© 2022 г. С.А. ШУМСКИЙ, канд. физ.-мат. наук
(serge.shumsky@gmail.com)
(Московский физико-технический институт)

ADAM — МОДЕЛЬ ИСКУССТВЕННОЙ ПСИХИКИ[1]

Предложена модель искусственной психики ADAM, реализующая иерархическую архитектуру глубокого обучения с подкреплением. ADAM способен обучаться все более сложным и протяженным во времени поведенческим навыкам по мере увеличения количества управляющих уровней искусственной психики. Целенаправленное поведение формируется иерархической обучающейся системой с постепенным наращиванием числа уровней, где каждый иерархический уровень ответственен за свой временной масштаб поведения.

Ключевые слова: общий искусственный интеллект, глубокое обучение с подкреплением, иерархическая система управления.

DOI: 10.31857/S0005231022060034, **EDN:** ACIMYZ

1. Введение

Под искусственным интеллектом (ИИ) обычно понимают алгоритмы решения различных интеллектуальных (когнитивных) задач на уровне человека или лучше. В разные времена под "интеллектуальными" понимались разные типы задач. В 1950-е гг. таковыми считались "творческие" задачи, в которых невозможно предусмотреть заранее все варианты решений: игра в шахматы, доказательство теорем, машинный перевод. С течением времени область ИИ расширялась и пополнялась другими типами когнитивных задач, уже не связанными с логическим интеллектом, например задачи распознавания образов и моделирования целесообразного поведения животных [1]. Однако все современные системы машинного интеллекта имитируют каждая лишь какую-то одну очень узкую область человеческих способностей, т.е. являются *слабым ИИ*. Задача создания *сильного ИИ*, способного конкурировать с человеком во всех областях, до недавнего времени на практике даже и не ставилась. Считалось, что это проблема очень отдаленного будущего.

Перелом во взглядах ИИ-сообщества на сильный ИИ произошел в последние несколько лет после свершившейся в 2010-х *революции глубокого обучения*. В ходе этой (все еще продолжающейся) революции происходит смена

[1] Работа выполнена при частичной финансовой поддержке Центра компетенций Национальной технологической инициативы по направлению "Искусственный интеллект" при МФТИ.

основной парадигмы ИИ. Мейнстрим ИИ сместился в область обучения искусственных нейросетей, и место ИИ, основанного на человеческих знаниях, занимает теперь ИИ, основанный на машинном обучении, которому удается решать практически все задачи ИИ в единой методологии, причем с гораздо лучшим качеством, чем прежде [2]. Лидеры революции глубокого обучения сегодня предсказывают переход от моделирования систем бессознательного сенсорного интеллекта к по-настоящему разумным машинам, самостоятельно планирующим свое поведение и "понимающим", что и зачем они делают [3]. Появление таких разумных машин создаст новый массовый рынок автономных роботов, способных к обучению, в отличие от современных роботов с программируемым поведением.

Иными словами, главной задачей следующего этапа развития ИИ является синтез всех видов интеллекта — сенсорного, моторного, стратегического и других в единой *искусственной психике*, называемой в англоязычной литературе *общим ИИ — Artificial General Intelligence* (AGI). Именно такую цель — создание искусственной психики роботов, позволяющей им самостоятельно планировать достижение поставленных целей и осуществлять эти планы, адаптируясь к изменяющейся обстановке, — ставит перед собой лаборатория Когнитивных архитектур МФТИ.

2. Когнитивные архитектуры

Искусственная психика представляет собой целостную систему со своей *когнитивной архитектурой*, которая определяет все ее базовые свойства. Поэтому разработка искусственной психики, как и любой сложной системы, должна начинаться именно с проектирования ее архитектуры. Как и архитектура фон-Неймановских компьютеров, когнитивная архитектура подразумевает исполнение самых разных алгоритмов. В традиционных когнитивных архитектурах эти алгоритмы (включающие *знания* об устройстве мира и полезные *навыки* поведения в этом мире) были в основном рукотворными [4]. Мы ставим перед собой задачу, чтобы эти алгоритмы не закладывались извне в готовом виде, а возникали в процессе активного взаимодействия искусственной психики с внешней средой путем обучения с подкреплениями. Такая постановка задачи, по нашему мнению и согласно [5], наиболее близка к определению AGI. Отличительной чертой нашего подхода является попытка воспроизвести в нашей когнитивной архитектуре базовые черты вычислительной архитектуры человеческого мозга.

Несколько огрубляя, базовая схема современных когнитивных архитектур может быть суммирована в так называемой *стандартной модели интеллекта* [6], описывающей схему взаимодействия основных типов когнитивных модулей искусственной психики. Как видно из рис. 1, связующим звеном между всеми когнитивными модулями является оперативная рабочая память, соответствующая текущей активности в коре мозга. Содержимое рабочей памяти контролируется моделями поведения в базальных ганглиях, управляющих текущей активностью коры — операциями, хранящимися в долговременной

