

Российская Академия Наук
Институт проблем управления
им. В.А. Трапезникова

Д.А. НОВИКОВ, А.Г. ЧХАРТИШВИЛИ

РЕФЛЕКСИВНЫЕ ИГРЫ

СИНТЕГ
Москва – 2003

УДК 519
ББК 22.18
Н 73

Новиков Д.А., Чхартишвили А.Г. **Рефлексивные**
Н 73 **игры**. М.: СИНТЕГ, 2003. – 160 с.

ISBN 5-89638-63-1

Монография посвящена обсуждению современных подходов к математическому моделированию рефлексии. Авторы вводят в рассмотрение новый класс теоретико-игровых моделей – *рефлексивные игры*, описывающие взаимодействие субъектов (*агентов*), принимающих решения на основании иерархии представлений о существенных параметрах, представлений о представлениях и т.д.

Анализ поведения *фантомных агентов*, существующих в представлениях других реальных или фантомных агентов, и свойств *информационной структуры*, отражающей взаимную информированность реальных и фантомных агентов, позволяет предложить в качестве решения рефлексивной игры *информационное равновесие*, которое является обобщением ряда известных концепций равновесия в некооперативных играх.

Рефлексивные игры дают возможность:

- моделировать поведение рефлексизирующих субъектов;
- исследовать зависимость выигрышей агентов от рангов их рефлексии;
- ставить и решать задачи *рефлексивного управления*;
- единообразно описывать многие явления, связанные с рефлексией: скрытое управление, информационное управление через СМИ, рефлексии в психологии, художественных произведениях и др.

Книга адресована специалистам в области математического моделирования и управления социально-экономическими системами, а также студентам вузов и аспирантам.

*Рецензенты: д.т.н., проф. В.Н. Бурков,
д.т.н., проф. А.В. Щепкин*

УДК 519
ББК 22.18
Н 73

ISBN 5-89638-63-1

О Д.А.Новиков, А.Г. Чхартишвили, 2003

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
ГЛАВА 1. Информация в принятии решений	22
1.1. Индивидуальное принятие решений: модель рационального поведения	22
1.2. Интерактивное принятие решений: игры и равновесия	25
1.3. Общие подходы к описанию информированности	32
ГЛАВА 2. Стратегическая рефлексия	35
2.1. Стратегическая рефлексия в играх двух лиц	35
2.2. Рефлексия в биматричных играх	42
2.3. Ограниченность ранга рефлексии	58
ГЛАВА 3. Информационная рефлексия	61
3.1. Информационная рефлексия в играх двух лиц	61
3.2. Информационная структура игры	66
3.3. Информационное равновесие	72
3.4. Граф рефлексивной игры	77
3.5. Регулярные структуры информированности	83
3.6. Ранг рефлексии и информационное равновесие	92
3.7. Рефлексивное управление	103
ГЛАВА 4. Прикладные модели рефлексивных игр	107
4.1. Скрытое управление	107
4.2. СМИ и информационное управление	118
4.3. Рефлексия в психологии	122
4.3.1. Психология шахматного творчества	122
4.3.2. Трансакционный анализ	125
4.3.3. Окно Джохари	127
4.3.4. Модель этического выбора	129
4.4. Рефлексия в художественных произведениях	130
ЗАКЛЮЧЕНИЕ	138
ЛИТЕРАТУРА	143

- Пескари привольно резвятся, в этом их радость!
- Ты же не рыба, откуда тебе знать, в чем ее радость?
- Ты же не я, откуда тебе знать, что я знаю, а чего не знаю?

Из даосской притчи

- Дело, разумеется, в том, достопочтенный архиепископ, что Вы верите в то, во что Вы верите, потому что Вы были так воспитаны.
- Может быть, и так. Но остается фактом, что и Вы верите в то, что я верю в то, во что я верю, потому что я был так воспитан, по той причине, что Вы были так воспитаны.

Из книги Д. Майерса «Социальная психология»

ВВЕДЕНИЕ

Настоящая работа посвящена обсуждению современных подходов к математическому моделированию рефлексии и, в первую очередь, введению в рассмотрение нового класса теоретико-игровых моделей – рефлексивных игр, описывающих взаимодействие субъектов, принимающих решения на основании иерархии представлений о существенных параметрах, представлений о представлениях и т.д.

Рефлексия. Одним из фундаментальных свойств бытия человека является то, что наряду с природной («объективной») реальностью существует ее отражение в сознании. При этом между *природной реальностью* и ее образом в сознании (будем считать этот образ частью особой – *рефлексивной реальности*) существует неизбежный зазор, несовпадение.

Целенаправленное изучение этого феномена традиционно связано с термином «*рефлексия*», которому «Философский словарь» [84] дает следующее определение: «РЕФЛЕКСИЯ (лат. reflexio – обращение назад). Термин, означающий отражение, а также исследование познавательного акта».

Термин «рефлексия» введен Дж. Локком; в различных философских системах (у Дж. Локка, Г. Лейбница, Д. Юма, Г. Гегеля и др.) он имел различное содержание. Систематическое описание рефлексии с точки зрения психологии началось в 60-е годы XX века (школа

В.А. Лефевра). Кроме того, следует отметить, что существует понимание рефлексии в другом значении, имеющем отношение к рефлексу – «реакции организма на возбуждение рецепторов» [78; С. 1122]. В настоящей работе используется первое (философское) определение рефлексии.

Для прояснения понимания сути рефлексии рассмотрим сначала ситуацию с одним субъектом. У него есть представления о природной реальности, но он может и осознавать (отражать, рефлексировать) эти представления, а также осознавать осознание этих представлений и т.д. Так формируется рефлексивная реальность. Рефлексия субъекта относительно своих собственных *представлений* о реальности, принципах своей деятельности и т.д. называется *авторефлексией* или *рефлексией первого рода*. Отметим, что в большинстве гуманитарных исследований речь идет, в первую очередь, об авторефлексии, под которой в философии понимается процесс размышления индивида о происходящем в его сознании [55].

Рефлексия второго рода имеет место относительно представлений о реальности, принципах принятия решений, авторефлексии и т.д. других субъектов.

Приведем примеры рефлексии второго рода, иллюстрирующие, что во многих случаях правильные собственные умозаключения можно сделать, лишь если занять позицию других субъектов и проанализировать их возможные рассуждения.

Первым примером является классическая «задача о грязных лицах» (Dirty Face Game) [110], иногда ее называют «задачей о мудрецах и колпаках» [22] или «о мужьях и неверных женах» [132]. Опишем ее, следуя [22, С. 46].

«Представим себе, что в купе вагона Викторианской эпохи находятся Боб и его племянница Алиса. У каждого испачкано лицо. Однако никто не краснеет от стыда, хотя любой Викторианский пассажир покраснел бы, зная, что другой человек видит его грязным. Отсюда мы делаем вывод, что никто из пассажиров не знает, что его лицо грязное, хотя каждый видит грязное лицо своего компаньона.

В это время в купе заглядывает Проводник и объявляет, что в купе находится человек с грязным лицом. После этого Алиса покраснела. Она поняла, что лицо у нее испачкано. Но почему она поняла это? Разве Проводник не сообщил то, что она уже знала?

Проследим цепочку рассуждений Алисы. Алиса: Предположим, мое лицо чистое. Тогда Боб, зная, что кто-то из нас грязный, должен сделать вывод, что грязный он, и покраснеть. Раз он не краснеет, значит, моя посылка про мое чистое лицо ложная, мое лицо грязное и я должна покраснеть.

