



## 重复博弈策略：做好人，还是做坏人？



高黑  
Shanghai

关注他

271 人赞同了该文章

最近电视剧《虎啸龙吟》正在热播，大家在讨论司马懿黑化的问题，而在第一部《军师联盟》上映时，一堆人却在讨论司马懿是不是被洗白了，司马懿究竟是好人，还是坏人？

其实，人性是复杂的、变化的，讨论道德上的好人、坏人，是一个主观价值判断的问题，意义不大。但是我们如果仅仅看这个人的外在行为，选择好人策略，还是选择坏人策略，却有着强烈的现实意义。

这个问题用在职场上，可以表述为，你是更愿意和大家分享，还是一亩三分地把得牢牢的？你是选择广结善缘、一团和气，还是坚决站队、恩怨分明？你是相信强者恒强、赢家通吃，还是合作共赢、利益均摊？作为领导，你是选择当鹰派还是鸽派？

用在商场上，可以表述为，你的企业是倾向于遵守规则，还是打破规则？你是更倾向于选择蓝海，还是红海？是追求利润，还是坚决打价格战？和竞争对手是选择共生、分而治之，还是你死我活，势不两立？

甚至还可以表述为，你选择做慈母，还是虎妈。将来是好学生有出息，还是捣蛋鬼？

关于这个问题，从很多方面都可以研究，而科学家们利用计算机进行的博弈论模拟，在我看来，是最精彩的。

### 重复博弈中的策略选择

《自私的基因》写于1976年，道金斯（Richard Dawkins）在里面记载了阿克塞尔罗德（Robert Axelord）利用计算机进行的一系列关于囚徒困境的模拟比赛，读起来惊心动魄、荡气回肠。有兴趣的朋友可以读读原著，我这里就只说一下结果。



我的出牌		你的出牌	
		合作	背叛
	合作	比较好 相互合作的奖赏 300 美元	很坏 失败的代价 100 美元罚款
	背叛	很好 背叛的诱惑 500 美元	比较坏 相互背叛的惩罚 10 美元罚款

囚徒困境说的是，明明有合作双赢的选择，但是博弈的结果是都选择背叛，这是一个让所有人沮丧，但又无法辩驳的结果（具体论证不展开了）。

原来以为零和博弈，要选择损人利己，做坏人，这么说，非零和博弈，也要做坏人？

然而，这个博弈还有另一个版本：“重复博弈”的“囚徒困境”，而这其实更接近现实世界。由于不是一局定胜负，重复博弈则可以有了很多可能性，可以试探性的选择合作，看看对方的反应，不打不相识，慢慢摸索出双方的偏好和底线。单次博弈里只有两种策略，合作或是背叛，而背叛是唯一理性的，重复博弈的策略组合则基本取决于想象力了。那究竟什么样的策略会最优呢？

阿克塞尔罗德组织进行了三轮计算机模拟比赛。

第一轮

阿克塞尔罗德向全世界的博弈论的专家们征集了14个策略，他还加了第15个策略，取名为“随机”。阿克塞尔罗德将这15个策略翻译成一种常用的计算机语言，在一个大型计算机中设定这些策略互相博弈。每个策略轮流与其他策略（包括它自己）进行重复博弈。15个策略总共组成15×15=225个排列组合，在计算机上轮番进行。每一个组合需要进行200回合的博弈，所有输赢累积计算，以得出最终的赢家 -- 看看谁赢得最多的“钱”。

