

УЧРЕЖДЕНИЕ РОССИЙСКОЙ АКАДЕМИИ НАУК  
ИНСТИТУТ ПРОБЛЕМ УПРАВЛЕНИЯ  
им. В.А. Трапезникова РАН

**В.О. КОРЕПАНОВ**

**МОДЕЛИ  
РЕФЛЕКСИВНОГО ГРУППОВОГО  
ПОВЕДЕНИЯ И УПРАВЛЕНИЯ**

Москва – 2011

Корепанов В.О. Модели рефлексивного группового поведения и управления. – М.: ИПУ РАН, 2011. – 127 с.

ISBN 978-5-91450-085-3

Работа содержит результаты исследования метода рефлексивных разбиений множества рациональных агентов, принимающих решения децентрализованно, на подмножества агентов, обладающих различными рангами стратегической рефлексии, который позволяет:

- с точки зрения теории принятия решений – расширить класс моделей коллективного поведения интеллектуальных (рефлексирующих) 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 за счет выбора эффективного разбиения агентов на ранги рефлексии.

Сформулированы и аналитически решены задачи рефлексивного управления для ряда прикладных моделей коллективного поведения, для каждой из которых получены оценки максимального целесообразного ранга стратегической рефлексии.

Описаны результаты имитационного моделирования группового поведения с учетом различий рангов рефлексии агентов. Исследованы зависимости оптимальных решений задач рефлексивного управления от параметров моделей.

Работа рассчитана на специалистов по теории принятия решений, имитационному моделированию и мультиагентным системам.

Рецензент: д.ф.-м.н. А.Г. Чхартишвили

ISBN 978-5-91450-085-3



© Корепанов В.О., 2011

## СОДЕРЖАНИЕ

<b>Введение .....</b>	<b>4</b>
<b>Глава 1. Проблемы моделирования рефлексивного группового поведения 8</b>	
1.1. Модели информационной и стратегической рефлексии, рефлексивные игры.....	8
1.2. Модели коллективного поведения, теория команд .....	12
1.3. Мультиагентные системы как средство моделирования коллективного поведения .....	17
1.4. Модели эвакуации .....	30
<b>Глава 2. Метод рефлексивных разбиений в задачах коллективного поведения и управления .....</b>	<b>46</b>
2.1. Модель игровой ситуации .....	47
2.2. Рефлексивные разбиения.....	50
2.3. Рефлексивная модель коллективного поведения .....	51
2.4. «Игра» однородных агентов.....	57
2.5. Постановка задачи управления.....	58
2.6. Прикладные модели рефлексивного коллективного поведения и управления .....	62
<b>Глава 3. Имитационные модели рефлексивного коллективного поведения 78</b>	
3.1. Система агентного имитационного моделирования.....	78
3.2. Модели стратегического поведения .....	86
3.3. Модели тактического поведения .....	91
3.4. Имитационное моделирование эвакуации.....	97
3.5. Имитационное моделирование диффузной бомбы.....	107
<b>Литература.....</b>	<b>119</b>

## Введение

Традиционно в теоретико-игровых моделях и/или в моделях коллективного поведения используется одно из двух предположений о взаимной информированности агентов (ЛПР, программных и/или мобильных агентов и т.д.). Либо считается, что вся существенная информация и принципы принятия агентами решений всем им известны, всем известно, что всем это известно и т. д. до бесконечности (так называемая концепция общего знания, используемая, например, при определении равновесия Нэша). Либо предполагается, что каждый агент в рамках своей информированности следует некоторой процедуре принятия индивидуальных решений и почти «не задумывается» над тем, что знают и как ведут себя остальные агенты. Первый подход является каноническим для теории игр, второй – для моделей коллективного поведения. Но между двумя этими «крайностями» существует достаточно большое разнообразие возможных ситуаций, в которых различные агенты могут обладать различными рангами стратегической (используемой при принятии решений) рефлексии.

Разработка методов учета возможного распределения агентов по рангам рефлексии (предложенного в настоящей работе *метода рефлексивных разбиений*) актуальна, так как даст возможность не только адекватно описывать многие ситуации коллективного поведения, но и ставить и решать задачи группового управления за счет целенаправленного выбора рангов рефлексии тех или иных агентов.

Под *агентом* будем понимать объект/сущность, способный действовать, обладающий автономностью, активностью, целевым поведением и системой отражения реальности, а также обладающий свойством взаимодействия с другими агентами, то есть извне агента можно влиять на его действия (в т.ч. управлять) и также действия агента могут влиять на других агентов и внешнюю среду.

Понятие *агент* в сознании сразу связывается с вопросом «чей?», то есть агент проявляет себя в связке со своим «создателем» и/или с управляющим органом, Центром. Иными словами, эта связка отражается в том, что Центр может влиять на агента и приносить ему пользу (иначе агент не стал бы поддаваться влиянию ввиду своей автономности), а агент способен приносить пользу Центру, подчинённость означает, что Центр обычно поставлен в более выгодные условия, хотя существуют ситуации, когда агент может манипулировать Центром.

Базовая модель, исследуемая в настоящей работе, включает один управляющий орган – Центр и множество управляемых агентов, которые принимают решения одновременно, в условиях неполной информированности. Такое определение опирается на определение многоэлементной организационной системы с неопределённостью из теории управления организационными системами [44]. При этом

управление может быть децентрализованным, то есть реализовываться путем априорного определения Центром правил взаимодействия агентов.

Как видно, ЛПР автоматически попадает в категорию агентов. Другие возможные примеры: самые различные роботы, животные, некоторые виды компьютерных программ и т.п.

Группа агентов – совокупность агентов, объединённых общими характеристиками, деятельностью и т.п. Коллектив агентов – группа агентов, объединённых общей целевой задачей [40]. Коллективное поведение как вид группового поведения характеризуется тем, что поведение каждого из агентов в текущий момент времени определяется его собственным состоянием, а также состояниями других членов «группы» в предшествующий момент времени (реже – в предшествующие моменты времени). То есть специфика коллективного поведения состоит, в том числе, в том, что агенты взаимозависимы.

Везде, где существует некоторая группа живых или технических агентов, которые должны совместными усилиями выполнять некоторую работу или решать некоторую общую задачу, возникает коллективное поведение.

Рефлексивное же управление подразумевает, что агент обладает некоторой автономностью, активностью и системой отражения внешней реальности – информированностью об окружающей среде, на основе которой он принимает решения о своих действиях (см. напр. [65]). В силу такого его поведения и возникает возможность управления им через управление информированностью об информированности и принципах принятия решений других агентов.

Получается, что с точки зрения непосредственного управления поведением агентов имеется децентрализованная система управления, так как агенты автономны, но централизованная в плане формирования информированности агентов. В результате в какой-то степени уменьшается нагрузка на центр при сохранении подконтрольности агентов центру.

Информированность есть результат отражения природной («объективной») реальности в сознании. В общем смысле с этим (одним из многих) фундаментальным свойством бытия человека связан термин «рефлексия». При этом между природной реальностью и ее образом в сознании (будем считать этот образ частью особой – рефлексивной реальности) существует неизбежный зазор, несовпадение.

Целенаправленное изучение рефлексии (термин «рефлексия» введён Дж. Локком) велось в различных философских системах, у Дж. Локка, Г. Лейбница, Д. Юма, Г. Гегеля и других он имел различное содержание. Систематическое описание рефлексии с точки зрения психологии началось в 60-е годы XX века (школа В.А. Лефевра).

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

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, [14, 43, 113]), второй – для моделей коллективного поведения (см., например, [6, 32, 45]).

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.

Таким образом, агенты разбиваются на группы по уровням рефлексии. Исход игры зависит от количества (доли) агентов в каждой группе. Возникает вопрос – как поведение всей совокупности агентов зависит от их распределения по рангам рефлексии, то есть от того, какая доля агентов того или иного ранга? Если долями рефлексиирующих агентов можно управлять, то каковы эти доли, оптимальные с точки зрения общей групповой цели? Предлагаемый метод можно назвать *методом рефлексивных разбиений* [24, 25, 29] множества рациональных агентов.

Логически настоящая работа раскрывает три подхода к проблеме: общая теория, математическое моделирование тех или иных практически важных случаев и имитационное моделирование. Такая логика прослеживается в структуре обзора (Глава 1) и в последующем изложении трёх глав.

Обзорная Глава 1 содержит необходимые известные сведения для создания теории и моделирования. Во-первых, существенный методологический вклад при исследовании игр с различной взаимной информированностью, сделан в работах [43, 65], краткие результаты

которых представлены в разделе 1.1. Далее в разделе 1.2 рассматриваются результаты исследований коллективного поведения. Разделы обзора 1.3, 1.4 необходимы для описания результатов создания среды имитационного моделирования движений агентов на плоскости.

В Глава 2 дана постановка и обсуждены методы решения задачи рефлексивного управления. Рассматриваются предпосылки и связи с существующими подходами. Для агента подробно описывается его информированность и принцип принятия им решения на основе структуры информированности. Дана общая постановка задачи управления, для которой пока сложно найти аналитическое решение, рассмотрены частные случаи. Эти примеры также являются приложениями теории к известным проблемам в экономике и другим задачам коллективного поведения. Для некоторых задач (олигополия Курно, задача о консенсусе, модель эвакуации) найдено аналитическое решение, для других рассмотрены качественные результаты.

Глава 3 содержит результаты имитационного моделирования процесса эвакуации людей из здания. Описана система имитационного мультиагентного моделирования движения агентов на плоскости: архитектура агента, архитектура системы моделирования и принципы работы системы. Раздел 3.4 предлагает обсуждение полученных результатов с точки зрения предлагаемой теории: нахождение объяснимого поведения при заданных условиях, нахождение необходимой структуры рефлексивного разбиения, например, для минимального времени выхода из здания в процессе эвакуации. Раздел 3.5 содержит описание другой имитационной модели – коллективного проникновения через систему обороны.

## **Глава 1. Проблемы моделирования рефлексивного группового поведения**

В первой главе настоящей работы приводится обзор существующих подходов, которые могут помочь при моделировании рефлексивного коллективного поведения. Краткие результаты работ по методологии информационного управления и рефлексивным играм представлены в разделе 1.1. Важно, что в этих работах описан ряд рефлексивных игр, построенных на классических моделях коллективного поведения из экономики и теории игр. В разделе 1.2 рассматриваются подходы, применяемые для исследования моделей коллективного поведения. Далее, для анализа возможностей создания среды имитационного мультиагентного моделирования, разделы 1.3 и 1.4 содержат обзоры, соответственно, по теории мультиагентного моделирования и по существующим моделям эвакуации людей из здания.

### ***1.1. Модели информационной и стратегической рефлексии, рефлексивные игры***

В данном разделе рассматриваются результаты исследования рефлексивных игр и известные теоретико-игровые модели рефлексии. Описываемая методология работ [43, 65] будет использоваться на протяжении всей работы.

Социально-экономические системы включают в себя людей (отдельных индивидуумов и их группы), поэтому управление ими заключается в побуждении людей к требуемому поведению. Однако человек самостоятельно принимает решения, значит для того, чтобы влиять на его поведение, необходимо иметь модель принятия им решений.

Господствующая в науке на протяжении последнего полувека модель принятия субъектом решений (гипотеза рационального поведения) заключается в следующем: субъект стремится выбрать наилучшую в рамках имеющейся у него информации альтернативу. При этом в модель принятия решений входят, как минимум, множество альтернатив, из которого производится выбор, а также предпочтения субъекта на этом множестве, которые обычно описываются функцией полезности [14]. Стоит отметить, что при этом существенной является информированность субъекта – та информация, которой он обладает на момент принятия решений о допустимых альтернативах, их предпочтительности, последствиях выбора той или иной альтернативы и т.д. то есть о параметрах ситуации, в которой он находится.

В случае, когда имеется только один субъект, дело обстоит достаточно просто – считается, что он выбирает из множества допустимых альтернатив такую альтернативу, на которой достигается максимум его функции полезности (выигрыша, предпочтения и т.д.) [11, 14].

Если субъектов несколько и выигрыш каждого зависит от выборов всех,

то ситуация усложняется – для того чтобы выбрать собственное действие, субъект должен «предсказать», какие действия выберут его оппоненты. Моделями совместного принятия решений субъектами, интересы которых не совпадают, занимается теория игр ([14, 113], см. там же ссылки на другие работы), одной из основных задач которой является предсказание решения игры – устойчивого в том или ином смысле исхода взаимодействия рациональных субъектов (игроков, агентов).

Попробуем промоделировать ход рассуждений субъекта, принимающего решения. Пусть он считает, что его оппоненты выберут определенные действия. Тогда он должен выбрать свое действие, являющееся наилучшим при сложившейся обстановке. Если он считает своих оппонентов такими же рациональными, как и он сам, то он должен предположить, что при выборе своих действий они будут ожидать соответствующего выбора от него. Но тогда он должен учитывать и то, что оппоненты знают о том, что он считает их рациональными и так далее – получаем бесконечную цепочку «вложенных» рассуждений. Как же замкнуть эту бесконечную цепочку наилучших ответов, какое решение принять в ситуации выбора? Наиболее распространенным способом такого «замыкания» является концепция так называемого равновесия Нэша. Равновесие Нэша – это такая ситуация, от которой никому из участников игры невыгодно отклоняться в одностороннем порядке. Иными словами: «Если все оппоненты выбирают именно эту ситуацию, то и я ничего не выигрываю, отклоняясь от нее» – и так для каждого игрока.

В теории игр, как правило, предполагается, что все параметры игры являются общим знанием, то есть каждому агенту известны все параметры игры, а также то, что это известно всем агентам, и так далее до бесконечности. Такое предположение соответствует объективному описанию игры и дает возможность использовать концепцию равновесия Нэша [116] 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.

В рамках теоретико-игровой модели отказ от предположения о наличии среди агентов общего знания приводит к моделям рефлексивных игр [43, 65]. При этом равновесие Нэша «превращается» в более общее информационное равновесие, в рамках которого каждый агент осуществляет информационную рефлексивность – при принятии решений использует не только свою информацию о существенных параметрах, но и свои представления о представлениях других агентов об этих параметрах, представления о представлениях о представлениях и т.д. или иначе агент опирается на свою иерархию представлений [43].

Итак, в процессе рефлексии агент пытается предсказать

информированность и действия своих оппонентов с целью устранения игровой неопределённости [43] при принятии им решения и выборе наилучшего ответа на их действия. С точки зрения рефлексивного поведения равновесие Нэша демонстрирует бесконечную глубину рефлексии, поскольку является предельной точкой отображения наилучших ответов [14]

Разделяют стратегическую и информационную рефлексии [43].

*Информационная рефлексия* – процесс и результат размышлений агента о том, каковы значения неопределённых 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, с помощью информационной рефлексии агент старается ответить на вопрос «что знает другой агент», как результат такого размышления и появляется иерархия представлений агента. А с помощью стратегической рефлексии агент отвечает на вопрос «на каких принципах основано принятие решений другим агентом».

В рамках таких определений, можно сказать, что подробнее всего пока исследовался вариант информационной рефлексии [40, 42, 43, 65]: как с её помощью управлять результатом игры рефлексивных агентов при заданной модели стратегической рефлексии. В данных работах под рефлексивным управлением (точнее – информационным управлением) обычно и понималось управление информационной рефлексией или иерархиями представлений агентов. Авторы упомянутых работ сформулировали единые теоретические основы разработки и исследования моделей информационного управления. При этом в качестве модели стратегической рефлексии использовалась стандартная модель общего знания агентов о рациональности друг друга. В этих работах были введены определения о ключевых ситуациях в рефлексивной игре:

1. *Информационное равновесие* – расширение равновесия Нэша на случай рефлексивной игры.
2. *Стабильное информационное равновесие* – информационное равновесие, при котором каждый агент наблюдает в результате разыгрывания игры то, что он ожидал пронаблюдать в результате процесса своей рефлексии.
3. *Истинное/Ложное стабильное информационное равновесие* – равновесие (набор действий агентов), которое остаётся равновесием при наличии общего знания о состоянии природы.

Помимо важных методологических результатов можно выделить следующие результаты этих работ:

1. Разработана концепция информационной структуры, описывающей информированность субъектов о существенных параметрах ситуации и о представлениях оппонентов; определены и исследованы адекватность, взаимность, одинаковость информированности агентов, а также глубина и сложность структуры информированности;
2. Для описания зависимости между информационной структурой и набором действий участников игры предложена концепция информационного равновесия, являющегося обобщением равновесия Нэша в некооперативных играх;
3. Сформулированы и доказаны достаточные условия существования информационного равновесия;
4. Доказана теорема, о не сужении множества наилучших ответов агента при увеличении ранга рефлексии.
5. Доказано несколько утверждений об истинности стабильного информационного равновесия и об ограниченности информационной структуры при стабильном информационном равновесии (см. раздел 3.2 в [65])
6. Введено и исследовано понятие стационарности рефлексивного отображения.
7. Качественно исследованы варианты соотношения множества достижимости при рефлексивном управлении и множества допустимых действий агентов.

В работе [43] также была рассмотрена стратегическая рефлексия в играх двух лиц. Рассматривались различные модели представлений игрока о принципах принятия решений оппонентом. Логически рост ранга рефлексии ничем не ограничен, поэтому рассматривались два варианта ограничений для роста ранга рефлексии: когда увеличение ранга не приводит к появлению новых равновесий и когда учитывается ограниченность человеческих возможностей при переработке информации.

Для биматричных игр получены важные результаты об ограниченности ранга стратегической рефлексии количеством возможных вариантов действий игрока из постановки игры, а также что в игре выбора рангов существует не более двух равновесий [13]. А, исходя из ограниченности когнитивных способностей человека, был сделан вывод, что стратегический ранг рефлексии не может превышать 2.

Приведем также следующие определения из [43]. Информационным управлением будем называть воздействие на структуру информированности агентов, осуществляемое с целью изменения информационного равновесия. Задача рефлексивного управления может быть на качественном уровне сформулирована следующим образом: найти такую структуру информированности агентов, чтобы информационное равновесие их рефлексивной игры было наиболее предпочтительно с точки зрения центра – субъекта, осуществляющего управление.

Настоящая работа в основном опирается на методологию книг [43, 65], но предметом исследования является класс моделей рефлексивного управления, где под управлением понимается управление стратегической рефлексией агентов. В определённом смысле исследуется некая «чистая» модель стратегической рефлексии, противоположная модели общего знания агентов о рациональности друг друга: каждый агент считает себя «умнее» всех остальных, это выражается в том, что агент приписывает всем остальным агентам ранги рефлексии не выше и не равные своему. Важно отметить, что в модели нет информационной рефлексии (все параметры игры являются общим знанием).

В такой постановке исчезает игровая неопределённость при выборе действия и, соответственно, задача поиска информационного равновесия становится тривиальной, поскольку при любой структуре стратегической рефлексии агента процесс вычисления действий оппонентов не требует поиска равновесия Нэша – предельной точки отображения наилучших ответов, а вычисляется непосредственно из приписываемой им информированности.

Задачами исследования являются: определение целесообразности применения данной модели стратегической рефлексии в моделях группового поведения, нахождение множества достижимости и обоснование максимального ранга рефлексии для ряда примеров.

Итак, мы рассмотрели известные результаты по рефлексивным играм и рефлексивному управлению. Перейдём теперь к обзору моделей группового поведения и моделей команд – автономных, целенаправленных групп агентов.

## **1.2. Модели коллективного поведения, теория команд**

Данный раздел содержит необходимые общие понятия из теории коллективного поведения, затем приводятся результаты исследования моделей поведения команд – групп агентов, способных автономно достигать общей цели.

Общая модель коллективного поведения подразумевает существование системы, состоящей из множества взаимосвязанных агентов  $N = \{1, \dots, n\}$ , каждый из которых имеет цель (или целевую функцию) зависящую в общем случае от состояния системы и распоряжается выбором своего действия  $x_i$ . Состояние системы определяется набором действий агентов  $x = \{x_1, \dots, x_n\}$ .

В исследовании моделей коллективного поведения упор делается на динамике моделей (аналог в теории игр – динамические игры) – поведение системы в течение определённого промежутка времени, в течение нескольких «ходов» агентов. Данные исследования вплотную приближаются к постановке задач агентного имитационного моделирования и имеют с ними много общего.

Обычно в данных моделях, очередной свой ход агент делает согласно

*гипотезе индикаторного поведения*: каждый агент с течением времени изменяет значение собственной переменной по направлению к текущему положению своей цели.

Это означает, во-первых, что для каждого агента  $i$  при любом допустимом состоянии системы  $x$  существует единственная точка, являющаяся его целью:  $x_i^g = f_i(x)$ . Тогда в дискретном времени тактика индикаторного поведения может быть описана итерационной процедурой:

$$\forall i, k: x_i^{k+1} = x_i^k + \gamma_i^k (f_i(x^k) - x_i^k), \gamma_i^k \in [0, 1],$$

где  $x_i^k$  – действие агента  $i$  в  $k$ -м периоде.

Гипотеза индикаторного поведения является лишь одним из возможных вариантов описания коллективного поведения [32, 39, 41, 45], но её использование уместно, так как, с одной стороны, ее свойства исследованы наиболее подробно по сравнению с другими процедурами, а с другой стороны – как показывают имитационные эксперименты, она достаточно адекватно описывает поведение людей в имитационных и деловых играх [5, 45, 63].

Гипотеза индикаторного поведения описывает возможный механизм коллективного поведения агентов. В случае группового управления, из набора механизмов коллективного поведения нужно выбрать для каждого агента такой, чтобы действия агентов в совокупности привели к достижению поставленной цели.

Возможной целью такого управления является создание/формирование *команды* в терминах [44]: под *командой* понимается группа агентов (в оригинале «группа лиц»), способных достигать цели автономно и согласованно, при минимальных управляющих воздействиях. То есть команду можно понимать как совокупность группы агентов и реализованного в ней механизма коллективного поведения, позволяющего достигать определённое множество целей автономно.

Существенными в приведенном определении команды являются два аспекта. Первый – достижение цели, то есть конечный результат совместной деятельности является для команды системообразующим фактором. Второй аспект – автономность и согласованность деятельности – означает, что каждый из членов команды демонстрирует поведение, требуемое в данных условиях (позволяющее достичь поставленной цели), то есть то поведение, которого от него ожидают другие члены команды.

В зависимости от используемого аппарата моделирования можно выделить несколько исторически сложившихся (и зачастую сильно пересекающихся) направлений исследования моделей команд (см. Рис. 1):

- «задачи о назначении», использующие, в основном, аппарат оптимизации для решения задач формирования состава команд, распределения ролей и объемов работ (см. раздел 2.1 в 40);

- *теоретико-игровые модели*, использующие аппарат теории игр [6, 14, 113] для описания и исследования процессов формирования и функционирования команд.



Рис. 1. Классификация моделей команд

- «экспериментальные исследования» команд, включающие *имитационные эксперименты и деловые игры* (см. обзор в разделе 2.6 в [40]);

- «*рефлексивные модели*», использующие аппарат теории рефлексивных игр (см. [43, 65]) для описания взаимодействия членов команды, имеющих несовпадающие взаимные представления о существенных параметрах друг друга.

В [40] выделены следующие компоненты практически любой модели команды:

**1. Состав команды** (множество агентов, входящих в команду). Для того чтобы описывать команду, нужно, как минимум, задать ее состав. Существуют исследования «команд» с переменным составом – см. [3,21].

**2. Состояния агентов** (включая выполняемые ими *функции* (роли) и *объемы работ*) и множества допустимых состояний. Иногда в описание модели включаются уравнения, отражающие взаимосвязь между состояниями агентов и/или законы изменения состояний во времени.

**3. В зависимости от того, выбирают ли агенты свои состояния самостоятельно или считается, что они определяются извне** (в результате решения оптимизационных задач или устанавливаются некоторым руководящим органом), выделяют соответственно модели, учитывающие *активность* агентов, и модели команд с *пассивными* агентами. Активность поведения агентов обычно описывается в рамках теоретико-игровых моделей.

4. *Результат деятельности* команды, который зависит от состояний агентов (их индивидуальных действий).

5. *Целевые функции* агентов могут зависеть от их индивидуальных действий (состояний) и/или результата совместной деятельности. Причем целевые функции различных агентов могут как совпадать (тогда имеется одна целевая функция, отражающая единую *эффективность команды*), так и различаться.

6. *Информированность агентов* (информация, которой они обладают о существенных внешних и внутренних параметрах) может быть как одинаковой, так и различной. Кроме того, она может быть *тривиальной* (когда имеется *общее знание* – факт, о котором всем известно, всем известно, что всем это известно и т.д. – см. введение [43]) или *нетривиальной* (тогда необходимо учитывать эффекты рефлексии – представления агентов о представлениях друг друга).

В рамках проблематики настоящей работы стоит кратко описать исследованную в [40] модель формирования однородной команды (под однородностью здесь подразумевается, что вид деятельности у каждого агента команды одинаковый). Модель основывается на рассмотрении иерархий взаимных представлений агентов об их индивидуальных параметрах (так называемых *типах* агентов), определяющих эффективность индивидуальной деятельности. Предполагается, что автономность деятельности сформированной команды соответствует стабильному информационному равновесию игры агентов, в котором ожидания членов команды относительно поведения друг друга оправдываются.

Процесс формирования команды описывается динамикой взаимных представлений агентов о типах друг друга в зависимости от наблюдаемых результатов деятельности команды в целом и/или ее отдельных членов. В рамках своих представлений каждый агент может предсказать, какие действия выберут другие агенты, какие они понесут индивидуальные затраты и каковы будут суммарные затраты. Если выбор действий производится многократно, и наблюдаемая некоторым агентом реальность оказывается отличной от его представлений, то он вынужден корректировать свои представления и при очередном своем выборе использовать «новые» представления. То есть, командой в рамках «рефлексивного» описания принятия решений считается множество агентов, выборы которых согласованы с иерархией их взаимных представлений друг о друге.

Показано, что стабильность команды и слаженность ее работы может достигаться, в том числе, и при ложных представлениях членов команды друг о друге. Выход из ложного стабильного равновесия требует получения агентами дополнительной информации.

Среди общих результатов теории команд можно выделить, помимо моделей формирования и функционирования неоднородных команд, также следующие:

1. *Модели репутации и норм деятельности*, позволяющие описать и исследовать образование и функционирование команд.

2. Модель однородной команды, использующей *единый ресурс*, суммарные затраты на приобретение которого зависят от суммы действий, выбираемых членами команды. Условием устойчивого функционирования такой команды считается существование процедуры распределения ресурса, при которой возможен выбор агентами такого вектора ненулевых действий, который был бы одновременно устойчив по Нэшу (устойчив относительно индивидуальных отклонений агентов) и эффективен по Парето (выгоден для команды в целом). Качественно, основные результаты заключаются в следующем. Во-первых, показано, что, при гладких процедурах распределения затрат устойчивое функционирование команды невозможно. Во-вторых, доказано, что, если члены однородной команды таковы, что их можно упорядочить по эффективности деятельности, и это упорядочение не зависит от выбираемых ими действий (например, объемов производства), то устойчивое функционирование команды также невозможно. И, наконец, в-третьих, обосновано, что условием устойчивого функционирования команды является наличие синергетического взаимодействия ее членов.

3. Модели *адаптации команд* – процесса изменения действий (включая в общем случае функции и объемы работ), выбираемых членами команды, на основе текущей информации в изменяющихся внешних условиях функционирования команды. Выделены несколько вложенных уровней адаптации:

- изменение информированности о внешней среде;
- изменение поведения (действий, выбираемых на основе имеющейся информации);
- изменение параметров системы, позволяющее реализовывать более эффективное в изменившихся условиях поведение;
- целенаправленное изменение внешней среды (активная адаптация).

Показано, что специфика команд заключается, в частности, в том, что каждый агент в качестве информации для корректировки своих представлений о неопределенном параметре может использовать не только результаты наблюдения за внешней средой, но и результаты наблюдения за действиями и результатами деятельности других агентов, пытаясь «объяснить», почему они выбрали именно эти действия. Другими словами, если результат совместной деятельности зависит от действий всех агентов, то у каждого агента имеются, как максимум, четыре «источника информации» о внешней среде:

1) априорная частная информация, которой обладает каждый из агентов;

2) действия других агентов: наблюдая их и предполагая, что оппоненты действуют рационально, агент может осуществлять рефлекссию – оценивать ту информацию о внешней среде, на основании которой рационален выбор оппонентами именно этих действий;

3) выигрыши агентов – на основании этой информации агенты могут сделать вывод о тех состояниях внешней среды, при которых наблюдаемый результат приводит к наблюдаемым выигрышам;

4) множество состояний внешней среды, при которых наблюдаемый вектор действий агентов приводит именно к данному наблюдаемому значению результата.

4. Модели обучения членов команды в процессе работы. В рамках предположения о том, что объем уже выполненных агентом работ условно отражает накопленный им «опыт», сформулирована и решена задача об оптимальном обучении – выбора объемов работ, выполняемых агентами в те или иные промежутки времени.

Исследования в работе [40] интересны в связи с проблематикой настоящего исследования, так как ключевые результаты этой работы основываются на использовании динамики информационной рефлексии, как способа автономного формирования и функционирования команд, то есть по определению являются динамическими моделями рефлексивного группового управления с минимальными управляющими воздействиями.

На этом разделе заканчивается часть обзора, посвящённая теоретико-игровым исследованиям рефлексии и группового поведения. Перейдём к обзору теории мультиагентного моделирования. Данная теория развилась, в том числе, под влиянием теории игр и теории коллективного поведения (см. [66]).

### **1.3. Мультиагентные системы как средство моделирования коллективного поведения**

Так как нами моделируется рефлексивное управление группами взаимосвязанных агентов с использованием теоретико-игрового подхода, то выбрана мультиагентная методология моделирования, в которой постулируется сущность агента и соответствующие этому механизмы формирования групп агентов и манипулирования каждым агентом в отдельности.

Общая идея мультиагентного моделирования – исследование децентрализованных систем, динамика функционирования которых определяется не глобальными правилами и законами, а наоборот, эти глобальные правила и законы являются результатом индивидуальной активности агентов группы. Первая из целей мультиагентных моделей – получить представление об этих глобальных правилах, общем поведении системы, исходя из предположений об индивидуальном, частном поведении её отдельных агентов и взаимодействии агентов в системе – коррелирует с поиском равновесий в теории игр. Нахождение/построение группового управления – следующая по сложности задача мультиагентного моделирования.

Рассмотрим существующие подходы к построению мультиагентных

систем и использованию их как средства моделирования коллективного поведения.

### 1.3.1. Делиберативные агентные системы

Термин «*делиберативный агент*» (deliberate agent) (ДА) был введен в работе [88]. Делиберативную архитектуру принято определять как архитектуру агентов, содержащих точную символическую модель мира и принимающих решения на основе логического вывода. Теоретическим основанием для построения моделей ДА послужила гипотеза физических символьных систем, сформулированная Ньювеллом и Саймоном [83]: физическая символьная система по определению должна быть физически реализуемым множеством физических сущностей или символов, которые комбинируются в определенные структуры и способны запускать процессы, оперирующие на этих символах в соответствии с символически закодированными множествами инструкций. Данная гипотеза постулирует утверждение о том, что такая система способна на интеллектуальное поведение в достаточно общем понимании этого термина. Согласно трактовке М.Р. Genesereth, делиберативный агент должен обладать следующими свойствами:

- содержать эксплицитно представленную базу знаний, заполненную формулами в некотором логическом языке, представляющую его убеждения;
- функционировать в следующем цикле: восприятие обстановки (обсервация) – логический вывод – действие ...;
- принимать решения о действиях на основе методов логического вывода.

Вулдридж и Дженнингс [137] определяли ДА как «то, что содержит явно представленную символическую модель мира, и в котором решение (например, о том какие действия выполнять) принимается через логический (или, по крайней мере, псевдологический) вывод, основанный на сопоставлении с образом или символической манипуляции». В соответствии с наиболее общим делиберативным подходом, когнитивный компонент содержит по существу две части: *планировщик* и модель мира (Рис. 2).

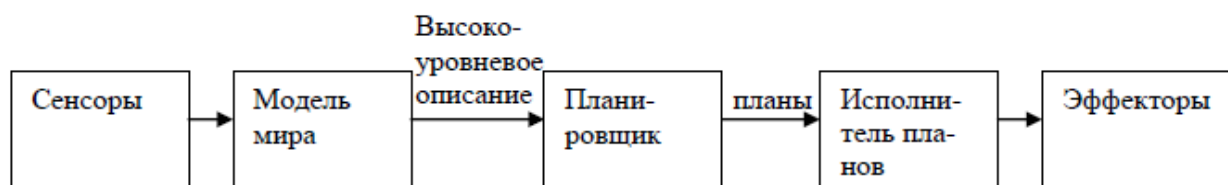


Рис. 2. Базовая архитектура делиберативного агента

Модель мира является внутренним описанием внешней для агента среды и иногда также включает описание самого агента.

Вторая часть когнитивного компонента — планировщик, использует модель мира для создания плана достижения агентом своей цели.

Планировщик действует следующим способом: задавая атомарные действия (операторы), которые агент способен выполнить, их предусловия и их результаты в мире (постусловие), и начальную и целевую ситуации, он ищет в пространстве последовательность операторов, пока не находит такую, которая преобразует начальное состояние в целевое состояние.

Готовый план есть список действий, который передается исполнителю планов, который будет выполнять эти действия, вызывая различные процедуры низкого уровня эффекторов.

Основы строгой формализации знаний и действий делиберативного агента заложены Куртом Конолигом (Kurt Konolige) в работе [102]. Он предложил иерархический метаязык для описания и логического вывода на знаниях и действиях вычислительных агентов.

Основные проблемы, выявленные в данной модели состоят в том, что трехуровневая иерархия устанавливается произволом автора; с вычислительной точки зрения модель Конолиге представляется неуправляемой из-за наличия метауровней; некоторые аксиомы, предложенные для общих фактов оказались неверными. Тем не менее, работы Конолиге послужили стимулом для последующего развития моделей ДА.

Несмотря на то, что в последние десятилетия было предложено несколько архитектур для ДА, большинство из них развертывались только в ограниченных искусственных средах, очень немногие были применены для решения реальных задач, и совсем ничтожное количество доведено до стадии реальных корпоративных приложений, отлаженных на конкретной предметной области. Одной из таких архитектур является Multi-Agent Reasoning System (dMARS) [89], основанная на более ранней системе Procedural Reasoning System (PRS) [90], и использующая концептуальную основу BDI-модели практического вывода: Beliefs, Desires and Intentions – у агента есть убеждения, желания и намерения [73].

Известна формализация ДА средствами языка AgentSpeak(L), разработанного Рао [123]. Этот язык является языком программирования, основанным на абстракции PRS-архитектуры. Далее d'Inverno и Luck построили формальную спецификацию AgentSpeak(L) [82], которая переопределяет оригинальное описание Рао в терминах состояний и операций над состояниями, позволяя перейти к реализации такой модели.

В последующих работах [109, 136] предложена развернутая формальная модель делиберативной архитектуры dMARS, использующая язык Z [131]. Z является модельно-ориентированным языком формальных спецификаций, основанным на теории множеств и логике первого порядка. Ключевыми компонентами Z-спецификации являются пространство состояний системы и множество возможных операций, переводящих одно состояние в другое.

Модели агентов в PRS и dMARS являются примерами наиболее популярной в настоящее время парадигмы, известной как BDI-подход. BDI-

архитектура, как правило, содержит четыре ключевые структуры данных: убеждения, желания (цели), намерения и библиотеку планов.

Агентные убеждения соответствуют информации агента о мире и могут быть неполными и некорректными. Обычно агенты в BDI-модели хранят убеждения в символической форме, подобно фактам в языке Prolog. Желания агентов (или цели) интуитивно соответствуют задачам, назначенным данным агентам. Для действующих BDI-агентов требуется, чтобы желания были логически непротиворечивы, хотя человеческие желания часто этому требованию не соответствуют. Агенты не могут, в общем случае, достичь всех своих желаний, даже если эти желания непротиворечивы. Агенты должны зафиксировать некоторое множество достижимых желаний и передать ресурсы для их достижения. Эти выбранные желания являются намерениями, и агент будет стремиться достигнуть намерения до тех пор пока его убеждения соответствуют желанию либо желание при данных убеждениях не является более достижимым.

Каждый агент в dMARS имеет библиотеку планов, определяющих варианты возможных действий, которые могут быть предприняты агентом для достижения его намерений. Планы, таким образом, реализуют процедурные знания агента. Каждый план содержит несколько компонентов. Триггер или условие вызова определяет обстоятельства при которых план должен рассматриваться как возможный для применения. План имеет контекст или предусловия, определяющие обстоятельства при которых выполнение плана может начаться. План может иметь также главное условие, которое должно быть истинным во время выполнения плана. План также имеет тело, которое может содержать цели и примитивные действия.

События, воспринимаемые агентом, помещаются в очередь событий. Внутренний интерпретатор агента непрерывно выполняет следующий цикл:

- 1) обзора мультиагентный мир и внутреннее состояние агента и изменяет очередь событий;

- 2) генерирует новые возможные желания (задачи), находя те планы, чьи триггеры событий включены;

- 3) выбирает из этого множества включенных планов один для выполнения;

- 4) помещает желаемое значение в существующий или новый стек, в соответствии с тем, имеется или нет подцель;

- 5) выбирает стек намерений, читает план, находящийся в вершине стека и выполняет следующий шаг из этого плана; если шаг есть действие – выполняет его, если это подцель – посылает эту подцель в очередь событий;

- 6) возврат к шагу 1.

Шаг 5 этого цикла реализует способ выполнения плана, когда подцель помещается в очередь событий, что вызывает активизацию очередного плана и т.д.

Содержательно модель агента в dMARS представляет собой ситуационную формальную систему.

Развитие моделей ДА идет по пути попыток формализации новых мотивационных свойств и отношений в комбинации с поведением и действиями агентов. Такой подход приводит к созданию абстрактных логических моделей, претендующих на строгое формальное описание всех релевантных свойств рациональных агентов в целях анализа, спецификации и верификации МАС.

Одной из наиболее разработанных моделей такого типа является модель, предложенная в работах Линдера, Хоека, Мейера (B. Van Linder, W. van der Hoek, J.-J. Ch. Meyer) [71, 108]. Главное свойство ДА, которое рассматривают эти ученые – знание. В представлении знаний ими используется общий подход эпистемической логики при котором формула  $K_i\psi$  означает тот факт, что агент  $i$  знает  $\psi$ .

Формула интерпретируется в стиле семантики возможных миров Крипке. В данной модели выделяются два уровня свойств: уровень утверждений, где рассматриваются предпочтения и цели, и практический уровень, где определяются обязательства. В аспекте действий агента исследуются результаты, способности и возможности. Авторами проводятся достаточно тонкие различия между этими терминами: способности агента выражают его ментальные и физические мощности, моральные характеристики, человеческие и физические возможности, в то время как возможность выполнения действий описывается в терминах «возможности по обстоятельствам» (circumstantial possibility).

С неформальной точки зрения цели в модели Линдера, Хоека, Мейера определяются как неосуществленные предпочтения агента, но не все неосуществленные предпочтения агента квалифицируются как цели. Так как они не осуществлены, то цели должны быть реалистичными, то есть неосуществленное предпочтение является целью для агента, только если агент знает, что как-то может осуществить предпочтение.

Цель тогда определяется как неосуществленное, реалистичное, известное предпочтение. Основные ограничения в плане программно-технической реализации данной модели заключаются в том, что сами авторы позиционируют свою модель как предназначенную для теоретического анализа возможных моделей агентных систем, а также в том, что в их построениях проводятся «тонкие», буквально на философском уровне постулируемые различия между способностью и возможностью агента выполнить некоторое действие.

Построение делиберативных архитектур требует решения таких проблем, как построение адекватного символического описания реального мира, учитывающего сложность происходящих во времени процессов и действующих объектов; организация логического вывода из имеющихся знаний, который должен приводить к определённым действиям агентов.

Чепмен [77] показал, что планирование неразрешимо в общем смысле, и вычислительно ведет к NP-полноте даже в простых формах. С другой стороны, традиционное планирование очень продолжительно, требует

экспоненциального поиска в потенциально неограниченном проблемном пространстве.

Достоинством делиберативных моделей и архитектур является возможность применения строгих формальных методов и хорошо отработанных технологий традиционного ИИ, позволяющих относительно легко представлять знания в символьной форме и переносить их в АОС. В то же время создание полной и точной модели некоторой предметной области реального мира, формализация ментальных свойств агентов и процессов рассуждения в этих когнитивных структурах представляют существенные трудности для технической реализации.

### 1.3.2. Реактивные агенты и архитектуры

Поиски путей разрешения проблем, возникающих при использовании в агентных системах классических методов ИИ, привели к появлению нового класса – реактивных архитектур. Основоположником этого направления принято считать Р. Брукса (R. Brooks [74]), который так сформулировал ключевые идеи бихевиористического взгляда на интеллект:

- интеллектуальное поведение может создаваться без явного символьного представления знаний;
- интеллектуальное поведение может создаваться без явного абстрактного логического вывода;
- интеллект является внезапно возникающим свойством некоторых сложных систем [74].

В реальном мире интеллект не является экспертной системой или машиной логического вывода, а интеллектуальное поведение возникает как результат взаимодействия агента со средой. Вместо моделирования мира и планирования реактивные агенты (РА) должны иметь коллекцию простых поведенческих схем, которые реагируют на изменения в среде в форме «стимул – реакция» (Рис. 3).

