Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение высшего профессионального образования

«Уральский федеральный университет
имени первого Президента России Б. Н. Ельцина»
Институт социальных и политических наук
Департамент философии
Кафедра онтологии и теории познания

Когнитивная модель агента в качестве интеллектуальной системы, взаимодействующей со средой

Допустить к защите:

Зав. кафедрой
к.ф.н.,
профессор
Кислов А.Г.

Магистерская диссертация
Студента 2курса
Оглуздина К.А.

Научный руководитель:
д.ф.н.,
профессор,
Анкин Д.В.

Екатеринбург
2015

Оглавление

Введение.....3

Глава 1. Моделирование интеллектуальных систем.....6

1.1. Проблема определения понятия интеллект.....6

1.2. Когнитивный агент как интеллектуальная система.....21

Глава 2. Когнитивные агенты и их взаимодействия.....41

2.1 Архитектуры интеллектуальных агентов.....41

2.2 Проблема эффективности взаимодействия когнитивных агентов в многоагентных системах.....54

Список литературы.....71

Введение

Актуальность исследования

На протяжении многих веков, человек изучал тайны природы с целью понять её устройство и использовать предоставляемые ею возможности для удовлетворения своих многообразных потребностей. Однако, изучая природу вокруг себя, человек оставался в неведенье относительно своих собственных возможностей. Наблюдая за движением звёзд и устанавливая его закономерности, человек был не способен понять устройство движения своих собственных мыслей. Исследование человеческого разума лишь недавно встало на прочный научный фундамент, хотя попытки понять человеком устройство своего разума не прекращались на протяжении многих веков.

Природа человеческого разума охватывает широкий спектр дисциплин, в рамках которых он изучается. Философия, психология, биология, лингвистика, компьютерные науки и многие другие дисциплины, так или иначе, связаны с данной проблемной областью. Так, психологи изучают ментальное пространство человека, биологи занимаются исследованиями мозга, лингвистов интересует взаимосвязь языка и мышления и т.д. Одним и самых захватывающих проектов XX века, был связан с разработкой особого типа машин, которые могли бы осуществлять мыслительную деятельность, сопоставимую с человеческой. Этот проект был назван искусственный интеллект и до сих пор не прекращаются попытки построить машины с квазичеловеческим интеллектом. На этом пути были достигнуты как значительные результаты, двинувшие науку вперёд, так и сокрушительные фиаско, которые, однако, не охладили пыл отдельных энтузиастов. Мы смогли получить ценные сведенья касательно границ рационального мышления, разработать многочисленные технологии, которые просто не могли возникнуть в рамках других направлений. Постепенно произошёл переход от создания моделей локальных интеллектуальных процедур, к созданию сложных интегративных моделей, объединяющих в себе множество интеллектуальных способностей и даже способных взаимодействовать с

окружающим их миром. Однако, многочисленные исследования в этой области, так и не привели к созданию полноценной модели человеческого интеллекта.

И хотя искусственный интеллект на данном этапе развития всё таки уступает человеческому, он может помочь нам понять работу последнего. Чем больше усилий мы направляем на построение модели интеллекта, тем больше мы продвигаемся в понимании работы оригинала. Поэтому данная сфера исследования и, посей день, не утратила своей актуальности, и притягивает новые поколения исследователей. Ещё одной причиной этой притягательности выступает отсутствие у нас понимания общих принципов работы нашего разума. И, несмотря на весьма внушительные результаты, касательно отдельных функций и способностей, мы всё ещё далеки от понимания работы всей системы в целом. И для того чтобы приблизиться к созданию подобной системы, нам необходимо развивать наши технологии моделирования интеллектуальной деятельности.

Степень разработанности темы

В рамках нашего исследования мы попытаемся осветить такую тему как моделирование когнитивных агентов. Данное направление является активно развивающимся в наши дни, и потому было выбрано нами в качестве направления исследования. Исчерпывающие представления об области искусственного интеллекта можно получить из работы С. Рассела, П. Норвинга [39]. В совместной работе Гаазе-Рапопорт М. Г., Поспелов Д. А. можно посмотреть развитие агентного направления [11]. Современные разработки в области бионического подхода к искусственному интеллекту можно найти в работах Редько В. Г. [40], Шамис А.Л.[50], Тарасов В.Б.[44]. Также весьма большой интерес представляют работы М. Минского[23], А. Жданов[16], Дж. Хокинса [48]. Для ознакомления с основными направлениями исследования интеллекта, можно обратиться к монографии Холодная М. А. [49].

Объектом исследования – искусственные интеллектуальные системы как объект философского исследования

Предметом исследования – искусственные интеллектуальные агенты.

Целью нашей работы – выявить основные свойства искусственных интеллектуальных агентов

Для достижения поставленной цели, нами были поставлены следующие задачи:

1.Произвести анализ понятия интеллект и выявить его специфику относительно искусственных и естественных систем.

2. Проанализировать основные модели агента как интеллектуальной системы.

3. Рассмотреть основные архитектуры интеллектуальных агентов.

4. Рассмотреть особенности взаимодействия интеллектуальных агентов между собой.

Основными методами в ходе нашего исследования являются концептуальный и сравнительный анализ когнитивных агентов.

Наша работа состоит из двух частей, каждая из которых поделена на два раздела. В первой части работы, в первом блоке нами изучено понятие интеллекта и его особенностей относительно естественных и искусственных систем, а второй блок посвящен анализу понятия агент, рассмотрению классификаций искусственных агентов в целом, и когнитивных агентов в частности. Во второй части работы, в первом блоке мы рассматриваем архитектуры когнитивных агентов, а также проведён их сравнительный анализ. Второй блок посвящен проблеме взаимодействия агентов в многоагентной системе.

Глава 1. Моделирование интеллектуальных систем

1.1. Проблема определения понятия интеллект

Часто в повседневной жизни, мы используем, фрагментарные и неточные определения, когда пытаемся выразить сущность того или иного явления. Подобной участи не избежал и такой феномен как интеллект. Ниже мы представим несколько подобных определений, которые нам хорошо известны из личного опыта.

1. Интеллект – способность решать задачи с использованием символьной информации.
2. Интеллект – способность читать и понимать текст.
3. Интеллект – способность находить устойчивые закономерности окружающего мира и использование их в своих целях
4. Интеллект – это способность к обучению
5. Интеллект – это способность успешно приспосабливаться к изменениям сред.

Ряд такого рода определений можно продолжать до бесконечности. Также нередко мы прибегаем к разнообразным классификациям на предмет включения каких-либо задач или каких-либо видов деятельности в разряд интеллектуальных. Вследствие такого рода классификации, мы определяем интеллектуальность того или иного субъекта в зависимости от того, выполняет или нет он те или иные действия, выполняет ли он те или иные задания. Такая чисто бихевиористская позиция относительно интеллекта приводит к размыванию смысла данного понятия.

В качестве иллюстрации нашей мысли, обратимся к тестам IQ, призванные выявлять уровень интеллектуального развития испытуемого. Испытуемый выполняет ряд тестовых заданий. В качестве таковых могут выступать: назвать все представленные вещи одним словом, обнаружить

наличие закономерности в последовательности картинок, запомнить и воспроизведении ряда слов и т. п.

Данные тесты разрабатывались психологами с целью измерения умственных способностей. И в ходе проведения исследований и их анализа сложились два диаметрально противоположных направления, связанных в первую очередь с интерпретацией тестов, так как методы исследования и результаты тестов были, в принципе, сопоставимыми. В рамках первого направления, было принято считать, что существует некий общий показатель интеллектуальной активности. Представители второго направления считали, что мы не измеряем какого-либо общего параметра, а измеряем несколько независимых параметров. К основателям первого направления относят К. Спирмена, в то время как, основателем второго принято считать Л. Тэрстоуна

Спирмен, основываясь на наличие положительной корреляции между выполнениями различных интеллектуальных тестов, выдвинул идею о наличии общего начала интеллектуальной деятельности (фактор g). Также он выделял фактор s, который характеризовал особенности при выполнении конкретных заданий. Данную теорию называют «двухфакторной теорией интеллекта». [47] Таким образом, фактор g являлся тем самым общим параметром интеллекта, о котором шла речь выше.

Трестоун в своих исследованиях выделял наличие 7 групповых факторов, которые он относил к «первичным умственным способностям:
S – пространственный;
P – восприятие;
N – вычислительный;
V – вербальное понимание;
F – беглость речи;
M – память;
R – логическое рассуждение.

Данная концепция получила названия «многофакторная теория интеллекта». [47] Здесь мы видим, что вместо одного общего показателя интеллекта используется набор характеристик, которые являются независимыми друг от друга

Однако, по мере нарастания эмпирического материала было обнаружено, что в каком бы направлении не двигался исследователь, не одно

из один из этих путей не даёт к желаемого результата. Исследователям, которые искали общий показатель интеллекта, с завидным постоянством удавалось находить взаимосвязь между смежными заданиями, а, следовательно, наличие нескольких способностей. Те, кто шёл по пути классификации умственных способностей по группам, неизменно сталкивались с возможность нахождения общего во многом. Не меньшей проблемой оказался подбор тестовых заданий и их проверка. Какие задания должны быть включены в тест, а какие нет? Какие типы заданий наиболее репрезентативны относительно той или иной способности? Является ли скорость выполнения показателем уровня интеллекта? К тому же тесты совершенно игнорировали влияние внешних и внутренних факторов, таких как социальный статус испытуемого, его эмоциональное состояние, личностных особенности. И хотя подобного рода корреляции установить было можно, однако, понять, влияют ли эмоции на процесс решения задания или же наоборот, исходя из самих тестов, было совершенно невозможно. Требовался нечто большее, чем просто набор признаков. Таким образом, мы приходим к вполне естественному выводу, что свести интеллект к решению каких-либо задач, является крайне бесперспективным предприятием.

Давайте попробуем обратиться к более содержательным психологическим теориям. В монографии Холодной М. А. [47] представлены основные направления в психологии интеллекта:

1. Феноменологический подход (интеллект как содержание ментального пространства).
2. Генетический подход (интеллект как следствие эволюционного развития, адаптационный механизм, выработанный при взаимодействии человека и среды).
3. Социо-культурный подход (интеллект как результат социализации и воздействия культуры в целом).
4. Процессуально-деятельностный подход (интеллект как особая форма человеческой деятельности).
5. Образовательный подход (интеллект как результат целенаправленного обучения).

6. Информационный подход (интеллект как схема обработки информации мышлением).
7. Функционально-уровневый подход (интеллект как система разноуровневых умственных способностей).
8. Регуляционный подход (интеллект как ведущий фактор психической регуляции).

Давайте рассмотрим все из этих направления более детально. Феноменологическое направление восходит своим корнями к интроспективным методам описания собственных ментальных состояний. Подобного рода описания мы можем найти практически во всех сферах культуры. Основной проблемой данного подхода является излишний субъективизм получаемых данных. Тем не менее, представители гештальт-психологии, Вундт и многие другие выдающиеся психологи использовали этот метод в своих исследованиях интеллекта. Наиболее интересным для нас оказывается направление гештальт-психологии[8]. Представители данного направления считали, что во время поиска решения задачи происходит перестройка гештальта, и когда найдено решение, происходит инсайт или озарение. Под гештальтами они понимали некоторые устойчивые психические образования, характеризующиеся целостностью. Именно, перестройка этих образований и есть интеллектуальная деятельность.

Генетический подход объясняет природу интеллекта через призму адаптации. В рамках данного подхода уместно упомянуть теорию интеллекта Ж. Пиаже. [29] В рамках его подхода, интеллект выступает как организующее начало взаимодействия организма со средой. Интеллект, по мнению Пиаже, является наивысшей формой адаптации организма. Интеллект осуществляет взаимодействие со средой посредством двух взаимосвязанных процессов: ассимиляции (формирование в уме организма когнитивной схемы внешней среды) и аккомадации (преобразование этой схемы под воздействием среды). При этом в ходе этого можно выделить две его составляющие: когнитивную и аффективную. Суть этих двух аспектов отражена в следующем высказывании: «чувство дает действию необходимую

энергию, в то время как знание налагает на поведение определенную структуру».[28]

Также Пиаже была предложена концепция стадий развития интеллекта. При этом, он полагал что единственным фактором, влияющим на формирование интеллекта является действия с предметами внешнего мира. Всего он выделял пять стадий:

1. Стадия сенсомоторного интеллекта.
2. Стадия символического (допонятийного) интеллекта
3. Стадия наглядно образного (интуитивного) интеллекта
4. Стадия конкретных операций
5. Стадия формальных операций

На каждой стадии развития происходит усложнение навыков операциональной деятельности, в плоть да формирования способности манипулировать формальными понятиями.

Социокультурный подход основывается на идее, что главным фактором в развитии интеллекта является социум или же влияние культуры. Таким образом, для того чтобы понять мышление человека следует исходить того в каком обществе он живет, какие культурные ценности он разделяет. Ярким представителем этого направления является Л. С. Выготский. [10] Для него все высшие психические функции, а значит и интеллект, являются следствиями использования семиотических орудий (языки, чертежи, счисления). Эти орудия были созданы человеком для управления психическими процессами. Также в рамках его теории разграничиваются природный и социальный интеллект. Первый выступает следствием биологической эволюции, а второй – исторического развития.

Процессуально-деятельностный подход в сущности своей представляет динамическую теорию интеллекта. В рамках этого подхода, интеллект нельзя понять иначе как в процессе его работы. Поэтому необходимо изучать интеллектуальную активность, то, как происходит процесс мышления. Такая интерпретация позволяет провести параллель с исследователями мозга, которые пытаются понять природу человеческого разума, изучая

функционирование нейронов. Последователи С. Л. Рубенштейна[39], который является выдающимся представителем данного направления, разграничивают интеллект как способность и мышление как процесс.

Функционально-уровневый подход основан на представлении интеллекта как сложно организованной структуры разнообразных психических способностей. Особенно здесь интересен подход Б. Г. Ананьева [2], который построил схему билатерального регулирования. Суть этой схемы сводиться к тому, что интеллект имеет вертикальную иерархическую структуру «ощущение – восприятие – представление – мышление», к тому же на каждом уровне существуют свои собственные горизонтальные связи. Так же в рамках его исследований были выявлено влияние низших уровней на высшие уровни и наоборот, что, по сути, является показателем наличия обратных связей. Таким образом, развитие интеллекта оказывает сопряжено с интеграцией различных познавательных функций и ростом корреляционных связей между ними. Помимо этого им выделялось наличие общих свойств, присущих все уровням познания, например, объем, координация и т.д.