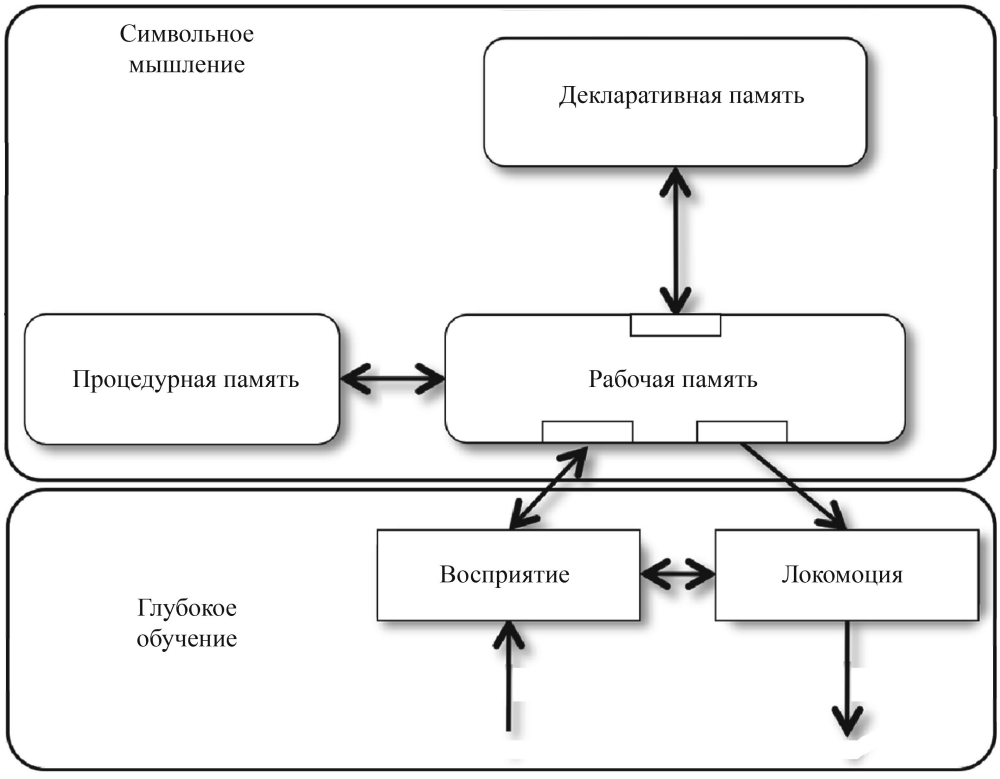

Рис. 1. Стандартная модель интеллекта (Standard Model for the Mind [6]).

процедурной памяти. Элементы долговременной декларативной памяти коры при активации поступают в рабочую память. Все когнитивные архитектуры, объединяемые стандартной моделью, используют знания, представленные в символьной форме (факты и правила). Интерфейс символьной рабочей памяти с векторным физическим пространством обеспечивается специальными кодирующими и декодирующими модулями — соответственно сенсорными и моторными, в качестве которых могут выступать современные глубокие нейросети.

Важнейшим элементом стандартной модели является идея *когнитивного акта*, стандартной операции выбора и исполнения одного из правил процедурной памяти. Любое сколь угодно сложное поведение состоит из таких элементарных когнитивных актов длительностью в десятые доли секунды. Вся сложность нашего мышления и поведения возникает в результате правильно подобранных цепочек элементарных когнитивных актов. Стандартная модель суммирует наши знания о механизмах работы мозга и структуре нашей психики. Но она не дает ответа, каким именно образом выстраиваются неимоверно длинные осмысленные цепочки когнитивных актов, как организовано планирование нашего поведения на больших масштабах времени — от минут до дней, месяцев и даже лет (см. рис. 2).

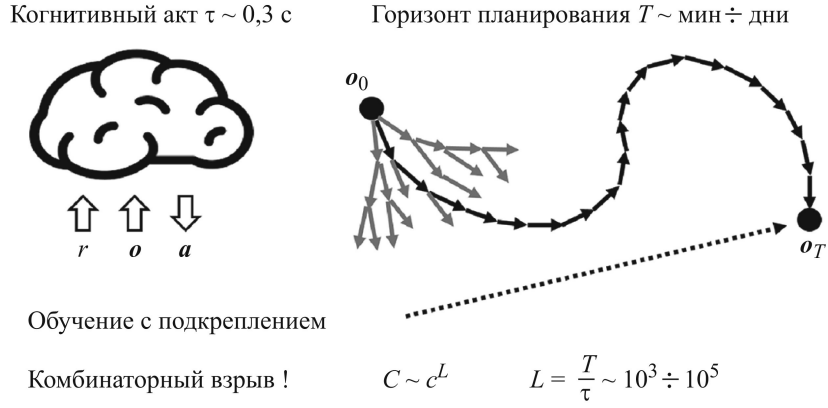

Рис. 2. Основная проблема машинного мышления — переход от единичного когнитивного акта к большим горизонтам планирования. Здесь: r – подкрепления, $\boldsymbol{o}$ – наблюдения, $\boldsymbol{a}$ – действия, C – разнообразие возможных цепочек действий, характеризующее сложность поиска оптимальной стратегии, c – разнообразие действий на каждом шаге, L – количество шагов на горизонт планирования T.

Планы в стандартной модели могут задаваться в виде иерархии правил процедурной памяти, где отдельные действия могут содержать в себе различные этапы. Но эти иерархии правил закладываются в них вручную, а не возникают автоматически, в отличие от иерархии признаков, автоматически возникающих в результате обучения глубоких нейросетей. Как пишет автор классического современного учебника по ИИ Стюарт Рассел: “В настоящее время все существующие методы иерархического планирования опираются на сгенерированные человеком иерархии абстрактных и конкретных действий. Мы еще не понимаем, как такие иерархии могут быть получены путем обучения” [7].

