

**Мультиагентный алгоритм обоснования символов
конвенционального языка на основе ситуативно обусловленного развития
нейрокогнитивной архитектуры***

**М. И. Анчёков¹, К. Ч. Бжихатлов¹, С. А. Канкулов², З. В. Нагоев¹,
О. В. Нагоева², И. А. Пшенокова¹**

¹ Кабардино-Балкарский научный центр Российской академии наук
360010, Россия, г. Нальчик, ул. Балкарова, 2

² Институт информатики и проблем регионального управления –
филиал Кабардино-Балкарского научного центра Российской академии наук
360000, Россия, г. Нальчик, ул. И. Арманд, 37-а

Аннотация. Цель данного исследования состоит в разработке методов и алгоритмов обоснования символов естественного языка в ходе его освоения интеллектуальным агентом при взаимодействии с другими агентами. Разработана мультиагентная нейрокогнитивная архитектура интеллектуального агента, погруженного в коммуникативную среду, формирующая функциональные системы на основе согласованного взаимодействия агентов-нейронов в ее составе. Разработан мультиагентный алгоритм обоснования символов упрощенного конвенционального языка, состоящего из односимвольных слов, использующий такие функциональные системы для формализации семантики с учетом контекста коммуникации.

Ключевые слова: интеллектуальный агент, обоснование символов, мультиагентные системы, когнитивные архитектуры, нейроподобные элементы

Поступила 01.12.2022, одобрена после рецензирования 12.12.2022, принята к публикации 14.12.2022

Для цитирования. Анчёков М. И., Бжихатлов К. Ч., Канкулов С. А., Нагоев З. В., Нагоева О. В., Пшенокова И. А. Мультиагентный алгоритм обоснования символов конвенционального языка на основе ситуативно обусловленного развития нейрокогнитивной архитектуры // Известия Кабардино-Балкарского научного центра РАН. 2022. № 6(110). С. 48–60. DOI: 10.35330/1991-6639-2022-6-110-48-60

MSC 68T42

Original article

**A multi-agent algorithm for substantiating
the symbols of a conventional language based on situational development
of neurocognitive architecture***

**M.I. Anchekov¹, K.Ch. Bzhikhatlov¹, S.A. Kankulov², Z.V. Nagoev¹,
O.V. Nagoeva², I.A. Pshenokova¹**

¹ Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences
360010, Russia, Nalchik, 2 Balkarov street

² Institute of Computer Science and Problems of Regional Management –
branch of Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences
360000, Russia, Nalchik, 37-a I. Armand street

© Анчёков М. И., Бжихатлов К. Ч., Канкулов С. А., Нагоев З. В. и др., 2022

* Работа выполнена при поддержке Российского научного фонда, грант № 22-19-00787

* The work was supported by the Russian Science Foundation, grant no. 22-19-00787

Abstract. The purpose of this study is to develop methods and algorithms for justifying natural language symbols in the course of its development by an intelligent agent when interacting with other agents. A multi-agent neurocognitive architecture of an intelligent agent immersed in a communicative environment has been developed, which forms functional systems based on the coordinated interaction of agents-neurons in its composition. A multi-agent algorithm for substantiating symbols of a simplified conventional language, consisting of single-character words, has been developed, using such functional systems to formalize semantics, taking into account the context of communication.

Keywords: intelligent agent, symbol substantiation, multi-agent systems, cognitive architectures, neuron-like elements

Submitted 01.12.2022,

approved after reviewing 12.12.2022,

accepted for publication 14.12.2022

For citation. Anchekov M.I., Bzhikhatlov K.Ch., Kankulov S.A., Nagoev Z.V., Nagoeva O.V., Pshenokova I.A. A multi-agent algorithm for substantiating the symbols of a conventional language based on situational development of neurocognitive architecture. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2022. No. 6(110). Pp. 48–60. DOI: 10.35330/1991-6639-2022-6-110-48-60

1. ВВЕДЕНИЕ

В данной работе формулируются основные принципы решения проблемы обоснования символов языка с помощью метафоры проектирования *интеллектуального агента* под управлением *когнитивной архитектуры* [1].

Сущность *проблемы «обоснования символов»* (Symbol Grounding Problem) состоит в том, что интерпретация смыслов естественно-языковых высказываний в формальных системах строится на основе некоторых других символов, а содержательная часть этих высказываний теряется.

Интеллектуальный агент рассматривается как *рациональный агент искусственной жизни* (Artificial Life) под управлением *мультиагентной рекурсивной нейрокогнитивной архитектуры* [2].

Мультиагентная нейрокогнитивная архитектура определяется как *рекурсивная когнитивная архитектура*, функциональные узлы которой состоят из программных *агентов-нейронов* (*агнейронов*) разной степени сложности и объединены в составе т.н. *инварианта организационной структуры принятия решений на основе нейрокогнитивной архитектуры* (*инвариант когнитивной архитектуры*) [2, 3]. Этот инвариант включает в себя последовательность обязательных операций *распознавания состояний, идентификации и оценки проблемных ситуаций, синтеза целевых состояний и путей их достижения*. Каждая из этих операций выполняется с помощью мультиагентного алгоритма, основанного на обмене сообщениями между агнейронами различных типов, расположенными в соответствующих функциональных узлах нейрокогнитивной архитектуры.

В целом мультиагентная нейрокогнитивная архитектура выполняет функции системы управления *поведением* интеллектуального агента. Задача *синтеза поведения* интеллектуального агента представляет собой задачу синтеза пути из начальной вершины т.н. *графа проблемной ситуации* (*графа решений*), описывающей текущее состояние системы «интеллектуальный агент – среда», в конечную, описывающую некоторое состояние этой системы в будущем, характеризующееся более высоким значением *целевой функции* интеллектуального агента. В качестве такой функции выбирается некоторая суперпозиция оценок состояния параметра *энергии*, характеризующего потенциал активности (способности совершать действия) агента в среде [2].