结果如下：

- 1、本来大家以为胜出者会是一个绝顶聪明的策略，没想到却是一个看起来最不聪明，而且超级简单的策略：**以牙还牙 (Tit for Tat)**。这个策略在第一回合时采取合作行动，然后在接下来的所有步骤里，只是简单复制对手上一步的行动，一报还一报，你合作，我合作，你背叛，我背叛，你悔改，我也悔改。
- 2、其它的策略，包括冷酷策略（一旦发现对方背叛，永不再合作）、下马威策略（从背叛开始）、不定期主动背叛策略（时不时的作恶一下）、永远背叛策略、永远合作策略等，都排在后面。
- 3、总体来说，偏“善良”类的策略，也就是不主动背叛，其表现优于“恶意”类的策略，后者倾向于主动背叛；“宽容”类的策略表现优于“刻薄”类的策略，后者失去了再次合作的可能性。
- 4、最差的策略是随机，其次是一个超级复杂的策略，没人看得懂它想干什么（当时大家嘲笑说，这一定是来自美国国防部，或者CIA）。
- 5、阿克塞尔罗德后来自己还想出了第16个策略，宽容版的以牙还牙 -- 两牙还一牙，你背叛两次，我才报复一次。这个策略表现更优，甚至超过了以牙还牙。

第二轮

这个乌托邦式的结论，让人有些惊讶。只是重复博弈，就一下子从丛林法则进入美好新世界了？

于是，阿克塞尔罗德又举办了第二次竞赛。他把上述所有结果都公之于众，然后再次征集策略。这次他收到了62个策略，再加上随机策略，总共63个。这一次，博弈中的回合数不再固定为200，而改为开放式的不定数。

结果如下：

- 1、很惊讶，两牙还一牙针对宽容版的
- ▲ 赞同 271 ▼
- 31 条评论
- 分享
- ♥ 喜欢
- ★ 收藏
- ...



计出了三报还一报的超级宽容策略，但也没有成为最终的胜利者。

2、所以，马上就可以有第二个结论，你的策略优劣和对手的出牌是相关的，也就是说外部环境很重要。

3、善意策略再次战胜了恶意策略，前15名中只有一个是恶意策略，而倒数15名中只有一个善意策略。

4、一旦比赛次数为固定次数，则倾向于单次博弈，永远背叛策略会获胜，如果比赛次数是不固定、或者无限次，那就是以牙还牙等善良策略胜出。

### 第三轮

作为生物学家，道金斯更关心的是这个模拟怎么才最接近自然选择。于是在他建议下，阿克塞尔罗德又模仿自然选择，进行了第三场竞赛。这次，他并没有邀请新的参赛者，而只是使用了第二次竞赛中的63个策略，但是赢家不再得到“金钱”或者“分数”，而是与其完全相同的“后代”，这样世世代代如此传递下去。

结果如下：

1、那些恶意策略，有一些开始蓬勃发展，迅速崛起，比如永远背叛、小偷、骗子，但它们的繁荣只是昙花一现，大部分后来被以牙还牙群殴，像极了好莱坞的英雄片，坏人终被惩罚。大部分恶意策略都在150代之内就灭绝了，唯一活过200代的一个策略叫做“哈灵顿”（Harrington），它靠欺负老实人获利，但随着老实人绝种，它也在1000代之后终于完全灭绝。

2、以牙还牙在6次模拟中有5次得了第一，重复其在第一、二次竞赛时的运气。另外5个虽善良但也会报复的策略几乎和以牙还牙一样成功（在种群数目上），还有一个策略甚至赢了第6次博弈。

3、当所有恶意策略都灭绝后，到了1000代形成了稳态，所有的善良策略与以牙还牙都无法辨认彼此了，所有人都只是简单的合作到底。

4、在一个被永远背叛策略控制的群体，任何其它策略都没法独立存活，但是只要很少数量的以牙还牙策略者聚集在一起，就会形成合作，并慢慢变大，进而超越永远背叛者，哪怕这个时间很长，很长。正义可能会迟到，但从来不会缺席。