Действительно, хотя в последнее десятилетие ручное программирование правил поведения и уступает место глубокому обучению с подкреплением, обеспечив тем самым прорыв в уровне стратегического игрового интеллекта, иерархическое планирование в глубоком обучении до сих пор отсутствует. Так, успех известной программы AlphaZero обеспечивается потрясающей интуицией ее глубокой нейросети, обученной правильно оценивать любую игровую позицию и находить в ней наилучшие варианты ходов. Однако глубокая нейросеть AlphaZero способна генерировать варианты своих ходов лишь на один шаг вперед. Для выбора наилучшего варианта на каждом шаге AlphaZero производит просчет очень объемного дерева вариантов на десятки ходов вперед [8]. Это обеспечивает отличное качество игры, но очень дорогой ценой из-за комбинаторного взрыва числа возможных комбинаций, перебираемых методом грубой силы.

Человеческое мышление устроено по-другому. Мы не перебираем в уме все возможные варианты цепочек когнитивных актов, что было бы практи-

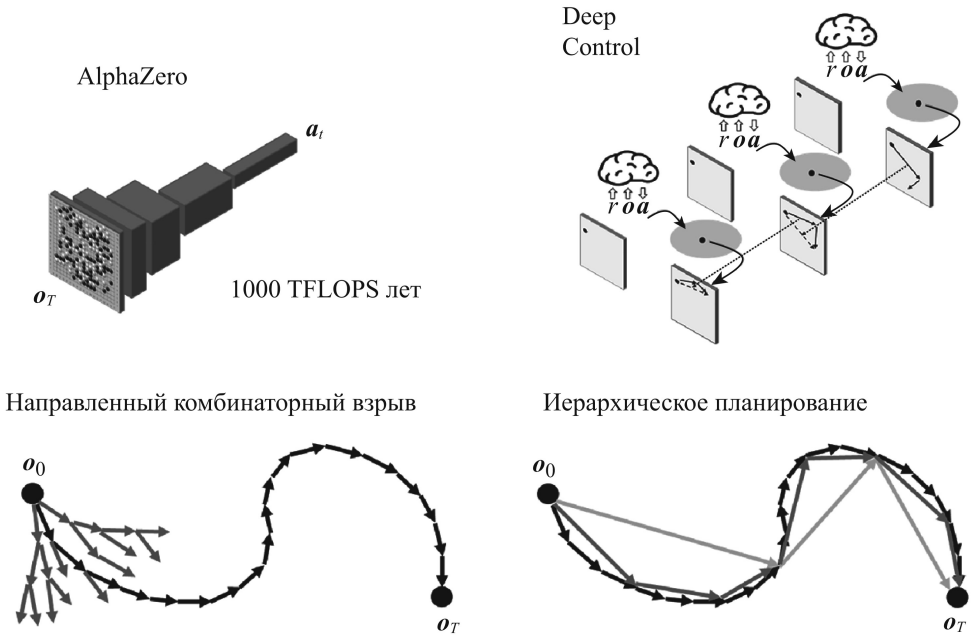

Рис. 3. Пошаговый просчет оптимальной траектории программой AlphaZero (слева) и иерархическое планирование поведения от общего замысла ко все более детальным планам в архитектуре Deep Control (справа).

чески невозможно. Вместо этого мы используем иерархии планов: от крупномасштабного замысла достижения цели — ко все более подробным планам его достижения. При этом разнообразие вариантов выбора на каждом уровне планирования относительно невелико, а детализируются лишь те этапы, которые реализуются в данный момент (см. рис. 3, справа внизу). Именно так планируют свое поведение люди, и именно так устроено планирование в предложенной автором когнитивной архитектуре Deep Control [9], где в процессе обучения автоматически формируется иерархия правил поведения по аналогии с глубоким обучением, автоматически формирующим иерархии признаков при распознавании образов. Тем самым решается проблема формирования разумного поведения с большим горизонтом планирования, т.е. перекидывается мостик от простейшей психики животных к сложно организованному символьному мышлению человека. В отличие от стандартной модели интеллекта в архитектуру Deep Control иерархичность встроена в явном виде, отражая иерархичность, присущую кортико-стриарной системе нашего мозга, управляющей нашим поведением и обучающейся с помощью дофаминовых подкреплений [10] (см. рис. 4).

Другим определяющим принципом архитектуры Deep Control является предиктивное управление, согласно которому наш мозг постоянно предсказывает будущее в контексте своих собственных управляющих воздействий

