

УДК 159.9

МОДЕЛИРОВАНИЕ ПРОСТРАНСТВА СТРАТЕГИЙ ЗАДАЧИ «ДИЛЕММА ЗАКЛЮЧЕННОГО»¹

Д.Ю. Баланев^а

^а Томский государственный университет, 634050, Россия, Томск, пр. Ленина 36

Представлены результаты эксперимента, направленного на трансформацию стратегий решения задачи итерационной дилеммы заключенного людьми, вовлеченными в искусственно созданные условия социализации. Процесс изменения стратегий показан с точки зрения взаимодействия их эталонных образцов в игровой информационной среде, обобщенного на уровень нейронной сети. Рассматривается возможность использования нейронной сети в качестве классификатора, решающего задачу распознавания неявно выраженных стратегий, реализуемых человеком в различных социальных условиях.

Ключевые слова: стратегии решения задач; информационное моделирование; «дилемма заключенного»; нейронная сеть; социализация; игра.

Введение

Итерационная версия игры «Prisoner's Dilemma» (IPD) применяется в качестве модели кооперации во многом благодаря широкому спектру стратегий, которые могут использовать ее субъекты. Стратегиями называют такие формы поведения участника, которые основаны на получении выгоды с условием учета действий оппонента.

Проблема эффективности стратегий решения итерационной дилеммы заключенного чаще рассматривается с точки зрения информационных моделей, где стратегии решения не «отягощены» отношениями, которые возникают при игре реальных людей.

Некоторые из этих стратегий очевидны, другие оказываются зависимыми от социального контекста. К числу первых, безусловно, можно отнести стратегии «всегда предавать» (ALLD) и «всегда сотрудничать» (ALLC). М. Nowak и К. Sigmund сделали эти две стратегии полюсами континуума, куда вошли 16 стратегий, расположенных с учетом отношения порядка. На девятой позиции в этом списке находилась стратегия TFT, на десятом – PAVLOV [1]. При этом учитывались также «стратегии с мутациями». К элементарным можно отнести стратегию случайного выбора возможностей. Несмотря на иррациональность такого подхода он легко может быть

¹ Результаты были получены в рамках выполнения государственного задания Минобрнауки России, проект № 0721-2020-0040.

объяснен с точки зрения психологических особенностей некоторых людей. Учитывая опыт использования различных когнитивных тестов, возможно также указать на распространенность стратегии поиска оптимального решения путем «проб и ошибок». В этом случае какие-то действия могут предприниматься не потому, что они выгодны с точки зрения выигрыша, но как способ получить дополнительные знания эмпирическим путем. Ключевым понятием в этом случае оказывается репутация оппонента – действия игроков в значительной степени зависят от нее. Исключить репутацию оппонента возможно тем, что игрок имеет дело не с конкретным человеком, а с группой людей. При этом он может идентифицировать себя с этой группой в различной степени. Можно взять две группы, одна будет in-group, другая out-group. Особенно интересны первые ходы, когда человек еще не определился со своей стратегией. В какой-то степени бессознательные реакции могут быть описаны «чистыми» стратегиями. Стратегии могут меняться по ходу игры. Если мы возьмем достаточно большой промежуток, он может оказаться заполненным образцами различных стратегий. Люди также склонны доверять «интуиции» как метафоре, не поддающейся рационализации. Возможность «наказания» некооперативного игрока приводит к появлению стратегии «око за око» – TFT и ее вариантов, например Generous TFT – GTFT [1]. Эти стратегии долгое время оценивались как оптимальные и оценивались в различных социальных средах [2–11].

Дальнейшее развитие темы стратегии решения задач итерационной дилеммы заключенного получило благодаря открытию эффекта смены стратегий. Например, оказалось, что стратегия PAVLOV, известная ранее как одна из неэффективных, становится более сильной после того, как до нее работала стратегия TFT [8, 12–14]. Помимо варианта с двумя игроками, возможны более сложные случаи, когда рассматривается система с тремя и более игроками. Например, М. Matsushima и Т. Ikegami рассматривают эволюцию стратегии в IPD с тремя игроками [15].

Несмотря на разработанность проблемы, она далеко не исчерпана. Анализ эволюции проблемы стратегий дилеммы заключенного в последние годы стал самостоятельной темой исследования [16–18]. Одно из наиболее перспективных направлений в развитии изучения стратегий IPD – применение искусственных нейронных сетей. Нейронные сети могут быть использованы и как средство моделирования, и в качестве игровой среды [19].

Основная цель нашей работы – построение информационной модели, предсказывающей поведение как отдельного человека, так и группы людей в ситуации решения задачи социальной дилеммы. При этом учитываются социальные отношения, в том числе и вызываемые экспериментальным воздействием, гендерными различиями, индивидуальными различиями в стратегии решения познавательных задач. Модель должна демонстрировать переход индивидуальных действий в социально детерминированное поведение.

Использование процессуального подхода позволяет учесть кратковременный эффект социализации, вызванный процедурой игры в IPD. В целом

ряде указанных выше работ показано, что с повышением номера итерации вероятность кооперации в IPD снижается. Это объясняется действием мотива «наказания» некооперативного игрока, как это типично для стратегии TFT. Иногда этот мотив смягчается «прощением», что характерно для стратегий GTFT, TF2T, RTFT. Оценка эффекта социализации, связанной с самой процедурой игры, позволяет получить дополнительную информацию об особенностях экспериментального воздействия. Здесь мы исходим из предположения о том, что экспериментально вызванная социализация снижает эффект «негативной социализации», вызванной процедурой игры.

Исследовательские вопросы:

1. Определить единицу анализа стратегии игрока IPD в ситуации с несколькими игроками, объединенными в группу, идентичность с которой может рассматриваться как переменная величина.

2. Определить влияние экспериментально сформированной групповой идентичности (социализации) на изменение предпочитаемых стратегий в социальных дилеммах.

3. Оценить возможности нейронных сетей второго поколения как средства категоризации результатов итерационной дилеммы заключенного с точки зрения применяемой игроком стратегии, а также социальных факторов, таких как место жительства, вид социализации, гендерная идентичность, групповая гендерная идентичность.

4. Оценить, возможно ли использовать социализацию как механизм повышения степени определенности действий игроков IPD при использовании моделей на основе нейронной сети.

