УДК 519 *Автоматика и телемеханика. 2012. № 1. С. 3 – 18.*

Обзоры

Д.А. Новиков, чл.-к. РАН
(Институт проблем управления им. В.А. Трапезникова РАН, Москва)

МОДЕЛИ СТРАТЕГИЧЕСКОЙ РЕФЛЕКСИИ

Проведен обзор основных теоретических и экспериментальных подходов теории игр и теории коллективного поведения к моделированию и исследованию стратегической рефлексии взаимодействующих агентов.

MODELS OF STRATEGIC BEHAVIOR

D.A. Novikov
(Trapeznikov Institute of Control Sciences, RAS, Moscow, novikov@ipu.ru)

The paper contains the survey of main theoretical and experimental approaches to modeling and research of strategic behavior in game theory and collective behavior theories.
Keywords: decision-making, collective behavior, game theory, reflexive games, strategic and informational reflexion, *k*-level players, cognitive hierarchies, reflexive control.

1. ВВЕДЕНИЕ

Традиционно в теоретико-игровых моделях и/или в моделях принятия коллективных решений используется одно из двух предположений о взаимной информированности агентов. Либо считается, что вся существенная информация и принципы принятия агентами решений всем им известны, всем известно, что всем это известно и т. д. до бесконечности (так называемая концепция *общего знания*, используемая, например, при определении равновесия Нэша [1]). Либо предполагается, что каждый агент в рамках своей информированности следует некоторой процедуре принятия индивидуальных решений и почти «не задумывается» над тем, что знают и как ведут себя остальные агенты. Первый подход является каноническим для *теории игр* (см., например, [1-3]), второй – для моделей *коллективного поведения* (см., например, [4-6]). Но между двумя этими «крайностями» существует достаточно большое разнообразие возможных ситуаций. Предположим, что некоторый агент осуществил акт *рефлексии* – попытался спрогнозировать поведение других агентов и выбирает свои действия с учетом этого прогноза (будем считать, что такой агент обладает первым рангом рефлексии). Другой агент (обладающий вторым рангом рефлексии) может знать о существовании агентов первого ранга и прогнозировать их поведение. И так далее. Возникает ряд вопросов: «Как поведение коллектива агентов зависит от их распределения по рангам рефлексии, т. е. от того, сколько в коллективе имеется агентов того или иного ранга? Если долями рефлексирующих агентов можно управлять, то каковы эти доли, оптимальные с точки зрения того или иного критерия эффективности, определенного на множестве действий агентов?»

В «классических» теоретико-игровых моделях предполагается, что в игре в нормальной форме агенты выберут равновесные по Нэшу действия. Однако исследования в области *экспериментальной экономики*[1] (experimental economics) свидетельствуют, что это далеко не всегда так (см., например, [7] и обзор [8]). Возможных объяснений отличиям поведения, наблюдаемого в экспериментах, от предсказанного теорией, может быть несколько:

[1] В России сегодня существуют несколько лабораторий экспериментальной экономики в вузах и академических институтах, например: МФТИ-ВЦ РАН – www.eelab.ru, ГУ ВШЭ – http://epee.hse.ru, РЭШ – http://fir.nes.ru/~abremzen/experiments, ЦЭМИ РАН – http://www.cemi.rssi.ru/structure/science_divisions/lab101.php.

– ограниченность когнитивных возможностей агентов [1, 9, 10] (вычисление, тем более децентрализованное, равновесия Нэша трудоемко[2] [11]);

– необходимость уверенности каждого агента в том, что все его оппоненты могут вычислить равновесие Нэша и сделают это;

– неполная информированность;

– наличие нескольких равновесий.

Таким образом, существуют как минимум два основания (описанных выше – «теоретическое» и «экспериментальное») для рассмотрения моделей коллективного поведения агентов, обладающих различными рангами рефлексии.

Структура изложения материала настоящего обзора следующая: во втором разделе кратко описаны «крайности» – подходы теории игр и теории коллективного поведения к построению моделей рационального поведения и коллективного принятия решений. Третий и четвертый разделы содержат изложение «промежуточных» концепций – соответственно информационной и стратегической рефлексии. Пятый раздел посвящен структуризации предметной области, которую можно условно назвать «рефлексия в теории игр и моделях коллективного поведения», а также проведению параллелей между стратегической и информационной рефлексией. Шестой раздел содержит классификацию и краткий обзор частных моделей стратегической рефлексии. В седьмом разделе формулируется задача рефлексивного управления – выбора оптимального разбиения агентов на ранги стратегической рефлексии, в восьмом разделе описаны результаты решения этой задачи для некоторых прикладных ситуаций. В заключении обсуждаются основные результаты и перспективы дальнейших исследований.

2. МОДЕЛИ РАЦИОНАЛЬНОГО ПОВЕДЕНИЯ

Описание модели. Рассмотрим множество $N = \{1, \ldots, n\}$, состоящее из *n агентов*. Агент i выбирает свое *действие* $x_i \in \Re^1$ (для простоты здесь и ниже, если не оговорено особо, считается, что ограничения на действия агентов отсутствуют). Вектор $x = (x_1, \ldots, x_n) \in \Re^n$ действий агентов, называемый *ситуацией игры* [2], определяет их выигрыши, задаваемые *целевыми функциями* $F_i(x)$, где $F_i(\cdot)$: $\Re^n \to \Re^1$, $i \in N$.

Рациональность поведения агента заключается в стремлении к максимизации своей целевой функции выбором собственного действия [2]:

(1) $x_i \in BR_i(x_{-i}) = \operatorname{Arg} \max\limits_{y \in \Re^1} F_i(y, x_{-i}),\ i \in N,$

где $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \Re^{n-1}$ – *обстановка игры* для i-го агента, $BR_i(\cdot)$ – его *наилучший ответ* (best response) [3], $i \in N$. Предположим, что функции $F_i(\cdot)$ таковы, что для любого агента при любой обстановке игры существует единственный наилучший ответ.

Теория игр. Из выражения (1) следует, что наилучший ответ каждого агента зависит в общем случае от обстановки игры, поэтому трудно однозначно сказать априори, какое действие выберет конкретный агент, так как он вынужден предсказывать поведение оппонентов – устранять так называемую *игровую неопределенность* [13]. Основным предметом *теории игр* [1-3] является поиск *равновесия* (решения игры), определяемого как устойчивый в том или ином (оговариваемом в каждом конкретном случае) смысле исход взаимодействия агентов – вектор их равновесных действий [2].

Введение определенных предположений об *информированности* агентов (о той информации, которой они обладают на момент выбора действий) приводит к соответствующим концепциям равновесия. Так, например, если считать, что каждый агент ориентируется на наихудшую для него обстановку игры, то получим равновесие в *гарантирующих стратегиях* [2]. Если считать, что все описание игры (множество агентов, их целевые функции и множества допустимых действий) являются *общим знанием* [1] среди агентов (для таких агентов в англоязычной литературе используется термин super-intelligent players), принимающих решения однократно, одновременно и независимо, то можно использовать концепцию *равновесия Нэша* в игре $\Gamma = \{N, F_i(\cdot)_{i \in N}\}$ в нормальной форме. Равновесием

[2] Равновесие Нэша не всегда адекватно описывает реальное поведение агентов в лабораторных экспериментальных одношаговых играх, в том числе потому, что агенты не успевают «исправить» свои неправильные представления о существенных параметрах игры [12] – например, концепция рационализуемых стратегий Д. Бернхейма требует от агентов неограниченной рациональности (высоких когнитивных возможностей).

Нэша называется такой вектор действий x^N, что действие каждого агента $i \in N$ является наилучшим ответом на Нэшевскую обстановку x_{-i}^N:

(2) $x_i^N \in BR_i(x_{-i}^N)$.

Коллективное поведение. В отличие от теории игр *теория коллективного* (*группового*) *поведения* [4-6] занимается исследованием динамики поведения рациональных агентов при достаточно слабых предположениях относительно их информированности. Так, например, не всегда требуется наличие среди агентов общего знания относительно множества агентов, множеств допустимых действий и целевых функций оппонентов. Или считается, что агенты не предсказывают поведение всех оппонентов, как это имеет место в теории игр (см. выше). Более того, зачастую агенты, принимая решения, могут «не знать о существовании» некоторых других агентов или иметь о них агрегированную информацию.

Наиболее распространенной моделью динамики коллективного поведения является *модель индикаторного поведения* [1, 5, 6], суть которой заключается в следующем. Предположим, что каждый агент в момент времени t наблюдает действия всех агентов $\{x_i^{t-1}\}_{i \in N}$, выбранные ими в предыдущий момент времени $t - 1$, $t = 1, 2, \ldots$ (начальный вектор действий $x^0 = (x_1^0, \ldots, x_n^0)$ считается заданным).

Каждый агент может рассчитать свое *текущее положение цели* – такое его действие, которое максимизировало бы его целевую функцию при условии, что в текущем периоде все агенты выбрали бы те же действия, что и в предыдущем:

(3) $w_i(x_{-i}^{t-1}) = \arg\max\limits_{y \in \Re^1} F_i(y, x_{-i}^{t-1})$, $t = 1, 2, \ldots, i \in N$.

