第 37 卷 第 3 期 西南师范大学学报(自然科学版) 2012 年 3 月
Vol. 37 No. 3 Journal of Southwest China Normal University (Natural Science Edition) Mar. 2012

文章编号:1000-5471(2012)03-0128-05

基于无限重复博弈的管理决策建模分析[①]

赖 水 香

赣南师范学院 商学院,江西 赣州 341000

摘要:分析了无限重复博弈成功应用的条件,并以此结论建立博弈模型,确定收益矩阵参数,该模型在实际应用中能得到合理的解释,其博弈结果是双赢的博弈平衡.

关 键 词:重复博弈;Nash 均衡;囚徒困境;建模;条件策略

中图分类号:O225 **文献标志码**:A

传统的博弈论定义为"研究决策主体的行为发生直接相互作用时候的决策以及这种决策的均衡问题,因此又称对策论",而 Aumann 定义的博弈论概念则强调理性主体在决策时必须考虑到对方的反应,在 Aumann 的理论体系中,博弈论即"交互的决策论"[1]. 在 Aumann 的重复博弈理论中,理性行为个体重复同样结构的博弈,不但博弈的个体之间是交互的,博弈的不同阶段也是相互依赖的. 理性行为个体的决策不仅受其过去经历的影响,其决策还要受未来可能的影响. 重复博弈中,博弈者不但可以观察到其他博弈者的行动,还可以观察到以前的博弈结果,理性个体由此选择自己的最佳策略. 在重复博弈中,博弈者有更多的机会学习调整自己的决策,由此能更好地避免"囚徒困境".

1 引入条件策略的企业竞争博弈

1.1 企业竞争博弈模型

先讨论两个竞争企业的博弈模型[2]. 在该模型中,收益矩阵能直接从实际应用中得到.

假定 A,B 两个企业共同竞争一个特定市场. 两个企业均发现,在其它条件不变的情况下增加广告宣传费用将提高本企业的市场份额,由此带来更高的利润,反之,如果对方企业增加了广告宣传费而本企业保持不变,则本企业市场份额将下降. 为简化模型,假定企业的广告费投入有高和低两种选择,则按照利润计算,两家公司的收益矩阵如表 1,表中向量的第 1 个分量为公司 A 的利润,第 2 个分量为公司 B 的利润.

表 1 企业竞争博弈收益矩阵

公司 A	公司 B	
	高	低
高	(4, 4)	(6, 3)
低	(3, 6)	(5, 5)

从收益矩阵上看,对两个公司来说,最好的局面是自己高广告投入而对方低广告投入,最坏的局面则是两个公司都采用高广告投入,而两个公司都采用低广告投入则能实现双赢.

很不幸,如果该博弈只进行一次,则两公司(高广告投入,高广告投入)的博弈状态是唯一的 Nash 均衡,也就是说在一次博弈中,双方都采用高广告投入是唯一的理性选择. 这是一个典型的"囚徒困境"式博弈. 为了实现双赢的目标,采用无限重复博弈,并引入条件策略.

1.2 引入两条条件策略

条件策略是指博弈者的行动建立在其他博弈者的行动之上,由此,当一个博弈者做出偏离帕累托效率的选择时,其他博弈者可以采用条件策略予以惩罚. 合理的条件策略有助于帕累托改善,达到双赢(多赢).

---

① 收稿日期:2011-11-09
作者简介:赖水香(1977-),女,江西安远人,讲师,硕士,主要从事管理博弈论与教育心理学的研究.

在上述模型中，拟引入两条条件策略，即公司 A,B 均声称：

条件策略①：如果我采用低广告投入而你采用高广告投入，则我下一年转为高广告投入并将一直保持高广告投入.

条件策略②：如果你采用低广告投入，则我也采用低广告投入；

条件策略必须是可信的，才会对决策结果产生影响. 策略①显然是可信的，因为它与（高广告投入，高广告投入）的 Nash 均衡一致，Nash 均衡策略总是可以信任的策略.

策略②的可信性比较复杂，在一次博弈中，策略②是不可信的. 因为当对方采用低广告投入时，自己采用高广告投入比低广告投入收益大，此时选择高广告投入才是理性的.

在无限重复博弈中，策略②的可信性需进一步分析.

