

УДК 612.8+519.85

## КОГНИТИВНАЯ АРХИТЕКТУРА ПОЗНАВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ ПРИ ЕЕ МОДЕЛИРОВАНИИ И ПСИХОФИЗИОЛОГИЧЕСКОЙ ОЦЕНКЕ

© 2023 г. О. М. Разумникова\*

Новосибирский государственный технический университет, г. Новосибирск, 630073 Россия

\*e-mail: razoum@mail.ru

Поступила в редакцию 20.12.2022 г.

После доработки 14.01.2023 г.

Принята к публикации 09.03.2023 г.

Описаны основные подходы к моделированию познавательной деятельности человека и нейронных механизмов, лежащих в ее основе. Приведена систематизация когнитивных архитектур и дан обзор таких популярных моделей как ACT-R, SOAR, CLARION и CHREST с примерами их практического применения в психологии и нейрофизиологии. Разработанные модели когнитивных функций позволяют давать прогнозы эффективности восприятия и селекции информации, какие знания и процедуры требуются для оптимального решения задачи, ожидаемую частоту ошибок при выполнении задания и какая функциональная система мозга используется для организации поведения. Совершенствование и дополнение существующих моделей когнитивной архитектуры рассматривается как перспектива развития когнитивной нейронауки, понимания закономерностей формирования естественного интеллекта и разработки искусственного интеллекта.

*Ключевые слова:* когнитивная архитектура, восприятие, память, принятие решения, эмоциональная регуляция, искусственный интеллект

**DOI:** 10.31857/S0301179823030074, **EDN:** GLHKRN

Активно развивающаяся разработка технологий искусственного интеллекта опирается на знания, полученные в когнитивной психологии, представляя их в формализованном виде как систему: когнитивную архитектуру. Эти знания включают результаты изучения таких психологических процессов как восприятие, распознавание образов, внимание, память, речь и принятие решения при конкуренции информационных потоков.

Понятие “когнитивная архитектура” согласно информации, представленной в Pubmed, в 2010 г. упоминалось в 211 работах, в 2020 – в 659, а в 2022 – уже в 861. Столь широкое применение понятия “когнитивная архитектура”, по-видимому, обусловлено разнообразием во взаимосвязи компонентов мыслительной деятельности: внимания, памяти и принятия решения, представляющих как естественный, так и искусственный интеллект. Приведем примеры когнитивной архитектуры как подходов моделирования искусственного интеллекта с учетом закономерностей, выявленных при изучении психофизиологических механизмов мышления. Модели когнитивной архитектуры искусственного интеллекта имитируют процессы естественного интеллекта и поведение человека в различных областях деятельности, позволяя таким образом проверить соответствие

выявленных нейробиологических закономерностей и вычислительных конструктов. В библиотеке “e-library” поиск с ключевым словом “когнитивная архитектура” хотя и дает менее 40 публикаций, однако охватывает статьи от методологических вопросов разработки искусственного интеллекта до прикладных задач применения когнитивных архитектур, например, в оптимизации транспортной системы или при освоении знаний английского или химии и т.д.

Развитие работ по созданию когнитивных архитектур привело к созданию интегрированных систем для имитации интеллектуального поведения при решении разнообразных задач. Их обзор, выполненный в 2016 г., показал существование 55 активных когнитивных архитектур [39]. Широкий спектр архитектур связан с тем, что каждая из них предназначена для выполнения различных ментальных процессов, включая память, обучение, решение проблем, эмоции, перцептивно-моторный контроль, язык, инженерную психологию и приложения для решения биологических или нейронных проблем. На рис. 1 приведена общая систематизация когнитивных архитектур и отдельные примеры символических, эмерджентных и гибридных архитектур, различающихся механизмами организации памяти и обучения [20].



Рис. 1. Общая систематизация когнитивных архитектур.

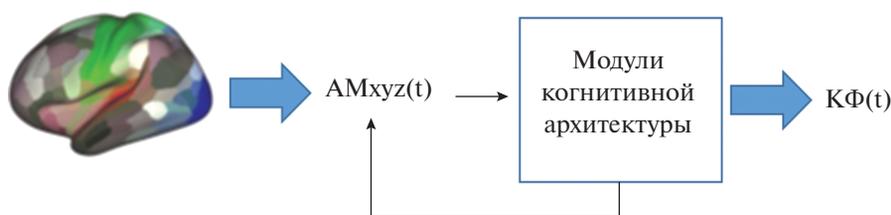


Рис. 2. Схематическое представление разработки когнитивной архитектуры искусственного интеллекта, имитирующего временную динамику когнитивных функций (КФ) на основе закономерностей активности мозга (АМ) для разных структур мозга.

Одновременное моделирование поведенческих данных и нейронных механизмов рассматривается как перспектива развития когнитивной нейронауки [72].

Принцип построения когнитивных архитектур искусственного интеллекта опирается на нейрофизиологические характеристики естественного интеллекта (функциональную специализацию структур мозга для организации когнитивных функций), как это схематически показано на рис. 2.

Одной из наиболее широко известных когнитивных архитектур является **ACT-R** (Adaptive Control of Thought-Rational – Адаптивный контроль мысли - рациональное). Ряд ее версий развивается с 70-х годов за счет расширения и преобразования модулей на основе данных когнитивной психологии и нейрофизиологии, в частности ассоциативной памяти и теории рационального анализа [8, 10].

ACT-R представлена набором программируемых механизмов обработки информации, которые можно использовать для прогнозирования и

объяснения поведения человека, включая познание и взаимодействие с окружающей средой. Это гибридная когнитивная архитектура, так как ее правила и декларативная память являются символическими, а процессы изменений: как и когда элементы памяти будут использоваться, описаны субсимволически.

Начавшись с разработки механизмов декларативной и процедурной памяти, когнитивная архитектура ACT-R для реализации теории компиляции знаний была дополнена GRAPES (Goal-Restricted Production System) [9]. Компиляция знаний – это процесс интеграции расширенных вычислений (применение нескольких правил и/или выборки из памяти) в одно производственное правило [56]. Изначально память задач кодировалась как декларативное знание посредством обучения. Однако при повторяющейся практике компоненты задачи (например, воспоминания, используемые для выполнения задачи, и знание того, какие воспоминания нужно извлечь и какие задачи выполнять) постепенно объединяются в единую процедурную память или “компилируются”.

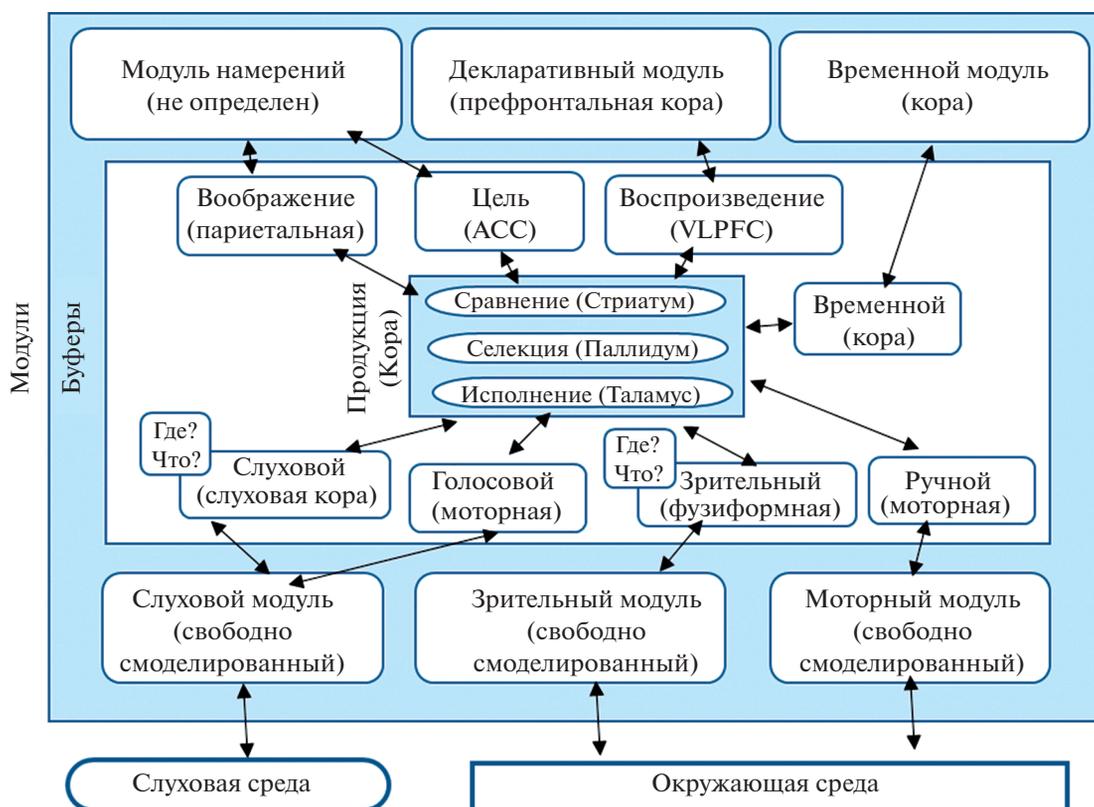


Рис. 3. Схематическая диаграмма когнитивной архитектуры ACT-R6.

