国家社会科学基金（15BGL014）
国家自然科学基金（No.71072128）
教育部人文社科基金（No.12YJCZH226）

复杂网络上的博弈

谢逢洁　著

清华大学出版社

西安邮电大学学术专著出版基金资助出版

复杂网络上的博弈

谢逢洁 著

清华大学出版社
北 京

内容简介

本书从博弈论和复杂网络的基本概念及基础理论出发，介绍相关研究领域，重点讲解博弈论和复杂网络的交叉领域的基本框架和研究思路。在此基础上，以不同的复杂网络结构（规则网络、小世界网络、无标度网络、社区网络、动态网络）作为章节划分的原则，结合具体的研究内容，配套相应的MATLAB仿真程序及注释说明，详细讲解该研究领域中具体问题的建模思路和仿真方法。

本书主要适合研究生的学习和研究。

图书在版编目（CIP）数据

复杂网络上的博弈/谢逢洁著. --北京：清华大学出版社，2015

ISBN 978-7-302-40262-6

Ⅰ.①复… Ⅱ.①谢… Ⅲ.①计算机网格—研究 Ⅳ.①TP393

中国版本图书馆CIP数据核字（2015）第106321号

责任编辑：王玉玲
装帧设计：王　军
责任校对：王凤芝
责任印制：刘海龙
出版发行：清华大学出版社

网　　址：http：//www.tup.com.cn，http：//www.wqbook.com
地　　址：北京清华大学学研大厦A座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
课件下载：http：//www.tup.com.cn

印 装 者：三河市金元印装有限公司
经　　销：全国新华书店
开　　本：185mm×230mm　**印　张**：13.5　**字　数**：240千字
版　　次：2016年1月第1版　**印　次**：2016年1月第1次印刷
定　　价：47.00元

产品编号：063409-01

引言

在众多的社会和经济组织中，个体都面临着自身利益和集体利益的选择冲突。这种选择冲突在博弈论中被称为社会两难选择（social dilemma），在社会学和经济学领域中有着很长的研究历史。比如，国家之间的贸易谈判、城市污染、知识共享、公共资源使用等都属于社会两难选择问题。如果个体在两难冲突中选择了集体利益，则称其采取了合作行为；相反，如果个体选择了自身利益，则称其采取了背叛行为。背叛行为的选择尽管给个体带来了经济利益的有效性，但必然导致集体经济利益的无效。

在人类文明的整个发展历程中，尽管存在着各种各样的冲突和争斗，但合作一直是人与人之间的主流行为，是人类社会得以稳定发展、进步的原动力，其范围和深度是任何其他动物无法比拟的。因此，人类的合作行为为何能够发生并得以维持是人类认识自我、了解自我的一个重要课题。以博弈理论为研究方法，研究者们通过不懈努力，尝试从各种不同的角度对人类的合作行为进行解释，形成了 5 个主要的研究分支，包括亲缘合作、互惠合作、声誉合作、网络合作和团队合作等。

随着 1998 年、1999 年小世界网络和无标度网络的相继发现，复杂网络理论研究引起了各个学科领域的广泛关注并形成了一个快速发展的交叉学科，在近 15 年内取得了令人瞩目的成果。复杂网络上的博弈研究正是将复杂网络理论和博弈论相结合，从人与人之间复杂的相互作用关系的视角来解释人类的合作行为，属于网络合作研究分支，产生了非常多的研究成果。但是，由于该研究方向是复杂网络理论和博弈理论的交叉研究，且产生的时间也较短，国内还没有一本系统介绍该研究方向的论著，以至于想进入该领域的学者及学生都面临着一道技术上的难关。本书旨在对复杂网络上的博弈研究进行全面的介绍，详细地讲解该学科的研究内容、建模思路和仿真方法，结合相应的 MATLAB 仿真程序，一步一步地教会读者，轻松地进入该研究领域的大门。

