实验课 8：博弈论实验-重复囚徒困境

一、实验大纲

1. 博弈论理论背景
2. 囚徒困境博弈
3. 重复囚徒困境实验

二、博弈论理论背景

博弈论(Game Theory)，有时也称为对策论，或者赛局理论，是研究具有对抗或竞争性质现象的理论和方法，它是应用数学的一个分支，既是现代数学的一个新分支，也是运筹学的一个重要学科。目前在生物学、经济学、国际关系学、计算机科学、政治学、军事战略和其他很多学科都有广泛的应用。主要研究公式化了的激励结构（游戏或者博弈（Game））间的相互作用，是研究具有对抗或竞争性质现象的数学理论和方法，也是运筹学的一个重要学科。

具有竞争或对抗性质的行为成为**博弈行为**。在这类行为中，参加斗争或竞争的各方各自具有不同的目标或利益。为了达到各自的目标和利益，各方必须考虑对手的各种可能的行动方案，并力图选取对自己最为有利或最为合理的方案。比如日常生活中的下棋，打牌等。博弈论就是研究博弈行为中斗争各方是否存在着最合理的行为方案，以及如何找到这个合理的行为方案的数学理论和方法。

博弈论的基本概念

(1)局中人（players）：在一场竞赛或博弈中，每一个有决策权的参与者成为一个局中人。只有两个局中人的博弈现象称为”两人博弈”，而多于两个局中人的博弈称为 “多人博弈”。

(2)策略(strategiges)：一局博弈中，每个局中人都有选择实际可行的完整的行动方案，即方案不是某阶段的行动方案，而是指导整个行动的一个方案，一个局中人的一个可行的自始至终全局筹划的一个行动方案，称为这个局中人的一个策略。如果在一个博弈中局中人都总共有有限个策略，则称为”有限博弈”，否则称为”无限博弈”。

(3)得失(payoffs)：一局博弈结局时的结果称为得失。每个局中人在一局博弈结束时的得失，不仅与该局中人自身所选择的策略有关，而且与全局中人所取定的一组策略有关。所以，一局博弈结束时每个局中人的”得失”是全体局中人所取定的一组策略的函数，通常称为支付（payoff）函数。

(4)次序（orders）：各博弈方的决策有先后之分，且一个博弈方要作不止一次的决策选择，就出现了次序问题；其他要素相同次序不同，博弈就不同。

(5)博弈涉及到均衡：均衡是平衡的意思，在经济学中，均衡意即相关量处于稳定值。在供求关系中，某一商品市场如果在某一价格下，想以此价格买此商品的人均能买到，而想卖的人均能卖出，此时我们就说，该商品的供求达到了均衡。所谓纳什均衡，它是一稳定的博弈结果。

目前经济学家们现在所谈的博弈论一般是指非合作博弈，由于合作博弈论比非合作博弈论复杂，在理论上的成熟度远远不如非合作博弈论。非合作博弈又分为：完全信息静态博弈，完全信息动态博弈，不完全信息静态博弈，不完全信息动态博弈。与上述四种博弈相对应的均衡概念为：纳什均衡(Nash equilibrium)，子博弈精炼纳什均衡（subgame perfect Nash equilibrium），贝叶斯纳什均衡(Bayesian Nash equilibrium)，精炼贝叶斯纳什均衡(perfect Bayesian Nash equilibrium)。

纳什均衡，又称为非合作赛局平衡，是博弈论的一个重要概念，以约翰·纳什命名。纳什均衡(Nash Equilibrium)：在一策略组合中，所有的参与者面临这样一种情况，当其他人不改变策略时，他此时的策略是最好的。也就是说，此时如果他改变策略他的支付将会降低。在纳什均衡点上，每一个理性的参与者都不会有单独改变策略的冲动。纳什均衡点存在性证明的前提是”博弈均衡偶”概念的提出。所谓”均衡偶”是在二人零和博弈中，当局中人 A 采取其最优策略 a*，局中人 B 也采取其最优策略 b*，如果局中人仍采取 b*，而局中人 A 却采取另一种策略 a，那么局中人 A 的支付不会超过他采取原来的策略 a*的支付。这一结果对局中人 B 亦是如此。