Предполагается, что этот механизм подобен тому, как люди улучшают выполнение процедурной задачи посредством повторной практики.

Развитие ACT-R в 6-ой версии представлено введением инструктируемой производственной системы и дополнениями на основе использования данных картирования функций мозга [7]. Схематическая диаграмма когнитивной архитектуры ACT-R6 показана на рис. 3.

Приведенная схема ACT-R6 включает сенсорные модули для обработки слуховой, зрительной и моторной информации, двигательные модули для организации действий, модуль для долгосрочных декларативных знаний и модуль целей. Каждый модуль имеет буфер памяти, представляющий кратковременную память архитектуры ACT-R. При обучении новым знаниям или правилам решений модуль декларативных знаний пополняется в случае успешного результата деятельности и модифицируется вследствие ошибки. Буферы имеют такие характеристики, как скорость «распада» памяти и какие действия могут быть предприняты между буферами.

Модель работает на основе правил «если-то», состоящих из шаблонов для сопоставления (если) и действий, которые необходимо предпринять (тогда). Структуры, которые могут принимать шаблоны, определяются архитектурой. В ACT-R

эти шаблоны соответствуют объектам (или их отсутствию) в буферах, составляющих механизмы ACT-R. Они используются для представления знаний о процедурных задачах.

Элементы декларативной памяти: это символы, представляющие объекты декларативной памяти, которые могут иметь атрибуты и ценности. Они используются для представления фактов о мире и могут отражать информацию о внутреннем состоянии, например, целях деятельности. Наряду с символическими компонентами имеется субсимвольная информация, согласно которой модерируется поведение и доступность декларативной памяти.

Модули и буферы ACT-R соответствуют функциональному значению тех областей мозга, которые вовлечены в процессы восприятия информации, ее запоминания и организации деятельности согласно данным, полученным в ходе энцефалографических и томографических исследований. Данные фМРТ применяются для информирования и ограничения архитектуры, и в свою очередь, когнитивную архитектуру ACT-R можно использовать для создания прогнозов и принципиальной интерпретации данных фМРТ.

Процессы восприятия включают не только результаты ощущений различных видов, но и их интеграцию в единое оценивание или описание

окружающей ситуации, которые могут быть использованы другими когнитивными процессами. Чувствительные элементы модулей восприятия дают неточную и частичную картину среды с наложенным на нее шумом. Динамические характеристики окружающей среды еще более усложняют этот процесс из-за того, что агент должен отслеживать изменения, которые происходят иногда с большой скоростью. Эти проблемы могут быть компенсированы с помощью перцептивных знаний о том, какие чувствительные элементы могут их вызывать и какие выводы являются правдоподобными. Архитектура может приобретать и улучшать эти знания путем обучения на основе предыдущего опыта восприятия.

Восприятие информации и ее обработка в анализаторных системах мозга создает сенсорную модель мира, положения в ней человека и организации его деятельности [3]. Большая часть взаимодействия с окружающей средой осуществляется с помощью зрительной и двигательной системы. Ранее полученные и новые результаты психофизических и нейрофизиологических исследований механизмов зрительного восприятия [2, 4, 6] широко используются для моделирования когнитивной архитектуры.

Представленный в АСТ-R модуль зрения, показанный на рис. 3, включает в себя два буфера: один для информации “Где” (местоположение) и один для информации “Что” (значение объекта в визуальной сцене). Эти две системы аналогичны дорсальному (“где”) и вентральному (“что”) потокам в зрительной системе человека [1–3, 8, 52].

Зрительная система согласно АСТ-R представляет собой имитацию фиксации и движений глаз и может следить за положением на экране и перемещать свое внимание (фиксировать взгляд на объекте), чтобы выполнять следящие движения глаз. Объединение фрагментов расположения объектов и их семантических контекстов становится доступным для центральной производственной системы, чтобы принять решение и привести к поведению: например, переключению визуального внимания.

В модели периодического внимания для классификации изображений предлагается обучение нейронной сети с использованием вознаграждения в форме отрицательной перекрестной энтропии, повышающего информативность сигнала подкрепления [47]. Полезным инструментом для анализа изображений при разработке систем искусственного интеллекта является обучение сверточных нейронных сетей с использованием признаков, характерных для биологических систем [45, 48].

Успешное применение модели зрительного внимания, реализованной в когнитивной архитектуре АСТ-R, продемонстрировано согласно

результатам анализа изменений амплитуды когерентности тета ритма, вычисленных при многоканальной регистрации ЭЭГ [73]. Показано, что динамика ресурса рабочей памяти согласно модели представлена тета осцилляциями в теменной коре, тогда как динамика декларативной памяти ресурсов – в задних областях. Предсказания межмодульной коммуникации в процессах идентификации стимула и консолидации цели ассоциируются с избирательным увеличением когерентности в прогнозируемые моменты времени.

Для распознавания и категоризации информации в АСТ-R предусмотрено ее сохранение и сравнение в памяти. Во время цикла принятия решений продукционные правила используются для анализа шаблонов во всех буферах, которые могут представлять внутреннюю и внешнюю информацию. Эти правила состоят из условий и действий в соответствии с архитектурными обязательствами относительно представлений внутренней и внешней информации в буферах и в симулированном “уме”. Правила, соответствующие текущей цели, сопоставляются с набором буферов, подключенных к центральному процессору и на основе сравнения выбирается и применяется то, которое является допустимым или соответствует наибольшей полезности. Правило может иметь несколько действий, и они могут передавать команды восприятию или двигательному модулю и/или изменять значения любого буфера, обновляя декларативную или целевую память или помещая новую команду в зрительный или двигательный модули. Улучшение выбора решения осуществляется на основе обучения новым альтернативам.

Модуль намерений может выполнять обновление содержимого любого слота, а также его очистки и сохранения отдельного фрагмента цели в декларативной памяти или создание нового фрагмента цели по запросу. Его буфер “Цель” содержит текущее состояние элемента управления и используется для поддержания задачи модели и служит источником активации для декларативного извлечения памяти. Модуль цели ассоциируется с передней цингулярной корой (рис. 3, АСС), а процедурный модуль – с базальными ганглиями.

Декларативная архитектура памяти в АСТ-R осуществляет имитацию человеческой памяти. Предусмотрены правила, учитывающие скорость извлечения воспоминаний и частоту их использования. Эффект “веера” (насколько хорошо связаны разные следы памяти) и порог воспоминаний влияют на время извлечения данного воспоминания, которое задается в модели условиями поиска. Затем полученная память используется для создания поведения [11]. В противном случае, если нужный след памяти не может быть извле-

чен, модель должна иметь стратегии для решения этой проблемы, иначе она перестанет работать. Чтобы избежать этого предлагаются механизмы, в которых связанные воспоминания могут быть сопоставлены или объединены (“смешивание воспоминаний”) или следование частичному соответствию правил текущему состоянию знаний [64]. Механизм с наибольшим обучающим эффектом в АСТ-R — усиление, укрепление и снижение времени поиска декларативной памяти по мере извлечения и использования воспоминаний.

Для измерения субъективной умственной нагрузки, которая сопровождает эффект гистерезиса при переходах между задачами и требованиями на основе модели АСТ-R были разработаны условия выполнения мультизадачи с разной нагрузкой и выявлены связанные с ней различия в нейронных сетях, вычисленных согласно ЭЭГ на альфа и бета частотах [38]. Обнаружен эффект гистерезиса, возникающий на нескольких уровнях анализа: поведенческом, субъективной умственной нагрузки и эффективной связи сетей мозга, в частности связей префронтальной и цингулярной коры. Авторы заключают, что результаты, касающиеся нейронных механизмов эффектов гистерезиса в многозадачных средах, могут быть использованы в будущих нейрокогнитивных моделях для прояснения направления и силы причинно-следственной связи Грейнджера при изменении спроса, расширяя понимание когнитивной обработки, интерпретируемой с помощью данных ЭЭГ. Не только сложность или тип задачи, или объем памяти, но и такие психические функции как эмоциональная и мотивационная регуляция ментальных процессов может быть представлена как суперпозиция низкочастотных и высокочастотных ритмов ЭЭГ, отражающая функциональное взаимодействие структур мозга, в том числе их резервы [62].