Проводник добавил к информации, известной Алисе, информацию о знаниях Боба. До этого она не знала, что Боб знает, что кто-то из них испачкан. Короче, сообщение проводника превратило знание о том, что в купе есть человек с грязным лицом, в общее знание».

Второй хрестоматийный пример – «задача о скоординированной атаке» (Coordinated Attack Problem) [112]; существуют близкие к ней задачи об оптимальном протоколе обмена информацией – Electronic Mail Game [138] и др. (см. обзоры в [105, 113]).

Ситуация выглядит следующим образом. На вершинах двух холмов расположены две дивизии, а в долине расположился противник. Одержат победу можно, только если обе дивизии нападут на противника одновременно. Генерал – командир первой дивизии – посылает генералу – командиру второй дивизии – гонца с сообщением: «Атакуем на рассвете». Так как гонец может быть перехвачен противником, то первому генералу необходимо дожидаться от второго генерала сообщения о том, что первое сообщение получено. Но так как второе сообщение также может быть перехвачено противником, то второму генералу необходимо получить от первого подтверждение, что тот получил подтверждение. И так далее до бесконечности. Задача заключается в том, чтобы определить, после какого числа сообщений (подтверждений) генералам имеет смысл атаковать противника. Вывод следующий – в описанных условиях скоординированная атака невозможна, а выходом является использование вероятностных моделей [130, 131].

Третья классическая задача – «задача о двух брокерах» [94] (см. также модели спекуляций в [66]). Предположим, что у двух брокеров, играющих на фондовой бирже, имеются собственные экспертные системы, которые используются для поддержки принятия решений. Случается так, что сетевой администратор нелегально копирует обе экспертные системы и продает каждому брокеру экспертную систему своего оппонента. После этого администратор пытается продать каждому из них следующую информацию – «Ваш оппонент имеет Вашу экспертную систему». Потом администратор пытается

продать информацию – «Ваш оппонент знает, что Вы имеете его экспертную систему», и т.д. Вопрос заключается в том, как брокерам следует использовать информацию, получаемую от администратора, а также какая информация на какой итерации является существенной?

Завершив рассмотрение примеров рефлексии второго рода, обсудим в каких ситуациях рефлексия является существенной. Если единственный рефлексирующий субъект является экономическим агентом, который стремится максимизировать свою целевую функцию, выбирая одно из этически допустимых действий, то природная реальность входит в целевую функцию как некий параметр, а результаты рефлексии (представления о представлениях и пр.) аргументами целевой функции не являются. Тогда можно сказать, что авторефлексия «не нужна», так как она не изменяет действия, выбираемого агентом.

Заметим, что зависимость действий субъекта от рефлексии может иметь место в ситуации, когда действия этически неравноценны, то есть наряду с *утилитарным* аспектом существует деонтологический (*этический*) – см. [92, 120-122]. Однако экономические решения, как правило, этически нейтральны, поэтому рассмотрим взаимодействие нескольких субъектов.

Если субъектов несколько (ситуация принятия решения является *интерактивной*), то в целевую функцию каждого субъекта входят действия других субъектов, то есть эти действия являются частью природной реальности (хотя сами они, разумеется, обусловлены рефлексивной реальностью). При этом рефлексия (и, следовательно, исследование рефлексивной реальности) становится необходимой. Рассмотрим основные подходы к математическому моделированию эффектов рефлексии.

Теория игр. Формальные (математические) модели поведения человека создаются и изучаются уже более полутора веков (см. обзор в [1]) и находят все большее применение как в теории управления, экономике, психологии, социологии и т.д., так и при решении конкретных прикладных задач. Наиболее интенсивное развитие наблюдается начиная с 40-х годов XX века – момента появления *теории игр*, который обычно датируют 1944 годом (выход первого издания книги Джона фон Неймана и Оскара Моргенштерна «Теория игр и экономическое поведение» [53]).

Под *игрой* в данной работе будем понимать взаимодействие сторон, интересы которых не совпадают (отметим, что возможно и другое понимание игры – как «вида непродуктивной деятельности, мотив которой заключается не в ее результатах, а в самом процессе» [78, С. 475] – см. также [87], где понятие игры трактуется гораздо более широко).

Теория игр – раздел прикладной математики, исследующий модели принятия решений в условиях несовпадения интересов сторон (*игроков*), когда каждая сторона стремится воздействовать на развитие ситуации в собственных интересах [21]. Далее для обозначения субъекта, принимающего решения (игрока), используется термин «*агент*». В настоящей работе рассматриваются некооперативные статические игры в нормальной форме, то есть игры, в которых агенты однократно, одновременно и независимо выбирают свои *действия*.

Таким образом, основная задача теории игр заключается в описании взаимодействия нескольких агентов, интересы которых не совпадают, а результаты деятельности (выигрыш, полезность и т.д.) каждого зависят в общем случае от действий всех [21, 132]. Итогом подобного описания является прогноз разумного исхода игры – так называемого *решения игры (равновесия)*.

Описание *игры* заключается в задании следующих параметров:

- *множества агентов*;
- *предпочтений агентов* (зависимостей выигрышей от действий): при этом предполагается (и этим отражается целенаправленность поведения), что каждый агент заинтересован в максимизации своего выигрыша;
- *множеств допустимых действий агентов*;
- *информированности агентов* (той информации, которой они обладают на момент принятия решений о выбираемых действиях);
- *порядка функционирования (порядок ходов* – последовательность выбора действий).

Условно говоря, множество агентов определяет, кто участвует в игре. Предпочтения отражают, что хотят агенты, множества допустимых действий – что они могут, информированность – что они знают, а порядок функционирования – когда они выбирают действия.

Перечисленные параметры задают игру, но они недостаточны для того, чтобы предсказать ее исход – решение игры (или равновесие игры), то есть множество рациональных и устойчивых с той или иной точки зрения действий агентов [15, 21, 22]. На сегодняшний день в теории игр не существует универсальной концепции равновесия – принимая те или иные предположения о принципах принятия агентами решений, можно получать различные решения. Поэтому основной задачей любого теоретико-игрового исследования (включая настоящую работу) является построение равновесия. Так как рефлексивные игры определяются как такое интерактивное взаимодействие агентов, в котором они принимают решения на основе иерархии своих представлений, то существенной является информированность агентов. Поэтому остановимся на ее качественном обсуждении более подробно.

Роль информированности. Общее знание. В теории игр, философии, психологии, распределенных системах и других областях науки (см. обзор в [111, 131]) существенны не только *представления* (beliefs) агентов о существенных параметрах, но и их представления о представлениях других агентов и т.д. Совокупность этих представлений называется *иерархией представлений* (hierarchy of beliefs) и в настоящей работе моделируется деревом информационной структуры рефлексивной игры (см. раздел 3.2). Другими словами, в ситуациях интерактивного принятия решений (моделируемых в теории игр) каждый агент перед выбором своего действия должен предсказать поведение оппонентов. Для этого у него должны быть определенные представления о видении игры оппонентами. Но оппоненты должны проделать то же самое, поэтому неопределенность относительно той игры, которая будет разыграна, порождает бесконечную иерархию представлений участников игры.