如果某情况下无一参与者可以独自行动而增加收益，则此策略组合被称为纳什均衡点。

这样，”均衡偶”的明确定义为：一对策略 a*(属于策略集 A)和策略 b*（属于策略集 B）称之为均衡偶，对任一策略 a(属于策略集 A)和策略 b（属于策略集 B)，总有：偶对（a， b*）≤偶对(a*，b*)≥偶对（a*，b)。

对于非零和博弈也有如下定义：一对策略 a*（属于策略集 A）和策略 b*（属于策略集 B）称为非零和博弈的均衡偶，对任一策略 a(属于策略集 A）和策略 b（属于策略集 B)，总有：对局中人 A 的偶对（a， b*） ≤偶对(a*，b*);对局中人 B 的偶对（a*，b）≤偶对(a*，b*)。

有了上述定义，就立即得到纳什定理： 任何具有有限纯策略的二人博弈至少有一个均衡偶。这一均衡偶就称为纳什均衡点。如果某情况下无一参与者可以独自行动而增加收益，则此策略组合被称为纳什均衡点。

三、囚徒困境博弈

囚徒困境是博弈论的非零和博弈中具代表性的例子，反映个人最佳选择并非团体最佳选择。虽然困境本身只属模型性质，但现实中的价格竞争、环境保护等方面，也会频繁出现类似情况。

1. 囚徒困境模型

1950 年，由就职于兰德公司的梅里尔·弗勒德（Merrill Flood）和梅尔文·德雷希尔（Melvin Dresher）拟定出相关困境的理论，后来由顾问艾伯特·塔克（Albert Tucker）以囚徒方式阐述，并命名为”囚徒困境”。经典的囚徒困境如下：

警方逮捕甲、乙两名嫌疑犯，但没有足够证据指控二人入罪。于是警方分开囚禁嫌疑犯，分别和二人见面，并向双方提供以下相同的选择：

若一人认罪并作证检控对方（相关术语称”背叛”对方），而对方保持沉默，此人将即时获释，沉默者将判监 10 年。

若二人都保持沉默（相关术语称互相”合作”），则二人同样判监半年。

若二人都互相检举（互相"背叛"），则二人同样判监 2 年。

用表格概述如下：

	甲沉默	甲认罪
乙沉默	二人同服刑半年	乙服刑 10 年，甲即时获释
乙认罪	甲服刑 10 年，乙即时获释	二人同服刑 2 年

囚徒困境假定每个参与者（囚徒）都是利己的，即都寻求最大自身利益，而不关心另一参与者的利益。参与者某一策略所得利益，如果在任何情况下都比其它策略要低的话，此策略称为”严格劣势”，理性的参与者绝不会选择。另外，没有任何其它力量干预个人决策，参与者可完全按照自己意愿选择策略。

囚徒到底应该选择哪一项策略，才能将自己个人的刑期缩至最短？两名囚徒由于隔绝监禁，并不知道对方选择；而即使他们能交谈，还是未必能够尽信对方不会反口。就个人的理性选择而言，检举背叛对方所得刑期，总比沉默要来得低。试设想困境中两名理性囚徒会如何作出选择：

- 若对方沉默、我背叛会让我获释，所以会选择背叛。
- 若对方背叛指控我，我也要指控对方才能得到较低的刑期，所以也是会选择背叛。
- 二人面对的情况一样，所以二人的理性思考都会得出相同的结论——选择背叛。背叛是两种策略之中的支配性策略。因此，这场博弈中唯一可能达到的纳什均衡，就是双方参与者都背叛对方，结果二人同样服刑 2 年。