Сопоставление картирования функциональной активности мозга с применением фМРТ при реализации разных задач и когнитивной архитектуры АСТ-R продемонстрировало хорошее соответствие предсказаний модели для модуля проблемного состояния, меньшее — для декларативной памяти и лучшие различия решения двух типов проблем при почти идентичном ответе согласно данным фМРТ [13].

Учитывая значимый вклад эмоций в регуляцию когнитивных процессов, в частности, памяти, ее модель дополняется компонентами аффективной оценки информации, основанной на результатах психологических и нейрофизиологических исследований. Для выполнения задачи требуется интеграция оценок, воплощенных в соответствующих модулях: позитивных эмоций согласно функциональной активности вентральных отделов стриатума и паллидума, как структур, опосредующих награду выбранной деятельности, и негативных эмо-

ций (опасности) согласно функциональной активности инсулы [49, 63]. Предложенная авторами система эмоциональной регуляции эпизодической памяти с одной стороны создает модель проверки более обширного взаимодействия структур мозга в аффективной регуляции когнитивных функций и поведения, а с другой — расширяет когнитивную архитектуру виртуальных агентов и роботов, улучшая их обучение процессам планирования и принятия решения за счет эмоциональной оценки ситуации [70].

Например, с использованием данных фМРТ показано, что в эмоциональную оценку предъявленных эмоциогенных стимулов (стимулов из атласа аффективных состояний — IAPS) вовлекаются не только структуры лимбической системы и префронтальной коры, но и система мозга по умолчанию (Default Mode Network), и система когнитивной переоценки (Cognitive Reappraisal Network), что указывает на широкое индивидуальное разнообразие нейронных сетей эмоциональной регуляции [53].

Согласно когнитивной архитектуре, основанной на теории двойных процессов Канемана [37], долговременная память моделируется как нейронная сеть, развивающаяся при взаимодействии с окружающей средой. Рабочая память представляет буфер, содержимое которого преобразуется в соответствии с правилами, хранящимися в долговременной памяти. В этой архитектуре этапы символического и субсимвольного рассуждений могут быть объединены, а вычисления с ограниченными ресурсами представлены в диапазоне от формальных доказательств до ассоциативных цепочек [71]. Традиционные вычислительные модели включают символические модели, такие как автоматические средства доказательства теорем [31], субсимвольные модели (например, искусственные нейронные сети (ИНС) [65] и вероятностные модели, такие как байесовские сети [60]).

Теория двойного процесса в моделировании когнитивной архитектуры используется, чтобы сохранить важное, но часто игнорируемое различие между реальностью (восприятием) и воображением. Архитектура основана на двух системах памяти: долговременной памяти (представляет собой автономную систему знаний, которая автоматически развивается посредством взаимодействия с окружающей средой) и рабочей памяти (как система памяти, используемая для текущих ограниченных по ресурсам операций). Эти операции включают моделирование арифметических вычислений, доказательства высказываний, генерации последовательностей, классификации, ассоциации и предположения. Таким образом, символическая и субсимвольная обработка могут сосуществовать и взаимодействовать друг с другом в этой монолитной архитектуре.

Символические структуры в АСТ-R представляют когнитивные процессы в функциональных системах мозга как процедурные правила операций со знаниями, т.е. правила извлечения и проверки накопленных знаний и представления о средствах и путях получения новых знаний.

Субсимволические системы исполняют математические операции для функциональных систем в каждом модуле, включая абстрактные репрезентации элементов (chunks), правил и скорости их выполнения [57].

Модуль двигателя (моторный) получает запросы команд через буфер двигателя, и эти команды указывают, как выводить действия во внешний мир. В АСТ-R имеется небольшой набор команд для набора текста на клавиатуре и перемещения мыши в окне, а некоторые исследователи расширили этот подход дополнительным кодом для управления автомобилями или роботами.

АСТ-R может запечатлевать процедурные воспоминания (т.е. накопленных правил выполнения предыдущих действий). В версиях АСТ-R 6 и 7, если два правила срабатывают достаточно близко друг к другу, они могут быть объединены в одно. В свою очередь время срабатывания этого правила часто зависит от извлечения декларативной памяти, т.е. механизмы декларативного и процедурного знания взаимодействуют. Дальнейшее улучшение перцептивно-моторных возможностей АСТ-R планируется на основе воспоминаний, включающих синтез ощущений (например, запаха и звука) или эмоций и семантики [64].

Инструменты для работы с АСТ-R даны в общем репозитории: <http://act-r.psy.cmu.edu>. Поведение может описываться с применением нескольких высокоуровневых языков, которые позволяют представлять иерархические или последовательные задачи [64]. Эти языки, включая Herbal, обладают всеми преимуществами других языков высокого уровня, используемых в информатике, включая плавную интеграцию в системы, созданные на таких широко распространенных языках, как Java; поддерживаемые обширными библиотеками и инструментами; высокой степенью модульности реализации и, следовательно, возможностью более легкого исследования изменений и расширений существующих когнитивных архитектур, а также выполнения сравнительного анализа и их реконструкции за счет меньшего количества операций с данными.

Когнитивная архитектура АСТ-R является одновременно и психологической теорией, и структурной моделированием для построения когнитивных моделей, которые придерживаются принципов теории. Модели АСТ-R дают множество прогнозов, включая элементы ответа, какие знания требуется использовать, частоту ошибок при выполнении задания или какая функциональная

система мозга используется для организации поведения, обучения или забывания. АСТ-R можно использовать в комбинации с данными фМРТ двумя способами: 1) данные фМРТ применять для оценки и ограничения модели в АСТ-R с помощью предопределенного анализа области интереса (ROI) и 2) прогнозы из АСТ-R модели могут использоваться для обнаружения нейронных коррелятов модельных процессов и представлений с помощью основанного на моделях анализа фМРТ или 3) использовать АСТ-R в сочетании с данными фМРТ [13].

Так как человеку приходится обычно решать множество задач, выбирая среди них приоритетные, то в развитие АСТ-R предлагается ввести параметр “сложность задания” и модуль переключения задач дополнительно к модулям прерывания и запуска выполнения задания [17, 57]. Чем труднее задача, тем больше времени на нее затрачивается и больше вероятность отказа и переключения на другую деятельность при сохранении этой ситуации в памяти. Поэтому для выполнения сложной задачи требуется учитывать ее трудность, отличительные черты, интерес, приоритет и затрачиваемое время [74]. Одним из психофизиологических показателей трудности задачи является степень расширения зрачка и соответствующее переключение деятельности [36]. При разработке нейросетевых архитектур для решения таких задач предлагается обучение с динамически настраиваемой структурой вычислений по принципу наименьшего действия [46, 48].

Фактором переключения могут быть эмоции, возникшие вследствие рассогласования сложности задачи и способности ее решить. Орбитофронтальная кора играет ключевую роль в этом процессе и имеет взаимосвязи с базальными ганглиями, которые представлены в АСТ-R процедурным модулем, выполняющим функции стриатума (рис. 3). Эти связи представлены несколькими субсимвольными уравнениями, чтобы количественно оценить дисбаланс между сложностью текущей задачи и способностью исполнителя. Дополнительно эти уравнения введены в процедурный модуль, так как текущая версия АСТ-R не имеет модуля, отражающего функции орбитофронтальной коры в регуляции эмоций [57].

Следующий шаг в поведении модели должен представлять решение о выполнении другой задачи с соответствующим освобождением буфера текущей или ее продолжении согласно доминирующей цели. Для этого АСТ-R содержит алгоритмы правила “переключить решение” или “переключить задачу”. Сравнительный анализ переключения деятельности в зависимости от сложности игровой задачи выявил ограничения, связанные с индивидуальным непостоянством решения и большим эффектом консерватизма у людей, что

предлагается использовать в дальнейшем для улучшения разработанной модели [57].