Цель исследования состоит в том, чтобы сформировать основные принципы решения научной *проблемы обоснования символов* в системах компьютерной лингвистики, основанных на метафоре проектирования *рационального агента* под управлением *мультиагентной нейрокогнитивной архитектуры*.

Для экспериментальной проверки сформулированных гипотез необходимо решить *задачу* разработки минимально функциональной системы имитационного моделирования

рассуждений на основе мультиагентных нейрокогнитивных архитектур и организовать ее самостоятельное обучение некоторому очень простому конвенциональному (условному) языку при взаимодействии в коммуникационной среде, выполняемое аналогично обучению ребенка, осваивающего естественный язык.

Проблема обоснования символов тесно связана с вопросом о возникновении символической коммуникации как таковой [4]. Согласно Де Соссюру, символы соотносят внутренние репрезентации значений с внешними сигналами [5]. Предпринимались многочисленные попытки формализации таких соотношений знаменитыми «семантическими треугольниками» (Ч. С. Пирс, Ф. Фреге, Ч. Огден и А. Ричардс, Ч. У. Моррис), представляющих особенности различных подходов к моделированию знаков [6–10].

Когнитивисты построили представление о возможности репрезентации смысла с помощью т.н. *физических символических систем*, способных интерпретировать символы в соответствии с некоторыми встроенными алгоритмами обработки этих внутренних репрезентаций [11]. Этот подход породил целый ряд постановок проблемы соотношения символов при интерпретации высказываний – проблема фреймов (Д. Маккарти и П. Хайес) [12], проблема китайской комнаты (Д. Сирл), проблема обоснования символов (С. Харнад). По мнению Вогта (P. Vogt), символы при использовании когнитивистского подхода не являются ни ситуативными, ни связанными с телесным опытом (телесно обоснованными) (*embodied*) [5].

С целью преодоления этой проблемы Р. А. Брукс выдвинул гипотезу о *физическом обосновании символов* на основе динамического сопоставления сенсорных и моторных потоков, возникающих при физическом взаимодействии агента (робота) с окружающей его средой [13].

Для обозначения процесса формирования знаков, значения которых строятся при взаимодействии агента со средой, Ч. С. Пирс ввел понятие *семиозис* [14]. Известны многочисленные попытки придания системам *телесного познания* [5], основанным на семиозисе экологической целевой функции [15, 16].

Актуальность данного исследования определяется необходимостью решения *проблемы обоснования символов*, сдерживающей развитие теоретических и прикладных исследований в области формализации семантики естественного языка.

2. ЗАДАЧА ОБОСНОВАНИЯ ОДНОБУКВЕННЫХ ВЫСКАЗЫВАНИЙ КОНВЕНЦИОНАЛЬНОГО ЯЗЫКА

В качестве модельной базовой задачи мы рассматриваем процесс «освоения» (семантизации) отдельных буквенных символов, представляющих собой высказывания простого конвенционального языка. Пока мы исследуем вопрос создания интеллектуального агента, который погружен в реальную среду с помощью сенсоров и эффекторов, обеспечивающих интерфейс с пользователями (социальное окружение), которые «разговаривают» с ним на примитивном языке, состоящем только из токенов – символов, не обремененном синтаксисом и грамматикой, и используют для коммуникации сообщения, обмен которыми проходит только посредством дисплея и клавиатуры. Такое упрощение не затрагивает основные требования к имитационной модели интеллектуального агента, основное из которых – ситуативно детерминированное интерактивное формирование у него семантических онтологий с использованием процесса освоения языка.

Для обучения интеллектуального агента элементам некоторого языка необходимо выполнить несколько условий. Во-первых, такой агент должен обладать гомеостатическим циклом, реализация которого должна составлять цель синтеза его поведения. Поведение должно синтезироваться агентом в среде, в которую он погружен с помощью сенсоров и эффекторов. С этой средой интеллектуальный агент должен быть связан гомеостатически – именно из нее он должен получать энергию, в ней же должен эту энергию тратить.

Воздействие интеллектуального агента на среду реализовано путем отправки команд различным исполнительным устройствам интеллектуального агента.

Каналы обмена информацией могут включать в себя подсистемы распознавания видео-, аудио-, тактильных или других образов.

Среда должна содержать других интеллектуальных агентов (люди, программные агенты, роботы), которые владеют этим и имеют техническую возможность обмениваться с интеллектуальным агентом высказываниями на нем.

Интеллектуальный агент должен обладать мультиагентной нейрокогнитивной архитектурой, которая управляет его поведением. Поведение интеллектуального агента должно синтезироваться как набор действий, выполняемых им с помощью эффекторов. Управляющая мультиагентная нейрокогнитивная архитектура должна расти и развиваться в процессе синтеза поведения интеллектуального агента в составе коллектива в направлении формирования нейроморфологического обеспечения взаимосвязанного комплекса коммуникативно-онтологических функциональных систем, задействованных в освоении и использовании языка.

Рассмотрим общую структуру такого интеллектуального агента (рис. 1). У него есть простой гомеостатический цикл, связанный с тем, что ему на поддержание жизнедеятельности требуется определенное количество энергии в единицу дискретного времени.

