

Методология коннекционизма и проблема имитации ментальных феноменов искусственными нейронными сетями¹

В искусственном интеллекте существуют два основных подхода к имитации ментальных феноменов: структурный и функциональный. Структурный подход наиболее ярко представлен коннекционистской² методологией, связанной с использованием искусственных нейросетей для динамической имитации когнитивных процессов. Основопологающей для коннекционистской методологии является идея моделирования нейронов довольно простыми автоматами. В рамках этой модели вся сложность мозга и гибкость его функционирования определяется связями между нейронами. При этом постулируется, что широкие возможности систем связей компенсируют упрощенность элементной базы и повышают надежность системы. Коннекционистская парадигма, позволяя перебросить концептуальный мост между микроскопическим уровнем нейронов мозга и макроскопическим феноменом сознания, предоставляет исследователям многообещающий инструментарий для развития когнитивных наук.

Философско-методологические основания традиционного подхода к моделированию ментальных феноменов берут свое начало в направлении функционализма, игнорирующем принципы физической реализации функции и выдвигающем на первый план функциональные отношения в системе. Целью функционального моделирования выступает достижение одинаковой реакции компьютерной модели и человека, в то время как моделирование механизмов или алгоритмов мышления отодвигается на второй план. При этом критерий выбора применяемых для достижения цели методов и моделей формулируется, исходя из сравнения результатов функционального моделирования с результатами решения тех же интеллектуальных задач человеком. Классическим примером данного критерия может выступить тест Тьюринга на лингвистическую компетентность.

В свою очередь, в коннекционизме представлена идеология бионической парадигмы, объединяющей в своих рамках, наряду с нейронными сетями, методы эволюционного моделирования (например, генетические алгоритмы, клеточные автоматы, фрактальные структуры) и мобоитизм. Во главу угла коннекционистской методологии поставлен *принцип структурной доминанты*, согласно которому функция системы является вторичной и определяется ее структурой. При этом реализация в технических устройствах таких способностей мозга, как адаптивность, обучаемость, ассоциативность, самоорганизация, большая вычислительная мощность, надежность функционирования и т.п., предполагает имитацию структуры, а также принципов представления и обработки информации мозгом. Таким образом, исходной целью коннекционизма выступает как можно более точная имитация структуры, принципов представления и обработки информации мозгом для достижения когнитивных способностей, демонстрируемых биологическим прототипом. Критерием выбора применяемых для достижения цели методов и моделей является, в конечном счете, их нейрофизиологическая и биологическая адекватность.

¹ Работа выполнена при поддержке БРФФИ, договор № G09M-051 от 15.04.2009.

² Термин «коннекционизм» является производным от английского «connection» — связь, соединение.

Направление функционального моделирования в традиционном искусственном интеллекте дополняется так называемой компьютерной метафорой, интерпретирующей когнитивные процессы с позиций критерия эффективной вычислимости. В данном контексте познавательные процессы представимы в виде совокупности эффективно вычислимых функций, т.е. функций, для которых существуют алгоритмы, вычисляющие их за конечное число шагов, используя, скажем, универсальную машину Тьюринга. Таким образом, компьютерная метафора состоит в проведении аналогии между познавательными процессами и переработкой информации в универсальном вычислительном устройстве. В рамках данной аналогии познание понимается как обработка информации с целью создания внутренних моделей мира и их преобразования в соответствии с требованиями конкретной задачи.

В свою очередь, в коннекционистской методологии при трактовке природы познавательных процессов имеет место другая расстановка акцентов: вместо компьютерной метафоры используется *метафора работы мозга* как ориентир для создания нейроморфных вычислительных устройств. Коннекционизм заменяет вычислительно-репрезентационный подход к познавательным процессам, свойственный функциональному моделированию, на динамически вычислительный. Таким образом, манипулированию символами и принципу эффективной вычислимости предложена альтернатива: нейросетевая динамическая обработка информации на основе принципа эмергентной вычислимости.

