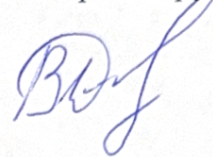


АО «НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ ВЫЧИСЛИТЕЛЬНЫХ
КОМПЛЕКСОВ ИМ. М. А. КАРЦЕВА»

На правах рукописи



Демидов Владислав Анатольевич

**ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА АВТОМАТИЧЕСКОЙ
АДАПТАЦИИ АРХИТЕКТУРЫ МНОГОАГЕНТНЫХ СИСТЕМ НА
ОСНОВЕ ЗООСОЦИАЛЬНОЙ МОДЕЛИ ПОВЕДЕНИЯ АГЕНТОВ**

Специальность 2.3.1 – Системный анализ, управление и обработка
информации, статистика

Диссертация на соискание ученой степени кандидата технических наук

Научный руководитель:
доктор технических наук, доцент
Чеканин Владислав Александрович

Москва - 2026

Содержание

ВВЕДЕНИЕ.....	5
ГЛАВА 1. СИСТЕМНЫЙ АНАЛИЗ СПОСОБОВ ОБУЧЕНИЯ	
ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ В ИЕРАРХИЧЕСКИХ МАС	14
1.1 Анализ архитектуры агентов	14
1.2 Анализ архитектуры иерархической МАС	22
1.3 Анализ способов организации совместных действий для множества интеллектуальных агентов	36
1.4 Формулировка научной проблемы и частных задач исследования. Определение критериев эффективности	47
1.5 Выводы.....	50
ГЛАВА 2. АНАЛИЗ МЕТОДОВ АВТОМАТИЧЕСКОЙ АДАПТАЦИИ	
ОРГАНИЗАЦИОННОЙ СТРУКТУРЫ ИЕРАРХИЧЕСКИХ МАС	51
2.1 Анализ типов организационных структур МАС	51
2.2 Анализ связей агентов в рамках организационных структур МАС	65
2.3 Автоматическая адаптация организационной структуры МАС путем модификации множества правил поведения агентов и функционирования, а также иерархического атрибутированного ориентированного графа МАС. Постановка задачи	77
2.3.1 Метод, основанный на модификации множества правил поведения агентов и функционирования МАС (реорганизация на макроуровне).....	79
2.3.2 Метод, основанный на модификации иерархического атрибутированного ориентированного графа МАС (реорганизация на микроуровне)	87
2.4 Выводы.....	93

ГЛАВА 3. ЗООСОЦИАЛЬНЫЙ МЕТОД АВТОМАТИЧЕСКОЙ АДАПТАЦИИ ОРГАНИЗАЦИОННОЙ СТРУКТУРЫ МАС..... 95

3.1 «Инстинкты» агентов в системах, основанных на зоосоциальных принципах	95
3.1.2 «Инстинкты саморазвития».....	99
3.1.3 «Ролевые инстинкты».....	102
3.2 Правила «доминирования»	104
3.3 Правила формирования связей	111
3.3.1 Правила формирования вертикальных связей.....	111
3.3.2 Правила формирования горизонтальных связей.....	120
3.4 Формирование множества правил, описывающих структуру и функционирование МАС, основанной на зоосоциальных принципах.....	122
3.5 Выводы.....	128

ГЛАВА 4. АРХИТЕКТУРА И АЛГОРИТМЫ РАБОТЫ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ОПТИМИЗАЦИИ ОРГАНИЗАЦИОННОЙ СТРУКТУРЫ МАС. ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА ЭФФЕКТИВНОСТИ ПРЕДЛОЖЕННОГО МЕТОДА И АЛГОРИТМОВ..... 129

4.1 Алгоритм адаптации организационной структуры МАС на макроуровне, основанный на зоосоциальном методе	129
4.1.1 Структура хромосомы.....	129
4.1.2 Функция приспособленности.....	130
4.1.3 Генетические операторы.....	134
4.1.4 Аппроксимация «функции стратегии».....	137
4.1.5 Структура алгоритма.....	140
4.2 Алгоритм оптимизации организационной структуры МАС в условиях неоднородной окружающей среды. Архитектура интеллектуальной системы оптимизации организационной структуры МАС.....	143

4.2.1 Алгоритм оптимизации организационной структуры МАС в условиях неоднородной окружающей среды.....	143
4.2.2 Архитектура интеллектуальной системы формирования и автоматической адаптации организационной структуры МАС.....	147
4.3 Описание параметров эксперимента.....	149
4.3.1 Глобальная функция качества.....	149
4.3.2 Описание параметров окружающей среды.....	151
4.3.3 Описание параметров МАС.....	152
4.3.4 Описание решаемой задачи.....	154
4.4 Описание результатов эксперимента.....	155
4.5 Экспериментальная оценка эффективности применения предложенного метода и алгоритмов при решении задачи доставки малогабаритных грузов.....	162
4.6 Выводы.....	166
ЗАКЛЮЧЕНИЕ	168
Словарь терминов.....	170
Список использованных источников	172
Приложение А. Акты о внедрении результатов диссертационного исследования	188
Приложение Б. Свидетельство о регистрации программы для ЭВМ....	192

ВВЕДЕНИЕ

Актуальность темы исследования. Актуальность темы обусловлена критической зависимостью производительности многоагентных систем (МАС) от выбора организационной структуры и её способности адаптироваться к изменениям окружающей среды. В современных прикладных задачах (управление распределёнными ресурсами, координация групп автономных устройств, динамическое планирование и др.) [100] условия функционирования многоагентных систем быстро и существенно меняются, поэтому статичная или слабо адаптивная организационная структура МАС становится узким местом.

Существующие подходы к адаптации архитектуры МАС обладают принципиальными ограничениями: первый подход, на основе корректировки связей при фиксированных базовых правилах, даёт лишь ограниченную гибкость и не позволяет оперативно перестраивать фундаментальные механизмы взаимодействия агентов. Второй подход (более гибкий), на основе адаптации базовых правил взаимодействия, реализован в настоящее время в относительно небольшом числе работ и обладает существенными недостатками (в частности, ограничен множеством моделируемых типов организационных структур и предопределённости условий их применения).

Это формирует потребность в разработке нового метода автоматической адаптации организационной структуры МАС, способного корректировать базовые правила взаимодействия агентов, в том числе, в режиме реального времени при существенных изменениях параметров среды. Перспективным направлением для реализации такого метода является использование биоинспирированных подходов – механизмов самоорганизации, наблюдаемых в зоосоциальных сообществах животных, которые могут служить основой для создания более гибких и устойчивых правил формирования и трансформации организационных структур МАС.

Степень её разработанности. Вопросы реализации многоагентных систем, в том числе иерархических МАС, получили широкое освещение в научной литературе. Фундаментальные аспекты теории МАС представлены в работах В. Б. Тарасова, М. Месаровича, Д. Мако, И. Такахары, С. Рассела и П. Норвига.

Проблема формирования и автоматической адаптации организационных структур МАС рассматривается в исследованиях ряда отечественных и зарубежных авторов: А. А. Никонов, А. В. Домышев, Д. Н. Сидоров, Bryan Horling, Victor Lesser, Ling Yu, Zhiqi Shen, Chunyan Miao, Luca Capezzuto, Danesh Tarapore, Sarvapali D. Ramchurn, Ghrieb N., Mokhati F., Tahar G., Douae Ahmadoun, Elise Bonzon, Cédric Louis Remane Buron, Pavlos Moraitis, Rahwan T., Michalak T. P., а также в трудах Р. Л. Дафта по теории организаций.

В настоящее время в области адаптации архитектуры МАС сложились два основных подхода:

1. **Корректировка множества связей между агентами** при неизменных базовых правилах их взаимодействия. В рамках этого подхода трансформируется древовидная структура связей в зависимости от состава агентов, поставленной задачи и текущих параметров среды, но фундаментальные правила (например, условия возникновения или разрыва связей) остаются фиксированными. Данный подход применяется, например, в работах Ling Yu, Zhiqi Shen, Chunyan Miao и Victor Lesser. Его ключевое ограничение – минимальные возможности оперативной корректировки структуры в процессе работы системы.

2. **Корректировка базовых правил взаимодействия агентов**, позволяющая автоматически менять фундаментальные условия функционирования МАС (включая процедуры образования и разрыва связей, число уровней иерархии и т. п.) в ходе её работы. Этот подход потенциально более универсален, однако число соответствующих исследований сравнительно невелико (Ling Yu, Zhiqi Shen, Chunyan Miao, Lesser V.,

Bistaffa F., Farinelli F., Cerquides J., Rodríguez-Aguilar J., Ramchurn S. D., Rahwan T., Michalak T. P., Ghrieb N., Mokhati F., Tahar G., Mahani M. N). При этом существующие работы характеризуются рядом существенных недостатков: ограниченным набором моделируемых типов организационных структур и жёсткой предопределённостью условий применения тех или иных структур.

Таким образом, несмотря на значительный объём исследований в области МАС и отдельные работы по адаптации их организационных структур, задача разработки универсального метода, обеспечивающего гибкую и оперативную корректировку базовых правил взаимодействия агентов (в том числе, в реальном времени), остаётся недостаточно проработанной.

Объектом исследования являются интеллектуальные многоагентные иерархические системы.

Предметом исследования являются методы формирования и автоматической адаптации архитектуры и организационной структуры интеллектуальных многоагентных иерархических систем.

Цель и задачи исследования. Целью диссертационного исследования является повышение эффективности функционирования интеллектуальных иерархических централизованных МАС за счет снижения временной сложности процедуры поиска оптимальной архитектуры или организационной структуры в условиях нестационарной и не полностью наблюдаемой внешней среды.

Для достижения поставленной цели необходимо решить следующие научные задачи:

1. Провести системный анализ способов обучения интеллектуальных агентов в иерархических МАС.
2. Провести анализ методов автоматической адаптации организационной структуры иерархических МАС.

3. Разработать зоосоциальный метод автоматической адаптации архитектуры и организационной структуры МАС.

4. Разработать архитектуру и алгоритм работы интеллектуальной системы оптимизации организационной структуры МАС.

5. Провести экспериментальную оценку эффективности предложенного метода и алгоритмов.

Методология и методы исследования. Теория и методы системного анализа, методология функционального моделирования IDEF0, теория многоагентных систем, математическая логика, нечеткая логика, теория биоинспирированных алгоритмов, теория обучения с подкреплением.

Научная новизна. В диссертации получены следующие результаты, характеризующиеся научной новизной:

– разработан зоосоциальный метод автоматической адаптации архитектуры и организационной структуры МАС. Отличие предложенного метода от аналогов заключается в следующем: обеспечивается возможность автоматического изменения множества связей между агентами в процессе работы без изменения базовых правил организации МАС; обеспечивается возможность оперативной корректировки организационной структуры за счёт изменения базовых правил организации (архитектуры); обеспечивается возможность автоматического выбора одного из следующих типов организационных структур в процессе работы МАС: «коалиция», «команда», «иерархическая структура», «федерация», «конгрегация» (п. 4, п. 9 паспорта научной специальности).

– разработан алгоритм адаптации организационной структуры МАС, основанный на зоосоциальном методе. Отличие предложенного алгоритма от аналогов заключается в следующем: генетический алгоритм используется для формирования множества хромосом, кодирующих параметры базовых правил, описывающих архитектуру МАС, а стратегия выбора конфигурации аппроксимируется ИНС и обучается с помощью алгоритма REINFORCE (п. 4 паспорта научной специальности).

– разработана архитектура интеллектуальной системы формирования и автоматической адаптации организационной структуры МАС, отличающаяся от аналогов применением базы знаний, хранящей информацию в виде кортежей, включающих параметры окружающей среды и решаемой задачи, хромосомы и значения показателей эффективности. База знаний формируется синтетически – с помощью предложенного алгоритма адаптации, и фактически – накапливается в процессе функционирования МАС. Накопленные данные используются для обучения ИНС, определяющей базовые правила, наиболее эффективные в текущих условиях (п. 10 паспорта научной специальности).

Теоретическая значимость заключается в развитии теории иерархических многоагентных систем. В ходе работы выявлены недостатки существующих подходов к формированию и адаптации архитектуры и организационной структуры иерархической МАС при решении задач мониторинга состояния окружающей среды. Кроме того, сформулированы гипотезы и идеи по оптимизации методов и алгоритмов формирования и адаптации архитектуры иерархической МАС, направленные на повышение их гибкости и автоматизации. Также проведена модернизация архитектуры интеллектуальной системы формирования и адаптации организационной структуры МАС - разработаны метод и алгоритмы, которые обеспечивают повышенную устойчивость МАС к колебаниям значений параметров окружающей среды.

Практическая значимость работы заключается в том, что применение разработанного метода и алгоритмов позволяет моделировать широкий спектр разновидностей организационных структур МАС за счёт использования различных комбинаций фиксированного множества параметров, отображающих базовые правила формирования и трансформации организационной структуры системы. При этом полученная МАС способна автоматически корректировать множество связей между агентами с учётом состава агентов, поставленной задачи и текущих

параметров окружающей среды. Кроме того, разработанный метод и алгоритмы могут быть использованы при решении разнообразных задач управления группами мобильных роботов в частности, для мониторинга состояния и картографирования инфраструктурных объектов и территорий сельскохозяйственного назначения, а также для доставки малогабаритных грузов.

Достоверность определяется строгостью приводимых математических доказательств, в получении которых был использован научно-методологический аппарат системного анализа, теория многоагентных систем, математическая логика, нечеткая логика, теория биоинспирированных алгоритмов, теория обучения с подкреплением.

Положения, выносимые на защиту:

1. Зоосоциальный метод автоматической адаптации архитектуры и организационной структуры МАС, предусматривающий реализацию множества правил поведения агентов и функционирования МАС в соответствии с принципами, характерными для социальных животных.

2. Алгоритм адаптации организационной структуры МАС на базе зоосоциального метода с комбинированным применением эволюционного и обучающего подходов, реализующий процедуру выбора конфигурации структуры в условиях неопределенности.

3. Архитектура интеллектуальной системы адаптации организационной структуры МАС с формализованным представлением знаний и механизмом их использования для обучения модели выбора организационной структуры.

Апробация работы. Основные положения и результаты работы докладывались и обсуждались на международных научно-практических конференциях, в том числе:

1. VI Международной научно-практической конференции преподавателей, аспирантов и научно-технических работников, посвящённой 65-летию со дня образования АМТИ. Краснодар, 2025 (2 доклада).

2. VII Международной научно-практической конференции – «Интеграция наук – 2025». Краснодар, 2025 (4 доклада).

Реализация и внедрение результатов работы. Результаты диссертации внедрены в практическую деятельность ООО «Аэромар-Краснодар» (г. Краснодар), в ООО «Олимп-Связь» (г. Краснодар) и АО «НИИВК им. М. А. Карцева» (г. Москва), а также в образовательный процесс ФГБОУ ВО «КубГТУ».

Соответствие паспорту специальности. Содержание диссертации соответствует п. 4 «Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта», п. 9 «Разработка объектно-ориентированных систем управления, принятия решений и оптимизации технических объектов», п. 10 «Методы и алгоритмы интеллектуальной поддержки при принятии управленческих решений в технических системах» паспорта специальности 2.3.1 Системный анализ, управление и обработка информации, статистика.

Публикации. По материалам диссертации опубликовано 14 печатных работ, в том числе в изданиях, рекомендованных ВАК РФ – 7 печатных работ, 1 свидетельство о регистрации программы для ЭВМ. В работах, опубликованных в соавторстве, лично автором получены следующие результаты: в работах [66, 64, 58] – зоосоциальный метод формирования и автоматической адаптации архитектуры и организационной структуры МАС, позволяющий на этапе формирования системы определять наиболее эффективную архитектуру и организационную структуру в зависимости от типа решаемой задачи и параметров окружающей среды, а также в случае изменения значений параметров окружающей среды автоматически выполнять их адаптацию в режиме реального времени; в работах [52, 57, 53] – архитектура интеллектуальной системы формирования и автоматической адаптации архитектуры и организационной структуры МАС, построенная на базе зоосоциального метода и разработанного алгоритма адаптации.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, заключения, списка использованных источников (содержит 102 наименования). Диссертация изложена на 187 страницах основного текста, включает 28 рисунков.

В первой главе проведен системный анализ способов обучения интеллектуальных агентов в иерархических МАС. В частности, выполнен анализ архитектур агентов и иерархических МАС, а также способов обучения интеллектуальных агентов совместным действиям. Сформулирована научная проблема и частные задачи исследования, определены критерии эффективности.

Во второй главе выполнен анализ типов организационных структур МАС, а также возможных вариантов связей между агентами. Приводится типовое описание методов автоматической адаптации организационной структуры иерархических МАС на макро и микроуровнях путем модификации множества правил поведения агентов и функционирования МАС.

В третьей главе приводится описание разработанного зоосоциального метода автоматической адаптации организационной структуры МАС, в частности перечислены возможные «инстинкты» агентов, описаны правила «доминирования» и формирования связей. Выполнено формирование множества правил, описывающих структуру и функционирование МАС, основанной на зоосоциальных принципах.

В четвертой главе приводится описание разработанной архитектуры и алгоритмов работы интеллектуальной системы оптимизации организационной структуры МАС. Приводится описание параметров и результатов экспериментальной оценки эффективности разработанного метода и алгоритмов.

В приложениях содержатся свидетельство о регистрации программы для ЭВМ и акты о внедрении в практическую деятельность ООО «Аэромар-Краснодар» (г. Краснодар), в ООО «Олимп-Связь» (г. Краснодар) и АО

«НИИВК им. М. А. Карцева», а также в образовательный процесс ФГБОУ ВО «КубГТУ».

ГЛАВА 1. СИСТЕМНЫЙ АНАЛИЗ СПОСОБОВ ОБУЧЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ В ИЕРАРХИЧЕСКИХ МАС

1.1 Анализ архитектуры агентов

Согласно [98], «агент представляет собой открытую систему, помещенную в некоторую среду, обладающую собственным поведением, удовлетворяющим некоторым экстремальным принципам». При этом для интеллектуального агента необходим следующий набор базовых характеристик: «способность к организации и реализации действий (активность)», «способность воспринимать состояние среды (реактивность)», «относительная независимость от окружающей среды или наличие некоторой «свободы воли»», способность к решению задач совместно с другими агентами («общительность»), «наличие собственных источников мотивации (целенаправленность)» [98]. Для реализации обозначенных характеристик агент должен обладать архитектурой, включающей «рецепторы» (средства, через которые осуществляется восприятие внешней и внутренней среды), «эффекторы» (средства, через которые осуществляется воздействие на окружающую среду), подсистема переработки информации, подсистема памяти, источник энергии [98]. Архитектура интеллектуального агента в общем виде представлена на рисунке 1.1 [98, 91].

На рисунке 1 представлены следующие элементы:

- 1 – внешняя среда;
- 2 – подсистема рецепторов;
- 3 – блок рецепторов внешней среды;
- 4 – блок рецепторов внутренней среды;
- 5 – подсистема эффекторов;
- 6 – подсистема памяти;
- 7 – блок фильтров;
- 8 – модель внешней среды;

Ниже приводятся типы элементов, представленных на рисунках 1.1, 1.2, рассматриваемые в данной работе:

1) Внешняя среда: некоторая реальная или виртуальная область, включающая статические и динамические трехмерные объекты (населенный пункт, территория предприятия, коммуникационная шахта, здание), описывается функцией $f_E(X)$, где X – параметры внешней среды.

2) Подсистема рецепторов:

а) Блок рецепторов внешней среды – набор устройств, позволяющих формализовать следующие характеристики объектов внешней среды: цветность, отображение в трехмерном виде, удаленность от агента (стереокамера, лазерное сканирующее устройство). Данный блок в момент времени t на вход получает некоторое восприятие внешней среды $O_t \sim f_E(X)$. На выходе блока имеем формализованное состояние внешней среды $S_{out} = f_{out}(O_t)$.

б) Блок рецепторов внутренней среды – набор датчиков, позволяющих фиксировать в момент времени t характеристики структурных элементов агента (P_t^{mem} – показатели функционирования подсистемы памяти, P_t^{cpu} – показатели функционирования подсистемы переработки информации, P_t^s – показатели функционирования источника энергии, P_t^{ex} – показатели функционирования подсистемы эффекторов). На выходе блока имеем формализованное состояние внутренней среды агента $S_{in} = f_{in}(P_t^{ex}, P_t^{mem}, P_t^{cpu}, P_t^s, P_t^{rec})$, где P_t^{rec} – параметры системы рецепторов.

3) Подсистема памяти (постоянная память):

а) Блок фильтров – база образов F_t , по которым агент осуществляет идентификацию элементов внешней среды.

б) Модель внешней среды – включает множество знаний об элементах внешней среды (внешние признаки – цвет, форма, габариты; поведенческие признаки; координаты; описание возможных поведенческих паттернов агента

как реакции при обнаружении того или иного объекта). Знания формализованы в виде кортежей E_t , включающих ретроспективные показатели параметров S_{out} , S_{in} , а также сгенерированных в ответ на них управляющих воздействий A_t .

в) Модель агента – включает описание внутренней структуры агента, формализованное в виде кортежей α_t , сформированных на основании S_{in} .

4) Подсистема переработки информации – вычислительная система агента, включающая микропроцессор и оперативное запоминающее устройство. На основании S_{out} , S_{in} идентифицирует текущее состояние агента $S_t = f_t(S_{out}, S_{in})$, формирует управляющее воздействие для подсистемы эффекторов $A_t = f_{action}(S_t, F_t, E_t, \alpha_t)$.

5) Подсистема эффекторов – транспортная платформа агента (колесная или гусеничная). На основании A_t реализует воздействие на внешнюю среду $I_t = f_{exem}(A_t)$.

6) Источник энергии – аккумуляторная батарея или топливный бак. При этом изменение текущего уровня ресурсов описывается функцией $f_s(p_t^{cpu}, p_t^{mem}, p_t^{ex}, p_t^{rec})$.

Поведение интеллектуального агента является целенаправленным [98], соответственно, агенту должно транслироваться некоторое описание цели – T , включающее следующие элементы:

- исходная модель внешней среды E , описываемая функцией $f_E(X)$;
- описание целевого состояния T относительно внешней среды – $T_E = f_E(T)$;
- функция для идентификации состояния агента относительно цели T – $f_{out}^T(O_t)$ – формирует пространство состояний S_{out}^T ;
- стандарт производительности агента P_T [91] – множество показателей, позволяющих оценить качество действий агента, направленных на достижение целевого состояния T ;

- множество фильтров F_T (при этом агент также имеет множество базовых фильтров F_{base} , $F_t = F_T \cup F_{base}$);
- функция для отображения пространства состояний во множество управляющих воздействий для эффекторов – $A_t = f_{action}^T(S_t, F_t, E_t, \alpha_t)$.

Таким образом, описание цели агента (поставленной задачи) может быть представлено в виде кортежа:

$$T = \langle T, f_E, f_{out}^T, P_T, F_T, f_{action}^T \rangle, (1.1)$$

Таким образом, интеллектуальный агент может быть представлен в виде кортежа:

$$\alpha = \langle T, f_{in}, F_{base}, \alpha_t, f_{exem}, f_s \rangle, (1.2)$$

Концептуальная модель интеллектуального агента, реализованная путем применения методологии IDEF0, представлена на рисунке 1.3.

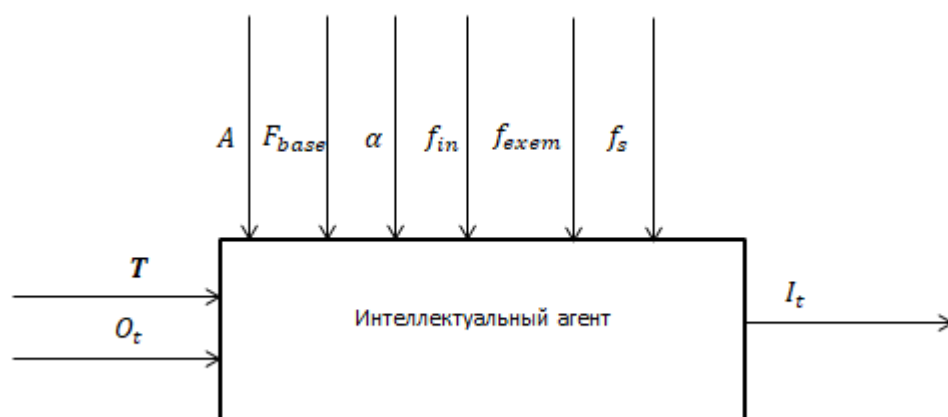


Рисунок 1.3 – Концептуальная модель интеллектуального агента

На рисунке 1.3 A – множество допустимых управляющих воздействий для агента α , α – модель агента α .

Декомпозиция модели, представленной на рисунке 1.3, иллюстрируется на рисунке 1.4.

На рисунке 1.4 блок 1 осуществляет сопоставление задачи T с моделью агента и множеством допустимых управляющих воздействий, по результатам которого активирует блоки 2, 3, либо отклоняет задачу T как невозможную для выполнения данным агентом по техническим причинам.

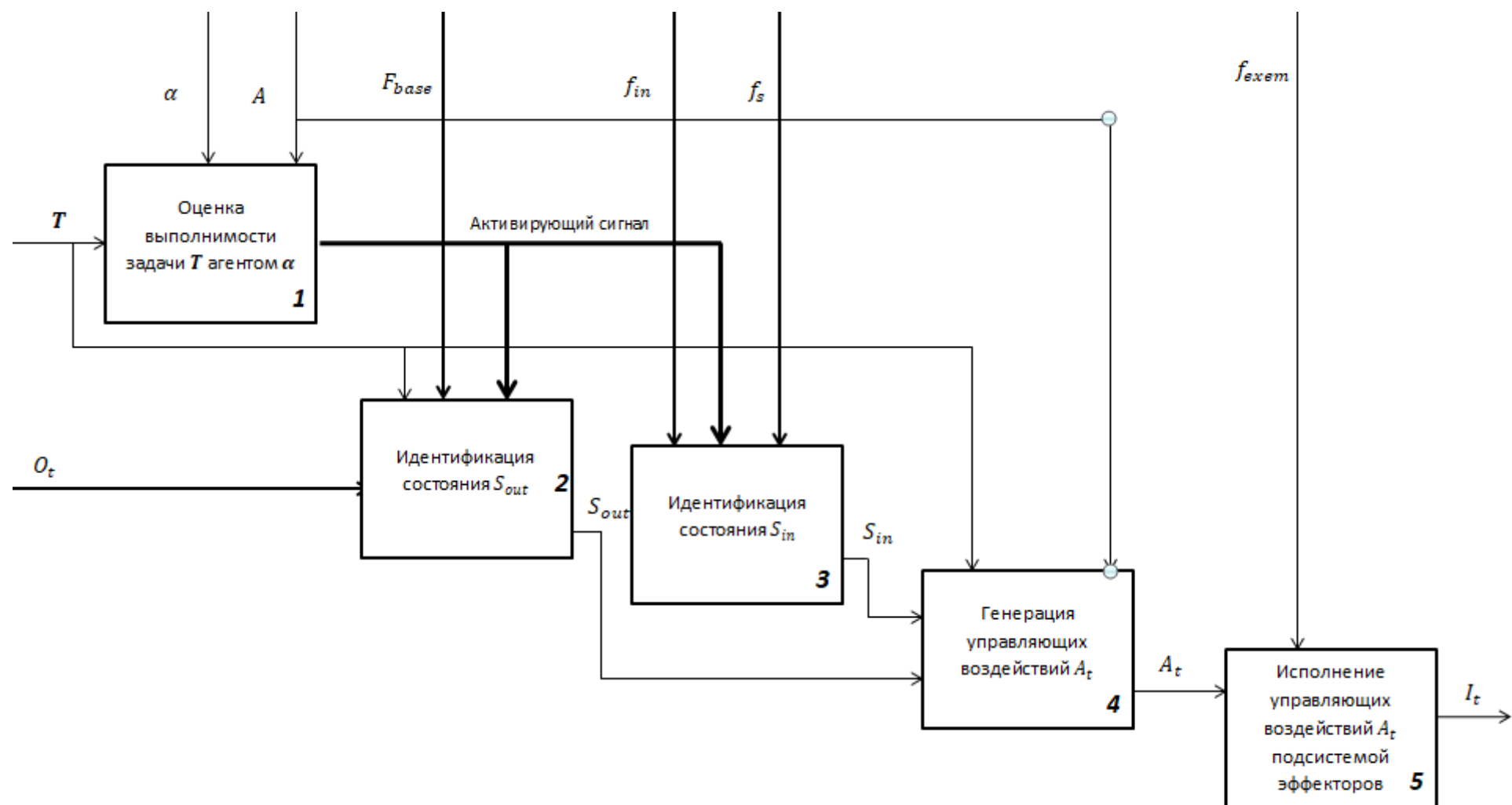


Рисунок 1.4 – Декомпозиция модели, представленной на рисунке 1.3

Рассматриваемый интеллектуальный агент обладает возможностью взаимодействия с подобными ему агентами, исходя из этого, выполним декомпозицию блока 4 (рисунок 1.5).

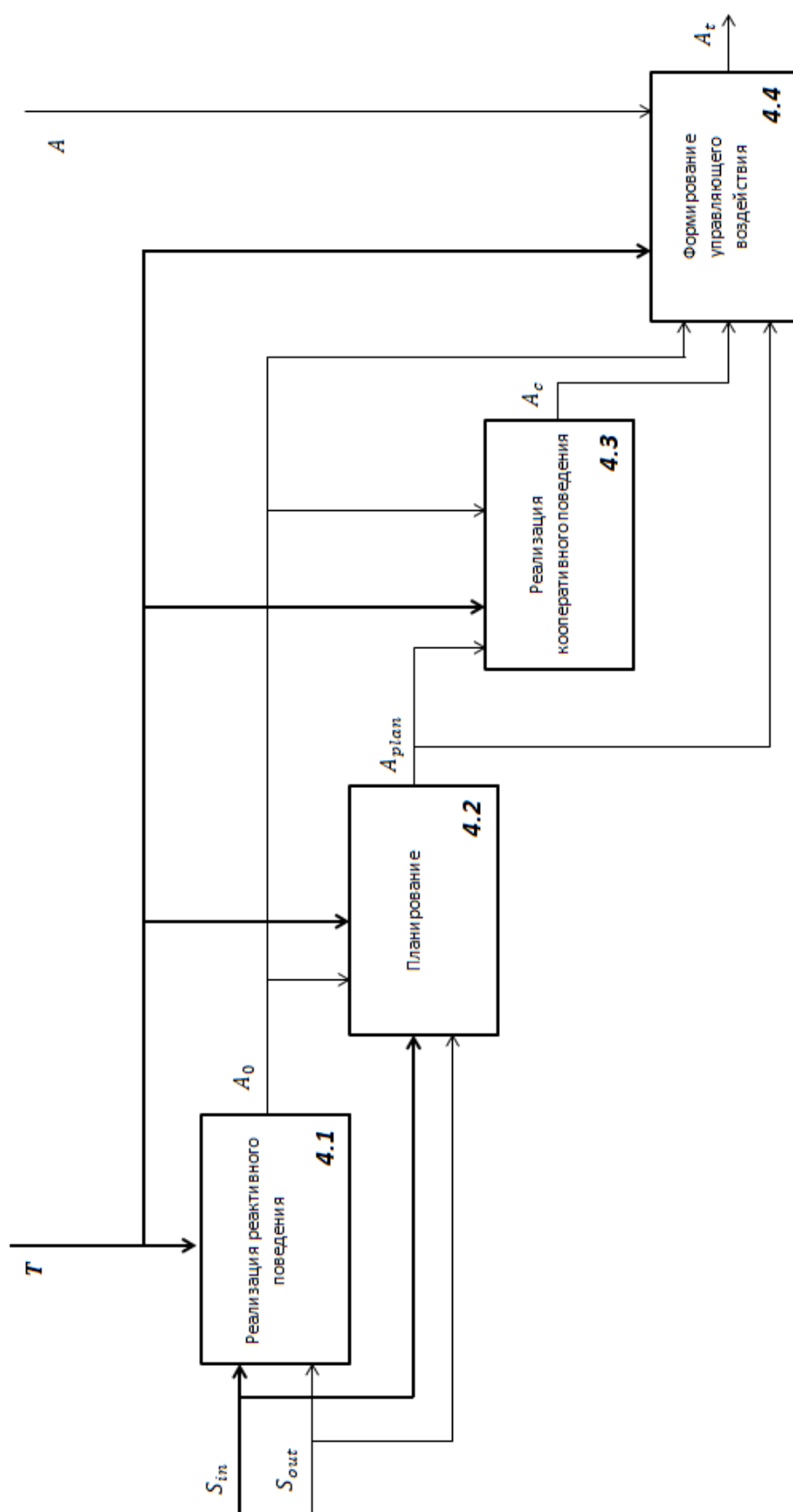


Рисунок 1.5 – Декомпозиция блока 4 (генерация управляющих воздействий A_t)

Декомпозиция блока 4 (рисунок 1.5) выполнена в соответствии с подходом к реализации вертикальной архитектуры интеллектуальных агентов «InteRRap». Она предусматривает наиболее полный набор функциональных возможностей в сравнении с аналогами («WILL-архитектура», «Touring Machine», «IDS-архитектура», «композиционная архитектура»), а именно распределение информации, координация и взаимодействие, коммуникация с другими агентами, анализ и планирование, организация реактивного и кооперативного поведения, ментальные функции агента [98, 85].

На рисунке 1.5 представлены следующие элементы:

1) Работы.

а) Реализация реактивного поведения – «использует базовые возможности агента по реактивному поведению», «позволяет агенту в стандартных ситуациях не обращаться к подсистеме планирования на основе знаний» [98].

б) Планирование – «содержит механизм планирования, позволяющий строить локальные планы агента, т.е. индивидуальные планы, не связанные с кооперативным поведением» [98].

в) Реализация кооперативного поведения – «создание планов совместного поведения агентов для достижения некоторых общих целей или выполнения ими обязательств перед другими агентами» [98].

г) Формирование управляющего воздействия – выбор управляющего воздействия A_t из множества A в соответствии с выходными данными блоков 4.1–4.3.

2) Стрелки.

а) A_0 – прототип управляющего воздействия, реализующего обработку базовых ситуаций (объезд препятствия и т.д.).

б) A_{plan} – перспективная (планируемая) последовательность действий агента.

в) A_c – план действий агента в рамках кооперации с иными агентами.

Отдельные интеллектуальные агенты могут быть объединены в многоагентную систему (МАС). Основные характеристики МАС рассматриваются в следующем подразделе.

1.2 Анализ архитектуры иерархической МАС

При подготовке текста п. 1.2 был использован авторский материал из опубликованных научных работ [58, 102].

Многоагентная система (МАС) – это «объединение отдельных интеллектуальных систем, основанных на знаниях» [98, 36, 33, 21].

Согласно [98], МАС состоит из следующих компонентов:

1) «Множество системных единиц» – $U = A \cup O$, включающее подмножества активных единиц A (агентов вида (1.2)), пассивных единиц O (объектов).

2) Внешняя среда E – «некоторое пространство, в котором существуют агенты и объекты».

3) Множество задач T и ролей R , назначаемых агентам.

4) Множество отношений R между агентами:

$$R = \{R_{\alpha_1} \times R_{\alpha_2} \times \dots \times R_{\alpha_{A_N}}\},$$

где $R_{\alpha_1}, R_{\alpha_2}, \dots, R_{\alpha_{A_N}}$ – множества отношений агентов $\alpha_1, \alpha_2, \alpha_{A_N}, A_N$ – число агентов в МАС, $A_N \in \mathbb{N}$.

5) Множество конфигураций C (организационных структур), формируемых агентами (множество ролей и отношений агентов, правил, регламентирующих их взаимодействие [19]).

6) Множество действий агентов A .

Для элементов E, T, R может быть сформировано множество состояний $S = f(E, T, R)$.

Согласно [98], основными направлениями в разработке МАС являются:

1) «Децентрализованный искусственный интеллект» – «исследуется деятельность автономного агента в динамической многоагентной среде».

2) «Распределенный искусственный интеллект» - «МАС образуется для решения задачи». «Осуществляется централизованное управление и координация действий нескольких интеллектуальных агентов».

Таким образом, по степени централизации может быть выделено два типа МАС:

1) Децентрализованные МАС. Агенты являются равноправными, однако могут вступать в кооперативные связи с целью максимизации индивидуального выигрыша. Структура агента децентрализованной МАС соответствует рисункам 1.3–1.5.

2) Централизованные МАС. Согласно [98], типичная последовательность действий при распределённом решении задач в рамках МАС выглядит следующим образом:

- «агент-субординатор проводит декомпозицию исходной проблемы на отдельные задачи, распределяемые между агентами-исполнителями» (при этом агент-субординатор производит подбор наиболее подходящего агента-исполнителя для реализации конкретной задачи);

- «каждый агент - исполнитель решает свою задачу, подчас также разделяя её на отдельные подзадачи»;

- «для получения общего результата производится композиция, интеграция частных результатов, осуществляемая агентом-интегратором» (функции агента-интегратора зачастую выполняет агент-субординатор).

Таким образом, в структуре централизованной МАС может быть выделено два уровня: уровень управления (агент-субординатор, агент-интегратор, предназначены для решения задачи координации [93, 50]), уровень исполнителей (агент-исполнитель). Формальное представление двухуровневой централизованной МАС, выполненное в соответствии с [87], приводится на рисунке 1.6.

На рисунке 1.6 приняты следующие обозначения:

α^c – агент-координатор (объединяет функции агента-субординатора и агента-интегратора);

$\alpha_1^p, \alpha_2^p, \dots, \alpha_{A_N}^p$ – агенты-исполнители;

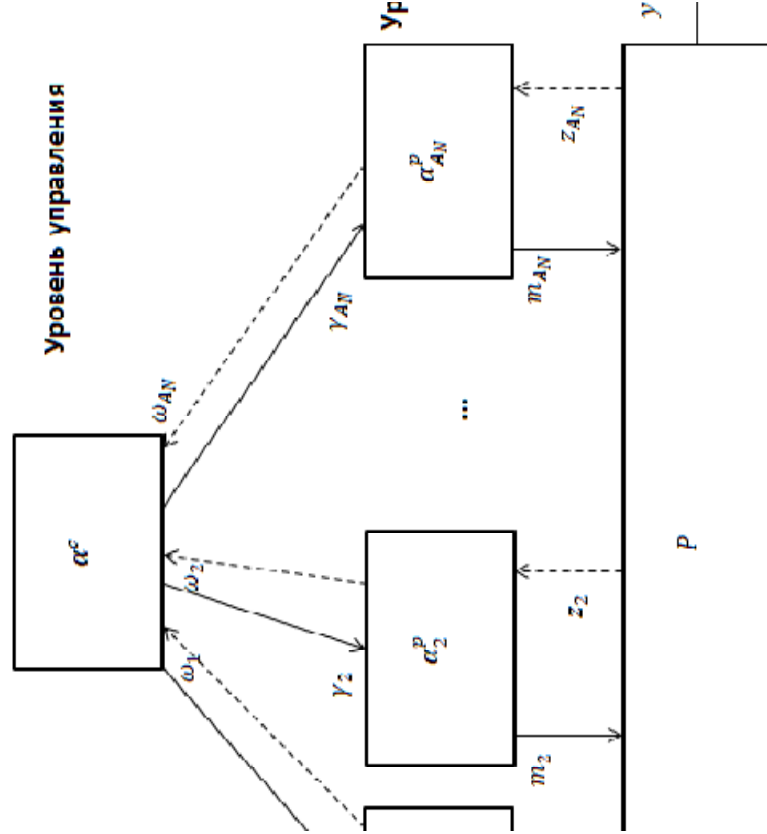


Рисунок 1.6 – Формальное представление управляемой централизованной МАС

P – управляемый процесс, имеющий множество выходов y , а также подвергаемый воздействию возмущений внешней среды x ;

$\gamma_1, \gamma_2, \dots, \gamma_{A_N}$ – координирующие воздействия [87], поступающие от агента-координатора к агентам-исполнителям;

$\omega_1, \omega_2, \dots, \omega_{A_N}$ – информационные сигналы [87], поступающие от агентов-исполнителей к агенту-координатору, при этом $\gamma_i^t = f(\omega_i^{t-1})$, $i = \overline{1, A_N}$;

$m_1, m_2, \dots, m_{A_N}$ – управляющие сигналы [87], поступающие от агентов-исполнителей к управляемому процессу P ;

$z_1, z_2, \dots, z_{A_N}$ – информационные сигналы [87], поступающие от управляемого процесса P к агентам-исполнителям, при этом $m_i^t =$

$f(\gamma_i^t, z_i^{t-1})$, $\omega_i^t = f(z_i^t)$, $V(\omega_i^t) < V(z_i^t)$, $i = \overline{1, A_N}$, V – некоторая функция, позволяющая оценить объем информации.

Структура агентов централизованной МАС имеет следующие особенности:

1) Структура агента-координатора полностью соответствует структуре, представленной на рисунках 1.3–1.5.

2) Структура агента-исполнителя аналогична структуре агента-координатора за исключением блока 4.3 (рисунок 1.5), отвечающего за выработку решений в рамках кооперативного поведения. Задача выработки решений в рамках кооперативного поведения возлагается на агента-координатора.

«Иерархия возникает, когда системы, которые функционируют как целое на одном уровне, функционируют как части системы более высокого уровня, становясь подсистемами этой системы» [19], «любая иерархия состоит из вертикально соподчиненных подсистем» [87], следовательно, система, представленная на рисунке 1.6, является иерархической.

Таким образом, рассматриваемые централизованные иерархические МАС являются объектом исследования в данной работе.

Концептуальная модель централизованной МАС, реализованная путем применения методологии IDEF0, представлена на рисунке 1.7.

На рисунке 1.77 представлены следующие элементы:

1) Входные стрелки.

а) T – поставленная задача, имеющая формат, соответствующий (1.1).

б) O_{t-1} – множество восприятий состояния управляемого процесса в момент времени $t - 1$ (соответствует $z_1, z_2, \dots, z_{A_N}$ на рисунке 1.6).



Рисунок 1.7 – Концептуальная модель централизованной MAC

в) R_{perf}^T – множество ролей уровня исполнителей (например, агент-исследователь, агент-сборщик проб и т. п.) зависит от характера выполняемой задачи T , поэтому вводится вместе с её описанием. Исходя из этого, дополним формулу (1.1) следующим образом:

$$T = \langle T, f_E, f_{out}^T, P_T, F_T, f_{action}^T, R_{perf}^T \rangle, \quad (1.3)$$

где R_{perf}^T – множество ролей уровня исполнителей, которые могут возлагаться на агентов-исполнителей для достижения общей цели T [93].

2) Стрелки «управление».

а) A – множество агентов.

б) C – множество допустимых конфигураций. Организационная структура: «коалиция», «команда», «конгрегация», «федерация» и т.д. [19].

Понятие «организационная структура» подробно рассматривается в работах [19, 56], где централизованным MAC в большей степени соответствуют конфигурации типа «команда». Агентам назначаются определенные роли, комбинация которых приводит к достижению поставленной цели, данная конфигурация предусматривает максимизацию группового выигрыша). «Федерация» (множество агентов делегирует часть

своих полномочий конкретному «выдающемуся» агенту, в число таких полномочий могут входить: декомпозиция поставленной задачи, распределение сформированных подзадач, мониторинг их выполнения, связь с лицом, принимающим решения) [19].

в) R_{coord}^C – множество ролей уровня управления (агент-координатор, агент-субординатор, агент-интегратор).

г) A – множество примитивных действий, доступных для выполнения агентами множества A (может быть обусловлено техническими возможностями агентов).

3) Выходные стрелки:

m_t – множество управляющих сигналов, воздействующих на управляемый процесс, сгенерированных как ответ на восприятия O_{t-1} (соответствует $m_1, m_2, \dots, m_{A_N}$ на рисунке 1.6).

На рисунке 1.8 представлена декомпозиция модели, показанной на рисунке 1.7.

На рисунке 1.8 представлены следующие элементы:

1) Стрелки.

а) $C_{T_n}^A$ – сформированная конфигурация МАС для решения задачи T с помощью агентов A , n – номер текущей модификации конфигурации МАС, $n = \overline{1, C_{T_N}^A}$, $C_{T_N}^A$ – общее число сформированных конфигураций за время решения задачи T , $C_{T_N}^A \in \mathbb{N}$. При этом множество $C_{T_n}^A$ имеет следующий вид [19]:

$$C_{T_n}^A = \langle \{ \langle \alpha_i, R_j \rangle \}, R, r \rangle, \quad (1.4)$$

где $\langle \alpha_i, R_j \rangle$ – множество пар «агент-назначенная роль» [29, 4], $\alpha_i \in A$, $R_j \in R$, $i = \overline{1, A_N}$, $j = \overline{1, R_N}$, A_N – число агентов МАС, R_N – допустимое число возможных ролей.

б) $C_{T_{n+1}}^A$ – новая конфигурация МАС (в формате (1.4)), сформированная при смене режима работы (с режима обучения на режим максимизации эффективности).