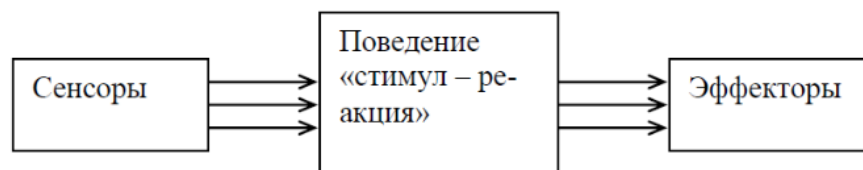


Рис. 3: Базовая архитектура реактивного агента

Наиболее спорным из принципов Брукса является принцип, касающийся репрезентации. Он утверждает, что явные представления мира не являются необходимыми для реализации эффективных агентов. Вместо этого агент должен использовать «... мир как его собственную модель – непрерывное обращение к собственным сенсорам лучше, чем к внутренней модели мира» [74].

Реактивные агенты, по крайней мере в нескольких экспериментах, доказали способность решать ограниченное число простых задач в областях реального мира. Однако они сталкиваются с проблемами при выполнении задач, требующих знаний о мире, которые получаются логическим выводом или из памяти.

Таковыми заданиями являются проблемы понимания ситуации на основе объективных перспектив, предсказание поведения других агентов, креативность. Более того, реактивные агенты часто сделаны «жестко» и не имеют никаких способностей к обучению. Тот факт, что каждое поведение должно быть отдельно закодировано в агенте ведет к сложным проблемам, как во время разработки, так и во время выполнения.

Проанализируем более детально наиболее известные модели реактивных агентов и распределенных реактивных архитектур.

В работах таких ученых как К. Cetnarowicz, Е. Nawarecki, М. Zabinska [76, 124], предложена М-агентная архитектура МАС. Эти авторы определяют мультиагентный мир (МА-мир), включающий агентов, агентное пространство, агентную среду, отношения между агентами и средой и отношения среди агентов.

В рассматриваемой модели М-агентной архитектуры заложена ориентация на работу агентов с ресурсами, причем связь агентов и пространства точно не определена. В этой модели отсутствует иерархия агентов, не определяется логическая взаимосвязь агентов, не предусматривается возможность логического вывода о состояниях МА-мира и взаимоотношениях среди агентов. Поведение агента в М-агентной архитектуре по сути является борьбой за определенный ресурс, так как цель агента формулируется как функция достижения ресурса, при этом не учитывается предыстория развития системы, не предлагается конструктивного механизма реализации стратегий.

Как показали Вулдридж и Дженнингс [136, 137] известные спецификации моделей ИА определяют те свойства, которые должны у агентов быть, но не направлены на конструктивное определение поведения агентов. Главным следствием из этой ситуации является невозможность связать высокоуровневые спецификации с программной реализацией таких моделей в МАС. Поэтому для моделей РА характерно использование концепции состояний и переходов и механизмов поведения типа «стимул-реакция». Ярким примером такой реактивной модели является модель Дж. Фербера [85].

Попыткой синтеза делиберативных и реактивных моделей является модель, предложенная в работах С. Амброшкевича, С. Билки, Я. Комара [68]. Авторы данной модели стремятся представить модель МАС так, чтобы фундаментальные понятия теории, такие как знания агента, убеждения, намерения, мотивы, модели логического вывода, коммуникация, переговоры, обязательства, координация и агентная организация имели ясную и простую семантику.

В данной модели выделяется пять базовых модулей агентной архитектуры: перцепция, тип агента, знания и убеждения, коммуникационный модуль, модуль вывода.

Агент ощущает окружающий мир, а поскольку этот мир есть динамическая система, то полная информация о мире есть в текущем глобальном состоянии динамической системы. Однако, агенту недоступна полная информация о глобальном состоянии.

Динамическая система определяет мир существования агентов, который состоит из множества глобальных состояний, с начальным состоянием; множества агентных позиций; множества действий, приписываемых каждой позиции, таких что агент, занимающий позицию может сделать только одно действие в любой дискретный момент времени; передаточной функции, определяющей следующее глобальное состояние системы, задаваемое текущим состоянием системы и действиями, сделанными агентами; продолжительностью существования системы.

Здесь, по мнению в [66], состояние локального мира агента совмещается с состоянием самого агента, а различие между ними должно быть, так как агент есть выделенная сущность мира и имеет свои состояния.

В модели Амброшкевича поведение агента определяется исключительно задачей, причем задача трактуется как множество действий, не рассматривается последовательность во времени, не анализируются логические взаимосвязи, не выполняется оценка состояний МА-мира. Процедурный аспект данной модели соответствует модели конечного автомата. По признанию самих авторов данная агентная модель не полна, так как поведение агента должно основываться на убеждениях подвергнутых ревизии на основе информации перцепции и коммуникации, а эти переходы не определены.

Модель реактивного агента, функционирующего в реальном масштабе времени построена и исследована в работах В.В. Денисова, Д.В. Пузанкова, М.Г. Пантелеева, Г.Г. Колосова [15, 22]. ИА определяется через совокупность пространств ментальных состояний и отображений между ними.

### **1.3.3. Гибридные агенты и архитектуры**

Реактивный подход позволяет эффективно использовать множество простых сценариев поведения агентов в рамках установленных реакций на определённые события окружающей среды, но его ограниченность проявляется в практической невозможности полного ситуативного анализа всех возможных активностей агентов. Поэтому в большинстве проектов и действующих систем используются гибридные архитектуры.

В последнее время ряд исследователей признают, что агент должен обладать высокоуровневым выводом и низкоуровневыми реактивными способностями. На этом пути можно использовать реактивную способность для текущих задач и добавить возможность логического вывода для более сложных долговременных задач.

Достаточно давно были выделены [93] две категории гибридных агентных архитектур. Однородная архитектура использует одну общую репрезентацию и схему управления для реакций и рассуждений, в то время как многослойные архитектуры используют различные представления и алгоритмы (реализованные в отдельных слоях) для выполнения этих функций (Рис. 4).

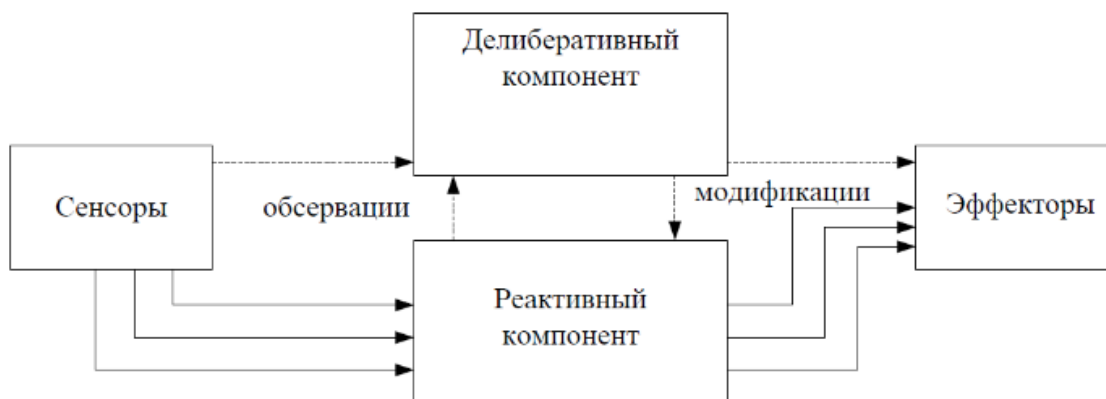


Рис. 4. Базовая архитектура многослойного гибридного агента

Реактивный компонент отображает перцептуальные стимулы на примитивные действия. Делиберативный компонент выполняет символический вывод для управления поведением реактивного компонента, например, изменяет множество правил ситуация – действие. В некоторых архитектурах делиберативный компонент прямо связан с сенсорами и эффекторами агента, в других нет.

Когда проектируется архитектура многоуровневого гибридного агента необходимо получить ответы на следующие критические вопросы:

- 1). Достаточно только одного реактивного и одного делиберативного уровня, или должны быть введены дополнительные уровни? Как когнитивная рабочая нагрузка должна быть разделена между уровнями?
- 2). Как должны взаимодействовать компоненты различных уровней?
- 3). Когда должен агент действовать и когда должен размышлять, т. е. как определить алгоритм расписания?

I.A.Ferguson предложил архитектуру циклической машины, основанную на модели машины А.М. Тьюринга [86] включающая три уровня: реактивный уровень, который представляет собой множество правил «ситуация – действие», планирующий уровень, чьим главным компонентом является иерархический, неполный планировщик, и моделирующий уровень. Задача реактивного уровня – обеспечить быструю

реакцию на события в реальном масштабе времени. Главная задача планирующего уровня – сгенерировать и выполнить планы для достижения долговременных целей агента. Наконец, задача моделирующего уровня – обнаружить и предсказать ситуации потенциального конфликта целей между агентами и затем предложить действия для выхода из этих конфликтов (Рис. 5).

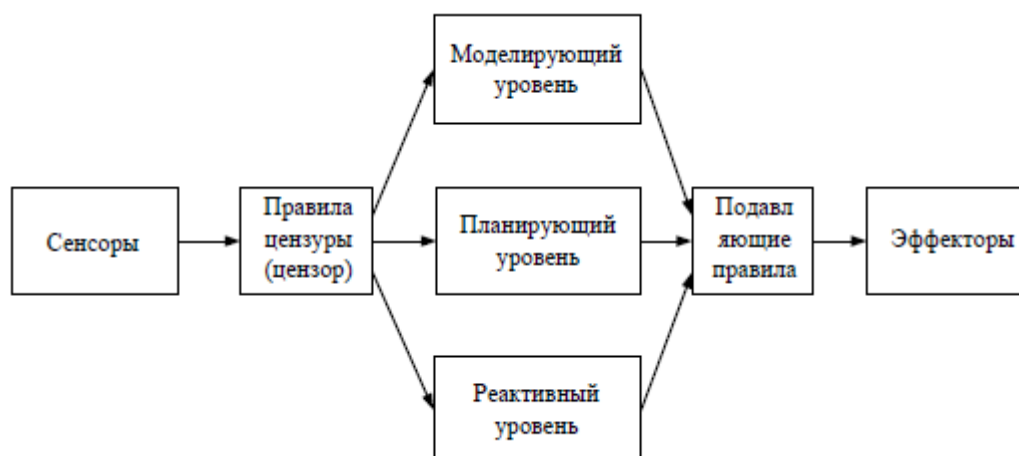


Рис. 5. Базовая архитектура агента циклической машины

Каждый уровень независимо связан с сенсорами и эффекторами и действует так, как если бы он управлял агентом самостоятельно. В результате этого действия будут часто конфликтовать друг с другом. Конфликты разрешаются подавляющими правилами. На стороне рецепторов имеются правила цензуры, которые фильтруют сенсорные данные, так что каждый уровень получает соответствующую часть сенсорных данных.

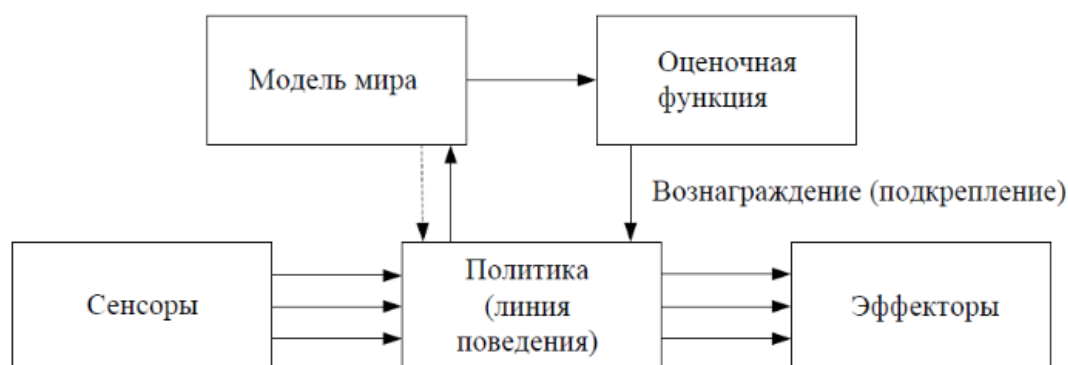
Главная идея агента «планировщик-реактор», предложенного в [110] состоит в том, что планировщик является системой, которая взаимодействует с реактором таким же способом, каким реактор взаимодействует с миром (Рис. 6). Планировщик содержит: модель мира, описание реактора и установку от пользователя на цели реактора и ограничения, которым должно удовлетворять поведение реактора.



*Рис. 6 Базовая архитектура системы «планировщик-реактор»*

Планировщик функционирует непрерывно, отслеживая поведение реактора с использованием перцепции, и воздействует на реактор в соответствии с установкой, данной пользователем. Он не имеет доступа к полному содержанию реактора, т.е. к рядам сенсорных данных, его обзор ограничен выходом процессов перцепции. Реактор и Планировщик рассматриваются как полностью независимые параллельные системы.

Архитектура DYNA, предложенная Sutton [125, 134] может рассматриваться как гибридная, но не вполне многоуровневая, интегрирующая реактивность и планирование. Она ориентирована на обучение с подкреплением и включает три главных компонента: модуль политики, который является реактивной системой, модель мира и оценочную функцию, которая отображает состояния в величины поощрений (Рис. 7).



*Рис. 7. Базовая архитектура DYNA-агента*

Вследствие закрытого отношения для обучения с подкреплением требуется, чтобы модель мира производила подкрепление после каждого выполненного действия.

Планирование перекрывается с действием не полностью. Они не могут выполняться параллельно, потому что во время планирования DYNA отключает восприятие реального мира и действие. В течение фазы планирования она не может реагировать на события реального мира, что делает ее очень уязвимой.

Существенным недостатком гибридных архитектур является то, что они делаются ad-hoc, т.е. их структура достаточно хорошо обоснована с конструкторской точки зрения, но не ясно, как ее обосновать со стороны любой глубокой теории. В основе BDI-архитектур лежит строгая теоретическая концепция, но они недостаточно приспособлены для реального проектирования ресурсно-ограниченных и целенаправленных агентных приложений. Многослойные архитектуры хорошо поддерживают моделирование различных уровней абстракции, ответственности и сложности представления знаний, но слишком сложны для формального назначения свойств агентов и во многом зависят от интуиции разработчика.

Гибридные архитектуры позволяют гибко сочетать возможности реактивного и делиберативного подходов, таким образом, в таких архитектурах можно совместить логические и автоматические подходы в управлении без строгого подчинения одного другим. Но такие архитектуры в большинстве случаев сильно зависят от специфики приложений, для которых они разрабатываются.

Итак, можно сделать следующие выводы. Два варианта объединения разных подходов к моделированию агентов это однородные архитектуры и многослойные. Многослойные архитектуры подразумевают, что для разных уровней абстракции используемых при принятии решений агентом (действия, тактики, стратегии, намерения, желания, цели и т.п.), существуют различные представления и алгоритмы.

В работах по созданию многослойных моделей агентов, функционирующих в реальном времени, делиберативный компонент (ДК) «внедряется» в архитектуру агента как управляющий реактивным компонентом (РК) и способный его изменять. Для этого, во-первых, ДК должен содержать модель РК в символьном виде, во-вторых, реактивный компонент должен быть внешне открыт для изменения. Это требуется для выполнения условий адекватного времени срабатывания агента в реальном времени. ДК должен уметь оценить время размышлений и реактивного срабатывания: так организовать деятельность, чтобы произвести как можно больше размышлений и в то же время успеть отреагировать на ситуацию. Можно сказать, ДК в такой архитектуре используется как модель авторефлексии над реактивным уровнем, причём не только пассивно, но и подстраивая РК для поставленных ДК задач.

В однородной архитектуре ДК и РК полностью сливаются, иначе говоря, логика размышлений ДК способна делать вывод, используя понятия реактивного уровня, и выводы ДК есть реактивные действия, которые временно сохраняются для сокращения размышлений в похожих ситуациях.

Но тогда, в плане соотношения ДК и РК вариант однородной архитектуры есть предельный вариант многослойной архитектуры, когда ДК способен полностью управлять РК. Это требует слишком сложной логики для ДК (совмещение примитивов и абстракций), что, как отмечено в [93], в работах по данному направлению приводит к неполной формализации модели, а описанию только «каркаса» («framework»), а не архитектуры агента. С другой стороны, как указывает Непейвода Н.Н. в [37], достижения современной логики (как науки) говорят о том, что нет «универсальной» логики (как модели рассуждений), различные логики истинности, времени, денег и др. не совместимы друг с другом изначально и нужно понимать в каких условиях какую применять. То есть логики на разных уровнях абстракции сложно совместимы.

Есть другой предельный вариант многослойной архитектуры – разделение по уровням абстракции, так что на любом уровне может быть задействован произвольно как ДК, так и РК. Такие модели зачастую строятся для имитационного моделирования и используют для уровней только РК, не используя ДК, т.к. среда в которой действуют такие агенты обычно замкнута в определениях [66], к тому же реализовать РК и наладить его работу проще и быстрее. Но, возможно, из-за подобного уклона в сторону реактивности такой вариант даже не упоминается в большинстве обзоров [66, 86, 137]. Данный подход может быть успешно применен в области имитационного моделирования, где ограничения на скорость реакции практически нет (можно запустить на суперкомпьютере и подождать месяц), а компоненту каждого уровня достаточно ставить формальные задачи для его уровня абстракции и получать ответ из определённого множества возможных ответов. Собственно обычно ответ более высокого уровня есть задача для более низкого (например, обобщённая иерархическая структура [20]) или же ответы разных уровней просто «складываются», т.е. каждый уровень отвечает только за себя.

Таким образом, получается интересный побочный вывод, что авторефлексия неявно используется в построении многослойных архитектур агентов как средство управления, способное решать задачи, требующие учёта времени исполнения. При этом используется только первый ранг авторефлексии, возможно ввиду требуемой скорости работы агента. Возможная модель авторефлексии второго ранга могла бы иметь ещё один ДК или РК, управляющий связкой ДК→РК.

Рефлексия же второго рода, так как относится к моделям других агентов, может быть реализована без каких-либо связей с выбранной архитектурой агента, на что указывает требование для агентов МАС наличия способности коммуникации и группового взаимодействия.

Итак, можно утверждать, что для построения имитационного моделирования, в основном достаточно использовать многослойную архитектуру с разделением по уровням абстракции.

Таким образом, мы описали важнейшие общие принципы построения

МАС и нашли, вероятно, более подходящее решение для построения имитационного моделирования.

Перейдём к описанию подходов моделирования эвакуации и реализованных имитационных моделей.

#### **1.4. Модели эвакуации**

В моделях эвакуации моделируется движение людей, находящихся в здании, к выходам из него. Объект моделирования – люди и их движение. Известные модели оперируют двумя уровнями абстракции для моделирования движения людей – людской поток и отдельный человек.

В соответствии с уровнем абстракции, используются и определённые модели пространства-здания. Для людских потоков участки пути разбиваются на взаимосвязанные элементарные участки, размером не менее метр на метр, на основе которых и считается движение людского потока, чаще всего это прямоугольная сетка (как для подхода клеточных автоматов, но элементарные участки имеют больший размер). Графовая модель здания может быть отнесена к этому же классу. Модель пространства при уровне абстракции отдельного человека чаще всего – непрерывная плоскость с геометрическими ограничениями в виде стен здания. Такая модель соответствует возможности человека двигаться в произвольном направлении и с произвольной скоростью, не превышающей максимальную. Исключение составляет только подход клеточных автоматов, где люди перемещаются скачками из клетки в клетку.

Рассмотрим последовательно основные известные модели эвакуации.

##### **1.4.1. Модель людских потоков**

Группа одновременно движущихся в одном направлении людей образует людской поток. Людской поток, грубо говоря, является прямоугольником, такой вид потока определяется характером движения людей в здании – обычно это идущие в прямоугольном коридоре люди.

Пути движения людских потоков служат проходы в помещениях различного назначения и специальные помещения: кулуары, коридоры, фойе, лестницы, пандусы. Пути движения характеризуются свободной длиной  $L$  и шириной  $b$ . Они имеют сужения за счёт различных выступающих из плоскости стен элементов. В данном подходе считается, что длина пути  $L$  в проёмах относительно невелика и не учитывается.

Пути движения делятся на участки, различающиеся по характеру (горизонтальный, наклоны и проёмы) и параметрам  $L$ ,  $b$ . Проёмы считаются самостоятельными участками пути. Механические средства передвижения людей (эскалаторы, лифты, движущиеся полы и т.п.) считаются самостоятельными участками пути.

Людской поток характеризуется числом людей  $N$ , плотностью  $D$  и скоростью движения  $v$  и имеет два измерения – длина  $l$  и ширина  $\delta$ . Путь характеризуется пропускной способностью, т.е. числом людей проходящих в

единицу времени через поперечное сечение пути (этот подход проанализирован в книге [47], далее краткий анализ модели людских потоков в основном базируется на ней).

#### **(i) Исследования Беляева и их развитие**

Под руководством С.В. Беляева в институте архитектуры Всероссийской академии художеств (ВАХ) в 1937 г. [4] было проведено около 200 наблюдений за движением потоков в общественных местах.

Исследования дали численные значения для различных простых ситуаций движения людского потока (поток-ряд, минимальная пропускная способность потока, поток на лестнице и т.п.). На основе результатов этих исследований были даны нормы расчёта и нормирования эвакуации людей из зданий. Основным научным достижением исследований ВАХ считается установление зависимости  $v = f(D)$ .

К серьёзным недостаткам этих исследований относят относительно небольшое число натуральных наблюдений и неудачно выбранные места для наблюдений – тротуары улиц, трамвайные вагоны, где движение людей сильно отличается от движения в зданиях; изображение двигающейся массы людей в виде элементарных потоков, не отвечающее действительному расположению людей в потоке и т.д.

Главным же недостатком работы ВАХ считают установление постоянных значений в качестве нормативных и расчётных скоростей движения.

Исследования М.Я. Ройтмана в Высшей школе МВД СССР [50, 51] в теоретическом отношении базируются на принципах С.В. Беляева, однако в ней были устранены некоторые недостатки работы ВАХ.

Проведённые в 1946-1948 гг. А.И. Милинским во Всесоюзном научно-исследовательском институте противопожарной обороны (ВНИИПО) исследования [35, 36] отличались от предыдущих широким масштабом и более совершенной методикой.

Было проведено более 6000 натуральных наблюдений в различных зданиях общественного назначения. На основании наблюдений были установлены средний размер взрослых и подростков (на уровне груди) в зимней и летней одежде, а также площади их горизонтальных проекций.

Плотность людского потока было предложено относить к площади пола и выражать либо числом человек, приходящихся на единицу площади, либо суммой площадей их горизонтальных проекций на ту же единицу.

Определены наиболее вероятные плотности потока на путях движения. Одновременно выявлена возможность образования повышенных плотностей в дверных проёмах и остановки движения.

Наблюдениями подтверждена зависимость скорости людского потока от плотности и установлены средние скорости для различных плотностей потоков. Важное значение имеет установление амплитуды колебаний скоростей движения при разных величинах плотности, а также вероятностные значения скоростей движения.

Впервые при определении пропускной способности проёмов введено понятие удельной пропускной способности проёмов:  $q = D v$ .

Исследованиями ВНИИПО сделан новый шаг вперёд в понимании процессов движения людских потоков. Большую ценность представляют экспериментальные данные о плотностях и скоростях движения людских потоков и о скорости как функции плотности. Они выполнены в таком объёме, что дали надёжную базу для дальнейших теоретических обобщений.

Однако работами ВНИИПО не были вскрыты объективные связи, отношения и закономерности, которые могли бы стать основой для построения теории движения людских потоков.

Основной целью исследований, проведённых В.М. Предтеченским на кафедре архитектуры Московского инженерно-строительного института им. В.В. Куйбышева, было создание теоретических основ процесса движения людских потоков и на их базе разработка более совершенного метода расчёта [47, 48, 49].

Для натуральных наблюдений был предложен и впервые применён фотокинометод, позволяющий оценить не только количественные, но и качественные характеристики процесса. Наблюдения (общим числом около 800) показали достоверность полученных ВНИИПО экспериментальных данных. Последующая обработка и теоретические исследования дали следующие результаты.

Прежде всего, была предложена классификация видов движения и установлены основные расчётные случаи. Результаты натуральных замеров скоростей движения как функции плотности, полученные ВНИИПО и МИСИ, обработаны при помощи методов математической статистики, и получены средние значения  $v = f(D)$  для обычных условий движения людских потоков. На основании опытных данных о пропускной способности проёмов, рассчитаны скорости движения в проёмах (как функция плотности).

Для всех других видов путей исследована величина удельной пропускной способности (интенсивности движения)  $q = D v$ , установлены закономерности её изменения относительно плотности.

Получены соответствующие аналитические выражения для определения параметров потока на каждом последующем участке пути по параметрам каждого предшествующего. Это впервые позволило рассматривать движения людского потока как единое целое.

Установлены закономерности образования скопления людей и выявлено влияние экстремальных значений  $q$  как фактора, определяющего эти скопления. Получены закономерности изменения параметров движения потоков при их слиянии или расчленении, а также при переформировании, т.е. при изменении плотности и скорости по ходу движения.

На основании найденных закономерностей предложен метод, позволяющий по расчётным предельным состояниям учитывать движение

не только в нормальных условиях, но и при вынужденной эвакуации. Предложена также графическая интерпретация процесса движения, позволяющая рекомендовать графоаналитическую методику расчёта.

В 60-е и 70-е годы происходило развитие данной работы. Ряд работ был посвящён исследованиям особых случаев движения людских потоков: закономерностей движения длительно существующих людских потоков (Р.Г. Григорьянц) и движения пересекающихся потоков (В.С. Гвоздяков). Были выполнены работы по исследованию закономерностей движения людских потоков в зданиях, где движение является одним из основных функциональных процессов.

Наконец были проведены исследования в области нормирования вынужденной эвакуации людей из зданий, поскольку действующие общие противопожарные нормы проектирования зданий и сооружений недостаточно обоснованы с научной точки зрения.

### **(ii) Зарубежные работы**

Среди зарубежных работ по исследованию людских потоков можно отметить работу Эдинга [117], посвящённую исследованию движения людских потоков в зависимости от времени и места движения. Им предложена своеобразная и спорная с научной точки зрения классификация видов движения (смешанное движение пешеходов различных профессий; целеустремлённый поток людей на промышленных предприятиях и т.п.).

Управлением транспорта Лондона [84] проведено исследование движения людских потоков в сооружениях метрополитена, в результате которого была получена зависимость скорости и интенсивности движения людского потока от плотности. Людские потоки создавались искусственно и были однородны по физическим данным людей, но искусственно созданные потоки не отражают действительной картины процесса движения.

Основным недостатком известных зарубежных исследований движения людских потоков исследователи считают отсутствие теории, позволяющей описать процесс движения людского потока от момента его формирования до завершения.

### **(iii) Алгоритмизация**

Работа [47] носит уже комплексный характер, где подробно описаны достигнутые к тому времени результаты: виды и особенности движения людей, параметры движения людских потоков, основные расчётные случаи движения потоков, а также некоторые особые случаи. В данной книге собраны не только количественные данные, но и систематизированы качественные показатели, достаточно внимания уделено особым случаям и эффектам, возникающим при движении, хотя и указывается, что влияние их на практический результат (расчётное время выхода) невелико.

В этой книге подробно описан алгоритм для процедурном языке АЛГОЛ-60 для нахождения расчётного времени эвакуации людей из здания. Дана алгоритмическая схема и сам текст программы.

Здесь также впервые даётся описание графоаналитического метода (о

котором речь пойдет ниже).

В работе [53] отмечено, что эмпирическая база натуральных наблюдений людских потоков в зданиях различного назначения, на которую ориентировались теоретические построения [46,47] получившие широкое признание за рубежом, хотя и была самой обширной в мире в 60-х годах, но составляла лишь четверть той, которая была накоплена к концу 70-х годов. Количественное разнообразие результатов проведенных серий натуральных наблюдений остро поставило проблему теоретического обоснования наблюдаемых зависимостей между параметрами людских потоков. Это – во-первых. Во-вторых, несмотря на кажущуюся простоту, графоаналитический метод расчета людских потоков [46, 47] был трудоемок для проектной практики и недостаточно полно отображал существующее словесное описание этого процесса.

Диктуемый этими проблемами огромный объем научно-исследовательских работ был проделан в кратчайшие сроки. Как итог этих работ впервые имеем [59] :

- статистический анализ однородности выборочных совокупностей значений скорости людских потоков в 69 сериях натуральных наблюдений и экспериментов;
- описание людского потока как случайного процесса;
- общий вид зависимости скорости людского потока в виде случайной

$$\text{функции: } V_{D,j} = V_{0,j} \left( 1 - a_j \ln \frac{D_i}{D_{0,j}} \right);$$

- математическую модель свободного движения людского потока (SDLP), основанную на соотношениях теории вероятностей;
- стохастическую имитационную модель анализа движения людских потоков (ADLPV) и ее детерминированный вариант (ADLP).

Теперь подробнее остановимся на практических результатах этих работ.

#### **(iv) Графоаналитический метод**

В отсутствие ЭВМ наиболее точным и наглядным методом математического моделирования людских потоков является графоаналитический [46, 47].

А.И. Милинский разработал графоаналитический метод расчета, «позволяющий учитывать особенности планировки здания, плотности и скорости движения потоков людей на всех участках для любого момента времени, отражая вместе с тем всю картину процесса и его продолжительность» [35, с. 145]. Для расчета параметров движения потока на последующем

участке он предложил использовать соотношение:  $D_{i+1} = \frac{D_i \cdot b_i}{b_{i+1}}$ .

Однако использование данного соотношения приводит к тому, что «... плотность на участках перед проемом и после проема при равной ширине участков оказывается одинаковой даже в том случае, когда пропускная способность проема меньше пропускной способности предшествующего

проему участка. Из расчета также следует, что... при одном и том же количестве людей, но при разных ширинах предшествующего проему участка, плотность в проеме не меняется» [48, с. 70-71].

Для решения этой проблемы, под руководством проф. В.М. Предтеченского были проведены натурные наблюдения движения людей перед проёмами или сужениями пути. Объектами наблюдения были выбраны станции московского метрополитена. Основываясь на результатах этих наблюдений и анализе исследований ВНИИПО, проф. В.М. Предтеченский вывел формулу для удельной пропускной способности, названной им интенсивностью движения для пересечения границы смежного участка пути:

$$q_{i+1} = \frac{q_i \cdot b_i}{b_{i+1}}$$

Профессор В.М. Предтеченский модернизировал графоаналитический метод расчета. Благодаря модифицированному графоаналитическому методу удалось промоделировать незамеченные ранее особенности процесса: перестроение, растекание потоков, образование и рассасывание скоплений.

Расчет графоаналитическим методом даёт более точные результаты по сравнению с расчётами по ГОСТам, однако значительно более трудоемок, ведет к почти 15 % увеличению времени эвакуации по сравнению с результатами моделирования с помощью алгоритма ADLPV.

Графоаналитический метод – это уникальный метод, аналогов у него нет и, скорей всего, уже не будет. Ни в одной стране мира (судя по опубликованным работам), кроме нашей, никогда не было законченного «ручного» метода расчета эвакуации людей. Но он уже недостаточно точен.

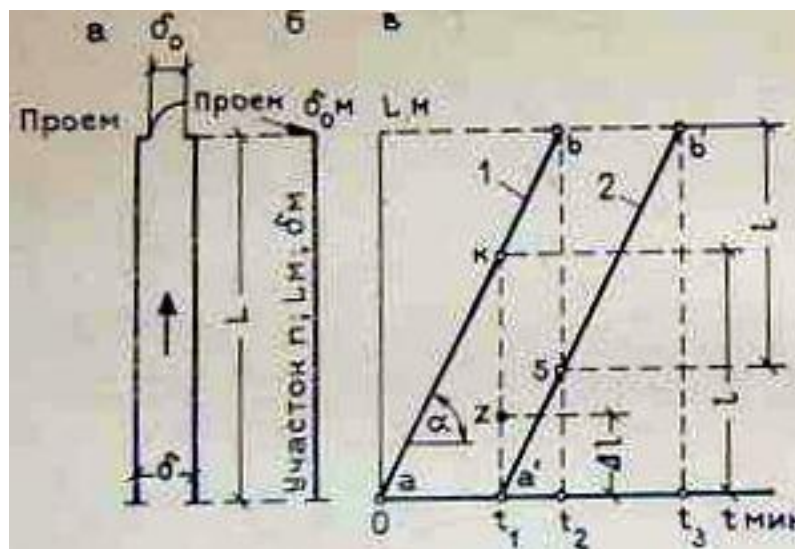


Рис. 8. Графоаналитический метод

Графическое изображение движения людского потока позволяет наглядно проиллюстрировать весь процесс. На Рис. 8 представлено движение людского потока по горизонтальному или наклонному пути. Слева (Рис. 8,а) показан в плане участок длиной  $L$  и шириной  $\delta$  горизонтального пути,

заканчивающегося проёмом. Движение людей происходит в направлении, показанном стрелкой; расчётная схема того же участка дана в виде прямой линии (Рис. 8,б). На Рис. 8,в дана координатная система с осью ординат  $L$  (длина пути) и осью абсцисс  $t$  (время).

Пусть в начальный момент при  $t = 0$  на участке появляется головная часть потока. Через время  $t_1$  весь поток входит в зону участка, включая его замыкающую часть. Рассматриваемый поток состоит из  $N$  человек и имеет плотность  $D$ .  $t_2 = \frac{L}{v}$ . Порядок построения графика движений следующий.

Зная плотность потока  $D$ , из расчётной таблицы находится скорость движения  $v$ , соответствующую этой плотности и условиям движения (аварийное, нормальное, комфортное). По скорости  $v$  и длине участка  $L$  определяется время  $t_2$ , мин., необходимое для прохождения данного участка:

Пересечение откладываемых на графике значений  $t_2$  и  $L$  в точке  $b$  соединяется с началом координат прямой  $ab$ , так как скорость движения на этом участке постоянная (поскольку по условию одинакова его ширина и плотность потока). Прямая  $ab$  выражает характер движения головной части потока на данном участке пути.

На основании условия, что через время  $t_1$  весь поток входит в зону участка, и учитывая, что скорость движения остаётся неизменной, можно без предварительных вычислений из точки  $a'$ , соответствующей началу пути и времени  $t_1$ , провести прямую  $a'b'$ , параллельную  $ab$ , до ординаты, соответствующей длине  $L$ . Прямая  $a'b'$  выражает движение замыкающей части потока. Более подробно про графоаналитический метод расчёта можно узнать в [47, 48].

#### (v) Модель ADLPV

Название программы есть аббревиатура: Анализ Движения Людских Поток, Вероятность.

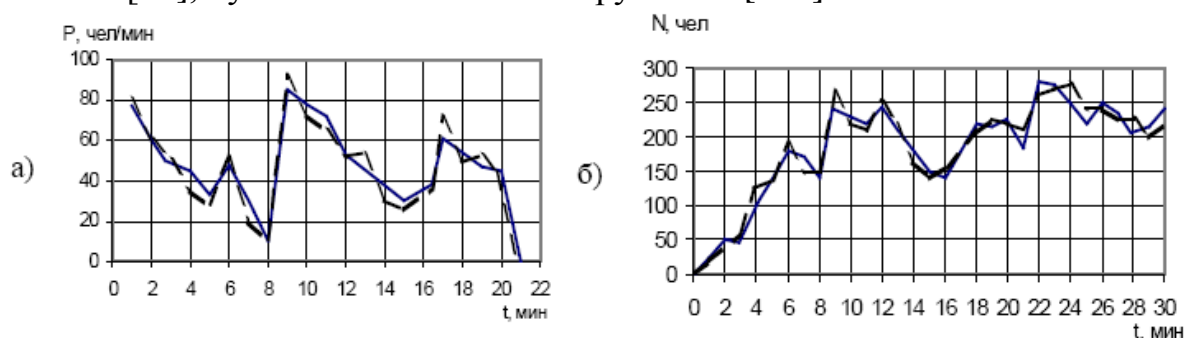
ADLPV – это стохастическая имитационная модель анализа движения людских потоков. Данная модель является одним из итогов работ [58, 59]. При использовании модели ADLPV производится разметка эвакуационных путей на «элементарные» участки и намечаются маршруты движения людских потоков, а затем производится одновременное моделирование их движения.

Людские потоки могут рассматриваться на разных уровнях: на более общем (макро) или, наоборот, на более детальном (микро). На макро уровне людские потоки рассматриваются, например, при проектировании структуры города как пассажирские и пешеходные потоки. На микро уровне людской поток анализируется при моделировании индивидуального поведения и, соответственно, передвижения людей в нем.

Каждый из уровней использует свои характеристики людских потоков. Чем выше уровень, тем более генерализированные характеристики в нем используются. Но характеристики более высокого уровня должны развертываться в систему показателей более низкого уровня. Так, например, данные о

пассажиропотоках в больших городах детализируются в почасовую нагрузку на станции метрополитена, которая и служит исходными данными для определения ожидаемой величины людских потоков при проектировании станций и пересадочных узлов метрополитена [19]. Также и показатели индивидуального передвижения людей в потоке при его микро моделировании должны воспроизводить в своей совокупности статистику наблюдаемых закономерностей людского потока.

В отсутствии ЭВМ наиболее точным и наглядным методом математического моделирования людских потоков был графоаналитический [47]; с появлением ЭВМ – модели ADLPV и SDLP с соответствующими им программами, которые модернизируются по мере развития компьютерной техники и её программного обеспечения. Но при любом способе и технике моделирования оценкой качества модели всегда была и будет степень воспроизведения ею моделируемого реального процесса. В мировой литературе чаще заявляется, чем демонстрируется степень соответствия моделей и натурных наблюдений людских потоков. В отличие от западных специалистов представители российской высшей школы не коммерциализируют, а весьма детально, даже в учебных пособиях [46, 47, 60] излагают разработанные ими методы и полученные результаты. В свете изложенного, приведенные на Рис.9 графики реально наблюдаемых людских потоков и полученные при моделировании с теми же исходными данными, уникальны. Они взяты из [19], публиковались также за рубежом [100].



*Рис.9 Результаты моделирования и реальные данные*

Приведенные графики позволяют наглядно оценить высокую степень соответствия моделируемого и реального количества людей, проходящих в сечение в пути в последовательные моменты времени: наблюдаемые значения лежат в границах доверительных интервалов моделирования, а сложная динамика распределения – практически идентична. Столь высокое качество модели ADLPV позволяет принять её результаты в качестве эталонных при сравнении с результатами, полученными графоаналитическим методом и по формулам тех основных положений, которые вошли в ГОСТ 12.1.004.

Результаты этих отечественных разработок использованы также в методиках Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий для определения расчетных величин пожарного риска [34].

Приведённые результаты говорят о проработанности модели потоков для практических нужд обеспечения безопасности эвакуации, т.к. дают близкое к реальности время выхода людей из здания по определённым заранее маршрутам. Для теоретических нужд исследования поведения групп людей данный подход годится только в качестве эталона с которым можно сравнить разрабатываемые модели поведения, поскольку он абстрагируется от движения отдельного человека и концентрируется на описании того как будет происходить процесс движения потоков при заданных заранее траекториях движения. Рассмотрим подходы, в которых присутствует сущность отдельного человека.

#### **1.4.2. Клеточные автоматы**

Клеточные автоматы – дискретные системы, то есть параметры, описывающие пространство и время, принимают значения из конечного набора значений. Клеткой называется узел пространственной решетки, ближайшие ячейки называются соседями. Каждому узлу присваивается некоторый набор значений, описывающий его текущее состояние, которое может изменяться по заданным правилам в зависимости от состояния соседей.

В случае описания движения людей, пространство (комната, здание) разбивается на набор ячеек, составляющих прямоугольную сетку, по размеру сопоставимых с размерами человека. Ячейки могут иметь разное состояние, описывающее состояние моделируемого объекта, например: свободная ячейка, препятствие (стена), ячейка занята человеком и т.п. Люди переходят из ячейки в ячейку основываясь на заданных правилах, одинаковых для всех ячеек. Например, зарубежные системы, основанные на клеточных автоматах: Egress [67] и Pedroute [92], Exodus [119]. Работ по моделированию эвакуации и пешеходов с помощью клеточных автоматов достаточно много (см. например обзор в [101] и [106, 107, 135]). Известно, что существующие модели на клеточных автоматах слабы в ситуациях с пересечением и встречным движением людских потоков. Кроме того, так как размер клеток и ассоциированные ограничения нуждаются в настройке при создании модели, результаты таких моделей сильно зависят от квалификации пользователя.

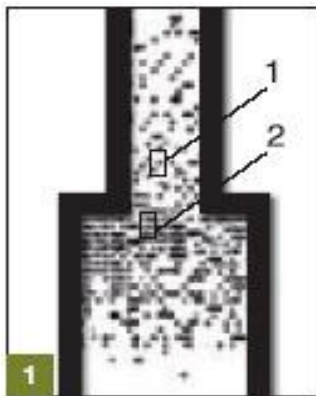
В нашей стране серьезно этот вопрос прорабатывал сотрудник физфака МГУ М.Е. Степанцов, который в 1997 году занялся созданием модели динамики толпы и опубликовал несколько статей по этой тематике. Его работа была поддержана Российским фондом фундаментальных исследований и ГНТП «Безопасность». Материалы по этому подходу можно найти в [55, 56]

Особое внимание Степанцов уделял проблеме практически неизбежной паники среди людей в случае возникновения чрезвычайных ситуаций в общественном месте. Главной задачей исследователя было смоделировать ситуацию с беспорядочным движением большой неорганизованной группы людей (то есть толпы), чтобы выявить те элементы и особенности конструкций, которые могут спровоцировать давку или «пробку». Степанцов полагает, что помимо общего потока пассажиров в проектировании городских пешеходных коммуникаций стоит также учитывать данные детального

мониторинга движения потоков на отдельных участках.