Одно из фундаментальных положений функционального моделирования выражается в виде требования максимальной формализации и точного представления знаний на основе логических и синтаксических манипуляций над символическими репрезентациями. При этом постулируется принципиальная возможность моделирования интеллектуального поведения через использование логических и синтаксических манипуляций над символами. Чаще всего для моделирования познавательных процессов используются предикативные модели представления знаний. В этом случае стандартная модель познавательного процесса представляет собой логическое манипулирование внутренними репрезентациями, в качестве которых могут выступать фреймы, планы и сценарии и т.п. Таким образом, внутренние репрезентации являются символическими структурами, которые характеризуются семантической значимостью, а познавательные процессы трактуются как синтаксические манипуляции над данными структурами.

В свою очередь, коннекционизм предлагает новый подход к воспроизведению механизмов представления информации в мозге, разрабатывая *распределенную версию ментальных репрезентаций*. В ее рамках формирование внутренних репрезентаций связывается с распределенными нейросетевыми паттернами активности и их динамикой. Композициональному характеру семантики классицистов коннекционисты противопоставляют семантически не значимый характер отдельных нейронов и отсутствие семантически значимых составляющих в распределенных паттернах нейронной активности. Таким образом, коннекционизм предлагает новое видение репрезентационных процессов, отказываясь от классической трактовки ментальных репрезентаций как символических в пользу субсимвольного, распределенного варианта репрезентаций и заменяя статические, логические операции с квазилингвистическими последовательностями на динамическую, чувствительную к структуре обработку информации.

На аппаратном уровне функциональное моделирование представлено фон неймановской компьютерной архитектурой с использованием последовательной и алгоритмизированной обработки информации. К числу определяющих характеристик фон неймановской компьютерной архитектуры относятся: необходимость программной алгоритмизации предписываемых вычислительной машине дей-

ствий, использование четко определенного формата данных, невысокая надежность функционирования, адресное кодирование данных в памяти вычислительной машины и т.п.

В свою очередь, аппаратное обеспечение нейросетевого моделирования характеризуется высокой степенью эффективного параллелизма вследствие использования *нейронной параллельной архитектуры*. Одна из определяющих характеристик данной архитектуры — отсутствие программной алгоритмизации предписываемых нейронной сети действий вследствие наличия процедуры обучения, позволяющей сети адаптироваться к данным обучающей выборки. Для нейросетевого моделирования познавательных процессов характерна также аналоговая распределенная обработка информации массивно параллельным образом при высокой взаимосвязанности элементов. Кроме того, на смену адресному кодированию данных приходит широкое использование структур ассоциативной памяти с адресацией по содержанию. Следует также отметить, что, в противоположность фон неймановской компьютерной архитектуре, нейросетевая параллельная архитектура характеризуется повышенной надежностью и устойчивостью функционирования вследствие возможности структурной самоорганизации нейронной сети.

Проблема имитации структурной и функциональной организации мозга в коннекционисткой методологии.

На сегодняшний день коннекционистская методология ориентируется на разработку аналогии в области механизмов представления и обработки информации между моделями когнитивных процессов и их прототипами. В частности, в коннекционистских моделях учтены следующие принципы и механизмы представления и обработки информации, свойственные информатике мозга: малая глубина информационных процессов в сочетании с их высокой параллельностью, совместное использование систем параллельной и последовательной обработки данных, распределенный характер хранения информации в мозге, аналоговые представления и обработка информации (определяющие помехоустойчивость и надежность нейронных структур мозга, а также размытый, неточный характер хранимой информации), наличие структур ассоциативной памяти с адресацией по содержанию, принцип контекстности хранения информации, наличие механизмов обнаружения сходства, самоорганизация — и ряд других свойств [1].

В то же время, исследователи часто указывают на недостаточную степень воспроизведения особенностей структурной и функциональной организации мозга в коннекционистских моделях. Основной мишенью критиков выступает чрезмерная упрощенность данных моделей, зачастую игнорирующих нейрофизиологические данные, касающиеся структурной и функциональной организации мозга [2].