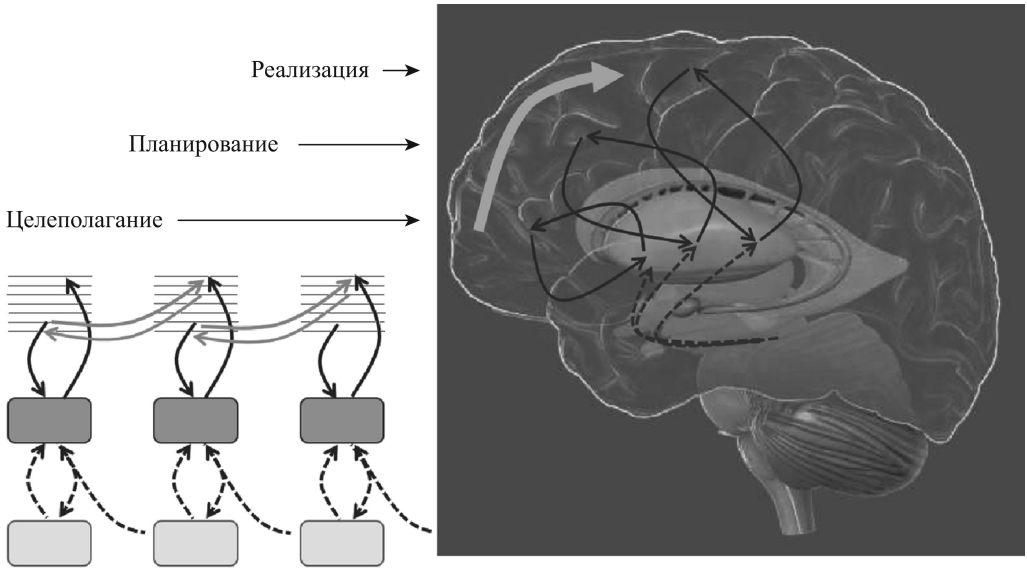

Рис. 4. Кортико-стриарная система мозга, управляющая целесообразным поведением, устроена иерархически. Программа поведения формируется в мозге от мотивации к планированию и далее — к реализации. Все уровни иерархии имеют одинаковый набор модулей, представленных в различных частях коры, базальных ганглий и дофаминовой системы среднего мозга.

(рис. 4, серые стрелки). Эта особенность нашего мозга хорошо изучена и положена в основу многих теоретических моделей мышления [11, 12], в отличие от которых нас интересует действующая модель искусственной психики, являющаяся результатом обратного инжиниринга архитектуры мозга.

3. Архитектура Deep Control: иерархическое планирование поведения

В архитектуре Deep Control проблема, о которой говорит Стюарт Рассел, — обучение роботов иерархическому планированию поведения — решается с использованием оригинальной технологии *глубокого структурного обучения* [9], а именно: управление поведением на разных временных масштабах осуществляется в разных вычислительных слоях. Чем выше слой, тем большим временным масштабом он оперирует, решая, по существу, одну и ту же типовую задачу, как показано на рис. 5. Каждый слой управляет взаимодействием с внешним миром, предсказывая свое очередное дискретное состояние, кодирующее на своем временном масштабе *сенсомоторную* информацию — как входящую (*наблюдения*), так и исходящую (*действия*), т.е. любой план действий сопровождается соответствующими предсказаниями наблюдений, которые постоянно сравниваются с реальностью, поставляя материал для обучения даже в отсутствие подкрепляющих сигналов, что выгодно отличает Deep Control от обычного глубокого обучения с подкреплением. На рис. 5

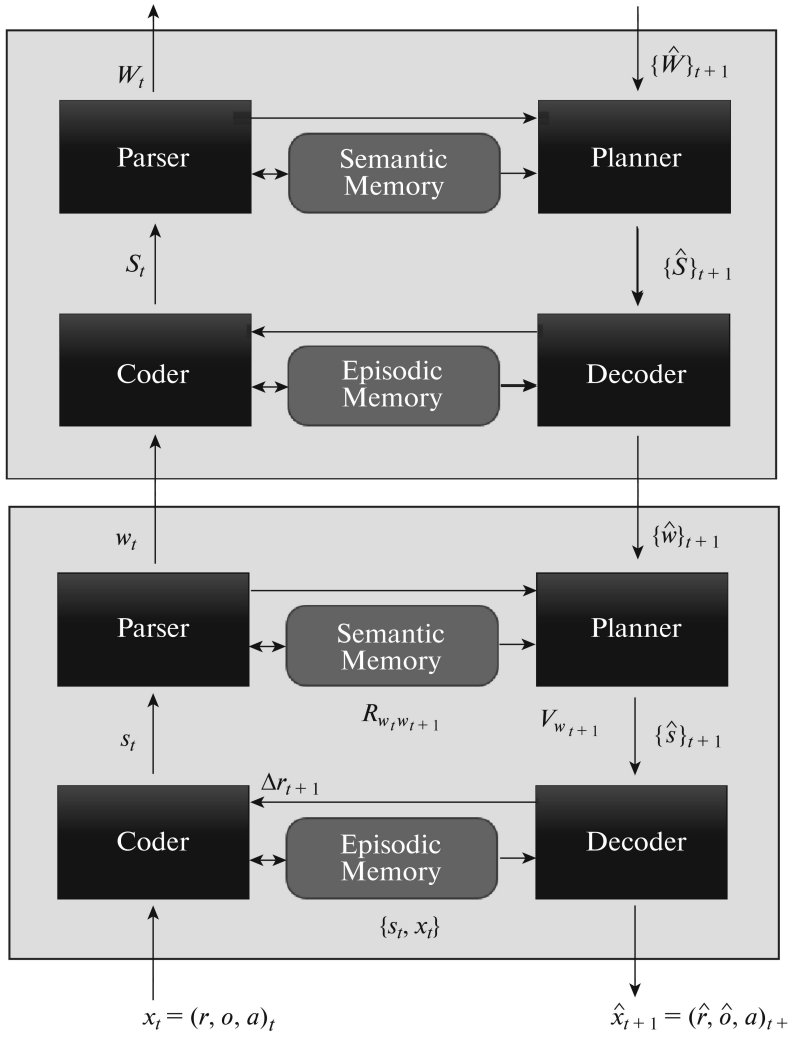

Рис. 5. Архитектура Deep Control представлена однотипными вычислительными слоями, осуществляющими управление поведением каждый на своем временном масштабе. Чем выше слой, тем бóльшим временны́м масштабом он оперирует. Каждый слой находит на своем масштабе решение локальной задачи, поставленной для него более высоким слоем, разбивая ее на подзадачи для более низкого слоя.

