

Классификация и условия применения алгоритмов автоматической онтологизации пространства состояний агента общего искусственного интеллекта под управлением нейрокогнитивной архитектуры*

**З. В. Нагоев¹, И. А. Пшенокова², М. И. Анчёков¹, К. Ч. Бжихатлов¹,
Б. А. Аталиков², С. А. Канкулов², А. З. Энес²**

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

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

Аннотация. Целью исследования является формирование алгоритмической базы онтологизации пространства состояний агента общего искусственного интеллекта. Основными задачами работы являются классификация алгоритмов автономного синтеза онтологий системы «агент – среда» и выбор условий их применения. Определены форма и содержание основных алгоритмов онтологизации пространства состояний интеллектуального агента общего искусственного интеллекта и построена их классификация. Обоснованы механизмы формирования концептов функциональной репрезентации сущностей пространства состояний интеллектуального агента на основе динамического порождения «по требованию». Построен мета-алгоритм выбора частных алгоритмов онтологизации в зависимости от выделенных условий применения.

Ключевые слова: автоматическое построение онтологий, общий искусственный интеллект, мультиагентные системы, нейрокогнитивные архитектуры

Поступила 02.12.2023, одобрена после рецензирования 08.12.2023, принята к публикации 10.12.2023

Для цитирования. Нагоев З. В., Пшенокова И. А., Анчёков М. И. и др. Классификация и условия применения алгоритмов автоматической онтологизации пространства состояний агента общего искусственного интеллекта под управлением нейрокогнитивной архитектуры // Известия Кабардино-Балкарского научного центра РАН. 2023. № 6(116). С. 210–225. DOI: 10.35330/1991-6639-2023-6-116-210-225

MSC: 68T42, 68T05

Review article

Classification and conditions of application of algorithms for automatic ontologization of the state space of a general artificial intelligence agent under the control of neurocognitive architecture*

**Z.V. Nagoev¹, I.A. Pshenokova², M.I. Anchekov¹, K.Ch. Bzhikhatlov¹,
B.A. Atalikov², S.A. Kankulov², A.Z. Enes²**

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

²Institute 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

© Нагоев З. В., Пшенокова И. А., Анчёков М. И., Бжихатлов К. Ч., Аталиков Б. А., Канкулов С. А., Энес А. З., 2023

* Работа выполнена при финансовой поддержке РФФИ (грант № 22-19-00787).

* The work was carried out with the financial support of the Russian Science Foundation (grant No. 22-19-00787).

Abstract. The purpose of the study is to form an algorithmic base for ontologizing the state space of a general artificial intelligence agent. The main objectives of the work are the classification of algorithms for the autonomous synthesis of ontologies of the “agent-environment” system and the selection of conditions for their application. The form and contents of the main ontologization algorithms for the state space of an intelligent agent of artificial general intelligence are determined and their classification is constructed. The mechanisms for the formation of concepts of functional representation of entities in the state space of an intelligent agent based on dynamic generation “on demand” are substantiated. A meta-algorithm for selecting private ontologization algorithms has been constructed depending on the selected application conditions.

Keywords: automatic construction of ontologies, artificial general intelligence, multi-agent systems, neurocognitive architectures

Submitted 02.12.2023,

approved after reviewing 08.12.2023,

accepted for publication 10.12.2023

For citation. Nagoev Z.V., Pshenokova I.A., Anchekov M.I. et al. Classification and conditions of application of algorithms for automatic ontologization of the state space of a general artificial intelligence agent under the control of neurocognitive architecture. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2023. No. 6(116). Pp. 210–225. DOI: 10.35330/1991-6639-2023-6-116-210-225

ВВЕДЕНИЕ

Проблема онтологизации пространства состояний системы общего искусственного интеллекта (ОИИ) характеризуется некоторыми существенными особенностями, отличающими ее от задач построения онтологий, решаемых специализированными интеллектуальными системами (системами так называемого «слабого искусственного интеллекта» в конкретных предметных областях) [1]. К их числу можно отнести, в частности, универсальную направленность, так как система ОИИ, предназначенная для решения универсального спектра задач, функционирует в пространстве состояний огромной размерности – в реальной среде¹.

Для обозначения такого пространства в [2] введено понятие системы «интеллектуальный агент – среда», основанное на процессе измерений сенсорной подсистемой интеллектуального агента интра- и экстрасенсорных физических данных. Для решения сложной проблемы эффективной работы в таком пространстве состояний естественные системы ОИИ выработали уникальный механизм так называемого *исследовательского поведения*.

Наличие феноменологического комплекса управления исследовательским поведением является одним из наиболее эффективных эволюционных приобретений разумных многоклеточных организмов. Суть этого комплекса состоит в формировании внутренних мотивационных стимулов для реализации контура управления идентификацией состояний дефицита онтологического знания относительно некоторых фрагментов системы «агент – среда» и синтеза поведения интеллектуального агента, направленного на устранение этого дефицита.

Отличительной этологической особенностью этой подсистемы является именно наличие проактивного контура внутренней мотивации, что связано с отложенностью результата действия, его направленностью на применение построенных онтологий в будущих возможных ситуациях. В этом смысле эта подсистема управления принципиально отличается от реактивных контуров управления, действующих в парадигме «стимул – реакция», и даже от проактивных контуров управления, ориентированных на достижение пусть отложенного, но некоторого конкретного результата.

Основным патентом природы, алгоритмическая сущность которого описана в [3], является формирование подсистемы внутренней мотивации, не связанной с внешними стиму-

¹https://www.researchgate.net/publication/271390398_Artificial_General_Intelligence_Concept_State_of_the_Art_and_Future_Prospects

лами поведения интеллектуального агента. Т.е. возникновение новой, ранее не встречавшейся ситуации вне зависимости от наличия непосредственных стимулов (в настоящем или будущем), связанных с увеличением значений целевой функции, приводит к тому, что в нейрокогнитивной архитектуре выполняется алгоритм, определяющий внутренний стимул в виде вознаграждения, заданного в терминах значений локальных целевых функций агентов соответствующих нейрокогнитивных, мотивирующий интеллектуального агента к синтезу и реализации исследовательского поведения, направленного на онтологизацию данной новой ситуации в интересах будущих применений.

Системный (биологический) смысл такой опережающей онтологизации, когда новые события (ситуации) онтологизируются как бы «впрок», состоит в том, что интеллектуальный агент в случае наступления такой ситуации в будущем будет готов к тому, каким образом действовать в ней некоторым субоптимальным по критериям целевой функции способом.

Из вышеизложенного, в частности, следует, что достроенная *онтология* в управляющей нейрокогнитивной архитектуре интеллектуального агента в общем случае представляет собой систему отношений между агентами в составе динамической функциональной системы (по Анохину). Такое представление об онтологии, формируемой интеллектуальным агентом для описания пространства состояний системы «агент – среда», демонстрирует глубокую сущностную связь с философской категорией бытия. Потребности управления бытием («онтос»), жизненным (функциональным) циклом интеллектуального агента определяют формы и содержание процессов онтологизации, реализуемых мультиагентной нейрокогнитивной архитектурой.