В рамках гипотезы индикаторного поведения каждый агент в каждый момент времени будет делать «шаг» от своего предыдущего действия к текущему положению цели:

(4) $x_i^t = x_i^{t-1} + \gamma_i^t\, [w_i(x_{-i}^{t-1}) - x_i^{t-1}]$, $i \in N$, $t = 1, 2, \ldots$,

где $\gamma_i^t \in [0; 1]$ – «величины шагов». Такое коллективное поведение можно условно назвать «оптимизационным», подчеркивая тем самым его отличие от игрового. Очевидно, что если $\gamma_i^t \equiv 0$, то динамика отсутствует; если $\gamma_i^t \equiv 1$, то каждый агент на каждом шаге выбирает свой наилучший ответ (см. (1)), однако в последнем случае соответствующая динамика может быть неустойчивой. Условия сходимости процедуры (4), области притяжения равновесий, условия на величины шагов $\{\gamma_i^t\}$, обеспечивающие сходимость, и т. д. можно найти в [1, 5, 6, 14].

Подходы теории коллективного поведения и теории игр согласованы в том смысле, что и та, и другая исследуют поведение рациональных агентов (ср. (1) и (4)), а равновесия игры, как правило, являются и равновесиями динамических процедур коллективного поведения (например, равновесие Нэша (2) является равновесием динамики (4) коллективного поведения).

Для полноты картины отметим, что в теории коллективного поведения существует и другой (выходящий за рамки настоящей работы) подход – *эволюционная теория игр* [15], которая исследует «поведение больших однородных групп (популяций) индивидуумов в типичных повторяющихся конфликтных ситуациях, причем каждую стратегию применяет множество игроков, а функция выигрыша характеризует успех отдельных стратегий, а не отдельных участников взаимодействия» [16, с. 296]. Русскоязычный обзор базовых результатов теории эволюционных игр можно найти в [16].

Таким образом, теория игр зачастую использует, условно говоря, максимальные предположения об информированности агентов (например, гипотезу о существовании общего знания), а теория коллективного поведения – минимальные. Промежуточное место занимают рефлексивные модели, поэтому перейдем к обсуждению роли рефлексии – информационной и стратегической – в принятии агентами решений.

3. ИНФОРМАЦИОННАЯ РЕФЛЕКСИЯ

Стратегическая и информационная рефлексия. *Рефлексивной* является *игра*, в которой информированность игроков не является общим знанием. С точки зрения теории игр и рефлексивных моделей принятия решений целесообразно разделять стратегическую и информационную рефлексию [1].

Информационная рефлексия – процесс и результат размышлений игрока о том, каковы значения неопределенных параметров, что об этих значениях знают и думают его оппоненты (другие игроки). При этом собственно «игровая» компонента отсутствует, так как никаких решений игрок не принимает.

Иными словами, информационная рефлексия относится к информированности агента о природной реальности (какова игра) и о рефлексивной реальности (какой видят игру другие).

Информационная рефлексия логически предшествует рефлексии несколько иного рода – стратегической рефлексии.

Стратегическая рефлексия – процесс и результат размышлений игрока о том, какие принципы принятия решений используют его оппоненты (другие игроки) в рамках той информированности, которую он им приписывает в результате информационной рефлексии.

Таким образом, информационная рефлексия имеет место только в условиях неполной информированности, и ее результат используется при принятии решений (в том числе при стратегической рефлексии). Стратегическая рефлексия имеет место даже в случае полной информированности, предваряя принятие игроком решения о выборе действия. Другими словами, информационная и стратегическая рефлексии могут изучаться независимо, однако в условиях неполной информированности обе они имеют место.

Рассмотрим сначала информационную рефлексию, в рамках которой взаимная информированность агентов описывается структурой информированности, а равновесием их игры является информационное равновесие.

Структура информированности. Если при взаимодействии агентов присутствует неопределенный параметр $\theta \in \Omega$ (будем считать, что множество Ω является общим знанием), от значения которого зависят целевые функции $\{f_i(\theta, x)\}_{i \in N}$ агентов, то *структура информированности* I_i (как синоним употребляют термины «информационная структура» и «иерархия представлений») i-го агента включает в себя следующие элементы [1]. Во-первых, представление i-го агента о параметре θ – обозначим его θ_i, $\theta_i \in \Omega$. Во-вторых, представления i-го агента о представлениях других агентов о параметре θ – обозначим их θ_{ij}, $\theta_{ij} \in \Omega$, $j \in N$. В-третьих, представления i-го агента о представлении j-го агента о представлении k-го агента – обозначим их θ_{ijk}, $\theta_{ijk} \in \Omega$, j, $k \in N$. И так далее.

Таким образом, структура информированности I_i i-го агента задается набором всевозможных значений вида $\theta_{ij_1 \ldots j_l}$, где l пробегает множество целых неотрицательных чисел, $j_1, \ldots, j_l \in N$, а $\theta_{i_1 \ldots i_l} \in \Omega$.

Аналогично задается *структура информированности I игры* в целом – набором значений $\theta_{i_1 \ldots i_l}$, где l пробегает множество целых неотрицательных чисел, $j_1, \ldots, j_l \in N$, а $\theta_{ij_1 \ldots j_l} \in \Omega$. Подчеркнем, что структура информированности I «недоступна» наблюдению агентов, каждому из которых известна лишь некоторая ее часть (а именно – I_i).

Обозначим через Σ_+ множество всевозможных конечных последовательностей индексов из N, через Σ – объединение Σ_+ с пустой последовательностью.

Наряду со структурами информированности I_i, $i \in N$, можно рассматривать структуры информированности I_{ij} (структура информированности j-го агента в представлении i-го агента), I_{ijk} и т. д. Отождествляя структуру информированности с характеризуемым ею агентом, можно сказать, что, наряду с *n реальными* агентами (*i-агентами*, где $i \in N$) со структурами информированности I_i, в игре участвуют *фантомные агенты* (*Ψ-агенты*, где $\Psi \in \Sigma_+$, $|\Psi| \geq 2$) со структурами информированности $I_\Psi = \{\theta_{\Psi\sigma}\}$, $\sigma \in \Sigma$. Фантомные агенты, существуя в сознании реальных и других фантомных агентов, влияют на их действия [1].

Помимо структуры I имеется счетное множество структур I_Ψ, $\Psi \in \Sigma_+$, среди которых можно при помощи отношения тождественности выделить классы попарно нетождественных структур. Количество этих классов естественно считать *сложностью структуры информированности*. Будем говорить, что структура информированности I имеет конечную сложность ν, если существует такой конечный набор из ν попарно нетождественных структур, что для любой структуры найдется тождественная ей структура из этого набора.

Таким образом, структура информированности – бесконечное n-дерево, вершинам которого соответствует конкретная информированность реальных и фантомных агентов. *Рефлексивной игрой Γ_I* называется игра, описываемая следующим кортежем: $\Gamma_I = \{N, f_i(\cdot)_{i \in N}, \Omega, I\}$, где N – множество реальных агентов, $f_i(\cdot)$: $\Omega \times \Re^n \to \Re^1$ – целевая функция i-го агента, $i \in N$, Ω – множество возможных значений неопределенного параметра, I – структура информированности [1]. Рефлексивная игра является обобщением понятия игры в нормальной форме, задаваемой кортежем $\{N, f_i(\cdot)_{i \in N}\}$, на случай, когда информированность агентов отражена иерархией их представлений (информационной структурой I). В рамках принятого определения «классическая» игра в нормальной форме является частным случаем рефлексивной игры – игрой с общим знанием.

Информационное равновесие. Если задана структура I информированности игры, то тем самым задана и структура информированности каждого из агентов (как реальных, так и фантомных). Выбор Ψ-агентом своего действия x_Ψ в рамках гипотезы рационального поведения определяется его структурой информированности I_Ψ, поэтому, имея перед собой эту структуру, можно смоделировать его рассуждения и определить это его действие. Выбирая свое действие, агент моделирует действия других агентов (осуществляет рефлексию). Поэтому при определении исхода игры необходимо учитывать действия как реальных, так и фантомных агентов.

Набор действий x_Ψ^*, $\Psi \in \Sigma_+$, называется *информационным равновесием* [1], если выполнены следующие условия:

1) структура информированности I имеет конечную сложность ν;

2) $\forall \lambda, \mu \in \Sigma \ \ I_{\lambda i} = I_{\mu i} \Rightarrow x_{\lambda i}^* = x_{\mu i}^*$;

3) $\forall\ i \in N, \forall\ \sigma \in \Sigma \quad x_{\sigma i}^* \in \operatorname{Arg}\max\limits_{x_i \in X_i} f_i(\theta_{\sigma i}, x_{\sigma i 1}^*, \dots, x_{\sigma i, i-1}^*, x_i, x_{\sigma i, i+1}^* \dots, x_{\sigma i, n}^*)$.

Первое условие в определении информационного равновесия означает, что в рефлексивной игре участвует конечное число реальных и фантомных агентов. Второе условие отражает требование того, что одинаково информированные агенты выбирают одинаковые действия. И, наконец, третье условие отражает рациональное поведение агентов – каждый из них стремится выбором собственного действия максимизировать свою целевую функцию, подставляя в нее действия других агентов, которые оказываются рациональными с точки зрения рассматриваемого агента в рамках имеющихся у него представлений о других агентах.

Таким образом, при рассмотрении информационной рефлексии отказ от предположения о наличии среди агентов общего знания приводит к моделям *рефлексивных игр* [1]. При этом равновесие Нэша «превращается» в более общее *информационное равновесие*, в рамках которого каждый агент осуществляет информационную рефлексию – при принятии решений использует не только свою информацию о существенных параметрах, но и свои представления о представлениях других агентов об этих параметрах, представления о представлениях о представлениях и т. д.