показаны два первых слоя Deep Control, на примере которых мы поясним, как именно происходит управление поведением в этой архитектуре.

Входная информация поступает в управляющую систему (искусственный мозг робота) из внешнего мира в виде единого сенсомоторного вектора $\boldsymbol{x}_t = (r, \boldsymbol{o}, \boldsymbol{a})_t$, объединяющего показания всех сенсоров $\boldsymbol{o}$ и актуаторов $\boldsymbol{a}$ управляемой системы (тела робота) в данный момент дискретного времени. Показания одного из сенсоров выделены в отдельный *подкрепляющий сигнал* r, который служит для обучения системы и обрабатывается особым

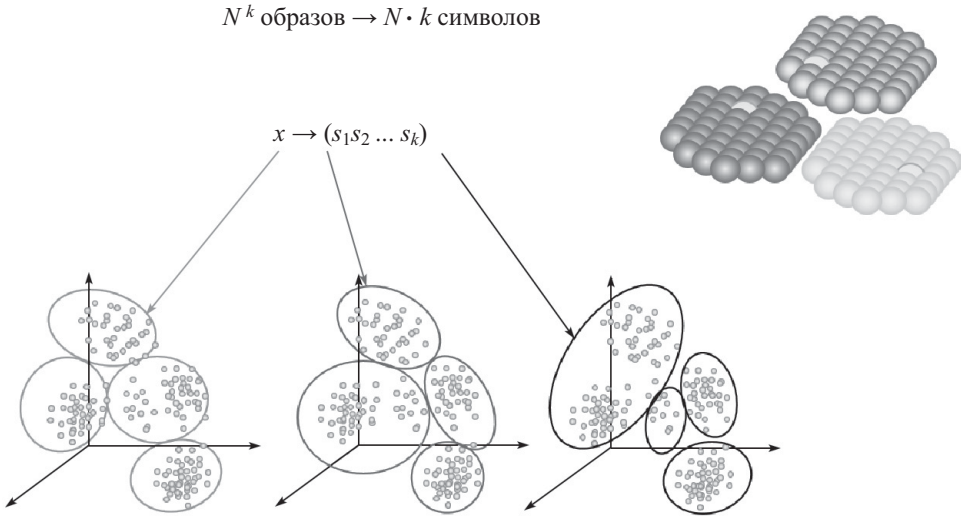

Рис. 6. Дискретное кодирование в Кодере.

образом. В ответ управляющая система выдает прогноз следующего сенсомоторного вектора в очередной момент времени $\hat{\boldsymbol{x}}_{t+1} = (\hat{r}, \hat{\boldsymbol{o}}, \boldsymbol{a})_{t+1}$, а именно: прогноз показаний всех сенсоров, которые она не контролирует, и *реальные* управляющие сигналы для всех актуаторов управляемой системы, которые она контролирует.

Входным элементом каждого слоя является *Кодер*, который кодирует непрерывный векторный сигнал набором дискретных символов $\boldsymbol{x}_t \rightarrow \boldsymbol{s}_t$, т.е. осуществляет дискретное кодирование. В простейшем варианте Кодер состоит из нескольких модулей, каждый из которых производит свой вариант кластеризации входных векторов, сохраненных в *эпизодической памяти*. Как показано на рис. 6, разнообразие дискретных кодов возрастает экспоненциально с числом модулей, так что требуемого для управления разнообразия всегда можно добиться даже с небольшим числом модулей. Так, например, 7 модулей по 30 кластеров в каждом достаточно для кодирования более чем 10^{10} образов (верхняя оценка количества когнитивных актов за всю человеческую жизнь).

В мозге дискретное кодирование производится гиперколонками неокортекса, так как из-за взаимной конкуренции колонок активной в каждой гиперколонке может быть лишь одна колонка. Такие гиперколонки размером порядка 1 мм^2 были экспериментально исследованы Маунткаслом [13], а их теоретическая модель предложена Кохоненом [14] (см. рис. 7).

Дискретный сигнал из Кодера передается в *Парсер*, аналог рабочей памяти в стандартной модели. Парсером обычно называют синтаксический анализатор, структурирующий входные данные. В нашем случае Парсер производит анализ временно́го ряда из поступающих к нему символов, выделяя в нем характерные цепочки символов, *морфемы* $\mathbf{w}_t$. Таким образом, Парсер фор-