Метод когнитивного моделирования использовался для понимания процессов, лежащих в основе динамического принятия решений в сложной системе с применением модели ACT-R [61]. Модель решает задачу категоризации, сначала апробируя однопризнаковые стратегии как начальное обучение, а затем, в результате многократной отрицательной обратной связи, переключаясь на двухпризнаковые стратегии или обратное обучение вследствие непредвиденных обстоятельств. Данные моделирования имели большую дисперсию и более низкую общую производительность по сравнению с данными человека, что может быть связано с предпочтениями восприятия или дополнительными знаниями и правилами, применяемыми участниками. Из-за больших индивидуальных различий в эффективности принятия решений информация о лежащих в их основе когнитивных процессах, полученная на основе поведенческих и нейрофизиологических исследований, может помочь оптимизировать будущие приложения модели, чтобы улучшить соответствие результатов и использовать для понимания других областей когнитивной деятельности.

Еще одной популярной когнитивной архитектурой является **SOAR** (State, Operator And Result) [43]. Цель **SOAR** – осуществить рациональное поведение, поддерживая все возможности, требуемые от обычного интеллектуального агента. Архитектура **SOAR** состоит из представления 1) структуры задач и подзадач, 2) долговременной памяти, 3) кратковременной памяти, 4) механизма генерирования целей и 5) механизма обучения. В совокупности эта архитектура предоставляет средства, облегчающие реактивное принятие решений, ситуационную осведомленность, рассуждения, понимание, планирование и обучение.

**SOAR** поддерживает три типа долговременной памяти: процедурную, семантическую и эпизодическую. Процедурная память представляет собой практические знания, семантическая – факты, а эпизодическая – опыт. Долговременная память **SOAR** представлены в виде операторов программы и правил типа *if/then/else*. Эти правила предварительно определены во входном файле, но могут быть изменены или созданы во время выполнения. Долговременная память внутренне организована в соответствии с проблемным пространством и доступна для поиска.

Кратковременная или рабочая память, – это знания о текущей ситуации. Рабочая память создается в ответ на изменения в восприятии окружающей среды. Она содержит как данные о воспринимаемом состоянии, так и знания, извлеченные из долговременной памяти. Решения принимаются на основе доступных знаний рабочей памяти или

правил, которые обеспечивают путь перехода от текущего состояния к некоторому будущему состоянию. Например, для предсказания эффективности зрительно-пространственной памяти могут быть полезны метрики, основанные на гештальт-принципе близости, реализованном в виде окрестностей фон Неймана и Мура [12].

Цель системы состоит в том, чтобы в конечном итоге перейти в состояние, которое завершает поставленную задачу. Если прямой переход недоступен из-за пробела в знаниях, **SOAR** пытается принять решение на основе имеющихся. После принятия решения результат или эффективность перехода сохраняется в долговременной памяти. Если переход никогда ранее не выполнялся, новое правило автоматически генерируется и сохраняется, чтобы заполнить пробел в знаниях в долговременной памяти.

Для обеспечения реалистичного поведения требуются обширные базы знаний. Например, **TacAir-Soar** представляет собой модель пилота-истребителя [33] и содержит более 7500 правил для полуреалистичного управления смоделированным самолетом в ходе тренировочных упражнений в реальном времени, что делает ее одной из крупнейших полевых экспертных систем. Решения о взлете не компилируются и не предопределяются заранее, а принимаются на основе соответствующих долгосрочных и практических знаний.

**SOAR** может рассматриваться как прямой преемник **GPS** (общий решатель задач [55]), созданный на основе предположения, что механизмы решения задач являются ключом к человеческому познанию и интеллекту. Целью модели **TacAir-Soar** было создание нескольких смоделированных пилотов, которые могут участвовать в крупномасштабных симуляциях боевых действий вместе с пилотами-людьми и другим персоналом. Основным критерий успеха заключается в том, действуют ли модели как настоящие пилоты и, следовательно, вносят свой вклад в опыт, верный для людей, участвующих в моделировании.

**TacAir-Soar** имеет впечатляющие характеристики и достижения. Программа способна выполнять широкий спектр миссий и может общаться с диспетчером и другими пилотами, используя естественный (хотя и ограниченный) язык. Система была успешно опробована в крупномасштабном синтетическом моделировании театра военных действий [33], в котором участвовало 3700 машин с компьютерным управлением, среди которых примерно 100 самолетов управлялись моделями **TacAir-Soar**. Успех модели определялся несколькими составляющими. Во-первых, использовалась иерархия целей с разбиением более крупных целей (например, перехват самолета) на более мелкие цели (например, изменение пеленга). Для поддержки этой целевой структуры

система содержала более 8000 правил производства, которые были получены от экспертов в предметной области, а также стандартной доктрины и тактики. Производственные правила в SOAR предлагают, определяют и реализуют операторы умственных или физических действий.

**Когнитивная архитектура CLARION** (Connectionist Learning with Adaptive Rule Induction ON-line) – коннекционистское обучение с адаптивным вводом правил в режиме онлайн) – была разработана, чтобы различать явный и неявный уровни поведения и их взаимодействие. В ее основе лежат представления о взаимодействии локализованных микроструктур, которые возбуждают и тормозят друг друга в параллельно осуществляемых операциях психических процессов [65]. CLARION состоит из четырех подсистем: ориентированной на действие (ACS), не ориентированной на действие (NACS), метакогнитивной (MCS) и мотивационной (MS).

ACS обрабатывает процедурные знания, а NACS – декларативные (факты, ассоциации, следы памяти). MS отвечает за работу с мотивациями и заботится об основных потребностях и вытекающих из них целях, которые предназначены для процессов ACS. MCS сочетает цели и подкрепление, наблюдает и выбирает между разными алгоритмами и параметрами системы [69]. Выходные данные этого модуля используются в качестве входных данных для ACS (подсистемы, отвечающей за выбор действий), так что на процесс принятия решений могут влиять мотивы текущего агента.

Каждая из этих систем включает имплицитный и эксплицитный уровень. Эксплицитный содержит набор взаимосвязанных элементов (chunks), каждый из которых представляет концепт и отвечает за явные знания (символические). Имплицитный – это нейронная сеть, обучающаяся с использованием алгоритмов глубокого обучения на основе градиента. Нейроны имплицитного уровня служат атрибутами элементов, представляя таким образом разные неявные знания (субсимвольные). Ввод данных в CLARION осуществляется через пары размер – значение, которые представляют, соответственно, тип информации и соответствующее ключу значение.

Как и все модули CLARION, его эксклюзивный модуль MS разделен на два уровня: верхний и нижний. Нижний уровень составляют побуждения как внутренние и психологические потребности организма и могут изменяться во времени, поэтому агент должен удовлетворить эти потребности, чтобы вернуться в состояние гомеостаза и расслабления. После того, как побуждения вычислены на нижнем уровне, они отправляются в MCS (мета-когнитивную подсистему), и на основе их значений MCS может добавлять новые цели

в обозначенную структуру целей верхнего уровня MS (ACS может добавлять новые цели непосредственно в эту структура тоже). Цели более конкретны и более явны, чем побуждения, и они хранятся в структуре целей, которая обеспечивает способ представления различных мотиваций и их взаимодействий в агенте. Модуль ACS затем может использовать цели, хранящиеся в структуре целей, для выбора действия.

Элементы ACS представляют концепты, которые комбинируются в правила условий и действий. Если условие соответствует текущей ситуации, рекомендуется действие. Имеется три типа правил: 1) фиксированные (рефлексы и моральные установки, которые не могут быть изменены), 2) с выбором улучшения и 3) правила на основе обучения (генерируются, изменяются и специализируются). Базовый уровень активации системы отражает частоту и время использования определенного правила. Полезность действия определяется его успешностью и ценой, а дальнейший выбор осуществляет MCS.

Элементы, представляющие NACS – это факты, ассоциации и установки. Они могут быть активированы за счет сети элементов, действий ACS, ассоциаций, их атрибутов на имплицитном уровне или подобий. Сила каждого из элементов определяется суммой этих составляющих и ослабевает или исчезает из памяти в случае неиспользования. Совершенствование NACS включает частоту распределения разных ситуаций и механизм подкрепления для улучшения обучения, а для мотивационной подсистемы – процесс оптимизации целей и предотвращение несостоятельного поведения вследствие сравнения новой цели и текущей [68].

В CLARION вероятность организации соответствующего поведения определяется базовым уровнем активации, который в свою очередь зависит от частоты ранее использованного правила или элемента. При этом имеется несколько путей достигнуть разных репрезентаций одной и той же проблемы: за счет уровней ее представлений, временных параметров или контроля MCS.

