

Алгоритм нейрокогнитивного обучения мультиагентной системы эволюционного моделирования экспрессии генов по данным ПЦР-анализа растений

З. В. Нагоев, М. И. Анчёков, Ж. Х. Курашев, А. А. Хамов

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

Аннотация. Работа нацелена на создание методологии применения систем общего искусственного интеллекта для управления процессом создания новых гибридов растений с заданным набором хозяйственно полезных признаков. Разработаны основные принципы создания имитационных моделей растений на основе мультиагентного моделирования на базе укрупненных условных агентов-клеток, синтез поведения которых выполняется управляющей нейрокогнитивной архитектурой. Разработаны основные принципы создания системы автоматического сбора данных для эволюционного машинного обучения интеллектуальных экспертных систем селекции и семеноводства на основе роботизированного цифрового фенотипирования и генетических данных. Разработан алгоритм обучения децентрализованной системы управления ростом и развитием имитационных моделей растений на основе идентификации фенотипических характеристик процессов роста и развития, детерминированных экспрессией генов растения.

Ключевые слова: общий искусственный интеллект, мультиагентные системы, нейрокогнитивные архитектуры, селекция растений, экспрессия генов, машинное обучение, цифровое фенотипирование

Поступила 04.12.2023, одобрена после рецензирования 08.12.2023, принята к публикации 10.12.2023

Для цитирования. Нагоев З. В., Анчёков М. И., Курашев Ж. Х., Хамов А. А. Алгоритм нейрокогнитивного обучения мультиагентной системы эволюционного моделирования экспрессии генов по данным ПЦР-анализа растений // Известия Кабардино-Балкарского научного центра РАН. 2023. № 6(116). С. 179–192. DOI: 10.35330/1991-6639-2023-6-116-179-192

MSC: 68T42

Original article

Neurocognitive learning algorithm for a multi-agent system for evolutionary modeling of gene expression according to PCR analysis of plants

Z.V. Nagoev, M.I. Anchekov, Zh.Kh. Kurashev, A.A. Khamov

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

Abstract. The work is aimed at creating a methodology for using general artificial intelligence systems to manage the process of creating new plant hybrids with a given set of economically useful traits. The basic principles for creating plant simulation models based on multi-agent modeling based on enlarged conditional cell agents, the synthesis of whose behavior is carried out by the controlling neurocognitive architecture, have been developed. The basic principles for creating an automatic data collection system

for evolutionary machine learning of intelligent expert systems for breeding and seed production based on robotic digital phenotyping and genetic data have been developed. An algorithm has been developed for training a decentralized system for controlling the growth and development of plant simulation models based on the identification of phenogenotypic characteristics of growth and development processes determined by the expression of plant genes.

Keywords: artificial general intelligence, multi-agent systems, neurocognitive architectures, plant breeding, gene expression, machine learning, digital phenotyping

Submitted 04.12.2023,

approved after reviewing 08.12.2023,

accepted for publication 10.12.2023

For citation. Nagoev Z.V., Anchekov M.I., Kurashev Zh.Kh., Khamov A.A. Neurocognitive learning algorithm for a multi-agent system for evolutionary modeling of gene expression according to PCR analysis of plants. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2023. No. 6(116). Pp. 179–192. DOI: 10.35330/1991-6639-2023-6-116-179-192

ВВЕДЕНИЕ

Задача создания интеллектуальных экспертных информационных систем селекции и семеноводства, основанных на моделях биологических процессов, актуализируется в связи с необходимостью повышения эффективности соответствующего производственного цикла с целью создания новых сортов полезных растений. Основу такого роста в данном случае может составить существенная экономия времени и ресурсов на формирование и реализацию плана гибридизации, составляемого с учетом результатов работы интеллектуальных моделей прогнозирования состава и свойств хозяйственно полезных признаков новых гибридов.

В [1] была предложена концепция коллаборативной селекционной системы на основе консорциума интеллектуальных агентов, реализующих управление селекционными, семеноводческими, технологическими и логистическими процессами в интересах комплексной автоматизации.

В [2] рассмотрены возможности применения для создания таких интеллектуальных агентов подхода на основе использования эволюционного моделирования для организации процесса многопоколенной оптимизации управляющих мультиагентных нейрокогнитивных архитектур.

Данное исследование представляет собой дальнейшее развитие идей, сформулированных в процитированных выше работах, в части, касающейся непосредственно интеллектуализации самого процесса вегетации растения. Его *целью является разработка методологии применения систем общего искусственного интеллекта для управления процессом создания новых гибридов растений с заданным набором хозяйственно полезных признаков.*

Как известно, рост и формирование растения происходят на основе фундаментальных процессов образования, развития и дифференциации новых клеток в его составе. Эти процессы в основе своей детерминированы экспрессией различных генов из состава генотипа этого растения. На уровне абстрактного рассмотрения этот процесс подобен выполнению мультиагентного алгоритма, в котором агенты-клетки выполняют программы, закодированные в их геномах, взаимодействуя друг с другом и с внешней окружающей средой.

В идеале задача селекционера состоит в том, чтобы научиться управлять процессами экспрессии таких генов в интересах создания условий для формирования целевых показателей параметров, определяющих ценность хозяйственно полезных признаков растения, как правило, являющихся предельно достижимыми или ранее к моменту планирования селекционного цикла никогда не достигавшимися. Сегодня возможности такого управле-

ния крайне ограничены. В основном это управление структурой генотипа, выполняемое в рамках традиционного процесса гибридизации [3, 4, 5, 6].

