

博弈生存

(第三版)



第三版序言

因人口的增多，我们每个人平均所拥有的资源越来越少；因科技的发达，我们与他人的距离越来越近。我们与他人生活在一个世界之中，这是一个既冲突又合作的世界。

博弈论学科的诞生与发展、以及博弈思想的流行，与20世纪的时代背景分不开。20世纪是一个人类“大冲突”和“大合作”的世纪。腥风血雨的两次世界大战是人类的大冲突，在这个大冲突中人类损失巨大；第二次大战后两个敌对集团间的冷战游戏是特殊的大冲突，冷战中尽管充满危机，但冷战的双方充满克制，这是冲突中的合作；20世纪末叶席卷全球的经济一体化、全球化则是大合作，人类的文明在不同国家、文化之间的相互交融中进入新的阶段。

21世纪人类的冲突和合作进入新的时期，这不是一个没有冲突的世纪。发生于2001年的“911”事件告诉人们，冲突依然存在。“911”意味着目前的主要冲突是西方世界与恐怖主义分子之间的冲突。尽管这不是实力对等的战争——西方世界处于强势、恐怖主义分子处于弱势，但是强势的一方并不感到安全，弱的一方并没有感到完全失败。人类是健忘的，也许只有通过血的教训，人类才能够认识到正确解决问题的途径是通过对话、交流从而合作，而不是通过武器从而消灭对方。人类要善于从过去的历史中学习，以面对未来。

博弈论是研究利益依存的多个理性人之间的互动。这样的理论不仅有助于我们理解社会，同时有助于我们更好地与他人冲突中或合作中共存。

本书于2002年出版，2004年再版，因读者的厚爱，本书成了畅销书和长销书。自本书出版以来，本人收到读者不计其数的



邮件：或探讨本书涉及到的博弈论的问题，或指出本书的错误……来函无不体现出读者对博弈论对知识的热诚，无不体现出对本书的喜爱。这里，本人对所有来函的读者表示感谢。因精力有限，本人没有能够对来函一一回复，请读者谅解。

本书试图用博弈论来分析各种社会现象，这些现象有合作的，也有竞争的。在中国“博弈”一词已经逐渐成为一个流行术语，本书在一定程度上促进了这种流行。因本书的畅销，本书出版后，模仿本书的许多作品充斥于市场，甚至有些书的书名中直接用了或含有“博弈生存”的字样。其中一些书不乏有抄袭之嫌疑，也存在许多对博弈论的误读之处。当然其中一些书还是有价值的：其价值在于，它们将博弈论作为理论工具，用之来分析或说明如中国历史或世界历史中的现象。

本书虽然是科普，但本人力图提供给读者的是原汁原味的博弈论思想，第三版修改时也秉承这个原则。这一版在第二版的基础上做了大幅度的调整：增加了三章新内容——“对话博弈与意见的一致性”、“制定博弈规则：机制设计”和“合作还是不合作”；对原有章节进行增或删，如在“公平分配”一章中增加了“《塔木德》中的遗产分配困惑与破产问题”一节；等等。

本书一方面力争将博弈论发展的最新内容加了进来，如补充了2005年诺贝尔奖获得者奥曼和谢林的工作；另外一方面，本书在保持科普风格的同时，尽可能地标上所用文献来源，这样做，旨在帮助年轻学子接近一流学者的工作，以促进博弈论的发展。当然，本版与前两版一样包含了本人对博弈、社会以及对博弈论本身的最新思考。

潘天群

2007年7月于南京龙江

第二版序言

人是行动的动物，因而是策略选择的动物。我们每天的所见所闻无时不体现出人行为的策略选择性。社会中的每个人都在努力用自己的理性选择策略来达到自己的生存目的。这也是研究多个理性人互动决策的博弈论能够用于解释社会现象的原因。掌握了博弈论的理论，我们就能自觉地用它来分析我们所见到的社会现象。我们熟知的是，革命导师马克思、恩格斯在《共产党宣言》中号召无产阶级起来革命，他们豪迈地宣称“让统治阶级在共产主义革命面前发抖吧！无产者在这个革命中失去的只是锁链，他们获得的将是整个世界！”如果用博弈论术语来分析，革命导师是在告诉无产者：“革命”行动与“不革命”行动相比，“革命”是“占优的”行动策略。“革命”将使无产者失去锁链并得到整个世界，而如果“不革命”，无产者仍然是一无所有的奴隶！