5. Определить закономерности изменения стратегии игрока IPD до и после социализации. Представляют интерес следующие вопросы: Насколько склонны реальные люди использовать «чистые» стратегии? Будут ли игроки повторяться, используя одну стратегию или последовательность действий до и после экспериментального воздействия, связанного с формированием социальной идентичности? В какой степени будут мутировать «чистые» стратегии?

Процедура

В ходе исследования оценивался ряд категориальных переменных, использованных для построения классификационных моделей:

1. Социализация, вызванная долговременными социальными отношениями (социализация 1). Данная переменная имеет два уровня: люди, ранее не знакомые между собой (незнакомцы) и поддерживающие дружеские связи на протяжении не менее одного года (друзья).

2. Экспериментально сформированная социализация (социализация 2). Здесь также задано два уровня: уровень социальных отношений до тренинга «социализации» и после его воздействия.

3. Гендерная дифференциация. Переменная, заданная двумя уровнями: мужчины, женщины. Факт половой принадлежности устанавливался по самоидентификации, самоназванию участников в условиях свободного выбора дихотомии male–female.

4. Групповой аспект гендерной дифференциации. Эта переменная оценивалась с точки зрения преимущественного количества мужчин и женщин, входящих в группу из шести человек. Здесь выделено три уровня: женская группа, смешанная группа, мужская группа.

5. Место проживания участников. Переменная, имеющая два уровня: Москва, Томск.

В экспериментальной серии с уровнем переменной «социализация 2» – «незнакомцы» приняли участие 180 человек – студенты Московского физико-технического института и Томского государственного университета.

Набор проводился при помощи объявлений в социальной сети vk.com. Для каждого исследования в серии требовалось 12 участников, незнакомых друг с другом. Контроль этого условия осуществлялся с учетом факультета, учебной группы, родного города участника и возможных родственных связей.

Эксперимент с «незнакомцами» состоял из трех этапов. Первый этап определен как диагностический. Участникам исследования необходимо было выполнить задачу итерационной дилеммы заключенного по схеме получения прибыли (рис. 1).

		Игрок 2	
		Кооперация (C)	Отказ (D)
Игрок 1	Кооперация (C)	5 / 5	0 / 10
	Отказ (D)	10 / 0	1 / 1

Рис. 1. Схема социальной «дилеммы заключенного». Числами показаны условные баллы, набираемые игроками в каждом определенном инструкции случае

Проводилось десять игровых итераций между случайно выбранными игроками. В качестве результата измерения данной переменной использовался относительный показатель доли кооперативных ходов, выраженный в процентном отношении к общему количеству ходов. На втором этапе проводилось изменение уровня независимой переменной, определяемой как второй вариант социализации. Для этого участники совместно выполняли ряд заданий, назначение которых состояло во включении их в социальные взаимодействия. Смысл третьего этапа – измерение уровня независимой переменной – эффекта социализации. Средством измерения, как и на первом этапе, служила игра «Дилемма заключенного». Результатом измерения данной переменной также стал относительный показатель доли кооперативных ходов, выраженный в процентном отношении к общему их количеству. В отличие от первого этапа участники играли только внутри подгрупп, образованных в ходе экспериментально сформированной социализации. Участники также каждый период случайным образом разделялись на пары, при этом они были проинформированы, что точно играют с участ-

ником из своей подгруппы, но неизвестно, с кем конкретно. Игра на этом этапе состояла из 10 периодов. Полученные баллы также учитывались в сумме итогового выигрыша, который переводился в выплачиваемое участнику денежное вознаграждение [20].

В серии экспериментов с «друзьями» приняли участие 120 человек – студенты Московского физико-технического института и Томского государственного университета, набранные с помощью объявлений в социальной сети vk.com.

Для каждого исследования в серии требовалось 12 участников, состоящих из двух подгрупп по шесть человек. Особенностью этих подгрупп являлся высокий уровень социального взаимодействия (устойчивые, предвзвешенно социализированные группы). Отбор подгрупп «друзей» производился через непосредственное общение с представителями подгрупп. В качестве критериев устойчивой дружеской связи были выбраны: продолжительное (более года) и постоянное общение. Также учитывались наличие совместной деятельности и общие интересы.

Исследование проводилось по плану, в целом аналогичному для групп «незнакомцев». Отличие состояло в содержании второго этапа, в ходе которого участники разделялись на подгруппы из шести друзей, в составе которых они были приглашены на исследование. Общим на втором этапе было то, что каждой подгруппе «друзей» давалось задание найти 5 общих характеристик и придумать название подгруппе.

В качестве зависимых переменных учитывались показатели результативности деятельности, а также процессуальные характеристики. Результативность измерялась количеством кооперативных и некооперативных исходов в двух сериях итеративной дилеммы заключенного – до процедуры экспериментальной социализации и после нее. Процессуальные характеристики представлены категориальными величинами в виде строк, описывающих последовательность действий участников эксперимента и их оппонентов в каждом раунде игры до и после процедуры экспериментальной социализации.

Для анализа результатов использовались две формы записи, описанные М. Nowak и К. Sigmund [8]. Первая форма учитывает действия одного игрока и состоит из последовательности символов «C» и «D» – cooperate и defect соответственно. Вторая форма отражает действия пары игроков и представлена последовательностью символов, каждый из которых может принимать четыре значения. Если оба игрока кооперируют, результат записывается как «R» (reward), если оба отказываются от кооперации, то «P» (punishment). В случае, если кооперирует только один игрок, то получаем «T» (temptation) или «S» (sucker). В случае, если первый игрок кооперирует, а второй нет, результатом будет «S». Таким образом, мы можем описывать процессуальные характеристики IPD в виде паттерна, отображающего действия пользователей во всех раундах в виде строк однотипных символов. Одна такая строка кодирует действия пользователя до процедуры экспериментальной социализации, другая – после. Строки имеют вид «RRPPTSPTTP».

Результаты исследования

В итерационной дилемме заключенного очень важен первый ход. В математических моделях этот момент определяется только свойством конкретной стратегии. Например, стратегия TFT обязательно начинается с кооперации. В игре с реальными людьми первый ход может зависеть от множества обстоятельств, диктуемых как ситуативно, так и с точки зрения личностных особенностей: предыдущий жизненный опыт игрока, репутация партнеров за пределами игры, особенности восприятия инструкции, отношение к экспериментатору и т.д.

