

А.Д. Редозубов

Фонд поддержки научных исследований механизмов работы мозга, лечения его заболеваний, нейромоделирования
имени академика Натальи Петровны Бехтеревой
galdrd@gmail.com

РЕАЛИЗАЦИЯ МЕХАНИЗМОВ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ В КОНТЕКСТНО-СМЫСЛОВОЙ МОДЕЛИ РАБОТЫ МОЗГА

Рассматривается биологически обоснованная математическая модель, показывающая возможные пути реализации мозгом процессов мышления и поведения. В основе модели лежат принципы обучения с подкреплением. Показывается, что миниколонки коры могут являться независимыми процессорами, выполняющими контекстные вычисления; что является смыслом информации и как в таком подходе мозг оперирует смыслом информации. Приводится механизм того, как реализация поведения осуществляется за счет параллельного рассмотрения доступных альтернатив в пространстве сформированных у агента контекстов.

Ключевые слова: обучение с подкреплением, мозг, кора мозга, пространство контекстов, смысл информации, контекстно-смысловая модель

Введение

В обучении с подкреплением есть существенный ряд фундаментальных вопросов, которые непосредственно влияют на практические схемы реализации. Например, как достоверно определить, что ситуация, в которой оказался агент, аналогична ранее встречавшейся? Как связать подкрепление с его истинной причиной? Как избежать комбинаторного взрыва при построении стратегий поведения в реальных задачах? И многие подобные вопросы. При этом мы видим, что биологический мозг успешно решает задачи формирования поведения, по-видимому, используя принцип обучения с подкреплением.

На сегодня сформулирована контекстно-смысловая модель работы биологического мозга [1]. В рамках этой модели сделаны предположения об используемых мозгом путях реализации обучения с подкреплением.

Информация и контекст

Для представления информации требуется задать знаковую систему. Задание знаковой системы подразумевает введение набора знаков (понятий), которые будут использоваться для записи информации и определение грамматики, которая задаст механизм, позволяющий связать запись с неким смыслом. Часто, например, в естественных языках, оказывается, что используемые понятия многозначны. В таких случаях, для правильного понимания информации требуется не только грамматически разобрать фразу, но и понять в каком из значений используются составляющие ее понятия.

Для разрешения неоднозначности используется представление о контексте. В случае естественного языка под контекстом обычно понимают законченный отрывок письменной или устной речи, общий смысл которого позволяет уточнить значение входящих в него отдельных слов, предложений, и т. п.

В теории формального анализа понятий [2] контекст определяют, как тройку $K = (G, M, I)$, где G — множество объектов, M — множество признаков, а отношение $I \subseteq G \times M$ говорит о том, какие объекты какими признаками обладают. В таком подходе в разных контекстах одним и тем же понятиям могут соответствовать разные признаковые описания. При этом подразумевается, что признаковые описания неким образом отражают суть, которую приобретают понятия в указанном контексте.

В [1] и [3] предложен несколько иной подход к пониманию контекстов. Будем полагать, что контекст соответствует некой сущности, которой может быть сопоставлено определенное понятие. Предположим, что эта сущность проявляет себя тем, что по определенным правилам изменяет текущее описание I на иное описание $I' (I)$, оставаясь при этом в той же знаковой системе.

$$I' = K(I) \quad (1)$$

Будем называть I исходным информационным описанием, а I' его трактовкой (интерпретацией) в контексте K .

Рассмотрим простой пример. Предположим у нас есть растровое изображение. Каждый пиксель может принимать только два значения. Можно ввести систему понятий, перечислив все пиксели изображения. Каждый пиксель — самостоятельное понятие. Также, можно ввести простую грамматику, в которой каждой картинке будет соответствовать простое перечисление активных пикселей. В такой системе контекстом может быть фиксированный сдвиг изображения. Например, может быть контекст K : сдвиг на 2 пикселя вниз. В таком контексте исходное описание — это исходная картинка, а трактовка — это она же, но сдвинутая на 2 точки вниз. Контекст

К при этом определяется правилами перехода одних понятий в другие. То есть, каждое понятие исходного описания переходит в понятие, соответствующее точке, находящейся на два пикселя ниже чем точка, соответствующая исходному понятию. Если точка выходит за пределы картинка, то такое понятие просто исчезает из описания.

Несложно представить, как такие правила могут быть получены в процессе обучения, когда есть множество примеров одного и того же сдвига.

Пространство контекстов и смысл

Можно не ограничиваться одним контекстом, а создать систему (пространство) контекстов. Например, для пиксельного изображения можно перечислить все возможные смещения не превышающие размеры картинка. Таким способом можно создать пространство комбинаций всех допустимых сдвигов по горизонтали и вертикали. Теперь предположим, что у нас есть память, которая хранит набор предварительно отцентрированных изображений. Продублируем эту память в каждом контексте.

