

# МЕТОД РЕФЛЕКСИВНЫХ РАЗБИЕНИЙ В МОДЕЛЯХ ГРУППОВОГО ПОВЕДЕНИЯ И УПРАВЛЕНИЯ

В.О. Корепанов, Д.А. Новиков

Предложен метод рефлексивных разбиений множества рациональных агентов, осуществляющих совместную деятельность, на подмножества, соответствующие различным рангам их рефлексии. Приведены примеры информационного управления в различных задачах принятия коллективных решений.

**Ключевые слова:** теория игр, коллективное поведение, принятие решений, рефлексия, информационное управление.

## 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. (так называемая концепция общего знания). Либо предполагается, что каждый агент следует некоторой процедуре принятия решений и «не задумывается» над тем, что знают и как ведут себя остальные агенты. Первый подход традиционен для теории игр (см., например, работы [1–3]), второй — для моделей коллективного поведения (см., например, работы [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 он обладает первым рангом рефлексии). Другой агент (обладающий вторым рангом рефлексии) может предположить существование агентов первого ранга и прогнозировать их поведение. И так далее. Возникает вопрос — 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 рефлексиирующих агентов можно управлять, то каковы эти доли, оптимальные в смысле того или иного критерия эффективности, определенного на множестве действий агентов?

Для ответа на эти и другие вопросы в статье рассматривается метод рефлексивных разбиений множества рациональных агентов, осуществляющих совместную деятельность, на подмножества, соответствующие различным рангам их рефлексии.

## 1. МОДЕЛИ РАЦИОНАЛЬНОГО ПОВЕДЕНИЯ

**Описание модели.** Рассмотрим множество  $N = \{1, 2, \dots, n\}$ , состоящее из  $n$  агентов. Агент  $i$  выбирает свое действие  $x_i \in \mathcal{X}^1$  (для простоты здесь и далее, если не оговорено особо, считается, что ограничения на действия агентов отсутствуют). Вектор  $x = (x_1, x_2, \dots, x_n) \in \mathcal{X}^n$  действий агентов, называемый *ситуацией игры* [1], определяет их выигрыши, задаваемые *целевыми функциями*  $F_i(x)$ , где  $F_i(\cdot): \mathcal{X}^n \rightarrow \mathcal{R}^1$ .

Рациональность поведения агента заключается в стремлении к максимизации своей целевой функции выбором собственного действия [1]:

$$x_i \in BR_i(x_{-i}) = \operatorname{Arg} \max_{y \in \mathcal{X}^1} F_i(y, x_{-i}), \quad i \in N, \quad (1)$$

где  $x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \mathcal{X}^{n-1}$  — *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. Основным предметом *теории игр* — см. работы [1–4] — 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 [1]. Введение определенных предположений об *информированности* агентов (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* [1]. Если считать, что все описание игры (множество агентов, их целевые функции и множества допустимых действий) является *общим знанием* [2] среди агентов, принимающих решения однократно, одновременно и независимо, то можно использовать концепцию *равновесия Нэша*, в рамках которой равновесием называется такой вектор действий  $x^N$ , что действие каждого агента является наилучшим ответом на нэшевскую обстановку:

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

**Коллективное поведение.** В отличие от теории игр, *теория коллективного (группового) поведения* — см. работы [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 агенты не обязаны предсказывать поведение всех оппонентов, как это имеет место в теории игр. Более того, зачастую агенты, принимая решения, могут «не знать о существовании» других агентов или иметь о них агрегированную информацию. Поэтому далее при рассмотрении рефлексивных моделей группового поведения и управления будем ориентироваться на случай, когда каждый агент в общем случае имеет агрегированную информацию о результатах деятельности своих оппонентов.

Наиболее распространенной моделью динамики коллективного поведения служит *модель индикаторного поведения* [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 в предыдущий момент времени  $t-1$ ,  $t = 1, 2, \dots$  (начальный вектор действий  $x^0 = (x_1^0, x_2^0, \dots, x_n^0)$  считается заданным).

Каждый агент может рассчитать свое *текущее положение цели* — такое его действие, которое максимизировало бы его целевую функцию при усло-

вии, что в текущем периоде все агенты выбрали бы те же действия, что и в предыдущем:

$$w_i(x_{-i}^{t-1}) = \arg \max_{y \in \mathcal{R}^1} F_i(y, x_{-i}^{t-1}), \quad i \in N. \quad (3)$$

В частном случае агент может иметь агрегированную информацию  $Q(x^{t-1})$  о действиях других или всех агентов. В этом случае выражение (3) примет вид  $w_i(x_{-i}^{t-1}) = \arg \max_{y \in \mathcal{R}^1} F_i(y, Q(y, x_{-i}^{t-1}))$ ,  $i \in N$ .

В рамках гипотезы индикаторного поведения каждый агент в каждый момент времени будет делать «шаг» от своей предыдущей стратегии к текущему положению цели:

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

где  $\gamma_i^t \in [0; 1]$  — «размеры шагов». 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. выражение (1)), однако в последнем случае соответствующая динамика может быть неустойчивой. Условия сходимости процедуры (4), области притяжения равновесий, условия на размеры шагов  $\{\gamma_i^t\}$ , обеспечивающие сходимость, и ряд других можно найти в книгах [5, 6].

Подходы теории коллективного поведения и теории игр согласованы в том смысле, что и та, и другая исследуют поведение рациональных агентов (сравните выражения (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).

**Рефлексия.** Различают стратегическую и информационную рефлексии [2]. *Информационная рефлексия* — процесс и результат размышлений агента об информированности других агентов, *стратегическая рефлексия* — процесс и результат размышлений агента о принципах принятия решений другими агентами.

В рамках теоретико-игровой модели отказ от предположения о наличии среди агентов общего знания приводит к моделям *рефлексивных игр* [2]. При этом равновесие Нэша «превращается» в более общее *информационное равновесие*, в рамках которого каждый агент осуществляет информационную рефлексии — при принятии решений использует не только свою информацию о существенных параметрах, но и свои представления о представлениях других агентов об этих параметрах, представлениях о представлениях и т. д. [2]. Поэтому интересно рассмотреть «по аналогии» эффектов стратегической рефлексии.



## 2. РЕФЛЕКСИВНЫЕ РАЗБИЕНИЯ

Отметим, что в рамках гипотезы индикаторного поведения 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 рефлексии), 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:

$$w_i(x_{-i}^t) = \arg \max_{y \in \mathcal{N}^1} F_i(y, x_{-i}^t),$$

где  $x_{-i}^t$  определяется выражением (4). Будем полагать, что рефлексирующий агент первого ранга считает всех остальных нерефлексирующими (что соответствует традиции рефлексивных игр, в рамках которой считается, что агент, имеющий некоторый ранг стратегической рефлексии, считает всех остальных имеющими ранг на единицу меньше его собственного [2]).

Аналогично можно рассматривать агентов и более высоких рангов рефлексии. Для этого определим  $\mathcal{N} = \{N_0, N_1, \dots, N_m\}$  — разбиение множества агентов  $N$ , где  $N_i$  — множество агентов  $i$ -го ранга рефлексии,  $i = \overline{0, m}$ ,  $m$  — максимальный ранг рефлексии. Назовем  $\mathcal{N}$  *рефлексивным разбиением*.

Будем считать, что агент некоторого ранга рефлексии  $k$  достоверно знает множества (или долю — см. далее) агентов всех более низких рангов  $k'$  (где  $k' < k - 1$ ) и считает всех агентов своего и больших рангов ( $k'' \geq k$ ) имеющими ранг на единицу меньше его ранга (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.

