，可能会决定性地影响人们的行动(考虑当政的最后几个月或宗教意图劝服他们的信仰者存在“来世”)。

MJO 首先，对于一个成分博弈的大集合，在相关的有限和无限次重复博弈的结果间没有不连续性(参看第 8.10 节)。其次，在某些不连续性确实存在的情形中，它确实是不吸引人的。如果面对已知固定距离边界的人们

如同边界是无限的那样行动，那么这就应该是有固定有限边界模型的预测；如果不是那样，就应怀疑在别的背景下子博弈精炼均衡概念的合理性了。

8.3　无限次重复博弈：定义

无限次重复博弈模型抓住了一种参与人重复进行一个战略博弈 G 的情形，我们称 G 为成分博弈(constituent game)。从始至终我们将注意力限 137
于每个参与人的行动集合是紧的和每个参与人的偏好关系是连续的博弈。对 G 进行的次数没有限制，在每个场合参与人同时选择他们的行动。当采取行动时，一个参与人知道以前被所有参与人所选择的行动。我们将这一个情形模型化为一个完全信息(和同时行动)扩展博弈如下。

▶定义 137.1　令 $G=\langle N,(A_i),(\succsim_i)\rangle$ 为一战略博弈，令 $A=\times_{i\in N}A_i$。G 的一个**无限次重复博弈**(infinitely repeated game of G)是一完全信息和同时行动扩展博弈 $\langle N,H,P,(\succsim_i^*)\rangle$，其中

- $H=\{\phi\}\cup(\cup_{t=1}^{\infty}A^t)\cup A^{\infty}$(这里 ϕ 是初始历史，A^{∞} 是 G 中无穷行动组合序列 $(a^t)_{t=1}^{\infty}$ 的集合)
- $P(h)=N$，对于每个非终点历史 $h\in H$
- $\succsim_i^*$ 是 A^{∞} 上的偏好关系，它在这样的意义下将偏好关系 $\succsim_i$ 扩展，即它满足下列的弱可分性(Weak separability)条件：如果 $(a^t)\in A^{\infty}$，$a\in A$，$a'\in A$，和 $a\succsim_i a'$ 那么对所有 t 值

$$(a^1,\cdots,a^{t-1},a,a^{t+1},\cdots)\succsim_i^*(a^1,\cdots,a^{t-1},a',a^{t+1},\cdots)。$$

一段历史是终点当且仅当它是无限的。在任一非终点历史之后每个参与人 $i\in N$ 选择 A_i 中的一个行动。因此参与人 i 的一个战略是对 G 中每个有限结果序列赋 A_i 中的一个行动的函数。

我们现在将约束加在参与人的偏好关系及弱可分性上。我们假定从始至终参与人 i 在重复博弈中的偏好关系 $\succsim_i^*$ 是基于代表了他在 G 中的偏好关系 $\succsim_i$ 的一个支付函数 u_i。我们假定是否 $(a^t)\succsim_i^*(b^t)$ 仅依赖于 G 中相应支付序列 $(u_i(a^t))$ 和 $(u_i(b^t))$ 间的关系。

我们考虑偏好的三种形式，它们中的第一个定义如下。

- **贴现**(Discounting)：存在某个数 $\delta\in(0,1)$(贴现因子)使得实数序列

138 (v_i^t)与(w_i^t)至少一样好，当且仅当$\sum_{t=1}^{\infty}\delta^{t-1}(v_i^t-w_i^t)\geqslant 0$。

根据这个准则一个参与人对某个贴现因子$\delta\in(0,1)$用$\sum_{t=1}^{\infty}\delta^{t-1}v_i^t$来评估支付序列$(v_i^t)$。(因为我们已经假定参与人的支付值在一有限集合中，这个和是确定的。)当参与人的偏好采取这种形式时，我们称组合$((1-\delta)\sum_{t=1}^{T}\delta^{t-1}v_i^t)_{i\in N}$为在重复博弈中与成分博弈中的支付组合序列$(v^t)_{t=1}^{\infty}$相联系的**支付组合**(payoff profile)。

具有贴现因素的偏好对时期的处理是不一样的，给定的所得价值会随时间而减小。我们现在确定两个可选择的对称地处理时期的准则。在一个准则中参与人通过它的极限均值$\lim_{T\to\infty}\sum_{t=1}^{T}v_i^t/T$来本质地评价支付序列$(v_i^t)$。不过，这个极限经常不存在(对$t$期的平均支付会随$t$增加而连续地摆动)，我们所讨论的准则定义如下。(用偏好关系这个说法可便利地定义这个准则。)

- **均值极限**(Limit of means)：实数序列(v_i^t)优于序列(w_i^t)当且仅当lim inf 时$\sum_{t=1}^{T}(v_i^t-w_i^t)/T>0$(即当且仅当对所有只是有限的周期$T$，存在$\epsilon>0$使得$\sum_{t=1}^{T}(v_i^t-w_i^t)/T>\epsilon$)。

当参与人的偏好采取这种形式我们称组合$\lim_{T\to\infty}(\sum_{t=1}^{T}v_i^t/T)_{i\in N}$(如果它存在的话)为在重复博弈中与成分博弈中支付组合序列$(v^t)_{t=1}^{\infty}$相联系的**支付组合**(payoff profile)。

注意如果根据均值极限序列(v_i^t)优于序列(w_i^t)则有一个与1足够近的贴现因子，使得(v_i^t)根据贴现准则优于(w_i^t)。

在贴现准则下，单个时期中支付的一个变动可能有重大影响，而在均方差极限准则下任何有限数量时期中支付的差异都没有重大影响。一个偏好满足均方差极限的参与人为了增加他最终获得的支付流而准备在一个有限数量时期内牺牲任何损失。例如，支付流$(0,\cdots,0,2,2,\cdots)$根据均方差极限准则优于常数流$(1,1,\cdots)$，而与在第一个流中参与人首先得到2的时期的序号无关。乍一看该性质可能似乎陌生，不过，考虑决策支付牺牲短期而将绝大部分重点放在长期上是不困难的(考虑民主主义者的斗争)。

我们现在介绍一个准则，它对称地处理所有时期前将重点放在长期上，但同时对单个时期中支付的变化又是敏感的。(再一次我们用严格偏好关系来确定准则。)

139 - **超越**(Overtaking)：序列(v_i^t)优于序列(w_i^t)当且仅当 lim inf $\sum_{t=1}^{T}$

$(v_i^t - w_i^t) > 0$。

下面的例子说明了三个准则间的一些差异。对任一 $\delta \in (0,1)$ 由贴现准则序列$(1, -1, 0, 0, \cdots)$优于$(0, 0, \cdots)$，但按照其它两个准则，这两个序列是无差别的，序列$(-1, 2, 0, 0\cdots)$根据超越准则优于$(0, 0, \cdots)$，但根据均值极限准则两个序列是无差别的，序列$(0, \cdots 0, 1, 1, \cdots)$，在其中 M 个 0 后是一个常数的序列，根据均值极限准则对每个 M 值它优于$(1, 0, 0, \cdots)$，但对每个 δ 存在M^*足够大，使得对所有 $M > M^*$，根据贴现准则对 δ 值后者优于前者。

令 $G = \langle N, (A_i), (\succsim_i) \rangle$ 为一战略博弈且对每个 $i \in N$ 令 u_i 为代表了 $\succsim_i$ 的支付函数。我们定义$\langle N, (A_i), (u_i) \rangle$的 **$\delta$－贴现的无限次重复博弈**($\delta$-discounted infinitely repeated game)为成分博弈是 G 和每个参与人 $i \in N$ 的偏好次序 $\succsim_i^*$ 派生于对每个参与人使用一个贴现因子为 δ 的贴现准则的支付函数 u_i 的无限次重复博弈。同理我们可定义$\langle N, (A_i), (u_i) \rangle$的**均方差极限无限次重复博弈**(limit of means infinitely repeated game)和$\langle N, (A_i), (u_i) \rangle$的**超越无限次重复博弈**(overtaking infinitely repeated game)。

我们用 $u(a)$表示支付$(u_i(a))_{i \in N}$。定义一个向量 $v \in \mathbb{R}^N$为$\langle N, (A_i), (u_i) \rangle$的一个**支付组合**，如果有一结果 $a \in A$ 满足 $v = u(a)$。我们称向量 $v \in \mathbb{R}^N$为$\langle N, (A_i), (u_i) \rangle$的一个**可行支付组合**(feasible payoff profile)；如果它是 A 中结果的支付组合的凸组合：即，如果对具有$\sum_{a \in A} \alpha_a = 1$的非负有理数 α_a 的某一集合$(\alpha_a)_{a \in A}$有 $v = \sum_{a \in A} \alpha_a u(a)$。(在经典文献中系数 α_a 允许为任一实数，而不必为有理的，当加上一点内容就是完成其证明的推广。)注意$\langle N, (A_i), (u_i) \rangle$的一个可行支付组合不必是$\langle N, (A_i), (u_i) \rangle$的一个支付组合。

?练习 139.1　考虑一个无限次重复博弈，在其中参与人的偏好派生于在使用不同的贴现因子的成分博弈中他们的支付。试证明这样一种重复博弈中的一个支付组合可能是成分博弈的一个可行支付组合。

8.4 作为机器的战略 140

正如我们在本部分介绍中所讨论的，重复博弈理论的主要成就是使我

们洞察个人相互作用时的行动结构。本部分我们开发一种语言用来便利地描述我们找到的均衡结构。我们先定义一个意图作为过程抽象的机器，通过该过程参与人实施重复博弈中的战略。$\langle N,(A_i),(\succsim_i)\rangle$在一个无限次重复博弈中的一个**机器**(a machine)(or automation)有下列组成部分。

- 一个集合 Q_i(**状态集合**)。
- 一个元素 $q_i^0 \in Q_i$(**初始状态**)。
- 一个函数 $f_i: Q_i \to A_i$，它对每个状态赋一个行动(**输出函数**(output function))。
- 一个函数 $\tau_i: Q_i \times A \to Q_i$，它对每一个包含一个状态和一个行动组合的二元组赋一个状态(**转移函数**(transition function))。

集合 Q_i 是没有限制的，状态的名称当然没有什么意义(例如，我们称一个状态是合作的这事实，并不意味着与它相联系的行动符合它的名字)。在第1期机器的状态是 q_i^0 并且机器选择行动 $f_i(q_i^0)$。只要机器在某一状态 q_i 中，则它选择与那个状态相应的行动 $f_i(q_i)$。转移函数 τ_i 确定了机器如何从一个状态转移到另一个状态：如果机器在状态 q_i 中并且所选的行动组合是 a，那么它的状态转移到 $\tau_i(q_i,a)$。

注意：转移函数的输入包括当前的状态和所有参与人当前的行动组合。将当前状态和由别的参与人所选择的行动表作为输入是更自然的。这符合将一个“行动准则”战略自然地描述为一个如何在所有与某人的计划相一致的环境中行动的计划。不过，因为博弈论的定义要求一个战略对所有可能的历史确定一个行动，包括这些与参与人自己的战略不相一致的历史，所以我们得将参与人自己的行动作为一个输入归入转移函数。

为了阐明机器的概念，我们现在对在重复囚徒困境博弈(图134.1)中的一个参与人给出四个机器的例子。

◇例 141.1 (关于“冷酷”战略的机器)下列定义的机器 $\langle Q_i, q_i^0, f_i, \tau_i\rangle$ 是完成(“冷酷”)战略最简单的一机器。在冷酷战略中只要两个参与人在已往的每个时期中都选 C 则选 C，否则选 D。