Имея зависимость информационного равновесия от структуры информированности, можно ставить и решать задачи *информационного управления* – формирования допустимых структур информированности, приводящих к наиболее выгодным для управляющего органа информационным равновесиям [17].

Ниже будем считать, что имеет место полная информированность агентов (неопределенные параметры отсутствуют), т. е. сконцентрируем внимание только на стратегической рефлексии[3]. В то же время достаточно полно исследованная на сегодняшний день логика анализа информационной рефлексии может служить «точкой отсчета» и источником аналогий при разработке гораздо менее изученных моделей стратегической рефлексии.

4. СТРАТЕГИЧЕСКАЯ РЕФЛЕКСИЯ: РЕФЛЕКСИВНЫЕ РАЗБИЕНИЯ И ОБЩАЯ МОДЕЛЬ РЕФЛЕКСИВНОГО КОЛЛЕКТИВНОГО ПОВЕДЕНИЯ

Рефлексивные разбиения. В рамках гипотезы индикаторного поведения [5, 6] неявно предполагается, что агент, выбирая свои действия в соответствии с процедурой (4), не задумывается о том, что и другие агенты действуют так же. Если бы он об этом задумался (осуществил рефлексию), то ему следовало бы искать, принимая решения в момент времени t, наилучший ответ на прогнозируемые им в рамках выражения (4) действия других агентов. Т. е. положение цели определялось бы уже не выражением (3), а следующим образом:

(5) $w_i(x_{-i}^t) = \arg\max\limits_{y \in \Re^1} F_i(y, x_{-i}^t)$,

[3] Совместное рассмотрение информационной и стратегической рефлексии представляется перспективным направлением будущих исследований.

где x_{-i}^t определяется выражением (4). При этом можно полагать, что рефлексирующий агент первого ранга считает всех остальных нерефлексирующими (что соответствует принятой в [1, 17-19] традиции считать, что агент, имеющий некоторый ранг стратегической рефлексии, считает всех остальных имеющими ранг на единицу меньше его собственного). Другие возможные предположения о представлениях агентов о рангах рефлексии оппонентов обсуждаются ниже (например, можно считать, что интеллектуальные агенты одного ранга рефлексии разыгрывают равновесие Нэша [20] и т. д.).

Аналогично можно рассматривать агентов и более высоких рангов рефлексии (в англоязычной литературе почти как синонимы термина «ранг рефлексии k» используются термины «step k player», «level k player», «k-level player», «smart k-player» [20-26]). Для этого определим $\aleph = \{N_0, \dots, N_m\}$ – разбиение множества агентов N, где N_i – множество агентов i-го ранга рефлексии, $i = \overline{0, m}$, *m – максимальный ранг рефлексии*, $n_i = |N_i|$, $i \in N$, $\sum_{i=0}^{m} n_i = n$. Назовем $\aleph$ *рефлексивным разбиением* [27].

Будем пока считать, что агент некоторого ранга рефлексии k достоверно знает множества (или долю – см. ниже) агентов всех более низких рангов k' (где $k' < k - 1$) и считает всех агентов своего и бóльших рангов ($k'' \geq k$) имеющими ранг на единицу меньше себя (т. е. ранг $k - 1$). Этим отражается предположение, что агент не допускает существования агентов, имеющих такой же или более высокий ранг рефлексии, чем он сам. При этом агент может неправильно оценивать множества агентов k – 1-го, k-го и более высоких рангов рефлексии.

Пусть задан вектор x^0 *«начальных» действий* агентов. Рассмотрим следующую динамическую модель рефлексивного принятия ими решений, параллельно помня при этом, что соответствующие выражения для одношаговой «игровой» модели могут быть получены как частный случай, в котором решения принимаются однократно при $\gamma_i^1 \equiv 1$, $i \in N$.

Нулевой ранг рефлексии. Будем считать, что агенты с нулевым рангом рефлексии (принадлежащие множеству N_0) выбирают свои действия, считая, что действия остальных агентов будут такими же, что и в предыдущем периоде. Тогда из (4) следует, что

(6) $x_i^t = x_i^{t-1} + \gamma_i^t \, [w_i(x_{-i}^{t-1}) - x_i^{t-1}]$, $i \in N_0$, $t = 1, 2, \dots$.

Если рефлексирующих агентов нет ($N_0 = N$), то в итоге все агенты пронаблюдают *реализованную траекторию* $(x^0, \dots, x^t, \dots)$ векторов действий агентов, определяемых (6).

Первый ранг рефлексии. Агент j, обладающий первым рангом рефлексии ($j \in N_1$), считает всех остальных агентов обладающими нулевым рангом рефлексии и в соответствии с (6) «предсказывает» их выбор. Поэтому его собственный выбор $x1_j^t$ будет ориентирован на наилучший ответ на ту обстановку, которая с его точки зрения должна сложиться:

(7) $x1_j^t = x1_j^{t-1} + \gamma_j^t \, [w_j(x_{-j}^t) - x1_j^{t-1}]$, $j \in N_1$.

Для агента $j \in N_1$ *прогнозируемой* является *траектория* $(x^0, \dots, (x1_j^t, x_{-j}^t), \dots)$, а на самом деле реализуется траектория $(x^0, \dots, (x1_{j \in N_1}^t, x_{i \in N_0}^t), \dots)$. Т. е. реализованная траектория может не совпадать с траекториями, прогнозируемыми как агентами нулевого, так и первого рангов рефлексии. О возможном несовпадении прогнозируемой и реализованной траекторий (и последствиях такого несовпадения) речь пойдет ниже – см. «условие стабильности рефлексивного разбиения» в седьмом разделе.

В качестве отступления отметим, что всюду в настоящей работе считается, что агенты любого ранга рефлексии достаточно «интеллектуальны» – они выбирают действия, стремясь максимизировать свои целевые функции. Можно допустить наличие и менее интеллектуальных агентов – *агентов-имитаторов* (условно обладающих минус первым рангом рефлексии), действия которых определяются известной функцией от текущих или прошлых действий других агентов (примеры: выбор действия, равного среднему арифметическому действий остальных агентов; или агентов, связанных с данным; или некоторого другого фиксированного агента). Наверное, подобные модели могут адекватно описывать такое явление, как диффузия инноваций и др.

Второй ранг рефлексии. Будем считать, что каждый агент j, обладающий вторым рангом рефлексии ($j \in N_2$), знает достоверно множество N_0 и считает всех агентов из множества $N_1 \cup N_2 \setminus \{j\}$ обладающими первым рангом рефлексии (отметим, что в общем случае, когда имеются несколько агентов второго ранга рефлексии, данный агент ошибочно приписывает им первый ранг). В силу этого он может «прогнозировать» поведение всех своих оппонентов. Поэтому его выбор будет наилучшим ответом на ту обстановку, которая с его точки зрения должна сложиться:

(8) $x2_j^t = x2_j^{t-1} + \gamma_j^t\ [w_j(x_{i\in N_0}^t,\ x1_{l\in N_1\cup N_2\setminus\{j\}}^t) - x2_j^{t-1}], j \in N_2$.

Для агента $j \in N_2$ *прогнозируемой* является *траектория* $(x^0, \ldots, (x2_j^t,\ x1_{l\in N_1\cup N_2\setminus\{j\}}^t,\ x_{i\in N_0}^t), \ldots)$, а на самом деле реализуется траектория $(x^0, \ldots, (x2_{j\in N_2}^t,\ x1_{l\in N_1}^t,\ x_{l\in N_0}^t), \ldots)$.

***k*-й ранг рефлексии** ($k \le m$). Поведение агентов k-го ранга рефлексии описывается аналогично рассмотренным выше трем случаям (нулевого, первого и второго рангов рефлексии) с учетом следующей *рефлексивной структуры* агентов. Обозначим через $\aleph_{jk}$ – *субъективное рефлексивное разбиение* – представления агента j, обладающего k-м рангом рефлексии, о разбиении всех агентов на ранги рефлексии:

(9) $\aleph_{jk} = (\underbrace{N_0, N_1, \ldots, N_{k-2}, N_{k-1} \cup N_k \cup \ldots \cup N_m \setminus \{j\}}_{k}, \{j\}, \underbrace{\varnothing, \ldots, \varnothing}_{m-k-1}), j \in N_k$.

Агент k-го ранга будет выбирать действия в соответствии с процедурой

(10) $xk_j^t = xk_j^{t-1} + \gamma_j^t\ [w_j(x_{l\in N_0}^t,\ x1_{l\in N_1}^t, \ldots,\ x[k-1]_{l\in N_{k-1}\cup N_k\cup\ldots\cup N_m/\{j\}}^t) - xk_j^{t-1}], j \in N_k$.

В «статическом» случае агент k-го ранга выберет действие

(11) $xk_j^*(\aleph_{jk}) = \arg\max_{y\in\Re^1} F_j(y,\ x_{l\in N_0}^1,\ x1_{l\in N_1}^1, \ldots,\ x[k-1]_{l\in N_{k-1}\cup N_k\cup\ldots\cup N_m/\{j\}}^1), j \in N_k$.