## 3. РЕФЛЕКСИВНАЯ МОДЕЛЬ КОЛЛЕКТИВНОГО ПОВЕДЕНИЯ

Пусть задан вектор  $x^0$  начальных действий агентов. Рассмотрим следующую динамическую рефлексивную модель 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 (принадлежа-

щие множеству  $N_0$ ) выбирают свои действия, считая, что действия остальных агентов будут такими же, что и в предыдущем периоде. Другими словами, из выражения (4) следует, что

$$x_i^t = x_i^{t-1} + \gamma_i^t [w_i(x_{-i}^{t-1}) - x_i^{t-1}], \\ i \in N_0, \quad t = 1, 2, \dots \quad (5)$$

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 выражением (5).

**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 выражением (5) «предсказывает» их выбор. Поэтому его собственный выбор  $x_1^t$  будет ориентирован на наилучший ответ на ту обстановку, которая с его точки зрения должна сложиться:

$$x_1^t = x_1^{t-1} + \gamma_j^t [w_j(x_{-j}^t) - x_1^{t-1}], \quad j \in N_1. \quad (6)$$

Для агента  $j \in N_1$  *прогнозируемой* является *траектория*  $(x^0, \dots, (x_1^t, x_{-j}^t), \dots)$ , а на самом деле реализуется траектория  $(x^0, \dots, (x_1^t, x_{-j}^t), \dots)$ ; 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 далее — см. § 5, «условие стабильности рефлексивного разбиения».

**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:

$$x_2^t = x_2^{t-1} + \gamma_j^t [w_j(x_{i \in N_0}^t, x_{l \in N_1 \cup N_2 \setminus \{j\}}^t) - x_2^{t-1}], \quad j \in N_2. \quad (7)$$

Агенты нулевого и первого рангов рефлексии будут вести себя в соответствии с выражениями (5) и (6) соответственно.

Для агента  $j \in N_2$  *прогнозируемой* служит *траектория*  $(x^0, \dots, (x_2^t, x_{l \in N_1 \cup N_2 \setminus \{j\}}^t, x_{i \in N_0}^t), \dots)$ , а на

самом деле реализуется траектория  $(x^0, \dots, (x2_{j \in N_2}^t, x1_{l \in N_1}^t, x_{l \in N_0}^t), \dots)$ .

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

$$\aleph_{jk} = (\underbrace{N_0, N_1, \dots, N_{k-2}, N_{k-1} \cup N_k \cup \dots \cup N_m \setminus \{j\}}_k, \underbrace{\{j\}, \emptyset, \dots, \emptyset}_{m-k-1}), \quad j \in N_k. \quad (8)$$

Отметим, что в рамках выражения (8) агент, обладающий  $k$ -м рангом рефлексии, обладает правильными представлениями о рангах рефлексии всех агентов, имеющих строго меньшие ранги рефлексии.

В частном случае — когда агенты однородны (имеют одинаковые целевые функции, множества допустимых действий и т. д.) — иногда можно ограничиться предположением о том, что рефлексивный агент знает лишь общее число (или долю) агентов, имеющих те или иные меньшие ранги рефлексии.

*Структура информированности* определяется совокупностью субъективных рефлексивных разбиений всех агентов. Если предположить, что представления агентов о рангах рефлексии друг друга описываются выражением (8), то структура информированности однозначно задается рефлексивным разбиением  $\aleph$ .

Таким образом, в рамках предложенной *рефлексивной модели коллективного поведения* (РМКП) последовательность векторов действий агентов однозначно определяется заданием кортежа  $(N, \{F_i(\cdot)\}_{i \in N}, \aleph)$ , состоящего из множества агентов  $N$ , их целевых функций  $\{F_i(\cdot)\}_{i \in N}$  и рефлексивного разбиения  $\aleph$ .

Введенная здесь общая РМКП 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 работе [7]), позволяющих описывать и прогнозировать групповое поведение (людей, мобильных роботов, программных агентов) в разнообразных ситуациях — см., например, рефлексивные имитационные модели эвакуации [8], рефлексивные модели транспортных потоков [9] и

примеры в § 6 настоящей работы. Аналитические результаты можно надеяться получить, вводя те или иные упрощающие предположения. Например, далее рассматривается «игровая» модель с однородными агентами и агрегированным влиянием обстановки игры на выигрыш каждого из них.

#### 4. «ИГРА» ОДНОРОДНЫХ АГЕНТОВ

Пусть:

— все агенты из множества  $N$  имеют одинаковые целевые функции ( $F_i(\cdot) = f(\cdot)$ ,  $i \in N$ );

— целевая функция  $i$ -го агента зависит от его действия  $x_i$  (причем она непрерывна и вогнута по этой переменной) и от *агрегированной ситуации*  $Q(x)$ , где  $Q(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}^1$  — симметричная функция своих аргументов;

— агенты принимают решения однократно (поэтому верхний индекс, соответствующий первому периоду времени, далее будем опускать);

— начальный вектор действий  $x^0$  и рефлексивное разбиение  $\aleph$  фиксированы.

Агенты нулевого ранга рефлексии в соответствии с выражением (5) выберут действия

$$x_i = \arg \max_{y \in \mathbb{R}^1} f(y, Q(y, x_{-i}^0)), \quad i \in N_0. \quad (9)$$

Агенты первого ранга рефлексии в соответствии с выражением (6) выберут действия

$$x1_j = \arg \max_{y \in \mathbb{R}^1} f(y, Q(y, x_{-j})), \quad j \in N_1.$$

Агенты второго ранга рефлексии в соответствии с выражением (7) выберут действия

$$x2_j = \arg \max_{y \in \mathbb{R}^1} f(y, Q(y, x_{l \in N_0}, x1_{l \in N_1 \cup N_2 \setminus \{j\}})), \quad j \in N_2.$$

И так далее. Агенты  $m$ -го ранга рефлексии в соответствии со своим субъективным рефлексивным разбиением (8) выберут действия

$$xm_j = \arg \max_{y \in \mathbb{R}^1} f(y, Q(y, x_{l \in N_0}, x1_{l \in N_1}, \dots, x[m-1]_{l \in N_{m-1} \cup N_m \setminus \{j\}})), \quad j \in N_m.$$

Вектор действий агентов зависит от функций  $f(\cdot)$  и  $Q(\cdot)$ , а также вектора начальных действий  $x^0$  и рефлексивного разбиения  $\aleph$ . Если первые три параметра фиксированы, то вектор действий агентов зависит только от рефлексивного разбиения:

$$x^*(\aleph) = (x_{l \in N_0}, x1_{l \in N_1}, x2_{l \in N_2}, \dots, xm_{l \in N_m}). \quad (10)$$

Следовательно, изменяя рефлексивное разбиение, можно менять действия агентов, т. е. осуществлять *рефлексивное управление* (см. также теоретико-игровые модели рефлексивного управления [2, 10]).