当然，博弈论有它的局限。在某些博弈中，人的实际行为不完全符合博弈论的理论结果（实验经济学力图用实验来验证博弈论的演绎结果），它们之间的偏差有一些不能用理性的有界性来解释，这使人们怀疑博弈论的理性人假设是否能够站得住脚。许多哲学家认为人是自由的，而博弈论专家假定人是理性的。人是自由的还是受理性法则所支配？如果人是理性的，那么他便不是自由的。理性人假设下的人是一个“机械装置”。一个博弈是多个这样的装置的互动：对于某个博弈参与人来说，如果给定他的策略空间、目标函数以及博弈参与人的个数、信息分布等，他能够确定自己的最优策略，或者至少能够排除掉某些策略。此时人便不是自由的——理性的行为选择预设了某个正确的行动方向，而自由是不预设任何方向的。理性法则规定无产者必须选择“革



命”，而无产者作为自由的主体，尽管他知道“革命”的结果优于“不革命”的结果，也完全可以不选择“革命”，如果他愿意做奴隶的话。

物理学是不完美的，这并不妨碍我们用它来观察和解释我们周围的世界，博弈论的局限也并不妨碍我们用它来分析和解释社会现象。《博弈生存》是一本深受读者欢迎的博弈论普及读本，自出版以来在全国许多书店（包括网上书店）长期位于畅销书排行榜之上，其畅销程度是本人写作时所未能料及的。本书的读者群体非常广泛，从中学生到大学教授，其中大多数读者是大学生。许多读者来信或来电，与本人探讨书中涉及博弈论的问题，并提出一些建设性的意见……本人对读者的厚爱表示感谢。

鉴于《博弈生存》一书引起的广泛关注，以及博弈论本身的重要意义，本人决定对原来的版本进行修订。最新的版本对第一版的部分章节作了修改和删减，纠正了第一版中的错误和不确切之处。同时，增加了“诸葛亮、周瑜的掌中之‘火’”和“将军的困境”两小节，以及新的一章——“公平分配可能吗？”，并且补充了附录——论述纳什均衡定义和纳什均衡存在定理。本书力图通过比较浅显的例子普及博弈论的知识和方法，阐发博弈论的思想，为普及博弈论作出绵薄之贡献，因而修订版仍保持科普风格。如果读者想借用博弈论工具进行学术研究或解决实际问题，请查阅有关的博弈论学术著作。

感谢中央编译出版社领导、编辑和发行人员，尤其是贾宇琰编辑，感谢他们为本书的出版和发行所付出的辛勤劳动。

潘天群

2004年8月于南京

目 录

Contents

第三版序言	1
第二版序言	1
导言	1
一、失火了，你往哪个门跑？	3
二、博弈论能解释所有社会现象吗？	6
第一章 博弈论基础知识	9
一、基本术语	11
二、博弈的类型	20
三、博弈案例	21
1. 徒博弈与我国的应试教育困境	21
2. 斗鸡博弈与古巴导弹危机	23
3. 美苏武器竞赛与骑虎难下博弈	26
4. 警察与小偷的故事——混合策略问题	29
5. 《三国演义》中的空城计与信息不对称的博弈	33
6. 约定与协调博弈	39
第二章 公共知识与集体行动的均衡	43
一、公共知识与博弈	45
二、一个寓言——村庄里的大屠杀	48



三、帽子：红色的还是白色的？	52
四、皇帝新装的新解读	55
五、“教一学”均衡的公共知识条件 ——教育的公共知识结构分析	56
六、诸葛亮与周瑜的掌中之火	58
七、将军的困境与公共知识形成悖论	59
八、交流与公共知识的形成	61
第三章 你的权力有多大？	65
一、“独裁的”妻子	67
二、一个国家的权力分配故事与班扎夫权力指数	69
三、夏普里—舒比克权力指数	75
四、聪明的股东：班扎夫权力指数的应用	78
五、逻辑结构与投票影响度	81
六、民主社会中为什么很多人不投票？	85
七、一群体中有多少种可能的权力结构？	87
第四章 公平分配：可能吗？	91
一、什么是公平分配？	93
二、8个金币的故事	94
三、贡献与收益相等：夏普里值在分配中的应用	95
四、海盗分宝石：公平规则下的不公平	97
五、所罗门的智慧：公平不是平均	100
六、从分蛋糕到财产分割与边界争端的解决： 双赢的分配	101
七、《塔木德》中的遗产分配困惑与破产问题	106

第五章 信息、言语与行动	111
一、知识与博弈	113
二、信息与战争博弈	115
三、声称的策略与实际的策略	120
四、“威胁”与“承诺”是否可信	122
五、空口声明博弈	125
第六章 对话博弈与意见的一致性	127
一、曼定理与意见一致性问题	129
二、突然考试悖论中的意见不一致问题	131
三、不一致的观点与交易	131
四、不一致观点的消除：成功的劝说	132
五、不一致的保持：不成功的劝说	133
六、诸葛亮舌战群儒	135
第七章 博弈中的学习机制问题	139
一、明天太阳必然能升起吗？——归纳推理的合理性 问题	141
二、归纳的暴力：偷东西的人应该叫小偷吗？	145
三、行动中的两种推理：归纳与演绎	146
四、酒吧问题与“百花奖”的评选	148
五、少数者博弈及其应用：股票买卖与交通路线的选择	154
六、归纳学习：有效与无效	158
七、对称的人群为什么会出现破缺的行动？ ——盲目的学习	160
第八章 博弈规则与机制设计	163
一、三个和尚何以能够有水吃	165