这场博弈的纳什均衡，显然不是顾及团体利益的帕累托最优解决方案。以全体利益而言，如果两个参与者都合作保持沉默，两人都只会被判刑半年，总体利益更高，结果也比两人背叛对方、判刑 2 年的情况较佳。但根据以上假设，二人均为理性的个人，且只追求自己个人利益。均衡状况会是两个囚徒都选择背叛，结果二人判监均比合作为高，总体利益较合作为低。这就是”困境”所在。

上述例子可能显得不甚自然，但现实中，无论是人类社会或大自然都可以找到类似囚徒困境的例子，将结果划成同样的支付矩阵。社会科学中的经济学、政治学和社会学，以及自然科学的动物行动学、进化生物学等学科，都可以用囚徒困境分析，模拟生物面对无止境的囚徒困境博弈。囚徒困境可以广为使用，说明这种博弈的重要性。

2. 重复囚徒困境-合作机制

一种特殊的博弈，在博弈中，相同结构的博弈重复多次，甚至无限次。其中，每次博弈称为”阶段博弈”。在每个阶段博弈中，参与人可能同时行动，也可能不同时行动。因为其他参与人过去的行动的历史是可以观测的，因此在重复博弈中，每个参与人可以使自己在每个阶段选择的策略依赖于其他参与人过去的行为。重复博弈具有三个基本特征：

- 重复博弈的阶段，博弈之间没有”物质”上的联系，即前一个阶段博弈并不改变后一个阶段的博弈的结构；
- 再重复博弈的每一个阶段，所有参与人都观测到该博弈过去的历史；

- 参与人的总收益是所有阶段博弈的收益的贴现值之和或加权平均数。

在重复博弈中，参与人存在着短期利益和长远利益的均衡，有可能为了长远利益牺牲短期利益而选择不同的均衡策略，因此，重复博弈为现实中的许多合作行为和社会规范提供了解释。

罗伯特·阿克塞尔罗德在其著作《合作行为的发生》(Axelrod ，1984)中，探索了经典囚徒困境情景的一个扩展，并把它称作“重复的囚徒困境”（Iterated Prisoner’s Dilemma，IPD）。阿克塞罗德是密歇根大学政治系教授，他以下面的试验为基础提出了重复的囚徒困境：他邀请了不同学科的学者参加一个竞赛，参赛者需要与他人玩一个多次博弈的游戏，每一个参赛者在每次博弈时可选择与对方合作或者不合作。在二人博弈的情况下，每一次游戏结果的得失矩阵与经典囚徒困境基本相同。不同的是在这个例子中，囚徒 A 和囚徒 B 需改成参赛者 A 和 B ，不坦白须改成合作，坦白须改成背叛。由于阿克塞罗德把哈丁的 n 人一次性模型改成了一个 2 人 m 次模型，所以参赛者可以根据自己不同的知识背景来设计整个博弈过程的行为策略。有人可能在 m 次博弈中每次都选择合作，有人可能会每次都选择背叛，有人甚至会让计算机帮他(她) 随机选择策略，不一而足。阿克塞罗德想知道比赛结束后，采取哪一类策略的人总得分最高。结果是，当博弈次数很大(或者说博弈链很长) 时，总得分最高的策略并不是我以上所列举的那些策略，而是一种被称为“一报还一报”(tit-for-tat) 的策略。这种策略很简单，就是第一次总是采取合作，以后就跟着对方的上一次策略走。如果对方上一次采取的是合作，那么我这一次仍然采取合作;如果对方上一次采取的是不合作，那么我这一次就采取不合作。可以看出， “一报还一报”的策略与做一个有原则的好人的原则差不多，即做人首先要与人为善，如果对方对你做了坏事你绝不能姑息养奸;但对方一旦承认错误，你就应该原谅他。