## 5. ЗАДАЧА УПРАВЛЕНИЯ

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

Рассмотрим теперь собственно задачу управления. Пусть предпочтения управляющего органа — *центра* — описываются его действительной целевой функцией  $F_0(Q(x^*))$ , заданной на множестве агрегированных ситуаций, т. е.  $F_0(\cdot): \mathfrak{N}^1 \rightarrow \mathfrak{N}^1$ . Тогда, воспользовавшись выражением (10), *эффективность рефлексивного разбиения*  $\mathfrak{N}$  можно определить как

$$K(\mathfrak{N}) = F_0(Q(x^*(\mathfrak{N}))).$$

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

$$K(\mathfrak{N}) \rightarrow \max_{\mathfrak{N} \in \mathfrak{Z}}. \quad (11)$$

Обозначим  $n_i = |N_i|$  — число агентов  $i$ -го ранга рефлексии,  $i = \overline{0, m}$ . В силу однородности агентов и симметричности функции  $Q(\cdot)$ , при введении ряда дополнительных предположений (например, об одинаковости начальных действий всех агентов и т. п.) эффективность рефлексивного разбиения определяется  $(m+1)$ -м числом:  $n_0, n_1, \dots, n_m$ . Следовательно, задача (16) примет вид:

$$K(\mathfrak{N}) \rightarrow \max_{\left\{ n_i \geq 0 \mid \sum_{i=0}^m n_i = N \right\}}. \quad (12)$$

При большом числе агентов задачу (12) иногда удобно формулировать в терминах долей агентов  $q_i = n_i/N$ , имеющих тот или иной ранг рефлексии  $i = \overline{0, m}$ .

Обозначим  $K_m$  — максимальное значение критерия эффективности в задаче (11) при фиксированном максимальном ранге рефлексии  $m$ . По аналогии с тем, как это делалось для моделей стратегической рефлексии в работе [2], можно сформулировать задачу о *максимальном целесообразном ранге рефлексии* — таком, больше которого центру (в смысле задачи управляемости или (и) эффек-

тивности рефлексивного управления) использовать не имеет смысла:

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

Обсудим согласованность субъективных рефлексивных разбиений агентов. Предположим, что каждый агент наблюдает только агрегированную ситуацию. Как ранее отмечалось при рассмотрении общей РМКП, 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 (см. работу [10]) можно ввести *условие стабильности рефлексивного разбиения* — потребовать, чтобы агрегированная ситуация для реализованной траектории совпадала с прогнозируемыми агрегированными ситуациями для всех агентов.

При фиксированном рефлексивном разбиении  $\mathfrak{N} \in \mathfrak{Z}$  реализуется вектор действий, определяемый выражением (10). Соответственно, реализуется агрегированная ситуация  $Q(x^*(\mathfrak{N}))$ .

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

$$\begin{aligned} \hat{x}_{jk}(\mathfrak{N}_{jk}) &= (x_{l \in N_0}, x_{l \in N_1}, x_{l \in N_2}, \dots, \\ &x_{[k-1] \in N_{k-1} \cup N_k \cup \dots \cup N_m \setminus \{j\}}, x_{k_j}), \\ &j \in N_k, \quad k = \overline{0, m}. \end{aligned}$$

Условие стабильности рефлексивного разбиения  $\mathfrak{N} \in \mathfrak{Z}$  примет вид:

$$Q(\hat{x}_{jk}(\mathfrak{N}_{jk})) = Q(x^*(\mathfrak{N})), \quad j \in N_k, \quad k = \overline{0, m}. \quad (13)$$

Задачу рефлексивного управления (11) 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 результатов «игры» не имеет оснований усомниться в справедливости своих представлений о рангах рефлексии оппонентов.

В заключение настоящего параграфа кратко обсудим, каким образом центр может управлять разбиением агентов по рангам рефлексии. На сегодня в литературе описаны два возможных подхода. Один из них 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. статью [11]) структуры информированности (см. также примеры в работах [2, 10]). Другой 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» [12]. В целом, разработка моделей формирования структур информированности агентов под влиянием поступающей к ним информации представляется чрезвычайно перспективным направлением будущих исследований.

## 6. ПРИМЕРЫ

Рассмотрим ряд примеров, иллюстрирующих эффекты рефлексии в моделях группового поведения.

### 6.1. Олигополия Курно

В модели олигополии Курно [13] агенты принимают решения об объеме выпускаемой ими продукции в условиях, когда ее рыночная цена является известной убывающей функцией суммарного предложения (объема выпуска, объема производства):  $P(x) = a - bQ(x)$ , где  $Q(x) = \sum_{i \in N} x_i$ ,  $a$  и  $b$  — известные неотрицательные константы.

Целевая функция  $i$ -го агента представляет собой разность между выручкой от продаж (равной произведению цены на объем производства) и квадратичными затратами на производство:

$$f_i(x_i, Q(x)) = (a - bQ(x))x_i - (x_i)^2/2, \quad i \in N.$$

Если бы целевые функции агентов были среди них общим знанием, то равновесию Нэша их игры соответствовали бы одинаковые действия:

$$x_i^N = \frac{a}{1 + b + nb}, \quad i \in N,$$

которые приводили бы к равновесному объему выпуска  $Q(x^N) = \frac{na}{1 + b + nb}$  и равновесной цене  $P(x^N) = \frac{a(1 + b)}{1 + b + nb}$ . Точке Парето, максимизирующей сумму целевых функций агентов, соответствуют действия:

$$x_i^P = \frac{a}{1 + 2nb}, \quad i \in N,$$

которые приводят к эффективному объему выпуска  $Q(x^P) = \frac{na}{1 + 2nb}$  и эффективной цене  $P(x^P) = \frac{a(1 + nb)}{1 + 2nb}$ .

При этом  $f(x^P) = \frac{a^2}{2(1 + 2nb)} \geq f(x^N) = \frac{a^2(1 + 2b)}{2(1 + b + nb)^2}$ , т. е.

выигрыш каждого агента в точке Парето не меньше, чем в точке Нэша.

Рассмотрим числовой пример. Пусть  $n = 10$ ,  $a = 2,1$ ,  $b = 0,1$ ,  $\gamma_i^t = 0,5$ . Тогда  $x_i^N = 1$ ,  $Q(x^N) = 10$ ,  $P(x^N) = 1,1$ ,  $x_i^P = 0,7$ ,  $Q(x^P) = 7$ ,  $P(x^P) = 1,4$ ,  $f(x^P) = 0,735 > f(x^N) = 0,6$ .

Проанализируем динамику коллективного поведения. Пусть фиксирован вектор  $x^0$  начальных объемов

производства. В соответствии с выражением (4) изменение во времени действий, выбираемых агентами, будет описываться следующим выражением:

$$x_i^t = x_i^{t-1} + \gamma_i^t \left[ \frac{a - b \sum_{j \neq i} x_j^{t-1}}{1 + 2b} - x_i^{t-1} \right], \quad i \in N, \quad t = 1, 2, \dots \quad (14)$$

В соответствии с выражением (14) действия агентов будут сходиться к равновесию Нэша.

Перейдем теперь к рефлексивному случаю. При заданном векторе начальных действий  $x^0$  агенты нулевого ранга рефлексии выберут действия  $x_i = A + Bx_i^0$ ,  $i \in N_0$ ,

где  $A = \frac{a - bQ(x^0)}{1 + 2b}$ ,  $B = \frac{b}{1 + 2b}$ . Агенты первого ранга рефлексии выберут действия

$$x_{1j} = A_1 + B^2 x_j^0, \quad j \in N_1,$$

где  $A_1 = \frac{a(1 + 3b) - bna + b^2 Q(x^0)(n - 2)}{(1 + 2b)^2}$ .