Активно развиваются генно-инженерные подходы, позволяющие интенсифицировать процесс гибридизации за счет геномного редактирования [7].

В глобальном методологическом аспекте можно выделить две ключевые проблемы эффективного управления генетическими процессами с целью достижения высоких фенотипических показателей – проблема понимания и проблема реализации.

Первая проблема состоит в том, что процессы экспрессии генов крайне сложны, разнообразны, многочисленны, синхронны, децентрализованы и взаимозависимы, что препятствует построению их моделей и затрудняет понимание характера, определение ключевых параметров, идентификаторов и переменных управления.

Вторая проблема вытекает из первой и состоит в том, что управление этими процессами даже при достижении высокого уровня их понимания требует разработки методик влияния на генетические процессы, параметры управления которых актуализируются на молекулярном уровне.

Данная работа направлена на решение первой из обозначенных проблем. Средством, которое мы избрали для продвижения на пути к пониманию процессов экспрессии генов, является компьютерное моделирование. В перспективе мы планируем разработать полнофункциональную имитационную модель генотипа кукурузы, состоящую из имитационных моделей геномов клеток различных специализаций в составе полнофункциональной мультиклеточной модели растения. В этой модели будут реализованы алгоритмы деления, специализации, роста и развития клеток на основе интерпретации генетической программы, закодированной в модели ДНК. Мы надеемся разработать мультиагентные алгоритмы взаимодействия имитационных моделей клеток в составе мультиклеточного организма, согласованные в рамках представления макроскопических онтофилогенетических процессов. Такие алгоритмы будут основаны на модели поведения живой клетки, построенной на основе концепции синтеза поведения под управлением *мультиагентной нейрокогнитивной архитектуры* [8].

Имитационная модель процесса гибридизации будет построена на основе эволюционного алгоритма над хромосомными наборами, кодирующими структуру и состав таких управляющих мультиагентных нейрокогнитивных архитектур для всех агентов в составе компьютерной симуляции растения.

Основной задачей данной работы является разработка алгоритма нейрокогнитивного обучения мультиагентной системы эволюционного моделирования экспрессии генов по данным лабораторных исследований.

1. ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ЭВОЛЮЦИОННОГО МОДЕЛИРОВАНИЯ ЭКСПРЕССИИ ГЕНОВ РАСТЕНИЙ

В данной работе в целях создания и обучения вышеописанной системы имитационного моделирования мы ориентируемся на метод контроля экспрессии генов, основанный на анализе клеточного материала частей растения, подверженных интенсивному росту. Исходя из того, что такой рост является генно-детерминированным процессом, его проявление служит точным маркером наличия подлежащих процессов экспрессии комплексов генов, содержащихся в геномах клеток именно тех тканей организма растения, которые и претерпевают такой интенсивный рост.

В [9] описаны стратегические перспективы применения роботов для выполнения некоторых ключевых операций селекционно-семеноводческого цикла. В перспективе операцию забора биологического материала в различных частях растения, в которых наблюдается интенсивный рост тканей, смогут выполнять те же роботы, которые осуществляют ежедневное цифровое фенотипирование растений.

Интеллектуальная система имитационного моделирования процессов роста и развития растений должна быть построена как открытая система. Необходимо обеспечить входные потоки данных из всех доступных источников данных. В основу способности концептуализируемой имитационной модели аппроксимировать процессы генно-детерминированного роста и развития растения будут положены такие свойства моделирующей мультиагентной нейрокогнитивной архитектуры, как децентрализованность и распределенность, коммуникативная связность, гетеромасштабность, обучаемость на всех уровнях идентификации исследуемых процессов, архитектурное соответствие им и другие.

Такая имитационная модель будет включать в себя от нескольких миллионов до нескольких миллиардов программных агентов, что вполне соотносится как с возможностями современных вычислительных систем, так и с открытым характером системы, основанной на методах аппроксимации законов взаимодействия структурно-функциональных элементов моделируемых объектов, построенных на применении самоорганизующихся мультиагентных рекурсивных нейрокогнитивных архитектур.

Свойство гетеромасштабности проявляется в том, что программные агенты могут моделировать поведение «естественных» «природных» агентов, существующих и функционирующих в составе объекта моделирования (например, растения) на уровнях различных масштабов – от наноскопического до макроскопического – и при этом взаимодействовать не только с программными агентами, равными по масштабу, но и с агентами, действующими на уровнях как меньшего, так и большего масштабов. Например, нуклеотиды, кодоны, управляющие последовательности, открытые рамки считывания, гены, ДНК, хромосомы, аминокислоты, белки, части растений, сами растения, объекты окружающей растению среды и т.д.

При этом основным инструментом абстрагирования, позволяющим выделить ключевые для целей моделирования свойства частей объекта на любом уровне масштаба, является возможность построения агента как субъекта синтеза своего поведения, выполняемого под управлением локальной автономной системы управления. Такая система управления в общем случае представляет собой мультиагентную нейрокогнитивную архитектуру, реализующую процедуры синтеза поведения программного агента с использованием мультиагентных баз знаний, содержащихся во внутренних агентах нейрокогнитивной архитектуры (актерах) в узлах различной функциональной специализации в составе этой когнитивной архитектуры.

