

Рекомендованная ссылка

Нейросетевые механизмы когнитивной гибкости [Электронный ресурс] / Терехин А.Т. [и др.] // Психологические исследования: электрон. науч. журн. 2009. N 4(6). URL: <http://psystudy.ru> (дата обращения: чч.мм.20гг).

[Дата обращения в формате число-месяц-год (чч.мм.гггг.) – дата, когда вы обращались к странице сайта и она была доступна]

Нейросетевые механизмы когнитивной гибкости

Терехин А.Т.¹, Будилова Е.В.², Карпенко М.П.³, Качалова Л.М.⁴, Чмыхова Е.В.⁵

Московский государственный университет им. М.В.Ломоносова, Москва, Россия

Современная гуманитарная академия, Москва, Россия

Когнитивная гибкость нейронной сети определяется как актуальный диапазон изменения гладкости ее функции Ляпунова – большая гладкость облегчает нахождение стратегически эффективных направлений решения задачи, а меньшая необходима для проработки его деталей. В качестве модели используется градуальная сеть Хопфилда, в которой изменение гладкости функции Ляпунова обеспечивается изменением параметра гладкости сигмоидной функции активации.

Ключевые слова: когнитивные функции, когнитивная гибкость, мозг, нейросетевое моделирование, нейронная сеть, сеть Хопфилда, аттрактор, когнитивный энергетический ландшафт

Neural network mechanisms of cognitive flexibility

Terekhin A.T., Budilova E.V., Karpenko M.P., Kachalova L.M., Chmykhova E.V.

Lomonosov Moscow State University, Moscow, Russia

Modern University for the Humanities, Moscow, Russia

Cognitive flexibility of a neural network is defined as the range of smoothness of its Lyapunov function: greater smoothness makes easier finding strategically more effective solutions of problems whereas lesser one is necessary for elaborating details. We use as a model the gradual Hopfield network ensuring the variation of smoothness of the Lyapunov function by changing the activation function smoothness parameter.

Keywords: cognitive functions, cognitive flexibility, brain, neural network modeling, neural network, Hopfield network, attractor, cognitive energy landscape

СВЕДЕНИЯ ОБ АВТОРАХ

¹ *Терехин Анатолий Тимофеевич.* Доктор биологических наук, профессор кафедры общей экологии, биологический факультет, Московский государственный университет им. М.В.Ломоносова, Ленинские горы, 1, стр. 12, 119991 Москва, Россия

E-mail: terekhin_a@mail.ru

Персональная web-страница: http://ecology.genebee.msu.su/3_SOTR/CV_Terekhin.htm

² *Будилова Елена Вениаминовна.* Кандидат технических наук, ст. научн. сотр. кафедры общей экологии, биологический факультет, Московский государственный университет им. М.В.Ломоносова, Ленинские горы, 1, стр. 12, 119991 Москва, Россия.

E-mail: evbudilova@mail.ru

³ *Карпенко Михаил Петрович.* Доктор технических наук, профессор, президент Современной гуманитарной академии; Современная гуманитарная академия, ул. Нижегородская, 32, 109029 Москва, Россия.

E-mail: rectorat@muh.ru

⁴ *Качалова Лариса Андреевна.* Кандидат биологических наук, директор Института когнитивной нейрологии, Современная гуманитарная академия, ул. Нижегородская, 32, 109029 Москва, Россия.

E-mail: lefi@muh.ru

⁵ *Чмыхова Екатерина Витальевна.* Кандидат социологических наук, зам. проректора Современной гуманитарной академии (СГА), директор Департамента науки и инноваций СГА, зав. кафедрой практической и экспериментальной психологии СГА; Современная гуманитарная академия, ул. Нижегородская 32, 109029 Москва, Россия.

E-mail: niipo@muh.ru

Постановка задачи

Как в повседневной жизни, так и в серьезных психологических исследованиях важной характеристикой мышления человека считается степень его гибкости (флексibility) – свойства, противоположного когнитивной жесткости (ригидности) [Залевский, 1993]. На нейробиологическом уровне обнаруживается обратная зависимость гибкости мышления от активности катехоламиновых систем мозга [Noradrenergic modulation ... , 1999]; снижение когнитивной гибкости наблюдается при некоторых психических заболеваниях – шизофрении [Twamley et al., 2003], аутизме [Botting, Conti-Ramsden, 2003]. Гибкость мышления тесно связана с его креативностью [Guilford, 1967; Torrance, 1972]. Задача данного исследования состоит в выявлении механизмов, обеспечивающих гибкость мышления на информационном, нейросетевом уровне. В качестве нейросетевой модели, описывающей базовые аспекты когнитивной деятельности мозга, используется градуальная сеть Хопфилда с сигмоидной функцией активации, содержащей изменяемый параметр гладкости. При неизменных синаптических весах увеличение параметра гладкости приводит к увеличению гладкости функции энергии сети. В свою очередь, степень гладкости функции энергии определяет масштаб переходов между аттракторами сети: большая гладкость обеспечивает возможность перехода между более удаленными аттракторами сети. Соответственно, в рамках данной модели, мы определяем когнитивную гибкость сети как диапазон изменения степени гладкости функции энергии, но полагаем, что, в общих чертах, этот подход применим и к пониманию механизмов обеспечения гибкости мыслительного процесса реального мозга.