Итак, *рефлексивная структура* определяется совокупностью субъективных рефлексивных разбиений всех агентов. Если предположить, что представления агентов о рангах рефлексии друг друга описываются выражением (9), то рефлексивная структура однозначно задается рефлексивным разбиением $\aleph$. Отметим, что рефлексивная структура (9), введенная в рамках рассмотрения стратегической рефлексии, в некотором смысле аналогична структуре информированности, используемой в моделях информационной рефлексии – см. выше и [1].

Вектор действий агентов

(12) $x^*(\aleph) = \{ xk_j^*(\aleph_{jk})\}_{j\in N_k,\, k=\overline{0,m}}$

назовем *рефлексивным равновесием* игры $\Gamma_\aleph = \{N, F_i(\cdot)_{i\in N}, \aleph\}$ [28]. В силу введенных выше предположений о существовании и единственности наилучших ответов рефлексивное равновесие всегда существует. Отметим, что рефлексивное равновесие достаточно экзотично в том смысле, что в нем действия агентов в общем случае не являются наилучшими ответами на действия, выбираемые оппонентами.

Изменяя рефлексивное разбиение, можно менять действия агентов, т. е. осуществлять *рефлексивное управление* (см. также теоретико-игровые модели рефлексивного управления в [1, 17]).

Вспомним, что при определении рефлексивного равновесия были введены следующие предположения:

– вектор x^0 начальных действий агентов фиксирован и является общим знанием;

– агент ранга k имеет правильные представления о рангах рефлексии всех агентов, имеющих ранг, меньший его собственного;

– агент ранга k считает всех агентов своего и более высоких рангов имеющими ранг $k-1$.

Можно определять представления агентов о рангах рефлексии оппонентов по-другому (см. подраздел 6.1 ниже); при этом соответствующим образом изменятся выражения для наилучшего ответа, текущего положения цели и рефлексивного равновесия (12), однако общая схема построения последнего сохранится.

Приведенная в настоящем разделе общая модель рефлексивного коллективного поведения вряд ли допускает получение в ее рамках каких-либо столь же общих аналитических выводов. Тем не менее она может служить базисом для создания частных аналитических или общих имитационных моделей (например, в соответствии с классификацией, приведенной в [29]), позволяющих описывать и прогнозировать групповое поведение (людей, мобильных роботов, программных агентов) в разнообразных ситуациях – см., например, рефлексивные имитационные модели эвакуации в [30], рефлексивные модели транспортных потоков [31] и примеры из других прикладных областей в [27].

5. РЕФЛЕКСИЯ В ТЕОРИИ ИГР И МОДЕЛЯХ КОЛЛЕКТИВНОГО ПОВЕДЕНИЯ: СТРУКТУРА ПРЕДМЕТНОЙ ОБЛАСТИ

Теория игр и теория коллективного поведения изучают модели взаимодействия рациональных агентов. Подходы и результаты этих теорий можно рассматривать с точки зрения трех взаимосвязанных гносеологических уровней (соответствующих различным функциям моделирования) – см. рис. 1:

– феноменологического уровня, на котором модель строится с целью описать поведение исследуемой системы (коллектива агентов);

– прогностического уровня (цель – прогноз поведения исследуемой системы);

– нормативного уровня (цель – обеспечение требуемого поведения системы).

Для теории игр традиционной является схема, когда сначала описывается «модель игры» (феноменологический уровень); затем выбирается концепция равновесия, определяющая, что понимается под устойчивым исходом игры (прогностический уровень); после чего может формулироваться та или иная задача управления – выбора значений управляемых «параметров игры», приводящих к реализации требуемого равновесия (нормативный уровень) – см. рис. 1.

Учет информационной рефлексии приводит к необходимости построения и анализа структур информированности, что в итоге дает возможность определить информационное равновесие и в дальнейшем ставить и решать задачи информационного управления – см. рис. 1.

Учет стратегической рефлексии приводит к аналогичной цепочке, выделенной на рис. 1 жирными линиями: «модели стратегической рефлексии» – «рефлексивная структура» – «рефлексивное равновесие» – «рефлексивное управление».

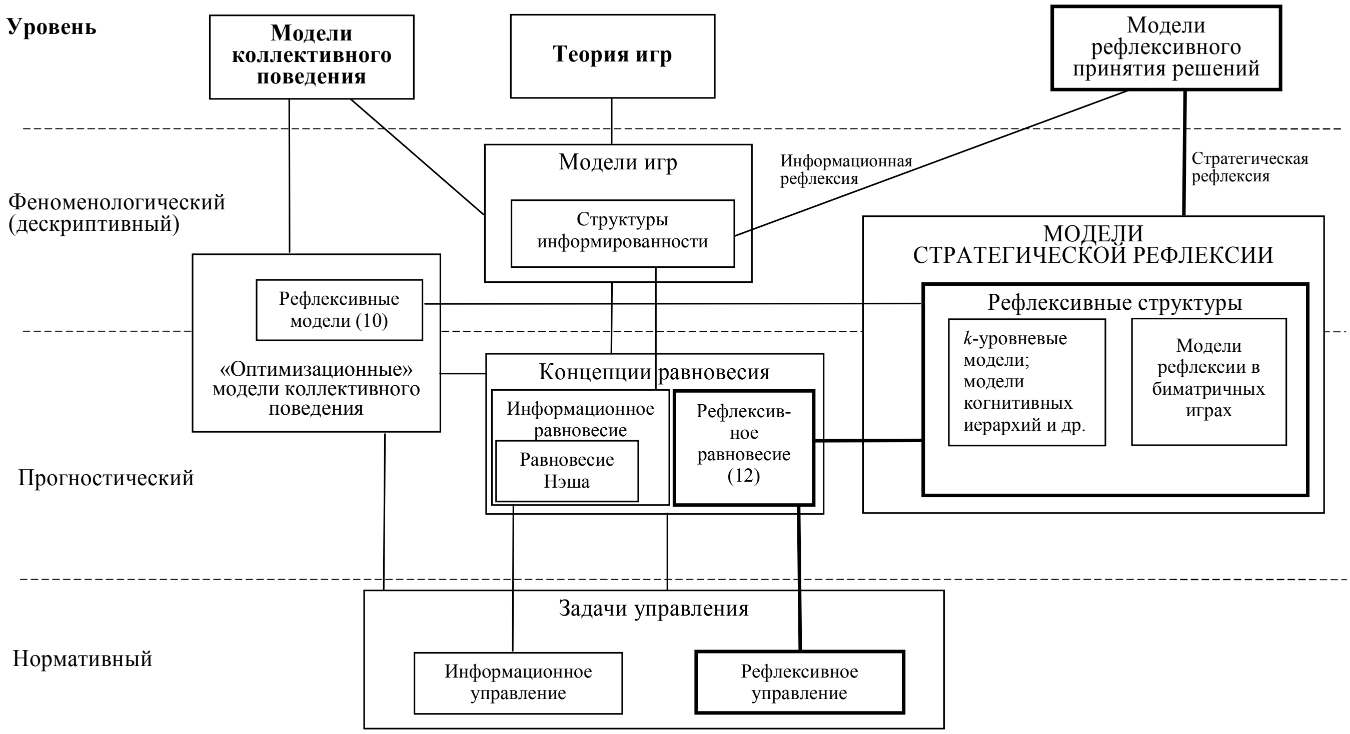

Рис. 1. Дескриптивные и нормативные модели информационной и стратегической рефлексии

Сравнение подходов к моделированию информационной и стратегической рефлексии проводится в следующей таблице 1.

Табл. 1

ПАРАМЕТР	Стратегическая рефлексия	Информационная рефлексия
«Модель игры»	Рефлексивная структура	Структура информированности
Равновесие	Рефлексивное равновесие	Информационное равновесие
Управление	Рефлексивное управление	Информационное управление

6. КЛАССИФИКАЦИЯ И КРАТКИЙ ОБЗОР ЧАСТНЫХ МОДЕЛЕЙ

6.1. Классификации моделей стратегической рефлексии

Обозначим через n_{ijl} представления агента i, обладающего j-м рангом рефлексии, о ранге рефлексии l-го агента. Для случая однородных агентов обозначим через q_{ijk} представления i-го агента j-го ранга рефлексии о доле агентов, имеющих ранг k, $q_k = n_k / n$ – «истинная» доля агентов k-го ранга.

Общий **постулат**, принимаемый практически во всех моделях рефлексивного коллективного поведения: агент некоторого ранга рефлексии «не знает» о существовании других агентов его ранга или более высоких рангов, т. е. $\forall\ k > j\ q_{ijk} = 0$, $q_{ijj} = 1 / n$.

Основания системы классификаций моделей стратегической рефлексии.

1) Множество возможных действий агента (конечно или «бесконечно» – например, отрезок $\Re^1$).

2) Принцип выбора действий агентами нулевого ранга рефлексии:

– фиксированные (априори заданные) действия, например фокальная точка [17];

– наилучший ответ на некоторые фиксированные (априори заданные) действия (например, результаты прошлого периода);

– случайные в соответствии с заданным распределением (как правило, равномерным).

3) Агенты одинаковые (однородные, т. е. различаются только рангами рефлексии) или различные (отличаются еще и целевыми функциями).

4) Распределение (объективное) агентов по рангам рефлексии:

– произвольное фиксированное;

– случайное (в соответствии с вероятностным распределением Пуассона $q_k = e^{\tau}\ \tau^{k} / k!$, где $\tau > 0$ – параметр распределения Пуассона).