Пусть в рассматриваемом числовом примере все начальные действия агентов одинаковы:  $x_i^0 = 0,5$ ,  $i \in N$ . Тогда  $x_i = 31/24 = 1,291(6)$ ,  $x_{1j} = 103,5/144 = 0,71875$ , что гораздо ближе к Парето-эффективным действиям. Варьируя число агентов первого уровня, можно менять сумму действий агентов от  $\sim 7,2$  до  $\sim 12,9$ . Этому диапазону принадлежат равновесные по Нэшу действия, но не принадлежит точка Парето; т. е. при векторе начальных действий  $x_i^0 = 0,5$ ,  $i \in N$ , наличия агентов первого ранга рефлексии недостаточно для реализации путем рефлексивного управления Парето-оптимальной точки. Но вполне достаточно для реализации соответствующего равновесию Нэша суммарного объема производства — для этого доля рефлексивных агентов первого уровня должна быть около 49 %.

Возможность реализации точки Парето зависит от вектора начальных действий: например, первый ранг рефлексии является максимальным целесообразным для реализации точки Парето при векторе начальных действий  $x_i^0 = 0,2$ ,  $i \in N$ . Тогда  $x_i = 1,5$ ,  $x_{1j} = 0,55$ , и при доле агентов первого ранга рефлексии, равной примерно 84 %, на рынке установится эффективная цена  $P(x^P)$ . Однако такая ситуация не будет стабильной в смысле условия (13).

Если все начальные действия агентов одинаковы, то рефлексивное разбиение задается лишь числом агентов с соответствующим рангом рефлексии, поэтому, опуская индексы, соответствующие номерам агентов, можно записать, что агенты второго ранга рефлексии выберут действия

$$x_2 = \frac{1}{1 + 2b} [a - bn_0 x - b(n - n_0 - 1)x_1].$$

Итого, получаем, что в зависимости от рефлексивного разбиения реализуется суммарное действие  $Q(n_1, n_2) = (n - n_1 - n_2)x + n_1 x_1 + n_2 x_2 = \frac{1}{1 + 2b} [(a - b(n - 1)x^0) \times$

$$\times (n - n_1 - n_2) + \frac{1}{1+2b} [a(1+3b) - nab + b^2 x^0 (n-1)^2] n_1 + \\ + \frac{1}{1+2b} [a - bn_0(n - n_1 - n_2)x] n_2].$$

Исследуем зависимость объема выпуска  $Q(n_1, n_2)$  от числа рефлексорирующих агентов первого и второго ранга и зададимся вопросом, при каких значениях  $(n_1, n_2)$  суммарный объем выпуска соответствует равновесному по Нэшу, т. е. когда выполняется  $Q(n_1, n_2) = Q(x^N)$  в зависимости от начальных действий агентов  $x^0$ .

Для рассматриваемого примера кривая  $AB$  пересечения графика  $Q(n_1, n_2)$  и «нэшевской» плоскости  $Q = 10$  приведена на рис. 1. Оказывается, что эта кривая не зависит от  $x^0$  — ее формула в плоскости  $Q = 10$ :

$$n_1 = 1 - n_2 - \frac{n-1}{\left(1 + \frac{b}{b+1}n\right)\left(\frac{b}{2b+1}n_2 - 1\right)}.$$

Из рис. 1 видно, что введение даже только агентов первого ранга увеличивает суммарный объем производства.

Отметим, что в смысле «стабильности», если имеет место динамика, то если на первом шаге агенты попадают в точку Нэша, то и в дальнейшем ни у одного из них (ни у нерефлексорирующего, ни у рефлексорирующего) нет оснований для изменения своих действий.

Если же мы ищем такое число рефлексорирующих агентов, чтобы объем производства был равен отличному от  $Q(x^N)$  значению, например объему, соответствующему Парето-оптимальной ситуации, то кривая  $AB$  будет меняться в зависимости от  $x^0$ . Оказывается что в рассматриваемом примере кривая пересечения  $Q(n_1, n_2)$  с любой плоскостью — это кривая второго порядка.

Сформулируем теперь задачу следующим образом — выбором рефлексивного разбиения реализовать требуемый суммарный объем производства, например, равный 12 (больше  $Q(x^N)$ ). Предположим, что в начальный момент времени агенты не осуществляли производства ( $x^0 = 0$ ). Достичь требуемого объема можно — см. рис. 2.

Если  $x^0 \approx 0,305$ , то кривая  $AB$  касается плоскости  $n_1 = 0$  (точка  $C$  на рис. 3), т. е., в этом случае при наличии только агентов нулевого и второго рангов рефлексии можно достичь требуемого суммарного объема производства.

Таким образом, в модели олигополии Курно введение рефлексорирующих агентов позволяет увеличить суммарный объем производства и (или) реализовать его Парето-эффективное значение.

### 6.2. Задача о консенсусе

Содержательная интерпретация «задачи о консенсусе» следующая: действиям агентов соответствуют их положения на прямой (координаты в пространстве, мнения и др. — см. обзоры в работах [14–16]), агрегированной ситуации — среднее значение координат агентов:

$$Q(x) = \frac{1}{n} \sum_{i \in N} x_i. \text{ Целевой функцией агента будем считать}$$

его «отклонение» от агрегированной ситуации:

$$(29) f(x_i, Q(x)) = -(x_i - Q(x))^2, \quad i \in N.$$

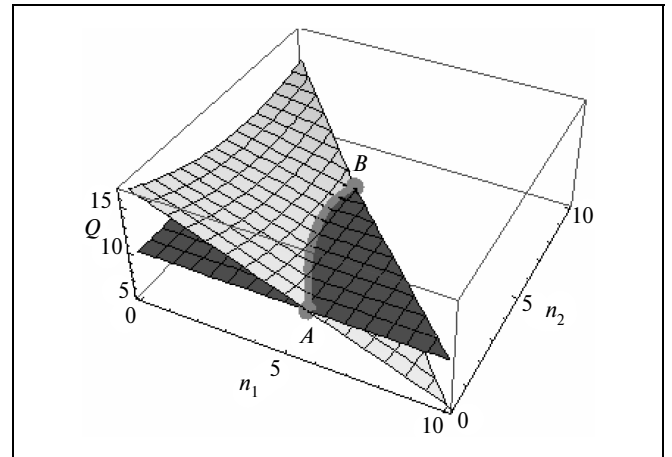


Рис. 1. Кривая  $AB$  «реализует» равновесие Нэша

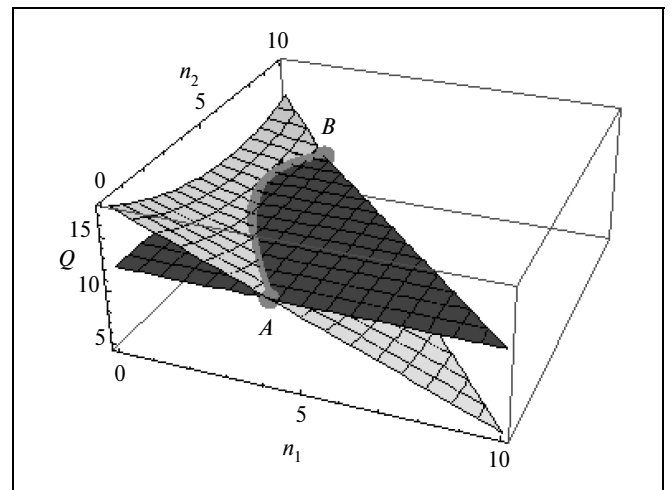


Рис. 2. «Реализация» требуемого суммарного объема производства

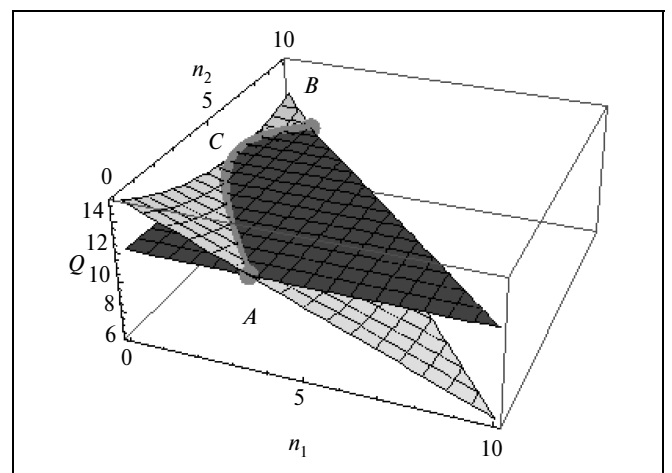


Рис. 3. «Реализация» требуемого суммарного объема производства в отсутствие агентов первого ранга (точка  $C$ )