Как уже указывалось, и сегодня многое о поведении таких «природных» агентов, как макромолекулы, клетки и клеточные комплексы, ткани и органы, отдельные организмы, еще не известно науке и является предметом научного исследования. Поэтому такая гетеромасштабная мультиагентная модель должна по определению носить фрагментарный, открытый характер и достраиваться на основе обучения функций систем управления отдельных агентов путем анализа прецедентов и модификации мультиагентных знаний на их основе.

Предлагается строить имитационную мультиагентную модель растения как систему взаимодействующих между собой имитационных моделей отдельных программных агентов – укрупненных (условных) клеток различной функциональной специализации. Каждая

такая клетка в свою очередь моделируется мультиагентной системой искусственной жизни (Artificial Life), представляющей собой рационального программного агента, максимизирующего многокритериальную целевую функцию.

С учетом оценки количества клеток в сформировавшемся растении кукурузы порядка 10^{11} теоретически можно с применением современных параллельных вычислительных систем построить имитационную модель такого растения, которой вычислительный агент будут соответствовать каждой клетке растения. Однако на практике, конечно, в этом нет необходимости, так как, во-первых, клетки в различных структурно и функционально значимых частях растений достаточно однородны, а, во-вторых, для оценки исследуемых фенотипических закономерностей совершенно не требуется такая точность, так как все оценки хозяйственно полезных признаков связаны с макроскопическими параметрами.

Поэтому под агентом-клеткой мы понимаем имитационную модель множества клеток (кластер клеток) растения, демонстрирующих структурную и функциональную однородность в некотором заданном диапазоне существенных параметров. Соответственно, такой условной «клетке» мы приписываем мезомасштабные свойства, количественная оценка значений которых коррелирует с оценкой совокупности клеток реального растения, относимых к данному кластеру на основе классифицирующего признака структурно-функциональной однородности. К таким свойствам можно отнести, например, пространственные размеры, массу, соотношение содержания тех или иных веществ, скорость роста и т.п. мезомасштабные характеристики.

В то же время такая мезомасштабная структурно-функциональная единица – условная «клетка» – может стать и основой для моделирования генетических процессов, так как эти процессы даже в большей степени, чем макроскопические, характеризуются структурной и функциональной однородностью.

Таким образом, мезомасштабный агент-клетка рассматривается нами как основной элемент имитационной фенотипической модели растения. Выбор конкретных значений масштаба зависит от целей моделирования, доступных вычислительных ресурсов, применяемых конкретных имитационных алгоритмов и вычислительных моделей и, в общем случае, должен быть одним из результатов машинного обучения по итогам выполнения серий вычислительных экспериментов на конкретных имитационных моделях растений.

В перспективе по мере накопления знаний в системе, роста вычислительной производительности, развития методов имитационного моделирования, идентификации и машинного обучения масштаб кластера клеток может стремиться к реальному.

2. ОСНОВНЫЕ ПРИНЦИПЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ АГЕНТОВ-КЛЕТОК РАСТЕНИЙ

Агент-клетка, будучи системой искусственной жизни, синтезирует свое поведение в дискретном времени. В данном случае под поведением мы понимаем смену состояний системы «агент – среда», в которой определен такой агент-клетка. Синтез поведения агента направлен на максимизацию целевой функции путем разрешения т.н. *проблемных ситуаций* – последовательностей событий, идентифицируемых управляющей нейрокогнитивной архитектурой агента-клетки, в составе которой есть события, ведущие к существенным приобретениям или к уменьшению значений целевой функции.

В результате реализации процесса идентификации и решения проблем агент-клетка переходит через определенную последовательность состояний системы «агент – среда». В

конечном счете агент-клетка, набрав определенное количество энергии, делится на двух новых агентов-клеток, прекращая тем самым свое существование как таковое, но опосредованно продолжая его в этих двух агентах-клетках.

Примем, что для обозначения отправки порции энергии (выражает значения целевой функции агента-клетки) Δe будем использовать оператор $' \rightarrow '$. Например, факт отправки интеллектуальным агентом \aleph_i^{γ} (многоклеточное растение) порции энергии Δe агенту-клетке $\mathfrak{I}_k^{ij(\gamma-1)}$ запишем:

$$\aleph_i^{\gamma} \rightarrow \mathfrak{I}_k^{ij(\gamma-1)} : \Delta e.$$

Агент-клетка $\mathfrak{I}_{kT_j}^{ij(\gamma-1)}$ типа T_j , расположенный в k -й позиции j -го функционального узла интеллектуального агента, представляет собой оптимизационную машину, максимизирующую свою локальную *целевую функцию* энергии:

$$\begin{aligned} Z_{\tau_b}^{\tau_c} \left(\mathfrak{I}_{kT_j}^{ij(\gamma-1)} \right) &= \Delta v_{kT_j}^{ij} - \Delta v_{KL}^{ij} * \Delta \tau_b^c - N \left(\Delta_{h\tau_b}^{ijk\tau_c} \right) - \Delta N_{ijk}^{ixy} + \\ &+ \Delta N_{ixy}^{ijk} + \Delta N^r \left(\Delta_{h\tau_b}^{ijk\tau_c} \right), \end{aligned}$$

