

利他之谜及其博弈论分析

The Altruism Puzzle and an Analysis on the Game Theory

刘鹤玲

(华中师范大学科学史与科学哲学研究中心, 副教授 武汉 430070)

利他现象的普遍存在

达尔文进化论确立以后,人们特别关注生物进化中的生存竞争和适者生存。但是,生物乃至社会集团、民族和国家之间即使在利益部分冲突的情况下也存在着利他现象。为了研究这一现象,1964年,伦敦大学学院的生物学家汉密尔顿创立亲缘选择理论。1976,牛津大学生物学家道金斯用基因观点对这个理论作了发展。但是,亲缘选择理论不能解释非近亲甚至非同一物种个体之间的利他现象。1971年,哈佛大学生物学家特里弗斯提出互惠理论解释利他现象。1981年,密歇根大学政策科学家阿克塞尔罗德与汉密尔顿合作,分析了策略在合作进化过程中的性质。同时,有许多实验生物学家对生命有机体的回报行为进行了实验室和野外观察。此后,关于互惠理论的研究工作大量涌现,成果迭出,文献量数以千计,对诸多领域产生了深刻的影响。这种局面的形成,在很大程度上得益于博弈论工具尤其是囚徒困境分析方法。

互惠利他主义的条件

从博弈论来看,利益部分冲突的两个个体之间的博弈是非零和博弈,博弈者之间采取的行动有合作和背叛两种。互惠理论的基本思想是,利益部分冲突的个体之所以采取合作行为,是因为它可能在今后与受惠者相遇时得到回报。威斯康星大学哲学家斯蒂芬斯给出了互惠利他行为及其进化满足的必要条件。其中,互惠利他行为的必要条件有四个。这就是:(1)该行为必须减少施惠者的与某个自私的抉择有关的适合度;(2)受惠者的适合度相对于非受惠者必须得到提高;(3)该行为的完成必须不依赖于某个直接利益的接受;(4)条件(1)、(2)和(3)必须适用于参与互惠帮助的两个个体。这里,条件(1)和(2)是使该行为互惠的条件,条件(3)把互惠利他主义与互助主义区别开来,条件(4)使利他主义互惠。这四点分开是必要条件,合起来是充分

条件。斯蒂芬斯并指出,互惠利他主义要能够进化,还必须满足另外两个条件,这就是:(5)必须存在察觉“骗子”的机制;(6)必须存在交换帮助的大量(不定)机会。条件(5)保证利他主义者有惩罚不合作的有机体的办法,条件(6)保证博弈的局数不定。

互惠理论把有机体采取一个步骤(合作或背叛)得到的结果称为适合度收益,在博弈论中叫做支付。自己背叛对方合作叫做背叛的诱惑,其收益用 T 表示;双方都合作叫做对合作的奖励,其收益用 R 表示;双方都背叛叫做对背叛的惩罚,其收益用 P 表示;自己合作对方背叛叫做对傻瓜的欺骗,其收益用 S 表示。如果对手合作自己背叛结果最好,对手背叛自己合作结果最差,双方合作比双方背叛结果要好,那么支付值的大小顺序就是 $T > R > P > S$ 。这称为指令条件。如果合作的奖励比对背叛的诱惑和对傻瓜的欺骗的平均效果好,则 $R > (T + S)/2$ 。这称为反剥削条件。鉴于这两个条件刻画了关于囚徒困境的假说,所以人们把它们定义的互惠利他主义称为标准囚徒困境模型。

这里,反剥削条件是为了保证困境的存在,因为对合作的奖励比对背叛的诱惑和对傻瓜的欺骗的平均效果好。但是在迭演(就是博弈多次进行)的情况下如果 $R \leq (T + S)/2$,困境仍然存在。 $R \leq (T + S)/2$ 意味着一对有机体轮流获得 T (诱惑)和 S (欺骗)回报的效果不会比轮流获得 R (奖励)的效果差,这是一个允许博弈者交替剥削和滞后合作的条件。去掉反剥削条件我们得到 $T > R > P > S$ 且 $R \leq (T + S)/2$,满足这两个条件的互惠利他主义模型称为修正的囚徒困境博弈模型。