5) Информированность агента k-го ранга относительно общего числа (множества) агентов:

– знает множество N достоверно и считает, что эта информация является общим знанием;

– имеет свои представления относительно общего числа (множества) агентов.

Отметим, что практически все известные на сегодняшний день модели рефлексивного коллективного поведения используют первое предположение.

6) Информированность агента k-го ранга относительно агентов более низких рангов (от 0 до $k - 1$ включительно):

– знает достоверно (или с некоторой погрешностью – см. в [32] модель для $m \leq 2$);

– предполагает, что эти агенты распределены по рангам рефлексии от 0 до $k - 1$ включительно в соответствии с некоторым нормированным ($\forall\ k < j\ q_{ijk} = \frac{n_j}{\sum_{l=0}^{j-1} n_l}(n-1)$) вероятностным распределением – как правило, распределением Пуассона) – см., например, [21];

– считает что все (!) остальные агенты имеют ранг $k - 1$ – см., например, модели k-level reasoning [20, 33], а также развитие и дальнейшее экспериментальное исследование этих моделей в [25, 34].

7) Информированность агента k-го ранга относительно других агентов своего и более высоких рангов:

– считает их всех принадлежащих нулевому рангу;

– считает их всех принадлежащих $k - 1$-му рангу;

– предполагает, что эти агенты распределены по рангам рефлексии от 0 до $k - 1$ включительно в соответствии с некоторым вероятностным распределением (как правило – распределением Пуассона);

– знает ранги их рефлексии (отметим, что при этом введенный выше «постулат» не выполнен) и при выборе своего действия устраняет неопределенность относительно их поведения, рассчитывая на выбор ими наихудших для него действий.

Как отмечалось выше, при любых значениях признаков данной системы классификаций рефлексивное равновесие строится по общей схеме, приведенной в четвертом разделе.

Все известные на сегодняшний день зарубежные (см. подразделы 6.2 и 6.3) и отечественные (см. подраздел 6.4) модели стратегической рефлексии укладываются в рамки предложенной системы их классификаций.

6.2. Краткий обзор теоретических зарубежных исследований

За рубежом в середине 90-х г. XX века появились работы по моделям стратегической рефлексии – так называемые *k-уровневые модели* (level k models) [32, 33, 35]. Затем в 2004 году возникло их обобщение – *модель когнитивных иерархий* (Cognitive Hierarchies Model – CHM) [21].

В обзоре [8] выделены четыре основные подхода к построению и исследованию моделей стратегической рефлексии в рамках теории игр и экспериментальной экономики (ниже указаны основополагающие работы, ссылки на их последователей можно найти в обзоре [8]):

– k-уровневый [25];
– равновесий квантового наилучшего ответа [36].
– квантовый k-уровневый [35];
– когнитивных иерархий [21].

Рассмотрим эти четыре подхода. Во всех моделях, во-первых, все агенты одинаковые (имеют одинаковые целевые функции и множества допустимых действий) и отличаются только своими рангами рефлексии. Во-вторых, агент k-го ранга не знает о существовании (не допускает существования) других агентов того же или более высоких рангов рефлексии. В-третьих, агенты нулевого ранга выбирают свои действия случайным образом (в соответствии с равномерным распределением на множестве своих допустимых действий). В ряде экспериментальных работ (см., например, [33, 37-41]) учитывается структура взаимодействия агентов – кто из агентов о чьих действиях имеет информацию.

***k*-уровневые модели.** Поставим в соответствие каждому агенту $i \in N$ ранг его рефлексии – число $k_i \in \{0; 1; 2; \dots\}$. Агент k-го ранга считает всех (!) остальных агентов имеющими ранг $k - 1$ и выбирает свой наилучший ответ на их действия. Если его наилучший ответ не единствен, то он выбирает равновероятно любой из них.

Одним из вариантов k-уровневой модели является так называемая *L_k-модель*, в которой предполагается, что существуют только агенты нулевого, первого и второго рангов, и что каждый агент ранга $k > 0$ характеризуется вероятностью ε_k «ошибиться» и выбрать неоптимальное действие – равновероятно любое из не принадлежащих соответствующему наилучшему ответу на действия агентов $k - 1$-го уровня [8]. Таким образом, L_k-модель описывается четырьмя параметрами (на сегодняшний день подробно исследованы только двухуровневые модели этого класса) – долями q_1 и q_2 агентов первого и второго рангов рефлексии и их «ошибками» ε_1 и ε_2.

Равновесие квантового наилучшего ответа. Для случая конечных множеств допустимых действий агентов *квантовый наилучший ответ* (QBR – Quantal Best Responce) [36] $QBR_i(p_{-i}; \lambda)$ i-го агента определяют как его смешанную стратегию с вероятностями $p_i(x_i) = \dfrac{\exp(\lambda\, F_i(x_i, p_{-i}(\cdot)))}{\sum_{y_i} \exp(\lambda\, F_i(y_i, p_{-i}(\cdot)))}$, где параметр λ отражает чувствительность агентов к изменениям их целевых функций. Используя концепцию квантового наилучшего ответа, можно определить *квантовое равновесие* [36] (QRE) как профиль смешанных стратегий $p^* = (p_1^*, \dots, p_n^*)$, таких, что $\forall\ i \in N \quad p_i^* = QBR_i(p_{-i}^*(\cdot), \lambda)$. В [36] доказано, что для любой игры в нормальной форме для любого неотрицательного λ квантовое равновесие существует.

Квантовые *k*-уровневые модели (QL_k) являются «пересечением» QRE и L_k моделей [35] (причем в отличие от L_k-моделей каждый агент считает, что агенты более низких уровней иерархии «ошибок» не делают) и описываются, соответственно, пятью параметрами – q_1, q_2, ε_1, ε_2 и λ.

CHM-модели. Агент выбирает наилучший ответ, зависящий от распределения всех агентов более низких рангов рефлексии по этим рангам. Если предполагается, что распределение агентов по рангам рефлексии описывается распределением Пуассона с параметром τ, то Poisson-CHM описывается единственным параметром τ.

В [8] перечисленные четыре класса моделей сравниваются на основании экспериментальных данных авторов соответствующих моделей. Делается вывод о неадекватности предположений о целесообразности использовании распределений пуассоновского типа для описания рангов рефлексии агентов, а также статистически обосновывается сравнительно более высокая эффективность QL_k-модели.

6.3. Краткий обзор экспериментальных зарубежных исследований

Типовая схема организации экспериментального исследования следующая: берется некоторая игра, как правило достаточно простая, для которой выбирается теоретическая модель, описывающая предполагаемое поведение участников; затем проводится серия розыгрышей – участниками обычно являются либо сами исследователи, либо студенты или аспиранты; после чего полученные экспериментальные результаты обрабатываются с использованием статистических методов с целью идентификации параметров модели.

Почти все исследователи признают, что существенным является тип аудитории, среди которой проводится эксперимент: так, например, студенты, изучавшие теорию игр, демонстрируют в среднем более высокий ранг рефлексии [42] (сравнение результатов игры «конкурс красоты» для различных групп агентов можно найти в [24]).

Игра «конкурс красоты». Игра «конкурс красоты» (beauty contest game) является одним из хрестоматийных примеров, иллюстрирующих и модель когнитивных иерархий [21] (современный обзор зарубежных экспериментальных исследований по СНМ можно найти в [43]), и другие модели стратегической рефлексии [8, 33, 44].

Пусть каждый из агентов (которых конечное число n) выбирает число от 0 до 100. Обозначим через $x_i \in \{0, \ldots, 100\}$ действие i-го агента. Победитель – агент, выбравший число, наиболее близкое к заданной доле $\alpha \in (0; 1]$ от среднего действия всех агентов $\frac{1}{n}\sum_{i=1}^{n} x_i$, – получает приз, остальные агенты не получают ничего. Если «победителей» несколько, то приз делится между ними поровну.

Если $\alpha = 1$ и агентов больше двух, то любой вектор одинаковых действий агентов является равновесием Нэша. Ситуация усложняется, если $\alpha < 1$, при этом единственное равновесие Нэша – выбор всеми агентами нулевых действий.

В моделях когнитивных иерархий предлагаются следующие рассуждения для определения «равновесного» исхода [21]. Пусть действия агентов нулевого ранга распределены равномерно и $\alpha < 1$. Тогда их среднее действие равно 50. Или можно предположить, что агенты нулевого ранга выберут действие 50 как фокальную точку [45]. Или в случае повторяющейся игры можно считать, что агенты нулевого ранга выберут свои действия на основе наблюдаемых действий агентов в прошлом периоде. Наилучшим ответом (см. (1)) агентов первого ранга рефлексии будет действие 50 α, агентов второго ранга – 50 α^2, третьего – 50 α^3 и т. д. Результаты различных экспериментов, подтверждающих наличие этих характерных точек, можно найти в [21, 24, 43], анализ и сравнение различных моделей, а также исследование их соответствия экспериментальным данным – в [8].

Игра «11-20». В игре «11-20» участвуют два агента, которые одновременно и независимо выбирают число от 11 до 20. Каждый агент получает сумму денег, равную названному им числу, кроме того, если некоторый агент назвал число, на единицу меньшее выбора оппонента, то первый получает дополнительно 20 единиц [42].

Разумно считать действием агентов нулевого ранга выбор числа 20. Наилучшим ответом на такой выбор, т. е. действием игроков первого ранга, будет 19, второго – 18 и т. д.