Недавние разработки в области исследования моделирования функциональных связей расширяют понимание данных нейровизуализации человека: от выявления изменений в региональных амплитудах активации до детального картирования крупномасштабных сетей мозга. Связывание сетевых и познавательных процессов требует развития теоретических основ, алгоритмов и применяемых экспериментальных подходов для перехода от описательного к объяснительному механизму когнитивных эффектов. Вопросы сетевой нейробиологии охватывают исследование когнитивной значимости сетей в состоянии покоя, спонтанную динамику сети или вызванную зада-

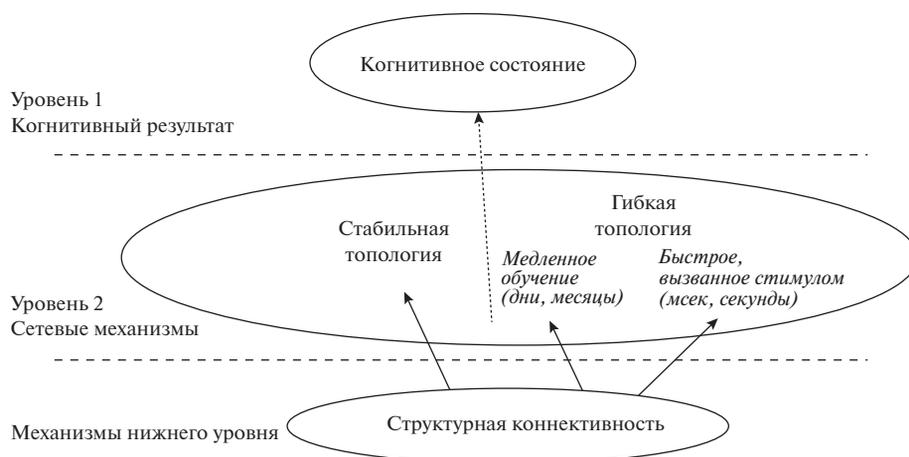


Рис. 4. Механизмы связи различных когнитивных состояний.

чей, а также как идентифицировать направленные или "эффективные" связи и как применять многомерный анализ паттернов на сетевом уровне. Моделирование когнитивной структуры позволяет выделить механистическое взаимодействие сетевых компонентов, которые остаются "стабильными" в разных предметных областях, и более "гибкие" компоненты, связанные с реконфигурацией по заданию. Применение различных аналитических подходов и надежных экспериментов дает возможность сопоставления формализованного когнитивного разума и крупномасштабных сетевых механизмов человеческого мозга [51].

Механизм сетевого уровня связывает взаимодействие между "стабильными" и "гибкими" пространственными компонентами с возникновением разных когнитивных состояний. "Гибкий" пространственный компонент сетевого механизма может быть разделен функционально различными временными компонентами, работающими в "медленных" и "быстрых" временных масштабах (связанных с обучением и реакцией, вызванной стимулом, соответственно) (рис. 4). Связи между стабильными, "медленными" гибкими и "быстрыми" гибкими сетевыми компонентами соответствуют различным лежащим в их основе нейрофизиологическим механизмам (механизмам нижнего уровня). Все прямые (а не косвенные) функциональные связи подкрепляются прямыми структурными связями с разделением на пространственно устойчивые и гибкие компоненты, обусловленные отсутствием или наличием динамики, соответственно.

Сетевые механизмы, формализованные таким образом, представляют прогресс по сравнению с описательными сетевыми отображениями состояния покоя.

Ключевой вопрос, привлекающий внимание при разработке когнитивной архитектуры: обеспечение гибкого выбора стратегии поведения на

основе заложенных правил. Эта же проблема рассматривается в рамках электрофизиологических и топографических исследований функциональных сетей мозга: их интернальных или экстернальных изменений в ходе обучения, или принятия решения, или специализации при решении задач разной сложности и разного информационного содержания.

Поведенческие исследования демонстрируют значение инструкции, вызывающей реконфигурацию нейронных сетей, обеспечивающую рабочую память и внимание для эффективного выполнения предъявленной задачи [27]. Подобным образом правило, заложенное в модели, определяет каскад операций, выполняемых в искусственной нейронной сети.

Согласно нейрофизиологическим данным намерение выполнить инструкцию может вызывать автоматическую двигательную активацию [22] или вовлекать различные области мозга для координации новых стимулов и формирования ответа [19, 29, 30, 58, 5]. При этом изменяется нейронная репрезентация инструктируемого контента в лобно-теменной сети (FPN) [14, 54], которая выполняет быстрый доступ и настройку реализации и высокоэффективного состояния мозга для успешного выполнения задачи [14, 29, 30, 54, 58, 59, 75]. Для моделирования лобно-теменной функции предлагается многоступенчатый процесс последовательного кодирования с начальным кодированием проинструктированной информации в декларативный код. Когда эта информация становится релевантной для поведения, декларативные представления FPN преобразуются в состояние реализации, состоящее из процедурных кодов для упреждающего связывания релевантной перцептивной и двигательной информации, оптимизированного для конкретных требований задачи [15].

Когнитивная гибкость человека отражает способности преобразовывать сложные символические

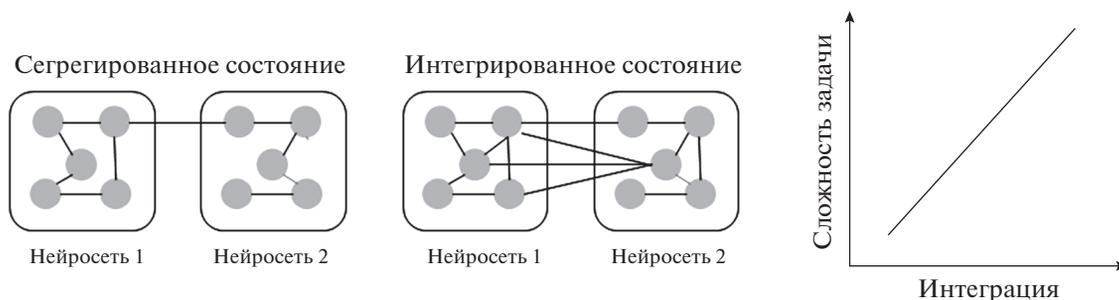


Рис. 5. Примеры изменений топологии сетей и повышения их интеграции в зависимости от сложности задачи.

инструкции в новое поведение. Это преобразование поддерживается двумя нейрокогнитивными состояниями: начальным декларативным знанием о задаче и состоянием реализации, необходимым для оптимального ее выполнения. Решающую роль в этом процессе выполняют лобные и теменные области мозга. Согласно результатам фМРТ в эксперименте с поиском новых согласно инструкции ассоциаций выявлены независимые процедурные и декларативные их репрезентации в лобно-теменных областях. Важно отметить, что степень процедурной активации предсказывала последующие поведенческие характеристики, что свидетельствует о важном вкладе лобно-теменных областей в нейронную архитектуру, которая регулирует когнитивную гибкость [27, 28].

С течением времени мозг переходит через разные состояния, и топология сети мозга, указывающая на интеграцию информации, связана с лучшим выполнением когнитивной задачи [66]. С применением теории графов [16] или кластеризации  $k$ -средних [67] обнаружено два таких состояния: с более обособленной структурой сетей при большей связности внутри каждой сети или с большей связностью между сетями, что способствует интеграции информации в мозге, усиливающейся при повышении сложности задачи (рис. 5).

Сдвиги функционального состояния в значительной степени обусловлены расширением возможностей взаимодействия лобно-теменной, дорсальной, цингуло-оперкулярной и зрительной сетей. Сдвиг организации сети в сторону интеграции чувствителен к требованиям задач: более интегрированная сетевая структура отмечена при повышении их сложности, и степень этой интеграции определяется не только параметром сложности, но и тем, насколько хорошо участники выполняют задачу и связанным с этим уровнем активации мозга [67].

С применением топической модели когнитивных функций выявлено пространственное соответствие между когнитивными функциями и сетью модулей, каждый из которых выполняет дискретную когнитивную функцию. Так как активность в

локальных узлах внутри модулей не увеличивается при выполнении задач, требующих большего количества когнитивных функций, это демонстрирует автономность функций модулей. Однако соединительные узлы, расположенные там, где активность мозга связана со множеством различных когнитивных функций, проявляют повышенную активность, когда в задаче задействовано больше когнитивных функций. Режимы коннектора поддерживают модульные функции мозга, интегрированные при реализации когнитивных функций [50].