Критерием эффективности будем считать «дисперсию» положений агентов (в данном примере целевая функция центра зависит не только от агрегированной ситуации игры, но и от всего вектора действий агентов):

$$F(x) = -\frac{1}{n} \sum_{i \in N} (Q(x) - x_i)^2. \quad (15)$$

С теоретико-игровой точки зрения ситуация тривиальна — если бы целевые функции агентов были бы среди них общим знанием, то агенты легко вычислили бы, что равновесием Нэша является любой вектор одинаковых действий. Отметим, что при этом полностью отсутствует конфликт интересов агентов, а любое равновесие Нэша однопериодной игры одновременно максимизирует и критерий эффективности (15). Однако в случае (даже одношагового) коллективного поведения агентов в условиях неполной их информированности все не так просто.

**Ранг 0.** При заданных начальных положениях агентов  $x^0$   $i$ -й агент в соответствии с выражением (9) выберет действие

$$x_i = \frac{1}{n-1} \sum_{j \neq i} x_j^0 = \frac{1}{n-1} (nQ(x^0) - x_i^0), \quad i \in N, \quad (16)$$

равное среднему положению остальных агентов. Следующий вывод остается в силе и в случае, когда целевые функции агентов зависят не от агрегированной ситуации, а от *агрегированной обстановки*:  $g(x_i, Q_i(x_{-i})) = -(x_i - Q_i(x_{-i}))^2$ , где  $Q_i(x_{-i}) = \frac{1}{n-1} \sum_{j \neq i} x_j$ ,  $i \in N$ .

Из выражения (16) следует, что  $Q(x) = Q(x^0)$ , т. е. среднее значение координат агентов не изменяется, а значение критерия эффективности возрастает в  $(n-1)^2$  раз:  $F(x) = \frac{1}{(n-1)^2} F(x^0)$ .

**Ранг 1.** Пусть имеются  $n_1$  агентов, обладающих первым рангом рефлексии, а остальные  $n_0 = n - n_1$  агентов имеют нулевой ранг. Агенты нулевого ранга рефлексии выберут действия, определяемые выражением (16), а агенты первого ранга — следующие действия:

$$x1_j = \frac{nQ(x) - x_j}{n-1} = \frac{n^2(n-2)Q(x^0) + x_j^0}{(n-1)^2}, \quad j \in N_1.$$

Если все агенты обладают первым рангом рефлексии, то  $Q(x1_{j \in N}) = Q(x) = Q(x^0)$ , т. е. среднее значение координат агентов не изменяется (такой случай является идеальным в смысле стабильности рефлексивного разбиения — все агенты наблюдают ожидаемые значения). Значение критерия эффективности возрастает еще в  $(n-1)^2$  раз:  $F(x1_{j \in N}) = \frac{1}{(n-1)^2} F(x) = \frac{1}{(n-1)^4} F(x^0)$ .

Рассмотрим пример — пусть  $n = 2$ . Получаем, что в зависимости от своих рангов рефлексии агенты выберут действия, представленные в табл. 1. Видно, что:

— вектор действий обоих агентов, обладающих вторым рангом рефлексии, совпадает с вектором действий нерелексирующих агентов;

— при одинаковых рангах рефлексии обоих агентов значение критерия эффективности не зависит от ранга;

Таблица 1

Действия агентов при  $n = 2$ 

| Начальные действия |   | Агент 1 | Агент 2 |
|--------------------|---|---------|---------|
|                    |   | $x_1^0$ | $x_2^0$ |
| Ранг рефлексии     | 0 | $x_2^0$ | $x_1^0$ |
|                    | 1 | $x_1^0$ | $x_2^0$ |
|                    | 2 | $x_2^0$ | $x_1^0$ |

— все четыре возможные комбинации действий агентов исчерпываются нулевым и первым рангами их рефлексии;

— максимальное (равное нулю) значение критерия эффективности (15) достигается в случае, когда один из агентов (любой) имеет нулевой ранг рефлексии, а другой агент — первый ранг.

Следовательно, в рассматриваемом примере максимальный целесообразный ранг рефлексии равен единице.

### 6.3. Активная экспертиза

Рассматриваемый здесь пример свидетельствует, что наличие рефлексирующих агентов может приводить к последствиям, негативным, условно говоря, с точки зрения группы в целом (см. также модели формирования команд в работе [17]).

Содержательная интерпретация модели активной экспертизы следующая: имеются  $n$  экспертов — агентов, сообщающих информацию организатору экспертизы — центру. Центр принимает решение  $Q(x) = \frac{1}{n} \sum_{i \in N} x_i$ , равное среднему арифметическому мнений агентов.

Пусть на сообщения агентов наложено требование неотрицательности. Целевой функцией агента будем считать «отклонение» итогового мнения от его начального (истинного) мнения [18]:

$$f(x_i, Q(x)) = -(x_i^0 - Q(x))^2, \quad i \in N.$$

Пусть агенты упорядочены по возрастанию их начальных мнений:  $x_1^0 < x_2^0 < \dots < x_n^0$ .

С теоретико-игровой точки зрения, если бы целевые функции агентов были среди них общим знанием, то агенты легко вычислили бы равновесие Нэша:  $x_i^N = 0$ ,  $i = \overline{1, n-1}$ ,  $x_n^N = nx_n^0$ .

Определим множество агентов  $M(x^0) = \left\{ i \in N | x_i^0 \geq \frac{1}{n} \sum_{l \neq i} x_l^0 \right\}$ .

**Ранг 0.** При заданных начальных мнениях агентов  $x^0$ , они в соответствии с выражением (9) выберут действия

$$x_i = \max \left\{ nx_i^0 - \sum_{l \neq i} x_l^0, 0 \right\}, \quad i \in N. \quad (17)$$



Таблица 2

## Действия агентов

| №        | 1 | 2 | 3 | 4 | 5    | 6  | 7  | 8  | 9  | 10 |
|----------|---|---|---|---|------|----|----|----|----|----|
| $x^0$    | 1 | 2 | 3 | 4 | 5    | 6  | 7  | 8  | 9  | 10 |
| $Q(x^0)$ |   |   |   |   | 5,5  |    |    |    |    |    |
| $x$      | 0 | 0 | 0 | 0 | 0    | 11 | 22 | 33 | 44 | 55 |
| $Q(x)$   |   |   |   |   | 16,5 |    |    |    |    |    |
| $x1$     | 0 | 0 | 0 | 0 | 0    | 0  | 0  | 0  | 0  | 0  |

$$\text{Вычислим } Q(x) = \sum_{i \in M(x^0)} x_i - \frac{1}{n} \sum_{i \in M(x^0)} \sum_{l \neq i} x_l^0.$$

**Ранг 1.** Пусть имеются  $n_1$  агентов, обладающих первым рангом рефлексии, а остальные  $n_0 = n - n_1$  агентов имеют нулевой ранг. Агенты нулевого ранга рефлексии выберут действия, определяемые выражением (17), а агенты первого ранга — следующие действия:

$$x1_j = \max \left\{ nx_j^0 - \sum_{l \in M(x^0) \setminus \{j\}} x_l; 0 \right\}, \quad j \in N_1. \quad (18)$$

Рассмотрим числовой пример. Пусть имеются 10 агентов, чьи начальные мнения равны их номеру. Действия агентов приведены в табл. 2.