Как показано в [4], основным механизмом онтологизации отдельных сущностей является алгоритм порождения функциональных элементов мультиагентной нейрокогнитивной архитектуры интеллектуального агента «по требованию». Однако в случае агента общего искусственного интеллекта, действующего в пространстве гигантской размерности в интересах эффективного управления своим жизненным циклом, необходима не только онтологизация сущностей, но и их влияние на состояние целевой функции агента, а также построение системы причинно-следственных связей между концептами, описывающими различные фрагменты системы «агент – среда».

Если интеллектуальный агент в процессе исследования такой системы сталкивается с некоторым новым событием (например, с событием идентификации нового объекта), то прямой алгоритм онтологизации предполагает, что необходимо сначала создать онтологическое описание этого события (событие-причина), а затем – онтологическое описание влияния (потенциально возможного влияния) этого события на изменение состояния системы «агент – среда», в которых будут идентифицироваться новые, пока не наступившие (либо вообще ранее не наступавшие) события (события-следствия) в будущем.

Алгоритмы построения причинно-следственных зависимостей должны в этом случае автоматически построить необходимые ассоциативные связи как к событиям прошлого (причины), так и возможным событиям будущего (следствия).

В целях идентификации алгоритмов онтологизации системы «агент – среда» для их рационального ситуативно-детерминированного выбора при автоматическом применении в процессе обучения интеллектуального агента общего искусственного интеллекта необходимо четко определить функциональные различия и особенности их применения.

Таким образом, проблема классификации и выбора условий применения алгоритмов автономной онтологизации пространства состояний программного агента общего искусственного интеллекта является актуальной.

Целью настоящего исследования является формирование алгоритмической базы онтологизации пространства состояний агента общего искусственного интеллекта.

Основными задачами работы являются классификация алгоритмов автономного синтеза онтологий системы «агент – среда» и выбор условий их применения.

1. ПРОБЛЕМА КЛАССИФИКАЦИИ И ВЫБОРА УСЛОВИЙ ПРИМЕНЕНИЯ АЛГОРИТМОВ АВТОМАТИЧЕСКОЙ ОНТОЛОГИЗАЦИИ ПРОСТРАНСТВА СОСТОЯНИЙ АГЕНТА ОИИ

У агента ОИИ процесс онтологизации может быть ориентирован на построение причинно-следственных связей от события-причины к событию-следствию или, наоборот, от события-следствия к событию-причине. Порядок в данном случае определяет последовательность идентификации и обработки событий, связанных причинно-следственной связью.

В целях настоящего исследования *событием* мы считаем формирование некоторого мультиагентного факта. Например: «Некоторый объект существует» либо «Объект находится передо мной», либо «Объект вкусный» и т.п. Этот процесс достаточно подробно описан в [5]. Его суть состоит в том, что любая информация, описывающая систему «агент – среда» в нейрокогнитивной архитектуре интеллектуального агента, представляется так называемыми мультиагентными знаниями, организованными в формате так называемого *мультиагентного факта*, часть которого представлена знаниями о *теме* (сущность которой посвящена информации), а другая часть – *реме* – непосредственно знанием о некоторых свойствах этой сущности или действиями с ее участием (рис. 1).

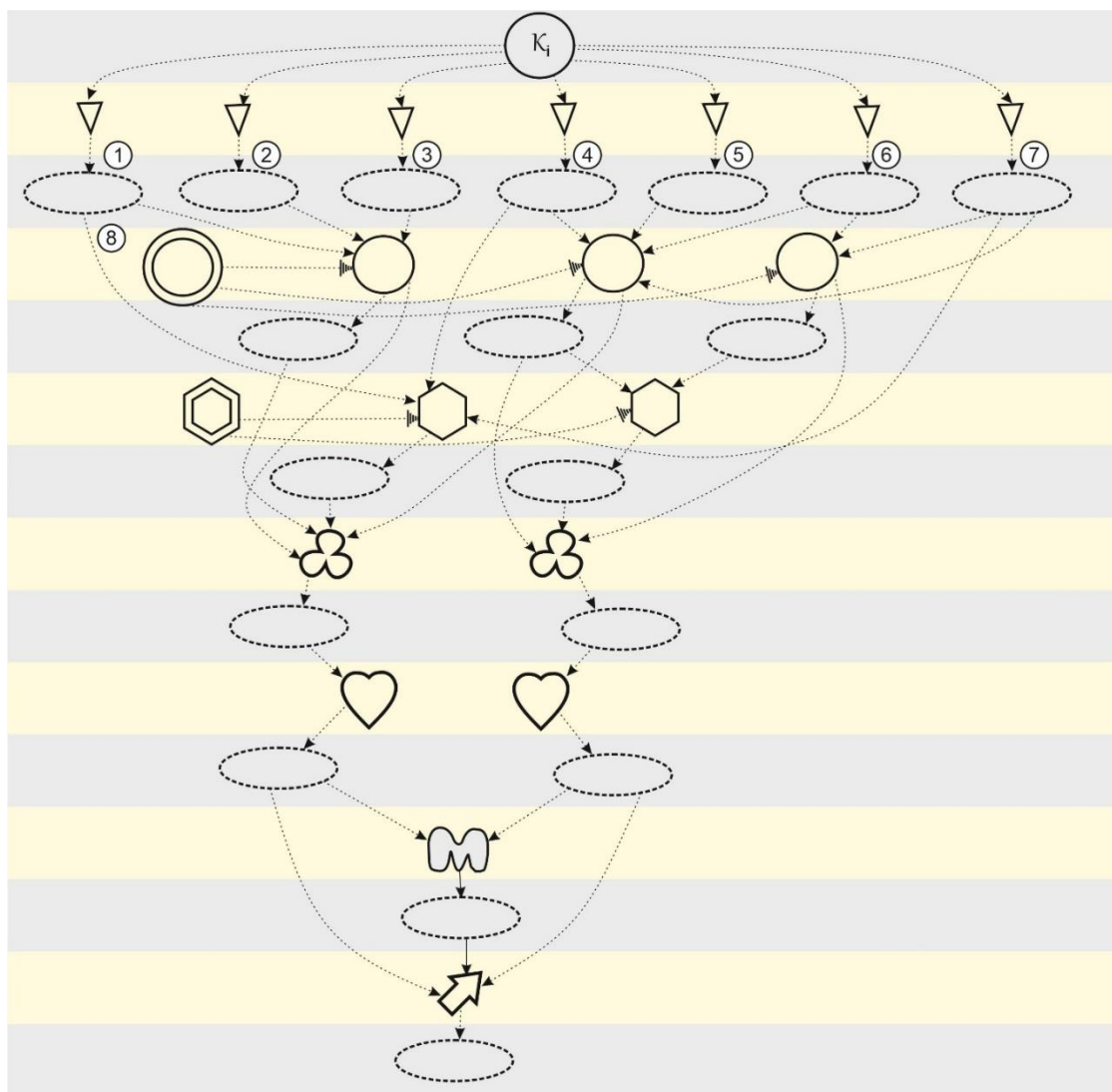


Рис. 1. Схематическое изображение функциональной репрезентации событий в пространстве состояний интеллектуального агента с помощью мультиагентных фактов

Fig. 1. Schematic representation of the functional representation of events in the state space of an intelligent agent using multi-agent facts

Под *сущностью* в данном случае мы понимаем некоторый *объект* или объекты, которые в явном виде выделяются сенсорной системой интеллектуального агента при наблюдении системы «агент – среда» и для функциональной репрезентации которых в мультиагентной нейрокогнитивной архитектуре агента выделяются (существуют) специальные агнейроны.