- $Q_i = \{\mathcal{C}, \mathcal{D}\}$
- $q_i^0 = \mathcal{C}$
- $f_i(\mathcal{C}) = C$ 且 $f_i(\mathcal{D}) = D$
- $\tau_i(\mathcal{C},(C,C)) = \mathcal{C}$ 且若 $(\mathcal{X},(Y,Z)) \neq (\mathcal{C},(C,C))$ 则 $\tau_i(\mathcal{X},(Y,Z)) = \mathcal{D}$。

这个机器如图141.1所示。每个方框对应于一个状态，在每个方框中是状态名，其后(符号后)是在那个状态中机器采取的行动。粗边框对应于初始状态。箭头对应于转移；同每个箭头相邻的是导致转移的结果集合。

141

图141.1 一个与囚徒困境博弈中的冷酷战略相对应的机器

◇例141.2 如图141.2所示参与人1的机器M_1，只要参与人2采取C则参与人1采取C；在三个时期它选择D，并且如果当参与人2该选C时而选了D，则它返回C。(我们可以认为在三个时期另一个参与人因为选D而“被惩罚”，然后“被饶恕”。)注意一个机器为了实施这个战略必须有至少四个状态。

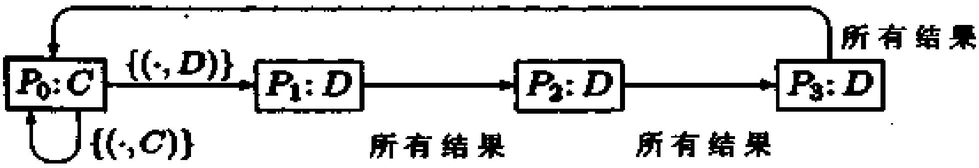


图141.2 机器M_1。在囚徒困境博弈中参与人1的机器，只要参与人2不进行C，则参与人1进行C并且在三个时期通过选择D来惩罚参与人2选择D(我们用$\{(\cdot,X)\}$来表示参与人2的行动是X的所有结果的集合。)

◇例141.3 如图142.1所示的参与人2的机器M_2由进行C开始，且在别的参与人选D条件下继续这样做。如果别的参与人选C，则它转向 142
D，一直到别的参与人再次选D，这时它回到进行C。

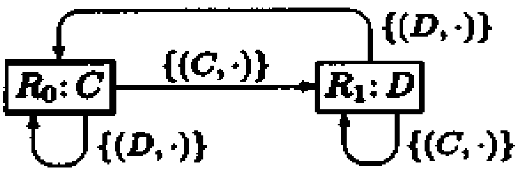


图142.1 机器M_2。在这个囚徒困境博弈中，参与人2的机器由进行C开始，当参与人1选择C，则转向D，仅当参与人选D才返回C

为了阐述当每个参与人的战略被一台机器实行时重复博弈中的每个博

弈进程，假设在囚徒困境博弈中参与人 1 使用机器 M_1 和参与人 2 使用机器 M_2。

机器各自从状态 P_0 和状态 R_0 开始。在第 1 期结果是 (C, C)，因为 M_1 的输出函数赋行动 C 给状态 P_0，且 M_2 的输出函数赋行动 C 给状态 R_0。以后各期的状态由转移函数决定。M_1 的转移函数使机器在状态 P_0 中，因为第 1 期的结果是 (C, C)，而对应于这个输入 M_2 的转移函数使机器从 R_0 移至 R_1。因此在时期 2 状态二元组为 (P_0, R_1)。输出函数决定了在时期 2 结果为 (C, D)，所以现在 M_1 从 P_0 转至 P_1，而 M_2 停留在 R_1。博弈如图 142.2 中的表所示继续至时期 5，在时期 6 状态二元组同在时期 1 一样；相应地状态和结果按前五个时期的格式循环。周期被产生这一事实
143 不是本例所特有的：只要每个参与人用有限多个状态去使用机器，则循环最终会产生，尽管在时期 1 不一定会这样。这从每台机器仅把前时期中的行动作为输入这一事实可很好理解(即它是“马氏的”)。

时期	M_1的状态	M_2的状态	结果	支付
1	P_0	R_0	(C,C)	3,3
2	P_0	R_1	(C,D)	0,4
3	P_1	R_1	(D,D)	1,1
4	P_2	R_0	(D,C)	4,0
5	P_3	R_0	(D,C)	4,0
6	P_0	R_0	(C,C)	3,3

图 142.2　当参与人 1 使用图 141.2 中的机器 M_1 和参与人 2 使用图 142.1 中的机器 M_2 时，在重复囚徒困境博弈中前六个时期的结果

☑练习 143.1　试证明在一个无限次重复博弈中并非每个战略都可由一台有有限个状态的机器来实施。

8.5　触发战略：纳什无名氏定理

现在研究无限次重复博弈纳什均衡结果集合。我们要证明这个集合包括了这样的结果，即它并非是成分博弈纳什均衡的重复。为了支持这一结果，每个参与人必须由“惩罚”来阻止偏离。这种惩罚可能采取多种形式。一种可能性是每个参与人使用一个“触发战略”：任何偏离都导致他实行一

个永远的惩罚性的行动。在本部分我们所研究的均衡中每个参与人都使用这一战略。

令$\langle N,(A_i)(\succsim_i)\rangle$为一战略博弈,对每个参与人$i\in N$令$u_i$为一代表了偏好次序$\succsim_i$的支付函数。记得我们将$G=\langle N,(A_i),(u_i)\rangle$的一个可行支付组合定义为一个凸组合$\sum_{a\in A}\alpha_a u(a)$。其中系数$\alpha_a$是有理数。令$w=\sum_{a\in A}\alpha_a u(a)$为这样的一个组合且假设对每一$a\in A$有$\alpha_a=\beta_a/\gamma$,这里每个$\beta_a$是一整数且$\gamma=\sum_{a\in A}\beta_a$。那么在包含了一个不确定的长度为$\gamma$的循环(其中每个$a\in A$进行$\beta_a$个时期)的重复博弈中,结果序列产生关于循环的平均支付组合,且在w的所有重复博弈中亦如此。

定义在G中参与人i的**最小最大支付**(此后记为v_i)为别的参与人可强加给参与人i的最小支付:

$$v_i=\min_{a_{-i}\in A_{-i}}\max_{a_i\in A_i}u_i(a_{-i},a_i)。\qquad(143.2)$$

满足对所有$i\in N$有$w_i\geqslant v_i$的支付组合w被称为**可实施的**(enforceable);若对所有$i\in N$有$w_i>v_i$则w是**严格可实施的**①。如果$a\in A$是G的一个满足$u(a)$是(严格)可实施的结果,则我们称a为G的一个**(严格)可实施的结果**。用$p_{-i}\in A_{-i}$表示在(143.2)右边最小化问题的一个解。对每个行动组合$a\in A$令$b_i(a_{-i})\in A_i$是参与人i在G中对a_{-i}的一个最优反 144
应行动(即$b_i(a_{-i})\in B_i(a_{-i})$)。(注意$p_{-i}$和函数$b_i$仅依赖于参与人关于$A$的偏好,而不依赖于这些偏好的支付表示。)在很多经典文献中用“个人理性的”一词来代替“可实施的”。行动p_{-i}的集合是别的参与人在G中可给予参与人i的最严厉的“惩罚”。(注意我们限制惩罚为决定性的。在一些经典文献中惩罚者被允许去随机化,可能会使他们的行动与时间有关;这就改变了可行的支付组合集合和可实施的支付,而不是重复博弈均衡集合的结构。)

在随后的两个结论中我们将说明,参与人用均方差极限来评价支付流的无限次重复博弈的纳什均衡,支付组合集合是成分博弈的所有可行的可实施支付组合集合。第三个结论说明当参与人用一个接近1的贴现因子去贴现将来支付时也同样是几乎正确的。

■命题144.1 $G=\langle N,(A_i),(u_i)\rangle$的均方差极限无限次重复博弈的每一个纳什均衡,支付组合是G的一个可实施支付组合。对任一$\delta\in(0,$

① 在很多文献中,经常使用术语“个人理性”,而不是“可实施的”。

1)，G 的用 δ 贴现的无限次重复博弈的每一个纳什均衡支付组合也是 G 的一个可实施支付组合。

证明：假设 w 是 G 的均方差无限次重复博弈的一个支付组合；假设 $w_i < v_i$。那么 w 不是重复博弈的一个纳什均衡支付组合，因为对任一战略组合 s 由 $s'_i(h) = b_i(s_{-i}(h))$（对每段历史 h）所定义的参与人 i 的战略 s'_i 在每个时期给参与人 i 一个至少为 v_i 的支付。同样的证明可应用于任一 G 的用 δ 贴现的无限次重复博弈。

下面的练习请你用我们在第 8.4 节所开发的机器语言来表达在这个证明中参与人 i 的战略。

?练习 144.2　考虑一个两人的无限次重复博弈。对任一个给定的参与人 2 的机器试构造一个给参与人 1 的至少产生支付 v_1 的机器。

■命题 144.3　（均方差极限准则下的纳什无名氏定理）$G = \langle N, (A_i), (u_i)\rangle$ 的每一个可行的可实施支付组合是 G 的均方差极限无限次重复博弈的一个纳什均衡支付组合。

145 证明：令 $w = \sum_{a \in A}(\beta_a/\gamma)u(a)$ 为一个可行的可实施支付组合，这里对每一个 $a \in A$，β_a 是一个整数，$\gamma = \sum_{a \in A}\beta_a$ 且令 (a^t) 为行动组合的循环序列，对于它循环（长度为 γ）包含了 β_a 个 a 的重复（对每个 $a \in A$）。令 s_i 是参与人 i 在重复博弈中这样的战略，即在每个时期 t 都选择 a_i^t，除非存在一个较前的有单个非 i 的参与人偏离 $a^{t'}$ 的时期 t'，在此情形中它选 $(p_{-j})_i$，这里 j 是在第一个这样的时期 t' 中的偏离者。战略组合 s 是重复博弈的一个纳什均衡，因为偏离的一个参与人 j 在每一个随后的时期里获得最多为他的最小最大支付 v_j；由 s 所产生的支付组合是 w。 □

?练习 145.1　试构造一台实行本证明中参与人 i 的均衡战略 s_i 的机器。

在此证明中的战略 s_i 是一个触发战略。很多别的战略可被用来证明这个结论（例如在命题 146.2 的证明中所使用的战略）。

下列是对于具有贴现的无限次重复博弈类似于命题 144.3 的结论。证明类似于上面结论的证明；我们把它留给读者。

■命题 145.2　（贴现准则下的纳什无名氏定理）令 w 为 $G = \langle N,$

$\langle(A_i),(u_i)\rangle$的一个严格可实施可行支付组合。对所有$\epsilon>0$存在$G$的用$\delta$贴现的无限次重复博弈有一个纳什均衡支付组合$\omega'$，满足$|w'-w|<\epsilon$。

为了阐述每个参与人使用一个触发战略的均衡特性，考虑两个无限重复博弈：一个为成分博弈是囚徒困境，我们对它用G_1表示(参看图134.1)，另一个为成分博弈是如图146.1所示的博弈G_2。在G_1和G_2中每个参与人的最小最大支付是1且通过采用D每个参与人使别人的支付维持在这个水平($p_{-1}=p_{-2}=D$)。