Прогресс в разработке систем гибкой реконфигурации модулей, встроенных в когнитивную архитектуру, связан с полученными в последнее время результатами, описывающими механизмы интеграции или сегрегации нейронных сетей мозга в зависимости от сложности задач и необходимости переключения стратегий: от менее к более эффективным. Так, например, фМРТ картирование активности мозга при выполнении задачи Струпа с изменяющимся соотношением конгруэнтных и неконгруэнтных стимулов показало большую реакцию в активации дорсальной сети внимания (Dorsal Attention Network – DAN) и фронто-париетальной сети (Fronto-Parietal Network – FPN) на неконгруэнтные предъявления стимулов, чем на конгруэнтные. По мере увеличения доли неконгруэнтных проб в блоке функциональная связность усиливалась в пределах цингуло-оперкулярной (Cingulo-Opercular Network – CON), FPN, сети по умолчанию (Default Mode Network – DMN), DAN и сети значимости (Salience Network – SN), и сопровождалась усилением связанности между CON и DAN/SN и между DMN и CON/DAN/SN, но связь ослаблялась между FPN и вентральной сетью внимания (Ventral Attention Network – VAN). Усиление связности внутри CON, внутри FPN и между CON и SN было положительно связано с улучшением поведенческих характеристик [44].

Сила когнитивных архитектур в том, что их реализация в виде компьютерных программ обеспечивает достаточно высокую степень точности выполнения изучаемых задач, что дает доказательства их



Рис. 6. Ключевые компоненты CHREST архитектуры.

полезности в проектировании искусственного интеллекта и постановке новых экспериментов в нейрофизиологических исследованиях. Успех, достигнутый при моделировании реального человеческого поведения, можно оценить с помощью таких показателей, как движения глаз, время реакции и количество ошибок. Адекватность соотношения психометрических показателей выполнения разных заданий и регионарной специфики активации соответствующих областей коры показана, например, сопоставлением нейронной архитектуры исполнительных функций и реализации ответной реакции или ее торможения [21].

**Архитектура CHREST** (Chunk Hierarchy and REtrieval Structures – иерархия элементов и структуры извлечения) [23–26, 41] моделирует данные в ряде областей, например, в экспертном поведении в настольных играх, решении проблем в физике, овладение первым языком или имплицитное обучение. CHREST включает структуры кратковременной памяти (STM), долговременной памяти (LTM) и моделирует познание как продукт взаимодействия перцептивного обучения, извлечения памяти и процессов принятия решений. Центральная идея заключается в том, что “познание есть восприятие” [18]. То есть, архитектура постулирует тесное взаимодействие между восприятием, обучением и памятью: знание в модели CHREST направляет внимание и восприятие, а восприятие, в свою очередь, направляет изучение нового знания.

Все познавательные операции, выполняемые системой CHREST, имеют параметры с фиксированными временными параметрами, что позволяет сравнивать человеческое и смоделированное поведение (подробности см. в [18]; техническую спецификацию CHREST можно найти на сайте [www.CHREST.info](http://www.CHREST.info)). Поведение CHREST ограничено объемом кратковременной зрительной памяти (3 фрагмента), относительно медленной скоростью заучивания новых элементов (8 секунд для создания нового фрагмента) и временем, кото-

рое требуется для передачи информации от долговременной к кратковременной памяти (50 мс). Акцент на когнитивных ограничениях контрастирует с такими архитектурами, как SOAR (State, Operator And Result), где упор делается на выполнение сложного интеллектуального поведения без наложения многих ограничений на архитектуру (например, SOAR обладает неограниченным объемом рабочей памяти). Таким образом, по сравнению с другими архитектурами CHREST может показаться очень строгой системой. Однако это мощная динамическая система, управляемая встроенными возможностями CORE Metadata, citation и подобными документами на [core.ac.uk](http://core.ac.uk), что позволяет охватить широкий спектр моделей поведения при взаимодействии с окружающей средой.

Три основных компонента CHREST показаны на рисунке 6: механизмы и структуры взаимодействия с внешним миром (модуль входа); модули кратковременной памяти, которые содержат информацию из различных модальностей ввода; модули долговременной памяти, в которых хранится информация, находящаяся в “сети фрагментации”.

Сеть фрагментации – это дискриминационная сеть, динамический рост которой является совместной функцией предыдущих состояний системы и входных данных из окружающей среды. Каналы визуального ввода-вывода (и связанные с ними элементы кратковременной памяти) были исследованы на моделях шахматного мастерства, где моделировались реальные движения глаз мастеров и новичков [18]. Такой смоделированный глаз имеет решающее значение для понимания взаимодействия между информацией низкого уровня, такой как визуальные стимулы, и познанием высокого уровня, таким как концепции. Слуховые каналы были исследованы в модели приобретения словарного запаса [32], которая имитирует то, как дети учат слова, используя фонематическую информацию.

Процесс, называемый дискриминацией, создает новые “узлы” в сети элементов (chunks), а процесс ознакомления постепенно добавляет информацию к существующим узлам. В системе предусмотрен механизм создания высокоуровневых структур из перцептивной информации с помощью формирования шаблона, для чего используется как стабильная информация (для создания ядра схемы), так и переменная информация (для создания ее слотов). Важным нововведением является создание латеральных связей (связей сходства, производственных связей, связей эквивалентности и порождающих связей) между узлами [24]. CHREST объясняет развитие опыта приобретением большого количества структур знаний и установлением связей между ними.

Решение CHREST состоит из трех частей, каждая из которых ведет к уменьшению объема информации. Во-первых, ограниченная емкость поля зрения исключает значительное количество информации, поступающей из окружающей среды. Во-вторых, знание того, что система приводит в действие еще больше ограничивает объем обрабатываемой информации. Дальнейшее сокращение информации обусловлено ограниченным объемом памяти. Таким образом, CHREST избирательна, как и когнитивная система человека.

В CHREST связь между восприятием и познанием может быть настолько тесна, что различие между этими двумя наборами механизмов практически исчезает. Фокус внимания определяет, какая информация будет усвоена. Затем движения глаз и, следовательно, внимание будут направляться предыдущими знаниями, что повышает вероятность того, что система обратит внимание на те функции, которые были важны в прошлом и приводили к обучению, и, соответственно, должны быть важны в будущем. Этот цикл “восприятие—обучение—восприятие” является еще одним способом, с помощью которого CHREST решает проблему фрейма, поскольку он приводит к избирательности внимания, что еще больше снижает объем информации, извлекаемой из окружающей среды, и позволяет реагировать в режиме реального времени.

Предсказания поведения, сделанные с использованием CHREST, подтверждены эмпирическими данными: например, припоминанием шахматных позиций, которые создаются путем случайной замены фигур игровой позиции на доске, или моделированием эмпирических явлений при раннем усвоении синтаксических категорий [41].

CHREST пригодна для моделирования экспериментов по категоризации. В первых моделях CHREST в основном предлагались отдельные узлы, имитирующие концепции, используемые людьми. Тип сетей фрагментации, созданный в CHREST,

содержит и другие возможности. Естественная интерпретация этих сетей заключается в том, что понятия не отображаются в отдельные фрагменты (или даже в отдельные шаблоны), а скорее соответствуют подмножеству узлов, в той или иной степени связанных между собой латеральными связями. С этой точки зрения концепты гораздо более распределены, чем в стандартных символических моделях, и приобретают черты того, как концепты представлены в естественных нейронных сетях [25]. Преимущество использования CHREST заключается в предоставлении механизмов, объясняющих, как узлы в ИНС и связи между ними приобретаются автономно в режиме реального времени, как они связаны с перцептивным вводом, как на них влияет структура окружающей среды и как они могут интегрировать информацию в различных модальностях.

Практическое применение в психологии и нейрофизиологии когнитивных архитектур, смоделированных в рамках разных подходов, не подразумевает каких-либо ограничений, так как все они в разной степени опираются на те компоненты механизмов когнитивной деятельности, основные закономерности которых были выявлены при изучении познавательной деятельности человека. Постепенно сокращается разрыв между символическими и нейронными сетями, разрабатываются комбинированные архитектуры, сочетающие модули разных когнитивных архитектур (например, Leabra и ACT-R [42]). Цель такого объединения заключается в достижении баланса в требованиях ограничения возможных вариантов человеческого поведения и информационного объема моделей.

Описанные примеры наиболее широко известных моделей когнитивной архитектуры позволяют дать прогнозы эффективности восприятия и селекции информации, какие знания и процедуры требуются для оптимального решения задачи, ожидаемую частоту ошибок при выполнении задания и какая функциональная система мозга используется для организации поведения. Совершенствование и дополнение существующих моделей когнитивной архитектуры следует рассматривать как перспективу развития когнитивной нейронауки, понимания закономерностей формирования и организации нейронных систем естественного интеллекта и наиболее перспективных путей разработки искусственного интеллекта.