首先假定未来收益是贬值的. 设 δ（$0<\delta<1$）为贴现率，由表 1 的收益矩阵，企业 A 在维持（低广告投入，低广告投入）状态时的收益 U_A 为

$$U_A(\text{低，低}) = 5 + 5\delta + 5\delta^2 + 5\delta^3 + \cdots\cdots$$

这是一个等比数列，因为 $0<\delta<1$，所以 $U_A(\text{低，低}) = 5/(1-\delta)$.

假设第一年 A 企业转到高广告投入，则第一年两企业为(高广告投入，低广告投入)状态；第二年开始，由于企业 B 也会理性地选择高广告投入，两企业转为(高广告投入，高广告投入)的 Nash 均衡状态并保持下去，所以有：

$$U_A(\text{高，低}) = 6 + 4\delta + 4\delta^2 + 4\delta^3 + \ldots\ldots = 6 + 4\delta/(1-\delta)$$

由计算可得，当贴现率 $\delta>1/2$ 时，$U_A(\text{低，低})>U_A(\text{高，低})$，此时，维持自己的低广告投入为理性选择.

经过上面的分析，当贴现率 $\delta\geqslant 1/2$ 时，条件策略②也是可信的. 所以在无限重复博弈中，企业 A,B 可以采用两种决策：

（Ⅰ）采用低广告投入并保持，用条件策略①警告对方；

（Ⅱ）采用低广告投入并保持，向对方承诺条件策略②.

1.3　关于理性与人的行为模式的问题

在实际应用中，人的决策过程会受到心理和行为模式的影响，所以人的决策过程不可能是纯理性的，这也一直是博弈论应用的难题. 我们建模时需要考虑到人的心理和行为模式. 上面的模型分析发现的一个问题是：条件策略①很容易让人相信，而条件策略②则需经过较复杂的分析过程才能确定可信. 越是简单的道理越容易被人接受，这是一个常识. 在企业竞争博弈模型中，选择条件策略时要考虑到博弈者是否有足够的耐心去分析该条件策略的可信性，所以采用决策(Ⅰ)比采用决策(Ⅱ)合理.

2　参数待定的辅导员学生管理博弈

在企业竞争博弈模型中，收益矩阵是可以直接得到的. 但是在很多应用案例中，都没有这样的直接数据. 在下面这个辅导员学生管理博弈模型中，收益矩阵的值是待定的. 一旦采用数学方法确定了这个收益矩阵，则意味着制度的制定者可以联系收益矩阵的值来修改奖惩的措施和力度.

2.1　辅导员学生管理博弈模型的初步设定

和其它管理博弈过程一样，辅导员和学生之间也存在类似的博弈过程. 在这个模型中，辅导员有两种纯策略，即高强度管理(严厉)、低强度管理(活泼)，学生也有两种纯策略，即高违纪率(散漫)、低违纪率(自觉). 将高强度管理和高违纪率策略用 H 表示，将低强度管理和低违纪率策略用 L 表示. 我们采用与企业竞争博弈模型相似的模型.

这个模型的合理性在于：辅导员收益可以理解为学生整体利益得到维护以及辅导员个人得到奖惩，高强度管理可以带来较好的辅导员收益，但同时也意味着辅导员时间精力的高投入；学生收益可以理解为学生的各种个体利益，高违纪在一定程度上张扬了学生个人的个性，但是也意味着侵害了其他同学的利益和辅导员的利益. 表 2 是辅导员学生管理博弈模型的收益矩阵.

表 2　辅导员学生管理博弈收益矩阵

学生	辅导员	
	高强度管理	低强度管理
高违纪	(a_2, b_2)	(a_4, b_1)
低违纪	(a_1, b_4)	(a_3, b_3)

借鉴企业竞争博弈模型，令 $a_1<a_2<a_3<a_4$，其合理性可解释为：

1) 因为学生总是倾向于张扬自己的个性和个体自由，所以在同等管理强度时，定义学生高违纪的收益比低违纪的收益要高，即 $a_1<a_2$，$a_3<a_4$.

2) 双赢总比双输好，所以定义学生在(高违纪，高强度管理)的收益比(低违纪，低强度管理)低，即 $a_2<a_3$.

类似的道理，令 $b_1<b_2<b_3<b_4$.

与企业竞争博弈模型类似，如果该博弈只进行一次，则（高违纪，高强度管理）是唯一的 Nash 均衡，这是一种“囚徒困境”. 从管理角度来看，(高违纪，高强度管理）的结果是最差的，所以本文建立一个无限重复博弈模型以期得到一个更成功的博弈平衡.