Под *действием* будем понимать некоторое преобразование, выполняемое интеллектуальным агентом над определенным объектом (объектами) (над определенным фрагментом системы «агент – среда»), находящимся в некотором состоянии, в результате которого, как правило, наступают изменения в составе, свойствах или состояниях этих и/или других объектов (интеллектуальный агент переходит в другое состояние пространства состояний системы «агент – среда»).

Можно считать, что действие объединяет два или более событий, так как состояния, между которыми переходит система «агент – среда» в результате его выполнения, описываются мультиагентными фактами, а они, как мы выяснили, как раз и есть события.

Алгоритмы онтологизации направлены на создание мультиагентного нейрокогнитивного описания текущего события, формирование таких же описаний событий, уже имевших место в прошлом, послуживших причинами наступления данного события, а также событий, которые, возможно, произойдут в будущем. Таким образом, алгоритмы онтологизации системы «агент – среда» направлены на формирование цепочки причинно-следственных зависимостей, связывающих текущее событие с событиями-причинами и событиями-следствиями (рис. 2).

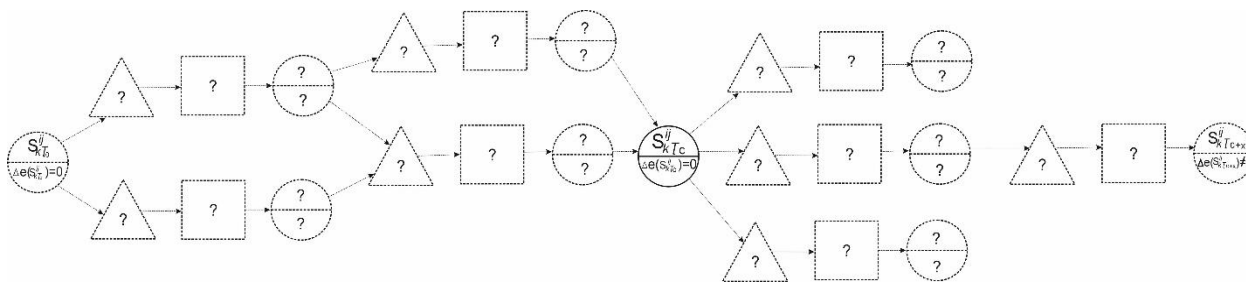


Рис. 2. «Пустое» дерево решений проблемной ситуации

Fig. 2. "Empty" tree of solutions to a problem situation

На рисунке приведено изображение т.н. *дерева решений текущей проблемной ситуации*, которое задает графическое представление процесса синтеза поведения интеллектуального агента. Кругами в нем представлены состояния системы «агент – среда», треугольниками – действия интеллектуального агента, а квадратами – ответные действия среды. Из рисунка следует, что в данный момент времени интеллектуальный агент однозначно идентифицировал только те события, которые описывают текущую ситуацию (обозначена символами в круге, оконтуренном сплошной жирной линией). Остальные (возможные) элементы дерева пока интеллектуальному агенту неизвестны, поэтому обозначены тонкой пунктирной линией. Заполнение этого дерева интеллектуальный агент может выполнить только на основе существующих, полученных из предыдущего опыта, сформированных на основе исследовательского поведения или приобретаемых «по случаю» («по требованию») в данном эпистемологическом контексте.

По нашей гипотезе, различные алгоритмы онтологизации всегда выполняются совместно, так как целесообразная картина мира, которая формируется в управляющей нейрокогнитивной архитектуре интеллектуального агента, должна быть логически целостной, не содержащей необусловленных онтологических единиц. Наличие таких единиц является ситуацией, формирующей внутреннюю мотивацию интеллектуального агента к синтезу и выполнению *исследовательского поведения*.

Как было показано в [2], ответить на вопрос о пользе или вреде, который может принести данное событие, интеллектуальный агент может, только получив соответствующие знания. Знания же эти он может получить только тремя основными способами: а) экспериментальным; б) аналитическим; в) коммуникативным.

Экспериментальный способ предполагает, что интеллектуальный агент выполняет над объектами, участвовавшими в формировании рассматриваемого мультиагентного факта, некоторые действия и регистрирует события, которые наступили в результате. Такой эксперимент позволяет агенту с достаточной степенью правдоподобия установить, как именно изменяется состояние системы «интеллектуальный агент – среда»: от начального состояния, частью описания которого является стартовое событие, в результате выполнения некоторого конкретного действия, – к некоторому конечному состоянию, в которое переходит система (рис. 3).

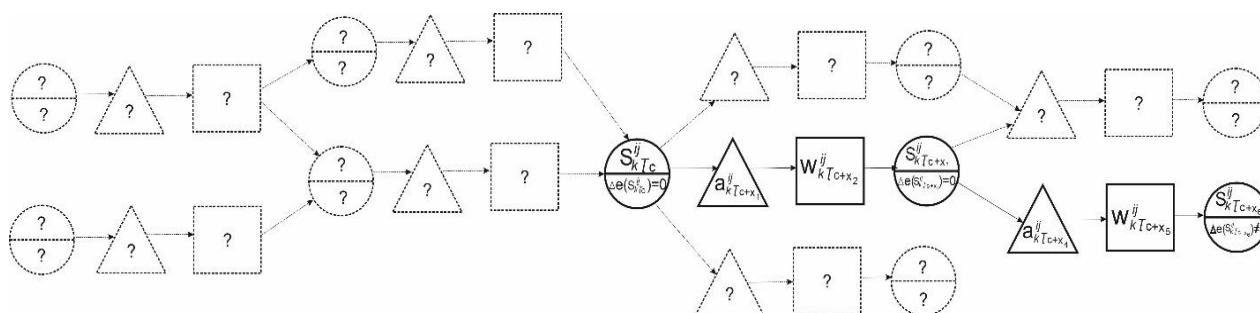


Рис. 3. Результат выполнения алгоритма прямой экспериментальной онтологизации

Fig. 3. Result of executing the direct experimental ontologization algorithm

На рисунке показано, что интеллектуальный агент заполняет один из потенциально возможных сценариев пространства альтернатив, переходя от текущего состояния к будущим (показаны кругами с обозначениями, оконтуренными жирными линиями), в результате выполнения фактического эксперимента.

Экспериментальный способ отличается достоверностью, однако, во-первых, для его реализации требуются определенные условия – такие, например, как управление монопольными ресурсами и структурными элементами интеллектуального агента (части тела, энергия), наличие времени и некоторых обязательных конкретных условий для выполнения эксперимента.