В рамках регулятивного и образовательного подхода, рассматриваются две важнейших свойства интеллекта. Регулятивная составляющая интеллекта говорит нам о том, что интеллект является не только инструментом познания, но и средством регуляции поведения. В случае необходимости, он может тормозить некоторые реакции организма, например, если действие не соответствует сложившейся обстановке. Образовательная составляющая показывает нам, что интеллект является не просто застывшей структурой, он также является динамичной структурой способной изменяться со временем. В рамках исследования регулятивных механизмов можно обратиться к исследованиям Р. Стренберга. [47] Им была выдвинута концепция троичной структуры интеллекта: субтеория компонентов, субтеория контекста и субтеория опыта. Первая субтеория отвечает за регуляцию внутренних ментальных процессов, вторая субтеория отвечает за управление

взаимодействием с внешней средой, а третья субтеория позволяет работать со шкалой «стереотипность – новизна».

Информационный подход, основан на идее «компьютерной метафоры», которая предполагает анализ информационных процессов лежащих в основе мыслительных процессов. Также этот подход предполагает использование переноса компьютерной терминологии в область исследования психологии. Так в психологии появляются такие понятия как оперативная и долговременная память, начинаются строиться схемы обработки информации и т.п. Так в теории поведения представленной в работе Д. Миллера, Ю. Галантера и К. Пибрама [21], активно используются термины кодирования, декодирование информации, классификация, распознавание конфигурации и пр. По сути дела, когнитивный подход в психологии весь основан на интерпретации психологических терминов в виде процессов обработки информации. Также важно отметить, что переход на компьютерную аналогию служит связующим звеном, объединяющим искусственный интеллект и когнитологию в целом, а также когнитивную психологию в частности.

Подведем итог нашего обзора психологических теорий интеллекта. Обобщая все представленные подходы, мы получаем некоторую общую картину того, что представляет собой интеллект. Во-первых, интеллект является результатом длительной эволюции человека, как в биологическом, так и в социальном аспектах. Во-вторых, интеллект обладает процессуальным характером, в ходе которого обрабатывается внешние и внутренние виды информации, генерируется принятие решений, формируются стратегии поведения и т.д. И, в-третьих, интеллект есть сложноорганизованная структура, состоящая из разнообразных познавательных процессов, которые имеют высокую степень корреляции между собой. И самое главное, это то, что интеллект является результатом адаптивного взаимодействия организма со средой.

Вместе с когнитивным подходом в психологии, исследования интеллекта стали осуществляться в рамках такого направления как

искусственный интеллект. В рамках данного направления, исследователи пытались с помощью вычислительных машин, создать модель мышления человека. В частности одной из таких моделей стала машина Тьюринга, которая в работах Х. Патнэма [28] рассматривается как полноценная модель человеческого мышления. С самого начала своего развития это направление распалось на две самостоятельные ветви: рационалистическая традиция и бионическая традиция. Первая из этих ветвей была связанна с разработкой думающих машин на основе имитации мыслительных процессов, которые основывались на манипуляции символьной информацией. В то время как бионическая ветвь развития, была представлена такими разработками как нейронные сети и генетические алгоритмы. Представители этого подхода пытались строить свои модели на основе имитации физиологических процессов в организме, а также эволюционных процессов.

По мере своего развития, в рамках рационалистической парадигмы, удалось достичь весьма впечатляющих результатов. Машины научились играть в шахматы, шашки, доказывать теоремы, производить сложные вычисления, удалось получить хорошие алгоритмы поиска решений. Однако, при всех свои успехах было выявлено явное несовершенство этого подхода. Первая причина этого несовершенства была связана с тем, что задачки, с которыми так хорошо справлялись машины либо уже имели решения, либо они были довольно простыми. Программы, которые создавались в рамках этого направления, работали с небольшими классами объектов и, соответственно, небольшим числом ветвлений. И хотя подобные вещи производили огромные впечатления на неискушенную публику, для профессионального сообщества они представлялись не более, чем детскими забавами. Открытие эффекта комбинаторного взрыва показало принципиальную ограниченность данного направления, так как стало очевидно, что при высокой сложности задачи машина просто не сможет решить её за какое- либо приемлемое время. Тем не менее, результатами

поисков стали экспертные системы, а также разработки в области машинного перевода и автоматической обработке текста.

В плане разработки систем машинного перевода и автоматической обработке текста стоит отметить, что они были инициированы с одной стороны знаменитым тестом Тьюринга, с другой стороны – социальным заказом. Алан Тьюринг[37], в своё время предложил тест на интеллект, в основе которого лежал не просто список каких-либо параметров, а, так сказать, «игра в имитацию». Суть теста сводилась к тому, чтобы машина могла вести диалог с человеком таким образом, чтобы невозможно было определить машина эта или человек. Однако, первые попытки в данном направлении связанные с чисто синтаксическими манипуляциями не оправдали возлагавшихся на них надежд. Попытки разработки автоматических систем перевода показали, что помимо чисто грамматических правил требуется ещё и знание объектного мира, а также особенностей ситуации говорения. Дополнительную сложность представляет также постоянное изменение языка, что требует наличие у программы способности к обучению, а также многозначность и нечеткость синтаксиса естественного языка. Исследования в этой области, показали, что овладение языком требует от наших машин достаточно высокой степени интеллектуализации. Однако, здесь стоит заранее оговориться, что высокая интеллектуальность, которая требуется для овладения языком отнюдь не означает что мы производить отождествление мышления с языком. Так в своей работе «Язык как инстинкт» Пинкер[30] отстаивает независимость языка и интеллектуальной деятельности. Однако, подробное изложение данного вопроса выходит за рамки нашего исследования.

Отдельного рассмотрения, в рамках нашего исследования, заслуживают экспертные системы, так как данные системы вполне можно отнести к моделям интеллекта. [33] И так что такое экспертная система? Данная программа работает в режиме «вопрос-ответ». Пользователь с помощью устройства ввода отправляет запрос системе, та на основе своей базы знаний

и уточняющих вопросов должна выдать ответ на поставленный ей вопрос. Например, выставить диагноз больному, на основе полученных анализов или указания симптомов болезни. Или же на основе данных о совершенном преступлении соотнести его с соответствующей санкцией. При разработке подобных систем производят опрос экспертов, узнавая у них, как они делают выводы, чем руководствуются при принятии решений, а также сюда же добавляют основные знания из данной области и правила вывода. Далее все это переводиться в удобоваримую, для компьютера, форму и данная система становится способной приступить к работе. Однако, применение подобных систем крайне ограниченно в связи с тем, что для успешного функционирования подобной системы требуется четкая формализация базы знаний, что далеко не всегда удается сделать. Например, в области инженерии не редко приходиться руководствоваться плохо формализуемыми «ноу-хау», что делает создание экспертной системы в этой области крайне затруднительным, если вообще возможным делом. Экспертные системы воплощают ещё одну основополагающую идею касательно интеллекта, а, именно, что интеллект есть определенная система знаний, однако, этот подход также оказывается не убедительным, так как само по себе наличие знаний не является показателем интеллектуальности. Более того, знания могут оказаться ещё и не востребованными и просто лежать мёртвым грузом в закромах памяти. Хотя хорошо организованная база знаний и является, как мы увидим в дальнейшем, необходимым условием интеллектуальной деятельности, однако, этим понятие интеллекта отнюдь не исчерпывается.

Энтузиасты бионического направления занимались разработками и совершенствованием нейронных сетей, генетических алгоритмов, а позднее созданием роботов и проблемами машинного зрения. И хотя на начальных этапах своего развития из-за малых мощностей исследования в этой области особо не финансировались, однако, по мере совершенствования технологии, удалось достичь весьма неплохих результатов. Если первая модель персептрона как модели нейрона не справлялась даже с законом

исключенного третьего, то более мощные многослойные нейронные сети способны решать эти, а также намного более сложные задания с высокой степенью эффективности. Главной особенностью нейронных сетей является их способность к обучению. Посредством процесса обучения, мы можем постепенно сделать нашу сеть способной распознавать образы, проводить классификацию, решать логические задачки и тому подобное. Давайте рассмотрим принцип работы нейронной сети на примере одного нейрона.[20] Нейрон может находиться в одном из двух состоянии: возбуждение, которое мы обозначим «1» или торможения, которое мы обозначим «0». На вход нейрону поступает некий набор входных сигналов **X.** У каждого входного сигнала есть свой вес **W**, характеризующий степень воздействия. Далее, эти сигналы суммируются с учетом весов и преобразуются в сигнал **Y.** Потом функция пороговой активации **F** преобразует сигнал **Y** в сигнал **Z –** сигнал выхода. Функция **F** может быть представлена в двух видах:

1) Линейная функция – $Z = K(Y)$, где К — постоянная;

2) Пороговая функция – $Z = 1$, если $Y > a$, $Z = 0$ в противном случае, где а – постоянная пороговая величина.

Такая система обучается посредством коррекции весов входных сигналов, которая определяется расхождением выхода от заданного результата. Принцип работы многослойных сети остается таким же, с тем лишь отличием, что входные сигналы проходят через несколько слоев нейронов с результирующим выходом. Также для обучения нейронных сетей используется принцип обратного распространения ошибки, который также выполняет функцию коррекции. Однако, подобная система, опирающаяся на принципы работы реальных нейронов, оказывается способна решать поставленные перед ней задачи, не позволяет нам понять каким образом был получен результат, который нам выдала система. Из-за этого возникает главная проблема нейронных сетей: они демонстрируют интеллектуальное поведение, но при этом, ни слова не говорят нам об устройстве нашего мышления. Аналогичной выглядит ситуация и в нейробиологии. Огромное

количество работ по исследованиям работы мозга, непрекращающиеся по сей день, так и не приблизили нас к пониманию мышления, более того к пониманию работы мозга как единого целого.

Обращаясь к работе Дж. Хокинса «Об интеллекте», [46] мы обнаруживаем ряд очень значимых идей касательно устройства нашего мозга. Первое, на что обращает внимание автор – на однообразие организации нейронов, на всех уровнях. По мнению автора, нейронные сети живого мозга организованны как иерархическая система с обратными связями, которые присутствуют на каждом уровне. Этим, например, достигается высокая эффективность в обработке информации от сенсорных данных. Также в работе изложена концепция «память-предсказание», которая выступает в качестве основного принципа работы мозга. Наш интеллект осуществляет свою деятельность на основе предсказаний ожидаемых входных данных. Таким образом, мы получаем ясно представление о том, что интеллектуальная деятельность предполагает упреждающий характер. И хотя данная модель, безусловно, не может претендовать на полноценное объяснение работы мозга и мышления, однако, она показывает нам, что интеллектуальная деятельность связана с принципом обратных связей.

Рассмотрение бионического направления не может быть полноценным без рассмотрения идей генетических алгоритмов. Генетические алгоритмы призваны имитировать процесс естественной эволюции. И так, вначале мы задаем некоторую популяцию, которая может быть интерпретирована как множество вариантов решения. Далее посредством скрещивания и селекции, мы выявляем наиболее приспособленные алгоритмы и отбираем их, пока входе повторения этой процедуры мы не получим необходимое решение. Такого рода модель может быть использована при моделировании развития мышления, чтобы показать, как из простейших умственных операций формируются более сложные. Данный подход также успешно реализуется в рамках направления «искусственная жизнь», представители которого пытаются изучать развитие жизни посредством компьютерных моделей.

Однако, в плане объяснения феномена интеллекта, эти алгоритмы вряд ли дадут нам исчерпывающую информацию.

Теперь мы переходим к рассмотрению такого направления в моделировании мышления как «Адаптивное поведение». Данное направление отчасти является продолжением кибернетики Виннера[9], в котором органично переплетаются различные наработки из таких областей как психофизиология, этология и прочие смежные с ними дисциплины. Основной моделью этого направления служит модель «организм-среда». Как мы уже видели, обращаясь к работам Пиаже, здесь интеллект представлен как показатель степени адаптации к среде. Организм получает информацию от внешней среды и выдает некоторую реакцию на эту информацию. Суть адаптивного процесса сводиться к оптимизации взаимодействия со средой, построение модели наиболее эффективного поведения. Стоит отметить, что данное направление берет своё начало из бихевиоризма и поначалу опиралось, на модель «черного ящика», т. е. рассмотрение модели можно было свести таблице функциональных связей стимулов с реакциями. По мере развития данного направления, организм стал обладать целями, потребностями, жизненными параметрами и т.п. Именно, в рамках этого направления, мы впервые сталкиваемся с агентом, однако, рассмотрение понятия агент и его возможные модели будут рассмотрены нами позднее. В данном русле, также происходит разработки адаптивных роботов.

Теперь подведём итоги нашего рассмотрения области искусственного интеллекта. Мы должны отметить, что в настоящее время развитие этой области связанно с интеграцией представленных нами ранее направлений. Активно начинаются применять гибридные системы, объединяющие нейронные сети, генетические алгоритмы и элементы моделирования рассуждений, пришедшие из рационалистической парадигмы. [49] При этом интеграция происходит не только в рамках самого направления, но и в плане его взаимодействия с другими науками. Если раньше искусственный интеллект был изолированной областью исследований, то теперь в него

активно проникают достижения психологии и биологии, развивается математический аппарат исследований. При моделировании используются методы математической статистики, теории вероятности, нечеткой логики. Так что можно сказать, что данное направление оформилось в самостоятельное научное направление, носящее междисциплинарный характер.

Г. Саймон, в своей работе «Науки об искусственном» [], вводит разграничение наук на те, что изучают явления, происходящие в природе, т.е. естественные науки, и на те, что изучают искусственные, созданные человеком объекты или науки об искусственном. Саймон установил четыре признака отличающие искусственное от естественного:

1. Искусственные объекты конструируются человеком.

2. Искусственные объекты могут внешне походить на естественные объекты, но существенно отличаться от последних в одном или нескольких аспектах.

3. Искусственные объекты можно охарактеризовать через их функции, цели, степень приспособленности к среде.

4. Искусственные объекты часто, особенно на уровне проектирования, рассматривают не только в описательных категориях, но и в нормативных терминах.

При этом, искусственный интеллект попадает в разряд наук об искусственном. Таким образом, нам приходиться переосмыслить саму суть искусственного интеллекта, который становится мыслящим артефактом. О необходимости разграничения исследований искусственного и естественного интеллекта говорится и в работе Рассела, Норвинга[37]. Исходя из этого, мы можем сделать вывод, что по мере развития направления искусственный интеллект был произведен скачок, от имитационных моделей человеческого интеллекта к качественно новым объектам, а, именно, интеллектуальным системам.

В таком случае, обращаясь к известному разграничению на слабую и сильную трактовку искусственного интеллекта, мы должны добавить ещё одну: искусственный интеллект есть самостоятельный объект изучения, который не является ни подобием, ни точной моделью человеческого мышления. Однако, это несколько не умаляет значение компьютерных моделей при исследовании человеческой психики. Мы просто должны с самого начала определиться с тем, что мы хотим исследовать: модель мышления реального человека или же модель мышления искусственного агента.

