



НЕЙРОННЫЕ СЕТИ КАК МОДЕЛЬ ЧЕЛОВЕЧЕСКОГО СОЗНАНИЯ: ОТ ФУНКЦИОНАЛЬНОГО СХОДСТВА К АРХИТЕКТУРНЫМ ПАРАЛЛЕЛЯМ

Кузнецов Владимир Николаевич

Студент кафедры сравнительной грамматики, Московский государственный
университет
г. Москва Россия

Аннотация

Данная статья посвящена междисциплинарному анализу возможностей искусственных нейронных сетей (ИНС) в качестве вычислительной и теоретической модели для понимания механизмов человеческого сознания и познания. Рассматривается историческая и методологическая преемственность от коннекционизма 1980-х годов до современных архитектур глубокого обучения (Deep Learning). Центральное внимание уделяется принципам сходства между биологическими нейронными сетями (БНС) и ИНС, включая параллельную обработку информации, распределенное представление знаний и способность к обучению на основе опыта через изменение весов (синаптическая пластичность). Анализируются конкретные архитектурные параллели, такие как использование сверточных нейронных сетей (CNN) в моделях зрительной коры и рекуррентных сетей (RNN/LSTM) для моделирования языковой и временной обработки информации. Обсуждаются фундаментальные ограничения ИНС как модели сознания, включая проблему квалиа (субъективного опыта) и проблему нехватки общих знаний (Common Sense). Сделан вывод о том, что, несмотря на функциональный успех и архитектурное сходство, современные ИНС остаются мощным инструментом для моделирования отдельных когнитивных процессов, но не могут претендовать на полное воспроизведение или объяснение феномена человеческого сознания без преодоления фундаментального разрыва между вычислениями и субъективным опытом.

Ключевые слова: Искусственные нейронные сети (ИНС), Модель сознания, Коннекционизм, Глубокое обучение, Биологические нейронные сети, Параллельная обработка, Распределенное представление, Квалиа, Зрительная кора, Когнитивные науки.

Введение

Вопрос о природе человеческого сознания и его вычислительной основе является одной из величайших загадок науки.

Начиная с середины XX века, вычислительная парадигма рассматривает мозг как сложную систему обработки информации, а познание — как результат вычислений. В рамках этой парадигмы искусственные нейронные сети (ИНС), вдохновленные архитектурой и функциями биологического мозга, стали не только наиболее успешным инструментом в области искусственного интеллекта (ИИ), но и мощной теоретической моделью для нейронаук и когнитивной психологии. ИНС, и особенно их современные многослойные формы (глубокое обучение), предлагают правдоподобный механизм, объясняющий, как простые вычислительные элементы (нейроны) могут объединяться для решения сложных задач восприятия, классификации и памяти.

Коннекционистская гипотеза, лежащая в основе ИНС, утверждает, что знания в мозге представлены не символически, а распределенно, в виде паттернов активности множества связей (весов). Эта идея кардинально отличается от классического, символического ИИ. Современные достижения, такие как способность глубоких сетей к распознаванию изображений, генерации текста и принятию решений, в некоторых задачах уже превосходят человека, что ставит вопрос: являются ли эти системы просто инструментами, или они действительно моделируют некоторые фундаментальные аспекты человеческого сознания? Цель данной статьи — провести сравнительный анализ принципов функционирования ИНС и биологических нейронных сетей (БНС), выявить архитектурные и функциональные параллели, а также критически оценить ограничения ИНС как полноценной модели человеческого сознания, особенно в свете проблемы субъективного опыта.

Функциональные и Методологические Параллели

Сходство между ИНС и БНС лежит не столько в деталях реализации каждого отдельного "нейрона", сколько в общих принципах архитектуры и обучения.

Параллельная и Распределенная Обработка: ИНС, как и биологический мозг, являются системами параллельной и распределенной обработки. В отличие от последовательного выполнения инструкций классическими компьютерами, ИНС обрабатывают входные данные (например, пиксели изображения, слова текста) одновременно через множество узлов. В БНС знания, воспоминания и концепции не хранятся в одном месте, а представлены в виде уникальных паттернов активации тысяч взаимосвязанных нейронов. Этот принцип распределенного представления обеспечивает системе высокую надежность и устойчивость к локальным повреждениям (*gradual degradation*), что является характерной чертой как биологического, так и эффективно спроектированного искусственного интеллекта.

Обучение на Основе Опыта (Синаптическая Пластичность): Ключевой принцип обучения в ИНС — это механизм обратного распространения ошибки (*Backpropagation*), который корректирует весовые коэффициенты (силы связей) между нейронами на основе разницы между желаемым и фактическим

результатом. Этот механизм является вычислительной аналогией **синаптической пластичности** — фундаментального биологического процесса, при котором прочность синаптических связей между нейронами изменяется в зависимости от их совместной активности (правило Хебба). Таким образом, как в БНС, так и в ИНС, обучение не является "программированием", а происходит через **самоорганизацию и модификацию связей** в ответ на внешний опыт, что позволяет системе адаптироваться и приобретать новые навыки.

Архитектурные Параллели и Моделирование Когнитивных Функций

Современные архитектуры глубокого обучения были разработаны с учетом конкретных особенностей структуры мозга, ответственных за специфические когнитивные функции.