Варьируя число рефлексующих агентов первого ранга (от 0 до 10), центр может менять результаты экспертизы (одиннадцать возможных точек) от 0 до 16,5. Отметим, что этот диапазон шире, чем интервал истинных мнений экспертов (сравните с результатами анализа информационной рефлексии в задачах экспертизы [19]), т. е. центр, осуществляя рефлексивное управление, имеет значительные возможности по манипулированию результатами экспертизы.

**Ранг 2.** Агенты второго ранга рефлексии выберут действия:

$$x2_j = \max \left\{ nx_j^0 - \sum_{l \in N_1 \cup N_2 \cap M(x^0) \setminus \{j\}} x1_l - \sum_{l \in N_0 \cap M(x^0) \setminus \{j\}} x_l; 0 \right\}, \quad j \in N_2. \quad (19)$$

Пусть центр использует следующее рефлексивное разбиение:  $N_0 = \{1, 2, 3, 4, 5\}$ ,  $N_1 = \{6, 7, 8, 9\}$ ,  $N_2 = \{10\}$ . Тогда в соответствии с выражениями (17) — (19) все агенты, кроме десятого, выберут нулевые действия, а десятый агент — действие, равное 100; т. е. в рассматриваемом примере второго ранга рефлексии достаточно, чтобы получить ситуацию, совпадающую с равновесной по Нэшу.

#### 6.4. Транспортные потоки и модель эвакуации

Рассмотрим помещение, в котором находятся  $n$  агентов. В помещении имеются два выхода, условно назовем их «левым» ( $L$ ) и «правым» ( $R$ ). Время выхода определяется моментом времени, когда из данного выхода вышел последний агент, направившийся к нему. Каждый агент однократно принимает решение, из какого выхода он будет выходить. Скорости движения всех агентов в отсутствие «пробок» примем одинаковыми. Обозна-

чим  $n_L$  ( $n_R$ ) — число агентов, направившихся к левому (правому) выходу,  $n_L + n_R = n$ .

Пусть известна зависимость  $T(k)$  времени выхода в зависимости от числа агентов  $k \geq 0$ . Зависимость эту будем считать непрерывной, выпуклой (отражение эффекта «пробок») и равной нулю в нуле (когда имеется один агент, «пробки» отсутствуют, и он покидает помещение без задержек). Обозначим через  $T_L$  ( $T_R$ ) время движения агента до левого (правого) выхода, причем  $T_L > T_R$ , т. е. правый выход расположен ближе левого. Итак, полное время выхода налево  $T(n_L) = T_L + T(n_L)$ , направо  $T(n_R) = T_R + T(n_R)$ .

Оптимальное в смысле времени эвакуации  $T^*$  — покидания помещения последним из агентов (а именно этот критерий используется в моделях эвакуации) — распределение агентов по направлениям движения ( $n_L^*$ ;  $n_R^*$ ) является решением следующей системы уравнений (см. также рис. 4):

$$\begin{cases} T(n_L^*) + T_L = T(n_R^*) + T_R, \\ n_L^* + n_R^* = n. \end{cases}$$

Минимальное время эвакуации

$$T^* = T(n_L^*) + T_L = T(n_R^*) + T_R. \quad (20)$$

Рассмотрим теперь коллективное поведение агентов, считая, что каждый из них стремится покинуть помещение как можно скорее. Агенты нулевого ранга рефлексии будут выбирать правый выход (до него они в рамках введенных предположений доберутся быстрее), агенты первого ранга рефлексии, прогнозируя, что в правом выходе агенты нулевого ранга создадут пробку, выберут левый выход.

Время выхода в зависимости от числа агентов первого ранга рефлексии (рис. 5)

$$T1(n_1) = \max\{T(n_1) + T_L; T(n - n_1) + T_R\}. \quad (21)$$

Видно, что как малое, так и очень большое число рефлексующих агентов плохо, так как увеличивает

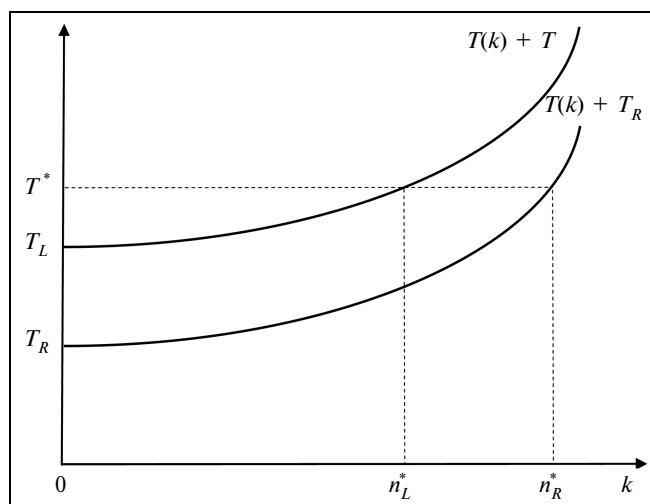


Рис. 4. Зависимость времени эвакуации от числа агентов, выбирающих правый или левый выход

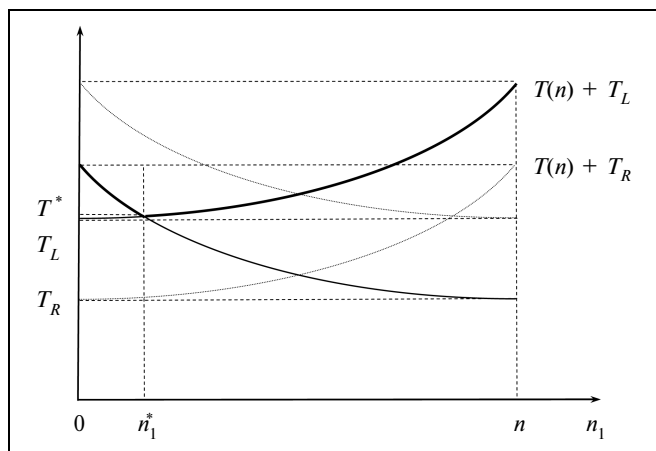


Рис. 5. Зависимость времени эвакуации от числа агентов первого ранга рефлексии

время эвакуации (см. рис. 5); т. е. существует оптимальное число рефлексизирующих агентов, при котором время эвакуации минимально.

Из свойств функции  $T(\cdot)$  и предположения  $T_L > T_R$  следует, что минимум выражения (21) достигается при числе агентов первого ранга рефлексии  $n_1^*$ , определяемом из соотношения

$$T(n_1^*) + T_L = T(n - n_1^*) + T_R.$$

Последнее условие совпадает с условием (20), т. е.  $n_1^* = n_L^*$ ,  $T_1(n_1^*) \equiv T^*$ , значит, первый ранг рефлексии является максимальным целесообразным в рамках рассматриваемой модели.

