

2002.04.010. БЕНДОР Дж., СВИСТАК П. ЭВОЛЮЦИЯ НОРМ.  
 BENDOR J., SWISTAK P. The evolution of norms //Am. j. of sociology, –  
 Chicago, 2001. – Vol. 106, №6 – P.1493-1545.

Авторы данной статьи ищут ответ на один из наиболее фундаментальных вопросов социальных наук: «почему существуют нормы?», применяя теорию эволюционных игр.

Под нормами Дж. Бендор и П. Свистак понимают «правила поведения, подкрепленные санкциями» (с. 1494). Наиболее подробно авторы анализируют характеристики и причины появления «социальных норм», то есть правил поведения, которые включают в себя применение санкций третьей стороной. Долгое время считалось, что существование такого типа норм сложно объяснить с позиции рационального выбора. Однако авторы полагают, что теория эволюционных игр способна доказать, что социальные нормы напрямую связаны с принципами ограниченно рационального выбора и необходимы для того, чтобы стабилизировать поведение людей внутри социальных групп и институтов.

В современных социальных науках можно выделить два типа объяснения существования норм. Первый тип связан с традицией методологического индивидуализма, с представлением о человеке, как *homo economicus*, действия которого описываются с точки зрения индивидуальных выборов и целей. Второй тип основывается на традиции структурного функционализма. Человек здесь рассматривается, как *homo sociologicus*, поведение которого определяется групповым влиянием и его функциональным значением для группы или общества.

Оба подхода не лишены недостатков. Например, структурно-функциональный подход объясняет существование норм их функциональностью для группы. К недостаткам структурно-функционального подхода относят нерешенные концептуальные вопросы (в определении понятия «функциональность»), его статичность в объяснении существования норм, нежелание искать причины появления норм, а изучение их, как данности.

Авторы полагают, что теория эволюционных игр помогает снять противоречия данных подходов, соединить микро- и макро-уровни социальной реальности и избежать статичности анализа.

Согласно классической «дилемме узника» во взаимодействие вступают два игрока, которые могут либо сотрудничать, либо нет. Если у игры всего один раунд, то результаты распределяются следующим

образом:  $T$  (для игрока, который не сотрудничает, по отношению к игроку, который сотрудничает)  $> R$  (оба игрока сотрудничают)  $> P$  (оба игрока не сотрудничают)  $> S$  (для игрока, который сотрудничает, в то время как другой отказывается от сотрудничества). Отказ от сотрудничества оказывается индивидуальным рациональной стратегией поведения, но невыгодной для группы. Однако если взаимодействие продолжается и повторяется некоторое неопределенное количество раз, то считается, что совместное взаимодействие оказывается выгоднее эксплуатации одного игрока другим:  $2R > T + S$ . Следовательно, кооперация становится стабильной только в том случае, если будущее является достаточно значимым для взаимодействующих игроков.

Подобное равновесие достигается во многих играх с повторяющимся взаимодействием, в которых будущее является достаточно значимым. Например, использование двумя игроками так называемой стратегии «зуб за зуб» (сотрудничать в первый период игры, а в последующий любой  $k$ -й период игры делать то, что другой игрок сделал в период игры  $k-1$ ), при условии, что будущее является достаточно значимым, приводит к устойчивому равновесию Неша, при котором ни один игрок, зная стратегию другого игрока, не намерен менять своей стратегии поведения.

Однако это равновесие не предполагает, что в любом случае возникнет кооперация (уровень сотрудничества может колебаться от 0 до 100%). Всегда важно учитывать представления игрока о стратегии оппонента. Дополнительная информация о поведении другого поступает за счет взаимодействия с социальным окружением. Теория эволюционных игр как раз и учитывает изменения индивидуального поведения, при условии, что люди узнают о поведении других членов группы. Эволюционная «игра представляет собой парные взаимодействия индивидов в группе» (с. 1505). «Узнав о стратегиях поведения других людей, игроки могут захотеть изменить свои стратегии поведения на те, что оказались более успешными» (с. 1505).

В эволюционных играх рациональность игроков ограничена, поскольку они способны совершать ошибки. Динамический процесс, который управляет копированием и принятием тех или иных стратегий поведения, является функцией общей суммы результатов всех парных взаимодействий и их частот.

Фундаментальным постулатом теории эволюционных игр является то, что эволюционный процесс представляет собой динамику, при

которой растет общая сумма результатов парных взаимодействий игроков.

