

Нейродинамическая система как одна из основ методологии моделирования ментальных ресурсов человека¹

Горбунов Иван Анатольевич

Санкт-Петербургский государственный университет, Санкт-Петербург

В современной психофизиологии часто используется методология математических моделей нейронных сетей (ММНС) для описания механизмов различных психических явлений и их моделирования. Можно сказать, что психофизиологами и психологами принимается новая «нейросетевая» метафора для понимания работы мозга с точки зрения механизмов психических явлений. С 60-х по 90-е годы в психологии доминировала компьютерная метафора, базирующаяся на понимании психики, как реализуемой мозгом компьютерной программы. Хотя когнитивные психологи того времени достаточно много знали об устройстве мозга и его функциях, все же наиболее популярное понимание его работы на тот момент базировалось на знаниях о работе компьютера. С 80-х годов началось бурное развитие математического моделирования нейросетей, проявившееся сначала в интересе к «ассоциативным» сетям Хопфилда, а затем в разработке алгоритма обратного распространения ошибки Руммельхартом и соавторами. Стало понятно, что работа нейронной сети, ее самообучение и другие функции, существенно отличаются от компьютерных алгоритмов, зато очень похожи по динамике и результатам на психические явления. Стало ясно, что в основе обучения нейросети, лежит избегание ошибок отклонения от целевой функции. В классической психофизиологии также множество теорий базируется на идее научения как минимизации ошибки прогноза средовых событий (Hull С.Л., Анохин П.К, Rescorla R. Wagner A., Швырков В.Б., Friston К.). При знакомстве с особенностями работы математических моделей нейронных сетей и схожих алгоритмов часто встает вопрос: а, что, упорно, на протяжении многих лет изучают многие когнитивные психологи, не проявления ли свойств реальных мозговых нейронных сетей, модели которых давно разработаны и изучены математиками? Некоторые из этих особенностей можно перечислить:

1. Одновременность работы множества элементов модели, качественно увеличивающая быстроедействие при решении сложных задач, требующих работы с многомерными данными.
2. Невозможность однозначного управления нейросетью, что не позволяет легко ее перепрограммировать в отличие от типичного компьютера.
3. Возможность самообучения, то есть оптимизации работы на основе перестройки параметров нейросети с целью минимизации ошибок, возникающих при работе сети.

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-06-00182

4. Алгоритм работы нейросети модифицируется в процессе ее функционирования. Данные, заложенные в нее прошлым опытом обучения, в дальнейшем меняются, что представляет из себя процесс изменения синаптических связей между нейронами (весов межнейронных связей). Процесс управления сетью, в основном, упирается в организацию предъявляемых ей данных, то есть процессом обучения.
5. При выполнении задач, сопровождающихся самообучением, наблюдается определенная динамика. Она характеризуется скачкообразным изменением эффективности работы (минимизации функции ошибок), что сопровождается, например, эффектом возникновения одних и тех же повторяющихся ошибок, связанным с преодолением локальных минимумов поверхности ошибки при обучении. (Хайкин С. 2008).
6. В процессе самообучения (при определенных условиях организации опыта обучения ММНС) можно наблюдать феномен дифференциации образов, а при работе с семантическим материалом дифференциации понятий (Rogers T. et al. 2004).
7. ММНС работает по принципу ассоциации между различными образами, подаваемыми на ее вход, что позволяет на основе этого ассоциативного принципа применять правила, оптимизированные и приобретенные на одном материале, к другим данным, что в некоторых случаях создает эффект порождения новых знаний (Rogers T. T. et al. 2004).

Здесь приведены только некоторые из особенностей ММНС, которые существенно отличаются от возможностей компьютеров, но похожи на проявления человеческой психики. На данный момент разработано много ММНС, обучающихся на разных принципах, большинство из которых схожи по структуре и функциям с различными подструктурами мозга и нервной системы (Arbib M.A. 2003). Это позволяет говорить о больших перспективах применения нейросетевой метафоры в психологии.