В рассматриваемой модели можно добавлять агентов второго, третьего и более высоких рангов рефлексии, однако это вряд ли целесообразно, так как не позволит улучшить уже достигнутое путем введения агентов первого ранга значение времени эвакуации (20). Описание имитационных моделей транспортных потоков и эвакуации можно найти в работах [8, 9, 20].

#### 6.5. Фондовый рынок: пример возможных расширений модели

Обсудим возможные расширения описанного метода рефлексивных разбиений, а именно — на примере частной модели фондового рынка рассмотрим стратегическую рефлексию агентов «над» их равновесными по Нэшу стратегиями. Фондовый рынок представляет собой объект моделирования, для которого наиболее часто используют «рефлексивные» рассуждения — см., например, работы [21, 22]. В работе [23] рассмотрена теоретико-игровая модель фондового рынка, в которой каждый агент в каждый момент времени обладает некоторым количеством (для которого выполняются динамические балансовые ограничения) денег и актива, который он может приобретать или продавать по сложившейся на рынке цене. Последняя зависит как от тренда  $\theta$  (внешний фактор, являющийся общим знанием), так и от соотношения между спросом и предложением — с ростом спроса рыночная цена на актив растет, с ростом предложения — падает. В указанной работе показано, что в условиях общего знания агентов обо всех параметрах

игры структура равновесия Нэша такова: либо все агенты приобретают актив на все имеющиеся у них средства (если они тем самым «увеличивают» относительную цену актива), либо все агенты продают все имеющиеся у них активы (если они тем самым «уменьшают» относительную цену актива).

Рассмотрим следующую модель. Пусть каждый агент обладает в начальный момент времени суммой  $u_0 \geq 0$  и активом  $x_0 \geq 0$ . В соответствии с результатами работы [23] в начальный момент времени у агента имеются две альтернативы: либо приобрести актив на всю сумму  $u_0$ , либо продать все  $x_0$  единиц актива (рынок при этом не ограничен).

В зависимости от действий  $x$  агента сложится следующая цена: если все агенты приобретают актив, то цена  $p$  будет  $p^+ = p_0 + \theta + \alpha n x_0$ ; если агенты продают актив, то цена  $p$  будет  $p^- = p_0 + \theta - \alpha n x_0$ , где  $\alpha$  — коэффициент зависимости цены от спроса-предложения.

Начальное значение целевой функции агента  $u_0 + x_0 p_0$ , конечно:

- $(x_0 + u_0/p_0)p^+ - u_0$ , если актив приобретается с намерением последующей продажи;
- $u_0 + x_0 p_0$ , если актив продается;
- $u_0 + x_0(p_0 + \theta)$ , если агент не предпринимает никаких действий.

Для того чтобы выяснить, какое из трех действий (покупать, продавать или ничего не делать) предпримет рациональный агент, необходимо сравнить три полученные величины. Получаем, что, если имеет место положительный тренд ( $\theta \geq 0$ ) или если тренд отсутствует ( $\theta = 0$ ), то актив следует приобретать. При отрицательном тренде ( $\theta < 0$ ) дело обстоит сложнее, а именно актив следует приобретать при условии

$$\theta \geq \frac{p_0 u_0}{p_0 x_0 + u_0} - \alpha n x_0, \quad (22)$$

которое означает, что если агенты, приобретая актив и повышая тем самым его цену в следующем периоде, могут «перебороть» отрицательный тренд, то актив следует приобретать. В противном случае актив им следует продавать.

Если подходить более корректно и исследовать все соотношения между параметрами, т. е. для каждого из трех действий найти условия, при которых данное действие оптимально, то получим, что рациональный агент должен придерживаться следующего алгоритма: приобретать актив, если выполнено условие (22), и продавать его, если верно обратное соотношение. Интересно, что пассивное поведение — не предпринимать никаких действий — невыгодно ни при одной комбинации параметров модели.

Качественный вывод из проведенного анализа следующий. Существование постоянного тренда цены актива относительно «стоимости» денег приводит к тому, что, если этот тренд положительный, то следует вкладывать все деньги в приобретение актива. Если тренд отрицательный, то наоборот — целесообразно избавляться от актива. Возможность влияния агентами на цену актива путем своих действий (покупки или продажи) приводит к тому, что приобретать актив в случае отрицательного тренда имеет смысл только в том случае, если этими действиями можно «преодолеть» тренд.



Итак, мы описали равновесие Нэша агентов. Рассмотрим теперь рассуждения рефлексизирующего агента первого ранга. Если выполнено условие (22), то он может спрогнозировать, что все агенты нулевого ранга будут приобретать актив. Если условие (22) не выполнено, то он может спрогнозировать, что все агенты нулевого ранга будут продавать актив (цена на него упадет) и ему выгодно действовать так же. Получаем, что действия рефлексизирующих агентов будут такие же, как и нерефлексизирующих, т. е. в рассмотренной модели добавление рефлексизирующих агентов любого ранга не меняет рыночной цены.

Сделанный вывод следует из того, что мы рассмотрели достаточно «интеллектуальных» нерефлексизирующих агентов. Действительно, предполагалось, что они способны прогнозировать изменение рыночной цены в зависимости от своих действий.

Рассмотрим другую модель с менее «интеллектуальными» агентами нулевого ранга, а именно предположим, что они ориентируются лишь на знак тренда. Тогда при положительном тренде агенты нулевого ранга будут приобретать актив, в результате чего его цена будет расти, и рефлексизирующим агентам лишь остается следовать их примеру. Ситуация меняется при отрицательном тренде — агенты нулевого ранга будут продавать актив, в результате чего цена «еще более снизится». Но, рефлексизирующие агенты могут попытаться своими действиями (приобретая актив) «переломить тренд». Для этого, правда, им необходимо быть уверенными, что доля  $q$  рефлексизирующих агентов является среди них общим знанием и что эта доля достаточна для того, чтобы цена выросла. Последнее условие по аналогии с условием (22) можно записать в виде:

$$\theta \geq \frac{p_0 u_0}{p_0 x_0 + u_0} + \alpha n(1 - 2q)x_0,$$

т. е.

$$q \geq q^* = \frac{1}{2} + \frac{1}{2\alpha n x_0} \left[ \frac{p_0 u_0}{p_0 x_0 + u_0} - \theta \right]. \quad (23)$$

Отметим, что критическая доля  $q^*$  рефлексизирующих агентов составляет не менее половины от общего числа агентов (условие  $q^* \leq 1$  эквивалентно условию (22)). Рассмотрим числовой пример. Пусть  $n = 100$ ,  $u_0 = 1000$ ,  $p_0 = 10$ ,  $x_0 = 100$ ,  $\alpha = 0,001$ ,  $\theta = -1$ . Условие (22) выполнено. Из выражения (23) находим  $q^* = 53\%$ .

Подчеркнем, что предположение о том, что доля рефлексизирующих агентов является среди них общим знанием, противоречит введенному ранее предположению о структуре субъективных рефлексивных разбиений (см. выражение (8)), так как последнее предполагает, что рефлексизирующие агенты «не знают о существовании» других агентов того же ранга рефлексии (и более высоких рангов). К росту рыночной цены при отрицательном тренде будет приводить любое рефлексивное разбиение, при котором доли рефлексизирующих агентов любых рангов (кроме нулевого) в сумме превышают  $q^*$ , и эта информация является общим знанием среди рефлексизирующих агентов соответствующих уровней. Данное утверждение, имеющее прозрачные содержательные интерпретации, свидетельствует, что структура субъективных рефлексивных разбиений, определяемая выражением (8), не является единственно возможной и адекватной всем моделям, представляющим интерес для практи-

ки; т. е., перспективным направлением будущих исследований представляется рассмотрение и других структур субъективных рефлексивных разбиений.