где $\Delta_{h\tau_b}^{ijk\tau_c} = \left\{ \delta_{h\tau_u}^{ijk\tau_q}, b \leq u \leq q \leq c \right\}$ – множество h -х действий (операций, функций, планов) $\delta_{h\tau_u}^{ijk\tau_q}$, которые клетка $\mathfrak{I}_{kT_j}^{ij(\gamma-1)}$ выполняет на отрезке времени (на протяжении своей «жизни») $[\tau_b \dots \tau_c]$, а $N \left(\Delta_{h\tau_b}^{ijk\tau_c} \right)$ – количество энергии, которое он соответственно на это затрачивает:

$$N \left(\Delta_{h\tau_b}^{ijk\tau_c} \right) = \sum_{h=1}^{\left| \Delta_{h\tau_b}^{ijk\tau_c} \right|} \Delta v_h^{ijk}, \mathfrak{I}_k^{ij(\gamma-1)} \rightarrow W : \Delta v_h^{ijk},$$

рассеивая его в среду W и выполняя, таким образом, некоторое поведение, состоящее из действий, входящих в $\Delta_{k\tau_b}^{ijk\tau_c}$;

$$\Delta N_{ijk}^{ixy} = \sum_{\tau=\tau_b}^{\tau_c} \sum_{x=1}^{n \left(\aleph_i^{\gamma}, \mathfrak{I}_{yT_x}^{ix(\gamma-1)} \right)} \Delta v_{ijk}^{ixy}$$

– общее количество энергии, которое агент-клетка $\mathfrak{I}_{kT_j}^{ij(\gamma-1)}$ отправил в качестве вознаграждения другим агнейронам $\mathfrak{I}_{yT_x}^{ix(\gamma-1)} \in \aleph_i^{\gamma}$ в обмен на предоставляемую ими информацию или на обязательства, в соответствии с которыми эти агнейроны должны выполнить некоторые действия, $n \left(\aleph_i^{\gamma}, \mathfrak{I}_{yT_x}^{ix(\gamma-1)} \right)$ – функция, возвращающая количество таких агнейронов;

$$\Delta N_{ixy}^{ijk} = \sum_{\tau=\tau_b}^{\tau_c} \sum_{x=1}^{n \left(\aleph_i^{\gamma}, \mathfrak{I}_{yT_x}^{ix(\gamma-1)} \right)} \Delta v_{ixy}^{ijk}$$

– общее количество ноотропа, которое агент-клетка $\mathfrak{I}_{kT_j}^{ij(\gamma-1)}$ получил в качестве вознаграждения от других агнейронов $\mathfrak{I}_{yT_x}^{ix(\gamma-1)} \in \aleph_i^{\gamma}$ в обмен на предоставляемую им информа-

цию, или на обязательства, в соответствии с которыми этот агент-клетка должен выполнить некоторые действия;

$$\Delta N^r(\Delta_{h\tau_b}^{ijk\tau_c}) = \sum_{h=1}^{|\Delta_{h\tau_b}^{ijk\tau_c}|} \Delta v_h^{ijk\tau_c}, \mathfrak{J}_k^{ij(\tau-1)} \rightarrow W: \Delta v_h^{ijk\tau_c}$$

– общее количество ноотропа, которое агент-клетка получает в результате выполнения некоторого поведения, состоящего из действий, входящих в $\Delta_{h\tau_b}^{ijk\tau_c}$, от интеллектуального агента \mathfrak{N}_i^{τ} за выполнение действий $\Delta_{h\tau_b}^{ijk\tau_c}$. Предполагается, что действия агнейрона $\mathfrak{J}_{kT_j}^{ij(\tau-1)}$ выгодны интеллектуальному агенту \mathfrak{N}_i^{τ} , направлены на достижение некоторых общесистемных результатов.

Таким образом, агент-клетка расходует энергию на поддержание своей жизни, выполнение действий (дополнительных), отправку энергии другим агентам-клеткам в оплату услуг по контрактам, а также в результате перехода в определенные (негативные) состояния, в связи выполнением определенных действий. В свою очередь он получает энергию от других агентов-клеток в оплату услуг, которые в соответствии с контрактами он оказывает им, а также от самого интеллектуального агента в связи с переходами в некоторые состояния в результате выполнения агентом-клеткой определенного поведения.

Агент-клетка так же, как и интеллектуальный агент, является проактивным агентом, синтезирующим свое поведение в системе «агент – среда», только в качестве «агента» выступает он сам, а в качестве внешней по отношению к нему среды рассматривается не окружающий мир W , а интеллектуальный агент \mathfrak{N}_i^{τ} , т.е. все растение. Поэтому агент-клетка в свою очередь состоит из внутренних структурных агентов – *акторов*, обеспечивающих синтез необходимых для построения своего локального графа решений состояний и действий.

Так как мультиагентная нейрокогнитивная архитектура рекурсивная, а основная функция агента-клетки – обнаружение и решение проблем в системе «агент – среда» на основе синтеза своего поведения – аналогична основной функции интеллектуального агента, в качестве организационного паттерна когнитивной архитектуры агента-клетки также избран инвариант мультиагентной когнитивной архитектуры. На рисунке 1 приведена когнитивная архитектура агента-клетки, включающая в себя акторов (изображены объемными геометрическими фигурами в квадратных рамках) в составе инварианта его когнитивной архитектуры.