目录

第一章 博弈论的基本概念及合作行为

对于目标和偏好存在潜在冲突的相互作用的参与者，博弈论提供了一整套具有解释力和预测力的分析工具和解的概念，从而使得它在许多的学科领域得到成功应用，如经济学、演化生物学、计算科学和运筹学、政治和军事学、人类学和伦理学、物理学等。

博弈论的诞生可以追溯到 1944 年普林斯顿大学出版社出版的书籍《Theory of Games and Economic Behaviour》[1]。这是第一本较为全面描述博弈理论的著作，通常被认为是博弈理论形成的标志。但实际上，在大多数学科领域，博弈论的萌芽时间远远早于此。例如，法国经济学家 Augustin Cournot 于 1838 年使用具有限制条件的纳什均衡概念求解垄断市场的定量选择问题。他的理论在 1883 年得到 Joseph Bertrand 的进一步扩展，形成价格竞争理论。Ysidro Edgeworth 在 1881 年提出了合作博弈理论的概念。Emile Borel 于 1921 年提出了混合策略和二人零和博弈的概念。博弈理论中的一个非常重要的里程碑是 John Nash 提出的非合作博弈策略均衡概念[2]。一个博弈的纳什均衡是参与此博弈的个体的一个策略组合，这个策略组合的特点是没有任何一个博弈参与者会单方面背离当前的策略而选择其他的策略。纳什均衡的基本概念及其后续的相关研究构成了非合作博弈的理论基础。Poundstone 于 1992 年出版书籍《Prisoner's Dilemma》，其中给出了 Paul Walker 整理形成的关于早期博弈理论发展历史的详细列表[3]。

在众多的社会和经济组织中，个体都面临着自身利益和集体利益的选择冲突。这种选择冲突在博弈论中被称为社会困境（social dilemma）。比如，国家之间的贸易谈判、城市污染、知识共享、公共资源使用等都属于社会两难选择问题。如果个体在两难冲突中选择了集体利益，则称其采取了合作行为；相反，如果个体选择了自身利益，则称其采取了背叛行为。背叛行为的选择尽管给个体带来了自身经济利益的有效性，但必然导致集体经济利益的无效。博弈理论最为核心的任务就是为求解社会困境提供方法指导，并且解释在没有计划者和控制者的条件下，微观层面的自利个体通过相互作用如何产生宏观层面的合作以获得更加有效的社会结果。在长期的发展历程中，博弈论形成了两个主要的研究方法：基于完全理性个体假设的经典博弈理论和基于有限理性个体假设的演化博弈理论。

经典博弈理论关注于研究完全理性的博弈参与者为了获取博弈收益或效用的最大化应该如何进行行为决策。“完全理性”不仅要求博弈参与者始终以自身最大利益为目标，具有在确定和非确定性环境中追求自身利益最大化的判断和决策能力，还要求他们在博弈环境中具有完美的判断和预测能力；不仅要求参与者自身有完美的理性，还

要求参与者信任博弈对手的理性。但是，现实中的许多情况下，“完全理性”的概念并没有准确地表达真实个体的行为特征。事实上，现实社会中的真实个体在大多数比较复杂的决策问题中表现出来的理性都无法满足“完全理性”的要求。

演化博弈理论的基本假设是博弈参与者是有限理性的[4]。有限理性首先意味着博弈参与者往往不能或不会采用完全理性条件下的最优策略，其次意味着博弈方之间的策略均衡往往是学习调整的结果而不是一次性选择的结果，而且即使达到了均衡也可能再次偏离。因此，演化博弈理论适用于分析有限理性状态下的博弈参与者所组成的特定群体内成员间的反复博弈。例如大群体中成员之间随机配对的反复博弈，相当于现实经济中对象不固定的大量个体之间较长期的经济、交易关系，或者小群体中相邻个体的相互学习和某种形式的相互反应。