Условия игры (выплаты в зависимости от комбинаций действий агентов) можно модифицировать тем или иным образом или строить по аналогии похожие игры (например, «91-100») [42]

Другие «экспериментальные» игры:

– «сороконожка» (centipede game) с игроками нулевого, первого и второго рангов [46-48];

– «вход на рынок» (market entry game) [22, 49].

Некоторые результаты экспериментальных исследований. Наиболее часто в СНМ используется распределение Пуассона, в том числе по следующим причинам: оно достаточно простое (с ним легко производить вычисления), доля соответствующих агентов быстро убывает (когнитивные ограничения) с ростом ранга (причем q_k / q_{k-1} убывает как $1 / k$) и, наконец, это распределение хорошо соответствует экспериментальным данным – см. [21], где доказано, что при $\tau \to +\infty$ вектор действий агентов в Poisson-CHM сходится к равновесию Нэша, получающемуся последовательным исключением слабо доминируемых действий [3] (в общем случае – при непуассоновском распределении – этот результат не имеет места).

В дискретном распределении Пуассона $q_k = e^{\tau} \tau^k / k!$ параметр τ является и средним, и дисперсией, кроме того $\frac{q_{k-2}}{q_{k-1}} = \frac{\tau}{k-1}$. Поэтому его «среднее» экспериментальное значение может интерпретироваться как «средний ранг» рефлексии агентов. Большинство экспериментов свидетельствуют, что это значение лежит от 1 до 2, обычно – около 1,5 (см. в [8, 21] результаты экспериментов в игре «конкурс красоты»).

В [42] для игры «11-20» (для 108 участников) получено следующее распределение агентов по рангам рефлексии:

Табл. 2

«Ранг»	0	1	2	3	4	5	6	7	8	9
Действие	20	19	18	17	16	15	14	13	12	11
Доля агентов (%)	6	12	30	32	6	1	6	3	0	4

В [50] описаны результаты экспериментов, в соответствии с которыми от 40 % до 60 % агентов имеют ненулевой ранг рефлексии (выбирают действия, отличные и от действий агентов нулевого ранга, и от равновесия Нэша). В [23] приводится и обсуждается следующая сводка распределений агентов по рангам рефлексии:

Табл. 3

Ранг	**Доля агентов** [32]	**Доля агентов** [51]	**Доля агентов** [25]
0	0,25	0,42	0,21
1	0,12	0,44	0,21
2	0,12	0,11	0,27
3	0,12	0,03	0,19
4	0,12	0,01	0,09
5	0,12	0	0,03
6	0,12	0	0,01
7 и выше	0	0	0

Стоит подчеркнуть, что делаемые на основании экспериментов выводы о наличии (и доле) агентов пятого, шестого и других более высоких «рангов» рефлексии представляются сомнительными (противоречащими здравому смыслу). В целом, получаемые в результате экспериментов распределения агентов по рангам рефлексии различны и, наверное, сильно зависят как от самой игры, так и условий проведения эксперимента, состава участников и т. д.

6.4. Стратегическая рефлексия в биматричных играх

В исследованиях советских и российских авторов основной акцент при изучении стратегической рефлексии был сделан на теорию биматричных игр.

Рассмотрим игру двух агентов, число действий каждого из которых конечно. Как известно, такие игры называются *биматричными* [2, 3] и целевые функции первого и второго агентов в них обычно задаются матрицами $\mathbf{A} = (a_{ij})$ и $\mathbf{B} = (b_{ij})$, вместе составляющими матрицу игры $(\mathbf{A}, \mathbf{B}) = (a_{ij}, b_{ij})$.

Обозначим $X_1 = \{1, 2, ..., r_1\}$ – множество действий первого агента, $X_2 = \{1, ..., r_2\}$ – множество действий второго агента. Введем следующие предположения:

– матрицы выигрышей не содержат доминируемых стратегий[4];

– для каждого агента существует однозначное отображение, ставящее в соответствие любой стратегии оппонента единственный наилучший ответ данного агента; кроме того, максимальный гарантированный результат каждого агента достигается ровно на одном действии.

Поясним, что будет пониматься под *рангом рефлексии* (точнее, под рангом стратегической рефлексии). В рефлексивных биматричных (и не только биматричных [1]) играх выбор стратегий агентами осуществляется на основании знания рангов рефлексии оппонента. Ранги рефлексии определяются

[4] Если отказаться от этого предположения, то приведенные в настоящем разделе результаты останутся в силе, если под r_1 и r_2 понимать число соответствующих недоминируемых стратегий.

следующим образом. «Игрок имеет нулевой ранг рефлексии, если он знает только матрицу платежей. Игрок обладает первым рангом рефлексии, если он считает, что его противники имеют нулевой ранг рефлексии, т. е., знают только матрицу платежей. Вообще, игрок с k-м рангом рефлексии предполагает, что его противники имеют k – 1-й ранг рефлексии. Он проводит за них необходимые рассуждения по выбору стратегии и выбирает свою стратегию на основе знания матрицы платежей и экстраполяции действий своих противников» [19].

В [52] доказано, что в биматричных играх неограниченное увеличение ранга рефлексии заведомо нецелесообразно, т. е. существует так называемый *максимальный целесообразный ранг рефлексии* [1], превышение которого не приводит к новым действиям агентов. Максимальный целесообразный ранг рефлексии не превышает $\max \{\min \{r_1, r_2 + 1\}, \min \{r_2, r_1 + 1\}\}$ [52].

Из последнего утверждения следует, что множество допустимых действий по выбору ранга конечно. Поэтому можно перейти из исходной игры к **игре рангов**, в которой действием агента является выбор ранга стратегической рефлексии. Верхняя оценка количества возможных попарно различных пар действий составляет $r = r_1 \times r_2$. Тогда исходную биматричную игру можно преобразовать в биматричную игру $r \times r$. Ясно, что некоторые строки и столбцы этой новой матрицы могут совпадать (это означает, что выбор агентами разных рангов приводит к одному и тому же действию в исходной игре). Отождествив совпадающие строки и столбцы, получают матрицу новой игры, которую будем называть *игрой выбора ранга стратегической рефлексии*, или для краткости *игрой рангов*. В [53] доказана справедливость следующих утверждений:

– матрица выигрышей в игре рангов является подматрицей матрицы исходной биматричной игры;

– для произвольной биматричной игры переход к игре рангов не приводит к появлению новых равновесий;

– в игре рангов существует не более двух равновесий.

7. ЗАДАЧА УПРАВЛЕНИЯ

Рассмотрим рефлексивное разбиение в качестве управляемого параметра. Можно сформулировать *задачу управляемости*: пусть задано множество $\Im$ допустимых рефлексивных разбиений; требуется найти множество $X(\Im) = \bigcup_{\aleph \in \Im} x^*(\aleph)$ векторов действий агентов, которые могут быть реализованы в результате *рефлексивного управления* – выбора рефлексивного разбиения $\aleph$. Обратной является задача поиска «минимального» в том или ином смысле множества допустимых рефлексивных разбиений, позволяющего реализовать заданный вектор действий агентов как рефлексивное равновесие.

Рассмотрим теперь собственно задачу управления. Пусть предпочтения управляющего органа – *центра* – описываются его действительнозначной целевой функцией $F_0(Q(x^*))$, заданной на множестве *агрегированных ситуаций игры* (Q: $\Re^n \to \Re^1$), т. е. $F_0(\cdot)$: $\Re^1 \to \Re^1$. Тогда, воспользовавшись выражением (12), *эффективность рефлексивного разбиения* $\aleph$ можно определить как $K(\aleph) = F_0(Q(x^*(\aleph)))$.

Следовательно, формально *задачу рефлексивного управления* (в терминах рефлексивных разбиений) можно сформулировать в виде

(13) $K(\aleph) \to \max_{\aleph \in \Im}$.

Обозначим через K_m максимальное значение критерия эффективности в задаче (13) при фиксированном максимальном ранге рефлексии m. По аналогии с тем, как это делалось для моделей стратегической рефлексии в [1], можно сформулировать задачу о *максимальном целесообразном ранге рефлексии* – таком, больше которого центру (с точки зрения задачи управляемости или/и эффективности рефлексивного управления) использовать не имеет смысла:

$$m^* = \min \{m \mid m \in \operatorname{Arg} \max_{w=0,1,2,\dots} K_w\}.$$

Обсудим согласованность субъективных рефлексивных разбиений агентов. Предположим, что каждый агент наблюдает только агрегированную ситуацию игры. Как отмечалось выше в четвертом разделе при рассмотрении общей модели рефлексивного коллективного поведения, прогнозируемые агентами траектории могут отличаться от реализованной. Это может служить для агентов основанием для того, чтобы усомниться в правильности своих субъективных рефлексивных разбиений. Если агенты наблюдают (помимо собственных действий) только агрегированную ситуацию игры, то по аналогии с условием стабильности информационного управления [17] можно ввести *условие стабильности рефлексивного разбиения* – потребовать, чтобы агрегированная ситуация для реализованной траектории для всех агентов совпадала с прогнозируемыми ими агрегированными ситуациями. Отметим, что стабильность рефлексивного разбиения тесно связана с обучением в играх [40] – наблюдая поведение оппонентов, отличное от прогнозируемого, агенты могут менять свои представления о рангах рефлексии оппонентов или могут переходить на более высокие уровни рефлексии – см. частные эксперименты в [26] (наиболее ярко эти эффекты проявляются в динамических играх – см. dynamic level-k games в [47]).