Кроме того, существуют работы, исследующие не только возможности создания, обучения и оптимизации таких моделей, но изучающие динамику их работы, как в сфере эффективности обучения, так и мониторинга их поведения на микро и мезоуровнях, то есть на уровне одной ячейки или суммарной работы множества ячеек (нейронов). Все это говорит о том, что представляет интерес и динамика работы нейросетей в ключе теорий динамического хаоса (Arecchi F. T. 2003, Рабинович М. И. и др. 2010). Сам факт, что замкнутая сама на себя модель нейросети может породить даже «хаотический» аттрактор, говорит о сходстве ее работы со сложными системами, к которым мы относим и психику. Соответственно, изучение динамических характеристик работы мозга и поведения по сравнению с динамикой функционирования их нейросетевых моделей, позволяет соотносить психические явления с их нейросетевыми моделями, тем самым уточняя структуру их гипотетических механизмов.

Принципы работы различных типов моделей нейросетей существенно различаются. В последнее десятилетие появляется все большее количество работ (Fera, F., Weickert, T. W., Goldberg, T. E., Tessitore, A., Hariri, A., Das, S., Gluck, M. A. 2005, Oh J. et al. 2015, Mnih V. et al. 2015, Zhao J. et al. 2015) в которых несколько разных по принципу функционирования ММНС комбинируются в одной сети, в конечном счете, выполняющей более успешно те функции, которые были на нее возложены. Этот подход позволяет комбинировать различные ММНС для более успешного достижения целей, поставленных перед разработчиками. Мы предлагаем использовать его для изучения мозговых механизмов психики.

Такой подход позволяет предложить понимание работы мозга и нервной системы в целом как громадную рекомбинацию различных по принципам работы нейросетей (подсетей), выполняющих свои функции различными способами. Для того, чтобы отдифференцировать все подобные подсети в мозге, отнюдь не достаточно анатомического принципа, потому что в тех или иных условиях один и тот же элемент в мозге может участвовать в одной такой «подсети», а в других в другой. В первую очередь при такой дифференциации может помочь эффект распространяемой по сети ошибки, связанной с обучением. Ведь в разных подсетях за модификацию межнейронных связей могут отвечать разные виды ошибки (отклонений от целевой функции), распространяемые разными способами. Например, эти сигналы могут распространяться различными нейромедиаторами. Сами виды ошибки, распространяемые в нейросетях могут быть связаны:

1. С отклонениями в различных физиологических константах внутренней среды организма (связанные с вегетативной нервной системой) или с отклонениями в прогнозе их изменений со временем.
2. С ошибками прогноза результата наших целенаправленных движений и действий (связанные с соматической нервной системой)
3. С ошибками предсказаний внешних событий и спонтанных движений. (связанные с центральной нервной системой)

Все эти подсети в реальном мозге функционируют в единой нейродинамической системе в которой реализуется их взаимодействие, переключение, выделение различных ресурсов и другие функции, которые можно назвать саморегуляцией. Нейродинамическая система это совокупность различных по функциям нейросетей мозга, анатомически состоящих из нейронов различных мозговых модулей. Дифференциация этих нейросетей может быть проведена не столько анатомически, сколько за счет видов ошибки, участвующих в модификации межнейронных соединений (обучении), Они могут быть математически смоделированы из отдельных известных ММНС разных типов, определенным образом взаимодействующих таким образом, чтобы результаты поведения и

динамика изменений модели совпадала по определенным параметрам с нейродинамикой и поведением человека, в психике которого функционируют явления, представленные данной нейродинамической системой,

Исходя из этих идей можно предложить новый подход к изучению и моделированию психических явлений. Каждое психическое явление можно рассматривать как результат работы комбинации нейросетей, определенным образом взаимосвязанных, обладающих определенными ресурсами (количество нейронов, связей средний уровень их активации и т.п.), обучающихся на определенных видах ошибки, распространяемой на те или иные подсети и связанной с теми или иными внешними и внутренними событиями и их прогнозами. Функционирование этих подсетей должно иметь определенную нейродинамику, параметрами которой будет определенная сложность (в терминах нелинейных динамических систем), упорядоченность (в терминах энтропии), предсказуемость и другие.