Подсчет частоты первых ходов до и после экспериментального формирования социализации показал следующее соотношение: до – 133 кооперативных и 167 некооперативных ходов; после – 204 кооперативных и 96 некооперативных. Применение критерия X-squared показывает результат 33,17 (с поправкой Йетса) при $p = 8,439 \times 10^{-9}$. Таким образом, анализ только первого хода в итерационной дилемме заключенного с экспериментальным формированием групповой идентичности позволяет отвергнуть нулевую гипотезу о незначимости этого влияния на кооперативное поведение игроков.

Что касается изменения частоты различных вариантов взаимодействия двух игроков во время первого хода, то они распределились следующим образом:

- до экспериментального воздействия: 'S': 73, 'P': 94, 'R': 60, 'T': 73;
- после экспериментального воздействия: 'S': 54, 'P': 42, 'R': 150, 'T': 54.

Игровое разнообразие

Для анализа использовались только первые 10 ходов как в варианте до социализации, так и после. Такое решение обусловлено тем, что 10 ходов – самая короткая игровая последовательность в проведенных нами исследованиях.

Интересно отметить, что за редким исключением последовательность действий в игре до экспериментального воздействия является уникальной. Для установления этого факта мы предприняли анализ частот всех строк, описывающих стратегии, и обнаружили всего восемь случаев, когда результаты двух игроков и их партнеров повторяются два раза. Шесть таких повторений относятся к ситуации встречи игрока, руководствующегося стратегией «всегда предавать» (ALLD) с другим игроком, предпринимаящим попытки кооперации. Например, «PPTPTTPPP» или «TPPPPTTPPP». Только два повторения из восьми имеют в начале хотя бы один кооперативный ход: «SSPPPPPPPP» и «STPPTTPPP». Попытка обнаружить игрока, отказывающегося от кооперации на каждом ходу и при этом встречающегося также с некооперативными партнерами, показала только один такой случай.

После экспериментального воздействия результаты изменилась существенным образом. Вариант «RRRRRRRRRR» встречается 52 раза. Таким образом, ситуация, когда два игрока придерживаются в отношении друг друга стратегии «только кооперировать», приобрела значительный харак-

тер. При этом случаев совпадения результатов игр с иным исходом мы обнаружили всего шесть. Пять из них однотипны – первый игрок, кооперируя во время каждой итерации, однократно сталкивался с единственным игроком, предпринявшим некооперативный ход. Формальная разница в одну итерацию при сохранении общего смысла встречи только кооперирующих игроков позволяет отнести эти повторы к устойчивой группе кооператоров, «встречающихся» друг с другом. В двух случаях игрок отказывается от кооперации только после того, как безуспешно кооперировал в течение семи итераций – «TTTTTTTPPP»

Зафиксировано также три одинаковых случая, когда только один из игроков всегда предавал, а его оппоненты кооперировали.

Если рассматривать действия игроков с точки зрения того, кто из них придерживался однозначной стратегии ALLD или ALLC, то до экспериментального воздействия вариант с предательством во всех раундах повторился 61 раз, а с полной кооперацией всего 4. После процедуры экспериментального формирования социализации 102 человек придерживались стратегии ALLC и только 39 – ALLD.

Если использовать динамику изменения количества повторений стратегии ALLD и ALLC до и после экспериментального формирования групповой идентичности, то в первом случае получаем $X\text{-square} = 5,29$ ($p = 0,021$), во втором $X\text{-square} = 107,81$ ($p < 2,2e-16$).

Дополнительно была оценена частота случаев, в которых стратегия «всегда предавать» или «всегда кооперировать» нарушалась в одной из итераций. Такое «нарушение» можно интерпретировать как случайный ход или единичный маркер какой-то другой стратегии. В турнирах ботов, реализующих различные стратегии, такие случаи рассматриваются как стратегии с мутациями [18]. Мы использовали обозначения ALLDM и ALLCM для идентификации таких стратегий. Первая из них встретилась в наших результатах 32 раза до социализации и 17 после, вторая – 3 раза до социализации и 11 после. Если учитывать количество стратегий ALLD и ALLDM в сумме, то изменение этого показателя до и после экспериментального воздействия оказывается достаточно надежным средством для доказательства его влияния: $X\text{-square} = 11,57$ ($p = 0,0007$). В случае с одновременным учетом ALLC и ALLCM $X\text{-square} = 114,84$ ($p < 2,2e-16$).

Несмотря на то, что количество случаев ALLD, в том числе и с мутациями, уменьшается после экспериментального воздействия не настолько радикально, как увеличивается частота ALLC, и тот и другой параметр может использоваться в качестве показателя эффективности процедуры экспериментального формирования групповой идентичности. Еще один вывод, который мы здесь можем сделать, касается уменьшения частоты проявления любых других стратегий, кроме уже рассмотренных, в игре, проводимой после экспериментального воздействия.

Выявление фиксированных форм игрового поведения

Различия в изменении стратегий реальных игроков по сравнению с моделями, предусматривающими рациональное поведение, содержат ключ к

поведению людей в социальных дилеммах. Вероятность того, что человек воспроизведет свое поведение в двух раундах, представляется достаточно высокой в том случае, если речь идет о стратегиях ALLC или ALLD. Действительно, 20 участников придерживались стратегии ALLD как до, так и после экспериментального формирования социальной идентичности. В то же время четыре человека, которые до экспериментального воздействия придерживались стратегии ALLC, сохранили ее и во второй игровой серии.

Особый интерес в этом отношении представляет форма игрового поведения, демонстрирующая резкий переход от стратегии ALLD к стратегии ALLC в различных игровых сериях. Мы обнаружили 17 таких случаев, при том что обратных переходов участники нашего эксперимента не проявили вовсе.

Еще раз уточним, что некоторые из игроков последовательно придерживаются фиксированной формы игрового поведения, а другие после экспериментального формирования групповой идентичности демонстрируют в данном отношении значительную гибкость. При этом две эти группы, очевидно, равны по величине ($X^2 = 0,115$, $p\text{-value} = 0,73$). На этом основании возможно выдвинуть гипотезу о наличии некоторой «суперстратегии», зависящей от личностных характеристик игроков, например, такой как психологическая ригидность, и проявляющейся в ситуации социально значимого воздействия на игроков. В данном случае мы видим крайние проявления набора таких стратегий в распределении результатов итерационной дилеммы заключенного.

Если в описанной картине учесть дополнительно семь пар, составленных из стратегий ALLDM–ALLC, то и тогда можно говорить о примерном равенстве упомянутых групп: $X^2 = 1,98$ ($p = 0,159$).