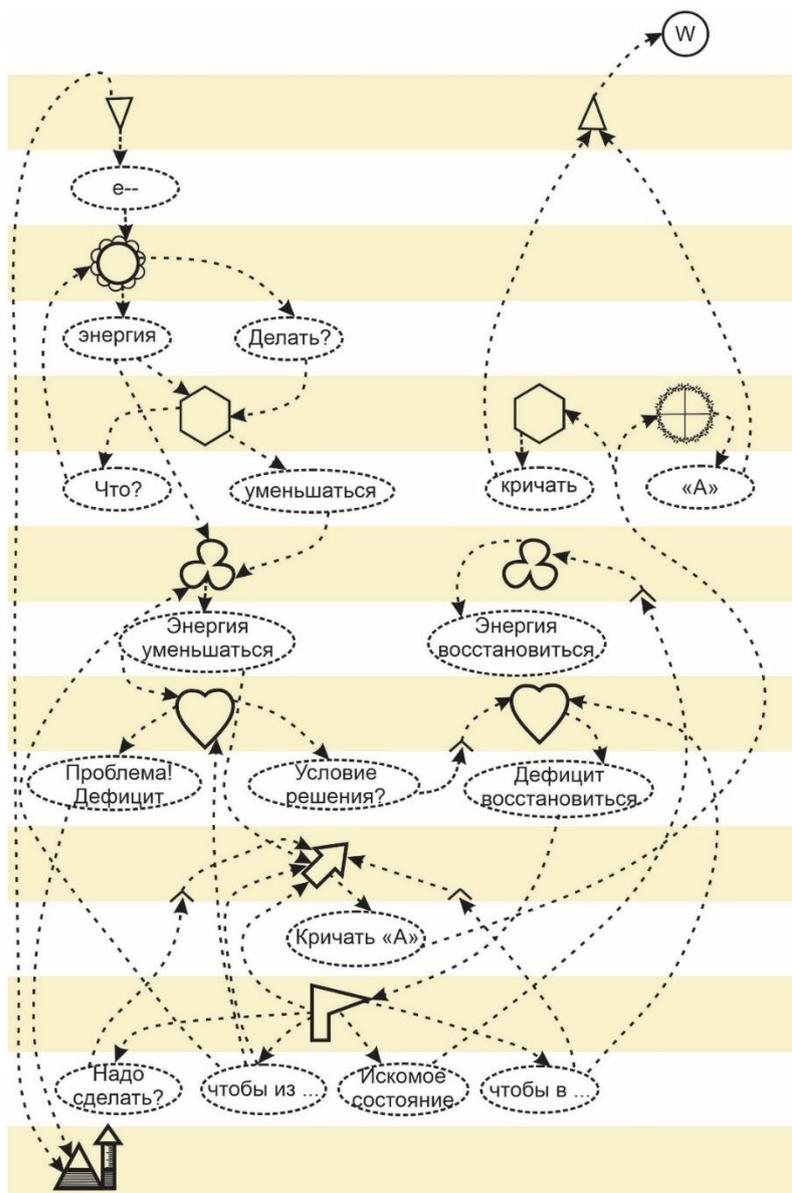


Рис. 1. Минимальная мультиагентная нейрокогнитивная архитектура коммуникативного интеллектуального агента

Fig. 1. Minimal multi-agent neurocognitive architecture of communicative intelligent agent

Он стремится всегда держать уровень энергии в резервуаре, изображенном на этом рисунке пиктограммой треугольного наполовину заполненного резервуара, соединенного с индикаторной «мачтой», на максимальной отметке. Сверху мачты изображен интероцептор энергии, который измеряет и отсылает по месту требования текущее значение уровня энергии в резервуаре энергии.

У интеллектуального агента есть коммуникативный рецептор, через который он получает символичные сообщения от пользователей из чата (рецептор чата), а также эффектор чата, с помощью которого он отправляет сообщения в чат. Пользователи обмениваются сообщениями с интеллектуальным агентом, в том числе естественно-языковыми высказываниями, с помощью чата.

Первоначально нейрокогнитивная архитектура не располагает значительным арсеналом специфических поведенческих реакций. Как и у новорожденного ребенка, скорее, целесообразно обеспечить модель искусственной жизни некоторым ограниченным арсеналом универсальных реакций для организации процесса приобретения специфических реакций путем обучения в конкретных ситуациях ее взаимодействия со средой своего обитания. На рисунке 1 приведен состав агнейронов, задействованных в реализации управления выполнением простой поведенческой реакции.

Логично, что в качестве первой универсальной реакции выступает актуализация способности привлекать внимание к себе путем отправки социальному окружению сообщений, которые его участники с большой вероятностью не смогут игнорировать. В случае нашей упрощенной модели в качестве таких сообщений могут рассматриваться символы алфавита естественного языка. Снабдим интеллектуального агента возможностью отправлять одиночные символы и их последовательности в тот же самый чат, через который интеллектуальный агент воспринимает сообщения от социального окружения. Такая способность может служить неким отдаленным аналогом наличия артикуляционного аппарата и возможности его активного применения в ситуации обучения и самообучения путем подражания, проб и ошибок у новорожденного ребенка, осваивающего естественный язык.

Для интерпретации высказывания необходимо снабдить мультиагентную нейрокогнитивную архитектуру агнейронами, выполняющими функциональную репрезентацию понятий, используемых в естественно-языковом высказывании. Они применяются для формирования описания текущего состояния системы «интеллектуальный агент – среда» на основе идентификации фактов, устанавливаемых интеллектуальным агентом об этой системе.

Функциональная репрезентация таких фактов требует наличия в мультиагентной нейрокогнитивной архитектуре средств описания функциональных репрезентаций всех элементов предикативных конструкций. Наиболее распространенной формой предиката, представленной в функциональной репрезентации интеллектуального агента, является утверждение о наличии у выделяемых этим агентом в системе «интеллектуальный агент – среда» объектов определенных свойств или отношений с другими объектами. В мультиагентной нейрокогнитивной архитектуре должны быть средства для функциональной репрезентации как объектов, так и свойств, а также и отношений.

Эти задачи в составе мультиагентной нейрокогнитивной архитектуры выполняют специализированные агнейроны, реагирующие на специализированные сигнатуры, позволяющие распознавать эти объекты, свойства и отношения, – *агнейроны-объекты, агнейроны-свойства, агнейроны-отношения*. Эти агнейроны взаимодействуют друг с другом с целью согласованного воздействия на процесс синтеза поведения интеллектуального агента. Следовательно, формирование фактов о мире в управляющей нейрокогнитивной архитектуре интеллектуального агента представляет собой мультиагентный процесс, в котором задействованы различные агнейроны, взаимодействующие друг с другом с помощью мультиагентного обмена сообщениями [17].