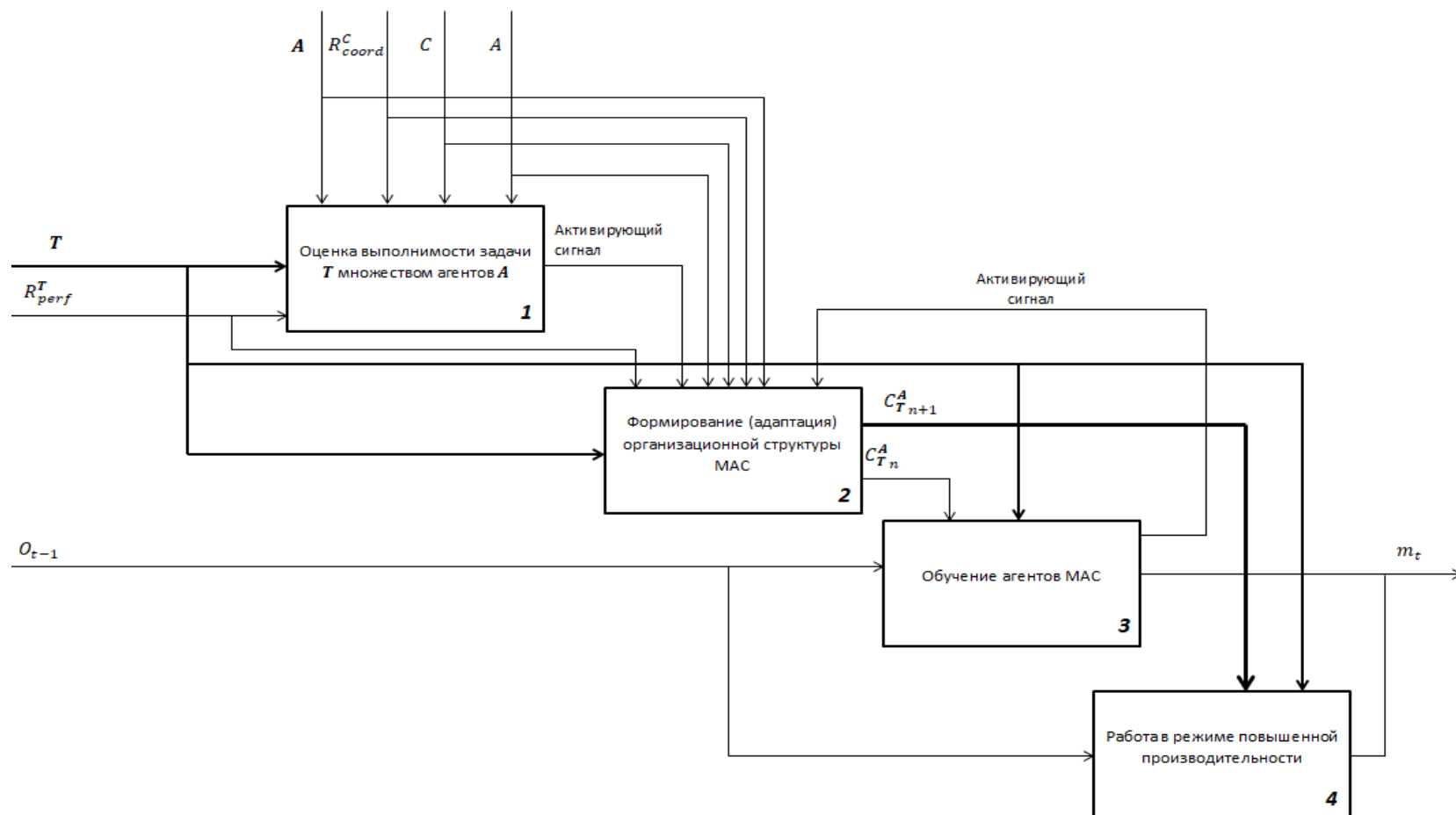


Рисунок 1.8 – Декомпозиция модели централизованной MAC

2) Работы.

а) Оценка выполнимости задачи T множеством агентов A .

Осуществляет проверку выполнимости соответствия техническим характеристикам агентов множества A требованиям задачи T . В случае, если агенты обладают такими возможностями, на блок 2 передается активирующий сигнал, в ином случае задача отклоняется.

б) Формирование (адаптация) организационной структуры МАС.

Осуществляет формирование либо адаптацию конфигурации (организационной структуры) МАС для решения поставленной задачи T . При этом могут выполняться следующие виды адаптации:

- «Количественная адаптация» (рассматривается в работе [67]), предусматривающая модификацию множества R , в меньшей степени - $\langle \alpha_i, R_j \rangle$ (см. формулу (1.4)).

- «Качественная адаптация» (рассматривается в работах [60, 61, 49]), предусматривающая модификацию множества r (см. формулу (1.4)).

В рамках данной работы также выполняется декомпозиция задачи T на отдельные подзадачи t_k и их распределение между агентами $\alpha_i \in A$, $k = \overline{1, t_N}$, t_N – число подзадач, на которые была разбита задача T .

в) Обучение агентов МАС.

В работе МАС может быть выделено два основных режима: режим обучения, режим повышенной производительности. Подробно данный вопрос рассматривается в работах [60–61]. В частности, в них описывается метод, направленный на модификацию конфигурации МАС с целью максимизации эффективности процесса обучения, либо общей производительности системы. В данном блоке реализуется обучение МАС, имеющей адаптированную для этих целей конфигурацию $C_{T_n}^A$. При достижении заданных критериев эффективности обучения (стандарт производительности P_T , формула (1.1)) происходит модификация конфигурации $C_{T_n}^A$ и активация блока 4.

г) Работа в режиме повышенной производительности.

Данная работа направлена на максимизацию производительности МАС. С этой целью на блоке 3 формируется специально адаптированная для этого конфигурация $C_{T_{n+1}}^A$. В случае выполнения определенных условий (существенное изменение параметров управляемого процесса) происходит активация блока 2 с целью адаптации конфигурации МАС для перехода в режим обучения (блок 3).

При выполнении декомпозиции работы 3 (обучение агентов МАС) необходимо учитывать двухуровневую структуру МАС:

1) Целью обучения агентов уровня управления является максимизация общего выигрыша по стандарту производительности P_T , достигнутая в результате оптимального распределения подзадач t_k между агентами α_i .

2) Целью обучения агентов уровня исполнителей является максимизация общего выигрыша по стандарту производительности P_T , достигнутая в результате выработки оптимальных последовательностей управляющих сигналов m_t к управляемому процессу.

Примечание: стандарт производительности P_T должен содержать показатели и критерии для оценки эффективности как отдельного агента, так и МАС в целом.

С учетом обозначенных выше особенностей выполним декомпозицию работы «Обучение агентов МАС» (рисунок 1.9).

На рисунке 1.9 представлены следующие обозначения:

1) Стрелки.

а) γ_{t-1} – множество координирующих воздействий к агентам уровня исполнителей.

б) r_{t-1}^{coord} – оценки эффективности действий агентов уровня управления.

в) r_{t-1}^{perf} – оценки эффективности действий агентов уровня исполнителей (соответствуют информационным сигналам $\omega_1, \omega_2, \dots, \omega_{A_N}$, см. рисунок 1.6).

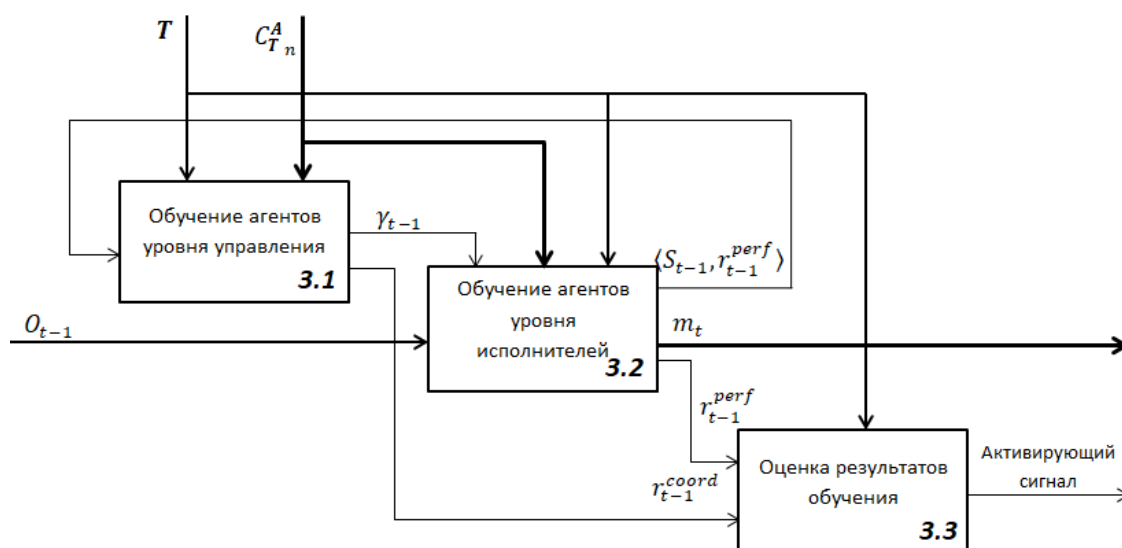


Рисунок 1.9 – Декомпозиция работы «Обучение агентов МАС»

г) S_{t-1} – множество состояний агентов-исполнителей.

2) Работы.

а) Обучение агентов уровня управления (см. рисунок 1.10).

б) Обучение агентов уровня исполнителей (см. рисунок 1.11).

в) Оценка результатов обучения – выполняет оценку множеств r_{t-1}^{coord} , r_{t-1}^{perf} с целью принятия решения о переходе к режиму повышенной производительности (в случае, если МАС достаточно обучена).

На рисунке 1.10 приняты следующие обозначения:

1) Стрелки.

а) S_{t-1} – состояние, в котором находится агент-координатор в момент времени $t - 1$ (должно учитывать состояние агентов-исполнителей). В случае если S_{t-1} имеет вид $S_{t-1}^{\alpha_1}, S_{t-1}^{\alpha_2}, \dots, S_{t-1}^{\alpha_{A_N}}$ ($\alpha_1, \alpha_2, \dots, \alpha_{A_N-1}$ – агенты-исполнители, A_N – число агентов во множестве A), размерность пространства состояний экспоненциально возрастает, $S_{t-1} = f_S(f_{out}^T, f_{in})$, f_S – функция для идентификации текущего состояния агента-координатора; f_{out}^T – функция для идентификации состояния агента-исполнителя в рамках внешней среды относительно цели поставленной задачи T (передается вместе с описанием задачи); f_{in} – функция для идентификации состояния внутренней среды

агента-исполнителя (его техническое состояние), интегрирована в прошивку агента.

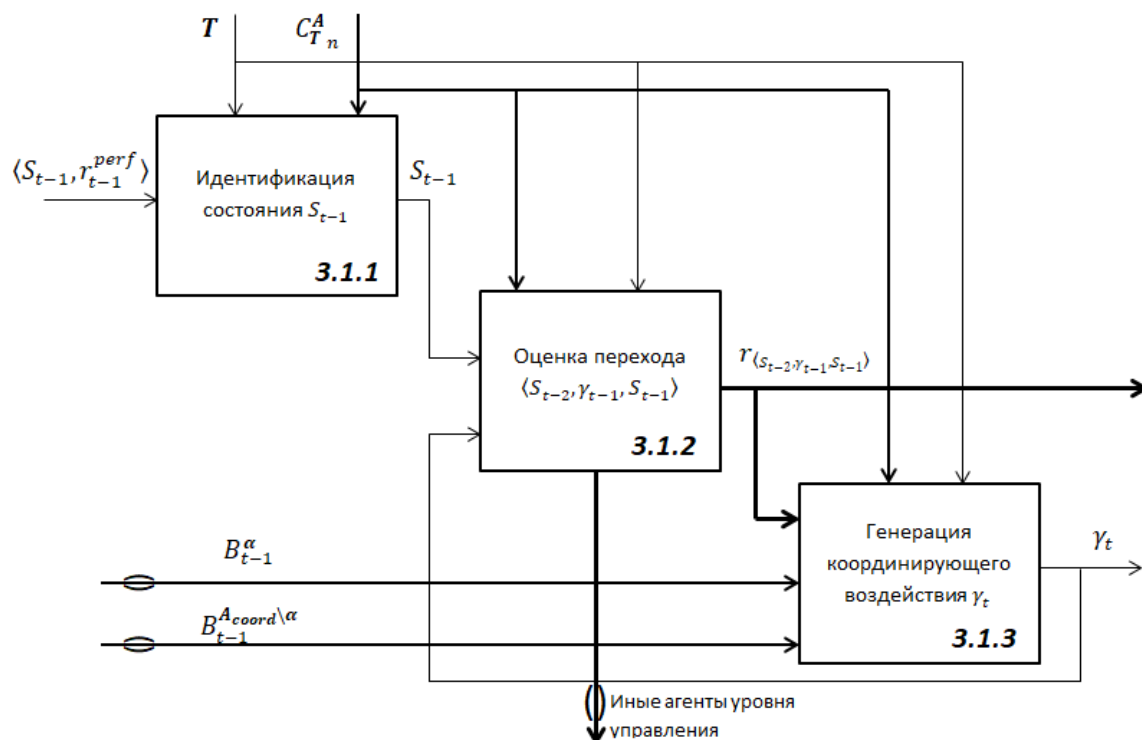


Рисунок 1.10 – Декомпозиция работы «Обучение агентов уровня управления»

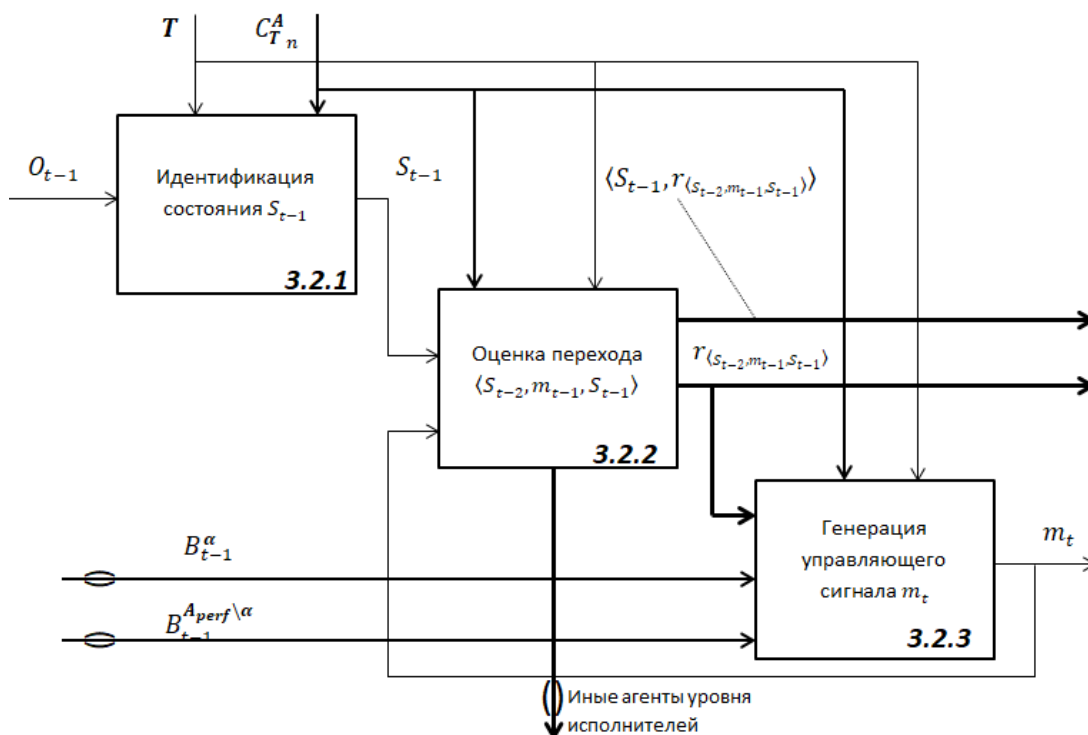


Рисунок 1.11 – Декомпозиция работы «Обучение агентов уровня исполнителей»

б) $r_{\langle S_{t-2}, \gamma_{t-1}, S_{t-1} \rangle}$ – оценка выигрыша от осуществления перехода $\langle S_{t-2}, \gamma_{t-1}, S_{t-1} \rangle$ (выигрыш может быть как положительный, так и отрицательный) (соответствует r_{t-1}^{coord}).

в) B_{t-1}^α – «буфер опыта» данного агента-координатора α .

г) $B_{t-1}^{A_{coord} \setminus \alpha}$ – «буфер опыта» иных агентов-координаторов МАС (множество $A_{coord} \subset A$).

2) Работы.

а) Идентификация состояния S_{t-1} – идентификация состояния агента-координатора.

б) Оценка перехода $\langle S_{t-2}, \gamma_{t-1}, S_{t-1} \rangle$ – генерация оценки эффективности (r_{t-1}^{coord}) действия агента-координатора γ_{t-1} , выполнение которого привело к осуществлению перехода $S_{t-2} \rightarrow S_{t-1}$. При этом под S_{t-2} , S_{t-1} следует понимать оценку общего состояния агентов множества A , координируемого данным агентом $\alpha \in A_{coord}$.

в) Генерация координирующего воздействия γ_t – выработка координирующего воздействия γ_t для множества агентов-исполнителей.

На рисунке 1.11 приняты следующие обозначения:

1) Стрелки.

а) S_{t-1} – текущее состояние агента-исполнителя, определяемое как $S_{t-1} = f_S(f_{out}^T, f_{in})$.

б) $r_{\langle S_{t-2}, m_{t-1}, S_{t-1} \rangle}$ – оценка выигрыша от осуществления перехода $\langle S_{t-2}, m_{t-1}, S_{t-1} \rangle$ (выигрыш может быть как положительный, так и отрицательный) (соответствует r_{t-1}^{perf}).

в) B_{t-1}^α – «буфер опыта» данного агента-исполнителя α .

г) $B_{t-1}^{A_{perf} \setminus \alpha}$ – «буфер опыта» иных агентов-исполнителей МАС (множество $A_{perf} \subset A$).

2) Работы.

а) Идентификация состояния S_{t-1} – идентификация состояния агента-исполнителя.

б) Оценка перехода $\langle S_{t-2}, m_{t-1}, S_{t-1} \rangle$ – генерация оценки эффективности (r_{t-1}^{perf}) действия агента-исполнителя m_{t-1} , выполнение которого привело к осуществлению перехода $S_{t-2} \rightarrow S_{t-1}$.

в) Генерация управляющего сигнала m_t – выработка управляющего сигнала m_t для управляемого процесса.

Процесс генерации воздействий γ_t, m_t регламентируется функциями $f_{action_{coord}}^T = \gamma_t = f(\{\omega_{t-1} = \langle S_{t-1}, r_{t-1}^{perf} \rangle_i\})$, $f_{action_{perf}}^T = m_t = f(\gamma_t, O_{t-1})$, процесс обучения – функциями $f_{edu_{coord}}^T(B_{t-1}^\alpha, B_{t-1}^{A_{coord} \setminus \alpha}, \langle S_{t-2}, \gamma_{t-1}, S_{t-1} \rangle, r_{t-1}^{coord})$, $f_{edu_{perf}}^T(B_{t-1}^\alpha, B_{t-1}^{A_{perf} \setminus \alpha}, \langle S_{t-2}, m_{t-1}, S_{t-1} \rangle, r_{t-1}^{perf})$. Исходя из этого, перепишем формулу (1.3) в следующем виде:

$$\mathbf{T} = \langle T, f_E, f_{out}^T, P_T, F_T, f_{action}^T = \langle f_{action_{coord}}^T, f_{action_{perf}}^T \rangle, R_{perf}^T \rangle, \quad (1.4)$$

а также добавим к блокам 3.1.2, 3.2.2 (рисунки 10–11) соответствующие стрелки «управление» – $f_{edu_{coord}}^T, f_{edu_{perf}}^T$.

Примечание: блоки 3.1, 3.2 (рисунок 1.9) предназначены для реализации обучения на основе «ситуационного опыта» (применяется при решении поставленных задач), при этом «базовый опыт» (применяется для выполнения элементарных действий, таких как обход препятствий, применяется в блоке 4.1, см. рисунок 1.5).

Функции $f_{edu_{coord}}^T, f_{edu_{perf}}^T$ реализуют методы обучения агентов-координаторов и агентов-исполнителей, анализ которых выполняется далее, в следующем подразделе.

1.3 Анализ способов организации совместных действий для множества интеллектуальных агентов

При подготовке текста п. 1.3 был использован авторский материал из опубликованных научных работ [58, 52, 47].

«Агент обучается, если он улучшает свою производительность в результате наблюдения за миром» [92]. Анализ структуры МАС, проведенный в подразделе 1.2, позволил выделить компоненты агентов уровня управления, предназначенные для обучения решению задачи координации [93], которая в общем виде может быть сформулирована следующим образом:

$$\gamma_t = f(\{\omega_{t-1}\}): \vartheta(\gamma_t, P_T) \rightarrow \max,$$

где ϑ – оценка γ_t по стандарту производительности P_T .

Соответственно, в данном подразделе выполним анализ основных методов обучения, применимых к решению задачи координации [79, 101].

В [92] выделяют три основных типа обучения:

1) Обучение с учителем [37, 17]. В [92] задача обучения с учителем формулируется следующим образом:

«При наличии обучающего набора данных в виде пар входных и выходных значений (или прецедентов)

$$(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N), \quad (1.5)$$

где N – число пар «вход-выход», сгенерированных неизвестной функцией $y = f(x)$, найти функцию h (гипотеза о мире), аппроксимирующую истинную функцию f ».

Основным недостатком данного подхода, ограничивающим его применение к решению задачи координации, является необходимость наличия пар вида (1.5), что не представляется возможным, учитывая огромное количество возможных комбинаций

$$(\omega_{t-1}, \gamma_t)^{T_{N*TN}}, \quad (1.6)$$

где ω_{t-1} – множество информационных сообщений от агентов-исполнителей. $\omega_{t-1} = \{\omega_{t-1}^1, \omega_{t-1}^2, \dots, \omega_{t-1}^{|A_{perf}|}\}$, $|A_{perf}|$ – число агентов-исполнителей. T_N – число задач, решаемых данной МАС. $\overline{\gamma}_N$ – среднее арифметическое числа воздействий γ_t , применяемых за время решения задачи T . γ – множество координирующих воздействий от агента-координатора $\gamma_t = \{\gamma_{t-1}^1, \gamma_{t-1}^2, \dots, \gamma_{t-1}^{|A_{perf}|}\}$.

Однако в данном случае нет необходимости генерировать все возможные варианты пар (1.6). Достаточно лишь получить некоторое количество характерных образцов для каждой типовой ситуации s_k , описываемой набором $\{\omega_{t-1}^1 \pm \Delta\omega, \omega_{t-1}^2 \pm \Delta\omega, \dots, \omega_{t-1}^{|A_{perf}|} \pm \Delta\omega\}_k$, где k – индекс типовой ситуации, $k = \overline{1, s_N}$, s_N – число типовых ситуаций, $\Delta\omega$ – некоторое отклонение информационного сообщения, позволяющее относить его к типу s_k . В результате обучения будет получена аппроксимирующая функция f , позволяющая обобщить пары (1.6). Но в этом случае возникает необходимость определения всех возможных типовых ситуаций s_k (получение и разметка обучающей выборки). Отдельным направлением, примыкающим к обучению с учителем, является имитационное обучение, при котором агент формирует стратегию поведения на основе подражания демонстрациям, полученным от эксперта или эталонной системы [26, 18, 41, 27]. Применение имитационного обучения к решению задачи координации сопряжено с теми же трудностями, что и обучение с учителем: требуется наличие размеченных примеров целесообразного поведения агентов, получение которых в условиях не полностью наблюдаемой внешней среды затруднено.

2) Обучение без учителя [34, 13]. Для определения типов s_k возможно применение методов обучения без учителя. Данный подход заключается в «выявлении во входных примерах потенциально полезных закономерностей» [92], однако он обладает рядом существенных недостатков:

- необходимость проверки полученных кластеров, поскольку обнаруженные закономерности не всегда могут быть полезны;
- большая вычислительная сложность.

3) Обучение с подкреплением [24, 2, 23]. Основным принцип данного подхода поясняется на схеме, представленной на рисунке 1.12 [95]. Агент, находясь в состоянии S_t , в момент времени t выполняется некоторое действие A_t , изменяя, таким образом, состояние внешней среды на S_{t+1} , получая при этом подкрепление R_{t+1} .



Рисунок 1.12 – Иллюстрация принципа обучения с подкреплением

В [92] выделяют следующие типы обучения с подкреплением:

а) «Обучение с подкреплением на основе модели». Предполагает, что выработка сигналов вознаграждения агента реализуется на основе «модели перехода среды» (например, при реализации игры в шахматы, когда правила известны). Системы обучения с подкреплением на основе данной модели часто изучают функцию ценности $V(S)$ («функция ценности состояния S при стратегии π , обозначаемая $V_\pi(S)$ – это ожидаемый доход, когда агент начинает работу в состоянии S и в дальнейшем следует стратегии π », при этом под стратегией понимается «отображение множества воспринимаемых состояний среды на действия, предпринимаемые в этих состояниях») [93, 92, 95].

б) «Обучение с подкреплением без модели». «В этих подходах агент изначально не знает и не изучает модель перехода для окружающей среды. Вместо этого он учится непосредственно представлению о том, как себя вести» [92]. Разновидностью обучения с подкреплением без модели является подход, основанный на изучении полезности действий (Q-обучение) [92]. «Основывается на изучении функции ожидаемой полезности действия $Q(S, A)$, определяющую сумму вознаграждений от состояния S и далее, если будет выполнено действие A . При известной Q-функции агент может выбрать, что ему делать в состоянии S , посредством поиска действия с самым высоким значением ожидаемой полезности Q » [92].

Главным преимуществом данного подхода является отсутствие необходимости в предоставлении агенту обучающей выборки, ему транслируется некоторая функция вознаграждения, позволяющая оценить успешность его действий. «Функция вознаграждения часто лаконична и проста», «от разработчиков не требуется быть экспертами, способными предоставить информацию о правильном действии в любой ситуации, как в случае, когда применяется обучение с учителем» [92].

Таким образом, применение парадигмы обучения с подкреплением в данном случае представляется более целесообразным. Очевидно, что при решении задачи координации действий агентов-исполнителей также возможны ситуации, при которых внешняя среда является неопределенной, частично наблюдаемой (например, при обследовании некоторого инфраструктурного объекта в отсутствии карты). В этом случае более обоснованным видится применение «обучения с подкреплением без модели». Далее рассматриваются основные подходы к применению данной парадигмы в рамках МАС.

Пусть $A = \{\alpha_i\}$ – множество агентов некоторой централизованной МАС MS , $i = \overline{1, |A|}$, $|A|$ – мощность множества A , $A = \{a_j | a_j \in A_{\alpha_i}\}$ – множество примитивных действий, доступных для выполнения агентами α_i ,

$S_t = \{S_t^{\alpha_1}, S_t^{\alpha_2}, \dots, S_t^{\alpha_{|A|}}\}$ – множество состояний агентов α_i в момент времени t , $S_t^{\alpha_i} \in S$, S – множество допустимых состояний МАС. Переход к задаче многоагентного обучения с подкреплением предполагает, что состояние системы MS в момент времени t (глобальное состояние) складывается из состояний отдельных агентов множества A [89]:

$$S_t^{MS} = \{S\} = \{S_t^{\alpha_1}, S_t^{\alpha_2}, \dots, S_t^{\alpha_{|A|}}\}.$$

Аналогичным образом имеем:

$$A_t^{MS} = \{A_t^{\alpha_1}, A_t^{\alpha_2}, \dots, A_t^{\alpha_{|A|}} | A_t^{\alpha_i} \in A\}.$$

Таким образом, пространство совместных действий $\{\langle S_t^{MS}, A_t^{MS} \rangle\}$ представляет декартово произведение $A \times S \times A$. Следовательно, имеем одну из основных проблем многоагентного обучения с подкреплением (МОП) – проблему размерности пространства совместных действий $\{\langle S_t^{MS}, A_t^{MS} \rangle\}$. Очевидно, что в этом случае процедура поиска оптимального решения обладает неприемлемо высокой пространственно-временной сложностью.

Для задачи многоагентного обучения с подкреплением также характерна проблема не стационарности агентов относительно друг друга, которая влечет необходимость учета факторов: степень информированности агентов о состоянии и намерениях друг друга, порядок действий агентов, обработка коллизий (конфликтные ситуации между агентами, когда они, например, планируют переход в одну и ту же область внешней среды).

В [35] также автор выделяет следующие проблемы:

- частичная наблюдаемость внешней среды (агенты не могут получить доступ к глобальному состоянию и должны принимать решения на основе локальных наблюдений);
- сложность определения индивидуального вклада каждого агента в общий выигрыш.

Для решения проблем не стационарности и частичной наблюдаемости в [35, 31] предлагается использовать «центральный контроллер», который производит сбор информации от каждого агента, а затем определяет

действия, которые они должны выполнить. Также известны методы, основанные на централизованном обучении и децентрализованном выполнении поставленной задачи, обмене опытом между агентами [35].

В работе [15] применяется подход, основанный на градиентных методах изучения стратегии. При этом во время обучения используется «централизованный критик», действия которого зависят от опыта выполнения совместных действий, в то время как стратегия каждого агента зависит только от его собственного опыта. Также используются разностные вознаграждения, в которых каждый агент учится на сформированном вознаграждении, которое сравнивает глобальное вознаграждение с вознаграждением, полученным, когда действие этого агента заменяется действием по умолчанию. При этом «централизованный критик» используется для вычисления «функции преимущества» для конкретного агента, которая сравнивает предполагаемый выигрыш от текущего совместного действия с некоторым базовым уровнем, который исключает действие одного агента, сохраняя при этом фиксированными действия других агентов.

В работе [47–48] для решения обозначенных проблем предлагается использовать метод, основанный на принципах «иерархического обучения с подкреплением» (разновидность – «феодалное обучение с подкреплением» [9]), предполагающий разделение МАС на «группы близкорасположенных агентов» (вне зависимости от степени централизации МАС).

Определим «группу близкорасположенных агентов», как множество агентов, объединенных связями одного из следующих типов [47, 72]:

– Вертикальная связь (R_{vert}) – реализует отношение подчинения, агент-менеджер, расположенный на верхнем уровне иерархии координирует действия нижестоящих агентов (агенты-подчиненные), назначая им задачи, оценивая результаты их выполнения, разрешая спорные ситуации (не обязательно, эта функция может быть также реализована на уровне агентов-подчиненных). При этом для получения роли «агент-менеджер» необходимо

выполнение некоторых условий, например, достижение определенного уровня производительности (следовательно, в том случае, если агент-менеджер перестал удовлетворять этим требованиям, соответствующая роль может быть снята с него). Агенты-подчиненные, имеющие одного агента-менеджера, образуют группу.

– Горизонтальная связь (R_{hor}) устанавливается между равнозначными агентами (в плане иерархического положения, т.е. имеющими одну и ту же роль) с целью реализации обмена полученным опытом, информацией о планируемых действиях (включая разрешение спорных ситуаций). Группа может представлять коалицию агентов, имеющих единую коллективную цель, соединенных набором горизонтальных связей.

Таким образом, группа может быть представлена в виде структуры следующего вида:

$$G = \{A, R\}, \quad (1.7)$$

где A – множество агентов МАС. $A = \{a_i\}$, $i = \overline{1, N_A}$, N_A – число агентов МАС. R – множество связей (вертикальных или горизонтальных), $R = \{R_j\}$, $j = \overline{1, N_R}$, N_R – число связей в группе G , $R_j = \{A_{R_j}, p_k^R\}$, A_{R_j} – множество агентов, участвующих в связи, $A_{R_j} \subset A$, p_k^R – множество параметров связи, в число которых могут входить: тип связи (вертикальная или горизонтальная). Цель заключения связи (координация действий, разрешение конфликтных ситуаций, обмен опытом), $k = \overline{1, N_{R_j}}$, N_{R_j} – число параметров связи R_j . При этом связь R_j между агентами $\langle a_i, a_k \rangle \subset A$, $k = \overline{1, N_A}$, $k \neq i$, может быть заключена в случае выполнения некоторых условий D_{R_j} , которые могут состоять в следующем:

1) Соответствие технических параметров (T) агентов набору критериев $\alpha \in D_{R_j}$.

2) Соответствие переменных параметров (S), характеризующих текущее состояние агентов, набору критериев $\beta \in D_{R_j}$.

3) Соответствие параметров окружающей среды (E) некоторому набору критериев $\gamma \in D_{R_j}$.

Таким образом, для установления связи R_j между агентами $\langle a_i, a_k \rangle$ необходимо выполнение набора условий D_{R_j} , имеющих следующий вид:

$$f(P, \varphi) = \{true|false\}, \quad (1.8)$$

где параметр $P \in \{T, S, E\}$, критерий $\varphi \in \{\alpha, \beta, \gamma\}$, f – функция для определения факта удовлетворения значения параметра P критерию φ .

Таким образом, регулирование процесса установления связей между агентами возможно через корректировку критериев φ .

Очевидно, что набор условий, подобных D_{R_j} , может быть определен и для иных аспектов формирования и функционирования МАС, например, назначение и исполнение заданий, назначение и исполнение ролей, познавательная активность (выражается через потребность агентов в изучении окружающей среды).

При этом под организационной структурой МАС будем понимать множество групп вида (1.7).

Исходя из этого, определим, что организационная структура МАС есть функция от множества правил (1.8), а именно:

$$O = F(D), \quad (1.9)$$

где D – множество правил вида (1.8).

Пусть F_e – функция эффективности МАС, очевидно, что в условиях динамической окружающей среды $F_e(O) \neq const$. Значение $F_e(O)$ может варьироваться в зависимости от флуктуаций параметров окружающей среды. Исходя из этого, может быть сделан вывод о необходимости корректировки (адаптации) организационной структуры O . Это будет в том случае, если изменение параметров окружающей среды привело к снижению $F_e(O)$. Однако изменение параметров окружающей среды может быть настолько существенным, что потребует корректировки (адаптации) множества базовых правил (1.8), которое далее будем называть архитектурой МАС. Различным

аспектам архитектуры МАС посвящены работы [29, 4, 30, 3, 11, 5, 10, 6, 14]. Таким образом, МАС, исследуемые в данной работе, могут быть классифицированы как иерархические централизованные с адаптивной архитектурой и организационной структурой. Вопросам формирования МАС подобного рода посвящены следующие работы [8, 16, 1, 12, 25, 40].

В работе [8] предложен алгоритм, который позволяет агентам динамически формировать коалиции для выполнения задач, учитывая пространственные, временные, ресурсные ограничения. Алгоритм обеспечивает динамическое перераспределение агентов по коалициям в ответ на появление новых задач и изменение условий.

В работе [16] предложены методы и алгоритмы, позволяющие выполнять перераспределение задач от «дефектных» агентов (по определенным причинам не способны в полной мере выполнять свои обязанности, например по причине загрузки вычислительной системы, неполадок аппаратной части) к полнофункциональным. При этом изменение структуры системы выполняется «агентом-мастером организации», перераспределение задач производится «агентом обслуживания». Достоинства: возможность динамической адаптации организационной структуры МАС. Недостатки: необходимость в выделении специальных агентов, выполняющих мониторинг состояния МАС и инициирующих изменения в её структуре; процесс адаптации выполняется исключительно для организационной структуры МАС, изменение архитектуры не предполагается, поддержка архитектуры только одного типа.

В работе [1] авторы предлагают метод, позволяющий агентам самостоятельно формировать коалиции для выполнения взаимозависимых задач, без централизованного координатора. Предложенный подход непосредственно реализует динамическую адаптацию организационной структуры МАС. Агенты не привязаны к жесткой структуре, а перераспределяются в коалиции «на лету» в зависимости от конкретного набора взаимосвязанных задач. Модель ищет такое разбиение агентов на

группы (коалиции), которое максимизирует общую эффективность выполнения всего пакета взаимозависимых задач, а не просто назначает одну задачу одному агенту.

В работе [12] авторы представляют структуру управления энергосистемой как граф, где агенты — это узлы, а каналы связи — ребра. Вводится целевая функция, которая зависит от конфигурации этого графа. Оптимизация проводится с учетом реальных технических ограничений энергосистемы

В работе [25] рассматривается проблема оптимизации организационной структуры МАС в рамках архитектуры типа «коалиция». При этом организационная структура представлена в виде графа. Для решения задачи генерации коалиционной структуры применяется гибридный алгоритм, основанный на «древовидном алгоритме поиска» (CFSS) и «алгоритме динамического программирования» (DuPE). Достоинства: способность к обучению. Недостатки: поддержка архитектуры только одного типа (при этом её параметры не подлежат настройке), организационная структура может быть изменена только на следующей итерации.

В работе [40] предлагается представлять МАС в виде набора иерархических деревьев, при этом агенты, расположенные в корневых узлах, транслируют задачи агентам, расположенным в дочерних узлах и контролируют процесс их выполнения, обмен информацией осуществляется только в вертикальном направлении. Набор иерархических деревьев преобразуется в массив фиксированной длины с целыми числами — хромосома. Далее производится определение оптимального набора деревьев с помощью «генетического алгоритма». В качестве функции приспособленности используется метрика, основанная на времени межуровневого обмена информацией между агентами (в формате «запрос-ответ»). Достоинства: способность к обучению. Недостатки: возможность работы только с архитектурой типа «федерация», отсутствие возможности

настройки её параметров; организационная структура может быть изменена только на следующей итерации.

При этом под обучением будем понимать процесс, описываемый следующим множеством:

$$\langle F_D, F_O, D, O, \Delta F_e \rangle,$$

где F_D – функция для определения (первоначального, т.е. перед началом выполнения задачи; адаптационного, т.е. в процессе решения задачи) архитектуры МАС, $D = F_D = f(A, E, T)$, A – множество агентов, E – параметры окружающей среды, T – поставленная задача, $D: F_e \rightarrow \max$; F_O – функция для определения (первоначального или адаптационного) организационной структуры $O = F_O$ (см. формулу (1.9)), $O: F_e \rightarrow \max$; $\Delta F_e = F_{e_t} - F_{e_{t-1}}$, $\Delta F_e > 0$, F_{e_t} – значение функции эффективности МАС после применения метода М на временном интервале t , $F_{e_{t-1}}$ – на временном интервале $t - 1$, $t \in \{t_i\}$, $t_i = t_{i-1} + \Delta t$, $i = \overline{1, N}$, Δt – приращение показателя t_i , N – номер временного интервала, соответствующего событию «достижение цели поставленной задачи».

В таблице 1.1 приводятся результаты оценки решений [8, 16, 1, 12, 25, 40] по следующим критериям:

- 1) Способность к обучению.
- 2) Возможность динамической адаптации организационной структуры.
- 3) Возможность динамической адаптации архитектуры.
- 4) Число поддерживаемых организационных структур.

Таблица 1.1 – Результаты оценки решений [8, 16, 1, 12, 25, 40]

Работа	Номер показателя			
	1	2	3	4
[54]	+	+	-	1
[9]	-	+	-	1
[72]	+	+	-	1
[30]	-	+	-	1
[3]	+	-	-	1
[11]	+	-	-	1

Общим недостатком рассмотренных работ является выполнение адаптации исключительно организационной структуры, а также малое число поддерживаемых организационных структур.

В следующем подразделе приводится формулировка научной проблемы, цели и задач исследования, приводятся критерии эффективности для оценки разрабатываемых методов и алгоритмов.

1.4 Формулировка научной проблемы и частных задач исследования. Определение критериев эффективности

При подготовке текста п. 1.4 был использован авторский материал из опубликованных научных работ [58, 52, 47].

В качестве областей применения иерархических централизованных МАС с адаптивной архитектурой и организационной структурой, исследуемых в данной работе, рассматриваются:

1) Мониторинг состояния (техническое состояние, факты несанкционированного доступа третьих лиц) некоторого объекта или территории (постройки и территория промышленного предприятия, складское помещение).

2) Управление коллективными действиями ботов в спортивных симуляторах.

3) Формирование оптимальных конфигураций МАС (данная задача подробно рассматривается в [56, 67, 60, 61]). При этом под конфигурацией будем понимать пару $\langle D, O \rangle$ (архитектура и организационная структура).

Далее, с учетом обозначенных областей применения, а также с учетом критериев оценки аналогичных решений, сформулируем перечень основных показателей, используемых для оценки эффективности иерархических централизованных МАС с адаптивной архитектурой и организационной структурой:

1) Число поддерживаемых организационных структур как показатель адаптивных свойств МАС, C_o [19].

2) Время выполнения поставленной задачи как показатель оперативности и ресурсоемкости [90], T .

Поскольку обозначенные показатели эффективности не являются равнозначными, далее выполняется решение задачи многокритериальной оптимизации с помощью метода последовательных уступок:

1) Оптимизация показателя C : $C \rightarrow \max$.

2) Оптимизация показателя T : $T \rightarrow \min$, при ограничении $C \geq C_{\max} - \Delta C$, $\Delta C \in \mathbb{N}$.

Результаты проведённого анализа позволили выявить противоречие между необходимостью эффективного функционирования иерархических централизованных МАС, способных к обучению, динамической адаптации архитектуры и организационной структуры и реализации двух и более типов организационных структур («коалиция», «команда», «иерархия», «общество», «конгрегация», «федерация» [19]), в условиях нестационарной, не полностью наблюдаемой внешней среды и отсутствием методического аппарата, обеспечивающего построение таких МАС. В процессе функционирования МАС возможно возникновение событий, требующих модификации её архитектуры: изменение состава множества агентов, снижение значений функции качества системы вследствие изменения параметров окружающей среды.

Таким образом, имеет место **проблема** автоматической адаптации архитектуры МАС в процессе её функционирования в ответ на отклонение параметров окружающей среды от прогнозируемых значений под влиянием возмущающих воздействий, а также на изменение множества агентов, принимающих участие в решении задачи.

Выполним математическую постановку задачи исследования.

$$M: \langle o, A, T, E, \Omega, Y \rangle \rightarrow \Delta f_{GQ}^T,$$

где M – метод автоматической адаптации архитектуры МАС к текущим условиям окружающей среды; o – архитектура МАС (на её основе, согласно (1.9) формируется организационная структура); A – множество агентов; T – решаемая задача, транслируемая ЛПР; E – окружающая среда; Ω – возмущающие воздействия, оказывающие влияние на состояние окружающей среды и, как следствие, на управляемый процесс, в качестве которого рассматривается реализация глобальной задачи T ; Y – результат решения глобальной задачи T ; $\Delta f_{GQ}^T = f_{GQ}^T_{after} - f_{GQ}^T_{before}$, $f_{GQ}^T_{before}$ – значение глобальной функции качества (эффективности) решения задачи T до применения метода M , $f_{GQ}^T_{after}$ – после применения метода M , $\Delta f_{GQ}^T > 0$.

Целью применения метода M является поиск $o: f_{GQ}^T \rightarrow \max$ в пространстве

$$U = A \times T \times E \times D_o,$$

где D_o – множество правил, описывающих архитектуру o . Очевидно, что размерность множества U (а также временная сложность применяемых алгоритмов) увеличивается экспоненциально в зависимости от $|A|, |T|, |E|, |D_o|$.

Для решения обозначенной научной задачи необходимо решение следующих частных задач:

- 1) Анализ методов автоматической адаптации организационной структуры иерархических МАС.
- 2) Разработка зоосоциального метода адаптации архитектуры и организационной структуры МАС.
- 3) Разработка архитектуры и алгоритмов интеллектуальной системы оптимизации организационной структуры МАС. Экспериментальная оценка эффективности предложенных методов и алгоритмов.

1.5 Выводы

При подготовке текста п. 1.5 был использован авторский материал из опубликованных научных работ [58, 52].

В данной главе выполнен анализ структуры интеллектуального агента, иерархической централизованной МАС, выполнен системный анализ способов организации совместных действий для множества интеллектуальных агентов.

Выявлена проблема автоматической адаптации архитектуры МАС в процессе её функционирования при отклонении параметров окружающей среды от прогнозируемых значений под влиянием возмущающих воздействий, а также при изменении множества агентов, принимающих участие в решении задачи.

Следовательно, может быть поставлена цель — снижение временной сложности процедуры поиска оптимальной архитектуры или организационной структуры МАС в условиях нестационарной, не полностью наблюдаемой внешней среды.

ГЛАВА 2. АНАЛИЗ МЕТОДОВ АВТОМАТИЧЕСКОЙ АДАПТАЦИИ ОРГАНИЗАЦИОННОЙ СТРУКТУРЫ ИЕРАРХИЧЕСКИХ МАС

2.1 Анализ типов организационных структур МАС

При подготовке текста п. 2.1 был использован авторский материал из опубликованных научных работ [58, 52, 57, 47].

Пусть дана многоагентная система (МАС) $\mathbf{S}$, функционирующая в окружающей среде $\mathbf{E}$ (реальный или виртуальный объект – населенный пункт, некоторый инфраструктурный объект, например, здание, шахта, канализационный коллектор). Она имеет следующие характеристики: динамичность, нечетко определенная, не полностью наблюдаемая.

Многоагентная система $\mathbf{S}$ включает множество гетерогенных агентов (мобильных роботов) α_i , $\alpha_i \in A_{\mathbf{S}}$ ($A_{\mathbf{S}}$ – множество агентов, входящих в МАС $\mathbf{S}$), $i = \overline{1, \alpha_N}$, $\alpha_N = |A_{\mathbf{S}}|$, оснащенных подсистемами машинного зрения и связи, при этом для каждого агента α_i определена некоторая окрестность n_{α_i} , которая может соответствовать, например, области покрытия подсистемы связи агента α_i .

Лицо, принимающее решения (ЛПР), транслирует МАС $\mathbf{S}$ некоторую задачу $\mathbf{T}$ [30] (*глобальная решаемая задача*, например, достижение целевого состояния, соответствующего приводимому описанию, формирование карты или мониторинг окружающей среды $\mathbf{E}$), содержащую следующие компоненты:

- 1) Тип задачи $\mathbf{T}$ (например, поиск целевого состояния, формирование карты, мониторинг окружающей среды).
- 2) Характеристики цели ($C_{\mathbf{T}}$), которые для различных типов задач могут заключаться в следующем:

а) Для задачи поиска целевого состояния (задача А) – описание, которому должно соответствовать целевое состояние σ_T . В процессе функционирования каждый агент α_i сопоставляет свое текущее состояние в момент времени t (S_t) с описанием σ_T , если

$$\Delta_{\sigma_T} = |\sigma_T - S_t| < \delta, \quad (2.1)$$

где δ – требуемый уровень точности при идентификации целевого состояния, передается МАС вместе с задачей, $\delta \geq 0$, то считается, что решение глобальной задачи T достигнуто.

Примечание: далее будем применять к функциям типа (2.1) термин *глобальная функция качества* [87] и обозначение $f_{GQ}^T(S_t)$.

б) Для задачи формирования карты окружающей среды (задача В) – описание критерия останова k_T ($k_T \equiv f_{GQ}^T(S_t) \in \delta$, где δ – некоторый интервал), который может заключаться в достижении границ окружающей среды (описание σ_T , по которому будет выполняться идентификация состояния, соответствующего границе окружающей среды). Также могут быть использованы дополнительные критерии (ak_T), например, параметры конкретных областей интереса (т.е., где конкретно нужно искать границу), полнота поиска (вся ли область, гипотетически принадлежащие окружающей среде были обследованы).

в) Для задачи мониторинга состояния окружающей среды (задача С) – достижение заданных значений показателей эффективности в единицу времени τ – θ_τ ($\theta_\tau = f_{GQ}^T(u_\tau^T)$, где u_τ^T – последовательность действий, произведенных агентами к моменту времени τ), например, обследуемая территория, число обнаруженных состояний, соответствующих σ_T .

При этом в качестве примеров показателей, применяемых для оценки эффективности решения задач обозначенных типов, могут быть отмечены:

1) Показатели, характеризующие способность системы к достижению целевых состояний (например, коэффициент выполнения заказов [65], коэффициент завершения миссии [38]).

2) Показатели, характеризующие объем ресурсов (например, времени), потраченный на достижение целевых состояний (например, скорость комплектования заказов в час [67], количество поступивших заказов за определенный период времени [65], время обучения [20]).

Примечание: допустим, что МАС одновременно может выполнять только одну глобальную задачу.