得到这一结果后，阿克塞罗德把它公布了出来，并邀请各学科的学者再次参加同样一个比赛。不同的是，这一次大家都知道“一报还一报”的策略在上一次比赛中获得了第一名，于是专门针对这种策略去设计策略，就像打擂台一样，想办法打败它。尽管如此，第二次比赛的结果还是“一报还一报”策略的总得分最高。于是阿克塞罗德就下了这样一个结论:由于“一报还一报”这一策略显然是一个最优策略，因此，当博弈次数很大时，只要博弈的一方有理性学习能力，或者说虽然博弈的一方没有理性学习能力，但其“总得分”的高低能够转化成某种进化压力，迫使其行为产生“突变”的话，那么，在多次博弈过程中，博弈各方的行为就会趋同，即逐渐在博弈中采取合作这一策略。

阿克塞尔罗德发现，当这些对抗被每个选择不同策略的参与者一再重复了很长时间之后，从利己的角度来判断，最终"贪婪"策略趋向于减少，而比较"利他"策略更多地被采用。他用这个博弈来说明，通过自然选择，一种利他行为的机制可能从最初纯粹的自私机制进化而来。

通过分析高分策略，阿克塞尔罗德指出了策略获得成功的几个必要条件。

友善

最重要的条件是策略必须"友善"，这就是说，不要在对手背叛之前先背叛。几乎所有的高分策略都是友善的。因此，完全自私的策略仅仅出于自私的原因，也永远不会首先打击其对手。

报复

但是，阿克斯洛德主张，成功的策略必须不是一个盲目乐观者。要始终报复。一个非报复策略的例子是始终合作。这是一个非常糟糕的选择，因为"下流"策略将残酷地剥削这样的傻瓜。

宽恕

成功策略的另一个品质是必须要宽恕。虽然它们不报复，但是如果对手不继续背叛，它们会一再退却到合作。这停止了报复和反报复的长期进行，最大化了得分点数。

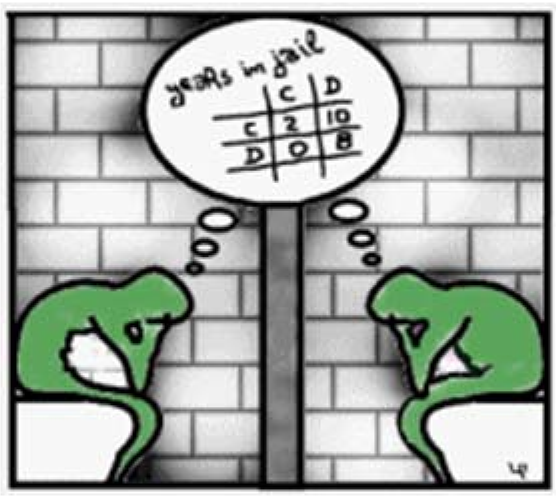
不嫉妒

最后一个品质是不嫉妒，就是说不去争取得到高于对手的分数（对于"友善"的策略来说这也是不可能的，也就是说"友善"的策略永远无法得到高于对手的分数）。

阿克塞尔罗德关于重复囚徒困境的研究的重要结论是用于解释人类合作的行为机制。

由于人类社会的基础就是各类合作性的集体行动，因此阿克塞罗德认为，他的这个实验揭示了人类合作行为得以形成和发展的基本动力。阿克塞罗德认为，只要博弈的次数足够多，即使是敌对的双方也会走向合作。一个比较典型的例子是第一次世界大战时欧洲战场上的壕沟战。那时壕沟挖得比较浅，在壕沟中走路时如果不小心把头抬起来就会遭到敌方士兵的冷枪。战争开始时，法、德两国的士兵互相用这一方法打死了对方许多人。但是到战争后期，双方士兵的行为都起了变化:如果一方的士兵头抬得太高，对方就朝他头的上方打一枪警告一下，而不是把他打死。其中的逻辑是，这次如果我把人家打死了，下次我不经意间把头露出战壕时，人家也会把我打死。按照这一逻辑，双方士兵在敌对的场合下达成了实际的合作。阿克塞罗德强调，他的模拟结果不但能够用来解释人类社会中合作现象的产生，而且能够用于解释生物的合作性行为的产生。的确，生物学中有一个重要现象，即协同进化。其中一种协同进化现象是，生物之间的捕食与被捕食关系在进化过程中可能发生如下演变:从捕食与被捕食关系(A 消灭了 B) ，逐渐转变为寄生与被寄生关系(A 仍然容许 B 的存活) ，并最后转化为共生关系(A 和 B 互相为对方提供增进存活的条件) 。