本章以经典博弈理论和演化博弈理论为线索，首先给出经典博弈论中的一些基本概念；其次介绍演化博弈理论的研究方法；最后总结一些基于博弈论的合作行为研究，其中的网络合作行为研究是复杂网络上博弈研究的基础。因此，本章的内容为后续讲解复杂网络上的博弈研究奠定了基础，并且有助于读者对合作问题的理解。

一、完全理性下的经典博弈论

1. 博弈的基本构成

任何一个博弈都包括以下三个组成部分：①至少两位决策者（decision - maker），或比赛者（player），或代理人（agent）参与博弈，可统称为博弈参与者；②博弈参与者有自己的博弈策略，所有参与者的策略构成策略空间；③博弈参与者按照相应的博弈规则进行博弈，获得收益。

在博弈过程中，博弈参与者以自身利益最大化为目标，即以此为指导原则进行策略选择。如果策略空间和收益函数对所有的参与者都是公开的，这个博弈就是完全信息博弈，否则被称作不完全信息博弈。如果博弈的参与者只能同时或者独立的选择策略而不是参考其他人的策略，则称为静态博弈。相反，如果博弈参与者的决策活动是依次选择行为而不是同时选择行为，而且后选择行为者能够看到先选择行为者的选择内容，则称为动态博弈。在静态博弈和动态博弈的基础上，重复博弈描述了静态博弈和动态博弈的反复进行过程，即基本博弈（静态博弈或动态博弈）重复两三次或其他有限的次数。下面仅以静态博弈说明一些重要的概念，其他详细内容可以参考文献[5]。

一个 2×2 标准形式的两人对称博弈通常可以用收益矩阵加以描述，如图 1－1 所

示。这个博弈有两个参与者，具有A和B两个策略。

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

图1-1　2×2对称博弈矩阵

参与者的收益依赖于自身的策略选择和对手的策略选择。比如，当参与者1（左侧参与者）选择A策略时，参与者2（上侧参与者）选择A策略（或B策略），参与者1的博弈收益为a（或b）；当参与者1选择B策略，参与者2选择A策略（或B策略），参与者1的博弈收益为c（或d）。那么，如果$a>c$并且$b>d$，我们说策略A是此博弈的占优策略，即任何一个参与者都会选择A策略。反过来，如果$a<c$并且$b<d$，策略B是占优策略。

通过这个简单的例子可以看出，博弈参与者的策略选择是针对其博弈对手策略的最佳对策，即获取自身博弈的最大收益。具有这种性质的策略组合是非合作博弈理论中最重要的一个解概念，即博弈中的"纳什均衡（Nash Equilibrium）"。

2. 纳什均衡

对于一般形式的博弈G，有n个博弈参与者，每个博弈参与者全部的可选策略集合称为"策略空间"，分别用S_1，…，S_n表示；$s_{ij}\in S_i$表示博弈参与者i的第j个策略，其中j可取有限个值（有限个策略），也可取无限个值（无限个策略）；博弈参与者i的收益用u_i表示，u_i是各博弈参与者策略的函数。将n个博弈参与者的博弈G表达为$G=\{S_1, \dots, S_n; u_1, \dots, u_n\}$，纳什均衡的定义如下：

在博弈$G=\{S_1, \dots, S_n; u_1, \dots, u_n\}$中，如果博弈参与者的策略组成的策略组合（$s_1^*, \dots, s_n^*$）中，任何一个参与者的策略$s_i^*$，都是对其余参与者策略组合（$s_1^*, \dots, s_{i-1}^*, s_{i+1}^*, \dots, s_n^*$）的最佳对策，即$u_i(s_1^*, \dots, s_{i-1}^*, s_i^*, s_{i+1}^*, \dots, s_n^*)\geqslant u_i(s_1^*, \dots, s_{i-1}^*, s_{ij}, s_{i+1}^*, \dots, s_n^*)$对任意$s_{ij}\in S_i$都成立，则称（$s_1^*, \dots, s_n^*$）为$G$的一个"纳什均衡"。