## СПИСОК ЛИТЕРАТУРЫ

1. Бондарко В.М., Данилова М.В., Красильников Н.Н., Леушина Л.И., Невская А.А., Шелепин Ю.Е. Пространственное зрение. СПб: Наука, 1999. 224 с.
2. Глезер В.Д. Механизмы опознания зрительных образов. Л: Наука, 1966. 203 с.

3. Глезер В.Д. Зрение и мышление. СПб: Наука, 1993. 248 с.
4. Красильников Н.Н. Передача, прием и восприятие изображений. М.: Радио и связь, 1986. 246 с.
5. Малашин Р.О. Принцип наименьшего действия в динамически конфигурируемых системах анализа изображений // Оптический журнал. 2019. Т. 86. № 11. С. 5–13.
6. Шелепин Ю.Е. Введение в нейроиконику: СПб: Троицкий мост, 2017. 352 с.
7. Anderson J.R. How can the human mind exist in the physical universe? New York, NY: Oxford University Press. 2007.
8. Anderson J.R., Bothell D., Byrne M.D., Douglass S., Lebiere C., Qin Y. An integrated theory of the mind // Psychological Review. 2004. V. 111. № 4. P. 1036–1060
9. Anderson J.R., Farrell R., Sauers R. Learning to program in LISP // Cognitive Science. 1984. V. 8. P. 87–129.
10. Anderson J.R., Lebiere C. The Atomic Components of Thought. Mahwah, NJ: Erlbaum. 1998.
11. Anderson J.R., Reder L.M. The fan effect: New results and new theories // J. Experimental Psychology: General. 1999. V. 128. P. 186–197.
12. Bakaev M., Razumnikova O. What makes a UI Simple? Difficulty and complexity in tasks engaging visual-spatial working memory // Future Internet. 2021. V. 13. № 1. Ar. 21.  
<https://doi.org/10.3390/fi13010021>
13. Borst J.P., Nijboer M., Taatgen N.A., van Rijn H., Anderson J.R. Using data-driven model-brain mappings to constrain formal models of cognition // PLoS One. 2015. V. 10. № 3. Ar. e0119673.  
<https://doi.org/10.1371/journal.pone.0119673>
14. Bourguignon N.J., Braem S., Hartstra E., De Houwer J., Brass M. Encoding of novel verbal instructions for prospective action in the lateral prefrontal cortex: evidence from univariate and multivariate functional magnetic resonance imaging analysis // J. Cogn. Neurosci. 2018. V. 30. P. 1170–1184.
15. Brass M., Liefoghe B., Braem S., De Houwer J. Following new task instructions: Evidence for a dissociation between knowing and doing // Neurosci. Biobehav. Rev. 2017. V. 81. P. 16–28.
16. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems // Nat Rev Neurosci. 2009. V. 10. № 3. P. 186–198.
17. Czerwinski M., Horvitz E., Wilhite S. A diary study of task switching and interruptions // In Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. 2004. P. 175–182.
18. De Groot A.D., Gobet F. Perception and memory in chess: Heuristics of the professional eye. Assen: Van Gorcum. 1996. 346 p.
19. Demanet J., Liefoghe B., Hartstra E., Wenke D., De Houwer J., Brass M. There is more into ‘doing’ than ‘knowing’: the function of the right inferior frontal sulcus is specific for implementing versus memorising verbal instructions // Neuroimage. 2016. V. 141. P. 350–356.
20. Duch W., Oentaryo R., Pasquier M. Cognitive Architectures: Where do we go from here? Frontiers in Artificial Intelligence and Applications Conference: Artificial General Intelligence. 2008. V. 171. P. 122–136.
21. Engelhardt L.E., K. Harden P., Tucker-Drob E.M., Church J.A. The neural architecture of executive functions is established by middle childhood // Neuroimage. 2019. V. 185. P. 479–489.
22. Everaert T., Theeuwes M., Liefoghe B., De Houwer J. Automatic motor activation by mere instruction // Cogn. Affect. Behav. Neurosci. 2014. V. 14. P. 1300–1309.
23. Gobet F. Chunking Models of Expertise: Implications for Education // Appl. Cognit. Psychol. 2005. V. 19. P. 183–204.
24. Gobet F., Lane P.C.R., Croker S., Cheng P.C.H., Jones G., Oliver I., Pine J.M. Chunking mechanisms in human learning // Trends in Cognitive Sciences. 2001. V. 5. P. 236–243.
25. Gobet F., Lane P.C.R., Lloyd-Kelly M. Chunks, Schemata, and Retrieval Structures: Past and Current // Computational Models Front Psychol. 2015. V. 6. Ar. 1785.  
<https://doi.org/10.3389/fpsyg.2015.01785>
26. Gobet F., Simon H.A. Five seconds or sixty? Presentation time in expert memory // Cognitive Science. 2000. V. 24. № 4. P. 651–682.
27. González-García C., Formica S., Wisniewski D., Brass M. Frontoparietal action-oriented codes support novel instruction implementation // NeuroImage. 2021. 226. Ar. 117608
28. González-García C., Formica S., Liefoghe B., Brass M. Attentional prioritization reconfigures novel instructions into action-oriented task sets // Cognition 2020. V. 194. Ar. 104059.  
<https://doi.org/10.1016/j.cognition.2019.104059>
29. González-García C., Mas-Herrero E., de Diego-Balaguer R., Ruz M. Task-specific preparatory neural activations in low-interference contexts // Brain Struct. Funct. 2016. V. 221. № 8. P. 3997–4006.
30. Hartstra E., Waszak F., Brass M. The implementation of verbal instructions: Dissociating motor preparation from the formation of stimulus–response associations // Neuroimage. 2012. V. 63. P. 1143–1153.
31. Harrison J. Handbook of practical logic and automated reasoning. Cambridge University Press, 2009. 1305 p.
32. Jones G., Gobet F., Pine J.M. Linking working memory and long-term memory: a computational model of the learning of new words. Dev Sci. 2007. V. 10. № 6. P. 853–873.
33. Jones R.M., Laird J.E., Nielsen P.E., Coulter K., Kenny P., Koss F. Automated Intelligent Pilots for Combat Flight Simulation // AI Magazine. 1999. V. 20. № 1. P. 27–42.
34. Jones G., Gobet F., Pine J.M. Linking working memory and long-term memory: A computational model of the learning of new words // Developmental Science. 2007. V. 10. № 6. P. 853–873.
35. Jones G., Gobet F., Pine J.M. Computer simulations of developmental change: The contributions of working memory capacity and long-term knowledge // Cognitive Science. 2008. V. 32. № 7. P. 1148–1176.