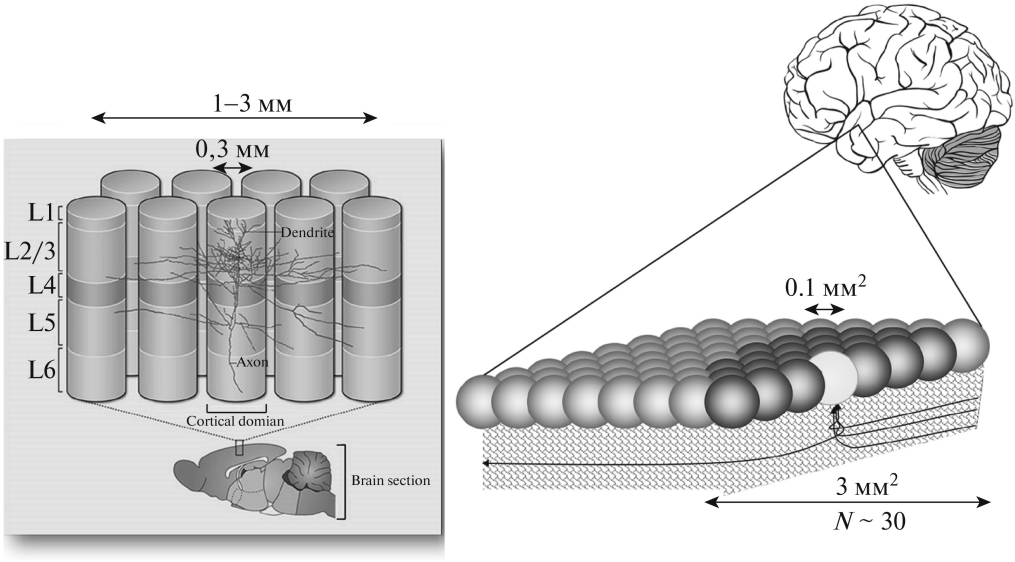

Рис. 7. Дискретное кодирование в неокортексе.

мирует укрупненное описание текущего контекста уже в виде цепочек морфем, а не символов. По мере накопления опыта Парсер обучается выявлять все более крупные морфемы, рекурсивно объединяя между собой наиболее часто встречающиеся пары более коротких морфем, начиная с единичных символов. Выявление и запоминание морфем в мозге может осуществляться гипотетическими *рекурсивными модулями* коры, отличающимися тем, что они "смотрят" сами на себя и поэтому способны кодировать временны́е последовательности (рис. 8).

Силы ассоциативных связей между кодами морфем, отражающие то, как часто морфема $\mathbf{w}'$ следует за морфемой $\mathbf{w}$, образуют *Семантическую память* $R_{\mathbf{ww}'}$, названную так потому, что она отражает характер употребления морфем, от которого только и зависят значения действий. Действия, осуществляемые в сходных ситуациях, т.е. перед и после определенных действий, имеют, очевидно, сходные назначения аналогично тому, как значения слов в языке определяются контекстами их употребления. Семантическая память $R_{\mathbf{ww}'}$ помнит, какие следующие морфемы (т.е. цепочки сенсомоторных состояний) и насколько часто встречались в данном контексте. Как и Кодер, Семантическая память разбита на независимые модули — *головы*, работающие каждая со своим входным алфавитом символов, поступающих от соответствующих модулей Кодера. Модульный дизайн матрицы $R_{\mathbf{ww}'}$ существенно снижает сложность и время вычислений. Это аналог нашей модели мира, хранящейся в неокортексе, предположительно — в рекурсивных модулях.

К этой модели мира надо добавить еще и модель своих собственных предпочтений — насколько желанны для нас различные состояния мира в контексте наших действий. Эти предпочтения, выявляемые в процессе обучения

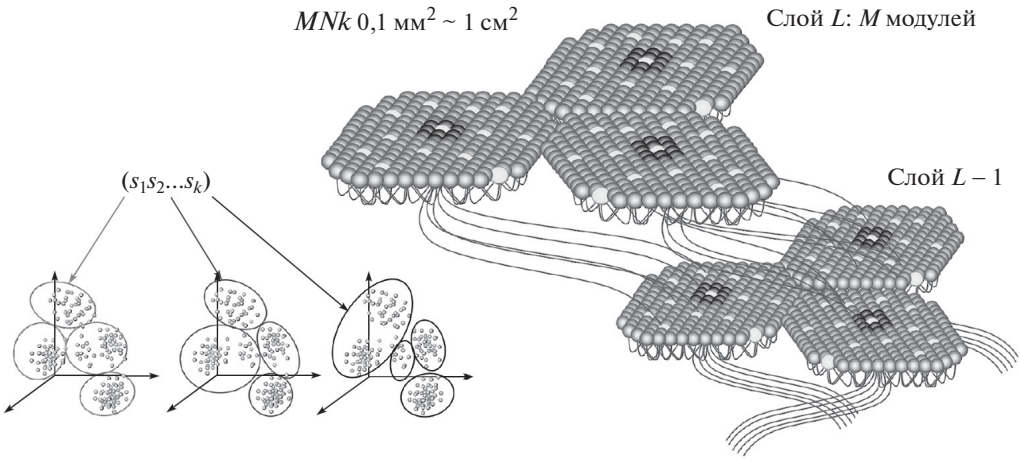

Рис. 8. Гипотетические рекурсивные модули неокортекса, соответствующие архитектуре Deep Control. Центральная гиперколонка каждого модуля с глобальными связями кодирует свою входную информацию активностью одной из своих кортикальных колонок. Окружающие ее гиперколонки с локальными связями кодируют последовательности таких символов в центральной гиперколонке, т.е. морфемы.

с подкреплением, задаются отдельной *функцией ценности* $V_{\mathbf{w}}$, за которую в мозге отвечают базальные ганглии. Семантическая память и функция ценности обучаются Парсером, который распознает и порядок следования морфем, и соответствующие им подкрепления. В простейшем случае используется алгоритм обучения SARSA [15]:

$$V(\mathbf{w}) \leftarrow V(\mathbf{w}) + \alpha(r + \gamma V(\mathbf{w}') - V(\mathbf{w})).$$

Модель мира и модель наших предпочтений используются *Планировщиком* для планирования поведения. Если вычислительный слой является самым верхним, то информация о текущем контексте, распознанным Парсером, передается непосредственно Планировщику. Зная последнюю распознанную морфему, какие морфемы могут следовать за ней и ценность каждой такой морфемы-кандидата, Планировщик выбирает оптимальную в данном контексте следующую морфему, которая и становится его текущим планом.