	A	D
A	2,3	1,5
D	0,1	0,1

图146.1 博弈G_2

在两个博弈中，命题144.3的证明所使用的触发战略都涉及每个参与人对应任一从均衡路径的偏离而永远转向D。在G_1中行动D优于行动C，所以对每个参与人来说选D是一个稳定次序。因此对某个相信偏离意味着现在稳定次序的终止的惩罚者有某种理由在将来选行动D。相反，在G_2中(D,D)的经常重复不是一个稳定次序，因为对参与人1来说A严格优于D。因此参与人1忍受他给予他对手的惩罚，使得他去惩罚偏离的威胁不可信且怀疑采用这种触发战略的均衡的合理性。我们被引致去研究子博弈精炼均衡的概念，它排除这种战略，因为它要求每个参与人的行动在每段历史后是最优的。 146

?练习146.1 考虑无限次重复博弈，参与人的偏好由贴现准则所代表，共同贴现因子是$\frac{1}{2}$，且成分博弈是图146.1中的博弈。试证明$((A,A),(A,A),\cdots)$不是一条子博弈精炼均衡结果的路径。

8.6 在一段有限长的时间里惩罚：均值极限准则下的精炼无名氏定理

在命题144.3中用于产生任意的可实施支付组合的战略不确定地惩罚

一个偏离者。这种惩罚不一定是急躁的：仅在消除他从偏离（一个时期）的所得的足够长时期里，一个偏离者的支付才需要被降至最小最大水平。如果参与人的偏好满足差方均准则，那么在惩罚后回到均衡路径的战略同均衡路径本身一样能给惩罚者产生同一支付，所以参与人没有理由不去接受它。因此在差方均准则下仅对有限个时期惩罚的团体规范是重复博弈的子博弈精炼均衡。

■命题 146.2 （均值极限准则下精炼无名氏定理）

$G=\langle N,(A_i),(u_i)\rangle$的每一个可行的严格可实施支付组合是 G 的均值极限无限次重复博弈的一个子博弈精炼均衡支付组合。

147 证明：令 $w=\sum_{a\in A}(\beta_a/\gamma)u(a)$为 G的一个可行的严格可实施支付组合，且令$(a^k)_{k=1}^{\gamma}$为包含了 β_a 个 a（对每个 $a\in A$）的重复的行动组合序列。

我们现在构造一个战略组合，它在 G 中产生一个包含循环$(a^k)_{k=1}^{\gamma}$的不确定重复的行动组合序列。每个参与人仅在有限个时期惩罚任一偏离。为便利起见，确定每个参与人的战略使得任一惩罚仅在随循环完成后的时期开始。如果某个参与人在某个没有人应被惩罚的时期偏离，那么这个参与人，比如说是 i，就被认为是该受惩罚的；下一个循环的第一期开始别的参与人通过选择 p_{-i}在一个足够长以至取消掉他的任何可能所得的时期内惩罚 i。随后惩罚者返回均衡，在循环的始点开始。（多个参与人的同时偏离被忽略了。）给定参与人的偏好满足均值极限准则的条件，在每段可能的历史之后支付组合是 w。

为了精确地定义战略令 g^* 为任一参与人在 G 中偏离任一行动组合所能获得的最大数量。也就是，令 g^* 为对所有 $i\in N$, $a'_i\in A_i$ 和 $a\in A$ 的 $u_i(a_{-i},a'_i)-u_i(a)$的最大值。因为 $w_i>v_i$ 所以存在 γ 的一个整数倍的整数 $m^*\geqslant 1$ 使得对所有 $i\in N$ 有 $\gamma g^*+m^*v_i\leqslant m^*w_i$，每个参与人 i 的战略在 m^* 个时期惩罚任一偏离，该战略由下列机器给定。

- 状态集合：$\{(Norm^k,d)$：要么 $k=1$ 且 $d=0$ 要么 $2\leqslant k\leqslant\gamma$ 且 $d\in\{0\}\cup N\}\cup\{P(j,t):j\in N$ 且 $1\leqslant t\leqslant m^*\}$。（状态$(Norm^k,0)$意为我们在循环的第 k 个时期且没有人应受惩罚。状态$(Norm^k,j)$意为我们在循环的第 k 个时期且参与人 j 应受惩罚。状态 $P(j,t)$意为参与人 j 正受惩罚且仍有 t 个应受惩罚的时期。
- 初始状态：$(Norm^1,0)$。
- 输出函数：对任一 $d\in\{0\}\cup N$ 在$(Norm^k,d)$中选 a_i^k；在 $P(j,t)$中，

若 $i \neq j$ 则选 $(p_{-j})_i$，若 $i = j$ 则选 $b_i(p_{-i})$。

- 转移函数：
- 从 $(Norm^k, d)$ 移到① $(Norm^{k+1(\mathrm{mod}\ \gamma)}, d)$ 除非：
  - $d = 0$ 且只有参与人 j 从 a^k 偏离，在此情形中，若 $k \leqslant \gamma - 1$ 则移至 $(Norm^{k+1}, j)$；若 $k = \gamma$ 则移至 $P(j, m^*)$。 148
  - $d = j \neq 0$，在此情形中若 $k \leqslant \gamma - 1$ 则移至 $(Norm^{k+1}, j)$；若 $k = \gamma$ 则移至 $P(j, m^*)$。
- 从 $P(j, t)$ 移至：若 $2 \leqslant t \leqslant m^*$ 则为 $P(j, t-1)$；若 $t = 1$ 则为 $(Norm^1, 0)$。

如此定义的战略是子博弈精炼均衡的证明，我们留给读者。 □

在此证明中我们所给定的战略并不在偏离之后立即开始惩罚，而是要等到一个循环的结束。我们这样确定战略是为了方便计算阻止偏离所需的惩罚长度：如果在偏离后惩罚立即开始，那么我们将不得不考虑当我们计算所需的惩罚长度时，偏离者的支付在循环的剩余部分是很低的这一可能性，所以他从终止循环中可获得额外的支付。

?练习 148.1　(一个既有长期，又有短期参与人的博弈)考虑一个无限边界扩展博弈，在其中战略博弈 G 在参与人 1(长期的)和一个参与人无限序列间进行，序列中的每个参与人仅存在一期且被告知在以往各期中所采取的行动。参与人 1 通过均值极限评价支付序列，别的参与人仅在他存在的那期对他所得的支付感兴趣。

a. 试找出当 G 是囚徒困境博弈(参看图 134.1)时博弈的子博弈精炼均衡集合。

b. 试证明：当 G 是这样的囚徒困境博弈时，在其中给 (C, D) 的参与人 2 的支付为 0，那么对每个有理数 $x \in [1,3]$ 存在一个子博弈精炼均衡，其中参与人 1 的平均支付是 x。

考虑成分博弈由图 146.1 所给定的无限次重复博弈。在此博弈中 $v_1 = v_2 = 1$。考虑命题 146.2 的证明中所定义的战略组合来支持结果序列 (a^t)，其中对所有 t，$a^t = (A, A)$ 采取下列形式：每个参与人在每个时期(循环长度为 1)选 A 除非别的参与人在以前的时期偏离，在此情形中他在 $m^* = 2$ 个时期选 D 然后回到 A。

---

① 我们定义 $m(\mathrm{mod}\ \gamma)$ 为满足 $1 \leqslant q \leqslant \gamma$ 的整数 q，且对某个整数 ℓ 满足 $m = \ell\gamma + q$ (所以，特别地，$\gamma(\mathrm{mod}\ \gamma) = \gamma$)。

当参与人的偏好由超越准则或贴现准则所代表时，这个战略组合不是
149 无限次重复博弈的一个子博弈精炼均衡。在参与人 2 的一个偏离之后，每个参与人被假设为在返回 A 之前在两个时期选 D。参与人 1 选 A 比惩罚参与人 2 更好，因为支付序列 $(1,1,2,2,\cdots)$（在两个准则下都优于序列 $(0,0,2,2,\cdots)$。（在均值极限准则下两个序列是无差别的）为了支持在每个时期结果为 (A,A) 的路径为一子博弈精炼均衡，如果参与人 1 不完成她惩罚参与人 2 的义务，则参与人 2 不得不惩罚参与人 1。进一步说如果参与人 2 因为参与人 1 没惩罚他而不惩罚参与人 1，则参与人 2 必受惩罚。如此等等。在下面两部分我们用具有这些性质的战略去证明当参与人的偏好由超越和贴现准则所表示时的精炼无名氏定理。

8.7 惩罚惩罚者：超越准则下的精炼无名氏定理

下面是一个在超越准则下类似于命题 146.2 的结论；它表明了当参与人的偏好用超越准则表示时，不同于那些用于证明均值极限准则下精炼无名氏定理的战略如何能支持想要的结果。为了简单，我们仅就均衡路径包括惟一（严格可实施的）结果的重复这一情形构造一个战略组合。

■**命题 149.1** （超越准则下精炼无名氏定理）对 $G=\langle N,(A_i),(u_i)\rangle$ 的任一严格可实施结果 a^* 都存在 G 的超越无限次重复博弈的一个子博弈精炼均衡，它产生路径 (a^t)，在其中对所有 t，有 $a^t=a^*$。

证明：令 M 为对所有 $i\in N$ 和 $a\in A$ 的 $u_i(a)$ 的最大值。考虑每个参与人 i 使用下列机器的战略组合。

• 状态集合：$\{Norm\}\cup\{P(j,t):j\in N$ 且 t 是一个正整数$\}$。（在状态 $P(j,t)$ 参与人 j 还要在 t 个时期受惩罚。）

• 初始状态：$Norm$。

• 输出函数：在 $Norm$ 中选择 a_i^*。在 $P(j,t)$ 中若 $i\neq j$ 选 $(p_{-j})_i$，若 $i=j$ 则选 $b_i(p_{-i})$。

• 对应于一个结果 $a\in A$ 的转移。

150 • 从 $Norm$，一直呆在 $Norm$ 除非对某个参与人 j 我们有 $a_{-j}=a_{-j}^*$ 和 $a_j\neq a_j^*$（即，j 是惟一的从 a^* 偏离者），在这情形下移至 $P(j,t)$，这里 t 是

使$M+tv_j<(t+1)u_j(a^*)$的最小的整数。

∘ 从$P(j,t)$：

• 如果$a_{-j}=p_{-j}$或对至少两个参与人l有$a_l\neq(p_{-j})_l$(即所有的参与人都惩罚或至少两人不这样做)那么若$t=1$则移至$Norm$，若$t\geqslant 2$则移至$P(j,t-1)$。

• 如果$a_{-j}\neq p_{-j}$且若$l\neq j^*$有$a_l=(p_{-j})_l$(即j^*是惟一不惩罚的惩罚者)那么移至$P(j^*,T(j,t))$，这里$T(j,t)$足够大使得在状态$P(j,t)$中j^*的支付与他在随后的$T(j,t)$个时期若他不偏离的支付和大于他在偏离中的支付加上$T(j,t)v_j^*$。(这样一个数$T(j,t)$存在因为在t个时期后参与人被认为回到均衡结果a^*且$u_{j^*}(a^*)>v_{j^*}$))。

在此战略组合下，在任一段历史(包括一段在其后惩罚被认为会发生的历史)后，一个参与人通过单边偏离去增加他支付的任何努力都被别的参与人随后的惩罚抵消了，再一次我们留给读者去证明该战略组合是一子博弈精炼均衡。 □

8.8 回报惩罚的参与人：贴现准则下精炼无名氏定理

在命题149.1的证明中所定义的战略组合，在其中参与人因为没有给予他们被分配到的惩罚而被惩罚，可确定当参与人的偏好由贴现准则表示时，该战略可能不会是一个子博弈精炼均衡。理由如下：在此战略组合下，一个没有加入假定为持续(比如说是)t个时期惩罚的参与人，他自己被惩罚(比如说是)t^*个时期，这里t^*可能远远大于t。进一步的偏离可能需要更长的惩罚，从而有战略应被设计用来完成无限长的惩罚。这一结论不管贴现率有多小，可能因此会有导致永远不能被恢复的损失的某个惩罚。结果是，如果参与人的偏好由贴现准则表示，则战略组合可能不会是一子博弈精炼均衡。