四、重复囚徒困境实验



4.1 实验目的

本次实验为综合性实验，要求基于LSL脚本语言，通过重复囚徒困境策略竞赛实验，掌握游戏中NPC自动合作的机制，实现具有主动合作意识的智能机器人或NPC。

4.2 实验场景

本次试验场景是把经典的囚徒困境改成 2 人多次的博弈策略竞赛。首先，把经典的囚徒困境策略得失矩阵转化为一个得分矩阵，具体如下：

得分矩阵		B 的选择	
		合作（不招认）	不合作（招认）
A 的选择	合作（不招认）	双方好的结果 2 个人各的 3 分	对 B 有利的结果 A 得 0 分，B 得 5 分
	不合作（招认）	对 A 有利的结果 A 得 5 分，B 得 0 分	双方坏的结果 2 个人各的 1 分

在实验中，会给出 5 种博弈策略供 Agent A 和 Agent B 进行选择，一旦 2 个 Agent 选定了各自的博弈策略，就开始了多次重复博弈。每一轮博弈会进行 200 次迭代，每一次迭代 2 个 Agent 各自所获得的分数会进行累积，因此博弈竞赛的目标就是看何种策略能获得尽量多的分数。每二种策略的竞赛做五轮，然后取每轮得分的平均值来进行最后比较。

而且为了能够更好的对比 5 种博弈策略，可以让 5 种博弈策略两两进行循环赛。

重复囚徒困境博弈策略：　　C-合作，D-不合作

- 策略 1 总是选择不合作

以 A 为例，无论 B 做出什么选择，A 总选择和 B 不合作。

Prisoner A using All D	D D D D D D D D DD D D D D
Prisoner B	C C D D C C D C C C D D C C

- 策略 2 总是选择合作

以 A 为例，无论 B 做出什么选择，A 总选择和 B 合作。

Prisoner A using All C	C C C C C C C C C C C C C C
Prisoner B	C C D D C C D C C C D D C C

- 策略 3 Tit For Tat

从合作开始，之后每次选择对方前一阶段的行动，以牙还牙，以眼还眼
以 A 为例，第一次 A 的选择是和 B 合作，接下来，若在本轮中 B 选择的是合作，那么 A 在下一轮的选择是合作，若在本轮中 B 选择不合作，那么 A 在下一轮的选择是不合作。

Prisoner A using TFT	C C C D C D C C C C C D D C
Prisoner B	C C D C D C C C C C D D C D

- 策略 4 Tit For Two Tats

以 A 为例，A 的选择总是合作，一旦 B 连续 2 次出现不合作，则 A 也开始选择不合作，直到 B 不在出现连续的不合作为止

Prisoner A using TF2T	C C C C D D C C C C C D D D
Prisoner B	C C D D C C D C C C D D D D

- 策略 5 随机选择合作或者不合作

概率设为 50%，以 A 为例，每次 A 有 50%的概率选择合作或者不合作

Prisoner A using Random	C D C C D D C C D C D C C D
Prisoner B	C C D D C C D C C C D D D D