Таблица 1

**Частота совпадений действий игроков до и после социализации
(без учета ALLC, ALLD, ALLDM, ALLCM)**

№	Кол-во совпадений ходов	Кодирование CD		Кодирование RPTS (CC DD DC CD)	
		до социализации	после социализации	до социализации	после социализации
1	10	CDDDDDCDDD	CDDDDDCDDD	SPPTPPSPPP	SPPPPPSPTP
2	9	CCDCDCDCD	CCDCDCDCD	RSRTTRTSPST	RSRTSPRP RR
3	6	CCCCCDDCDD	CCCCCCCCC	RRSSSPSTP	RSRRRSRRRR
4	6	CCCCCDDDCD	CCCCCCCCC	RRSRSPPPRP	RRRSRRRRRR
5	6	CCCCCDDDDD	CCCCCCCCC	SRRSSPPPPP	RRSRRRRRRS
6	6	CCCCCDDDDD	CCCCCCCCD	RRSSRSPTPP	RRRRRSRTT
7	5	CCCCDCDDDD	CCCCCCCCC	RRSRSTPTT	RRRRRRRRRR
8	5	CCCCCDDDDD	CCCCCCCCC	RSRRSPPTP	RRSRSRRRR
9	5	CCDCCCCDD	CCDCCDDCDD	SSTSSRSST	RSTSTTSR
10	5	CDDDDDCDDDD	CDDDDDDDD	STPPPSPPPP	STPPPPPTPT

Примечание. Жирным шрифтом выделены ходы тех игроков, которые совершали одинаковые последовательности действий в начале игры до и после социализации.

Только один случай показывает полное совпадение действий игрока до и после социализации в случае более сложного игрового поведения (табл. 1, строка 1). Однако если взять только первые 9 ходов, то есть еще один случай полного совпадения действий игрока до и после экспериментального воздействия (см. табл. 1, строка 2). Здесь мы видим, что первые три итерации до и после социализации – кооперация. Примечательно, что первое «предательство» игрока в данном случае следует не сразу после аналогичного хода оппонента, а через один ход, т.е. с некоторой задержкой. Дебют, который повторялся бы в двух игровых сериях для 7 и 8 ходов, мы не обнаружили, но зато отметили четыре случая повторения первых шести и пяти ходов. Почти всегда это были ситуации переключения с последовательности кооперативных ходов на последовательность некооперативных после того, как оппонент «эксплуатировал» кооператора в течение двух, трех или даже четырех ходов. Каждый раз «переключение» происходило после очередного хода «кооперация–предательство». Все игроки, продемонстрировавшие такого рода стратегию на первой игровой серии, выступали после экспериментального воздействия как последовательные «кооператоры», даже если количество некооперативных ходов их оппонентов превышало пороговую для первой серии величину (см. табл. 1, строка 8).

Поиск стратегий Grudger и OppositeGrudger

Следующим шагом мы выявляем наиболее очевидные стратегии. Под очевидными подразумеваются наиболее вероятные с точки зрения реализации игроком и такие, которые в истории развития итерационной дилеммы заключенного были описаны прежде других. Одной из первых таких стратегий стала GRIM (Grudger). Эта стратегия состоит в том, что игрок начинает с кооперации и придерживается этого действия до тех пор, пока оппонент не допускает первое предательство. После этого игрок всегда «предает». Существует инверсная версия этой стратегии – OppositeGrudger, когда игрок начинает с «предательства», но после первого же кооперативного хода оппонента всегда «кооперирует». Обе эти стратегии относятся классу «триггерных» стратегий и особенно удобны для анализа поведения потому, что переключение с одного устойчивого состояния на другое происходит однократно.

Мы обнаружили стратегию Grudger в эталонном виде только один раз до введения экспериментального воздействия и три раза после (табл. 2). Что касается стратегии OppositeGrudger, то она не встретила нам ни в одном случае.

Тем не менее формальные признаки и той и другой стратегии можно встретить в целом ряде случаев (см. табл. 2). Эта таблица была построена из всех случаев, когда игрок однократно менял свои действия в течение игры до или после экспериментального воздействия. Здесь мы почти всегда встречаемся с эффектом «отсроченного переключения» действий пользователя.

Еще одно основание позволяет отнести описываемые случаи к классу стратегий Grudger – переключение почти всегда происходит после хода «S» (в 17 случаях из 22).

После первого такого хода, когда игрок кооперирует, а оппонент «предает», мы обнаруживаем переключение с «С» на «D» в двух случаях, после второго – 5, после третьего – 4, после четвертого – 1. Возможна также задержка, например в один ход (см. табл. 2, строка 1). После социализации переход с «С» на «D» происходит несколько реже: в трех случаях после одного хода «S», в двух случаях, после двух ходов, в одном случае после трех и даже шести ходов (см. табл. 2). Можно считать все эти случаи «мягким» вариантом проявления стратегии OppositeGrudger.

Таблица 2

Примеры действий игроков, однократно меняющих действия в течение игры до или после экспериментального воздействия

№	Кодирование CD		Кодирование RPTS (CC DD DC CD)	
	до социализации	после социализации	до социализации	после социализации
1	CCCCCDDDD	CCCCCCCDD	RRSSRSPTTP	RRRRRRSRTT
2	CCCCCDDDD	CCCCDCDDDC	RRSSRSPPPP	RSSSPSPPTS
3	CCCCCDDDD	CCCCCCCCC	RSRRSPPTTP	RRRSRSRRRR
4	CCCCCDDDD	DCCDCCCCC	SRSSSPPTTP	TSSTRRRRRR
5	CCCCDDDDDD	CCCCCCCCC	RSSSPPPPT	RRRRRRRRRR
6	CDDDDDDDD	DDDDDDDDCC	STPPTTPPPP	PTPPTPPPSR
7	DCCDDDDDD	DDDDDDDDCC	PSRRTPTTTP	TTPTPPPPSS
8	DDCCDDDDDD	CCCCCCCDD	PPPSRPPPPP	SRRRRSRPP
9	DCDDCDDDD	CCCCCCCDD	TSPTSPPPPP	RSSSRSSPP
10	DDDDDDDDDD	CCCCCCCDD	PPTPPTTPPP	RRRSRSPPPP
11	DDDDDDDDDD	DDDDCCCCC	PPPPPPPTTP	TPTPSRRSRS

Примечание. Жирным выделены места изменения стратегии в кодировании, учитывающей действия второго игрока.