На рисунке 1 видно, что подобно агентам-клеткам в составе когнитивной архитектуры интеллектуального агента акторы в составе когнитивной архитектуры агента-клетки располагаются на определенных уровнях, соответствующих когнитивным узлам инварианта мультиагентной нейрокогнитивной архитектуры. Эти уровни для когнитивной архитектуры агента-клетки мы называем *акторкогнитонами*. Актор $\alpha_{hT_l}^{ijkl}$ типа T_l , располагается в l -м акторкогнитоне когнитивной архитектуры агента-клетки $\mathfrak{J}_{kT_j}^{ij(\tau-1)}$. Индекс h обозначает номер актора.

Таким образом, агент-клетка $\mathfrak{J}_{kT_j}^{ij(\tau-1)}$ содержит в себе акторкогнитоны $K_{IT_l}^{ijk}$, которые также подразделяются на типы T_l , но в отличие от функциональных узлов интеллектуального агента $\mathfrak{J}_{jT_j}^i$ в их состав входят не агнейроны $\mathfrak{J}_{kT_j}^{ij(\tau-1)}$, а акторы $\alpha_{hT_l}^{ijkl}$.

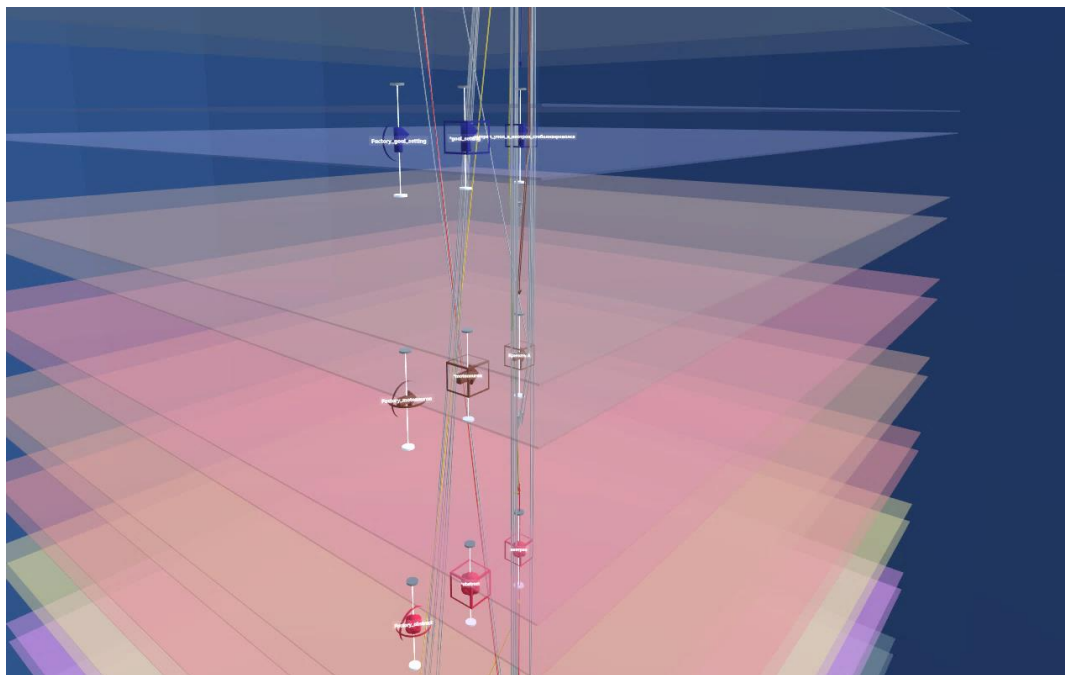


Рис. 1. Изображение когнитивной архитектуры агента-клетки в разрабатываемой программе автоматизированного проектирования

Fig. 1. Image of the cognitive architecture of the cell agent in the developed computer-aided design program

Основное различие между агентами-клетками и актерами состоит в том, что актер является минимальным агентом, на уровне которого рекурсия мультиагентной когнитивной архитектуры заканчивается.

Таким образом, агент-клетка синтезирует свое поведение, решая оптимизационную задачу максимизации вышеуказанной целевой функции.

Источником пополнения энергии служит внешняя по отношению к этому агенту-клетке среда – имитационная модель всего растения.

В агенте-клетке можно выделить две основные функциональные подсистемы – тело и систему управления. В свою очередь, как было показано в [3], система управления состоит из мультиагентной когнитивной архитектуры и *генома*.

Все внутренние функциональные элементы агента-клетки также имеют возможность взаимодействовать друг с другом, однако они сами по себе уже не являются агентами искусственной жизни. Условно можно считать, что они представляют собой функциональные аналоги комплексов некоторых органелл в составе клетки.

Таким образом, в нашей модели агент-клетка содержит в себе мультиагентную систему имитационных моделей некоторых условных органелл живой клетки – т.н. акторов, которые также представляют собой программных агентов с более примитивным внутренним устройством и более простыми целевыми функциями. К основным элементам тела агента-клетки относятся: распределенное устройство управления – децентрализованная мультиагентная нейрокогнитивная архитектура (включающая в себя сенсоры и эффекторы), полезные «агреганты» – имитационные модели макромолекулярных комплексов, представляющих на внутриклеточном уровне макроскопические параметры растений, характеризующие их хозяйственно-полезные признаки (хлорофилл, крахмал, клетчатка, вода и т.п.), а также геном агента-клетки.