Приведем пример иерархии представлений. Предположим, что имеются два агента – А и Б. Каждый из них может иметь собственные нерефлексивные представления о неопределенном параметре q , который мы будем в дальнейшем называть *состоянием природы* (state of nature, state of the world). Обозначим эти представления q_A и q_B соответственно. Но каждый из агентов в рамках процесса *рефлексии первого ранга* может задуматься о представлениях оппонента. Эти представления (*представления второго порядка*) обозначим q_{AB} и q_{BA} , где q_{AB} – представления агента А о представлениях агента Б,

q_{BA} – представления агента Б о представлениях агента А. Но этим дело не ограничивается – каждый из агентов в рамках процесса дальнейшей рефлексии (*рефлексии второго ранга*) может задуматься над тем, каковы представления оппонента о его представлениях. Так порождаются представления *третьего порядка* – q_{ABA} и q_{BAB} . Процесс порождения представлений более высоких порядков может продолжаться до бесконечности (никаких логических ограничений увеличению ранга рефлексии не существует). Совокупность всех представлений – $q_A, q_B, q_{AB}, q_{BA}, q_{ABA}, q_{BAB}$ и т.д. – образует иерархию представлений.

Частным случаем информированности – когда все представления, представления о представлениях и т.д. до бесконечности совпадают – является *общее знание*. Более корректно, термин «общее знание» (common knowledge), введен в [123] для обозначения факта, удовлетворяющего следующим требованиям:

- 1) о нем известно всем агентам;
- 2) всем агентам известно 1;
- 3) всем агентам известно 2 и т.д. до бесконечности

Формальная модель общего знания предложена в [96] и получила развитие во множестве работ – см. [97, 99, 106, 107, 108, 113, 116, 130, 140 и др.].

Моделям информированности агентов – иерархии представлений и общему знанию – в теории игр посвящена, фактически целиком, настоящая работа, поэтому приведем примеры, иллюстрирующие роль общего знания в других областях науки – философии, психологии и др. (см. также обзор [105]).

С точки зрения философии общее знание анализировалось при изучении *соглашений* [123, 142]. Рассмотрим следующий пример. В Правилах Дорожного Движения записано, что каждый участник дорожного движения должен соблюдать эти правила, а также вправе рассчитывать на то, что их соблюдают другие участники дорожного движения. Но другие участники дорожного движения также должны быть уверены в том, что остальные соблюдают правила, и т.д. до бесконечности. Следовательно, соглашение «соблюдать ПДД» должно быть общим знанием.

В психологии существует понятие *дискурса* – «(от лат. discursus – рассуждение, довод) – опосредованное прошлым опытом речевое мышление человека; выступает как процесс связанного логического

рассуждения, в котором каждая последующая мысль обусловлена предыдущей» [76, С. 99)]. Роль общего знания в понимании дискурса иллюстрируется в [104, 105] следующим примером.

Два человека выходят из кинотеатра. Один спрашивает другого: «Как тебе фильм?». Для того чтобы второй человек понял вопрос, он должен понять, что его спрашивают о том фильме, который они только что вместе посмотрели. Кроме того, он должен понимать, что это понимает первый. Задающий вопрос, в свою очередь, должен быть уверен, что второй поймет, что речь идет о том фильме, который они посмотрели, и т.д. То есть для адекватного взаимодействия (общения) «фильм» должен быть общим знанием (люди должны достичь соглашения об использовании языка [123]).

Взаимная информированность агентов является существенной также в распределенных вычислительных системах [106, 108, 113], в искусственном интеллекте [112, 127] и других областях.

В теории игр, как правило, предполагается, что все¹ параметры игры являются *общим знанием*, то есть каждому агенту известны все параметры игры, а также то, что это известно всем агентам, и т.д. до бесконечности. Такое предположение соответствует *объективному описанию игры* и дает возможность использовать концепцию *равновесия Нэша*² [134] как прогнозируемого исхода некооперативной игры (то есть игры, в которой невозможны переговоры между агентами с целью создания коалиций, обмена информацией, совместных действий, перераспределения выигрышей и т.д.). Таким образом, предположение об общем знании позволяет утверждать, что все агенты знают, в какую игру они играют, и их представления об игре совпадают.

Вместо действия агента можно рассматривать нечто более сложное – его *стратегию*, то есть отображение имеющейся у агента информации во множество его допустимых действий. Примерами могут служить: стратегии в многошаговой игре, смешанные стратегии, стратегии в метаиграх Ховарда [117, 118] (см. также информа-

¹ Если в исходной модели присутствуют неопределенные факторы, то используются процедуры устранения неопределенности, которые позволяют получить детерминированную модель.

² Вектор действий агентов является равновесием Нэша, если никому из них не выгодно одностороннее (то есть при условии, что остальные агенты выбирают соответствующие компоненты равновесия) отклонение от равновесия – см. корректное определение ниже.

ционные расширения игр [18, 39, 40]). Однако и в этих случаях правила игры являются общим знанием. Наконец, можно считать, что игра выбирается случайным образом в соответствии с некоторым распределением, которое является общим знанием – так называемые *Байесовы игры* [109, 114, 132].

В общем случае каждый из агентов может иметь собственные представления о параметрах игры, каждому из которых соответствует некоторое *субъективное описание игры* [18]. При этом оказывается, что агенты участвуют в игре, но объективно не знают в какой, или по-разному представляют разыгрываемую игру – ее правила, цели, роли и информированность оппонентов и т.д. Универсальных подходов к построению равновесий при недостаточном общем знании на сегодняшний день в теории игр не существует.

С другой стороны, в рамках «рефлексивной традиции» гуманитарных наук для каждого агента окружающий его мир содержит (включает) остальных агентов, и представления о других агентах отражаются в процессе рефлексии (различия представлений могут быть обусловлены, в частности, неодинаковой информированностью). Однако до настоящего момента конструктивных формальных результатов в этой области получено не было.

Следовательно, возникает необходимость разработки и исследования математических моделей игр, в которых информированность агентов не является общим знанием и агенты принимают решения на основе иерархии своих представлений. Этот класс игр назовем *рефлексивными играми* (формальное определение приведено в разделе 3.2 настоящей работы).

Следует признать, что термин «рефлексивные игры» был введен В.А. Лефевром в 1965 г. в [42]. Однако в этой работе, а также в работах [43-47, 122] того же автора содержится, в основном, качественное обсуждение эффектов рефлексии во взаимодействии субъектов, и никакой общей концепции решения для этого класса игр предложено не было. То же замечание справедливо и для [19, 24-26, 66, 79], в которых рассматривался ряд частных случаев информированности участников игры.

Таким образом, актуальным является изучение рефлексивных игр и построение для них единой концепции равновесия, что и мотивирует настоящее исследование.

Прежде чем переходить к изложению основного содержания работы, обсудим на качественном уровне основные используемые ниже подходы.

Основные подходы и структура работы. В первой главе «Информация в принятии решений», носящей, в основном, обзорный и вводный характер, приводятся модели индивидуального и интерактивного принятия решений, проводится анализ информированности, необходимой для реализации тех или иных известных концепций равновесия, а также обсуждаются известные модели общего знания и иерархии представлений.

Как определено выше, рефлексивной является игра, в которой информированность агентов не является общим знанием³ и агенты принимают решения на основе иерархии своих представлений. С точки зрения теории игр и рефлексивных моделей принятия решений целесообразно разделять стратегическую и информационную рефлексию.