Этот план затем пошагово транслируется *Декодеру*. В первом слое Декодер переводит его из символьной формы в векторную — предсказывает следующий сенсомоторный вектор, т.е. непосредственно взаимодействует с внешним миром. В остальных слоях Декодер формирует с помощью своей эпизодической памяти планы для Планировщика нижележащего слоя — ранжированный список возможных морфем-кандидатов, из которого последний выбирает оптимальную для текущего момента морфему. Тем самым каждый слой выбирает наилучший вариант исполнения планов вышележащего слоя с учетом поступающей от нижележащего слоя входной информации.

Информация от нижележащего слоя передается в вышележащий слой Парсером в момент распознания им очередной морфемы. Эта морфема

передается Кодеру вышележащего слоя в форме семантического вектора для его последующего дискретного кодирования. Семантический вектор $\mathbf{X_w} = R_{...\mathbf{w}} R_{\mathbf{w}...}$ каждой морфемы определяется частотами морфем, предшествующих и следующих за данной морфемой. Таким образом, морфемы, употребляемые сходным образом, будут иметь близкие семантические векторы и соответственно будут закодированы Кодером следующего слоя близкими дискретными кодами с большим числом одинаковых компонент.

Итак, мы описали в общих чертах, как в архитектуре Deep Control происходит формирование иерархии планов поведения, вложенных друг в друга и постоянно адаптирующихся к изменяющимся внешним обстоятельствам. Заметим, что от нас не требовалось задавать никаких правил поведения. Все паттерны поведения, кодируемые морфемами на каждом временном масштабе, появляются автоматически в процессе активного взаимодействия системы с внешним миром на основе полученных при этом данных. Искусственный мозг с такой архитектурой способен автоматически формировать свою картину мира и постоянно совершенствовать свое поведение, нацеленное на максимизацию ожидаемого потока подкреплений. Он не ограничен решением какой-то одной определенной задачи и может накапливать опыт решения разных задач в разных контекстах, постоянно накапливая знания о мире и своем опыте взаимодействия с ним. Можно сказать, что он обладает свободой воли, так как он исполняет лишь свои собственные планы, вырабатываемые на верхнем (мотивационном) слое иерархии. Управление его поведением, например нацеливание на решение определенной задачи, происходит не директивно, а через управление подкреплениями, например увеличением соответствующей награды, чтобы повлиять на его текущую мотивацию.

4. ADAM: прототип искусственной психики роботов

ADAM (Adaptive Deep Autonomous Machine) представляет собой действующий прототип искусственной психики с архитектурой Deep Control, разрабатываемый в лаборатории Когнитивных архитектур МФТИ с целью проверить работоспособность предложенного подхода и апробировать различные алгоритмы обучения для новой архитектуры [16]. Разработанная и отлаженная в рамках проекта ADAM архитектура сильного ИИ может использоваться в самых разных сервисах и продуктах, требующих креативного машинного мышления. Мы надеемся, что эта архитектура будет положена в основу будущих операционных систем роботов и специализированных чипов их искусственного мозга.

ADAM представляет собой программу на языке Julia, управляющую поведением робота или программного агента в симуляторе реальности:

$$\textbf{Environment}\,(\mathbf{a}, \mathbf{o})_t \rightarrow (r, \mathbf{o})_{t+1}.$$

Псевдокод ADAM может быть представлен в следующем виде:

ADAM(parameters,$(r, \mathbf{o}, \mathbf{a})_{1:t}$) $\rightarrow (r, \mathbf{o}, \mathbf{a})_{t+1}$
track_record[1] $\leftarrow$ **explore_environment**() # gather initial learning data
adam $\leftarrow$ **new_adam**(parameters) # create new adam
adam.layer[L=1] $\leftarrow$ **new_layer**(adam, track_record[1]) # adam's first layer
while(stop_criteria)
$\hat{r}, \hat{\mathbf{o}}, \mathbf{a} \leftarrow$ **predict!**(adam, $r, \mathbf{o}, \mathbf{a}$) # generate new action/prediction
create next layer if needed:
if(expand_criteria) adam.layer[L+1] $\leftarrow$
$\leftarrow$ **new_layer**(adam, track_record[L+1])

Код ADAM тестируется на задачах из коллекции OpenAI Gym, а также используется в прикладных задачах, в частности — в автоматической биржевой торговой системе, разрабатываемой в МФТИ. Предварительные результаты показывают, что ADAM действительно обучается достигать решения задач за все более короткое время, как показано на рис. 9. В частности, в задаче CartPole требуется научиться балансировать обратный маятник в те-