И так, рассмотрев основные теории интеллекта и ключевые подходы к его моделированию, мы можем заключить, что интеллектуальная система и естественный интеллект, это различные объекты исследования, но при этом они обладают пожими структурами. В обоих случаях мы имеем сложноорганизованную структуру мыслительных процессов. В принципе, в обоих случаях интеллект обладает процессуальным характером. Основным отличием этих систем является способ их происхождения. Естественный интеллект является следствием биологической эволюции и исторического развития социума, а искусственный интеллект является мыслящим артефактом, сконструированным человеком. В следующем разделе, мы рассмотрим понятие агента, а также коснемся вопросов, которые связаны с конструированием и классификацией искусственных агентов.

1.2. Когнитивный агент как интеллектуальная система

И так, что же такое агент? Для начала представим себе двуполярную шкалу. На одной стороне этой шкалы находиться субъект, а на другой стороне – объект. Если мы попробуем идти к пониманию агента от полюса субъект, то тогда агент у нас будет выступать как представитель субъекта. Если же начать движение от объектного полюса, тогда наш агент окажется «активным объектом», т.е. таким объектом, который способен осуществлять манипуляцию с другими объектами, а если мы немного расширим его

возможности, то он будет способен формировать собственные модели поведения, которые будут определяться собственными потребностями и направляться на достижение собственных целей. В первом случае агент преимущественно будет выступать как помощник человека, который обладает некоторым уровнем свободы, регулируемым пользователем. Во втором случае агент становиться автономным роботом, способным осуществлять самостоятельное решение задачи в проблемной среде. Теперь давайте рассмотрим, какие существуют трактовки понятия «агент» в информатике.

Первая трактовка восходит к трудам Дж. Холланда [40], где агент интерпретируется как искусственный организм, который осуществляет своё развитие в популяции себе подобных и стремиться к обучению и адаптации в целях выживания. Данная трактовка опирается на подходы и модели, разработанные в таких направлениях как эволюционное моделирование и искусственная жизнь. При данном подходе, имитируются такие процессы, как мутация и трансформация агентов, борьба за существование, принцип естественного отбора, а также самосохранение и репликация искусственных организмов. Данная трактовка нашла своё отражение в робототехнике, при проектировании интегральных роботов или групп роботов.

Вторая трактовка, основана на метафоре агента как помощника пользователя, а в дальнейшем агент стал трактоваться как посредник между пользователем и средой, в которой он работает. Данная трактовка соответствует движению от субъектного, при котором агент как бы замещает субъекта в виртуальной среде. Также к данной трактовке можно отнести также интерпретацию агента как автономного или как полуавтономного программного модуля, способного осуществлять взаимодействие с пользователем, а также подстраиваться под его запросы. Это позволяет создавать удобную форму пользовательского интерфейса, который становится не только дружественным, но и персональным. Данная идея берет своё начало из идеи работы человека и компьютера в режиме диалога, а также в разработках по созданию интеллектуальных интерфейсов.

Третья трактовка, относит агента к классу активных объектов, способных осуществлять манипуляции с объектами (например, создавать, уничтожать, перемещать), а также взаимодействовать со средой и другими агентами. Таким образом, агент представляет собой введение нового класса объектов, а значит, мы имеем дело с расширением объектно-ориентированного подхода в программировании. Так, с одной стороны, мы вводим неоднородные объекты, а также задаем построение иерархии программных модулей через их активность и автономность. В итоге, мы получаем расширение классической концепции объектно-ориентированного программирования. А с другой, наличие протоколов коммуникации между различными агентами, позволяет конкретизировать и детализировать данный подход.

Помимо таких общих трактовок, можно привести и более частные. Так, например, С. Рассел и П. Норвинг считают, что в качестве агента может выступать любая систем, способная воспринимать окружающую среду посредством датчиков и осуществлять воздействие на неё с помощь исполнительных механизмов. [39] Общая схема представления агента может быть задана следующим образом. Итак, агент представляет собой систему, которая способна получать информацию о внешнем мире или среде посредством сенсоров и оказывать соответствующее воздействие посредством эффекторов. К данной схеме можно также добавить два внутренних блока, а именно блок обработки информации от сенсоров и блок хранения информации или память. Таким образом, простейший агент ничем не отличается от простейшей схемы «организм – среда» [11]. Однако, следует ввести ряд дополнительных свойств необходимых для полноценного описания агента, а именно:

1. активность, выраженная в способности планировать свои действия и осуществлять их, согласно этим планам;
2. автономность, выраженная в самостоятельном осуществлении своей деятельности;

3. коммуникативность, которая характеризует способность агента к общению с другими агентами;
4. целенаправленность, которая характеризует поведение агента как направленное на достижение цели.

Теперь мы можем немного уточнить нашу модель. Совокупность сенсоров формирует систему восприятия, которая должна осуществлять прием и передачу первичных данных от среды. Источником сигналов может выступать как внешний мир, так и внутренние состояния агента. Затем, обработанная информация переходит в блок памяти. При этом, система восприятия может оказывать регуляторное воздействие на поведение посредством сопоставления текущих и ожидаемых состояний. Далее может осуществляться корректировка действий в зависимости от величины расхождения. Также должна присутствовать система фильтров, которая позволить выделять наиболее значимую информацию.

В блоке памяти должна храниться информация о типичных реакциях на внешние стимулы, также там должны содержаться данные о состоянии эффекторов и имеющихся в наличии ресурсов, программа, перерабатывающая входные данные в команды для эффекторов, записи о реакции на то или иное внешнее воздействия. Помимо всего выше сказанного агент может обладать собственной моделью внешнего мира и моделью самого себя. В зависимости от того, какой объем знаний храниться в памяти агента, от того насколько развитой окажется его модель мира и того насколько он способен к рефлексии, определяется возможность осуществлять им более сложные манипуляции с внешним миром и демонстрировать более сложные модели поведения.

Процессоры задействуют лежащие в памяти программы переработки данных и осуществляют их интеграцию, вырабатывают реакции на среду и генерируют решения. Выбор соответствующей последовательности действий при заданных ограничениях является одной из самых важных характеристик для агента любой степени сложности.

Эффекторы должны осуществлять воздействие на внешнюю среду, а также стремиться поддерживать стабильность во внутренней среде. Также для полноценного функционирования агента требуется так же наличие источника питания, который позволяет поддерживать и даже воспроизводить жизненный цикл агента.

В качестве подобных агентов можно представить интегральных и интеллектуальных роботов [16]. Такие машины, обычно обладают широким набором сенсорных датчиков (зрительные, звуковые, тактильные, кинестетические) и искусственных манипуляторов. Также эти роботы обладают определенным уровнем мобильности, который достигается за счет присоединения к роботу колес, гусениц, шагающих систем передвижения. Активность, автономность и интеллектуальность робота обеспечивается наличием у него системы целепологания, генератора последовательности действии, решателя задач, а также развитыми средствами коммуникации и средствами обработки знаний. Набор сенсоров может быть расширен за счет использования температурных датчиков, определителей влажности, радиоактивности и прочих физических датчиков.

Можно представить ряд основных требований к функциональным возможностям систем подобного рода:

1. Способность осуществлять планирование поведения при наличии нескольких возможных направлении решения.
2. Способность к приращению собственного знания и обобщению наличного опыта, последующий перенос его в новые условия функционирования.
3. Соотнесения различных целей с соответствующими способами их достижения.
4. Выбор стратегии поведения наиболее релевантной данной ситуации.
5. Активное влияние на внешний, по отношению к агенту, мир и возможность свободно перемещаться в нем на значительные расстояния.

Теперь перейдем к рассмотрению некоторых конкретных проектов по созданию искусственных агентов. Одним из таких проектов является модель гиромата, предложенная в трудах Д. А. Поспелова и В. Н. Пушкина [34]. Гиромат выступает как элементарная модель целесообразного поведения, которая способна подстраиваться к условиям решаемой задачи. Данная модель уже включала в себя некоторые, важные для агента, блоки: блок мотивов, блок содержащий опыт модели, блок выработки гипотез, блок активного опыта, блок времени, блок модели внешнего мира и блок селекции, который выполняет функцию отбора.

Функционирование гиромата осуществляется в специальной среде, которая была названа дискретной ситуационной сетью [34]. В этой среде могли появляться, исчезать и перемещаться некоторые объекты. Сама система описывалась как граф, имеющий вершины трех типов: истоки (S), решатели (D) и стоки (F). Все три указанных множества являются конечными. Любая из вершин графа может выступать в качестве представителя одного из данных множеств. Также вводиться некоторое множество объектов O, а также некоторое множество атрибутов AT. Истоки осуществляют генерацию некоторого множества объектов внутри дискретной ситуационной сети, задачей стока является удаление объектов из сети, решатели же могут находиться в одном из двух возможных состояний: активном или пассивном. Если решатель пассивен, то это приводит к задержке на несколько тактов. Если же решатель находиться в активном состоянии, то происходит, либо смена атрибутов объекта, либо объект продолжает движение, либо осуществляются оба предложенных варианта.

Общий принцип работы гиромата можно представить следующим образом. Информация о текущей ситуации в пространстве состояния ситуационной сети поступает в блок селекции. Отсюда происходит распределение по двум следующим блокам: блок модели внешнего мира, в котором происходит классификация ситуации, и в блок выработки гипотез, в котором делаются предположения относительно взаимосвязи между

ситуациями. Далее, во временной блок поступает информация о времени фиксации сигнала в сети. Впоследствии эти данные позволяют проанализировать повторяемость конкретных ситуаций, а также отследить изменения во времени. Модели внешнего мира является аналогом опыта дискретной ситуационной сети, который был получен в ходе её функционирования. Эта база данных выступает источником для блока модельного опыта, в котором осуществляет подборку информации для принятия решений, которая попадает в генератор решений, задающий программу блоку активного опыта. Если ситуацию типичная то программа задается сразу, если же нет, то происходит дополнительное обращение к блоку гипотез. Кроме того, блок активного опыта может непосредственным образом фиксировать реакцию, полученную от среды с последующей передачей в блок гипотез, что позволяет осуществлять корректирующее воздействие.

Также стоит рассмотреть две отечественных разработки искусственных агентов с автономным поведением. Это Транспортный Автономный Исследовательский Робот или, сокращенно, ТАИР, разработанный школой Амосова Н.М. [1]. И программа «Животное», разработка которой осуществлялось под руководством М. М. Бонгардом. [3] Управление сетью ТАИР осуществлялось с помощью искусственной М-сети, описание которой представлено ниже. Эта система была создана с опорой на исследования по устройству и функционированию нейронных систем живых организмов. Разработчики системы основывались на допущении, что любое возможное состояние искусственного организма может быть создано посредством отдельных блоков, которые впоследствии будут объединены в единую структуру. Каждый такой блок получил название i-модели, эти блоки также могли осуществлять процедуру, которая производит объединение паттернов в единый образ. Взаимодействие между блоками осуществлялось посредством обмена сообщениями, которые интерпретировались как

возбуждение или как торможение. В итоге сеть из этих блоков получила название М-модели.

Для осуществления передвижения в пространстве, данная система производил разбиение на восемь секторов, которые задаются системой координат «вперед – назад», «влево – вправо». Таким образом, для определения местоположения цели системе было необходимо учесть расстояние и нахождение в одном из секторов. Система же управления представляла собой шесть блоков, являющихся участками М-сети, и набор различных рецепторов. Первоначально в систему входили 14 рецепторов, набор блоков представлял собой следующий список:

1. блок распознавания;
2. блок оценки;
3. блок решения;
4. блок маневра;
5. блок навигации;
6. блок руления.

Блок распознавания представлял собой множество i-моделей, каждая из которых представляла некоторую типовую ситуацию внешнего мира, например, «цель находиться слева за стеной», «стена справа», «тупик» и т.д.

Блок оценок содержал все возможные оценки, которые машина могла вынести относительно состояний внутри себя и состояний окружающей среды, относительно имеющихся ресурсов и осуществленных действий. Решение о том, какие следует предпринять действия, определялось опытом оценок полученных на прошлых шагах, на основе данных блока распознавания.

В программе «Животное» помимо имитации самого агента осуществлялась также имитация среды его действия. При этом набор показаний рецепторов впоследствии использовался для построения обобщенных описаний ситуации. Среда обитания представляла собой плоскость, разбитую на отдельные ячейки. В рамках модели среды,

функционировали разнообразные законы, представлявшие собой имитацию природных жизненных условии (ограничения на число объектов, на скорость передвижения агента). В арсенале «Животного» имелись рецепторы, необходимые для получения информации о среде, а также интеррецепторы, с помощью которых осуществлялось воздействие на среду. Показания рецепторов формировались в обобщенное представление ситуации следующим образом. Каждая ситуация определялась набором связанных с ней предикатов, которые оказывались в истинном состоянии. Так, предикат P(x, a) оказывался истинным при условии, что x > a. Например, число объектов x, которые могут являться пищей для агента, выше минимального порога a. Таким образом, информация, полученная от рецепторов и соотнесенная с данным предикатом, описывает ситуацию утоления голода. Предикат Q(y) – представляющий собой опасного врага, добавляющийся к предикату P(x, a), при установлении приоритета Q(v) (убегать) над P(x, a) (утоления голода), то «Животное» определяло ситуацию как опасную и покидало данную местность. В случае, если программа находилась в состоянии голода, то она осуществляла свои действия с поправкой на значение соответствующих предикатов.

Каждый поведенческий акт программы был связан с преобразованием исходной ситуации в желательную ситуацию. Данная модель «ситуация – действие», была создана с помощью системы правдоподобных продукционных правил, которые функционировали с учётом изменения оценки достоверности продукций.

Далее мы рассмотрим несколько примеров многоагентных систем, которые являются альтернативой автономным агентам. Так К. Хьюитт [42], разработав систему PLANNER, предназначенную для доказательства теорем, раскрыл всю важность коммуникации и управления в плане их влияния на организацию и понимание рассуждений. Он использовал распределенную структуру работы с информацией, которую рассматривал в качестве «образцов прохождения сообщении». Эти сообщения свободно

циркулировали между активными объектами, которые были названы акторами. Актор определялся Хьюиттом как программный агент, имеющий свой почтовый адрес и обладающий собственной манерой поведения.

В результате, была предложена и реализована концепция разработки программ, имитирующих деятельность сообщества экспертов, которые вместе решают какую-нибудь задачу и в ходе её решения активно коммуницируют посредством передачи друг другу посланий. Интересно, что в данной модели, происходило столкновение разнообразных позиций по отношению к выполняемому заданию. В итоге, родилось семейство языков акторов, которые стали первым средством реализации многоагентных систем. В них общение между агентами осуществляется посредством посылки асинхронных сообщений.