В частности, ряд исследователей придерживается того мнения, что важные особенности реального *нейрона* как элемента структурной организации мозга и функциональной единицы нейроинформационной обработки не отражены в достаточной степени в коннекционистской методологии, использующей слишком упрощенную схему функционирования искусственного нейрона Мак Каллока-Питтса. В числе приводимых ими аргументов указываются следующие: в данной модели не учитываются задержки во времени, оказывающие значительное влияние на динамику биологического прототипа; за рамки модели вынесена синхронизирующая функция биологического нейрона и т.д. [3]

По мнению критиков, слишком упрощенная модель нейрона дополняется также малой рекурсивностью нейросетевых моделей когнитивных процессов; биологически маловероятными комбинациями возбуждающих и ослабляющих связей; нереалистичными обучающими алгоритмами, обособленными от общего осмысления работы нервной клетки и часто игнорирующими важную роль молекулярных механизмов в биологическом обучении. Кроме того, коннекционистские модели ког-

нитивных процессов не имеют аналогов таким элементам, как нейропередатчики и гормоны, оказывающим значительное влияние на процесс познания [4].

Соглашаясь с аргументами критиков, отметим, что, пытаясь воплотить в нейросетевой модели все морфологические, структурные и функциональные особенности отдельного нейрона и т.п., мы рискуем столкнуться с проблемой чрезвычайной сложности модели и трудностью ее технической реализации. В этой связи, необходимо провести достаточно четкое разграничение существенных и несущественных для моделирования элементов в работе биологической нейронной сети. Решение данной задачи, в свою очередь, предполагает наличие достаточно четкой модели работы мозга, скорректированной с учетом возможностей, предоставляемых современными технологиями для моделирования.

И здесь мы сталкиваемся со значительными трудностями. Дело в том, что на сегодняшний день существует относительно непротиворечивая система взглядов о принципах переработки информации в нейронных структурах, но говорить о наличии единой теории мозга, на основе которой можно было бы установить четкие приоритеты для построения корректных нейросетевых моделей когнитивных процессов, пока преждевременно.

Следовательно, слишком упрощенный подход к моделированию нервной ткани, в котором часто упрекают коннекционистскую методологию, свидетельствует, прежде всего, об упрощенности нейробиологических представлений: современные нейросетевые модели проектируются исходя из нейробиологических представлений о работе мозга, и потому ограничения, присущие нейробиологическим моделям, автоматически наследуются и коннекционистскими моделями [5]. В этой связи, нейросетевое моделирование заинтересовано как в дальнейшем развитии нейробиологических представлений о работе мозга, так и в осмыслении достижений нейробиологии с позиций их продуктивности для развития модельных представлений.

При рассмотрении проблемы имитации структурной и функциональной организации мозга в коннекционистском моделировании большое значение имеет вопрос об *элементарных структурных и функциональных единицах организации мозга*.

Изначальной догмой нейронауки, нашедшей свое воплощение и в коннекционистских моделях когнитивных процессов, явилось постулирование нейрона в качестве элементарной единицы структурной организации мозга и, одновременно, в качестве функциональной единицы нейроинформационной обработки. На сегодняшний день данная догма все чаще подвергается сомнению со стороны ряда исследователей, предлагающих других претендентов на роль структурной и функциональной единицы нейроинформационной обработки.

Например, в качестве исходной функциональной единицы сегодня все чаще рассматривается не отдельный нейрон, а скоординированная во времени и пространстве, функционально однородная популяция нейронов¹. Постулируется, что популяционные взаимодействия в нейрофизиологических системах обработки информации с большей вероятностью, чем взаимодействия между отдельными нейронами, отражают характерные свойства системы и ее поведения.

Если сторонники популяционного кодирования в поисках нового претендента на роль функциональной единицы нейроинформационной обработки пытаются выйти вовне, за пределы отдельного нейрона, то сторонники *мультикеллюлярной модели нейрона* обращаются вглубь структуры отдельного нейрона. В частности, А.В. Савельев отмечает: «Огромная сложность реальных нейронов и широчайший диапазон вариаций их свойств, морфологических особенностей, физических и физиологиче-

¹ Например, в рамках селекционистской теории Дж. Эделмена и М. Маунткласа, в качестве функциональных элементов нейроинформационной обработки выступают нейронные колонки — скоординированные во времени и пространстве функционально-однородные группы нейронов.

ских параметров не укладывается в рамки уницеллюлярной модели» [6], — трактующей нейрон как самостоятельную отдельную клетку и элементарную единицу структурной организации мозга. Признание мультицеллюлярной структуры некоторых высокоорганизованных нейронов позволяет объяснить ряд феноменов, выходящих за рамки уницеллюлярной модели. Например, мультицеллюлярная модель нейрона позволяет систематизировать и упорядочить текущие морфодинамические перестройки, а также открывает возможности для развития новой «морфологической парадигмы» обучения нейронных сетей, связанной с появлением или исчезновением ряда структурных элементов нейронной сети.