Наиболее уместным решением в данном случае является использование класса математических моделей, известных как «клеточные автоматы», которые активно применяются, например, в газодинамике. Степанцов считает, что моделирование ситуаций с активным участием людей на данном этапе научного прогресса практически невозможно, так как исследователю приходится иметь дело с критически большим количеством скрытых факторов. Однако по мере увеличения количества участвующих в модели людей роль рациональных и иррациональных факторов, описывающих поведение отдельного человека, снижается, и поведение группы может быть описано вероятностным образом, то есть спрогнозировано. Это характерная иллюстрация действия закона больших чисел. Даже если исследователь не учитывает причин, которые могут заставить индивидуума действовать нестандартным образом, его поведение вряд ли скажется на действиях общей группы.

В данном случае клеточный автомат имеет два состояния клетки (наличие/отсутствие в ней человека) и учитывает две составляющие движения: хаотичное и направленное. Интересно, что данные условия схожи с уже существующей моделью, описывающей диффузионные процессы (окрестность Марголуса). Степанцов модифицировал правила этой модели, добавив к диффузионной составляющей движения направленную и представив движение частицы (человека) как суперпозицию случайного и направленного перемещения.



*Рис.10 Выход из узкого перехода*

Далее была программно реализована работа клеточного автомата и исследованы некоторые модельные задачи. Так, на Рис.10 отражена ситуация, в которой люди выбегают из узкого выхода. При этом измеряется временная зависимость плотности числа людей до сужения и после. В данном случае плотность в широкой части прохода выше, а значит, данный выход может спровоцировать затор.

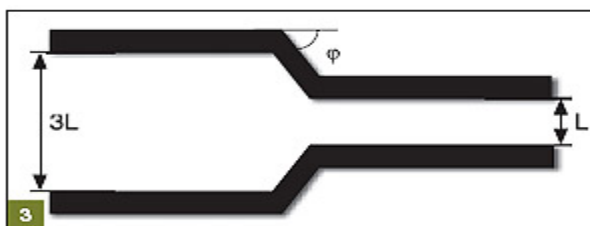


*Рис.11 Модифицированный проход*

Этого не произошло при изменении конфигурации прохода (Рис.11). Таким образом, модель позволяла описывать возникновения эффекта пробки и «отрицательной вязкости» (более быстрое движение у границ прохода).

Впоследствии Степанцов дважды вносил изменения в свою модель – в 1999 и 2003 гг. Главной целью доработки был учет решений человека, который стремится выбрать оптимальный путь к выходу, даже находясь в толпе. Раннюю версию исследователь называл моделью «слепых котят», так как выбор способа движения в заданном направлении отдавался на волю случая.

Для привнесения элементов анализа ситуации людьми выдвигается предположение, что каждый индивид старается передвигаться в нужном направлении и если видит непреодолимое препятствие или значительное скопление других людей на своем пути, то меняет траекторию движения в область с минимальными затруднениями. Для этого в модель вводится понятие «просмотра» на некое расстояние (добавляется параметр  $r$ ). Если в направлении движения на этом расстоянии встречается клетка с препятствием, то она и все лежащие за ней клетки считаются занятыми. Тогда человек начинает смещаться в сторону, стараясь придерживаться общего вектора (для этого просто нужно увеличить вероятность движения в направлении выхода).



*Рис.12 Оптимальное значение угла сужения*

Эффективность модели также была проверена на решении тестовых задач. Так, Степанцов попробовал оценить различные профили сужения коридора и выделить оптимальный угол, при котором длина сужения будет минимальной, однако достаточной для того, чтобы избежать возникновения заторов. После моделирования оказалось, что оптимальным значением угла является 45 градусов. В этом случае пробка будет возникать только при

плотности потока, приближающейся к максимально возможной (Рис.12).

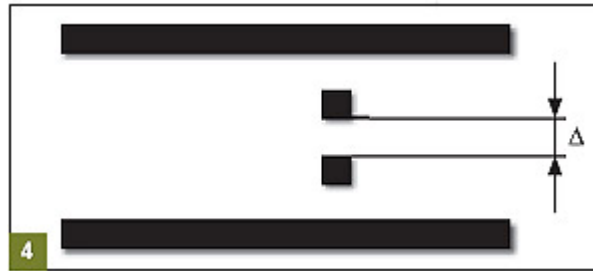


Рис.13 Оптимальное расположение двух конструкций

Другая задача состояла в том, чтобы определить наиболее удачное расположение двух конструкций, занимающих 40 % ширины прохода. Вместо угла в качестве искомого параметра фигурировало расстояние между этими конструкциями, а условием оптимизации было минимизирование времени прохождения людей. Оптимальным оказалось расположение, при котором конструкции делят проход на равные отрезки (Рис.13).

Степанцов рекомендовал эту модель в качестве одного из средств проверки безопасности перемещения толпы в условиях пространственных ограничений, то есть для небольших проблемных участков перемещения людей.

### 1.4.3. Силовой подход

Идея моделей этой группы заключается в применении второго закона Ньютона. Движение каждого человека происходит под действием нескольких сил. Это могут быть не только настоящие физические силы, но и некоторое психологическое взаимодействие. Главное, что каждая из этих сил заставляет человека двигаться с ускорением (одно из предположений модели).

Таким образом, человек является «частицей», на которую действуют разного рода силы: движущая сила, сила отталкивания между пешеходами, сила отталкивания от препятствий, временные аттракторы (витрины, информационные стенды и т.п.), стремление к объединению (очередь, огибание препятствий) и т.п. Далее составляются уравнения, описывающие движение каждого человека (2-й закон Ньютона), которые выглядят, например как в [94, 95]:

$$m_i \frac{dv_i}{dt} = m_i \frac{v_i^0(t)e_i^0(t) - v_i(t)}{\tau_i} + \sum_{j \neq i} f_{ij} + \sum_W f_{iW}$$

Поскольку такое уравнение движения имеет место для каждого человека ( $i = 1, 2, \dots, N$ ), мы на самом деле имеем дело с огромной системой связанных уравнений.

Что с этими уравнениями делать? Аналитически решить эту систему не удастся, поэтому приходится обращаться к численным методам. В результате мы получаем динамическую картину движения толпы, которую можно отобразить и графически.

В численном моделировании результатом можно называть явление,

устойчивое по отношению к небольшим изменениям параметров, то есть неизменно возникающее при различном наборе коэффициентов. При моделировании толпы, в данной работе и в других подобных удаётся получить эффекты, возникающие в реальности при выходе паникующей толпы.

Во-первых, был обнаружен переход от простого неупорядоченного движения к толпе и возникновению давки около выхода. Пока «желаемая» скорость покидания помещения была меньше 1,5 м/с, движение людей было более-менее организованным. При больших скоростях сила, «толкающая» человека к выходу, превышала взаимное психологическое отталкивание между людьми, в игру вступал непосредственный физический контакт, возникала давка, лавинообразно образовывалась толпа.

Затем, наблюдался эффект, названный авторами «чем быстрее, тем медленнее». Это значит, чем быстрее люди хотят покинуть комнату (т.е. чем выше «желаемая» скорость), тем медленнее толпа просачивалась через дверь: люди, находящиеся непосредственно у выхода, мешали друг другу.

В-третьих, когда «желаемая» скорость возрастала еще больше и сила взаимодействия людей превышала критическую, давление в толпе стало приводить к травмам отдельных людей. В рамках этой модели, травмированный человек превращался в неподвижное препятствие, это приводило к еще большим заторам, и в результате количество человек, успевших покинуть комнату за определенное время, резко уменьшалось.

Кроме того, в зависимости от конкретной ситуации (пожар, задымленная комната, комната непрямоугольной формы) были получены и другие предсказания.

Этот подход также применим для теоретических исследований скорости/времени движения людей на определённых небольших участках эвакуации, а также для практических исследований движения людей по заранее заданным траекториям.

Собственно чисто клеточный и силовой подходы пытаются построить модель только для обхода препятствий людьми. Вопросами стремления к цели, целеполагания не занимаются, поэтому могут быть использованы как уровень поведения агента, отвечающий за его тактическое поведение: обход препятствий, движение по заданной траектории.

Работы в этом направлении в основном ведут зарубежные учёные, например, уже упоминавшийся Dirk Helbing с группой учёных [94], Taras Lakoba [104]. Результаты работы группы Helbing-а можно найти на сайте [121]. Существует похожий подход, но основанный на магнитных силах, например модель Shigeoyuki Okazaki [118]. Ещё одна модель, использующая физическую аналогию – модель диффузии газа (см. напр. [99]), стандартно применяемая для нахождения путей роботами [127].

#### **1.4.4. Агентный подход**

Применение агентного подхода логично для теоретического исследования поведения людей при эвакуации или вообще движении: описать такое поведение агентов, которое приводит к наблюдаемому

поведению людей, затем исследовать поведение людей в нестандартных ситуациях или влияние присутствия агентов с нестандартным поведением на время эвакуации. К тому же в агентном подходе можно задавать глобальные цели агента, стратегию достижения цели, тактику, а также использовать наработки МАС.

Можно заметить, что многие из работ, использующие подход клеточных автоматов или силовых алгоритмов называют свой подход агентным ([69, 80, 135]), в какой-то степени это верно, так как используется «микроскопический» масштаб моделирования – каждый человек по отдельности. Плюс добавляются алгоритмы поиска глобального пути.

Также в рамках агентного подхода могут использоваться реактивные модели, реализующие силовые алгоритмы или различные эвристические алгоритмы поведения человека и групп людей в эвакуации [81, 103, 112]. Но есть разработки по использованию гибридных агентных моделей: [114, 120, 122, 128].

Например, в концепции, описанной в [120], поведение толпы – комплексный феномен, который в рамках этой работы раскрывается тремя уровнями поведения:

1. Уровень индивидуума
  - a. Следование инстинктам
  - b. Следование опыту
  - c. Рационализм
2. Взаимодействия между индивидуумами
  - a. Социальная идентификация (понятие ролей индивида в социуме)
  - b. Персональное пространство
  - c. Социальная коррекция (когда человек не знает, что делать и доверяется движению толпы, например, сюда относится «стадное поведение»)
3. Группа
  - a. Плотность толпы
  - b. Окружающие ограничения (прежде всего геометрия здания)
  - c. Восприятие эмоций и напряжённости

Конечно, данные уровни зависимы и могут перекрывать друг друга.

Подобная иерархия понятий – толпа, группа, индивид также используется в работе [114].

Рассмотрим отдельно подход, предложенный в [128]. Поведение человека есть результат действия трёх блоков: восприятие, поведенческое управление, разумное управление.

Восприятие включает в себя: определение уровня земли, распознавание статичных объектов, распознавание подвижных объектов, расположение объектов на карте, интерпретация сложных ситуаций.

Поведенческое управление включает в себя три уровня: базовое реактивное поведение, навигация (выбор пути), мотивационное поведение.

На блок разумного управления попадает знание о виртуальном мире, планирование, выбор действий в зависимости от мотиваций, память и стыковка поведенческого и разумного уровней.

Отдельно стоит отметить проработанный блок восприятия и последовательно структурированное базовое реактивное поведение.

Предложенная гибридная архитектура хорошо проработана, сильно нацелена на моделирование именно человека, делиберативный уровень связан с реактивным (блок разумного управления должен контролировать взаимодействие с блоком поведенческого управления). Есть мотивации агента, но нет явно представленных целей. Между блоками поведения нет заранее определённых потоков информации, в общем модель недостаточно хорошо структурирована, но общие идеи хороши и известны. Ещё один практический «минус» в том, что данная работа опирается на платные закрытые разработки по трёхмерному моделированию движения людей.

Основная проблема агентных моделей эвакуации – узкая их специализация на эвакуацию, при которой поведение агента если и можно изменить, то только в пределах параметров модели. Запрограммировать своё поведение сложно в виду этой специализации или закрытости платного программного обеспечения (ПО). Таким образом, агентные модели эвакуации не используют наработки теории МАС, а общие подходы МАС обычно слишком далеки от приложений эвакуации.

По итогам главы 1 можно сделать следующие выводы. Во-первых, на базе анализа известных результатов исследования рефлексивных игр и моделей информационного управления можно построить модель рефлексивного разбиения – осуществить переход от общего описания информированности к модели информированности групп агентов с определённым рангом (см. главу 2).

Во-вторых, теория коллективного поведения и теория команд позволяют классифицировать модели рефлексивного разбиения, а также использовать полученные в этих разделах модели как базу для дальнейшего исследования. В этих моделях можно добавить в принципы принятия агентами решений рефлексивную информированность и рефлексивные размышления.

В-третьих, по результатам обзора по мультиагентному моделированию следует отметить, что в настоящее время существует возможность реализовать теоретические модели рефлексивного коллективного поведения в современном имитационном подходе – гибридных агентных системах. Например, описание реактивной имитационной модели [76, 124] походит на теоретико-игровое описание моделей коллективного поведения, а гибридность модели подразумевает возможность добавления логического вывода в архитектуру агентов.

В-четвёртых, обзор по моделям эвакуации показал, что для обеспечения безопасности и минимизации времени эвакуации для конкретного здания с

заранее определёнными путями эвакуации в настоящее время хорошо подходят потоковые модели, основанные на реальных измерениях движения людских потоков. Существующие агентные модели эвакуации используют разные эмпирические предположения о поведении людей и не рассчитаны на определение совершенно разных моделей поведения. Кроме того, результаты агентных моделей пока неточны – время эвакуации не равно реальному времени. Но существует теоретическая и практическая тенденция для общего агентного моделирования эвакуации, при котором можно исследовать разные модели поведения, в том числе тактические (физика движения групп людей) и стратегические (каким путём идти). Последнее актуально для реализации моделей рефлексивного поведения.

Перейдём далее к описанию и исследованию свойств общей модели рефлексивного поведения.

## **Глава 2. Метод рефлексивных разбиений в задачах коллективного поведения и управления**

В данной главе вводится модель рефлексивного поведения агентов. Агенты разбиваются на группы по уровню рефлексии. Устанавливаемое после действий агентов состояние системы зависит от количества (доли) агентов в каждой группе. Рассматривается возможность управления состоянием с помощью варьирования распределения агентов по уровням рефлексии, анализируется ряд прикладных моделей. Исследуются свойства оптимальных управлений, приводится обсуждение полученных результатов.

Классическая теория игр [14, 113] занимается моделями совместного принятия решений рациональными субъектами, интересы которых не совпадают, а в качестве решения игры предлагает различные концепции равновесия [14, 57, 111, 113]. Распространенной является концепция так называемого равновесия Нэша. Равновесие Нэша – это такая ситуация, от которой никому из участников игры не выгодно отклоняться в одностороннем порядке. Но само по себе равновесие является лишь математической абстракцией – оно не имеет смысла без указания механизма его реализации посредством действий агентов. В настоящее время существует обширная литература по механизмам реализации концепций равновесия (см. обзор в [96]).

Однако давно известно, что реальные действия экономических агентов зачастую не объясняются ни равновесием Нэша, ни даже стремлением к максимизации функции полезности [97]. Так, в реальности (например, на финансовых рынках) далеко не всегда наблюдается сходимости установившейся на рынке цены к фундаментальной, справедливой цене даже в условиях, близких к свободной конкуренции (при большом количестве участников рынка). В рамках концепции ограниченной рациональности агентов [130] - в условиях недостатка времени или большого объёма информации (большого числа агентов) - равновесие Нэша заменяется более простыми алгоритмами поведения.

Рефлексивное поведение является одной из концепций ограниченной рациональности. Рефлексия с этой точки зрения понимается как попытка агента предсказать информированность и действия своих оппонентов с целью устранения игровой неопределённости [43] при принятии им решения и выборе наилучшего ответа на их действия. С точки зрения рефлексивного поведения равновесие Нэша демонстрирует бесконечную глубину рефлексии, поскольку является предельной точкой отображения наилучших ответов [14]. Ограничение ранга рефлексии при принятии решений экономическими агентами является одним из проявлений ограниченной рациональности.

Ограниченная рациональность порождает проблему неэффективности. Тогда как рыночное равновесие эффективно по Парето (что традиционно служит одним из доводов в пользу свободной конкуренции), рефлексивные

субъекты в силу неадекватной информированности о механизмах ценообразования и своего иррационального поведения могут принимать решения, приводящие к неэффективным исходам [91]. Следовательно, на рынке, где действуют агенты с ограниченной рациональностью, возникает потребность в регулировании. Поскольку причина неэффективности лежит в неадекватной информированности и ограниченной рациональности агентов, адекватным методом рыночного регулирования представляется т.н. информационное управление [43, 65].

## **2.1. Модель игровой ситуации**

В данном разделе определим структуру агента и задаваемое такой структурой разбиение множества агентов. Затем описывается механизм выбора агентом действия при заданной структуре. Это общее описание механизма игры, взаимодействия участников. В следующих разделах предлагается модель рефлексивного поведения, которая позволяет агенту определить обстановку в игре (какие действия выберут остальные агенты) и выбрать своё действие.

Базой модели является игровая ситуация, когда целевая функция каждого агента зависит от агрегированной информации по действиям всех агентов. Задача состоит в исследовании зависимости результата игры от определённой информированности агентов, а с точки зрения управления, Центру нужно достичь определённого результата игры, изменяя структуру информированности агентов. В модели агент может вообще не учитывать влияние других игроков (функция агрегирования – константа). Такая модель игры более общая, чем в [43, 65], но предлагается модель рефлексивного поведения, которая задаёт более узкую структуру информированности, чем в [43, 65].

Начнём с общего описания игры.

### **2.1.1. Описание игры**

Пусть имеется некоторая система, состояние которой описывается вектором  $\chi \in X \subset R^p$ ,  $p \geq 1$ .

Дано множество агентов  $N = \{1, \dots, n\}$ ,  $n \in N$ , как правило,  $n \gg 1$ . Агенты могут выбирать различные действия  $a \in \mathfrak{R}^1$ , для простоты будем считать, если не оговорено особо, что ограничения на действия агентов отсутствуют. Вектор  $a = (a_1, a_2, \dots, a_n) \in \mathfrak{R}^n$  действий агентов, называется ситуацией игры [14].

Взаимосвязь системы и агентов описывается следующим образом. От действий агентов (ситуации игры) зависит состояние системы по закону агрегирования:  $\chi = \chi(a) = \chi(a_1, \dots, a_n)$ , а агенты выбирают действия, основываясь на своей информации о состоянии системы и ситуации игры.

Действия агенты выбирают одновременно и независимо на конечном количестве шагов игры.

Выбор действия осуществляется рационально, что поясняет

следующий параграф.

### 2.1.2. Рациональное поведение агентов

Состояние системы  $\chi$ , зависящее от действий агентов, определяет их выигрыши, задаваемые целевыми функциями  $f_i(y, \chi)$ , где  $f_i(\cdot): \mathfrak{R}^1 \times X \rightarrow \mathfrak{R}^1$ .

Рациональность поведения  $i$ -го агента заключается в стремлении к максимизации своей целевой функции выбором собственного действия [14]:

$$a_i \in BR_i(\chi(a)) = \text{Arg max}_{y \in \mathfrak{R}} f_i(y, \chi(y, a_{-i})),$$

где  $a_{-i} = (a_1, a_2, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$  называется *обстановкой игры*,  $BR_i(\cdot)$  называется *наилучшим ответом* (best response) [113], в терминах теории игр это наилучший ответ на обстановку игры. Предположим, что функция  $f(\cdot)$  такова, что для любого агента при любом состоянии системы существует единственный наилучший ответ. Через  $F_i$  обозначим целевую функцию агента  $i$  от ситуации игры:  $F_i(a) = f_i(a_i, \chi(a))$ .

Для обозначения наилучшего ответа допускаются следующие записи:

$$(1) a_i \in BR_i(a_{-i}) = BR_i(\chi(a_i, a_{-i})).$$

**Теория игр.** Поскольку наилучший ответ каждого агента зависит в общем случае от обстановки игры, трудно однозначно сказать априори, какое действие выберет конкретный агент. Основным предметом *теории игр* – см. [6, 14, 43, 113] – 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 [14]. 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* [14]. Если считать, что все описание игры (множество агентов, их целевые функции и множества допустимых действий) является *общим знанием* [43] среди агентов, принимающих решения однократно, одновременно и независимо, то можно использовать концепцию *равновесия Нэша*, в рамках которой равновесием называется такой вектор действий  $a^N$ , что действие каждого агента является наилучшим ответом на Нэшевскую обстановку:

$$(2) a_k^N \in BR_k(a_{-k}^N), k \in N.$$

**Коллективное поведение.** В отличие от теории игр, *теория коллективного (группового) поведения* – см. [6, 32, 45] – 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 (см. выше).

Более того, зачастую агенты, принимая решения, могут «не знать о существовании» других агентов или иметь о них агрегированную информацию. Поэтому ниже 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 является *модель индикаторного поведения* [32, 45], суть которой заключается в следующем. Предположим, что каждый агент в момент времени  $t$  наблюдает действия всех агентов  $\{a_i^{t-1}\}_{i \in N}$ , выбранные в предыдущий момент времени  $t-1$ ,  $t=1, 2, \dots$  (начальный вектор действий  $a^0 = (a_1^0, a_2^0, \dots, a_n^0)$  считается заданным).

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

$$(3) \quad w_i(a_{-i}^{t-1}) = BR_i(a_{-i}^{t-1}), i \in N.$$

Если более подробно то, поскольку агент имеет агрегированную информацию  $\chi_i(a^{t-1})$  о действиях всех агентов, выражение (3) примет вид  $w_i(a_{-i}^{t-1}) = \arg \max_{y \in \mathbb{R}^1} f_i(y, \chi(y, a_{-i}^{t-1}))$ ,  $i \in N$ .

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

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

где  $\gamma_i^t \in [0; 1]$  – «величины 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, однако в последнем случае соответствующая динамика может быть неустойчивой. Условия сходимости процедуры (4), области притяжения равновесий, условия на величины шагов  $\{\gamma_i^t\}$ , обеспечивающие сходимость, и т. д. можно найти в [32, 45].

Подходы теории коллективного поведения и теории игр согласованы в том смысле, что и та, и другая исследуют поведение рациональных агентов (ср. 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 рефлексия [43]. *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*.

Стратегическая рефлексия агента описывает его информированность о принципах принятия решений (действий) остальными агентами. Ограниченность информированности может проявляться в ранге рефлексии: агент  $k$ -го ранга считает, что ранг всех остальных агентов может быть только из множества меньших рангов:  $\{0, 1, \dots, (k-1)\}$ . При этом равновесие Нэша «превращается» в более общее *информационное равновесие*, в рамках которого каждый агент осуществляет не информационную, а стратегическую рефлексия – при принятии решений использует не только информацию о своих принципах принятия решения, но и свои представления о представлениях других агентов об их принципах, представления о представлениях о представлениях и т. д. Модель индикаторного поведения можно рассматривать как стратегическую рефлексия: каждый агент при выборе цели считает, что остальные агенты не изменят своего решения (но двигается к цели с некоторым шагом, всё-таки опасаясь изменения).

## 2.2. 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 стратегическую рефлексия), 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 (с учётом упрощения (1)):

$$(5) w_i(a_{-i}^t) = BR_i(a_{-i}^t),$$

где  $a_{-i}^t$  определяется выражением (4). При этом будем полагать, что рефлексиярующий агент первого ранга считает всех остальных нерелфлексияющими (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 [43]).

Аналогично можно рассматривать агентов и более высоких рангов рефлексии. Для этого определим рефлексивные разбиения.

Рефлексивным разбиением  $\aleph$  множества агентов  $N$ , назовём следующее множество подмножеств  $N$ :

$$(6) \aleph = \{ {}^0N, {}^1N, \dots, {}^mN \},$$

где  ${}^kN$  – множество агентов  $k$ -го ранга рефлексии,  $k = \overline{0, m}$ ,  $m$  – максимальный ранг рефлексии.

По аналогии с вектором действий всех агентов, введём вектор действий агентов  $k$ -го ранга в момент времени  $t$ :

$${}^k a^t = \left( {}^k a_j^t \right)_{j \in {}^k N}.$$

Каждый  $i$ -й агент  $k$ -го ранга рефлексии имеет своё представление о рефлексивном разбиении множества агентов – субъективное рефлексивное разбиение:

$$(7) {}^k \mathfrak{N} = \{ {}^{k0}N, {}^{k1}N, \dots, {}^{k(k-1)}N, \{i\}, 0, \dots, 0 \}.$$

Из вида  ${}^{kk}N$  понятно, что агент  $i$  считает, что только он имеет ранг  $k$ .

Определим также количество агентов ранга  $h$ :  $n_h = |{}^h N|$ ,  $k$ -субъективное количество агентов ранга  $h$ :  $n_{kh} = |{}^{kh} N|$ .

В некоторых случаях удобно вместо количества использовать доли  $q_{kh}$  – меры множеств  ${}^{kh} N$ : при большом количестве агентов и для удобства анализа модели. Тогда соответственно  $0 \leq q_k, q_{kh} \leq 1$  и  $q_{kh} = {}^{kh} n/n$  (в случае конечного множества агентов),  $\sum_{h=0}^{\infty} q_h = \sum_{h=0}^{\infty} q_{kh} = 1, \forall k$ .

Примем гипотезу **1** (адекватность рефлексии): Агент ранга  $k$  приписывает агенту ранга  $l < k$  его истинный ранг  $l$ .

Гипотеза 1 говорит о том, что представление агента ранга  $k$  о всех агентах более низких рангов истинно, при этом агент может неправильно оценивать только ранг агентов  $k$ -го,  $(k+1)$ -го и более высоких рангов рефлексии и соответственно добавлять их к множествам агентов нижнего ранга. Из чего следует:  $\forall l < k : {}^l N \subseteq {}^k N$ .

Этим отражается предположение, что агент не допускает существования агентов, имеющих такой же или более высокий ранг рефлексии, чем он сам.

Следствием гипотезы 1 является истинность следующих неравенств:

$$(8) \quad \forall h < k: q_{kh} \geq q_h \ \& \ n_{kh} \geq n_h$$

### 2.3. Рефлексивная модель коллективного поведения

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

**Нулевой ранг рефлексии или «обычные» агенты.** Будем считать, что «обычные» агенты (принадлежащие множеству  ${}^0 N$ ) выбирают свои действия в соответствии с гипотезой индикаторного поведения. Из (4) следует, что

$$(9) \quad {}^0 a_i^t = {}^0 a_i^{t-1} + \gamma_i^t [w_i(a_{-i}^{t-1}) - {}^0 a_i^{t-1}], i \in {}^0 N, t = 1, 2, \dots$$

Если рефлексизирующих агентов нет ( ${}^0 N = N$ ), то в итоге все агенты пронаблюдает реализованную траекторию  $(a^0, \dots, a^t, \dots)$  векторов действий агентов, определяемых выражением (9).

Отметим, что «обычные» агенты могут опираться только на информацию о предыдущих действиях агентов. При этом если функция состояния системы  $k$ -го «обычного» агента не зависит от действий остальных агентов, то его цель  $w_i$  не зависит от времени, а при  $\gamma_k^t \equiv 1$  и его

действия не зависят от времени. Таким способом можно моделировать, например, агентов, выбирающих фиксированное действие на каждом шаге.

Вообще говоря, за «обычных» агентов можно брать агентов, поведение которых не опирается на рефлексию, например: вероятностное поведение, выбор среднего от предыдущих действий, выбор наилучшей стратегии в предыдущей игре, то есть действия таких агентов опираются максимум только на фактическую информацию о действиях других агентов в прошлом периоде.

**Первый ранг рефлексии.** Агент  $j$ , обладающий первым рангом рефлексии ( $j \in {}^1N$ ), считает всех остальных агентов обладающими нулевым рангом рефлексии и в соответствии с выражением (9) «предсказывает» их выбор. Поэтому его собственный выбор  ${}^1a_j^t$  будет ориентирован на наилучший ответ на ту обстановку, которая с его точки зрения должна сложиться:

$$(10) \quad {}^1a_j^t = {}^1a_j^{t-1} + \gamma_j^t [w_j({}^0a_{-j}^t) - {}^1a_j^{t-1}], j \in {}^1N.$$

Для агента  $j \in {}^1N$  прогнозируемой является траектория  $(a^0, \dots, ({}^1a_j^t, {}^0a_{-j}^t), \dots)$ , а на самом деле реализуется траектория  $(a^0, \dots, ({}^1a^t, {}^0a^t), \dots)$ . То есть реализованная траектория может не совпадать с траекториями, прогнозируемыми как агентами нулевого, так и первого рангов рефлексии. О возможном несовпадении прогнозируемой и реализованной траекторий (и последствиях такого несовпадения) речь пойдет ниже – см. «условие стабильности рефлексивного разбиения».

**Второй ранг рефлексии.** Будем считать, что каждый агент  $j$ , обладающий вторым рангом рефлексии ( $j \in {}^2N$ ), имеет субъективное рефлексивное разбиение  ${}^2\aleph$  и знает как себя ведут агенты нулевого и первого рангов рефлексии (формулы (9) и (10)). 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:

$$(11) \quad {}^2a_j^t = {}^2a_j^{t-1} + \gamma_j^t [w_j({}^{20}a^t, {}^{21}a^t) - {}^2a_j^{t-1}], j \in {}^2N,$$

где  ${}^{kl}a^t = \{ {}^l a_i^t \}_{i \in {}^{kl}N}$ , т.е. это действия реальных агентов ранга  $l$ , какие им приписывает агент  $k$ -го ранга рефлексии, в соответствии со своим субъективным рефлексивным разбиением.

Агенты нулевого и первого рангов рефлексии будут вести себя в соответствии с выражениями (9) и (10) соответственно.

Для агента  $j \in {}^2N$  прогнозируемой является траектория  $(a^0, \dots, ({}^2a_j^t, {}^{21}a^t, {}^{20}a^t), \dots)$ , а на самом деле реализуется траектория  $(a^0, \dots, ({}^2a_j^t, {}^1a^t, {}^0a^t), \dots)$ .

**$k$ -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 субъективных рефлексивных разбиений:

$${}^k a_j^t = {}^k a_j^{t-1} + \gamma_j^t [w_j({}^{k0} a^t, {}^{k1} a^t, \dots, {}^{k(k-1)} a^t) - {}^k a_j^{t-1}]$$

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

Рефлексивную модель коллективного поведения совместно с рефлексивным разбиением назовём **моделью рефлексивных разбиений** (МРР).

Вектор действий агентов зависит от представлений агентов друг о друге (субъективных рефлексивных разбиений) и от рефлексивного разбиения  $\aleph$ . Если зафиксировать субъективные разбиения то вектор действий агентов  $a$  однозначно задаётся рефлексивным разбиением  $\aleph$ :

$$(12) \quad a^t(\aleph) = ({}^0 a^t, {}^1 a^t, \dots, {}^m a^t)$$

Следовательно, изменяя рефлексивное разбиение, можно менять действия агентов, то есть, осуществлять *рефлексивное управление* (см. также теоретико-игровые модели рефлексивного управления в [42, 43, 65]).

**Рефлексивным равновесием** игры  $\Gamma_\Sigma = \{N, F_i(\cdot)_{i \in N}, \aleph\}$  назовём вектор действий агентов

$$(13) \quad a^*(\aleph) = ({}^0 a^*, {}^1 a^*, \dots, {}^m a^*).$$

Отметим, что с точки зрения агента ранга  $k$ , реализуется следующее рефлексивное равновесие:

$$(14) \quad a^*({}^k \aleph) = ({}^{k0} a^*, {}^{k1} a^*, \dots, {}^{kk} a^*).$$

Предложенная модель эквивалентна модели, предложенной в [75] и названной «*когнитивная иерархия*» (*Cognitive Hierarchy*, CH). С.Ф. Camerer и его коллеги вводят сначала общую модель похожую на МРР, но сразу же переходят к конкретной модели и исследуют её адекватность на результатах игры *p-beauty contest*. Предложенные ими «обычные» агенты (0-step players) выбирают действия случайно, рефлексивное разбиение представляет собой однопараметрическое распределение Пуассона, параметр  $\tau$  есть среднее значение распределения. Соответственно в рефлексивном разбиении используются доли, а не количества. Агент ранга  $k$  ( $k$ -step players) распределяет всех агентов своего ранга и выше по рангам  $0, \dots, k-1$  пропорционально доли агентов соответствующего ранга:  ${}^k \aleph = \{{}^{kh} N\}_{h=0}^\infty$ ,  $\{q_h\}_{h=0}^\infty$  – дискретное распределение Пуассона:

$$q_{kh} = \begin{cases} 0 & , h > k \\ 1/n & , h = k \\ q_h / \sum_{j=0}^{k-1} q_j, & h = 0, \dots, k-1 \end{cases} .$$

Такую модель называют Poisson-CH.

Отличия подходов состоят в направлении исследований. Зарубежные исследователи пока не ставят задачи управления системой, а пытаются найти такую модель рационального поведения, которая предсказывала бы поведения людей в реальных экспериментах. Поэтому сначала идентифицируют данную модель. Для одношаговых игр она показала пригодность для анализа и лучшие результаты по предсказанию, чем равновесие Нэша. Авторы отмечают, что нужно дополнительно исследовать её применимость для многошаговых игр. Также есть небольшое отличие в описании модели игры, целевая функция явно зависит от действий остальных игроков – обстановке, тогда как в МРР целевая функция зависит от агрегированной информации об обстановке игры.

В статьях, связанных с развитием данной модели, можно отметить проведённые практические экономические эксперименты: со студентами, с читателями газет и журналов, с читателями интернет изданий. Зарубежные исследователи активно исследуют применимость частных моделей МРР и некоторых других на практике, описанные ими эксперименты можно использовать и для проверки МРР.

Например, в работе [115] «обычные» агенты выбирают действие произвольно. А агент ранга  $k$  имеет следующее субъективное рефлексивное разбиение –  ${}^k\mathfrak{N} = \{0, 0, \dots, N, 1, 0, \dots, 0\}$ , где единица стоит на  $k$ -ом месте, то есть он считает ранг всех остальных агентов равным  $(k-1)$ . Далее эксперименты в [72] показали, что 64% участвовавших аргументировали своё действие используя эту модель, причём встречались только модели с  $k$  равным 0,1,2,3 и  $\infty$ .

В работах [24, 132] исследована статическая игра, «обычные» агенты выбирают фиксированное действие. Субъективное рефлексивное разбиение агента  $i$  следующее:  ${}^k\mathfrak{N} = \{{}^{k0}N, {}^{k1}N, \dots, {}^{k(k-1)}N \cup {}^{kk}N \setminus \{i\} \cup {}^{k(k+1)}N \cup \dots \cup {}^{km}N, 1, 0, \dots, 0\}$ .

Работа [29] предлагает ещё одно из возможных разбиений:  ${}^k\mathfrak{N} = \{{}^{k0}N \cup {}^{kk}N \setminus \{i\} \cup {}^{k(k+1)}N \cup \dots \cup {}^{km}N, {}^{k1}N, \dots, {}^{k(k-1)}N, 1, 0, \dots, 0\}$ .

Рассмотрим случай, когда множество однородных агентов может быть бесконечно. Это означает, во-первых, что агенты одного ранга действуют одинаково, во-вторых, адекватно будет использовать понятие доли (меры множества) агентов одного ранга, а не их количество. Тогда под субъективным рефлексивным разбиением  $\sum_{jk}$  достаточно понимать вектор  $q^k = (q_{k0}, q_{k1}, \dots, q_{kk})$ . Реальное же разбиение есть бесконечная последовательность  $q = (q_0, q_1, \dots)$ . Рассмотрим банахово пространство  $l_1$  бесконечных последовательностей с нормой  $\|x\| = \sum_{i=0}^{\infty} |x_i|$ . Поскольку

$$(15) \quad \sum_{i=0}^{\infty} |q_i| = \sum_{i=0}^{\infty} |q_{ki}| = 1,$$

то  $q^k, q \in l_1$  ( $q^k$  дополняем нулями до бесконечной последовательности). Тогда, с учётом выражения (8),  $\|q^k - q\| = \sum_{i=0}^{\infty} |q_{ki} - q_i| = \sum_{i=0}^{k-1} (q_{ki} - q_i) + \sum_{i=k}^{\infty} q_i$

$$= \sum_{i=0}^{k-1} q_{ki} + \sum_{i=0}^{\infty} q_i - 2 \sum_{i=0}^{k-1} q_i = 2 - 2 \sum_{i=0}^{k-1} q_i \rightarrow 0.$$

Обозначим через  ${}^k A$  вектор действий всех оппонентов с точки зрения агента ранга  $k$ , обладающего субъективным рефлексивным разбиением  $q^k$ :  ${}^k A = ({}^{k0} a, {}^{k1} a, \dots, {}^{k(k-1)} a)$ , с долей агентов, обладающих рангом  $h$ :  $q_{kh}$ . По гипотезе 1 имеем (8), а из (8) и (15) следует, что с ростом  $k$ :  $q_{kh} \rightarrow q_h$ . Следовательно,  ${}^k A$  покоординатно (слабо) сходится по  $k$  к вектору истинных действий агентов  $a$ :  ${}^{kh} a \rightarrow {}^h a$ .

Предположим, что  $j$ -й агент может произвольно выбирать свой ранг рефлексии –  $k$ , на заданном шаге игры  $t$ . После выбора ранга, его действие определяется в соответствии с  $q^k$ .

Ясно, что в реальности для повышения своего ранга нужно провести ряд вычислений, количество которых растёт с ростом ранга – это будут затраты агента на выбор ранга. Определим неотрицательную, возрастающую функцию затрат  $j$ -го агента на выбор ранга  $k$  –  $G_j(k)$ ,  $G_j(0)=0$ .

Определим ожидаемый выигрыш агента с учётом затрат:

$$F_j(k) = f_j({}^k a_j, \chi({}^k a_j, {}^k A)) - G_j(k)$$

Тогда справедливо следующее утверждение, сформулированное словесно и также обоснованное в [75]:

**Утверждение 1.** Если  $\chi(a)$  вполне непрерывна,  $f_j(y, \chi)$  непрерывна по паре аргументов  $(y, \chi)$ ,  $G_j(k)$  – строго возрастающая, неограниченная функция, тогда  $j$ -му агенту невыгодно бесконечно увеличивать свой ранг рефлексии при фиксированных рангах оппонентов.

**Доказательство.** Для доказательства в случае бесконечного количества агентов нужно показать, что функции  $\{F_j(k)\}$  убывают с некоторого  $k_0 \in \mathbb{N}$ . Для этого достаточно показать, что последовательность  $\{f_j^k = f_j({}^k a_j, \chi({}^k a_j, {}^k A))\}_k$ , есть ограниченная последовательность, что будет верно, если это сходящаяся последовательность при  $k \rightarrow \infty$ .

По определению,  ${}^k a_j$  доставляет максимум функции  $f_j(y, \chi(y, {}^k A))$  по  $y$ , то есть  $f_j({}^k a_j, \chi({}^k a_j, {}^k A)) = \max_{y \in \mathbb{R}} f_j(y, \chi(y, {}^k A))$ . Из непрерывности функции  $f_j(y, \chi)$  по паре аргументов  $(y, \chi)$ , по известной теореме сравнительной статики следует:

$$\max_{y \in \mathbb{R}} f_j(y, x_k) \rightarrow \max_{y \in \mathbb{R}} f_j(y, x),$$

при  $k \rightarrow \infty$ , для любой последовательности  $\{x_k\}$ :  $x_k \rightarrow x$ .

А из того, что функция  $\chi()$  вполне непрерывна, следует:

$$\chi(y, {}^k A) \rightarrow \chi(y, A),$$

при  $k \rightarrow \infty$ .

Отсюда получаем:

$$f_j^k \rightarrow \max_{y \in \mathbb{R}} f_j(y, \chi(y, A)),$$

при  $k \rightarrow \infty$ , что и требовалось доказать.

Если же количество агентов конечно и ранги оппонентов агента  $j$  фиксированы на данном шаге  $t$ , тогда существует максимальный ранг рефлексии  $k_0 \in \mathbb{N}$ , для которого  $\aleph = {}^{(k+1)} \aleph$ . Получаем, что  $\forall k > k_0$  :

$F_j(k) = -G(k)$  – убывающая функция, то есть агенту  $j$  невыгодно увеличивать далее свой ранг.

Пример №1. Пусть имеется  $n$  агентов, выходящих из помещения в два выхода. Действие агента – выбор выхода:  $a_i \in \{0, 1\}$  и состояние системы – двумерный вектор – количество выходящих в первый и во второй выходы соответственно:  $x = (u_0 = \sum(1-a_i), u_1 = \sum a_i)$ . Агенты минимизируют целевую функцию – время своего выхода:  $T(a_i, x = (u_0, u_1)) = u_{a_i} \rightarrow \min$ .

Пусть имеются «обычные» агенты, всегда выбирающие выход 0, то есть, например, их функция состояния системы  $\chi_i(\cdot) = (0, n)$ . Агенты первого ранга рефлексии считают всех остальных «обычными», тогда они считают, что состояние системы  $x_1(a) = (n-a, a)$ , и, минимизируя время выхода, выбирают  $a = 1$ .