3. 博弈模型及其相互关系

在博弈论的发展过程中出现了非常多的博弈模型，无法在此一一陈述。下面介绍在复杂网络上的博弈研究中应用较广泛的三个博弈模型，并说明其相互关系。

1）囚徒困境模型（Prisoner's Dilemma Game） 有两个人犯了案子，被警察抓到。

这两个嫌疑犯被警察分别关在两个不同的屋子里审讯，互相不可能有任何的信息沟通。他们两个各自的律师却分别替他们做了这样一个分析：如果他们两个人都坦白了（即不合作策略或 D 策略），那么各判刑 5 年；如果两个人都抵赖（即合作策略或 C 策略），死不承认，可能因为警察及诉讼官的证据不足而只能以一个小罪名各判 1 年；如果其中一个坦白了，另一个抵赖，那么坦白的那个人可能因为被迫从犯且坦白从宽的原则而不作刑事起诉，仅作些口头教育便释放了，而那个选择抵赖的人依据抗拒从严的原则，将被判 8 年监禁。这里指的合作是两个嫌疑犯之间的合作，而不是嫌疑犯与警察的合作。根据以上描述，囚徒困境的博弈矩阵如图 1 – 2 所示。那么，每个嫌疑犯根据图 1 – 2 所给出的博弈收益会这样考虑自己的策略：如果对方选择不坦白（合作 C 策略），我选择坦白（背叛 D 策略）的收益为 0，比选择不坦白（合作 C 策略）的收益 –1 更好，因此我应该选择坦白（背叛 D 策略）；反过来，如果对方选择坦白（背叛 D 策略），我选择坦白（背叛 D 策略）的收益为 –5，以比选择不坦白（合作 C 策略）的收益 –8 更好，因此我应该选择坦白（背叛 D 策略）。因此，两个嫌疑犯都会选择坦白（背叛 D 策略），此博弈的纳什均衡解为（D，D）策略组合。

$$
\begin{array}{c c}
 & \begin{array}{cc} C & D \end{array} \\
\begin{array}{c} C \\ D \end{array} & \begin{pmatrix} -1 & -8 \\ 0 & -5 \end{pmatrix}
\end{array}.
$$

图 1 – 2　囚徒困境博弈矩阵

2）雪崩博弈模型（Snowdrift Game） 两个司机在暴风雪中被困于一个大雪崩的两侧，他们现在有两种选择：要么下车开始铲雪（合作策略 C），要么待在车上什么也不做（背叛策略 D）。如果两个司机都愿意下车铲雪，则两者因为得以顺利回家而人均得到数量为 b 的收益，共同承担铲雪所付出的劳动代价为 c，即每个人因为相互合作得到了报酬 $(b-c)/2$。如果两个司机都待在温暖的车上，则他们因不能按时回家而没有任何收益，即收益为 0。如果其中的一个下来铲雪，则两人也都能顺利回家，但是铲雪的司机（合作者）将独自承担铲雪的工作，因而得到的收益为 $b-c$，而待在车上不劳而获的司机（背叛者）没有付出劳动也回了家，故其获得最大收益 b。根据雪崩博弈模型描述的实际意义，其博弈矩阵如图 1 – 3 所示，其中 $b>c>0$。那么，每个司机根据图 1 – 3所给出的博弈收益会这样考虑自己的策略：如果对方选择铲雪（合作策略 C），我选择待在车上（背叛策略 D）的收益为 b，比选择铲雪（合作策略 C）的收益

$(b-c)/2$ 更好，因此我应该待在车上（背叛策略 D）；如果对方选择待在车上（背叛策略 D），我选择铲雪（合作策略 C）的收益为 $b-c$，比选择待在车上（背叛策略 D）的收益0 更好，因此我应该选择铲雪（合作策略 C）。因此，雪崩博弈的纳什均衡解有两个，一个是（D，C）策略组合；另一个是（C，D）策略组合，即对方选择合作策略时自己选择背叛策略，对方选择背叛策略时自己选择合作策略。