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

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

Таким образом, информационная рефлексия обычно связана с недостаточной взаимной информированностью, и ее результат используется при принятии решений (в том числе – при стратегической рефлексии). Стратегическая рефлексия имеет место даже в случае полной информированности, предвзято принимая решение о выбранном действии. Другими словами, информационная и стратегическая рефлексии могут изучаться независимо, однако в условиях неполной и недостаточной информированности обе они имеют место.

³ Если в рассматриваемой модели информированность является общим знанием, то все результаты исследования рефлексивных игр переходят в соответствующие классические результаты теории игр – см. ниже.

Стратегическая рефлексия рассматривается во второй главе настоящей работы. Оказывается, что если предположить, что агент, моделируя поведение оппонентов, приписывает им и себе определенные ранги рефлексии, то исходная игра превращается в новую игру, в которой стратегией агента является выбор ранга рефлексии.

Если рассмотреть процесс рефлексии в новой игре, то получим новую игру и т.д. При этом, даже если в исходной игре множество возможных действий было конечно, то в новой игре множество возможных действий – число различных рангов рефлексии – бесконечно. Следовательно, основной задачей, решаемой при исследовании стратегической рефлексии, является определение максимального целесообразного ранга рефлексии. Ответ на этот вопрос получен во второй главе для биматричных игр (раздел 2.2) и моделей, учитывающих ограниченность возможностей человека по переработке информации (раздел 2.3).

Приведем пример стратегической рефлексии – «Пенальти» (см. также примеры «Игра в прятки» и «Снос на мизере» в разделе 2.2). Агентами являются игрок, бьющий по воротам, и вратарь. Предположим для простоты, что у игрока есть два действия – «бить в левый угол ворот» и «бить в правый угол ворот». У вратаря также есть два действия – «ловить мяч в левом углу» и «ловить мяч в правом углу». Если вратарь угадывает, в какой угол бьет игрок, то он ловит мяч.

Промоделируем рассуждения агентов. Пусть вратарю известно, что данный игрок обычно бьет в правый угол. Следовательно, ему нужно ловить мяч в правом углу. Но, если вратарь знает, что игроку известно, что вратарь знает, как обычно поступает игрок, то вратарю следует моделировать рассуждения игрока. Он может думать так: «Игроку известно, что я знаю его обычную тактику. Поэтому он ожидает, что я буду ловить мяч в правом углу и может ударить в левый угол. В этом случае мне надо ловить мяч в левом углу». Если игрок обладает достаточной глубиной рефлексии, то он может догадаться о рассуждениях вратаря и попытаться его перехитрить, ударив в правый угол. Эту же цепочку рассуждений может провести и вратарь и на этом основании ловить мяч в правом углу.

И игрок, и вратарь, могут увеличивать глубину рефлексии до бесконечности, проводя рассуждения друг за друга, и ни один из них не имеет рациональных оснований остановиться на некотором конечном шаге. Следовательно, в рамках моделирования взаимных

рассуждений нельзя априори определить исход рассматриваемой игры. Сама игра, в которой у каждого из агентов есть по два возможных действия, может быть заменена на другую игру, в которой агенты выбирают ранги рефлексии, приписываемые оппоненту. Но и в этой игре нет разумного решения, так как каждый агент может моделировать поведение оппонента, рассматривая «дважды рефлексивную» игру, и т.д. до бесконечности.

Единственно, чем можно помочь в рассматриваемой ситуации агентам, так это ограничить глубину их рефлексии, подметив, что начиная со второго ранга рефлексии (в силу конечности исходного множества возможных действий) ситуация начинает повторяться – находясь как на нулевом, так и на втором (и, вообще, на любом четном) уровне рефлексии, игрок будет бить в правый угол. Следовательно, вратарю остается угадать четность уровня рефлексии игрока.

Максимальный ранг рефлексии, который следует иметь агенту для того, чтобы охватить все многообразие исходов игры (упуская из виду некоторые стратегии оппонента, агент рискует уменьшить свой выигрыш), назовем *максимальным целесообразным рангом рефлексии*. Оказывается, что во многих случаях этот ранг конечен – соответствующие формальные результаты приводятся в разделах 2.2 и 3.6). В примере «Пенальти» максимальный целесообразный ранг рефлексии агентов равен двум.

В случае отсутствия у вратаря информации о том, куда обычно бьет нападающий, действия последнего симметричны (левый и правый углы «равноценны»). Однако остаются возможности искусственно внести асимметрию, чтобы попытаться ею воспользоваться в своих целях. Например, вратарь может сдвинуться в сторону одного из углов, как бы приглашая нападающего ударить в другой (и бросается именно в тот, «дальний» угол). Более сложная стратегия состоит в следующем. Игрок команды вратаря подходит к нему и показывает, куда собирается бить нападающий, причем делает это так, что нападающий это видит (после чего в момент удара вратарь ловит мяч не в том углу, на который демонстративно показал ему товарищ по команде, а в противоположном). Заметим, что оба описанных приема взяты «из жизни» и оказались успешными. Первый имел место в международном матче сборной СССР, второй – в финале Кубка СССР по футболу в серии послематчевых пенальти.

Третья глава посвящена исследованию формальных моделей информационной рефлексии. Так как ключевым фактором в рефлексивных играх является информированность агентов – иерархия представлений, то для ее формального описания вводится понятие *информационной структуры* – дерева (в общем случае – бесконечного), вершинам которого соответствует информация (представления) агентов о существенных параметрах, представлениях других агентов и т.д. (см. пример иерархии представлений выше).

Понятие структуры информированности (информационной структуры) позволяет дать формальное определение некоторых интуитивно ясных понятий, таких как: адекватная информированность одного агента о другом, взаимная информированность, одинаковая информированность и др.

Одним из ключевых понятий, применяемых в данной работе для анализа рефлексивных игр, является понятие *фантомного агента*. Обсудим его на качественном уровне (отложив строгое математическое определение до раздела 3.2).

Пусть в некоторой ситуации взаимодействуют два агента – А и Б. Вполне естественно, что в сознании каждого из них имеется некий образ другого: у А имеется образ Б (назовем его АБ), а у Б – образ А (назовем его БА). Эти образы могут совпадать с реальностью, а могут отличаться от нее. Иными словами, агент, например, А может иметь адекватное представление о Б (этот факт можно записать в виде тождества $АБ = Б$), а может и не иметь.

Тут сразу возникает вопрос – а может ли в принципе выполняться тождество $АБ = Б$, ведь Б – это реальный агент, а АБ – лишь его образ? Не вдаваясь в обсуждение этого философского, по сути, вопроса, отметим следующие два обстоятельства. Во-первых, речь идет не о всецелом понимании личности во всей ее полноте, а о ее моделировании в данной конкретной ситуации. На обыденном, житейском уровне человеческого общения мы постоянно сталкиваемся с ситуациями как адекватного, так и неадекватного восприятия одним человеком другого.

Во вторых, в рамках формального (теоретико-игрового) моделирования человеческого поведения агент – участник ситуации – описывается относительно небольшим набором характеристик. И эти характеристики могут быть полностью известны другому агенту в той же мере, в какой они известны исследователю.