上述两个模型都满足指令条件, R 大于 P 和 S ,就是说,除了自己背叛对方合作的情况之外,双方都合作的效果最佳。但是,观察结果显示,情况并非总是如此。有时同步合作的收益(R)大于同步背叛的收益(P)。这就像过多的厨师会弄糟一锅汤一样。于是,就有了以下三种情况,即 $T > P > R > S$ 且 $Z > (Y + X)/2$ 、 $T > P > R > S$ 且 $P \leq (T + S)/2$ 、 $T > P > S > R$ 且 $P \leq (T + S)/2$ 。这就是斯蒂芬斯提出的“厨师

困境博弈”。不过,在这三种情况下,虽然同步合作不会使博弈者受益,但是非同步合作还是会相互受益,所以互惠利他主义仍然存在。在上述三个厨师困境模型中,第一个模型是不稳定的厨师困境博弈,因为两个博弈者由于同步背叛而做得更好。第二个模型中 $R \geq P$,是强厨师困境博弈。最后一个模型称为弱厨师困境博弈。

上述模型对一些实际的观察结果,如鲈鱼的性角色变换、猎神狒狒交配中的雄性个体联盟、野蝙蝠的血液反哺等均作出了很好的解释。

合作进化的链条

以上条件给出了互惠利他主义及其进化的可能性。但是,现实世界中的情况是复杂的。这些条件并不意味着,利他者总是采取合作行为,因为采取这种行为的有机体若是面对总是采取背叛行为的有机体是会败下阵来的。这样,采取合作行为的有机体就会被淘汰而不会进化。就是说,利他主义者也可以在必要时采取背叛行为。博弈论中把支配博弈者采取合作和背叛行为的规则称为策略。策略有许多种。以上讲的任何时候都合作和任何时候都背叛分别称为全合作和全背叛策略。有一种交替策略,随机地选择合作或背叛,但采取合作和背叛的机会各占一半的次数。全报复策略决不首先背叛,但是只要对方有一次背叛,就从此一直背叛下去。一报还一报策略是第一步采取合作行为,然后采取对方上一步采取的做法,就是说,如果对方背叛它就背叛,如果对方合作它就合作,但是第一步它不知道对方的做法时就采取合作行为。两报还一报策略是第一步合作,然后如果对方连续两步背叛它就背叛一次,在其他情况下都采取合作。这些都是简单、常见的策略。

阿克塞尔罗德把策略的进化分为三个阶段,每一个阶段都对应于一个性质完全不同的问题,从而把利他之谜放到进化链条中并更加精细化。这些阶段和问题分别是:

1. 起始阶段:一个利他的策略如何能够在一个不合作占优势的环境中取得立足之地? 这个问题称为策略的初始成活性问题。

2. 中间阶段:什么类型的策略可以在一个由采用其他多种复杂策略构成的多样化环境中繁荣起来? 这个问题称为策略的强健性问题。

3. 最后阶段:在什么条件下,这样的策略一旦完全建立就能抵抗变异策略的侵犯? 这个问题称为策略的稳定性问题。这个问题中涉及到的变异策略、策略侵犯和稳定策略概念是梅纳德·史密斯提出来的。可以这样理解这些概念:假设一个群体中的成员除了个别变异个体之外都采用某一特定策

略,而这个变异个体采用的另外的不同策略就是变异策略。如果这个变异个体所得收益高于群体成员的平均值,则称这个变异策略能侵犯这个群体采用的策略;反之,如果一个策略不能被其他策略侵犯,则称这个策略是集体稳定的。假设一个群体的成员采用几种策略,如果其中一种策略的收益值大于其他策略的收益值,则称这种策略是进化稳定的。