Управляющая мультиагентная нейрокогнитивная архитектура синтезирует поведение агента-клетки на основе алгоритмов, описанных в [8]. Имитационная модель генома агента-клетки представляет собой структурированную мультиагентную систему, описанную в [3]. В общем случае каждый такой агент также является рациональным и имеет возможность взаимодействовать с другими агентами-клетками с помощью обмена сообщениями.

3. МУЛЬТИАГЕНТНЫЙ АЛГОРИТМ ОБУЧЕНИЯ НЕЙРОКОГНИТИВНЫХ СИСТЕМ ИНТЕЛЛЕКТУАЛЬНОГО УПРАВЛЕНИЯ ПО ФИЛОГЕНЕТИЧЕСКИМ ДАННЫМ

Связь генетической программы, заложенной в геноме клетки, и интерпретатора этой программы реализуется на основе алгоритмов машинного обучения по данным измерений наблюдаемых онтофилогенетических процессов.

Для идентификации этих процессов с целью извлечения и накопления данных для обучения интеллектуальной нейрокогнитивной архитектуры планируется использовать непрерывное цифровое фенотипирование и доступную генетическую информацию, в том числе результаты генетических анализов.

Кроме того, в перспективе планируется также применить для развития и обучения мультиагентных онтофилогенетических моделей методологию приобретения и интерпретации знаний, полученных интеллектуальным программным агентом, обеспечивающим управление имитационными экспериментами, с использованием системы понимания высказываний на естественном языке. Применение этого механизма позволит, в частности, организовать высокопроизводительный процесс накопления знаний, выраженных в абстрактной знаковой форме посредством таких мощных инструментальных средств, как диалоговые и читающие системы [10].

Цель обучения системы состоит в развитии мультиагентной базы знаний о поведении отдельных агентов-клеток в интересах достижения продуктивной полноты знаний, достаточной для того, чтобы выполнять имитационное моделирование процессов экспрессии генов с точностью, необходимой для применения селекционером в процессе разработки нового гибрида растения.

Интеллектуальная система будет строить мультиагентные модели роста и развития растений по данным геномов, аппроксимируя соответствующие функции на основе измерений. Для обучения управляющих мультиагентных нейрокогнитивных архитектур будет применяться обучающая выборка, состоящая из результатов цифрового фенотипирования и данных генетических анализов (в том числе ПЦР-анализа). Цикл обучения будет выполняться в непрерывном режиме. Укрупненная блок-схема алгоритма обучения приведена на рисунке 2.

Неотъемлемой частью алгоритма обучения является идентификация режимов развития растений с помощью роботизированной системы цифрового фенотипирования растений. Автономный мобильный робот проводит ежедневное сканирование посадок с помощью системы распознавания и идентификации изменений в фенотипе растений. Далее интеллектуальная экспертная система определяет изменения и их интенсивность. По команде от нее робот выполняет забор биологического материала в тех частях растений, которые продемонстрировали активный рост. Транспортный робот доставляет материалы в лабораторию.