Аналитический способ основан на самостоятельном (внутреннем) моделировании последствий данного события в процессе «мысленного эксперимента». В ходе такого моделирования выполняется мультиагентный алгоритм, имитирующий процесс рассуждения на заданную тему, в котором основная функциональная роль по предсказанию последствий (развитию ситуации, сценария событий) отводится агнейронам из нейрокогнитона моделирования, ансамблям событийных агнейронов, выполняющих в нейрокогнитивной архитектуре функциональную репрезентацию взаимозависимых фрагментов системы «интеллектуальный агент – среда», и т.н. *агнейронам действия*, которые, соответственно, выполняют в нейрокогнитивной архитектуре функциональную репрезентацию действий, реализуя которые интеллектуальный агент переходит между такими фрагментами. При этом недостающие знания производятся самой управляющей нейрокогнитивной архитектурой интеллектуального агента на основе моделирования и прогнозирования. Такой способ представляется гораздо более экономичным, однако существенной проблемой является точность (качество) знаний, продуцируемых таким способом (рис. 4).

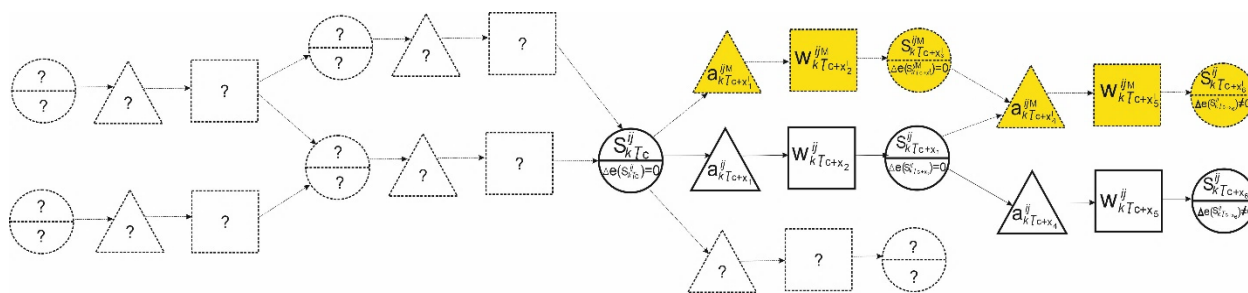


Рис. 4. Результаты выполнения алгоритма прямой аналитической онтологизации

Fig. 4. Results of executing the direct analytical ontologization algorithm

На рисунке результаты выполнения алгоритма прямой аналитической онтологизации показаны закрашенными элементами дерева решений проблемной ситуации с обозначениями, оконтуренными тонкой пунктирной линией (так как это воображаемое будущее).

На рисунке 5 показано, что аналитический метод вполне применим и к событиям прошлого.

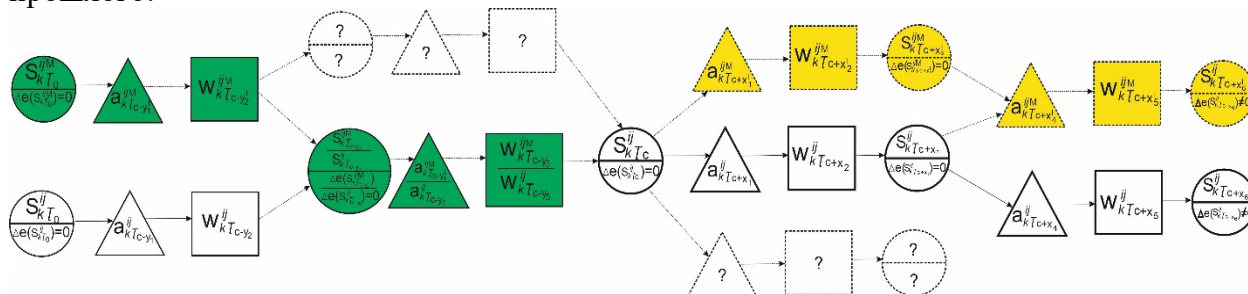


Рис. 5. Результаты совместной работы прямого и обратного алгоритмов аналитической онтологизации

Fig. 5. Results of joint work of forward and reverse algorithms analytical ontologization

При коммуникативном способе интеллектуальный агент не затрачивает вообще никакие ресурсы на проведение ни мысленных, ни физических экспериментов. Вместо этого он просто обращается за необходимой информацией (знаниями) к другим интеллектуальным агентам, используя доступные коммуникативные каналы. В этом случае полученные знания могут характеризоваться различной степенью точности. Безусловным преимуществом является то, что такие знания выражены в абстрактной знаковой форме, что позволяет существенным образом повысить удельную интенсивность обучения, передавая значительно большее количество информации в единицу времени (рис. 6).

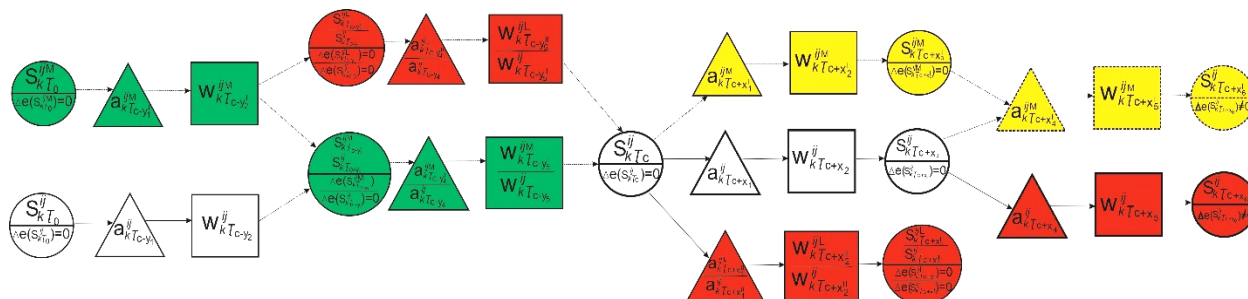


Рис. 6. Результаты применения прямого и обратного алгоритмов коммуникативной онтологизации

Fig. 6. Results of applying the forward and inverse algorithms communicative ontologization

Сохраняя сильные стороны двух предыдущих способов, коммуникативное обучение преодолевает их недостатки, так как, с одной стороны, в передаваемом, как правило с помощью естественного языка, сообщении описываются и начальное событие, и действия, которые необходимо выполнить для достижения семантически значимого результата, и конечные события, характеризующие ситуацию, в которой эти результаты должны проявиться. При этом в нейрокognитивной архитектуре строятся и сохраняются все необходимые модели, которые могут, например, в дальнейшем понадобится для имитационного моделирования самых разнообразных ситуаций с участием задействованных в контексте обучения событий.

С другой же стороны, при реализации коммуникативного способа нет никаких ограничений ни на ресурсные затраты, необходимые для проведения физического эксперимента, ни на качество знаний, которые могли бы, например, проявиться при обучении на основе самостоятельного внутреннего моделирования.

Помимо классификации способов обучения *по критерию источника знаний*, получаемых интеллектуальным агентом, введем также классификацию *по критерию направления процесса верификации* нового знания.

Для того чтобы оценить такое направление, рассмотрим понятие *ситуации*, идентифицируемой управляющей мультиагентной нейрокognитивной архитектурой интеллектуального агента, введенное в [2]. Под этим термином мы будем понимать множество событий, связанных причинно-следственными и/или темпоральными зависимостями, одно или более из которых характеризуются существенным изменением значения целевой функции интеллектуального агента. У каждой ситуации есть начало и окончание, в явном виде связанные с наступлением определенных событий, а также некоторая внутренняя логика, позволяющая рассматривать множество входящих в эту ситуацию событий – от начальных через промежуточные к завершающим – в качестве комплекса событий, с достаточной степенью вероятности повторяющихся при наступлении данной ситуации.