Репрезентацию таких мультиагентных фактов в нейрокогнитивной архитектуре выполняют т.н. событийные агнейроны, которые интегрируют сигналы от агнейронов, описывающих

объекты и действия в формате мультиагентной предикативной конструкции. На рисунке 1 агнейроны-объекты показаны круглыми пиктограммами, агнейроны-действия – шестиугольными, а агнейроны-события – трилистниками с мягкими лопастями. На рисунке видно, что агенты взаимодействуют друг с другом, обмениваясь символическими сообщениями.

Основным механизмом, используемым нами в описываемой модели для имитации динамики мультиагентной нейрокогнитивной архитектуры интеллектуального агента при формировании функциональных систем, является алгоритм формирования причинно-следственных зависимостей между агнейронами, выполняющими в нейрокогнитивной архитектуре функциональную репрезентацию взаимосвязанных событий. Это алгоритм т.н. *онтонейроморфогенеза* – ситуативно-детерминированного процесса роста и деградации аксо-дендрональных связей [18]. Его содержательной основой является поведение событийных нейронов, направленное на поиск контрагентов среди таких же событийных нейронов, которые стали бы, с одной стороны, производить и «продавать» другим агнейронам (за вознаграждение в виде энергии) информацию о некотором внешнем событии, а с другой стороны (на стороне другого агнейрона – агнейрона-покупателя), – конструировать («покупать») эту информацию, «оплачивая» ее наличной энергией. Устойчивая связь, образуемая таким способом между агнейроном-покупателем и агнейроном-продавцом, как правило, отражает наличие причинно-следственной зависимости между событиями в системе «интеллектуальный агент-среда», нейроморфологическим коррелятом наступления которых и выступают агнейрон-продавец и агнейрон-покупатель. В [3] предложено обозначать такую связь термином «*мультиагентный контракт*», или просто «*контракт*» между агнейронами на обмен информацией и энергией.

Необходимым условием образования такой причинно-следственной зависимости является последовательное наступление события, репрезентируемого в нейрокогнитивной архитектуре нейроном-продавцом (событие-причина), и события, репрезентируемого в нейрокогнитивной архитектуре нейроном-покупателем (событие-следствие), в течение которого, достаточно малого промежутка времени.

На рисунке 1 видно, что наступление события «сказал»: «А» (артикулировал «А» – отправил символ «А» в чат) приводит к регистрации этого действия специальным рецептором чата и соответствующим событийным агнейроном. До образования причинно-следственной зависимости связь между событийными нейронами отсутствует. Это означает, что до наступления события-следствия – «Слышу «А» (в чате – в коммуникационной среде) никакая нейронная группа не активизируется, и интеллектуальный агент в целом не сможет получить информацию об ожидаемом наступлении события-следствия до момента его наступления. После того как причинно-следственная связь образована, активация нейроморфологического коррелята события-причины (активация событийного нейрона-причины) приводит к активации нейроморфологического коррелята события-следствия (активация событийного нейрона-следствия).

Таким образом, каждый раз, когда интеллектуальный агент артикулирует некоторый символ, он обучается тому, что вслед за артикуляцией этого символа он ощущает присутствие этого символа («слышит» его) в коммуникационной среде.

При этом часть нейрокогнитивной группы, выполняющая функциональную репрезентацию исходного состояния, конечного состояния и действия, по нашей гипотезе, на основе принципов и мультиагентного алгоритма онтонейроморфогенеза, формирует внутри себя и вместе с агнейроном целеполагания множественное мультиагентное отношение. Оно формируется путем заключения контрактов между отдельными агнейронами в его составе.

На рисунке 1 представлен агнейрон управления действиями (фигурой в форме широкой стрелки), отвечающий за формирование дуги в графе проблемной ситуации, представляющей действие, найденное мультиагентной нейрокогнитивной архитектурой в результате работы по синтезу плана поведения, направленного на решение текущей проблемы. Партнерами агнейрона управления действиями, которым он направляет команды, описывающие мультиагентный факт (предикативная конструкция из сообщений), описывающий действие, которое необходимо включить в план и выполнить, являются агнейроны объектов и

действий. Эти агнейроны затем формируют необходимые команды и направляют их на эффекторы интеллектуального агента, что уже непосредственно приводит к выполнению соответствующего действия в среде.

Функция агнейрона целеполагания, представленного на рисунке 1 в виде пиктограммы в форме флажка, состоит в том, чтобы, используя итеративный опрос событийных агнейронов, агнейронов управления действиями и агнейронов оценки (представлены на рисунке пиктограммами в форме «сердечек») путем отправки запросов, сначала найти целевое состояние в графе проблемной ситуации, а затем подобрать такие действия, совокупность которых и составила бы искомый план поведения интеллектуального агента, направленный на решение исходной проблемы путем перехода из начального состояния в целевое.

Данные для обучения мультиагентной нейрокогнитивной архитектуры формируются в результате взаимодействия этого интеллектуального агента с пользователями с применением программного обеспечения в коммуникативной ситуации, связанной с диалогами, направленными на обеспечение его простейшего пищевого поведения.

3. МУЛЬТИАГЕНТНЫЙ АЛГОРИТМ ОБОСНОВАНИЯ СИМВОЛОВ

Допустим, уровень энергии в резервуаре интеллектуального агента снизился до некоторого достаточно низкого значения. При этом значении энергии срабатывает интероцептор энергии, что запускает дальнейшую активность агнейронов в составе мультиагентной нейрокогнитивной архитектуры, направленную на компенсацию наличных и возможных последствий события уменьшения энергии.

На рисунке 2 приведена последовательность действий по отправке сообщений рецепторами и агнейронами различного вида и порождению т.н. *нейрофабриками* – специальными акторами (изображены на рисунках фигурами с двойными контурами), выполняющими добавление в нейрокогнитивную архитектуру структурно-функциональных элементов по требованию, – новых агнейронов. Эта последовательность и представляет собой искомый мультиагентный алгоритм, который синтезируется нейрокогнитивной архитектурой с целью и в процессе решения задачи обоснования символов при освоении конвенционального языка.