Модель

Сеть Хопфилда [Hopfield, 1984] характеризуется тем, что все ее элементы (нейроны) связаны со всеми остальными и есть только два ограничения на веса синаптических связей: веса прямых и обратных связей равны между собой, а вес связи нейрона с самим собой равен нулю. Состояние каждого нейрона i , $i = 1, 2, \dots, n$, в момент времени t задается переменной состояния $x_i(t)$, принимающей значение от -1 до 1 . Динамика сети определяется правилами изменения состояний ее нейронов и весов $w_{ij}(t)$ синаптических связей, соединяющих нейроны i и j . Если веса $w_{ij}(t)$ и состояния $x_i(t)$ сети в момент времени t заданы, то в следующий момент $t+1$ состояние нейрона i изменится в соответствии с правилом

$$x_i(t+1) = \frac{2}{1 + e^{-\frac{1}{S} \sum w_{ij} x_j(t)}} - 1,$$

где S – параметр гладкости сигмоидной функции, задаваемой правой частью уравнения (при $S = 0$ сигмоидная функция превращается в пороговую, а при $S \rightarrow \infty$ приближается к горизонтальной прямой, совпадающей с осью абсцисс). Что касается синаптических весов, то при обучении сети они изменяются для всех $j \neq i$ в соответствии с правилом Хебба

$$w_{ij}(t+1) = w_{ij}(t) + x_i(t)x_j(t),$$

то есть вес увеличивается на единицу, если нейроны i и j находятся в одинаковых состояниях и уменьшается на единицу, если – в разных [Hebb, 1949].

Поведение сети Хопфилда очень удобно интерпретировать в терминах функции Ляпунова (функции энергии), вид которой в момент t определяется значениями весов связей и которая задается формулой

$$E(x_1, \dots, x_n, t) = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n x_i(t) w_{ij}(t) x_j(t) + S \sum_{i=1}^n \ln[(1+x_i(t))^{1+x_i(t)} (1-x_i(t))^{1-x_i(t)}].$$

Можно показать, что изменение состояния сети может происходить только в направлении уменьшения функции энергии. Этот факт позволяет наглядно представлять изменение состояния сети как его движение по поверхности функции энергии. Все пространство состояний распадается на области, образующие так называемые бассейны притяжения аттракторов, и динамика сети, установленной в некоторое начальное состояние, полностью зависит от того, в бассейн притяжения какого из аттракторов она попала.

Определенная таким образом сеть обладает рядом интересных свойств с точки зрения ее использования как модели биологических когнитивных процессов (см., напр.: [Веденов, 1988; Amit, 1989]). Пусть вначале все синаптические веса сети равны нулю и ее функция энергии представляет собой «горизонтальную» гиперплоскость $E(x_1, \dots, x_n, t) = 0$. Если на входы нейронов сети поступит некоторая конфигурация сигналов – образ и нейроны установятся в состояния, соответствующие этому образу, то, в соответствии с правилом Хебба, сформируются новые веса сети. Если после этого предъявить сети любое другое сочетание сигналов и разрешить ей свободно эволюционировать во времени, то через небольшое число шагов состояние сети стабилизируется на ранее предъявленном образе. Можно сказать, что сеть его «запомнила». Что касается функции энергии, то в точке, соответствующей запомненному образу при запоминании образуется локальный минимум.

Предъявление других образов внесет свой вклад в изменение синаптических весов и если затем предъявить сети образ, похожий на один из прежних, то сеть придет в состояние, соответствующее именно этому образу, то есть воспоминание имеет ассоциативный характер – сеть эволюционирует в направлении локального минимума, в бассейн притяжения которого она попала. При запоминании образов (обучении сети) могут образоваться и дополнительные локальные минимумы, соответствующие образам, как бы составленным из частей запоминаемых образов, то есть своего рода «химеры» [см., напр.: Терехин, Будилова, 1995]. Возможность появления таких химер наделяет сеть, в некотором смысле, способностью к «творческой фантазии» и «обобщению», то есть действительно сеть Хопфилда описывает базовые свойства когнитивной деятельности мозга и, следовательно, может служить основой для моделирования феномена когнитивной гибкости.