Например, если утром выпало много снега (1-е событие), в обед температура поднялась выше нуля (2-е событие), к вечеру часть снега подтаяла (3-е событие), а ночью температура опустилась ниже нуля (4-е событие), то, вероятнее всего, на тротуаре будет гололедица (5-е событие), и, соответственно, перемещаться по улице будет некомфортно – от человека (интеллектуального агента) потребуются гораздо больше усилий, чем обычно (6-е, завершающее, событие). В этой ситуации часть событий связана друг с другом причинно-следственными отношениями, а часть – только темпоральными. Однако каждое из этих событий существенным образом влияет на вероятность появления результирующего, завершающего события, наступление которого для интеллектуального агента было бы связано со значительными дополнительными потерями энергии.

Таким образом, любая ситуация однозначно идентифицируется составом и последовательностью входящих в нее событий. При этом часть этих событий сами по себе, взятые отдельно, могут быть нейтральными с точки зрения своего непосредственного влияния на значение целевой функции, однако внутренняя логика, объединяющая эти события в ситуацию, неизбежно ведет их последовательность к событию, характеризующемуся существенным изменением (позитивным или негативным) значения целевой функции.

Так как ключевым фактором преимущества интеллектуальных систем в эффективности является *проактивное принятие решений*, принципиальное значение для реализации этого преимущества имеет *длина горизонта планирования*. Соответственно, чем раньше интеллектуальным агентом будут идентифицированы первые события некоторой ситуации, тем больше шагов (элементов последовательности возникающих событий) в ее составе останется до наступления события, в котором интеллектуальный агент может испытать существенные изменения значений целевой функции. Такая ранняя идентификация, следовательно, дает больше шансов для проактивного принятия решений, направленных либо на

устойчивое движение к результирующему событию, если это событие позитивно, либо на активное противодействие его наступлению в случае, когда такое событие негативно.

С учетом того, что реальная среда, в которой функционирует интеллектуальный агент, является неструктурированной, неопределенной, эпизодической и частично наблюдаемой, априорная вероятность встречи интеллектуального агента с любым из событий в составе некоторой ситуации примерно одинакова. Следовательно, агент может либо сначала пройти через череду нейтральных событий и лишь в конце ситуации встретить результирующее событие, либо сразу столкнуться с результирующим событием, что, соответственно, приведет к потере или приобретению энергии. В зависимости от того, по какому из этих двух сценариев идет ситуация, – от нейтральных событий в составе ситуации к результирующему событию или наоборот, – мы и классифицируем алгоритмы онтологизации по направлению верификации, выделяя соответственно *прямой и обратный алгоритмы онтологизации*.

Необходимо отметить, что все процессы формирования функциональных связей между агнейронами в составе мультиагентной нейрокогнитивной архитектуры выполняются на основе т.н. *алгоритма онтонейроморфогенеза* [6], определяющего гипотетический механизм формирования связей в динамических функциональных системах головного мозга на основе процессов ситуативно-детерминированного роста и деградации аксо-дендрональных связей между биологическими нейронами.

Так как перебор всех возможных потенциальных вариантов ситуаций, в которых рассматриваемое событие может «встретиться» (возникнуть) как составная часть, является практически нереализуемой задачей, необходим механизм, который позволял бы обучающемуся интеллектуальному агенту (ребенку, программному агенту, роботу) непосредственно «на месте» первоначального возникновения этого события сразу же найти такие ситуации, в которых данное событие несло бы в себе угрозу либо было бы полезно.

Таким механизмом в социальной среде является интерактивное взаимодействие с наставниками (родителями, учителями, коллегами, друзьями), в ходе которого интеллектуальный агент запрашивает и получает необходимую информацию на основе прямой или косвенной коммуникации (рис. 7).

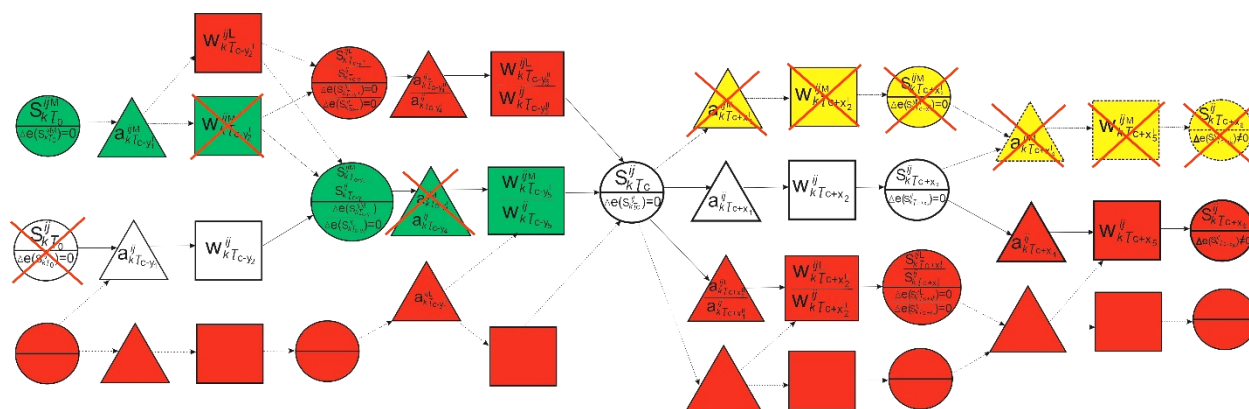


Рис. 7. Развитие и частичная деградация дерева проблемной ситуации за счет многократного применения алгоритмов прямой и обратной онтологизации

Fig. 7. Development and partial degradation of the problem situation tree due to repeated use of direct and reverse ontologization algorithms

Следует обратить внимание, что так же, как и при применении аналитических алгоритмов онтологизации, возможно многократное применение коммуникативных алгоритмов для развития дерева решений проблемной ситуации, а также для его прореживания путем удаления элементов, потерявших актуальность.

Таким образом, в результате анализа процессов формирования функциональных связей на основе алгоритма онтонейроморфогенеза в управляющей нейрокогнитивной архитектуре интеллектуального агента можно выделить следующие способы онтологизации: *алгоритм порождения концептов, прямой экспериментальный алгоритм, обратный экспериментальный алгоритм, прямой аналитический алгоритм, обратный аналитический алгоритм, прямой коммуникативный алгоритм, обратный коммуникативный алгоритм.*

В процессе реализации каждого из этих алгоритмов интеллектуальный агент выполняет исследовательское поведение, направленное на достройку недостающих знаний с помощью соответствующего мультиагентного алгоритма синтеза, реализующего формирование соответствующей цели, синтез необходимых для ее достижения действий, их выполнение и обучение интеллектуальной системы управления на основе измерений по факту исполнения этих действий.

Как следует из вышеизложенного, цель онтологизации пространства состояний интеллектуального агента – формирование системы знаний, необходимых и достаточных для эффективной реализации функционального назначения, может быть реализована только с помощью комплексного, совместного применения всех типов этих алгоритмов. Каковы же существенные условия выбора каждого из этих алгоритмов при реализации сложного исследовательского поведения интеллектуального агента, направленного на автономный синтез онтологий системы «агент – среда».