Таким образом, для рассмотренных видов задач имеем $C_T^A = \{\sigma_T, \delta, f_{GQ}^T(S_t)\}$, $C_T^B = \{k_T, ak_T, \sigma_T, \delta, f_{GQ}^T(S_t)\}$, $C_T^C = \{\theta_t^p, c_\theta, f_{GQ}^T(u_t^T)\}$ (c_θ – множество дополнительных параметров, используемых при определении показателей эффективности θ).

Очевидно, что для решения обозначенных задач агенту также необходима функция $f_{S_t}(e_{\langle \alpha_i, t \rangle}^E)$, позволяющая идентифицировать его текущее состояние S_t относительно окружающей среды E в рамках характеристик C_T , где $e_{\langle \alpha_i, t \rangle}^E$ – восприятие агентом α_i окружающей среды E в дискретный момент времени t . Например, при решении задачи составления карты некоторого инфраструктурного объекта результатом f_{S_t} может являться географическое положения агента относительно некоторой точки отсчета.

Введем показатель τ_{lim} – предельное время, за которое должна быть решена глобальная задача T , при достижении которого задача будет считаться не выполненной.

Таким образом, имеем

$$T = \{T, C_T, f_{S_t}, \tau_{lim}\}. \quad (2.2)$$

Агенты обладают минимальным набором априорных знаний об окружающей среде – это описание целевого состояния, а также функция для идентификации текущего состояния агента относительно окружающей среды в рамках поставленной задачи. Закон функционирования окружающей среды f_E им неизвестен.

Для достижения цели глобальной задачи T агентам множества A_S необходимо выполнить последовательность действий $u_T = \{u_t\}$, где u_t – множество действий, выполняемых агентами $\alpha_i \in A_S$ в дискретный момент времени t , $t = t_0, t_0 + \Delta t, \dots, t_N$, t_0 – время начала выполнения глобальной задачи T , t_N – время завершения выполнения глобальной задачи T , $t_N \leq \tau_{lim}$, Δt – приращение параметра t , $u_t = \{u_t^{\alpha_i}\}$, $u_t^{\alpha_i}$ – действие, выполненное агентом α_i в дискретный момент времени t .

Пусть U_T – множество возможных решений глобальной задачи T , $U_T = \{u_{T_k}\}$, $k = \overline{1, |U_T|}$. Очевидно, что решения $u_{T_k} \in U_T$ будут дифференцированы по множеству показателей эффективности θ_T (например, время решения задачи, вероятность достижения цели). Таким образом, целесообразно определить решение, удовлетворяющее следующему условию:

$$u_{T_k}: \theta_T \rightarrow \max, \text{ при } \forall e_t = f_E, \quad (2.3)$$

где e_t – параметры, описывающее состояние окружающей среды E в дискретный момент времени t .

Очевидно, что для достижения условия (2.3) агенты $\alpha_i \in A_S$ должны выполнить обучение – исследование поведения окружающей среды с целью аппроксимации функции f_E , а также определения пар $\langle e_t, u_t \rangle$, удовлетворяющих условию (2.3). При этом вывод о достижении целей обучения может быть сделан на основании значений $\Delta\theta_T = \theta_{Tj} - \theta_{Tj-1}$, где j – номер итерации обучения, $j = \overline{1, j_N}$, j_N – заданное (или выполненное) число эпох обучения. Таким образом, условие достижения цели обучения может быть записано в следующем виде:

$$\theta_T > h_{\theta_T} \vee \forall \langle j, j-1, \dots, j-d \rangle: \Delta\theta_T < \delta \vee j = j_N, \quad (2.4)$$

т.е. необходимо выполнение хотя бы одного из условий: достигнуто некоторое пороговое значение показателей эффективности θ_T ; приращение показателей эффективности $\Delta\theta_T$ меньше заданного порога δ в течение $d + 1$

эпох обучения («выход на плато»); достигнуто заданное число эпох обучения j_N .

Необходимо отметить, что для рассматриваемой ситуации применение обучения с учителем [92, 37, 17] нецелесообразно. Имеющихся данных об окружающей среде (особенно с учетом её характеристик: динамическая, нечетко определенная) недостаточно для генерации релевантного множества образцов $\langle \mathbf{e}_j, u_j \rangle \rightarrow \theta_{Tj}$ ($j = \overline{1, j_N}$, j_N – число образцов в обучающем множестве). С учетом обозначенных обстоятельств целесообразным является применение парадигмы обучения с подкреплением, позволяющей агентам получать опыт в процессе решения поставленной задачи (при этом агентам назначается вознаграждение в зависимости от успешности их действий) [92, 24, 2, 23].

Примечание: для случая применения парадигмы обучения с подкреплением перепишем выражение (2.4) следующим образом:

$$\theta_T > h_{\theta_T} \vee \forall \langle j, j-1, \dots, j-d \rangle: \Delta\theta_T < \delta \vee t = t_N, \quad (2.5)$$

где t – дискретный момент времени, $t = t_0, t_0 + \Delta t, \dots, t_N$, t_0 – время начала обучения, t_N – время завершения обучения, Δt – приращение параметра t .

Далее, после достижения целей обучения (2.5), возможен переход МАС в иной режим работы (основной режим), позволяющий преобразовать полученные в результате обучения знания в решение задачи u_{Tk} , удовлетворяющее условию (2.3).

Таким образом, очевидно, что МАС S должна функционировать, как минимум, в двух режимах: режим обучения, основной режим.

Как уже ранее было обозначено, рассматриваемая окружающая среда E является динамической, следовательно, её параметры $\mathbf{e}_{t_0, t_N} \neq const$, где t_0, t_N – временной интервал, на котором МАС S осуществляла функционирование. При этом допускаются относительно продолжительные периоды времени, на которых параметры окружающей среды могут быть

стабильны. Следовательно, в момент запуска (t_0) МАС S активирует режим обучения. При этом после достижения условия (2.5) в момент времени t_1 переходит в основной режим с целью максимизации показателей эффективности (выражение (2.3)), в котором находится до изменения параметров окружающей среды на Δe (вследствие влияния некоторых возмущающих воздействий ω) в момент времени t_2 : $\Delta e = e_{t_2} - e_{t_1} > h_e$, где h_e - пороговое значение Δe , достижение которого требует повторной активации режима обучения. Таким образом, процесс функционирования МАС S представляет собой последовательность пар «режим обучения->основной режим», при этом переход из режима обучения в основной режим осуществляется при выполнении условия (2.5), а повторная активация режима обучения при следующей ситуации:

$$\Delta e = e_{t_j} - e_{t_{j-1}} > h_e, \quad (2.6)$$

где j – номер дискретного момента времени, на котором осуществлялась смена режима работы МАС S , $j = \overline{1, j_N}$, j_N – общее число изменений режима работы МАС S .

Согласно [98], МАС S включает следующие элементы:

- 1) Множество системных единиц U (активные единицы – агенты A , пассивные единицы – объекты O).
- 2) Окружающая среда E .
- 3) Множество задач T , поручаемых агентам.
- 4) Множество организационных структур o , формируемых агентами.
- 5) Множество действий агентов u_A («различных операций над объектами»).

Таким образом, МАС может быть представлена в следующем виде:

$$S = \langle U = \{A, O\}, E, T, o, u_A \rangle \quad (2.7)$$

Как уже ранее отмечалось, рассматриваемая МАС будет работать в двух режимах: режим обучения, основной режим. Определим элементы МАС, изменяемые в зависимости от режима работы:

1) Очевидно, что элементы U , E , T не имеют зависимости от режима работы МАС и их возможное изменение обусловлено иными причинами (например, действия ЛПР – изменение условий глобальной задачи T , отзыв агента на базу; законы функционирования окружающей среды; воздействие на E внешних факторов и т.п.).

2) Множество действий агентов u_A (например, ехать вперед, назад, направо, налево и т.п.) также останется постоянным (поскольку обусловлено техническими особенностями агентов, однако при этом может быть регламентировано условиями глобальной задачи T , например, на некоторые действия может быть введен запрет).

3) Согласно [19], организационная структура МАС – это «совокупность ролей и связей агентов, компонентов, осуществляющих управление их поведением, а также правил, регламентирующих взаимодействие элементов системы».

В [19] выделяются следующие типы организационных структур МАС:

1) «Общества» («Societies») - мягкая форма объединения отдельных агентов либо групп агентов, имеющих различные цели, структуру, но консолидированных соблюдением некоторых базовых норм и законов, за нарушение которых может быть назначен штраф, влияющий на «репутацию» агента. Структура данного типа может быть представлена в следующем виде:

$$O_{\text{societies}} = \langle A, T_A, R_b, r_b, \varphi_{A_b} \rangle$$

где A – множество агентов, $A \subset U$; T_A – множество целей отдельных агентов множества A , $T_A \subset T$, R_b – множество базовых правил поведения агентов и функционирования системы (например, алгоритмы разрешения конфликтов [18], репликации (обмена) опыта в рамках «трансферного обучения» [92], некоторые базовые правила взаимодействия с окружающей средой – получение отрицательного вознаграждения за столкновение с препятствием), r_b – множество связей между агентами, с ЛПР, реализуемых

в рамках правил R , φ_{A_b} – множество данных о «репутации» отдельных агентов множества A .

2) «Коалиции» («Coalitions») – множество равноправных агентов, координирующих свои действия в рамках достижения общей цели (задачи), максимизирующих при этом индивидуальные выигрыши:

$$O_{\text{coalitions}} = \langle A, T_A, C, R, r, \varphi_A \rangle,$$

где C – коалиции, построенные над множеством A в соответствии с некоторыми правилами, определенными в $R_s \subset R$, $\forall C_i \in O_{\text{coalitions}}: C_i \subset A$, $i = \overline{1, |C|}$ (при этом один агент может являться членом нескольких коалиций); $R = \{R_b, R_s\}$ – базовые и дополнительные (например, правила вступления в коалицию (выхода из коалиции), протокол обмена сообщениями для координации действий агентов в рамках решения поставленной задачи [92]) правила поведения агентов и функционирования системы; $r = \{r_b, r_s\}$ – множество связей между агентами, с ЛПР, реализуемых в рамках правил R_b, R_s ; $\varphi_A = \{\varphi_{A_b}, \varphi_{A_s}\}$ – множество данных о «репутации» отдельных агентов множества A в рамках выполнения правил R_b, R_s .

3) «Конгрегации» – предусматривают создание долговременных групп (в отличие от коалиций), не привязанных к одной конкретной задаче. Агенты стремятся максимизировать индивидуальные выигрыши, а также имеют возможность свободного выхода из «конгрегации», в случае, если получают слишком низкие выигрыши.

4) «Команды» («Teams») – формирование отдельных команд на основе общности намерений агентов, членам команды назначаются роли, ограниченные таким образом, что для выполнения задачи необходима определенная их комбинация (т.е. выполнена декомпозиция общей задачи, возложенной на команду, на отдельные подзадачи, назначенные отдельным агентом - членом команды), действия агентов направлены на максимизацию группового выигрыша:

$$o_{\text{teams}} = \langle A, T_A, K, T_{AK}, R, r, \varphi_A \rangle,$$

где K – множество команд, построенных над множеством A в соответствии с некоторыми правилами, определенными в $R_s \subset R, \forall K_i \in o_{\text{coalitions}}: K_i \subset A, i = \overline{1, |K|}; T_{AK}$ – множество подзадач (ролей), возложенных на отдельных агентов - членов команд K (задачи T_{AK} получены путем декомпозиции задач T_A , назначенных командам в целом, при этом отстающим агентам могут быть назначены новые подзадачи); $R = \{R_b, R_s\}$ – базовые и дополнительные (например, правила формирования команд, назначения групповых вознаграждений) правила поведения агентов и функционирования системы.

5) «Рыночные структуры» («Markets») – «агенты конкурируют друг с другом в достижении поставленной цели, в общем случае выделяют два типа агентов: «агенты-продавцы» (предлагают свои услуги, выставя при этом «цену» за их вербовку), «агенты-вербовщики» (нанимают «агентов-продавцов» для реализации поставленных задач, «выплачивают» агентам-продавцам вознаграждение по итогам выполнения работы). Структура данного типа может быть представлена в следующем виде:

$$o_{\text{markets}} = \langle A, T_A, A_{\text{rec}}, A_{\text{sel}}, R, r, \varphi_A \rangle,$$

где A_{rec} – множество агентов-вербовщиков, $A_{\text{rec}} \subset A$; A_{sel} – множество агентов-продавцов (предлагают свои услуги по выполнению задач T_A агентам A_{rec}), $A_{\text{sel}} \subset A$; $R = \{R_b, R_s\}$ – базовые и дополнительные (например, правила формирования команд, назначения групповых вознаграждений) правила поведения агентов и функционирования системы.

6) «Федерации» («Federations») – «группа агентов, уступивших часть своих прав одному агенту-делегату», выступающему в качестве посредника между ними и, например, ЛПР или агентами-делегатами других групп:

$$o_{\text{federations}} = \langle A, T_A, A_D, A_C, R, r, \varphi_A \rangle,$$

где A_D – множество агентов-делегатов, $A_D \subset A$; A_C – множество агентов (агентов-клиентов), уступивших часть своих полномочий агентам-

делегатам, $A_C \subset A$; $R = \{R_b, R_s\}$ – базовые и дополнительные (например, правила назначения (избрания) агентов-делегатов, их взаимодействия между собой, а также с агентами-клиентами, алгоритмы выполнения обязанностей, которые могут быть вверены агентам-делегатам, например, декомпозиция задач, распределение, мониторинг их выполнения агентами-клиентами) правила поведения агентов и функционирования системы.

7) «Холархии» («Holarchies») – множество агентов подразделяется на ряд неоднородных (имеющих определенный уровень иерархии и автономии в рамках выполнения назначенных им задач), взаимодействующих (как по вертикали, так и по горизонтали) функциональных единиц («холонов» – отдельных агентов или групп):

$$o_{\text{holarchies}} = \langle A, H, D_H, T_H, R, r, I_H, \varphi_A \rangle,$$

где H – множество «холонов», $\forall H_i \in o_{\text{holarchies}}: H_i \subset A, i = \overline{1, |H|}$; D_H – множество обязанностей, которые могут быть возложены на «холоны» (например, назначение подкреплений [28], координация действий агентов [16], репликация опыта в рамках «трансферного обучения» [92]); T_H – множество задач, назначенных «холонам»; I_H – возможные уровни иерархии «холонов».

8) «Матричные организации» («Matrix Organizations») – в отличие от «холархии» на группу агентов-подчиненных могут оказывать влияние несколько агентов-управленцев.

9) «Составные организации» («Compound Organizations») – предполагает разделение агентов на группы по видам реализуемых задач (при этом отдельные группы могут иметь различную организационную структуру).

На рисунке 2.1 рассмотренные виды организационных структур расположены в порядке возрастания их сложности (сверху-вниз).

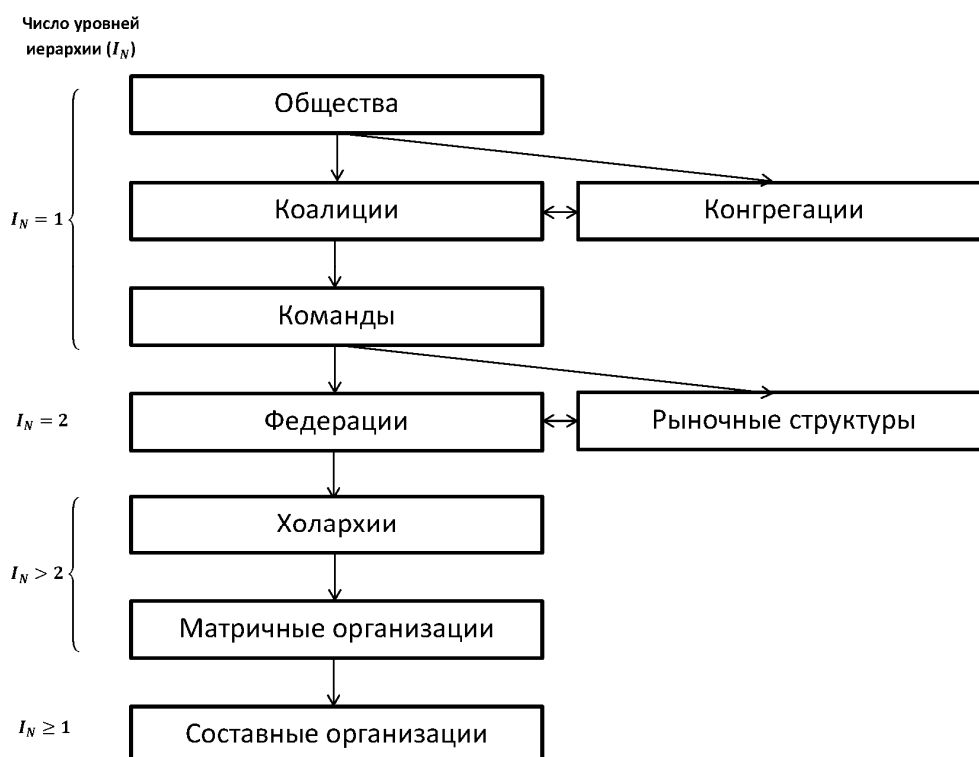


Рисунок 2.1 – Иллюстрация возрастания сложности рассмотренных организационных структур

При этом *иерархической* будем считать систему, имеющую следующие характеристики [87, 97]:

- а) «последовательное вертикальное расположение подсистем, составляющих данную систему (вертикальная декомпозиция)»;
- б) «приоритет действий или право вмешательства подсистем верхнего уровня»;
- в) «зависимость действий подсистем верхнего уровня от фактического исполнения нижними уровнями своих функций».

При этом принцип осуществления вертикального взаимодействия между уровнями иерархии иллюстрируется на рисунке 2.2 [87].

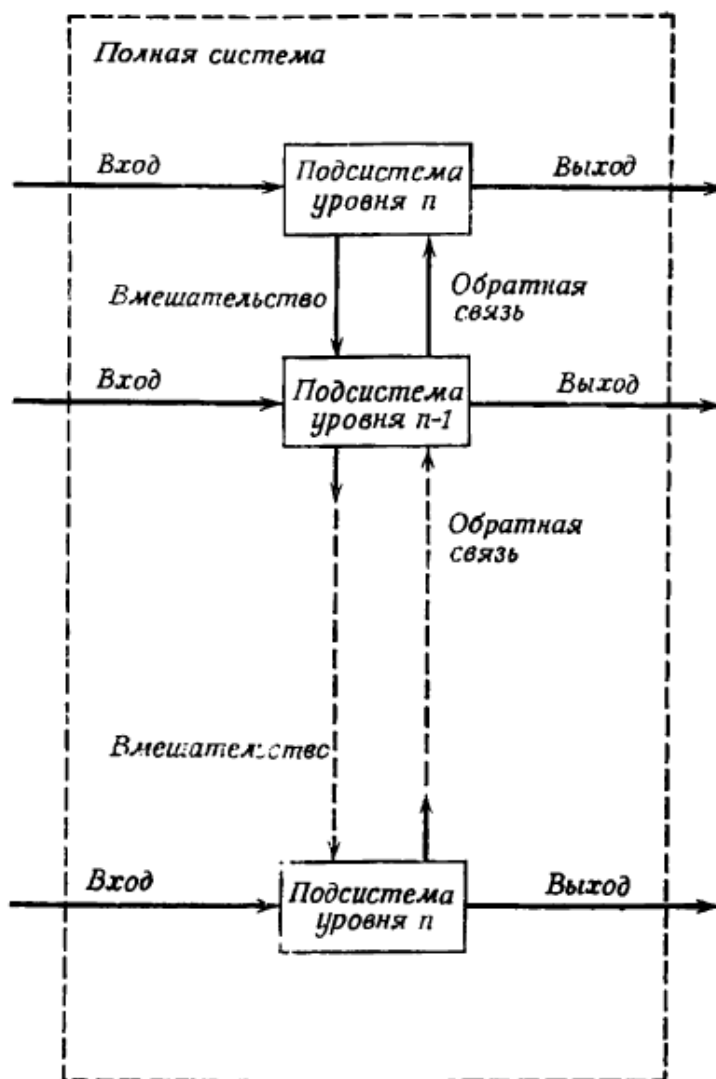


Рисунок 2.2 – Вертикальное взаимодействие между уровнями иерархии

Применительно к рассматриваемой МАС под *вмешательством* (вмешательство до принятия решения агентами нижестоящего уровня) будем понимать назначение / отмену некоторой задачи (предназначенной для выполнения агентами нижестоящего уровня), управляющие воздействия в рамках арбитража конфликтов (межуровневых, когда не совместимы глобальная и локальная цели; внутриуровневых, когда «минимальных локальных затрат одним локальным решающим элементом препятствует другому элементу в достижении минимальных локальных затрат» [87]), вознаграждение (поощрение либо санкции) [11], назначенное по результатам выполнения задачи агентами нижестоящего уровня; под обратной связью –

отчет агентов нижестоящего уровня о результатах выполнения поставленной задачи, запрос о необходимости разрешения конфликта.

В данном случае под иерархией понимается иерархия *эшелонов* [87]. Такое понятие иерархии имеет следующие характеристики [87]:

- а) «система состоит из семейства четко выделенных взаимодействующих подсистем»;
- б) «некоторые из подсистем являются принимающими решения (решающими) элементами»;
- в) «принимающие решения элементы располагаются иерархически в том смысле, что некоторые из них находятся под влиянием или управляются другими решающими элементами» (под принимающими решения элементами будем понимать агентов).

В *многоуровневых многоцелевых системах* (многоэшелонных системах) «элементы верхнего уровня хотя и обуславливают целенаправленную деятельность элементов нижних уровней, но не полностью управляют ею» [87], т.е. у элементов нижних уровней имеется определенная степень свободы в решении поставленных задач.

Соответственно, под системами, представленными на рисунке 2.1 и имеющими число уровней иерархии $I_N = 1$, будем понимать *одноуровневые одноцелевые системы*, под системами, для которых $I_N > 1$ – *многоуровневые многоцелевые системы*.

Ниже представлены критерии, которые были использованы при оценке сложности организационных структур, представленных на рисунке 2.1:

- 1) *Число уровней иерархии (I_N)*. Структуру, в которой агенты неравноправны, будем считать более сложной по следующим причинам:
 - необходимость наличия протоколов общения не только между агентами одного уровня иерархии, но и между агентами нижестоящего и вышестоящего уровней (далее – межуровневый протокол общения);
 - необходимость наличия алгоритма распределения агентов по уровням иерархии (алгоритм распределения агентов по уровням).

Примем также, что с увеличением числа уровней иерархии сложность структуры еще более возрастает, поскольку межуровневые протоколы общения, а также алгоритмы распределения агентов по уровням в зависимости от непосредственно взаимодействующих уровней иерархии могут отличаться. Таким образом, при наличии, например, трехуровневой структуры, число межуровневых протоколов общения может составить два, а алгоритмов распределения агентов по уровням – три.

2) *Число агентов вышестоящего уровня иерархии, которым может подчиняться один агент, стоящий в иерархии на ступень ниже (Sub_N).* Увеличение числа таких агентов вышестоящего уровня иерархии свидетельствует о росте сложности структуры, поскольку влечет усложнение межуровневых протоколов общения, а также алгоритма распределения агентов по уровням.

3) *Необходимость вычисления группового вознаграждения ($Gr = \{true, false\}$)* влечет усложнение структуры, поскольку требует выработки алгоритма, позволяющего оценить индивидуальный вклад каждого агента в общий результат.

4) *Наличие функциональных единиц, отличающихся по типу внутренней организационной структуры ($Hs = \{true, false\}$).* Организационную структуру, удовлетворяющую данной характеристике будем считать более сложной.

Уровень важности перечисленных характеристик иллюстрируется следующим выражением:

$$Hs > I_N > Sub_N > Gr. \quad (2.8)$$

Примечание: организационные структуры типов «коалиции», «конгрегации», а также «федерации», «рыночные структуры» расположены на рисунке 2.1 на одних уровня иерархии, поскольку результаты их оценки критериям сложности организационных структур полностью совпадают. В случае типов «коалиции», «конгрегации» они отличаются только по времени существования объединений. При этом тип «рыночные структуры» может

рассматриваться как разновидность типа «федерации», поскольку также предполагает делегирование части полномочий (а именно, назначение задач, оценка результатов их выполнения) некоторым отдельным агентам (агентам-продавцам).

2.2 Анализ связей агентов в рамках организационных структур МАС

При подготовке текста п. 2.2 был использован авторский материал из опубликованных научных работ [58, 52, 57].

Увеличение сложности организационных структур МАС происходит за счет добавления новых правил поведения агентов и функционирования системы, а также формируемых на их основании связей между отдельными агентами.

Ниже рассматриваются основные группы правил поведения агентов и функционирования системы (множество R) [29, 4, 14]:

1) *Правила взаимодействия с окружающей средой (R_E).*

2) *Правила обучения (R_L)* – в случае применения парадигмы обучения с подкреплением могут быть объединены с правилами предыдущей группы. Например, в рамках парадигмы обучения с подкреплением могут быть реализованы в виде функций следующего вида:

$$r = f(S_{before}, A, S_{after}), \quad (2.9)$$

где S_{after} – состояние, в которое перешел агент из состояния S_{before} в результате выполнения некоторого действия A , r – подкрепление (вознаграждение), полученное агентом за выполнение перехода $S_{before} \rightarrow S_{after}$ [92] (обеспечивают формирование базового опыта [69]). Могут включать также правила репликации опыта [65, 51] между агентами в рамках трансферного обучения [92].

3) *Координирующие правила (R_C)* в аспектах самоорганизации и управления, например, правила декомпозиции поставленных задач, правила распределения задач или ролей между агентами, правила определения вознаграждений за решение поставленных задач (обеспечивают формирование ситуационного опыта [69]), правила вступления или выхода агентов из объединений (коалиций, команд, федераций, холархий и т.д.), правила арбитража конфликтов, правила формирования корректирующих воздействий [87, 14].

Следовательно, имеем

$$R = \{R_E, R_L, R_C\}. \quad (2.10)$$

Примечание: под координацией будем понимать «сферу деятельности или задачу вышестоящей управляющей системы, в ходе которой она пытается добиться, чтобы нижестоящие системы управления функционировали согласованно» [87]. Выделим «*координируемость по отношению к задаче, решаемой вышестоящей управляющей системой*», а также «*координируемость по отношению к решаемой в настоящий момент глобальной задаче*» [87].

Далее множество правил поведения агентов и функционирования системы (множество R) будем называть *множеством законов МАС*.

Рассмотрим общую структуру иерархической МАС, имеющей число уровней $I_N = 2$ (рисунок 2.3) [87].

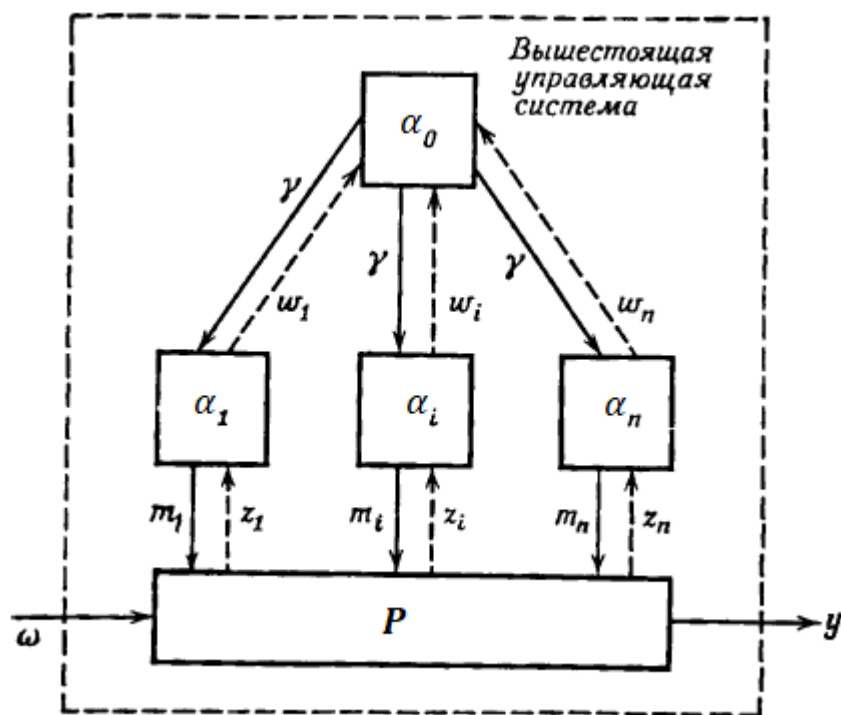


Рисунок 2.3 – Общая структура двухуровневой иерархической MAC

На рисунке 2.3 приняты следующие обозначения [87]:

$\alpha_0, \alpha_1, \alpha_i, \alpha_n$ – агенты MAC S , $\{\alpha_0, \alpha_1, \alpha_i, \alpha_n\} \subset A_S$, $i = \overline{1, n-1}$, $n = |A_S|$, α_0 – агент координатор [76], $\alpha_1, \alpha_i, \alpha_n$ – агенты нижестоящего уровня, находящиеся у него в подчинении (при этом непосредственная возможность воздействия на управляемый процесс имеется только у агентов нижнего уровня – $\alpha_1, \alpha_i, \alpha_n$), $\alpha_i: Z \times \Gamma \rightarrow M$;

γ – координирующие сигналы, поступающие от агента вышестоящего уровня (агента-координатора) α_0 , $\gamma \in \Gamma$;

w_1, w_i, w_n – информационные сигналы обратной связи, поступающие от агентов нижестоящего уровня, $w_i \in W$, $W \rightarrow \Gamma$, $\Gamma \times Z \times M \rightarrow W$; P – управляемый процесс (реализация глобальной задачи T , например, формирование карты некоторого инфраструктурного объекта), $P: M \times \Omega \rightarrow Y$;

m_1, m_i, m_n – управляющие сигналы, направленные от агентов нижестоящего уровня к управляемому процессу P , $m_i \in M$;

z_1, z_i, z_n – информационные сигналы (сигналы обратной связи), поступающие от управляемого процесса, $z_i \in Z$, $M \times \Omega \times Y \rightarrow Z$; ω –

возмущающие воздействия от окружающей среды E , оказывающие влияние на процесс P , $\omega \in \Omega$;

y – выходы процесса P , $y \in Y$; при этом имеют место следующие соотношения:

$$\begin{cases} Z \times \Gamma \rightarrow M \\ \Gamma \times Z \times M \rightarrow W \\ W \rightarrow \Gamma \\ M \times \Omega \times Y \rightarrow Z \\ M \times \Omega \rightarrow Y \end{cases}. \quad (2.11)$$

Пусть $X = \{X_j\}$ – множество решений некоторой глобальной задачи T (см. формулу (2.2)), $X_j = \{\langle \gamma, m_i, z_i, w_i \rangle_k, Y_j\}$, при $p(\Delta = Y_j - C_T \leq \delta, \tau_{lim}) = \{0,1\}$ – некоторый предикат, определяющий факт решения задачи T (при $p=1$ – задача считается решенной, в ином случае – задача не решена), при этом в рамках парадигмы обучения с подкреплением (см. формулу (2.9)) $\alpha_0: A = \Gamma$ (здесь A – это последовательность координирующих сигналов для α_i); $\alpha_i: A_{\alpha_i} = m_i$ (здесь A_{α_i} соответствует управляющему воздействию агента α_i на процесс P), $\alpha_0: S_{after} = f_{S_t}(w_i)$, $\alpha_i: S_{after} = f_{S_t}(z_i)$. Из (2.11) следует, что на множество выходов Y , значения которых представляют результат решений X , оказывают влияние возмущающие воздействия Ω , влияние которых может быть компенсировано путем модификации связей M, Γ, W, Z , при этом из них M, Z представляют связи между агентами нижестоящего уровня и управляемым процессом P , а Γ, W – между агентами двух различных уровней иерархии (далее – *вертикальные связи*), т.е. $r = \{\Gamma, W\}$.

Согласно [87], управляемый процесс P может быть представлен как «управляемая система, к которой поступают управляющие воздействия от систем управления нижнего уровня» (т.е. от агентов α_i), имеющей два типа входных сигналов (возмущающие воздействия Ω (поступают из окружающей среды E), управляющие сигналы M) и один выход Y , следовательно, $M \times \Omega \rightarrow Y$. Исходя из $n - 1 > 1$ (число агентов α_i), представим M как декартово произведение $M = \prod_{i=1}^{n-1} M_i$, где M_i – множество управляющих воздействий

от агентов α_i к управляемому процессу P . Процесс P может быть представлен как множество подпроцессов P_j ($j = \overline{1, P_N}$, P_N – число подпроцессов, на которые декомпозирован процесс P),

$$P_j: M_{P_j} \times Y_{P_j} \rightarrow U_{P_j}, \quad (2.12)$$

где M_{P_j} – множество управляющих воздействий от агентов α_i , поступивших на подпроцесс, $M_{P_j} \subset M$, Y_{P_j} – множество выходов подпроцесса P_j , $Y_{P_j} \subset Y$, U_{P_j} – множество сигналов (*связующие сигналы*), через которые подпроцесс P_j связывается с другими подпроцессами, $U_{P_j} \subset U$, где U – множество связующих сигналов всех $P_j \in P$.

Необходимо отметить, что множество агентов α_i также является управляемым процессом относительно агента α_0 :

$$P = \{\alpha_i\}, \quad (2.13)$$

выход Y соответствует множеству управляющих воздействий M , информационные сигналы Z соответствуют W , множество M соответствует Γ , а динамический тип окружающей среды допускает непосредственное влияние возмущающих воздействий Ω на агентов α_i (соответствуют подпроцессам P_j), таким образом, множество связующих сигналов U будет представлять множество связей между агентами α_i , которые обозначим как *горизонтальные связи*, следовательно, $r = \{\Gamma, W, U\}$.

Таким образом, могут быть выделены следующие типы связей непосредственно между агентами:

1) *Вертикальные связи* (r_v):

а) *Координирующие сигналы*, поступающие от агента вышестоящего уровня (Γ) – предназначены для обеспечения согласованности действий агентов нижестоящего уровня в рамках решения глобальной задачи, включая арбитраж конфликтов (как межуровневых, так и внутриуровневых).

б) *Информационные сигналы обратной связи*, поступающие от агентов нижестоящего уровня (W) – передача агентам вышестоящего уровня отчета о реализации ранее переданных координирующих воздействий.

в) *Межуровневые репликационные сигналы* (R_v) – предназначены для трансляции агентам нижестоящего уровня (в рамках трансферного обучения) множества кортежей опыта следующего вида (см. (2.9)) [69]:

$$\langle S_{before}, A, S_{after} \rangle. \quad (2.14)$$

2) *Горизонтальные связи* (r_h):

а) *Связующие сигналы* (U) – обеспечивают связность действий агентов одного уровня в рамках решения глобальной задачи.

б) *Внутриуровневые репликационные сигналы* (R_h) – предназначены для обмена опытом, формализованным в виде кортежей (2.14), агентами одного уровня (в рамках трансферного обучения).

Следовательно, множество r принимает вид:

$$r = \{r_v = \{G, W, R_v\}, r_h = \{U, R_h\}\}. \quad (2.15)$$

Очевидно, что в зависимости от числа уровней иерархии структура множества r будет различна:

1) $I_N = 1 \rightarrow r = r_h$, т.е. в этом случае в качестве вышестоящего уровня иерархии может рассматриваться непосредственно ЛПР, транслирующее МАС глобальную задачу T .

2) $I_N > 1 \rightarrow r = \{r_v, r_h\}$, т.е. в этом случае на вышестоящих уровнях иерархии будет осуществляться декомпозиция глобальной задачи, транслированной ЛПР, на локальные задачи (подзадачи), которые будут отображаться во множество координирующих сигналов γ :

$$T_{\alpha_0} \rightarrow \Gamma_{\alpha_i}^{T_{\alpha_0}}, \quad (2.16)$$

где T_{α_0} – локальная задача (подзадача), $T_{\alpha_0} \subset T$, назначенная агенту верхнего уровня α_0 ; $\Gamma_{\alpha_i}^{T_{\alpha_0}}$ – множество координирующих сигналов, передаваемых агентам нижнего уровня α_i для реализации локальной задачи

T_{α_0} . Однако агенты α_i также могут являться вышестоящими по отношению к некоторым агентам α_j , $j = \overline{1, \alpha_N^L}$, где $\alpha_N^{L=I_N}$ – число агентов на уровне L , к которому принадлежат агенты α_j . Соответственно, агенты α_i также будут генерировать подзадачи T_{α_i} (согласно некоторого правила декомпозиции задач $R_{decT} \in R_C$) и транслировать их агентам α_j в виде множеств координирующих сигналов $\Gamma_{\alpha_j}^{T_{\alpha_i}}$. Таким образом, для α_i имеем $\Gamma_{\alpha_i}^{T_{\alpha_0}} = \{T_{\alpha_i} | T_{\alpha_i} \subset T_{\alpha_0}\}$. Отметим также, что и характер управляющих сигналов γ также может регламентироваться некоторыми правилами, например $\{R_{Task}, R_{Limit}\} \subset R_C$, где R_{Task} – правила координации путем изменения целей (состоит в изменении *локальных функций качества* [87] задачи T , $f_{LQ}^{T_{\alpha_0}|T_{\alpha_i}}$), R_{Limit} – правила координации путем изменения ограничений (состоит в определении некоторых ограничений на множестве решений) [87]. Очевидно, что реализация обозначенных правил R_C может быть осуществлена путем генерации специфических управляющих сигналов $\gamma \in \Gamma$ (например, отличающихся по формату передаваемых данных), следовательно, имеет место ситуация

$$r: R_C \rightarrow r, \quad (2.17)$$

т.е. реализация конкретных правил координации возможна путем применения отдельных типов связей между агентами. Следовательно, реализация отдельных типов организационных структур, рассмотренных ранее, возможна путем комбинирования связей различных видов между агентами.

Таким образом, организационная структура МАС может быть представлена в виде ориентированного графа

$$G = \langle V, E \rangle, \quad (2.18)$$

где V – множество вершин, $V = \{V_i\}$, $V_i = \{\alpha_i, \varphi_{\alpha_i}, L_j\}$, α_i – некоторый агент, $\alpha_i \in A_S$, $i = \overline{1, |A_S|}$, L_j – уровень иерархии агента α_i , $j = \overline{1, I_N}$, I_N – количество уровней иерархии в МАС S , φ_{α_i} – множество атрибутов агента,

характеризующих его состояние в рамках МАС $\mathcal{S}$ (например, «репутация» [15], «доход» [67] и т.п.), $\varphi_{\alpha_i} = \{\varphi_k\}$, $\varphi_k = \{v, V\}$, v – значение атрибута, $v \in V$, V – множество допустимых значений атрибута, $k = \overline{1, |\varphi_{\alpha_i}|}$; E – множество ребер, $E = \{E_k\}$, $E_k = \{c, \langle b, e \rangle, a\}$, b – агент, от которого исходит связь (начало ребра), e – агент, на которого направлена эта связь (конец ребра), $\langle b, e \rangle \subset A_{\mathcal{S}}$, c – тип связи между агентами $\langle b, e \rangle$ – вертикальная (координирующие сигналы, информационные сигналы обратной связи, межуровневые репликационные сигналы), горизонтальная (связующие сигналы, внутриуровневые репликационные сигналы), a – множество атрибутов связи, $a = \{a_j\}$, $a_j = \{a, v, V\}$, $a = (1|0)$ – индикатор активности атрибута (активен, отключен), v – значение атрибута, $v \in V$, V – множество допустимых значений атрибута, $j = \overline{1, |a|}$, $k = \overline{1, |E|}$.

Различным аспектам формирования организационной структуры МАС (включая адаптацию графа (2.18) к изменениям окружающей среды) посвящены работы [62, 71, 59, 68, 74, 88, 75, 73, 70, 55].

Будем называть далее графы вида (2.18) иерархическими атрибутизированными ориентированными графами (ИАОГ) [80].

Примечание: ребро E_k будет иметь направление только в том случае, если обозначает вертикальную связь.

Таким образом, организационная структура МАС (ОС МАС) может быть представлена в следующем виде:

$$o = \langle A, T, R, G \rangle, \quad (2.19)$$

где A – множество агентов МАС, T – глобальная решаемая задача, включающая глобальную функцию качества (см. (2.2)); R – множество законов МАС; G – ИАОГ (см. (2.18)), описывающий организационную структуру МАС $\mathcal{S}$. При этом процесс развёртывания ОС МАС осуществляется следующим образом:

- 1) ЛПР транслирует МАС некоторую глобальную задачу T .

2) МАС через некоторый вычислительный элемент, например, сервер или отдельный агент осуществляет выбор из множества R набора правил, по которым будет организовано функционирование МАС в целом, а также поведение отдельных агентов при решении задачи T . Таким образом, $T \xrightarrow{f} R$, где f – некоторая функция, которая может осуществлять выбор правил, например, на основании имеющегося опыта решения подобных задач, либо (в случае отсутствия указанного опыта), выбирать правила произвольным образом (при этом должны быть сформированы все три множества, составляющие R , а именно R_E, R_L, R_C , см. формулу (2.10)), т.е. $\langle A, T \rangle \rightarrow R$.

3) Формирование графа G на основании $R_C \in R$ (см. формулу (2.17)), т.е.

$$R \rightarrow G, \quad (2.20)$$

следовательно,

$$\langle A, T \rangle \xrightarrow{M_{deploy}} G, \quad (2.21)$$

где M_{deploy} – метод, позволяющий выполнить первоначальное развертывание организационной структуры МАС на основании множества агентов A и глобальной задачи T .

Представленная выше последовательность действий позволяет сформировать ОС МАС непосредственно перед началом выполнения поставленной задачи. Однако, как уже ранее, отмечалось, МАС функционирует в двух режимах (обучения и основной). Цели функционирования МАС в этих режимах отличаются, в частности:

1) Для режима обучения целью является изучение окружающей среды и адаптация к ней МАС для создания условий последующего перехода в основной режим.

2) Целью основного режима является максимизация производительности МАС.

Исходя из этого, можно предположить, что и организационная структура МАС при функционировании в двух этих обозначенных режимах также будет отличаться (в частности, на это указывается в работе [46]).

Пусть o_{Learn}^T – организационная структура МАС S в режиме обучения при решении задачи T (которая заключается, например, в обследовании некоторого инфраструктурного объекта), при достижении условия (2.5) она трансформируется в o_{Main}^T (организационная структура МАС S в основном режиме). Т.е. в процессе обучения агенты МАС достаточно изучили функцию f_E (описывает поведение окружающей среды S), чтобы в дальнейшем иметь возможность её аппроксимировать с приемлемой точностью, а также сформировали множества вида $\langle e_j, u_j \rangle \rightarrow \theta_{T_j}$, где e_j – значения параметров окружающей среды, u_j – произведенные действия, θ_{T_j} – достигнутые значения показателей эффективности, $j = \overline{1, j_N}$, j_N – число образцов в обучающем множестве. Однако при функционировании в основном режиме также возможно изменение параметров окружающей среды e_j под воздействием возмущающих воздействий Ω (новые значения параметров окружающей среды – $e_j \pm \omega, \omega \in \Omega$), в этом случае возможно возникновение следующих ситуаций:

1) В базе знаний МАС отсутствует множество $\langle e_j \pm \omega, u_j | \omega \in \Omega \rangle \rightarrow \theta_{T_j}$.
Необходим переход МАС в режим обучения, т.е. выполнение следующей трансформации

$$o_{Main}^T \xrightarrow{M_{modify}} o_{Learn}^T, \quad (2.22)$$

где M_{modify} – некоторый метод, позволяющий сформировать модифицирующее воздействие $\Delta o_{Main \rightarrow Learn}^T$ для преобразования организационной структуры o_{Main}^T в o_{Learn}^T .

2) Текущая структура o_{Main}^T может не обеспечить функционирование МАС с максимальной эффективностью, следовательно, необходимо модифицировать организационную структуру o_{Main}^T в $o_{Main}'^T$:

$$\mathbf{e}_j \pm \omega \xrightarrow{M_{modify}} \Delta o_{Main}^T = o_{Main}^T - o_{Main}^{T'} \wedge f_{GQ}^T \rightarrow \max, \quad (2.23)$$

где M_{modify} – некоторый метод, позволяющий отобразить имеющиеся отклонения в параметрах окружающей среды в модифицирующее воздействие Δo_{Main}^T .

Триггером к модификации ОС МАС также могут послужить различные ситуации, возникающие в рамках множества агентов:

1) *Триггер «изменение множества агентов»*. В процессе решения глобальной задачи T возможен выход из строя некоторых агентов ΔA_S ($\Delta A_S = A_{S_{before}} - A_{S_{after}}$, где $A_{S_{before}}$ – изначальный состав множества агентов МАС S , $A_{S_{after}}$ – состав множества агентов МАС S после выхода из строя агентов ΔA_S), что может повлечь перераспределение локальных задач и перестройку графа G (см. формулу (2.18)). Таким образом, имеет место ситуация:

$$\Delta A_S \xrightarrow{M_{modify}} \Delta o_{Main}^T = o_{Main}^T - o_{Main}^{T'} \wedge f_{GQ}^T \rightarrow \max, \quad (2.24)$$

где M_{modify} – некоторый метод, позволяющий отобразить изменения во множестве A_S в модифицирующее воздействие Δo_{Main}^T , приводящее к преобразованию организационной структуры o_{Main}^T в $o_{Main}^{T'}$.

2) *Триггер «снижение значений функции качества системы вследствие изменения параметров окружающей среды»*. В момент времени t под влиянием возмущающего воздействия ω_t величина зафиксированных параметров окружающей среды ($\mathbf{e}_t'$) отклонилась от прогнозируемого значения $\mathbf{e}_t$, т.е. имеем $\mathbf{e}_t' = \mathbf{e}_t \pm \omega_t$, при этом значение глобальной функции качества $f_{GQ_{t+1}}^T \downarrow$,

$$f_{GQ_t}^T - f_{GQ_{t+1}}^T > \Delta f_{GQ}^T, \quad (2.25)$$

где $\Delta f_{GQ}^T < 0$ – параметр чувствительности МАС к изменению величины функции f_{GQ}^T . Таким образом, возникает необходимость в генерации с помощью метода M_{modify} модифицирующего воздействия

Δo_{Main}^T , преобразующего текущую организационную структуру МАС o_{Main}^T в $o_{Main}^{T'}$, для которой условие (2.25) не выполняется:

$$\Delta f_{GQ}^T \xrightarrow{M_{modify}} \Delta o_{Main}^T = o_{Main}^T - o_{Main}^{T'} \wedge f_{GQ}^T \rightarrow \max. \quad (2.26)$$

Примечание: Выше приведено описание триггера «снижение значений функции качества системы вследствие изменения параметров окружающей среды» для глобальной функции качества. Однако данный триггер также может быть реализован и для локальной функции качества.