Если  $n_0, n_1$  – соответственно количества «обычных» агентов и агентов первого ранга. То состояние системы при отсутствии других агентов:  $x = (n_0, n_1)$ .

То есть за счёт выбора долей агентов  $n_0 = n_1 = n/2$  можно "уравновесить" время выхода.

Добавим  $n_2$  агентов второго ранга, пусть в случае (а) они относят агентов своего ранга к «обычным», в случае (б) к агентам первого ранга, а в случае (в) половину считает «обычными», а другую половину агентами первого ранга. Тогда состояние системы для них равно:

- (а)  $x_2(a) = (n_0 + n_2 - a, n_1 + a)$ ,
- (б)  $x_2(a) = (n_0 - a, n_1 + n_2 + a)$ ,
- (в)  $x_2(a) = (n_0 + n_2/2 - a, n_1 + n_2/2 + a)$ ,

при условии, что  $n = n_0 + n_1 + n_2$ .

Каждый агент второго ранга выберет выход-0, если

- (а)  $n_0 < n_1 - n_2$ .
- (б)  $n_0 < n_1 + n_2$ .
- (в)  $n_0 < n_1$ .

Следовательно, состояние системы можно вычислить по формуле:

- (а)  $x = \begin{cases} (n_0 + n_2, n_1), n_0 \leq n_1 - n_2 \\ (n_0, n_1 + n_2), n_0 > n_1 - n_2 \end{cases}$ ,
- (б)  $x = \begin{cases} (n_0 + n_2, n_1), n_0 \leq n_1 + n_2 \\ (n_0, n_1 + n_2), n_0 > n_1 + n_2 \end{cases}$
- (в)  $x = \begin{cases} (n_0 + n_2, n_1), n_0 \leq n_1 \\ (n_0, n_1 + n_2), n_0 > n_1 \end{cases}$

Графики времени выхода агентов в выход-0 и выход-1 представлены на Рис. 14. График времени выхода агента второго ранга обозначен зелёной точечной линией на Рис. 14 в случае (а).

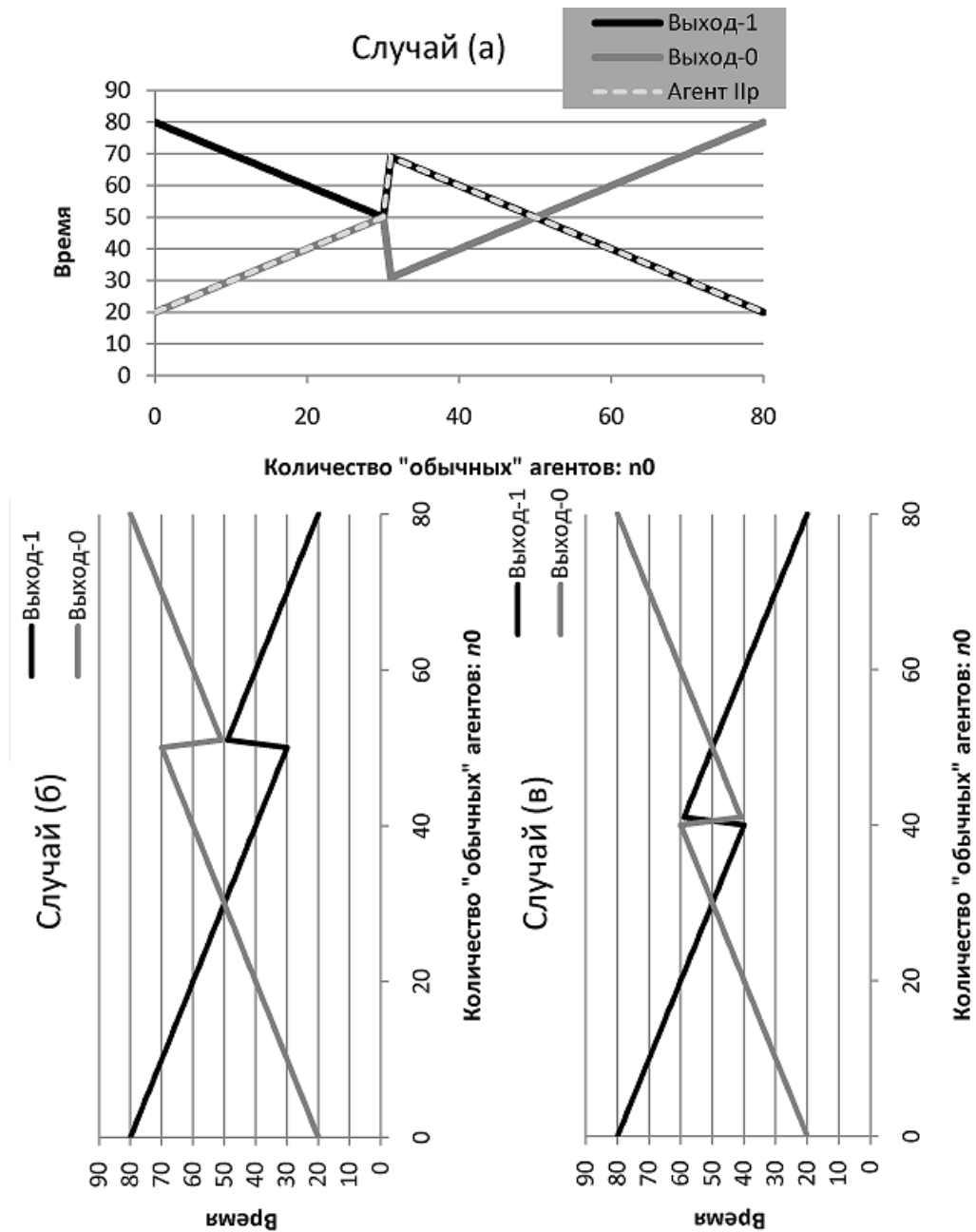


Рис. 14. Три случая для времени выхода

## 2.4. «Игра» однородных агентов

Пусть:

- 1) все агенты из множества  $N$  имеют одинаковые целевые функции (одинаковые типы):  $f_i(\cdot) = f(\cdot)$ ,  $i \in N$ ;
- 2) целевая функция  $i$ -го агента зависит от его действия  $a_i$  (причем она непрерывна и вогнута по этой переменной) и от состояния системы  $\chi(a)$ , где  $\chi(\cdot): \mathfrak{R}^n \rightarrow \mathfrak{R}^1$  – симметричная функция своих аргументов;
- 3) агенты принимают решения однократно (поэтому верхний индекс, соответствующий первому периоду времени, далее будем опускать),  $\gamma_i^t \equiv 1$ ;

4) начальный вектор действий  $a^0$  и рефлексивное разбиение  $\aleph$  – фиксированы.

Агенты нулевого ранга рефлексии в соответствии с выражением (9) выберут действия

$$(16) \quad {}^0a_i = \arg \max_{y \in \mathfrak{R}^1} f(y, \chi(y, a_{-i}^0)), i \in {}^0N.$$

Агенты первого ранга рефлексии в соответствии с выражением (10) выберут действия

$$(17) \quad {}^1a_j = \arg \max_{y \in \mathfrak{R}^1} f(y, \chi(y, {}^0a_{-j})), j \in {}^1N.$$

Агенты второго ранга рефлексии в соответствии с выражением (11) выберут действия

$$(18) \quad {}^2a_j = \arg \max_{y \in \mathfrak{R}^1} f(y, \chi(y, {}^{20}a, {}^{21}a)), j \in {}^2N.$$

И так далее. Агенты  $m$ -го ранга рефлексии в соответствии со своим субъективным рефлексивным разбиением (7) выберут действия

$$(19) \quad {}^ma_j = \arg \max_{y \in \mathfrak{R}^1} f(y, \chi(y, {}^{m0}a, {}^{m1}a, \dots, {}^{m(m-1)}a)), j \in {}^mN.$$

Вектор действий агентов зависит от функций  $f(\cdot)$  и  $\chi(\cdot)$ , а также вектора начальных действий  $a^0$  и рефлексивного разбиения  $\aleph$ . Если первые три параметра фиксированы, то вектор действий агентов зависит только от рефлексивного разбиения, как и в формуле (12), но ввиду однородности агентов, от их перестановки в рефлексивном разбиении  $\aleph$  ничего не зависит, то есть  $a^*(\aleph)$  зависит только от общего количества рангов и количества агентов каждого ранга рефлексии.

## 2.5. Постановка задачи управления

Управляемым параметром является рефлексивное разбиение агентов – их реальные ранги рефлексии.

Можно сформулировать задачу управления в форме достижимости: пусть задано множество  $\mathfrak{Z}$  допустимых рефлексивных разбиений; требуется найти множество

$$(20) \quad \mathfrak{R}(\mathfrak{Z}) = \bigcup_{\aleph \in \mathfrak{Z}} a(\aleph)$$

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  $F_0(\chi(a))$ , заданной на множестве состояний системы, т. е.  $F_0(\cdot): X \rightarrow \mathfrak{R}^1$ . Тогда, воспользовавшись выражением (12), эффективность рефлексивного разбиения  $\aleph$  можно определить как

$$(21) \quad K(\aleph) = F_0(\chi(a^t(\aleph))).$$

Следовательно, формально задачу рефлексивного управления (в

терминах рефлексивных разбиений) можно сформулировать в виде:

$$(22) K(\aleph) \rightarrow \max_{\aleph \in \mathfrak{Z}}$$

*Методом рефлексивных разбиений* называется метод управления агентами, с помощью варьирования разбиения множества агентов на ранги рефлексии в соответствии с моделью рефлексивных разбиений.

Пусть  $m$  – максимальный ранг рефлексивного разбиения  $\aleph$ . Тогда при однородности агентов, то есть когда у всех агентов одинаковый тип, эффективность рефлексивного разбиения определяется  $(m+1)$ -им числом:  $n_0, n_1, \dots, n_m$ . Следовательно, задача (22) примет вид:

$$(23) K(\aleph) \rightarrow \max_{n_i \geq 0, \sum_{i=0}^m n_i = n}$$

При большом числе агентов задачу (23) иногда удобно формулировать в терминах долей агентов  $q_i = n_i / n$ , имеющих тот или иной ранг рефлексии  $i = 0, \dots, m$ .

Обозначим  $K_m$  – максимальное значение критерия эффективности в задаче (22) при фиксированном максимальном ранге рефлексии  $m$ . По аналогии с тем, как это делалось для моделей стратегической рефлексии в [43], можно сформулировать задачу о *максимальном целесообразном ранге рефлексии* ([43], сс.15, 49) – таком, больше которого центру (с точки зрения задачи управления в форме достижимости или/и эффективности рефлексивного управления) использовать не имеет смысла:

$$(24) m^* = \min \{ \text{Arg} \max_{w=0,1,2,\dots} K_w \}.$$

Обсудим далее согласованность субъективных рефлексивных разбиений агентов. Предположим, что каждый агент наблюдает только состояние системы после действий всех агентов. Очевидно, прогнозируемое агентами состояние может отличаться от наблюдаемого. Это может служить для агентов основанием для того, чтобы усомниться в правильности своих субъективных рефлексивных разбиений. Если агенты наблюдают (помимо собственных действий) только состояние системы, то по аналогии с условием стабильности информационного управления (см. [65]) 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  $\aleph \in \mathfrak{Z}$  реализуется вектор действий, определяемый выражением (12). Соответственно, реализуется агрегированная ситуация  $\chi(a^t(\aleph))$ .

С точки зрения агента, обладающего  $k$ -м рангом рефлексии, реализуется вектор  $a^t(k, \aleph)$ .

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

$$(25) \chi(a^t(\aleph)) = \chi(a^t(k, \aleph)), k = 0, \dots, m.$$

Задачу рефлексивного управления (22) 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 раздела кратко обсудим, каким образом центр может управлять разбиением агентов по рангам рефлексии. На сегодняшний день в литературе описаны два возможных подхода. Первый подход предполагает, что агенты безусловно верят центру и воспринимают сообщаемую им информацию как истинную, независимо от своих первоначальных представлений. Тогда центр, последовательно сообщая ту или иную информацию различным группам агентов, может формировать различные (но не любые! – см. [52]) структуры информированности. Второй подход заключается в том, что агенты не просто заменяют свои представления теми, которые сообщает центр, а сообщения центра лишь снижают для агентов неопределенность – сокращают множество возможных с их точки зрения «миров» [64]. В целом, разработка моделей формирования структур информированности агентов, под влиянием поступающей к ним информации, представляется чрезвычайно перспективным направлением будущих исследований. Например, повысить ранг можно сообщением определённой информации о доле агентов нижних рангов, сообщением соответствующих вычислений и рациональных рассуждений. Понизить ранг  $j$ -го агента также можно сообщением, что он выбрал слишком высокий ранг, на самом деле максимальный ранг среди остальных агентов гораздо ниже его ранга. Ещё один возможный способ – ограничение времени, которое даётся на принятие решения. Прикладные технологии управления рефлексией также изучаются и применяются в социальной психологии, нейролингвистическом программировании, психотерапии и т.д.

Итак, с нормативной точки зрения модель рефлексивного группового поведения позволяет разбить множество агентов на группы и управлять системой за счёт различных вариантов разбиения.

В общем виде вряд ли задача допускает получение общих аналитических результатов. При введении ограничения на ранг рефлексии, условий на свойства функции  $F_0$  и заданном количестве агентов, для решения задачи (22) возможно применение оптимизационных методов и методов комбинаторной оптимизации. Многое также зависит от простоты формул (10), (11), (12).

При решении конкретных задач, находятся формулы (12), с помощью которых возможно решение задачи управления в форме достижимости (20) (стр.58). В примере с олигополией из раздела 2.6 возможно нахождение аналитических выражений для формул (12), а также решение задачи управления в форме достижимости.

Далее, если необходимо, ищется решение задачи (22).

Если задача управления в форме достижимости решена, то есть найдено множество  $\mathfrak{R}(\mathfrak{T})$ , то задачу (22) можно решить с её помощью, найдя максимум функции  $F_0$  на множестве  $\chi(\mathfrak{R}(\mathfrak{T}))$ :

$$f^* = \max_{h \in \chi(\mathfrak{R}(\mathfrak{T}))} F_0(h),$$

и решив обратную задачу к задаче управления в форме достижимости, то есть найти множество рефлексивных разбиений, которое доставляет максимум  $F_0$ :

$$\left\{ \mathfrak{N} \in \mathfrak{T} \mid a(\mathfrak{N}) \in \left\{ y : F_0(\chi(y)) = f^* \right\} \right\}.$$

Основания системы классификаций моделей стратегической рефлексии.

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

2) Принцип выбора действий агентами нулевого ранга рефлексии:

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

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

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

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

4) Распределение (объективное) агентов по рангам рефлексии:

– произвольное фиксированное;

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

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

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

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

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

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

– знает достоверно (или с некоторой погрешностью);

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

– считает что все (!) остальные агенты имеют ранг  $k - 1$ .

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

– считает их всех принадлежащих нулевому рангу;

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

– предполагает, что эти агенты распределены по рангам рефлексии от 0 до  $k - 1$  включительно в соответствии с некоторым вероятностным

распределением (как правило – распределением Пуассона);

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

При любых значениях признаков данной системы классификаций рефлексивное равновесие строится по общей схеме, приведенной ниже. Все известные на сегодняшний день зарубежные и отечественные модели стратегической рефлексии укладываются в рамки предложенной системы их классификаций.

Введенная в настоящей главе общая рефлексивная модель управления может служить базисом для создания частных аналитических математических моделей или общих имитационных моделей, позволяющих описывать и прогнозировать групповое поведение (людей, мобильных роботов, программных агентов) в разнообразных ситуациях. Например, рефлексивные имитационные модели эвакуации и транспортных потоков [26, 27], рефлексивные модели в олигополии [29, 79, 87, 126] и другие примеры в работах [72, 75, 78, 132, 133]. Аналитические результаты можно надеяться получить, вводя те или иные упрощающие предположения. Например, модели с однородными агентами. В следующем разделе 2.6 рассматриваются приложения предложенной рефлексивной модели управления в различных математических моделях.

## **2.6. Прикладные модели рефлексивного коллективного поведения и управления**

Рассмотрим применение МРР к различным известным прикладным моделям – рынка, теории принятия решений и т.п.

### **2.6.1. Олигополия Курно**

Олигополия Курно – экономическая модель рыночной конкуренции. Основные положения модели (см. [16]):

- на рынке действует фиксированное число  $n > 1$  агентов (фирм), выпускающих одинаковый продукт;
- вход на рынок новых фирм и выход из него отсутствуют;
- фирмы действуют без кооперации, одновременно выбирая объемы выпуска;
- фирмы максимизируют свою прибыль.

Общее количество фирм на рынке  $n$  предполагается известным всем участникам. Функции издержек фирм могут быть различными, и также предполагаются известными всем участникам. Функция спроса представляет собой линейную убывающую функцию от цены продукта. Обозначим объем производства  $i$ -й фирмы-агента через  $a_i$ .

Рассмотрим модель олигополии Курно с квадратичными функциями издержек. Тогда целевые функции (функции прибыли) агентов можно

записать в виде:

$$f(a_i, r_i, x) = a_i x - c(a_i) = a_i x - a_i^2/2;$$

$$x = \chi(a_1, \dots, a_n) = A - bQ(a),$$

где  $x$  – цена, устанавливаемая на рынке в зависимости от суммарного выпуска  $\sum a_j$ ;  $c(a_i) = a_i^2/2$  – издержки агента на производство заданного объема продукции;  $A$  – максимальная цена на рынке;  $b$  – коэффициент, отражающий «эластичность» цены от суммарного действия агентов.

### 2.6.2. Олигополия с «совершенной» конкуренцией

Рассмотрим статическую (одношаговую) модель, при  $Q(a) = \frac{1}{n} \sum_{j=1}^n a_j$ .

Тогда

$$x = \chi(a_1, \dots, a_n) = A - \frac{b}{n} \sum_{j=1}^n a_j.$$

Также рассматриваем большое количество агентов, что позволяет использовать *гипотезу слабого влияния* [44, с. 162]: агенты не учитывают собственного влияния на состояние системы.

В равновесии Нэша действия всех агентов равны  $a^N = \frac{A}{1 + b + b/n}$ , а

равновесная цена равна  $x^N = A \frac{1 + b/n}{1 + b + b/n}$ .

В условиях совершенной конкуренции, когда агентов достаточно много, агент пренебрегает своим влиянием на установившуюся цену  $x$ , оптимальное действие  $i$ -го агента при цене  $x$  можно записать как:

$$(26) \quad a = \arg \max_y f(y, x) = x.$$

Эта формула верна только при выполнении гипотезы слабого влияния, то есть если  $x$  не зависит от действия  $i$ -го агента.

То есть единицы измерения действий агентов подобраны таким образом, что, максимизируя свою целевую функцию, они выбирают действия, численно равные рыночной цене. При этом из уравнения  $x = A - b \sum_{j=1}^n a_j^*/n$  можно вычислить цену рыночного равновесия  $x_p = A/(1 + b)$ .

Рассмотрим следующую модель рефлексивного поведения.

«Обычные» агенты имеют некоторое фиксированное представление  ${}^0x$  о цене, соответственно и  ${}^0a = {}^0x$ .

Субъективное рефлексивное разбиение агента ранга  $k$  следующее:

$${}^k \mathfrak{X} = \{ {}^{k0}N \cup {}^{kk}N \setminus \{i\} \cup {}^{k(k+1)}N \cup \dots \cup {}^{km}N, {}^{k1}N, \dots, {}^{k(k-1)}N, 1, 0, \dots, 0 \}.$$

Соответствующие вычисления можно посмотреть в работе [29], итоговые формулы для определения вектора действий (12) следующие:

$$(27) \quad \begin{cases} {}^k a_i = {}^{k-1} x_\phi \\ {}^0 x_\phi = A - b x^0 \\ {}^k x_\phi = b q_k (x^0 - {}^{k-1} x_\phi) + {}^{k-1} x_\phi \end{cases}$$

При этом  ${}^k x_\phi$  – цена, фактически реализующаяся на рынке – зависит от

представлений  $x^0$  «обычных» агентов о цене, а также от распределения агентов по рангам рефлексии:  $q_0, q_1, \dots, q_k$ .

Рефлексивное управление со стороны управляющего органа – Центра – в рассматриваемой модели состоит в изменении рангов рефлексии всех или части агентов, то есть, влиянии на доли  $q_0, q_1, \dots, q_k$ .

Множество достижимости системы в данной модели:

$$D(x^0) = \bigcup_i D_i(x^0), \text{ где } D_i(x^0) = [\min_{\Lambda_i} x, \max_{\Lambda_i} x],$$

где  $\Lambda_i = (q_0, q_1, \dots, q_i)$ . Отметим, что представления  $x^0$  «обычных» агентов считаются фиксированными. В работе [29] доказано следующее утверждение:

### Утверждение 2.

1. В рамках описанной выше модели множество достижимости  $D(x^0)$  представляет собой отрезок с границами

$$(28) D(x^0) = \begin{cases} [x^0 - dm; x^0 + dM] & \text{при } d > 0 \\ [x^0 - dM; x^0 + dm] & \text{при } d < 0 \end{cases}$$

где  $d = A - bx^0 - x^0$ ;  $M = (\frac{b}{i_M} - 1)^{i_M}$ ;  $m = (\frac{b}{i_m} - 1)^{i_m}$ ;

$$i_M = \arg \max_{i=2k, k \in \mathbb{N}} (\frac{b}{i} - 1)^i; i_m = \arg \max_{i=2k+1, k \in \mathbb{N}_0} (\frac{b}{i} - 1)^i.$$

Длина отрезка  $D(x^0)$  экспоненциально растет с ростом параметра  $b$ , более того,  $|M| \sim e^b$ ;  $|m| \sim e^b$ . Максимальный целесообразный ранг рефлексии  $m^*$  (с. 59) равен:

$$(29) m^* = \max\{i_M, i_m\}.$$

2. Равновесная рыночная цена при рефлексивном управлении достижима при любом  $x^0$ , для этого достаточно взять

$$(30) q_1^p = \frac{1}{1+b} \in (0,1]$$

рефлексивных агентов ранга 1, а всех остальных – «обычных».

3. Фактическая цена по формуле (28) может быть отрицательной, поэтому из формул (27) следует, что отрицательны могут быть действия агентов по выпуску продукта. Это удобно для аналитических расчётов, но не для практических результатов. С учётом неотрицательности объёма производства агентов-фирм формула (26) должна быть заменена следующей (31)  $a = \arg \max_{y \geq 0} f(y, x) = \max\{x, 0\}$ .

Введём необходимые обозначения:

$$k^* = \frac{\ln c}{W(1/e) + 1}, \text{ где } W(x) - \text{W-функция Ламберта.}$$

$$\begin{cases} k = \text{even}(k^*), d < 0 \\ k = \text{odd}(k^*), d > 0 \end{cases} \Rightarrow \begin{cases} k = \text{even}(k^*), d < 0 \\ k = \text{odd}(k^*), d > 0 \\ q = \frac{1}{b} \left( (x^0/d)^{1/k} + 1 \right) \end{cases}$$

где  $even(x)$  – ближайшее к  $x$  целое чётное число, а  $odd(x)$  – нечётное.

$$x_{\max} = \max\{bq_{k+1}x^0 \mid q_{k+1} \in [0, 1-kq]\} = bx^0(1-kq),$$

$$(32) \exists j \leq i_m : (-1)^j \left(\frac{b}{i_m} - 1\right)^j < -x^0,$$

$$(33) \exists j < i_M : (-1)^j \left(\frac{b}{i_M} - 1\right)^j < -x^0.$$

Тогда, если  $D(x^0)$  содержит отрицательные значения, то соответствующее (31) значение множества достижимости  $D^*(x^0)$  уменьшится, также как и максимальный целесообразный ранг рефлексии  $m^+$ :

$$D^*(x^0) = \begin{cases} [0, x_{\max}], & \text{при выполнении условий (28) и (29)} \\ [0, x^0 + dM], & \text{при условии (28) и нарушении условия (29)} \end{cases}$$

$$D^*(x^0) \subseteq D(x^0)$$

$$m^+ = k + 1 \leq m^*.$$

То есть формулы (28) и (29) есть верхние оценки.

Содержательно, первый результат утверждения 2 говорит о том, что Центру для максимальной управляемости рынка достаточно формировать  $m^*$  рангов рефлексии агентов, кроме «обычных». Этого достаточно, чтобы цену на рынке можно было сделать любой из множества достижимых цен  $D(x^0)$ .

Ширина множества достижимости  $|D(x^0)|$  экспоненциально увеличивается с ростом  $b$ , отражающего «эластичность» цены на рынке от действий агентов.

Второй результат содержательно говорит о полезности рефлексивного управления: Центр способен привести такую экономическую систему к равновесной рыночной цене за счёт использования только «обычных агентов» и агентов первого ранга. Стоит отметить, что  $q_1^P$  не зависит от  $x^0$ .

В [43] со ссылкой на [61] описывается психологический эксперимент, проведенный изучавшим психологию владельцем компании, импортирующей в США говядину. «Торговые агенты позвонили, как обычно, постоянным клиентам компании – закупщикам говядины для супермаркетов и других точек, торгующих продуктами в розницу, и одним из трех способов предложили им сделать заказ.

Одни клиенты услышали предложение, сделанное в стандартной форме. Другим клиентам дополнительно была предоставлена информация о том, что поставки импортной говядины будут сокращены в ближайшие несколько месяцев. Третья группа клиентов получила те же сведения, что и вторая группа, а также информацию о том, что мало кто узнает о предстоящем сокращении поставок, так как эти сведения поступили из надежного, но засекреченного источника. ...

По сравнению с клиентами, которым было сделано торговое предложение в стандартной форме, те клиенты, которым было также сказано о дефиците говядины, заказали ее в два раза больше... Клиенты, которые решили, что владеют «исключительной» информацией ... приобрели в шесть раз больше говядины, чем клиенты, которым было сделано торговое

предложение в стандартной форме».

Несмотря на то, что в данном случае речь идет о конкуренции не поставщиков, а потребителей, аналогия с рассмотренной выше моделью олигополии Курно очевидна. Клиенты, получившие предложение в стандартной форме, в своих закупках будут ориентироваться на стандартные цены. Их, таким образом, можно считать «обычными» агентами. Клиенты третьей группы считают себя обладателями эксклюзивной информации о рыночной цене и полагают, что правильно предсказывают поведение «обычных» агентов. В описанной выше модели такое поведение характерно для агентов первого ранга рефлексии. Информация, сообщенная клиентам второй группы, в терминах рассматриваемой модели – это рефлексивное управление частью «обычных» агентов, оставляющее их на нулевом ранге рефлексии, но меняющее их прогноз цены  $x^0$ .

Специфика рассматриваемой модели состоит в её статичности и большом количестве агентов (действия одного агента не влияют на общий результат). Описанная в предыдущем разделе структура информированности достаточно проста, что и позволяет ее исследовать аналитически.

### 2.6.3. Олигополия с «несовершенной» конкуренцией

Рассмотрим другую модель – без гипотезы слабого влияния. Цена считается по формуле:  $x = \chi(a_1, \dots, a_n) = A - bQ(a_1, \dots, a_n)$ , но  $Q(a_1, \dots, a_n) = \sum_{i \in N} a_i$ .

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

$$(34) a_i^N = \frac{A}{1+b+nb}, i \in N,$$

которые приводили бы к равновесному объему выпуска  $Q(a^N) = \frac{nA}{1+b+nb}$  и

равновесной цене  $P(a^N) = \frac{A(1+b)}{1+b+nb}$ . Точке Парето, максимизирующей сумму целевых функций агентов, соответствуют действия:

$$(35) a_i^P = \frac{a}{1+2nb}, i \in N,$$

которые приводят к эффективному объему выпуска  $Q(a^P) = nA/(1+2nb)$  и эффективной цене  $P(a^P) = A(1+nb)/(1+2nb)$ . При этом  $f(a^P) = A^2/(2(1+2nb)) \geq f(a^N) = A^2(1+2b)/(2(1+b+nb)^2)$ , то есть выигрыш каждого агента в точке Парето не меньше, чем в точке Нэша.

**ПРИМЕР №2.** Рассмотрим числовой пример. Пусть  $n = 10$ ,  $A = 2.1$ ,  $b = 0.1$ ,  $\gamma_i^t = 0.5$ . Тогда  $a_i^N = 1$ ,  $Q(a^N) = 10$ ,  $P(a^N) = 1.1$ ,  $a_i^P = 0.7$ ,  $Q(a^P) = 7$ ,  $P(a^P) = 1.4$ ,  $f(a^P) = 0.735 > f(a^N) = 0.6$ .

Проанализируем динамику коллективного поведения. Пусть фиксирован вектор  $a^0$  начальных объемов производства. В соответствии с выражением (4) изменение во времени действий, выбираемых агентами, будет описываться следующим выражением:

$$(36) \quad a_i^t = a_i^{t-1} + \gamma_i^t \left[ \left( A - b \sum_{j \neq i} a_j^{t-1} \right) / (1 + 2b) - a_i^{t-1} \right], \quad i \in N, t = 1, 2, \dots$$

В соответствии с выражением (36) действия агентов будут сходиться к равновесию Нэша.

Перейдем теперь к рефлексивному случаю (этот пример из [24]). Субъективные рефлексивные разбиения выглядят следующим образом:

$${}^k \mathfrak{R} = \{ {}^{k0}N, {}^{k1}N, \dots, {}^{k(k-1)}N \cup {}^{kk}M \setminus \{i\} \cup {}^{k(k+1)}N \cup \dots \cup {}^{km}N, 1, 0, \dots, 0 \}.$$

При заданном векторе начальных действий  $a^0$  агенты нулевого ранга рефлексии выберут действия  $a_i = C + D a_i^0$ ,  $i \in N_0$ , где  $C = \frac{A - bQ(a^0)}{1 + 2b}$ ,

$D = \frac{b}{1 + 2b}$ . Агенты первого ранга рефлексии выберут действия

$${}^1 a_j = C_1 + D^2 a_j^0, \quad j \in N_1, \quad \text{где } C_1 = \frac{A(1 + 3b) - bnA + b^2Q(a^0)(n - 2)}{(1 + 2b)^2}.$$

Пусть в рассматриваемом числовом примере все начальные действия агентов одинаковы:  $a_i^0 = 0.5$ ,  $i \in N$ . Тогда  $a_i = 31/24 = 1.291(6)$ ,  ${}^1 a_j = 103.5/144 = 0.71875$ , 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. То есть при векторе начальных действий  $a_i^0 = 0.5$ ,  $i \in N$ , наличия агентов первого ранга рефлексии недостаточно для реализации за счет рефлексивного управления Парето-оптимальной точки. Но вполне достаточно для реализации соответствующего равновесию Нэша суммарного объема производства – для этого доля рефлексивующих 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  $a_i^0 = 0.2$ ,  $i \in N$ . Тогда  $a_i = 1.5$ ,  ${}^1 a_j = 0.55$ , и при доле агентов первого ранга рефлексии, равной примерно 84 %, на рынке установится эффективная цена  $P(a^P)$ . Однако такая ситуация не будет стабильной с точки зрения условия (25).

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

$${}^2 a = \frac{1}{1 + 2b} [A - b n_0 a - b (n - n_0 - 1) {}^1 a].$$

Итого, получаем, что в зависимости от рефлексивного разбиения реализуется суммарное действие

$$(37) \quad Q(n_1, n_2) = (n - n_1 - n_2) a + n_1 {}^1 a + n_2 {}^2 a = \frac{1}{1 + 2b} [(A - b(n-1)a^0)(n - n_1 -$$

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

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, то есть когда выполняется  $Q(n_1, n_2) = Q(x^N)$  в зависимости от начальных действий агентов  $a^0$ .

Для рассматриваемого примера кривая  $AB$  пересечения графика  $Q(n_1, n_2)$  и «нэшевской» плоскости  $Q=10$  приведена на Рис. 15. Оказывается, что эта кривая не зависит от  $a^0$  – её формула в плоскости  $Q=10$ :

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

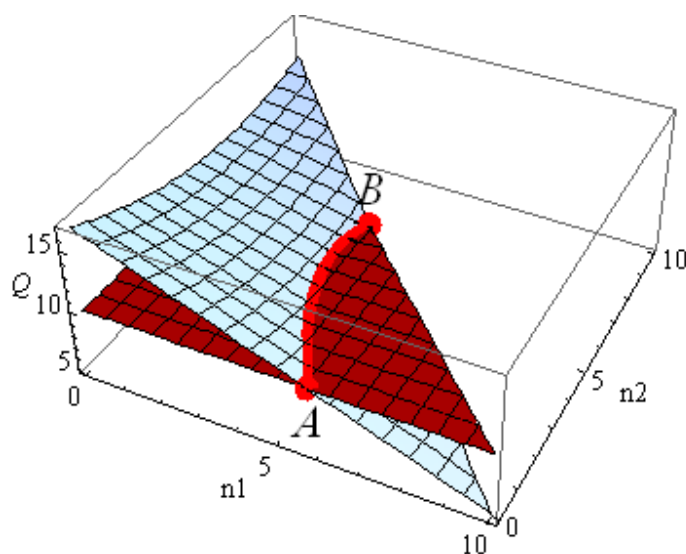


Рис. 15. Кривая  $AB$  «реализует» равновесие Нэша

Из Рис. 15 видно, что введение даже только агентов первого ранга увеличивает суммарный объем производства.

Отметим, что с точки зрения «стабильности», если имеет место динамика, то если на первом шаге агенты попадают в точку Нэша, то и в дальнейшем ни один из них (ни нерефлексорирующий, ни рефлексорирующий) не имеют оснований для изменения своих действий.

Если же мы ищем такое число рефлексорирующих агентов, чтобы объем производства был равен отличному от  $Q(a^N)$  значению, например объёму, соответствующему Парето-оптимальной ситуации, то кривая  $AB$  будет меняться в зависимости от  $a^0$ . 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  $Q(a^N)$ ). Предположим, что в начальный момент времени агенты не осуществляли производства ( $a^0 = 0$ ).

Достичь требуемого объёма можно – см. Рис. 16.

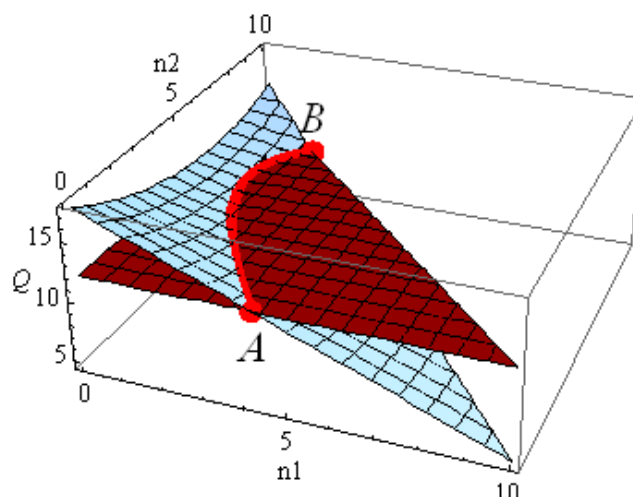


Рис. 16. «Реализация» требуемого суммарного объёма производства

Если  $a^0 \approx 0.305$ , то кривая  $AB$  касается плоскости  $n_1 = 0$  (точка  $C$  на Рис. 17), то есть, в этом случае можно только агентами нулевого и второго ранга рефлексии достичь требуемого суммарного объёма производства.

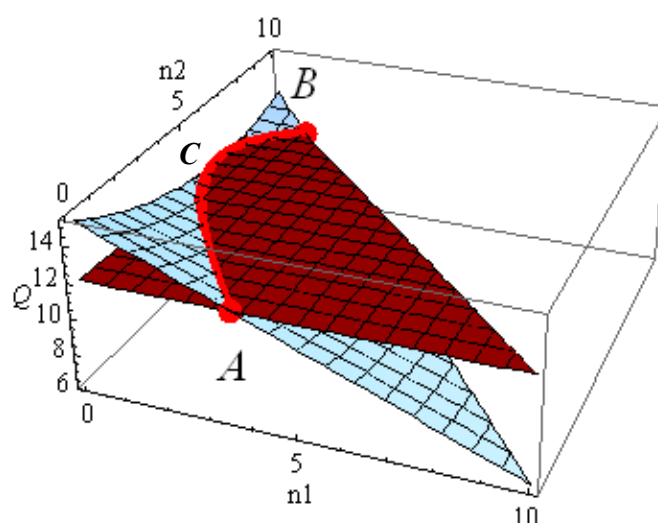


Рис. 17. «Реализация» требуемого суммарного объёма производства в отсутствии агентов первого ранга (точка  $C$ )

Таким образом, в модели олигополии Курно введение рефлекслирующих агентов позволяет увеличить суммарный объём производства и/или реализовать его Парето-эффективное значение.

#### 2.6.4. Задача о консенсусе

Содержательная интерпретация «задачи о консенсусе» следующая: действиям агентов соответствуют их положения на прямой (координаты в пространстве или мнения и т.д. – см. обзоры в [62, 129]), состояние системы – среднее значение координат агентов:  $\chi(a) = \frac{1}{n} \sum_{i \in N} a_i$ . Целевой функцией агента будем считать его «отклонение» от состояния системы:

$$f(a_i, \chi(a)) = -(a_i - \chi(a))^2, i \in N.$$

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

$$(38) \quad F(a) = -\frac{1}{n} \sum_{i \in N} (\chi(a) - a_i)^2.$$

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 (38). Однако в случае (даже одношагового) коллективного поведения агентов в условиях неполной их информированности все не так просто.

Ранг 0. При заданных начальных положениях агентов  $a^0$   $i$ -й агент в соответствии с выражением (16) выберет действие

$$(39) \quad a_i = \frac{1}{n-1} \sum_{j \neq i} a_j^0 = \frac{1}{n-1} (n\chi(a^0) - a_i^0), i \in {}^0N,$$

равное среднему положению остальных агентов,  $i \in N$ . Сделанный вывод остается в силе и в случае, когда целевые функции агентов зависят не от агрегированной ситуации, а от *агрегированной обстановки*:

$$g(a_i, \chi_i(a_{-i})) = -(a_i - \chi_i(a_{-i}))^2, \text{ где } \chi_i(a_{-i}) = \frac{1}{n-1} \sum_{j \neq i} a_j, i \in N.$$

Из выражения (39) следует, что  $\chi(a) = \chi(a^0)$ , то есть среднее значение координат агентов не изменяется, а значение критерия эффективности возрастает в  $(n-1)^2$  раз:  $F(a) = \frac{1}{(n-1)^2} F(a^0)$ .

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 (39), а агенты первого ранга – следующие действия:

$${}^1a_j = \frac{n\chi(a) - a_j}{n-1} = \frac{n^2(n-2)\chi(a^0) + a_j^0}{(n-1)^2}, j \in {}^1N.$$

Если все агенты обладают первым рангом рефлексии, то  $\chi({}^1a) = \chi(a) = \chi(a^0)$ , т. е. среднее значение координат агентов не изменяется (такой случай является идеальным с точки зрения стабильности рефлексивного разбиения – все агенты наблюдают ожидаемые значения). Значение критерия эффективности возрастает еще в  $(n-1)^2$  раз:

$$F({}^1a) = \frac{1}{(n-1)^2} F(a) = \frac{1}{(n-1)^4} F(a^0).$$

Рассмотрим пример – пусть  $n = 2$ . Получаем, что в зависимости от своих рангов рефлексии агенты выберут следующие действия:

	Агент 1	Агент 2
<b>Начальные действия</b>	$a_1^0$	$a_2^0$

<b>Ранг рефлексии</b>	<b>0</b>	$a_2^0$	$a_1^0$
	<b>1</b>	$a_1^0$	$a_2^0$
	<b>2</b>	$a_2^0$	$a_1^0$

Видно, что, во-первых, вектор действий обоих агентов, обладающих вторым рангом рефлексии, совпадает с вектором действий нерелексирующих агентов. Во-вторых, при одинаковых рангах рефлексии обоих агентов значение критерия эффективности не зависит от ранга. В-третьих, все четыре возможные комбинации действий агентов исчерпываются нулевым и первым рангами их рефлексии. В-четвертых, максимальное (равное нулю) значение критерия эффективности (38) 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.

### 2.6.5. Активная экспертиза

Рассматриваемый в настоящем параграфе пример свидетельствует, что наличие релексирующих агентов может приводить к последствиям, негативным, условно говоря, с точки зрения группы в целом.

Содержательная интерпретация модели активной экспертизы следующая: имеются  $n$  экспертов – агентов, сообщающих информацию организатору экспертизы – центру. Центр принимает решение  $\chi(a) = \frac{1}{n} \sum_{i \in N} a_i$ , 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 [44]:

$$f(a_i, \chi(a)) = -(a_i^0 - \chi(a))^2, i \in N.$$

Пусть агенты упорядочены по возрастанию их начальных мнений:  $a_1^0 < a_2^0 < \dots < a_n^0$ .

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

Определим множество агентов  $M(a^0) = \{i \in N \mid a_i^0 \geq \sum_{l \neq i} a_l^0 / n\}$ .

**Ранг 0.** При заданных начальных мнениях агентов  $a^0$ , они в соответствии с выражением (16) выберут действия

$$(40) \quad a_i = \max \{n a_i^0 - \sum_{l \neq i} a_l^0, 0\}, i \in N.$$

Вычислим  $\chi(a) = \sum_{i \in M(a^0)} a_i - \frac{1}{n} \sum_{i \in M(a^0)} \sum_{l \neq i} a_l^0$ .