Все это позволяет заключить, что описание стратегий, демонстрируемых игроками-людьми, невозможно описать при помощи простых марковских цепей, как это делается в случае простых реактивных стратегий [18].

До социализации мы видим только стратегию Grudger – переход с «С» на «D». Однако после социализации встречается три случая перехода «D» в «С» (см. табл. 2). Все эти три случая выглядят парадоксально – переключение следует после двух или трех случаев хода «Т» – встречи с «кооперативным игроком», которая приносит максимальный выигрыш. Можно полагать, что возникающее чувство вины диктует игроку смену поведения, и такой исход характерен для игрока, находящегося под влиянием экспериментального воздействия. Тип рациональности «Наибольший выигрыш» сменяется типом «Ценность кооперации».

Случай 1 из табл. 2 – самый интересный, так как показывают игровое поведение единственного участника эксперимента, меняющего стратегию однократным переключением с кооперации на «отказ». Но до социализации он сделал это на шестом ходу, а после – на восьмом.

Поиск стратегий Win-Stay Lost-Shift

Стратегия Win-Stay Lost-Shift (WSLS) также известна как стратегия PAVLOV [1]. Действия игрока продиктованы здесь сохранением состояния, которое приносит наибольшее количество баллов – в нашем случае 5 или 10. Если действие оппонента на текущем ходе приводит к «проигрышу», следующий делается с противоположным значением.

Таблица 3

Повторяемость стратегии WSLS до и после социализации

№	Кодирование CD		Кодирование RPTS (CC DD DC CD)	
	До социализации	После социализации	До социализации	После социализации
1	CCDCCDCDC	CCDCDCCDC	RSPRSSPRPS	RSPSPSPRS
2	CCDCDDDDCD	CDDCDDCDD	RSPSTPPSP	STPSPSPSP
3	DCDCDDDCDC	DCCCCDDDC	PRPSTPPSPS	TRRRRSTTR

В чистом виде мы находим эту стратегию только один раз: «RRSTPSPSPS» в первой серии игры, до социализации, и один раз после: «RSTPRRRST». Однако если снизить требования к «чистоте» проявления действий игрока, предусмотрев возможность одной «мутации», то таких случаев мы обнаруживаем уже 5 в каждой серии игр. Если пойти еще дальше и допустить две «мутации», то общее число случаев возрастет до 18 и 17 соответственно. При этом доля случаев, когда такая стратегия, пусть и не в «чистом» виде, встречалась до и после социализации, составляет 1% (табл. 3)

Поиск стратегий TFT

Можно с уверенностью говорить о том, что самой популярной стратегией у исследователей IPD стала стратегия TFT. Об этом можно судить и по исходу уже ставшего историческим турнира, проведенного Р. Аксельродом в 1981 г., а также по тому количеству вариаций, которые появились на основе этой стратегии: GTFT, PTDT, ATFT, TF2T, T2FT, STFT, HTFT, CTFT, RTFT, OTFT [21].

Основной вариант этой стратегии описывается следующим образом: первая итерация – «С», затем в зависимости от исхода предыдущего хода – если «R» или «Т», то «С», если «S» или «P» – то «D».

В чистом виде стратегия TFT встречается в наших данных 1 раз до социализации и 5 раз после. Снижение требований к алгоритмической чистоте приводит к увеличению подходящих случаев. Для допуска в одну мутацию мы находим 8 случаев до социализации и ни одного случая после. В случае с двумя мутациями – 12 до социализации и 7 после. Найдено также два варианта, при которых игрок до и после социализации использовал TFT. При этом один из игроков в последней игре сделал это точно по алгоритму: «SSTRSPPRST» и «RRSPPTRSP». Следует также отметить, что в половине случаев игроки, демонстрировавшие использование TFT до экспериментального воздействия, действовали затем как последовательные кооператоры. Аналогичный результат мы получили ранее и для стратегии

WSLS. Но мы видим, что после социализации количество игроков, придерживающихся этих стратегий, значительно не изменилось. Это объясняется тем, что пул игроков, использующих после экспериментального воздействия стратегии GRIM, WSLS и TFT, пополнился за счет перехода с других стратегий. При этом в половине случаев первой стратегией являются ALLD и ALLDM. Стратегия TFT здесь является исключением – только два игрока, использующие стратегию «всегда предавать», стали использовать во второй серии игр стратегию «око за око».

На примерах со стратегиями GRIM и WSLS мы видели, что их «буквальные» проявления встречаются у реальных игроков только в единичных случаях. В случае с TFT мы видим несколько иную картину – влияние экспериментального воздействия с очевидностью повышает вероятность применения данной стратегии. Причину этого мы склонны искать в том факте, что действия, подразумеваемые данной стратегией, укоренились в культуре гораздо сильнее каких-либо других моральных законов. Последнее утверждение носит, скорее, гипотетический характер, так как для его статистической поддержки нам не хватает полученных данных. Однако даже в статусе гипотезы оно кажется достаточно ценным для того, чтобы целенаправленно работать над его подтверждением в дальнейшей работе.

Повышение доли кооперативных ходов в игре, проводившейся после экспериментального воздействия, приводит к еще одному интересному эффекту. Мы проанализировали частоту ходов, которые можно интерпретировать как «Прощение», «Наказание» и «Отказ от сотрудничества», и обнаружили, что после экспериментального воздействия их величины во всех трех случаях достаточно заметно снизились: в случае с «Прощением» – с 405 до 251 случая, с «Наказанием» – с 336 до 202 случаев, «Отказ от сотрудничества» – с 126 до 85. «Прощение» оппонента мы фиксировали в том случае, если предыдущий ход «Р» или «Т», то следующий – «R» или «S», «Наказание» – если предыдущий ход «S», то следующий – «Т» или «Р», «Отказ от сотрудничества» – предательство после сотрудничества – если предыдущий ход был «R», то следующий «Р» или «Т».

Попытка определить, в какой степени « типовые » стратегии охватывают решения, продемонстрированные участниками нашего эксперимента, позволила объяснить 145 из 300 решений до экспериментального воздействия и 211 из 300 после него. Здесь мы имеем возможность говорить о том, что экспериментальное воздействие привело к существенному повышению предсказуемости действий игроков. Однако остается актуальной проблема идентификации тех стратегий, которые оказались неопознанными при помощи основных стратегий. Для повышения вероятности распознавания каждого случая мы воспользовались классификацией, построенной на основе нейронной сети – трехслойного перцептрона с одним скрытым слоем.