假设一个群体开始都采用全背叛策略,而且这个策略是进化稳定的。有两个机制可以使基于合作的策略成活。第一个机制是亲缘关系的作用。当这个群体中的一员因为亲缘关系而对某一个近亲成员采取利他行为时,合作就会出现。这样,合作就会逐步使有近亲关系的一个小群体受益。当远亲成员对利他行为不回报时便会得到背叛的反应。这样,合作就会逐步扩散到无亲缘关系的群体之中去。在两个个体再次相遇的机会足够大时,基于回报的合作就会繁荣并且稳定下来。第二个机制是采用基于合作的策略的小群体成员之间的相互作用。只要这个小群体的成员相互作用的比例和它们相遇的机会足够多,这种策略就会成活并且稳定。阿克塞尔罗德组织了两次计算机竞赛,分别有 14 和 62 种策略参赛。后来,他又用生态模拟策略竞争过程。两次竞赛和生态模拟的结果都是一报还一报获胜。这说明基于合作的策略可以是进化稳定的。当然,上述两种机制还可以相互作用。总之,合作可以在一个无条件背叛的世界里产生,以相互回报合作为宗旨的小群体之间,一旦有交往的可能,合作便会出现。这样,基于回报的策略能够在许多不同类型的策略组成的环境里成长起来,并且能够抵抗其他不太合作的策略的侵犯。

稳定策略的理论命题

阿克塞尔罗德系统总结了在标准迭演囚徒困境情况下集体稳定策略的特征,把它们概括为 8 个理论命题,并且对之作了证明。

有两个概念对于理解这些命题是重要的,一是不同步骤博弈的收益值,二是策略相遇时的收益值。前者是从历时维度看,后者是从共时维度看。先介绍前一个概念。博弈者之间能合作是因为他们有可能再次相遇。但是,现在比未来更为重要,因为一般认为未来所得的收益值会随着时间的推移而减少,并且现在的博弈是现实,而再次相遇只是可能,迭演博弈可能会终结于对手的职业改变、居住地迁移、死亡或者破产。所以,同样的行为相遇导致的收益,下一步与当前一步相比并不相同,前者要打一个折扣。在计算迭演囚徒困境中的收益值时,我们用一个折扣系数 w 表示下一步相对于当前一步的权重或重要性,它的值在 0 与 1 之间。后一个概念

前面曾间接提到。如果策略 A 与策略 B 相遇,策略 A 的收益值用 $V(A/B)$ 表示;同时策略 B 也要与自身作用,策略 B 与自身作用的收益值用 $V(B/B)$ 表示。就是说,如果一个采用 B 的群体中有个别个体采用 A 的话,采用 B 的个体除了要与采用 A 的个体打交道之外,还要与同样采用 B 的其他个体打交道。如果策略 A 侵犯策略 B,就可以表示为 $V(A/B) > V(B/B)$ 。

现在我们分别讨论这些命题。

首先我们看是否存在无条件的最佳策略。答案是否定的。假设对方采用全背叛策略,那么这一方的最优策略就只能是全背叛,因为合作没有任何意义。再假设对方采用全报复策略,那么这一方的最优策略就只能是全合作,因为对方不首先背叛但又决不饶恕任何一次背叛。可见,策略的优劣,依情况而论。就是说,只要折扣系数 w 足够大,就不存在独立于对方策略的最优策略,这是**命题 1**。

命题 2 是,一报还一报是集体稳定的,当,且仅当 w 至少大于 $(T-R)/(T-P)$ 和 $(T-R)/(R-S)$ 二者中的较大者。这个命题的证明稍复杂一些。其大体思路是,全背叛和交替策略不能侵犯它,所以没有其他策略能够侵犯它,因此它是集体稳定的。通俗地讲,这个命题的意义就是,如果一个群体中的每个个体都采用一报还一报策略,只要未来足够重要,那么这个群体中就不会有哪个个体采用其他策略而能更多地受益。