说到这里，似乎已经有了结论，利他和合作可以从自私中诞生出来，并一步步壮大。最终我们会越来越善良，好人终将有回报。

果真如此吗？

### 第四轮

很多科学家都对重复博弈的模拟感兴趣，诺瓦克（Martin A. Nowak）就是其中之一。10几年后，他进行了新的实验。诺瓦克发现了阿克塞尔罗德的一个问题，就是现实世界不是计算机，所有的信息不可能这么精准无误，存在着噪声，比如对方可能不是主动背叛，而是发出了一个错误信号，或者被误解了，以牙还牙对于这种情况没有容错机制，而是直接报复，如果对手也是类似策略，则容易陷入本不需要的缠斗，从而血海深仇。与恶龙缠斗，最终变成恶龙。

诺瓦克进行了一次新的比赛，还是模拟自然选择。这次他针对对方的背叛，设计了几种不同的应对，有的是永不合作，有的是以牙还牙，有的是两报还一报，还有三报，乃至五报，但选择报复还是合作却是随机的，不是确定的。

结果出来了：

1、以牙还牙没能再次成为最终的赢家。

2、宽容版以牙还牙成为了最终的赢家，具体是在三次背叛中会选择一次进行合作，但究竟是哪次是随机的。可以理解为三报还两报。或者说，**永远不会忘记对方的好意，偶尔原谅对方的恶意。**

3、其它结果和以前类似，恶意策略开始占优，但后来被善意策略超越，最终大家都变得面目模糊，简单的合作。

▲ 赞同 271 ▼ ● 31 条评论 ➤ 分享 ♥ 喜欢 ★ 收藏 ...



这个精彩的研究成果发表于1992年的英国《自然》杂志，值得一提的是前面讲的阿克塞尔罗德的  
研究成果发表在了1981年的美国《科学》杂志上，都是科学界顶级的权威刊物。

诺瓦克发现，这里还有个问题，不管是以牙还牙，还是三报还两报，最后善意策略一统江湖，再演化下去，就会变成永远合作者主导天下。这时，一旦有一个来自外部的永远背叛的恶意策略，就可以摧枯拉朽般的横扫天下，迅速扩张。都是合作者的群体无法抵抗恶意入侵。这变成了一个循环，从永远背叛，到以牙还牙，到宽容版的以牙还牙，再到永远合作，最后又是永远背叛。

这倒是像极了盛极而衰的人类社会和各种组织。有没有什么办法破解吗？

## 第五轮

诺瓦克在后续的模拟中，无意间发现了一个更优的策略 -- **赢定输移 (Win stay, Lose shift)**。

这个策略描述起来既有些复杂，其实又很简单：

如果我们在上一轮中都合作，那么我会再次合作；如果都背叛，那么我会（以某一概率）进行合作；如果你合作，我背叛，那么我会继续背叛；如果你背叛，我合作，那么我也会背叛。

感觉越说越复杂，其实这个策略最简单的描述是：**如果我现在收益良好，我就会重复以前的策略，如果收益不好，我就改变策略。我的策略是否改变不取决于对手什么策略，而取决于我自己的收益。**

这个策略和以牙还牙相以及三报还两报比，在我合作对方背叛上是一样的，都是选择报复，所以不会在和恶意策略搏斗中吃亏。但是当双方都背叛时，以牙还牙选择的是缠斗到底，三报还两报是随机的试探性合作，但是赢定输移策略更倾向于迅速化干戈为玉帛，只在少数情况下才继续报复。不过真正的不同在于，一旦出现对方合作我背叛的时候，这个策略会继续坏下去，不会因为对方合作而选择合作。

这样，这个策略就具备了有两个优势，第一，不会陷入缠斗，一旦双方都是背叛，下一轮随机性的选择合作，既能威慑对方，又能迅速的收手。第二，一旦出现对方合作我背叛的局面，装傻，让子弹飞一会儿，让利益最大化。

实战中，这个策略比以牙还牙和宽容版的以牙还牙活得更长，在善意策略一统江湖后继续存活很久，它能够在不被报复的情况下对合作者进行盘剥。而这个时候，一旦有恶意者入侵时，它会立马组织起报复，从而最大限度的延长了种群的寿命。