Построение модели нейронной сети для определения стратегии

Для построения модели классификатора мы использовали библиотеку Axelrod-python, содержащую реализацию более 100 различных стратегий. На первом этапе мы выбрали только восемь из них. Были использованы

стратегии Cooperator, Defector, TitForTat, WinStayLoseShift, Prober, Random ($p = 0,5$), Grudger, OppositeGrudger.

Стратегия Prober предлагает следующий алгоритм действий: игрок начинает с ходов (D, C, C), затем кооперация (C) всегда, если противник сотрудничал на втором и третьем шаге, иначе применяет стратегию «око за око». Стратегия Random: игрок случайным образом выбирает между сотрудничеством и отказом.

При помощи специально написанного для этих целей скрипта в процессе проведения турнира между ботами, оперирующими этими стратегиями вносились мутации по схеме: от 0,0 до 0,09 с шагом 0,1. Для получения достаточного разнообразия решений турнир проводился достаточно большое количество раз – не менее 100 000 ходов. Результаты игры кодировались в соответствии со схемой «RSTP». Варианты с повторяющимися значениями отбрасывались, так как они могли быть источником неопределенности вследствие того, что становились результатом встречи различных стратегий. Например, ALLC и ALLC, а также ALLC и TFT дадут идентичный результат. На выходе получаем файл с набором строк, включающих результат игры и указание того, какие стратегии сыграли.

Таблица 4

Результат работы классификатора на основе нейронной сети

№	Стратегии	До социализации		После социализации	
		частота	%	частота	%
1	Defector	129	43,00	68	22,67
2	Random: 0,5	75	25,00	36	12
3	Win-Stay Lose-Shift	33	11,00	26	8,67
4	Tit For Tat	31	10,33	34	11,33
5	Prober	14	4,67	11	3,67
6	Cooperator	11	3,67	120	40
7	Grudger	6	2,00	4	1,33
8	Opposite Grudger	1	0,33	1	0,33

Таблицу с результатами игры использовали для построения нейронной сети. Для этого используется Statistica For Windows, модуль Net, позволяющий автоматизировать процесс подбора оптимальной модели. Из полученных результатов выбиралась модель с наилучшими параметрами, которая затем становилась основанием для классификации данных реальных пользователей. Для тренировки сети мы использовали 80% результатов, 20% привлекались для тестирования. Наиболее эффективной оказалась сеть с 13 скрытыми слоями, она показала вероятность классификации 83,68% обучающей выборки и 82,17 тестовой. Относительно невысокий процент вероятности распознавания обучающей выборки объясняется тем, что результаты работы некоторых стратегий в условиях большого количества мутаций могут совпадать. Особенно это заметно на коротких последовательностях ходов, как это и было в нашем случае. Результаты классификации показаны в табл. 4. Очевидно, что состав стратегий, используемых для построения модели классификации, может быть изменен и расширен.

Обсуждение результатов

Использование информации о стратегии решений задач IPD наряду со сведениями о результативности действий (количество кооперативных и некооперативных ходов) повышает эффективность модели предсказания действий пользователя на основании его социальных характеристик.

В качестве единицы анализа эффективно использование даже минимальной последовательности, состоящей из двух ходов.

Социализация повышает частоту кооперативных стратегий. В большей степени это касается увеличения доли стратегии ALLC.

Мы показали, что экспериментально вызванная социализация может быть рассмотрена как механизм повышения степени определенности в выборе стратегий при решении IPD. Модели на основе нейронных сетей оказываются более работоспособны после экспериментально вызванной социальной идентичности в группе из шести человек.

Этот результат можно также обсуждать с точки зрения разработки технологии повышения предсказуемости действий людей, решающих социальную дилемму типа IPD. Такой вывод может стать точкой роста в дальнейших исследованиях с целью разработки технологий повышения предсказуемости действий

Гендерная идентичность может быть спрогнозирована в незначительной степени (не более 66% событий в тестовой выборке категоризируется правильно), однако социализация повышает процент успешно прогнозируемых событий. Кроме того, высокий уровень (около 90%) прогнозов в тренировочной выборке позволяет надеяться на улучшение результатов в тестовой выборке по мере увеличения количества испытуемых, привлекаемых в эксперимент.

Наименее эффективны модели на основе нейронных сетей в случае прогноза принадлежности испытуемого к гендерной группе. Здесь результаты не очень отличаются от случайных.

IPD может рассматриваться как своеобразный психодиагностический инструмент, способный оценить степень организации мышления человека, а также влияние на эту организацию социального контекста. Учитывая дефицит психодиагностических средств, относящихся к темам «решение задач», «социальный интеллект», развитие IPD как средства психологической диагностики может быть рассмотрено как актуальное направление развития технологий использования социальных дилемм.

Решение задачи IPD человеком не может быть адекватно описано марковскими процессами вследствие того, что решение принимается на основании не только результата предыдущего хода, но и ряда других факторов. Человек может принимать во внимание результаты нескольких предыдущих ходов. Мы видим, что причиной отказа от кооперации может быть негативный опыт, полученный в двух, трех, или даже четырех итерациях. Очевидно, решение нужно искать в обобщенных марковских моделях.

При решении задач IPD реальными людьми возможно говорить о некоторой «суперстратегии», учитывающей не только эволюционные свойства различных «чистых» стратегий, но и различные социальные факторы. Некоторыми такими факторами мы можем управлять, конкретизируя свойства подобных «суперстратегий».