## ЗАКЛЮЧЕНИЕ

Таким образом, 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 устойчивые исходы взаимодействия агентов; соответственно, в рамках задач управления — расширить область управляемости;
- с нормативной точки зрения — ставить и решать задачи группового управления путем подбора структуры информированности агентов.

Анализ рассмотренных примеров позволяет констатировать, что наличие рефлексизирующих агентов может изменять групповое поведение самым разным образом.

В примере «Олигополия Курно» (п. 6.1) при определенном диапазоне значений начальный действия агентов можно реализовать эффективные по Парето или равновесные по Нэшу уровни производства путем введения агентов первого и второго рангов рефлексии.

В примере «Задача о консенсусе» (п. 6.2) введение рефлексизирующих агентов расширяет множество векторов действий, выбираемых агентами, и приводит к росту значения критерия эффективности.

В примере «Активная экспертиза» (п. 6.3) наличие рефлексизирующих агентов даже первого ранга существенно расширяет диапазон возможных результатов экспертизы. Второй ранг рефлексии позволяет реализовать равновесие Нэша.

В примере «Транспортные потоки и модель эвакуации» (п. 6.4) наличие рефлексизирующих агентов первого ранга позволяет достичь минимального (оптимального с «централизованной» точки зрения) времени эвакуации.

В примере «Фондовый рынок» (п. 6.5) показано, что изменить ситуацию (по сравнению с нерефлексивным принятием решений) может только определенная «критическая масса» рефлексизирующих агентов.

В заключение отметим, что в настоящей работе:

- почти не рассматривались агенты со вторым и более высокими рангами рефлексии (либо они

превышают максимальный целесообразный ранг, либо соответствующие модели получаются слишком сложными для получения аналитических выводов);

- считалось, что агенты любого ранга рефлексии достаточно «интеллектуальны» — они выбирают действия, стремясь максимизировать свои целевые функции. Можно допустить наличие и менее интеллектуальных агентов — *агенто-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, такие модели могут адекватно описывать такое явление как диффузия инноваций и др.;
- явно недостаточное внимание уделено условиям стабильности.

Представляется перспективным установление соответствия и совместное развитие метода рефлексивных разбиений с теорией когнитивных иерархий (в которой рангам рефлексии соответствуют когнитивные уровни и используется вероятностная модель — игрок некоторого уровня считает остальных распределенными по более низким уровням в соответствии с распределением Пуассона) — направление, активно развиваемое в экспериментальной экономике и поведенческой теории игр (см., например, работы [24, 25]).

Кроме того, задачи управления, поиска максимального целесообразного ранга рефлексии и ряд других можно и нужно ставить и решать в рамках и иных (отличных от рассмотренных здесь) модификаций предложенной рефлексивной модели коллективного поведения, что представляется перспективным направлением будущих исследований. В первую очередь, это — задачи *активного прогноза* [26], в рамках которого агенты по информации центра о будущем состоянии системы «восстанавливают» текущее состояние и на основании этой новой информации принимают решения. Здесь введение рефлексивных разбиений выглядит весьма многообещающим.

## ЛИТЕРАТУРА

1. Губко М.В., Новиков Д.А. Теория игр в управлении организационными системами. — М.: СИНТЕГ, 2002.
2. Новиков Д.А., Чхартишвили А.Г. Рефлексивные игры. — М.: СИНТЕГ, 2003.
3. Myerson R. Game theory: analysis of conflict. — London: Harvard Univ. Press, 1991.
4. Васин А.А. Некооперативные игры в природе и обществе. — М.: МАКС Пресс, 2005.
5. Малишевский А.В. Качественные модели в теории сложных систем. — М.: Наука, 1998.

6. Опоицев В.И. Равновесие и устойчивость в моделях коллективного поведения. — М.: Наука, 1977.
7. Новиков Д.А. «Когнитивные игры»: линейная импульсная модель // Проблемы управления. 2008. — № 3. — С. 14—22.
8. 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., 2009. — Т. 2. — С. 69—71.
9. Корепанов В.О. Влияние рефлексивных агентов на транспортный поток / Тр. междунар. науч.-практ. конф. «Теория активных систем». Т. II. — М.: ИПУ РАН, 2009. — С. 122—125.
10. Чхартишвили А.Г. Теоретико-игровые модели информационного управления. — М.: ПМСОФТ, 2005.
11. Романько А.Д., Чхартишвили А.Г. Моделирование информационных воздействий в рефлексивных играх: простые сообщения // Сб. тр. ВГАСУ. 2006. — С. 157—167.
12. Чхартишвили А.Г. Рефлексивные игры: трансформация структур информированности // Проблемы управления. — 2008. — № 5. — С. 43—48.
13. Mas-Collel A., Whinston M., Green J. Microeconomic theory. — N.-Y.: Oxford Univ. Press, 1995.
14. Чеботарев П.Ю., Агаев Р.П. Согласование характеристик в многоагентных системах и спектры лапласовских матриц орграфов // Автоматика и телемеханика. — 2009. — № 3. — С. 136—151.
15. Shoham Y., Leyton-Brown K. Multiagent systems: Algorithmic, Game-Theoretical and Logical Foundations. — Cambridge: Cambridge University Press, 2009.
16. Wei R. Consensus Seeking, Formation Keeping and Trajectory Tracking in Multiple Vehicle Cooperative Control. PhD Dissertation. — Brigham Young University, 2004.
17. Новиков Д.А. Математические модели формирования и функционирования команд. — М.: Физматлит, 2007.
18. Новиков Д.А. Теория управления организационными системами / 2-е изд. — М.: Физматлит, 2007.
19. Новиков Д.А., Чхартишвили А.Г. Прикладные модели информационного управления. — М.: ИПУ РАН, 2004.
20. Корепанов В.О. Модель рефлексивного группового управления в организационных системах / Тр. VII молодежной школы-конференции «Управление большими системами». — Пермь: ПГТУ, 2010. — Т. 1. — С. 287—294.
21. Ерешко Ф.И. Моделирование рефлексивных стратегий в управляемых системах. — М.: ВЦ РАН, 2001.
22. Сорос Д. Алхимия финансов. — М.: ИНФРА-М, 1999.
23. 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.
24. Explaining Focal Points: Cognitive Hierarchy Theory versus Team Reasoning / N. Bardsley, J. Mehta, C. Starmer, R. Sugden // CeDEx Discussion Paper Series / The University of Nottingham. — Nottingham, 2008. — CeDEx Discussion Paper No 17. — 56 p.
25. Camerer C., Ho T., Chong J. A Cognitive hierarchy model of games // The Quarterly Journal of Economics. — 2004. — № 8. — P. 861—898.
26. Новиков Д.А., Чхартишвили А.Г. Активный прогноз. — М.: ИПУ РАН, 2002.

Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым.

Корепанов Всеволод Олегович — ст. математик,  
☎ (495) 334-90-51, ✉ moskvo@gmail.com,

Новиков Дмитрий Александрович — чл.-корр. РАН,  
зам. директора по научной работе,  
☎ (495) 334-75-69, ✉ novikov@ipu.ru,

Институт проблем управления РАН им. В.А. Трапезникова,  
г. Москва.