为了对参与人的偏好由贴现准则所表示的情形建立一个类似于命题149.1的命题，我们构造一个新战略。在这个战略中惩罚了由战略所确定 151

的偏离者的参与人随后会得到补偿，从而使得他们完成指派是值得的。如在前面部分一样，我们仅就均衡路径包含惟一（严格可实施的）结果重复的情形构造一个战略组合。该结论需要对博弈集合有个常被称为满维(full dimensionality)的限制。

■命题 151.1　（贴现准则下精炼无名氏定理）令 a^* 为 $G=\langle N,(A_i),(u_i)\rangle$ 的一个严格可实施结果。假设有 G 的一个严格可实施结果族 $(a(i))_{i\in N}$ 使得对每个参与人 $i\in N$ 我们有 $a^*\succ_i a(i)$，且对所有 $j\in N\setminus\{i\}$ 有 $a(j)\succ_i a(i)$。那么存在 $\underline{\delta}<1$ 使得对所有 $\delta>\underline{\delta}$ 有 G 的用 δ 贴现的无限次重复博弈的一个子博弈精炼均衡，它产生路径 (a^t)，在其中对所有 t 有 $a^t=a^*$。

证明：每个参与人使用下列机器的战略组合是在每个时期支持结果 a^* 的一个子博弈精炼均衡。该机器有三种类型。在状态 $C(0)$ 由参与人所选择的行动组合是 a^*。对每个 $j\in N$ 状态 $C(j)$ 是一个在参与人 j 的任一惩罚完成后被进入的“修改”状态；在此状态中被选择的行动组合是 $a(j)$。对每个参与人 j 和在 1 与某个后来我们将确定的数 L 间的时间 t，状态 $P(j,t)$ 是一个参与人 j 被认为仍有 t 期被惩罚的状态；在此状态中每个非 j 的参与人 i 采取行动 $(p_{-j})_i$，它使得 j 降到他的最小最大支付。如果在任一状态中任一参与人 i 偏离，则有一至状态 $P(i,L)$ 的转移（即，别的参与人计划在 L 个时期惩罚参与人 i）。如果在惩罚的 L 个时期中没有任何时期存在单个参与人的偏离，则状态变为 $C(i)$。状态集合 $\{C(i)\}$ 作为一个惩罚在惩罚时期不执行规定的参与人的系统：如果参与人 i 不像所认为的那样去惩罚参与人 j，那么状态就变为 $C(j)$ 了，基本结果是 $a(j)$，参与人 i 在 L 个时期被惩罚，在此之后状态变为 $C(i)$，在其中结果是 $a(i)\prec_i a(j)$

简言之，参与人 i 的机器定义如下：这里为了方便我们写成 $a(0)=a^*$；后面我们再确定 L。

152 • 状态集合：$\{C(j):j\in\{0\}\cup N\}\cup\{P(j,t):j\in N$ 且 $1\leqslant t\leqslant L\}$。

• 初始状态：$C(0)$。

• 输出函数：在 $C(j)$ 中选 $(a(j))_i$，在 $P(j,t)$ 中若 $i\neq j$ 选 $(p_{-j})_i$，若 $i=j$ 则选 $b_i(p_{-i})$。

• 对应于一个结果 $a\in A$ 的转移是：

◦ 从 $C(j)$：一直呆在 $C(j)$ 除非单个参与人 k 从 $a(j)$ 偏离，在此情形中移到 $P(k,L)$。

◦ 从 $P(j,t)$：

• 如果单个参与人 $k\neq j$ 从 p_{-j} 偏离，那么移至 $P(k,L)$；

• 否则若 $t \geq 2$ 移至 $P(j, t-1)$或若 $t=1$ 移至 $C(j)$。

我们现在确定$\underline{\delta}$ 和 L 的值。同前，令 M 为对所有 $i \in N$ 和 $a \in A$ 的 $u_i(a)$的最大值，我们选 L 和$\underline{\delta}$ 足够大使得所有可能的偏离被阻止。为了阻止任一参与人在任一状态 $C(j)$中的偏离，我们取 L 足够大使得对所有 $i \in N$，和所有 $j \in \{0\} \cup N$ 有 $M - u_i(a(j)) < L(u_i(a(j)) - v_i)$并且选 $\delta > \delta^*$，这里 δ^* 非常靠近 1，使得对所有 $\delta > \delta^*$ 我们有

$$M - u_i(a(j)) < \sum_{k=2}^{L+1} \delta^{k-1}(u_i(a(j)) - v_i).$$

[这个条件是充分的，因为对 $j \neq i$，有 $u_i(a(j)) > u_i(a(i))$。]对 $j \neq i$ 如果一个参与人 i 偏离$P(j, t)$那么他在由 L 个 $v_i < u_i(a(i))$的时期相随的偏离时期内最多获得 M 且随后获得 $u_i(a(i))$。如果他不偏离则对于 1 和 L 间的时期他获得 $u_i(p_{-j}, b_j(p_{-j}))$且随后获得 $u_i(a(j))$。因此为了阻止偏离下列做法是充分的：选$\underline{\delta} > \delta^*$ 足够靠近 1 使得对所有 $\delta > \underline{\delta}$ 我们有

$$\sum_{k=1}^{L} \delta^{k-1}(M - u_i(p_{-j}, b_j(p_{-j}))) < \sum_{k=L+1}^{\infty} \delta^{k-1}(u_i(a(j)) - u_i(a(i)))。$$

(这样一个$\underline{\delta}$ 值存在因为我们假设了若 $i \neq j$ 则 $u_i(a(j)) > u_i(a(i))$。) □

☒练习 152.1　考虑三人对称无限次重复博弈，在其中每个参与人的偏好由贴现准则表示且成分博弈是$\langle \{1,2,3\}, (A_i), (u_i) \rangle$，这里对 $i = 1, 2, 3$ 我们有 $A_i = [0, 1]$且对所有$(a_1, a_2, a_3) \in A_1 \times A_2 \times A_3$ 有 $u_i(a_1, a_2, a_3) = a_1a_2a_3 + (1-a_1)(1-a_2)(1-a_3)$。

a. 试找出成分博弈的可实施支付集合。 153

b. 试证明对任一贴现因子 $\delta \in (0, 1)$在重复博弈的任一子博弈精炼均衡中任一参与人的支付至少为$\frac{1}{4}$。

c. 用命题 151.1 来修改这些结论。

8.9 贴现准则下子博弈精炼均衡结构

在命题 151.1 的证明中所构造的每个参与人在均衡中的战略(它仅关心贴现因子靠近 1 的博弈)，具有当任一参与人偏离时随后的行动组合序列，仅依赖于偏离者的认定而不依赖于偏离前的历史这一特殊性质。在这

部分我们要证明对任一共同贴现因子支持任一子博弈精炼均衡结果的这种战略的一个组合都能被找到。

我们以两个引理开始,其中第一个将引理98.2中对有限扩展博弈证明了的一次偏离性质扩展到贴现的无限次重复博弈。

■引理153.1　一个战略组合是G的用δ贴现的无限次重复博弈的一个子博弈精炼均衡,当且仅当没有参与人在任一段历史之后的单个时期内通过偏离能有所得。

?练习153.2　证明这个结论。

下一个结论证明了在我们的假设下任一用δ贴现的无限次重复博弈的子博弈精炼均衡,支付组合集合是闭的。

■引理153.3　令$(w^k)_{k=1}^{\infty}$为一个G的用δ贴现的无限次重复博弈的子博弈精炼均衡支付组合的序列,它收敛到w^*。那么w^*是这个重复博弈的一个子博弈精炼均衡支付组合。

证明:对k的每个值令s^k为重复博弈的产生支付组合w^k的一个子博弈精炼均衡。我们构造一个战略组合s,我们将证明它为一个子博弈精炼均衡且产生支付组合w^*。我们通过对历史h长度的归纳,来定义G的一个行动组合$s(h)$和一个附加的序列(s^k)的无限子序列(r^k),它的性质是在历史h之后的子博弈中,由子序列元素所产生的支付组合有一极限且行动
154 组合$r^k(h)$收敛到$s(h)$。假设我们对所有长度为T或更小的历史已这样做了,并考虑长度为$T+1$的一段历史(h,a),这里h是长度为T的一段历史。令(r^k)为我们对历史h所选择的战略组合序列,且令$s(h)$为我们对那段历史所选择的行动组合。对$a=s(h)$为(h,a)选择(r^k)的一个子序列(r'^k),使得$(r'^k(h,a))$收敛且令$r'^k(h,a)$收敛到的行动组合是$s(h,a)$。显然我们已选择的子序列的收敛支付组合同(r^k)的是一样的。当$a\neq s(h)$时,为(h,a)选择(r'^k)的一个子序列(r''^k)使得支付组合序列和序列$(r''^k(h,a))$都收敛,且令$r''^k(h,a)$收敛到的行动组合是$s(h,a)$。

没有参与人i通过在历史h后改变他的行动且导致一个代替$s(h)$的某个结果a,从s_i的偏离中有所得,因为如果是这样的话,那么对足够大的k,他能从r_i^k偏离中获利,这里(r^k)是我们为(h,a)所选择的序列。进一步说,s的支付组合是w^*。　□

由此推论任一参与人i在重复博弈中的子博弈精炼均衡支付集合是闭

的;因为它有界所以有一最小值,我们用 $m(i)$表示它。令$(a(i)^t)$为参与人 i 的支付是$m(i)$的子博弈精炼均衡的结果。

■命题 154.1 令(a^t)为 $G=\langle N,(A_i),(u_i)\rangle$的用 δ 贴现的无限次重复博弈的一个子博弈精炼均衡结果。那么每个参与人 i 使用下列机器的战略组合是具有相同结果(a^t)的一个子博弈精炼均衡。

- 状态集合:$\{Norm^t: t$ 是一正整数$\}\cup\{P(j,t): j\in N$ 且 t 是一正整数$\}$。
- 初始状态:$Norm^1$。
- 输出函数:在状态 $Norm^t$ 进行a_i^t,在状态 $P(j,t)$进行 $a(j)_i^t$。
- 转移函数:
    - 在状态 $Norm^t$ 移至$Norm^{t+1}$除非恰好有一参与人,比如说 j,从a^t 偏离,在此情形中移至 $P(j,1)$。
    - 在状态 $P(j,t)$:移至 $P(j,t+1)$除非恰好有一参与人,比如说j',从 $a(j)^t$ 偏离,在此情形中移至 $P(j',1)$。

证明:利用引理 153.1,可直接证明这确定了一个符合所要求性质的子博弈精炼均衡。 □

8.10 有限次重复博弈 155

8.10.1 定义

我们现在研究有限次重复博弈。有限次重复博弈的正式表达非常类似于无限次重复博弈:对任一正整数 T,战略博弈$\langle N,(A_i),(\succsim_i)\rangle$的一个 T 期有限次重复博弈是满足定义 137.1 中的条件的一个完全信息扩展博弈,此时符号∞由 T 代替。我们将注意力限于在有限次重复博弈中每个参与人 i 的偏好关系$\succsim_i^*$ 由函数$\sum_{t=1}^{T}u_i(a^t)/T$ 代表的情形,这里 u_i 是代表参与人i 在成分博弈中的偏好的一个支付函数。我们称此博弈为$\langle N,(A_i),(u_i)\rangle$的 **$T$ 期重复博弈**(T-period repeated game)。