Таким образом, имеет место *проблема* первоначального развертывания организационной структуры МАС, а также её автоматической модификации (адаптации) в процессе функционирования в ответ на отклонение параметров окружающей среды от прогнозируемых значений под влиянием возмущающих воздействий, а также на изменение множества агентов, принимающих участие в решении задачи.

В итоге имеем:

$$M = \{M_{deploy}, M_{modify}\}: \langle S, o, T, E, \Omega, Y \rangle \rightarrow \Delta f_{GQ}^T, \quad (2.27)$$

где M – методология синтеза организационной структуры МАС S ; M_{deploy} – метод первоначального развертывания организационной структуры МАС S – o (имеет вид (2.19)); M_{modify} – множество методов модификации организационной структуры o (имеет вид (2.19)) в ответ на следующие ситуации: отклонение параметров окружающей среды под влиянием возмущающих воздействий; изменение множества агентов, принимающих участие в решении задачи; T – решаемая задача, транслируемая ЛПР (имеет вид (2.2)); E – окружающая среда; Ω – возмущающие воздействия, оказывающие влияние на состояние окружающей среды и, как следствие, на управляемый процесс, в качестве которого рассматривается реализация глобальной задачи T ; Y – результат решения глобальной задачи T ; $\Delta f_{GQ}^T = f_{GQ}^T_{before} - f_{GQ}^T_{after}$, $f_{GQ}^T_{before}$ – значение глобальной функции качества

решения задачи T до применения методологии M , f_{GQ}^T – после применения методологии M .

2.3 Автоматическая адаптация организационной структуры МАС путем модификации множества правил поведения агентов и функционирования, а также иерархического атрибутированного ориентированного графа МАС. Постановка задачи

При подготовке текста п. 2.3 был использован авторский материал из опубликованных научных работ [58].

Пусть R – множество допустимых правил, регламентирующих функционирование МАС, координацию действий агентов, её составляющих, её взаимодействие с окружающей средой (см. (2.10)) (далее – *множество правил поведения агентов и функционирования МАС*), $R = \{R_i\}$, $i = \overline{1, |R|}$, $|R|$ – мощность множества R ;

R_S – *множество правил поведения агентов и функционирования некоторой конкретной МАС S* , $R_S = \{R_{Sj}\}$, $j = \overline{1, |R_S|}$, $|R_S|$ – мощность множества R_S ;

G_S – *иерархический атрибутированный ориентированный граф МАС* (см. формулу (2.18)), включающий множество вершин V (соответствует множеству агентов МАС, $V = A_S$), множество ребер E (описывает связи между агентами множества A_S).

Согласно (2.20), R_S отображается в G_S , при этом очевидно, что характер связей между агентами (множество E , вид которого определяется в R_S) оказывает влияние на значение глобальной функции качества f_{GQ}^T (например, если агентам запрещено обмениваться данными о намерениях друг друга, то впоследствии могут возникнуть конфликтные ситуации, приводящие в т.ч. к физическому повреждению агентов вследствие столкновения друг с другом).

Также очевидно, что характер связей между отдельными агентами (например, $E_{\langle \alpha_x, \alpha_y \rangle}$ – связь между агентами α_x и α_y ($x \neq y$), $\alpha_x \in A_S$, $\alpha_y \in A_S$, $x = \overline{1, |A_S|}$, $y = \overline{1, |A_S|}$) оказывает влияние как на локальные функции качества $f_{LQ}^{T_{\alpha_x}}$, $f_{LQ}^{T_{\alpha_y}}$ (T_{α_x} , T_{α_y} – локальные задачи, назначенные агентам α_x , α_y), так и на глобальную функцию качества f_{GQ}^T , где $T_{\alpha_x} \subset T$, $T_{\alpha_y} \subset T$.

Таким образом, выделим следующие уровни управления организационной структурой МАС (o):

1) *Макроуровень* – управление организационной структурой o осуществляется путем применения модифицирующих воздействий U_R для множества правил R_S , действующих для всей МАС, что приводит к изменению организационной структуры на Δo . Таким образом,

$$\Delta o = o - o', \quad (2.28)$$

$$\Delta R_S = R_S - R_S', \quad (2.29)$$

$$\Delta R_S \xrightarrow{U_R} \Delta o, \quad (2.30)$$

где R_S' – множество правил поведения агентов и функционирования МАС S после выполнения модифицирующих воздействий U_R , ΔR_S – разность между исходным (R_S) и новым видом (R_S') множества правил поведения агентов и функционирования МАС S , o' – организационная структура МАС S , построенная после применения множества R_S' , Δo – разность между исходной (o) и новой (o') организационной структурой S .

2) *Микроуровень* – через модификацию множество ребер E (в рамках графа G), обозначающего имеющиеся связи между отдельными агентами. В частности, к некоторому множеству ребер $E_{sub} = \{e_k\}$, $E_{sub} \subset E$, $k = \overline{1, |E_{sub}|}$ может быть применено множество модифицирующих воздействий U_E . Для каждого ребра $e_k \in E_{sub}$ применяется модифицирующее воздействие $u_{E_k} \in U_E$, приводящее к изменению его значения на Δe_k , $\Delta e_k = e_k - e_k'$, где e_k' – новое значение ребра e_k . Таким образом, имеем $\Delta E_{sub} = E_{sub} - E_{sub}'$, $\Delta E = E - E'$. Соответственно, $\Delta E \xrightarrow{U_E} \Delta o$, где Δo – см. формулу (2.28). В

рассмотренных аналогичных работах [8, 16, 1, 12, 25, 40, 39] (см. подраздел 1.3) управление выполняется на микроуровне.

Путем применения воздействий U_R или U_E организационная структура МАС S может быть приведена к одному из вариантов, приводимых на рисунке 2.1.

Примечание: сначала выполняется реорганизация на макроуровне, затем происходит дальнейшая настройка организационной структуры на микроуровне.

Далее методы реорганизации структуры МАС на макро и микроуровне рассматриваются более подробно.

2.3.1 Метод, основанный на модификации множества правил поведения агентов и функционирования МАС (реорганизация на макроуровне)

С помощью множества функций $F_R = \{f_{R_i}\}$ выполним отображение правил $R = \{R_i\}$ в некоторые интервалы $\Delta_R = \{\delta_{R_i} | \delta_{R_i} = [\delta_{R_{i_begin}}, \delta_{R_{i_end}}]\}$, где $\delta_{R_{i_begin}}$ – начало интервала δ_{R_i} , $\delta_{R_{i_end}}$ – конец интервала δ_{R_i} , $\delta_{R_{i_begin}} \in \mathbb{R}$, $\delta_{R_{i_end}} \in \mathbb{R}$.

Для каждого правила R_i опишем множество $d_{R_i} = \{d_{R_{i_k}}\}$, где $d_{R_{i_k}}$ – различные варианты правила R_i , $k = \overline{1, |d_{R_i}|}$, например,

R_i – правило установления связей между агентами, имеет следующие варианты реализации:

$d_{R_{i_1}}$ – агенты не могут устанавливать связи друг с другом;

$d_{R_{i_2}}$ – агенты могут получать только координирующие воздействия от агентов, находящихся на более высоком уровне иерархии;

$d_{R_{i_3}}$ – агенты могут обмениваться только информацией о намерениях с целью арбитража *коллизий* (разрешения спорных ситуаций, конфликтные ситуации между агентами, когда они, например, планируют переход в одну и ту же область окружающей среды);

$d_{R_{i_4}}$ – агенты могут обмениваться только полученным опытом;

$d_{R_{i_5}}$ – агенты могут обмениваться информацией о намерениях, а также полученным опытом;

$d_{R_{i_6}}$ – агенты могут обмениваться информацией о намерениях, полученным опытом, а также получать координирующие воздействия от агентов, находящихся на более высоком уровне иерархии.

Примечание: варианты $d_{R_{i_5}}$, $d_{R_{i_6}}$ содержат различные комбинации вариантов $d_{R_{i_2}}-d_{R_{i_4}}$. Варианты вида $d_{R_{i_2}}-d_{R_{i_4}}$ будем называть *простыми*, вида $d_{R_{i_5}}$, $d_{R_{i_6}}$ – *составными*.

Следовательно, имеем множество

$$d_{R_i} = \{d_{R_{i_1}}, d_{R_{i_2}}, d_{R_{i_3}}, d_{R_{i_4}}, d_{R_{i_5}}, d_{R_{i_6}}\}. \quad (2.31)$$

Примечание: последовательность вариантов правила R_i во множестве d_{R_i} должна быть упорядочена по некоторым критериям $C = \{c_j\}$ с помощью множества функций $F_C = \{f_{c_j}\}$, $j = \overline{1, |C|}$, например:

1) Средний объем информации, передаваемой агентами друг другу на некотором интервале времени (данный критерий взят за основу в рассмотренном примере, см. формулу (2.31)).

2) Степень влияния на свободу выбора агента. С учетом данного критерия множество d_{R_i} должно иметь следующий вид:

$$d_{R_i} = \{d_{R_{i_1}}, d_{R_{i_4}}, d_{R_{i_2}}, d_{R_{i_3}}, d_{R_{i_5}}, d_{R_{i_6}}\}. \quad (2.32)$$

Множество d_{R_i} (см. формулу (2.32)) сформировано по принципу убывания свободы агента в результате введения тех или иных правил. В частности, правило $d_{R_{i_4}}$ (возможность обмена опытом) предполагает, что

агент получит знания от других агентов, которые в дальнейшем могут оказать влияние на принимаемые им решения (с точки зрения их потенциальной успешности, на которую указывает опыт других агентов). Координирующие воздействия, поступающие от агента верхнего уровня иерархии ($d_{R_{i_2}}$), могут ограничить действия агента определенными рамками, в которых он будет иметь свободу выбора. Арбитраж коллизий ($d_{R_{i_3}}$) еще более ограничивает свободу действий агентов в том случае, если отдельным правилом (*правило арбитража коллизий*, обозначим его как $R_{d_{R_{i_3}}}$) не регламентируется необходимость поиска компромиссного решения, т.е. уступить должен только один из агентов. Однако из этого также следует необходимость в наличии некоторых ограничений (r) на вид множеств (2.31), (2.32), например:

1) *Корректирующие ограничения* (r_{corr}). Ограничения, корректирующие функцию $f_{c_j}(d_{R_i}) = d_{R_{i_{c_j}}}$, где d_{R_i} – множество вариантов правила R_i , $d_{R_{i_{c_j}}}$ – множество вариантов правила R_i , в результате применения функции f_{c_j} упорядоченное в определенной последовательности. Ограничения данного типа имеют следующий вид:

$$d_{R_{child}} = 1 \rightarrow f_{c_{j_a}} \vee d_{R_{child}} = 0 \rightarrow f_{c_{j_b}}, \quad (2.33)$$

где $d_{R_{child}}$ – некоторый вариант дочернего правила, например $R_{d_{R_{i_3}}}$, $d_{R_{child}}$ – соответствующий ему вариант родительского правила, например $d_{R_{i_3}}$, $\{f_{c_{j_a}}, f_{c_{j_b}}\}$ – виды функции f_{c_j} , применяемые в зависимости от статуса (активен, неактивен) варианта $d_{R_{child}}$.

Примечание: Корректирующие ограничения соответствуют связи типа «зависимость» языка UML [32].

2) *Аннулирующие ограничения* (r_{annul}). Имеют следующий вид:

$$d_{R_{i_k}} = 0 \rightarrow R_{d_{R_{i_k}}} = 0, \quad (2.34)$$

т.е., если вариант $d_{R_{i_k}}$ – неактивен, то связанное с ним правило $R_{d_{R_{i_k}}}$ также аннулируется (например, если выбран вариант $d_{R_{i_2}}$, не предполагающий обмен опытом, то бессмысленно учитывать соответствующее правило $R_{d_{R_{i_4}}}$).

Примечание: Аннулирующие ограничения соответствуют связи типа «композиция» языка UML [32].

Примечание: таким образом, правила $R_i \in R$ могут быть классифицированы по уровню иерархии. Соответственно, родительские правила (например, рассмотренное ранее *правило установления связей между агентами*) обладают более высоким уровнем иерархии, чем дочерние (например, *правило арбитража коллизий*), см. рисунок 2.4.

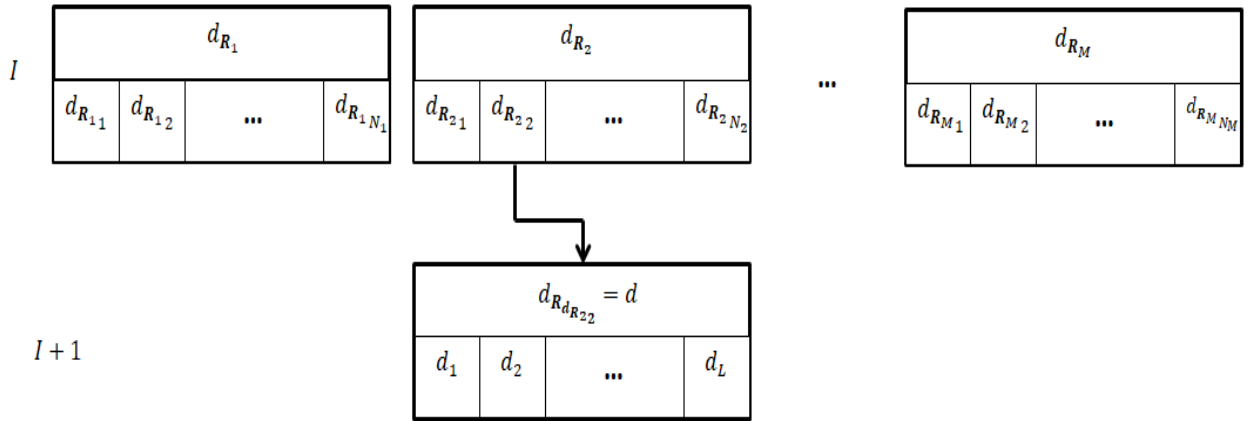


Рисунок 2.4 – Примеры множеств d_{R_i}

На рисунке 4: $N_1 = |d_{R_1}|$, $N_2 = |d_{R_2}|$, N_M – число множеств R_i уровня иерархии I , $L = |d|$.

Отообразим с помощью функции f_{R_i} множества $d_{R_{i_{c_j}}}$ в интервал δ_{R_i} :

$$d_{R_{i_{c_j}}} \xrightarrow{f_{R_i}} \delta_{R_i}, \quad (2.35)$$

На основании (2.35) сформируем пары $\langle d_{R_{i_{c_j k}}}, x_{R_{i_k}} \rangle$, где $x_{R_{i_k}}$ – некоторое числовое выражение варианта $d_{R_{i_{c_j k}}}$ на интервале δ_{R_i} , $x_{R_{i_k}} \in \delta_{R_i}$.

Определим на интервалах δ_{R_i} нечеткие множества [82] $\overline{\delta_{R_i}}$ с характеристическими функциями $\mu_{\overline{\delta_{R_i}}}$:

$$\overline{\delta_{R_i}} = \left\{ \left(x_{R_{i_k}}, \mu_{\overline{\delta_{R_i}}}(x_{R_{i_k}}) \right), x_{R_{i_k}} \in \delta_{R_i} \right\}. \quad (2.36)$$

Характеристическая функция при этом принимает следующий вид:

$$\mu_{\overline{\delta_{R_i}}} = \sum_{k=1}^{|d_{R_{i_{c_j}}}|} \frac{\mu_{\overline{\delta_{R_i}}}(x_{R_{i_k}})}{x_{R_{i_k}}}. \quad (2.37)$$

Таким образом, в случае срабатывания одного из триггеров, описанных в подразделе 1.2, возможны следующие действия:

1) *В основном режиме.* После идентификации текущего состояния МАС (S_t , где t – текущий момент времени, $t = t_0, t_0 + \Delta t, \dots, t_N$, t_0 – время начала работы МАС, t_N – время окончания работы МАС) из базы знаний извлекаются кортежи

$$K = \{K_n = \langle S, X, Q \rangle\}, \quad (2.38)$$

где S – некоторое состояние МАС, $S = S_t$ (либо $S \cong S_t$), $X = \{X_i\}$ – множество нечетких переменных ($X_i \in \overline{\delta_{R_i}}$), описывающих состав множества R_S (множество правил поведения агентов и функционирования некоторой конкретной МАС S), достижение которого путем корректирующего воздействия ΔX ($\Delta X = X_t - X$, X_t – множество нечетких переменных, описывающих состав R_S в момент времени t) позволит перевести состояние МАС в некоторое состояние S' : $f_{GQ}^T(S') > f_{GQ}^T(S)$ ($f_{GQ}^T(S')$, $f_{GQ}^T(S)$ – значения глобальной функции качества при нахождении МАС в состояниях S' и S соответственно), Q – ценность действия [95] «формирование организационной структуры, описываемой X », при нахождении в состоянии S , $n = \overline{1, |K|}$. При этом конкретный кортеж K_n выбирается с некоторой вероятностью $P(K_n)$, определяемой как отображение Q с помощью Softmax-функции [95]. Если $K = \emptyset$, то МАС переходит в режим обучения.

2) *В режиме обучения.* Целью данного режима является формирование кортежей вида (2.38). Однако, в отличие, от основного режима, с

вероятностью ε может быть выбрано неизвестное решение (либо не имеющее максимальное значение Q). Ниже представлена последовательность шагов, выполняемая в режиме обучения:

а) Идентификация текущего состояния S_t .

б) Выбор корректирующего воздействия ΔX , осуществляемый следующим образом:

– с вероятностью $1 - \varepsilon$ выбирается решение (ΔX), для которого $Q \rightarrow \max$;

– с вероятностью ε выбирается случайное решение.

в) Обновление Q (в соответствии с методом SARSA, см. [95]) в кортеже K_n (см. формулу (2.38)), при этом также сохраняется множество $X = X_t - \Delta X$.

Примечание: в состав множества S_t , описывающего состояние МАС, входят следующие элементы:

а) *Постоянные параметры:* глобальная задача T , описание окружающей среды E (необходимы для оценки применимости опыта для решения подобных задач).

б) *Варьируемые параметры.* Фиксируются в момент возникновения события, а также в момент срабатывания триггера. При этом под *событием* понимается возникновение ситуации (изменения параметров окружающей среды, числа агентов), потенциально способной привести к срабатыванию *триггера* (срабатывание триггера происходит в момент фиксации падения глобальной функции качества до некоторого порогового значения). В число варьируемых параметров входят:

– структурные параметры МАС: число агентов различных типов, координаты центра тяжести множества A_S в рамках окружающей среды E (координаты соответствующей точки в пространстве), множество X (множество нечетких переменных, описывающих состав множества R_S (множество правил поведения агентов и функционирования некоторой конкретной МАС S , см. формулу (2.38)));

- значение глобальной функции качества;
- параметры окружающей среды (зависят от типа окружающей среды, в данном случае приводится примерный вариант для окружающей среды типа «лабиринт с динамически изменяющимся положением препятствий»): число объектов различных типов, координаты центра тяжести множества этих объектов в рамках окружающей среды.

Примечание: для формирования корректирующего воздействия ΔX могут применяться следующие сценарии:

1) Изменение значения переменной $X_i \in \overline{\delta_{R_i}}$, $X_i \in X$ (соответствует некоторому $d_{R_{i_k}} \in d_{R_i}$, см. (2.31), (2.32)) таким образом, чтобы новое значение X_i соответствовало $d_{R_{i_{k-1}}}$ или $d_{R_{i_{k+1}}}$. При этом, если $d_{R_{i_k}}$ является родительским вариантом, то вместо $d_{R_{i_k}}$ варьируемым параметром становятся варианты $d_{R_{d_{R_{i_k}m}}}$ соответствующего дочернего правила $R_{d_{R_{i_k}}}$. Варианты родительского правила R_i вновь станут варьируемыми только после перебора всех вариантов дочернего правила $R_{d_{R_{i_k}}}$.

2) Получение X_i , соответствующего случайно сгенерированному значению функции $\mu_{\overline{\delta_{R_i}}}$.

3) Применение генетических алгоритмов [44]. Совокупность множеств d_{R_i} представляется в бинарном виде как хромосома, где «гены» соответствуют вариантам $d_{R_{i_k}}$ («ген» равный 1 указывает, что соответствующий вариант активен).

Предложенный метод, основанный на модификации множества правил поведения агентов и функционирования MAC (реорганизация на макроуровне), иллюстрируется на рисунке 2.5 (в виде диаграммы IDEF0).

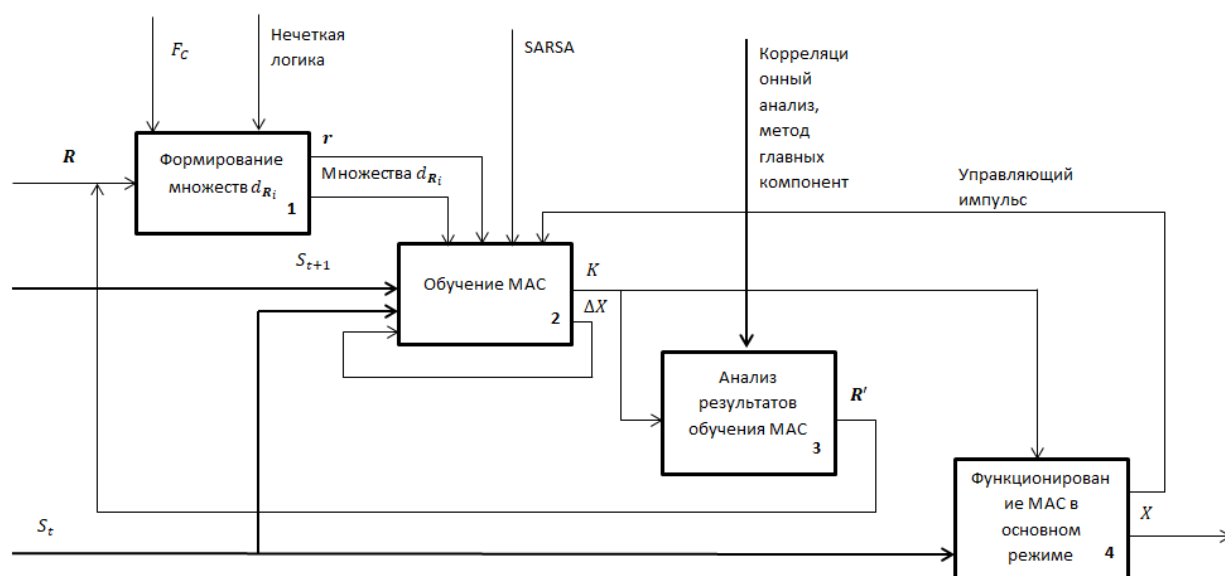


Рисунок 2.5 – Иллюстрация метода, основанного на модификации множества правил поведения агентов и функционирования MAC (реорганизация на макроуровне), в виде диаграммы IDEF0

На рисунке 2.5 имеются следующие компоненты, роль которых нуждается в пояснении:

– Блок «Анализ результатов обучения MAC» (блок 3) - выполняет анализ кортежей K и множества R на предмет выявления правил, не оказывающих влияние (или слабо коррелирующих) на глобальную функцию качества f_{GQ}^T , а также снижение размерности множества R . Применяемые методы: корреляционный анализ [89], метод главных компонент [77].

– R' - новое множество, полученное после оптимизации множества R на блоке 3 (исключены правила, не оказывающие влияния на f_{GQ}^T , состав оптимизирован с помощью метода главных компонент);

– «управляющий импульс» - сигнал, активизирующий блок 2 в случае необходимости перехода из основного режима в режим обучения.

2.3.2 Метод, основанный на модификации иерархического атрибутированного ориентированного графа МАС (реорганизация на микроуровне)

Согласно (2.18), граф G (иерархический атрибутированный ориентированный граф МАС) содержит множество вершин V (соответствует множеству агентов A_S) и множество ребер E (множество связей между агентами). Путем воздействия на граф G может быть реализована более гибкая настройка организационной структуры МАС, поскольку множество E может одновременно содержать связи различных типов (например, вертикальные: координирующие сигналы, информационные сигналы обратной связи, межуровневые репликационные сигналы; горизонтальные: связующие сигналы, внутри уровневые репликационные сигналы, см. подраздел 1.2). Однако, применение подобного подхода также сопряжено с определенными сложностями. В частности, число ребер в графе G может достигать величины $|E| = \frac{|V|*(\beta)}{2}$, где $\beta = |V| - 1$ – степень вершин графа при условии, что $\forall E_j: E_j = \langle V_a, V_b | V_a \neq V_b \rangle$, $j = \overline{1, |E|}$, $a = \overline{1, |V|}$, $b = \overline{1, |V|}$, соответственно, для определения

$$G': f_{GQ}^T \rightarrow \max \quad (2.39)$$

может потребовать перебрать $N_{|E|}^{|r_v \cup r_h|} = \frac{|E|!}{|r_v \cup r_h|(|E| - |r_v \cup r_h|)}$ сочетаний,

где r_v – множество вертикальных связей, r_h – множество горизонтальных связей. При этом, согласно [43], число агентов, которые могут принимать участие в процессе оптимизации коллективных действий не может превышать следующего значения:

$$N_{\max} = \frac{\tau_p}{N_{\text{iter}}(t_{\text{поу}} + t_{\text{пи}})},$$

где τ_p – «время изменения ситуации в системе группа агентов-окружающая среда» [43], N_{iter} – число итераций, $t_{\text{поу}}$ – «время реализации оптимизации действий некоторым агентом группы» [43], $t_{\text{пи}}$ – «время

осуществления процедуры передачи информации остальным агентам группы» [43].

Очевидно, что при большом числе агентов во множестве A_S может иметь место следующая ситуация:

$$N_{|E|}^{|r_v \cup r_h|} \gg N_{max}. \quad (2.40)$$

Таким образом, очевидным решением является разбиение агентов множества A_S на отдельные группы g_k , $k = \overline{1, |g|}$, удовлетворяющие условию (2.40), с помощью алгоритмов кластеризации, приводимых в [43]. Однако подобный подход также обладает некоторыми недостатками, в частности

1) Величина параметра N_{max} (в зависимости от технических возможностей агентов) также может быть достаточно большой.

2) В группах g_k , имеющих большое число агентов, распределенных (не всегда равномерно) по некоторой области окружающей среды, могут быть выделены подгруппы агентов $g_k^{sub}_a$, $a = \overline{1, N_{g_k}}$ (N_{g_k} – число подгрупп, которые могут быть выделены в группе g_k), расположенных близко к некоторым агентам группы g_{k+1} , составляющим подгруппу $g_{k+1}^{sub}_b$, $b = \overline{1, N_{g_{k+1}}}$ ($N_{g_{k+1}}$ – число подгрупп, которые могут быть выделены в группе g_{k+1}). Очевидно, что по причине территориальной близости необходимо наличие связующих сигналов между агентами подгрупп $g_k^{sub}_a$ и $g_{k+1}^{sub}_b$, однако для агентов $g_k \setminus g_k^{sub}_a \cup g_{k+1} \setminus g_{k+1}^{sub}_b$ такой необходимости нет.

Следовательно, в структуре МАС необходимо выделение множеств, объединяющих нескольких агентов (например, по территориальному признаку), однако не тождественных группе. Далее к таким множествам будет применяться термин «микрокластер». Ниже представлены определения понятий «микрокластер» и группа.

Микрокластер – это небольшое множество агентов (как правило, содержащее 3–4-х агентов), сформированное на основании некоторых критериев (например, территориальный признак), в рамках которого может

осуществляться оптимизация непосредственных связей между конкретными агентами, в него входящими. Микрокластеры одной группы могут устанавливать связи (за исключением вертикальных) с микрокластерами другой группы. Агенты не могут покидать микрокластеры. Для иллюстрации см. рисунок 2.6.

Группа – это множество агентов (количество не должно превышать N_{max} или 3–4 микрокластера), сформированное на основании некоторых критериев, включающее несколько микрокластеров, имеющее централизованное управление (т.е. выделяется агент-координатор [76]). Микрокластеры могут выходить из групп, например, используя алгоритм, приводимый в [63], основанный на определении привлекательности агента-координатора, а также удовлетворенности агентов-подчиненных (недовольство и желание покинуть группы растет, если агент-подчиненный территориально отделился от агента-координатора или не получает от него задачи). При этом может выполняться оптимизация связей только между микрокластерами, а не между агентами, их составляющими. Группы подразделяются по уровню иерархии. Группы с низшим уровнем иерархии формируются из микрокластеров, группы с более высоким уровнем иерархии состоят из других групп. Для иллюстрации см. рисунок 2.6.

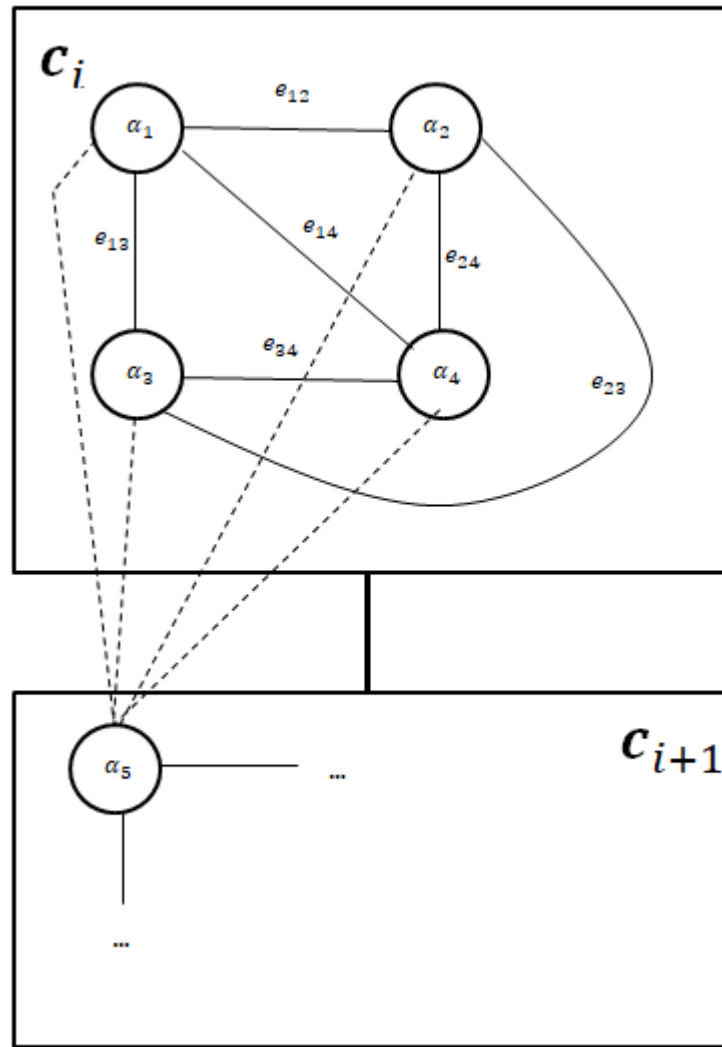


Рисунок 2.6 – Пример микрокластера

Микрокластер c_i , представленный на рисунке 2.6, описывается ИАОГ $G_{c_i} = \langle V_{c_i} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}, E_{c_i} = \{e_{12}, e_{13}, e_{14}, e_{34}, e_{23}, e_{24}\} \rangle$, где V_{c_i} – множество вершин, содержит агентов, составляющих микрокластер c_i , $V \subset A_S$, E_{c_i} – множество ребер (связей между агентами), $i = \overline{1, N_c}$, N_c – число микрокластеров, выделенных на A_S . Соответственно, оптимизации подлежит именно множество ребер E_{c_i} .

Микрокластер c_i связан с микрокластером c_{i+1} . Связь микрокластеров означает, что все агенты, их составляющие, также связаны между собой (тип связи при этом одинаков). На рисунке 2.7 связи между агентами двух микрокластеров показаны пунктирной линией.

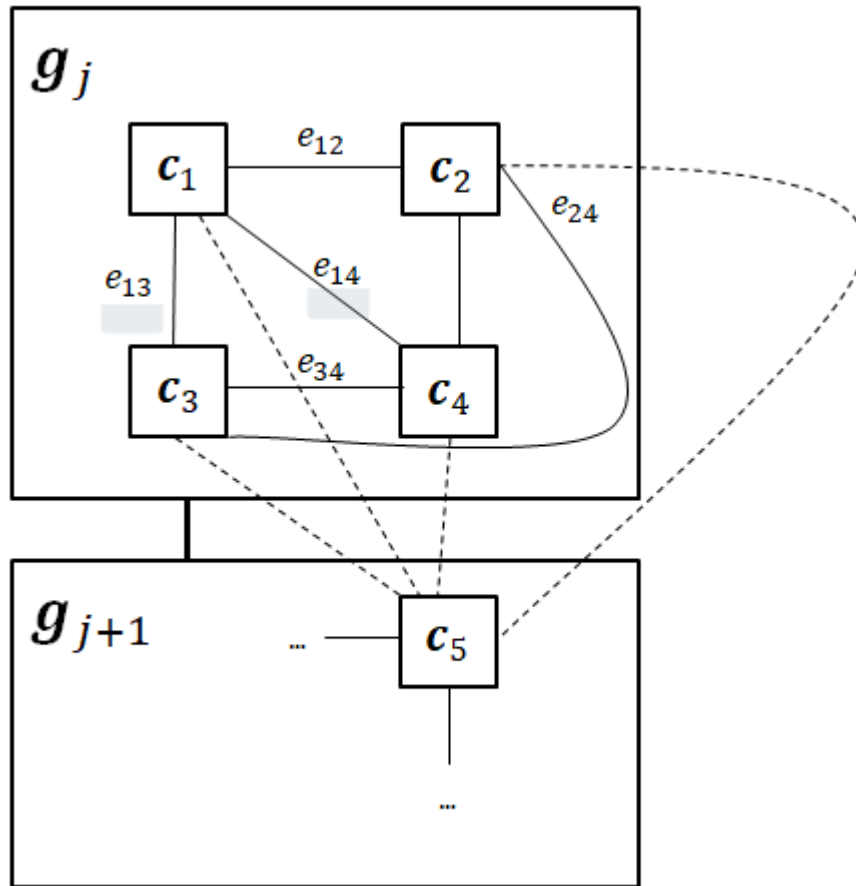


Рисунок 2.7 – Пример группы

На рисунке 2.7 представлена группа g_j , описываемая графом $G_{g_j} = \langle V_{g_j} = \{c_1, c_2, c_3, c_4\}, E_{g_j} = \{e_{12}, e_{13}, e_{14}, e_{34}, e_{23}, e_{24}\} \rangle$. В данном случае оптимизации подлежит именно множество ребер E_{g_j} .

Примечание: на рисунках 2.6–2.7 представлены общие примеры микрокластера и группы, поэтому обозначения направлений для ребер ИАОГ опущены.

Таким образом, процедура оптимизации графа G для достижения условия (2.39) выполняется для отдельных групп (различных уровней иерархии) и микрокластеров, см. рисунок 2.8.

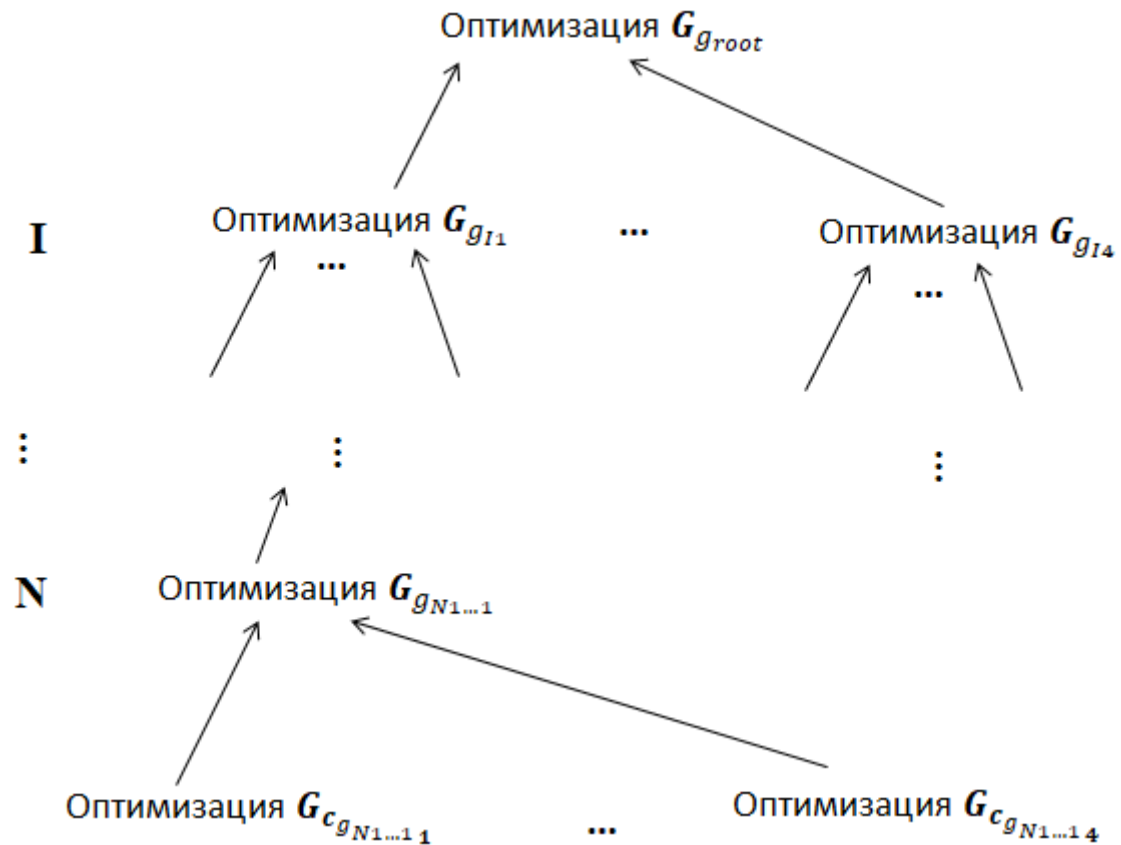


Рисунок 2.8 – Иллюстрация процедуры оптимизации графа G для отдельных групп и микрокластеров

На рисунке 2.8 представлены следующие элементы:

- g_{root} – корневая группа, объединяющая группы первого уровня иерархии;
- $I, \dots, N$ – уровни иерархии;
- $c_{g_{N1...1}1}, \dots, c_{g_{N1...1}4}$ – микрокластеры группы $g_{N1...1}$.

Непосредственно процедура оптимизации может быть реализована на основе генетических алгоритмов [44], при этом хромосома имеет следующий вид (см. рисунок 2.9).



Рисунок 2.9 – Вид применяемой хромосомы

На рисунке 2.9 представлены следующие элементы:

– $e_{12} = e_1, \dots, e_{ML} = e_{K=|E|}$ – ребра ИАОГ соответствующего элемента (микрокластер, группа), M – номер вершины, из которой исходит ребро, L – номер вершины, принимающей ребро;

– $a_{11}, a_{12}, \dots, a_{1N}, \dots, a_{K1}, a_{K2}, \dots, a_{KN}$ – «гены» хромосомы, обозначающие типы связи, разрешенные для данного ребра, при этом первая цифра в индексе обозначает номер соответствующего ребра, вторая цифра – номер типа связи.

В качестве функции приспособленности рассматривается локальная функция качества соответствующего элемента – $f_{LQ}^{T_C}$, где C – индекс элемента (микрокластер или группа), T_C – назначенная ему локальная подзадача.

2.4 Выводы

Предложены подходы к управлению организационной структурой МАС на макро- и микроуровнях:

1) автоматическая адаптация архитектуры и организационной структуры МАС на макроуровне основывается на модификации множества правил поведения агентов и функционирования МАС. Подход характеризуется возможностью воздействия на организационную структуру МАС через множество параметров, являющихся отображением множества допустимых правил, регламентирующих функционирование МАС, координацию действий агентов, её составляющих, взаимодействие с окружающей средой.

2) автоматическая адаптация архитектуры и организационной структуры МАС на микроуровне основывается на модификации иерархического атрибутированного ориентированного графа (ИАОГ) МАС. Подход характеризуется выполнением декомпозиции множества агентов на

микрокластеры, что позволяет минимизировать затраты машинного времени на определение оптимальной структуры ИАОГ.

ГЛАВА 3. ЗООСОЦИАЛЬНЫЙ МЕТОД АВТОМАТИЧЕСКОЙ АДАПТАЦИИ ОРГАНИЗАЦИОННОЙ СТРУКТУРЫ МАС

3.1 «Инстинкты» агентов в системах, основанных на зоосоциальных принципах

При подготовке текста п. 3.1 был использован авторский материал из опубликованных научных работ [52, 57, 47].

В подразделе 2.3 отмечалось, что существует некоторое множество допустимых правил (R), регламентирующих функционирование МАС, координацию действий агентов, её составляющих, её взаимодействие с окружающей средой, отображающееся в иерархический атрибутированный ориентированный граф МАС G_S (см. формулы (2.18), (2.20)), который, наряду с R , является основой организационной структуры МАС (o , см. формулу (2.19)). Далее было определено, что адаптация организационной структуры МАС (как ответ на некоторые внутренние (изменение множества агентов A) или внешние стимулы (изменение параметров окружающей среды E)) возможна на макро и микроуровнях (путем воздействия на множества R , G_S , соответственно). Таким образом, для задания o (и дальнейшей настройки), прежде всего, необходимо определить состав множества R , при этом он должен предоставлять возможность описания следующих критериев, в совокупности позволяющих моделировать все типы организационных структур, описанные в подразделе 2.1;

- связей между агентами – вертикальных и горизонтальных, постоянных и переменных [98];
- способ формирования организационной структуры – «заранее заданная», «спроектированная организация или спонтанно возникающая в результате взаимодействия между агентами» [98];
- «характер распределения функций, ролей, ответственности» [98];
- «вид и геометрия организационных единиц» [98];

- «тип управления (субординация, координация)» [98];
- «структура принятия решений (моноцентризм, полицентризм)» [98];
- «характер коммуникации агентов (непосредственная или опосредованная, ближняя или дистанционная)» [98].

Одним из примеров существующих (и успешно функционирующих) многоагентных систем являются сообщества, образуемые животными, назовем такие МАС основанными на «зоосоциальных принципах». Соответственно, применим для формирования множества R некоторые принципы и правила организации социальных структур, выделяемые в зоологии [78, 83]:

3.1.1 «Витальные инстинкты»

1) «Пищевой инстинкт» [83], выражается в стремлении агентов к получению поощрений («пищи») за успешное выполнение поставленных задач:

$$F_{\alpha_i} \rightarrow \max,$$

$$F_{\alpha_i} = \sum_{t=t_0}^{t_N} r_T k_{hunger}(t), \quad (3.0)$$

где t – параметр времени, $t = t_0, t_0 + \Delta t, \dots, t_N$, t_0 – время начала функционирования МАС, Δt – некоторый временной лаг, t_N – время окончания процесса функционирования МАС, r_T – вознаграждение, полученное агентом α_i за выполнение задачи (или подзадачи) T (таким образом, при определении величины полученной «пищи» используются принципы парадигмы обучения с подкреплением [95]), $k_{hunger}(\tau) = \frac{1}{1+e^{-\beta_{hunger}\tau}}$ – «функция пищеварения», реализующая процесс снижения вклада вознаграждений, полученных ранее, в показатель F_{α_i} , $i = \overline{1, \alpha_N}$, $\alpha_N = |A_S|$, β_{hunger} – «скорость пищеварения» – настраиваемый параметр, с увеличением значения которого график функции становится менее пологим, $\beta_{hunger} > 0$.

Примечание: Агент, длительное время не получающий пищи («отверженный», $F_{\alpha_i} = 0$ на временном интервале продолжительностью t_{hunger}) может быть отозван на базу.

2) «Инстинкт экономии энергии» [83], выражается в стремлении агентов к минимизации энергетических затрат, может быть формализован в виде набора следующих параметров:

а) $P_{rest}(c_t^{bat})$ – «желание отдыха» – вероятность перехода в режим экономии энергии, c_t^{bat} – уровень заряда аккумуляторной батареи в момент времени t , $P_{rest}(c_t^{bat}) = 1 - \frac{1}{1 + e^{\beta_{fatigue} c_t^{bat}}}$, $c_t^{bat} \in [c_{t_{max}}^{bat} \rightarrow (-L), c_{t_{min}}^{bat} \rightarrow L | L \in \mathbb{N}]$, $\beta_{fatigue}$ – «скорость наступления усталости» – настраиваемый параметр, с увеличением значения которого график функции становится менее пологим, $\beta_{fatigue} > 0$. При этом под переходом в режим экономии энергии будем понимать минимизацию энергопотребления за счет снижения производительности аппаратной части агента;

б) P_{min_path} – «показатель трудолюбия» – вероятность выбора для исполнения менее энергозатратного решения. Например, на пути к целевому состоянию агент будет выбирать кратчайший из известных ему путей; агент будет получать большое глобальное подкрепление в случае определения кратчайшего пути к цели.

Примечание: Значение параметра P_{min_path} во многом зависит от решаемой задачи, например, при решении задачи патрулирования территории постоянный выбор кратчайшей траектории может повлечь потерю контроля над частью окружающего пространства, находящегося под надзором.

3) «Защитный инстинкт» [83], может быть выражен через параметр «стремление к риску» (вероятность выбора действий, приводящих к повреждению или выходу из строя агента) (P_{risk}). При этом под действиями,

приводящими к повреждению или выходу из строя агента (далее такие действия будем называть *рискованными*), будем понимать кортежи

$$\langle S_{start}, A, S_{end} \rangle, (3.1)$$

формирующие «табу-список», где A – действие, выполнение которого привело к переходу агента из состояния S_{start} в состояние S_{end} , что повлекло повреждение агента или выход из строя.

Примечание: Очевидно, что «табу-список» может пополняться по мере функционирования МАС, однако при этом необходимо учитывать следующее:

- возможны ситуации, при которых переход в состояние S_{end} может и не приводить к повреждению или выходу агента из строя;
- негативное воздействие на агента при нахождении в состоянии S_{end} может иметь случайный характер.

Соответственно, полный запрет любых рискованных действий (включая действия с небольшой вероятностью потери или повреждения агента) – контрпродуктивен. Таким образом, необходимо определить критерий, при выполнении которого действия могут считаться опасными для агента:

$$P(\text{destroy} | S_{start} \xrightarrow{A} S_{end}) > \sigma,$$

где σ – некоторое пороговое значение, $\sigma \in [0,1]$, $P(\text{destroy} | S_{start} \xrightarrow{A} S_{end})$ – апостериорная вероятность выхода агента из строя (или повреждения) в результате выполнения конкретного рискованного действия - перехода $S_{start} \xrightarrow{A} S_{end}$.