二、卖力还是磨洋工	166
三、拍卖的规则	167
四、群体决策及其规则的确定	169
五、“帽子”与群体决策	172

第九章 理性的界限 177

一、蜈蚣博弈悖论	179
二、最后通牒博弈中理性的困境	181
三、陈水扁上台与民主选举的不可根除的缺陷	183
四、通过民主的方式能使少数人支持的候选人赢吗	187
五、投票悖论与阿罗不可能性定理	191
六、革命：另外一种投票方式	194

第十章 合作还是不合作 199

一、囚徒困境中的合作与不合作	201
二、冲突中的合作：谢林的思想	203
三、联盟与合作	206
四、中国传统文化中的联盟思想	207

第十一章 集体行动的悲剧 209

一、哈定悲剧	211
二、中国北方草原荒漠化问题	213
三、发展经济与对自然破坏	215
四、光天化日之下的违法行为为什么总能成功？	216

第十二章 集体行动悲剧的克服与文明 223

一、囚徒困境的消解与文明	225
二、一个关于猴子的故事	228



三、人类原初的状态就是囚徒困境	229
四、国家和道德是两种克服囚徒困境的机制	231
五、更为基本的道德调节	235
六、道德作为“模码”的遗传与变异	237
七、国家的作用及局限	239
八、文明新解	242
九、当前的悲剧与解决途径：世界政府的可能性与 可行性分析	244
第十三章 集体行动的未来能否预测？	249
一、弑父娶母的俄狄浦斯的传说	251
二、回到过去是原则上不可能的	254
三、我们不知道未来我们会知道什么	255
四、社会在什么意义上是可预测的？	257
五、社会科学的作用	258
结语	261
附录	265
参考文献	266

导言

一、失火了，你往哪个门跑？

一天晚上，你参加一个派对，屋里有很多人，你玩得很开心。这时候，屋里突然失火，火势很大，无法扑灭。此时你想逃生。你的面前有两个门，左门和右门，你必须在它们之间选择。但问题是，其他人也要争抢这两个门出逃。如果你选择的门是很多人选择的，那么你将因人多拥挤、冲不出去而烧死；相反，如果你选择的是较少人选择的，那么你将逃生。这里我们不考虑道德因素，你将如何选择？

你的选择必须考虑其他人的选择，而其他人的选择也考虑你的选择。你的结果——博弈论称之为支付，不仅取决于你的行动选择——博弈论称之为策略选择，同时取决于他人的策略选择。你和这群人构成一个博弈（game）。

这就是博弈论（game theory）——对这一类多人互动决策进行研究的理论。

上述博弈是一个叫张翼成的中国人在1997年提出的一个博弈论模型，被学界称之为少数者博弈或少数派博弈（Minority Game）。当然，原来的博弈形式不是这么简单，这里我把它简化了，我们在第三部分论述归纳推理时还要谈这个博弈模型。现在很多学者在研究这个模型。

生活中博弈的案例很多，你会见到很多例子。只要涉及到人群的互动（interaction），就有博弈。

什么叫博弈？博弈的英文为 game，我们一般将它翻译成“游戏”。而在西方，game 的意义不同于汉语中的游戏。在英语中，game 即是人们遵循一定规则下的活动，进行活动的人的目的是使自己“赢”。如：Olympic Games 翻译成奥林匹克运动会而不是奥林匹克游戏大会。在英文中，game 有竞赛的意思，进行 game 的



失火了，你往哪个门跑？

人是很认真的，不同于汉语中游戏的概念。在汉语中，游戏有儿戏的味道。因此将关于 game 的理论，即 game theory 翻译成博弈论或者对策论，是恰当的。本书下面统称 game theory 为博弈论。

博弈论作为一门学科的出现有 60 多年的历史。博弈论的开创者为诺意曼与摩根斯坦，他们 1944 年出版了《博弈论与经济行为》，该书的出版被公认为博弈论这门学科诞生的标志。诺意曼是著名的数学家，他同时对计算机的发明作出了巨大贡献，他去世时博弈论还未对经济学产生广泛影响，否则经济学的诺贝尔奖的获得者中必定有他的名字，因为诺贝尔奖有规定，只颁发给在世的学者。谈到博弈论，不能忽略博弈论天才纳什（John Nash）。纳什的开创性论文《n 人博弈的均衡点》（1950）、《非合作博弈》（1951），给出了纳什均衡的概念和均衡存在定理。今天博弈论已发展成一个较完善的学科。