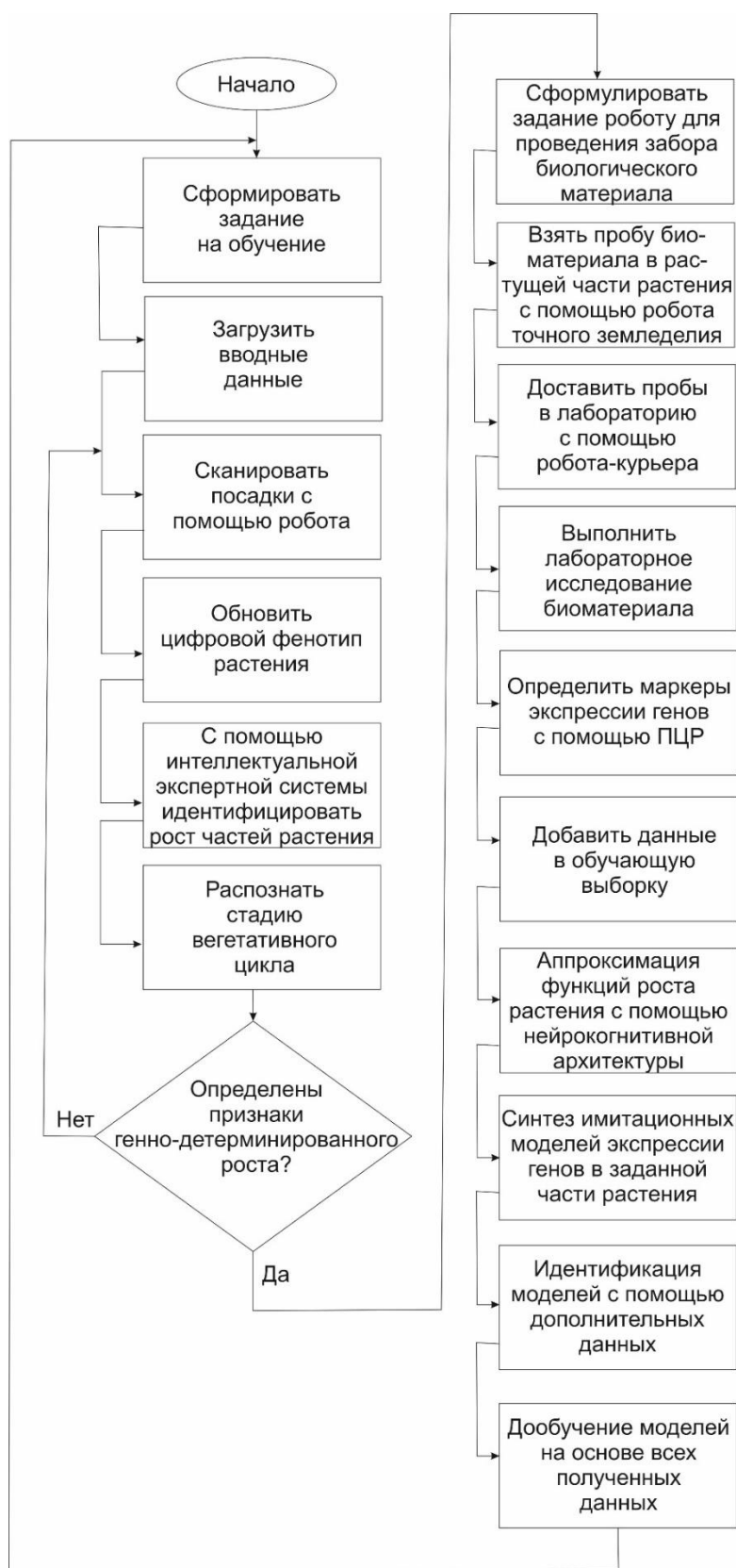


Рис. 2. Алгоритм непрерывного обучения имитационной мультиагентной нейрокогнитивной модели растения

Fig. 2. Algorithm for continuous training of simulation multi-agent neurocognitive plant model

Далее выполняется генетический анализ с применением заданного набора праймеров – определяются локусы, сайты, аллели, ответственные за активизацию роста, проводится анализ белков и аминокислот. Интеллектуальная экспертная система идентифицирует процессы экспрессии генов по выявленным молекулярным маркерам.

Полученные данные о зафиксированных фенотипических изменениях и соответствующих процессах экспрессии генов добавляются в обучающую выборку – мультиагентные нейрокогнитивные системы управления агентов-клеток выполняют процедуру аппроксимации функций роста анализируемой части растений. На основе полученных функций строятся и уточняются имитационные модели экспрессии генов и генно-детерминированного роста заданной части растения. Выполняется идентификация полученной модели на доступных дополнительных экспериментальных данных и переобучение системы – модель достраивается по принципу открытой системы, непрерывно повышая свою прогностическую силу.

Интеллектуальные системы управления агентов-клеток стремятся достичь таких состояний системы «агент – среда», в которых наблюдается минимальное рассогласование между фактическими и ожидаемыми значениями комплекса целевых параметров, характеризующих на клеточном уровне макроскопические свойства, описывающие хозяйственно полезные признаки растения. Соответственно, значения их целевых функций увеличиваются при переходе такой системы в состояние, в котором это рассогласование минимально, и, наоборот, уменьшаются при переходе в состояние, в котором рассогласование увеличивается.

Множество действий, которые такие агенты-клетки могут выполнить для того, чтобы перейти в лучшие состояния, характеризующиеся меньшими значениями рассогласования и, соответственно, большими значениями вознаграждения, включает в себя инкремент и декремент значений соответствующих параметров и возможность варьирования шага таких изменений. Фактически такой агент будет решать многокритериальную оптимизационную задачу, направленную на минимизацию комплексной ошибки на основе аппроксимации функционала, состоящего из функций, описывающих изменение значений целевых параметров в зависимости от времени, с помощью обучения на наборах экспериментальных данных.

Так как процессы роста и развития растения хорошо наблюдаемы по достаточно широкому спектру параметров и масштабов (геномный, внутриклеточный, клеточный, макроскопический уровень), формирование обучающих выборок для решения задач такой аппроксимации представляется сложной технической, но принципиально понятной и реализуемой задачей. Для ее решения необходимо задействовать все доступные организационно-технические средства, включая визуальное наблюдение и измерение макропараметров, применение робототехнических комплексов цифрового фенотипирования [9], системы распознавания образов, интеллектуальные экспертные системы [3], лабораторные исследования внутриклеточного состава, ПЦР-анализ, полногеномное секвенирование и др.

ЗАКЛЮЧЕНИЕ

Разработана концепция создания имитационных моделей растений на основе мультиагентного моделирования на базе укрупненных условных агентов-клеток, синтез поведения которых выполняется управляющей нейрокогнитивной архитектурой.

Разработаны основные принципы создания системы автоматического сбора данных для эволюционного машинного обучения интеллектуальных экспертных систем селекции и семеноводства на основе роботизированного цифрового фенотипирования и генетических данных.

Разработан алгоритм обучения децентрализованной системы управления ростом и развитием имитационных моделей растений на основе идентификации феногенотипических характеристик процессов роста и развития, детерминированных экспрессией генов растения.

СПИСОК ЛИТЕРАТУРЫ

1. Анчёков М. И., Боготова З. И., Пшенокова И. А. и др. Коллаборативная селекционная система на основе консорциума гетерогенных интеллектуальных агентов // Известия Кабардино-Балкарского научного центра РАН. 2022. № 5(109). С. 25–37. DOI: 10.35330/1991-6639-2022-5-109-25-37.