Величина $P(\text{destroy} | S_{start} \xrightarrow{A} S_{end})$ определяется по формуле Байеса:

$$P(\text{destroy} | S_{start} \xrightarrow{A} S_{end}) = \frac{P(S_{start} \xrightarrow{A} S_{end} | \text{destroy}) P(\text{destroy})}{P(S_{start} \xrightarrow{A} S_{end})},$$

где $P(\text{destroy})$ – априорная вероятность выхода агента из строя (или повреждения), $P\left(S_{start} \xrightarrow{A} S_{end}\right)$ – вероятность выполнения перехода $S_{start} \xrightarrow{A} S_{end}$, $P\left(S_{start} \xrightarrow{A} S_{end} | \text{destroy}\right)$ – вероятность перехода $S_{start} \xrightarrow{A} S_{end}$ как причины выхода агента из строя (или повреждения).

Примечание: Очевидно, что такие параметры «желание отдыха», «стремление к риску» определяют физическую целостность и существование агента и их величина должна быть близка к единице. Однако, в зависимости от условий задачи, ценность агента может быть существенно ниже выгоды, получаемой путем достижения конечной цели, на пути к которой частью агентов можно пожертвовать.

3.1.2 «Инстинкты саморазвития»

1) «*Инстинкт новизны*» [83]:

– $P_{research}$ – «показатель любопытства» – вероятность выбора действий, ведущего к переходу в неизвестное состояние (побуждает агента к дальнейшему исследованию окружающей среды);

2) «*Исследовательский инстинкт*» [83]:

– P_{not_best} – «показатель любознательности» – вероятность выбора действия, выполнение которого не ведет, как показывает практика (опыт агента), к оптимальному результату (побуждает агента к верификации опыта, как собственного, так и полученного от других агентов).

3) «*Подражательный инстинкт*» [83]. Рассмотрим это понятие в двух аспектах:

- подражание как слепое повторение действий других агентов;
- подражание как воспроизведение опыта, полученного ранее другими агентами.

Исходя из вышеуказанного определим следующие параметры:

а) P_{imit} – «*желание подражать*» – вероятность повторения агентом действий иных агентов в зависимости от их положения: в зоне видимости, в личной зоне, в социальной зоне, в публичной зоне [78]. Приведенные ранее термины нуждаются в пояснении:

– *Зона видимости* – территория, на которой агент способен идентифицировать иных агентов через средства визуального восприятия (видеокамера).

– *Личная зона* – территория, непосредственно прилегающая к агенту, имеющая радиус $R_{private}$. Для агентов, находящихся в личных зонах друг друга риск возникновения коллизий (конфликтные ситуации между агентами, когда они, например, планируют переход в одну и ту же область внешней среды) считается максимальным. Соответственно, личная зона может быть определена как территория, доступ на которую потенциально опасен для целостности агента.

– *Социальная зона* – территория радиусом R_{social} , на которой должны располагаться агенты для формирования горизонтальных связей, $R_{private} < R_{social}$.

– *Публичная зона* – территория радиусом R_{public} , на которой должны располагаться агенты для формирования вертикальных связей, $R_{private} < R_{public}$.

Разделим параметр P_{imit} на следующие составляющие: P_{imit}^{vis} – «*желание подражать агентам, находящимся в зоне видимости*», $P_{imit}^{private}$ – «*желание подражать агентам, находящимся в личной зоне*», P_{imit}^{social} – «*желание подражать агентам, находящимся в социальной зоне*», P_{imit}^{public} – «*желание подражать агентам, находящимся в публичной зоне*». Таким образом, имеем:

$$P_{imit} = \{P_{imit}^{vis}, P_{imit}^{private}, P_{imit}^{social}, P_{imit}^{public}\}. \quad (3.2)$$

Примечание: Очевидно, что в социальной и публичной может находиться гораздо больше агентов, чем в личной зоне. Эти агенты могут

одновременно выполнять действия различного характера, поэтому возникает противоречие относительно выбора «образца для подражания» (некоторый агент). Таким образом, возникает необходимость в ограничении, согласно которому «желание подражать» может быть активировано только в случае возникновения некоторого события (например, сигнал об опасности), при выполнении агентами лишь ограниченного множества действий, при выборе определенного «образца для подражания», которым может быть агент, находящийся на более высоком уровне иерархии. Следовательно, перепишем (3.2) в следующем виде:

$$P_{imit} = \{P_{imit}^{vis}(D_{event}^{vis}), P_{imit}^{private}(D_{event}^{private}), P_{imit}^{social}(D_{event}^{social}), P_{imit}^{public}(D_{event}^{public})\},$$

где D_{event}^{vis} , $D_{event}^{private}$, D_{event}^{social} , D_{event}^{public} – множества кортежей вида (3.1), содержащих описание действий, разрешенных для подражания, а также возможные «образцы для подражания» (например, агенты-координаторы – «вожаки стаи»).

Примечание: Параметр P_{imit} может быть использован для реализации роевых структур.

б) P_{exper} – «желание делиться своим опытом» – вероятность обмена агентами полученным опытом. Как и в предыдущем случае, имеет место целесообразность в пространственном распределении данного параметра:

$$P_{exper} = \{P_{exper}^{vis}, P_{exper}^{private}, P_{exper}^{social}, P_{exper}^{public}\},$$

где P_{exper}^{vis} , $P_{exper}^{private}$, P_{exper}^{social} , P_{exper}^{public} – вероятности обмена опытом между агентами, расположенными в зоне видимости, либо в личной, социальной, публичной зоне друг друга.

Примечание: Под опытом будем понимать кортежи вида:

$$\langle S_{start}, A, S_{end}, r \rangle, \quad (3.3)$$

где S_{start}, A, S_{end} – см. (3.1), r – вознаграждение, полученное агентом за выполнение перехода $S_{start} \xrightarrow{A} S_{end}$.

Примечание: Параметр P_{exper} позволяет воздействовать на объем сетевого трафика агентов. Например, если окружающая среда уже

исследована, то агентам лучше сосредоточится на решении поставленной задачи (например, патрулирование территории), чем на обмене уже не столь актуальным опытом.

4) «*Инстинкт свободы*» [83]:

– $P_{movement}$ – «*желание двигательной активности*» – вероятность, с которой агенты будут получать штраф к вознаграждению, если их двигательная активность (свобода действий) заблокирована (например, агент уперся в стену).

5) «*Коммуникативный инстинкт*»:

– $P_{horizontal}$ – «*общительность*» – вероятность, с которой агенты будут формировать горизонтальные связи (см. подразделы 2.2, 3.3.2).

3.1.3 «Релевые инстинкты»

1) «*Стадный инстинкт*» [83].

а) P_{subor} – «*желание подчиняться*» – вероятность установления агентом вертикальных связей, в которых он будет являться подчиненной стороной.

При этом под вертикальной связью будем понимать координирующие сигналы. Они направлены на нижний уровень иерархии и предназначены для назначения агентам-подчинённым задач и вознаграждений по результатам их выполнения, а также (опционально) для арбитража коллизий. Информационные сигналы обратной связи направлены на верхний уровень иерархии и содержат отчёт о выполнении поставленной задачи либо сигнал о возникновении конфликтной ситуации — коллизии (см. подраздел 2.2).

б) P_{free} – «*желание свободы*» – вероятность разрыва подчиненным агентом вертикальной связи.

2) «*Территориальный инстинкт*» [83].

Индивидуальная дистанция (её наличие или отсутствие, а также выраженность). Может выражаться в наличии запрета на сближение агентов

(в зависимости от роли и групповой принадлежности) в пространстве и определяться на основе ранее упоминаемых «пространственных радиусов», например, «публичная зона», «социальная зона», «личная зона» [78]. Перечислим параметры, позволяющие формализовать правило территориальной дистанции:

а) R_{terr} – «радиус индивидуальной дистанции»,

$$R_{terr} = \{0 | R_{vis} | R_{private} | R_{social} | R_{public}\}, \quad (3.4)$$

где R_{vis} – радиус зоны видимости, $R_{private}$ – радиус «личной зоны» агента, R_{social} – радиус «социальной зоны» агента, R_{public} – радиус «публичной зоны» агента.

Примечание: При формировании множества R_{terr} необходимо учитывать следующие ситуации:

- при $R_{social} \in R_{terr}$ фактически устанавливается запрет на формирование горизонтальных связей между агентами, следовательно, все параметры, регулирующие коммуникативную активность агентов, в этом случае становятся неактивными;

- при $R_{public} \in R_{terr}$ фактически устанавливается запрет на образование групп, следовательно, все параметры, регулирующие процесс формирования вертикальных связей, также становятся неактивными.

б) $A_{allowed}$ – «*принципы территориальности*» – множество типов агентов, присутствие которых в индивидуальной зоне отдельного агента является терпимым (например, агент-координатор, агенты из одного микрокластера или группы, все агенты).

Определения понятий «микрокластер» и «группа» ранее приводилось в подразделе 2.3.2, однако они нуждаются в уточнении.

Микрокластер – множество агентов, между которыми установлены непосредственные горизонтальные связи (для этого агенты должны располагаться в «социальных зонах» друг друга).

Группа – это множество агентов, установивших вертикальную связь с одним и тем же агентом-координатором (для этого «свободный агент» и агент-координатор должны находиться в публичной зоне друг друга).

Свободный агент – агент, не имеющий агента-координатора.

3.2 Правила «доминирования»

При подготовке текста п. 3.2 был использован авторский материал из опубликованных научных работ [58, 52, 57].

Отражают авторитет агента, на основании которого принимаются решения об установлении вертикальных связей (распределение ролей агент-координатор – «вожак стаи», агент-подчиненный – «добытчик пищи»), о разрешении конфликтных ситуаций (коллизий), порядок распределения задач.

Формализуем понятие «авторитет агента» в виде параметра

$$\varepsilon_{\alpha_i} = \frac{\sum_{t=t_0}^{t_N} F_{\alpha_i t} k_{oblivion}(t)}{t_N}, \quad (3.5)$$

где $F_{\alpha_i t}$ – уровень «пищи» в момент времени t (см. формулу (3.0)), добытой агентом α_i , $F_{\alpha_i t} : k_{hunger} |_{t \in [t_0, t]}$, $k_{oblivion}$ – «коэффициент забвения»,

$$k_{oblivion} = f(t) = \frac{1}{1 + e^{-\beta_{oblivion} t}}, \quad (3.6)$$

где $\beta_{oblivion}$ – «скорость забывания», t – параметр времени, $t = t_0, t_0 + \Delta t, \dots, t_N$, t_0 – время начала функционирования МАС, Δt – некоторый временной лаг, t_N – время окончания процесса функционирования МАС (в процессе функционирования МАС параметр t_N учитывается как t_{now} – текущий момент времени).

Формула (3.5) позволяет в большей степени учитывать прошлые «заслуги» агента, а добавление «коэффициента забвения» позволяет постепенно снижать их вклад в текущий уровень «авторитета».

При этом очевидно, что «авторитет» агента-координатора имеет два аспекта в [67, 63] об этом также идет речь. Однако вместо термина «авторитет» применяется термин «ресурс»:

- *«авторитет перед агентами-подчиненными»;*
- *«авторитет как агента-подчиненного»*, имеющего вертикальную связь с более высоким уровнем иерархии (т.е. агент-координатор, в свою очередь, может являться агентом-подчиненным для другого агента-координатора, далее к таким агентам будет применяться термин *«зависимый агент-координатор»*);
- *«авторитет перед другими агентами-координаторами».*

Очевидно, что при определении для «зависимого агента-координатора» уровня «авторитета как агента-подчиненного» необходимо учитывать, что он также является исполнителем воли (формализованной в виде поставленной задачи) агента верхнего уровня иерархии (*«доминирующий агент»*). Однако в отличие от *«простого агента-подчиненного»* (т.е. не имеющего вертикальных связей, в которых он находится в доминирующем положении) «зависимый агент-координатор» перепоручает поставленные ему задачи своей группе (его агентам-подчиненным). При этом он также выполняет определенную работу, результаты которой должны получить соответствующую оценку от «доминирующего агента»: декомпозиция поставленной задачи на отдельные подзадачи и их распределение между агентами-подчиненными, оценка результатов выполнения подзадач, арбитраж коллизий (опционально). Таким образом, деятельность «зависимого агента-координатора» может быть оценена по формуле (3.5), соответственно, в данном случае будем воспринимать первые два понятия как тождественные, т.е. авторитет агента-координатора перед агентами-подчиненными будет зависеть от того, насколько хорошо он выполняет поставленные ему задачи, что определяется по результатам работы его группы (его агентов-подчиненных).

Очевидно, что уровень «авторитета перед другими агентами-координаторами» не должен зависеть только от качества выполнения поставленных задач (т.е. «пищи», добытой его подчиненными). В живой природе авторитет особи может поддерживаться путем «демонстрации силы» [78, 83]. В данном случае также применим этот принцип, выделив следующие факторы, которые также должны учитываться при определении величины рассматриваемого показателя:

– $|A_{\alpha_{coord}}|$ – общее число агентов подчиненных (позволяет утверждать о способности агента-координатора завербовать и, что самое главное, удержать существенное число подчиненных), где α_{coord} – агент-координатор, $A_{\alpha_{coord}}$ – множество агентов-подчиненных α_{coord} ;

– $|A_{\alpha_{coord}t_{came}}^{came}|$ – число недавно завербованных агентов-подчиненных (если агент-координатор вербует новых агентов, то это свидетельствует о его возрастающей силе), где t_{came} – величина периода времени, за который учитывается число завербованных агентов, $A_{\alpha_{coord}t_{came}}^{came}$ – множество агентов, завербованных α_{coord} ;

– $|A_{\alpha_{coord}t_{gone}}^{gone}|$ – число недавно потерянных агентов-подчиненных (потеря агентов свидетельствует о деградации агента-координатора), где t_{gone} – величина периода времени, за который учитывается число потерянных агентов, $A_{\alpha_{coord}t_{gone}}^{gone}$ – множество агентов, потерянных α_{coord} .

Обозначим параметр «авторитет перед другими агентами-координаторами» как «престиж» агента-координатора, определяемый по формуле:

$$\tau_{\alpha_{coord}} = w_1(\varepsilon_{\alpha_{coord}} \rightarrow [0,1]) + w_2(|A_{\alpha_{coord}}| \rightarrow [0,1]) + w_3(|A_{\alpha_{coord}t_{came}}^{came}| \rightarrow 0,1) + w_4|A_{\alpha_{coord}t_{gone}}^{gone}| \rightarrow 0,1, \quad (3.7)$$

где $\varepsilon_{\alpha_{coord}}$ – «авторитет» α_{coord} перед «доминирующим агентом»; w_1, w_2, w_3, w_4 – весовые коэффициенты соответствующих аргументов, $w_j \in [0,1]$, $j = \{1,2,3,4\}$.

Однако в окрестности агента-координатора может быть существенно меньше агентов, чем он способен завербовать, и в таких условиях возможность «демонстрации силы» с целью увеличения «престижа» ограничена. В связи с этим целесообразно установить зависимость размера территории («публичной зоны») агента-координатора, на которой может осуществляться вербовка, от его «престижа»:

$$R_{public}^{prestige}(\tau_{\alpha_{coord}}) = R_{public} + \frac{R_{public}}{1 + e^{-\beta_{prestige}\tau_{\alpha_{coord}}}}, \quad (3.12)$$

где $R_{public}^{prestige}$ – радиус «публичной зоны», определяемой исходя из уровня «престижа» агента-координатора, R_{public} – базовый радиус «публичной зоны», $\beta_{prestige}$ – «ценность престижа» – показатель, влияющий на вклад престижа в величину R_{public} .

Примечание: Очевидно, что в качестве «образца для подражания» (см. подраздел 3.1.2, параметр P_{imit} – «желание подражать») может быть выбран агент, обладающий наибольшим «авторитетом» (агент-подчиненный) или «престижем» (агент-координатор). Таким образом, параметр P_{imit} («желание подражать») может быть преобразован в экспоненциальное распределение вероятностей, определяемое по формуле:

$$P_{imit}(\theta) = \frac{1}{1 + e^{-\beta_{imit}\theta}}, \quad (3.8)$$

где $\theta = \{\varepsilon_{\alpha_i} | \tau_{\alpha_{coord}}\}$, ε_{α_i} – «авторитет» агента-подчиненного α_i , $\tau_{\alpha_{coord}}$ – «престиж» агента-координатора α_{coord} , β_{imit} – «степень уважения источника» – настраиваемый параметр, с увеличением значения которого возрастает вероятность подражания «авторитетным» или «престижным» агентам, $\beta_{imit} > 0$.

Примечание: Очевидно, что ограничение множества агентов-источников опыта на основании величины их «авторитета» или «престижа»

нецелесообразно, поскольку отрицательный опыт, который будет чаще всего поступать от агентов с небольшим уровнем данных показателей, также является ценным. С другой стороны, подобный опыт ранее мог быть получен от агентов, имеющих в настоящее время высокий уровень «авторитета» или «престижа», соответственно, его дублирование нецелесообразно. Следовательно, также необходим параметр, позволяющий ограничить вероятность восприятия опыта в зависимости от доверия к его источнику:

$$P_{learn}(\theta) = \frac{1}{1+e^{-\beta_{learn}\theta}}, \quad (3.9)$$

где P_{learn} – «желание учиться» – вероятность восприятия агентом полученного (от другого агента) опыта в зависимости от «авторитета» или «престижа» источника, θ – см. описание к формуле (3.8), β_{learn} – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)). «Желание учиться» (P_{learn}) должно быть включено в число параметров, описывающих «инстинкт саморазвития».

Примечание: При описании параметров P_{imit} , P_{learn} в указанном виде (см. формулы (3.8), (3.9)) возникает противоречие, ограничивающее процесс обмена опытом на начальном этапе функционирования МАС, пока «авторитет» или «престиж» агентов недостаточно высок (соответственно, вероятность его восприятия также невелика). Для его устранения введем «правило наивности агента», состоящее в следующем:

- 1) К кортежам опыта, имеющим вид (3.3), добавляется величина «авторитета» (ε_{α_i}) или «престижа» ($\tau_{\alpha_{coord}}$) агента.
- 2) Агенты обмениваются кортежами опыта в соответствии с имеющимся у них «подражательным инстинктом».
- 3) При этом величины $\varepsilon_{max\alpha_j}$, $\tau_{max\alpha_j}$ – максимально возможный «авторитет» или «престиж» с точки зрения агента α_j ($j = \overline{1, \alpha_N}$, $\alpha_N = |A_S|$, A_S – множество агентов МАС), являющегося получателем кортежа опыта – определяются следующим образом:

– агент выполняет ранжирование кортежей опыта, находящихся в его памяти по критерию $\varepsilon_{\alpha_i} \rightarrow \max$ (или $\tau_{\alpha_{coord}} \rightarrow \max$);

– величину «авторитета» ε_{α_i} (или «престижа» $\tau_{\alpha_{coord}}$), удовлетворяющую обозначенным ранее критерием, агент α_j будет воспринимать максимально возможной и, исходя из этого, определять вероятность P_{learn} .

Примечание: Исходя из приведенных выше выкладок, введем параметр «наивность агента» ($p_{naivety}$), принимающий значения «истина» («правило наивности» учитывается при расчете P_{learn}), «ложь» («правило наивности» неактивно).

Далее рассматриваются основные «привилегии», которые может получить агент благодаря уровню авторитета:

– $T_{distribut}$ – «приоритет доступа к пище» – порядок распределения задач: «дискриминация» (задачи в первую очередь получают агенты с высоким авторитетом), «обратная дискриминация» (задачи в первую очередь получают агенты с низким авторитетом).

Примечание: Получение задачи означает приобретение доступа к «пище». Агент, длительное время не получающий «пищи», становится «отверженным», т.е. выводится из МАС. Таким образом, параметр $T_{distribut}$ позволяет регулировать общее число агентов МАС ($|A_S|$) следующим образом: исключение отстающих и сокращение числа агентов; сохранение отстающих на случай существенных изменений в состоянии окружающей среды, адаптация к которым может потребовать дополнительных ресурсов (агентов).

– T_{coll} – «приоритет в арбитраже коллизий» – порядок разрешения конфликтных ситуаций (коллизий) между агентами (одного или разных уровней иерархии, агентов-подчиненных, агентов-координаторов, свободных агентов, т.е. не имеющих вертикальных связей): «право сильного» (*strong*) – коллизии разрешаются в пользу агента с большим авторитетом, «защита

слабых» (*weak*) - коллизии разрешаются в пользу агента с меньшим авторитетом; «*право протекции*» (*protection*) – «право сильного» или «право слабого», при этом учитывается среднее арифметическое «авторитета» агента-подчиненного и его агента-координатора; «*право вожака*» (*leader*) – приоритетом пользуется агент-координатор; «*третейский судья*» (*arbitration*) – решение принимает агент-координатор, в случае если оба агента являются подчиненными и принадлежат к одной группе. При этом также определим параметры «*преобладание права сильного*» (P_{power}) (вероятность разрешения коллизий по «праву сильного»), «*преобладание права вожака*» (P_{leader}) (вероятность разрешения коллизий по «праву вожака»). Таким образом, имеем

$$T_{coll} = \begin{cases} \left\{ \begin{array}{l} protection \rightarrow \begin{cases} P_{power}(1 - P_{leader}) \rightarrow strong \wedge \forall \langle \alpha_i, \alpha_j \rangle: \frac{\varepsilon_{\alpha \rightarrow [0,1]} + \tau_{\alpha_{leader} \rightarrow [0,1]}}{2} \\ (1 - P_{power})(1 - P_{leader}) \rightarrow weak \wedge \forall \langle \alpha_i, \alpha_j \rangle: \frac{\varepsilon_{\alpha \rightarrow [0,1]} + \tau_{\alpha_{leader} \rightarrow [0,1]}}{2} \end{cases} \\ not\ protection \rightarrow \begin{cases} P_{power}(1 - P_{leader}) \rightarrow strong \wedge \forall \langle \alpha_i, \alpha_j \rangle: \varepsilon_{\alpha} \\ (1 - P_{power})(1 - P_{leader}) \rightarrow weak \wedge \forall \langle \alpha_i, \alpha_j \rangle: \varepsilon_{\alpha} \end{cases} \end{array} \right. , \\ \begin{array}{l} leader \rightarrow P_{leader}, \alpha_i \in c_k \wedge \alpha_{leader} \in c_k \\ arbitration, \alpha_i \in c_k \wedge \alpha_j \in c_k \end{array} \end{cases}$$

где ε_{α} – «авторитет» агента-подчиненного, $\tau_{\alpha_{leader}}$ – «престиж» агента-координатора, стоящего над агентом α , $\alpha_i \in A_S$, $\alpha_j \in A_S$ (A_S – множество агентов, входящих в MAC S), $i = \overline{1, \alpha_N}$, $j = \overline{1, \alpha_N}$, $\alpha_N = |A_S|$.

Примечание: С учетом наличия территориального инстинкта необходимо внести уточнение в определение понятия «коллизия», приводимое в главе 2:

Коллизии – конфликтные ситуации между агентами, когда они, например, планируют переход в одну и ту же область окружающей среды, в том случае если существует риск физического столкновения, либо, если согласно «принципу территориальности» агенты данного типа не допускаются в личные зоны друг друга.

Очевидно, что главные «привилегии», которые может приобрести агент благодаря доминированию (обуславливается уровнем «авторитета» или «престижа»), заключаются в возможности установления вертикальных связей

и получения роли агента-координатора (или перехода на более высокий уровень иерархии, когда другие агенты-координаторы могут быть завербованы в качестве агентов-подчиненных), см. подраздел 3.3.1.

3.3 Правила формирования связей

При подготовке текста п. 3.3 был использован авторский материал из опубликованных научных работ [58, 52, 57, 47].

3.3.1 Правила формирования вертикальных связей

Формирование между агентами вертикальной связи делает одного из них (доминирующего агента) агентом-координатором, при этом потенциально возможны следующие процедуры приобретения этой роли («правила становления вожака»):

1) «Авторитет» свободного агента достигает определённого уровня. Такой сценарий наиболее вероятен в начале работы МАС, когда группы еще не были сформированы, а агенты еще не заработали высокий «авторитет».

2) При высоком «авторитете» агента-подчиненного его агент-координатор имеет низкий престиж, соответственно, «сильный» агент-подчиненный может отобрать часть группы у «слабого» агента-координатора. «Слабость» агента-координатора может быть связана, в т.ч., с технической неисправностью. Его постепенная «агония» потенциально может привести к «побегу» агентов-подчиненных в другие группы. В результате может быть потеряна плотность агентов-подчиненных, необходимая в данной области окружающей среды. Избежать такого сценария поможет оперативная передача управления группой другому агенту.

3) Агент-координатор α_i вербует другого агента-координатора α_j , при этом агент α_i находится на более высоком уровне иерархии, чем α_j (или

повышает уровень иерархии в результате вербовки α_j), $i = \overline{1, \alpha_N}$, $j = \overline{1, \alpha_N}$, $\alpha_N = |A_S|$.

Как уже ранее отмечалось, доминирующий агент получает привилегии, одной из которых является возможность стать агентом-координатором. Однако остается непонятным, какие преимущества несет свободным агентам возможность получения роли агента-подчиненного. Т.е. помимо мер принудительного характера, должны быть дополнительные стимулы, побуждающие агентов объединяться в группы. Для разрешения обозначенного противоречия введем следующее правило (*«правило пищи, пригодной к употреблению»*): количество «пищи» будет определяться только на основании вознаграждений, полученных за выполнение задач, назначенных непосредственно агентом-координатором или ЛПР. Т.е. агент в процессе решения поставленной задачи может получать локальные вознаграждения, однако уровень его «насыщения» будет зависеть только от глобального вознаграждения (т.е. полученного за решение этой задачи). При этом уровень сложности задач, назначаемых свободным агентам и агентам-координаторам, будет одинаковым. Из вышеперечисленного могут быть определены следующие особенности функционирования агентов рассматриваемых типов:

1) Свободные агенты:

- свободному агенту назначается некоторая задача T , которую он разбивает на отдельные подзадачи T_k , $k = \overline{1, N}$, N – количество подзадач, на которые была разбита задача T ;

- свободный агент самостоятельно (в течение времени t_T^{free}) выполняет поставленную задачу путем решения составляющих её подзадач T_k (время решения каждой подзадачи - $t_{T_k}^{free}$, $t_T^{free} = \sum_{k=1}^N t_{T_k}^{free}$);

- свободный агент получает «пищу» $F_{\alpha_{free}}$ (вознаграждение) за решение задачи T (но не подзадач T_k).

2) Агенты-подчиненные $\alpha_i \in A_{\alpha_{coord}}$, $i = \overline{1, |A_{\alpha_{coord}}|}$, объединенные в группу под руководством агента-координатора α_{coord} ($A_{\alpha_{coord}}$ – множество агентов-подчиненных, входящих в группу под руководством α_{coord}):

– агент-координатор, получив задачу T , выполняет её декомпозицию на отдельные подзадачи T_k , которые распределяет между агентами-подчиненными α_i ;

– агенты-подчиненные α_i выполняют подзадачи T_k , каждый за время $t_{T_k}^{\alpha_i}$, за что получают «пищу» F_{α_i} от агента-координатора, при этом весьма вероятно, что время выполнения задачи $t_T^{coord} < \sum_{k=1}^N t_{T_k}^{\alpha_i} \wedge t_T^{coord} < t_T^{free}$;

– агент-координатор получает пищу $F_{\alpha_{coord}}$ от его агента-координатора, находящегося на верхнем уровне иерархии (или непосредственно от ЛПР).

Таким образом, агент α_{coord} за меньший промежуток времени получает одинаковый (или больший, если при расчёте вознаграждения учитывать время выполнения задачи) объем «пищи», чем свободный агент. При этом агенты-подчиненные также быстрее получили «пищу», чем в случае, если бы они действовали как свободные агенты.

В начале данного подраздела отмечалось, что агент-координатор может завербовать другого агента-координатора (при сохранении за ним имеющихся вертикальных связей с агентами-подчиненными). Такого агента-координатора будем называть «агент-координатор k -го уровня иерархии», причем значение k определяется следующим образом:

$$k = |v_{to\ subordinates}| + 1, \quad (3.10)$$

где $v_{to\ subordinates}$ – множество вершин ИАОГ (агентов-координаторов нижних уровней иерархии), отделяющих агента α_{coord} на максимальном пути (построенном на ИАОГ) до «простого агента-подчиненного». При этом под «простым агентом-подчиненным» будем понимать агента, не имеющего других агентов в подчинении (т.е. не имеющего вертикальные связи, в которых он занимает доминирующее положение).

Примечание: В соответствии с (3.10) низшим будет считаться 1-й уровень иерархии. Находящиеся на нем агенты-координаторы имеют в подчинение только простых агентов-подчиненных. Соответственно, 2-й уровень иерархии будет считаться более высоким, а расположенные на нем агенты-координаторы будут иметь в подчинении агентов-координаторов 1-го уровня иерархии.

Соответственно, рассмотренный принцип иерархии может быть распространен на получаемый «престиж»: «престиж» агента-координатора 1-го уровня иерархии, «престиж» агента-координатора 2-го уровня иерархии и т.д.

При этом потенциальный агент-координатор k -го уровня должен удовлетворять следующему критерию: «авторитет» (или «престиж») агента не должен опускаться ниже заданного порогового значения на некотором интервале времени:

$$\forall t: \varepsilon_{\alpha_{i_t}} > \varphi \vee \tau_{\alpha_{i_t}} > \varphi_k, \quad (3.11)$$

где $t \in t_{coord}, t_{coord} \in [t_{coord_1}, t_{coord_2} = t_{coord_1} + \Delta t, \dots, t_{coord_{\omega_{coord}}}]$, $|t_{coord}| = \omega_{coord}, t_{coord} \subset [t_0, t + \Delta t, \dots, t_N]$, t_0 – время начала функционирования МАС, t_N – время окончания функционирования МАС, Δt – лаг, с которым выполняется фиксация показателя времени, ω_{coord} – величина временного интервала, на котором производится мониторинг показателя «авторитета» $\varepsilon_{\alpha_{i_t}}$ (или «престижа» $\tau_{\alpha_{i_t}}$), φ – пороговое значение, ниже которого не должен опускаться уровень «авторитета» агента α_i (простой агент-подчиненный или свободный агент), φ_k – пороговое значение, ниже которого не должен опускаться уровень «престижа» агента-координатора k -го уровня иерархии.

Процесс первоначального получения роли агента-координатора и восхождения по уровням иерархии (для 1-го и 2-го уровней иерархии) может быть описан следующим образом:

1) Свободный агент α_i^{free} при выполнении левой части условия (3.11) становится потенциальным агентом-координатором 1-го уровня.

2) В публичной зоне агента α_i^{free} находится другой свободный агент α_j^{free} , с которым может быть установлена вертикальная связь с доминированием агента α_i^{free} , $i = \overline{1, |A_S|}$, $j = \overline{1, |A_S|}$, A_S – множество агентов МАС. Агент α_j^{free} заносится во множество потенциальных агентов-подчиненных α_i^{free} ($A_{\alpha_i^{free}}$).

3) Если $|A_{\alpha_i^{free}}| \geq \vartheta_k$, то все агенты $\alpha_j^{free} \in A_{\alpha_i^{free}}$ становятся агентами-подчиненными α_i^{free} , следовательно, α_i^{free} становится агентом-координатором 1-го уровня иерархии α_i^{coord} , множество $A_{\alpha_i^{free}}$ преобразуется во множество агентов-подчиненных α_i^{coord} ($A_{\alpha_i^{coord}}$), ϑ_1 – «минимальное число агентов в группе» для 1-го уровня иерархии, $\vartheta_k \neq 1, k = 1$.

4) Если для агента α_i^{coord} выполняется правая часть условия (3.11), то он становится потенциальным агентом-координатором 2-го уровня иерархии.

5) Если в публичной зоне агента-координатора 1-го уровня иерархии α_i^{coord} расположен другой агент-координатор 1-го уровня иерархии α_j^{coord} , с которым может быть установлена вертикальная связь, то он заносится во множество $A_{\alpha_i^{coord}}^{k+1}$ (множество потенциальных агентов-подчиненных α_i^{coord} как агента-координатора 2-го уровня иерархии), $k = 1$.

6) Если $|A_{\alpha_i^{coord}}^{k+1}| \geq \vartheta_{k+1}$, то α_i^{coord} становится агентом-координатором 2-го уровня иерархии с множеством агентов-подчинённых $A_{\alpha_i^{coord}}^k$, $k = 2$.

7) Пункты 4–6 выполняются и для последующих уровней иерархии.

Примечание: Один агент может совмещать роли агента-координатора разных уровней, например, «агента-координатора 1-го уровня иерархии» и «агента-координатора 2-го уровня иерархии» (в том случае, если в его

подчинении одновременно находятся простые агенты-подчиненные и агенты-координаторы 1-го уровня иерархии). В этом случае для него одновременно определяется показатель «престижа» 1-го уровня иерархии и «престижа» 2-го уровня иерархии.

Примечание: Исходя из необходимости начисления «престижа» для различных уровней иерархии, аналогичный подход должен быть применен к определению радиуса «публичной зоны» агента-координатора. Перепишем формулу (3.12) в следующем виде:

$$R_{public\ k}^{prestige}(\tau_{\alpha_{coord\ k}}) = R_{public\ k-1\ max}^{prestige} + \frac{R_{public\ k-1\ max}^{prestige}}{1 + e^{-\beta_{prestige\ k} \tau_{\alpha_{coord\ k}}}}, \quad (3.13)$$

где $R_{public\ k}^{prestige}$ – радиус «публичной зоны» агента-координатора k -го уровня иерархии, $R_{public\ k-1\ max}^{prestige}$ – радиус «публичной зоны» данного агента на нижнем уровне (уровень $k - 1$) в момент приобретения им роли агента-координатора k -го уровня иерархии, $\beta_{prestige\ k}$ – «ценность престижа» (см. формулу (3.12)).

Приводимые ранее выкладки раскрывают сценарии реализации первого и третьего «правил становления вожака», приведенные в начале данного подраздела. Однако они не дают ответа на вопрос, каким образом агент-подчиненный может перехватить управление группой (или её частью) у агента-координатора (второе правило). Для этого необходимо определить критерии, позволяющие лишить агента-координатора его статуса:

1) «Критерий низкого престижа»:

$$\forall t: \tau_{\alpha_{i_t}} > \gamma_k, \quad (3.14)$$

где $t \in t_{nc}, t_{nc} \in [t_{nc1}, t_{nc2} = t_{nc1} + \Delta t, \dots, t_{nc\ \omega_{not\ coord}}]$, $|t_{coord}| = \omega_{not\ coord}$, $t_{nc} \subset [t_0, t + \Delta t, \dots, t_N]$, t_0 – время начала функционирования МАС, t_N – время окончания функционирования МАС, Δt – лаг, с которым выполняется фиксации показателя времени, $\omega_{not\ coord}$ – величина временного интервала, на котором производится мониторинга показателя

«престижа» $\tau_{\alpha_{i_t}}$, γ_k – пороговое значение, ниже которого должен держаться уровень «престижа» агента-координатора k -го уровня иерархии в течение времени t , $\gamma_k < \varphi_k$, φ_k – см. формулу (3.11).

2) «Критерий потери агентов»:

$$|A_{\alpha_i^{coord}}^k| < \vartheta_k,$$

где ϑ_k – «минимальное число агентов в группе».

По аналогии с «критерием низкого престижа» определим «критерий слабого агента-координатора» и «критерий сильного агента-подчиненного»:

1) «Критерий слабого агента-координатора»:

$$\forall t: \tau_{\alpha_{i_t}} > \mu_k,$$

где параметр t (см. формулу (3.14)) фиксируется на временном интервале ω_{weak} , μ_k – пороговое значение, ниже которого должен держаться уровень «престижа» агента-координатора k -го уровня иерархии в течение времени t , $\mu_k > \gamma_k$, $\mu_k < \varphi_k$.

2) «Критерий сильного агента-подчиненного» соответствует (3.11), т.е. сильным будет считаться агент-подчиненный, готовый стать агентом-координатором k -го уровня иерархии.

При выполнении перечисленных критериев агент-подчиненный может стать агентом-координатором с вероятностью:

$$P_{sub \rightarrow coord} = P_{dfp} * P_{free},$$

где P_{free} – «желание свободы» (вероятность разрыва подчиненным агентом вертикальной связи, см. подраздел 3.1.3, «стадный инстинкт»), $P_{dfp}(\Delta)$ – «желание перехватить власть» – распределение вероятности перехвата власти сильным агентом-подчиненным:

$$P_{dfp}(\Delta) = \frac{1}{1 + e^{-\beta_{sub \rightarrow coord} \Delta}}, \quad (3.15)$$

где $\Delta = \{\tau_{\alpha_{i_{k-1}}} | \varepsilon_{\alpha_i}\} - \tau_{\alpha_{i_k}}$, $\beta_{sub \rightarrow coord}$ – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)).

Все приводимые выше выкладки рассматривали сценарии формирования вертикальных связей со стороны потенциального агент-координатора, т.е. сверху. Далее этот процесс рассматривается снизу, с точки зрения потенциального агента-подчиненного:

1) *Вступление свободного агента в группу* регламентируется следующими правилами:

а) Потенциальные агент-координатор α_{coord} и агент-подчиненный α_{sub} должны находиться в публичных зонах друг друга, при этом вероятность установления вертикальной связи зависит от расстояния между ними («дальность агента-координатора»):

$$P_{a_dist}(d) = \frac{1}{1+e^{-\beta_{a_dist}d}}, \quad (3.16)$$

где d – расстояние между агентами α_{coord} и α_{sub} , $d = f(\alpha_{coord}, \alpha_{sub})$, $d < R_{public}$, β_{a_dist} – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)).

б) «Привлекательность агента-координатора» [67, 63] зависит от уровня его «престижа»:

$$P_{c_pr}(\tau_{\alpha_{coord}}) = \frac{1}{1+e^{-\beta_{c_pr}\tau_{\alpha_{coord}}}}, \quad (3.17)$$

где $\tau_{\alpha_{coord}}$ – «престиж» агентами α_{coord} , β_{c_pr} – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)).

в) Параметр «желание подчиняться» (P_{subor}), см. подраздел 3.1.3, «стадный инстинкт».

Таким образом, вероятность вступления свободного агента в группу определяется по формуле [84]:

$$P(joining|a_dist, c_prest) = P(joining)P(a_dist|joining)P(c_pr|joining), \quad (3.18)$$

где $P(joining) = P_{subor}$, $P(a_dist|joining) = P_{a_dist}$ (см. (3.16)), $P(c_pr|joining) = P_{c_pr}$ (см. (3.17)).

2) *Выход агента-подчинённого из группы*. Вероятность данного события описывается параметром «недовольство агента-подчиненного» [67, 63]:

$$P_{diss} = \frac{P_{free}P_{a_auth}}{P_{c_pr}P_{a_dist}}, \quad (3.19)$$

где P_{free} – «желание свободы» (см. подраздел 3.1.3, «стадный инстинкт»), P_{a_dist} – «дальность агента-координатора» α_{coord}^{now} (см. формулу (3.16)), P_{c_pr} – «привлекательность агента-координатора» α_{coord}^{now} (см. формулу (3.17)), P_{a_auth} – распределение вероятности выхода агента-подчиненного α_i из группы в зависимости от уровня его собственного «авторитета» ε_{α_i} ,

$$P_{a_auth}(\varepsilon_{\alpha_i}) = \frac{1}{1 + e^{-\beta_{a_auth}\varepsilon_{\alpha_i}}},$$

где β_{a_auth} – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)).

3) *Вступление свободного агента-подчиненного в другую группу.* Для определения вероятности прямого перехода агента-подчиненного α_i из группы агента α_{coord}^{now} в группу агента α_{coord}^{future} (минуя статус свободного агента) также необходимо учитывать «недовольство агента-подчиненного» (см. формулу (3.19)). Таким образом, вероятность перехода агента-подчиненного α_i из группы агента-координатора α_{coord}^{now} в группу агента-координатора α_{coord}^{future} определяется по формуле:

$$P_{\alpha_i: \alpha_{coord}^{now} \rightarrow \alpha_{coord}^{future}} = P(joining|a_dist, c_prest)P_{diss},$$

где $P(joining|a_dist, c_prest)$ – см. формулу (3.18), P_{diss} – см. формулу (3.19).

Примечание: Группа агентов представляет собой коллектив, направленный на решение конкретной задачи, в этой связи территориальная дифференциация агентов-подчиненных должна быть минимальна. Соответственно, введем правило «*территориальной дифференциации*», ограничивающее поле деятельности агентов-подчиненных «*публичной зоной*» из агента-координатора.

Примечание: Дополнительной «привилегией» для «авторитетного» простого агента-подчиненного может быть расширение его поля деятельности за счет увеличения радиуса его «*публичной зоны*» по аналогии с (3.13):

$$R_{public}^{authority}(\varepsilon_{\alpha_i}) = R_{public}^{authority} + \frac{R_{public}^{authority}}{1 + e^{-\beta_{public}\varepsilon_{\alpha_i}}}, \quad (3.20)$$

где $R_{public}^{authority}$ – радиус «публичной зоны» простого агента-подчиненного α_i , β_{public} – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)). Распространим действие данного правила и на свободных агентов.

3.3.2 Правила формирования горизонтальных связей

В подразделе 2.2 выделяют следующие типы *горизонтальных связей* (r_h):

а) *Связующие сигналы* (U) обеспечивают связность действий агентов одного уровня в рамках решения глобальной задачи.

б) *Внутриуровневые репликационные сигналы* (R_h) предназначены для обмена опытом, формализованным в виде кортежей (2.14), агентами одного уровня (в рамках трансферного обучения).

Параметры формирования связей типа R_h описываются в подразделе 3.1.2, «подражательный инстинкт».

При этом под связующими сигналами (U) будем понимать обмен между агентами информацией о намерениях, преследующий следующие цели:

- выбор агентами действий для исполнения с учетом положения и намерений других агентов;
- арбитраж коллизий (см. подраздел 3.2).

Очевидно, что такой обмен наиболее целесообразен между близко расположенными (соседними) агентами. Поскольку они с наибольшей вероятностью способны оказать влияние на деятельность друг другу, а также с целью ограничения объемов используемых вычислительных ресурсов. Следовательно, он может быть ограничен «социальной зоной» агента (R_{social}). «Пространственные зоны» [78] также учитываются при регулировании коммуникационной активности агентов в рамках обмена

опытом. Исходя из вышеуказанного, можно сделать вывод о необходимости расширения «социальной зоны» для успешных агентов, имеющих высокий «авторитет» или «престиж» по аналогии с «публичной зоной» (см. (3.20)):

$$R_{social}^{authority}(\varepsilon_{\alpha_i}) = R_{social}^{authority} + \frac{R_{social}^{authority}}{1+e^{-\beta_{social}\varepsilon_{\alpha_i}}}, \quad (3.21)$$

где $R_{social}^{authority}$ – радиус «социальной зоны» агента-подчиненного, $\beta_{prestige_k}$ – параметр, аналогичный β_{imit} (см. описание к формуле (3.8));

$$R_{social}^{prestige_k}(\tau_{\alpha_i}) = R_{social}^{authority}_{k-1max} + \frac{R_{social}^{authority}_{k-1max}}{1+e^{-\beta_{social_k}\tau_{\alpha_i}}}, \quad (3.21)$$

где $R_{social}^{prestige_k}$ – радиус «социальной зоны» агента-координатора k -го уровня иерархии, $R_{social}^{prestige}_{k-1max}$ – радиус «социальной зоны» данного агента на нижнем уровне (уровень $k-1$) в момент приобретения им роли агента-координатора k -го уровня иерархии, β_{social_k} – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)).

Вероятность формирования горизонтальных связей определяется по формуле:

$$P_{h_r} = P_{horizontal}P_{h_{dist}},$$

где $P_{horizontal}$ – вероятность формирования агентами горизонтальных связей («общительность», см. подраздел 3.1.2, «коммуникативный инстинкт»), $P_{h_{dist}}$ – распределение вероятностей формирования горизонтальных связей в зависимости от расстояния между агентами:

$$P_{h_{dist}}(d) = \frac{1}{1+e^{-\beta_{h_{dist}}d}}, \quad (3.16)$$

где d – расстояние между агентами α_i и α_j , $d = f(\alpha_i, \alpha_j)$, $d < R_{social_{\alpha_i}} \wedge d < R_{social_{\alpha_j}}$, $\beta_{h_{dist}}$ – параметр, аналогичный β_{imit} (см. описание к формуле (3.8)).

3.4 Формирование множества правил, описывающих структуру и функционирование МАС, основанной на зоосоциальных принципах

При подготовке текста п. 3.4 был использован авторский материал из опубликованных научных работ [57, 47].

Для определения оптимальной конфигурации параметров МАС, основанной на зоосоциальных принципах, применим метод, основанный на модификации множества правил поведения агентов и функционирования МАС (см. подраздел 2.3.1). Ниже представлены основные параметры (см. подразделы 3.1–3.3), определяющие структуру и функционирование МАС:

1) «Пищевой инстинкт»:

а) F_{α_i} – «пища» (варьируемый опосредованно через параметр «функция пищеварения», $\beta_{\text{hunger}}, \beta_{\text{hunger}} > 0$);

б) t_{hunger} – продолжительность «голодания» агента ($F_{\alpha_i} = 0$), приводящая к его исключению из МАС (отзыв на базу) (варьируемый напрямую, $t_{\text{hunger}} > 0$).

2) «Инстинкт экономии энергии»:

а) P_{rest} – «желание отдыха» (варьируемый опосредованно через параметр «скорость наступления усталости», $\beta_{\text{fatigue}}, \beta_{\text{fatigue}} > 0$);

б) $P_{\text{min_path}}$ – «показатель трудолюбия» (варьируемый напрямую, $P_{\text{min_path}} \in [0,1]$).

3) «Защитный инстинкт»:

а) P_{risk} – «стремление к риску» (варьируемый напрямую, $P_{\text{risk}} \in [0,1]$);

б) σ – пороговое значение вероятности выхода агента из строя (или его повреждения) в результате выполнения некоторых действий, после которого вероятность их применения определяется параметром P_{risk} (варьируемый напрямую, $\sigma \in [0,1]$).

4) «Инстинкт новизны»: P_{research} – «показатель любопытства» (варьируемый напрямую, $P_{\text{research}} \in [0,1]$).

5) «Исследовательский инстинкт»: P_{not_best} – «показатель любознательности» (варьируемый напрямую, $P_{not_best} \in [0,1]$).