**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 (40), а агенты первого ранга – следующие действия:

$$(41) \quad {}^1a_j = \max \{na_j^0 - \sum_{l \in M(a^0) \setminus \{j\}} a_l; 0\}, j \in {}^1N.$$

Рассмотрим числовой пример. Пусть имеются 10 агентов, чьи начальные мнения равны их номеру. Действия агентов приведены в следующей таблице:

№	1	2	3	4	5	6	7	8	9	10
$a^0$	1	2	3	4	5	6	7	8	9	10
$Q(a^0)$	5,5									
$a$	0	0	0	0	0	11	22	33	44	55
$Q(a)$	16,5									
${}^1a$	0	0	0	0	0	0	0	0	0	0

Варьируя число рефлексорирующих 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 (ср. с результатами анализа информационной рефлексии в задачах экспертизы в [42]), то есть, центр, осуществляя рефлексивное управление, имеет значительные возможности по манипулированию результатами экспертизы.

Ранг 2. Агенты второго ранга рефлексии выберут действия:

$$(42) \quad {}^2a_j = \max \{na_j^0 - \sum_{l \in {}^1N \cup {}^2N \cap M(a^0) \setminus \{j\}} {}^1a_l - \sum_{l \in {}^0N \cap M(a^0) \setminus \{j\}} a_l; 0\}, j \in {}^2N.$$

Пусть центр использует следующее рефлексивное разбиение:  ${}^0N = \{1, 2, 3, 4, 5\}$ ,  ${}^1N = \{6, 7, 8, 9\}$ ,  ${}^2N = \{10\}$ . Тогда в соответствии с выражениями (40)-(42) все агенты, кроме десятого, выберут нулевые действия, а десятый агент – действие, равное 100. То есть в рассматриваемом примере второго ранга рефлексии достаточно, чтобы получить ситуацию, совпадающую с равновесной по Нэшу.

### 2.6.6. 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. Обозначим  $n_L(n_R)$  – 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, возрастающей 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  $T_L > T_R$ , то

есть правый выход расположен ближе левого. Итак, полное время выхода налево равно  $T(n_L) = T_L + T(n_L)$ , направо:  $T(n_R) = T_R + T(n_R)$ .

*Время эвакуации* – время покидания помещения последним из агентов. Именно этот критерий используется во всех моделях эвакуации.

Оптимальное, с точки зрения *времени эвакуации*  $T^*$ , распределение агентов по направлениям движения  $(n_L^*; n_R^*)$  является решением следующей системы уравнений (см. также Рис. 18):

$$(43) \quad \begin{cases} T(n_L^*) + T_L = T(n_R^*) + T_R, \\ n_L^* + n_R^* = n. \end{cases}$$

Минимальное время эвакуации равно

$$(44) \quad T^* = T(n_L^*) + T_L = T(n_R^*) + T_R.$$

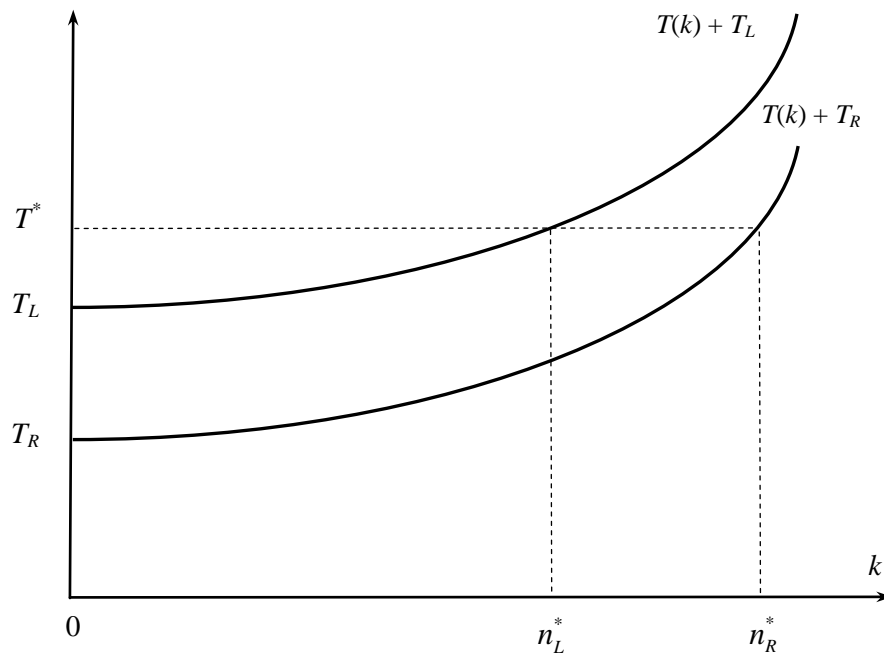


Рис. 18. Зависимость времени эвакуации от числа агентов, выбирающих правый или левый выход

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 (которыми они считают всех остальных агентов) создают пробку, выберут левый выход.

Время выхода в зависимости от числа агентов первого ранга рефлексии (см. Рис. 19) равно

$$(45) \quad T_1(n_1) = \max \{T(n_1) + T_L; T(n - n_1) + T_R\}.$$

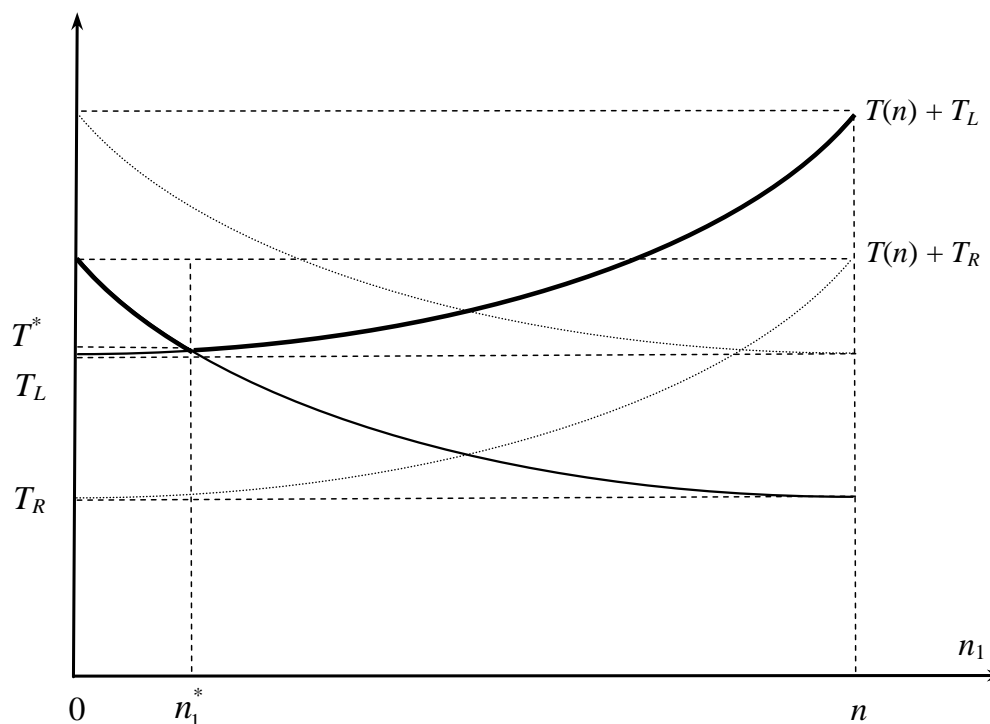
Видно, что как малое, так и очень большое число рефлекслирующих агентов плохо, так как увеличивает время эвакуации (см. Рис. 19). Т. е. существует оптимальное число рефлекслирующих агентов, при котором

время эвакуации минимально.

Из свойств функции  $T(\cdot)$  и предположения  $T_L > T_R$  следует, что минимум выражения (45) достигается при числе агентов первого ранга рефлексии  $n_1^*$ , определяемом из следующего соотношения:

$$(46) \quad T(n_1^*) + T_L = T(n - n_1^*) + T_R.$$

Последнее условие совпадает с условием (44), т. е.  $n_1^* = n_L^*$ ,  $T1(n_1^*) \equiv T^*$ , значит, первый ранг рефлексии является максимальным целесообразным в рамках рассматриваемой модели.



*Рис. 19. Зависимость времени эвакуации от числа агентов первого ранга рефлексии*

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 за счет введения агентов первого ранга значение времени эвакуации (44). Описание имитационных моделей транспортных потоков и эвакуации можно найти в [26, 27].

### **2.6.7. Фондовый рынок: пример возможных расширений модели**

В настоящем параграфе обсуждаются возможные расширения описанного выше метода рефлексивных разбиений, а именно – на примере частной модели фондового рынка рассматривается стратегическая рефлексия агентов «над» их равновесными по Нэшу стратегиями. Фондовый рынок является объектом моделирования, для которого наиболее часто

используют «рефлексивные» рассуждения – см., например, [17, 54]. В работе [18] 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 [18] 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 агента сложится следующая цена: если все агенты приобретают актив, то цена  $p$  будет равна  $p^+ = p_0 + \theta + \alpha n x_0$ ; если агенты продают актив, то цена  $p$  будет равна  $p^- = p_0 + \theta - \alpha n x_0$ , где  $\alpha$  – коэффициент зависимости цены от спроса/предложения.

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

- ♦  $(x_0 + u_0 / p_0) p^+ - u_0$ , 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

$$(47) \quad \theta \geq \frac{p_0 u_0}{p_0 x_0 + u_0} - \alpha n x_0.$$

Последнее условие 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, то есть 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 (47), 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 за счет 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 (47), то он может спрогнозировать, что все агенты нулевого ранга будут приобретать актив. Если условие (47) 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, то есть в рассмотренной модели добавление рефлексизирующих агентов любого ранга не меняет рыночной цены.

Сделанный вывод является следствием 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, во-первых, что доля  $q$  рефлексизирующих агентов является среди них общим знанием, а во-вторых, что эта доля достаточна для того, чтобы цена выросла. Последнее условие по аналогии с условием (47) можно записать в виде:

$$(48) \quad \theta \geq \frac{p_0 u_0}{p_0 x_0 + u_0} + \alpha n (1 - 2q) x_0,$$

т. е.,

$$(49) \quad q \geq q^* = \frac{1}{2} + \frac{1}{2\alpha n x_0} \left[ \frac{p_0 u_0}{p_0 x_0 + u_0} - \theta \right].$$

Отметим, что критическая доля  $q^*$  рефлексизирующих агентов составляет не менее половины от общего числа агентов (условие  $q^* \leq 1$  эквивалентно условию (47)). Рассмотрим числовой пример. Пусть  $n = 100$ ,  $u_0 = 1000$ ,  $p_0 = 10$ ,  $x_0 = 100$ ,  $\alpha = 0.001$ ,  $\theta = -1$ . Условие (47) выполнено. Из выражения (49) находим  $q^* = 53\%$ .

Подчеркнем, что предположение о том, что доля рефлексизирующих агентов является среди них общим знанием, противоречит введенному выше предположению о структуре субъективных рефлексивных разбиений (см. выражение (7)), 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 (7), не является единственно возможной и адекватной всем моделям, представляющим интерес для практики. То есть, перспективным направлением будущих исследований представляется рассмотрение и других структур субъективных рефлексивных разбиений.

Таким образом, в настоящей главе:

- сформулирован общий метод описания рефлексивных структур, в которых агенты, обладающие различными рангами стратегической рефлексии, могут иметь в общем случае несогласованные представления о рангах оппонентов;
- предложена концепция рефлексивного равновесия как устойчивого исхода взаимодействия агентов, осуществляющих совместную деятельность и обладающих различными рангами стратегической рефлексии;
- сформулирована общая задача рефлексивного управления – формирования допустимой рефлексивной структуры, приводящей к наиболее выгодному для управляющего органа рефлексивному равновесию;
- для прикладных моделей "Олигополия Курно", "Задача о консенсусе", "Транспортные потоки и модель эвакуации" и др. получены оценки максимальных целесообразных рангов рефлексии и найдены оптимальные рефлексивные структуры.

## **Глава 3. Имитационные модели рефлексивного коллективного поведения**

После описания модели рефлексивного коллективного поведения, приведенного во второй главе, в данной главе рассмотрим задачи её приложения в практических случаях, в которых уместно рассмотреть агентов с рефлексивным поведением, а именно – задачу об эвакуации и задачу о диффузной бомбе. В отдельную главу данные задачи вынесены потому, что в связи со сложностью их теоретического анализа для исследования используется метод имитационного моделирования,.

Для моделирования игроков (агентов, подвижных объектов) используется агентный подход в моделировании, в качестве инструмента разработки и анализа использовалась система агентного имитационного моделирования, разработанная автором (см. разделы 3.1, 3.2, 3.3 и 3.4), а также и система имитационного моделирования AnyLogic (см. раздел 3.5).

В разделах 3.2, 3.3 описаны модели стратегического и тактического поведения агентов, после чего, в разделе 3.4 приводятся результаты моделирования эвакуации людей из здания. Качественно исследуются параметры эвакуации при заданном коллективном поведении агентов, затем приводится два примера эвакуации с добавлением рефлексивных агентов и выводы по их влиянию на время эвакуации.

В разделе 3.5 рефлексивное поведение исследуется в задаче о «диффузной бомбе» – проникновении группы агентов к цели через систему противодействия.

### **3.1. Система агентного имитационного моделирования**

Назначение созданной автором системы имитационного моделирования – моделирование движения групп агентов в здании или, если более абстрактно, на плоскости с заданной геометрией.

В описании системы агентного моделирования остановимся на качественных сторонах: архитектура и ключевые блоки системы – параграф 3.1.1; архитектура агента – параграф 3.1.2; модель здания – параграф 3.1.3.

#### **3.1.1. Архитектура системы моделирования**

Система состоит из множества взаимосвязанных модулей:

- а. **соответствующих общей методологии** моделирования: агент, эксперимент, статистика, пакетный загрузчик, визуализатор;
- б. **соответствующих предметной области** моделирования: здание, помещение, переход, выход, подвижный агент, размещение агентов в здании;
- в. **соответствующих назначению системы** моделирования: модели целенаправленного движения, модели эвакуации, модели группового движения;

Ключевой модуль системы моделирования – «*Эксперимент*» –

содержит полное описание проводимого с помощью моделирования конкретного исследования: параметры процесса имитации (шаг времени, точность и т.п. параметры компьютерного моделирования), конкретное здание, типы агентов (поведение), количество и размещение агентов в здании, собираемая статистика.

Однократное моделирование процесса эвакуации называется *тестом*. Для автоматизации сбора данных, в эксперименте выделяется *параметр исследования*, например, ширина первого перехода, и диапазон его изменения в течение времени эксперимента. Совокупность тестов для одного значения параметра исследования назовём «*исследованием*». Количество тестов задано и одинаково для каждого «исследования».

Структура выполнения проводимого эксперимента основывается на многократном моделировании процесса эвакуации для текущего значения параметра исследования и перехода к следующему его значению и так для всего заданного диапазона его изменения – смотрите Рис. 20 для примера.

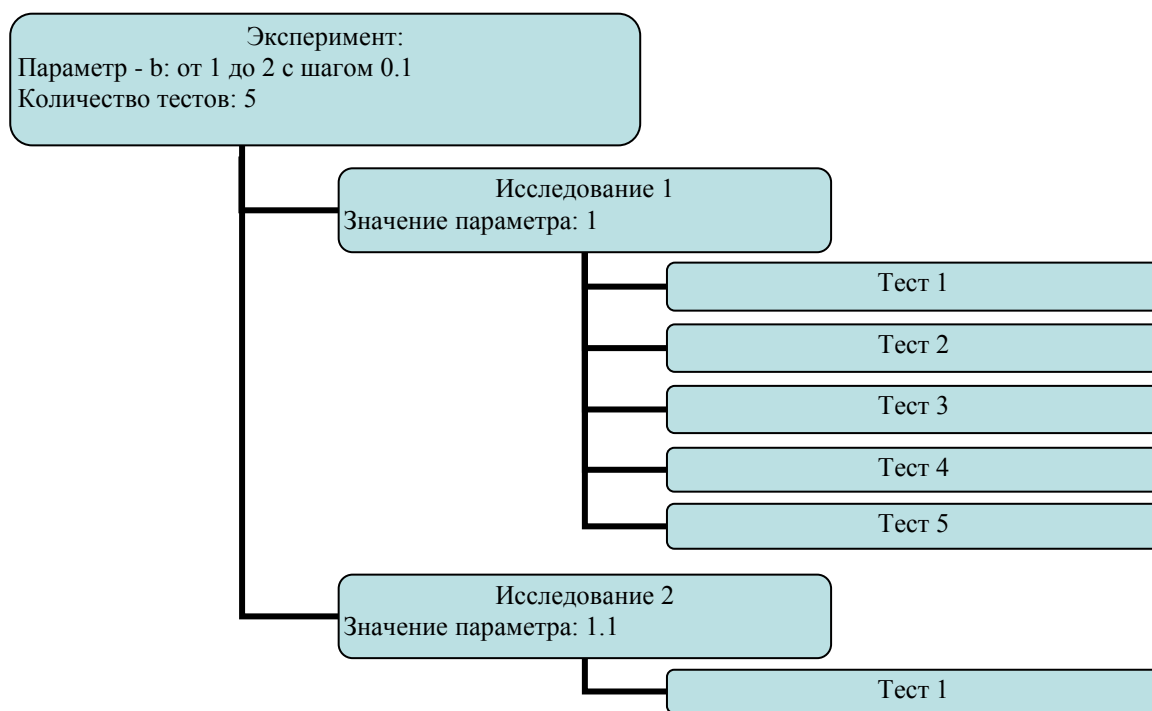


Рис. 20. Структура выполнения эксперимента

Следующий по важности модуль – модуль статистики. Он связан со сбором и обработкой получаемой в процессе моделирования информации. Собранные данные связаны со структурой выполнения эксперимента и, соответственно, разделены по исследованиям и тестам.

На Рис. 21 показана архитектура системы в целом. Тип агента определяется набором используемых алгоритмов поведения. Размещение агентов содержит тройки: (тип агентов, количество, помещение).

Два оставшихся модуля «Пакетный загрузчик» и «Визуализатор» отвечают за следующие функции. Пакетный загрузчик, собственно, запускает эксперимент на выполнение, это основной вариант использования

системы моделирования. Визуализатор интерактивно отображает процесс эвакуации (один тест) для стартового значения параметра эксперимента без сбора статистики. Есть возможность остановки/запуска процесса эвакуации и сброса.

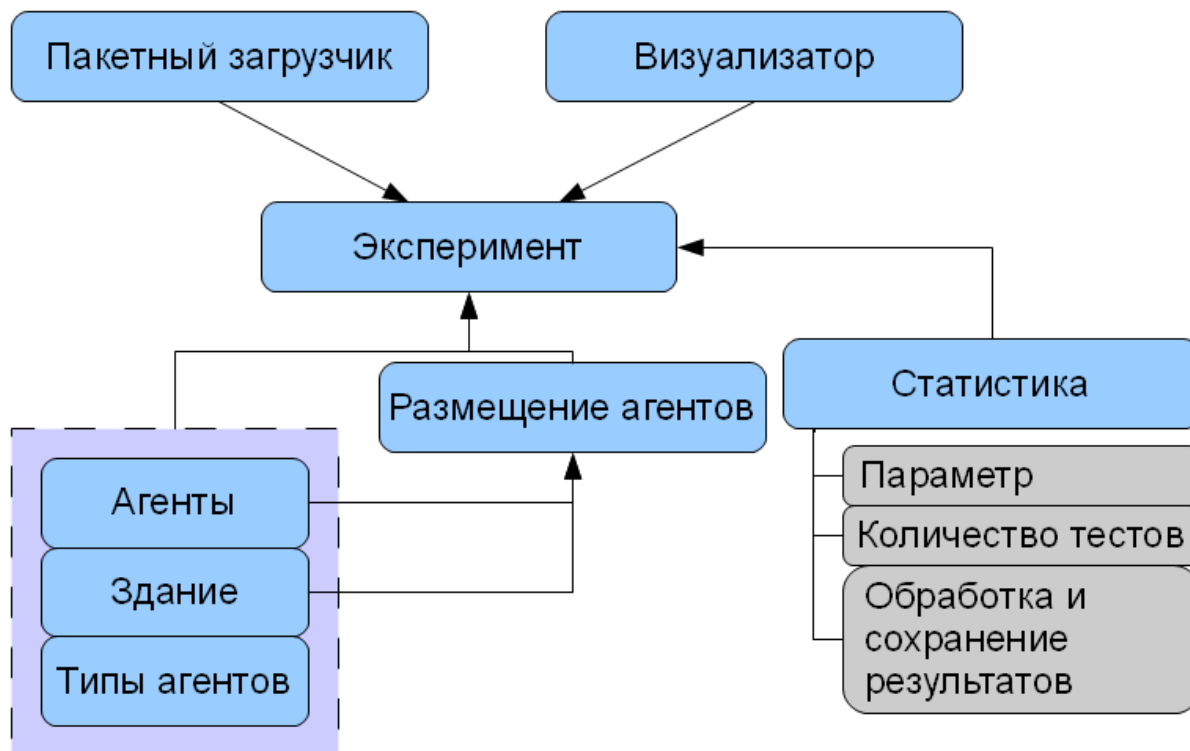


Рис. 21. Архитектура системы моделирования

Перейдём теперь к описанию архитектуры агента.

### 3.1.2. Архитектура агента

Агент как программная сущность, в системах агентного моделирования, в числе прочего ([66, 137]), имеет способности целеполагания, самостоятельности и рациональности в выборе своего поведения.

В рассматриваемой системе поведением агента является, главным образом, движение от начальной позиции до конечной. Предполагаем, что своё движение в здании агент рассматривает с разных уровней абстракции: в каком помещении здания находится агент и в какое помещение ему надо попасть, в каком месте внутри помещения и в какое место ему надо, как далек его вектор скорости от оптимального. Для объединения алгоритмов поведения для этих уровней абстракции, агент содержит *иерархический механизм поведения*, а также набор параметров, необходимых для функционирования этого механизма:

1. Параметры:

а. *Цель* – объект (например, выход в здании), к которому

- стремится придти агент. По умолчанию Цель – любой выход;
- в. *Мишень* – объект (например, ближний дверной проём), к которому идёт агент в настоящее время внутри помещения, т.е. можно сказать, что мишень это подцель;
  - с. Координаты агента, соответствующие уровням механизма поведения. Например, декартовы координаты и координата-помещение (в каком помещении находится агент), скорость, ускорение и т.д.
2. Структура иерархического механизма поведения:
- а. *Уровень целеполагания* – отвечает за определение Цели агента;
  - б. *Стратегический уровень* – определение Мишени, необходимый параметр – Цель агента;
  - с. *Тактический уровень* отвечает за движение к Мишени;
  - д. *Исполнительный уровень* отвечает за механические свойства агента как движущегося в здании объекта: ограничивает вектор скорости и его изменение исходя из физических характеристик агента, напр. максимальная скорость, ускорение, угловая скорость, а также обработка столкновений и т.п.

Следует отметить, что выделяемые в данном разделе уровни поведения (см. также [31]) не соотносятся жёстко с предлагаемым в работе методом рефлексивных разбиений и вообще с рефлексивным поведением агентов. Вводимые уровни просто достаточно удобный инструмент разбиения сложного поведения агентов на части, используемый во многих моделях поведения людей и роботов (см. например [105, 114, 120]).

Схематически архитектура агента представлена на Рис. 22, её реализация на Рис. 23.

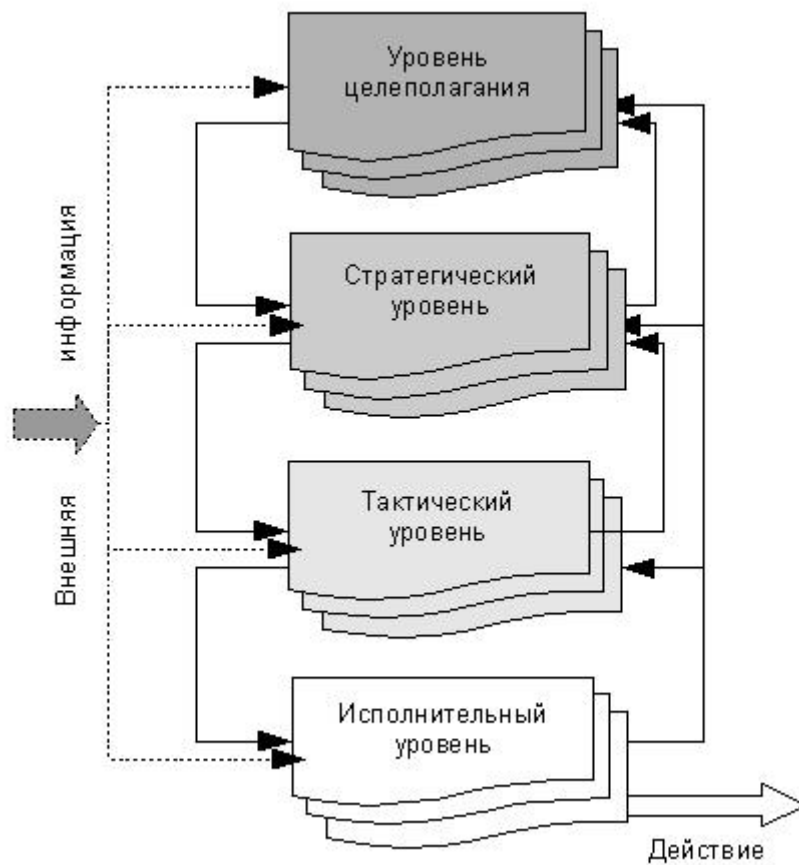


Рис. 22. Архитектура агента

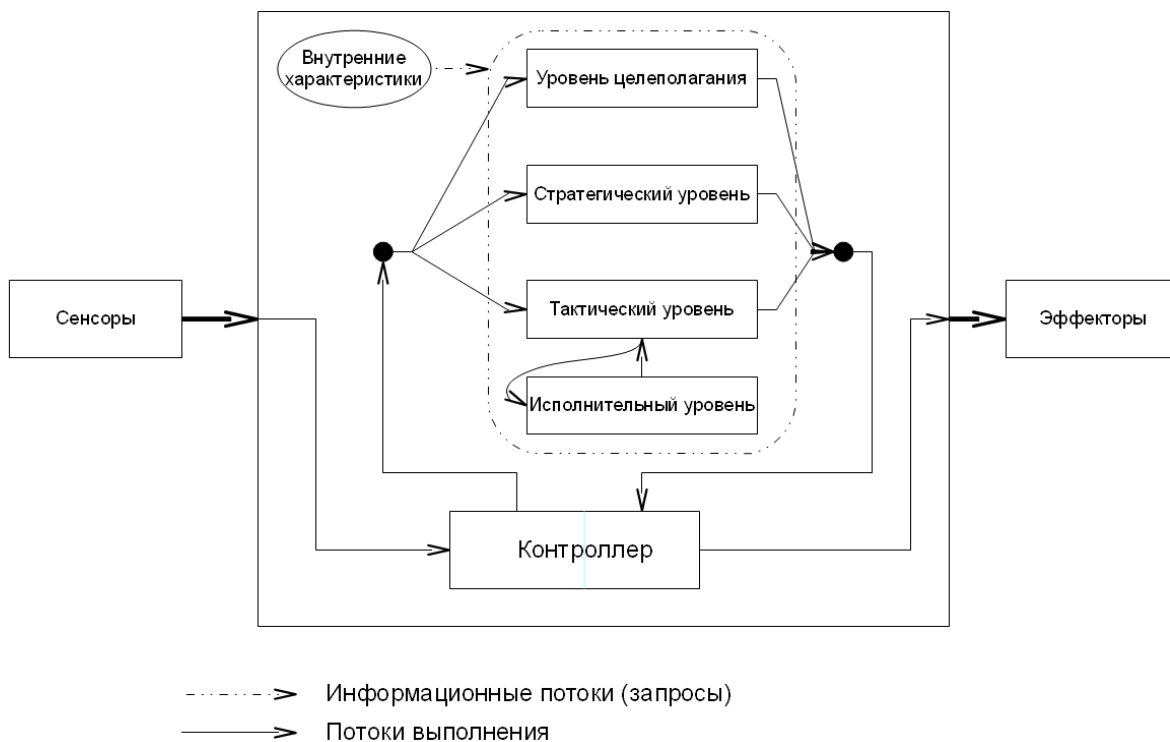


Рис. 23. Реализация архитектуры агента

Начнём описание механизма «снизу». Исполнительный уровень отвечает за реализацию механических свойств движения агента. Например, для исполнительного уровня, имитирующего человека, это: минимальные и максимальные ограничения вектора скорости, изменение вектора скорости с учётом величины ускорения, зависимость поворота вектора скорости от величины его модуля, возможно, реализация примитивов движения (по прямой, по дуге, остановка и т.п.). Команды имеют качественный вид: замедлиться/ускориться до определённой скорости или с заданным ускорением за промежуток времени, повернуть налево, повернуть направо на заданный угол и т.п. Исполнительный уровень используется напрямую в реализациях тактического поведения (см. Рис. 23).

Тактический уровень обеспечивает движение к Мишени внутри помещения, с заданными критериями эффективности.

- Входные данные: агент, помещение, в котором он находится, его Мишень и другие агенты в помещении;
- Результатом работы механизма тактического поведения является вектор скорости, который нужно выбрать для достижения Мишени, и который используется для выработки команд для оперативного уровня;

Стратегический уровень отвечает за выбор оптимальной последовательности помещений в здании, через которые агент должен пройти для достижения своей Цели.

- Входные данные: Цель агента, помещение в котором находится агент, здание (и его параметры: размеры помещений и переходов и т.п.), другие агенты и их модели стратегического поведения, представления самого агента;
- Результатом является дискретное распределение вероятности по имеющимся в помещении переходам. Мишень – переход, через который агент может пройти в следующее помещение для дальнейшего движения к Цели, выбирается как случайная величина с такой дискретной функцией распределения. С точки зрения теории игр, далее будем говорить, что стратегический уровень возвращает смешанную стратегию агента ([14] с.58) на множестве имеющихся в помещении переходов.

Вероятностный выбор помещения позволяет реализовывать как недетерминированное (неоднозначное поведение людей), так и детерминированное (вероятность равна 1 только для одного перехода) поведение агентов.

На уровне целеполагания, механизм задаёт правила выбора выхода, к которому идёт агент, то есть результатом этого уровня является Цель. Здесь реализованы простые механизмы: фиксированный выход и фиксированный набор выходов, в частности один или все.

Прежде чем перейти к рассмотрению конкретных механизмов поведения, представим используемую модель здания.

### 3.1.3. Модель здания

Будем считать, что моделируемое здание состоит из элементарных узлов двух типов: переход и помещение, с помощью которых можно представить все элементы здания. Например: двери, соединение коридора с холлом – это переход; сам коридор, холл и комната – помещения.

Учитывая наиболее распространенный в настоящее время метод строительства с использованием прямоугольных помещений и прямых углов, при анализе переходов между ними, будем далее считать следующее.

1. Помещения имеют строго прямоугольную форму, переходы между ними имеют ширину прохода, но не имеют длины.
2. Переходы ведут из помещения в помещение из центра стороны одного помещения к центру стороны другого помещения.
3. Помещения связаны между собой не более чем четырьмя переходами в направлениях, обозначаемых как «север», «юг», «запад», «восток» для удобства ориентирования. Обратим также внимание на то, что данные обозначения направлений не служат для обозначения сторон света.

Легко видеть, что большинство планов зданий представимо в виде приведенной схемы. В случае если не выполняется условие 2, помещение может быть разбито на несколько составляющих с вырожденными (шириной во все помещение) переходами. В некоторых же случаях совершенно неважно, в какой части стены помещения находится переход (например, в случае выхода из большого зала).

Таким образом, у перехода, как элемента модели, есть основной параметр «ширина». Помещение описывается такими базовыми параметрами, как: длина, ширина, количество человек в помещении. На их основе рассчитываются, например, плотность людского потока и объём помещения, координаты углов помещения для моделирования движения в нём агентов.

Для определённости далее выберем в качестве «ширины» расстояние между северным и южным переходом в помещении, а в качестве «длины» – между западным и восточным.

Так как здание разбивается на помещения и переходы, целям моделирования взаиморасположения помещений в здании отвечает нагруженный двудольный граф. Известно, что двудольным называется граф, у которого вершины могут быть двух типов (в данном случае: «помещение» и «переход»); типы вершин строго чередуются. Под «нагрузкой» графа понимается наличие у каждой вершины величины или набора величин, необходимых для отражения каких-либо свойств решаемой задачи.

Модель здания, в отличие от моделей произвольных графов, должна описывать взаимосвязь между помещениями, а также иметь геометрическую интерпретацию в силу представления здания, описанного здесь ранее.

Пусть модель здания представляется в виде графа  $G = (U, Q)$ , где  $U$  –

множество вершин графа типа «помещение»,  $Q$  – множество вершин типа «переход». Ребра графа ввиду его особой структуры являются частью вершин типа «помещение». Любая вершина  $u_i \in U$  графа  $G$  есть четвёрка  $u_i = (n_i, l_i, b_i, K_i)$ , где:  $n_i$  – множество агентов в помещении;  $l_i$  – «длина» помещения в обозначениях на стр.84;  $b_i$  – «ширина» помещения в тех же обозначениях;  $K_i = (k_i^{(3)}, k_i^{(B)}, k_i^{(C)}, k_i^{(IO)})$  – вектор направлений (запад – восток – север – юг), значения каждой компоненты которого есть индекс связанного перехода в множестве  $Q$  или ноль, если данное помещение не имеет перехода в соответствующем направлении.

Любой переход  $q_j \in Q$  графа  $G$  есть двойка  $q_j = (b_j, S_j)$ , где:

$b_j$  – «ширина» перехода;

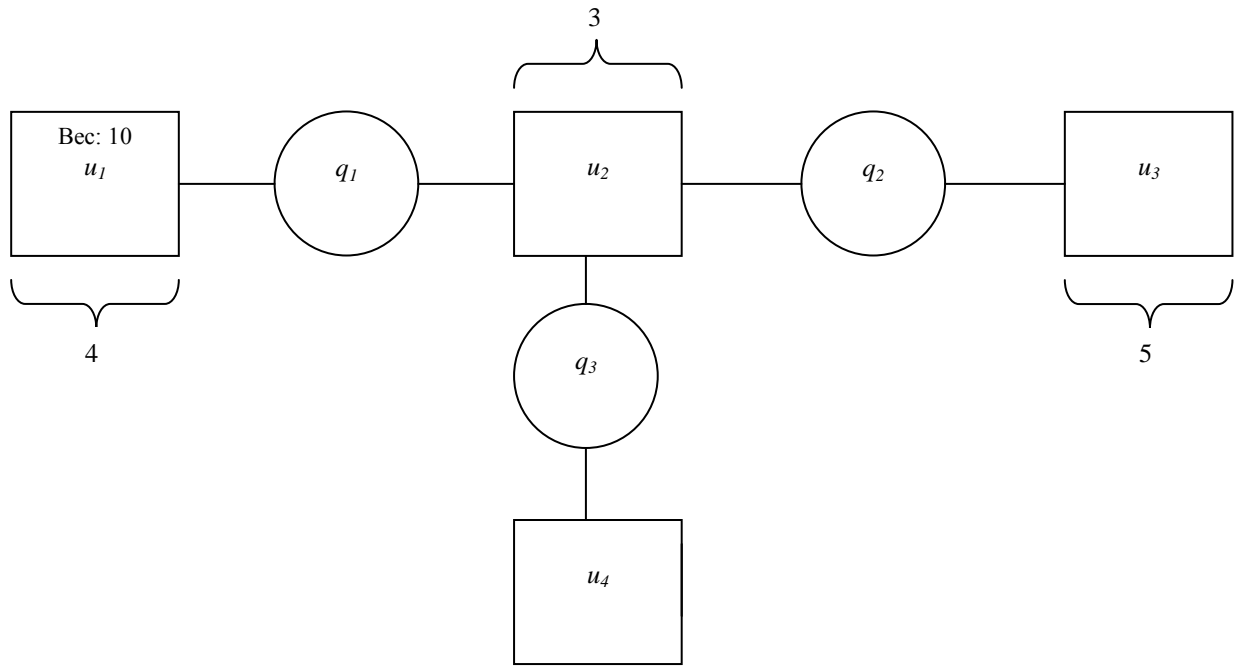
$S_j = (s_j^{(1)}, s_j^{(2)})$  – вектор ссылок на соседние помещения, значения каждой компоненты вектора есть индекс связанного перехода в множестве  $U$ ; принимать значение ноль компонента не может из-за связи графа с моделью здания, которая не предусматривает существование вершины-«перехода» со степенью равной единице (степень вершины – число рёбер, исходящих из вершины).

Основным свойством данного графа является то, что количество вхождений любой вершины-«перехода»  $q_j$  во множество компонент  $K_i$  всех вершин  $u_i$  равно двум. Это обусловлено тем, что у каждого перехода могут быть только две связи с помещениями, которые он объединяет. Причём, на одну вершину-«переход» индекс-переходы  $k_i^{(\dots)}$  ссылаются могут только парами:  $k_i^{(3)}-k_k^{(B)}$ ;  $k_i^{(C)}-k_k^{(IO)}$ .

Формально:

$$\forall q_j \in Q \exists! u_a, u_b \in U : (k_a^{(3)} = j = k_a^{(B)}) \vee (k_a^{(C)} = j = k_a^{(IO)}).$$

Пример двудольного графа в принятых выше обозначениях приведён на Рис.24.



*Рис.24. Пример двудольного графа для описания модели здания:  
 $u_i$  – вершины типа «помещение»,  $q_i$  – вершины типа «переход».*

Таким образом, мы описали модель здания в виде двудольного графа. Перейдём теперь к описанию архитектуры агента, используемой в системе моделирования.

### **3.2. Модели стратегического поведения**

После описания системы моделирования, перейдём к моделям и алгоритмам поведения агентов, используемым в прикладных имитационных задачах. В данном разделе представлены модели стратегического поведения агентов в здании.

#### **3.2.1. Статическая модель (от геометрических параметров здания)**

Особенность этой модели в том, что расчёт смешанных стратегий для каждого помещения можно провести до начала моделирования эвакуации (конкретного теста), так как геометрические параметры здания в тесте не меняются. Соответственно она и является статической, поскольку вероятности не зависят от изменяющихся в процессе моделирования эвакуации параметров.

Если рассматривать такое поведение с точки зрения реальности, то можно считать, что люди приблизительно знают здание и положение всех выходов в здании, поэтому могут дать оценку времени эвакуации через каждый выход, не учитывая влияние других агентов, захламлённости, задымлённости и т.п. факторов на время своего выхода. Соответственно была предложена модель, зависимость вероятности выбора данного перехода в которой, обратно пропорциональна квадрату расстояния от перехода до

выхода.

Алгоритм расчёта вероятностей для выбора перехода, подразумевает нахождение минимальных расстояний до выходов от каждого перехода в помещении, а потом считает итоговые вероятности. Рассмотрим алгоритм вычисления расстояния между двумя вершинами графа (помещениями).

### Алгоритм вычисления расстояния

Пусть зафиксирована некоторая вершина графа  $u_m$  типа «помещение». Определим расстояние  $\rho_i^{(m)}$  от вершины  $u_m$  до любого другого помещения графа –  $u_i$ , следующим рекурсивным образом.

Пусть функция  $I(u_i)$  – возвращает множество, содержащее индексы вершин типа «помещение», связанных с  $u_i$  через имеющиеся «переходы», информация о которых имеется в  $K_i$ :

$$I(u_i) = \{j \in U : (k_j^{(3)} = k_i^{(B)}) \vee (k_j^{(B)} = k_i^{(3)}) \vee (k_j^{(C)} = k_i^{(IO)}) \vee (k_j^{(IO)} = k_i^{(C)})\}$$

Пусть далее, функция  $H(u_i, u_m)$  – вычисляет расстояние от центра помещения  $u_i$  до центра помещения  $u_m$  соседнего с  $u_i$ :

$$H(u_i, u_m) = \begin{cases} l_i/2 + l_m/2, & \text{если } k_i^{(3)} = k_m^{(B)} \text{ или } k_i^{(B)} = k_m^{(3)} \\ b_i/2 + b_m/2, & \text{если } k_i^{(IO)} = k_m^{(C)} \text{ или } k_i^{(C)} = k_m^{(IO)}. \end{cases}$$

Тогда формула для  $\rho_i^{(m)}$  примет вид:

$$(50) \quad \rho_i^{(m)} = \begin{cases} 0, & i = m; \\ \min_{k \in I(u_i)} (\rho_k^{(m)} + H(u_i, u_k)), & i \neq m. \end{cases}$$

То есть расчет расстояния от первого помещения до второго выполняется по рекурсивной схеме.

Например, рассчитаем расстояние до помещения  $u_1$  из помещения  $u_3$  (Рис.24) при условии, что:  $l_1 = 4, l_2 = 3, l_3 = 5, b_3 = 3$ .

Согласно формуле (40)  $\rho_3^{(3)} = 0$ .

Рассчитаем расстояние  $\rho_2^{(3)}$ . Так как перемещение выполняется с востока на запад, то в вычислении участвует только длина помещения:

$$\rho_2^{(3)} = \rho_3^{(3)} + l_3/2 + l_2/2 = 0 + 5/2 + 3/2 = 4.$$

Рассчитаем расстояние  $\rho_1^{(3)}$ :

$$\rho_1^{(3)} = \rho_2^{(3)} + l_2/2 + l_1/2 = 4 + 3/2 + 4/2 = 7,5.$$

Вероятность выйти из данного помещения через данный переход, как упоминалось ранее, считается обратно пропорционально квадрату расстояния пути через данный переход до ближайшего выхода.