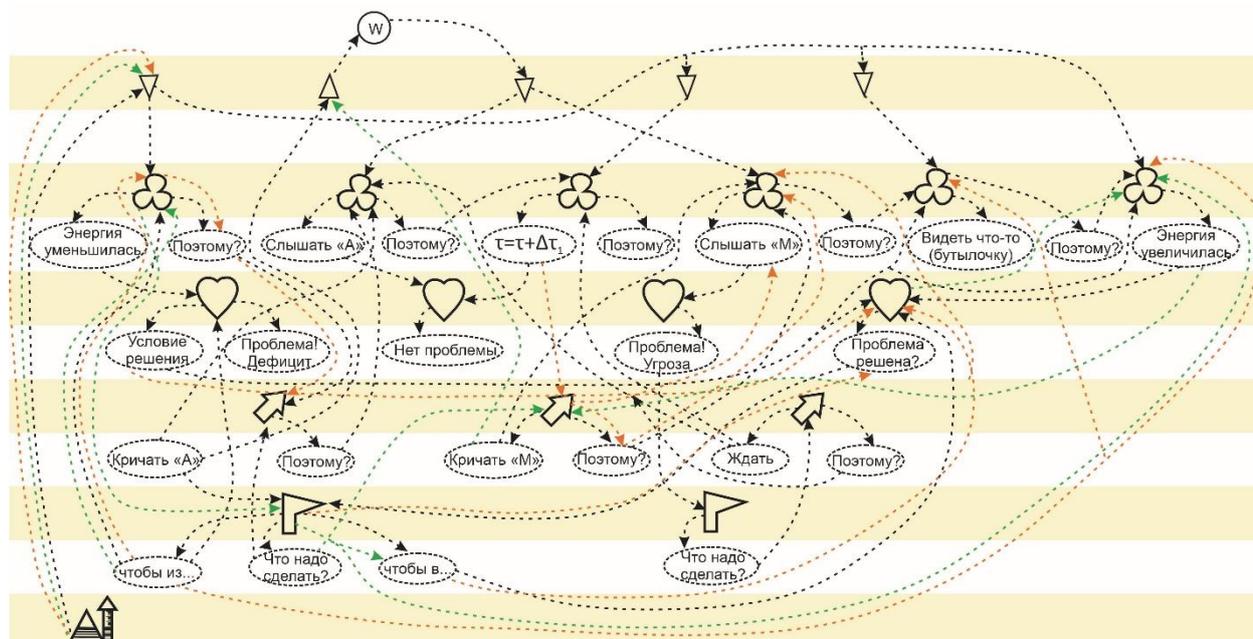


Рис. 2. Мультиагентный алгоритм обоснования символов в процессе обучения нейрокогнитивной архитектуры

Fig. 2. Multi-agent symbol substantiation algorithm in the process of learning of neurocognitive architecture

После того как интероцептор энергии посылает сообщение агнейронам управляющей нейрокогнитивной архитектуры интеллектуального агента, отвечающим за распознавание текущей ситуации, происходит идентификация этого сообщения объектным агнейроном, представляющим в когнитивной архитектуре понятие «энергия» (агнейрон «энергия»).

Одновременно этот сигнал также отправляется агнейрону действия, представляющему в когнитивной архитектуре понятие «равняется» (агнейрон «равняется»). Этот агнейрон, в свою очередь, распознает сигнал и отправляет сообщение агнейрону «энергия».

Получив эти сигналы, агнейроны «энергия» и «действие» отправляют сообщения агнейрону события, который, в свою очередь, формирует и отправляет сообщение агнейрону оценки для того, чтобы тот оценил степень значимости данного события.

Агнейрон оценки, получив это сообщение, используя продукционные правила из своей базы знаний, генерирует сообщение о том, что данное событие является проблемным, что оно входит в состав проблемной ситуации, требующей решения. Это сообщение отправляется другим агнейронам оценки, а также агнейрону целеполагания.

После того как цель определена в виде состояния в будущем, в котором происходит событие пополнения энергии, агнейрон целеполагания дает команду агнейронам управления действиями, которые далее должны будут «включить» некоторое поведение интеллектуального агента, направленное на достижение цели.

Если, обработав запрос от агнейрона целеполагания, агнейрон управления действиями на основании правила, содержащегося в его базе знаний, определит, что именно он является контрагентом по данному запросу, то он ответит агнейрону целеполагания сообщением, которым уведомит его о том, что соответствует запросу.

Для процесса решения задачи этот ответ определяет локальный выбор дуги в графе проблемной ситуации, который в децентрализованном стиле выполнили агнейроны локальной функциональной системы. Если, как в рассматриваемом случае, весь искомый путь состоит всего из двух вершин и одной связывающей их дуги, то на этом поиск останавливается и целевой агнейрон дает команду на исполнение выбранного действия.

В рассматриваемом случае, идентифицировав проблемную ситуацию «Энергия уменьшилась», интеллектуальный агент в качестве целевого определяет состояние «Энергия в норме». Затем он выполняет единственное доступное ему пока действие – «кричит», посылая последовательности сообщений (команд) на агнейроны управления действиями, отвечающие за отправку на дисплей символов «А», «Б» и т.д. Эти команды попадают на т.н. «символьные» агнейроны, изображенные на рисунке 2 пиктограммой круга с лучиками по всей длине окружности, пересеченного внутри перпендикулярными друг другу отрезками, проходящими через центр, а также на агнейрон действия «написать». Это приводит к тому, что на эффектор «чата» – специализированный эффектор, отвечающий за отправку символов в чат между интеллектуальными агентами и пользователями, – отправляются соответствующие последовательности символов.

На рисунке 2 видно, что оба события – отправка символа в чат и считывание его из чата – идентифицируются в когнитивной архитектуре функциональной системой восприятия, в частности, активируются соответствующие событийные агнейроны. Событийный агнейрон (в форме округлого трилистника), выполняющий в нейрокогнитивной архитектуре функциональную репрезентацию события «написать в чат [символ] «А», констатирует факт отправки символа «А» в чат.