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} (b-c)/2 & b-c \\ b & 0 \end{pmatrix} \end{array}$$

图 1－3　雪崩博弈矩阵

3）牡鹿捕捉模型（Hunt Stag Game）牡鹿捕捉博弈起源于著名的法国思想家卢梭对狩猎者两难选择的描述。狩猎者有两个可选择的行为：合作捕捉牡鹿（合作策略 C）或单独捕捉兔子（背叛策略 D）。对于两个狩猎者来说，其博弈矩阵如图 1－4 所示。那么，每个狩猎者根据图 1－4 所给出的博弈收益会这样考虑自己的策略：如果对方选择合作捕捉牡鹿（合作 C 策略），我选择合作捕捉牡鹿（合作 C 策略）的收益为 5，比选择单独捕捉兔子（背叛 D 策略）的收益 2 更好，因此我应该选择合作捕捉牡鹿（合作 C 策略）；如果对方选择单独捕捉兔子（背叛 D 策略），我选择单独捕捉兔子（背叛 D 策略）的收益为 1，比选择合作捕捉牡鹿（合作 C 策略）的收益 0 更好，因此我应该选择单独捕捉兔子（背叛 D 策略）。因此，牡鹿捕捉博弈的纳什均衡解有两个，一个是（C，C）策略组合，共同捕捉牡鹿；另一个是（D，D）策略组合，各自捕捉兔子，即对方选择合作策略时自己选择合作策略，对方选择背叛策略时自己选择背叛策略。

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} 5 & 0 \\ 2 & 1 \end{pmatrix} \end{array}$$

图 1－4　牡鹿捕捉博弈矩阵

4）博弈模型之间的关系通过对前面三个不同博弈模型的基本描述可以看出，这三个博弈模型都具有典型的社会困境（Social Dilemma）的特点，但其纳什均衡解并不相同。社会困境可以用两人博弈的形式进行描述，博弈中有两个可以选择的策略——合

作（C）或背叛（D）。合作的含义包括诚实、坦白、互相帮助等符合人类道德标准的行为，背叛的含义包括撒谎、欺骗、剽窃等违背人类道德标准的行为。在两人博弈里，合作和背叛行为可能产生4种策略组合，即（C，C）、（D，D）、（C，D）、（D，C）。令 R 和 P 分别为两人共同合作和共同背叛时个体得到的收益，S 为（C，D）策略组合中选择 C 策略的个体收益，T 为（D，C）策略组合中选择 D 策略的个体收益，一个社会困境的博弈矩阵如图 1－5 所示。

$$
\begin{array}{cc}
 & \begin{array}{cc} C & D \end{array} \\
\begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix}
\end{array}
$$

图1－5　社会困境

在一个社会困境中，策略组合产生的收益关系必须满足以下4个条件。

（1）$R>P$。博弈个体偏好共同合作（C，C）而非共同背叛（D，D）。

（2）$R>S$。博弈个体偏好共同合作（C，C）而非单方面合作（C，D）。

（3）$2R>T+S$。共同合作（C，C）产生的集体收益大于单方面合作（C，D）或单方面背叛（D，C）所产生的集体收益。

（4）$T>R$ 或 $P>S$。博弈个体偏好单方面背叛（D，C）而非共同合作（C，C），或者博弈个体偏好共同的背叛（D，D）而非单方面合作（C，D）。$T>R$ 描述了个体的贪婪特性，$P>S$ 描述了个体对博弈对手选择背叛策略的担心或害怕。

这4个条件清晰地刻画出社会困境中个体利益和集体利益之间的冲突。这种冲突表现为：从集体的角度出发，共同合作（C，C）策略组合优于单方面合作（C，D）或单方面背叛（D，C）；从个体的角度出发，共同合作（C，C）优于共同背叛（D，D）。然而，在个体贪婪特性（$T>R$）或者对博弈对手选择背叛策略的担心（$P>S$）的作用下，个体选择背叛策略，从而产生了与集体利益相冲突的结果。