那么,包括一报还一报在内的所有集体稳定策略有没有共同特征呢?有。**命题 3** 给出了以合作开始的策略集体稳定的条件。这就是:任何可能首先合作的策略,只有当 w 足够大时,才可能是集体稳定的。有人说,商业中最有力的道德执法者是持续的关系。如果供应商濒于破产,那么其客户往往以种种理由要求拒付贷款。下次选举中可能落选的议员很难与其他议员进行立法交易。这都是因为 w 不够大而导致首先合作的策略不稳定。

命题 4 给出了善良策略集体稳定的条件。所谓善良策略就是不首先背叛的策略。一个善良策略要成为集体稳定的,它就必须能被对方的第一次背叛所激怒。就是说,当对手第一次背叛之后,他也必须以自己的背叛来报复对手,否则,由于其收益值减少就会被背叛的策略侵犯。

当然,在所有集体稳定的策略当中,也有无条件的稳定策略。**命题 5** 告诉我们,全背叛策略总是集体稳定的。这个命题是容易明白的,因为群体成员中的每个个体每一步都背叛,都将得到 P ,而任何采取合作行为的个体的收益总是 S ,而且永远别想会得到回报。可见,如果采用其他策略的单个个体进入“小人”(相对于“君子”而言的)世界的话,这个世界完全可以抵抗这个个体采用的任何策略。正所

谓“秀才遇到兵,有理讲不清”。

不过,可以使我们不至于太悲观的是,如果采取其他策略并进入全背叛的“小人世界”中的不是单个人而是一个小群体,即使这个小群体与采取全背叛策略的大群体相比显得微不足道,但是只要采取这种策略的个体之间相互作用的比例足够大,这个新策略就可以侵犯全背叛策略。在满足这个条件的情况下,侵犯全背叛策略的那些策略当中,对采取新策略的个体的相互作用的要求是不同的。**命题 6** 告诉我们,能以最小 p 值的一小群体侵犯全背叛策略的是那些具有最大识别力的策略,如一报还一报。这里的 p 是采取侵犯全背叛策略的新策略的个体之间相遇的比例。如果采取一报还一报的个体与采取全背叛的个体相遇,它的第一步是合作,但此后都是背叛;如果它与采取相同策略即一报还一报的个体相遇,就会总是合作下去。这就很容易识别采用两类不同策略的个体。这种识别力可以使采用它的小群体侵犯“小人”世界。

接下来的问题自然是,善良策略会不会像全背叛策略那样,虽然不能被某单一个体侵犯,但却能被这类个体组成的小群体侵犯呢?**命题 7** 就是对这个问题的否定回答。这个命题说,如果一个善良策略不能被某单一个体侵犯,那么它也不能被这类个体的小群体侵犯。这种情况正好与全背叛策略的情况相反。全背叛策略可以抵抗任何由单独个体采取的其他策略的侵犯,但却不能抵抗这样的小群体的侵犯。善良策略只要能抵抗由单独个体采取的新策略的侵犯,它就能抵抗这样的小群体的侵犯。这是善良策略在进化过程中优于全背叛策略之处。政治学中这方面的例子甚多。

策略的侵犯可以像上述讨论中的情况那样,由采取某种新策略的个体或小群体进入一个采取老策略的大群体来实现,也可以在这样的个体或者小群体并不进入采取老策略的大群体的情况下实现。

博弈者的活动领域称为领地,它可以是地理或物理空间如居住地等,也可以是抽象的空间如产品类别、政治见解等。一个领地的主人可以不进入其邻居的领地,而使自己采取的策略侵犯其邻居领地或者被其邻居的策略所侵犯。**命题 8** 的内容就是,如果一个策略是集体稳定的,那么它也是领地稳定的。这个命题可以理解为,领地系统中的博弈者抵抗其他策略的侵犯并不比博弈者相遇的情况下抵抗其他策略的侵犯更难。

囚徒困境中的策略有很多,我们可以从现实生活中总结,可以逻辑地设计,还可以通过对生命、有机体的行为进行观察获得。但是,各种策略的竞争,都会满足上述 8 个命题。

[编者注:参考文献(主要是外文)全部删去,需者请与作者联系]
(责任编辑 蔡德诚)