Рассмотрим подробнее случай, когда между Б и АБ имеется различие (это различие может проистекать, говоря формально, из неполноты информации А о Б, либо из доверия к ложной информации). Тогда А, принимая решение о каких-либо своих действиях, имеет в виду не Б, а тот его образ, который у него имеется, то есть АБ. Можно сказать, что субъективно А взаимодействует с АБ. Поэтому АБ можно назвать фантомным агентом. Его нет в реальности, но он присутствует в сознании *реального агента* А и, соответственно, влияет на его действия, то есть на реальность.

Приведем простейший пример. Пусть А считает, что они с Б друзья, а Б, зная об этом, является врагом А (эту ситуацию можно описать словом «предательство»). Тогда, очевидно, в ситуации имеется фантомный агент АБ, которого можно описать так: «Б, являющийся другом А»; в реальности такой субъект отсутствует. Отметим, что при этом Б адекватно информирован об А, то есть $BA = A$.

Таким образом, помимо реальных агентов, фактически участвующих в игре, предлагается рассматривать фантомных агентов, то есть агентов, которые существуют в сознании реальных и других фантомных агентов. Реальные и фантомные агенты в рамках своей рефлексии наделяют фантомных агентов определенной информированностью, которая отражается в информационной структуре.

Участвующих в игре реальных и фантомных агентов может быть бесконечно много, что означает потенциальную бесконечность осуществления актов рефлексивного отражения (бесконечную глубину дерева структуры информированности). Действительно, даже в простейшей ситуации возможно бесконечное развертывание рассуждений вида «я знаю...», «я знаю, что ты знаешь...», «я знаю, что ты знаешь, что я знаю...», «я знаю, что ты знаешь, что я знаю, что ты знаешь...» и т. д. Однако на практике такая «дурная бесконечность» не имеет места, поскольку начиная с некоторого момента представления «стабилизируются», и увеличение ранга рефлексии не дает ничего нового. Таким образом, в реальных ситуациях структура информированности имеет конечную *сложность*: у соответствующего дерева имеется конечное число попарно различных поддеревьев.

ев. Иными словами, в игре участвует конечное число реальных и фантомных агентов⁴.

Введение понятия фантомных агентов позволяет определить рефлексивную игру как игру реальных и фантомных агентов, а также определить *информационное равновесие* как обобщение равновесия Нэша на случай рефлексивной игры, в рамках которого предполагается, что каждый агент (реальный и фантомный) при вычислении своего *субъективного равновесия* (равновесия в той игре, в которую он со своей точки зрения играет) использует имеющуюся у него иерархию представлений об объективной и рефлексивной реальности [89].

Удобным инструментом исследования информационного равновесия является *граф рефлексивной игры*, в котором вершины соответствуют реальным и фантомным агентам, и в каждую вершину-агента входят дуги (их число на единицу меньше числа реальных агентов), идущие из вершин-агентов, от действий которых в субъективном равновесии зависит выигрыш данного агента. Граф рефлексивной игры может быть построен и без конкретизации целевых функций агентов. При этом он отражает если не количественное соотношение интересов, то качественное соотношение информированности рефлексизирующих агентов, и является удобным и выразительным средством описания эффектов рефлексии (см. раздел 3.4).

Для описанного выше примера двух агентов граф рефлексивной игры имеет вид: $B \leftarrow A \leftrightarrow AB$ – реальный агент Б (предатель) адекватно информирован об агенте А, который взаимодействует с фантомным агентом АБ (Б, являющимся другом А).

Приведем еще один пример графа, который отражает рефлексивное взаимодействие (хотя и не является формально графом рефлексивной игры в смысле введенного выше определения). На обложку настоящей книги вынесена картина Э. Берн-Джонса «Смертоносная голова», написанная в 1886-1887 гг. по мотивам мифа о Персее и Андромеде. В ситуации участвуют три реальных агента: Персей (обозначим его буквой П), Андромеда (А) и горгона Медуза (М). Кроме того, имеются следующие «фантомные» агенты:

⁴ В предельном случае – когда присутствует общее знание – фантомный агент первого уровня совпадает со своим реальным прообразом и дерево имеет единичную глубину (точнее, все остальные поддеревья повторяют деревья более высокого уровня).

отражение Персея (ОП), отражение Андромеды (ОА) и отражение Медузы (ОМ). Граф приведен на рисунке 1.

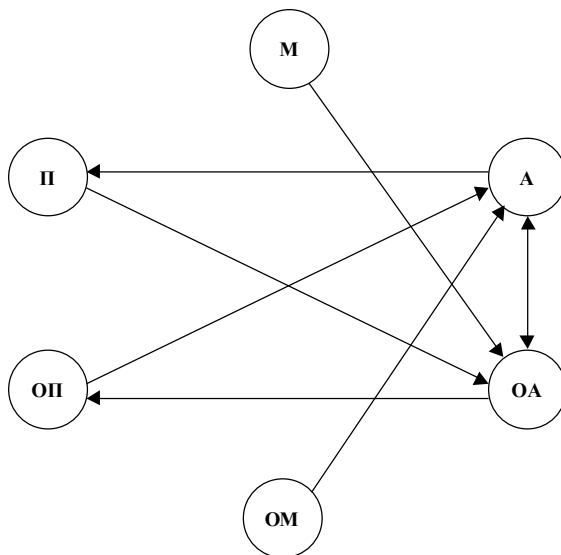


Рис. 1. Граф картины Э. Берн-Джонса «Смертоносная голова» (см. ниже)



Информированность реальных агентов в рассматриваемом примере следующая: Персей видит Андромеду; Андромеда не видит Персея, но видит его отражение, свое отражение и отражение горгоны Медузы; отражение Персея видит отражение Андромеды; отражение Андромеды видит всех реальных агентов. К счастью, самому горгону Медузу никто из реальных агентов не видит.

Введение информационной структуры, информационного равновесия и графа рефлексивной игры, во-первых, позволяет с единых методологических позиций и с помощью единого математического аппарата описывать и анализировать разнообразные ситуации коллективного принятия решений агентами, обладающими различной информированностью, исследовать влияние рангов рефлексии на выигрыши агентов, изучать условия существования и реализуемости информационных равновесий и т.д. Многочисленные примеры прикладных моделей приведены ниже.

Во-вторых, предложенная модель рефлексивной игры дает возможность изучать влияние рангов рефлексии (глубины информационной структуры) на выигрыши агентов. Полученные в разделах 2.2, 3.5 и 3.6 настоящей работы результаты свидетельствуют, что при минимальных предположениях можно показать ограниченность максимального целесообразного ранга рефлексии. Другими словами, во многих случаях неограниченное увеличение ранга рефлексии нецелесообразно с точки зрения выигрышей агентов.

В-третьих, наличие модели рефлексивной игры позволяет определить условия существования и свойства информационного равновесия, а также конструктивно и корректно сформулировать **задачу рефлексивного управления**, заключающуюся в поиске управляющим органом такой информационной структуры, что реализующееся в ней информационное равновесие наиболее выгодно с его точки зрения. Задача рефлексивного управления ставится и решается для ряда случаев в разделе 3.7. Теоретические результаты ее решения используются в ряде приводимых в четвертой главе прикладных моделей – скрытое управление, информационное управление через СМИ и др.

И, наконец, в-четвертых, язык рефлексивных игр (информационные структуры, графы рефлексивной игры и др.) является удобным для описания эффектов рефлексии как в психологии (что иллюстрируется на примере шахматной игры, транзакционного анализа,

моделей этического выбора и др.), так и в художественных произведениях – см. четвертую главу настоящей работы.