Анализ

Рассмотрим изменение свойств функции энергии при изменении параметра S , определяющего степень гладкости сигмоидной функции активации. Когда значение S очень мало, то есть сигмоидная кривая близка к пороговой функции, то второй член в правой части формулы энергии, обозначим его E_2 , становится пренебрежимо малым и в ней доминирует первый член – обозначим его E_1 . На рис. 1 для случая сети из двух нейронов с синаптическими весами $w_{12} = w_{21} = 2,5$ и параметром сглаженности $S = 1$ представлены графики E_1 , E_2 и $E_1 + E_2$ в сечении $x_1 = x_2$. Мы видим, что минимумы функции энергии наиболее выражены (и соответствуют крайним состояниям нейронов) при малых S (либо при больших абсолютных значениях синаптических весов), а при увеличении гладкости S (или уменьшении весов) минимумы становятся все менее и менее

выраженными и соответствуют значениям состояний нейронов, приближающимся к нулю. При некоторой достаточно большой гладкости S минимумы исчезают полностью, а состояния нейронов обращаются в нуль.

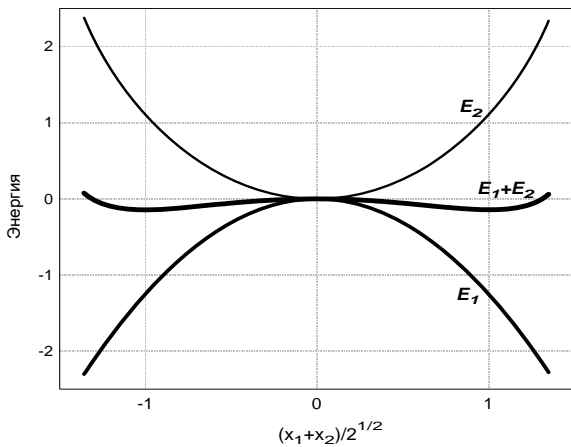


Рис. 1. Графики функции энергии E_1+E_2 и ее компонент E_1 и E_2 в сечении x_1+x_2 .

При числе нейронов n , большем двух, характер расположения минимумов функции энергии и его изменение при изменении S аналогичны случаю $n = 2$. Подобно тому, как в двумерном случае при малых S минимумы функции энергии находятся в противоположных вершинах квадрата и при увеличении S перемещаются к центру квадрата вдоль диагонали, соединяющей эти вершины, в n -мерном случае они находятся в противоположных вершинах n -мерного куба и при увеличении S перемещаются к центру куба вдоль соединяющей эти вершины диагонали. Поскольку различные минимумы могут иметь разную глубину, то при монотонном увеличении S , начиная с нуля, они будут постепенно исчезать — сначала самые мелкие, затем более глубокие, пока, в конечном итоге, не исчезнут все.

В качестве иллюстрации на рис. 2 представлен график уменьшения числа аттракторов первоначально обученной сети из 81 нейрона при увеличении параметра сглаженности S .

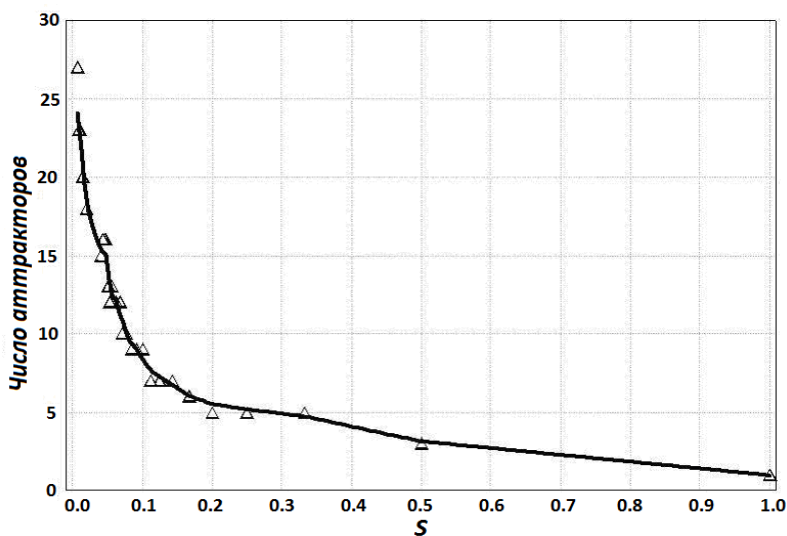


Рис. 2. Зависимость числа аттракторов от значения S для сети из 81 нейрона.

Обсуждение

Таким образом, при увеличении параметра S происходит сглаживание функции энергии – можно назвать ее когнитивным энергетическим ландшафтом. Этот эффект наглядно проиллюстрирован на рис. 3. При больших S сеть как бы не замечает мелких минимумов и ее состояние эволюционирует в направлении более глубокого минимума, что можно интерпретировать как нахождение стратегически более эффективного решения задачи.