8.10.2 纳什均衡

对无限次重复博弈无名氏定理的直观证明是,一个共同想要的结果可

通过一种稳定的社会安排支持,在这种安排中,一个参与人由于他若偏离则要被惩罚这一威胁,使得他被阻止偏离。通过修改同样的证明也可应用于一大类有限次重复博弈。需要修改的根源在于这样的事实,即在任一有限次重复博弈的任一纳什均衡的最后一期的结果,必须是成分博弈的一个纳什均衡,这是个对博弈的其余部分起预示作用的事实。对于每个参与人的支付在成分博弈的每个纳什均衡中都等同于他的最小最大支付(如同在囚徒困境中一样)这一特殊情形,这个预示最长。在下列情形中无名氏定理之后的直观证明无效:在每个时期的结果必须是成分博弈的一个纳什均衡,因为如果有某个时期结果不是这样的一均衡,那么在最后的这种时期中某个参与人能不受惩罚地偏离,下面的结论系统表达了这个证明。

■**命题 155.1**　如果在战略博弈 G 的每个纳什均衡中支付组合是 G 中最小最大支付组合(v_i),那么对任一 T 值,G 的 T 期重复博弈的每个纳什均衡的结果($a^1, \cdots, a^T$)具有性质:对所有 $t=1, \cdots, T$, a^t 是 G 的一个纳什均衡。

证明:令 $G=\langle N,(A_i),(u_i)\rangle$ 且令 $a=(a^1, \cdots, a^T)$ 为 G 的 T 期重复博弈的一个纳什均衡 s 的结果。假定对每个时期 a^t 不是 G 的一纳什均衡。
156 令 $t\geqslant 1$ 为 a^t 不是 G 的一纳什均衡的最后时期;假定 $u_i(a^t_{-i}, a_i)>u_i(a^t)$。考虑参与人 i 的在每段最长为 $t-2$ 的历史后符合 s_i 规程的战略 $\hat{s}_i$,在时期 t 选 a_i,在每个随后的时期给定由别的参与人这样的行动选择一个产生他的最小最大支付的行动。($s_{-i}, \hat{s}_i$)的结果是一段终点历史 $\hat{a}$,直到 $t-1$ 期它与 a 是一样的;参与人 i 对 $\hat{a}^t$ 的偏好 L 优于 a^t 且对 $s\geqslant t+1$ 在 $\hat{a}^s$ 和 a^s 间是无差别的。所以参与人 i 的偏好 $\hat{a}$ 优于偏好 a,违背了我们关于 s 是重复博弈的一个纳什均衡的假设。□

该结论只适用于一个很小的博弈集合。如果同结论的假设相反,成分博弈有一纳什均衡,其中某个参与人的支付超过他的最小最大支付,那么只要在最后一期的结果是 a^*,那个参与人在倒数第二个时期会因为偏离而被惩罚。参与人的最小最大支付与他在 a^* 中的支付间差别较小,则这个惩罚就可能不够大到阻止偏离。不过,常有某整数 L,使得如果在最后 L 个时期结果为 a^*,那么参与人在任一这种 L 段博弈序列开始前的时期进行的任何偏离都会由于在剩余的时期里强加给参与人他的最小最大支付这一威胁而被阻止。进一步说,L 值不依赖于博弈的长度,所以,如果对每个参与人成分博弈有一参与人的支付超过他们最小最大支付的均衡,那么对

足够大的 T，作为在 T 期重复博弈的一个纳什均衡中的平均支付组合，任一可行的严格可实施支付组合都能被近似获得。为了简便，我们仅就成分博弈有一个惟一的且其中的每个参与人的支付都超过他的最小最大支付的纳什均衡这一情形来叙述并证明这一结论。

■命题 156.1 （有限次重复博弈纳什无名氏定理）

如果 $G=\langle N,(A_i),(u_i)\rangle$ 有一纳什均衡 $\hat{a}$，其中每个参与人 i 的支付都超过了他们最小最大支付 v_i，那么对 G 的任一严格可实施结果 a^* 和任一 $\epsilon>0$ 存在一个整数 T^*，使得如果 $T>T^*$，则 G 的 T 期重复博弈有一纳什均衡，在其中每个参与人 i 的支付在 $u_i(a^*)\pm\epsilon$ 之间。

证明：考虑由下列机器所完成的参与人 i 的战略。状态集合包括对 $t=1,\cdots,T-L$ 的 $Norm^t$（L 在后面确定），$Nash$ 及对每个 $j\in N$ 的 $P(j)$，每个 157
参与人 i 在 $Norm^t$ 中对所有 t 值选 a_i^*，在 $Nash$ 中选 $\hat{a}_i$，且在 $P(j)$ 中通过选择 $(p_{-j})_i$ 惩罚参与人 j。如果有惟一参与人 j 在状态 $Norm^t$ 中偏离，那么就有一个至 $P(j)$ 的转移；否则若 $t<T-L$ 则有一至 $Norm^{t+1}$ 的转移且若 $t=T-L$ 则有一至 $Nash$ 的转移。一旦到达，状态 $P(j)$ 和 $Nash$ 就不会被离开。结果是在前 $T-L$ 期 a^* 被选择，在最后 L 期 $\hat{a}$ 被选择，简言之，参与人 i 的机器如下。

- 状态集合：$\{Norm^t:1\leqslant t\leqslant T-L\}\cup\{P(j):j\in N\}\cup\{Nash\}$。
- 初始状态：$Norm^1$。
- 输出函数：在 $Norm^t$ 选 a_i^*，在 $P(j)$ 选 $(p_{-j})_i$，且在 $Nash$ 中选 $\hat{a}_i$。
- 转移函数：

 ◦从 $Norm^t$ 移至 $Norm^{t+1}$ 除非或者 $t=T-L$，在此情形中移至 $Nash$ 或者恰好有一参与人，比如说是 j，从 a^* 偏离，在此情形中移至 $P(j)$。

 ◦对任一 $j\in N$ 的 $P(j)$ 和 $Nash$ 都是吸收的。

L 还有待确定。一个有利的偏离仅在状态 $Norm^t$ 中是可能的。为阻止这种偏离，我们要求 L 足够大使得对所有 $i\in N$ 有 $\max_{a_i\in A_i}u_i(a_{-i}^*,a_i)-u_i(a^*)\leqslant L(u_i(\hat{a})-v_i)$。最后，为了获得 $u(a^*)\pm\epsilon$ 间的一个支付组合，我们选择 T^*，使得对所有 $i\in N$ 有 $|[(T^*-L)u_i(a^*)+Lu_i(\hat{a})]/T^*-u_i(a^*)|<\epsilon$。 □

☑练习 157.1 将此结论扩展到在 G 的纳什均衡中参与人 i 的支付超过 v_i 可能依赖于 i 的这种情形。

8.10.3 子博弈精炼均衡

在有限次重复博弈的任一子博弈精炼均衡中，在任一段历史后（而不仅仅是在若每个参与人拥护他的战略才发生的历史后）最后时期中的结果是成分博弈的一个纳什均衡。所以用于证明纳什无名氏定理（命题156.1）战略中的惩罚，是与子博弈精炼均衡不一致的；确实，如果成分博弈有惟一的纳什均衡支付组合，那么不可能有惩罚。随后我们有下列结论。

■命题157.2 如果战略博弈 G 有惟一的纳什均衡支付组合，那么对 T 的任一值，在 G 的 T 期重复博弈的任一子博弈精炼均衡中，任一历史之
158 后所选择的行动组合是 G 的一纳什均衡。

证明：在从重复博弈的任一子博弈精炼均衡的时期 T 开始的任一子博弈的结果都是 G 的一纳什均衡。因此在博弈的最后时期每个参与人的支付与历史无关。相应地，在任一从 $T-1$ 时期开始的子博弈中，行动组合是 G 的一纳什均衡。用归纳法可完成证明。 □

	C	D	E
C	3,3	0,4	0,0
D	4,0	1,1	0,0
E	0,0	0,0	$\frac{1}{2},\frac{1}{2}$

图 158.1 一个修改了的囚徒困境博弈

如果成分博弈有不止一个纳什均衡支付组合，那么惩罚可嵌在一个子博弈精炼均衡战略组合中：参与人在博弈的最后时期的支付会依赖于他们在以前各期的行动。下例说明了该情形中引出的均衡。我们认为在图158.1战略博弈的 T 期重复博弈中有一个子博弈精炼均衡。对于它在最后三期结果是 (C,C)，在其余各期结果是 (D,D)，所以若 T 较大则平均支付组合靠近 $(3,3)$。在均衡中每个参与人使用下列战略：在每个直到 $T-3$ 期的时期选 C 除非有一个参与人在以前的某个时期选 D，此情形中在每个随后的时期选 E，而不管随后的结果，如果在直到 $T-3$ 期的结果是 (C,C)，那么在最后三期选 D。一段在以前各期结果是 (C,C) 的历史之后，在

任一未到 $T-3$ 期的时期偏离到 D 的参与人可在那个时期获得一单位支付，但是随后会失去至少 1.5 个单位，因为别的参与人在随后的每个时期选 E。也就是，随后进行 E 的威胁足够阻止任一偏离；这个惩罚是可置信的。因为 (E,E) 是成分博弈的一个纳什均衡。（注意到如果在成分博弈中支付组合 $(1/2,1/2)$ 由 $(0,0)$ 代替。则同样的战略组合也是一子博弈精炼均衡。 159
在此情形中成分博弈不同于囚徒困境之处仅在于每个参与人另有一个弱劣行动。）

这个例子表明了如果有两个成分博弈 G 的纳什均衡，其中一个优于另一个。那么任一支付组合，在其中每个参与人获得多于在较劣的 G 的纳什均衡中的支付，都可作为在 G 的 T 期（对足够大的 T）重复博弈的一个子博弈精炼均衡中的平均支付组合得到。实际上可建立一个更强的结论：任一严格可实施支付组合都可作为在重复博弈的一个子博弈精炼均衡中的平均支付得到。这一支付组合由这样的战略组合支持，即直到博弈的最后时期都类似于在证明命题 151.1 中所构造的战略组合。

证明（类似于在命题 151.1 和 156.1 证明中的思想）如下，令 a^* 为 G 的一个严格可实施结果。在产生一个当 T 较大时平均支付组合接近于 $u(a^*)$ 的结果序列的 T 期重复博弈中，一个战略组合有下列形式。共有三个阶段。在前两个阶段从始至终每个参与人 i 只要没有参与人偏离就选 a_i^*。在第三阶段参与人在没有偏离的情况下附着成分博弈纳什均衡的一个序列，对于它每个参与人的平均支付超过他在成分博弈中的最低纳什均衡支付。偏离按下列形式被惩罚。在第一阶段发生的偏离由别的参与人采用一个使偏离者降到他的最小最大支付的足够长以至消除他的所得的行动来惩罚。在这个惩罚完成后，一个“修改”的状态被进入足够长来回报为完成他们的指派而加入惩罚的参与人（参见在命题 151.1 证明中的战略）。由某个参与人 i 发生在第二期的偏离被忽略直到第三阶段的开始，在此期间对于参与人 i 来说最坏的纳什均衡在每个时期被执行。在最后一个阶段的偏离不需要受惩罚，因为在每个时期结果是成分博弈的一个纳什均衡。第二个阶段的长度被选为足够长使得对一个在第一阶段最后一个时期偏离的参与人来说，惩罚和随后的修改都可在第二阶段中完成。给定第二阶段的长度，第三阶段的长度被选为足够大使得在第二阶段的第一期偏离的参与人在给定他的从第三阶段的第一期开始的惩罚情况下境况更差。第二和第 160
三阶段长度的下边界不依赖于 T，所以对足够大的 T 由战略组合所导致的平均支付组合靠近 $u(a^*)$。