2. УСЛОВИЯ ПРИМЕНЕНИЯ АЛГОРИТМОВ АВТОМАТИЧЕСКОЙ ОНТОЛОГИЗАЦИИ

Как правило, различные способы онтологизации используются интеллектуальным агентом в процессе его индивидуального развития в соответствии с логикой, обусловленной текущим *эпистемологическим контекстом*. Под таким контекстом мы, собственно, и понимаем существенные условия и обстоятельства, в зависимости от которых интеллектуальный агент извлекает знания в процессе онтологизации системы «агент – среда» тем или иным способом. К существенным условиям, определяющим эпистемологический контекст, относятся комбинации таких обстоятельств, характеризующих текущую ситуацию, как степень новизны онтологизируемого события, направление онтологизации, наличие, возможность и сложность доступа к полному набору знаний, необходимых для построения комплекса онтологий, связанных с данным событием.

Если у интеллектуального агента есть возможность выбора способов онтологизации некоторого фрагмента системы «агент – среда», то он может выработать определенные предпочтения относительно того, в каких именно эпистемологических контекстах какие именно из этих способов применять.

После того, как алгоритмы распознавания регистрируют некоторую текущую ситуацию, производится оценка системной (биологической) значимости этой ситуации с помощью эмоциональных нейрокогнитивных. Так как ситуация неизвестная, в составе нейрокогнитивного тракта подсознания с высокой вероятностью отсутствуют агнейроны, которые отвечали бы за ее обработку. Поэтому рационально было бы предположить, что оценка входного события как неизвестного в этом когнитивном тракте должна приводить к постановке цели, достижение которой приведет к устранению этого пробела в знаниях интеллектуального агента. Как минимум, необходимо выяснить, несет ли это событие какую-либо угрозу или, наоборот, с ним можно связать некоторые надежды на достижение более высоких значений целевой функции интеллектуального агента в настоящей ситуации либо в возможных будущих ситуациях, в которых может возникнуть такое же (аналогичное) событие.

В условиях неопределенности нейрокогнитивной архитектурой интеллектуального агента используется априорная информация, связанная с базовым системным (биологическим) значением объектов, задействованных в новом событии, такая, например, как знание о том, пригоден ли предмет в пищу, каковы его экстеросенсорные свойства (форма, цвет, запах и др.).

Если за стартовую точку процесса онтологизации системы «агент – среда» принять начало функционирования интеллектуального агента в данной среде, то можно констатировать ситуацию практически полного отсутствия у него знаний об этой системе (исключая базовый предустановленный набор онтологий). Отсутствие знаний практически исключает возможность эффективного использования аналитического способа синтеза онтологий.

На рисунке 8 приведена блок-схема разработанного мета-алгоритма выбора интеллектуальным агентом частных алгоритмов онтологизации в зависимости от условий применения.

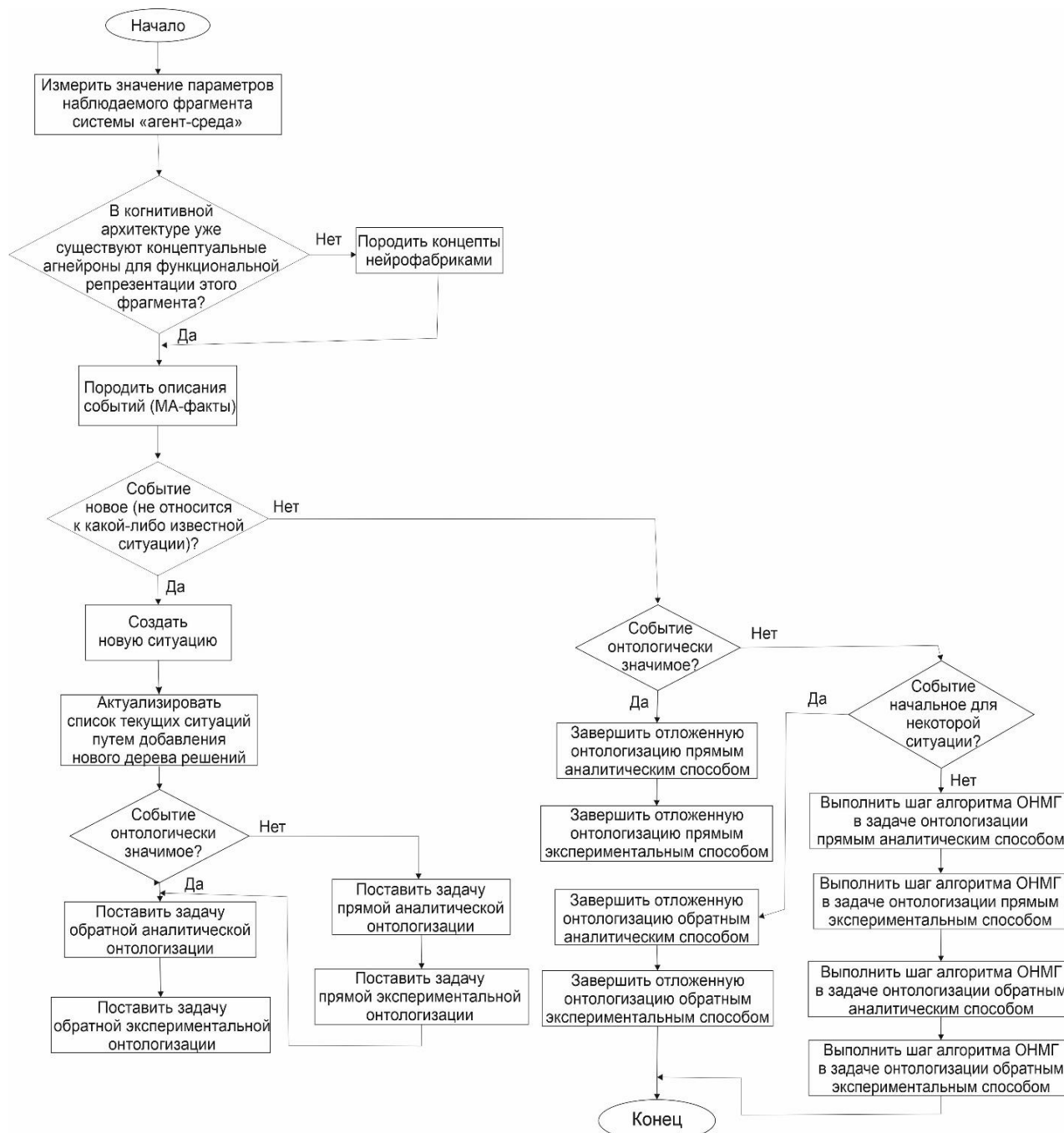


Рис. 8. Мета-алгоритм выбора алгоритмов онтологизации пространства состояний интеллектуального агента

Fig. 8. Meta-algorithm for selecting space ontologization algorithms intelligent agent states

Учитывая отсутствие опыта функционирования, можно также утверждать, что интеллектуальный агент пока не переходил (не попадал) ни в какие ситуации, т.е. система «агент – среда» еще не переходила ни в «хорошие» состояния (в которых значение целевой функции увеличивается), ни в «плохие» ситуации (в которых соответственно значение целевой функции уменьшается). Строго говоря, в начальный момент времени эта система вообще характеризуется опытом пребывания только в одном-единственном состоянии.