4.3 实验步骤

（1）实验场景建模

分别创建上述 5 种策略的 LSL 脚本。

（2）初始化 Agent

建立 2 个 Agent，分别选定一种博弈策略，开始比赛。可选 llDialog 方式，以对话框方

式进行 Agent 策略对话。

（3）重复的囚徒困境实验

a.每 2 种博弈策略进行循环赛

b.每次比赛共进行五轮，每轮进行 200 次迭代的囚徒困境博弈，得分矩阵如上表，累积五轮分数求平均值

c.最后给出一个总的结果表

策略名称	ALL-D	ALL-C	TFT	TF2T	Random
ALL-D		0：0	0：0	0：0	0：0
ALL-C			0：0	0：0	0：0
TFT				0：0	0：0
TF2T					0：0
Random					

4.4 扩展实验：

投资游戏：

- 瑞士苏黎世大学和圣加仑大学的科学家设计了一种投资游戏：每个小组有 4 名匿名的参与者，他们自行决定要往一个公共的罐子里“投资”多少钱。一个小组的总投资所获得的“利润”，是在 4 名参与者之间平均分配的。对精明的人来说，最有利的方法是一分钱也不投入，坐享其成。但如果参与者多投钱，就能增加小组的总收益。
- 每一轮游戏结束后，每个人的投资额对本小组成员公开。在一些游戏里，允许人们对本组中那些混水摸鱼者处以罚款，但其他成员也要为罚款机制付出一点钱，也就是说，人们并不能从惩罚别人的行为中得到直接的物质利益。有的游戏里则不允许罚款机制存在。
- 下一轮游戏中，参与者重新组合，保证每个人的新一轮合作者全都是陌生人，不能用上一轮的经验来判断其它成员是否可以信任。这样一轮一轮地进行下去，结果发现，在允许罚款的游戏里，合作进行得很好，投资额不断增长；而如果不允许惩罚，投资额就迅速减少，很快就进行不下去了。
- 游戏结果表明，如果允许人们惩罚不劳而获者，合作就能成功地进行下去，即使大家要为设立惩罚机制付出一定代价，人们也能够接受。而如果不允许惩罚机制存在，合作很快就会崩溃。这很好解释，即在特定情况下，人们追求公正的渴望能够压倒自私的想法，平均主义或许也是人类行为的基本特征之一，公众愿意为设立惩罚机制付出代价。希望不劳而获的人由于害怕被罚而不敢轻举妄动，而诚实的合作者则因为拥有惩罚的权力而获得安全感。

趣味游戏：合作心理实验-选美游戏

游戏说明：创建几个选美化身形象，由 Avatar 用户和 Agent 共同投票选择心目中的美女。如果化身 Avatar 的选择和大多数 Agent 一致，可加分。投票共进行 3 轮，每轮的规则可能不同。请化身每轮仔细阅读说明，进行投票。第 1 轮投票不显示任何投票信息，化身和 Agent 可自由投票；第 2 轮可显示本轮投票结果，在第 2 轮化身和 Agent 可根据投票统计情况，改变各自的投票意见；第 3 轮，除显示本轮投票结果外，还显示上一轮的投票结果和大家的真实偏好，然后化身改变投票。游戏结束，显示各游戏方的成绩统计。


选美博弈
请选择您心目中的美女
霍青桐
黄蓉
阿朱
周芷若
确定并进入

示例情境：

第 1 轮：


选美博弈（第一轮
此处显示游戏的进度，第几轮，每一轮您能看到的信息会不同！
选美博弈
现在是第 1轮投票
帮助
关于
退出
在游戏公告里会有本轮相关信息，请认真阅读！
阿朱
第一轮投票不显示任何提示信息
投票统计
霍青桐
黄 蓉
阿 朱
周芷若
Neway
游戏公告
请您参加选美博弈！
如果您的选择和多数投票一致，您的课程成绩将会获得加分！
投票将进行三轮！每一轮规则不一样！
请注意每一轮的游戏公告！