没有永远的策略，在进化中，赢定输移也会消失，但是它是在目前的模拟中发现的最能延长种群寿命的策略之一。

## 人类是超级合作者

上面讲的几位科学家其实都是生物学家，他们研究博弈论，是为了更好的解释自然界的进化论现象，尤其是利他是怎么从自私中产生的，弱肉强食、适者生存的条件下，为什么会出现合作。重复博弈里的以牙还牙、赢定输移等策略，在动物界都得到了验证，这说明实验和现实是吻合的。

达尔文在发表进化论时就说过，一个种群内如果有合作出现，这个种群就会变得比全是自私者的种群更加繁荣，从而会在种群之间的竞争中胜出。这个现象在人类社会里最为典型。

《人类简史》告诉我们，我们的直接祖先是一种叫智人的人，大约出现在15万年前，而在同期，地球上至少存在这6种历史更悠久的人类，其中就包括著名的尼安德特人。但是，之后的几万里，其它人类都消失了，只有智人活了下来，并走到了几乎地球上每一个角落，同时消失的还有许多大型的哺乳类动物，比如猛犸象等。究竟是什么不一样的特质，让智人战胜了所有的动物和其他人类，走到了今天呢？

显然不是人们通常意义上以为的用火和制造工具，因为其他人类也会，而且早在几十万年前就学会了。也显然不是自私自利、损人利己，这些连动物都会。赫拉利告诉我们是合作，是超级合作。

社会学领域有一个邓巴数的概念，说的是熟人社交的群体很难突破150人的规模。由于语言和抽象思维能力的突破

▲ 赞同 271 ▼ ● 31 条评论 ➤ 分享 ♥ 喜欢 ★ 收藏 ...

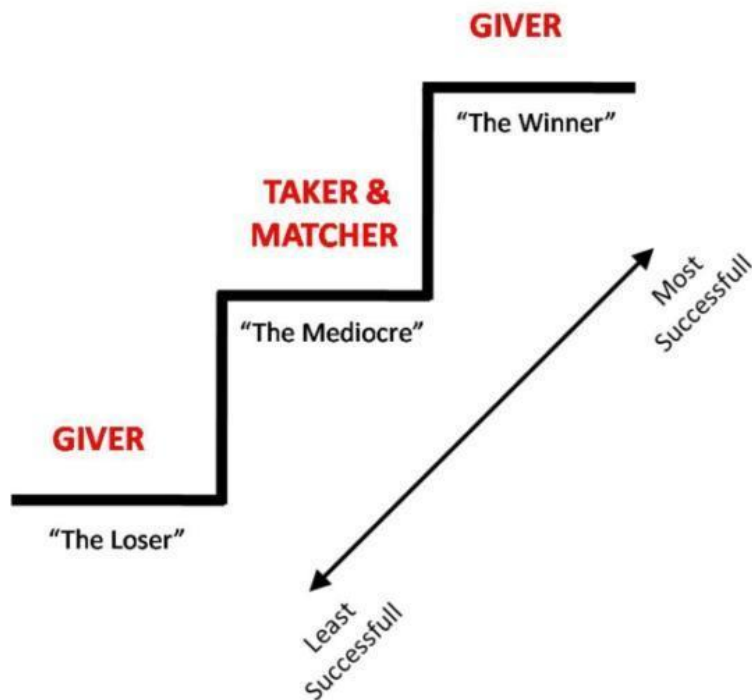


在远古时代，群体之间的战争，以及对大型哺乳动物的捕猎，规模起着决定性的作用。

看看我们今天的企业、军队、国家，到处都是巨型组织。沃尔玛是人数最多的企业，全球有230万人。无数重要的技术研发项目，都是在几百、数千甚至上万人的跨地区的合作中完成的。毫不夸张的说，是超级合作能力帮助人类走到了今天，并且，未来人类的合作能力，只会越来越强。