При фиксированном рефлексивном разбиении $\aleph \in \Im$ реализуется вектор действий, определяемый выражением (12). Соответственно реализуется агрегированная ситуация $Q(x^*(\aleph))$.

С точки зрения j-го агента, обладающего k-м рангом рефлексии, реализуется вектор

$$\tilde{x}_{jk}(\aleph_{jk}) = (x_{l \in N_0},\ x1_{l \in N_1},\ x2_{l \in N_2},\ \ldots,\ x[k-1]_{l \in N_{k-1} \cup N_k \cup \ldots \cup N_m \setminus \{j\}},\ xk_j),\ j \in N_k,\ k = \overline{0, m}.$$

Условие стабильности рефлексивного разбиения $\aleph \in \Im$ примет вид

$$Q(\tilde{x}_{jk}(\aleph_{jk})) = Q(x^*(\aleph)),\ j \in N_k,\ k = \overline{0, m}.$$

Задачу рефлексивного управления (13) можно ставить на множестве стабильных рефлексивных управлений (если таковое не пусто). Содержательно это будет означать, что центр формирует такое оптимальное разбиение агентов по рангам рефлексии, что ни один из агентов на основании наблюдения агрегированных результатов «игры» не имеет оснований усомниться в справедливости своих представлений о рангах рефлексии оппонентов.

В заключение настоящего раздела кратко обсудим, каким образом центр может управлять разбиением агентов по рангам рефлексии. На сегодняшний день в литературе описаны два возможных подхода. Первый подход предполагает, что агенты безусловно верят центру и воспринимают сообщаемую им информацию как истинную, независимо от своих первоначальных представлений. Тогда центр, последовательно сообщая ту или иную информацию различным группам агентов, может формировать различные (но не любые! – см. [54]) структуры информированности (см. также примеры в [1, 17]) и, соответственно, рефлексивные структуры. Второй подход заключается в том, что агенты не просто заменяют свои представления теми, которые сообщает центр, а сообщения центра лишь снижают для агентов неопределенность – сокращают множество возможных с их точки зрения «миров» [55]. В целом, разработка моделей формирования рефлексивных структур под влиянием поступающей к агентам информации представляется чрезвычайно перспективным направлением будущих исследований.

8. ПРИМЕРЫ ВОЗМОЖНЫХ ПРИЛОЖЕНИЙ

На сегодняшний день известен ряд прикладных моделей, иллюстрирующих эффекты стратегической рефлексии в моделях группового поведения (прикладные модели, учитывающие информационную рефлексию агентов, можно найти в [56]). Перечислим кратко основные результаты, которые свидетельствуют, что наличие рефлексирующих агентов может изменять групповое поведение существенным образом.

Модель **«Олигополия Курно»** (в качестве базовой используется классическая модель олигополии [57]) демонстрирует, что при определенном диапазоне значений начальных действий агентов можно реализовать эффективные по Парето или равновесные по Нэшу уровни производства или увеличить суммарный объем производства за счет введения определенного числа агентов первого и второго рангов рефлексии [27, 58].

Второй пример – «**Задача о консенсусе**», содержательная интерпретация которой следующая: действиям агентов соответствуют их положения на прямой (координаты в пространстве или мнения и т. д. – см. обзоры в [59-61]), агрегированной ситуации – среднее значение координат агентов. Целевой функцией агента является его «отклонение» от агрегированной ситуации, а критерием эффективности – «дисперсия» положений агентов (в данном примере целевая функция центра зависит не только от агрегированной ситуации игры, но и от всего вектора действий агентов). Оказывается, что в этом случае введение рефлексирующих агентов расширяет множество векторов действий, которые могут выбрать агенты, и приводит к росту значения критерия эффективности [27].

В модели «**Активная экспертиза**» наличие рефлексирующих агентов даже первого ранга существенно расширяет диапазон возможных результатов экспертизы [13], которыми может манипулировать ее организатор, т. е. наличие рефлексирующих агентов может приводить к последствиям, негативным, условно говоря, с точки зрения группы в целом (см. также модели формирования команд в [62]). Второй ранг рефлексии позволяет реализовать равновесие Нэша [27].

В модели «**Транспортные потоки и эвакуация**» наличие рефлексирующих агентов первого ранга позволяет достичь минимального (оптимального с «централизованной» точки зрения) времени эвакуации из здания группы агентов, принимающих решения децентрализовано [27]. В [30, 31, 58] можно найти описание имитационных моделей транспортных потоков и эвакуации, учитывающих рефлексию агентов.

В модели «**Диффузная бомба**» число агентов, достигших цели, монотонно растет с увеличением доли рефлексирующих агентов первого ранга [63].

Для модели «**Фондовый рынок**» (следует подчеркнуть, что именно фондовый рынок является объектом моделирования, для которого, наверное, наиболее часто используют «рефлексивные» рассуждения – см., например, [64-66]) показано, что изменить рыночную цену (по сравнению с нерефлексивным принятием решений) может только определенная «критическая масса» рефлексирующих агентов [27].

9. ЗАКЛЮЧЕНИЕ

Метод рефлексивных разбиений множества рациональных агентов на подмножества агентов, обладающих различными рангами стратегической рефлексии, позволяет:

- с точки зрения теории принятия решений – расширить класс моделей коллективного поведения интеллектуальных агентов, осуществляющих совместную деятельность в условиях неполной информированности и отсутствия общего знания;
- с дескриптивной точки зрения – расширить множество ситуаций, которые в рамках модели могут быть «объяснены» как рефлексивные равновесия – устойчивые исходы взаимодействия агентов; соответственно, в рамках задач рефлексивного управления – расширить область управляемости;
- с нормативной точки зрения – ставить и решать задачи рефлексивного группового управления.

Современный уровень исследований моделей стратегической рефлексии таков, что пока не существует универсального «аппарата» поиска аналитических решений для широких классов задач рефлексивного управления – все успехи ограничены набором частных и достаточно простых моделей (см. ссылки в восьмом разделе). Видятся два направления возможных будущих прорывов. Первое – экспериментальные исследования принятия агентами решений и поиск общих закономерностей на основе анализа результатов экспериментов. Второе направление (теоретическое) – разработка языка описания рефлексивных моделей, позволяющего достаточно просто и единообразно ставить и решать различные задачи рефлексивного управления.

Задачи рефлексивного управления, поиска максимального целесообразного ранга рефлексии и др. можно и нужно ставить и решать в рамках и других (отличных от рассмотренных выше) модификаций предложенной модели рефлексивного коллективного поведения, что представляется перспективным направлением будущих исследований. В первую очередь, это – задачи *активного прогноза* [67], в рамках которого агенты по информации центра о будущем состоянии системы «восстанавливают» текущее состояние (осуществляя информационную рефлексию) и принципы принятия оппонентами решений (осуществляя стратегическую рефлексию), а потом на основании этой новой информации принимают решения. Здесь введение рефлексивных разбиений выглядит очень многообещающим.

Почти все авторы, исследующие модели стратегической рефлексии в коллективах агентов, сходятся во мнении, что перспективной областью прикладного использования этих моделей являются *мультиагентные системы*. С одной стороны, так как эти системы искусственные, то наделение различных агентов теми или иными рангами рефлексии не представляет проблемы (в отличие от рефлексивного управления коллективами людей – см. краткое описание методов информационных воздействий выше) и может производиться на этапе проектирования мультиагентной системы. Задача «управления» при этом заключается в выборе оптимальной доли рефлексирующих агентов с учетом того, что повышение их «интеллектуальности» требует затрат на наделение этих агентов соответствующими вычислительными, коммуникационными и другими ресурсами. С другой стороны, анализ литературы по мультиагентным системам свидетельствует, что использование искусственных автономных агентов, наделенных способностью к рефлексии, не получило еще массового распространения (исключение составляют модели эвакуации, а также некоторые другие частные модели группового поведения – см. восьмой раздел и [27]).