36. *Katidioti I., Borst J.P., Taatgen N.A.* What happens when we switch tasks: Pupil dilation in multitasking // *J. Exp. Psychol. Appl.* 2014. V. 20. P. 380–396
37. *Kahneman D.A.* perspective on judgment and choice: mapping bounded rationality // *American psychologist.* 2003. V. 58. № 9. P. 697–720.
38. *Kim N.Y., House R., Yun M.H., Nam C.S.* Neural Correlates of Workload Transition in Multitasking: An ACT-R Model of Hysteresis Effect // *Front Hum Neurosci.* 2019. V. 12. Ar. 535.
39. *Kotseruba I., Tsotsos J.K.* A review of 40 years of cognitive architecture research: Focus on perception, attention, learning and applications // *arXiv preprint: 1610.08602.* 2016. <https://doi.org/10.48550/arXiv.1610.08602>
40. *Lane P.C.R., Gobet F., Smith R.L.* Attention Mechanisms in the CHREST Cognitive Architecture. In: *Palletta L., Tsotsos J.K.* (eds) *Attention in Cognitive Systems.* WAPCV 2008. Lecture Notes in Computer Science. 2009. V. 5395. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-00582-4\\_14](https://doi.org/10.1007/978-3-642-00582-4_14)
41. *Lane P.C.R., Gobet F.* Perception in chess and beyond: Commentary on Linhares and Freitas (2010)' // *New Ideas in Psychology.* 2011. V. 29. P. 156–161.
42. *Lebiere C., O'Reilly R., Jilk D., Taatgen N.A., Anderson J.R.* The SAL integrated cognitive architecture. In A. Samsonovich (Ed.). *Biologically inspired cognitive architectures: Papers from the AAAI 2008 Fall Symposium* Menlo Park, CA: AAAI Press. 2009. P. 98–104.
43. *Lehman J.F., Laird J.E., Rosenbloom P.A.* Gentle Introduction to Soar, An Architecture for Human Cognition // *An Invitation to Cognitive Science.* MIT Press. 2006. V. 4.
44. *Li Y., Wang Y., Yu F., Chen A.* Large-scale reconfiguration of connectivity patterns among attentional networks during context-dependent adjustment of cognitive control // *Hum Brain Mapp.* 2021. V. 42. P. 3821–3832.
45. *Malakhova E.Yu.* Information representation space in artificial and biological neural networks // *J. Optical Technology.* V. 87. № 10. P. 598–603.
46. *Malashin R.* Principle of least action in dynamically configured image analysis systems // *J. Optical Technology.* 2019. V. 86. № 11. P. 678–685.
47. *Malashin R.O.* Training an improved recurrent attention model using an alternative reward function // *J. Optical Technology.* 2021. V. 88. № 3. P. 127–130.
48. *Malashin D.O., Malashin R.O.* Efficient hardware implementation of neural networks // *Neural Networks and Neurotechnologies.* *Aleksandrov A.A., Alizade M.R., Andreeva G.O., Andreeva I.G.,* (Eds). St. Petersburg, 2019. P. 187–192.
49. *Martin L., Rosales J.H., Jaime K., Ramos F.* Affective episodic memory system for virtual creatures: The first step of emotion-oriented memory // *Comput Intell Neurosci.* 2021. Ar. 7954140. <https://doi.org/10.1155/2021/7954140>
50. *Maxwell A., Thomas B.V.T., Yeo and M. D'Esposito.* The modular and integrative functional architecture of the human brain // *PNAS.* 2015. Ar. E6798–E6807. <https://doi.org/10.1073/pnas.1510619112>
51. *Mill R.D., Ito T., Cole M. W.* From connectome to cognition: The search for mechanism in human functional brain networks // *Neuroimage.* 2017. V. 160. P. 124–139.
52. *Mishkin M., Ungerleider L.G., Macko K.A.* Object vision and spatial vision: two cortical pathways // *Trends in Neurosci.* 1983. V. 6. P. 414–417.
53. *Moreira J.F.G., McLaughlin K.A., Silvers J.A.* Characterizing the Network Architecture of Emotion Regulation Neurodevelopment // *Cerebral Cortex.* 2021. V. 31. P. 4140–4150.
54. *Muhle-Karbe P.S., Duncan J., Baene W.De, Mitchell D.J., Brass M.* Neural coding for instruction-based task sets in human frontoparietal and visual cortex // *Cereb. Cortex.* 2017. V. 27. P. 1891–1905.
55. *Newell A., Simon H.* GPS, A Program that Simulates Human Thought. In: *Computers and Thought,* *E.A. Feigenbaum, J. Feldman* (Eds.), R. Oldenbourg KG., 1963. P. 109–124.
56. *Neves D.M., Anderson J.R.* Knowledge compilation: Mechanisms for the automatization of cognitive skills. In: *J. R. Anderson* (Ed.) *Cognitive skills and their acquisition* Hillsdale, NJ: Erlbaum. 1981. P. 57–84.
57. *Oh H., Yun Y., Myung R.* Cognitive Modeling of Task Switching in Discretionary Multitasking Based on the ACT-R Cognitive Architecture // *Appl. Sci.* 2021. V. 11. Ar. 3967. <https://doi.org/10.3390/app11093967>
58. *Palenciano A.F., González-García C., Arco J.E., Pessoa L., Ruz M.* Representational organization of novel task sets during proactive encoding // *J. Neurosci.* 2019. V. 39. № 42. P. 8386–8397. <https://doi.org/10.1523/JNEUROSCI.0725-19.2019>
59. *Palenciano A.F., González-García C., Arco J.E., Ruz M.* Transient and sustained control mechanisms supporting novel instructed behavior // *Cereb. Cortex.* 2019. V. 29. P. 3948–3960.
60. *Pearl J.* Probabilistic reasoning in intelligent systems: networks of plausible inference. 1988. Elsevier Inc. 152 p.
61. *Prezenski S., Brechmann A., Wolff S., Russwinkel N.A.* Cognitive Modeling Approach to Strategy Formation in Dynamic Decision Making // *Front Psychol.* 2017. V. 8. Ar.1335. <https://doi.org/10.3389/fpsyg.2017.01335>
62. *Razumnikova O., Bakaev M.* Ontology of frequency-spatial organization of brain activity reflecting the cognitive reserves // *International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON).* 2019. P. 0950–0954.
63. *Richards J.M., Gross J.J.* Emotion regulation and memory: the cognitive costs of keeping one's cool // *J. Pers Soc. Psychol.* 2000. V. 79. № 3. P. 410–424.
64. *Ritter F.E. Tehranchi F., Oury J.D.* ACT-R: A cognitive architecture for modeling cognition WIREs // *Cogn Sci.* 2018. Ar. e1488 <https://doi.org/10.1002/wcs.1488>
65. *Rumelhart D., McClelland J.* Parallel distributed processing: Psychological and biological models. The MIT press. 1986. V. 2. 581 p.
66. *Schultz D.H., Cole M.W.* Integrated Brain Network Architecture Supports Cognitive Task Performance // *Neuron.* 2016. V. 92. P. 278–279.

67. *Shine J.M., Bissett P.G., Bell P.T., Koyejo O., Balsters J.H., Gorgolewski K.J., Moodie C.A., Poldrack R.A.* The dynamics of functional brain networks: Integrated network states during cognitive task performance // *Neuron*. 2016. V. 92. № 2. P. 544–554.
68. *Sun R.* Anatomy of the Mind: Exploring Psychological Mechanisms and Processes with the Clarion Cognitive Architecture. Oxford University Press. 2016. 481 p.
69. *Sun R.* The importance of cognitive architectures: An analysis based on clarion // *J. Experimental & Theoretical Artificial Intelligence*. 2007. V. 19. № 2. P. 159–193
70. *Slovic P.* Affect, Reason, Risk and Rationality // *Notas Económicas*. 2018. № 46. P. 1–10.
71. *Strannegård C., von Haugwitz R., Wessberg J., Balkenius C.A.* Cognitive Architecture Based on Dual Process Theory. In: *Kühnberger K.U., Rudolph S., Wang P.* (eds) *Artificial General Intelligence. AGI 2013. Lecture Notes in Computer Science*. 2013. V. 7999. Springer, Berlin, Heidelberg. Doi: [https://doi.org/10.1007/978-3-642-39521-5\\_15](https://doi.org/10.1007/978-3-642-39521-5_15)
72. *Turner B.M., Forstmann B.U., Love B.C., Palmeri T.J., Van Maanen L.* Approaches to Analysis in Model-based Cognitive Neuroscience // *J Math Psychol*. 2017. V. 76 (B). P. 65–79.
73. *van Vugt M.K.* Cognitive architectures as a tool for investigating the role of oscillatory power and coherence in cognition // *Neuroimage*. 2014. V. 85. Pt 2. P. 685–693.
74. *Wickens C.D., Gutzwiller R.S., Santamaria A.* Discrete task switching in overload: A meta-analysis and a model // *Int. J. Hum. Comp. Stud*. 2015. V. 79. P. 79–84.
75. *Woolgar A., Afshar S., Williams M.A., Rich A.N.* Flexible coding of task rules in frontoparietal cortex: an adaptive system for flexible cognitive control // *J. Cogn. Neurosci*. 2015. V. 27. P. 1895–1911.

## Cognitive Architecture of Cognitive Activity: Modeling and Psychophysiological Assessment

O. M. Razumnikova\*

*Novosibirsk State Technical University, Novosibirsk, 630073 Russia*

*\*e-mail: razoum@mail.ru*

**Abstract**—The main approaches to modeling human cognitive activity and the underlying neural mechanisms are described. The systematization of cognitive architectures is given, and such popular models as ACT-R, SOAR, CLARION and CHREST is overviewed with examples of their practical application in psychology and neurophysiology. The use of the developed models of cognitive functions makes it possible to predict the effectiveness of perception and selection of information, which knowledge and procedures are required for the optimal solution of the problem, the expected error rate while task performing, and what functional brain system is used to organize behavior. Improvement and addition of existing models of cognitive architecture is considered as a prospect for the development of cognitive neuroscience, understanding the patterns of intelligence formation and the development of artificial intelligence.

**Keywords:** cognitive architecture, perception, memory, decision making, emotional regulation, artificial intelligence