Так как ввиду полного отсутствия социального опыта в начальный момент времени интеллектуальный агент также не знает никакого языка, коммуникативный способ онтологизации, основанный на обмене абстрактными языковыми описаниями событий, ему также не доступен.

Больше того, в начальный момент времени потоки данных, описывающие систему «интеллектуальный агент – среда», в принципе являются неструктурированными, так как в мультиагентной нейрокогнитивной архитектуре интеллектуального агента пока еще нет агентов, выполняющих функциональную репрезентацию сущностей, выделяемых из этих потоков данных сенсорной системой интеллектуального агента на основе действия комплекса предустановленных (врожденных) априорных паттернов распознавания.

Следовательно, в начальный момент функционирования интеллектуальному агенту доступен только один способ онтологизации системы «агент – среда» – концептуализация входных потоков неструктурированных многомодальных данных [7, 8]. Такая онтологизация предполагает как порождение нейрофабриками агентов, выполняющих функциональную репрезентацию фрагментов системы «агент – среда» в различных категориях сущностной парадигмы концептов – объекты, действия, признаки и т.д., так и формирование мультиагентных фактов, выполняющих уже функциональную репрезентацию событий в системе «агент – среда», в составе которых ранее порожденные единичные концепты как раз и используются для передачи отношений между различными фрагментами наблюдаемой системы. Для функциональной репрезентации таких мультиагентных фактов в нейрокогнитивной архитектуре интеллектуального агента предусмотрены т.н. *событийные агенты* (рис. 1).

Непосредственно после начала процесса онтологизации системы «агент – среда» на основе концептуализации при помощи порождения нейрофабриками агентов и автоматического формирования в нейрокогнитивной архитектуре мультиагентных фактов интеллектуальному агенту становится доступен еще только один новый способ онтологизации – прямая экспериментальная онтологизация.

Этот эффект проявляется в силу того, что концептуальные агенты, порождаемые нейрофабриками в режиме реального времени, постепенно начинают заполнять соответствующие нейрокогнитоны, принимать участие в построении мультиагентных фактов, описывающих события в системе «агент – среда», а также активно и автономно взаимодействовать друг с другом с целью формирования динамических мультиагентных функциональных систем, ориентированных на решение текущих проблем на основе построения причинно-следственных зависимостей.

Сразу же после перехода в состояние, в котором значение целевой функции интеллектуального агента претерпевает существенное изменение («плохие» и «хорошие» состояния), для обозначения которого мы будем применять термин «онтологически значимые состояния» интеллектуального агента (имея в виду опосредованное влияние изменения значения целевой функции на степень функциональной целостности и готовности этого агента, его «онтос» – «бытие»), у интеллектуального агента возникает возможность использования метода обратной экспериментальной онтологизации. Необходимым условием для этого является еще и участие в формировании ситуации, завершающим событием которой и стал переход в онтологически значимое состояние, событий, ранее онтологизированных интеллектуальным агентом в составе некоторых других ситуаций.

После накопления некоторой базы знаний о системе «агент – среда» становятся принципиально возможными эпистемологические контексты, в которых интеллектуальный агент выполняет уже аналитический синтез онтологий, так как «машина рассуждений» в виде инварианта КА позволяет использовать совокупность некоторых (уже не пустых) нейрокогнитивных, таких, например, как моделирующий нейрокогнитивный потенциал, синтез целей, действий, оценки, управления для выполнения цикла внутреннего «ментального» мультиагентного нейрокогнитивного моделирования ситуаций, в которые входят те или иные определенные события.

На этом уровне развития нейрокогнитивной архитектуры интеллектуальный агент уже обладает всеми нужными функциональными элементами (нейроны-концепты, мультиагентные алгоритмы, реализующие те или иные методы рассуждений), которые необходимы для выполнения как прямых, так и обратных способов онтологизации как экспериментальным, так и аналитическим способом.

Строго говоря, учитывая необходимость постоянного совершенствования мультиагентных алгоритмов нейрокогнитивного моделирования системы «агент – среда», интеллектуальный агент на данном этапе своего развития, как правило, применяет все четыре вышеупомянутых способа конструирования онтологий [8–10].

Так, например, если сначала интеллектуальный агент идентифицирует новое событие, с которым ранее он не сталкивался, и если оно само по себе оценивается как нейтральное, то далее в этом эпистемологическом контексте интеллектуальному агенту доступен как прямой аналитический, так и прямой экспериментальный способ онтологизации.

В первом случае с помощью нейрокогнитивного моделирования когнитивная архитектура пытается построить возможные сценарии развития ситуации, начало которой может быть положено данным идентифицированным событием, которые представляют последствия данного события в виде цепочек воображаемых (еще не наступивших) событий в будущем.

Во втором интеллектуальный агент опирается на данные фактических измерений, регистрируемые его сенсорной системой по результату наступления событий, последовавших за первым идентифицированным событием (рис. 8).

Если же интеллектуальный агент идентифицирует некоторое нейтральное событие, не являющееся новым, но входящее (в черед других сопутствующих событий) в некоторую ранее имевшую место ситуацию (событие, ранее встречавшееся в некотором эпистемологическом контексте), а затем наступает некоторое новое для данного эпистемологического контекста онтологически значимое событие, можно говорить о запуске обратного экспериментального алгоритма онтологизации.

Способ обратной аналитической онтологизации запускается интеллектуальным агентом в работу в том случае, когда он идентифицирует некоторое известное событие, которое ранее регистрировалось в качестве неначального события некоторой известной ситуации. Констатация факта наступления известного события в таком новом эпистемологическом контексте приводит интеллектуального агента к необходимости учесть данное событие в составе некоторой новой ситуации. Агент может выполнить эту задачу только комбинацией применения нескольких способов – прямой аналитический или прямой экспериментальный способ он может использовать для достройки ситуации в направлении будущего онтологически значимого события (необходимо понять, к чему приведет наступление данного события). Соответственно, обратный аналитический и обратный экспериментальный способы будут применяться для достройки ситуации в направлении прошедших событий, так как всегда существует также необходимость понять, какое событие явилось стартовым для текущей ситуации.

Разница между обратным экспериментальным и обратным аналитическим способами состоит в том, что с помощью первого из них интеллектуальный агент может проверить участие

в текущей ситуации только тех событий, которые были идентифицированы в прошлом самим этим агентом в результате самостоятельно выполненных измерений соответствующих значенных параметров системы «агент – среда». Что же касается обратного аналитического способа, то с его помощью интеллектуальный агент может идентифицировать даже те события, которые по факту своего наступления в прошлом непосредственно не были зарегистрированы сенсорами интеллектуального агента. Как следует из краткого описания такого эпистемологического контекста, эти непосредственно не наблюдавшиеся события также восстанавливаются интеллектуальным агентом с помощью функциональной системы воображения, работа которой обеспечивается нейрокогнитивной архитектурой. Соответственно, будучи «воображаемыми в прошлом», эти события могли на самом деле (фактически) никогда и не иметь места в действительности. Для их восстановления с помощью описанного механизма, так же, как и для восстановления событий в прошлом, интеллектуальный агент использует агнейроны моделирующего когнитона нейрокогнитивной архитектуры [9].