在下列结论的叙述中，我们将注意力限于包含了成分博弈惟一结果的

重复的均衡路径(就像我们在上面的讨论中那样做)。我们忽略了证明,它可在 Krishna(1989)中找到。

■命题 160.1 (有限次重复博弈的精炼无名氏定理)令 a^* 为 $G = \langle N,(A_i),(u_i)\rangle$ 的一个严格可实施结果。假定(i)对于每个 $i \in N$ 有 G 的两个纳什均衡,它们的不同之处在于参与人 i 的支付;(ii)存在 G 的一个严格可实施结果 $(a(i))_{i\in N}$ 使得对每个参与人 $i \in N$,我们有 $a^* \succ_i a(i)$ 及对所有 $j \in N \setminus \{i\}$ 有 $a(j) \succ_i a(i)$。那么对于任一 $\epsilon > 0$ 存在一个整数 T^*,使得若 $T > T^*$ 则 G 的 T 期重复博弈有一子博弈精炼均衡,在其中每个参与人 i 的支付在 $u_i(a^*) \pm \epsilon$ 之间。

[注解]

较早关于重复博弈的概念和纳什无名氏定理之后的思想的讨论出现在 Luce 和 Raiffa(1957, pp. 97 – 105(特别是 p.102)及 Appendix 8), Shubik (1959b, Ch. 10(特别是 p.226))和 Friedman(1971)。均值极限准则下的精炼无名氏定理由 Aumann 和 Shapley 及 Rubinstein 在 70 年代中期建立;参看 Aumann 和 Shapley (1994)和 Rubinstein (1994)。超越准则下的精炼无名氏定理(命题 149.1)应归于 Rubinstein(1979)。贴现准则下的精炼无名氏定理(命题 151.1)应归于 Fudenberg 和 Maskin(1986);我们所给的证明基于 Abreu, Dutta, 和 Smith (1994),第 8.9 节基于 Abreu(1988)。命题 155.1 及有限次重复博弈的纳什和精炼无名氏定理(命题 156.1 和 160.1)应归于 Benoît 和 Krishna(1985, 1987)。(Luce 和 Raiffa(1957, 第 5.5 节)较早地讨论了对于囚徒困境命题 155.1 的结论。)

对于有限次与无限次重复博弈间差别的(第 8.2 节)的较早讨论可参看 Aumann(1959, 第 6 节)。对于一个详细的关于在结果上偏好关系的讨论可参看 Diamond(1965)。对于用机器语言对某些无名氏定理的表达参看 Ben-
161 Porath 和 Peleg (1987)。图 158.1 中的例子摘自于 Benoît 和 Krishna (1985);Friedman(1985)包含了一个相似的例子。练习 148.1 应归于 Fudenberg 和 Levine(1989)。练习 152.1 摘自于 Fudenberg 和 Maskin(1986)。

对于当参与人使用混合战略时所引起的问题的讨论可参看 Fudenberg 和 Maskin (1991)。如我们所知,一个重复博弈的均衡并不都是有效的;进

一步说，在一个偏离发生之后由一个均衡所产生的结果可能不是有效的，即使结果在无偏离的情况下是有效的。Pearce(1992，第 4 节)讨论了一个模型，它检验了允许参与人集合在任一段历史后从他们现存的状态组合转移至帕累托最优组合(即至“再谈判”)的结果。如果在一个重复博弈中一些或全部参与人不知道成分博弈的形式，那么很多新问题将产生。Zamir(1992)和 Forges(1992)都是该领域工作的概览。

Krishna(1989)，Sorin(1990，1992)，Fudenberg(1992)和 Pearce(1992)都是覆盖了本章中的材料及其扩展的概览。

163 # 第 9 章

重复博弈中复杂性的考虑

本章我们考察无限次重复博弈中均衡战略的结构，在此博弈中每个参与人都关心他的战略复杂性。

9.1 介绍

在前一章我们描述了被称为“无名氏定理”的一组结论的代表，它确定了在众多关于参与人偏好的假设下，一个范围很宽的支付与纳什均衡甚至与无限次重复博弈中子博弈精炼均衡相容。无名氏定理给出了产生所需结果的一些均衡的构造。它并不要求这些均衡战略在任何意义下是合理的；我们对证明中所使用的均衡战略性质的判断都是非正式的。在本章我们将利用前面各章中所描述的“机器”这一工具来进一步关注均衡战略的结构，而不是均衡支付集合。

表述分析所依赖的基本假设是参与人关心他们战略的复杂性。当选择一个战略时，参与人面临着一笔交易：一方面他希望他的战略尽可能好地服务于他的目标，另一方面他希望它尽可能简单。参与人有很多理由认为简单有价值：计划越复杂它越有可能破产，它越难学，它需更多时间去完成。在这里我们不研究这些理由，而简单地假设复杂性是有成本的，且复杂性在决策主体的控制之下。

164 我们探讨这一假设对无限次重复博弈均衡结果的影响，特别地要探究复杂性成本的引入如何影响模型的预测。尽管我们将注意力限于重复博弈，但在任一选择模型背景下复杂性考察都可被研究。包含这种决策“程序”方面的模型称为“有限理性”模型。

9.2 复杂性与机器博弈

为了简便，本章中我们将注意力限于参与人的偏好由贴现准则所表示的无限次重复博弈：我们在 $G=\langle\{1,2\},(A_i),(u_i)\rangle$ 的两人用 δ 贴现的无限次重复博弈中研究参与人的行为(参看第8.3节)。我们研究这种行为是通过分析一个机器博弈进行的，在其中每个参与人选择一台机器去进行无限次重复博弈。本章中我们定义参与人 i 的一台机器为一四元组 $\langle Q_i, q_i^0, f_i, \tau_i\rangle$，其中

- Q_i 是一状态集合
- $q_i^0\in Q_i$ 是初始状态
- $f_i:Q_i\to A_i$ 是输出函数
- $\tau_i:Q_i\times A_j\to Q_i$(这里 $j\neq i$)是转移函数。

这个定义不同于上一章所给定的(第8.4节)地方在于，参与人的转移函数描述了状态如何随着别的参与人的行为而改变，而不是随着战略博弈的一个结果(即行为二元组)而改变。如在上一章所定义的那样，一台机器对应于扩展博弈中战略的概念，它要求对于每一段历史(包括那些被战略本身所排除的历史)参与人的行动都要被确定。如一般所理解的那样，这里我们想希望一台机器对应于一个行动计划，且因此仅将别的参与人的行动作为一个输入放进参与人的转移函数。

每个机器二元组(M_1, M_2)导致了 G 中的一个结果序列$(a^t(M_1, M_2))_{t=1}^{\infty}$和状态二元组的一个序列$(q^t(M_1, M_2))_{t=1}^{\infty}$，其定义如下：对 $i=1,2$ 和 $t\geqslant 1$ 我们有

- $q_i^1(M_1,M_2)=q_i^0$
- $a_i^t(M_1,M_2)=f_i(q_i^t(M_1,M_2))$
- $q_i^{t+1}(M_1,M_2)=\tau_i(q_i^t)(M_1,M_2), a_j^t(M_1,M_2))$(这里 $j\neq i$)。

我们限制每个参与人只能选择一台有有限个状态的机器，对参与人 i 165
用$\mathcal{M}_i$ 表示所有这种机器的集合。因此机器博弈是一种两人战略博弈，在其中每个参与人 i 的行动集合是$\mathcal{M}_i$。为了完成该博弈的描述我们需要描述参与人的偏好，如果我们假定每个参与人仅关心他重复博弈中的支付 $\mathcal{U}_i(M_1,M_2)=(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u_i(a^t(M_1,M_2))$，那么我们获得像纳什无

名氏定理(命题 145.2)一样的结论,因为在该结论的证明中所使用的触发战略能被有限台机器执行。如果另一方面每个参与人既关心他在重复博弈中的支付,又关心他战略的复杂性,那么如我们将要看到的那样,我们获得非常不同于无名氏定理的结论。

定义机器的复杂性有很多方法。我们采用一种朴素的方法:机器 $M=\langle Q, q^0, f, \tau\rangle$ 的复杂性 $c(M)$ 是状态的个数(即:Q 的基数性(cardinality)。我们的分析对于我们使用的复杂性测度是敏感的。我们视此测度为关于战略情形的一条附加信息,它应该反映参与人在执行战略中相关的困境。从这个角度看,模型对复杂性测度的敏感性是值得要的,在不同环境中不同的测度可能是恰当的。

在下列定义中我们假设,在机器博弈中每个参与人的偏好与他在重复博弈中的支付正相关且与其机器的复杂性负相关。

▶定义 165.1 $\langle\{1,2\},(A_i),(u_i)\rangle$ 的用 δ 贴现的无限次重复博弈的一个**机器博弈**(machine game)是一战略博弈 $\langle\{1,2\},(\mathcal{M}_i),(\succsim_i)\rangle$,其中对每个参与人 i

• $\mathcal{M}_i$ 是在无限次重复博弈中所有参与人 i 的有限机器的集合

• $\succsim_i$ 是一偏好次序,它在参与人 i 在重复博弈中的支付上是递增的,而在他的机器复杂性上是递减的:只要 $\mathcal{U}_i(M_1, M_2) > \mathcal{U}_i(M'_1, M'_2)$ 且 $c(M_i)=c(M_i')$ 或者只要 $\mathcal{U}_i(M_1, M_2)=\mathcal{U}_i(M_1', M_2')$ 且 $c(M_i)<c(M_i')$ 就有 $(M_1, M_2)\succ_i(M_1', M_2')$。

一种特殊的情形是每个参与人的偏好是可加的:$\succsim_i$ 由 $\mathcal{U}_i(M_1, M_2)-\gamma c(M_i)$(对于某个 $\gamma>0$)所代表,在此情形中 γ 可被解释为机器的每一状
166 态的成本。另一种特殊情形是参与人的偏好是**字典式的**(lexicographic):每个参与人先关心在重复博弈中他的支付,其次才关心他的机器复杂性。这种情形特别有趣,因为字典式的偏好接近于在标准的不考虑复杂性的重复博弈模型中的偏好,那是个机器模型的原始模型。

◇例 166.1 假定博弈 G 是囚徒困境,支付在图 166.1 给出。考虑执行冷酷战略的两状态机器 M(见图 166.2)。如果参与人的共同贴现因子 δ 足够大,那么在 G 的用 δ 贴现的重复博弈中该机器是对它自己的一个最优反应。即使通过应用一台更复杂的机器,参与人 1 在重复博弈中也不能获得一个更高的支付。不过,当给定参与人 2 应用 M 没有一台参与人 1 的机器比 M 在重复博弈中获得更高的支付,那么就有一台参与人 1 的获得同一

支付且更不复杂的机器:即在其中有一 C 被选取的状态。状态 D 被设计用来允许参与人威胁他的对手,但在均衡中这个威胁是冗余的,因为每个参与人经常选 C。因此每个参与人造访状态 D 而不影响结果,所以(M,M)不是机器博弈的一个纳什均衡。

	C	D
C	3,3	0,5
D	5,0	1,1

图 166.1 囚徒困境

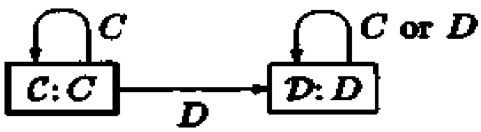


图 166.2 例 166.1 中的机器 M(一台完成重复囚徒困境博弈冷酷战略的机器)