Одним из первых ученых, начавший распространять представления о различных психических процессах (или ментальных состояниях) на искусственные объекты, и трактовать ментальную сферу как следствие взаимодействия между активными объектами, стал М. Минский [23]. Он дал развёрнутое описание целого ряда механизмов, которые приводили к проявлению интеллектуальных форм поведения в ходе возникновения конфликтных ситуации и сотрудничества между элементарными вычислительными единицами, которые именовались им агентами. Каждой подобной единице соответствует какое-либо ментальное состояние и при этом их взаимодействие имеет случайный характер, тем самым, исключалось наличие командного центра управления. Это позволило М. Минскому прейти к одному важному заключению, а именно, что осуществление мыслительных операции основывается не столько на выводах, базирующихся на символьной системе, а на принципах самоорганизации, которые проявляются в ходе взаимодействия автономных процессов.

Первые западные разработки практического характера, связанные с многоагентными системами, были преимущественно посвящены проблематике распределённого искусственного интеллекта и

интеллектуальным агентам. Они относятся к 70-м гг. и связанны с именами В. Лессера, К. Хьюитта и Д. Лената. [42] Так, работы В. Лессера, Л. Эрмана и Ф. Хэйс-Рота [42] с системой HEARSAY II привели к возникновению архитектуры «доска объявлений». А также эти работы позволили в дальнейшем углубить разработки, связанные с организацией коммуникативного пространства агентов. Идея «доски объявлений» базируется на исследованиях, которые предприняли авторы, изучая проблемы автоматического понимания речи. Решая любую проблему, мы вынуждены прибегать к помощи разнообразных источников знаний или экспертов в интересующей нас области, при этом эти обращения не являются заранее продуманными. Структура управления процессом коммуникации также заранее не задана. В данной модели источники знаний выполняют функции доставки, преобразования и извлечения объектов с доски объявлений, т.е. из зоны совместной работы в базе данных. Где модель предметной области структурирована как пространство гипотез и решений. Специальное устройство управления разрешает конфликты доступа к доске объявлений, возникающие между агентами и неявным образом организует их совместную работу.

Примерно в это же время, У. Корнфелд и К. Хьюитт [42], которые занимались разработкой системы ETHER, создали подобный подход к решению совместных задач, который был основан на метафоре научного сообщества. В этой системе распределенного искусственного интеллекта, «демоны», во многом похожие на источники знания, доставляют факты, гипотезы и доказательства в общую область, похожую на «доску объявлений». Отличительной чертой этой модели является разделение демонов на два класса – эксперты и критики. Задача экспертов заключается в поиске доказательств в пользу той или иной гипотезы. А критики должны осуществлять поиск контрпримеров, которые позволили опровергнуть эти гипотезы. Специальные управляющие агенты распределяют время, которое предоставляется каждому демону.

Другой тип организации управления взаимодействия агентов был предложен Д. Ленатом и К. Хьюиттом []. В середине 70-х гг., Д. Ленат, который впоследствии стал разработчиком системы EURISCO и выступил с инициативой проекта создания больших баз знаний CYC, создал систему под названием PU PS. В основе этой системы была заложена идея решения некоторой задачи группой агентов, которые именовались «существа». Эти агенты направлены на создание особого «существа», который может порождать концепции и при этом способен к самостоятельному решению поставленной задачи. Сами специалисты постоянно меняются в процессе решения задачи и поэтому не могут быть отнесены к разряду источников знания. Каждый из них создаёт своё собственное множество пар «атрибут-значение» и может обращаться за сведениями к другим специалистам, не зная их лично. В дальнейшем, Д. Ленат построил демонстрационный прототип системы с взаимодействующими «существами».

Р. Смитом [] была предложена своя модель распределённого решения задач, получившая название «контрактная сеть», которая и сегодня имеет большой резонанс. В данной модели, используется схема, основанная на ведение переговоров между автономными интеллектуальными агентами, в роли которых выступают узлы сети. Данная модель является своего рода имитацией рыночных торгов. В ней присутствуют три различных класса агентов. Первый из них, агент-менеджер, осуществляет распространение информации об объявлении и назначает базовую стоимость. Второй класс составляют агенты исполнители или покупатели, которые делают предложения своих услуг и выставляют собственные расценки, за счёт чего могут участвовать в формировании наилучшего предложения, связанного с исходным заданием. Далее агент-менеджер должен произвести отбор между кандидатами с наиболее приемлемыми предложениями, а затем заключить с избранниками контракты, после чего, те переходят в разряд подрядчиков (третий класс агентов).

На основе представленной выше модели, можно представить себе разнообразные схемы, по которым ведутся переговоры, которые будут представлять собой последовательность контпредложений, поступающих от агентов в ходе их взаимодействия. Этот процесс можно представить как последовательную цепь переходов во множестве возможных состояний.

Теперь мы попробуем осуществить классификацию существующих моделей агентов. Подобная классификация может быть основана на множестве разнообразных оснований. Одним из таких оснований может выступать разбиение агентов на две разновидности: естественные и искусственные. К естественным агентам будут относиться животные, человек, а также группы, которые они образуют. Соответственно, в разряд искусственных агентов у нас попадают всевозможные роботы, автоматы и компьютерные программы.

При этом всех искусственных агентов, мы также можем разбить на два группы. К первой группе относятся физически реализованные роботы, которые функционируют в режиме реального времени и также с реальным пространством. В качестве примера можно привести интегральные модели роботов, которые обладают «органами чувств», а также средствами для манипуляции с объектами.

Ко второй группе можно отнести таких агентов, которые способны существовать только в виртуальной среде. Данных агентов можно представить как своего рода роботов, осуществляющих работу с информацией. В качестве примера можно привести «программных роботов» или сокращенно софтботами.

В качестве основания для классификации могут также выступать такие параметры как подвижность и стационарность, сосредоточенность и распределенность. В качестве примеров для первой пары параметров могут выступать поисковый бот, осуществляющий поиск в компьютерной сети релевантной информации и промышленный робот-манипулятор. Мобильные софтботы выступают представителями распределенных агентов, которые не

обладают инструментами взаимодействия со средой в виде сенсоров и манипуляторов, однако они обладают развитым протоколом коммуникаций для взаимодействия с другими агентами, а также они могут свободно перемещаться виртуальном пространстве, выполняя свою задачу. Сосредоточенные или локализованные агенты ограничены в плане своего перемещения, не обладают знаниями о среде, а всё их общение носит в основном опосредовано дополнительными приборами.

Одним из важнейших оснований для осуществления классификации агентов может выступать способность агента к обучению или адаптации. Если агент способен обучаться, то его модель поведения будет опосредована предыдущим опытом.

По мнению Поспелова Д. А., [34] можно заниматься построением классификации агентов, а также подбирать необходимые средства для их формализации, основываясь на трёх параметрах: представление о среде, уровень развития автономии, степень социализации агентов. Если мы имеем дело с простыми замкнутыми средами, то в них вполне успешно функционируют простые автоматы, если же мы имеем усложнённый вариант замкнутой среды, то нам, соответственно, понадобиться использовать более сложные модели автоматов, которые обладают неким конечным списком правил и набором алгоритмов их применения. Подобные агенты активно применяются для работы в компьютерных сетях при использовании клиент-серверных технологий. В принципе, если среда имеет замкнутый тип, а значит и конечное описание, то взаимодействие агента со средой может представлено в виде релевантного логического исчисления. При работе с открытыми средами требуется использование сложных семиотических моделей.

Также можно разделить искусственных агентов на тех, которые основаны на использовании психологической метафоры, и на тех, которые основаны на биологической метафоре. В первом случае, мы интерпретируем агента как модель субъекта, который способен решать поставленные перед

ним задачи самостоятельно. Во втором случае, агент приравнивается к простому организму, который осуществляет взаимодействие с внешней средой, реагируя на внешние воздействия и адаптируясь к ним. На основе уподобления естественным агентам были построены М-автоматы Н. М. Амосова [11], а также были созданы так называемые «аниматы», которые выступали в качестве модели животного, адаптирующегося к изменениям окружающей среды с целью выживания. При этом аниматов удалось реализовать как в виртуальной среде, так и в виде роботов, работающих в реальных средах.

Подобное разделения основано на классической проблеме взаимодействия субъекта и объекта. Когда мы пытаемся построить модель субъективного уровня, то можно это сделать, опираясь исключительно на символьную форму представлений, необходимых для построения рассуждений. Или же наш агент будет работать в основном на уровне образных представлений, которые напрямую связанны с организацией собственного взаимодействия со средой с помощью сенсоров и эффекторов. В таком случае, мы способны также провести разграничение ещё по двум параметрам, а именно: а) уровень развития представлений агента касательно внешнего мира; б) то, каким образом устроено его поведение.

На основе первого разграничения, мы можем выделять агентов интеллектуального типа и агентов реактивного типа. Так, например, когнитивные агенты, относящиеся к первому типу, имеют хорошо развитую модель внешнего мира, которая также постоянно обновляется, что оказывается возможным за счёт наличия у агента базы знаний, алгоритмов принятия решения и наличия обратной связи для контроля и анализа действий. Ещё одним вариантом интеллектуального агента является рассуждающий, который способен на основе своей модели внешнего мира, имеющей символьную форму, осуществлять некоторые формы рассуждений в соответствии с заданными алгоритмами их построения. А затем, основываясь на полученных выводах, принимать решение и взаимодействовать с

окружающей его средой, которая, в свою очередь, способна постоянно меняться.

Интеллектуальные агенты могут различаться между собой в зависимости того, на какие когнитивные функции мы делаем основной упор, например, на необходимость сбора информации о среде или на построение некоторых рассуждений касательно будущих действий. Агенты коммуникативного типа, используют модель внешнего мира, в которую включены участники коммуникации и набор наиболее желательных результатов, для построения своей модели общения с другими агентами. Ну и наконец, ресурсный тип интеллектуальных агентов, обладает базой данных о ресурсах, которая содержит информацию об их структуре и их состоянии, что, в дальнейшем, оказывает непосредственное влияние на последующее поведение агента.

Если же мы намерены создать полноценного интеллектуального агента, то в таком случае он должен обладать всеми свойствами, которыми обладают все вышеперечисленные агенты, а именно, он должен обладать развитой картиной мира, алгоритмами построения рассуждений на основе символьной информации, обладать развитым протоколом коммуникации, а также владеть знаниями о ресурсах.

В свою очередь реактивные агенты [9] не обладают каким-либо развёрнутым представлением об окружающей среде, не умеют строить сложные цепочки рассуждений, а также не обладают большим объёмом собственных ресурсов. Таким образом, становиться понятным ещё одно принципиальное различие, которое имеется между агентами реактивного типа и агентами интеллектуального типа. Это различие в способности прогнозировать последующие изменения среды, а, следовательно, и своего будущего состояния. Реактивные агенты обладают весьма скудными способностями в плане планирование собственных действий и прогнозировании. Такие агенты реагируют на непосредственное воздействие, и структура их обратной связи полностью исключает использование

механизма предвидения. В отличие от них, интеллектуальные агенты, основываясь на сложной структуре своих представлений о внешнем мире и развитых методов рассуждения, способны сохранять в своей памяти и проводить развёрнутый анализ встречающихся ситуаций, предсказывать ответные действия со стороны внешней среды в зависимости от его поведения. В том числе, они могут делать выводы на основе обратной связи, которые могут быть использованы для осуществления дальнейших действий, тем самым он может планировать модель своего будущего поведения. Благодаря столь развитым внутренним ресурсам, подобные агенты оказываются способны конструировать виртуальное пространство, в рамках которого они осуществляют планирование.

У интеллектуальных агентов, имеющие высокую степень автономии по сравнению со своими реактивными собратьями, имеются разнообразные индивидуальные свойства, проявляется целесообразное поведение в ходе взаимодействия с другими агентами и, в тоже время, они могут использовать ресурсы, имеющиеся у других агентов, ради удовлетворения собственных потребностей. У реактивных агентов в этом плане индивидуальные черты почти отсутствуют, и они очень сильно зависимы от параметров внешней среды, т.к. их поведение основано на схеме «стимул-реакция».

Интеллектуальные агенты могут быть также подвергнуты разделению на интенциональный тип и рефлекторный тип, в то время как реактивные агенты имеют две разновидности: побуждаемые и трофические. Многих интеллектуальных агентов можно отнести к интециональному типу [13], потому что они обладают мотивированным поведением. В структуре поведения таких агентов большую роль играют такие элементы, как наличие убеждений, желаний, мотивов, которые оказывают непосредственное влияние на формирование целевой ситуации и подборе средств, позволяющих её достичь.

В тоже время, у рефлекторных агентов отсутствуют подобные регуляторы действий, что приводить к использованию ими примитивных

моделей вывода, которые полностью автоматизированы и жёстко детерминированы внешними стимулами от окружающей среды.

Это позволяет нам сделать вывод, что рефлекторные агенты являются вырожденным случаем когнитивного агента, которые способны выполнять преимущественно вспомогательные функции. Они сопоставимы по своим возможностям с акторами, поскольку способны отвечать на внешние запросы и даже выполнять поручения от других агентов, но при этом они абсолютно не способны вырабатывать свои собственные цели. К агентам подобного рода можно отнести такие системы как поиск нужной информации в некоторой базе данных или же простые логические регуляторы.

Однако, такие системы обладают некоторой формой знаний о необходимом поведении, которая представлена в сжатом виде. Они не создают внутри себя какой-либо «картины мира», а обходятся лишь относительно небольшим списком реакции на текущие внешние запросы, т.е. они функционируют на основе знаний о данном положении дел.

В зависимости от степени сложности ответных действий и от характера источников стимуляции реактивные агенты могут обладать либо импульсивной формой поведения, либо трофической формой поведения. Если речь идёт о трофической форме поведения, то тогда поведение целиком и полностью определяется пищевыми связями, что в свою очередь сводиться к набору ответов на стимулы в локализованной ситуации. К таким автоматам относятся клеточная их разновидность. В их работе участвует целый ряд параметров: радиус обзора у агента, объём единиц пищи, наличествующих во внешней среде, и количество затрачиваемой энергии на действия. Данные агенты ограничены небольшим объёмом правил, которые соответствуют сложившейся ситуации и которые задают ответное действие на стимулы. Например, если в поле зрения агента попадает единица пища, то он направляется к ней, а если таковой не обнаружено, необходимо передвинутся в любой из соседних квадратов в произвольном порядке.