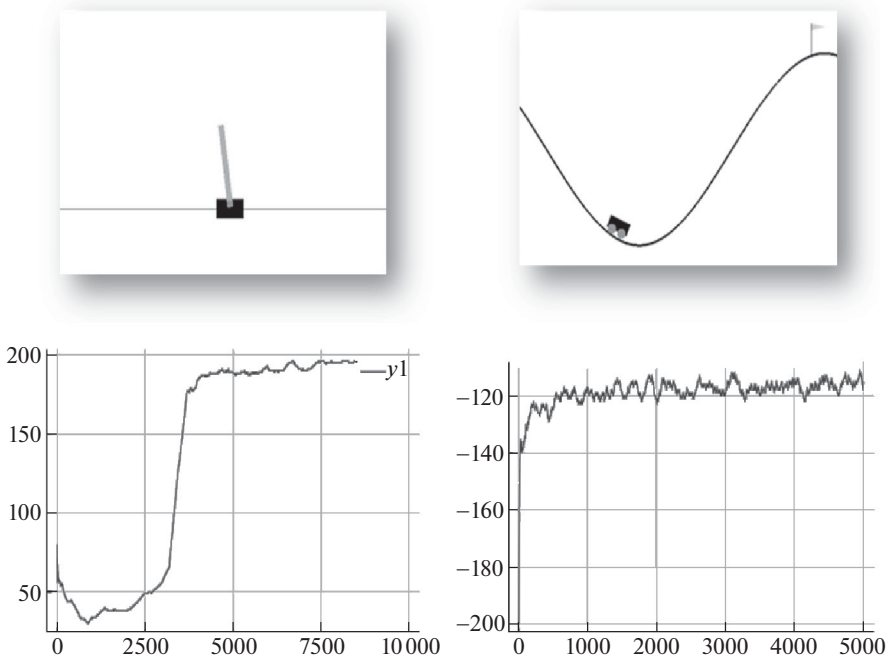

Рис. 9. Пример обучения кода ADAM с одним вычислительным слоем задачам CartPole и MountainCar из библиотеки OpenAI Gym. Ось абсцисс — количество эпизодов обучения, ось ординат — количество подкреплений, полученных в каждом эпизоде.

чение как минимум 200 тактов, а в задаче MountainCar — достигнуть флага, забравшись на склон быстрее, чем за 200 тактов. Причем в первом случае дается единичная награда за каждый такт удачного балансирования, а во втором — отрицательная единичная награда за каждый такт до достижения флага. Как видно из рис. 9, награды, полученные в каждом эпизоде, растут по мере обучения решения обеих задач.

5. Заключение

В данной статье описана модель искусственной психики ADAM с иерархической архитектурой глубокого обучения с подкреплением Deep Control. ADAM способен планировать свое поведение на многих масштабах времени, вписывая планы более низких уровней в планы более высоких и согласовывая поток планов, спускаемых сверху-вниз, с потоком сенсорной информации, поступающей снизу-вверх. По мере накопления опыта взаимодействия с внешней средой и роста числа слоев ADAM обучается целенаправленному поведению на все более долгих временных масштабах. Этот подход может быть использован при создании операционных систем автономных роботов, способных накапливать опыт обучения решению самых разных задач.

СПИСОК ЛИТЕРАТУРЫ

1. *Russell S., Norvig P.* Artificial Intelligence: A Modern Approach. (3rd Edition). Pearson, 2009.
2. *Николенко С., Кадурин А., Архангельская Е.* Глубокое обучение. Погружение в мир нейронных сетей. Питер, 2018.
3. *Bengio Y.* From System-1 Deep Dearning to System-2 Deep Learning // Thirty-third Conf. on Neural Information Processing Syst. 2019.
4. *Kotseruba I., Tsotsos J.K.* 40 Years of Cognitive Architectures: Core Cognitive Abilities and Practical Applications // Artificial Intelligence Review. 2020. V. 53. No. 1. P. 17–94.
5. *Silver D., Singh S., Precup D., Sutton R.S.* Reward is Enough // Artificial Intelligence. 2021. 103535.
6. *Laird J.E., Lebiere C., Rosenbloom P.S.* A Standard Model of the Mind: Toward a Common Computational Framework Across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics // AI Magazine. 2017. V. 38. No. 4. P. 13–26.
7. *Russell S.* Human Compatible: Artificial Intelligence and the Problem of Control. Viking, 2019.
8. *Silver D., et al.* Mastering chess and shogi by self-play with a general reinforcement learning algorithm // arXiv preprint arXiv:1712.01815. 2017.
9. *Шумский С.А.* Глубокое структурное обучение: новый взгляд на обучение с подкреплением // Сб. науч. тр. XX Всеросс. науч. конф. Нейроинформатика-2018. Лекции по нейроинформатике. С. 11–43. М.: 2018.
10. *Шумский С.А.* Реинжиниринг архитектуры мозга: роль и взаимодействие основных подсистем // Сб. науч. тр. XVII Всеросс. науч. конф. Нейроинформатика-2015. Лекции по нейроинформатике. С. 13–45. М.: 2015.

11. *Clark A.* Surfing Uncertainty: Prediction, Action, and the Embodied Mind. Oxford University Press, 2015.
12. *Friston K.J.* Waves of Prediction // PLoS Biology. 2019. V. 17. No. 10. e3000426.
13. *Mountcastle V.B.* The Columnar Organization of the Neocortex // Brain: a journal of neurology. 1997. V. 120. No. 4. P. 701–722.
14. *Kohonen T.* Self-organized Formation of Topologically Correct Feature Maps // Biological cybernetics. 1982. V. 43. No. 1. P. 59–69.
15. *Sutton R.S., Barto A.G.* Reinforcement learning: An introduction. MIT press, 2018.
16. *Шумский С.А., Басков О.В.* Программный Агент глубокого иерархического обучения с подкреплением ADAM Deep Control // Государственная регистрация программ для ЭВМ. RU 2021660307.

Статья представлена к публикации членом редколлегии О.П. Кузнецовым.

Поступила в редакцию 31.10.2021

После доработки 21.01.2022

Принята к публикации 26.01.2022