Для расчёта этих вероятностей в помещении  $u_i$  для каждого из его переходов последовательно вычисляются расстояния до всех имеющихся выходов и для каждого перехода выбирается минимальное из таких расстояний (то есть для каждого перехода может быть свой ближайший выход). Пусть  $r_i^0, r_i^1, r_i^2, r_i^3$  – минимальные расстояния до выходов через западный, восточный, северный и южный переходы соответственно. Для начала возведём их в (-2) степень, чтобы наименьшему пути соответствовали наибольшие вероятности с обратной квадратичной зависимостью:  $c^j = \frac{1}{(r_i^j)^2}, j = 0,1,2,3$ . Их сумма:  $R = \sum_{j=0}^3 c^j$ . Тогда

окончательные вероятности для «помещения»  $u_i$  вычисляются по следующей формуле:  $p_i^v = \frac{c^v}{R}$ ,  $\forall v \in \{0,1,2,3\}$ .

**Расширение модели.** Первым шагом к расширению данной модели было добавление идеи о стремлении агентов не к любому ближайшему выходу, а к конкретному, используя параметр агентов – «Цель», который указывал, к какому именно выходу двигается агент. Алгоритм вычисления был скорректирован: для каждого перехода в помещении сначала вычисляется самый короткий путь до Цели, итоговые вероятности считаются исходя из этих расстояний.

Если рассматривать типы данных то возникла структура, как на Рис. 25

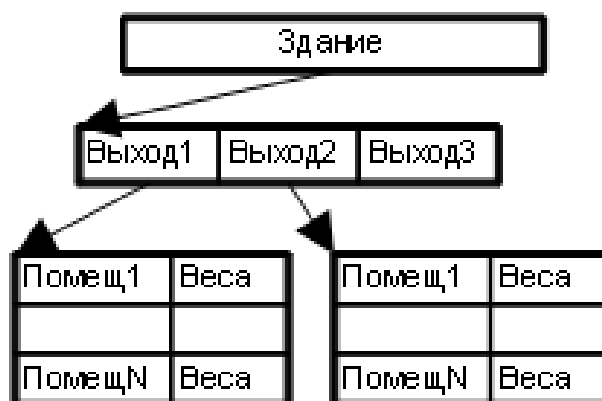


Рис. 25: Структура хранения вероятностей переходов

То есть для каждого выхода храним смешанную стратегию для каждого помещения. При этом в списке выходов есть специальный выход – общий, который является вариантом движения к ближайшему выходу, то есть содержит смешанные стратегии для базовой модели. Для получения весов агенту нужно передать два параметра: помещение, в котором он находится и его Цель.

Теперь «выходом» для агента можно назначить любое помещение, не обязательно по сути являющееся выходом.

Можно сделать ещё большее расширение: добавить различные модели поведения, зависящие от статических параметров, например не только от расстояния, но и от ширины проходов, освещённости, захламлённости и т.д.

Тогда структура будет такой, как приведена на Рис. 26:

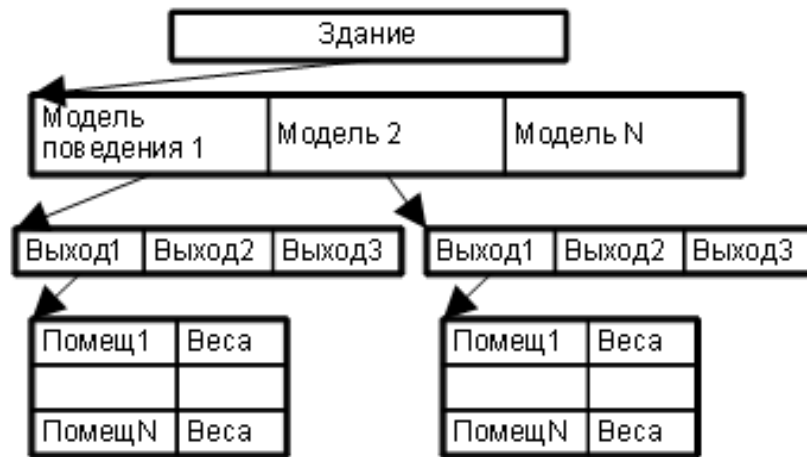


Рис. 26: Расширенная структура

### 3.2.2. Рефлексивная модель

Данная модель реализует модель рефлексивных размышлений для рефлексивного агента на стратегическом уровне. Данная модель будет использоваться для исследования влияния рефлексии при моделировании эвакуации в разделе 3.4.

Рефлексивная модель стратегического поведения не является статической, в отличие от предыдущей модели, так как вероятности выбора переходов зависят как от положения самого агента, принимающего решение, так и от положения остальных агентов в помещении на момент принятия решения. В виду сложности вычисления целевой функции, данная модель используется только в нескольких примерах, в которых целевая функция определяется путём сбора статистики для каждого конкретного случая.

Рефлексивность заключается в способности агента учитывать своё знание о поведении остальных агентов: агенты 0-го ранга используют «Статическую» модель (т.е. не рефлексивную), агент 1-го ранга считает, что все остальные агенты имеют 0-ой ранг, а агент ранга 2 знает количество агентов 1-го ранга и их модель поведения, а модель поведения всех остальных считает «Статической» и так далее для больших рангов.

Рефлексивная модель как любая модель стратегического поведения должна возвращать смешанную стратегию – дискретное вероятностное распределение на множестве возможных переходов для агента в каждый момент времени (в произвольном помещении и при произвольном расположении агентов).

Для расчёта вероятностей выбора переходов используется следующая информация: множество, количество и местоположение всех агентов на момент принятия решения –  $N, n, P = \{(x_i, y_i)_{i \in N}\}$ , а также то, куда пойдут все остальные агенты – распределение агентов по переходам  $\{n_3, n_B, n_C, n_{Ю}\}$ ,  $n_3 + n_B + n_C + n_{Ю} = n$ .

Действие  $a_i$  агент  $i$  выбирает из дискретного множества  $A = \{З, В, С, Ю\}$  – варианты выбора переходов в помещении – Западный, Восточный,

Северный и Южный.

Состояние системы в данной модели (в терминах общей игровой модели с. 47) – положение агентов в помещении и их распределение по переходам. Для агента, принимающего решение, состояние зависит от его выбора  $x(a_i) = (P, n_3(a_i), n_B(a_i), n_C(a_i), n_{Ю}(a_i))$ .

Положение агентов является доступной информацией, а распределение по переходам рефлексивный агент получает исходя из своих рефлексивных знаний: применяет известные ему модели поведения к остальным агентам в помещении в соответствии с рангами рефлексии, которые он им приписывает. Таким образом, он может предсказать (возможно неправильно) выбираемые остальными агентами переходы и получает значения для  $a_j, j \neq i$ . Тогда  $n_b = \#\{j \in N \mid a_j = b\}, \forall b \in A$ .

У агента должна быть возможность оценить своё время выхода в зависимости от выбранного действия и действий остальных агентов – целевая функция. В общем виде целевая функция и её аргументы выглядят так:  $f_i(a_i, x(a_i))$ . Зная эту функцию, рефлексивный агент рационально выбирает своё действие, см. параграф 2.1.2.

Чтобы получить функциональную зависимость для целевой функции в примерах раздела 3.4 в имитационной среде проводится сбор статистики, при следующих упрощениях модели: пусть состояние системы  $x(a_i) = (x, y, n_3(a_i), n_B(a_i), n_C(a_i), n_{Ю}(a_i))$ , где  $x$  и  $y$  – координаты агента  $i$ , а другие агенты располагаются в помещении по некоторому фиксированному двумерному закону распределения  $\mathcal{F}(\cdot, \cdot)$ . Таким образом, мы заменили  $P$ , на  $(x, y, \mathcal{F}(\cdot, \cdot))$ , что уместно при следующих условиях:

- $n \gg 1$ ,
- в начальный момент времени агенты расположены в помещении в соответствии с распределением  $\mathcal{F}(\cdot, \cdot)$ ,
- агенты принимают решение о дальнейшем пути к выходу только в начальный момент времени.

Тогда целевая функция должна устанавливать следующее соответствие для каждого помещения, в котором находятся агенты:

$$(51) \quad (a, (x, y, n_3(a), n_B(a), n_C(a), n_{Ю}(a))) \rightarrow t,$$

где  $t$  – время выхода агента, находящегося в местоположении  $(x, y)$  и выбравшего переход  $a \in A$ , при условии, что количества агентов, выбравших соответствующие переходы равны, соответственно,  $n_3(a), n_B(a), n_C(a), n_{Ю}(a)$ .

Поскольку для всех значений координат агента в помещении, получить зависимость (51) тестированием невозможно, необходима интерполяционная функция  $z(\cdot, \cdot)$ , для задания которой нужно определённое минимальное количество пар  $(x, y)$ , но чем больше, тем лучше. Пусть  $Z$  – такое количество пар.

Тогда минимальное количество тестов для задания целевой функции:  $C_{min} = C_{n+k-1}^n * Z / n$ , где  $C_{n+k-1}^n$  – число сочетаний из  $(n+k-1)$  по  $n$  – число размещений  $n$  агентов по  $k$  переходам. Делить на  $n$  нужно, так как каждый тест даёт результаты для  $n$  начальных положений агентов (поскольку

моделирует движение сразу  $n$  агентов).

Учитывая это, для случая с простым зданием (параграф 3.4.2)  $C_{min} = 101 * 225 / 100 = 227.25$ , где  $C_{100+2-1}^{100} = 101$  – количество размещений 100 агентов по 2 выходам, 225 – параметр  $Z$ , 100 – количество агентов. В реальности было проведено  $101 * 20 = 2020$  тестов, где 20 это количество тестов эвакуации, получаемые  $20 * 100$  координат использовались для задания интерполяционной функции  $z(\cdot, \cdot)$ .

Зависимостью целевой функции от начальных позиций остальных агентов –  $P$  (кроме начальной позиции принимающего решения), как упоминалось выше, решили пренебречь, используя определённый вид распределения случайных начальных положений агентов. Тогда время эвакуации зависит от вида распределения случайной величины или, иначе говоря, от конфигурации расположения остальных агентов: при равномерном распределении – толпа равномерной плотности, при нормальном распределении – толпа с сильным уплотнением в центре.

Ещё более простая модель без учёта начального положения агента была исследована в работах [26, 27]. Она показала особенность данной модели рефлексии, связанную с одинаковой информированностью агентов одного ранга: в этих моделях все рефлексизирующие агенты одного ранга идут в один выход, поскольку считают вероятность выхода одинаково – считают всех остальных агентов «глупее» себя. Это следствие отсутствия общего знания рефлексивных агентов друг о друге и неучёта начального положения.

### **3.2.3. Стратегия без возврата**

Стратегия представляет собой надстройку над любой другой стратегией (не используется отдельно). Смешанная стратегия для агента, использующего данную стратегию-надстройку, не содержит варианта возвращаться в помещение, из которого он непосредственно пришёл в текущее помещение. Соответственно, если в здании нет циклов, то ни один агент не будет «блуждать» по зданию. При наличии циклов в здании «блуждание» по нему не исключается.

Данная стратегия не статическая. Достаточно одного класса для всех агентов, потому что она хранит в себе таблицу соответствий агент-помещение, чтобы определять, в каком предыдущем помещении был агент.

## **3.3. Модели тактического поведения**

Выше мы описали модели стратегического поведения агентов, в данном разделе опишем базовые модели тактического поведения агентов, то есть правила, которыми они руководствуются при непосредственном движении к своим Мишеням внутри помещений. Описанные модели есть развитие и более подробное описание моделей, предложенных в [30].

Модели тактического поведения позволяют агентам преодолевать препятствия – столкновения с другими агентами в помещениях, а также эффективно двигаться к Мишени. Параметр, на который непосредственно воздействуют данные правила – вектор скорости агента. Тестирование не-

скольких моделей данного типа проведено в [28].

Общая схема моделей тактического поведения представлена на Рис. 27. На рисунке показаны исходные модели: «прямое нацеливание», «уход от столкновений» и «сканирующая тактика», а также составные: «простой тип», «базовый тип», порядок составления которых из исходных указан стрелками.

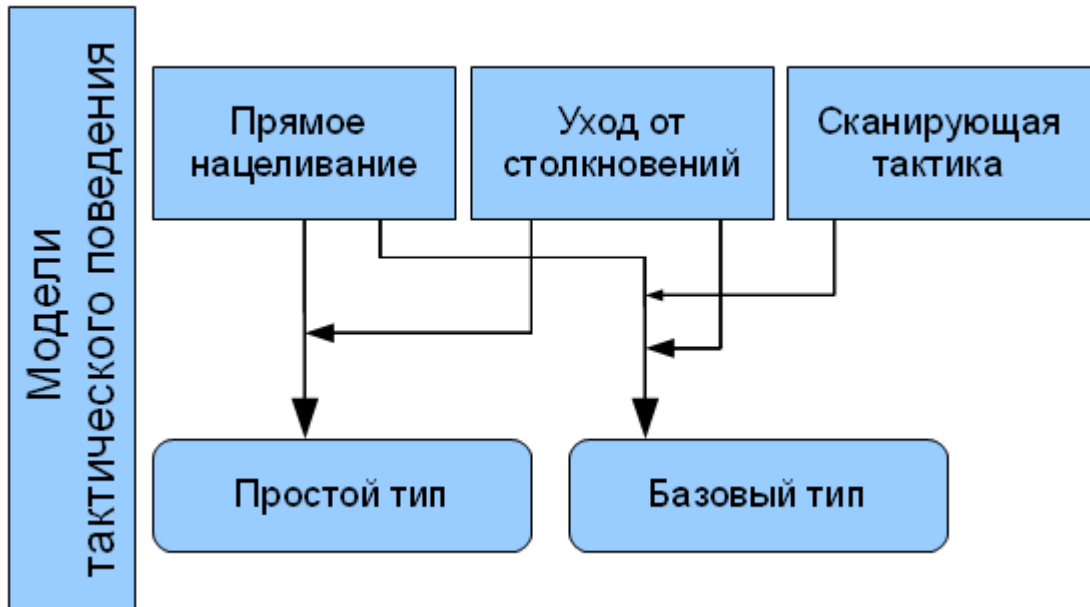


Рис. 27 Схема моделей тактического поведения

### 3.3.1. Прямое нацеливание

Данное поведение – просто прицеливание на выход агента. Рассмотрим агента в помещении на Рис. 28.

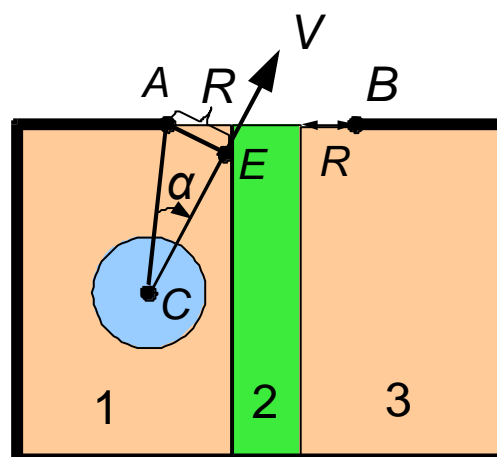


Рис. 28 Прямое нацеливание

Если точка  $C$  (его положение) находится в зоне, отмеченной зелёным цветом, то достаточно направить скорость к  $AB$  перпендикулярно.

В красной зоне агент сможет пройти в переход  $AB$ , если перпендикуляр к скорости  $V - AE$  будет равен радиусу агента (на самом деле  $AE$  должен быть не меньше, но мы рассматриваем оптимальный вариант). Поэтому берём направление скорости сначала на точку  $A$ , а потом отклоняем его на угол  $\alpha$ , синус которого находится с помощью известных величин  $AE$  и  $AC$  (знак  $\alpha$  выбирается в зависимости от выбранного перехода и того его конца, который агент стремится обойти).

*Алгоритм.*

Входные данные: векторы  $A(A_x, A_y)$ ,  $B(B_x, B_y)$ ,  $C(C_x, C_y)$

Выходные данные: вектор  $V(x, y)$

1. Определяем, в какой зоне находится агент. Пусть  $C'$  – проекция  $C$  на прямую  $AB$ . Если  $C'x < Ax + R$  то агент в зоне 1, Иначе, если  $C'x < Bx - R$  то в зоне 3, Иначе агент в зоне 2.
2. Действия для зоны 1.
  - а. Направляем вектор  $V$  на точку  $A$
  - б. Находим угол  $\alpha$ , зная его синус:  $\sin(\alpha) = R / |AC|$
  - с. Поворачиваем  $V$  на угол  $\alpha$
3. Действия для зоны 2.
  - а. Поворачиваем  $V$  перпендикулярно к  $AB$  (в направлении к  $AB$ )
4. Действия для зоны 3 аналогичны действиям для зоны 1.

### 3.3.2. Уход от столкновений

Данное поведение основывается на уходе от ближайших столкновений агентов. Дадим последовательно методику ухода: нахождение столкновения, определение типа столкновения и определение правил ухода.

#### *Нахождение столкновения*

Рассмотрим двух агентов, которые через некоторое время (при неизменных скоростях) столкнутся.

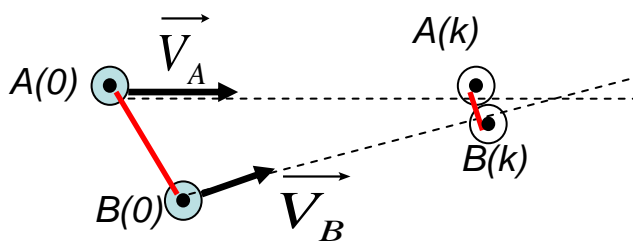


Рис. 29 Нахождение столкновения

Считаем скорость постоянной, потому что мы не знаем, как на самом деле поведут себя агенты и можем сделать только такое предположение. Необходимо найти время столкновения этих агентов (по которому можно найти координаты агентов при столкновении).

Нас интересует длина отрезка  $AB$  (см. Рис. 29) в моменты времени  $k$ , точнее длина отрезка  $A(k)B(k)$ .

Формулы зависимости координат  $A$  и  $B$  от  $k$  выглядят так:

$$A(k) = A(0) + k * \vec{V}_A; \quad B(k) = B(0) + k * \vec{V}_B$$

Тогда, для нахождения искомого времени решаем квадратное уравнение относительно  $k$  (после возведения в квадрат):

$$\|A(k)B(k)\|_{\mathbb{R}^2} = 2 * R$$

### Параметры столкновения

Для определения вида столкновения, а также для дальнейших методов ухода от них, выделим следующие ключевые параметры:

- Вектор  $d$  – перпендикулярен  $AB$ . Выбирается сонаправленным вектору скорости  $V$  агента  $A$  или  $B$ , рассматриваемого в контексте. Алгоритм (входные данные:  $AB(a, b)$ ,  $V$ , выходные: вектор  $d(d1, d2)$ ):  
 1 Пусть  $d = (b, -a)$ , тогда  $(d, AB) = 0$   
 2 Если  $(d, V) < 0$  то  $d = d * (-1)$
- Углы  $\alpha$  и  $\beta$  (между  $V_A$  и  $d$  и  $V_B$  и  $d$ , соответственно).
- Положение векторов  $V_A$  и  $V_B$  относительно отрезка  $AB$ : одностороннее (рис. Рис. 30.1а) или нет (рис. Рис. 30.1б) – логический параметр *OneSided*.
- $D_A, D_B$  – расстояния от точки  $B$  до прямой  $AC$  и от точки  $A$  до прямой  $BC$ , соответственно (Рис. 30.2). Вычисляются с помощью углов  $\alpha$  и  $\beta$ :  
 $D_A = 2 R \cos(|\alpha|)$ ;  $D_B = 2 R \cos(|\beta|)$ .

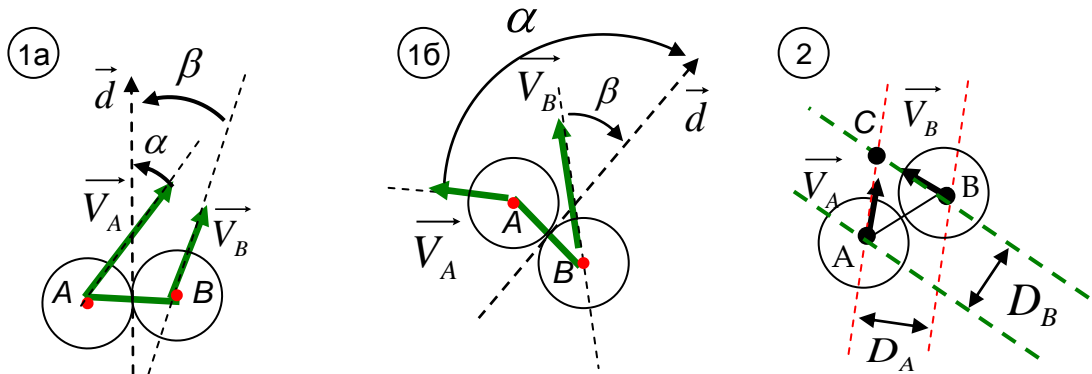


Рис. 30 Параметры столкновений

### Типы столкновений

При анализе столкновений, были выделены два их типа:

1. Неизбежно изменение скорости только одного агента. В этом случае впередиидущий может двигаться, в итоге оба смогут достичь до цели, Рис.31.1
2. Возможно изменение скоростей двух агентов, Рис.31.2а. В этом случае, без изменения скорости, оба агента могут остановиться, Рис.31.2б

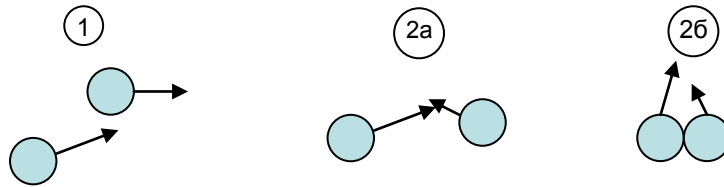


Рис.31 Типы столкновений

**Определение типа столкновения**

Необходимое и достаточное условие для определения столкновения 1-го типа:

$$\alpha * \beta \geq 0,$$

то есть  $\alpha$  и  $\beta$  одного знака

Необходимое и достаточное условие для определения столкновения 2-го типа:

$$\alpha * \beta < 0 \text{ (разных знаков)}$$

**Уход от столкновения 1-го типа**

Можно выделить две различные ситуации столкновения 1-го типа, они представлены на Рис.32

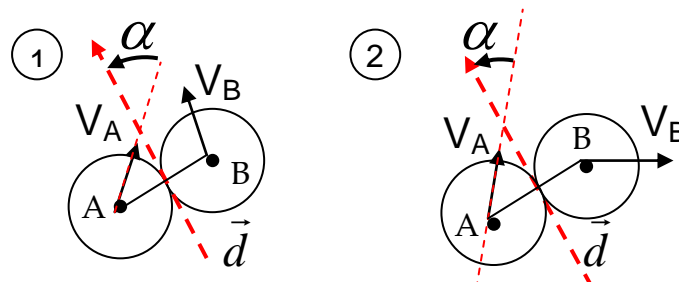


Рис.32 Столкновение 1-го типа

Первая ситуация – односторонняя. В этом случае, если  $D_A$  близко к  $2R$  целесообразно увеличивать  $V_A$  и/или отклоняться на  $\alpha$ , этим можно обеспечить своё преимущество, то есть перевести столкновение к ситуации 1-го типа, но с ведущим агентом  $A$  (но ситуация может перейти и к столкновению 2-го типа); если  $D_A$  близко к  $0$ , уменьшать  $V_A$  и/или отклоняться на угол  $\alpha - \text{sign}(\alpha)\pi$ , при такой тактике, мы пропускаем агента  $B$ .

Вторая ситуация – разносторонняя. Здесь, если  $D_A$  близко к  $2R$ , то целесообразно уменьшать  $V_A$  и/или отклоняться на  $\alpha$ , этим мы пропустим агента  $B$ ; если  $D_A$  близко к  $0$ , увеличивать  $V_A$  или отклоняться на угол  $\alpha - \text{sign}(\alpha)\pi$ , опять же, в таком случае можем стать ведущими в одностороннем столкновении 1-го типа, а можем привести столкновение ко 2-му типу.

**Уход от столкновения 2-го типа**

Опять можно выделить две различные ситуации столкновения 2-го типа, представленные на Рис.33

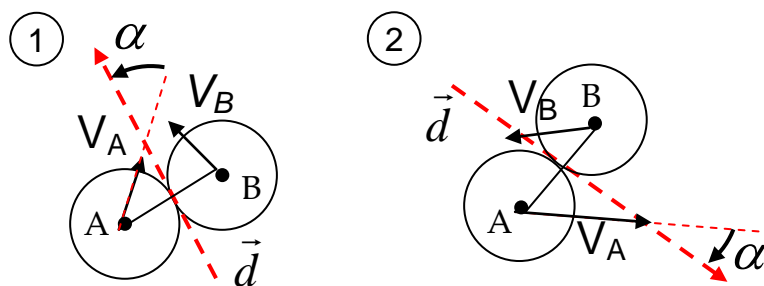


Рис.33 Столкновение 2-го типа

Будем говорить, что агент  $A$  имеет преимущество перед агентом  $B$ , если  $D_A > D_B$ .

Тогда в ситуации 1 – правила такие же как в столкновении 1-го типа, но возможность с увеличением скорости агент  $A$  реализует, если есть преимущество, иначе реализует пропускающее поведение.

В ситуации 2 – разумно отклоняться на  $\alpha$ . Увеличивать  $V_A$ , только если есть преимущество.

### 3.3.3. Сканирующая тактика

Предыдущее поведение позволяло уйти от непосредственного столкновения с другим агентом. Но в той модели агент не «видит» толпы, он видит только других агентов и реагирует на ближайшего. Введём правила, позволяющие агенту обходить скопления и выбирать некоторый оптимальный маршрут.

Рассмотрим Рис. 34

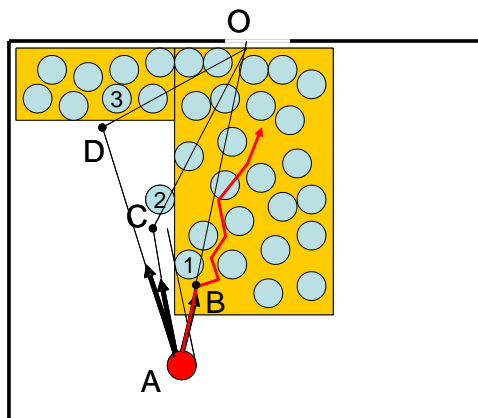


Рис. 34 Сканирующая тактика

Алгоритм:

1. Находим столкновение по текущему вектору скорости – это столкновение с агентом 1. Запоминаем точку  $B$  и скорость агента 1.
2. Отклоняем вектор скорости  $V$  так, чтобы столкновения с агентом 1 не было, находим ближайшее столкновение (с агентом 2), запоминаем точку  $C$  и скорость агента 2.
3. Далее находим ещё столкновения (с агентом 3)

4. Получаем пути  $ABO$ ,  $ACO$ ,  $ADO$  и так далее, ограничить поиск можно ограничением отклонения скорости (назад возвращаться не нужно), нахождением столкновения со стеной, а также учесть, что люди не склонны далеко обходить препятствие, пускай и времени на это уйдёт меньше (в ситуации эвакуации).
5. Время прохождения оцениваем по скорости (зная длины путей): скорость на участках пути  $A[B|C|D]$  – максимальна, скорость на  $[B|C|D]O$  – оцениваем по скорости агента, с которым сталкиваемся. Выбираем наименьший путь и соответствующее направление скорости.

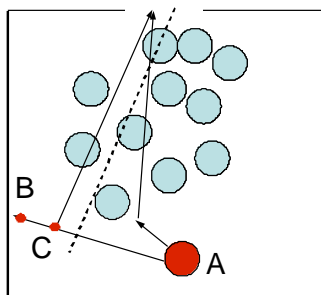


Рис. 35 Точки видимости

Для ограничения поиска (п.4 алгоритма) было сделано два ограничения: ограничение по углу обзора – только 180 градусов и алгоритм нахождения точек видимости – таких точек, при перемещении в которые можно дойти до мишени прямо: при достижении стены, выбираем не точку столкновения со стеной, а точку  $C$  от которой можно двигаться не задев агента мешавшего на предыдущем шаге (Рис. 35).

### 3.3.4. Составные тактические типы

В программе реализованы представленные модели, но сами по себе они не дают ценный результат, поэтому были добавлены составные:

Реализованы следующие составные типы тактик:

1. Базовый тип тактики. Структурно состоит из последовательного применения тактик прямого нацеливания, ухода от скоплений, ухода от столкновений с другими агентами (вероятностного).
2. Простой тип тактики. Последовательно применяется тактика прямого нацеливания, а потом – тактика ухода от столкновений (вероятностная).
3. Два остальных типа тактик «Базовая детерминированная» и «Простая детерминированная» – такие же, как первые два, но детерминированные. В них агент одним, определённым образом выбирает, как он будет избегать столкновения.

### 3.4. Имитационное моделирование эвакуации

В настоящем разделе проводятся исследования рефлексивного

поведения агентов на примере эвакуации людей из здания с помощью имитационного агентного моделирования. В параграфе 3.4.1 ставится общая задача для моделирования. В следующем параграфе 3.4.2 ставится эксперимент тестирования теоретического пример с простым зданием, результаты которого представлены в параграфе 3.4.3.

### **3.4.1. Постановка задачи**

Будет исследовано простое здание на предмет влияния рефлексивных агентов в нём при эвакуации.

Типов агентов всего три:

- «обычные» агенты, идущие к определённому (ближайшему) выходу;
- рефлексивные агенты первого ранга;
- рефлексивные агенты второго ранга.

«Обычные» агенты это агенты, использующие модель стратегического поведения «Статическая модель» (параграф 3.2.1) с надстройкой «Без возврата» (параграф 3.2.3). Рефлексивные агенты используют модель «Рефлексивная модель» (параграф 3.2.2) с надстройкой «Без возврата».

Ключевой критерий, по которому анализируются результаты – время эвакуации всех агентов из здания.

В каждом примере меняются доли агентов первого и/или второго рефлексивного ранга. Как и в каких помещениях производится изменение, будет описано в каждом конкретном случае.

Задачей исследования в данной модели является исследование влияния доли рефлексивных агентов на время эвакуации из здания. Попутной задачей является качественное исследование поведения агентов с предложенными алгоритмами тактического поведения – похожи ли они на поведение людей. Обнадёживающим опытом в этом направлении является работа [128], а в особенности недавняя работа [112], посвящённая исследованию похожих моделей поведения человека, основанных на рациональности в выборе скорости и направления движения.

### **3.4.2. Тестирование теоретического примера**

Исследуется пример, по условиям совпадающий с предположениями в математической модели эвакуации из параграфа 2.6.5. Напомним, что там предполагается присутствие двух путей эвакуации, один из которых ближе, чем другой. Теоретически показано, что применение агентов ранга выше 1 для минимизации времени эвакуации из здания нецелесообразно.

Итак, для моделирования возьмём здание, представленное на Рис. 36, параметры здания даны в Таблица 1.

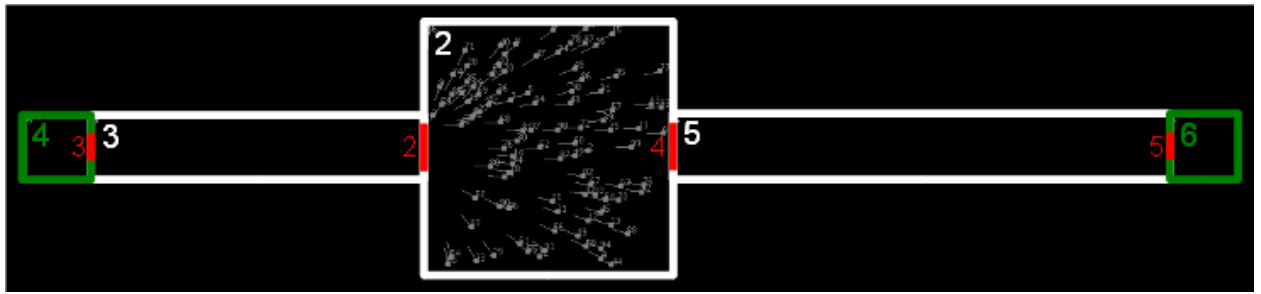


Рис. 36 Здание для тестирования теоретического примера

Таблица 1 Параметры здания для тестирования теоретического примера

Объект	длина,м × ширина,м
Размер здания	65×15
Помещение № 2	15×15
Помещение № 3	20×4
Помещение № 5	30×4
Переходы №№ 2,4(красным цветом)	3
Переходы №№ 3,5	1

Описанное здание характерно тем, что если 99 агентов идут в ближний выход, то одному агенту выгодней идти в дальний выход, поскольку все остальные создадут пробку в ближнем выходе. Это условие определяет поведение рефлексивных агентов 1 ранга: поскольку каждый такой агент считает всех остальных «обычными», постольку почти все они пойдут в дальний выход. Исключение составляют агенты 1-го ранга, близко расположенные к ближнему выходу – они успеют выйти в него до образования пробки.

Итак, поведение рефлексивных агентов 1-го и 2-го рангов известно. Все агенты размещаются в помещении № 2.

### Первый эксперимент

Цели эксперимента:

- i. Получение функций зависимости времени эвакуации от количества агентов для каждого выхода в отдельности;
- ii. Сбор информации по поведению агентов 0 («обычных») и 1 рангов;
- iii. Получение необходимых данных для задания поведения агентов 2 ранга.

В первом эксперименте используются только агенты первого ранга и «обычные». Всего агентов в помещении №2 – 100. Располагаются в случайном порядке внутри помещения.

Параметр исследования в данном эксперименте – доля рефлексивных агентов 1 ранга. Диапазон изменения – последовательно от 0 (все «обычные») до 100 (все 1 ранга), общее количество агентов постоянно и равно 100.

Собранные статистикой данные содержат координаты начального

положения агента в помещении № 2, а также время его выхода в определённый выход (для «обычных» - ближний, для агентов 1 ранга - дальний).

Результатом эксперимента получаем необходимые данные, для определения целевой функции – смотрите формулу (51) параграфа 3.2.2.

Полученные данные достаточны для определения поведения агентов второго ранга, которым известно соотношение долей агентов, то есть достигается цель iii. Также полученные данные позволяют достичь i и ii цели эксперимента – получение данных, характеризующих используемое здание и модели поведения агентов.

#### Второй эксперимент

Цель – показать нецелесообразность использования агентов 2 ранга для задачи минимизации времени эвакуации.

В данном эксперименте моделируется эвакуация при наличии трёх типов агентов: 0-, 1- и 2-го рангов.

Агентов в помещении №2 – 100 в каждом исследовании. Параметров исследования 2: количество агентов 1 ранга, обозначим его  $n_1$  и количество агентов 2 ранга –  $n_2$ . Область их изменения – треугольник:  $\{ (n_1, n_2) \in \mathbb{N}^2: n_1 \leq 100 - n_2 \}$ .

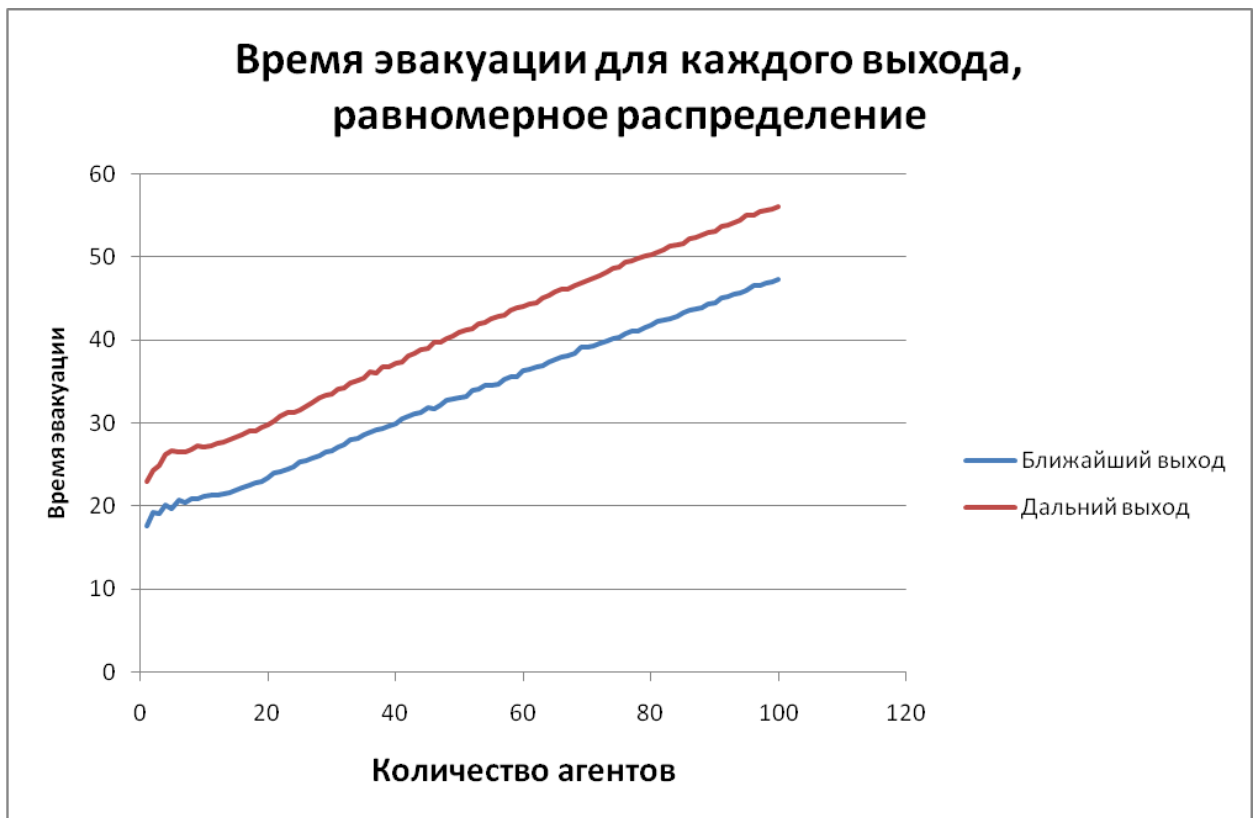
Собранные статистикой данные содержат времена эвакуации отдельно для ближайшего и дальнего выходов, соответственно, для каждого соотношения количества агентов. Взяв максимум от времён эвакуации в каждый выход, получим время общей эвакуации из здания для каждого соотношения количества агентов.

Результат эксперимента – график зависимости времени эвакуации от количества агентов 0, 1 и 2 рангов, показывающий, что цель эксперимента достигнута – с помощью агентов 2 ранга не удаётся уменьшить минимальное время эвакуации, достигнутое при использовании агентов 0 и 1 рангов.

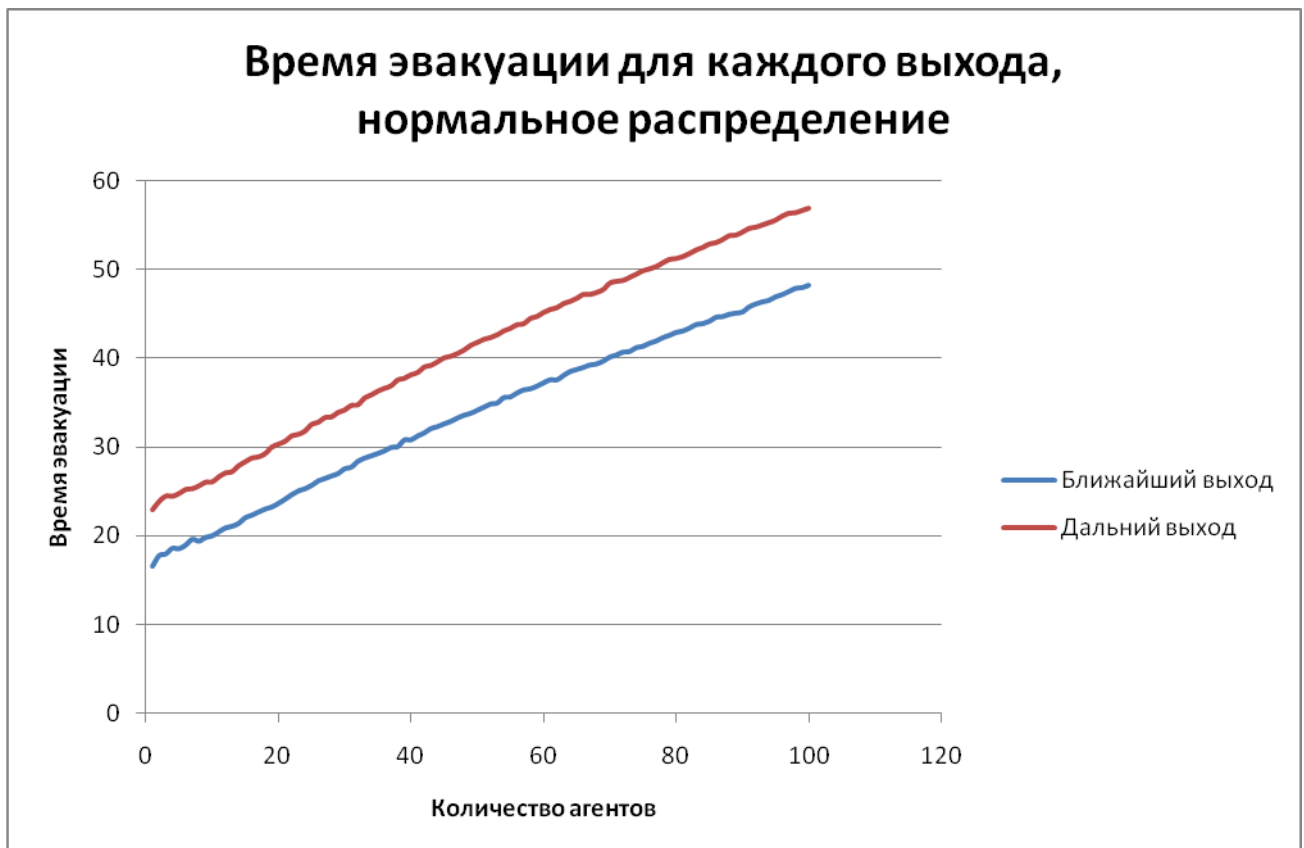
### **3.4.3. Результаты тестирования теоретического примера с простым зданием**

#### Первый эксперимент

Приведём полученные графики зависимости времени эвакуации от количества агентов для каждого выхода: при равномерном распределении – Рис. 37, при нормальном распределении – Рис. 38.