◇例 166.2 对囚徒困境(如在前一个例子中)令 M 为图 167.1 中的机器。这台机器产生的行为可被解释为以展示惩罚的能力开始。在此展示之后参与人开始一个合作阶段,在其中他进行 C 且威胁通过移回到初始状 167
态去惩罚偏离者。如果两个参与人都使用机器 M,那么在重复博弈中的支付序列是(1,1,1)及随后一无限三维序列。

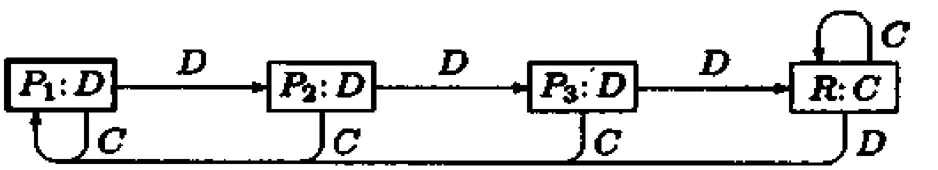


图 167.1 例 166.2 中的机器 M

我们断言如果参与人的共同贴现因子 δ 足够大,那么在参与人的偏好不给复杂性太多权重的情况下,(M,M)是机器博弈的一纳什均衡(就好像这样的情形,即他们的偏好是具有较小复杂性成本的字典式的或可加的)。证明如下,为了增加在重复博弈中他的支付,当他对手的机器在状态 R 时参与人必须在最少的次数里选 D。任一这种 D 的选择导致别的机器在至少三个时期里选 D,所以当 δ 足够靠近 1 时,参与人通过这样的一个行动无所收获(即对足够接近 1 的 δ 有 $5+\delta+\delta^2+\delta^3<3+3\delta+3\delta^2+3\delta^3$)。因此

对足够大的 δ,一个参与人通过无论多复杂的机器都不能提高他在重复博弈中的支付。

我们现在证明参与人通过使用一个更简单的机器不能获得同在重复博弈中一样的支付。当别的参与人的机器在状态 R 时,为了获得同样的支付他必须选择 C 至少一次。为了这样做他的机器必须有至少四个状态。为说明之,考虑第一期,比如说是 t,在其中 $f_i(q_i^t)=C$ 且 $q_j^t=R$。我们必须有 $f_i(q_i^{t-3})=f_i(q_i^{t-2})=f_i(q_i^{t-1})=D$ 且因此特别地有 $q_i^t\neq q_i^{t-1}$。进一步而言,当 $\tau_i(q_i^{t-1},D)=q_i^t$,因为 $\tau_i(q_i^{t-2},D)=q_i^{t-1}$,所以 $q_i^{t-2}\neq q_i^{t-1}$。类似地 $q_i^{t-3}\neq q_i^{t-2}$ 且 $q_i^{t-3}\neq q_i^{t-1}$。

在一个机器博弈中参与人不得不解决这样一个问题,即他要平衡获得一个高支付和应用一种简单机器的欲望。在某种意义上说,这个问题比在重复博弈中寻找一个最优战略更复杂,因为参与人必须考虑他的行为准则的复杂性;我们对参与人解决此问题的能力不加任何约束。

168 9.3 机器博弈均衡的结构

我们现在描述机器博弈纳什均衡的结构。首先对我们所作的关于例 166.1 的评述一般化,当 M_1 和 M_2 运行时,如果某个参与人 i 的机器 M_i 有一未被使用的状态,那么(M_1,M_2)就不是一个纳什均衡,因为该状态可被删除而不影响结果。并且参与人 i 偏好状态被删除的机器。

■引理 168.1 如果(M_1^*,M_2^*)是机器博弈的一个纳什均衡,那么对机器 M_i^* 的任一状态 q_i 存在一个时期 t 使得 $q_i^t(M_1^*,M_2^*)=q_i$。

我们下一个结论表明了在一纳什均衡中每台机器都有相同的状态数,并且机器博弈的任一纳什均衡都对应于重复博弈的一个纳什均衡。

■引理 168.2 如果(M_1^*,M_2^*)是机器博弈的一个纳什均衡,那么

• $c(M_1^*)=c(M_2^*)$

• 在重复博弈中与(M_1^*,M_2^*)相关的战略二元组是重复博弈的一个纳什均衡。

证明 对在重复博弈中参与人 j 的任一战略 s_j 和参与人 i 的任一机器 M_i，用 $\mathcal{U}_j(M_i, s_j)$ 表示当参与人 j 使用 s_j 和参与人 i 使用与 M_i 相对应的战略时，参与人 j 在重复博弈中的支付。因为 M_1^* 是有限的，参与人 j 寻找一个对重复博弈中的机器 M_1^* 的最优反应的问题 $\max_{s_j}\mathcal{U}_j(M_i^*, s_j)$（忽略复杂性）有一个解（参看 Derman(1970, p.23 定理 1)。令 $M_i^* = \langle Q_i, q_i^0, f_i, \tau_i\rangle$ 且对每个 $q \in Q_i$ 令 $\mathcal{V}_j(q) = \max_{s_j}\mathcal{U}_j(M_i^*(q), s_j)$，这里 $M_i^*(q)$ 与 M_i^* 不同之处仅在于初始状态是 q 的机器。对每个 $q \in Q_i$ 令 $A_j(q)$ 为下列问题的解集

$$\max_{a_j \in A_j}\{u_j(f_i(q), a_j) + \delta\mathcal{V}_j(\tau_i(q, a_j))\}.$$

那么在重复博弈中参与人 j 的一个战略是对与 M_1^* 相对应的战略的一个最优反应，当且仅当它所采取的行动是当参与人 i 的机器在状态 q 中时 $A_j(q)$ 的一个元素。特别地，对每个 $q \in Q_i$ 选择 $a_j^*(q) \in A_j(q)$，存在一个由下列机器所执行的最优反应，该机器有 $c(M_i^*)$ 个状态。

- 状态集合是 Q_i。
- 初始状态是 q_i^0。 169
- 输出函数 f_j 由 $f_j(q) = a_j^*(q)$ 定义。
- 转移函数 τ_j 定义为对所有 $x \in A_i$，$\tau_j(q, x) = \tau_i(q, f_j(q))$。

因为 (M_1^*, M_2^*) 是机器博弈的一个纳什均衡，从而有 $c(M_j^*) \leqslant c(M_i^*)$，且因此有 $c(M_1^*) = c(M_2^*)$。进一步而言，因为参与人 j 能应用一台具有 $c(M_i^*)$ 个状态的机器去获取一个在重复博弈中等价于 $\max_{s_j}\mathcal{U}_j(M_i^*, s_j)$ 的支付，从而由 (M_1^*, M_2^*) 所执行的战略二元组是重复博弈的一个纳什均衡。 □

? 练习 169.1 试给出一个三人博弈的例子，对于它相关的机器博弈有一纳什均衡，在均衡中参与人机器里的状态数是不一样的。

我们现在派生一个对机器博弈的纳什均衡集合有较强涵义的结论。为了获得结论的某些直观感受，考虑对于如图 169.1 所示的无限次重复囚徒困境博弈的机器二元组。这个机器二元组产生这样一条路径：初始具有结果为 (D, D) 的 $k \geqslant 2$ 个时期（参与人显示他们的威胁），在其之后结果为 (C, C)、(C, C)、(C, D) 和 (D, C) 的长度为 4 的循环无限次重复。任何一个参与人从循环中规定的行动的偏离都会导致他的对手的机器进入初始状

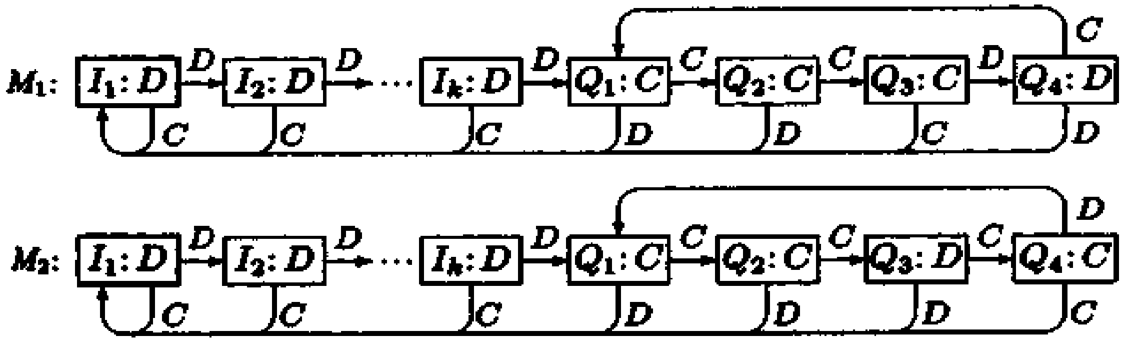


图 169.1 对无限重复囚徒困境博弈参与人 1 的机器 M_1 和参与人 2 的机器 M_2，二元组(M_1,M_2)产生了一条(D,D)发生 k 期且随后是序列(C,C)、(C,C)、(C,D)、(D,C)重复无限次的路径

态并在 k 个时期惩罚偏离者。正如你可检测一样，机器二元组当贴现因子 δ 足够接受 1 时是重复博弈的一个纳什均衡。不过它不是机器博弈的一个
170 均衡。为说明之，考虑 M_1，在 Q_1,Q_2,Q_3 这三个状态的每一个参与人 1 采取同一行动；她应用仅有三个状态知道什么时候选行动 D 这一事实。不过，她可以采取如下方法观察参与人 2 的行动来获得这一信息。假定她采用机器 M_1'，在其中三个状态 Q_1，Q_2 和 Q_3 均被惟一一个状态 Q 所代替，在状态 Q 中她选C，只要参与人 2 选 C 则状态保持在Q，若参与人 2 选 D 则状态移至Q_4，且若参与人 2 选 D 则从状态 I_k 转移到 Q，那么(M_1',M_2)同(M_1,M_2)一样产生相同的囚徒困境结果序列；因此在机器博弈中参与人 1 可以有利地偏离至 M_1'，因为它有比 M_1 更少的状态。

注意当参与人 2 的机器在 Q_1 或 Q_2 中时 M_1'并不控制参与人 2 的行动：如果参与人 2 在这些状态的某一个中选 D，那么 M_1'不是移到状态I_1而是移到 Q_4。如果参与人 1 使用机器 M_1'，那么参与人 2 可通过在状态 Q_3 选择 C 来利用这个特征。

该情形类似于下列的情形：一个伞兵在数到 100 后必须跳伞，而另一个参与人在数到 101 后必须跳伞。如果第二个伞兵数数，则他就可控制害怕跳伞的第一个伞兵。不过，在飞机紧张的环境中数数是有成本的，且第二个参与人可以简单地通过观察他的朋友且在她之后立即跳伞来避免数数的负担。但是，如果第二个伞兵不数数则第一个伞兵可利用这种控制的缺点并且不跳伞。

总而言之我们可以证明：如果一个纳什均衡机器二元组产生某个参与人在两个不同时期采取同一行动的结果，那么另一个参与人也在这两个时期采取同样的行动（与我们已讨论的例子中在时期 $k+2$ 和 $k+3$ 参与人的行动相矛盾）。

■引理 170.1 如果(M_1^*, M_2^*)是机器博弈的一纳什均衡，那么在由M_1^*和M_2^*所规定的参与人1与参与人2的行动之间存在着一个一一对应：如果对某个$t \neq s$有$a_i^t(M_1^*, M_2^*) = a_i^s(M_1^*, M_2^*)$，那么$a_j^t(M_1^*, M_2^*) = a_j^s(M_1^*, M_2^*)$。