6) «Подражательный инстинкт»:

а) $P_{imit} = \{P_{imit}^{vis}, P_{imit}^{private}, P_{imit}^{social}, P_{imit}^{public}\}$ – «желание подражать» (варьируемый опосредованно через параметр β_{imit} , $\beta_{imit} > 0$ (см. формулу (3.8)); $P_{imit}^{vis} \in [0,1]$, $P_{imit}^{private} \in [0,1]$, $P_{imit}^{social} \in [0,1]$, $P_{imit}^{public} \in [0,1]$ – желание подражать агентам в «зоне видимости», в «личной зоне», в «социальной зоне», в «публичной зоне»);

б) $P_{exper} = \{P_{exper}^{vis}, P_{exper}^{private}, P_{exper}^{social}, P_{exper}^{public}\}$ – «желание делиться своим опытом» (варьируемый напрямую; $P_{exper}^{vis} \in [0,1]$, $P_{exper}^{private} \in [0,1]$, $P_{exper}^{social} \in [0,1]$, $P_{exper}^{public} \in [0,1]$);

в) P_{learn} – «желание учиться» (варьируемый опосредованно через параметр β_{learn});

г) $p_{naivety} = \{0|1\}$ – «наивность агента» при восприятии опыта, см. подраздел 3.2 (варьируемый напрямую).

Дополнительные параметры: D_{event}^{vis} , $D_{event}^{private}$, D_{event}^{social} , D_{event}^{public} – множества кортежей вида (3.1), содержащих описание действий, разрешенных для подражания, а также возможные «образцы для подражания».

7) «Инстинкт свободы»: $P_{movement}$ – «желание двигательной активности» (варьируемый напрямую, $P_{movement} \in [0,1]$).

8) «Коммуникативный инстинкт»: $P_{horizontal}$ – «общительность» (варьируемый напрямую, $P_{horizontal} \in [0,1]$).

9) «Стадный инстинкт»:

а) P_{subor} – «желание подчиняться» (варьируемый напрямую, $P_{subor} \in [0,1]$);

б) P_{free} – «желание свободы» (варьируемый напрямую, $P_{subor} \in [0,1]$).

10) «Территориальный инстинкт»:

а) $\{R_{private}|R_{social}|R_{public}\}$ – радиус личной, социальной, публичной зоны соответственно (варьируемые напрямую, $R_{private} > 0$, $R_{social} > 0$, $R_{public} > 0$);

б) $R_{terr} = \{0|R_{vis}|R_{private}|R_{social}|R_{public}\}$ – «радиус индивидуальной дистанции» (варьируемый напрямую);

в) $A_{allowed}$ – «принципы территориальности» (варьируемый напрямую, возможные значения: агент-координатор, агенты из одного микрокластера, агенты из одной группы, все агенты);

г) $R_{public}^{prestige}$ – зависимость радиуса «публичной зоны» от «престижа» (варьируемый опосредованно через параметр $\beta_{prestige_k}$);

д) $R_{public}^{authority}$ – зависимость радиуса «публичной зоны» от «авторитета» (варьируемый опосредованно через параметр β_{public});

е) $R_{social}^{authority}$ – зависимость радиуса «социальной зоны» от «авторитета» (варьируемый опосредованно через параметр β_{social});

ж) $R_{social}^{prestige}$ – зависимость радиуса «социальной зоны» от «престижа» (варьируемый опосредованно через параметр β_{social_k}).

11) *Правила «доминирования»:*

а) ε_{α_i} – «авторитет» агента (варьируемый опосредованно через параметр $\beta_{oblivion}$ – «скорость забывания»);

б) $\tau_{\alpha_{coord}}$ – «престиж» агента-координатора (варьируемый опосредованно через весовые коэффициенты аргументов $w_j \in [0,1]$, $j = \overline{1,4}$, см. формулу (3.7));

в) $T_{distributed}$ – «приоритет доступа к пище» (варьируемый напрямую, возможные значения: «дискриминация», «обратная дискриминация»);

г) $T_{coll} = \{strong|weak|protection|leader|arbitration\}$ – «приоритет в арбитраже коллизий» (варьируется напрямую);

д) P_{power} – «преобладание права сильного» (варьируемый напрямую, $P_{power} \in [0,1]$);

е) P_{leader} - «преобладание права вожака» (варьируемый напрямую, $P_{leader} \in [0,1]$).

12) *Правила формирования вертикальных связей:*

а) $p_{good_food} = \{0|1\}$ – «правило пищи, пригодной к употреблению» (варьируемое напрямую);

б) ω_{coord} - величина временного интервала, на котором производится мониторинга показателя «авторитета» $\varepsilon_{\alpha_{i_t}}$ (или «престижа» $\tau_{\alpha_{i_t}}$) потенциального агента-координатора (варьируемый напрямую, $\omega_{coord} > 0$);

в) φ - пороговое значение, ниже которого не должен опускаться уровень «авторитета» потенциального агента-координатора α_i (простой агент-подчиненный или свободный агент) (варьируемый напрямую, $\varphi > 0$);

г) φ_k – аналог φ при приобретении роли агента-координатора k-го уровня иерархии;

д) ϑ_k – минимальное количество агентов в группе (варьируется напрямую, $\vartheta_k \geq 2$);

е) «критерий низкого престижа» - один из критериев лишения агента-координатора его статуса (варьируемый опосредованно через параметр γ_k – пороговое значение, ниже которого должен держаться уровень «престижа» агента-координатора k-го уровня иерархии в течение времени t , $\gamma_k < \varphi_k$);

ж) «критерий слабого агента-координатора» (варьируется опосредованно через параметр μ_k – пороговое значение, ниже которого должен держаться уровень «престижа» агента-координатора k-го уровня иерархии в течение времени t , $\mu_k > \gamma_k$, $\mu_k < \varphi_k$);

и) P_{dfp} - «желание перехватить власть» (варьируемый опосредованно через параметр $\beta_{sub \rightarrow coord}$);

к) «дальность агента-координатора» - учитывается при определении величины недовольства агента-подчинённого (варьируемый опосредованно через параметр β_{a_dist});

л) «привлекательность агента-координатора» (варьируемый опосредованно через параметр $\beta_{c_{pr}}$);

м) $P_{a_{auth}}$ - распределение вероятности выхода агента-подчиненного α_i из группы в зависимости от уровня его собственного «авторитета» ε_{α_i} (варьируемый опосредованно через параметр $\beta_{a_{auth}}$);

13) *Правила формирования горизонтальных связей*: $P_{h_{dist}}$ – распределение вероятностей формирования горизонтальных связей в зависимости от расстояния между агентами (варьируется опосредованно через параметр $\beta_{h_{dist}}$).

Очевидно, что для правил, представленных в подразделах 3.1–3.3 также может быть построена иерархическая структура (рисунок 3.1), подобная содержащейся на рисунке 2.4 (см. подраздел 2.3.1).

На рисунке 3.1 стрелка указывает на доминирующий параметр. Стрелки со сплошным контуром указывают на аннулирующие или корректирующие ограничения (соответствующие параметры исключаются полностью, либо для них ограничивается или изменяется область допустимых значений, см. подраздел 2.3.1), с пунктирным – на зависимые параметры (параметр, на который указывает стрелка, является аргументом при определении параметра, из которого стрелка выходит, соответствует «агрегации» UML [32]).

Примечание: Параметры, для которых имеются аннулирующие ограничения, также, в большинстве случаев, связаны и корректирующими ограничениями.

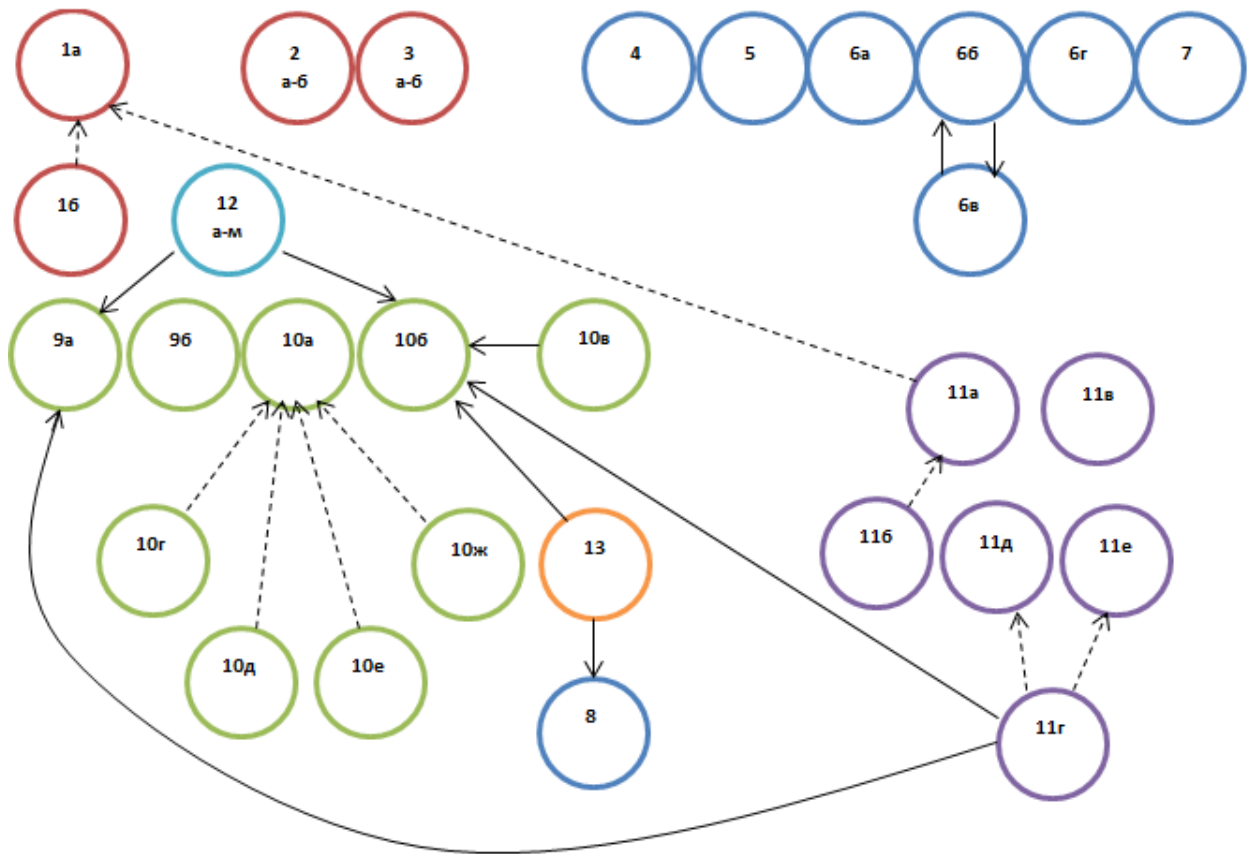


Рисунок 3.1 – Граф ограничений и зависимостей параметров МАС

Применим к заданному множеству параметров МАС (см. рисунок 3.1) метод, основанный на модификации множества правил поведения агентов и функционирования МАС (реорганизация на макроуровне), представленный в подразделе 2.3.1 (параметры 6г, 12а будут представлены в бинарном виде; исключение – параметры 10б, 10в, имеющие фиксированное множество возможных значений, остальные параметры будут представлены в непрерывном виде). Для подбора оптимальных параметров МАС будем использовать «генетический алгоритм» [42]. При этом хромосома [42] будет включать множество параметров 1–13, где каждый из них будет представлять собой отдельный «ген». При выполнении оператора мутации [42] новое значение гена будет определяться случайным образом (в пределах области допустимых значений соответствующего параметра).

3.5 Выводы

При подготовке текста п. 3.5 был использован авторский материал из опубликованных научных работ [57].

Предложен «зоосоциальный метод автоматической адаптации организационной структуры МАС» (является методом адаптации архитектуры МАС на макроуровне, см. подраздел 2.3.1), предусматривающий реализацию множества правил поведения агентов и функционирования МАС в соответствии с принципами, характерными для социальных животных, при этом для определения оптимальных параметров МАС применяется генетический алгоритм. Отличиями предложенного метода от аналогов являются:

- возможность представления всех типов организационных структур, перечисленных в подразделе 2.1, путем подбора различных комбинаций предложенных параметров МАС;
- возможность оперативной корректировки организационной структуры МАС путем изменения значений предложенных параметров МАС.

ГЛАВА 4. АРХИТЕКТУРА И АЛГОРИТМЫ РАБОТЫ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ОПТИМИЗАЦИИ ОРГАНИЗАЦИОННОЙ СТРУКТУРЫ МАС. ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА ЭФФЕКТИВНОСТИ ПРЕДЛОЖЕННОГО МЕТОДА И АЛГОРИТМОВ

4.1 Алгоритм адаптации организационной структуры МАС на макроуровне, основанный на зоосоциальном методе

При подготовке текста п. 4.1 был использован авторский материал из опубликованных научных работ [52, 53, 47].

4.1.1 Структура хромосомы

Структура применяемой хромосомы (подробно рассматривается в подразделе 3.4) имеет следующий вид:

$$H = \langle \beta_{hunger}, t_{hunger}, \beta_{fatigue}, P_{min\ path}, P_{risk}, \sigma, P_{research}, P_{not_best}, \beta_{imit}, P_{exper}^{vis}, P_{exper}^{private}, P_{exper}^{social}, P_{exper}^{public}, \beta_{learn}, p_{naivety}, P_{movement}, P_{horizontal}, P_{subor}, P_{free}, R_{private}, R_{social}, R_{public}, R_{terr}, A_{allowed}, \beta_{prestige_k}, \beta_{public}, \beta_{social}, \beta_{social_k}, \beta_{oblivion}, w_1, w_2, w_3, w_4, T_{distribut}, T_{coll}, P_{power}, P_{leader}, p_{good_food}, \omega_{coord}, \varphi, \varphi_k, \vartheta_k, \gamma_k, \mu_k, \beta_{sub \rightarrow coord}, \beta_{a_dist}, \beta_{c_pr}, \beta_{a_auth}, \beta_{h_dist} \rangle, \quad (4.1)$$

при ограничениях: $\beta_{hunger} > 0$, $t_{hunger} > 0$, $\beta_{fatigue} > 0$, $P_{min_path} \in [0,1]$, $P_{risk} \in [0,1]$, $\sigma \in [0,1]$, $P_{research} \in [0,1]$, $P_{not_best} \in [0,1]$, $\beta_{imit} > 0$, $P_{exper}^{vis} \in [0,1]$, $P_{exper}^{private} \in [0,1]$, $P_{exper}^{social} \in [0,1]$, $P_{exper}^{public} \in [0,1]$, $\beta_{learn} \in \mathbb{R}$, $p_{naivety} = \{0|1\}$, $P_{movement} \in [0,1]$, $P_{horizontal} \in [0,1]$, $P_{subor} \in [0,1]$, $P_{free} \in [0,1]$, $R_{private} > 0$, $R_{social} > 0$, $R_{public} > 0$, $R_{terr} = \{0|R_{vis}|R_{private}|R_{social}|R_{public}\}$, $A_{allowed} = \{ac|amc|ag|all\}$ (ac - агент-координатор, amc - агенты из одного микрокластера, ag - агенты из одной группы, all - все агенты), $\beta_{prestige_k} \in \mathbb{R}$, $\beta_{public} \in \mathbb{R}$, $\beta_{social} \in \mathbb{R}$, $\beta_{social_k} \in \mathbb{R}$, $\beta_{oblivion} \in \mathbb{R}$, $w_j \in [0,1]$ ($j = \overline{1,4}$), $T_{distribut} = \{disk|b_disk\}$ ($disk$ - «дискриминация», b_disk - «обратная

дискриминация»), $T_{coll} = \{strong|weak|protection|leader|arbitration\}$ (*strong* - «право сильного», *weak* - «защита слабых»; *protection* - «право протекции», *leader* - «право вожака», *arbitration* - «третейский судья»), $P_{power} \in [0,1]$, $P_{leader} \in [0,1]$, $p_{good_food} = \{0|1\}$, $\omega_{coord} > 0$, $\varphi > 0$, $\varphi_k > 0$, $\vartheta_k \geq 2$, $\gamma_k < \varphi_k$, $\mu_k > \gamma_k$, $\mu_k < \varphi_k$, $\beta_{sub \rightarrow coord} \in \mathbb{R}$, $\beta_{a_dist} \in \mathbb{R}$, $\beta_{c_pr} \in \mathbb{R}$, $\beta_{a_auth} \in \mathbb{R}$, $\beta_{h_dist} \in \mathbb{R}$, $P_{subor} = 0 \rightarrow vertRelation$, где $vertRelation = \{p_{good_food} = \emptyset, \omega_{coord} = \emptyset, \varphi = \emptyset, \varphi_k = \emptyset, \vartheta_k = \emptyset, \gamma_k = \emptyset, \mu_k = \emptyset, \beta_{sub \rightarrow coord} = \emptyset, \beta_{a_dist} = \emptyset, \beta_{c_pr} = \emptyset, \beta_{a_auth} = \emptyset\}$ (т.е. при $P_{subor} = 0$ учет перечисленных параметров теряет смысл, аннулирующее ограничение), $R_{terr} = R_{public} \rightarrow vertRelation$ (аннулирующее ограничение), $R_{terr} = R_{social} \rightarrow \beta_{h_dist} = \emptyset$ (аннулирующее ограничение), $R_{terr} = 0 \rightarrow A_{allowed} = \emptyset$ (аннулирующее ограничение), $P_{subor} = 0 \rightarrow T_{coll} = \{strong|weak\}$ (корректирующее ограничение), $R_{terr} = R_{public} \rightarrow T_{coll} = \{strong|weak\}$ (корректирующее ограничение), $P_{horizontal} = 0 \rightarrow \beta_{h_dist} = \emptyset$ (аннулирующее ограничение), $P_{exper}^{public} = 0 \rightarrow P_{learn} = \emptyset$, при $R_{public} > R_{social}$ (т.е., если агенты не делятся опытом в «публичной зоне», то величина «желания учиться» не имеет значения, аннулирующее ограничение), $P_{exper}^{social} = 0 \rightarrow P_{learn} = \emptyset$, при $R_{public} < R_{social}$ (ограничение, аналогичное предыдущему), $P_{learn} = 0 \rightarrow P_{exper}^{vis|private|social|public} = \emptyset$, (т.е., если агенты не желают учиться, то величина вероятности трансляции опыта не имеет значения, аннулирующее ограничение).

Примечание: Параметры, для которых установлено значение $\emptyset$, при выполнении операций скрещивания и мутации не учитываются.

4.1.2 Функция приспособленности

Соответственно, в рамках рассматриваемого алгоритма предполагается определение некоторой хромосомы H , для которой выполняется следующее условие:

$$f_{fitness}(H) \rightarrow \max, \quad (4.2)$$

где $f_{fitness}$ – функция приспособленности хромосомы.

Выделяются следующие режимы работы МАС: «режим обучения», «основной режим» (см. подраздел 2.1). Очевидно, что в вид функции при различных режимах работы МАС (режим обучения, основной режим, см. подраздел 2.1) будет отличаться:

$$f_{fitness} = \{f_{fitness}^{learn} | f_{fitness}^{main}\}, \quad (4.3)$$

где $f_{fitness}^{learn}$ – функция приспособленности хромосомы в режиме обучения, $f_{fitness}^{main}$ – функция приспособленности хромосомы в основном режиме. Применяемые виды функций $f_{fitness}^{learn}$, $f_{fitness}^{main}$ приводятся в подразделе 4.4.

Целью «режима обучения» является приобретение агентами знаний об окружающей среде, формирование кортежей вида (3.3). Таким образом, в «режиме обучения» предполагается формирование организационной структуры, способствующей повышению эффективности процесса приобретения агентами знаний об окружающей среде. Например, при решении задачи патрулирования некоторого объекта, агенты вначале исследуют окружающую среду, её структуру, поведение, как следствие, достижению поставленной цели может способствовать повышение «любопытности» агентов (увеличение вероятности выбора пары «состояние-действие», ценность которой не является максимальной), их желания к обмену опытом (чтобы ускорить обучение).

Целью «основного режима» является максимизация производительности МАС, т.е. предполагается, что агенты уже достаточно изучили окружающую среду, и организационная структура должна быть адаптирована для повышения производительности МАС. Например, при решении ранее упомянутой задачи патрулирования, данная цель может быть

достигнута путем снижения исследовательской активности агентов, интенсивности репликации опыта.

В результате формируется организационная структура, удовлетворяющая условию:

$$o_{rw}: f_{GQ}^T \rightarrow \max, \quad (4.4)$$

где $o_{rw} = \{o_{learn}|o_{main}\}$, o_{learn} - организационная структура МАС в «режиме обучения», развертываемая на основании параметров, содержащихся в «хромосоме» H_{learn} (см. (4.1)), o_{main} – организационная структура МАС в «основном режиме», развертываемая на основании параметров, содержащихся в «хромосоме» H_{main} (см. (4.1)), f_{GQ}^T – глобальная функция качества МАС при решении задачи T в рамках окружающей среды E (см. (2.27)), $f_{GQ}^T = \{f_{fitness}^{learn}|f_{fitness}^{main}\}$.

Достижение условия (4.4) возможно путем приложения методов $\{M_{deploy}, M_{modify}\}$ (см. формулу (2.27)), соответствующих разработанному «зоосоциальному методу автоматической адаптации организационной структуры МАС». При этом цель M_{deploy} – определение априорной $o_{main_{t_0}}$ или $o_{learn_{t_0}}$ (организационная структура МАС в начале её функционирования, в момент времени t_0 , сформированная на основании имеющихся знаний о поставленной задаче T и окружающей среде E), применение которой для пары $\langle T, E \rangle$ будет способствовать достижению условия (4.4). Целью M_{modify} (см. (2.22)) является формирование модифицирующего воздействия Δo_{main_t} или Δo_{learn_t} при наступлении условий (2.24)-(2.26) в момент времени t , $t > t_0$.

Таким образом, в обоих режимах предполагается формирование множества «кортежей опыта» следующего вида:

$$C = \langle E, T, H, f_{fitness} \rangle, \quad (4.5)$$

при этом примем $f_{fitness} \equiv f_{GQ}^T$.

При этом организационная структура будет сформирована в соответствии с $H \in \mathbf{C}_{best}$:

$$H_{\mathbf{C}_{best}} \xrightarrow{f_{H \rightarrow o}} o_{\mathbf{C}_{best}}, \quad (4.6)$$

где $\mathbf{C}_{best}$ – хромосома, удовлетворяющая условию (4.4), $f_{H \rightarrow o}$ – функция отображения хромосомы $H_{\mathbf{C}_{best}}$ в организационную структуру $o_{\mathbf{C}_{best}}$.

Замечание: Рассматриваемый подход не позволяет определить $\frac{\partial f_{GQ}^T}{\partial h_i}$ – скорость изменения (частную производную) глобальной функции качества f_{GQ}^T по параметрам h_i , где h_i – отдельный «ген» хромосомы H , $i = \overline{1, |H|}$. Таким образом, отсутствует возможность ранжирования «генов» h_i по степени корреляции с функцией f_{GQ}^T и, как следствие, снижение размерности хромосомы H .

Исходя из приведенного замечания, перепишем формулу (4.5) в следующем виде:

$$\mathbf{C}_E^T = \langle \mathbf{E}, \mathbf{T}, \boldsymbol{\tau} = \left\{ \boldsymbol{\tau}_k = \left\{ \mathbf{C}_i, \mathbf{C}_{i+1}, \dots, \mathbf{C}_{\tau_{kN}} \right\} \right\}, \pi: H \rightarrow f_{fitness} \rangle, \quad (4.7)$$

где $\mathbf{C}_E^T$ – «опыт» решения агентами МАС задачи $\mathbf{T}$ в окружающей среде $\mathbf{E}$, $\boldsymbol{\tau}$ – множество траекторий $\boldsymbol{\tau}_k$, содержащих кортежи $\mathbf{C}_i$, являющиеся урезанными версиями (4.5), $\mathbf{C}_i = \langle H, f_{fitness} \rangle$, при этом $H \in \mathbf{C}_i$, $H' \in \mathbf{C}_{i+1}$, хромосома H' является дочерней H , $k = \overline{1, |\boldsymbol{\tau}|}$, $i = \overline{1, \tau_{kN}}$, τ_{kN} – число кортежей траектории $\boldsymbol{\tau}_k$, $\pi_{\mathbf{E}, \mathbf{T}}$ – «функция стратегии», позволяющая определить вероятность выбора хромосомы H для реализации организационной структуры МАС в рамках окружающей среды $\mathbf{E}$ при решении задачи $\mathbf{T}$, $\pi: \langle \mathbf{d}_E, H, f_{fitness}(H) \rangle \rightarrow [0, 1]$, при этом для $H: f_{fitness}(H) \rightarrow \max$ величина $\pi_{\mathbf{E}, \mathbf{T}}(\mathbf{d}_E, H, f_{fitness}(H)) \cong 1$, где $\mathbf{d}_E$ – параметры окружающей среды $\mathbf{E}$, при которых МАС, построенная по «хромосоме» H , достигла показателей $f_{fitness}(H)$.

Однако определение функции $\pi_{\mathbf{E}, \mathbf{T}}$ невозможно при неизвестной $f_{fitness}$, поэтому для аппроксимации функции $\pi_{\mathbf{E}, \mathbf{T}}$ применим «метод градиента политики» [95, 45] (см. подраздел 4.1.4).

Пусть S_E^T – множество состояний некоторой МАС при решении задачи T в окружающей среде E , $S_E^T = \{S_j\}$, $j = \overline{1, |S_E^T|}$, при этом состояние S_j имеет следующую структуру:

- $f_{fitness}(E, T, H)$ – значение функции $f_{fitness}$, достигнутое при решении задачи T в окружающей среде E с помощью МАС, имеющей организационную структуру O_H ;
- d_E – описание окружающей среды E ;
- хромосома H .

Для перехода в новое состояние S_j' может быть выполнено действие $A_k = \Delta H$, $\Delta H = H - H'$, где H' – хромосома, являющаяся дочерней H , $k = \overline{1, |A|}$, A – множество действий, допустимых для осуществления перехода $S_j \rightarrow S_j'$. При этом действие A_k представляет собой некоторое множество новых «генов» $A_k = \{h_i\}$, $|A_k| \geq 1$, внедряемых в «хромосому» с помощью операторов скрещивания или мутации.

Применяемые виды операторов скрещивания и мутации рассматриваются в следующем подразделе.

4.1.3 Генетические операторы

Оператор скрещивания.

При выполнении оператора скрещивания для отбора пары хромосом $\langle H_i, H_j \rangle$ для «спаривания» будем использовать «метод рулетки» [42], где вероятность скрещивания $P_{cross} \propto f_{fitness}(H)$, $H \equiv \{H_i | H_j\}$, при этом $\langle H_i, H_j \rangle \subset H_{best}$, H_{best} – множество лучших хромосом, $\forall H \in H_{best}: f_{fitness}(H) > \theta_{fitness}$, $\theta_{fitness}$ – пороговое значение, определяющее минимальный уровень приспособленности хромосомы для отбора во множество H_{best} . В результате скрещивания хромосом формируется $\langle H_i, H_j \rangle N_{children}$ потомков. При этом с целью достижения

максимального разнообразия вариантов дочерних хромосом применим «сегментированное скрещивание», при этом число точек скрещивания выбирается случайным образом [42].

При подборе пар $\langle H_i, H_j \rangle$ будем использовать и «видовые стратегии» (препятствование скрещиванию между особями с близкими хромосомами) с целью поощрения исследования пространства состояний [42], т.е. $\langle H_i, H_j \rangle: D(H_i, H_j) > \theta_{kin}$, где D – мера сходства хромосом (например, «Евклидова метрика»), θ_{kin} – пороговое значение, определяющее минимальный уровень «родства» для выполнения скрещивания, $i = \overline{1, |\mathbf{H}_{best}|}, j = \overline{1, |\mathbf{H}_{best}|}, i \neq j$.

Оператор мутации.

Для хромосомы H_i применим «равномерную мутацию» [42], что также позволит рассмотреть наибольшее количество хромосом H_i' , для которых $D(H_i, H_i') < \theta_{kin}$:

$$r \leftarrow U[0,1], \quad (4.8)$$

$$H_i(k) \leftarrow \begin{cases} H_i(k) & \text{при } r \geq P_{mutation} \\ U[H_i(k) - \alpha(k), H_i(k) + \alpha(k)] & \text{при } r < P_{mutation} \end{cases}, \quad (4.9)$$

где U – «случайное число, равномерно распределенное на $[0,1]$ » [42], $\alpha(k)$ – параметр величины мутации, k – номер «мутировавшего» «гена», $k = \overline{1, |H_i|}$, $P_{mutation}$ – вероятность мутации, задается эмпирическим путем. Формулы (4.8)-(4.9) взяты из [42].

Примечание: На начальных этапах поиска с целью максимизации пространства поиска примем $\alpha(k) \rightarrow \max$, в свою очередь на завершающих этапах поиска примем $\alpha(k) \rightarrow \min$ для повышения точности найденного решения [81]. Исходя из этого функция для определения параметра величины мутации примет следующий вид:

$$\alpha(t) = mt + \alpha_{\max},$$

где m – коэффициент наклона графика функции, $m < 0$, α_{max} – значение α при $t = 0$, t – номер итерации алгоритма, $t = \overline{0, t_{end}}$, t_{end} – номер конечной итерации алгоритма.

В данном случае генетические операторы необходимы для формирования траекторий τ_k (см. формулу (4.7)) следующим образом (см. псевдокод ниже) [42]:

Алгоритм 4.1.

Задать t_{end} – число итераций генетического алгоритма, $N_{population}$ – размер популяции, $\theta_{fitness}$, θ_{kin} , $P_{mutation}$, $\mathbf{H} \leftarrow \emptyset$ – множество хромосом («популяция»), $\mathbf{C} \leftarrow \emptyset$ – множество урезанных кортежей (4.5) (см. также (4.7)), $\mathbf{H}_{best} \leftarrow \emptyset$, $\tau \leftarrow \emptyset$;

//создание первоначальной популяции

Для $i = 1, 2, \dots, N_{population}$ делать:

{

Для $j = 1, 2, \dots, |H_i|$ делать:

{

Если h_j – «ген», заполняемый непрерывными значениями, то:

$h_j \leftarrow U[h_{j_{min}}, h_{j_{max}}]$; // где $U[h_{j_{min}}, h_{j_{max}}]$ – «случайное число, равномерно распределенное на $[h_{j_{min}}, h_{j_{max}}]$ »

Иначе, если h_j заполняется дискретными значениями, то:

$h_j \leftarrow \{h_{j_1} | h_{j_2} | \dots | h_{j_{N_{h_j}}}\}$; // где N_{h_j} – число

элементов множества допустимых значений h_j

}

$\mathbf{H} \leftarrow \mathbf{H} \cup \{H_i\}$;

}

$t \leftarrow 1$;

$k \leftarrow 1$;

Пока $t \leq t_{end}$ делать:

{

//расчет значений функции приспособленности

Для $\forall H_i \in \mathbf{H}$ делать:

{

Получить значение $f_{fitness}(H_i)$;

```

 $C_i \leftarrow \langle H_i, f_{fitness}(H_i) \rangle;$ 
ЕСЛИ  $f_{fitness}(H_i) > \theta_{fitness}$ , ТО:
 $C \leftarrow C \cup \{C_i\}; H_{best} \leftarrow H_{best} \cup \{H_i\},$ 
}
 $H_{children} \leftarrow \emptyset;$ 
//выполнение скрещивания
Для  $i = 1, 2, \dots, |H_{best}|$  ДЕЛАТЬ:
{
    Label:  $\langle H_j, H_k \rangle = \text{Рулетка}(H_{best});$  // Выбрать
«родителей»  $\langle H_j, H_k \rangle$  «методом рулетки», где  $j = \overline{1, N_{population}},$ 
 $k = \overline{1, N_{population}}, j \neq k$ 
    ЕСЛИ  $D(H_j, H_k) > \theta_{kin}$ , ТО: перейти к метке Label;
 $H_i \leftarrow \text{Скрещиванию}(H_j, H_k);$ 
 $H_i \leftarrow \text{Мутация}(H_i, P_{mutation});$ 
 $\tau_k \leftarrow \{H_j, H_i\};$ 
 $\tau \leftarrow \tau \cup \{\tau_k\};$ 
 $k \leftarrow k + 1;$ 
 $\tau_k \leftarrow \{H_k, H_i\};$ 
 $\tau \leftarrow \tau \cup \{\tau_k\};$ 
 $k \leftarrow k + 1;$ 
 $H_{children} \leftarrow H_{children} \cup \{H_i\};$ 
}
 $H \leftarrow H_{children};$ 
 $H_{best} \leftarrow \emptyset$ 
 $t \leftarrow t + 1;$ 
}

```

В данном подразделе рассмотрен процесс генерации данных для дальнейшего обучения «функции стратегии», выполняемого с помощью «метода градиента политики» [95, 45] (подробное описание см. в следующем подразделе).

4.1.4 Аппроксимация «функции стратегии»

Аппроксимация «функции стратегии» осуществляется с помощью алгоритма «REINFORCE», основанного на «методе градиента стратегии» [45]:

Алгоритм 4.2.

Задать α – параметр скорости обучения, θ – множество весовых коэффициентов стратегии π ;

Для $k=1, 2, \dots, |\tau|$ делать:

{

Извлечь траекторию $\tau_k \in \tau$;

$\tau_{k_{norm}} \leftarrow \text{Нормализация_траектории}(\tau_k)$;

Задать $\nabla_{\theta} J(\pi_{E,T}) \leftarrow 0$;

Для $i = 1, 2, \dots, |\tau_{k_{norm}}|$ делать:

{

$R_i(\tau_{k_{norm}}) \leftarrow \sum_{i'=i}^{|\tau_{k_{norm}}|} \gamma^{i'-i} r_{i'}$;

$\nabla_{\theta} J(\pi_{E,T}) \leftarrow \nabla_{\theta} J(\pi_{E,T}) + R_i(\tau_{k_{norm}}) \nabla_{\theta} \log \pi_{E,T}(A_i | S_i)$;

}

$\theta \leftarrow \theta + \alpha \nabla_{\theta} J(\pi_{E,T})$;

}

Алгоритм 4.2 содержит следующие элементы:

$J(\pi_{E,T})$ – целевая функция, представляющая «абстрактную гиперповерхность с θ в качестве переменных» [45];

$\pi_{E,T}$ – «функция стратегии» (см. формулу (4.7)), аппроксимация которой осуществляется с помощью *искусственной нейронной сети (ИНС)*, имеющей *весовые коэффициенты* θ ;

$\tau_{k_{norm}}$ – нормализованная траектория τ_k , т.е. траектория вида

$$\tau_{k_{norm}} = \{\langle S_i, A_i, S_{i+1}, r_i \rangle, \langle S_{i+1}, A_{i+1}, S_{i+2}, r_{i+1} \rangle, \dots, \langle S_{N-1}, A_{N-1}, S_N, r_{N-1} \rangle\}, \quad (4.10)$$

где $N = |\tau_{k_{norm}}|$, S_i – состояние МАС, A_i – действие, выполнение которого приводит к переходу в состояние S_{i+1} (применяемый формат состояний отличается от предложенного в подразделе 4.1.2 и поясняется далее, см. формулу (4.13));

Приведение к виду (4.10) осуществляется с помощью функции «Нормализация_траектории»;

$|\tau_{k_{norm}}|$ – число кортежей вида $\langle S_i, A_i, S_{i+1}, r_i \rangle$, которые содержит траектория $\tau_{k_{norm}}$;

$\nabla_{\theta} J(\pi_{E,T})$ – «градиент стратегии» [95, 45];

$\gamma^{i'-i}$ – «коэффициент обесценивания» [95, 45], $\gamma^{i'-i} \in [0, 1]$;

r_i – величина вознаграждения, полученная за реализацию перехода

$$\mu_i = \langle S_i, A_i, S_{i+1} \rangle, \quad (4.11)$$

$r_i \in [-1, 1]$, $\mathbf{r}(\mu_i) = r_i$, $\mathbf{r}$ – функция вычисления вознаграждения, $\mathbf{r}: f_{fitness} \rightarrow [-1, 1]$;

Примечание: В подразделе 4.1.2 в формат состояния S_i включено значение «функции приспособленности» $f_{fitness}(H_i)$, $H_i \in S_i$, при этом очевидно, что $f_{fitness}(H_i) \propto r_i$. Исходя из этого, примем следующий формат состояния:

$$S_i \equiv \{\mathbf{d}_E, H_i\}, \quad (4.12)$$

где $\mathbf{d}_E$ – описание окружающей среды E , при этом также перепишем формулу (4.11) в следующем виде:

$$\mu_i = \langle S_i, A_i, S_{i+1}, r_i \rangle, \quad (4.13)$$

где $\mu_i \in \tau_k$.

Примечание: Входные значения ИНС, аппроксимирующей $\pi_{E,T}$, будут соответствовать структуре состояния (4.13). Однако при определении выходных значений ИНС, отображающих вероятность выбора того или иного действия (принятый формат действия см. в подразделе 4.1.2), необходимо учитывать следующие факторы:

- часть «генов» хромосомы представляет непрерывные значения;
- значения «генов» могут быть увеличены, так и уменьшены, при этом величина приращения или разности для двух соседних точек траектории не является фиксированной.

Таким образом, для «генов», имеющих непрерывные значения, выполним дискретизацию с шагом d , получив, таким образом, следующее множество выходов ИНС:

$$Y = \{y_k^{h_i}\}, \quad (4.14)$$

где $y_k^{h_i}$ – выход, обозначающий вероятность того, что дискретный (дискретный изначально, либо дискретизированный с шагом d) «ген» h_i некоторой хромосомы H примет значение k , $k \in U_{h_i}$, U_{h_i} – множество

допустимых значений «гена» h_i , $y_k^{h_i} \in [0,1]$. Таким образом, вероятность возможных значений дискретизированного «гена» h_i будет обозначать $\frac{h_{i\max} - h_{i\min}}{d}$ выходов, где $h_{i\max}$, $h_{i\min}$ – максимальное и минимальное значения «гена» h_i .

Таким образом, имеем следующую структуру ИНС:

$$ANN = \{X, Y, L\}, \quad (4.15)$$

где X – множество входов, $X \equiv S_i$ (см. (4.12)); Y – множество выходов (см. (4.14)), L – множество слоев ИНС.

Примечание: Пусть соседние кортежи μ_i , μ_{i+1} (см. формулу (4.13)) траектории τ_k (см. формулу (4.7)) содержат исходные состояния S_i , S_{i+1} , $S_{i+1} - S_i = \Delta S_{i+1,i} \propto A_i$, A_i – действие, выполнение которого привело к осуществлению перехода $S_i \rightarrow S_{i+1}$. Однако, согласно формату действий, приводимому в подразделе 4.1.2, реализация A_i может заключаться в изменении значений множества «генов» $h_H^{change} = \{h_i = h_{i_{t-1}} - h_{i_t}\}$, $h_H^{change} \subset H$, при этом $|h_H^{change}| \geq 1$, $t = \overline{1, |\tau_k|}$. В этом случае, P_{A_i} (вероятность выполнения такого составного действия $A_i \equiv h_H^{change}$) будет отображаться на множество выходов ИНС как $y_k^{h_i} = P_{A_i} \forall h_i \in h_H^{change} \wedge h_{i_t} \in k$.

4.1.5 Структура алгоритма

«Алгоритм адаптации организационной структуры МАС на макроуровне, основанный на зоосоциальном методе» иллюстрируется с помощью методологии IDEF0 на рисунке 4.1.



Рисунок 4.1 – Иллюстрация «алгоритма адаптации организационной структуры МАС на макроуровне, основанного на зоосоциальном методе

Предложенный алгоритм предназначен для определения оптимальной структуры множества **R**: базовые законы функционирования МАС - множество допустимых правил, регламентирующих функционирование МАС, координацию действий агентов, её составляющих, её взаимодействие с окружающей средой. Различают априорный подбор параметров — подбор с помощью ИНС (4.15) такого стартового набора параметров, формализованных в «хромосоме» H , который позволит в дальнейшем обеспечить функционирование МАС (в основном режиме или режиме обучения) с максимальной эффективностью при решении задачи T в окружающей среде E . При незначительных флуктуациях параметров окружающей среды E (прогнозируемых или случайных, но допустимых в рамках законов функционирования окружающей среды данного типа) в процессе решения задачи T изменения значений «генов» хромосомы H , отображаемой во множество R , не требуется, поскольку организационная структура МАС адаптируется автоматически благодаря применению зоосоциального метода.

Однако при существенных изменениях законов функционирования окружающей среды E в процессе решения задачи T (например, если в статической окружающей среде появились динамические элементы) также необходимо изменение базовых законов функционирования МАС, т.е. генерация новой хромосомы H с помощью ИНС (4.15) (и, как следствие, внесение изменений во множество R) путем подачи на её вход описания нового состояния окружающей среды d_E , а также текущей хромосомы H . На выходе ИНС будут получены рекомендации по изменению хромосомы H с целью максимизации глобальной функции качества МАС.

Примечание: Наиболее объективным индикатором необходимости реорганизации МАС являются текущие значения глобальной или локальной функции качества. Даже в случае существенного изменения параметров МАС её реорганизация не является целесообразной при сохранении показателей глобальной или локальной функции качества примерно на том же уровне.

Примечание: Переход между режимом обучения и основным режимом осуществляется путем изменения глобальной функции качества. Например, d в режиме обучения глобальная функция качества может соответствовать показателю $R(\tau_k)$ – получаемая агентами усредненная награда по всем траекториям τ_k , в основном режиме – время выполнения задачи, вероятность достижения цели и т.п. (см. подраздел 4.3).

Предложенный алгоритм позволяет определить вид хромосомы (4.1), отображение которой в организационную структуру МАС позволит на макроуровне максимизировать значения глобальной функции качества для указанного типа окружающей среды, состава множества агентов, решаемой задачи. Однако масштабы окружающей среды могут быть настолько велики, что законы функционирования отдельных её кластеров (подсистем) могут различаться (далее окружающая среда такого типа будет обозначаться как «неоднородная»). Очевидно, что в таких условиях правила функционирования отдельных составляющих МАС (на микроуровне) также должны отличаться, поскольку применение единых правил (оптимизация на

макроуровне), общих для всех агентов, скорее всего, позволит получить лишь некоторое компромиссное (усредненное) значение глобальной функции производительности, но не максимальное. Алгоритм оптимизации, применяемый для описанных случаев, приводится в следующем подразделе.

4.2 Алгоритм оптимизации организационной структуры МАС в условиях неоднородной окружающей среды. Архитектура интеллектуальной системы оптимизации организационной структуры МАС

При подготовке текста п. 4.2 был использован авторский материал из опубликованных научных работ [58, 47].

4.2.1 Алгоритм оптимизации организационной структуры МАС в условиях неоднородной окружающей среды

Пусть окружающая среда E включает кластеры (подсистемы) v_k^E ,

$$E = \bigcup_{k=1}^{N_E} v_k^E, \quad (4.16)$$

при этом процесс функционирования каждого кластера (подсистемы) v_k^E описывается множеством функций $F_{v_k^E}$, а их взаимодействие в рамках E описывается множеством функций F_E .

В качестве примера кластера (подсистемы) вида (4.16) могут быть рассмотрены сельскохозяйственные угодья, включающие поля, засеянные различными культурами, сады, виноградники. Очевидно, что для каждого такого поля имеет место условие

$$F_{v_a^E} \neq F_{v_b^E}, \quad (4.17)$$

где v_a^E – поле, засеянное культурой типа a , v_b^E – поле, засеянное культурой типа b . Пусть, например, культура a – это пшеница, культура b – виноградник. Вполне ясно, что условия работы на полях этих двух типов

будут отличаться. Примером задачи T , реализуемой в рамках описанной окружающей среды, может являться выполнение обработки от вредителей. В этом случае могут применяться агенты, имеющие гомогенное техническое исполнение. При этом по мере выполнения поставленной задачи вполне возможно перемещение агентов между полями двух типов (например, если скорость обработки полей существенно отличается). Исходя из (4.17) очевидно, что в данном случае необходимо три хромосомы вида (4.1):

- хромосома, описывающая поведение агентов, обрабатывающих поле с культурой a ;
- хромосома, описывающая поведение агентов, обрабатывающих поле с культурой b ;
- хромосома, регламентирующая процесс взаимодействия между агентами, работающим на разных полях.

Однако также нужно учитывать, что агентам априори могут быть неизвестны параметры окружающей среды, следовательно, выводы о её неоднородности могут быть сделаны в процессе функционирования МАС. Следовательно, решение о формировании своих хромосом для отдельных кластеров может быть принято в процессе решения поставленной задачи.

Таким образом, может быть сформирован *алгоритм оптимизации организационной структуры МАС в условиях неоднородной окружающей среды*, включающий следующие этапы:

1) *Кластеризация состояний*. Может производиться на основании информации, собранной агентами, а именно множество состояний S_i , в которых находились агенты, $S_i = \{s_k\}$, где s_k – некоторый параметр, характеризующий состояние S_i , $i = \overline{1, |S|}$, S – множество состояний, в которых могут находиться агенты в рамках окружающей среды E , $k = \overline{1, |S_i|}$. Для $\forall S_i$ назначается класс C_j , $C_j \in C$, $j = \overline{1, |C|}$, при этом множество классов C формируется с помощью метода «k-means» [86].

Примечание: Параметры s_k могут характеризовать, например, габариты объектов окружающей среды, особенности их размещения в рамках окружающей среды и т.п.

2) *Кластеризация окружающей среды.* На основании множества пар $\langle S_i, C_j \rangle$ производится выделение кластеров C_k , $k = \overline{1, |C|}$, C – множество кластеров, выделенных в окружающей среде E . Таким образом, исходим из того, что если структура состояний существенно отличается, а также есть тенденция к локализации состояний отдельных классов в определенных областях, то такая окружающая среда является неоднородной.

3) Для каждого кластера C_k ($C_k \in C$) с помощью алгоритма оптимизации организационной структуры МАС на макроуровне, основанного на зоосоциальном методе (см. подраздел 4.1), подбирается хромосома H_{C_k} . Таким же образом формируется хромосома H_E , регламентирующая взаимодействие агентов, относящихся к разным кластерам.

Чтобы минимизировать издержки, связанные с некорректным определением параметров s_k , используемых при формировании классов C_j , изменим структуру хромосомы, описываемой формулой (4.1), добавив параметр $P_{cluster}$ – вероятность выделения кластера, $P_{cluster} \in [0,1]$.

Примечание: Очевидно, что для определения факта неоднородности окружающей среды необходимо максимизировать возможности обмена опытом между агентами МАС, изменив соответствующим образом определенные «гены» хромосомы (4.1).

Ниже рассматриваемый алгоритм представлен в виде псевдокода:

Алгоритм 4.3.