这是我们的基因决定的，这就是自然选择的力量。所以，如果你发现自己是个好人，那恭喜你，说明你进化的比较好。

## Giver or Taker?



上面讲的所有东西，不是计算机模拟，就是定性的推理，虽然有道理，但是有没有实际数据支撑呢？

这个问题其实一直困扰着学术界，人们花了很多资源来研究手机的蓝光对睡眠有没有影响，碳水化合物是不是比肉类跟容易让人发胖，但是在我们究竟“要不要做一个好人”，或者“为什么很多成功者都是坏人”这些更重要的问题上，却几乎没有什么统计数据。

这个现象在2013年得到了改观，亚当·格兰特（Adam Grant）在他的畅销书《沃顿商学院最受欢迎的成功课》中把人分为三类：给予者（Giver），索取者（Taker）和互利者（Matcher）：给予者在工作中倾向于奉献、服从、服务；索取者相反，一切按我想要的办；互利者则会视对方的态度而选择自己的策略。

他做了一个规模达到3万人的调查，涵盖了各行各业和不同的岗位。以下是一部分结论：

- 1、在最低级别的岗位中，给予者们表现最差。Giver = Loser。
- 2、虽然如此，但是有众多Giver的团队，则具备更多的服务意识，其整体表现反而更好。所以，对于组织来说，需要考虑的是如何保护给予者。
- 3、Taker往往会在一个新的环境中，短期内占据优势，但是时间一长，劣势就清楚了他们的败子就会知道如何对

▲ 赞同 271 ▼ ● 31 条评论 ➤ 分享 ♥ 喜欢 ★ 收藏 ...



4、中层中Taker和Matcher居多，但是到了高层，情况发生了逆转，大部分成功人士反而是Giver。好人终究有了好报。

5、如果你打算爬到最高层，那么你的最好策略，是做一个Disagreeable Givers。你本质上是一个给予者，但你并不好打交道，你有原则，有选择，不会无条件的提供帮助。

## 总结

说了这么多，那么究竟什么策略是最优的生存策略呢？究竟应该做好人，还是做坏人呢？至少以下几点值得借鉴：

- 1、最差的策略是不清晰的策略，要么过于随机，要么过于复杂，因为别人无法对你的行为和反应形成稳定的预期。
- 2、当老好人不行。对于对方的恶意，你必须有清晰的对策，最简单的是一报还一报，该出手时就出手。稍微复杂点的策略也可以，但是不能过于复杂，要让对方很容易的判断的出你的出牌。
- 3、要先表示善意，不要上来就做坏人。先下手为强不适用与多数情况，你有很大的风险立马遭受报复，从而一开始就受损，且释放了不友好的合作信号。
- 4、尽可能制造重复博弈的机会，不要陷入单次博弈，不要陷入零和博弈的局面。社会一直在发展，一直在进步，没那么多事只有一次机会，需要一锤定音。
- 5、做坏人是需要实力的。如果你实在没有这个资本，或者没有这个能力，就尽可能选择一个良好的环境吧。
- 6、最后，要和自己比，不要和别人比。你的目标是让自己更好，而不是把别人比下去。只有采用这种真正理性的策略，才能让自己的利益最大化。不要因为别人收益比自己大而做坏人。不过说起来容易，做起来难。

最好的策略是，**我能够做坏人，但是我可以选择做还是不做，选择的标准是我自己能否变得更好。**

## 参考书籍和文章

《自私的基因》，理查德·道金斯

《超级合作者》，马丁·诺瓦克，罗杰·海菲尔德

Ted演讲 Are you a giver or a taker? 亚当·格兰特

Why it pays to be jerk, Jerry Useem, 《大西洋月刊》2015年6月

编辑于 2018-01-12

[博弈论](#) [职场](#) [进化](#)

## 推荐阅读

▲ 赞同 271 ▼ 31 条评论 分享 喜欢 收藏 ...