Данный подход реализуется в работах лаборатории психофизиологии СПбГУ и работах коллег и учащихся, проходящих стажировку и проводящих эксперименты на ее базе (Gorbunov I. Semenov P. 2009, Горбунов И.А., Коваль В.М., Балин В.Д., Першин И.И. 2014, Shcherbakova, O. V., Gorbunov, I. A., Golovanova, I. V., Kholodnaya, M. A. 2014, Tkacheva L. O., Gorbunov I. A., Nasledov A. D. 2015, Горбунов И. А., Зайнутдинов М. Р., Локоткова М. А. 2015).

Литература

1. Горбунов И. А., Зайнутдинов М. Р., Локоткова М. А. Моделирование процесса речепорождения с помощью математических моделей нейронных сетей у больных неврозами //Петербургский психологический журнал. – 2015. – №. 11.
2. Горбунов И.А., Коваль В.М., Балин В.Д., Першин И.И. Изучение зависимости семантического пространства человека от предыдущего опыта на основе нейросетевого моделирования //Материалы Шестой Международной научной конференции по когнитивной науке, 23-27.06.14. Калининград, 2014
3. Рабинович М. И., Мюезинолу М. К. Нелинейная динамика мозга: эмоции и интеллектуальная деятельность //Успехи физических наук. – 2010. – Т. 180. – №. 4.
4. Хайкин С. Нейронные сети: полный курс, 2-е издание. – Издательский дом Вильямс, 2008.
5. Arbib M. A. The handbook of brain theory and neural networks. – MIT press, 2003.
6. Arecchi F. T. Chaotic neuron dynamics, synchronization, and feature binding: quantum aspects //Mind and Matter. – 2003. – Т. 1. – №. 1
7. Fera F. et al. Neural mechanisms underlying probabilistic category learning in normal aging //The Journal of Neuroscience. – 2005. – Т. 25. – №. 49

8. Gorbunov I. Semenov P. Brain centers model and its applications to EEG Analysis// International conference on neural computation. Portugal, Funchal-Madeira, 2009
9. Mnih V. et al. Human-level control through deep reinforcement learning //Nature. – 2015. – Т. 518. – №. 7540
10. Oh J. et al. Action-conditional video prediction using deep networks in atari games //Advances in Neural Information Processing Systems. – 2015
11. Rogers T. T., McClelland J. L. Semantic cognition: A parallel distributed processing approach. – MIT press, 2004.
12. Shcherbakova, O. V., Gorbunov, I. A., Golovanova, I. V., Kholodnaya, M. A. The neural efficiency hypothesis: Further evidence from the EEG study of conceptual thinking //International Journal of Psychophysiology. – 2014. – Т. 2. – №. 94
13. Tkacheva L. O., Gorbunov I. A., Nasledov A. D. Reorganization of system brain activity while understanding visually presented texts with the increasing completeness of information //Human Physiology. – 2015. – Т. 41. – №. 1
14. Zhao J. et al. Stacked what-where auto-encoders //arXiv preprint arXiv:1506.02351. – 2015.

Neurodynamic system as one of the foundations of the methodology for modeling human mental resources.

Gorbunov Ivan Anatolyevich

Saint-Petersburg state university (Saint-Petersburg Russia)

The article proposes the use of a new methodology for studying and modeling human mental resources by understanding the brain as a complex neurodynamic system, understood as a combination of a simple neural networks trained on different principles. On the basis of a combination of mathematical models of these networks can be provided or generated model neurodynamic system, which has a specific dynamics and characteristics of the structure. It will be the basis for the construction of the experiment and clarify the mechanisms of psychic phenomena underlying mental human resources.

В статье предлагается использование новой методологии изучения и моделирования ментальных ресурсов человека на основе понимания мозга как сложной нейродинамической системы, понимаемой как комбинации более простых нейронных сетей, обучающихся по разным принципам. На основе комбинации математических моделей этих сетей может быть представлена или сформирована модель нейродинамической системы, обладающая определенной динамикой и особенностями структуры, и являющаяся основой для построения эксперимента и уточнения механизмов психических явлений, лежащих в основе ментальных ресурсов человека.