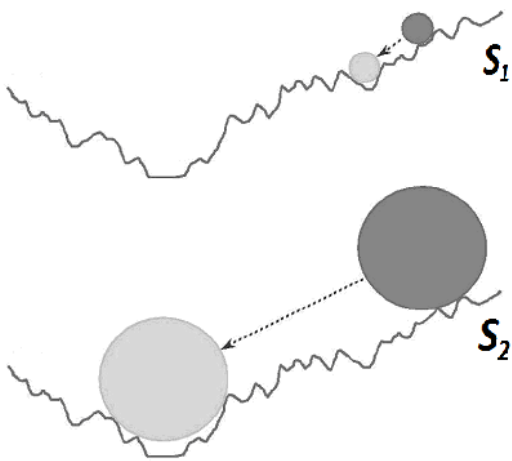


Рис. 3. Эффект сглаживания когнитивного энергетического ландшафта: $S_1 < S_2$.

В частности, в работе [Когнитивные преимущества ... , 1995], основываясь на том, что возрастные изменения мозга (например, снижение плотности дофаминовых рецепторов) приводят, в основном, к ослаблению межнейронных связей и, следовательно к сглаживанию когнитивного ландшафта, был сделан вывод о том, что при старении мозга наряду с отрицательными изменениями его когнитивных свойств, такими как ослабление памяти на детали повседневной жизни (см., напр.: [Блокирование памяти ... , 2009]), у него появляется ценная способность видеть проблемные ситуации с более общей точки зрения, выделять в них главное и, в конечном итоге, находить наиболее эффективные решения, то есть возникает способность, которую можно рассматривать как составляющую часть мудрости – когнитивного свойства, традиционно ассоциирующегося с пожилым возрастом.

Однако умение вникнуть в детали решаемой задачи также является ценным когнитивным качеством. В идеале, желательно, чтобы мозг обладал способностью произвольно и в полном диапазоне менять степень подробности анализа проблемы, подобно тому как человек с нормальным зрением (в отличие от близорукого или дальновзоркого) может без очков легко перестроиться с созерцания пейзажа на чтение текста, написанного мелким шрифтом, и наоборот. Можно назвать эту способность «трансфокальностью» мышления (по аналогии с функцией «зум» фотоаппарата, обеспечиваемой специальным устройством – трансфокатором).

В рамках данной модели можно определить когнитивную гибкость как диапазон изменения степени гладкости когнитивного ландшафта. Следует, однако учитывать, что диапазоны одинаковой ширины могут отличаться своим положением на шкале гладкости, поэтому исчерпывающе когнитивная гибкость определяется двумя значениями – нижней и верхней границами достигаемой гладкости.

Литература

Блокирование памяти: факты, проблемы, модели / Будилова Е.В. [и др.] // *Биофизика*. 2009. Т. 54, вып. 2. С. 287–292.

Веденов А.А. Моделирование элементов мышления. Москва: Наука, 1988.

Залевский Г.В. Психическая ригидность в норме и патологии. Томск: Изд-во ТГУ, 1993.

Когнитивные преимущества третьего возраста: нейросетевая модель старения мозга / Карпенко М.П. [и др.] // *Журнал высшей нервной деятельности*. 2009. N 59, вып. 2. С. 252–256.

Терехин А.Т., Будилова Е.В. Сетевые механизмы биологической регуляции // *Успехи физиологических наук*. 1995. Т. 26, N 4. С. 75–97.

Amit D.J. Modeling brain function. The world of attractor neural networks. New York: Cambridge University Press, 1989.

Noradrenergic modulation of cognitive flexibility in problem solving / Beversdorf D.Q. [et al.] // *Cognitive Neuroscience*. 1999. Vol. 10(13). P. 2763–2767.

Botting T., Conti-Ramsden G. Autism, primary pragmatic difficulties, and specific language impairment: can we distinguish them using psycholinguistic markers? // *Developmental Medicine and Child Neurology*. 2003. Vol. 45(8). P. 515–524.

Guilford J.P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Hebb D.O. The organization of behavior. New York: Wiley, 1949.

Hopfield J.J. Neurons with graded response have collective computational properties like those of two-state neurons // *Proc. Natl. Acad. Sci. USA*. 1984. Vol. 81. P. 3088–3092.

Torrance E.P. Can we teach children to think creatively? // *Journal of Creative Behavior*. 1972. Vol. 6. P. 236–262.

Twamley E.W., Jeste D.V., Bellack A.S. A Review of Cognitive Training in Schizophrenia // *Schizophrenia Bulletin*. 2003. Vol. 29(2). P. 359–382.

Поступила в редакцию 15 апреля 2009 г. Дата публикации 25 августа 2009 г.