Стратегия является эволюционно стабильной, если она достаточно распространена в группе, и поэтому способна противостоять вторжению в группу незначительного числа стратегий-мутантов. Это противостояние может быть более сильным, если после вторжения частота стратегий поведения, вторгшихся в группу, будет уменьшаться, или более слабым, если их частота уменьшаться не будет. При повторяющихся взаимодействиях достижима только слабая стабильность определенной стратегии, при которой частота данной стратегии не уменьшается в любой группе с ограниченным набором стратегий, в которой ее частота достаточно высока.

Определенная стратегия может быть стабильной при одной динамике и нестабильной при другой. Характер динамики напрямую связан с некоторыми фундаментальными социальными характеристиками поведения людей в группе. Например, тип *homo sociologicus* будет исходить из фактора конформизма, а тип *homo economicus* будет считаться только с размерами выигрыша.

Однако существуют такие стратегии, которые (слабо) стабильны при любых условиях, т.е. при учете всех эволюционных процессов. Авторы называют данные стратегии «постоянно стабильными». Именно такой тип стратегий связан с появлением норм.

Авторы статьи вывели ряд теорем, которые направлены на то, чтобы дать объяснение возникновению норм (в особенности, социальных).

Теорема 1 гласит: «В любой нетривиальной игре с повторяющимся взаимодействием, при учете, что будущее является достаточно значимым, никакая чистая парная стратегия не может быть постоянно стабильной» (с. 1512). Игра с одним раундом называется тривиальной, если она приводит к получению максимального выигрыша, независимо от действия другого игрока. Стратегия  $i$  называется парной, если она строит свое поведение относительно стратегии  $j$ , исходя из истории взаимодействия стратегий  $i$  и  $j$ . Информация о предыдущих взаимодействиях  $j$  с третьей стратегией  $k$  не влияет на поведение  $i$  относительно  $j$ .

Следствием из данной теоремы является следующий вывод: «Если чистая стратегия постоянно стабильна в игре с повторяющимся взаимодействием, при учете, что будущее является относительно значимым, то она

должна быть социальной стратегией» (с. 1512). Фактически, социальные стратегии представляют собой социальные нормы.

Таким образом, если какие-либо стратегии стабильны перед лицом эволюционных изменений, включая процессы, связанные с экономическим принуждением, то эти стратегии должны быть зафиксированы в определенном виде социальных норм.

Особый тип социальных норм связан со стратегией конформизма. Согласно этой стратегии, все участники игры разделяются на друзей и врагов. Сначала все игроки считаются друзьями. Если игрок является другом в период игры  $t$ , то он остается таковым в период  $t+1$ , тогда и только тогда, если он сотрудничал со всеми друзьями и избегал сотрудничества со всеми врагами в период  $t$ . Друг, нарушающий правило, становится врагом.

Следовательно, если будущее является достаточно значимым, стратегия конформизма является постоянно стабильной для «дилеммы узника» с повторяющимся взаимодействием. Можно сказать, что эта стратегия воплощает житейское правило: «друг твоего друга – твой друг, друг твоего врага – твой враг, враг твоего друга – твой враг».

Логика конформизма включает в себя три ключевых элемента: 1) разделение всех игроков на друзей и врагов; 2) социально желаемое действие и наказание в одном раунде; 3) правила игры социально желаемого действия по отношению к друзьям и наказания по отношению к врагам. Для симметричных игр, использующих такую логику важно учитывать два обстоятельства, выделенных еще Гоббсом: «договоры возможны; существуют мечи для наказания людей, нарушивших договоры» (с. 1516). То есть в симметричных играх должно существовать некоторое принуждение к кооперации.

Это подтверждается Теоремой 2, согласно которой «чистые постоянно стабильные стратегии существуют в повторяющихся симметричных играх с достаточно важным будущим, тогда и только тогда, когда в игре обеспечивается принуждение к кооперации» (с. 1517).

Основными характеристиками постоянно стабильных стратегий (в частности, стратегии конформизма) можно назвать следующие: постоянный учет взаимодействий с третьей стороной и контроль взаимодействий всех игроков без исключения. Таким образом, следует выделить две существенные характеристики данных стратегий: 1) стратегия называется «постоянно социальной», если она не становится парной (не учитывается третья сторона) после конечного числа периодов

взаимодействия; стратегия  $i$  называется «всесторонне социальной», если не существует стратегии  $j$  и периода  $t$  игры, при которых во всех периодах после  $t$ , действия  $i$  по отношению к другим игрокам в группе не зависят от их взаимодействия с  $j$ .