*Рис. 37* Время эвакуации для каждого выхода при равномерном распределении агентов



*Рис. 38* Время эвакуации для каждого выхода при нормальном распределении агентов

Во-первых, графики практически идентичны и не зависят от вида распределения начальных позиций.

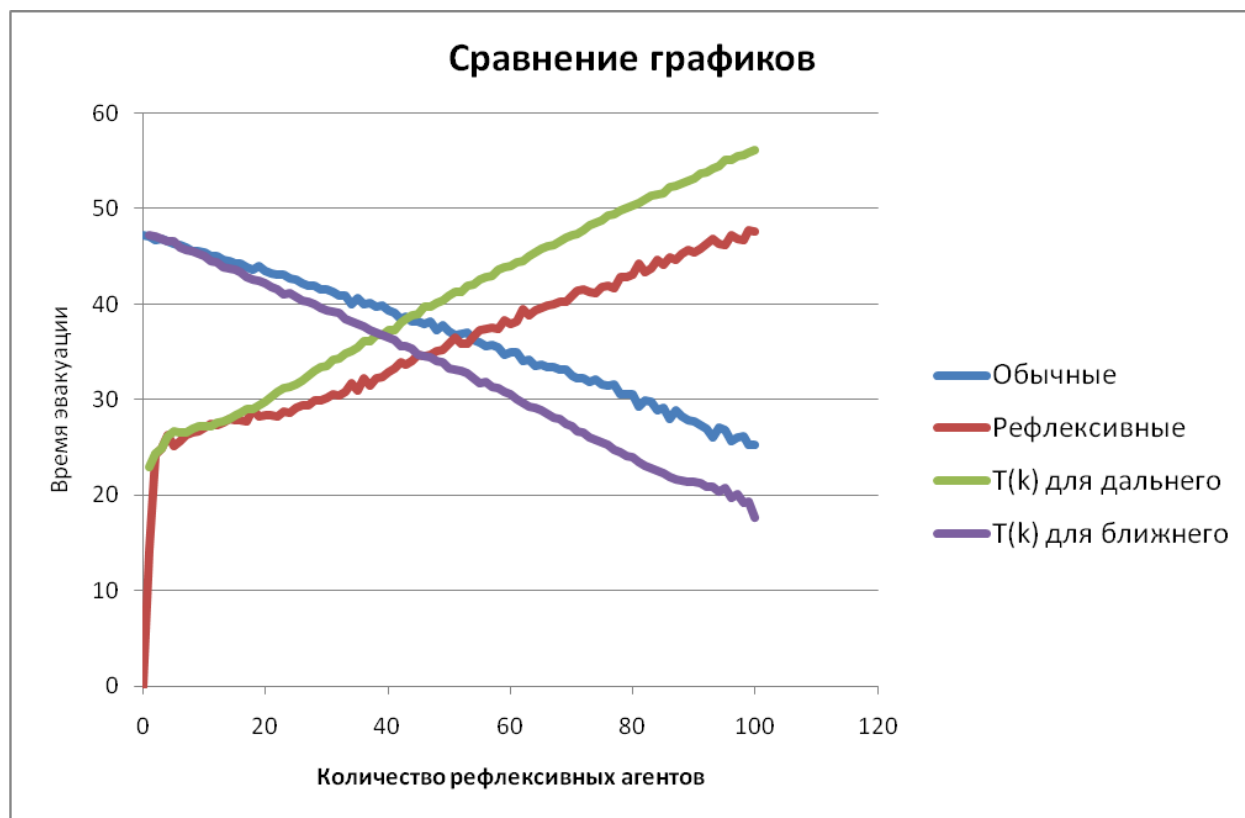
Во-вторых, сразу заметно отсутствие у зависимостей свойства выпуклости и вогнутости, точнее, сначала функции выпуклы, затем вогнуты. Также заметен рост разности между временами эвакуации двух выходов с ростом количества выходящих агентов.

Зная эти зависимости, в соответствии с формулой (46), найдём оптимальное число  $n_1^*$  рефлексивных агентов 1-го ранга по теоретическим предположениям из параграфа 2.6.5:

$$(52) \quad n_1^* = 38,$$

при равномерном распределении.

Результаты моделирования эвакуации с агентами 1-го ранга представлены на *Рис. 39* синей и красной линией. Зелёная и фиолетовая линии представлены для сравнения – это зависимости времени эвакуации для каждого выхода в отдельности в случае равномерного распределения (см. *Рис. 37*).



*Рис. 39* Зависимость времени эвакуации от числа рефлексивных агентов 1 ранга

По результатам данного моделирования получаем, что оптимальное количество агентов 1-го ранга находится в отрезке [51,54], что не соответствует значению (52).

Однако, как отмечалось в этом разделе, не все агенты 1-го ранга идут в дальний выход. Собранные данные позволяют определить графики времени

эвакуации для количества агентов 1-го ранга, выбравших дальний выход, полученные графики представлены на Рис. 40.

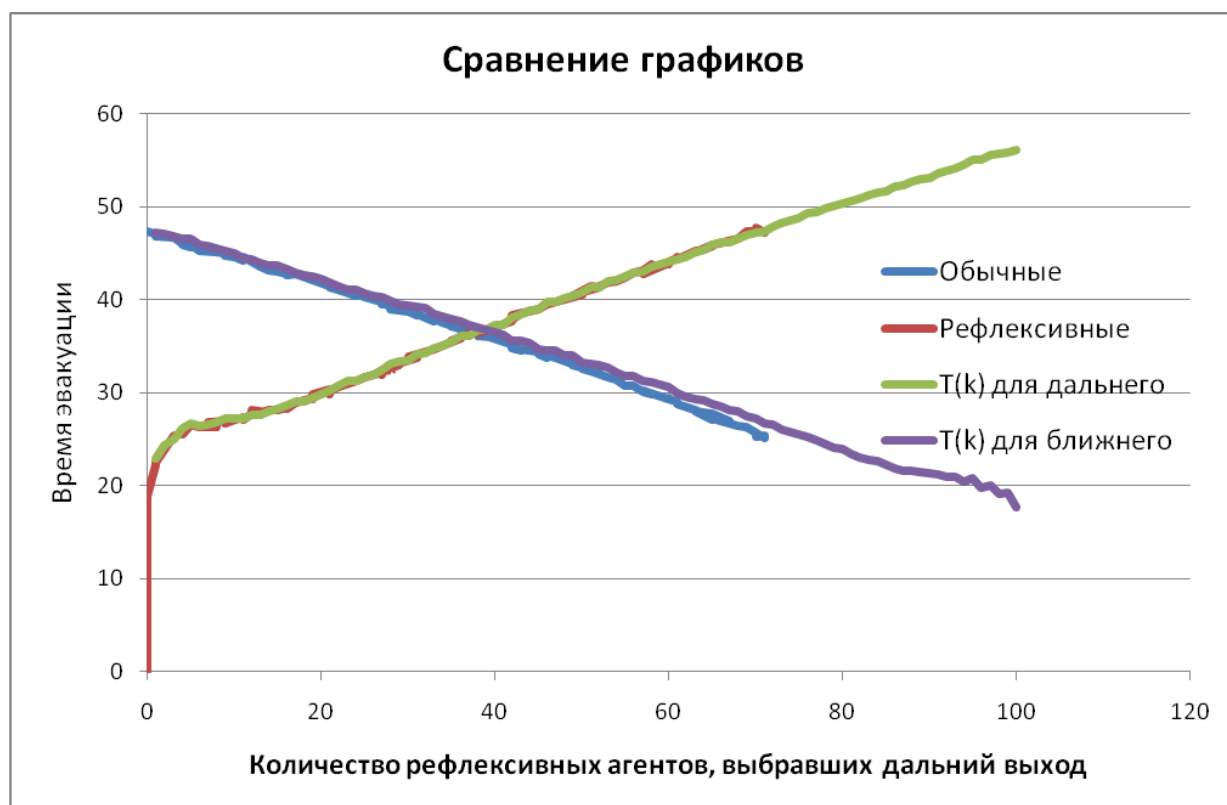
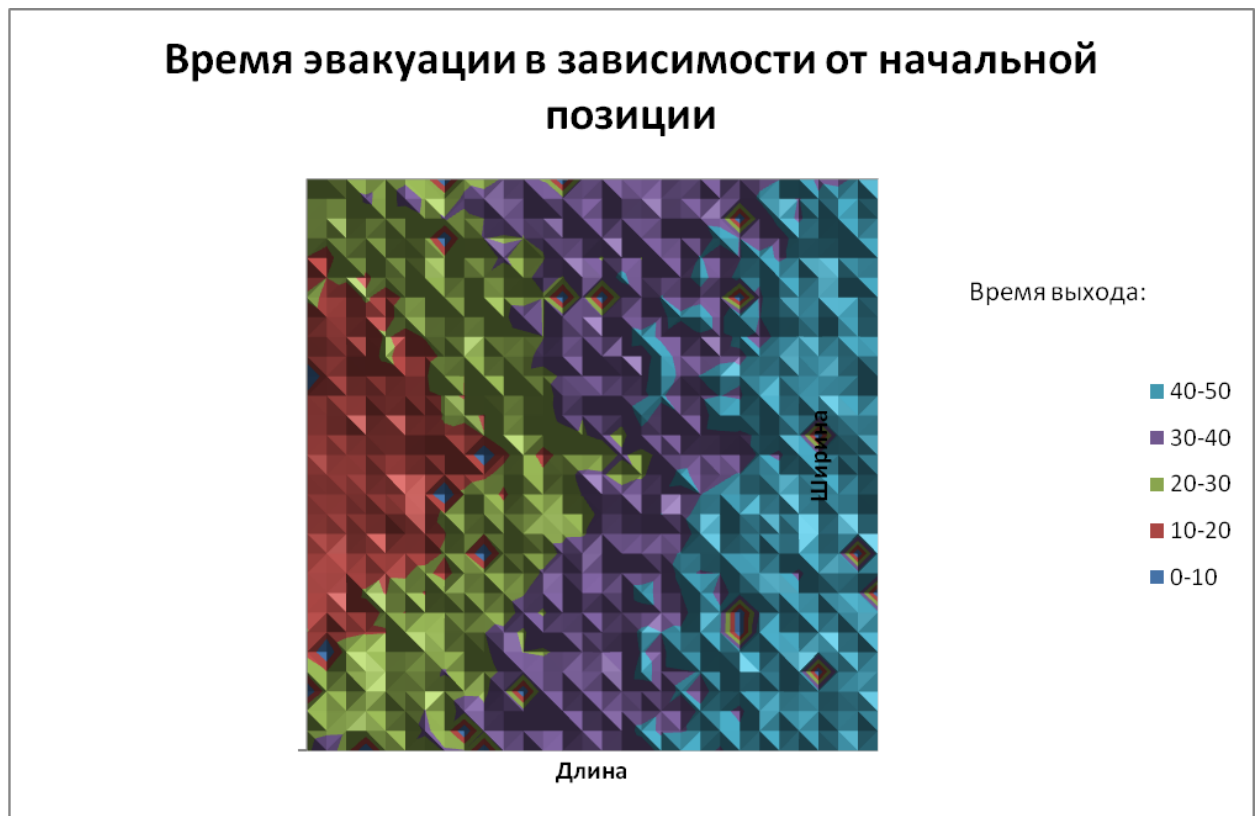


Рис. 40 Зависимость времени эвакуации от числа рефлексивных агентов 1 ранга, выбравших дальний выход

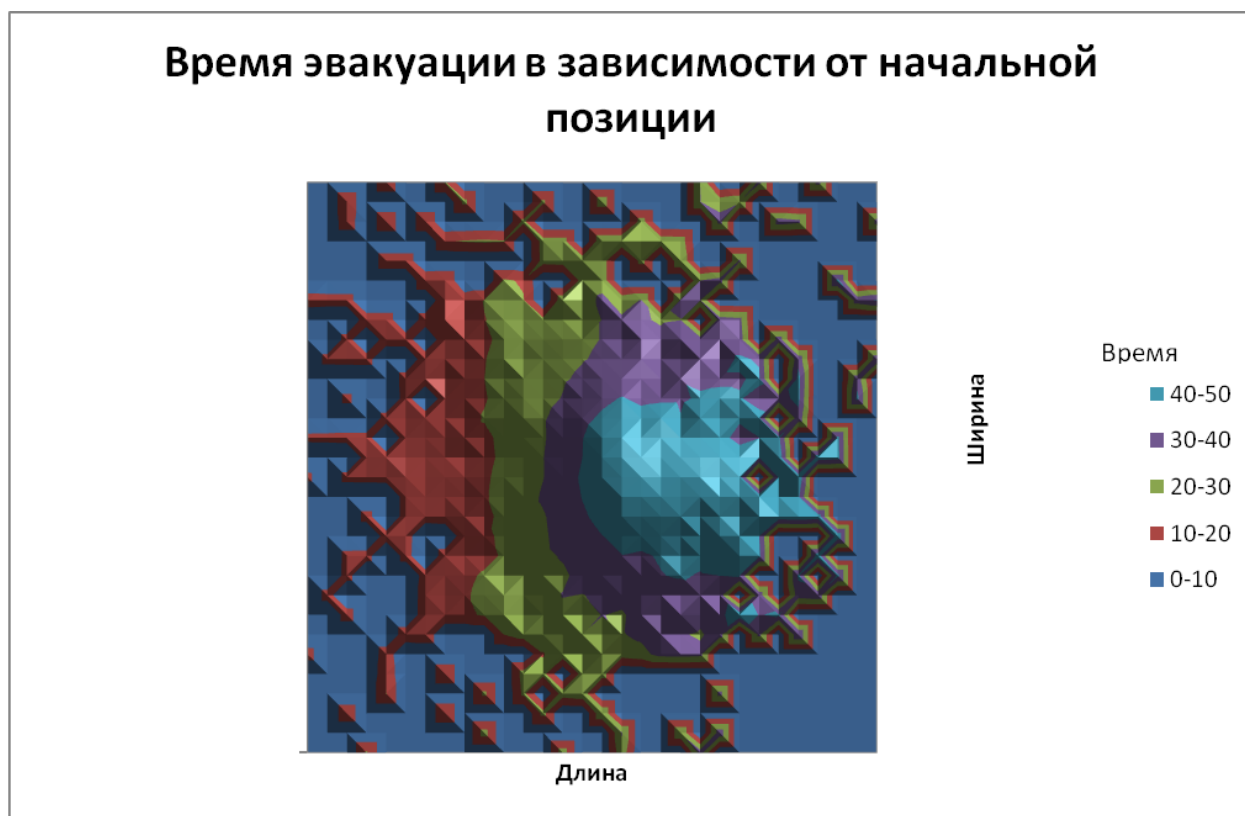
Очевидно теперь, что оптимальное количество агентов 1-го ранга, идущих в дальний выход близко с (52), значение же [51,54] определяет количество таких агентов в системе, необходимое для того, чтобы  $n_1^*$  агентов из них пошли в дальний выход. Таким образом, данный эксперимент показал соответствие теоретических и практических результатов.

Следующие графики отражают время эвакуации агента в ближайший выход в зависимости от его начального положения в помещении № 2, направление к выходу на рисунке слева. Случаи соответствуют: равномерному распределению – Рис. 41 и нормальному распределению с математическим ожиданием в центре помещения №2 – Рис. 42.



*Рис. 41* Время в зависимости от начальной позиции для равномерного распределения

Для случая равномерно распределённых агентов заметно, что время выхода близко к линейной зависимости от расстояния до выхода, об этом говорит практически равная толщина выделенных областей, а также их сферические контуры. Заметно также распрямление контура областей от сферического – до прямой вертикальной линии. Подобные картины известны для существующих моделей эвакуации, основанных на применении потоковых моделей [98, 99, 107], что говорит о качественном соответствии предложенной модели поведения агентов известным эмпирическим данным движения людских потоков. То есть, на основе моделирования отдельных «частиц» получились результаты, соответствующие макропараметрам людских потоков.



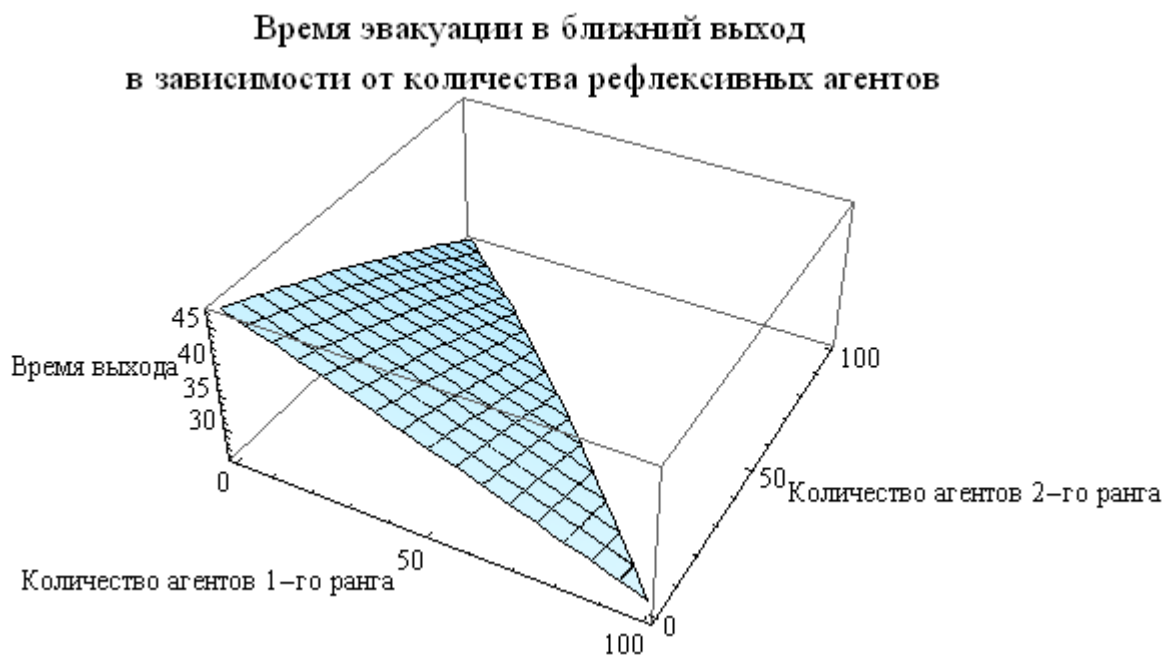
*Рис. 42 Время в зависимости от начальной позиции для нормального распределения*

Случай нормально распределённых агентов соответствует эвакуации толпы агентов из помещения. Качественно рисунок полностью другой, чем для равномерно распределённых агентов. Агенты, вышедшие позднее всех, находятся в той части толпы, которая начинается от её середины и, расширяясь, заканчивается у противоположной стены. Этот график показывает характер «рассасывания» толпы: раньше выбегают не только первые, но и крайние, расположенные по бокам от центра скопления. Подобные качественные картины для эвакуации толпы автор ещё не видел в литературе. Отличие графиков говорит о том, что параметры эвакуации зависят не только от количества человек в помещении, но и от распределения людей в помещении, но время эвакуации возможно слабо от этого зависит, поскольку максимальное время выхода для нормального и равномерного распределения близки. В принципе, можно ставить задачу поиска оптимального размещения людей в помещении для минимизации времени их эвакуации, но графики на Рис. 37 и Рис. 38 говорят о том, что хотя от вида распределения время выхода отдельных агентов зависит сильно, время эвакуации для двух различных распределений почти одинаково.

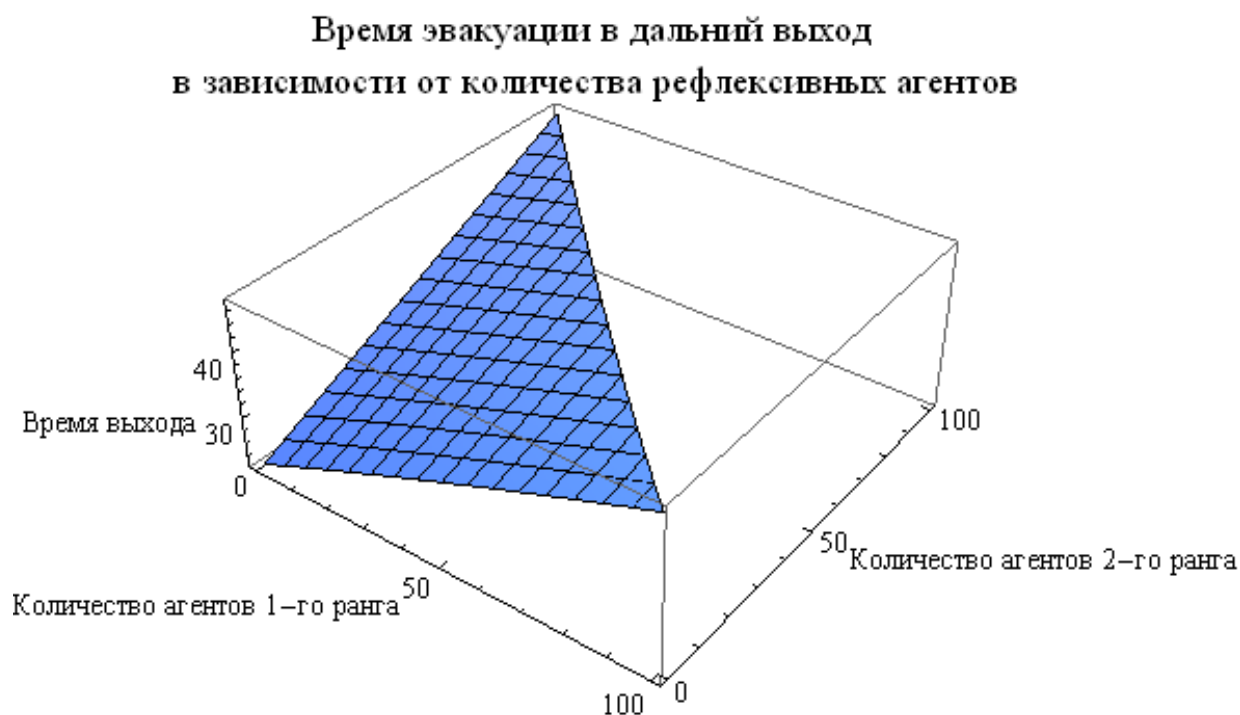
#### Второй эксперимент

Приведём полученные сглаженные графики зависимости времени эвакуации от количества агентов: для ближнего выхода – Рис. 43, для

дальнего выхода – Рис. 44.



*Рис. 43* Время эвакуации в ближний выход в зависимости от количества рефлексивных агентов



*Рис. 44* Время эвакуации в дальний выход в зависимости от количества рефлексивных агентов

Для ближнего выхода формула данной поверхности:

$$T_{\text{ближ}}(x, y) = 47.2397 - 0.185864 * x - 0.000119978 * x^2 - 0.179633 * y + 0.00084612 * x * y - 0.000360454 * y^2,$$

для дальнего:

$$T_{\text{даль}}(x, y) = 22.6752 + 0.247724 * x - 0.000178149 * x^2 + 0.247732 * y -$$

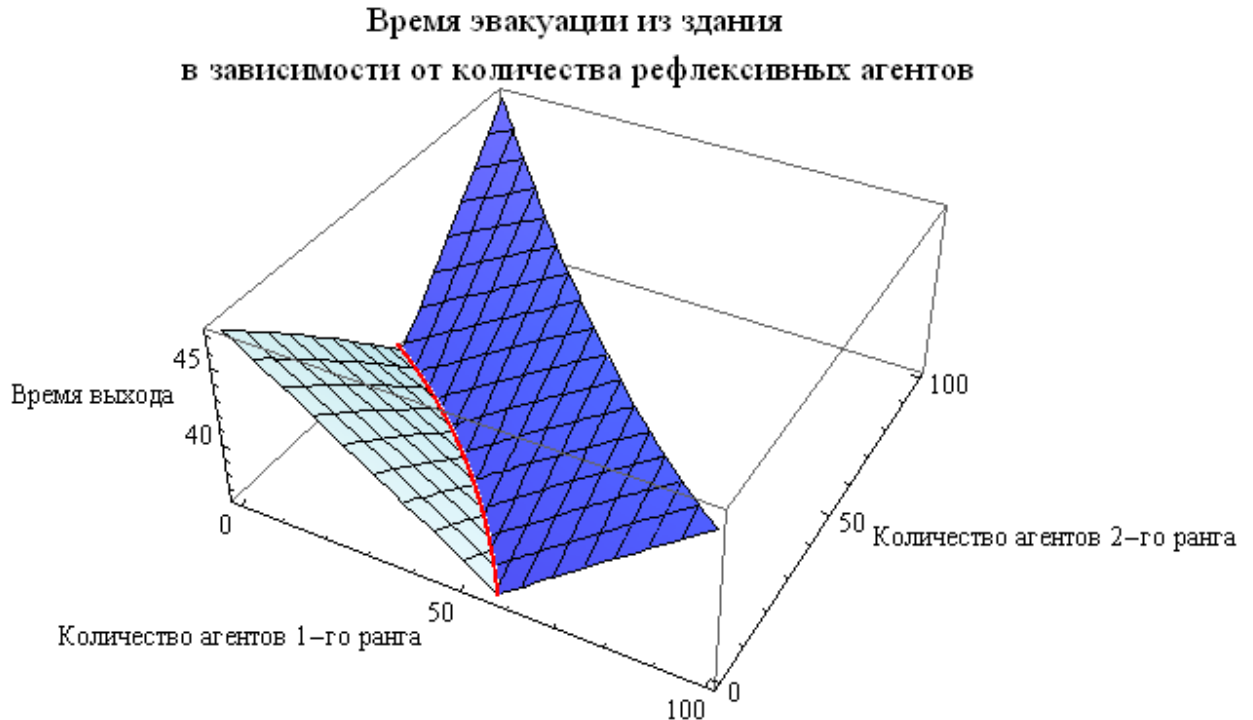
$$0.00141244*x*y - 0.000012856*y^2,$$

где  $x$  – количество агентов 1-го ранга,  $y$  – 2-го ранга.

Тогда время эвакуации из здания равно:

$$T_{\text{эвак}}(x, y) = \max(T_{\text{ближ}}(x, y), T_{\text{даль}}(x, y))$$

и имеет поверхность, изображённую на Рис. 45.



*Рис. 45* Время эвакуации из здания в зависимости от количества рефлексивных агентов 1-го и 2-го ранга

Минимум времени эвакуации достигается на кривой:

$$o_{\min} = \{(x, y) | T_{\text{ближ}}(x, y) = T_{\text{даль}}(x, y), x \in [0, 100], y \in [0, 100-x]\} = \{(3726.87 - 19.4132y - 19.5665\sqrt{35176.4 - 358.769y + y^2}, y) | y \in [0, 55.0171]\},$$

с учётом неточности данных, поскольку точно минимум  $T_{\text{эвак}}(x, y)$  достигается в точке  $(57.09, 0)$ , но изменение времени эвакуации на кривой  $o_{\min}$  соответствует погрешности данных:

$$T_{\text{эвак}}(o_{\min}) = [36.2374, 36.3687]$$

На Рис. 45 красным цветом обозначена кривая  $T_{\text{эвак}}(o_{\min})$ .

Таким образом, получаем, что введение рефлексивных агентов 2-го ранга для уменьшения времени эвакуации нецелесообразно – для достижения минимума достаточно использовать 57 агентов 1-го ранга рефлексии.

В данном примере найдено оптимальное количество агентов первого ранга, отклонение от которого увеличивает время эвакуации, показана нецелесообразность использования агентов второго ранга.

### **3.5. Имитационное моделирование диффузной бомбы**

Настоящий раздел посвящён т.н. задаче о диффузной бомбе (задаче о

групповом проникновении через систему обороны) [23]. Проводится имитационный сравнительный анализ пяти вариантов, различающихся «интеллектуальностью» поведения (адаптивность, способность к рефлексии, прогнозированию и т.д.) подвижных объектов (ПО). Показано, что наделение ПО возможностью учета параметров системы обороны и прогнозирования поведения других ПО повышает эффективность группового проникновения через систему обороны.

### 3.5.1. Планирование траекторий в условиях противодействия

Рассмотрим следующую задачу. Заданы начальные положения  $(x_j(0), y_j(0))$ ,  $j = \overline{1, K_0}$ , на плоскости  $K_0$  подвижных объектов. Их цель – оказаться в точке с координатами  $(x^*, y^*)$ . Положение  $j$ -го ПО в момент времени  $t \geq 0$  обозначим через  $(x_j(t), y_j(t))$ , его скорость – через  $v_j(t) = \sqrt{(\dot{x}_j)^2 + (\dot{y}_j)^2}$ , время первого попадания в точку  $(x^*, y^*)$  – через  $T_j$ .

Имеются  $N$  неподвижных сенсоров с координатами  $(a_i, b_i)$ ,  $i = \overline{1, N}$ , имеющих возможность суммировать приходящие на них в один и тот же момент времени сигналы. Расстояние от  $j$ -го ПО до  $i$ -го сенсора обозначим через  $\rho_{ij}(t) = \sqrt{(x_j(t) - a_i)^2 + (y_j(t) - b_i)^2}$ .

В общем случае *риск обнаружения*  $j$ -го ПО системой сенсоров описывается следующим функционалом:

$$(53) R_j = \int_0^{T_j} \sum_{i=1}^N \frac{(v_j(t))^m}{(\rho_{ij}(t))^k} dt,$$

где "сигнал" на сенсоре (слагаемое в выражении (53)) зависит от скорости ПО и расстояния от последнего до сенсора. Из вида функционала (53) следует, что риск обнаружения ПО зависит от значений "сигналов" на различных сенсорах. Величина показателя степени  $k$  является характеристикой физического поля, в котором осуществляется обнаружение [1], а величина показателя степени  $m$  характеризует зависимость уровня интенсивности излучаемого сигнала от скорости движения объекта (например, сигналов первичного гидроакустического поля [2]). Значения  $k = 1, 2, 3, 4$  соответствуют:

- $k = 1$  – первичному гидроакустическому полю в мелком море;
- $k = 2$  – тепловому полю, первичному электромагнитному полю и первичному гидроакустическому полю в глубоком море;
- $k = 3$  – магнитному полю;
- $k = 4$  – вторичному (при активном режиме обнаружения) электромагнитному и гидроакустическому полю.

Для случая одного сенсора и движения одного ПО на постоянной скорости решение задачи выбора оптимальной траектории получено в [138].

В [7, 8] доказано, что в случае одного ПО, движущегося с переменной скоростью, и одного сенсора оптимальная траектория и режим изменения

скорости таковы, что текущее значение суммарного "сигнала" постоянно во времени. В [1] это свойство оптимальной траектории и режима изменения скорости обобщено на случай одного ПО и нескольких сенсоров (см. также [2]). Анализ свойств функционала обнаружения, основывающийся на переходе в комплексную плоскость, осуществлен в [9]; в [10] исследованы свойства функционала обнаружения. "Сводка" результатов приведена в следующей таблице:

		НАБЛЮДАТЕЛИ		
		1	2	$\geq 3$
ПОДВИЖНЫЙ ОБЪЕКТ	Один	[138] – скорость ПО постоянная; [7, 8] – скорость ПО переменная	[9] – скорость ПО переменная	[138] – скорость ПО постоянная, [1, 2] – скорость ПО переменная
	Несколько	–	–	[70] – «взаимодействие» ПО не учитывается

Свойства оптимальных траекторий позволяют сформулировать эвристику, использующую идею о том, что плоскость может быть разбита на  $N$  областей (каждая соответствует своему сенсору), а оптимальные траектории должны проходить по границам этих областей. Другим примером может служить используемое в [70] разбиение на многогранники Вороного (причем при определении траекторий движения группы ПО условием является движение различных ПО по различным граням этого многогранника). Альтернативой является сеточное разбиение плоскости, что позволяет свести задачу к дискретной задаче поиска «кратчайшего» пути (где длина определяется вероятностью обнаружения ПО и/или длиной траектории, и/или энергетическими или временными критериями – см. [70, 138]).

Отметим, что ни в одной из перечисленных работ не рассматривается "взаимодействие" ПО, описание результатов имитационного моделирования которого представлено далее в следующих параграфах.

### 3.5.2. «Некооперативная» модель

Пусть все ПО движутся с постоянной по абсолютной величине скоростью  $v_0$ . Зная расположение сенсоров и их условные неотрицательные чувствительности  $\{c_i\}$ ,  $i = \overline{1, N}$ , мы можем по аналогии с выражением (1) для каждой точки  $(x, y)$  плоскости определить риск (вероятность обнаружения) обнаружения отдельного ПО, находящегося в этой точке:

$$(54) \quad r(x, y) = \min \left\{ \sum_{i=1}^N \frac{c_i}{\left( \sqrt{(x-a_i)^2 + (y-b_i)^2} \right)^k}; 1 \right\}.$$

Пусть время дискретно. Шаг времени обозначим через  $\tau$ , через  $p$  обозначим вероятность уничтожения обнаруженного ПО (для простоты будем считать, что эта вероятность не зависит от координат точки обнаружения, времени и скорости ПО – учет в будущих исследованиях этих зависимостей представляется перспективным), через  $\mathbf{e}(x, y) = (x^* - x, y^* - y) / \sqrt{(x - x^*)^2 + (y - y^*)^2}$  – единичный вектор направления на цель в точке  $(x, y)$ , через  $\rho((x, y); (q, w))$  – евклидово расстояние между точками  $(x, y)$  и  $(q, w)$ , через  $s_{\Delta}(x, y)$  – круг радиуса  $\Delta \geq 0$  с центром в точке  $(x, y)$ .

Рассмотрим несколько стратегий поведения ПО.

Вариант I. Первый (самый простой) вариант, когда каждый ПО движется по прямой, соединяющей его начальное положение с целью. Соответствующий ПО условно назовем *неинтеллектуальным*.

В рамках варианта I каждый ПО в каждый момент времени должен знать только свое текущее положение и положение цели.

Более «интеллектуальные» ПО должны учитывать текущие и/или будущие вероятности их обнаружения. Для описания их поведения определим множество таких точек, что: 1) в них ПО может оказаться, начав двигаться из точки  $(x, y)$  со скоростью  $v_0$ , через время  $\tau$ ; 2) вероятность обнаружения ПО не превышает *пороговой величины*  $\delta$ :

$$(55) \quad S_{v_0\tau}^{\delta}(x, y) = \{(q, w) \mid \rho((x, y); (q, w)) = v_0\tau; r(q, w) \leq \delta\}.$$

Геометрическое место точек плоскости, в которых вероятность обнаружения равна пороговой вероятности  $\delta$ , назовем *пороговой линией*.

Обозначим  $\text{Proj}_{S_{v_0\tau}^{\delta}}(x^*, y^*)$  – проекцию точки цели на множество  $S_{v_0\tau}^{\delta}$  (если такая проекция не единственна, выберем равновероятно любую из них).

Вариант II. Введем следующие правила планирования ПО своих траекторий (*алгоритм некооперативного поведения*).

Шаг 1. В каждый период времени ПО, находящийся в точке  $(x, y)$ , с вероятностью  $p$   $r(x, y)$  уничтожается системой обороны, а с вероятностью  $1 - p$   $r(x, y)$  продолжает движение.

Шаг 2. Продолжив движение, ПО к началу следующего периода времени окажется в точке  $(u, v)$ , где

$$(56) \quad (u, v) \in \begin{cases} \text{Proj}_{S_{v_0\tau}^{\delta}(x, y)}(x^*, y^*), & \text{если } (x, y) \notin \text{Proj}_{S_{v_0\tau}^{\delta}(u, v)}(x^*, y^*) \\ (x, y) + \min\{v_0\tau; \rho((x, y); (x^*, y^*))\} \mathbf{e}(x, y), & \text{иначе} \end{cases}.$$

Первый случай в выражении (56) соответствует неудалению от цели при условии, что риск обнаружения не превысит пороговый. Второй случай – «прорыв» по прямой к цели (в случае, если не удастся обеспечить выполнение условия непревышения вероятностью обнаружения порогового

значения, то есть когда следование этому условию приводит к удалению от цели).

В рамках алгоритма некооперативного поведения, ПО, находящийся в некоторой точке, должен иметь оценки риска обнаружения только для  $v_0 \tau$ -окрестности этой точки. То есть поведение ПО локально-оптимально, и требует только локальной информации. Отметим, что при активном режиме обнаружения ( $k = 4$ ) эта локальная информация может вычисляться посредством экстраполяции текущих измерений сигналов сенсоров. При пассивном режиме достаточно потребовать знания всеми ПО координат и чувствительностей сенсоров (что дает им возможность рассчитать риск (54) для произвольной точки плоскости).

Некооперативное поведение группы ПО будет описываться следующим образом: для каждого момента времени для каждого ПО выполнять последовательно Шаг 1 и Шаг 2 до тех пор, пока либо не будут уничтожены все ПО, либо все уцелевшие ПО не достигнут цели.

### 3.5.3. «Кооперативная» модель

Взаимодействие ПО будем учитывать следующим образом: предположим, что вероятность обнаружения данного ПО зависит не только от текущих расстояний от него до сенсоров, но и от того, насколько близко к нему расположены другие ПО (пример – рост эффективной поверхности рассеяния). То есть, условно можно считать, что ПО являются "сенсорами" друг для друга, и по мере их взаимного сближения растет вероятность обнаружения.

Обозначим через

$$(57) R_j(x_j, y_j) = \min \left\{ r(x_j, y_j) + \sum_{l \neq j} \frac{\alpha}{1 + \left( \sqrt{(x_j - x_l)^2 + (y_j - y_l)^2} \right)^k}; 1 \right\}$$

риск обнаружения  $j$ -го ПО, находящегося в точке  $(x_j, y_j)$ , с учетом его взаимодействия с другими ПО, где  $\alpha$  – неотрицательная константа.

Введём соответствующие варианты поведения.

Вариант III. ПО прорываются к цели по прямой, не учитывая и не прогнозируя вероятности их обнаружения. Данный вариант соответствует варианту I с точностью до замены риска (54) на риск (57). Информированность ПО при этом такая же, что и в варианте I.

Вариант IV. Алгоритм "кооперативного" поведения будет описываться Шагами 1' и 2', которые с точностью до замены риска (54) на риск (57) совпадают соответственно с Шагами 1 и 2, причем в выражении (57) суммирование ведется по тем ПО, которые к текущему моменту не были уничтожены. Вариант IV соответствует варианту II с точностью до замены риска (54) на риск (57).

В данном случае для планирования своей траектории каждый ПО, помимо информации, необходимой во втором варианте, должен знать текущие координаты всех ПО. Возможно обобщение "кооперативной" модели на случай, когда каждый ПО имеет свой фиксированный "радиус

обзора" и при планировании своей траектории имеет информацию и учитывает (в выражении типа (57)) только те другие ПО, которые находятся от него на расстоянии, не превышающем этот радиус.

### 3.5.4. Рефлексивная модель

Будем считать, что в группе присутствуют ПО двух типов. Первый тип – назовем их *нерефлексирующими* – действует в соответствии с алгоритмом "кооперативного" поведения (вариант IV). Второй тип – назовем их *рефлексирующими* – действует более сложным образом: каждый из них, считая всех остальных *нерефлексирующими*, прогнозирует их поведение. Другими словами, рефлексирующий ПО рассчитывает, где окажутся в следующий момент времени другие ПО (действующие в соответствии с вариантом IV) и выбирает направление своего движения с учетом прогнозируемых положений других ПО.

Определим Шаги 1'' и 2'', которые с точностью до замены риска (57) на прогнозируемый риск совпадают соответственно с Шагами 1' и 2'.

Вариант V. Алгоритм рефлексивного поведения группы ПО: для каждого момента времени для каждого *нерефлексирующего* ПО выполнять последовательно Шаги 1' и 2', а для каждого *рефлексирующего* ПО выполнять последовательно Шаги 1'' и 2'', пока все уцелевшие ПО не достигнут цели.

Отметим, что в рамках алгоритма рефлексивного поведения (вариант V) информированность каждого ПО должна быть такой же, что и в случае кооперативного поведения (вариант IV).