Завершив качественный обзор содержания работы, отметим, что можно предложить несколько подходов к ознакомлению с материалом настоящей книги. Первый – линейный, заключающийся в последовательном прочтении всех четырех глав. Второй рассчитан на читателя, интересующегося в большей степени формальными моделями, и заключается в прочтении второй и третьей глав и беглом ознакомлении с примерами в четвертой главе. Третий ориентирован на читателя, не желающего вникать в математические тонкости, и заключается в прочтении введения, четвертой главы и заключения.

ЛИТЕРАТУРА

- 1 Автономов В.С. Модель человека в экономической науке. СПб.: Экономическая школа, 1998. – 230 с.
- 2 Адельсон-Вельский Г.М., Арлазаров В.Л., Донской М.В. Программирование игр. М.: Наука, 1978. – 255 с.
- 3 Аткинсон Р. Человеческая память и процесс обучения. М.: Прогресс, 1980. – 528 с.
- 4 Берн Э. Игры, в которые играют люди. Люди, которые играют в игры. М.: Прогресс, 1988. – 400 с.
- 5 Бириштейн Б.И., Боршевич В.И. Стратегемы рефлексивного управления в западной и восточных культурах // Рефлексивные процессы и управление. 2002. Т. 2. № 1. С. 27 – 44.
- 6 Бириштейн Б.И., Боршевич В.И. Теория рефлексивности Дж. Сороса: опыт критического анализа // Рефлексивные процессы и управление. 2001. № 1. С. 88 – 101.
- 7 Бурков В.Н., Еналеев А.К., Новиков Д.А. Механизмы стимулирования в вероятностных моделях социально-экономических систем // Автоматика и Телемеханика. 1993. № 11. С. 3 – 30.
- 8 Бурков В.Н., Еналеев А.К., Новиков Д.А. Механизмы функционирования социально-экономических систем с сообщением информации // Автоматика и Телемеханика. 1996. № 3. С. 3 – 25.
- 9 Бурков В.Н., Кондратьев В.В. Механизмы функционирования организационных систем. М.: Наука, 1981. – 384 с.
- 10 Бурков В.Н., Новиков Д.А. Теория активных систем: состояние и перспективы. М.: Синтез, 1999. – 128 с.
- 11 Бэндлер Р., Гриндер Д. Структура магии. СПб.: Издательство «Белый кролик», 1996. – 496 с.
- 12 Бэндлер Р., Гриндер Д. Из лягушек – в принцы (нейро-лингвистическое программирование). Екатеринбург, 1998. – 206 с.
- 13 Варшавский В.И., Поспелов Д.А. Оркестр играет без дирижера. М.: Наука, 1989. – 208 с.
- 14 Васин А.А., Гурвич В.А. Коалиционные ситуации равновесия в метаиграх / Вестник МГУ. Вычислительная математика и кибернетика. 1980. № 3. С. 38 – 44.
- 15 Вилкас Э.Й. Оптимальность в играх и решениях. М.: Наука. 1990. – 256 с.
- 16 Воеводин А.И. Стратегемы – стратегии войны, манипуляции, обмана. М.: Белые альвы, 2002. – 256 с.

- 17 Волгин Л.Н. Принцип согласованного оптимума. М.: Советское радио, 1977. – 144 с.
- 18 Гермейер Ю.Б. Игры с непротивоположными интересами. М.: Наука, 1976. – 327 с.
- 19 Горелик В.А., Кононенко А.Ф. Теоретико-игровые модели принятия решений в эколого-экономических системах. М.: Радио и связь, 1982. – 144 с.
- 20 Грачев Г., Мельник И. Манипулирование личностью: организация, способы и технологии информационно-психологического воздействия. М.: Институт философии РАН, 1999. – 235 с.
- 21 Губко М.В., Новиков Д.А. Теория игр в управлении организационными системами. М.: Синтег, 2002. – 148 с.
- 22 Данилов В.И. Лекции по теории игр. М.: Российская экономическая школа, 2002. – 140 с.
- 23 Доценко Е.Л. Психология манипуляции: феномены, механизмы и защита. М.: ЧеРо, 2000. – 344 с.
- 24 Ерешко Ф.И., Лохныгина Ю.В. Рефлексивные стратегии в системах управления / Труды Юбилейной международной научно-практической конференции «Теория активных систем». Общая редакция – В.Н. Бурков, Д.А. Новиков. М.: Синтег, 1999. С. 211 –213.
- 25 Ерешко Ф.И., Лохныгина Ю.В. Исследование моделей рефлексивных стратегий в управляемых системах. М.: ВЦ РАН, 2001. – 37 с.
- 26 Ерешко Ф.И. Моделирование рефлексивных стратегий в управляемых системах. М.: ВЦ РАН, 2001. – 37 с.
- 27 Зимбардо Ф., Ляйппе М. Социальное влияние. СПб.: Питер, 2000. – 448 с.
- 28 Зинченко В.П. Рефлексивные процессы в интернет-взаимодействиях (на примере шахматных игр) // Рефлексивные процессы и управление. 2002. Т. 2. № 1. С. 90 – 95.
- 29 Информационное общество: Информационные войны. Информационное управление. Информационная безопасность / Под ред. М.А. Вуса. СПб.: Издательство С.-Петербургского университета, 1999. – 212 с.
- 30 Кабаченко Т.С. Методы психологического воздействия. М.: Педагогическое общество России, 2000. – 544 с.
- 31 Карнеги Д. Как завоевывать друзей и оказывать влияние на людей. М.: Прогресс, 1989.
- 32 Клыкков Ю.И. Ситуационное управление большими системами. М.: Энергия, 1974. – 136 с.
- 33 Козелецкий Ю. Психологическая теория решений. М.: Прогресс, 1979. – 504 с.

- 34 Кононенко А.Ф., Халезов А.Д., Чумаков В.В. Принятие решений в условиях неопределенности. М.: ВЦ АН СССР, 1991. – 197 с.
- 35 Крогиус Н.В. Личность в конфликте. Саратов: СГУ, 1976. – 144 с.
- 36 Крогиус Н.В. О психологии шахматного творчества М.: Физкультура и спорт, 1969. – 96 с.
- 37 Крогиус Н.В. Психология шахматного творчества. М.: Физкультура и спорт, 1981. – 183 с.
- 38 Крылов В.Ю. Методологические и теоретические проблемы математической психологии. М.: Янус-К, 2000.
- 39 Кукушкин Н.С., Морозов В.В. Теория неантагонистических игр. М.: МГУ, 1984. – 104 с.
- 40 Кукушкин Н.С. Роль взаимной информированности сторон в играх двух лиц с противоположными интересами // ЖВМ и МФ. 1972. Т. 12. № 4. С. 1029 – 1034.
- 41 Кульба В.В., Малюгин В.Д., Шубин А.Н., Вус М.А. Введение в информационное управление. С.Пб.: Изд-во С.-Петербургского Университета, 1999. – 116 с.
- 42 Лефевр В.А. Исходные идеи логики рефлексивных игр /Материалы конференции «Проблемы исследования систем и структур». М.: Издание АН СССР, 1965.
- 43 Лефевр В.А. Конфликтующие структуры. М.: Советское радио, 1973. – 158 с.
- 44 Лефевр В.А. Комический субъект. М.: Институт психологии РАН, 1997.
- 45 Лефевр В.А. Логика рефлексивных игр и рефлексивное управление / Принятие решений человеком. Тбилиси: Мецниереба, 1967.
- 46 Лефевр В.А. Формула человека. Контурсы фундаментальной психологии. М.: Прогресс, 1991. – 108 с.
- 47 Лефевр В.А. Элементы логики рефлексивных игр / Проблемы инженерной психологии. Вып. 4. Ленинград, 1966. С. 273 – 299.
- 48 Льюс Р., Райфа Х. Игры и решения. М.: Изд-во Иностран. Лит., 1961 – 642 с.
- 49 Лэйнг Р. Я и другие. М.: Эксмо-пресс, 2002. – 304 с.
- 50 Майерс Д. Социальная психология. СПб.: Питер, 1998. – 688 с.
- 51 Малявин В.В. (перевод с кит.) Тридцать шесть стратагем. М.: Белые альфы, 2000.
- 52 Миркин Б.Г. Проблема группового выбора. М.: Наука, 1974. – 256 с.
- 53 Нейман Дж. фон, Моргенштерн О. Теория игр и экономическое поведение. М.: Наука, 1970. – 708 с.
- 54 Нижегородцев Р.М. Теоретические основы информационной экономики. Владикавказ: Проект-Пресс, 1998. – 248 с.