证明 令$M_i^* = \langle Q_i, q_i^0, f_i, \tau_i \rangle$且对每个$q_i \in Q_i$定义与在引理168.2的证明中一样的$A_j(q_i)$。由引理168.2的第二部分机器$M_j^*$执行重复博弈中的一个战略，它是问题$\max_{s_j} \mathcal{U}_j(M_i^*, s_j)$的一个解。因而对所有$t$有$f_j(q_i^t(M_1^*, M_2^*) \in A_j(q_i^t(M_1^*, M_2^*))$。由此若对两个时期$t$和$s$，在其中$a_i^t(M_1^*, M_2^*) \neq a_i^s(M_1^*, M_2^*)$且$a_j^t(M_1^*, M_2^*) = a_j^s(M_1^*, M_2^*)$，那么存在参与人$j$的一个最优策略$a_j'$使得$a_j'(q_i^t(M_1^*, M_2^*)) = a_j'(q_i^s(M_1^*, M_2^*))$。也就是，只要参与人$i$的状态是$q_i^t(M_1^*, M_2^*)$或$q_i^s(M_1^*, M_2^*)$则参与人$j$采用同一行动。下列机器完成策略$a_j'$且有$c(M_i^*) - 1$个状态，这与引理168.2的第一部分相矛盾。 171

- 状态集合是$Q_i \setminus \{q_i^s\}$。
- 初始状态当$q_i^s \neq q_i^0$时为q_i^0，否则为q_i^t。
- 输出函数由$f_j(q) = a_j'(q)$确定。
- 转移函数定义如下：如果$\tau_i(q, f_j(q)) = q_i^s$，那么对所有$x \in A_i$有$\tau_i(q, x) = q_i^t$，否则对所有$x \in A_i$若$q \neq q_i^t$且

$$\tau_j(q_i^t, a_i) = \begin{cases} \tau_i(q_i^s, f_i(q_i^s)) & 若\ a_i = a_i^s(M_1^*, M_2^*) \\ \tau_i(q_i^t, f_j(q_i^t)) & 否则 \end{cases}$$

有$\tau_j(q, x) = \tau_i(q, f_j(q))$。这就完成了证明。 □

该结论对在任一每个参与人有两个行动的博弈中的均衡结果有一明显的内涵。例如，如果在重复囚徒困境中两个结果在均衡路径上出现，那么这对结果是$\{(C, C), (D, D)\}$或$\{(C, D), (D, C)\}$。

我们现在转向探讨均衡机器的结构。因为每个参与人的机器是有限的，所以有一个最小数t'使得对某个$t > t'$我们有$q_i^t = q_i^{t'}$；令t^*是这种t的最小值。在时期t'开始的状态二元组序列包含长度为$t^* - t'$的循环。我们称这个阶段为**循环阶段**(cycling phase)。在时期t'前的阶段为**导入阶段**(introductory phase)。

我们现在证明，一个参与人在循环和导入阶段所应用的状态集合是不相交的。进一步而言，在导入阶段每个状态只被进去一次，在循环阶段所应

用的参与人的每个状态在每一循环只出现一次。因此均衡中在机器里的参与人 1 和 2 的状态之间存在着一个一一对应，这一事实可被解释为在每个时期每台机器“知道”别的机器所处的状态。

■**命题 171.1** 如果(M_1^*, M_2^*)是一个机器博弈的均衡，那么存在一个时期 t^* 和一整数 $l<t^*$ 使得对 $i=1,2$，序列$(q_i^t(M_1^*, M_2^*))_{t=1}^{t^*-1}$中的状态是不一样的，且对 $t \geqslant t^*$ 有 $q_i^t(M_1^*, M_2^*) = q_i^{t-l}(M_1^*, M_2^*)$。

172 **证明** 令 t^* 为某一机器的某一状态第二次出现的第一个时期。即，令 t^* 为有一参与人 i 和一时期 $t_i<t^*$ 使得 $q_i^{t^*} = q_i^{t_i}$ 的最小时期。我们有 $a_i^{t^*} = a_i^{t_i}$ 且由引理 170.1 有 $a_j^{t^*} = a_j^{t_i}$，从而对所有 $k \geqslant 0$ 我们有 $q_i^{t^*+k} = q_i^{t_i+k}$，所以用引理 168.1，$c(M_i^*) = t^*-1$。通过参与人 i 的选择所有 M_j^* 的状态直到时间 t^*-1 都是不一样的，所以引理 168.2 的第一部分蕴含着存在 $t_j<t^*$ 使得 $q_j^{t_j} = q_j^{t^*}$。还需证明 $t_j = t_i$。假设其不成立，比如，$t_j > t_i$，那么参与人 j 使用一台这样的机器可获得相同的结果路径，在该机器里 $q_j^{t_i}$ 被从 $q_j^{t_j-1}$ 至 $q_j^{t^*}$ 的一个转移排除，从而省略 $q_j^{t_i}$。但与 M_j^* 的最优性相矛盾。 □

一个机器博弈是一战略博弈，所以分析没有考虑由子博弈精炼均衡的概念所模化的类型。为了加进这些考虑，我们可以修改解的概念并且要求在重复博弈中的每段历史之后机器二元组为机器博弈的一均衡。这样一个修改蕴含了机器的运行没有任何导入阶段：一旦循环阶段被达到，一个在进行过程中可以改变他的机器的参与人想省略任何导入状态。从而如练习 173.2 所阐述的那样，均衡路径集合被这个解的修改严格限制了。

9.4 字典式偏好的情形

上部分的结论比较有意义地约束了机器博弈的均衡集合。为了进一步约束均衡集合，我们需要在每个参与人的偏好关系中确定重复博弈中他的支付与他的机器复杂性之间的替代关系。本部分我们假设参与人的偏好是字典式的(复杂性是第二考虑，在重复博弈中的支付之后)；我们将注意力限

于成分博弈是囚徒困境(支付见图 166.1)的情形。

如我们上面注意到的，引理 170.1 蕴含了在均衡路径上所发生的结果集合是$\{(C,C),(D,D)\}$的一个子集或$\{(C,D),(D,C)\}$的一个子集。先考虑前一类型的均衡，令 n_C 和 n_D 为两个非负整数，它们中至少有一为正的，那么对足够接近于 1 的 δ 可证明有一长度为 n_C+n_D 的循环，在其中(C,C)出现 n_C 次且(D,D)出现 n_D 次，对于 $n_C=n_D=1$ 的情形，有一每个参与人用图 173
173.1 中机器 M 的对称均衡。(对于不是字典式的偏好二元组(M,M)只有在参与人的偏好对复杂性不加太大权重的情形下才是一个均衡。)

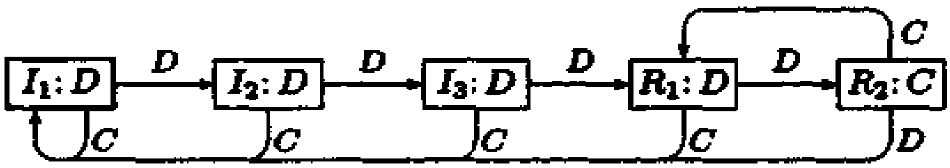


图 173.1　每个参与人在无限次重复囚徒困境博弈中的机器 M

? 练习 173.1

a. 试证明若 δ 足够接近 1 则机器二元组(M,M)是机器博弈的一个纳什均衡。

b. 试证明若机器 M 被这样修改：在 R_1 进行 C，在 R_2 进行 D，并且在 R_1 和 R_2 中的转移被反过来，那么新的机器二元组不是机器博弈的一纳什均衡。

在这些均衡中导入阶段是非空的，并且对任一支持(C,C)为一结果路径的均衡也是如此。

? 练习 173.2　试证明每一个以(C,C)为结果的均衡有一导入阶段。

现在考虑在均衡路径上的每个结果是(C,D)或(D,C)的均衡。一些这种均衡是循环的，没有任何导入阶段。精确点来说，对所有满足 $5n_i/(n_1+n_2)>1(i=1,2)$的正整数 n_1 和 n_2 存在 δ 足够大使得有一机器博弈的均衡，在其中循环包括 n_1 个(D,C)的进行和随后 n_2 个(C,D)的进行且无任何导入阶段。(这个在 n_1 和 n_2 上的条件保证了每个参与人的平均支付超过了他的最小最大支付 1。)

对于 $n_1=n_2=1$ 情形的一个均衡如图 174.1 所示。这个均衡的一个解释是参与人轮流对对方慷慨。人们可以认为(C,D)是一个参与人 1 给参与人 2 一件礼物的事件且(D,C)是一个参与人 2 给参与人 1 一件礼物

的事件。在均衡中一个参与人若他的对手不接受他的礼物则他并不介意
174 (即当他可能已选 D 且已接受礼物时选 C),但他坚持他的对手在他期望得到礼物之时给他礼物(进行 C)。如果他收不到礼物,那么他不会移到他是慷慨的那一状态。

在我们的全部分析中成分博弈是战略博弈。人们也可考察成分博弈是扩展博弈的情形。当重复博弈纳什均衡的分析不受影响(尽管子博弈精炼均衡的集合可能有些不一样),机器博弈纳什均衡的分析在此情形中如下列练习所示也很不一样。

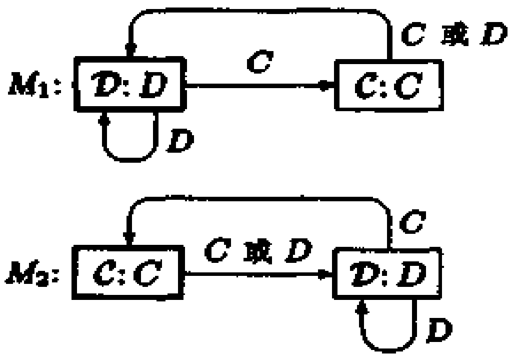


图 174.1 对于无限次重复囚徒困境博弈的参与人 1 的机器 M_1 和参与人 2 的机器 M_2 对足够大的 δ 二元组 (M_1, M_2) 是机器博弈的一个均衡;它产生包含循环 $((D,C),(C,D))$ 重复的路径

	A	B
A	3,1	1,3
B	2,0	2,0

图 174.2 练习 174.1 中重复博弈的成分博弈

? 练习 174.1 考虑成分博弈由图 174.2 所给的无限次重复博弈。

a. 试证明与机器博弈纳什均衡相关的路径集合只包含结果 (A, A) 和 (B, B)。

b. 试证明如果在机器博弈中参与人的偏好是字典式的,那么只包含结果 (A, A) 和 (B, B) 的对 δ 足够大的每一个有限序列是与某一机器博弈纳什均衡相关的路径的循环阶段。

175 c. 注意博弈是完全信息扩展博弈的战略形式。假定参与人进行的成分博弈是这个扩展博弈的无限次重复博弈,并在每轮结束时了解到所发生

的终点历史，试证明这个重复博弈的机器博弈有惟一纳什均衡，其中在每个时期支付组合是(2,0)。（提示：当参与人 1 选 B 时她不能控制，若她选 A 则参与人 2 计划选 A 还是选 B。）

[注解]

本章基于 Rubinstein(1986)和 Abreu and Rubinstein(1988)。证明的思路，特别是引理 170.1 的证明，是 Piccione(1992)对 Abreu 和 Rubinstein (1988)证明的一个修改。练习 174.1 基于 Piccione 和 Rubinstein(1993)。

在相关的一系列文献的思路中，一个参与人可利用的机器复杂性被认为是非常广的。这些研究的主要目的是证明不同于(D,D)重复的均衡结果可在有限次重复囚徒困境博弈中被支持；例子可参看 Neyman(1985)和 Zemel(1989)。