2.2 辅导员学生管理博弈模型的条件策略

在实际社会生活中，管理者常制订一些包含奖惩条款的管理制度，这些管理制度对被管理者的行为做了规范. 从博弈论的角度来看，这些管理制度实际上起到了条件策略作用. 而被管理者也常常会采用提意见、背后议论、发牢骚等方式向管理者发出一些带有警示性质的反馈信息，这些反馈信息也起到了条件策略作用.

在辅导员学生管理博弈模型中，我们将辅导员和学生各自宣示的信息简化为下面的条件策略.

条件策略③：辅导员声称如果学生出现高违纪的情况则在下一阶段马上采用高强度管理并将一直保持.

条件策略④：辅导员声称如果学生低违纪则自己也采用低强度管理.

条件策略⑤：学生声称如果辅导员采用高强度管理则在下一阶段以高违纪形式抗议并将一直保持.

条件策略⑥：学生声称如果辅导员采用低强度管理则自己将低违纪.

条件策略③,⑤均为可信，因为它们与(高违纪，高强度管理)的 Nash 均衡一致.

条件策略④,⑥的可信性在 2.3 分析.

2.3 博弈分析

表 2 的辅导员学生管理博弈收益矩阵中参数是待定的，我们希望确定这些参数，使得无限重复博弈的结果为(低违纪，低强度管理). 为了达到这个结果，我们希望达到下面的目标：

1) 在(低违纪，低强度管理)博弈状态，学生维持低违纪.

2) 在(高违纪，高强度管理)博弈状态，学生转变为低违纪.

在后面的分析中，我们可以证明，满足上面两个要求的参数，即能使无限重复博弈的结果为（低违纪，低强度管理).

设 S 表示学生，T 表示辅导员，u 为违纪程度(取值为高和低，分别用 H 和 L 表示)，v 为管理强度(取值为高和低，分别用 H 和 L 表示). $S(u,v)$ 和 $T(u,v)$分别表示学生和辅导员在(u,v)状态下的长期收益.

在重复博弈中，未来收益总是贬值的. 在辅导员学生管理博弈模型中，用 δ $(0<\delta<1)$表示贴现率，贴现率越高表示对未来的收益越关注. 只要 δ 满足 $0<\delta<1$，δ 的具体取值不影响理论分析.

先分析如何达到目标 1). 显然，如果能保证学生在(低违纪，低强度管理)状态下的长期收益高于(高违纪，低强度管理)下的长期收益，学生的理性选择将是低违纪.

由表 2 的辅导员学生管理博弈收益矩阵，δ 表示贴现率，则在无限重复博弈中，(低违纪，低强度管理)状态下学生的长期收益为：

$$S(\mathrm{L},\mathrm{L})=a_3+\delta\cdot a_3+\delta^2\cdot a_3+\delta^3\cdot a_3+\cdots=a_3/(1-\delta)$$

由条件策略③，前一个阶段的（高违纪，低强度管理）状态在接下来的阶段会转化为（高违纪，高强度管理)，所以有：

$$S(\mathrm{H},\mathrm{L})=a_4+\delta\cdot a_2+\delta^2\cdot a_2+\delta^3\cdot a_2+\cdots=a_4+\delta\cdot a_2/(1-\delta)$$

满足 $S(\mathrm{L},\mathrm{L})>S(\mathrm{H},\mathrm{L})$时，在(低违纪，低强度管理)状态下学生将维持低违纪，目标 1)实现，即：

$$(S(\mathrm{L},\mathrm{L})>S(\mathrm{H},\mathrm{L}))\Leftrightarrow(a_3/(1-\delta)>a_4+\delta\cdot a_2/(1-\delta))\Leftrightarrow(a_4-a_3<\delta\cdot(a_4-a_2)) \tag{1}$$

这样便得到如下结论 1.

结论 1 当式(1)成立时条件策略⑥是可信的.

同样的道理分析辅导员的收益，如果能保证辅导员在(低违纪，低强度管理)状态下的长期收益高于(低违纪，高强度管理)下的长期收益，辅导员的理性选择将是低强度管理.