用社会困境的一般表达形式来描述前面讲述的三个模型，对比他们在不同策略组合下的博弈收益关系，囚徒困境模型的博弈收益关系可以表达为 $T>R>P>S$，雪崩博弈模型的博弈收益关系可以表达为 $T>R>S>P$，牡鹿捕捉模型的博弈收益关系可以表达为 $R>T>P>S$。可以看出，在牡鹿捕捉模型中，尽管共同合作的收益大于单方面背叛的收益，即 $R>T$，但由于 $P>S$ 使得个体担心博弈对手不和自己合作捕捉牡鹿，进而产生了（D，D）纳什均衡。在雪崩博弈模型中，则是由于 $T>R$ 使得自利个体的贪

婪特性发生作用，进而产生了（D，C）纳什均衡。在囚徒困境博弈模型中，由于 $T>R$ 和 $P>S$ 同时满足，进而产生了最为糟糕的情形，即产生了唯一的（D，D）纳什均衡。

二、有限理性和演化博弈论

1. 有限理性及其对博弈的影响

在静态博弈、动态博弈和重复博弈中，尽管少数地方会考虑博弈参与者的理性局限和犯错误的可能性，但都是以博弈参与者具有完全理性为基础的。完全理性包括理性意识、分析推理能力、识别判断能力和准确行为能力等多方面的完美性要求，其中任何一方面不完美就不是完全理性。这对于现实中的决策行为者而言是很难满足的高要求。当社会经济环境和决策问题比较复杂时，人们表现出来的往往是有限理性，而不是完全理性。

有限理性意味着博弈参与者没有能力一开始就找到最优策略，而是会在博弈过程中进行不断学习，通过试错的过程不断寻找更好的策略；还意味着均衡是不断调整和改进而不是一次性选择的结果，而且即使达到了均衡也可能再次偏离。在有限理性的前提下，博弈分析的核心不再是博弈方的最优策略选择，而是有限理性博弈参与者的学习和策略调整过程、趋势和稳定性。这里的稳定性是指博弈参与者采用特定策略的比例不变，而非某个博弈方的策略不变。这也是演化博弈理论与经典博弈理论的一个重要区别，即均衡意义的转变。在演化博弈理论里，演化稳定策略是一个非常重要的概念，与经典博弈理论中的纳什均衡概念紧密相关，其具体含义在后续相关内容中详细讲解。

由前面的分析可以看出，在有限理性的前提下，博弈分析的关键是确定博弈参与者的学习和策略调整模式。由于有限理性博弈参与者可能有很多理性层次，学习和策略调整的方式和速度也就不尽相同。下面介绍两种重要的学习和策略调整模式：最优反应动态和复制动态。

2. 最优反应动态

最优反应动态所描述的学习和策略调整模式为：博弈参与者能够针对不同的策略结果进行比较和评估，并相应地调整自己的策略。也就是说，给定前面的博弈结果，各个博弈参与者都能找到和采用针对前期其他博弈参与者策略的最佳反应策略。下面以图 1－4 给出的牡鹿捕捉博弈为例，说明最优反应动态的学习和策略调整模式，并解