При этом у реактивных агентов всё же может присутствовать некоторая упрощенная схема мотивирования, которая заставляет их выполнять какие-либо задачи с целью поддержания определённых жизненных параметров на должном уровне. Это может быть поддержание энергетического уровня в указанных рамках или поддержание гомеостазиса или автопоэзиса, с целью обеспечить своё собственное выживания [21,51].

Когнитивные агенты, за счёт своего уровня сложности, а также благодаря обладанию определённым набором знаний и навыков рассуждения касательно собственного поведения или параметров окружающей среды, обладают высокой степенью автономии по сравнению с реактивными агентами. Они способны осуществлять самостоятельное поведение, которое имеет весьма гибкий характер. Однако, эта сложность, которой обладают автономные агенты, проявляющаяся в способности противодействовать внешним воздействиям. Это в свою очередь приводит к определённым сложностям при организации их взаимодействия между собой. Поэтому при создании многоагентной системы, составленной исключительно из агентов интеллектуального типа, следует ограничить количество входящих в неё автономных элементов.

В данном случае, реактивные агенты, обладающие более простой структурой и при этом жестко детерминированные внешней средой, обладают весьма ограниченными способностями, поэтому объём их возможностей весьма невелик и ограничен своими незначительными ресурсами. Из-за этого таким агентам выгоднее объединиться в группу, которая смогла бы проявить высокую адаптивность за счёт регуляции с помощью естественного отбора. Именно поэтому таких агентов интереснее изучать не на уровне индивидуумов, а на групповом уровне, причём их способность к адаптации проявляется и получает развитие на основе случайных индивидуальных взаимодействий. В результате этого, эти безликие агенты, объединенные в единую систему способны решать задачи высокого уровня сложности. Подобные многоагентные системы могут

образовываться как результат случайных взаимодействий, в которых отсутствует какая-либо дифференциация агентов по функционалу. Данные системы представляют собой так называемый «роевой интеллект»[40], который представляет собой совокупность большого числа мобильных, реактивных агентов, образующих при своём взаимодействии нечто вроде гигантского организма. Так как все члены данного роя приспосабливаются друг к другу и образуют единую сеть обратной связи, то в результате образуемая ими система обладает высокой степенью устойчивости и согласованности внутри себя. [40]

Разделение на реактивных и когнитивных агентов проистекает из двух различных направлений в исследованиях искусственного интеллекта, о которых уже упоминалось ранее, а именно рационалистический и коннекционистский подходы. Проведя сравнительный анализ интеллектуальных и реактивных агентов, мы приходим к выводу, что для обеспечения наибольшей автономии при разработке агентов, необходимо использовать смешанные варианты архитектуры, которые бы объединяли в себе достоинства представителей обоих направлений. Например, можно объединить между собой методологические принципы логического подхода в искусственном интеллекте и принципы построения нейронных сетей, чтобы получить интегрированную модель. Такие модели проявляют большую гибкость и эффективность. []

Интеллектуальные агенты могут иметь разную направленность относительно своего поведения в плане обмена информацией. Одни агенты могут стремиться поставлять исключительно правдивую информацию, другие же наоборот пытаются любым способом дезинформировать других агентов. Подобное разделение можно произвести, опираясь на то, как агент относиться к другим агентам, каковы его намерения по отношению к ним. При таком подходе мы можем выделять агентов-эгоистов и агентов-альтруистов, а также разделять их на злонамеренных и на благонамеренных агентов. Например, благонамеренным агентом является такой агент, который

всегда выполняет те требования, которые к нему предъявляют другие и при этом всеми возможными способами пытается избежать попадания в конфликтную ситуацию.

И теперь мы рассмотрим последний вариант классификации. В этой классификации помимо психологического аспекта и биологического аспект, рассматривается ещё и социальный аспект. В результате, мы получаем три разновидности агентов: реактивные, интенциональные и социальные. Первые характеризуются выполнением простых стереотипных действий в зависимости от поступающей информации, а также умением отправлять сигналы в среду и для других агентов. Агенты интенционального типа умеют формировать цели, продумывать свои действия наперёд, а также вступать в коммуникацию с целью достичь исполнения собственных желаний. Социальные агенты обладают помимо развитой модели окружающей среды, которая наличествует у предыдущего вида агентов, ещё и моделями других агентов, что говорит об их способности к рефлексивному управлению, а также обладают направленностью на создание коллектива агентов. Уровень сложности их внутренних моделей прямо пропорционален объёму опыта, накопленного агентами, а также от объёма накопленных им знаний о внешнем мире.

Итак, мы рассмотрели общие представления о том, что такое быть агентом, а также рассмотрели возможные классификации агентов и некоторые модель их конструирования.

Глава 2. Когнитивные агенты и их взаимодействия

2.1 Архитектуры интеллектуальных агентов

В этой части работы нами будут рассмотрены несколько разнообразных вариантов архитектуры программных агентов. А начнем мы наше рассмотрение с общей классификации архитектур. [44] Возьмём в качестве основания для классификации такой критерий как способ происхождения. На основании этого критерия можно выделить два типа архитектуры. Первый тип архитектуры основан на классических методах искусственного интеллекта, в которой интеллектуальная система является системой обработки и представления знаний. Второй тип архитектуры основан на формировании системой своего поведения в зависимости от того, какие события она наблюдает во внешней среде.

К архитектуре первого класса можно отнести таких агентов, которые построены на основе продукционной системы. Эта система может быть задана как набор правил, каждое из которых позволяет нам установить соответствие между условием из заданного списка и конкретным действием из заданного списка. При этом список условий непосредственно связан с имеющимся в системе набором фактов, в свою очередь, список действий состоит из набора простых действий, например, переместить объект из ячейки 1 в ячейку 2.

Любой интеллектуальный агент, который представляет собой продукционную систему, должен быть способен осуществлять такие операции как интерпретация и выполнение. Хотя эта система и имеет определённое сходство с экспертными системами, однако есть и существенные отличия. Во-первых, агент способен осуществлять взаимодействие с внешней средой, и для этих целей он снабжен такими инструментами как рецепторы, позволяющие получать информацию о состоянии окружающей среды, а также специальными устройствами для

манипуляции с внешними объектами, которые называются эффекторами. С помощью органов восприятия обеспечивается обновление базы знаний системы, что позволяет в дальнейшем учитывать эти данные в ходе работы системы правил. Во-вторых, в структуре интеллектуального агента также должны присутствовать ещё два важных модуля. Первый из них должен обеспечивать возможность коммуникации агента с другими агентами, второй модуль отвечает за выработку таких составляющих поведения как целевое состояние, которое должно быть достигнуто, а также формирование оценок ситуации, что в совокупности образует базис простой интенциональной системы.

Дж. Холландом были предложены системы особого типа, которые основывались на базе классификаторов[40]. В данных системах любой факт представлен в виде вектора, имеющего строго зафиксированный размер, правила, в качестве которых выступают классификаторы, представляющие собой пару векторов чисел. К этой паре добавляется некой дополнительное число, именуемое «джокером». Все правила обладают весовыми коэффициентами, которые определяют степень вероятности запуска некоторого правила.

Вся информация, которую система получила от среды, преобразуется в блоке восприятия в набор фактов. К ним осуществляется применение правил, в результате которого образуются новые факты. Эти факты могут, как привести в действие новые классификаторы, таки взаимодействовать с внешней средой с помощью исполнительной подсистемы.

В отличие от продукционных систем, системы, основанные на использовании классификаторов, имеют две дополнительные подсистемы. Первая из них осуществляет вознаграждение, а вторая запускает воспроизводство правил, которые базируются на генетических алгоритмах. Правила, позволившие достичь желаемого результата, подвергаются «вознаграждению», т.е. происходит увеличение весового коэффициента. А если правило привело к отрицательному результату, то тогда, соответственно,

коэффициент уменьшается. Правила, обладающие наибольшим коэффициентом весов, обладают большим приоритетом при запуске. Также такие системы способны осуществлять генерацию новых правил, которые создаются с использованием правил с наибольшим весом посредством особых операторов, которые способны осуществлять операции скрещивания, мутации и т.п. Подобные системы нередко используются для реализации адаптивной модели поведения, например, для адаптации аниматов к динамическим средам.

Следующая архитектура интеллектуального агента предполагает использование детализированной спецификации некоторой базы знаний. Подобная система состоит из нескольких уровней представления знаний, а также рабочей памяти, специального модуля, управляющего коммуникацией, и человеко-машинного интерфейса. Для того, чтобы агент мог эффективно решать стоящие перед ним задачи, он должен обладать как навыками планирования, проработки последовательности дальнейших действий, разделение общей задачи на подзадачи, так и уметь оперативно реагировать на поступление новых данных, а также на изменения в уже имеющихся. В данной системе интеллектуализация поведения достигается с помощью использования алгоритмов принятия решений, которые осуществляют выбор плана действий, распределяют задачи, а также благодаря алгоритмам кооперации, которые осуществляют выработку набора обязательств. Кроме этого система снабжена особым контролирующим устройством, способным мгновенно реагировать на любые изменения происходящие в рабочей памяти. Например, можно отслеживать информацию о появлении новых целей, получать своевременную информацию обо всех изменениях в наличествующих данных, а также об изменениях обязательств и задач.

Можно выделить три пласта знаний, которые играют наибольшую роль в данной архитектуре агента:

1. Знание конкретной предметной области.

2. Знания об организации взаимодействия. Сюда входят правила регулирующие поведение, а также правила, по которым происходит увеличение и изменение знаний в предметной области. Также эти правила можно разделить на те, которые позволяют генерировать решения в условиях неопределённости и те, благодаря которым осуществляется контроль над кооперацией агентов.

3. Управляющие знания, которые осуществляют применение знаний из области взаимодействия к знаниям о предметной области, за счёт чего обеспечивается процесс запоминания и преобразования информации в рабочей памяти.

В блоке рабочей памяти содержатся данные временного характера, которые могут туда поступать из различных источников, таких как блок управления, блок управления коммуникациями, а также непосредственно от пользователя. Там же содержатся необходимые данные о цели агента, о ходе выполнения текущих задач, о том, какие задачи завершены, вся поступившие и отправленные послания, а также имеющиеся в наличии обязательства. Рабочую память в этой системе можно уподобить глобальному рабочему пространству.

С помощью модуля управляющего коммуникацией производиться формирование и последующее отправление посланий, предназначенных для других агентов. Эти послания образуются с помощью набора примитивов, которые, в свою очередь, вырабатываются в блоках управления задачами и кооперацией. Любой примитив обладает своим собственным типом, а также специфическим содержанием. Также этот модуль предоставляет информацию о доставке сообщений другим агентам.

С помощью человеко-машинного интерфейса регулируется взаимодействие данной системы с пользователем. Таким образом, система оказывается способной к выполнению целого ряда специфических функций, однако её действия не обладают достаточной автономией. В большинстве случаев, пользователь получает отчёт о проделанной агентом работе и он же

решает необходимость подтверждения его результатов, до того как они будут переданы следующим агентам.

В архитектуру типичного представителя семейства интеллектуальных агентов с необходимостью должны входить следующие составляющие:

1. язык, предназначенный для общения с другими агентами, который основан на теории речевых актов и осуществляет передачу знаний в декларативном виде (KQML, KIF);

2. подсистему, отвечающую за распределение соответствующей информации между агентами, а также в неё должны быть встроены такие функции как модификация имеющихся у агентов убеждений, обработка знаний дедуктивными методами;

3. организационная модель, обеспечивающая всё необходимое для самостоятельной работы агента в многоагентной среде. Она должна содержать в себе представления агента о самом себе, о других известных ему агентах, об их роли в системе, о структуре коммуникации с ними и т.д.;

4. координирующая модель, вбирающая в себя все правила и условия обмена информацией между агентами при реализации их взаимодействия, например, в виде правил кооперации и общих оценок текущего положения дел;

5. подсистема разрешения конфликтов, с помощью которой агент способен осуществлять принятие решений в условиях противоречивой информации, которая поступает от других агентов;

6. подсистема управления знаниями, основывающаяся на дескриптивных логических моделях, способная осуществлять темпоральные рассуждения и иные формы обработки.

При создании агентов с использованием искусственных нейронных сетей, используется некоторое количество однородных элементов, которые выступают в качестве формальной модели нейрона. Подобная архитектура агентов позволяет обеспечить им высокую обучаемость. Используя нейронную сеть, состоящую из трёх слоёв, в которой рецепторы

взаимодействуют с эффекторами через промежуточный слой, мы способны создать модель реактивного агента, который может успешно адаптироваться к среде. Более широкий спектр возможностей можно обеспечить, если мы будем использовать сети с обратными связями, позволяющих на уровне одного слоя осуществлять передачу данных между нейронами.

Большой интерес представляют такие нейронные сети, которые способны работать с нечёткой информацией, а также способны обрабатывать нечёткие сигналы и работать с нечёткими весами. В тоже время можно использовать сети, сконструированные из нечётких нейронов, которые можно использовать при создании логических моделей агентов.

Помимо этого существует хаотическая разновидность нейронных сетей, которая представляет собой объединение в единую сеть нейронов с пороговой системой активации и нейронов с логической функцией выхода. С помощью данных сетей можно моделировать сложные режимы поведения агента в динамике, например, распознавать образы, которые раньше были неизвестны агенту.

Если говорить в общем, то в наше время существуют такие нейронные сети как, например, сети Кохонена, Фукушими, адаптивного резонанса, с помощью которых можно успешно реализовывать такие процессы как самоорганизация агентов и их самообучение.[39] Архитектура данных сетей должна иметь многоуровневый характер для того, чтобы служить адекватной моделью самоорганизации. Так, например, на нижней ступеньке мы будем полагать, что поведение нейронов имеет хаотический характер. На средней ступени будут образовываться ансамбли нейронов, направленные на выполнение конкретных функций. На высшей ступени возникают резонансные эффекты, которые приводят к возникновению адаптационных механизмов и механизмов кооперации.

Распространённым вариантом архитектуры, используемой для создания реактивных агентов, является модель «соподчинения». Данная модель предполагает наличие двухуровневой структуры, в которой на верхнем

уровне происходит управление, а нижний уровень при этом находиться в подчинении у верхнего. При этом процессы на обоих уровнях происходят параллельно друг другу. Все виды взаимодействия, возможные между уровнями, строго определены и находятся в соответствии заранее определённому принципу, согласно которому один уровень подчиняется другом. Так, например, если при одновременной работе уровней возникают противоречия между уровнями, то тогда верхний уровень может заблокировать работу нижнего.

Работа в многоуровневых архитектурах во многом зависит от способа организации взаимодействий между уровнями данной системы. Поэтому можно выделить две базовые разновидности подобной организации: вертикально организованная и горизонтально организованная.