2. Анчёков М. И., Бжихатлов К. Ч., Нагоев З. В. и др. Онтоэписоцифилогенетическое развитие систем общего искусственного интеллекта на основе мультиагентных нейрокогнитивных архитектур // Известия Кабардино-Балкарского научного центра РАН. 2022. № 6(110). С. 61–75. DOI: 10.35330/1991-6639-2022-6-110-61-75.

3. Анчёков М. И., Апишев А. З., Бжихатлов К. Ч. и др. Формальная модель генома агента общего искусственного интеллекта на основе мультиагентных нейрокогнитивных архитектур // Известия Кабардино-Балкарского научного центра РАН. 2023. № 5(115). С. 11–24.

4. Doursat R. Organically grown architectures: Creating decentralized, autonomous systems by embryomorphic engineering. *Organic Computing*. Ed. R.P. Würtz, Springer-Verlag. 2008. Pp. 167–200.

5. Fedor A., Zachar I., Szilágyi A. et al. Cognitive Architecture with Evolutionary Dynamics Solves Insight Problem // *Front. Psychol.* 2017. Vol. 8.

6. Werfel J., Nagpal R. Extended stigmergy in collective construction // *IEEE Intelligent Systems*. 2006. No. 21(2). Pp. 20–28.

7. Jinek M., Chilynski K., Fonfara I. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity // *Science*. 2012. Vol. 337. No. 6069. Pp. 816–821. DOI: 10.1126/science.1225829.

8. Нагоев З. В. Интеллектика, или Мышление в живых и искусственных системах. Нальчик: Издательство КБНЦ РАН, 2013. 232 с.

9. Анчёков М.И., Бжихатлов К.Ч., Лешкенов А.М. Высокопроизводительные системы фенотипирования сельскохозяйственных культур // Известия Кабардино-Балкарского научного центра РАН. 2022. № 5(109). С. 19–24.

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

REFERENCES

1. Anchekov M.I., Bogotova Z.I., Pshenokova I.A. et al. Collaborative breeding system based on a consortium of heterogeneous intelligent agents. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2022. No. 5(109). Pp. 25–37. DOI: 10.35330/1991-6639-2022-5-109-25-37. (In Russian)

2. Anchekov M.I., Bzhikhatlov K.Ch., Nagoev Z.V. et al. Ontoepisociophylogenetic development of general artificial intelligence systems based on multi-agent neurocognitive architectures. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2022. No. 6(110). Pp. 61–75. DOI: 10.35330/1991-6639-2022-6-110-61-75. (In Russian)
3. Anchekov M.I., Apshev A.Z., Bzhikhatlov K.Ch. et al. Formal genome model of a general artificial intelligence agent based on multi-agent neurocognitive architectures. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2023. No. 5(115). Pp. 11–24. (In Russian)
4. Doursat R. Organically grown architectures: Creating decentralized, autonomous systems by embryomorphic engineering. *Organic Computing*. Ed. R.P. Würtz, Springer-Verlag. 2008. Pp. 167–200.
5. Fedor A., Zachar I., Szilágyi A. et al. Cognitive Architecture with Evolutionary Dynamics Solves Insight Problem. *Front. Psychol.* 2017. Vol. 8.
6. Werfel J., Nagpal R. Extended stigmergy in collective construction. *IEEE Intelligent Systems*. 2006. No. 21(2). Pp. 20–28.
7. Jinek M., Chilynski K., Fonfara I. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012. Vol. 337. No. 6069. Pp. 816–821. DOI: 10.1126/science.1225829.
8. Nagoev Z.V. *Intellektika, ili Myshleniye v zhivyykh i iskusstvennykh sistemakh* [Intelligence, or thinking in living and artificial systems]. Nalchik: Izdatel'stvo KBNTS RAN, 2013. 232 p. (In Russian)
9. Anchekov M.I., Bzhikhatlov K.Ch., Leshkenov A.M. High-throughput crop phenotyping systems. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2022. No. 5(109). Pp. 19–24. (In Russian)
10. 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)

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>

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>

Курашев Жираслан Хаутиевич, зав. лабораторией «Молекулярная селекция и биотехнология», Кабардино-Балкарский научный центр РАН;

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

ORCID: <https://orcid.org/0000-0001-9442-6122>

Хамов Анзор Азаматгериевич, мл. науч. сотр. лаборатории «Молекулярная селекция и биотехнология», Кабардино-Балкарский научный центр РАН;

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

orpitnoe2014@mail.ru, ORCID: <https://orcid.org/0000-0003-3269-4572>

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>

Anchekov Murat Inusovich, Researcher of the Laboratory “Molecular Breeding 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>

Kurashev Zhiraslan Hautievich, Head of the Laboratory “Molecular Breeding and Biotechnology”, Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences;

360000, Russia, Nalchik, 224 Kirova street;

ORCID: <https://orcid.org/0000-0001-9442-6122>

Khamov Anzor Azamatgerievich, Junior Researcher of the Laboratory “Molecular Breeding and Biotechnology”, Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences;

360000, Russia, Nalchik, 224 Kirov street;

opitnoe2014@mail.ru, ORCID: <https://orcid.org/0000-0003-3269-4572>