释说明演化稳定均衡解的概念。

对于图 1-4 所示的牡鹿捕捉博弈，已经知道在二人静态博弈下存在两个纳什均衡解，即（C，C）策略组合和（D，D）策略组合。现在假定有 5 个人参与博弈，处于图 1-6 所示圆周上的 5 个位置，每个博弈参与者与各自的左右邻居进行博弈。假设 $x_i(t)$ 为在 t 时期博弈参与者 i 采用 D 策略的邻居数量，有 0，1，2 三个可能的值。相应地，采用 C 策略的邻居数量为 $2-x_i(t)$，也有 0，1，2 三个可能值。那么，针对第 t 期邻居的策略情况 $x_i(t)$，博弈参与者 i 采用 D 策略的收益为 $x_i(t)\cdot 1+[2-x_i(t)]\cdot 2$，采用 C 策略的收益为 $x_i(t)\cdot 0+[2-x_i(t)]\cdot 5$。根据最优反应动态机制，博弈参与者会比较不同策略所产生的收益结果，当 $x_i(t)\cdot 1+[2-x_i(t)]\cdot 2$ 大于 $x_i(t)\cdot 0+[2-x_i(t)]\cdot 5$，即 $x_i(t)>1.5$ 时（有两个以上的背叛邻居），博弈参与者 i 会在 $t+1$ 期采用 D 策略；相反，当 $x_i(t)<1.5$ 时（有一个或 0 个背叛邻居），博弈参与者 i 会在 $t+1$ 期采用 C 策略。

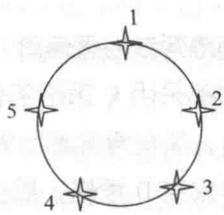

图 1-6　博弈参与者的位置分布

在有限理性的前提假设下，博弈参与者在初次进行博弈时可能采取 C 策略，也可能采取 D 策略。那么，初次博弈总共有 32 种可能的情况。这 32 种情况包括一种全部采用 C 策略和一种全部采用 D 策略的情况，其他都是两种策略均有人采用。以初次博弈仅有一个博弈方采用 C 策略为例，根据最优反应动态所描述的学习和策略调整规则，不难得出五个博弈参与者策略选择的全部过程。如图 1-7 所示，最终五个博弈参与者都选择了 C 策略。

通过对 32 种初次博弈情形进行分析，除了一种初次博弈全部采用 D 策略的情况外，其余 31 种在最优反应动态机制下都会收敛为全部采用 C 策略的状态。也就是说，所有博弈参与者都采用 C 策略和所有博弈参与者都采用 D 策略都是有限理性博弈参与者进行牡鹿捕捉博弈的均衡状态，但前一种状态更为重要，因为博弈方通过策略调整收敛到这种状态的可能性远远高于后一种状态。更为关键的是，如果达到所有博弈参

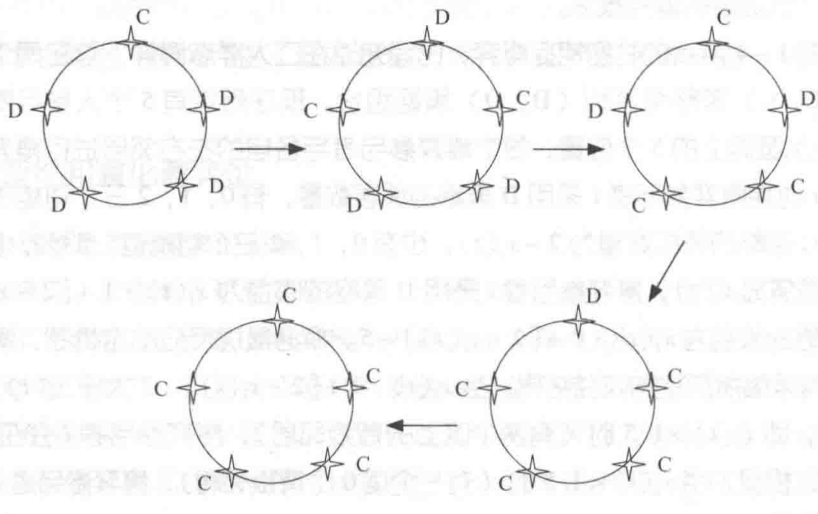

图 1－7　牡鹿捕捉博弈初次博弈为一个合作参与者的最优反应动态

与者都采用 C 策略的状态，即使有博弈参与者偏离 C 策略而选择 D 策略，最优反应动态会使博弈参与者的策略很快回到都采用 C 策略的状态。因此，所有参与者都采用 C 策略的均衡状态具有稳健性。相反，所有博弈参与者都采用 D 策略的均衡状态却不具有稳健性，因为一旦有博弈参与者偏离 D 策略，那么最优反应动态会使博弈参与者的策略离该均衡状态越来越远（如图 1－7 所示）。因此，在最优反应动态下，图 1－4 所描述的牡鹿捕捉博弈虽然存在两种稳定状态，但只有所有博弈参与者都采用 C 策略的均衡状态具有对偏离扰动的稳健性，这在演化博弈理论中被称为“演化稳定均衡解”。