Если в основе системы лежит горизонтальная организованная архитектура, то в такой системе на любом уровне наличествует взаимосвязь с блоками восприятия и действия, что позволяет всем уровням взаимодействовать друг с другом. Если же архитектура имеет вертикальную организацию, то тогда восприятие и действие возможны только на одном уровне, в то время как все остальные уровни могут взаимодействовать лишь с ближайшими соседями.

При горизонтальной организации возникает проблема координации работы всех уровней структуры, что связано со сложностью организации этого взаимодействия. Подобная проблема не является столь весомой для вертикальной организации, поскольку в ней передача информации с уровня на уровень всегда имеет конкретного адресата, к которому эта информация должна поступить. Существует два принципа, согласно которым будет происходить распределение функциональных модулей по уровням, если речь идёт о вертикальных структурах. В соответствии с первым принципом, каждый уровень, которым обладает агент, представляет собой отдельный уровень абстракции, которая надстраивается над некоторым списком характеристик функционирования. При втором способе организации, каждый

уровень системы представляет собой определённое функциональное свойство или же соответствует нескольким подобным свойствам. К недостаткам такого способа организации можно отнести излишняя нагрузка, которая имеется на уровне действий.

Теперь мы рассмотрим более детально конкретные модели основанные как на горизонтально организованной архитектуре, так и на вертикально организованной

Существует особый тип архитектуры агента, которая в своём базовом варианте рассчитана на одного агента, у которого имеется одна единственная цель. Данная архитектура именуется will-архитектурой, в которой используются идеи и метафоры, которые вполне традиционны для описания интеллектуальной деятельности человека, что, в свою очередь делает её понятной и привлекательной. В рамках данной подхода к созданию агента, полностью исключаются из рассмотрения такие аспекты деятельности агента как его коммуникация и кооперация с другими агентами. При этом агент направляет все свои усилия на достижения своих собственных целей.

Успешная деятельность агента в ходе его функционирования в рамках окружающей его среды предполагает наличие у него некоторого набора базовых характеристик, к которым относятся наличие у него сенсорной системы, а также блока действия. При will-архитектуре любая функция, которой обладает агент, должна быть реализована в виде отдельного блока. В качестве основных блоков, помимо сенсорного блока и блока действия должен присутствовать ещё блок, отвечающий за планирование. Основной проблемой здесь является то, каким образом должна быть организована их совместная работа, например, как скоординировать поток поступающей информации с поступающими командами от блока управления. Для того чтобы максимально скоординировать потоки информации используется схема «бродкастинага», при которой все имеющиеся входы отдельных модулей связаны со всеми имеющимися у них выходами. Это приводит к тому, что любой сигнал, источником которого может выступать совершенно

произвольный блок, является доступным для любого другого блока, входящего в эту систему. При этом все сгенерированные сигналы или сообщения попадают в особый блок, который выступает в качестве блока памяти. Обращаться к нему могут все имеющиеся модули, за исключением сенсорного, и, в тоже время, любой блок может сохранять в нём новую информацию, кроме блока действия.

Также данная система предполагает возможность порождать и модифицировать цели в «ментальном пространстве» агента, которые, в свою очередь, определены его «интересами». Под интересами подразумевается наличие у агента стремления находиться в некоторых определённых состояниях и избегать каких-либо других состояний. Каждое возможное состояние среды, в которой находиться агент, должно быть каким-либо образом соотнесено с его интересами и получить соответствующую оценку. Таким образом, когда какой-либо модуль обращается к блоку памяти, то в поле его зрения автоматический попадает та её часть, степень соответствия которой обладает наибольшей оценкой, а затем эта часть обрабатывается тем модулем, от которого был получен запрос. Стоит обратить внимание, что для того чтобы добиться оптимального функционирования агента, обладающего такой архитектурой, требуется чтобы он обладал высокой степенью развития самоорганизации.

Следующий вариант архитектуры, который мы рассмотрим, это так называемая InteRaP-архитектура. Суть такого рода архитектуры в том, что модель агента состоит из некоторого множества уровней, взаимосвязь которых обеспечивается с помощью специальной подсистемы управления, а также эта система обладает базой знаний общего типа. Такой агент обладает определённым интерфейсом, который позволяет ему общаться с внешним миром. Также он должен обладать подсистемой работающей по принципу «стимул-реакция», блоком планирования, специальным блоком, который позволяет организовывать взаимодействие с другими агентами в виде

кооперации и, наконец, базой знаний, которая имеет иерархическую структуру.

Его интерфейс включает в себя такие составляющие как специальные средства для регистрации объектов и событий во внешнем мире, средства для взаимодействия с этим миром и средства, обеспечивающие его коммуникацию с другими агентами.

Блок «стимул-реакция» обеспечивает возможность осуществлять реактивную форму поведения. Данный блок основывается на таком понятии как «фрагмент поведения». В него входят все, заранее заданные агенту, типовые формы поведения, которые соответствуют набору стереотипных ситуаций. За счёт этого блока, агенту не требуется обращаться к блоку планирования, если он оказался в некой типовой ситуации, а использовать уже имеющиеся у него алгоритмы поведения, обеспечивающие максимальную эффективность в заданных условиях.

В блоке планирования осуществляется разработка плана индивидуальных действий агента. Такой план может быть представлен в виде графа, у которого в качестве узлов могут выступать как конкретные действия, так и подпланы, которые требуют дальнейшего уточнения. Такая система планирования осуществляет активацию некоторой формы поведения, которая будет соответствовать некоторым целям. Этот же блок участвует в разработке кооперативной модели поведения агента, осуществляя её планирование.

Система, организующая кооперативное поведение агента, задействована в таких процессах как разработка плана взаимодействия агента с другими агентами, который направлен на достижения совместных целей, контролирует выполнение агентом обязательств по отношению к другим агентам, осуществление принятых соглашений.

База знаний агента разделена на три уровня иерархии и организована таким же образом, что и «доска объявлений». Каждому уровню этой иерархии соответствует определённый уровень абстракции знаний, при этом данная структура находиться в полном соответствии структуре управления.

На первом уровне располагаются все знания агента, которые связаны с его реактивным поведением. Именно этот уровень задействуется блоком «стимул-реакция». На втором уровне моделируется ментальное пространство агента, в котором присутствуют цели агента, его планы и интенции. Этот уровень задействуется блоком планирования наряду с нижестоящим уровнем. На третьем уровне располагается всё то, что относиться к знаниям агента относительно других агентов, к взаимодействию с ними, общие планы, цели и т.п. Этот уровень задействуется блоком организации кооперативного поведения, который также имеет доступ ко всем предыдущим уровням. В совокупности, все эти уровни задают модель мира агента, основываясь на которой он действует.

За счёт постоянного взаимодействия между уровнями иерархии знаний осуществляется полноценное управление действиями агента. Когда агент сталкивается с некоторой ситуацией, он соотносит её с имеющейся у него моделью мира. При этом происходит постепенный переход о нижнего уровня к верхнему, до тех пор, пока система не выйдет на тот уровень, который позволит эффективно действовать в сложившейся ситуации.

Исходя из этого, у агента имеется три альтернативных варианта ответа на возникшую ситуацию. Если решение достижимо на первом уровне то, тогда с помощью реактивного блока агент подбирает ответную реакцию релевантную данной ситуации. Если не удалось решить проблему на первом уровне, то мы переходим на следующий и используем блок планирования для того, чтобы спланировать ответное действие. И, наконец, мы можем перейти на третий уровень и задействовать блок кооперативного поведения. Таким образом, мы всегда имеем возможность подобрать оптимальную реакцию на возникшую извне ситуацию.

Теперь мы переходим к рассмотрению многоуровневой архитектуры агента по типу машины Тьюринга. Данная модель позволяет создать автономного агента с использованием некоторых моделей рассуждения на основе знаний, которые объединены с моделями поведения реактивного

характера. Она является одной из разновидностей горизонтально организованной архитектуры агентов, которая была специально создана для разработки подвижного робота. [39]

В данной модели предполагается, что агент (в качестве которого здесь выступает робот), способный вступать во взаимодействие с другими агентами, сталкивается с непредвиденными ситуациями в ходе своей деятельности, которые обладают фрагментарностью и неточностью. В подобных ситуациях, он должен осуществлять релевантные ответные реакции и принимать эффективные решения. Это означает, что в архитектуре агента должна быть учтена возможность его работы в условиях неопределённости, оперируя при этом с неточной, фрагментарной информацией. При этом его реакция на неожиданности должна опираться на набор относительно простых правил. Для реализации всех этих требований, была разработана специальная трёхуровневая структура агента, в которой каждый уровень представляет собой определённую разновидность способностей агента. В эту структуру входят следующие уровни:

1. уровень реакции на события R, который включает в себя способность мгновенно реагировать на некоторые события, определяемы на более высоком уровне, даже если эти события не были предвидены заранее;

2. уровень планирования P, который обеспечивает способность агента к разработке частичных планов, а также их исполнение и модификацию в динамическом режиме;

3. уровень предсказания или моделирования M, на котором создаётся модель внешней среды, в которой определен характер поведения объектов в мире и самого агента, и на её основе, прогнозируется изменение характера этого поведения с течением времени.

На любом из этих уровней имеется доступ к сенсорной системе и эффекторам, а также на любом из них может приниматься решение о необходимости реагировать на текущее положение дел. При этом решения не зависят друг от друга. Также при данной архитектуре имеется блок

управления, состоящий из набора правил, который активируется в зависимости от контекста и позволяет достичь оптимальной модели поведения даже при возникновении конфликтов уровней. Этот блок выполняет посреднические функции, работая с имеющимися данными на различных уровнях, а также может поставлять новую информацию на уровни или же удалять её с уровней. Он продолжает своё функционирование на протяжении всего времени деятельности агента и при этом никак не сказывается на работе уровней, которые продолжают свою работу, как бы ни замечая существования дополнительного блока. Для синхронизации входных и выходных сигналов уровней, также используется данный блок. С помощью его правил фильтруется информация, поступающая от сенсорной системы к системе уровней, а также та информация, которая исходит от уровней к эффекторам.

В заключении, рассмотрим многоагентную систему DESIRE, которая является воплощением композиционной архитектуры. Эта система позволяет сделать описание сложного агента максимально «прозрачным». С помощью неё можно также осуществлять интеграцию рассуждений и действий в едином логическом пространстве.

В данной системе, агент, осуществляющий какую-либо форму деятельности, способен производить следующие разновидности действий:

1. воспринимать и обрабатывать поступающие данные от среды;

2. строить на основе полученных данных выводы;

3. участвовать в коммуникации с другими агентами, стремясь достичь сотрудничества с ними;

4. формировать и трансформировать свои убеждения, на основе принятия или неприятия дополнительных посылок;

5. воздействовать на среду, намериваясь её трансформировать.

При этом агент обладает своеобразной базой знаний, которая состоит из представлений о внешнем мире, знаний о своём ментальном пространстве и подобном пространстве у других агентов, навыков взаимодействия со средой

и навыков, необходимых для осуществления коммуникации с другими агентами. Эта база знаний имеет два блока, в одном из которых информация остаётся в неизменном виде, несмотря на то, в каком состоянии находиться система. Вторая часть содержит переменные, которые изменяют свои значения с течением времени.

Основная идея, заложенная в эту архитектуру, заключается в том, что для создания модели сложного агента можно использовать отдельные блоки, каждый из которых представляет собой подсистему общей системы. Каждый компонент подобной системы имеет своё собственное описание и опирается на свою систему знаний, а также осуществляет решения определённого типа задач. Эффект сложного поведения достигается за счёт того, что в ходе сложного динамического взаимодействия составляющих этой систем, достигается возможность осуществления сложных действий и рассуждений. В качестве блоков в многоагентной систем используются отдельные автономные агенты.

2.2 Проблема эффективности взаимодействия когнитивных агентов в многоагентной системе.

Организация взаимосогласованных действий является первоочередной задачей при построении системы, в которой участвует коллектив агентов. В ходе её решения оказывается жизненно важным анализ того, как будут организованны сложные и многогранные способы взаимного сотрудничества между членами сообщества. Эти формы сотрудничества являются основой для построения определенной устойчивой структуры. В тоже время они являются следствием этой же самой структуры. Таким образом, взаимосогласованные действия являются необходимым условием для возникновения таких устойчивых образований как коллективы или целые сообщества агентов.

Сотрудничество предполагает не только наличие связности и взаимного влияния между различными агентами внутри сообщества, но и возможность

его трансформации, которая может быть выражена как в виде изменения самих членов сообщества, так и формы их взаимодействия. В качестве примера можно привести смена роли агента в рамках сообщества: он может перейти из разряда исполнителей поручений в разряд организатора выполнения работ, что автоматически приведет к изменению его взаимоотношении с другими членами коллектива.

Для полноценного сотрудничества необходимо, чтобы внутри системы, состоящей из множества различных агентов, наблюдалось наличие хорошо организованных и согласованных между собой действий элементов данной системы, а это в свою очередь порождает систему социальных ритмов, которые позволяют системе сохранять свою целостность и при этом быть способной к реорганизации.

Основными характеристиками, которые должны присутствовать при любом возможной форме взаимодействия, являются:

1. направленности;
2. избирательность;
3. степени интенсивности;
4. динамичности.

Всеми этими характеристиками также должны и обладают коллективы агентов или многоагентные системы.

В первую очередь любая форма взаимосогласованных действий агентов имеют некоторую направленность. Она может носить как положительный характер и приводить к усилению связей между агентами, так и отрицательный характер, что влечет за собой возникновения различных форм конфронтации внутри коллектива. Таким образом, результатом этой направленность может быть кооперация агентов ради выполнения некоторой общей задачи, так и конкуренция или даже конфликт. Если преобладает позитивная направленность, то мы видим рост консолидации между агентами, подавление собственных интересов во имя интересов сообщества, также наблюдается наличие приспособленности агентов друг к другу, умение

при необходимости задействовать умения и навыки друг друга. Если же доминирует негативная направленность, то в таком случае возникают конфликты и борьба интересов внутри сообщества, своекорыстные интересы начинают вытеснять общие цель, агенты начинают мешать друг другу в достижении собственных целей.

Кроме этих двух крайних форм взаимодействия агентов, возможно возникновение всевозможных промежуточных вариантов направленности. В качестве примера такого рода явлении может выступать антипатия двух агентов, в результате которой они стремятся избегать каких-либо форм сотрудничества и взаимодействия. Ещё одним вариантом данного явления выступает однонаправленное содействие, в ходе которого один агент оказывает помощь в достижении цели другому агенту, а тот в свою очередь всячески избегает какого-либо взаимодействия с первым. Подобная ситуация возникает в образовательном процессе, когда один агент, выступающий как учитель, пытается чему- то научить другого, в данном случае ученика, но тот упорно отказывается обучаться и лишь сохраняет видимость понимания материала, который ему преподноситься. Также возможна такая ситуация, что один агент оказывает противодействие другому агенту, а тот в свою очередь не желает вступать во взаимодействие с первым и стремиться избегать встречи с ним. И, наконец, могут возникать ситуации, в которых налицо явно противоречие. Например, один агент стремится помочь другому, а тот в ответ оказывает противодействие этим благородным начинаниям.