Задать S – множество состояний окружающей среды, известных агентам, $A_S = \{\alpha_i\}$ – множество агентов МАС S , $i = \overline{1, |A_S|}$, $iter_{max}$ – максимальное число итераций алгоритма, C – множество кластеров, J – множество значений

коэффициента Джини для кластеров множества $\mathbf{C}$, p – множество параметров разбиения состояний окружающей среды по группам, $S_E \leftarrow \emptyset$ – карта состояний окружающей среды E , $S \leftarrow \emptyset$ – множество возможных состояний окружающей среды E ;

Алгоритм_подраздел_4.1.5(A_S); //формирование хромосомы для всей МАС путем вызова алгоритма, представленного в подразделе 4.1.5

Пока ($iter \neq iter_{max}$) делать:

{

Для ($\forall \alpha_i \in A_S$) делать:

{

$S_E \leftarrow S_{\alpha_i}^{iter}$;

Если ($S_{\alpha_i}^{iter} \notin S$), то://формирование агентами карты состояний окружающей среды

{

$S \leftarrow S \cup \{S_{\alpha_i}^{iter}\}$; // $S_{\alpha_i}^{iter}$ – состояние агента α_i на итерации $iter$

}

}

$\mathbf{C}$ = Кластеризация_k-means (S, p); //определение возможных классов состояний

$\mathbf{C}$ = Кластеризация_k-means (S, S_E); //выделение кластеров на карте окружающей среды S_E

Для ($\forall \mathbf{C}_k \in \mathbf{C}, k = 1, |\mathbf{C}|$) делать://определение степени неоднородности полученных кластеров

{

$J \leftarrow J \cup \{\text{Коефф_Джини}(\mathbf{C}_k)\}$;

}

$J_{average} = \text{Средн_Арифм}(J)$;

Для ($\forall J_j \in J$) делать:

{

$\Delta J \leftarrow |J_{average} - J|$;

Если ($\Delta J > \delta$), то:

{

$is_cluster \leftarrow \text{Метод_рулетки}(P_{cluster})$;

Если ($is_cluster = true$), то:

{

$A_{C_j} \leftarrow \{\alpha_i | \alpha_i \in \mathbf{C}_k\}$;

```

    Алгоритм_подраздел_4.1.5( $A_{C_k}$ ); // формирование
хромосом для агентов, функционирующих в
отдельном кластере окружающей среды
}
}
}
}

```

В представленном выше псевдокоде для определения степени неоднородности кластеров используется «коэффициент Джини», вычисляемый по следующей формуле («формула Брауна») [7]:

$$J = 1 - \sum_{i=1}^N (X_k - X_{k-1})(Y_k - Y_{k-1}), \quad (4.18)$$

где X_k – «кумулятивная доля» состояний, имеющих класс C_j , $Y_k = \frac{n_{C_j} |S_{C_j}|}{|S_E|}$, S_{C_j} – множество состояний, имеющих класс C_j , S_E – карта состояний окружающей среды E , n_{C_j} – числовое обозначение класса C_j , $n_{C_j} = j$, $j = \overline{1, |C|}$.

В следующем подразделе приводится архитектура интеллектуальной системы оптимизации организационной структуры МАС, функционирование которой основано на алгоритме 4.2.1.

4.2.2 Архитектура интеллектуальной системы формирования и автоматической адаптации организационной структуры МАС

Архитектура интеллектуальной системы формирования и автоматической адаптации организационной структуры МАС иллюстрируется на рисунке 4.2.

На рисунке 4.2 представлены следующие элементы:

1 – база знаний, содержит кортежи

$$\langle E, T, H, f_{GQ}^T \rangle, \quad (4.19)$$

где E – параметры окружающей среды, T – поставленная задача, H – примененная хромосома, f_{GQ}^T – достигнутые показатели глобальной функции качества;

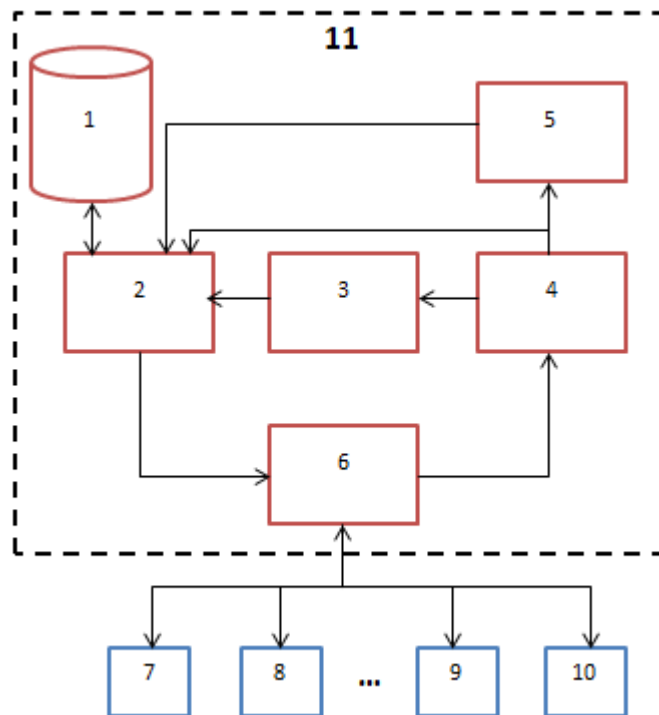


Рисунок 4.2 – Архитектура интеллектуальной системы формирования и автоматической адаптации организационной структуры МАС

2 – блок *формирования хромосомы*, реализуется на основании искусственной нейронной сети (ИНС) (4.15) с учетом текущего режима работы МАС (данные о режиме работы получаются от блока 3). В случае если база знаний 1 пуста (или имеющихся кортежей (4.19) недостаточно для обучения ИНС), сформированные хромосомы транслируются агентам (7–10) через блок связи 6. По результатам работы МАС примененная хромосома записывается в базу знаний 1, при этом данные ***E*** и ***T*** получаются от блока 4.

3 – блок *идентификации режима работы МАС* (основной или обучения), определяет текущий режим работы МАС путем мониторинга параметров окружающей среды, а также глобальной функции качества МАС (получаются от блока 4);

4 – блок *мониторинга параметров окружающей среды и МАС*, отслеживает изменения глобальной функции качества МАС, а также состояние окружающей среды;

5 – блок коррекции формата хромосомы. Определяет степень влияния отдельных «генов» хромосомы на глобальную функцию качества МАС. В результате корректирует формат хромосомы, исключая малозначащие «гены»;

6 – блок связи;

7–10 – агенты;

11 – сервер.

4.3 Описание параметров эксперимента

При подготовке текста п. 4.3 был использован авторский материал из опубликованных научных работ [52, 47].

4.3.1 Глобальная функция качества

Для режима обучения в качестве глобальной функции качества применим получаемое агентами усредненное вознаграждение по всем траекториям $\tau_k \in \tau$, $k = \overline{1, |\tau|}$ [95]:

$$R_{\tau} = \frac{R_k(\tau_k)}{|\tau|}, \quad (4.20)$$

где

$$R_k(\tau_k) = \sum_{i=1}^{|\tau_k|} \gamma(i)r_i, \quad (4.21)$$

где $\gamma(i)$ – «коэффициент обесценивания» [95, 45], $\gamma(i) \in [0,1]$, $\gamma_{min}(i) = 0$, $\gamma_{max}(i) = 1$, имеет экспоненциальный рост; r_i – величина вознаграждения, полученная за реализацию перехода.

Для основного режима глобальная функция качества будет представлять совокупность следующих показателей [69]: время выполнения поставленной задачи (T_{task}), вероятность достижения цели в заданные временные рамки (P_{target}), оценка вычислительной сложности алгоритма

($C_{algorithm}$). При этом перечисленные параметры неравнозначны и по степени важности распределены следующим образом:

$$P_{target} > T_{task} > C_{algorithm}.$$

Критерии перехода в режим обучения:

1) Существенное изменение параметров окружающей среды $\Delta E = E_t - E_{t-1}$:

$$\Delta E \rightarrow \Delta J = |J_{average} - J_t| > \delta_J, \quad (4.22)$$

где t – время функционирования МАС, $t = \overline{1, N_t}$, N_t – момент окончания работы МАС, J – «коэффициент Джини» [48] ($J_{average}$ – среднее значение на интервале $[0, t]$, J_t – текущее значение), δ_J – пороговое значение, превышение которого активирует режим обучения.

2) Существенное падение производительности МАС:

$$\Delta f_{GQ}^T = f_{GQ_t}^T - f_{GQ_{t-1}}^T < \delta_{f_{GQ}^T},$$

где $\delta_{f_{GQ}^T}$ – пороговое значение, преодоление которого активирует режим обучения, $\delta_{f_{GQ}^T} < 0$.

Критерий перехода в основной режим:

$$\Delta R_\tau < \delta_{R_\tau},$$

где δ_{R_τ} – пороговое значение, преодоление которого активирует основной режим, $\delta_{R_\tau} > 0$.

В качестве альтернативы «коэффициенту Джини» (возможный индикатор изменений в окружающей среде) может рассматриваться, например, индикатор снижения эффективности МАС, например, *среднее подкрепление* r_p , получаемое агентами a_i по всем траекториям p_{a_i} [66]. При этом под траекторией понимается множество «кортежей опыта», имеющих вид

$$s = \langle S_j, A_k, S_j', r \rangle,$$

сформированных агентом a_i к текущему моменту времени) (локальный показатель эффективности), где S_j – некоторое состояние $S_j \in S_E$, S_E –

множество возможных состояний агентов относительно окружающей среды E , $j = \overline{1, |S|}$, $S = f(E)$ (например, окружающая среда может быть разбита на множество проходимых или непроходимых клеток фиксированных размеров). После выполнения некоторого примитивного действия A_k (ехать вперед, назад, вправо, влево) состояние агента a_i меняется на S_j' . За осуществление указанного перехода агент получает подкрепление $r \in [-1, 1]$. При этом показатель r_p определяется по формуле:

$$r_p = \frac{\sum_{i=1}^{|A|} \left(\sum_{k=1}^{|p_{a_i}|} \gamma_{r_p}(k) r_k \right)}{|A|},$$

где $r_k \in S_k$, S_k – кортеж опыта, $S_k \in p_{a_i}$, γ_{r_p} – «коэффициент обесценивания» (позволяет снижать веса подкреплений, содержащихся в кортежах опыта, полученного достаточно давно), определяемый по формуле:

$$\gamma_{r_p}(k) = \begin{cases} \frac{1}{1 + e^{-\beta_{r_p} k}} & \text{при } k > \tau_0, \\ 0, & \text{при } k = \tau_0 \end{cases}$$

где $\beta_{r_p} \geq 0$ – «скорость обесценивания», τ_0 – время начала выполнения поставленной задачи T .

4.3.2 Описание параметров окружающей среды

Окружающая среда представляет собой лабиринт, имеющий следующие характеристики:

1) *Минимальная абстрактная единица* – квадратная клетка $p_{ij} = \{1|0\}$, где ij – пространственные координаты клетки, $i = \overline{1, N_i}$, $j = \overline{1, N_j}$, N_i – длина лабиринта, N_j – ширина лабиринта, $N_i \in \mathbb{N}$, $N_j \in \mathbb{N}$, при $p_{ij} = 1$ клетка считается проходимой, в ином случае – непроходимой.

2) *Доля непроходимых клеток* ($p_{p_{ij}=0}$), %.

3) Множество соседних клеток для p_{ij} - $\Gamma_{p_{ij}}$. При этом под соседней понимается клетка $p_{i'j'}$: $i' = \{i|i-1|i+1\}$, $j' = \{j|j-1|j+1\}$, $i' = i \rightarrow j' \neq j$, $j' = j \rightarrow i' \neq i$. $|\Gamma_{p_{ij}}| = 4$.

4) Вероятность изменения клеткой p_{ij} своего состояния на противоположное ($P_{\neg p_{ij}}$) на интервале времени $t_{\neg p_{ij}}$.

Примечание: Первоначальная карта лабиринта генерируется случайным образом. Новая итерация наследует карту лабиринта в состоянии, имевшем место в конце предыдущей итерации. При этом под *итерацией* понимается процесс решения агентами поставленной задачи до момента достижения конечного результата (цели задачи).

Примечание: При формировании лабиринта не допускается наличие замкнутых контуров, отделенных от основной части лабиринта:

$$\forall p_{ij}: |\Gamma_{p_{ij}}^{imp}| \leq 3, \quad (4.23)$$

где $\Gamma_{p_{ij}}^{imp} = \{p_{ij} | p_{ij} = 0\}$.

4.3.3 Описание параметров МАС

МАС представляет множество агентов, имеющих колесную транспортную платформу, подсистему машинного зрения, систему связи, вычислительную систему, обладает следующими характеристиками:

- 1) Число агентов МАС - $|A_S|$.
- 2) Пространственные координаты агентов α_i ($\langle i, j \rangle$).

Примечание: Координаты $\langle i_{t_0}, j_{t_0} \rangle$ генерируются случайным образом на каждой итерации, $\alpha_i \in A_S$, t_0 – время начала функционирования МАС.

- 3) Скорость передвижения агента (v_{α_i}), клеток / с.

4) Производительность вычислительной системы агента (P_{α_i}), операций / с. При этом под *операцией* (O_j) будем понимать одно из перечисленных действий (перечислены в порядке убывания приоритета):

а) Арбитраж коллизий (операция выполняется агентом-подчиненным α_i или агентом-координатором $\alpha_{i_{coord}}$ в зависимости от значения параметра T_{coll} – «приоритет в арбитраже коллизий»).

б) Назначение агенту-подчиненному α_i задачи T_k для выполнения, $T_k \subset T$, $k = \overline{1, |T|}$, T – глобальная задача, поставленная МАС в целом.

в) Назначение агенту-подчиненному α_i вознаграждения r по результатам выполнения поставленной задачи T_k .

г) Принятие решения о выполнении перехода $S_t \xrightarrow{A} S_{t+1}$, где t – текущий момент времени, S_t – текущее состояние, S_{t+1} – последующее состояние, A – выполняемое действие.

Примечание: Для параметра P_{α_i} определено следующее ограничение:

$$\forall \alpha_i: P_{\alpha_i} \in [P_{\alpha_{i_{min}}}, P_{\alpha_{i_{max}}}] \wedge ((P_{\alpha_i} = P_{\alpha_{i+1}}) \vee (P_{\alpha_i} \neq P_{\alpha_{i+1}})), \quad (4.24)$$

т.е. во множестве A_S могут присутствовать агенты, имеющие различные показатели производительности.

5) *Емкость аккумуляторной батареи* (E_{α_i}), условных единиц. Для данного параметра также имеет место условие, аналогичное (4.24).

6) *Уровень энергопотребления* (условных единиц / операция) – кортежи $\langle O_j, E_{O_j} \rangle$, где E_{O_j} – объем энергии, потребляемой агентом при выполнении операции O_j , $j = \overline{1, |O|}$, O – множество операций, допустимых для исполнения.

7) *Текущий уровень заряда аккумуляторной батареи* ($E_{\alpha_{i_t}}$), где t – текущий момент времени, $E_{\alpha_{i_t}} < E_{\alpha_i}$.

8) *Время подзарядки* t_{boost} – время, которое необходимо агенту для повышения текущего уровня заряда аккумуляторной батареи E_{α_i} на одну условную единицу (для подзарядки определяются специальные базовые клетки p_{ij}^{base}).

9) Число базовых клеток $N_{p^{base}}$. Под базой (или базовой клеткой) p_{ij}^{base} понимается некоторая клетка лабиринта, при нахождении в которой агент повышает уровень заряда собственной аккумуляторной батареи.

Примечание: Для агентов множества A_S информация о технических характеристиках друг друга является неизвестной.

4.3.4 Описание решаемой задачи

Рассматриваемая задача заключается в поиске целевых состояний в лабиринте (задача мониторинга) и имеет следующие параметры:

1) Общее число генерируемых целевых состояний (N_{targ}).

2) Число целевых состояний, существующих одновременно ($N_{targ,t}$),

где t – текущий момент времени, $N_{targ,t} < N_{targ}$.

Примечание: Под целевым состоянием понимается некоторая клетка $p_{ij} = 1$ (являющаяся проходимой). Пространственные координаты целевых состояний выбираются случайным образом, при этом $\forall (p_{ij}^{targ}, p_{xy}^{targ} | p_{ij}^{targ} \in p_t, p_{xy}^{targ} \in p_t): i \neq x \wedge j \neq y$, где $x = \overline{1, N_i}, y = \overline{1, N_j}$, p_t – множество целевых состояний, находящихся в лабиринте в момент времени t , $|p_t| = N_{targ,t}$.

3) Вероятность появления целевого состояния в некотором регионе лабиринта P_{targ}^{ireg} , где i_{reg} – номер региона. При этом под регионом понимается некоторая квадратная или прямоугольная область лабиринта $R_{ireg} = \{p_{ij} | i = \overline{1, N_{len}}, j = \overline{1, N_{width}}, |p_1| \geq 1\}$, p_1 – множество проходимых клеток, $i_{reg} = \overline{1, |R|}$, R – множество регионов лабиринта, $\forall R_{ireg}: N_{len} = const, N_j = const$, $p_1: p_1 = \{p_{ij} | p_{ij} = 1\} \wedge p_1 \subset R_{ireg}$ (т.е. регион должен включать хотя бы одну проходимую клетку), N_{len} – длина региона, N_{width} – ширина региона.

Примечание: Клетки равномерно распределяются по региону.

4) Вероятность изменения параметра «доля непроходимых клеток» ($P_{p_{ij}=0}$).

5) Вероятность пересчета параметра $P_{targ}^{i_{reg}}$ для всех регионов P_{reg} .

6) Глобальная цель задачи: достижение всех N_{targ} целевых состояний.

Процесс выполнения поставленной задачи состоит в выполнении следующего алгоритма:

1) Размещение целевых состояний в лабиринте (множество p_t) с учетом $P_{targ}^{i_{reg}}$, $p_{targ}^T \leftarrow p_t$, где p_{targ}^T – множество всех целевых состояний, сгенерированных за время решения задачи T .

2) Размещение в лабиринте агентов $\alpha_i \in A_S$.

3) Поиск агентами α_i целевых состояний $p_{ij}^{targ} \in p_t$. Если некоторое p_{ij}^{targ} , то $p_t \leftarrow p_t \setminus p_{ij}^{targ}$, $p_{found}^T \leftarrow p_{found}^T \cup p_{ij}^{targ}$, где p_{found}^T – множество обнаруженных целевых состояний.

4) Если $|p_t| < N_{targ,t} \wedge |p_{targ}^T| < N_{targ}$, то сгенерировать координаты нового целевого состояния p_{ij}^{targ} , $p_t \leftarrow p_t \cup p_{ij}^{targ}$, $p_{targ}^T \leftarrow p_t$.

5) Если $E_{\alpha_i} = \theta_E$, то агент α_i автоматически отправляется на базу для подзарядки, при этом под базой понимается некоторая клетка p_{ij}^{base} , координаты которой определяются случайным образом в начале выполнения задачи T .

6) Повторение шагов 3–5, пока не достигнута глобальная цель задачи - $|p_{found}^T| = N_{targ}$.

4.4 Описание результатов эксперимента

При подготовке текста п. 4.4 был использован авторский материал из опубликованных научных работ [52, 57, 47].

Компьютерное моделирование производилось с помощью программного обеспечения, реализованного в среде Unity [96] 2019 на базе

ПК со следующими характеристиками: процессор Intel(R) Core(TM) i3-8145U CPU @ 2.10GHz 2.30 GHz; ОЗУ 8,00 ГБ.

Приняты следующие значения параметров эксперимента:

1) Параметры окружающей среды: $N_i = 100$ клеток (длина лабиринта), $N_j = 100$ клеток (ширина лабиринта), $p_{p_{ij}=0} = 70\%$ (доля непроходимых клеток), $P_{\neg p_{ij}} = 0.3$ для $t_{\neg p_{ij}} \in \left[t_0, \frac{t_{end}}{4}\right]$, $P_{\neg p_{ij}} = 0.1$ для $t_{\neg p_{ij}} \in \left[\frac{t_{end}}{4}, \frac{t_{end}}{2}\right]$, $P_{\neg p_{ij}} = 0.5$ для $t_{\neg p_{ij}} \in \left[\frac{t_{end}}{2}, \frac{3t_{end}}{4}\right]$, $P_{\neg p_{ij}} = 0.1$ для $t_{\neg p_{ij}} \in \left[\frac{3t_{end}}{4}, t_{end}\right]$, где t_0 – время начала работы МАС, t_{end} – время окончания работы МАС (применяется усредненное значение по всем итерациям), $P_{p_{ij}=0} = 0.3$ (вероятность изменения параметра «доля непроходимых клеток» на новой итерации).

При расчете «коэффициента Джини» [48] применялась следующая методика:

а) Выделение следующих групп клеток по показателю проходимости: $\mathbf{g}_1: p_{ij}(t-1) = \emptyset \rightarrow p_{ij}(t) = 0$ (изменение состояния от неизвестного к непроходимому), $\mathbf{g}_2: p_{ij}(t-1) = \emptyset \rightarrow p_{ij}(t) = 1$ (изменение состояния от неизвестного к проходимому), $\mathbf{g}_3: p_{ij}(t-1) = 1 \rightarrow p_{ij}(t) = 0$ (изменение состояния от проходимого к непроходимому), $\mathbf{g}_4: p_{ij}(t-1) = 0 \rightarrow p_{ij}(t) = 1$ (изменение состояния от непроходимого к проходимому), $\mathbf{g}_5: p_{ij} = const$ (состояние не изменилось).

б) Выполнение классификации клеток p_{ij} по группам $\mathbf{g}_k$, т.е. формирование кортежей

$$\langle \mathbf{g}_k, N_{p_{ij}}^{g_k} \rangle, \quad (4.25)$$

где $N_{p_{ij}}^{g_k}$ – число клеток, принадлежащих группе $\mathbf{g}_k$, $k = \overline{1,5}$.

в) Добавление в кортежи (4.25) параметра $W_{g_k} = k_{g_k} N_{p_{ij}}^{g_k}$, где k_{g_k} – некоторое числовое значение, поставленное группе $\mathbf{g}_k$, $\forall \langle \mathbf{g}_k, \mathbf{g}_{k+1} \rangle: k_{g_k} \neq k_{g_{k+1}}, k_{g_k} \neq 0$.

г) Определение величины «коэффициента Джини» [48] для временного интервала $t - J_t$.

2) Параметры МАС: $P_{\alpha_{i_{sub}}} = 2$ операции / с. (производительность вычислительной системы агента-подчиненного), $P_{\alpha_{i_{coord}}} = 20$ операций / с. (производительность вычислительной системы агента-координатора); $|A_S| = 40$ (число агентов МАС), из них 3 агента-координатора (имеют вычислительную систему с производительностью $P_{\alpha_{i_{coord}}}$), 25 агентов-подчиненных (имеют вычислительную систему с производительностью $P_{\alpha_{i_{sub}}}$), 12 агентов имеют производительность вычислительной системы в интервале $(P_{\alpha_{i_{sub}}}, P_{\alpha_{i_{coord}}})$, т.е. могут быть как агентами-подчиненными, так и агентами-координаторами; $v_{\alpha_i} = 1$ клетка / с. (скорость передвижения агента), $E_{\alpha_i} = \frac{P_{\alpha_i} N_{targ} N_j N_i (1 - p_{p_{ij}=0})}{|A_S|}$ условных единиц (емкость аккумуляторной батареи агента), уровень энергопотребления – 1 условная единица / операция, $t_{boost} = 0,1$ с. (время подзарядки на 1 условную единицу), $N_{p^{base}} = 10$ (число базовых клеток – пунктов подзарядки),

3) Параметры решаемой задачи: $N_{targ} = 100$ (общее число генерируемых целевых состояний), $N_{targ,t} = 5$ (число целевых состояний, существующих одновременно); $P_{targ}^{i_{reg}}$ (вероятность появления целевого состояния в некотором регионе лабиринта) – задается случайным образом в начале эксперимента, может быть изменена на новой итерации с вероятностью P_{reg} , $N_{len} = 10$, $N_{width} = 10$.

Примечание: Резкий рост вероятности изменения значения клетки (параметр $P_{-p_{ij}}$) на 3-м временном интервале процесса решения поставленной задачи, а также параметры $P_{p_{ij}=0}$ и P_{reg} позволяют оценить адаптационные возможности МАС.

На рисунке 4.3 приводится пример сформированного лабиринта. Окружностями выделены агенты, ромбами – целевые состояния.

Результаты применения предложенной интеллектуальной системы оптимизации организационной структуры МАС (см. подраздел 4.2.2) по показателю эффективности «время выполнения поставленной задачи» (T_{task}) представлены на рисунке 4.4. Показатель фиксировался для всего периода решения задачи на интервале $t_0 \dots t_{end}$, а не только для основного режима; на графике показана эффективность лучшей хромосомы популяции на каждой итерации.

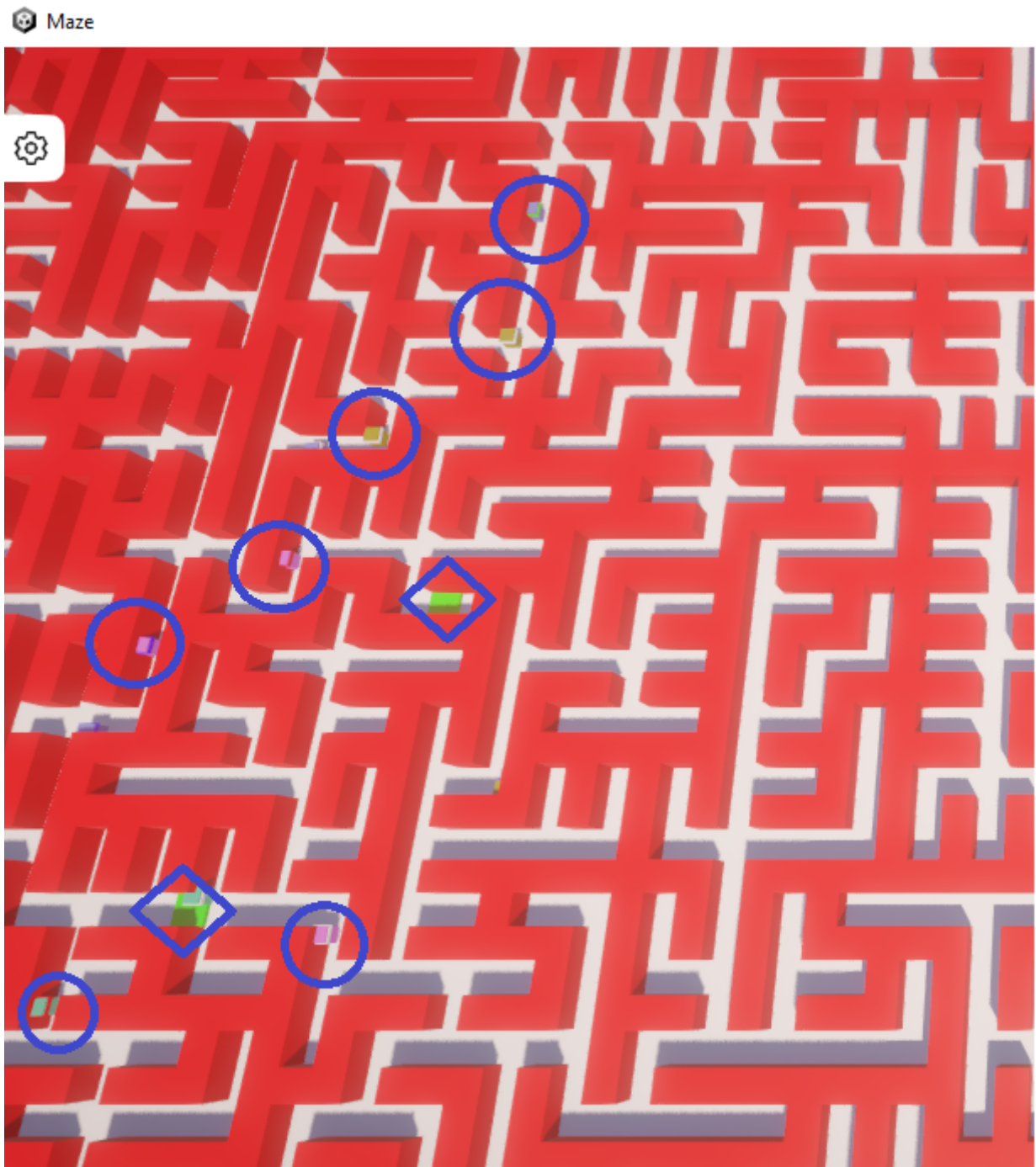


Рисунок 4.3 – Пример сформированного лабиринта

В качестве аналога приняты методы, представленные в работах [16, 40].

В работе [16] предложены методы и алгоритмы, позволяющие выполнять динамическое перераспределение задач от «дефектных» агентов к полнофункциональным (модифицируется организационная структура без изменения архитектуры МАС). При этом изменение структуры системы выполняется «агентом-мастером организации», перераспределение задач производится «агентом обслуживания».

В работе [40] предлагается представлять МАС в виде набора иерархических деревьев, набор которых преобразуется в массив фиксированной длины с целыми числами – «хромосома». Далее производится определение оптимального набора деревьев с помощью «генетического алгоритма» (организационная структура может быть изменена только на следующей итерации).

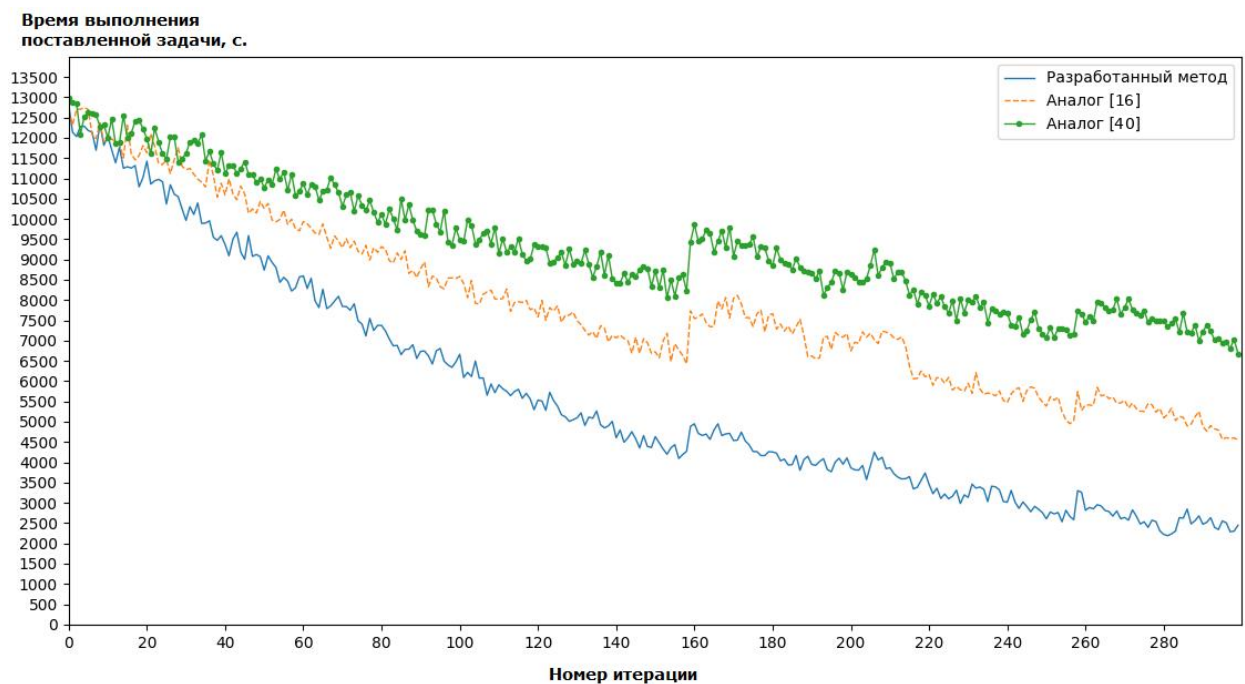


Рисунок 4.4 – Результаты применения предложенной интеллектуальной системы оптимизации организационной структуры МАС по показателю эффективности «время выполнения поставленной задачи» (T_{task})

По результатам экспериментальной оценки, представленным на рисунке 4.4, могут быть сделаны следующие выводы:

1) Имеет место большое количество резких скачков (например, на итерациях 160, 210, 260), которое обусловлено резким изменением показателя $p_{p_{ij}=0}$ (доля непроходимых клеток) и, как следствие, существенной перестройкой карты, которая становится менее знакомой для агентов. При этом МАС, основанная на разработанном методе и алгоритмах (далее МАС-3), показала большую устойчивость к изменению структуры лабиринта, чем МАС, основанные на методах, представленных в [16] (далее – МАС [16]) и [40] (далее – МАС [40]). При этом динамическая организационная структура МАС [16] показала лучшие результаты, чем статическая организационная структура МАС [40].

2) Более высокая скорость сходимости, характерная для МАС-3, уже явно прослеживается на промежутке до 100-й итерации (на 100-й итерации МАС-3 опережает МАС [16] по величине этого показателя на 25 %).

Обозначенные преимущества МАС-3 достигнуты за счет большей адаптивности, которая выражается в следующем:

1) Хромосома, оптимизируемая с помощью предложенного зоосоциального метода, содержит некоторый образ (описание архитектуры), который затем раскрывается в виде организационной структуры МАС с учетом характеристик поставленной задачи **T** и окружающей среды **E**. При этом в работах [16, 40] выполняется адаптация исключительно организационной структуры, базовые правила формирования и функционирования МАС (архитектурные параметры) не изменяются. Недостаток подобного подхода отчетливо виден на рисунке 4.4 на 160-й итерации (событие «существенное изменение параметров окружающей среды») – при падении эффективности МАС-3 (динамическая адаптация архитектуры) на 16 % показатели МАС [16] и МАС [40] снизились, на 23 % и 23,5 % соответственно.

2) Предложенный зоосоциальный метод охватывает существенно большее число параметров, влияющих на производительность МАС, чем

методы, представленные в [16, 40]. Это обстоятельство делает его более гибким в сравнении с аналогами.

Примечание: Значения показателя T_{task} (время выполнения поставленной задачи), представленные на рисунке 4.4., соответствуют величине функции приспособленности лучшей хромосомы в популяции.

Разработанный подход к представлению архитектуры МАС позволяет моделировать следующие типы организационных структур (далее указан минимальный набор показателей, необходимых для моделирования той или иной организационной структуры) [57]:

1) «Коалиция» при $P_{horizontal} > 0$ (вероятность формирования горизонтальных связей), $P_{subor} = 0$ (вероятность установления агентом вертикальных связей, в которых он будет являться подчиненной стороной), $R_{terr} < R_{social}$ (радиус индивидуальной дистанции между агентами должен быть меньше радиуса «социальной зоны»).

2) «Команда» при $P_{subor} > 0$, $P_{free} < 1$ (вероятность разрыва подчиненным агентом вертикальной связи), $R_{terr} < R_{public}$ (радиус индивидуальной дистанции между агентами должен быть меньше радиуса «публичной зоны»), $\varphi_k \rightarrow \infty$ и $\vartheta_k \rightarrow \infty$ при $k \neq 1$ (запрет на формирование нескольких уровней иерархии, т.е. в «команде» будут представлены только агенты-координаторы 1-го уровня иерархии).

3) «Иерархическая структура»: $P_{subor} > 0$, $P_{free} < 1$, $R_{terr} < R_{public}$, $\varphi_k < \varphi_{max}$ и $\vartheta_k < \vartheta_{max}$, где φ_{max} – максимально значение показателя «престижа», достижимое для агентов, ϑ_{max} – максимально возможное число агентов в группе, $k = \overline{1, N}$, N – максимально достижимое число уровней иерархии.

4) «Федерация»: $P_{horizontal} = 0$, $R_{terr} = R_{social}$ (запрет горизонтальных связей), $P_{subor} > 0$, $P_{free} < 1$, $R_{terr} < R_{public}$.

5) «Конгрегация»: $P_{subor} > 0$, $P_{free} < 1$, $R_{terr} < R_{public}$, $\varphi_k < \varphi_{max}$ и $\vartheta_k < \vartheta_{max}$, $P_{free} = 0$ (сформированные вертикальные связи являются постоянными).

При этом переход от одной организационной структуры к другой может быть выполнен путем изменения значений соответствующих показателей.

4.5 Экспериментальная оценка эффективности применения предложенного метода и алгоритмов при решении задачи доставки малогабаритных грузов

С целью расширения области экспериментальной оценки эффективности предложенного метода и алгоритмов за пределы задачи поиска целевых состояний (задача А, подраздел 2.1) дополнительно выполнена серия экспериментов по решению задачи доставки малогабаритных грузов. Серия реализована в той же программной модели, разработанной в среде Unity 2019 (см. подраздел 4.4), путем модификации глобальной функции качества и правил генерации целевых состояний: целевые состояния интерпретируются как пункты приема грузов, а каждый заказ – как пара «целевое состояние – пункт отправления (база)». Данный подход позволяет использовать все параметры окружающей среды и МАС, принятые в подразделе 4.3, без модификации методического аппарата.

Заказ считается выполненным, если соответствующий груз доставлен в целевое состояние до истечения времени жизни заказа. В качестве показателя эффективности используется «коэффициент выполнения заказов» [28]:

$$K = \frac{N_{\text{в}}}{N_{\text{общ}}},$$

где $N_{\text{в}}$ – число заказов, доставленных до истечения времени их жизни; $N_{\text{общ}}$ – общее число заказов, поступивших в МАС за время решения поставленной задачи.

Приняты следующие значения параметров серии:

- 1) лабиринт разбит на 4 региона; на каждой итерации генерируется пакет из 24 заказов, распределение заказов по регионам задается случайным образом и может изменяться на новой итерации с вероятностью 0,15;
- 2) время жизни заказа – 40 условных тактов; заказы считаются малогабаритными (один агент транспортирует один заказ без дополнительных операций погрузки-разгрузки);
- 3) грузоподъемность и энергообеспечение агентов соответствуют подразделу 4.3 (скорость передвижения агента, емкость аккумуляторной батареи, число пунктов подзарядки).

Совокупность изменяющихся параметров (доля непроходимых клеток – итерации 160, 210, 260; распределение заказов по регионам) позволяет оценить адаптационные возможности МАС в условиях двух типов нестационарности: нестационарности окружающей среды и нестационарности потока задач.

Результаты экспериментальной оценки по показателю «коэффициент выполнения заказов» представлены на рисунке 4.5 (на графике представлено усредненное значение по 5 запускам модели).

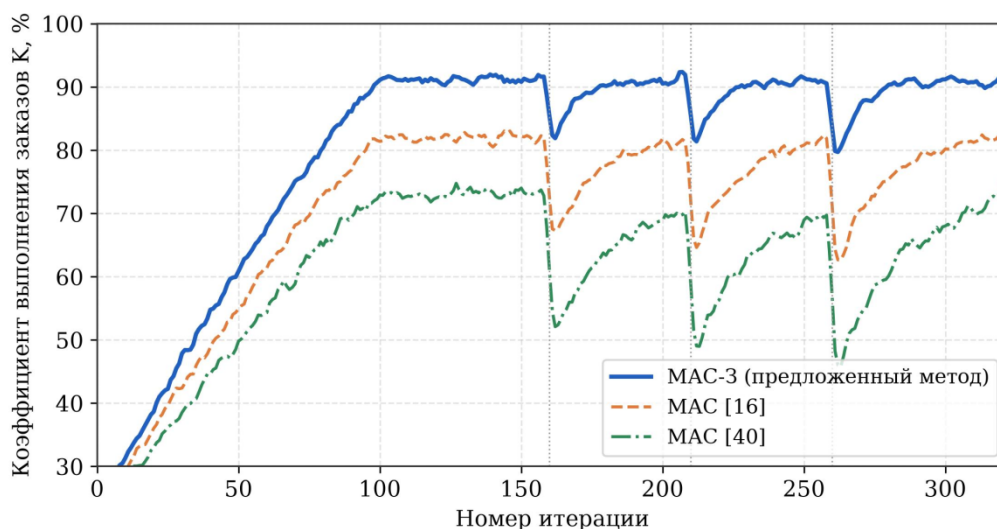


Рисунок 4.5 – Результаты применения предложенной интеллектуальной системы оптимизации организационной структуры МАС по показателю эффективности «коэффициент выполнения заказов»

По результатам экспериментальной оценки, представленным на рисунке 4.5, могут быть сделаны следующие выводы:

1) На стационарных временных интервалах МАС-3 выходит на значение показателя $K \approx 0,91$, что превышает результаты аналогов: $K \approx 0,82$ у МАС [16] и $K \approx 0,73$ у МАС [40]. Преимущество МАС-3 сохраняется во всем диапазоне изменения параметров окружающей среды.

2) Скачки показателя K на итерациях 160, 210, 260 коррелируют с изменением доли непроходимых клеток и перераспределением потока заказов по регионам. При этом глубина просадки показателя у МАС-3 минимальна (до 13–19 %), а восстановление достигнутого уровня происходит в среднем за 20–25 итераций против 40–50 итераций у МАС [16] и более 60 итераций у МАС [40]. Это обусловлено возможностью автоматического изменения базовых правил (архитектуры): МАС-3 оперативно перераспределяет подзадачи между регионами лабиринта, тогда как МАС [40] способна изменить организационную структуру только на следующей итерации, что приводит к накоплению невыполненных заказов.

Дополнительно выполнена серия экспериментов с имитацией выхода из строя части агентов в ходе решения задачи: в середине основного режима работы МАС (итерация 100) из состава множества агентов выводилось от 0 до 10 агентов из 40. Результаты представлены на рисунке 4.6.

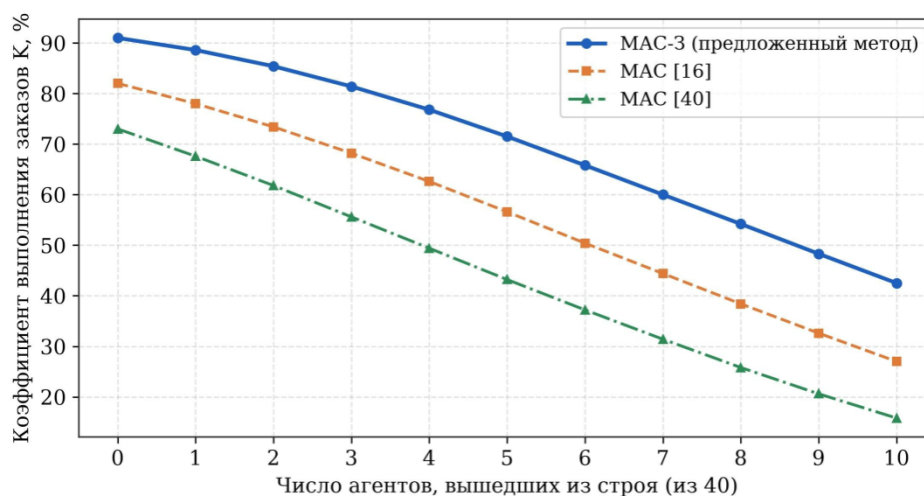


Рисунок 4.6 – Зависимость коэффициента выполнения заказов от числа агентов, вышедших из строя в ходе решения задачи

По результатам данной серии могут быть сделаны следующие выводы:

1) Снижение значения коэффициента K при нарастании числа отказов у МАС-3 носит практически линейный характер: в среднем около 2,4 п.п. на один вышедший из строя агент. Для МАС [16] и МАС [40] темп снижения ускоряется по мере увеличения числа отказов. При выходе из строя 10 из 40 агентов МАС-3 сохраняет значение $K \approx 0,66$ против $K \approx 0,53$ у МАС [16] и $K \approx 0,38$ у МАС [40], что соответствует преимуществу МАС-3 по данному показателю в 0,13–0,28.

2) Устойчивость МАС-3 к отказам агентов обусловлена тем, что после выхода агентов из строя происходит автоматическое перераспределение задач между оставшимися агентами и смена ролей агентов, обладающих промежуточной производительностью вычислительной системы (см. подраздел 4.3: 12 из 40 агентов способны выполнять роли, как агентов-подчиненных, так и агентов-координаторов). В МАС [16] компенсация ограничена перераспределением задач, выполняемым «агентом обслуживания», а в МАС [40] перестройка иерархических деревьев возможна только на следующей итерации, что приводит к временной потере управляемости распределением заказов.

Совокупный вывод по разделу: результаты подтверждают применимость предложенного зоосоциального метода и разработанного алгоритма к прикладной задаче из заявленной области практического использования (доставка малогабаритных грузов) и демонстрируют сохранение преимуществ МАС-3 при изменении типа решаемой задачи без модификации методического аппарата.

4.6 Выводы

При подготовке текста п. 4.5 был использован авторский материал из опубликованных научных работ [47].

1) Предложен алгоритм адаптации организационной структуры МАС на макроуровне, основанный на зоосоциальном методе, отличающийся

- применением генетического алгоритма для генерации множества хромосом, используемого для обучения ИНС;

- применением ИНС для определения хромосомы, позволяющей получить лучший результат для текущей пары «поставленная задача»-«окружающая среда»;

- выполнением аппроксимации «функции стратегии» (представлена ИНС) с помощью алгоритма «REINFORCE»;

- способностью к работе в условиях неоднородной окружающей среды.

2) Предложена архитектура интеллектуальной системы автоматической адаптации организационной структуры МАС, базирующейся на разработанном алгоритме оптимизации организационной структуры МАС на макроуровне, основанном на зоосоциальном методе.

3) Выполнено описание параметров эксперимента: глобальная функция качества, параметры окружающей среды, МАС, решаемой задачи.

4) Выполнена экспериментальная оценка разработанного алгоритма оптимизации организационной структуры МАС на макроуровне, основанного на зоосоциальном методе, по результатам которой сделаны следующие выводы:

- а) Имеет место большое количество резких скачков (например, на итерациях 160, 210, 260), которое обусловлено резким изменением показателя $p_{p_{ij}=0}$ (доля непроходимых клеток) и, как следствие, существенной перестройкой карты, которая становится менее знакомой для агентов. При этом МАС, основанная на разработанных методах и алгоритмах (далее МАС-3), показала большую устойчивость к изменению структуры

лабиринта, чем МАС, основанные на методах, представленных в МАС [16] и МАС [40]. При этом динамическая организационная структура МАС [16] показала лучшие результаты, чем статическая организационная структура МАС [40].

б) Более высокая скорость сходимости, характерная для МАС-3, уже явно прослеживается на промежутке до 100-й итерации (на 100-й итерации МАС-3 опережает МАС [16] по величине этого показателя на 25 %).

5) Обозначенные преимущества МАС-3 достигнуты за счет большей адаптивности, которая выражается в следующем:

а) хромосома, оптимизируемая с помощью предложенного зоосоциального метода, содержит некоторый образ (описание архитектуры), который затем раскрывается в виде организационной структуры МАС с учетом характеристик поставленной задачи **Т** и окружающей среды **Е**. При этом в работах [16, 40] выполняется адаптация исключительно организационной структуры, базовые правила формирования и функционирования МАС (архитектурные параметры) не изменяются. Недостаток подобного подхода отчетливо виден на рисунке 4.4 на 160-й итерации (событие «существенное изменение параметров окружающей среды») – при падении эффективности МАС-3 (динамическая адаптация архитектуры) на 16 % показатели МАС [16] и МАС [40] снизились, на 23 % и 23,5 % соответственно.