由表 2 的收益矩阵，能得到：

$$T(\mathrm{L},\mathrm{L})=b_3+\delta\cdot b_3+\delta^2\cdot b_3+\delta^3\cdot b_3+\cdots=b_3/(1-\delta)$$

$$T(\mathrm{L},\mathrm{H})=b_4+\delta\cdot b_2+\delta^2\cdot b_2+\delta^3\cdot b_2+\cdots=b_4+\delta\cdot b_2/(1-\delta)$$

显然当 $T(\mathrm{L},\mathrm{L})>T(\mathrm{L},\mathrm{H})$时，条件策略④是可信的，即：

$$(T(\mathrm{L},\mathrm{L})>T(\mathrm{L},\mathrm{H}))\Leftrightarrow(b_3/(1-\delta)>b_4+\delta\cdot b_2/(1-\delta))\Leftrightarrow(b_4-b_3<\delta\cdot(b_4-b_2)) \tag{2}$$

这样便可得如下结论 2.

结论 2　当式(2)成立时条件策略④是可信的.

再分析如何达到目标 2). 显然，如果能保证学生在(低违纪，高强度管理)状态下的长期收益高于(高违纪，高强度管理)下的收益，学生的理性选择将是低违纪.

由于策略④是可信的，(低违纪，高强度管理) 状态在下一阶段将转为(低违纪，低强度管理)状态并保持，所以(低违纪，高强度管理)状态下学生长期收益 $S(\mathrm{L},\mathrm{H})$为：

$$S(\mathrm{L},\mathrm{H})=a_1+\delta\cdot a_3+\delta^2\cdot a_3+\delta^3\cdot a_3+\cdots=a_1+\delta\cdot a_3/(1-\delta)$$

在(高违纪，高强度管理)状态下，学生长期收益 $S(\mathrm{H},\mathrm{H})$为：

$$S(\mathrm{H},\mathrm{H})=a_2+\delta\cdot a_2+\delta^2\cdot a_2+\delta^3\cdot a_2+\cdots=a_2/(1-\delta)$$

满足 $S(\mathrm{H},\mathrm{H})<S(\mathrm{L},\mathrm{H})$时，学生在(高违纪，高强度管理)状态时在下一阶段转为低违纪. 目标(2)实现，即：

$$(S(\mathrm{H},\mathrm{H})<S(\mathrm{L},\mathrm{H}))\Leftrightarrow(a_2/(1-\delta)<a_1+\delta\cdot a_3/(1-\delta))\Leftrightarrow(a_2-a_1<\delta\cdot(a_3-a_1)) \tag{3}$$

这样便得到如下结论 3.

结论 3　满足式(3)，则学生在(高违纪，高强度管理) 状态时在下一阶段转为低违纪.

为了得到具体的收益矩阵，取双输、双赢时双方的收益 a_2,a_3 为基准值，设 $a_2=3$，$a_3=5$(a_2,a_3 可联系对学生的奖惩来设定)，并假设贴现率 $\delta=0.5$(δ 的具体取值不影响分析过程). 由式(1)，

$$a_4-a_3<\delta\cdot(a_4-a_2)\Rightarrow$$
$$a_4-\delta\cdot a_4<a_3-\delta\cdot a_2\Rightarrow$$
$$a_4<(a_3-\delta\cdot a_2)/(1-\delta)$$

即

$$a_4<7 \tag{4}$$

同理，由式(3)

$$a_2-a_1<\delta\cdot(a_3-a_1)\Rightarrow$$
$$a_1-\delta\cdot a_1>a_2-\delta\cdot a_3\Rightarrow$$
$$a_1>(a_2-\delta\cdot a_3)/(1-\delta)$$

即

$$a_1>1 \tag{5}$$

由式(4)，取 $a_4=6$；再由式(5)，取 $a_1=2$. 当 $\delta=0.5$ 时，$a_1=2$，$a_2=3$，$a_3=5$，$a_4=6$ 就是满足要求的取值.

同样的道理，设 $b_2=4$，$b_3=6$ (b_2,b_3 可联系对辅导员的绩效考核设定)，有

$$b_4<(b_3-\delta\cdot b_2)/(1-\delta)\Rightarrow b_4<8 \tag{6}$$
$$b_1<(b_2-\delta\cdot b_3)/(1-\delta)\Rightarrow b_1>2 \tag{7}$$

由式(6)，取 $b_4=7$；再由式(7)，取 $b_1=3$. 当 $\delta=0.5$ 时，$b_1=3$，$b_2=4$，$b_3=6$，$b_4=7$ 就是满足要求的取值. 于是得到表 3.

表 3　辅导员学生管理博弈收益矩阵

学生	辅导员	
	高强度管理	低强度管理
高违纪	(3，4)	(6，3)
低违纪	(2，7)	(5，6)

2.4　博弈结论证明

由前面的分析，现在要证明如下结论.