В соответствии с Теоремой 3: «В любой повторяющейся игре с принуждением к кооперации, если будущее является достаточно важным, все постоянно стабильные стратегии должны быть постоянно социальными» (с. 1520).

Согласно Теореме 4: «В любой повторяющейся игре с принуждением к кооперации, если будущее является достаточно важным, все постоянно стабильные стратегии должны быть всесторонне социальными» (с. 1520).

Теорема 5 гласит «В любой повторяющейся игре с принуждением к кооперации, если будущее является достаточно важным, каждая постоянно стабильная стратегия должна обладать полностью метанормативной структурой» (с. 1520).

Таким образом, постоянно стабильные стратегии должны создавать социальные нормы следующим образом: контролировать и сдерживать каждого внутри сообщества, делать это постоянно и наказывать любое нарушение закона  $n$ -го порядка.

Девиация в подобных стратегиях имеет социальную функцию. Например, если для стратегии конформизма характерны два образца поведения: либо все следуют закону всеобщей солидарности, тогда стратегии-мутанты остаются в латентном состоянии, либо кто-то нарушает закон, проявляет себя, и, следовательно, его наказывают.

Однако было бы неверным говорить, что все постоянно стабильные нормы функциональны (в паретовском смысле), хотя все приведенные примеры доказывают, что постоянно стабильные нормы поддерживают оптимальный для коллектива результат.

Теорема 6 проясняет эту особенность социальных норм: «В любой повторяющейся игре с принуждением к кооперации, если будущее является достаточно важным, постоянно стабильное состояние сообщества при использовании социальных стратегий может поддерживать любой строго положительный уровень  $a(t)$ -кооперации» (с. 1524).

Следовательно, только полный отказ от кооперации может быть назван нестабильным состоянием. Данный факт подтверждает, что эволюционные процессы не всегда приводят к оптимальным для

сообщества результатам, и опровергает чрезмерно функционалистские утверждения.

Тем не менее, стоит отметить, что, в конце концов, преобладают более функциональные нормы.

Согласно Теореме 7: «В любой повторяющейся игре с принуждением к кооперации и очевидным наказанием, при достаточно значимом будущем, стратегии, которые являются нормативными и носят репрессивный характер» (стратегии типа конформизма), «нуждаются в минимально малой частоте проявления, чтобы стать постоянно стабильными» (с. 1526).

Теорема 8 гласит: «В любой повторяющейся игре с принуждением к кооперации и очевидным наказанием, при достаточно значимом будущем, чем эффективнее (максимально устойчивее) является принятая стратегия конформизма в сообществе, тем большее максимальное число вторжений чужих стратегий она может отразить» (с. 1527).

Все перечисленные выводы относились к сообществам с плотными взаимодействиями: все игроки взаимодействуют друг с другом в каждый период игры и по окончании каждого периода знают, что произошло в процессе всех взаимодействий в сообществе.

Однако социальные нормы необходимы и в сообществах с более редкими взаимодействиями. Поэтому возможно предположить (в качестве следствия из Теоремы 1), что «во всех повторяющихся симметричных играх со случайным взаимодействием отдельных партнеров и единичным свидетелем, при достаточно значимом будущем, любая чистая стратегия, являющаяся постоянно стабильной, должна быть социальной» (с. 1528).

Проводимый в статье анализ, естественно, является лишь идеальной моделью стратегий поведения людей, поскольку, в реальности, всегда присутствует некоторая степень неопределенности: игрок может неверно понимать действия другого или неверно осуществлять задуманные действия, вводя другого в заблуждение. Незначительный уровень неопределенности («шума»), который не способен создать большое количество стратегий-мутантов, не пошатнет стабильность принятой в сообществе стратегии. Однако если уровень неопределенности перейдет порог допустимости, возможны различные проблемы. В такой ситуации парные стратегии, которые обычно нестабильны, могут стать постоянно стабильными. Тем не менее, за социальными стратегиями останется эволюционное преимущество.

Таким образом, в ситуации, когда необходим отбор, социальные нормы имеют значительное эволюционное преимущество перед другими формами поведения. Социальные нормы копируются чаще и проще, чем другие формы поведения, если изменения в базовой модели поведения не производят слишком много девиантных стратегий.

*Д.А.Архипова*