При анализе взаимодействия агентов, а также групп агентов уместным является рассмотрение концепции групповой динамики, которая была предложена К. Левином [18]. Эта концепция основывается на теории поля того же автора. Исходя из данной теории, если взять одного агента и его непосредственное окружение, то мы получаем психологическое поле, а если мы возьмем за основу группу и то, что её окружает, то тогда мы имеем дело с социальным полем. Таким образом, групповое поведение может быть описано как функция общего состояния социального поля в любой момент

времени. Само это состояние напрямую зависит от конкурентных отношений между подгруппами данной группы и отдельными агентами в ней, а также от различных ограничении и путей общения. При описании групповой динамики в контексте данной теории, также используется такое понятие как валентность, которое оказывается весьма близким к понятию направленности. Валентность можно охарактеризовать через притяжение и отталкивание между отдельными агентами. Так, например, положительная валентность показывает нам, что некоторые агенты обладают направленностью в определенный район социального поля, в обратном случае, мы видим движение в противоположную сторону.

Давайте теперь рассмотрим такие виды взаимодействия как кооперация и конкуренция. Относительно кооперации можно выделить следующие подвиды данного типа отношении: союзничество, которое представляет собой аддитивный эффект при объединении агентов, и партнерство, которое является мультипликативным эффектом при объединении агентов. Важным параметром при оценке степени кооперации выступает её объём, т.е. какое количество агентов относиться к данной кооперации. Конкуренция же выражается через отрицательное влияние на степень объединения между агентами, в то время как конфликт может привести к нарушению гомеостаза внутри группы, что фактически означает её развал.

Теперь перейдем к следующей характеристике взаимодействия в группе – избирательности. Когда формируется система, состоящая из нескольких агентов, то в рамках данной системы взаимодействие между агентами осуществляется лишь в том случае, когда агенты некоторым образом соотносятся друг с другом, а также со стоящей перед ними задачей. При этом агенты могут быть взаимосвязаны в одних аспектах и независимы в других аспектах. Также некоторые агенты могут осуществлять взаимодействовать между собой с большей легкостью, нежели другие, и наоборот. Так, например, взаимодействие между агентами облегчается за счет их близости относительно пространства, за счет определенного сходства между ними,

наличием или отсутствием предварительного знакомства до вступления в группу и т.п.

При всем при этом взаимосвязи между агентами и их зависимость относительно друг друга нельзя просто свести к их наличию или отсутствию. Они также могут быть охарактеризованы посредством такого понятия как интенсивность, которое выступает в качестве показателя силы этих связей. Это позволяет нам установить, с какими агентами в данной группе интересующий нас агент взаимодействует в большей степени, с какими в меньшей степени. Обычно, каждый агент находиться во взаимодействии с несколькими агентами в один момент времени, при этом интенсивность этих взаимодействии может быть весьма разнообразной.

Теперь мы приступаем к обсуждению последнего из заявленных нами параметров, а именно к динамичности. В различных ситуациях, могут происходить изменения таких параметров как интенсивность и направленность в плане взаимодействия, а также можно наблюдать эскалацию или же наоборот дезэскалацию сотрудничества и конфликта. Подобного рода явления можно наблюдать, когда при наличии совместного пользования некоторыми ресурсами происходить усиление взаимозависимости между агентами, при условии что перед ними встаёт сложная задача, требующая их активного сотрудничества. Также посредством переговоров, можно достичь определенного компромисса и даже сближения целей, что в свою очередь повлечёт за собой изменения в режиме взаимодействия, например, переход от активного противодействия к интенсивному содействию.

Наличие норм относительно взаимодействия друг с другом, выступает в качестве важнейшей составляющей взаимодействия агентов. В качестве такой нормы может выступать некоторое соглашение, принятое членами какой-либо группы. Обычно подобные нормы обладают весьма устойчивым характером, так как соглашения обычно строятся на таких принципах, которые позволяют свести к минимуму возможность уклониться от их

выполнения. В качестве примеров подобного рода гарантии могут выступать всевозможные санкции, которые могут быть представлены как в виде поощрения, так и виде наказания.

Также стоит отметить, что когда мы имеем дело с эгоистичными агентами, то соглашения между ними обладают взаимовыгодным характером. Это достигается за счёт обеспечения максимизации этим соглашением функции полезности для всех агентов, которым оно предложено. Но при этом оказывается, что наличие взаимовыгодных условий соглашения является необходимым, но не достаточным условием для сохранения его устойчивости в группе агентов. Поэтому при построении многоагентной системы, необходимо вводить не только эгоистичных агентов, но и альтруистичных агентов, что позволит сделать модели их взаимодействия более гибкими, а значит и более устойчивыми. Альтруистичные агенты, которые при взаимодействии учитывают групповые интересы, намного превосходят своих эгоистичных собратьев, когда речь заходить об обеспечении автономии и устойчивости системы.

И так давайте теперь сформулируем основные задачи, которые необходимо решить при анализе взаимодействия агентов в рамках многоагентной системы:

1. Определить ситуации, в которой происходит взаимодействие.
2. Выявить основные роли агентов в системе и того как они распределены между ними.
3. Выявить количества агентов, находящихся во взаимодействии.
4. Установить какие разновидности взаимодействий существуют между агентами.
5. Построить формализованную модель соответствующих взаимодействий.
6. Выяснить какие стратегии поведения возможны в данной модели для агентов.
7. Определить набор коммуникативных действий.

В общем, необходим многоуровневый анализ всего многообразия возможных взаимодействий между агентами. Любая сложная реальная ситуация, в которой происходит взаимодействие, может быть представлена в виде набора более простых ситуаций. Так, например, необходимо разграничивать те ситуации, в которых учитываются все взаимодействия между агентами в многоагентной системе (макроситуации), и те ситуации, в которых имеют значения лишь небольшая группа локализованных взаимодействии (микроситуации). Такое разделение позволит нам определить значимость тех или иных видов взаимодействия, а также понять какое место они занимают в изучаемой нами системе агентов. Так, например, можно определить наличие или отсутствие корреляции между уровнем развития сотрудничества между агентами и локальным соперничеством между агентами.

Теперь мы должны рассмотреть вопрос о критериях, которые позволят определять различные типы ситуации. В качестве таковых могут выступать такие показатели как активность или пассивность агентов, ориентация на собственные цели или учёт целей других агентов, осуществление единичных или групповых действий.

Если действия агента активны, то это может привести либо к сотрудничеству с другими агентами, либо к конкурентной борьбе с ними. Для возникновения первой ситуации необходимо чтобы агент был способен к ориентации на совместные цели, и отдавал бы предпочтение совместным действиям, а не индивидуальным. Иначе возникнет ситуация конкурентных взаимоотношений, которые могут перерасти в конфликт. Если агент стремиться достигать только свои цели и не склонен к сотрудничеству, то конфликт переходит в активную конфронтацию агентов, которая будет вестись до полной победы одной из сторон. Если же агент способен принять во внимание цели других агентов, то тогда он, скорей всего, предпочтёт компромиссную стратегию, позволяющую достичь удовлетворения целей всех участников взаимодействия за счёт достижения общей цели.

Когда речь идет о пассивной форме поведения, то здесь подразумевается поведения направленное на избежание конфликта. К такому типу поведения можно отнести конформизм и приспособление. При таком типе поведении происходит отказ от собственных целей в пользу целей другого агента, наблюдается стремление максимально подстроиться под воздействия извне. Так, например, при выборе стратегии солидарности формируется такой тип взаимодействия, при котором групповые действия не взаимосвязаны с активным поведением агента и не приводят к образованию общих целей. А если речь идет об уклонении, то здесь налицо полнейший отказ от любого взаимодействия.

И так, представим себе такую ситуацию. Перед нами пара агентов, которые обладают набором несовпадающих целей, но при этом они способны осуществлять переговоры и приходит к соглашению за счёт принятия взаимных обязательств. В таком случае, мы имеем некоторое множество различных переговорных ситуаций, в зависимости от свойств которого, образуются следующие типы: симметричная кооперативная ситуация, ассиметричная кооперативно-компромиссная ситуация, симметричная компромиссная ситуация, конфликтная ситуация.

Если ситуация является симметричной кооперативной, то тогда во множестве переговорных ситуаций имеется подмножество соглашений, выгодных для обеих сторон. При таких условиях, каждому агенту из рассматриваемой нами пары осуществить достижение собственной цели проще посредством взаимодействия с партнером, в отличие от одиночных действий. Соответственно, выбирается стратегия сотрудничества между агентами.

Если же ситуация симметричная компромиссная, то тогда каждому агенту гораздо удобнее действовать самостоятельно и избегать взаимодействия. Но в силу принятых на себя обязательств или каких-либо иных внешних причин, они вынуждены каким-то образом сосуществовать

друг с другом, преимущественно за счёт компромиссных соглашений. При этом данная стратегия оказывается выигрышной для обоих агентов.

Если один агент считает, что их взаимодействие является сотрудничеством, а другой полагает, что это компромисс, то тогда перед нами ассиметрично кооперативно-компромиссная ситуация. Конфликтная же ситуация возникает в том случае, когда невозможно перейти в такую ситуацию, при которой возможно соглашение между агентами.

В третьем случае возникает элемент асимметрии: один агент расценивает взаимодействие как сотрудничество, а другой — как компромисс. Наконец, конфликтная ситуация возникает тогда, когда переговорное множество пусто: соглашения между агентами невозможны.

Теперь мы выделим основные критерии, необходимые нам для того, чтобы охарактеризовать процесс формирования взаимодействия в многоагентной системе. В качестве таковых мы выделяем четыре основополагающих критерия:

1. Совместимость целей или намерений агентов.
2. Отношение агентов к ресурсам.
3. Опыт агентов относительно некоторой проблемной области.
4. Обязательства агентов друг перед другом.

Теперь мы можем приступить к рассмотрению каждого из выше перечисленных критериев в отдельности. И начнем мы с рассмотрения целей, которыми обладают агенты, а также их совместимости или несовместимости. Например, в случае полной несовместимости, удовлетворение целей одного агента приводит к устранению возможности достижения аналогичного результата другим агентом. В частности, в ситуации «хищник – жертва», удовлетворение цели, преследуемой хищником, автоматически приводить к гибели жертвы. Таким образом, становиться очевидным важное условие для кооперации между агентами. Это условие можно сформулировать следующим образом: время жизни одного агента не должно уменьшаться в случае его взаимодействия с другим агентом.

Подобная модель может выступать, как показательны пример наличия в одном жизненном пространстве двух различных стратегий: сотрудничества и конфронтации. Хищники сотрудничают между собой для успешной охоты за жертвами, которые находятся в состоянии конфронтации с ними

Теперь давайте рассмотрим отношение к ресурсам. Ресурсом может выступать любое средство пригодное для осуществления какой-либо цели, например, выживание агента. Ограничение количества ресурсов приводить к возникновению конфликтов. Они возникают из-за того, что для некоторого числа агентов наличествуют потребности в одинаковых ресурсах в данный момент времени. В качестве примера можно привести «пробки» на дороге или конфликт двух одновременно запущенных программ. В такой ситуации взаимодействие агентов носит деструктивный характер, потому что один агент мешает другому.

Существуют разнообразные стратегии действий, направленные на устранение подобных ситуации или предупреждение их. Они опираются на различные способы координации действий или разрешения конфликтов. Одним из простейших механизмов разрешения конфликта касательно ресурсов сводиться к тому, что тот агент, который обладает большей силой относительно другого, получает в итоге ресурс. Менее агрессивными средствами являются переговоры, посредством которых достигается компромиссное решение.

Одной из важных причин, приводящих к взаимодействию между агентами, служит несоответствие между требованиями, предъявляемыми к задаче, и теми возможностями, что есть у агентов. В таком случае, возможны три альтернативных варианта:

1. Агент в состоянии справиться с поставленной задачей самостоятельно, т.к. его возможностей вполне достаточно для выполнения.

2. Агент в состоянии справиться с задачей сам, однако группа справиться с этим заданием эффективнее.

3. Агент не в состоянии справиться с возложенной на него задачей сам, потому что он не обладает необходимыми возможностями для этого.

Любая социальная система обладает некоторой структурой обязательств, в основном достаточно сложной. Эти обязательства могут быть представлены в виде договоров, соглашении и прочих подобных вещей, тем самым выступая в качестве необходимого основания для осуществления совместных действий.

Эти соглашения могут регламентировать любые формы взаимоотношений внутри социума. Например, они могут касаться целей, которые ставятся перед агентами («я должен перенести объект s в точку p»). Или же они могут определять то, каким образом может быть выполнено то или иное действие, роль, выполняемую в социуме, и т.п.

Отсюда вытекает необходимость наличия некоторых обязательств между агентами, в случае если мы рассматриваем многоагентную систему, состоящую из интеллектуальных агентов. Эта структура позволяет упорядочить взаимодействие агентов в рамках системы, тем самым ограничивая агента в плане свободы его поведения. За счёт этого взаимодействия между агентами утрачивают хаотический характер.

В рамках многоагентной системы, обязательства исполняют роль определителей связей, возникающих между агентами, что приводит к стабилизации взаимоотношении внутри системы и улучшению качества её функционирования. Благодаря наличию обязательств, агент способен эффективно прогнозировать поведение других членов сообщества, а также упорядочивать собственные действия и предвидеть грядущие события.

Все обязательства, которые имеются внутри многоагентной системы можно классифицировать следующим образом:

1. Обязательства одного агента перед другим.
2. Обязательства агента перед группой.
3. Обязательства группы перед агентом.
4. Обязательства агента перед самим собой.

Самым распространенным видом обязательств являются первый из указанных выше. В основном они связаны с тем, что один агент для другого должен осуществить некоторое действие, которое тот ему поручил или же с возложением на себя обязательств касательно достоверности передаваемой информации. Обязательства перед группой может носить двоякий характер. Это может быть как взятие на себе выполнения какой-либо задачи, поставленной группой, так и выполнение определенных функций, связанных с позицией занимаемой агентом в группе. В свою очередь, обязательства группы перед агентом предполагают обеспечение агента защитой от внешних угроз, поддержание и изменение его социального статуса, а также предоставление ему доступа к общим ресурсам.