结论 4　基于表 3 的博弈收益矩阵经无限重复博弈后总能达到(低违纪，低强度管理)的双赢状态.

证

1) 当博弈状态为(高违纪，高强度管理)时，由结论 3，有 $S(\mathrm{H},\mathrm{H})<S(\mathrm{L},\mathrm{H})$，则学生将在下一阶段选择低违纪，由此在下一阶段进入到 (低违纪，高强度管理) 状态.

2)当博弈状态为(低违纪，高强度管理)状态时，因为条件策略④可信(结论 2)，所以辅导员在下一阶段采用低强度管理，由此在下一阶段进入到(低违纪，低强度管理)状态.

3)当博弈状态为(高违纪，低强度管理)时，因为条件策略⑥可信(结论 1)，学生将在下一阶段转化为低违纪，由此在下一阶段进入到(低违纪，低强度管理)状态.

4)当博弈状态为(低违纪，低强度管理)时，因为条件策略④可信(结论 2)和条件策略⑥可信(结论 1)，所以(低违纪，低强度管理)将维持.

2.5 进一步分析

表 3 中，a_4 为学生在(高违纪，低强度管理)下的收益，由式(4)的结论，a_4 取较小值有助于博弈状态向(低违纪，低强度管理)转化. 在实际应用中，辅导员可以采用个别惩罚的方式降低高违纪学生的收益.

a_1 为学生在(低违纪，高强度管理)下的收益，由式(5)的结论，a_1 取较大值有助于博弈状态向(低违纪，低强度管理)转化. 在实际应用中，辅导员可以采用个别奖励的方式提高低违纪学生的收益.

辅导员学生管理博弈模型中，贴现率体现了对未来收益的关注，贴现率越高表示对未来的收益越关注. 由前面对企业竞争博弈模型的分析，对于确定值的收益矩阵，无限重复博弈能达到双赢的条件是贴现率要大于一定的数值，应用到现实中就是辅导员要引导学生多关注未来的收益，放眼未来. 辅导员自己也要放眼于长远发展，做好职业规划，兢兢业业，百年育人.

3 结　　论

重复博弈理论是 Aumann 最有代表性的成果，在博弈论发展史上也具有里程碑式的意义. 本文分析的企业竞争博弈模型和辅导员学生博弈模型都是具有广泛代表性的管理模型，重复博弈理论在其中的成功应用也说明这一理论具有很强的实用性.

参考文献：

[1] 朱艳丽，杨定华. 奥曼及其无限重复博弈理论——2005 年诺贝尔经济学奖述评 [J]. 云南财经大学学报，2005，20(6)：12－13.

[2] 格若赫姆·罗伯. 博弈论导引及其应用 [M]. 柯华庆，闫静怡，译. 北京：中国政法大学出版社，2005.

[3] 严太华，刘　贞，任玉珑. 基于 Mealy 自动机的多人重复博弈演化模型及仿真分析 [J]. 管理工程学报，2010，24(1)：158－162.

[4] 王春永. 博弈论的诡计 II：日常生活中的博弈策略(插图本) [M]. 北京：中国发展出版社，2009.

[5] MEHMET B，GUILLHERME C，HAMID S. Repeated Games with One-Memory [J]. Journal of Economic Theory，2008，144(1)：312－336.

[6] CHONG J K，COLIN F. A Learning-Based Model of Repeated Games with Incomplete Information [J]. Games and Economic Behavior，2006，55(2)：340－371.

[7] ELIOT M. Adaptation and Complexity in Repeated Games [J]. Games and Economic Behavior，2008，63(1)：166－187.

Modeling Analysis of Management Decision Based on Infinite Repeated Game

LAI Shui-xiang

Business College，Gannan Normal University，Ganzhou，Jiangxi 341000，China

Abstract：In a single game，patients are prone to fall in prisoner dilemma of Nash equilibrium. In a repeated game，however，patients can choose the best strategy according to their opponent′s actions and previous game results，and have more opportunities to avoid prisoner dilemma. In this paper，the limit conditions of successfully applying infinite repeated game theory are analyzed and obtained. Based on this，income matrix parameters are confirmed and a game model is established. A rational explanation is given for applying this game model in practice. The game result of this model is a win-win balance.

Key words：repeated game；Nash equilibrium；prisoner dilemma；modeling；trigger strategy

责任编辑　张　栒