Сверточные Нейронные Сети (CNN) и Зрительная Кора: Архитектура сверточных сетей напрямую имитирует иерархическую организацию зрительной системы млекопитающих. CNN состоят из слоев, которые выполняют свертку (фильтрацию) входного изображения. Первый слой (аналогично первичной зрительной коре V1) обнаруживает простые признаки, такие как края и углы. Последующие, более глубокие слои (аналогично V2, V4 и нижневисочной коре) постепенно объединяют эти простые признаки для распознавания все более сложных объектов и паттернов (лиц, предметов). Таким образом, CNN не просто успешно распознают изображения, но и делают это с помощью вычислительной архитектуры, которая демонстрирует поразительное сходство с принципами функционирования биологической зрительной системы, включая феномен рецептивных полей (локальные области обработки информации).

Рекуррентные Сети (RNN, LSTM) и Обработка Последовательностей: Рекуррентные нейронные сети, и их более совершенные варианты — сети с долгой краткосрочной памятью (LSTM) и Трансформеры — используются для моделирования функций, требующих обработки временных последовательностей, таких как речь, язык и принятие решений, основанных на контексте. Эти сети содержат обратные связи (петли), позволяющие сохранять и передавать информацию из предыдущих шагов в текущий, что является вычислительным аналогом рабочей памяти и способности мозга интегрировать информацию во времени. В частности, способность Трансформеров к механизму внимания (Attention Mechanism), позволяющему системе взвешивать важность различных частей входной последовательности, отражает нейробиологические процессы, лежащие в основе селективного внимания в мозге.

Фундаментальные Ограничения и Проблема Сознания

Несмотря на функциональное сходство и архитектурные параллели, ИНС сталкиваются с критическими философскими и вычислительными ограничениями, которые ставят под вопрос их способность моделировать полное человеческое сознание.

Проблема Квалиа и Субъективного Опыта: Самое серьезное ограничение — это проблема квалиа. Квалиа — это субъективные, феноменальные качества опыта (например, то, *как* чувствуется красный цвет, *как* звучит музыка). Современные ИНС могут выполнять сложные вычисления, распознавать цвет и классифицировать звук, но они не имеют субъективного "чувства" или внутреннего переживания этого процесса. ИНС — это лишь алгоритмы, и они не способны объяснить, почему обработка информации должна сопровождаться субъективным опытом. Дэвид Чалмерс называет это "трудной проблемой сознания", и современные ИНС, будучи моделями *функций* сознания, не дают ответа на вопрос о его *феноменологической* сущности.

Проблема Общих Знаний (Common Sense) и Обобщение: В отличие от людей, которые могут обучаться, используя минимальное количество примеров (обучение "одним выстрелом" — One-Shot Learning), ИНС требуют колоссальных объемов размеченных данных для достижения надежного уровня производительности. ИНС часто демонстрируют хрупкость и неспособность к обобщению за пределы узкого распределения обучающих данных. Человеческое сознание обладает общими знаниями (Common Sense) — интуитивным пониманием физического и социального мира, которое не является результатом явного обучения, а формируется через естественный, мультисенсорный опыт. Этот фундаментальный недостаток способности к широкому, интуитивному обобщению и пониманию причинно-следственных связей демонстрирует, что современные ИНС, по сути, являются высокоэффективными системами сопоставления паттернов, а не истинными моделями разума.

Заключение

Искусственные нейронные сети представляют собой наиболее убедительную и функционально эффективную вычислительную модель, которую наука на сегодняшний день имеет для изучения человеческого познания. Функциональные принципы параллельной обработки, распределенного представления и самоорганизующегося обучения через изменение весов демонстрируют глубокие параллели с работой биологического мозга. Архитектуры, такие как CNN и RNN, стали мощными инструментами для моделирования конкретных когнитивных доменов, например, зрительного восприятия и обработки языка. Однако, несмотря на их успехи, современные ИНС остаются мощным инструментом для моделирования функций, но не полноценной моделью феномена сознания. Преодоление "трудной проблемы сознания" и реализация истинной способности к обобщению (Common Sense) потребует разработки принципиально новых вычислительных парадигм, которые, возможно, будут основаны не только на коннекционистских принципах, но и на глубоком понимании биологических механизмов самоорганизации, метапознания и субъективного переживания.

Литература

1. Hassabis, D., Kumaran, D., Summerfield, C., Botvinick, M. *The neuroscience of artificial intelligence*. Neuron, 2017.
2. Rumelhart, D. E., McClelland, J. L. *Parallel distributed processing: Explorations in the microstructure of cognition*. MIT Press, 1986.
3. Koch, C. *The feeling of life itself: Why consciousness is widespread but can't be computed*. MIT Press, 2019.
4. Chalmers, D. J. *The conscious mind: In search of a fundamental theory*. Oxford University Press, 1996.
5. LeCun, Y., Bengio, Y., Hinton, G. *Deep learning*. Nature, 2015.
6. Черниговская, Т. В. *Человеческое сознание: нейробиологические и лингвистические аспекты*. Наука, 2023.
7. Дубровский, В. И. *Проблема сознания в современной философии и науке*. Канон+, 2022.