Литература

1. Nowak M., Sigmund K. Chaos and the evolution of cooperation // *Proceedings of the National Academy of Sciences*. 1993. 90 (June). P. 5091–5094. URL: <http://doi.org/10.1073/pnas.90.11.5091>
2. Baek S., Kim B. Intelligent tit-for-tat in the iterated prisoner's dilemma game // *Physical Review E*. 2008. Vol. 78. 011125. URL: <http://doi.org/10.1103/PhysRevE.78.011125>
3. Dugatkin L.A. Dynamics of the TIT FOR TAT strategy during predator inspection in the guppy (*Poecilia reticulata*) // *Behavioral Ecology and Sociobiology*. 1991. Vol. 29 (2). P. 127–132. URL: <http://doi.org/10.1007/BF00166487>
4. Eiser J.R., Bhavnani K. The effect of situational meaning on the behaviour of subjects in the Prisoner's Dilemma Game // *European Journal of Social Psychology*. 1970. Vol. 4 (1). P. 93–97. URL: <http://doi.org/10.1002/ejsp.2420040108>
5. Golbeck J. Evolving Strategies for the Prisoner's Dilemma // *Advances in Intelligent Systems, Fuzzy Systems, and Evolutionary Computation*. 2002. P. 299–306.
6. Kuhlman D.M., Marshello A.F. Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma // *Journal of Personality and Social Psychology*. 1975. Vol. 32 (5). P. 922–931. URL: <http://doi.org/10.1037/0022-3514.32.5.922>
7. Milinski M. TIT FOR TAT in sticklebacks and the evolution of cooperation // *Nature*. 1987. Vol. 325. P. 433–435. URL: <http://doi.org/10.1038/325433a0>
8. Nowak M., Sigmund K. Tit for tat in heterogenous populations // *Nature*. 1992. Vol. 355. P. 250–253. URL: <http://doi.org/10.1038/315250a0>
9. Segal U., Sobel J. Tit for tat: Foundations of preferences for reciprocity in strategic settings // *Journal of Economic Theory*. 2007. Vol. 136, is. 1. P. 197–216. URL: <https://doi.org/10.1016/j.jet.2006.07.003>
10. Smith N.S., Vernon C.R., Tarte R.D. Random Strategies and Sex Differences in the Prisoner's Dilemma Game // *Journal of Conflict Resolution*. 1975. Vol. 19 (4). P. 643–650. URL: <http://doi.org/10.1177/002200277501900405>
11. Wilson W. Reciprocation and other techniques for inducing cooperation in the Prisoner's Dilemma game // *Journal of Conflict Resolution*. 1971. Vol. 15 (2). P. 167–195. URL: <http://doi.org/10.1177/002200277101500205>
12. Borges P.S.S., Pacheco R.C.S., Barcia R.M., Khator S.K. A fuzzy approach to the prisoner's dilemma // *Biosystems*. 1997. Vol. 41 (2). P. 127–137. URL: [http://doi.org/10.1016/S0303-2647\(96\)01667-X](http://doi.org/10.1016/S0303-2647(96)01667-X)
13. Macy M.W. PAVLOV and the Evolution of Cooperation // *Social Psychology Quarterly*. 1995. Vol. 58 (June). P. 74–87.
14. Nowak M., Sigmund K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game // *Nature*. 1993. Vol. 364 (6432). P. 56–58. URL: <http://doi.org/10.1038/364056a0>
15. Matsushima M., Ikegami T. Evolution of strategies in the three-person iterated prisoner's dilemma game // *Journal of Theoretical Biology*. 1998. Vol. 195 (1). P. 53–67. URL: <http://doi.org/10.1006/jtbi.1998.0780>
16. Brunauer R., Andreas L., Mayer H.A., Mitterlechner G., Payer H. Evolution of Iterated Prisoner's Dilemma Strategies with Different History Lengths in Static and Cultural Environments // *Evolution*. 2007. P. 720–727. URL: <http://doi.org/10.1145/1244002.1244163>

17. Ishibuchi H., Namikawa N. Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing // IEEE Transactions on Evolutionary Computation. 2005. Vol. 9 (6). P. 552–561. URL: <http://doi.org/10.1109/TEVC.2005.856198>
18. Ishibuchi H., Ohyanagi H., Nojima Y. Evolution of strategies with different representation schemes in a spatial iterated prisoner's dilemma game // IEEE Transactions on Computational Intelligence and AI in Games. 2011. Vol. 3 (1). P. 67–82. URL: <http://doi.org/10.1109/TCIAIG.2011.2109718>
19. Chen Y.S., Lin H., Wu C.X. Evolution of prisoner's dilemma strategies on scale-free networks // Physica A: Statistical Mechanics and Its Applications. 2007. Vol. 385 (1). P. 379–384. URL: <http://doi.org/10.1016/j.physa.2007.06.008>
20. Яминов Р.И. Взаимосвязь стратегий участников лабораторных экспериментов при добавлении социальной составляющей с их психологическими характеристиками // Труды Московского физико-технического института. 2017. Т. 9, № 3 (35). С. 98–104.
21. Nowak M., Sigmund K. Game-dynamical aspects of the prisoner's dilemma // Applied Mathematics and Computation. 1989. Vol. 30 (3). P. 191–213. URL: [http://doi.org/10.1016/0096-3003\(89\)90052-0](http://doi.org/10.1016/0096-3003(89)90052-0)

Поступила в редакцию 15.08.2020 г.; принята 30.11.2020 г.

Баланев Дмитрий Юрьевич – кандидат психологических наук, заведующий лабораторией экспериментальной психологии, декан факультета психологии Томского государственного университета.
E-mail: balanevd@gmail.com

For citation: Balanev, D.J. Space Modeling of Problem Solving Strategies of “Prisoner's Dilemma”. *Sibirskiy Psikhologicheskii Zhurnal – Siberian journal of psychology*. 2020; 78: 99–116. doi: 10.17223/17267080/78/6. In Russian. English Summary

Space Modeling of Problem Solving Strategies of “Prisoner's Dilemma”

D.J. Balanev^a

^a *Tomsk State University, 36, Lenin Ave., Tomsk, 634050, Russian Federation*

Abstract

An iterated version of the game "Prisoner's Dilemma" is used as a model of cooperation largely due to the wide range of strategies that the subjects can use. The problem of the effectiveness of strategies for solving the Iterated Prisoner's Dilemma (IPD) is most often considered from the point of view of information models, where strategies do not take into account the relationship that arise when real people play. Some of these strategies are obvious, others depend upon social context. In our paper, we use one of the promising directions in the development of studying IPD strategies – the use of artificial neural networks. We use neural networks as a modeling tool and as a part of game environment.

The main goal of our work is to build an information model that predicts the behavior of an individual person as well as group of people in the situation of solving of social dilemma. It takes into account social relationship, including those caused by experimental influence, gender differences, and individual differences in the strategy for solving cognitive tasks. The model demonstrates the transition of individual actions into socially determined behavior. Evaluation of the effect of socialization associated with the procedure of the game provides additional information about the effectiveness and characteristics of the experimental impact.