第 2 轮：


选美博弈（第二轮）
这里是本轮结束的倒计时，在每一次页面刷新时更新！
选美博弈
现在是第 2轮投票 记时： 3:0
帮助
关于
退出
霍青桐
黄蓉
阿朱
周芷若
投票统计
霍青桐 0
黄 蓉 1
阿 朱 0
周芷若 0
请您投票
改投
您投给了黄蓉！但您仍可改投！
游戏公告
本轮将即时显示投票结果！在五分钟之内，您可以根据显示的结果改变您的投票！结果将以您最后的投票为准！
在这一轮里，5分钟之内可以根据投票统计，随时改变自己的投票！
投票统计显示本轮投票的统计结果！

第 3 轮：


选美博弈（第三轮）
选美博弈
现在是第 3轮投票
帮助
关于
退出
霍青桐
黄蓉
阿朱
周芷若
投票统计
霍青桐 1
黄 蓉 0
阿 朱 0
周芷若 0
请您投票
改投
您投给了霍青桐！但您仍可改投！
Neway
游戏公告
上一轮投票结果：
霍青铜 0 黄 蓉 1
阿 朱 0 周芷若 0
所有人的真实偏好：
霍青铜 1 黄 蓉 0
阿 朱 0 周芷若 0
本轮还显示上一轮的投票结果和大家的真实偏好！

参考文献：

[1]. 罗伯特•阿克塞尔罗德和威廉•汉米尔顿，《合作的进化》（”The Evolution of Cooperation”），出自《科学》（Science），1981 年，第 211 期：1390-1396 页。

[2]. 罗伯特•阿克塞尔罗德，《合作的进化》（The Evolution of Cooperation），Basic Books 出版社，ISBN 0465021212。

[3]. 罗伯特•阿克塞尔罗德，《合作的复杂性》（The Complexity of Cooperation），普林斯顿大学出版社，1997 年，ISBN 0691015678。

[4]. 理察•道金斯，《自私的基因》（The Selfish Gene），第二版——有两章关于合作的进化，1990 年，ISBN 0192860925。

[5]. 格罗夫曼和普尔，《重复囚徒困境博弈的贝叶斯模型》（”Bayesian Models for Iterated Prisoner's Dilemma Games”），1975 年，出自《一般系统》（General Systems），第 20 期：185--194 页。

[6]. 加勒特•哈丁，《公共地悲剧》（”The Tragedy of the Commons”），出自《科学》（Science），1968 年，第 162 期：1243-1248 页。

[7]. 科勒普斯、戴维、罗伯特•威尔逊、保罗•米尔格罗姆和约翰•罗伯茨，《有限重复囚徒困境中的理性合作》（”Rational Cooperation in the Finitely Repeated Prisoners' Dilemma”），出自《经济理论杂志》（Journal of Economic Theory），1982 年，第 27（2）期：245-252 页。

[8]. 保罗•米尔格罗姆，《阿克斯洛德的合作的进化》（”Axelrod's The Evolution of Cooperation”），出自《兰德经济学杂志》（Rand Journal of Economics），1984 年，第 15（2）期：30-59 页。

[9]. 威廉•庞德斯通，《囚徒困境：约翰•冯诺依曼、博弈论和炸弹之谜》（Prisoner's Dilemma: John von Neumann， Game Theory， and the Puzzle of the Bomb），Doubleday 出版社，1992 年，ISBN 0385415672。如标题所指出的，这是一个广泛流行的入门介绍。

[10]. 阿纳托尔•拉波波特和阿尔伯特•查马哈，《囚徒困境》（Prisoner's Dilemma），密西根大学出版社，1965 年。此书解释了许多实验，实验中进行了囚徒困境心理博弈。

[11]. 汤姆•费尔霍夫，《交易者困境：囚徒困境的连续版本》（”The Trader's Dilemma: A Continuous Version of the Prisoner's Dilemma”），《计算器科学讲义》（Computing Science Notes）93/02，1998 年，荷兰爱因霍芬科技大学，数学和计算器科学系。