После того как агнейрон управления действием «написать в чат [символ] «А» отправил команду на агнейрон символа «А» и на агнейрон действия «написать», а они в свою очередь сформировали команды для эффектора чата, и он выполнил эти команды – отправил в чат символ «А», информация для активации событийного агнейрона «написать в чат [символ] «А» порождается специальным сенсором эффектора чата, который изображен на рисунке 2 пиктограммой в форме равностороннего треугольника, связанного отрезком с актором чата и обращенного к нему (к пиктограмме актора чата) своим основанием. Этот сенсор пред-

ставляет собой интероцептор, информирующий заинтересованных агнейронов в составе когнитивной архитектуры о том, что эффектор чата «сработал» (активировался), и о том, как именно он активировался – какой именно символ он отправил в чат.

Как следует из рисунка 2, между некоторыми агнейронами в составе мультиагентной нейрокогнитивной архитектуры формируются контрактные отношения, выраженные каузативными вопросами «Почему?» и «Поэтому?». Такие отношения формируют причинно-следственные связи между агнейронами в составе мультиагентной нейрокогнитивной архитектуры на основе алгоритма онтонейроморфогенеза [4].

В результате в мультиагентной нейрокогнитивной архитектуре формируется система функциональных репрезентаций событий, действий и причинно-следственных связей между ними, отражающая опыт взаимодействия интеллектуального агента со средой, в которую он погружен. В частности, эта репрезентация отражает убеждение интеллектуального агента в том, что в системе «интеллектуальный агент – среда» событие «энергия уменьшилась» является причиной действия «написать «А», которое в свою очередь является причиной события «прочитать «А», что видно на рисунке 2.

После того как интеллектуальный агент привлекает внимание окружающих криком, пользователь как агент социального окружения включается в разрешение проблемы, помогает интеллектуальному агенту поставить правильную цель и сформировать путь в графе проблемной ситуации, ведущий от начального состояния к состоянию, описывающему эту цель. Для этого пользователь (мама) «кормит» интеллектуального агента (ребенка), например, выбирает с помощью специального программного обеспечения некоторый тип «пищи» (например, молоко) и дает ее ребенку, при этом, как правило (и это очень важно!), называя (с помощью диалогового окна чата) саму пищу на своем (в данном случае – на нашем простом, конвенциональном) языке, – например, она говорит: «[Это –] молоко». Пока токен «молоко» для большего упрощения заменим на токен «М», непосредственно доступный нашему (очень простому пока) интеллектуальному агенту для артикуляции и восприятия.

Получив пищу, ребенок насыщается, обретает состояние комфорта, достигается целевое состояние, общее для графов проблемных ситуаций мамы и ребенка, и «сдвоенная» кооперативная проблема разрешается для них обоих.

Естественно, все вышеперечисленные события и действия также находят свою функциональную репрезентацию в мультиагентной нейрокогнитивной архитектуре ребенка (интеллектуального агента) в виде мультиагентных коалиций, взаимодействующих между собой на основе контрактов агнейронов, в конечном счете связанных с событийными агнейронами и агнейронами управления действиями. Эти, последние, агнейроны в свою очередь, используя алгоритм онтонейроморфогенеза, формируют между собой мультиагентную нейрокогнитивную функциональную репрезентацию причинно-следственных отношений, которая дополняет ранее построенную.

Так, ранее построенная система причинно-следственных зависимостей дополняется следующими элементами. Состояние «[Я] прочитать [-ал сообщение] «молоко» (в коммуникационной среде появился символ «М») является причиной события «[Я] получить молоко», которое в свою очередь является причиной события «[Я] ощущать сытость» («Энергия в норме»), являющегося причиной завершающего события проблемной ситуации – «Проблема решена».

Таким образом, в результате наступления и выполнения всех вышеописанных событий и действий в мультиагентной нейрокогнитивной архитектуре интеллектуального агента формируется функциональная репрезентация соответствующего опыта участия этого агента в разрешении проблемной ситуации и причинно-следственных зависимостей, описывающих часть системы «интеллектуальный агент – среда», которая была вовлечена в этот процесс (рис. 2). Именно формирование новых и корректировка существующих функциональных систем и является концептуальной и алгоритмической основой машинного обучения, реализуемого на основе самоорганизации мультиагентной нейрокогнитивной архитектуры [2].

Допустим, у интеллектуального агента в нейрокогнитивной архитектуре представлено несколько функциональных систем, каждая из которых обеспечивает артикуляцию некоторого конкретного символа в общей коммуникационной среде (рис. 2). Предположим также, что изначально интеллектуальный агент не знает (в его нейрокогнитивной архитектуре не представлено знание), какой именно символ необходимо использовать (а выше мы установили, что в данном случае это символ «М») для того, чтобы в рассмотренной выше ситуации «голода» (падения значения энергии) быстро получить от пользователя «пищу» («молоко»).

Допустим, интеллектуальный агент «хочет» (формирует целевое состояние в указанном смысле), чтобы в чате появился символ «М», а вышеописанный механизм работы его мультиагентной нейрокогнитивной архитектуры произвел в чате символ «Б». Тогда, зарегистрировав с помощью рецептора чата результат испытания, нейрокогнитивная архитектура на основе мультиагентного алгоритма онтонейроморфогенеза перестраивает контрактные связи между агентами таким образом, чтобы в ситуации, когда в качестве целевого будет выбрано состояние, в котором в чате появляется символ «Б», именно функциональная система артикуляции символа «Б» активировалась бы в ответ, что, соответственно, привело бы к появлению в чате именно этого символа.