The paper defines the minimum unit of analysis of the IPD player's strategy in a group, the identity with which can be considered as a variable. It discusses the influence of the experimentally formed group identity on the change of preferred strategies in social dilemmas. We use the possibilities of neural networks as means of categorizing the results of the prisoner's iterative dilemma in terms of the strategy applied by the player, as well as social factors. We define the patterns of changes in the IPD player's strategy before and after socialization are determined. The paper discusses the questions of real player's inclination to use IPD solution strategies in their pure form or to use the same strategy before and after experimental interventions related to social identity formation.

It is shown that experimentally induced socialization can be considered as a mechanism for increasing the degree of certainty in the choice of strategies when solving IPD task. It is found out that the models based on neural networks turn out to be more efficient after experimentally evoked social identity in a group of 6 people; and the models based on neural networks are least effective in the case of predicting a subject's belonging to a gender group. When solving IPD problems by real people, it turns out to be possible to talk about generalized strategies that take into account not only the evolutionary properties of «pure» strategies, but also reflect various social factors.

Keywords: strategies of solving problems; information modeling; the prisoner's dilemma; neural network; socialization; a game.

References

1. Nowak, M. & Sigmund, K. (1993) Chaos and the evolution of cooperation. *Proceedings of the National Academy of Sciences*. 90. pp. 5091–5094. DOI: 10.1073/pnas.90.11.5091
2. Baek, S. & Kim, B. (2008) Intelligent tit-for-tat in the iterated prisoner's dilemma game. *Physical Review E*. 78. 011125. DOI: 10.1103/PhysRevE.78.011125
3. Dugatkin, L.A. (1991) Dynamics of the TIT FOR TAT strategy during predator inspection in the guppy (*Poecilia reticulata*). *Behavioral Ecology and Sociobiology*. 29(2). pp. 127–132. DOI: 10.1007/BF00166487
4. Eiser, J.R. & Bhavnani, K. (1970) The effect of situational meaning on the behaviour of subjects in the Prisoner's Dilemma Game. *European Journal of Social Psychology*. 4(1). pp. 93–97. DOI: 10.1002/ejsp.2420040108
5. Golbeck, J. (2002) Evolving Strategies for the Prisoner's Dilemma. In: Grmela, A. & Mastoraki, N.E. (eds) *Advances in Intelligent Systems, Fuzzy Systems, and Evolutionary Computation*. WSEAS press. pp. 299–306.
6. Kuhlman, D.M. & Marshello, A.F. (1975) Individual differences in game motivation as moderators of preprogrammed strategy effects in prisoner's dilemma. *Journal of Personality and Social Psychology*. 32(5). pp. 922–931. DOI: 10.1037/0022-3514.32.5.922
7. Milinski, M. (1987) TIT FOR TAT in sticklebacks and the evolution of cooperation. *Nature*. 325. pp. 433–435. DOI: 10.1038/325433a0
8. Nowak, M. & Sigmund, K. (1992) Tit for tat in heterogenous populations. *Nature*. 355. pp. 250–253. [Online] Available from: <http://doi.org/10.1038/315250a0>
9. Segal, U. & Sobel, J. (2007) Tit for tat: Foundations of preferences for reciprocity in strategic settings. *Journal of Economic Theory*. 136(1). pp. 197–216. DOI: 10.1016/j.jet.2006.07.003
10. Smith, N.S., Vernon, C.R. & Tarte, R.D. (1975) Random Strategies and Sex Differences in the Prisoner's Dilemma Game. *Journal of Conflict Resolution*. 19(4). pp. 643–650. DOI: 10.1177/002200277501900405
11. Wilson, W. (1971) Reciprocation and other techniques for inducing cooperation in the Prisoner's Dilemma game. *Journal of Conflict Resolution*. 15(2). pp. 167–195. DOI: 10.1177/002200277101500205

12. Borges, P.S.S., Pacheco, R.C.S., Barcia, R.M. & Khator, S.K. (1997) A fuzzy approach to the prisoner's dilemma. *Biosystems*. 41(2). pp. 127–137. DOI: 10.1016/S0303-2647(96)01667-X
13. Macy, M.W. (1995) PAVLOV and the Evolution of Cooperation. *Social Psychology Quarterly*. 58. pp. 74–87.
14. Nowak, M. & Sigmund, K. (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*. 364(6432). pp. 56–58. DOI: 10.1038/364056a0
15. Matsushima, M. & Ikegami, T. (1998) Evolution of strategies in the three-person iterated prisoner's dilemma game. *Journal of Theoretical Biology*. 195(1). pp. 53–67. DOI: 10.1006/jtbi.1998.0780
16. Brunauer, R., Andreas, L., Mayer, H.A., Mitterlechner, G. & Payer, H. (2007) Evolution of Iterated Prisoner's Dilemma Strategies with Different History Lengths in Static and Cultural Environments. *Evolution*. March. pp. 720–727. [Online] Available from: <http://doi.org/10.1145/1244002.1244163>
17. Ishibuchi, H. & Namikawa, N. (2005) Evolution of iterated prisoner's dilemma game strategies in structured demes under random pairing in game playing. *IEEE Transactions on Evolutionary Computation*. 9(6). pp. 552–561. DOI: 10.1109/TEVC.2005.856198
18. Ishibuchi, H., Ohyanagi, H. & Nojima, Y. (2011) Evolution of strategies with different representation schemes in a spatial iterated prisoner's dilemma game. *IEEE Transactions on Computational Intelligence and AI in Games*. 3(1). pp. 67–82. DOI: 10.1109/TCIAIG.2011.2109718
19. Chen, Y.S., Lin, H. & Wu, C.X. (2007) Evolution of prisoner's dilemma strategies on scale-free networks. *Physica A: Statistical Mechanics and Its Applications*. 385(1). pp. 379–384. DOI: 10.1016/j.physa.2007.06.008
20. Yaminov, R.I. (2017) Relations between the laboratory experiments participants' strategy in games with added social component and their psychological characteristics. *Trudy Moskovskogo fiziko-tekhnicheskogo instituta*. 3(35). pp. 98–104. (In Russian).
21. Nowak, M. & Sigmund, K. (1989) Game-dynamical aspects of the prisoner's dilemma. *Applied Mathematics and Computation*. 30(3). pp. 191–213. DOI: 10.1016/0096-3003(89)90052-0

Received 15.08.2020; Accepted 30.11.2020

Dmitry J. Balanев – Head of the Laboratory of Experimental Psychology, Dean of the Faculty of Psychology of Tomsk State University. Cand. Sc. (Psychol.).

E-mail: balanevd@gmail.com