- 55 Новиков А.М. Методология образования. М.: Эгвес, 2002. – 320 с.
- 56 Новиков Д.А. Механизмы стимулирования в динамических и многоэлементных социально-экономических системах // Автоматика и Телемеханика. 1997. № 6. С. 3 – 26.
- 57 Новиков Д.А. Механизмы функционирования многоуровневых организационных систем. М.: Фонд «Проблемы управления», 1999. – 150 с.
- 58 Новиков Д.А., Петраков С.Н. Курс теории активных систем. М.: Синтег, 1999. – 108 с.
- 59 Новиков Д.А. Стимулирование в социально-экономических системах (базовые математические модели). М.: ИПУ РАН, 1998. – 216 с.
- 60 Новиков Д.А., Цветков А.В. Механизмы функционирования многоэлементных организационных систем. М.: ИПУ РАН, 2001. – 188 с.
- 61 Новиков Д.А., Цветков А.В. Механизмы функционирования организационных систем с распределенным контролем. М.: ИПУ РАН, 2001. – 118 с.
- 62 Новиков Д.А., Чхартишвили А.Г. Активный прогноз. М.: ИПУ РАН, 2002. – 101 с.
- 63 Опойцев В.И. Равновесие и устойчивость в моделях коллективного поведения. М.: Наука, 1977. – 248 с.
- 64 Орловский С.А. Проблемы принятия решений при нечеткой исходной информации. М.: Наука, 1981. – 206 с.
- 65 Оуэн Г. Теория игр. М.: Мир, 1971. – 230 с.
- 66 Охрименко В.В. Простая модель экономической динамики со спекуляциями. М.: ВЦ РАН, 2002. – 31 с.
- 67 Петровский В.А. Опыт событийной транскрипции в рефлексии // Рефлексивные процессы и управление. 2001. № 1. С. 61 – 70.
- 68 Петросян Л.А., Зенкевич Н.А., Семина Е.А. Теория игр. М.: Высшая школа, 1998. – 304 с.
- 69 Пиз А. Язык телодвижений. Н. Новгород: Ай кью, 1992.
- 70 Пиндайк Р., Рубинфельд Д. Микроэкономика. М.: Дело, 2001 – 808 с.
- 71 Поддъяков А.Н. Исследовательское поведение: стратегии познания, помощь, противодействие, конфликт. Фак-т психологии МГУ им. М.В. Ломоносова, 2002. – 189 с.
- 72 Поспелов Д.А. Игры рефлексивные / Энциклопедия кибернетики. Т. 1. Киев: Гл. редакция УСЭ, 1974. С. 343.
- 73 Поспелов Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов. М.: Радио и связь, 1989.
- 74 Поспелов Д.А. Ситуационное управление: теория и практика. М.: Наука, 1986. – 288 с.
- 75 Почепцов Г.Г. Информационно-психологическая война. М.: Синтег, 2000. – 180 с.

- 76 Психологический словарь / Под ред. В.П. Зинченко. М.: Педагогика-пресс, 1996. – 400 с.
- 77 Саймон Г. Науки об искусственном. М.: Мир, 1972. – 147 с.
- 78 Советский энциклопедический словарь М.: Советская энциклопедия, 1988.
- 79 Сорос Д. Алхимия финансов. М.: ИНФРА-М, 1999. – 416 с.
- 80 Эндидж Ч., Фрайбургер В., Ротцолл К. Реклама: теория и практика. М.: Прогресс, 1989.
- 81 Таран Т. Логические модели рефлексивного выбора // Автоматика и Телемеханика. 2001. № 10. С. 103 – 117.
- 82 Таран Т.А. Рефлексивные модели в системах поддержки принятия решений / Труды 2-й Международной конференции «Когнитивный анализ и управление развитием ситуаций». М.: ИПУ РАН, 2002. Том 2. С. 117 – 135.
- 83 Томас Т.Л. Рефлексивное управление в России: теория и военные приложения // Рефлексивные процессы и управление. 2002. Т. 2. № 1. С. 71 – 89.
- 84 Философский энциклопедический словарь. М.: Советская энциклопедия, 1983.
- 85 Харрис Р. Психология массовых коммуникаций. СПб.: Прайм-Еврознак, 2002. – 448 с.
- 86 Харшаньи Д., Зельтен Р. Общая теория выбора равновесия в играх. СПб.: Экономическая школа, 2001. – 405 с.
- 87 Хэйзинга Й. Homo ludens. В тени завтрашнего дня. М.: Прогресс, 1992. – 464 с.
- 88 Чалдини Р. Психология влияния. СПб.: Питер, 2001. – 288 с.
- 89 Чхартишвили А.Г. Информационное равновесие / Управление большими системами. Сборник трудов молодых ученых. Общая редакция – Д.А. Новиков. Выпуск 3. М.: ИПУ РАН, 2003. С. 100 – 119.
- 90 Шейнов В.П. Психология обмана и мошенничества. М.: ООО «Издательство АСТ», 2002. – 512 с.
- 91 Шейнов В.П. Скрытое управление человеком (психология манипулирования). М.: ООО «Издательство АСТ», 2002. – 848 с.
- 92 Шеманов А.Ю. Самоидентификация на пороге «осевых времен» (к интерпретации модели рефлексии В. Лефевра) / От философии жизни к философии культуры. СПб., 2001. С. 137 – 158.
- 93 Шибутани Т. Социальная психология. Ростов-на-Дону: Феникс, 1998. – 544 с.
- 94 Щедровицкий Г.П. Принципы и общая схема методологической организации системно-структурных исследований и разработок / Системные исследования. М., 1981. С. 193 – 227.