Теперь мы можем попробовать построить более подробную классификацию типов взаимодействия между агентами, основываясь на трёх основных критериях, лежащих в основе образования групп из отдельных агентов, а именно: степень совместимости целей агентов, необходимость использования опыта других агентов, использование ресурсов группой.

Теперь приступим к рассмотрению типов взаимодействия. Простое сотрудничество представляет собой объединение опыта нескольких агентов посредством разделения задач и обмена опытом, при этом никаких дополнительных мер по координации их действий не требуется. Непродуктивное сотрудничество связано с отсутствием у агентов направленности на обмен знаниями, в результате чего при совместном распоряжении общими ресурсами они только мешают продуктивной работе друг друга. Координируемое сотрудничество предполагает высокую степень согласованности действий для эффективного распределения знаний и ресурсов, при этом возможно использование дополнительных агентов координаторов. Этот тип сотрудничества весьма сложен, ведь приходиться осуществлять как распределение задач между агентами, так согласовывать их действия.

Помимо случаев сотрудничества также можно рассматривать различные виды соперничества. Простейшим из таких типов взаимодействия является чистое индивидуальное соперничество. При таком взаимодействии агенты находятся в, примерно, одинаковых условиях и при этом получение доступа к ресурсам не является основанием для конфликта. В качестве варианта такого соперничества может выступать борьба за получение задания между агентами. В случае невозможности совместить цели, перед агентами встаёт необходимость вести переговоры. Также существует индивидуальное соперничество за ресурсы, которое базируется на стремлении агентов к монополизации имеющихся ресурсов. Помимо этого, мы можем иметь дело с чистым коллективным сотрудничеством, которое предполагает объединение агентов с небольшим опытом и различными индивидуальными целями в группы, после образования этих групп наступает соперничество между ними. И наиболее продуктивным типом соперничества является коллективное соперничество за ресурсы, которое органично сочетает в себе черты коллективного соперничества и индивидуального соперничества за ресурсы.

Теперь мы перейдём к теме кооперации агентов, так как именно за счёт эффекта кооперативного взаимодействия достигается максимизация результативности многоагентной системы. Кооперация предполагает объединение усилий отдельных агентов ради достижения общих целей. Однако, помимо интеграции совместных усилий, также происходит распределение обязанностей между агентами, специализация групп агентов. Для теории агентов понятие общей цели является весьма значимым, так как предполагает достижения состояния мира, соответствующего целям отдельных агентов и строящегося в согласии с ними.

Формирование совместных целей возможно в двух различных случаях: а) цели отдельных агентов близки или практически совпадают; б) цели, сформированные одним агентом, полностью принимается другими агентами.

В данном случае, мы полагаем, что агенты доброжелательны по отношению друг к другу или, во всяком случае, направлены на достижение компромисса в условиях использования общих ресурсов.

Если попытаться выделить составляющие такого явления как кооперация, то мы получим три основных компонента: сотрудничество агентов, координация их действий и разрешение спорных моментов.

Данный вид организации взаимодействия в подобной системе имеет ряд следующих преимуществ:

1. Решение задач, с которыми не мог бы справиться один агент, посредством группового взаимодействия.

2. Возрастание степень производительности и эффективность выполнения своих функций каждым агентом.

3. Рост качества принимаемых общих решений благодаря разделению общей задачи на подзадачи и высокой степени проработки этих подзадач, а также за счёт активной коммуникации между агентами в ходе выполнения задач.

4. Шанс найти решение многократно возрастает вследствие объединения различных подходов, алгоритмов решения и т.д.

Кооперацию можно изучать с двух различных точек зрения. Мы можем встать на позицию внешнего наблюдателя и изучать её со стороны. В противном случае, мы можем попытаться изучить её изнутри многоагентной системы, рассматривая её как структуру поведения агентов в ходе взаимодействия.

Если мы выберем первый вариант, то мы будем пытаться найти некоторые объективные показатели, которые можно измерить, полностью исключив из рассмотрения ментальные состояния агентов. При таком подходе, можно взять в качестве таких показателей эффективность деятельности агентов, как на групповом, так и на индивидуальном уровне, а также можно рассмотреть частотность такого явления в системе как

конфликты. При оценке эффективности кооперации следует учитывать следующие параметры:

1. время, затрачиваемое на решение задачи коллективом;

2. качество распределения подзадач между агентами;

3. эффективность интеграции различных подходов к решению.

Первый критерий позволяет нам, посредством сравнительного анализа с индивидуальной работой, определить эффективность использования временного ресурса, а благодаря второму критерию можно понять, каким образом происходить распределение задач в рамках системы. Также, используя третий критерий, мы можем оценить успешность согласования различных позиций, относительно решения поставленной задачи, и то, как они минимизируют возникновения конфликтов в системе.

Можно представить и более развёрнутый набор показателей, с помощью которых мы можем определить степень кооперации:

1. распределение функций, степени ответственности и любых ресурсов в системе;

2. степень развития координации взаимодействий между агентами;

3. совместимость задач, которые решаются различными агентами;

4. отсутствие повторов операций, исполняемых агентами;

5. уровень конфликтности внутри системы;

6. устойчивость системы относительно числа входящих в неё элементов, т.е. способность сохранять свою эффективность, несмотря на сокращения числа агентов.

Можно выделить ряд ограничений, которые оказывают непосредственное влияние на кооперацию между агентами. Примерами подобных ограничений могут служить: активность коммуникации между агентами, разница в степени автономности агентов, взаимозависимость решаемых задач, разделение агентов в пространстве и т.д.

Если же мы предпочитаем встать на позицию «взгляд изнутри», включив в наше рассмотрение ещё и намеренья агентов, то тогда кооперация

будет состоять из общей цели и обязательствами между агентами. При подобном взгляде, для оценки эффективности взаимодействия нужно обратить внимания на то, как происходит распределение между агентами в плане задач, а также на способ координации взаимодействий между агентами. Существуют три различных способа централизованного распределения задач:

1. командное управление, при котором существует некоторый центр, который производит распределение между исполнителями и контролирует исполнение в единственном числе;

2. распределение посредством конкурса, при котором происходит набор агентов на выполнение общего задания, предлагаемого центром, в компетенции которого остаётся контроль исполнения, а также выбор исполнителя тех или иных подзадач;

3. распределение с помощью соревнования между агентами, которые претендуют на исполнение конкретных задач, а центр осуществляет контроль над исполнением задач и согласование действий между агентами.

Заключение

Итак, мы переходим к подведению итогов нашего исследования. В результате проделанной нами работы, мы проанализировали современные подходы к понятию интеллект и выявили ряд его ключевых свойств. К этим ключевыми являются интегративная иерархическая структура интеллект, выполнение им функции адаптации, динамический характер взаимодействия мыслительных процессов, осуществления им контроля над поведением агента и его внутренними состояниями.

Изучая современные разработки в рамках агентного подхода, мы пришли к выводу, что интеллектуальный агент это автономная система, взаимодействующая с внешним миром, посредством сенсорной системы и эффекторов. Он способен осуществлять сложные формы поведения, владеет навыками рассуждения, планирования и прогнозирования. Альтернативой автономного когнитивного агента являются многоагентные системы, в которых интеллектуальное поведение обеспечивается за счёт взаимодействия множества агентов, обладающих специфическими характеристиками.

Также нами были изучены основные архитектуры интеллектуальных агентов, а также вопрос о сотрудничестве агентов. В ходе данного исследования, нами было выявлено, что все существующие ныне архитектуры связанны с трёмя основными направлениями в искусственном интеллекте: логицизм, нейронные сети, эволюционные модели. Однако в рамках многоагентных систем делается акцент на способность агента к сотрудничеству с другими агентами, наличием у него развитых коммуникативных навыков и автономии.

Исходя из выше сказанного, можно заключить, что исследования искусственных агентов на современном этапе их развития позволяют изучать как внутреннею структуру агента, так и различные сценарии его поведения при взаимодействии с другими агентами. Это позволяет сделать модели искусственных агентов весьма ценным орудием в изучении гуманистических систем.

Также стоит отметить, что для успешного функционирования многоагентной системы, требуется организовать её таким образом, чтобы достигнуть максимальной кооперации агентов. Для этого должна быть продумана система коммуникации между ними, а также характер их взаимодействия. Так, например, система, составленная исключительно из эгоистичных агентов или только из альтруистичных агентов, не может обеспечить устойчивой работы многоагентной системы. Необходимо задействовать агентов различного типа для достижения максимального результата. Также должен быть продуман механизм разрешения конфликтных ситуаций между агентами, структура распределения должностных обязанностей, а также осуществлён контроль над соблюдением агентами своих обязательств. Это может быть осуществлено как с помощью усложнения внутренней структуры агента, так и с помощью использования специальных контролирующих агентов. Важно понимать, что в хорошо организованной многоагентной системе, все агенты работают на достижение общих целей, которые удовлетворяют потребности всех агентов входящих в систему.

Итак, по нашему мнению, существует две возможных альтернативы разработки когнитивных агентов. В первом случае мы строим сложно организованного агента, обладающего широким набором когнитивных инструментов, и тем самым создадим искусственный разум. Во втором случае, мы будем создавать коллективы агентов, которые с помощью скоординированной совместной работы будут порождать искусственный разум. Какая из этих альтернатив приведёт нас к желаемому результату, покажет время, но на данный момент оба направления выглядят весьма перспективными и приводят к неплохим результатам. Однако, многоагентные системы всё таки выигрывают по ряду параметров, например, в скорости, поэтому мы искренне верим в блестящее будущее этого направления.

Список литературы

1. Амосов Н.М. Искусственный разум. – Киев: Наукова думка, 1969.
2. Ананьев Б. Г. Человек как предмет познания. – СПб.: Питер, 1980.
3. Бонгард М. М. Проблема узнавания. М.: Наука, 1967.
4. Брунер Дж. Психология познания. – М.: Прогресс, 1977.
5. Брушлинский А. В. Мышление и прогнозирование. – М.: Мысль, 1979.
6. Варшавский В.И., Поспелов Д.А. Оркестр играет без дирижера. – М.: Наука, 1984.
7. Венада В. Ф. Системы гибридного интеллекта. Эволюция, психология, информатика. – М.: Машиностроение, 1990.
8. Вертгеймер М. Продуктивное мышление. – М.: Прогресс, 1987.
9. Винер Н. Кибернетика или Управление и связь в животном и машине. – М.: Наука, 1983.
10. Выготский Л. С. Психология развития человека. – М.: Эксмо, 2005.
11. Гаазе-Рапопорт М. Г., Поспелов Д. А. От амебы до робота: модели поведения. – М. ; Наука, 1987.
12. Городецкий В.И. Многоагентные системы: основные свойства и модели координации поведения//Информационные технологии и вычислительные системы, № 1, 1998, с. 22-34.
13. Деннет Д. Виды психики: на пути к пониманию сознания. – М.:Идея-Пресс, 2004.
14. Докинз Р. Эгоистичный ген. – М.: Мир, 1993.
15. Дорф Р., Бишоп Р. Современные системы управления. – М.: Лаборатория Базовых Знаний, 2002.
16. Жданов А.А. Автономный искусственный интеллект. – М.: БИНОМ. Лаборатория знаний, 2008.
17. Кастлер Г. Возникновение биологической организации. – М.: Мир,1967.
18. Левин К. Теория поля в социальных науках. – СПб.: Речь, 2000.
19. Левитин К. Е., Поспелов Д. А. Будущее искусственного интеллекта – М.: Наука, 1991.

20. Мак-Каллок У. С., Питтс В. Логическое исчисление идей, относящихся к нервной активности//в сб. «Автоматы» под ред. К.Э. Шеннона и Дж. Маккарти. – М.: Иностранная литература, 1956.
21. Матурана У., Варела Ф. Древо познания: биологические корни человеческого понимания. – М.: Прогресс-Традиция, 2008
22. Миллер Д., Галантер Ю. Прибрам К. Планы и структуры поведения. – М.: Прогресс, 1965.
23. Минский М. Вычисления и автоматы. М.: Мир, 1971.
24. Налимов В.В. Вероятностная модель языка. О соотношении естественных и искусственных языков. – М.: Наука, 1979.
25. Нейсер У. Познание и реальность. – М.: Прогресс, 1981.
26. Новиков Д.А., Чхартишвили А.Г. Рефлексивные игры. – М.: СИНТЕГ,2003.
27. Нильсон Н. Искусственный интеллект: методы поиска решений. – М.: Мир, 1973.
28. Осипов Г. С. Методы искусственного интеллекта. – М.: Физматлит, 2011.
29. Патнэм Х. Философия сознания. – М.: Дом интеллектуальной книги, 1999.
30. Пиаже Ж. Избранные психологические труды. – М.: МПА, 1994.
31. Пинкер С. Язык как инстинкт. – М.: Едиториал УРСС, 2004.
32. Полани М. Личностное знание. – М.: Прогресс, 1985.
33. Поспелов Д.А. Нечёткие множества в моделях управления и искусственного интеллекта. – М.: Наука, 1986.
34. Поспелов Д. А., Пушкин В.Н. Мышление и автоматы. – М.: Наука,1972.
35. Поспелов Д. А. Экспертные системы: состояние и перспективы. – М.: Наука, 1989.
36. Поспелов И.Г. Моделирование экономических структур. – М.: Фазис, 2003.
37. Поспелов Э. В. Общение с ЭВМ на естественном языке. М.: Наука, 1982.
38. Прибрам К. Языки мозга. – М.:Прогресс, 1975.
39. Рассел С., Норвинг П. Искусственный интеллект: современный подход. – М.: Вильямс, 2006.

40. Редько В. Г. Эволюция, нейронные сети, интеллект: модели и концепции эволюционной кибернетики. – М.: ЛИБРОКОМ,2011.
41. Рубенштейн С.Л. Основы общей психологии. – СПб.: Питер, 2000.
42. Саймон Г. Науки об искусственном. – М.: Едиториал УРСС, 2004.
43. Симонов П. В. Эмоциональный мозг. – М.: Наука, 1981.
44. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: УРСС, 2002.
45. Тихомиров О.К. Психология мышления. – М.: Издательство Мгу, 1984.
46. Философия языка/под ред.-сост. Дж. Р. Сёрль. – М.: Едиториал УРСС, 2010.
47. Финн В. К. Интеллектуальные системы и общество. – М.: КомКнига, 2006.
48. Хокинс Дж., Блейксли С. Об интеллекте. – Вильямс, 2007.
49. Холодная М. А. Психология интеллекта: парадоксы исследования. – СПб.: Питер, 2002.
50. Шамис А.Л. Пути моделирования мышления: активные синергетические нейронные сети, мышление и творчество, формальные модели поведения и «распознавание с пониманием». – М.: КомКнига, 2006.
51. Эшби У. Р. Введение в кибернетику. – М.: Либрком, 2009.