СПИСОК ЛИТЕРАТУРЫ

1. *Новиков Д.А., Чхартишвили А.Г.* Рефлексивные игры. М.: СИНТЕГ, 2003.
2. *Губко М.В., Новиков Д.А.* Теория игр в управлении организационными системами. М.: Синтег, 2002.
3. *Myerson R.* Game Theory: Analysis of Conflict. London: Harvard Univ. Press, 1991.
4. *Васин А.А.* Модели динамики коллективного поведения. М.: Изд-во МГУ, 1989.
5. *Малишевский А.В.* Качественные модели в теории сложных систем. М.: Наука, 1998.
6. *Опойцев В.И.* Равновесие и устойчивость в моделях коллективного поведения. М.: Наука, 1977.
7. The Handbook of Experimental Economics / Ed. by *J. Kagel and A. Roth*. Princeton: Princeton University Press, 1995.
8. *Wright J., Leyton-Brown K.* Beyond Equilibrium: Predicting Human Behavior in Normal Form Games / Proc. of Conf. Associat. Advancement of Artificial Intelligence (AAAI-10), 2010. P. 461 – 473.
9. *Канеман* Д., *Словик П*., *Тверски* А. Принятие решений в неопределенности: Харьков: Гуманитарный центр, 2005.
10. *Поддъяков А.Н.* Исследовательское поведение: стратегии познания, помощь, противодействие, конфликт. М.: Факультет психологии МГУ им. М.В. Ломоносова, 2002.
11. *Nisan N.*, *Roughgarden T., Tardos E. et al.* Algorithmic Game Theory. Cambridge: Cambridge University Press, 2007.
12. *Bernheim D*. Rationalizable Strategic Behavior // Econometrica. 1984. V. 52. № 4. P. 1007 – 1028.
13. *Новиков Д.А.* Теория управления организационными системами. 2-е изд. М.: Физматлит, 2007.
14. *Беленький В.З., Волконский В.А., Иванков С.А. и др.* Итеративные методы в теории игр и программировании. М.: Физматлит, 1974.
15. *Weibull J.* Evolutionary Game Theory. Cambridge: MIT Press, 1995.
16. *Васин А.А.* Некооперативные игры в природе и обществе. М.: МАКС Пресс, 2005.
17. *Чхартишвили А.Г.* Теоретико-игровые модели информационного управления. М.: ПМСОФТ, 2005.
18. *Лефевр В.А.* Конфликтующие структуры. М.: Сов. радио, 1973.
19. *Поспелов Д.А.* Игры рефлексивные / Энциклопедия кибернетики. Т. 1. Киев: Гл. редакция УСЭ, 1974. С. 343.
20. *Stahl D.* Evolution of Smart_n Players // Games and Economic Behavior. 1993. № 5. P. 604 – 617.
21. *Camerer C., Ho T., Chong J.* A Cognitive Hierarchy Model of Games // The Quarterly J. of Economics. 2004. № 8. P. 861 – 898.
22. *Camerer C., Ho T., Chong J.* Behavioral Game Theory: Thinking, Learning, and Teaching. Working Paper. 2001. 70 p.
23. *Camerer C., Ho T., Chong J.* Cognitive Hierarchy: A Limited Thinking Theory In Games / Experimental Business Research, 2005. V. 3. P.203 – 228.
24. *Camerer C., Ho T., Chong J.* Models of Thinking, Learning and Teaching in Games // AEA Paper Proc. 2003. V. 92. № 2. P. 192 – 195.
25. *Costa-Gomes M., Broseta B*. Cognition and Behavior in Normal-Form Games: An Experimental Study // Econometrica. 2001. V. 69. № 5. P. 1193 – 1235.
26. *McCain R.* Learning Level-k Play in Noncooperative Games. Philadelphia: Drexel University, 2010. Working Paper. 23 p.
27. *Корепанов В.О., Новиков Д.А.* Метод рефлексивных разбиений в задачах группового поведения и управления // Проблемы управления. 2011. № 1. С. 21 – 32.
28. *Novikov D.A.* Reflexive Models of Collective behavior / Proceedings of 18 IFAC World Congress. Milan, 2011 (в печати).
29. *Новиков Д.А.* «Когнитивные игры»: линейная импульсная модель // Проблемы управления. 2008. № 3. С. 14 – 22.

30. *Корепанов В.О.* О влиянии рефлексивных агентов с общим знанием на транспортный поток / Тр. 52-й науч. конф. МФТИ «Современные проблемы фундаментальных и прикладных наук». Ч. I. Радиотехника и кибернетика. М.: МФТИ, 2009. Т. 2. С. 69 – 71.

31. *Корепанов В.О.* Влияние рефлексивных агентов на транспортный поток / Тр. междунар. науч.-практ. конф. «Теория активных систем». М.: ИПУ РАН, 2009. Т. II. С. 122 – 125.

32. *Stahl D., Wilson P.* On Players Models of Other Players: Theory and Experimental Evidence // Games and Economic Behavior. 1995. V. 10. P. 213 – 254.

33. *Nagel R.* Unraveling in Guessing Games: An Experimental Study // AER. 1995. V. 85. P. 1313 – 1326.

34. *Costa-Gomes M., Crawford V.* Cognition and Behavior in Two-Person Guessing Games: An Experimental Study // AER. 2006. V. 96. P. 1737 – 1768.

35. *Stahl D., Wilson P.* Experimental Evidence on Players' Models of Other Players // J. of Economic Behavior and Organization. 1994. V. 25. P. 309 – 327.

36. *McKelvey R., Palfrey T.* Quantal Response Equilibria for Normal Form Games // Games and Economic Behavior. 1995. V. 10(1). P. 6 – 38.

37. *Choi S.* A cognitive hierarchy model of learning in networks. London: Centre for Economic Learning and Social Evolution, 2006. Working Papers 238. http://eprints.ucl.ac.uk/14461/

38. *Choi S., Gale D., Kariv S.* Behavioral Aspects of Learning in Social Networks: An Experimental Study / Advances in Behavioral and Experimental Economics. Ed. by John Morgan, 2005. JAI Press.

39. *Choi S., Gale D., Kariv S.* A Quantal Response Equilibrium Analysis of Social Learning in Networks. Working paper. UCL, 2006.

40. *Fudenberg D., Levine D.* The Theory of Learning in Games. Massachusetts: MIT Press, 1998.

41. *Gale D., Kariv S.* Bayesian Learning in Social Networks // Games and Economic Behavior. 2003. V. 45(2). P. 329 – 346.

42. *Arad A., Rubinstein A.* The 11-20 Money Request Game: Evaluating the Upper Bound of k-Level Reasoning / Working Paper – http://ideas.repec.org/p/cla/levarc/661465000000000073.html.

43. *Burchardi K., Penczynski S.* Out Of Your Mind: Estimating The Level-k Model. –London: London School of Economics, 2010. Working Paper. 44 p.

44. *Moulin H.* Game Theory for Social Sciences . NY: New York Press, 1986.

45. *Schelling T.* The Strategy of Conflict. Cambridge: Harvard University Press, 1960.

46. *Bottero M.* Cognitive hierarchies and the centipede game. Stockholm: SSE/EFI Working Paper Series in Economics and Finance № 723. 2010. 38 p.

47. *Ho T., Su X.* A Dynamic Level-k Model in Games. Berkeley: University of California, 2010. Working Paper. 45 p.

48. *McKelvey R., Palfrey T.* An Experimental Study of the Centipede Game // Econometrica. 1992. V. 60. № 3. P. 803 – 836.

49. *Van Huyck J., Cook J., Battalio R.* Adaptive Behavior and Coordination Failure // J. of Economic Behavior and Organization. 1997. V. 32. P. 483 – 503.

50. *Sutter M., Czermak S., Feri F.* Strategic sophistication of individuals and teams in experimental normal-form games. University of Innsbruck. Working Paper. 55p. – http://econpapers.repec.org/paper/innwpaper/2010-02.htm

51. *Cooper D., Van Huyck J.* Evidence on the Equivalence of the Strategic and Extensive Form Representation of Games // J. of Economic Theory. 2003. V. 110. № 2. P. 290 – 308.

52. *Новиков Д.А.* Стратегическая рефлексия в биматричных играх / «Региональная экономика в информационном измерении: модели, оценки, прогнозы». Сб. научн. тр. под ред. Е.Ю. Иванова, Р.М. Нижегородцева. М.: Бизнес-Юнитек, 2003. С. 296 – 307.

53. *Губанов Д.А., Чхартишвили А.Г.* О стратегической рефлексии в биматричных играх // Управление большими системами. 2008. № 21. С. 49 – 57.

54. *Романько А.Д., Чхартишвили А.Г.* Моделирование информационных воздействий в рефлексивных играх: простые сообщения // Сб. тр. ВГАСУ. 2006. С. 157 – 167.

55. *Чхартишвили А.Г.* Рефлексивные игры: трансформация структур информированности // Проблемы управления. 2008. № 5. С. 43 – 48.

56. *Новиков Д.А., Чхартишвили А.Г.* Прикладные модели информационного управления. М.: ИПУ РАН, 2004.

57. *Mas-Collel A., Whinston M., Green J.* Microeconomic theory. N.-Y.: Oxford Univ. Press, 1995.

58. *Корепанов В.О.* Управление рефлексивным поведением агентов в модели олигополии Курно // Управление большими системами. 2010. № 31. С. 225 – 249.

59. *Чеботарев П.Ю., Агаев Р.П.* Согласование характеристик в многоагентных системах и спектры лапласовских матриц орграфов // АиТ. 2009. № 3. С. 136 – 151.

60. *Shoham Y, Leyton-Brown K.* Multiagent Systems: Algorithmic, Game-Theoretical and Logical Foundations. Cambridge: Cambridge University Press, 2009.

61. *Wei R.* Consensus Seeking, Formation Keeping and Trajectory Tracking in Multiple Vehicle Cooperative Control. PhD Dissertation. Brigham Young University, 2004.

62. *Новиков Д.А.* Математические модели формирования и функционирования команд. М.: Физматлит, 2007.

63. *Корепанов В.О., Новиков Д.А.* Задача о диффузной бомбе // Проблемы управления. 2011. № 5. С. 66 – 73.

64. *Ерешко Ф.И.* Моделирование рефлексивных стратегий в управляемых системах. М.: ВЦ РАН, 2001.

65. *Зинченко В.И., Новиков Д.А., Старостенко В.В.* Об одной теоретико-игровой модели фондового рынка // Тр. IV междунар. конф. «Современные сложные системы управления». Тверь, ТГТУ, 2004. С. 294 – 297.

66. *Сорос Д.* Алхимия финансов. М.: ИНФРА-М, 1999.

67. *Новиков Д.А., Чхартишвили А.Г.* Активный прогноз. М.: ИПУ РАН, 2002.