博弈论对于社会科学有着重要的意义，它正成为社会科学研究范式中的一种核心工具，以至于我们可称博弈论是“社会科学的数学”，或者说是“关于社会的数学”。从理论上讲，博弈论是研究理性的行动者（agents）相互作用的形式理论，而实际上它正深入到经济学、政治学、社会学等等，被各门社会科学所应用。甚至有学者声称要用博弈论重新改写经济学。诺贝尔经济学奖已两度授予博弈论专家：2005 年颁发给罗伯特·奥曼（R. Aumann）、托马斯·谢林（T. Schilling），1994 年颁发给纳什（J. Nash）、塞尔屯（R. Selten）、哈桑尼（J. Harsanyi）。而像 1985 年获得诺贝尔奖的公共选择学派的领导者布坎南，1995 年获得诺贝尔奖的理性主义学派的领袖卢卡斯（Lukas），其理论与博弈论都有着较深的联系。现在博弈论正渗透到各门社会科学，成为这些学科的通用工具，更重要的是它正深刻地改变着人们的思维。

二、博弈论能解释所有社会现象吗？

社会由不同的人群的集合体所构成。不同的人群集合体形成不同的结构，一个结构中的群体之间的相互作用（interaction）就构成一个博弈。这个博弈是广义上的。社会中有不同的文化，人类有文明、道德，如果说文明、文化、道德等是宏观的社会现象，那么还存在微观的社会现象，如：群体为什么有合作又有不合作？为什么人群之间或集团之间有“威胁”或“承诺”？这些都是博弈论研究的对象。本书下面将努力用博弈论的基本思想来解释社会中的这些现象。

博弈论对人的基本假定是：人是理性的（rational）。所谓理性的人是指行动者具有推理能力，在具体策略选择时其目的是使自己的利益最大化。博弈论研究的是理性的人之间如何进行策略选择的。

博弈论力图在这个最简单的假定下得到丰富的结论，正如我们下面所看到的，它确实做到了。博弈论专家的这种做法如同物理学家对自然的假定一样。大家知道，物理学家往往给出几个最基本的假设，这最基本的假设构成公设，其余的结论由它们推得，如爱因斯坦的狭义相对论只有两条公设：（1）物理定律在所有参考系中不变；（2）在所有参考系中光速保持常数。多么简单的公设！在这两个公设下得出了惊奇的结论，如运动的参照系中尺子收缩，时钟变慢，等等。相对论的这两条公设改变了物理学的整个构架，也改变了人们对自然的整个看法！博弈论当然不是关于自然的，它是关于社会的，它不能构成人们对自然看法的革命。博弈论的假定是非常简单的，它能得出令人惊奇的结论吗？它能改变人们关于社会的看法吗？——这些是伟大科学的要求！我们将发现，博弈论确实如此！



本书力图用“科学的”方法来解读社会现象。经济学是对社会现象进行研究的科学，社会学、政治学等等也都是社会科学。然而今天经济学已成了帝国主义，它的领域没有疆界。不仅仅经济行为是它的研究领域，它的研究范围是“经济的”行为，在经济学家看来人类的几乎所有的行为都是“经济的”，即在经济学家看来，其他社会科学无存在的必要。这里我们看到“经济的”等同于“理性的”。由此可见，研究互动中的理性人行为的博弈论在经济学家看来可解释几乎一切社会现象。

本书的写作方式是科普式的（popular science）。本书没有对博弈论作系统的论述，提到的理论有时缺少来龙去脉的交代。若读者对所涉及的理论感兴趣，请查找有关的书籍或资料。

这里讲一个小小的插曲。在与朋友聊博弈论时，他随意说了一句话给我启发很深。他说，中国人研究其他学问难说，但研究博弈论是有优势的。这句话是褒义，也是贬义。说它是褒义，是因为中国古代有很多这方面的著述与实践，春秋战国时期诸侯争雄，其实是谋士之间的谋略较量，而罗贯中的《三国演义》在今天看来就是一部博弈论教材！无论是兵书如《孙子兵法》、《三十六计》，还是现代流行的所谓“商战策略”、“公共关系”以及所谓“厚黑学”，都是关于如何赢得与人交往的胜利的，或者说关于在与他人的竞争或合作中如何获取成功的。说它是贬义，是因为，中国人走关系、相互算计是出了名的，中国人对博弈论有突然的悟性。我们常说的“世事洞明皆学问，人情练达即文章”：人与人之间的关系、社会交往均是学问，是实用的博弈学问。而中国很多“做人”的道理，道出了如何在人与人的博弈中获取成功，如：在任何场合下都不要得罪人，不要锋芒毕露（如“枪打出头鸟”），等等。不过，在中国文化传统中，人与人之间的所谓关系之学并不像西方那样是科学，而更像一门艺术。

本书下面将尝试着用博弈论解读人类的社会行动或集体行