С точки зрения процессов, протекающих в мультиагентной нейрокогнитивной архитектуре, принципиальное различие между экспериментальным и аналитическим способом онтологизации состоит в том, что при экспериментальном способе алгоритм онтонейроморфогенеза выполняется путем образования новых контрактных связей между агнейронами событийного нейрокогнитона, активизирующимся в результате выполнения функциональной репрезентации фактически наступивших событий, а при аналитическом способе – развитие новых аксо-дендрональных связей происходит между такими же событийными агнейронами уже на основе выполнения ими функциональной репрезентации воображаемых событий, активация которых наступает в результате работы прежде всего моделирующих агнейронов (агнейронов моделирующего нейрокогнитона).

ЗАКЛЮЧЕНИЕ

Таким образом, в результате выполнения данного исследования определены форма и содержание основных алгоритмов онтологизации пространства состояний интеллектуального агента общего искусственного интеллекта и построена их классификация. Обоснованы механизмы формирования концептов функциональной репрезентации сущностей пространства состояний интеллектуального агента на основе порождения «по требованию» с помощью нейрофабрик. Дано обоснование применения комплекса мета-алгоритмов исследовательского поведения таких агентов в условиях реальной среды для формирования систематизированных знаний, необходимых и достаточных для реализации эффективного функционирования по назначению применения. Построен мета-алгоритм выбора частных алгоритмов онтологизации системы «агент – среда» в зависимости от выделенных условий применения.

REFERENCES

1. Russell S., Norvig P. *Artificial Intelligence: A Modern Approach (AIMA)*. 2nd ed. Moscow: Williams, 2007. 1424 p.
2. *Нагоев З. В.* Интеллектика, или Мышление в живых и искусственных системах. Нальчик: Издательство КБНЦ РАН, 2013. 232 с.
3. Nagoev Z.V. *Intellektika, ili Myshleniye v zhivykh i iskusstvennykh sistemakh* [Intelligence, or Thinking in living and artificial systems]. Nalchik: Izdatel'stvo KBNTS RAN, 2013. 232 p. (In Russian)
3. Nagoev Z., Nagoeva O., Anchekov M. et al. The symbol grounding problem in the system of general artificial intelligence based on multi-agent neurocognitive architecture. *Cognitive Systems Research* [this link is disabled](#), 2023. No. 79. Pp. 71–84.

4. *Нагоев З. В., Нагоева О. В.* Обоснование символов и мультиагентные нейрокогнитивные модели семантики естественного языка. Нальчик: Издательство КБНЦ РАН, 2022. 150 с.

Nagoev Z.V., Nagoeva O.V. *Obosnovaniye simvolov i mul'tiagentnyye neyrokognitivnyye modeli semantiki yestestvennogo yazyka* [Symbol grounding and multi-agent neurocognitive models of natural language semantics]. Nalchik: Izdatel'stvo KBNTS RAN, 2022. 150 p. (In Russian)

5. *Апшев А. З., Аталиков Б. А., Канкулов С. А. и др.* Онтофилогенетические алгоритмы синтеза фенотипов интеллектуальных программных агентов для применения в задачах многопоколенной оптимизации управляющих нейрокогнитивных архитектур // Известия Кабардино-Балкарского научного центра РАН. 2022. № 6(110). С. 76–91. DOI: 10.35330/1991-6639-2022-6-110-76-91

Apshev A.Z., Atalikov B.A., Kankulov S.A. et al. Ontophylogenetic algorithms for the synthesis of phenotypes of intelligent software agents for use in multi-generation optimization problems of control neurocognitive architectures. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2022. No. 6(110). Pp. 76–91. DOI: 10.35330/1991-6639-2022-6-110-76-91. (In Russian)

6. *Нагоев З. В.* Онтонейроморфогенетическое моделирование // Известия Кабардино-Балкарского научного центра РАН. 2013. № 4(54). С. 46–56.

Nagoev Z.V. Ontoneuromorphogenetic modeling. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2013. No. 4(54). Pp. 46–56. (In Russian)

7. DiFilippo N.M., Jouaneh M.K. Using the soar cognitive architecture to remove screws from different laptop models. *IEEE Trans Autom Sci Eng*. 2018. No. 16(2). Pp.767–780.

8. Fittner M., Brandstatter C. How human inspired learning enhances the behavior of autonomous agents. *JCP*. 2018. No 13(2). Pp.154–160.

9. Gobet F., Lane P. The chrest architecture of cognition: The role of perception in general intelligence. *In: 3d Conference on Artificial General Intelligence (AGI-2010)*, Atlantis Press. 2010. <https://doi.org/10.2991/agi.2010.20>

10. González-Casillas A., Parra L., Martin L. et al. Towards a model of visual recognition based on neurosciences. *Proc Comput Sci*. 2018. Vol. 145. Pp. 214–231.

Информация об авторах

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>

Пшенокова Инна Ауесовна, канд. физ.-мат. наук, зав. лаб. «Интеллектуальные среды обитания», Институт информатики и проблем регионального управления – филиал Кабардино-Балкарского научного центра РАН;

360000, Россия, г. Нальчик, ул. И. Арманд, 37-а;
pshenokova_inna@mail.ru, ORCID: <https://orcid.org/0000-0003-3394-7682>

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>

Аталиков Борис Анзорович, 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;
atalikov10@gmail.com

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>

Энес Ахмед Зюлфикар, мл. науч. сотр. лаборатории «Компьютерная лингвистика», Институт информатики и проблем регионального управления – филиал Кабардино-Балкарского научного центра РАН;

360000, Россия, г. Нальчик, ул. И. Арманд, 37-а;
ahmedenes@mail.ru, ORCID: <https://orcid.org/0000-0003-3633-4910>

Information about the authors

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>

Pshenokova Inna Auesovna, Candidate of Physical and Mathematical Sciences, Head of the Laboratory “Intelligent Living Environments”, 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;
pshenokova_inna@mail.ru, ORCID: <https://orcid.org/0000-0003-3394-7682>

Anchekov Murat Inusovich, Researcher of the Laboratory “Molecular Selection and Biotechnology”, Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences;

360000, Russia, Nalchik, 224 Kirov street;
murat.antchok@gmail.com, ORCID: <https://orcid.org/0000-0002-8977-797X>

Bzhikhatlov Kantemir Chamalovich, Candidate of Physical and Mathematical Sciences, Head of the Laboratory “Neurocognitive Autonomous Intelligent Systems”, Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences;

360002, Russia, Nalchik, 2 Balkarov street;
haosit13@mail.ru, ORCID: <https://orcid.org/0000-0003-0924-0193>

Atalikov Boris Anzorovich, Trainee Researcher of the Laboratory “Intellectual Habitats”, 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;
atalikov10@gmail.com

Kankulov Sultan Akhmedovich, Trainee Researcher of the Laboratory “Intellectual Habitats”, 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;
skankulov@mail.ru, ORCID: <https://orcid.org/0000-0002-2996-7376>

Enes Akhmed Zulfikar, Junior Researcher of the Laboratory “Computational Linguistics”, 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;
ahmedenes@mail.ru, ORCID: <https://orcid.org/0000-0003-3633-4910>