Расходясь в вопросе о элементарных структурных и функциональных единицах организации мозга, а также в определении существенных для нейросетевого моделирования особенностей функционирования биологического прототипа, исследователи сходятся в том, что представления о структуре и механизмах переработки информации, лежащие в основе нейросетевого моделирования, должны быть пересмотрены.

Как нам представляется, дальнейшее развитие коннекционистской методологии с данных позиций предполагает непосредственное взаимодействие двух составляющих. Одна из них связана с дальнейшим изучением особенностей моделируемого прототипа и, прежде всего, с построением в рамках нейронаук единой теории мозга; в то время как вторая составляющая, связанная с будущим развитием вычислительных технологий, предоставляет расширенные возможности для воплощения первой составляющей в нейросетевом моделировании.

Необходимость создания и разработки нейрофизиологических моделей обуславливает взаиморекурсивные связи нейроинформатики и нейрофизиологии. В самом деле, переход нейрофизиологии на более высокий уровень понимания механизмов работы мозга является значимым также и для создания новой интеллектуальной вычислительной техники, функционирование которой будет подчинено биологическим принципам.

В свою очередь, развитие вычислительных технологий предоставляет возможность расширить рамки коннекционистского моделирования и более полно воссоздать в нейросетевой модели особенности моделируемого прототипа, в том числе и за счет сочетания коннекционистской и междисциплинарной методологий. Возможным вариантом подобной гибридизации может выступить использование в нейросетевом моделировании идей и методов, наработанных в рамках направления синергетики и диссипативного хаоса, теории фрактальных структур и т.п.

Согласно другой точке зрения, расширение коннекционистской методологии можно осуществить за счет использования математического аппарата нейронных полей, позволяющего объединить возможности коннекционистской и квантовой методологий [7]. Эвристический потенциал квантовой эпистемологии для коннекционизма связан с возможностью репрезентации семантического аспекта информации на основе нейро-квантовой когерентности. Кроме того, смещение акцента с нейронного на нейроквантовый уровень моделирования позволяет учесть субклеточный, квантовый и виртуальный уровни иерархической структуры мозга.

Итак, дальнейшее развитие коннекционизма в данном направлении послужит более тесному сближению биологических и технологических парадигм, а также формированию и развитию междисциплинарного языка, общего для нейрофизиологов и кибернетиков.

Литература

1. *Величковский Б.М., Зинченко В.П.* Методологические проблемы современной когнитивной психологии // Вопросы философии. 1979. № 7. С. 67—79; *Кругликов Р.И.* О структуре функциональной организации головного мозга в процессах памяти // Вопросы философии. 1978. № 1. С. 96—107.

2. *Koch C.* Methods in neuronal modeling: From ions to networks / Ed. C.Koch. Cambridge, MA: MIT Press, 1998.
3. *Hertz J., Krogh A., Palmer R.* Introduction to the Theory of Neural Computation. — Addison Wesley Publishing Company, 1991. 327 p.; *Dudai Y.* The Neurobiology of Memory: Concepts, Findings, Trends. Oxford: Oxford University Press, 1989.
4. *Либерман Е.А.* Молекулярная вычислительная машина клетки: Общие соображения и гипотезы // Биофизика. 1972. Т. 17, № 5. С. 932—943.
5. *Савельев А.В.* Философия методологии нейромоделирования: Смысл и перспективы // Философия науки. 2003. № 1. С. 50—51.
6. *Савельев А.В.* Модель нейрона как возможная мультиклеточная структура: (К вопросу о том, что все-таки мы моделируем) // Нейрокомпьютеры: разработка, применение. 2002. № 1—2. С. 18—19.
7. *Linkevich A.D.* Thinking, Language, Self and Attractors: Cognitive Studies and Synergetic Paradigm // Nonlinear Phenomena in Complex Systems. 1999. Vol. 2, № 1. P. 24—40; *Perus M.* Multi-Level Synergetic Computation in Brain // Nonlinear Phenomena in Complex Systems. 2001. Vol. 4, No 2. P. 157—193.