3. 复制动态

前面的最优反应动态分析了少数几个博弈参与者之间反复博弈的策略调整和稳定性，复制动态则是用来分析数量较多的成员进行随机配对博弈时，有限理性博弈参与者的策略调整及稳定性。

假定数量较多的有限理性博弈参与者组成了一个大群体，参与者之间没有差异并进行随机配对博弈，即博弈参与者的博弈对手不像图 1－6 那样以某种形式固定下来，而是每次博弈可能有不同的博弈对手。在这样的假定下，博弈参与者根本不知道下一次的博弈对手是谁，最优反应动态的策略调整规则也就不再合理和适用。在大群体的随机配对博弈情形下，复制动态很好地描述了博弈参与者通过模仿学习博弈对手的策略逐渐达到策略的稳定性。下面以图 1－5 给出的社会困境一般博弈模型详细说明复制

动态的分析过程。

首先，有限理性的博弈参与者无法在一开始就找到最佳策略。一般地，假设整个群体中采取合作（C）策略的博弈参与者比例是 x，那么背叛（D）策略的博弈参与者比例是 $1-x$。群体中博弈参与者进行随机配对博弈，每个博弈参与者既可能遇到 C 策略类型的博弈对手，也可能遇到 D 策略类型的博弈对手。那么，在一个大群体里，忽略掉博弈参与者自身，每个博弈参与者遇到 C 策略类型博弈对手的概率是 x，遇到 D 策略类型博弈对手的概率是 $1-x$。群体里采用 C 策略和 D 策略的博弈参与者的期望收益 u_C 和 u_D 分别为

$$u_C = x \cdot R + (1-x) \cdot S \tag{1-1}$$

和

$$u_D = x \cdot T + (1-x) \cdot P \tag{1-2}$$

那么，整个群体的平均期望收益 $\bar{u}$ 为

$$\bar{u} = x \cdot u_C + (1-x) \cdot u_D \tag{1-3}$$

根据以上收益结果可以看出，不同策略选择的博弈参与者的博弈收益存在着差异，其大小不仅与博弈模型的收益参数 R，S，T，P 有关，还与群体中选择某种策略的博弈参与者的数量紧密相关。只要博弈方有基本的直觉和经验判断能力，必然会发现不同策略选择所产生的收益差异，收益较差类型的博弈参与者会很自然地开始改变策略，模仿收益较好类型博弈参与者的策略。这意味着两种类型博弈参与者的比例 x 和 $1-x$ 不是固定不变的，而是随时间的变化而变化。采用 C 策略的博弈参与者的比例变化可以用动态微分方程进行描述，即

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x(u_C - \bar{u}) \tag{1-4}$$

这个动态微分方程就是复制动态方程，将 u_C 和 $\bar{u}$ 代入上述方程，则有

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x(1-x)(Rx - Sx - Tx + Px + S - P) \tag{1-5}$$

对于给定的收益参数 R，S，T，P，$\frac{\mathrm{d}x}{\mathrm{d}t}$是 x 的函数，可以将复制动态方程记为$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x)$。令 $F(x)=0$，可以解出比例 x 不再改变的状态，即合作者比例的稳定状态为 $x_1=0$，$x_2=1$，$x_3=\frac{P-S}{R-T+P-S}(0 \leqslant x_3 \leqslant 1)$。当 x_3 的值与 x_1 或 x_2 的值相同时，则只存在两个稳定状态。