По нашему мнению, непосредственно процедура синтеза пути должна выполняться в двунаправленном режиме – как от начального состояния к целевому, так и от целевого к начальному. В теории принятия решений эти методы известны соответственно как прямое и обратное планирование (*forward and backward chaining*). При этом на каждом шаге обоих видов планирования задача поиска локального действия, способного сформировать дугу в графе проблемной ситуации для перехода между двумя локальными вершинами, решается с помощью того же мультиагентного алгоритма, в котором задействованы агенты целеполагания, агенты управления действиями, агенты описания состояний, который был рассмотрен выше.

Этот алгоритм будет выполняться до тех пор, пока между начальной и целевой вершинами графа не будет построен связный путь, включающий в себя промежуточные вершины и дуги. После этого интеллектуальный агент выполняет действия, которые входят в этот путь и доступны ему для выполнения, чтобы реализовать построенный план. В этом случае действия интеллектуального агента могут рассматриваться как специфическая реакция, что до некоторой степени гарантирует значительное преимущество в эффективности в вышеописанном смысле. Следует обратить внимание на то, что среди всех приведенных на рисунке 2 действий интеллектуальный агент непосредственно сам может выполнить только одно – это действие «написать в чат «М», и именно оно и будет выполнено.

Если интеллектуальный агент не способен добиться того, чтобы непосредственно из состояния «Энергия уменьшилась» перейти в состояние «Получил молоко», то в состоянии «В чате появился символ «М» он перейти может, и для этого ему надо выполнить действие «написать в чат «М», что он и выполняет. Пользователь, получив от интеллектуального агента сообщение «М» («молоко»), конечно, гораздо быстрее поможет ему решить проблему. Построенное знание получает положительное подкрепление, символ «М» получает специфическую семантику и становится знаком в полном смысле слова, а интеллектуальный агент овладевает данным элементом того языка, который он осваивает.

Мультиагентный алгоритм автономного синтеза функциональных систем обоснования символов на основе роста и развития мультиагентных нейрокогнитивных архитектур, приведенный в данном разделе на примере простых односимвольных высказываний, по нашему мнению, универсален и может быть применен для обоснования символов различной природы, а также составных символов (например, слов) и высказываний на естественном языке.

4. ЗАКЛЮЧЕНИЕ

В результате выполнения исследования сформулированы основные принципы автономного построения онтологической семантики символов в процессе освоения интеллектуальным агентом естественного языка в коммуникативных ситуациях. Построенные на базе этих принципов мультиагентные алгоритмы основаны на процессах самоорганизации управляющих нейрокогнитивных архитектур в ситуациях идентификации и разрешения универсального спектра проблем в системе «интеллектуальный агент – среда» при взаимодействии в социальном окружении.

Разработаны основные принципы и универсальный мультиагентный алгоритм обоснования символов на основе формирования причинно-следственных отношений между агентами, выполняющими в мультиагентной нейрокогнитивной архитектуре функциональную репрезентацию событий, описывающих контекст коммуникативной ситуации (в каких случаях уместно высказывание), событий, описывающих предмет (тему и рему) высказывания (что именно является содержанием высказывания), и событий, описывающих само высказывание (какое именно высказывание, его конкретная форма).

Мы считаем, что совокупность сформулированных принципов и разработанных алгоритмов представляет собой решение фундаментальной научной проблемы обоснования символов, конструктивное с точки зрения его применимости, хорошей масштабируемости и открытости. Опора на метафору проектирования интеллектуального агента позволяет реализовать предлагаемое решение на самых различных программных и роботизированных платформах, а универсальность спектра проблем, решаемых такими системами под управлением мультиагентной нейрокогнитивной архитектуры, обеспечит вариативность контекстов коммуникативных ситуаций, необходимую для организации процесса обучения при освоении естественного языка.

REFERENCES

1. Russell S., Norvig P. *Artificial Intelligence: A Modern Approach*. Moscow: Williams, 2007. 1424 p. ISBN 0-13-790395-2. (In Russian)
Рассел С., Норвиг П. Искусственный интеллект: современный подход. 2-е издание. Москва: Вильямс, 2007. 1424 с. ISBN 0-13-790395-2.
2. Nagoev Z.V. *Intellektika, ili myshleniye v zhivykh i iskusstvennykh sistemakh* [Intelligence, or thinking in living and artificial systems]. Nalchik: KBSCof RAS Printing house, 2013. 232 p. (In Russian)
Нагоев З. В. Интеллектика, или мышление в живых и искусственных системах. Нальчик: Издательство КБНЦ РАН, 2013. 232 с.
3. Nagoev Z.V. Multiagent recursive cognitive architecture. *Biologically Inspired Cognitive Architectures. Proceedings of the third annual meeting of the BICA Society. Advances in Intelligent Systems and Computing series*. Springer, 2012. Pp. 247–248.
4. Deacon T. *The Symbolic Species*. New York: W. Norton and Co, 1997.
5. Vogt P. Language evolution and robotics: Issues on symbol grounding and language acquisition. *Artificial Cognition Systems*. Hershey, PA: Idea Group, 2006. Pp. 176–209.
6. Leichik V.M. *Terminovedeniye. Predmet, metody, struktura* [Terminology. Subject, methods, structure]. Moscow: KomKniga, 2005. 256 p. (In Russian)
Лейчик В. М. Терминоведение. Предмет, методы, структура. Москва: КомКнига, 2005. 256 с. (In Russian)
7. Morris Ch.W. Foundations of the theory of signs. *Semiotika: Antologiya* [Semiotics: Anthology]. Moscow: Yekaterinburg, 2001. P. 46. (In Russian)
Моррис Ч. У. Основания теории знаков // Семиотика: Антология. Изд. 2-е. Москва: Екатеринбург, 2001. С. 46.