б) Предложенный зоосоциальный метод охватывает существенно большее число параметров, влияющих на производительность МАС, чем методы, представленные в [16, 40]. Это обстоятельство делает его более гибким в сравнении с аналогами.

б) Разработанный подход к представлению архитектуры МАС позволяет моделировать следующие типы организационных структур: «коалиция», «команда», «иерархическая структура», «федерация», «конгрегация».

ЗАКЛЮЧЕНИЕ

В диссертационной работе получены следующие основные научные и практические результаты:

1) Выполнен анализ структуры интеллектуального агента, иерархической централизованной МАС. Выполнен системный анализ способов организации совместных действий для множества интеллектуальных агентов.

2) Предложены подходы к управлению организационной структурой МАС на макро- и микроуровнях:

а) автоматическая адаптация архитектуры и организационной структуры МАС на макроуровне основывается на модификации множества правил поведения агентов и функционирования МАС. Подход характеризуется возможностью воздействия на организационную структуру МАС через множество параметров, являющихся отображением множества допустимых правил, регламентирующих функционирование МАС, координацию действий агентов, её составляющих, взаимодействие с окружающей средой.

б) автоматическая адаптация архитектуры и организационной структуры МАС на микроуровне основывается на модификации иерархического атрибутированного ориентированного графа (ИАОГ) МАС. Подход характеризуется выполнением декомпозиции множества агентов на микрокластеры, что позволяет минимизировать затраты машинного времени на определение оптимальной структуры ИАОГ.

3) Предложен зоосоциальный метод автоматической адаптации организационной структуры МАС, предусматривающий реализацию множества правил поведения агентов и функционирования МАС в соответствии с принципами, характерными для социальных животных, при этом для определения оптимальных параметров МАС применяется генетический алгоритм. Практическая эффективность метода подтверждена

внедрением в ООО «Аэромар-Краснодар»: за счёт автоматической перестройки иерархии агентов и связей между ними при изменении условий среднее отклонение от целевого времени доставки сократилось на 34,8 % (с 23 до 15 минут).

4) Предложен алгоритм адаптации организационной структуры МАС на макроуровне, включающий генерацию множества хромосом генетическим алгоритмом для обучения искусственной нейронной сети и аппроксимацию функции стратегии алгоритмом REINFORCE. Алгоритм внедрён в ООО «Олимп-связь» в составе системы управления сетевыми ресурсами и мониторинга телекоммуникационной инфраструктуры: за счёт динамического изменения топологии связей и ролей агентов под текущую нагрузку и состояние сети оптимизированы маршруты передачи данных и повышена самоорганизация агентов при изменении условий работы сети.

5) Предложена архитектура интеллектуальной системы автоматической адаптации организационной структуры МАС, базирующейся на разработанном алгоритме оптимизации организационной структуры МАС на макроуровне, основанном на зоосоциальном методе.

6) Выполнена экспериментальная оценка разработанного метода и алгоритма адаптации организационной структуры МАС. Система продемонстрировала высокую скорость сходимости, опережая базовые методы адаптации по показателю эффективности на 25 %, а также высокую устойчивость к резким изменениям параметров внешней среды: при существенной перестройке условий её эффективность снижалась не более чем на 16 %, тогда как у базовых методов падение достигало 23,5 %).

7) Разработанный подход к представлению архитектуры МАС позволяет одновременно моделировать следующие типы организационных структур: «коалиция», «команда», «иерархическая структура», «федерация», «конгрегация».

Словарь терминов

Архитектура МАС – множество базовых правил, определяющих процессы формирования и дальнейшего функционирования МАС.

Организационная структура МАС – ориентированный граф с агентами в качестве вершин, а также связями между ними (горизонтальные, вертикальные) в качестве ребер. Формируется на основании базовых правил, определяемых архитектурой МАС.

Коалиция – тип организационной структуры, предполагающий объединение равноправных агентов, целью которых является максимизация индивидуальных, а не групповых выигрышей.

Команда – тип организационной структуры, предполагающий объединение равноправных агентов, целью которых является максимизация группового выигрыша.

Иерархическая структура - многоуровневая иерархическая организационная структура.

Федерация – тип организационной структуры, предполагающий выделение специальных агентов-делегатов, которым остальные члены группы уступают часть своих прав.

Конгрегация - в отличие от «коалиции» и «команды» предполагает создание долговременных вертикальных связей без привязки к какой-либо конкретной задаче.

Автоматическая адаптация архитектуры и организационной структуры МАС на макроуровне – подход, предполагающий адаптационное воздействие на организационную структуру путем модификации множества правил поведения агентов и функционирования МАС (применяется в данной работе).

Автоматическая адаптация архитектуры и организационной структуры МАС на микроуровне – подход, предполагающий адаптационное воздействие на организационную структуру МАС путем

модификации её иерархического ориентированного графа (агенты – вершины, связи между агентами - ребра).

Хромосома – структурированный набор параметров, описывающих архитектуру МАС. Состоит из генов, каждый из которых представляет отдельную переменную или параметр некоторого базового правила.

REINFORCE – алгоритм, определенный в рамках парадигмы обучения с подкреплением. Выполняет оптимизацию политики (стратегии) агента путем максимизации ожидаемой суммарной награды.

Список использованных источников

1. Ahmadoun D. Decentralized coalition structure formation for interdependent tasks allocation / D. Ahmadoun, E. Bonzon, C. Buron, P. Moraitis, P. Savéant, O. Shehory // 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI). – 2021. – P. 73–80. – Текст : электронный. – DOI: 10.1109/ICTAI52525.2021.00018. – URL: <https://doi.org/10.1109/ICTAI52525.2021.00018> (дата обращения: 10.02.2025).
2. Alkam T. Reinforcement learning in artificial intelligence and neurobiology / T. Alkam, A. H. Van Benschoten, E. Tarshizi // Neuroscience Informatics. – 2025. – Vol. 5, № 3. – Текст : электронный. – URL: <https://www.sciencedirect.com/science/article/pii/S2772528625000354> (дата обращения: 14.04.2026).
3. Amaral C. J. From goals to organisations: automated organisation generator for MAS / C. J. Amaral, J. F. Hübner // Engineering Multi-Agent Systems: 7th International Workshop (EMAS 2019) : Lecture Notes in Computer Science. – Vol. 12058. – Cham : Springer, 2020. – P. 25–42. – Текст : электронный. – DOI: 10.1007/978-3-030-51417-4_2. – URL: https://doi.org/10.1007/978-3-030-51417-4_2 (дата обращения: 15.03.2024).
4. Bernon C. The ADELFE methodology for an intranet system design / C. Bernon, M.-P. Gleizes, G. Picard, P. Glize // Agent-Oriented Information Systems (AOIS'02) at CAiSE'02. – Toronto, 2002. – Текст : электронный. – URL: <https://ceur-ws.org/Vol-57/id-6.pdf> (дата обращения: 14.03.2024).
5. Bouslimi I. The scalability impact on organization-based multi-agent systems / I. Bouslimi // Procedia Computer Science. – 2024. – Vol. 246. – P. 991–1000.
6. Broersen J. The BOID architecture: conflicts between beliefs, obligations, intentions and desires / J. Broersen, M. Dastani, J. Hulstijn, Z. Huang // Proceedings of the Fifth International Conference on Autonomous Agents (AGENTS'01). – 2001. – P. 9–16. – Текст : электронный. – DOI:

10.1145/375735.375766. – URL: <https://doi.org/10.1145/375735.375766> (дата обращения: 09.02.2024).

7. Brown M. C. Using Gini-style indices to evaluate the spatial patterns of health practitioners: theoretical considerations and an application based on Alberta data / M. C. Brown // *Social Science & Medicine*. – 1994. – Vol. 38, № 9. – P. 1243–1256.

8. Capezzuto L. Anytime and efficient coalition formation with spatial and temporal constraints / L. Capezzuto, D. Tarapore, S. D. Ramchurn // *arXiv preprint arXiv:2003.13806*. – 2020. – Текст : электронный. – DOI: 10.48550/arXiv.2003.13806. – URL: <https://arxiv.org/abs/2003.13806> (дата обращения: 07.05.2024).

9. Dayan P. Feudal reinforcement learning / P. Dayan, G. E. Hinton // *Advances in Neural Information Processing Systems (NIPS)*. – 1992. – Vol. 5. – P. 271–278.

10. Dignum F. Autonomous agents and norms / F. Dignum // *Artificial Intelligence and Law*. – 1999. – Vol. 7, № 1. – P. 69–79. – Текст : электронный. – DOI: 10.1023/A:1008315530323. – URL: <https://doi.org/10.1023/A:1008315530323> (дата обращения: 09.02.2024).

11. Dignum V. OMNI: introducing social structure, norms and ontologies into agent organizations / V. Dignum, J. Vázquez-Salceda, F. Dignum // *Programming Multi-Agent Systems: Second International Workshop (ProMAS 2004)*. – Berlin : Springer, 2005. – P. 181–198.

12. Domyshev A. Optimization of the structure of power system multi-agent control / A. Domyshev, D. N. Sidorov // *IFAC-PapersOnLine*. – 2022. – Vol. 55, № 9. – P. 250–255. – Текст : электронный. – DOI: 10.1016/j.ifacol.2022.07.044. – URL: <https://doi.org/10.1016/j.ifacol.2022.07.044> (дата обращения: 15.02.2025).

13. Esser P. Theoretical foundations of representation learning using unlabeled data: statistics and optimization / P. Esser, M. Fleissner, D. Ghoshdastidar // *arXiv preprint arXiv:2509.18997*. – 2025. – Текст :

электронный. – DOI: 10.48550/arXiv.2509.18997. – URL: <https://arxiv.org/abs/2509.18997> (дата обращения: 10.10.2025).

14. Ferber J. Agent/group/roles: simulating with organizations / J. Ferber, O. Gutknecht, F. Michel // Fourth International Workshop on Agent-Based Simulation (ABS 4). – Montpellier, 2003.

15. Foerster J. Counterfactual multi-agent policy gradients / J. Foerster, G. Farquhar, T. Afouras, N. Nardelli // Proceedings of the AAAI Conference on Artificial Intelligence. – 2018. – Vol. 32, № 1. – Текст : электронный. – DOI: 10.1609/aaai.v32i1.11794. – URL: <https://doi.org/10.1609/aaai.v32i1.11794> (дата обращения: 15.02.2025).

16. Ghrieb N. Maintaining organizational multi-agent systems: a reorganization-based preventive approach / N. Ghrieb, F. Mokhati, G. Tahar // Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART 2021). – 2021. – P. 384–389. – Текст : электронный. – DOI: 10.5220/0010314803840389. – URL: <https://doi.org/10.5220/0010314803840389> (дата обращения: 10.03.2025).

17. Hernandez-Garcia A. Rethinking supervised learning: insights from biological learning and from calling it by its name / A. Hernandez-Garcia // arXiv preprint arXiv:2012.02526. – 2020. – Текст : электронный. – DOI: 10.48550/arXiv.2012.02526. – URL: <https://arxiv.org/abs/2012.02526> (дата обращения: 12.11.2023).

18. Ho J. Generative adversarial imitation learning / J. Ho, S. Ermon // Advances in Neural Information Processing Systems (NIPS). – 2016. – Vol. 29. – P. 4565–4573. – Текст : электронный. – URL: <https://papers.nips.cc/paper/2016/hash/cc7e2b878868cbae992d1fb743995d8f-Abstract.html> (дата обращения: 18.08.2025).

19. Horling B. A survey of multi-agent organizational paradigms / B. Horling, V. Lesser // The Knowledge Engineering Review. – 2004. – Vol. 19, № 4. – P. 281–316.

20. Kandhasamy S. Decentralized motion planning for multi-robot navigation using deep reinforcement learning / S. Kandhasamy, V. B. Kuppusamy, T. V. Samak, C. V. Samak // arXiv preprint arXiv:2011.05605. – 2020. – Текст : электронный. – DOI: 10.48550/arXiv.2011.05605. – URL: <https://arxiv.org/abs/2011.05605> (дата обращения: 03.10.2023).

21. Khapre S. Agentic AI systems: a systematic survey of multi-agent architectures, cognitive foundations, interaction, explainability, security, and performance evaluation / S. Khapre et al. // Neurocomputing. – 2026. – Vol. 696. – Текст : электронный. – URL: <https://www.sciencedirect.com/science/article/abs/pii/S0925231226014475> (дата обращения: 12.05.2026).

22. Krnjaic A. Scalable multi-agent reinforcement learning for warehouse logistics with robotic and human co-workers / A. Krnjaic et al. // arXiv preprint arXiv:2212.11498. – 2022. – Текст : электронный. – DOI: 10.48550/arXiv.2212.11498. – URL: <https://arxiv.org/abs/2212.11498> (дата обращения: 15.02.2024).

23. Mali I. G. Self-play meta-reinforcement learning in multi-agent games / I. G. Mali // Acta Universitatis Sapientiae, Informatica. – 2026. – Vol. 18. – Текст : электронный. – DOI: 10.1007/s44427-026-00021-у. – URL: <https://doi.org/10.1007/s44427-026-00021-у> (дата обращения: 03.06.2026).

24. Neves D. E. Advances and challenges in learning from experience replay / D. E. Neves, L. Ishitani, Z. K. G. do Patrocinio Junior // Artificial Intelligence Review. – 2024. – Текст : электронный. – DOI: 10.1007/s10462-024-11062-0. – URL: <https://doi.org/10.1007/s10462-024-11062-0> (дата обращения: 25.09.2025).

25. Rahwan T. Coalition structure generation on graphs / T. Rahwan, T. P. Michalak // arXiv preprint arXiv:1410.6516. – 2015. – Текст : электронный. – DOI: 10.48550/arXiv.1410.6516. – URL: <https://arxiv.org/abs/1410.6516> (дата обращения: 22.01.2024).

26. Ross S. A reduction of imitation learning and structured prediction to no-regret online learning / S. Ross, G. J. Gordon, J. A. Bagnell // Proceedings of the

14th International Conference on Artificial Intelligence and Statistics (AISTATS). – 2011. – P. 627–635. – Текст : электронный. – URL: <https://proceedings.mlr.press/v15/ross11a.html> (дата обращения: 12.01.2025).

27. Rubio-Manzano C. Machine-imitative learning by using computational perceptions as labeled data-sets: a first empirical approximation / C. Rubio-Manzano, T. Lermenda, C. Martinez, A. Segura, C. Vidal // Joint Proceedings of the 19th World Congress of the International Fuzzy Systems Association (IFSA), the 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT) and the 11th International Summer School on Aggregation Operators (AGOP). – 2019.

28. Si J. Vehicle dispatching and routing of on-demand intercity ride-pooling services: a multi-agent hierarchical reinforcement learning approach / J. Si, F. He, X. Lin, X. Tang // arXiv preprint arXiv:2307.06742. – 2023. – Текст : электронный. – DOI: 10.48550/arXiv.2307.06742. – URL: <https://arxiv.org/abs/2307.06742> (дата обращения: 19.09.2023).

29. Sichman J. S. Agents organizations: a concise overview / J. S. Sichman, V. Dignum, C. Castelfranchi // Journal of the Brazilian Computer Society. – 2005. – Vol. 11, № 1. – P. 3–8. – DOI: 10.1590/S0104-65002005000200001.

30. Sims M. Knowledgeable automated organization design for multi-agent systems / M. Sims, D. Corkill, V. Lesser // Autonomous Agents and Multi-Agent Systems. – 2008. – Vol. 16, № 2. – P. 151–185. – Текст : электронный. – DOI: 10.1007/s10458-007-9023-8. – URL: <https://doi.org/10.1007/s10458-007-9023-8> (дата обращения: 09.02.2024).

31. Sun L. Multi-agent coordination across diverse applications: a survey / L. Sun, Y. Yang, Q. Duan, Y. Shi, C. Lyu, Y.-C. Chang, C.-T. Lin, Y. Shen // arXiv preprint arXiv:2502.14743. – 2025. – Текст : электронный. – DOI: 10.48550/arXiv.2502.14743. – URL: <https://arxiv.org/abs/2502.14743> (дата обращения: 06.03.2025).

32. Unified Modeling Language Specification. Version 2.5.1 : specification / Object Management Group. – 2017. – Текст : электронный. – URL: <https://www.omg.org/spec/UML/2.5.1> (дата обращения: 15.06.2026).

33. Wang Z. Multi-agent systems: from classical paradigms to large foundation model-enabled futures / Z. Wang, M. Gong, Q. Sun, J. Xu, S. Mao, X. Jin // IEEE/CAA Journal of Automatica Sinica. – 2026. – Vol. 13, № 5. – Текст : электронный. – URL: <https://ieeexplore.ieee.org/abstract/document/11543455> (дата обращения: 21.07.2026).

34. Watson D. S. On the philosophy of unsupervised learning / D. S. Watson // Philosophy & Technology. – 2023. – Текст : электронный. – DOI: 10.1007/s13347-023-00635-6. – URL: <https://doi.org/10.1007/s13347-023-00635-6> (дата обращения: 25.09.2025).

35. Wong A. Deep multiagent reinforcement learning: challenges and directions / A. Wong, T. Bäck, A. V. Kononova, A. Plaat // Artificial Intelligence Review. – 2023. – Vol. 56, № 6. – Текст : электронный. – DOI: 10.1007/s10462-022-10299-x. – URL: <https://doi.org/10.1007/s10462-022-10299-x> (дата обращения: 25.09.2025).

36. Wooldridge M. Intelligent agents: theory and practice / M. Wooldridge, N. R. Jennings // The Knowledge Engineering Review. – 1995. – Vol. 10, № 2. – P. 115–152.

37. Wulff P. Applying machine learning in science education research / P. Wulff, M. Kubsch, C. Krist. – Cham : Springer, 2025. – Текст : электронный. – URL: https://link.springer.com/chapter/10.1007/978-3-031-74227-9_4 (дата обращения: 19.11.2025).

38. Yang Y. Many-agent reinforcement learning : PhD thesis / University College London. – London, 2021. – Текст : электронный. – URL: https://discovery.ucl.ac.uk/id/eprint/10124273/12/Yang_10124273_thesis_revised.pdf (дата обращения: 05.03.2024).

39. Yu L. Genetic algorithm aided optimization of hierarchical multi-agent system organization / L. Yu, Z. Shen, C. Miao, V. Lesser // Proceedings of the

Tenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011). – Taipei, Taiwan, 2011. – Vol. 1–3.

40. Shen Z. An evolutionary approach for optimizing hierarchical multi-agent system organization / Z. Shen, L. Yu, H. Yu // arXiv preprint arXiv:1411.6202. – 2014. – Текст : электронный. – URL: <https://arxiv.org/abs/1411.6202> (дата обращения: 14.04.2024).

41. Ziebart B. D. Maximum entropy inverse reinforcement learning / B. D. Ziebart, A. Maas, J. A. Bagnell, A. K. Dey // Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI 2008). – 2008. – P. 1433–1438. – Текст : электронный. – URL: <https://cdn.aaai.org/AAAI/2008/AAAI08-227.pdf> (дата обращения: 08.02.2024).

42. Алгоритмы эволюционной оптимизации / Дэн Саймон ; пер. с англ. А. В. Логунова. – М. : ДМК Пресс, 2020. – 1002 с. – Текст : непосредственный.

43. Модели и методы управления большими группами роботов : учебное пособие / А. Р. Гайдук, И. А. Каляев, С. Г. Капустян, А. П. Кухаренко. – Ростов-на-Дону : Изд-во ЮФУ, 2014. – 106 с. – Текст : непосредственный.

44. Гладков Л. А. Генетические алгоритмы / Л. А. Гладков, В. В. Курейчик, В. М. Курейчик ; под ред. В. М. Курейчика. – 2-е изд., испр. и доп. – М. : Физматлит, 2006. – 320 с. – Текст : непосредственный.

45. Грессер Л. Глубокое обучение с подкреплением: теория и практика на языке Python / Л. Грессер, В. Л. Кенг. – СПб. : Питер, 2022. – 416 с. – Текст : непосредственный.

46. Дафт Р. Л. Теория организации : учебник / Р. Л. Дафт ; пер. с англ. под ред. Э. М. Короткова. – М. : ЮНИТИ-ДАНА, 2006. – 736 с. – Текст : непосредственный.

47. Демидов В. А. Метод синтеза и автоматической адаптации архитектуры и организационной структуры многоагентной системы / В. А. Демидов // Инженерный вестник Дона. – 2026. – № 3 (135). – Текст :

электронный. – URL: <https://cyberleninka.ru/article/n/metod-sinteza-i-avtomaticheskoy-adaptatsii-arhitektury-i-organizatsionnoy-struktury-mnogoagentnoy-sistemy> (дата обращения: 03.07.2026).

48. Коэффициент Джини / В. Г. Минашкин // Большая российская энциклопедия : в 35 т. / гл. ред. Ю. С. Осипов. – М. : Большая российская энциклопедия, 2004–2017.

49. Дубенко Ю. В. Автоматическая адаптация организационной структуры многоагентной системы путем модификации множества правил поведения агентов и функционирования / Ю. В. Дубенко, Е. Е. Дышкант, В. А. Демидов, М. Н. Фалевич // Современные электротехнические и информационные комплексы и системы : сборник материалов VI Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 151–157. – Текст : непосредственный.

50. Дубенко Ю. В. Алгоритм коллективного взаимодействия интеллектуальных агентов в централизованных многоагентных системах / Ю. В. Дубенко // Вестник компьютерных и информационных технологий. – 2022. – Т. 19, № 10 (220). – С. 30–42. – Текст : непосредственный.

51. Дубенко Ю. В. Алгоритм обучения с подкреплением для децентрализованных многоагентных систем, основанный на обмене опытом и обучении агентов случайному взаимодействию / Ю. В. Дубенко, Н. А. Рудешко // Вестник Воронежского государственного технического университета. – 2022. – Т. 18, № 4. – С. 30–36. – Текст : непосредственный.

52. Дубенко Ю. В. Алгоритм определения оптимальной архитектуры многоагентной системы / Ю. В. Дубенко, С. А. Подгорный, Е. Е. Дышкант, В. А. Демидов // Вестник Адыгейского государственного университета. Серия: Естественно-математические и технические науки. – 2025. – № 1 (356). – С. 26–37. – Текст : непосредственный.

53. Дубенко Ю. В. Алгоритм оптимизации архитектуры многоагентной системы, основанный на аппроксимации функции стратегии / Ю. В. Дубенко,

С. А. Подгорный, Е. Е. Дышкант, В. А. Демидов, А. Е. Литвинов // Научные труды КубГТУ. – 2025. – № 4. – С. 141–151. – Текст : непосредственный.

54. Дубенко Ю. В. Анализ иерархического обучения с подкреплением для реализации поведенческих стратегий интеллектуальных агентов / Ю. В. Дубенко, Е. Е. Дышкант, Д. А. Гура // Вестник компьютерных и информационных технологий. – 2020. – Т. 17, № 9 (195). – С. 35–45. – Текст : непосредственный.

55. Дубенко Ю. В. Анализ методов формирования организационной структуры иерархических многоагентных систем / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский // Современные электротехнические и информационные комплексы и системы : материалы IV Междунар. науч.-практ. конф. студентов, аспирантов и преподавателей. – Армавир, 2022. – С. 108–113. – Текст : непосредственный.

56. Дубенко Ю. В. Анализ процедуры синтеза организационной структуры многоагентной системы / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский // Молодые ученые в решении актуальных проблем науки : материалы XII Междунар. науч.-практ. конф. – Владикавказ, 2022. – С. 32–36. – Текст : непосредственный.

57. Дубенко Ю. В. Зоосоциальная модель представления архитектуры многоагентной системы / Ю. В. Дубенко, С. А. Подгорный, Е. Е. Дышкант, В. А. Демидов // Инженерный вестник Дона. – 2025. – № 7. – С. 26–37. – Текст : непосредственный.

58. Дубенко Ю. В. Метод автоматической адаптации архитектуры многоагентной системы, основанный на модификации множества базовых правил ее организации / Ю. В. Дубенко, Е. Е. Дышкант, В. А. Демидов // Научные труды КубГТУ. – 2025. – № 1. – С. 90–108. – Текст : непосредственный.

59. Дубенко Ю. В. Метод автоматической адаптации организационной структуры многоагентной системы в условиях стохастической окружающей среды / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский, М. А. Покасенко //

Научный потенциал вуза – производству и образованию : сборник материалов VII Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 30–36. – Текст : непосредственный.

60. Дубенко Ю. В. Метод автоматической реорганизации структуры многоагентных систем в условиях сложной, частично наблюдаемой, динамической окружающей среды / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский // Научные труды КубГТУ. – 2024. – № 1. – С. 65–78. – Текст : электронный. – DOI: 10.26297/2312-9409.2024.1.8. – URL: <https://ntk.kubstu.ru/data/mc/0101/4760.pdf> (дата обращения: 08.02.2025).

61. Дубенко Ю. В. Метод автоматической реорганизации структуры многоагентных систем, основанный на применении алгоритма симулированного закаливания / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2024. – № 1. – Текст : непосредственный.

62. Дубенко Ю. В. Метод адаптации многоагентных систем / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский, А. Д. Сидоркович // Современные электротехнические и информационные комплексы и системы : сборник материалов VI Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 118–123. – Текст : непосредственный.

63. Дубенко Ю. В. Метод оптимизации организационной структуры централизованных многоагентных систем в автоматическом режиме / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. – 2024. – № 1. – С. 46–55. – Текст : непосредственный.

64. Дубенко Ю. В. Метод первоначального формирования и последующей адаптации архитектуры многоагентной системы, основанный на применении генетического алгоритма / Ю. В. Дубенко, Е. Е. Дышкант, В. А. Демидов // Вестник Астраханского государственного технического

университета. Серия: Управление, вычислительная техника и информатика. – 2025. – № 2. – С. 49–59. – Текст : непосредственный.

65. Дубенко Ю. В. Метод повторного применения и обмена опытом при коллективном взаимодействии интеллектуальных агентов / Ю. В. Дубенко // Вестник Воронежского государственного технического университета. – 2022. – Т. 18, № 1. – С. 62–72. – Текст : непосредственный.

66. Дубенко Ю. В. Метод синтеза и автоматической адаптации архитектуры иерархической многоагентной системы / Ю. В. Дубенко, Е. Е. Дышкант, В. А. Демидов // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2025. – № 1 (73). – С. 40–54. – Текст : непосредственный.

67. Дубенко Ю. В. Метод синтеза организационной структуры централизованных многоагентных систем с возможностью ее автоматической адаптации в случае изменения параметров окружающей среды или поставленной задачи / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2023. – № 3 (67). – С. 18–30. – Текст : непосредственный.

68. Дубенко Ю. В. Метод формирования двухуровневой организационной структуры многоагентной системы, основанной на кластеризации множества агентов / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский, А. Д. Сидоркович // Современные электротехнические и информационные комплексы и системы : материалы V Междунар. науч.-практ. конф. – 2023. – С. 123–127. – Текст : непосредственный.

69. Дубенко Ю. В. Методология коллективного взаимодействия агентов интеллектуальных иерархических систем в процессе обучения с подкреплением при исследовании окружающего пространства : дис. ... д-ра техн. наук / Ю. В. Дубенко. – Дубна : Университет «Дубна», 2024. – 500 с. – Текст : непосредственный.

70. Дубенко Ю. В. Функциональное моделирование процедуры синтеза организационных структур многоагентных систем / Ю. В. Дубенко, Е. Е. Дышкант, А. А. Обозовский // Научный потенциал вуза – производству и образованию. – 2023. – С. 52–56. – Текст : непосредственный.

71. Дышкант Е. Е. Алгоритм адаптации фокуса функционирования многоагентных систем / Е. Е. Дышкант, Ю. В. Дубенко, А. А. Обозовский, А. Е. Гудченко // Современные электротехнические и информационные комплексы и системы : сборник материалов VI Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 137–142. – Текст : непосредственный.

72. Дышкант Е. Е. Анализ связей агентов в рамках организационных структур многоагентных систем / Е. Е. Дышкант, Ю. В. Дубенко, В. А. Демидов, М. А. Покасенко // Современные электротехнические и информационные комплексы и системы : сборник материалов VI Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 171–177. – Текст : непосредственный.

73. Дышкант Е. Е. Математическая модель организационной структуры многоагентной системы / Е. Е. Дышкант, Ю. В. Дубенко, А. А. Обозовский, А. Е. Гудченко // Современные электротехнические и информационные комплексы и системы : материалы V Междунар. науч.-практ. конф. – 2023. – С. 172–175. – Текст : непосредственный.

74. Дышкант Е. Е. Метод реализации механизма миграции агентов между группами, основанный на применении алгоритма биогеографической оптимизации / Е. Е. Дышкант, Ю. В. Дубенко, А. А. Обозовский, Н. А. Коломзаров // Современные электротехнические и информационные комплексы и системы : материалы V Междунар. науч.-практ. конф. – 2023. – С. 134–138. – Текст : непосредственный.

75. Дышкант Е. Е. Применение принципов теории организаций при адаптации структуры многоагентных систем в условиях изменения параметров окружающей среды / Е. Е. Дышкант, Ю. В. Дубенко, А. А. Обозовский, М. А. Покасенко // Современные электротехнические и

информационные комплексы и системы : материалы V Междунар. науч.-практ. конф. – 2023. – С. 165–169. – Текст : непосредственный.

76. Жданов А. А. Автономный искусственный интеллект / А. А. Жданов. – 3-е изд. (эл.). – М. : БИНОМ. Лаборатория знаний, 2012. – 359 с. – Текст : электронный. – URL: https://rusneb.ru/catalog/000199_000009_007488569/ (дата обращения: 08.10.2025).

77. Иберла К. Факторный анализ / К. Иберла ; пер. с нем. В. М. Ивановой. – М. : Статистика, 1980. – 398 с. – Текст : непосредственный.

78. Иваницкий В. В. Зоосоциология : конспект лекций. – Текст : электронный. – URL: <https://teach-in.ru/file/synopsis/pdf/zoosociology-M.pdf> (дата обращения: 17.10.2026).

79. Игнатьева Е. В. Аналитический обзор методов обучения интеллектуальных агентов в иерархических мультиагентных системах / Е. В. Игнатьева, В. А. Демидов // Интеграция наук – 2025 : сборник материалов VII Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 131–136. – Текст : непосредственный.

80. Касьянов В. Н. Методы и средства визуализации информации на основе атрибутированных иерархических графов с портами / В. Н. Касьянов // Сибирский аэрокосмический журнал. – 2023. – Т. 24, № 1. – С. 8–17. – Текст : непосредственный.

81. Козов А. В. Сравнение эффективности некоторых модификаций алгоритма эволюционной стратегии / А. В. Козов // Наука и образование. – 2018. – № 5. – Текст : электронный. – DOI: 10.18698/2541-8009-2018-05-309. – URL: <https://doi.org/10.18698/2541-8009-2018-05-309> (дата обращения: 05.04.2024).

82. Кофман А. Введение в теорию нечетких множеств / А. Кофман ; пер. с франц. – М. : Радио и связь, 1982. – 432 с. – Текст : непосредственный.

83. Лысов В. Ф. Этология животных / В. Ф. Лысов, Т. Е. Костина, В. И. Максимов ; под ред. В. И. Максимова. – М. : КолосС, 2010. – 296 с. – Текст : непосредственный.

84. Ма Вэй Цзи. Байесовские модели восприятия и действия / Вэй Цзи Ма, К. Кёрдинг, Д. Голдрайх ; пер. с англ. В. С. Яценкова. – М. : ДМК Пресс, 2023. – 458 с. – Текст : непосредственный.

85. Маслобоев А. В. Гибридная архитектура интеллектуального агента с имитационным аппаратом / А. В. Маслобоев // Вестник МГТУ. – 2009. – Т. 12, № 1. – С. 113–124. – Текст : непосредственный.

86. Мендель И. Д. Кластерный анализ / И. Д. Мендель. – М. : Финансы и статистика, 1988. – 176 с. – Текст : непосредственный.

87. Месарович М. Теория иерархических многоуровневых систем / М. Месарович, Д. Мако, И. Такахара. – М. : Мир, 1973. – 344 с. – Текст : непосредственный.

88. Обозовский А. А. Адаптация архитектуры многоагентной системы путем изменения степени выраженности признаков функциональной или дивизионной структур / А. А. Обозовский, Ю. В. Дубенко, Е. Е. Дышкант, М. Н. Фалевич // Современные электротехнические и информационные комплексы и системы : материалы V Междунар. науч.-практ. конф. – 2023. – С. 150–154. – Текст : непосредственный.

89. Петренко В. И. Классификация задач мультиагентного обучения с подкреплением / В. И. Петренко // Известия Кабардино-Балкарского научного центра РАН. – 2021. – № 3 (101). – С. 32–44. – Текст : непосредственный.

90. Петухов Г. Б. Методологические основы внешнего проектирования целенаправленных процессов и целеустремленных систем / Г. Б. Петухов, В. И. Якунин. – М. : АСТ, 2006. – 504 с. – Текст : непосредственный.

91. Рассел С. Искусственный интеллект : современный подход. В 4 т. Т. 1. Решение проблем: знания и рассуждения / С. Рассел, П. Норвиг ; пер. с

англ. – 4-е изд. – СПб. : Диалектика, 2021. – 704 с. – Текст : непосредственный.

92. Рассел С. Искусственный интеллект : современный подход. В 4 т. Т. 2. Знания и рассуждения в условиях неопределенности / С. Рассел, П. Норвиг ; пер. с англ. – 4-е изд. – СПб. : Диалектика, 2021. – 480 с. – Текст : непосредственный.

93. Рассел С. Искусственный интеллект : современный подход. В 4 т. Т. 3. Обучение, восприятие и действие / С. Рассел, П. Норвиг ; пер. с англ. – 4-е изд. – СПб. : Диалектика, 2022. – 640 с. – Текст : непосредственный.

94. Саати Т. Аналитическое планирование. Организация систем / Т. Саати, К. Кернс. – М. : Радио и связь, 1991. – 224 с. – Текст : непосредственный.

95. Саттон Р. С. Обучение с подкреплением: введение : практическое руководство / Р. С. Саттон, Э. Дж. Барто ; пер. с англ. А. А. Слинкина. – 2-е изд. – М. : ДМК Пресс, 2020. – 552 с. – Текст : непосредственный.

96. Свидетельство о государственной регистрации программы для ЭВМ № 2026685959 Российская Федерация. Программа для моделирования мультиагентного обучения с подкреплением : № 2026685919 : заявл. 22.07.2026 : опубл. 26.08.2026 / В. А. Демидов. – Текст : непосредственный.

97. Симанков В. С. Системный анализ в иерархических интеллектуальных многоагентных системах / В. С. Симанков, Ю. В. Дубенко // Вестник компьютерных и информационных технологий. – 2021. – Т. 18, № 3 (201). – С. 33–46. – Текст : непосредственный.

98. Тарасов В. Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М. : Эдиториал УРСС, 2002. – 352 с. – Текст : непосредственный.

99. Ферстер Э. Методы корреляционного и регрессионного анализа / Э. Ферстер, Б. Ренц ; пер. с нем. В. М. Ивановой. – М. : Финансы и статистика, 1983. – 304 с. – Текст : непосредственный.

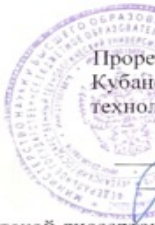
100. Шимширт Д. И. Применение мультиагентных систем в современных условиях развития / Д. И. Шимширт, В. А. Демидов // Интеграция наук – 2025 : сборник материалов VII Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 252–263. – Текст : непосредственный.

101. Шимширт Д. И. Современные методы обучения интеллектуальных агентов в мультиагентных системах: сравнительный анализ и практические реализации / Д. И. Шимширт, В. А. Демидов // Интеграция наук – 2025 : сборник материалов VII Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 243–252. – Текст : непосредственный.

102. Широ́в П. Н. Анализ архитектуры иерархической МАС / П. Н. Широ́в, В. А. Демидов // Интеграция наук – 2025 : сборник материалов VII Междунар. науч.-практ. конф. – Краснодар, 2025. – С. 263–270. – Текст : непосредственный.

Приложение А. Акты о внедрении результатов диссертационного исследования

Акт о внедрении в образовательный процесс КубГТУ



УТВЕРЖДАЮ
 Проректор по учебной работе
 Кубанского государственного
 технологического университета
 канд. техн. наук, доцент
 Е.М. Дьяконов
 2026 г.

АКТ
 об использовании результатов кандидатской диссертации
 Демидова Владислава Анатольевича на тему «Исследование и разработка методов обучения интеллектуальных агентов в иерархических многоагентных системах» по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика
 в учебном процессе Кубанского государственного технологического университета

Комиссия в составе:
 председатель – Макарян А.С., и.о. декана факультета информационных технологий и кибербезопасности Кубанского государственного технологического университета (КубГТУ), канд. техн. наук, доцент кафедры «Кибербезопасности и защиты информации» КубГТУ,
 члены комиссии:
 Дубенко Ю.В., заведующий кафедрой «Информатика и вычислительная техника» КубГТУ, д-р техн. наук, доцент;
 Хисамов Ф.Г., д-р техн. наук, профессор кафедры «Информатика и вычислительная техника» КубГТУ;
 Лубенцов В.Ф., д-р техн. наук, профессор кафедры «Автоматизация производственных процессов» КубГТУ.

составила настоящий акт о том, что результаты кандидатской диссертации ассистента кафедры информатики и вычислительной техники КубГТУ Демидова В.А., а именно:

- метод формирования и автоматической адаптации архитектуры и организационной структуры интеллектуальной иерархической многоагентной системы; - метод обмена опытом между интеллектуальными агентами, работающими в группе для достижения поставленных ЛПР целей и задач в динамических, стохастических, четко не определенных и неполностью наблюдаемых средах; - метод коллективного взаимодействия интеллектуальных агентов в централизованных и децентрализованных многоагентных системах, на основе иерархического обучения с подкреплением с использованием биоинспирированных алгоритмов;
- внедрены в учебный процесс в виде методических указаний и рекомендаций для обучения с подкреплением интеллектуальных агентов в многоагентных системах, используемых при проведении лабораторного практикума, а также в разделах выпускных квалификационных работ бакалаврами и магистрантами направлений 09.03.01, 09.04.01 «Информатика и вычислительная техника», профиль «Прикладные системы искусственного интеллекта».

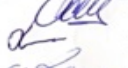
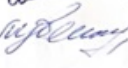
Методический материал диссертации Демидова В.А. в части разработки программного обеспечения, основанного на приводимых алгоритмах обучения интеллектуальных агентов для задачи поиска целевых функций в виртуальном лабиринте, использован в изданных методических указаниях по выполнению курсовой работы по дисциплине «Вычислительные системы» для студентов всех форм обучения направления 09.03.01 «Информатика и вычислительная техника».

Результаты диссертационного исследования использованы при подготовке лекционных курсов по дисциплинам «Вычислительные системы» направления 09.03.01, «Искусственный интеллект», «Прикладные интеллектуальные системы» направления 09.04.01.

Использование указанных материалов способствует приобретению выпускниками университета направления «Информатика и вычислительная техника» умения осуществлять программирование интеллектуальных агентов в иерархических многоагентных системах на основе подхода обучения с подкреплением.

Председатель

Члены комиссии:

 А.С. Макарян
 Ю.В. Дубенко
 Ф.Г. Хисамов
 В.Ф. Лубенцов

Акт о внедрении в ООО «Олимп-Связь»**ООО «Олимп-Связь»****ИНН 2308207851**

Юридический адрес: 350024, Краснодарский край, г. Краснодар, ул. им. Артюшкова, д. 15

АКТ

о внедрении результатов кандидатской диссертации
Демидова Владислава Анатольевича
на тему: «Исследование и разработка методов обучения
интеллектуальных агентов в иерархических многогенных системах»
в исследовательской и проектной деятельности
«ООО Олимп-связь»

Мы, представители ООО «Олимп-связь», настоящим актом подтверждаем, что результаты кандидатской диссертации ассистента кафедры информатики и вычислительной техники ФГБОУ ВО «Кубанский государственный технологический университет» Демидова Владислава Анатольевича использованы в исследовательской и проектной деятельности ООО «Олимп-связь» в виде алгоритмического обеспечения при проектировании и разработке иерархических интеллектуальных многоагентных систем поддержки и принятия решений по управлению группами интеллектуальных агентов.

Разработанные методы и алгоритмы применены в системе управления сетевыми ресурсами и мониторинга телекоммуникационной инфраструктуры ООО «Олимп-связь».

Созданные Демидовым В.А. методы и алгоритмы формирования и автоматической адаптации архитектуры и организационной структуры МАС позволяют динамически изменять топологию связей и роли агентов в зависимости от текущей нагрузки и состояния сети. Биоинспирированные алгоритмы позволили оптимизировать маршруты передачи данных, а также улучшили самоорганизацию агентов при изменении условий работы сети.

Директор ООО «ОЛИМП-СВЯЗЬ»



/И.Д.Лебедев/

Акт о внедрении в ООО «Аэромар-Краснодар»



Общество с ограниченной ответственностью
«Аэромар-Краснодар»
Г. Краснодар, Краснодарский край , 350912
Ул. Еядоки Бершанской дом 355
E-mail: dp@aeromar-krd.
ИНН/ОГРН 23112270648/1182375029374

АКТ № 1 от «11» августа 2026 г.
о внедрении результатов диссертационного исследования

г. Краснодар

Комиссия в составе:

Бердова Ольга Владимировна – заместитель генерального директора
Попов Владимир Николаевич – директор обособленного подразделения
Сидненко Лилия Викторовна – начальник оперативной службы

составила настоящий акт о том, что в ООО «Аэромар-Краснодар» внедрена методика автоматической адаптации архитектуры и организационной структуры в иерархических многоагентных системах, разработанная Демидовым Владиславом Анатольевичем в рамках диссертационного исследования на тему: «Интеллектуальная система автоматической адаптации архитектуры многоагентных систем на основе зоосоциальной модели поведения агентов».

В контур оперативного планирования производства и доставки бортового питания внедрён модуль динамической адаптации иерархии агентов: производственные линии (по типам продукции), бригады (по квалификации), транспорт (по маршрутам и приоритетам), слоты вылетов (окна времени, приоритеты рейсов).

Адаптация организационной структуры агентов реализуется посредством комбинированного применения методов:

- **генетического алгоритма (ГА)** — для подбора оптимальной конфигурации распределения заказов по производственным линиям и транспорту;
- **обучения с подкреплением (RL)** — для динамической подстройки приоритетов при изменении условий;
- **зоосоциального метода** — для перераспределения ролей (координатор, исполнитель, резерв) между агентами при сбоях и всплесках нагрузки.

В ходе пилотного внедрения (период: 01.07.2026–14.07.2026) на обозначенных исторических данных (архив заказов питания, статусы линий, расписания вылетов) получены следующие измеримые результаты:

Показатель	До внедрения (базовый уровень)	После внедрения (пилот)	Изменение
Доля рейсов, обеспеченных питанием в срок	87,2 %	93,0 %	+5,8 п.п.
Среднее отклонение от целевого времени доставки	23 мин	15 мин	–8 мин
Доля спецпитания без ошибок комплектации	90,5 %	95,7 %	+5,2 п.п.
Количество экстренных переназначений в день	6,8	4,9	–27,9 %

Эффект обусловлен автоматической перестройкой иерархии агентов и связей между ними при изменении условий (переносы слотов, задержки на линиях, опоздания транспорта), что соответствует научной новизне методики.

Настоящий акт подтверждает внедрение результатов диссертационного исследования в производственную деятельность ООО «Аэромар-Краснодар».

Подписи членов комиссии:



Акт о внедрении в АО «НИИВК им. М. А. Карцева»



ОРДЕНОВ ТРУДОВОГО КРАСНОГО ЗНАМЕНИ, АЛЕКСАНДРА НЕВСКОГО
АКЦИОНЕРНОЕ ОБЩЕСТВО
«НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ ВЫЧИСЛИТЕЛЬНЫХ КОМПЛЕКСОВ им. М.А. КАРЦЕВА»
АО «НИИВК им. М.А. КАРЦЕВА»
ИНН 7728032882; КПП 772801001; ОГРН 1037700128828; ОКПО 11478124
ул. Профсоюзная, дом 108, г. Москва, 117437. тел. (495) 330-09-29; факс (495) 330-56-30
postoffice@niivk.ru; www.niivk.ru

АКТ

о внедрении результатов диссертационного исследования в практическую деятельность АО «НИИВК им. М. А. Карцева»

Настоящий акт подтверждает применение в практической деятельности АО «Научно-исследовательский институт вычислительных комплексов им. М. А. Карцева» (г. Москва) результатов диссертационного исследования Демидова Владислава Анатольевича на тему «Интеллектуальная система автоматической адаптации архитектуры многоагентных систем на основе зоосоциальной модели поведения агентов», выполненного в АО «НИИВК им. М. А. Карцева». Применение этих результатов обеспечило:

- снижение времени перераспределения задач между роботами;
- сокращение времени простоя группы при выходе из строя робота-координатора за счёт автоматического перехода к децентрализованной структуре типа «коалиция» и назначения нового координатора на основе механизма авторитета;
- повышение устойчивости группы к колебаниям параметров среды (изменение площади зоны мониторинга, появление препятствий);
- расширение спектра моделируемых организационных структур, что позволило подобрать оптимальную структуру для различных сценариев применения робототехнических комплексов.

Результаты диссертационного исследования Демидова В.А. используются при разработке робототехнических комплексов и систем с искусственным интеллектом, в том числе для задач группового управления мобильными роботами, мониторинга и патрулирования. Это подтверждает практическую значимость результатов диссертационного исследования Демидова Владислава Анатольевича.

Генеральный директор

Технический директор

Начальник НТЦ № 2



В.В. Долгов

А.А. Рудин

В.Г. Дементьев

«08» октября 2026 г.

Приложение Б. Свидетельство о регистрации программы для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации программы для ЭВМ
№ 2026685959

**Программа для моделирования мультиагентного
обучения с подкреплением**

Правообладатель: *Демидов Владислав Анатольевич (RU)*

Автор(ы): *Демидов Владислав Анатольевич (RU)*

Заявка № **2026685919**
Дата поступления **22 июля 2026 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ **26 августа 2026 г.**



Руководитель Федеральной службы
по интеллектуальной собственности
Документ подписан электронной подписью
Сертификат 004775487 — 00571646881847933556
Владимир Зубов Юрий Сергеевич
Действителен с 04.07.2015 по 26.11.2028 **Ю. С. Зубов**