Итак, имеем пять вариантов поведения ПО:

	Номер варианта	Учет вероятности обнаружения	Учет положений других ПО	Прогноз поведения других ПО	Информированность
Некооперативная модель	I	НЕТ	НЕТ	НЕТ	В каждый момент времени каждый ПО должен знать только свое текущее положение и положение цели.
	II	ДА	НЕТ	НЕТ	Дополнительно к варианту I в каждый момент времени каждый ПО должен знать оценки риска обнаружения для $v_0 \tau$ -окрестности своего текущего положения
ративная	III	НЕТ	ДА	НЕТ	Как в варианте I.

	IV	ДА	ДА	НЕТ	Дополнительно к варианту II в каждый момент времени каждый ПО должен знать текущие координаты всех остальных ПО.
Рефлексивная модель	V	ДА	ДА	ДА	Как в варианте IV.
Адаптивная модель	VI ПО первого типа	ДА	НЕТ	НЕТ	Как в варианте II
	ПО второго типа	НЕТ	ДА	ДА	Дополнительно к варианту I в каждый момент времени каждый ПО второго типа должен знать текущие координаты всех ПО первого типа.

Возникает вопрос, как соотносятся между собой эффективности использования ПО тех или иных стратегий. Нахождение ответа на этот вопрос в общем аналитическом виде представляется вряд ли реализуемым, поэтому был выбран путь создания имитационной модели.

### 3.5.5. Результаты имитационного моделирования

Рассмотрим следующую имитационную модель, реализованную в среде AnyLogic. Выберем  $K_0 = 100$ ,  $N = 5$ ,  $c_i = 0,25$ ,  $p = 0.5$ ,  $\delta = 0.03$ . Начальные положения ПО, цель, сенсоры и линии уровня суммарного «сигнала» изображены на Рис. 46.

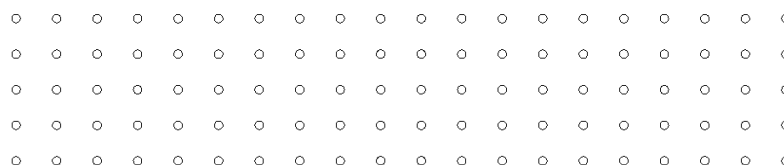
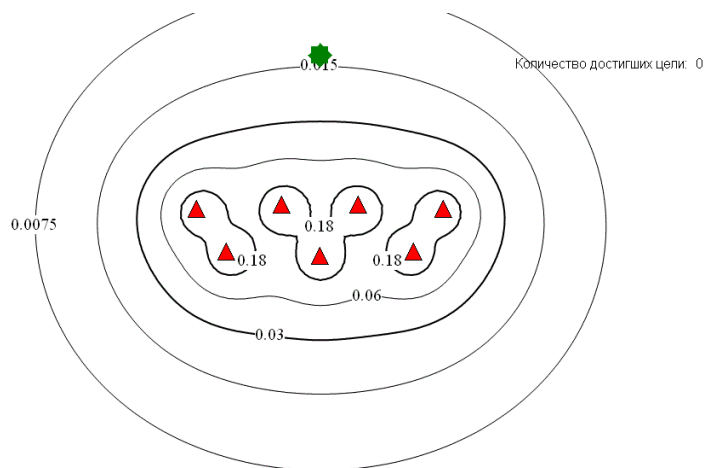


Рис. 46. Начальное расположение ПО, цель (звездочка), сенсоры (треугольники) и линии уровня суммарного "сигнала" (2)

Пример результатов группового проникновения через систему обороны для варианта II приведен на Рис. 47, где черными кружками обозначены уничтоженные ПО.

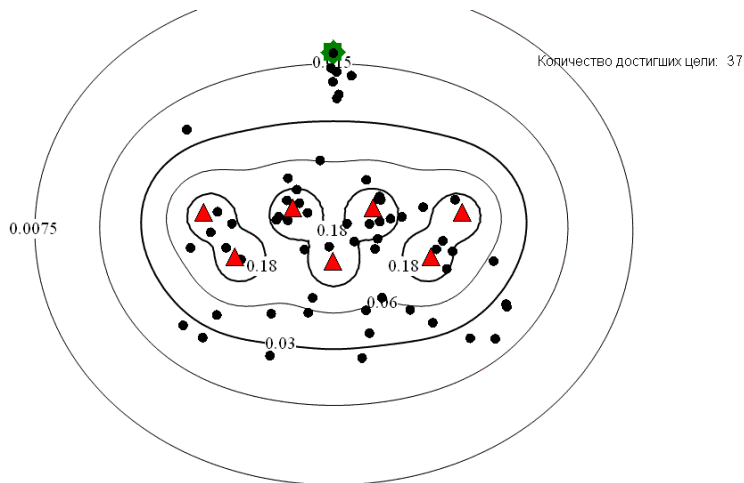


Рис. 47. Пример результатов группового проникновения через систему обороны для варианта II

На Рис. 48 для вариантов I и II приведены зависимости эффективности  $K$  (здесь и ниже каждая точка на графике эффективности является результатом усреднения по 200 испытаниям) действий группы ПО от вероятности  $p$  уничтожения обнаруженного ПО. Естественно, с ростом вероятности уничтожения эффективность уменьшается.

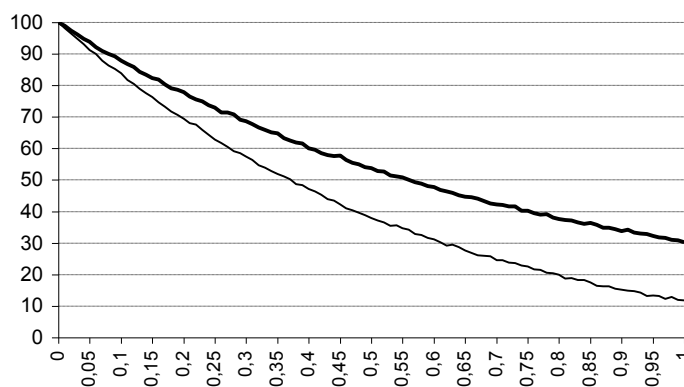


Рис. 48. Зависимость эффективности  $K$  действий группы ПО от вероятности  $p$  уничтожения обнаруженного ПО для вариантов I (тонкая линия) и II (жирная линия)

Видно, что переход от варианта I к варианту II, т.е. рост интеллектуализации ПО за счет их анализа вероятности уничтожения в  $\nu_0$   $\tau$ -окрестности текущего положения, существенно повышает эффективность преодоления системы обороны (например, при  $p = 0.5$  эффективность увеличивается с 38 до 53 – примерно на 40 %)

Отметим, что варианты I-II и III-V не сравнимы между собой, так как в последних учитывается взаимодействие ПО и вероятности их обнаружения выше. Поэтому приведем Рис. 49, содержащий для вариантов III-V зависимости эффективности  $K$  действий группы ПО от вероятности  $p$  (меняющейся в диапазоне от 0.4 до 0.6) уничтожения обнаруженного ПО при  $\alpha = 0.03$  (в варианте V считается, что рефлексирующими является половина ПО).

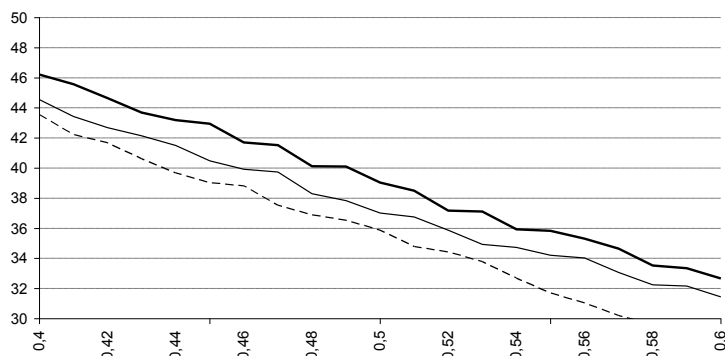


Рис. 49. Зависимость эффективности  $K$  действий группы ПО от вероятности  $p$  уничтожения обнаруженного ПО для вариантов III (пунктирная линия), IV (тонкая линия) и V (жирная линия)

Видно, что, опять же, рост интеллектуализации ПО повышает эффективность преодоления системы обороны (вариант V является самым эффективным, далее идет вариант IV, затем вариант III).

На Рис. 50 приведена зависимость эффективности  $K$  действий группы ПО от значений параметра  $\alpha$ , отражающего взаимовлияние ПО.



Рис. 50. Зависимость эффективности  $K$  действий группы ПО от значений параметра  $\alpha$

Обозначим через  $K^* \in \{0, 1, \dots, K_0\}$  число рефлексорирующих ПО. График зависимости  $K(K^*)$  при  $\alpha = 0.25$  приведен на Рис. 51.

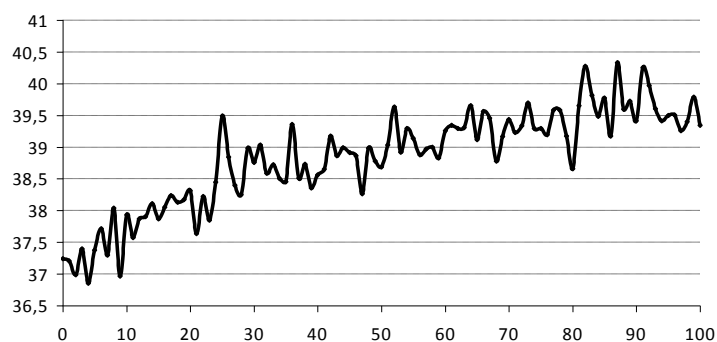


Рис. 51. Зависимость эффективности  $K$  действий группы ПО от числа рефлексорирующих ПО  $K^*$

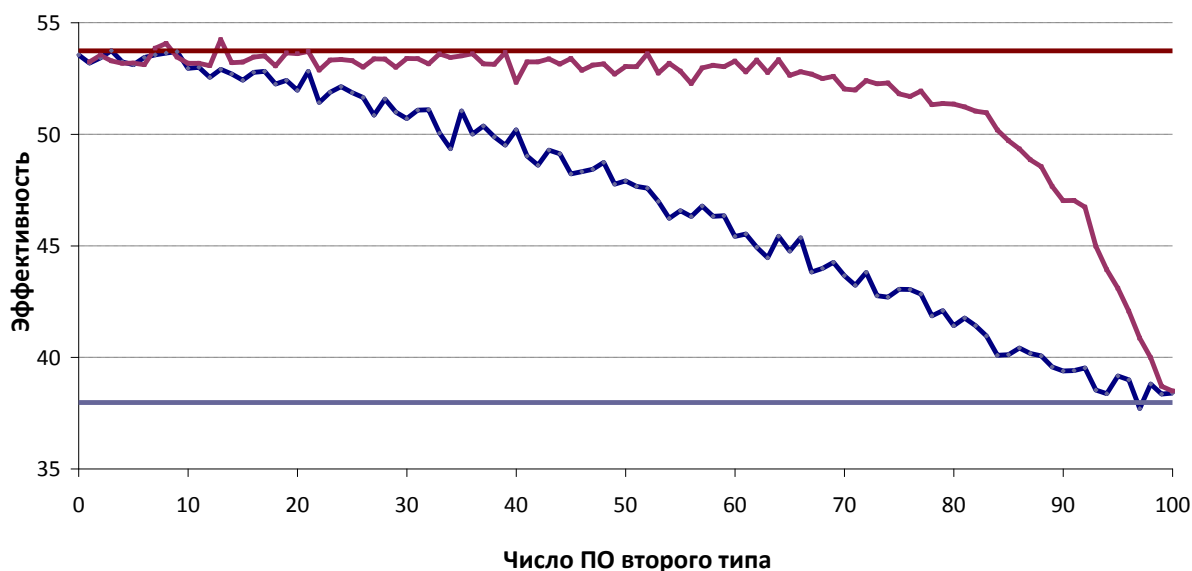
Видно, что с ростом доли рефлексорирующих ПО эффективность действий группы ПО увеличивается. Более того, «выживаемость» рефлексорирующих ПО выше – среднее число рефлексорирующих ПО, достигших цели, больше, чем нерефлексорирующих (причем в рассматриваемой имитационной модели, например, при 200 испытаниях и равном числе рефлексорирующих и нерефлексорирующих ПО эта оценка статистически значима).

В заключение настоящего раздела рассмотрим следующий вариант информированности и «интеллектуальности» ПО.

Вариант VI (адаптивное поведение). Пусть имеются ПО двух типов. Предположим, что все ПО в каждый момент времени знают свое текущее положение и положение цели. Дополнительно ПО первого типа в каждый момент времени знают оценки риска обнаружения для  $\nu_0 \tau$ -окрестности своего текущего положения, а ПО второго типа в каждый момент времени

знают (или могут измерить) текущие и предыдущие координаты всех остальных. ПО первого типа действуют в соответствии с вариантом II, а ПО второго типа на каждом шаге сначала на основании наблюдения за движением других ПО вычисляют оценку расположения пороговой линии. Затем они действуют в соответствии с вариантом II, подставляя в аналог выражения (56) свою текущую оценку пороговой линии. Другими словами, ПО второго типа ведут себя адаптивно (в смысле [41]). Условно можно назвать ПО первого типа «разведчиками» – они лучше информированы (и, наверное, дороже) и проводят разведку боем, добывая информацию о системе обороны (точнее – о пороговой линии) для других ПО (второго типа).

«Пределными» являются два случая – когда все ПО первого типа (тогда имеем вариант II) или когда все ПО второго типа (тогда имеем вариант I). Пусть вероятность уничтожения обнаруженного ПО равна 0,5. На *Рис. 52* представлена зависимость числа ПО, достигших цели, от числа ПО второго типа (горизонтальные линии соответствуют оценкам *Рис. 48* – эффективностям, равным 53,74 и 37,97) для двух случаев – когда все ПО движутся одновременно (нижняя кривая) и когда сначала оборону преодолевают ПО первого типа, а потом уже начинают двигаться ПО второго типа (верхняя кривая). Видно, что 20-30 % разведчиков обеспечивают в рассматриваемом примере почти такую же эффективность, что и использование только дорогостоящих ПО первого типа.



*Рис. 52. Зависимость эффективности K действий группы ПО от числа ПО второго типа*

Итак, в настоящем разделе для задачи о диффузной бомбе проведен имитационный сравнительный анализ шести вариантов, различающихся «интеллектуальностью» поведения ПО. Показано, что наделение ПО

возможностью учета параметров системы обороны и прогнозирования поведения других ПО повышает эффективность решения задачи о групповом проникновении через систему обороны. С другой стороны, понятно, что «платой за интеллектуальность» является рост массогабаритных характеристик, энергетических, вычислительных и других ресурсов, которыми должны обладать ПО. Поэтому при решении каждой конкретной задачи придется оптимизировать баланс между этими критериями и собственно эффективностью проникновения через систему обороны.

В данном примере введение любого количества рефлексивных ПО улучшает состояние системы – количество доходящих до цели ПО увеличивается.

## Литература

1. АБРАМЯНЦ Т.Г., МАСЛОВ Е.П., ЯХНО В.П. *Уклонение подвижного объекта от обнаружения группой наблюдателей* // Проблемы управления. – 2010. – № 5. – С. 73–79.
2. АБРАМЯНЦ Т.Г., МАСЛОВ Е.П., РУДЬКО И.М., ЯХНО В.П. *Уклонение подвижного объекта от обнаружения группой наблюдателей при малых отношениях сигнал/помеха* // Обработка информации и управление. – 2011. – № 2. – С. 2–7.
3. БАЛАШОВ В.Г., ЗАЛОЖНЕВ А.Ю., ИВАЩЕНКО А.А., НОВИКОВ Д.А. *Механизмы управления организационными проектами*. – М.: ИПУ РАН, 2003.
4. БЕЛЯЕВ С.В. *Эвакуация зданий массового назначения*. М., 1938.
5. БУРКОВ В.Н., ДЖАВАХАДЗЕ Г.С., ДИНОВА Н.И., ЩЕПКИН А.В. *Применение игрового имитационного моделирования для оценки эффективности экономических механизмов*. М.: ИПУ РАН, 2003. – 51 с.
6. ВАСИН А.А. *Некооперативные игры в природе и обществе*. – М.: МАКС пресс, 2005.
7. ГАЛЯЕВ А.А., МАСЛОВ Е.П., РУБИНОВИЧ Е.Я. *Об одной задаче управления движением объекта в конфликтной среде* // Известия РАН. Теория и системы управления. – 2009. – № 3. – С. 134–140.
8. ГАЛЯЕВ А.А., МАСЛОВ Е.П. *Оптимизация законов уклонения подвижного объекта от обнаружения* // Известия РАН. Теория и системы управления. – 2010. – № 4. – С. 43–53.
9. ГАЛЯЕВ А.А. *О задаче прорыва между двумя сенсорами при движении объекта в конфликтной среде* // АиТ. – 2010. – № 5. – С. 3–10.
10. ГАЛЯЕВ А.А. *О функционале обнаружения при движении объекта в конфликтной среде* // АиТ. – 2010. – № 4. – С. 100–105.
11. ГЕРМЕЙЕР Ю.Б. *Игры с противоположными интересами*. М.: Наука, 1976.
12. *Глоссарий по теории управления организационными системами* – <http://www.mtas.ru/Glossary.htm>
13. ГУБАНОВ Д.А., ЧХАРТИШВИЛИ А.Г. *О стратегической рефлексии в биматричных играх* // Управление большими системами. 2008. №21. С. 49–57.
14. ГУБКО М.В., НОВИКОВ Д.А. *Теория игр в управлении организационными системами*. – М.: Синтег, 2002.
15. ДЕНИСОВ В.В., КОЛОСОВ Г.Г., ПАНТЕЛЕЕВ М.Г., ПУЗАНКОВ Д.В. *Методы и средства построения интеллектуальных агентов реального времени* // Труды 7-й национальной конференции по искусственному интеллекту КИИ'2000. Т. 1–2. – М.: Изд-во физ.-мат. лит., 2000. – С. 805–813.
16. ДЮСУШЕ О.М. *Статичное равновесие Курно-Нэша и рефлексивные*

- игры олигополии: случай линейных функций спроса и издержек* / Экономический журнал Высшей школы экономики, Т.10, М.: ГУ ВШЭ, 2006. – С. 3–32.
17. ЕРЕШКО Ф.И. *Моделирование рефлексивных стратегий в управляемых системах*. – М.: ВЦ РАН, 2001.
18. 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.
19. ИСАЕВИЧ И.И. *Разработка основ многовариантного анализа планировочных решений станций и пересадочных узлов метрополитена на основе моделирования закономерностей движения людских потоков* : Дисс. канд. техн. наук. – М., 1990.
20. *Интеллектуальные роботы* / Под ред. И.А. КАЛЯЕВА. – М.: Машиностроение, 2007.
21. КАРАВАЕВ А.П. *Модели и методы управления составом активных систем*. – М.: ИПУ РАН, 2003.
22. КОЛОСОВ Г.Г., ПАНТЕЛЕЕВ М.Г. *Проектирование процессоров обработки производственных знаний на основе RETE-сети* / Искусственный интеллект. – 2003. – №3. – <http://www.iai.donetsk.ua>.
23. КОРЕПАНОВ В.О., НОВИКОВ Д.А. *Задача о диффузной бомбе* // Проблемы управления. – 2011. – № 5. – С. 66–73.
24. КОРЕПАНОВ В.О., НОВИКОВ Д.А. *Метод рефлексивных разбиений в моделях группового поведения и управления* // Проблемы управления. – 2011. – № 1. – С. 21–32.
25. КОРЕПАНОВ В.О. *Модель рефлексивного группового управления в организационных системах* / Труды VII молодежной школы-конференции «Управление большими системами». – Пермь: ПГТУ, 2010. – Том. 1. – С. 287–294.
26. КОРЕПАНОВ В.О. *Влияние рефлексивных агентов на транспортный поток* / Труды международной научно-практ. конф. «Теория активных систем». – М.: ИПУ РАН, 2009. – Т. II. – С. 122–125.
27. 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.
28. КОРЕПАНОВ В.О. *Имитационные модели тактического поведения агентов* // Управление большими системами. – 2009. – №26. – С. 145-157.
29. КОРЕПАНОВ В.О. *Управление рефлексивным поведением агентов в модели олигополии Курно* / Управление большими системами. Выпуск 31. – М.: ИПУ РАН, 2010. – С.225–249.
30. КОРЕПАНОВ В.О. *Модели движения агентов в замкнутом пространстве* // Труды IV Всероссийской школы-семинара молодых

- учёных «Проблемы управления и информационные технологии (ПУИТ '08)». – 2008. – с. 118
- 31.КОРЕПАНОВ В.О. *Реализация иерархической структуры агента* // Труды 51 Всероссийской молодёжной научной конференции «Современные проблемы фундаментальных и прикладных наук». – 2008. – Ч.1 – с. 161
  - 32.МАЛИШЕВСКИЙ А.В. *Качественные модели в теории сложных систем*. – М.: Наука, 1998.
  - 33.*Математическая Энциклопедия*. Т. 1 (А - Г). Ред. коллегия: И. М. Виноградов (глав ред) [и др.] – М., «Советская Энциклопедия». – 1977. – 1152 с.
  - 34.*Методика определения расчетных величин пожарного риска в зданиях, сооружениях и строениях различных классов функциональной пожарной опасности*. Утверждена приказом МЧС России № 382 от 30.06.2009 и зарегистрирована в Министерстве юстиции Российской Федерации (регистрационный № 14486 от 06 августа 2009 г).
  - 35.МИЛИНСКИЙ А.И. *Исследование процесса эвакуации зданий массового назначения*: Дисс. канд. техн. наук. – М., 1951.
  - 36.МИЛИНСКИЙ А.И. *К вопросу об исследовании процессов эвакуации*. Информационный сб. ЦНИИПО, М., 1950.
  - 37.НЕПЕЙВОДА Н.Н. *Прикладная логика*. Новосибирск: НГУ, 2000, 491 с.
  - 38.НОВИКОВ А.М. *Методология образования*. М.: Эгвес, 2002. – 320 с.
  - 39.НОВИКОВ Д.А. *Динамика поведения систем с большим числом целенаправленных элементов* // Автоматика и Телемеханика. 1996. № 4. – С. 187–189.
  - 40.НОВИКОВ Д.А. *Математические модели формирования и функционирования команд*. – М.: Изд-во физ.-мат. лит., 2008. – 184 с.
  - 41.НОВИКОВ Д.А. *Модели адаптации команд / Управление большими системами*. 2008. – № 20. – С. 57–76.
  - 42.НОВИКОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Прикладные модели информационного управления*. – М.: ИПУ РАН, 2004.
  - 43.НОВИКОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Рефлексивные игры*. – М.: Синтег, 2003.
  - 44.НОВИКОВ Д.А. *Теория управления организационными системами*. 2-е изд. – М.: Физматлит, 2007.
  - 45.ОПОЙЦЕВ В.И. *Равновесие и устойчивость в моделях коллективного поведения*. – М.: Наука, 1977.
  - 46.ПРЕДТЕЧЕНСКИЙ В.М., МИЛИНСКИЙ А.И. *Проектирование зданий с учетом организации движения людских потоков*. – М., Стройиздат. 1969.
  - 47.ПРЕДТЕЧЕНСКИЙ В. М., МИЛИНСКИЙ А. И. *Проектирование зданий с учетом организации движения людских потоков*. Изд. 2-е, перераб и доп. – М.: Стройиздат, 1979.
  - 48.ПРЕДТЕЧЕНСКИЙ В.М. *Людские потоки в зданиях*. Глава в учебнике

- «Архитектура гражданских и промышленных зданий, т. II, Основы проектирования, М. 1976.
49. ПРЕДТЕЧЕНСКИЙ В.М. *О расчете движения людских потоков в зданиях массового назначения // Архитектурно-строительное образование и научные основы проектирования.* – М.: Стройиздат, 1983.
50. РОЙТМАН М.Я. *Общие принципы нормирования процесса вынужденной эвакуации людей.* М. 1963.
51. РОЙТМАН М.Я. *Эвакуация людей из промышленных и гражданских зданий на случай пожара.* М. 1956.
52. РОМАНЬКО А.Д., ЧХАРТИШВИЛИ А.Г. *Моделирование информационных воздействий в рефлексивных играх: простые сообщения // Сборник трудов ВГАСУ.* 2006. С. 157–167.
53. САМОШИН Д. А. *Расчет времени эвакуации людей. Проблемы и перспективы // Пожаровзрывобезопасность.* – 2004. – Т. 13, № 1. – С. 33–46.
54. СОРОС Д. *Алхимия финансов.* – М.: ИНФРА-М, 1999.
55. СТЕПАНЦЕВ М.Е. *Модель движения группы людей на основе решетчатого газа с нелокальными взаимодействиями // Известия ВУЗов. Прикладная нелинейная динамика.* 1999. – Т. 7. – С. 44–46.
56. ТОФФОЛИ Т., МАРГОЛУС Н. *Машины клеточных автоматов.* М: Мир, 1991.
57. ХАРШАНЬИ Д., ЗЕЛЬТЕН Р. *Общая теория выбора равновесия в играх.* СПб.: Экономическая школа, 2001. – 405 с.
58. ХОЛЩЕВНИКОВ В.В. *Исследование людских потоков и методологии нормирования эвакуации людей из зданий при пожаре.* – М.: МИПБ МВД РФ, 1999.
59. ХОЛЩЕВНИКОВ В.В. *Людские потоки в зданиях, сооружениях и на территории их комплексов.: Дисс. доктора техн. наук – М., 1983.*
60. ХОЛЩЕВНИКОВ В.В., НИКОНОВ С.А., ШАМГУНОВ Р.Н. *Моделирование и анализ движения людских потоков в зданиях различного назначения.* М., МИСИ, 1986.
61. ЧАЛДИНИ Р. *Психология влияния.* СПб.: Питер, 2001.
62. ЧЕБОТАРЕВ П.Ю., АГАЕВ Р.П. *Согласование характеристик в многоагентных системах и спектры лапласовских матриц оргграфов // Автоматика и телемеханика.* 2009. – № 3. – С. 136–151.
63. ЧЕПРУНОВА О.Ю., ЩЕПКИН А.В. *Разработка экспериментов с моделями организационных систем // Автоматика и Телемеханика.* 1988. № 8.
64. ЧХАРТИШВИЛИ А.Г. *Рефлексивные игры: трансформация структур информированности // Проблемы управления.* 2008. – № 5. – С. 43–48.
65. ЧХАРТИШВИЛИ А.Г. *Теоретико-игровые модели информационного управления.* – М.: ПМСОФТ, 2005.
66. ШВЕЦОВ А.Н. *Агентно-ориентированные системы: от формальных*

- моделей к промышленным приложениям // Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы», 2008. – 101 с. – [http://window.edu.ru/window/library?p\\_rid=56179](http://window.edu.ru/window/library?p_rid=56179).*
67. AEA Technology, *A Technical Summary of the AEA EGRESS Code*, Technical Report, AEAT/NOIL/27812001/002(R), Issue 1, 2002.
  68. AMBROSZKIEWICS S., BYLKA S., KOMAR J. *Knowledge in a Dynamic Model of Multi-Agent System*. [Electronic Data] – Mode of access : <http://www.ipipan.waw.pl/mas/>
  69. BATTY M. *Agent-based pedestrian modelling*. Environment and Planning B: Planning and Design 28 321-326. – 2001.
  70. BEARD R., McLAINY T., GOODRICH M. *Coordinated Target Assignment and Intercept for Unmanned Air Vehicles / Proceedings of IEEE2002 International Conference on Robotics and Automation*. Washington DC, 2002. – P. 2581–2586.
  71. de BOER F. S., HINDRIKS K., van der HOEK W., MEYER J.-J. *A verification framework for agent programming with declarative goals* in Journal of Applied Logic, 5:2, pp. 277–302, 2007.
  72. BOSCH-DOMENECH A., MONTALVO J.G., NAGEL R., SATORRA A. *”One, Two, (Three), Infinity, ... Newspaper and Lab Beauty-Contest Experiments”*. American Economic Review, XCII (2002), P. 1687–1701.
  73. BRATMAN M.E. *Plans and resource-bounded practical reasoning*. Computational Intelligence / M. E. Bratman, D. Y. Israel, M. E. Pollack. – 1988. – №4. – P. 349–355.
  74. BROOKS R.A. *Intelligence without representation / Artificial Intelligence*. – 1991. – № 47. – P. 139–159.
  75. CAMERER C.F., HO T.-H., CHONG J.-K. *A Cognitive Hierarchy Theory of Games*. Quarterly Journal of Economics, 119(3), 861–898. –2004.
  76. CETNAROWICZ K., GRUER P., HILAIRE V., KOUKAM A. *A Formal Specification of M-Agent Architecture*. Ed. by B. Keplics, E. Nawarecki // Proc. of the Second Int. Workshop of Central and Eastern Europe on Multi-Agent Systems. - Poland, 2001. – P. 4150.
  77. CHAPMAN D. *Planning for conjunctive goals*. Artificial Intelligence , 32(2):333-378, 1987.
  78. CHONG J.-K., CAMERER C.F., HO T.-H. *Cognitive hierarchy: A limited thinking theory in games*. In Experimental Business Research, Volume III: Marketing, Accounting and Cognitive Perspectives, chapter 9. Kluwer Academic Press, 2005.
  79. CONLISK J. *Costly optimizers versus cheap imitators*, Journal of Economic Behavior and Organization 1, 275-293, 1980.
  80. DIJKSTRA J., JESURUN J., TIMMERMANS H. *A multi-agent cellular automata model of pedestrian movement*. In Pedestrian and evacuation dynamics, edited by M. Schreckenberg and S. Deo Sarma, 173-80. Berlin: Springer-Verlag. – 2002.

- 81.DIMAKIS N., FILIPPOUPOLITIS A., GELENBE E. *Distributed building evacuation simulator for smart emergency management*. Accepted for publication in *The Computer Journal* (2010).
- 82.D'INVERNO M., LUCK M. *A formal specification of AgentSpeak(L)* // *Journal of Logic and Computation*. – 2000. – V. 13. – P. 157–176.
- 83.EMERSON E.A, SHRINIVANAS J. *Branching time logic* : REX School-Workshop on Linear Time, Branching Time and Parial Order in Logics and Models for Concurrency (LNCS Vol. 354); Ed. by J. W. de Bakker, W.-P. de Roever, G. Rosenberg . – Springer-Verlag : Heidelberg, 1988. – P. 123–172.
- 84.*Engineering News Record*, vol. 111, 1933.
- 85.FERBER J. *Multi-agent systems: An introduction to distributed artificial intelligence*. – Addison-Wesley, 1999. – 528 p.
- 86.FERGUSON I.A. *Touring Machines: An Autonomous agents with attitudes*. *IEEE Computer*, 25(5): 51–55, 1992.
- 87.GALE D., ROSENTHAL R. W. *Experimentation, imitation, and stochastic stability*: Abendum, *Journal of Economic Theory* 97, 164–174, 2001.
- 88.GENESERETH M.R. *Logical Foundations of Artificial Intelligence* / M. R. Genesereth, N. Nilsson. – Los Altos : Morgan Kaufmann, 1987. – 405 p.
- 89.GEORGEFF M.P, RAO A.S. *A profile of the Australian AI Institute* // *IEEE Expert*, December 1996. – 11 (6). – P. 89–92.
- 90.GEORGEFF M.P., LANSKY A.L. *Reactive reasoning and planning* // In *Proceedings of the Sixth National Conference on Artificial Intelligence (AAAI – 87)*, Seattle. – WA. – 1987. – P. 677–682.
- 91.GREENWALD B., STIGLITZ J.E. *Externalities in economies with imperfect information and incomplete markets*. *Quarterly Journal of Economics* 101(2): 229–264. 1986.
- 92.Halcrow Group Limited, Pedroute, Sept. 2003. (Available at: [http://www.halcrow.com/pdf/urban\\_reg/pedrt\\_broch.pdf](http://www.halcrow.com/pdf/urban_reg/pedrt_broch.pdf))
- 93.HANKS S., FIRBY R.J. *Issues and architectures for planning and execution*. In *DARPA Workshop on Innovative Approaches to Planning, Scheduling and Control*, p. 59-70, San Mateo, CA, 1990. Morgan Kaufmann.
- 94.HELHING D. *Improved Fluid-Dynamic Model for Vehicular Traffic*, *Physical Review E*, 51,1995. – P. 3164–3169.
- 95.HELHING D., FARKAS I., VICSEK T., *Simulating dynamical features of escape panic*. *Nature* 407(2000); <http://ru.arxiv.org/abs/cond-mat/0009448>
- 96.JACKSON M.O. “*Mechanism Theory*”, in *Optimization and Operations Research*, edited by Ulrich Derigs, in the *Encyclopedia of Life Support Systems*, EOLSS Publishers: Oxford UK, 2003.
- 97.KAHNEMAN D. *Maps of Bounded Rationality: Psychology for Behavioral Economics* // *The American Economic Review*, Vol. 93, No. 5 (Dec., 2003), pp. 1449–1475
- 98.KAMMANN J., ANGERMANN M., LAMI B. *A new mobility model based on Maps*, in *VTC 2003*.

99. KHIDER M., KAISER S., ROBERTSON P., and ANGERMANN M. *A novel movement model for pedestrians suitable for personal navigation*, in Proceedings of the Institute of Navigation National Technical Meeting 2008, San Diego, California, USA, Jan. 2008.
100. KHOLSHEVNIKOV V.V., SHIELDS T.J., SAMOSHYN D.A., GALUSHKA M.G. *Modelling of pedestrian flows*. 4th International Seminar on Fire and Explosion Hazards, September 8-12 2003, University of Ulster, Londonderry, UK.
101. KLUGL F., RINDSFUSER G. *Large-scale agent-based pedestrian simulation*. In Muller, J. P., Petta, P., Klusch, M., and Georgeff, M., editors, Multi-Agent Technologies V, Proceedings of the MATES 2007, number 4687 in LNAI. Springer.
102. KONOLIGE K. *A first-order formalization of knowledge find action for multiagent planning system* : Machine Intelligence 10; Ed. by J. E. Hayes, D. Michie, Y. Pao. – Chichester: Ellis Horwood, 1982. – P.41–72.
103. KORHONEN T., HOSTIKKA S., HELIÖVAARA S., EHTAMO H., MATIKAINEN K. *Integration of an Agent Based Evacuation Simulation and the State-of-the-Art Fire Simulation*. Proceedings of the 7th Asia-Oceania Symposium on Fire Science & Technology, 20-22 September, 2007, Hong Kong.
104. LAKOBA T.I., KAUP D.J., FINKELSTEIN N.M. *Modifications of the Helbing-Molnár-Farkas-Vicsek Social Force Model for Pedestrian Evolution*, SIMULATION, Vol. 81, No. 5, P. 339–352, 2005.
105. LAMARCHE F., DONIKIAN S. *Crowd of virtual humans: A new approach for real time navigation in complex and structured environments*, Computer Graphics Forum 23 (3) (2004) –P. 509–518.
106. LI H., TANG W., SIMPSON D. *Behaviour Based Motion Simulation for Fire Evacuation Procedures*, Proceedings - Theory and Practice of Computer Graphics 2004, Bournemouth, United Kingdom, pp. 112–118, 2004.
107. LIGHTFOOT T.J., MILNE G.J. *Modelling emergent crowd behaviour*. Proc. of the 1<sup>st</sup> Australian Conf. on Artificial Life, Canberra. – 2003.
108. van LINDER B., van der HOEK W., and MEYER J.-J. *Formalising abilities and opportunities of agents*. Fundamenta Informaticae 34 (1998) P. 53–101.
109. LOPEZ F., LUCK M., D'INVERNO M. *A normative framework for agent based systems*, Computational & Mathematical Organization Theory, Volume 12(2 - 3), pages 227–250, 2005.
110. LYONS D.M., HENDRIKS A.J. *Exploiting patterns of interaction to achieve reactive behavior*, Artificial Intelligence, v.73 n.1-2, p.117–148, Feb. 1995.
111. MAS-COLLEL A., WHINSTON M.D., GREEN J.R. *Microeconomic theory*. N.Y.: Oxford Univ. Press, 1995. – 981 p.
112. MOUSSAID M., HELBING D., and THERAULAZ G. *How simple rules determine pedestrian behavior and crowd disasters*, Proceedings of the National Academy of Sciences 108 no. 17, (2011) 6884-6888.

113. MYERSON R.B. *Game theory: analysis of conflict*. London: Harvard Univ. Press, 2001. – 4th printing – 568 p.
114. MUSSE S., THALMANN D., *Hierarchical model for real time simulation of virtual human crowds*, IEEE Transactions on Visualization and Computer Graphics 7 (2) (2001) P. 152–164.
115. NAGEL R. *Unraveling in Guessing Games: An Experimental Study*, American Economic Review, 85, P. 1313–1326. – 1995.
116. NASH J.F. *Non-cooperative games* / Ann. Math. 1951. Vol. 54. P. 286–295.
117. OEDING D. *Verkehrsbelaufung und Dimensionierung von Gehwegen und anderen Anlagen des Fussgängerverkehrs, Strassenbau und Strassenverkehrstechnik*, Heft 22, Bonn, 1963.
118. OKAZAKIA S., MATSUSHITAA S. *A Study of Simulation Model for Pedestrian Movement with Evacuation and Queuing*, Proceedings of the International Conference on Engineering for crowd safety, 1993.
119. OWEN M., GALEA E., LAWRENCE P. *The EXODUS Evacuation Model Applied to Building Evacuation Scenarios*. <http://fseg.gre.ac.uk/exodus/>
120. PAN X., HAN C.S., DAUBER K., LAW K.H. *A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations*. October 2007, AI & Society, Volume 22 Issue 2. Publisher: Springer-Verlag.
121. Panic: A Quantitative Analysis // Web site is a supplement to results published in Dirk Helbing, Illés Farkas, and Tamás Vicsek: Simulating dynamical features of escape panic. Nature 407, 487-490 (2000). URL: <http://angel.elte.hu/~panic/> (дата обращения 12.08.2011).
122. PELECHANO N., BADLER N. 2006. “Modeling Crowd and Trained Leader Behavior during Building Evacuation.” IEEE Computer Graphics and Applications. vol. 26, no. 6, pp. 80–86, Nov/Dec.
123. RAO A.S. *AgentSpeak(L): BDI agents speak out in logical computable language*. Ed. by W. Van de Velde, J. W. Perram // Agents Breaking Away : Proc. of the Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World, (LNAI Vol. 1031). – Heidelberg, Germany : Springer-Verlag, 1996. – P. 42–55.
124. ROJEK G., CIKCIWA R., CETNAROWICZ K. *Algorithm of behavior evaluation in multi-agent system*. Computational Science - ICCS 2005 : 5th Int. Conference : Atlanta, GA, USA, May 22-25, 2005 : Pt. 3 eds. Vaidy S. Sunderam [et al.]. Lecture Notes in Computer Science ; LNCS 3516. Berlin; Heidelberg : Springer-Verlag, 2005.
125. SANTAMARIA Juan C., SUTTON R.C., RAM A. *Experiments with reinforcement learning in problems with continuous state and action spaces*. Adaptive Behavior, 6(2), 1998.
126. SCHIPPER B.C. *Imitators and optimizers in Cournot oligopoly*, University of Bonn, 2004.

127. SCHMIDT G.K., AZAM K. *Mobile robot path planning and execution based on a diffusion equation strategy*, in *Advanced Robotics*, Vol. 7, No. 5, pp. 479–490, 1993.
128. SHAO W., and TERZOPOULOS D. *Autonomous pedestrians*. *Graphical Models*, 69(5–6): P. 246–274, 2007.
129. SHOHAM Y, LEYTON-BROWN K. *Multiagent systems: Algorithmic, Game-Theoretical and Logical Foundations*. – Cambridge: Cambridge University Press, 2009.
130. SIMON H.A., EGIDI M., MARRIS R., & VIALE R. *Economics, bounded rationality, and cognitive revolution*. Edited by M. Egidi and R. Marris. Brookfield, VT: Edward Elgar, 1992.
131. SPIVEY, M. *The Z Notation (second edition)* - Hemel Hempstead, England: Prentice Hall International, 1992. – 260 p.
132. STAHL D.O. *The Evolution of Smart<sub>n</sub> Players*, *Games and Economic Behavior*, October 1993, 5(4), pp. 604–617.
133. STAHL D., WILSON P.W. *On Players' Models of Other Players: Theory and Experimental Evidence*, *Games and Economic Behavior*, July 1995, 10, P. 218–54.
134. SUTTON R.S. *Integrated architectures for learning, planning, and reacting based on approximate dynamic programming/* *Proceedings of the Seventh International Conference on Machine Learning*, P. 216–224, Morgan Kaufmann, 1990.
135. TOYAMA M.C., BAZZAN A. L.C., da SILVA R. *An agent-based simulation of pedestrian dynamics: from lane formation to auditorium evacuation*. In *AAMAS'06: Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 108–110, New York, NY, USA, 2006. ACM Press.
136. WOOLDRIDGE M. *An Introduction to Multi-Agent Systems*. – John Wiley and Sons. – 2002. – 376 p.
137. WOOLDRIDGE M., JENNINGS N. *Intelligent Agents: Theory and Practice* // *Knowledge Engineering Review*. – 1995. № 10 (2). – P. 115–152.
138. ZABARANKIN M., URYASEV S., PARDALOS P. *Optimal Risk Path Algorithms* // *Cooperative Control and Optimizatón*. Ch. 1 / Eds. Murphey R., Pardalos P. – Dordrecht: Kluwer Acad., 2002. – P. 271–303.