- 95** Ambroszkiewicz S. On the concepts of rationalizability in games // *Annals of Operations Research*. 2000. № 97. P. 55 – 68.
- 96** Aumann R.J. Agreeing to disagree // *The Annals of Statistics*. 1976. Vol. 4. № 6. P. 1236 – 1239.
- 97** Aumann R.J., Brandenburger A. Epistemic conditions for Nash equilibrium // *Econometrica*. 1995. Vol. 63. № 5. P. 1161 – 1180.
- 98** Aumann R.J., Heifetz A. Incomplete information . *Handbook of Game Theory*. Vol III. Chapter 43. Amsterdam, Elseiver (forthcoming).
- 99** Aumann R.J. Interactive epistemology I: Knowledge // *International Journal of Game theory*. 1999. № 28. P. 263 – 300.
- 100** Bernheim D. Rationalizable strategic behavior // *Econometrica*. 1984. № 5. P. 1007 – 1028.
- 101** Brams S.J. *Theory of moves*. Cambridge: Univ. of Cambridge, 1995. – 248 p.
- 102** Brandenburger A., Dekel E. Hierarchies of beliefs and common knowledge // *Journal of Economic Theory*. 1993. Vol. 59. P. 189 – 198.
- 103** Camerer C., Weigelt K. Information mirages in experimental asset markets // *Journal of Business*. 1991. Vol. 64. P. 463 – 493.
- 104** Clark H.H., Marshall C.R. Definite reference and mutual knowledge / *Elements of Discourse Understanding* (ed. By A.K. Joshi, B.L. Webber, I.A. Sag). Cambridge: Cambridge University Press, 1981.
- 105** Fagin R., Geanakoplos J., Halpern J.Y., Vardi M.Y. The hierarchical approach to modeling knowledge and common knowledge // *International Journal of Game Theory*. 1999. Vol. 28. P. 331 – 365.
- 106** Fagin R., Halpern J., Moses Y., Vardi M.Y. *Reasoning about knowledge*. Cambridge: MIT Press, 1995.
- 107** Fagin R., Halpern J., Moses Y., Vardi M.Y. Common knowledge revisited // *Annals of Pure and Applied Logic*. 1999. Vol. 96. P. 89 – 105.
- 108** Fagin R., Halpern J., Vardi M.Y. A model-theoretic analysis of knowledge // *Journal of Assoc. Comput. Mach.* 1991. Vol. 38. № 2. P. 382 – 428.
- 109** Fudenberg D., Tirole J. *Game theory*. Cambridge: MIT Press, 1995. – 579 p.
- 110** Gamov G., Stern M. *Puzzle Math*. N.Y.: Viking Press, 1958.
- 111** Geanakoplos J. *Common Knowledge / Handbook of Game Theory*. Vol. 2. Amsterdam: Elseiver, 1994. P. 1438 – 1496.
- 112** Gray J. *Notes on database operating system / Operating Systems: An Advanced Course*. Lecture Notes in Computer Science. Vol. 66. Berlin: Springer, 1978.
- 113** Halpern J., Moses Y.O. Knowledge and common knowledge in a distributed environment // *Journal of Assoc. Comput. Mach.* 1990. Vol. 37. № 3. P. 549 – 587.

- 114** Harsanyi J. Games with incomplete information played by "Bayesian" players // Management Science. Part I: 1967. Vol. 14. № 3. P. 159 – 182. Part II: 1968. Vol. 14. № 5. P. 320 – 334. Part III: 1968. Vol. 14. № 7. P. 486 – 502.
- 115** Heifetz A. Iterative and fixed point belief // Journal of Philosophical Logic. 1999. Vol. 28. P. 61 – 79.
- 116** Hintikka J. Knowledge and belief. Ithaca: Cornell University Press, 1962.
- 117** Howard N. Theory of meta-games / General systems. 1966. № 11. P. 187 – 200.
- 118** Howard N. "General" metagames: an extension of the metagame concept / Game theory as a theory of conflict resolution. Dordrecht: Reidel, 1974. P. 258 – 280.
- 119** Kripke S. A completeness theorem in modal logic // Journal of Symbolic Logic. 1959. № 24. P. 1 – 14.
- 120** Lefebvre V.A. Algebra of Conscience. Dordrecht, Holland: Reidel, 1982.
- 121** Lefebvre V.A. Psychological theory of bipolarity and reflexivity. Levinston: The Edwin Mellen Press, 1992.
- 122** Lefebvre V.A. Sketch of reflexive game theory / Proc. Of Workshop on Multi-Reflexive Models of Agent Behavior. Los Alamos, New Mexico, USA, 1998. P. 1 – 44.
- 123** Lewis D. Convention: a philosophical study. Cambridge: Harvard University Press, 1969.
- 124** Luft J. On Human interaction. Palo Alto, CA: National Press, 1969. – 177 p.
- 125** Luft J., Ingham H. The Johari window: a graphic model for interpersonal relations. University of California: Western Training Lab, 1955.
- 126** Mas-Collel A., Whinston M.D., Green J.R. Microeconomic theory. N.Y.: Oxford Univ. Press, 1995. – 981 p.
- 127** McCarthy J., Sato M., Hayashi T., Igarishi S. On the model theory of knowledge. Technical Report STAN-CS-78-657. Stanford University, 1979.
- 128** Mertens J.F., Zamir S. Formulation of Bayesian analysis for games with incomplete information // International Journal of Game Theory. 1985. № 14. P. 1 – 29.
- 129** Miller G. The magical number seven plus or minus two: some limits on capacity for information processing // Psychological Review. 1956. Vol. 63. № 1. P. 81 – 92.
- 130** Morris S. Approximate common knowledge revisited // International Journal of Game Theory. 1999. Vol. 28. P. 385 – 408.
- 131** Morris S., Shin S.S. Approximate common knowledge and coordination: recent lessons from game theory // Journal of Logic, Language and Information. 1997. Vol. 6. P. 171 – 190.
- 132** Myerson R.B. Game theory: analysis of conflict. London: Harvard Univ. Press, 1991. – 568 p.

- 133** Nagel R. Experimental results on interactive competitive guessing // American Economic Review. 1995. Vol. 85. № 6. P. 1313 – 1326.
- 134** Nash J.F. Non-cooperative games / Ann. Math. 1951. Vol. 54. P. 286 – 295.
- 135** Pearce D.G. Rationalizable strategic behavior and the problem of perfection // Econometrica. 1984. № 5. 1029 – 1050.
- 136** Rapoport A., Guyer M. A taxonomy of 2×2 games / General Systems: Yearbook of the Society for General Systems Research. 1966. № 11. P. 203 – 214.
- 137** Ross L., Greene D., House P. The “false consensus” effect: an egocentric bias in social perception and attribution // Journal of Experimental Social Psychology. 1977. Vol. 13. P. 279 – 301.
- 138** Rubinstein A. The electronic mail game: strategic behavior under “almost common knowledge” // American Economic Review. 1989. Vol 79. P. 385 – 391.
- 139** Sakovics J. Games of incomplete information without common knowledge priors // Theory and decision. 2001. № 50. P. 347 – 366.
- 140** Simon R.S. The difference of common knowledge of formulas as sets // International Journal of Game Theory. 1999. Vol. 28. P. 367 – 384.
- 141** Stahl D.O., Wilson P.W. Experimental evidence on players’ models of other players // Journal of Economic Behavior and Organization. 1994. Vol. 25. P. 309 – 327.
- 142** Vanderschraaf P. Knowledge, equilibrium and conventions // Erkenntnis. 1998. Vol. 49. P. 337 – 369.
- 143** Weber R. Behavior and learning in the “Dirty Face” game // Experimental Economics. 2001. Vol. 4. P. 229 – 242.
- 144** Wolter F. First order common knowledge logics // Studia Logica. 2000. Vol. 65. P. 249 – 271.