8. Sakhachev N.L. *O semiotike Ch.S. Pirsas* [On the semiotics of Ch. S. Pierce]. St. Petersburg: Nauka, 2003. P. 103. (In Russian)
Сухачев Н. Л. О семиотике Ч. С. Пирса. СПб.: Наука, 2003. С. 103.
9. Frege G. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*. No. 100. 1892. Pp. 25–50.
10. Ogden C.K., Richards I.A. *The Meaning of Meaning*. New York, Harcourt, Brace & World, Inc., 1923.
11. Newell A., Simon H. A. Computer science as empirical inquiry: Symbols and search. *Communications of the ACM* 19. 1976. Pp. 113–126.
12. McCarthy J. Hayes P.J. Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence* 4. 1969. Pp. 463–502.
13. Brooks R.A. Elephants don't play chess. *Robotics and Autonomous Systems* 6. 1990. Pp. 3–15.
14. Peirce C. S. *Collected Papers, Volume I-VIII*. Cambridge MA.: Harvard University Press. (1931–1958).
15. Pfeifer R., Scheier C. *Understanding Intelligence*. Cambridge MA.: MIT Press. 1999.
16. Ziemke T., Sharkey N.E. A stroll through the worlds of robots and animals: Applying Jakob von Uexküll's theory of meaning to adaptive robots and artificial life. *Semiotica*. 2001. 134(1–4). Pp. 701–746.
17. Nagoev Z.V., Nagoeva O.V., Pshenokova I.A. Formal model of the semantics of natural language statements based on multi-agent recursive cognitive architectures. *News of Kabardino-Balkarian Scientific Center of RAS*. 2017. No. 4(78). Pp. 19–31. (In Russian)
Нагоев З. В., Нагоева О. В., Пшенокова И. А. Формальная модель семантики естественно-языковых высказываний на основе мультиагентных рекурсивных когнитивных архитектур // Известия Кабардино-Балкарского научного центра РАН. 2017. № 4(78). С. 19–31.
18. Nagoev Z.V. Ontoneuromorphogenetic modeling. *News of Kabardino-Balkarian Scientific Center of RAS*. 2013. No. 4(54). Pp. 46–56. (In Russian)
Нагоев З. В. Онтонейроморфогенетическое моделирование // Известия Кабардино-Балкарского научного центра РАН. 2013. № 4(54). С. 46–56.

Информация об авторах

Анчёков Мурат Инусович, науч. сотр. лаборатории «Молекулярная селекция и биотехнология», Кабардино-Балкарский научный центр РАН;

360000, Россия, г. Нальчик, ул. Кирова, 224;

murat.antchok@gmail.com, ORCID: <https://orcid.org/0000-0002-8977-797X>

Бжихатлов Кантемир Чамалович, канд. физ.-мат. наук, зав. лабораторией «Нейрокогнитивные автономные интеллектуальные системы», Кабардино-Балкарский научный центр РАН;

360002, Россия, г. Нальчик, ул. Балкарова, 2;

haosit13@mail.ru, ORCID: <https://orcid.org/0000-0003-0924-0193>

Канкулов Султан Ахмедович, стажер-исследователь лаборатории «Интеллектуальные среды обитания», Институт информатики и проблем регионального управления – филиал Кабардино-Балкарского научного центра РАН;

360000, Россия, г. Нальчик, ул. И. Арманд, 37-а;

skankulov@mail.ru, ORCID: <https://orcid.org/0000-0002-2996-7376>

Нагоев Залимхан Вячеславович, канд. техн. наук, генеральный директор Кабардино-Балкарского научного центра РАН;

360000, Россия, г. Нальчик, ул. И. Арманд, 37-а;

zaliman@mail.ru, ORCID: <https://orcid.org/0000-0001-9549-1823>

Нагоева Ольга Владимировна, н. с. отдела «Мультиагентные системы», Институт информатики и проблем регионального управления – филиал Кабардино-Балкарского научного центра РАН;

360000, Россия, г. Нальчик, ул. И. Арманд, 37-а;

nagoeva_o@mail.ru, ORCID: <https://orcid.org/0000-0003-2341-7960>

Пшенокова Инна Ауесовна, канд. физ.-мат. наук, вед. науч. сотр. лаборатории «Нейрокогнитивные автономные интеллектуальные системы», Кабардино-Балкарский научный центр РАН; 360002, Россия, г. Нальчик, ул. Балкарова, 2; pshenokova_inna@mail.ru, ORCID: <https://orcid.org/0000-0003-3394-7682>

Information about the authors

Anchekov Murat Inusovich, Researcher of the laboratory “Molecular selection and biotechnology”, Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences; 360000, KBR, Nalchik, 224 Kirov street; murat.antchok@gmail.com, ORCID: <https://orcid.org/0000-0002-8977-797X>

Bzhikhatlov Kantemir Chamalovich, Candidate of Physical-Mathematical Sciences, Head of the Laboratory “Neurocognitive Autonomous Intelligent Systems” of the Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences; 360002, Russia, Nalchik, 2 Balkarov street;

Kankulov Sultan Akhmedovich, Trainee researcher of the Laboratory “Intellectual Habitats” of the Institute of Computer Science and Problems of Regional Management of the KBSC of the Russian Academy of Sciences; 360000, Russia, Nalchik, 37-a I. Armand street; skankulov@mail.ru, ORCID: <https://orcid.org/0000-0002-2996-7376>

Nagoev Zalimkhan Vyacheslavovich, Candidate of Technical Sciences, General Director of the Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences; 360000, Russia, Nalchik, 37-a I. Armand street; zaliman@mail.ru, ORCID: <https://orcid.org/0000-0001-9549-1823>

Nagoeva Olga Vladimirovna, Researcher of the Department of the Multiagent Systems of the Institute of Computer Science and Regional Management Problems – branch of the Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences; 360000, Russia, Nalchik, 37-a I. Armand street; nagoeva_o@mail.ru, ORCID: <https://orcid.org/0000-0003-2341-7960>

Pshenokova Inna Auesovna, Candidate of Physical and Mathematical Sciences, Leading Researcher of the Laboratory “Neurocognitive autonomous intelligent systems”, Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences; 360002, Russia, Nalchik, 2 Balkarov street; pshenokova_inna@mail.ru, ORCID: <https://orcid.org/0000-0003-3394-7682>