Предположим, что в этой памяти есть квадрат. Создадим новое изображение с этим квадратом, поместив его из центра в правый верхний угол. Теперь описание этого изображения подадим в пространство контекстов, то есть потребуем, чтобы каждый контекст сформировал свою интерпретацию. По сути, мы проварьерируем все возможные смещения. Неизбежно найдется контекст, который приведет квадрат из правого верхнего угла обратно в центр. Если в каждом из контекстов выполнить сравнение трактовки с памятью, то именно в том контексте, где квадрат окажется в центре, обнаружится совпадение трактовки с одним из элементов памяти.

Будем говорить, что когда трактовка оказывается похожа на память, то в таком контексте входная информация приобретает смысл. При этом будем полагать, что такая трактовка и есть смысл исходной информации. Такая схема вычислений представлена на рис. 1.

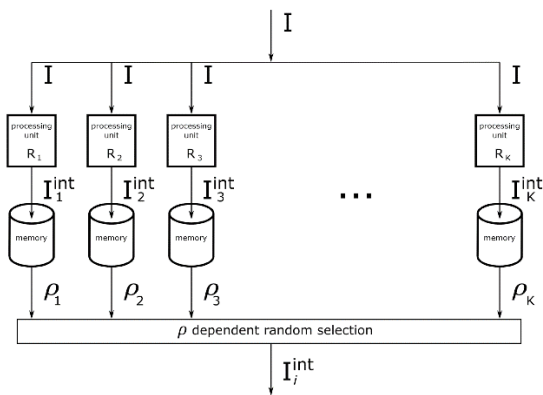


Рис. 1. Пространство контекстов

Контекстная-ориентированная схема вычислений, примененная к изображениям, оказывается очень похожа на идеологию сверточных сетей [4]. Знание о том, как применить ядро свертки к различным частям изображения эквивалентно построению правил контекстных преобразований. Память контекстов – аналог набора ядер свертки. Идеологическое различие можно описать, как кто к кому идет: «гора к Магомеду» или «Магомед к горе». Использование подхода, в котором не память приводится к текущему описанию, а текущее описание меняется в каждом из контекстов, создавая множество трактовок, позволяет применить этот метод к любому типу информации. Кроме того, если в сверточных сетях правила свертки задаются изначально, как отражение нашего априорного знания о геометрии мира, то в контекстной модели пространство контекстов может быть построено в результате наблюдения за входной информацией.

Миниколонки коры

Миниколонки коры или, как их еще принято называть, кортикальные модули — это группы из компактно расположенных, вертикально организованных нервных клеток (рис.2). Каждая миниколонка содержит порядка сотни сильно связанных между собой нейронов. Вся кора мозга - это совокупность миниколонк. Выделяют различные зоны коры. В пределах каждой зоны миниколонки морфологически однородны. При этом даже в различных зонах миниколонки достаточно схожи по своей организации. Более

того, миниколонки коры похожи у различных видов животных. Это позволяет предположить, что миниколонки играют фундаментальную роль в информационных процессах. Вернон Маунтклас [5] выдвинул гипотезу, что для мозга кортикальная колонка – это основная структурная единица переработки информации.

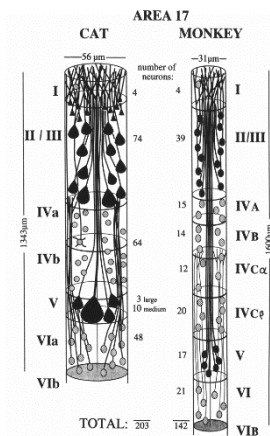


Рис. 2. Миниколонки первичной зрительной коры кошки (слева) и обезьяны (справа) [6]

Часто миниколонку рассматривают, как модуль, который позволяет узнать некое явление в различных его проявлениях. Например, из последних работ в этом направлении показательны капсульные сети Джофри Хинтона [7].

В [3] показывается, что миниколонки, из которых состоит кора, - это независимые вычислительные модули, выполняющие контекстную обработку поступающей информации в том смысле, который был описан выше. Контекстная обработка включает в себя создание интерпретации входящей информации и последующее определение того, насколько полученная интерпретация соответствует памяти, хранящейся в той же миниколонке (рис. 3).

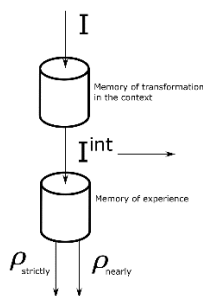


Рис. 3. Функциональная схема миниколонки коры

Для того, чтобы миниколонка коры могла выполнять функции контекстного процессора, ей необходимо обладать способностью осуществлять преобразование исходной информации в ее трактовку, причем миниколонка должна уметь на примерах обучаться правилам такого преобразования. Также требуется наличие у миниколонки достаточно емкой событийной памяти и механизмов работы с ней. В [1] и [3] были показаны биологически обоснованные механизмы, реализующие указанные свойства.

Обучение с учителем в комбинаторном пространстве

В предлагаемой концепции предполагается, что кодирование информации мозгом осуществляется не за счет существования «нейронов бабушки», то есть, нейронов, специфичных к определенным явлениям, а за счет того, что нейроны одной миниколонки могут своей совместной активностью формировать коды, соответствующие используемым мозгом понятиям [8]. При таком подходе каждая миниколонка может кодировать практически неограниченное число понятий. Показано, что возникающие в миниколонке коды могут приводить к распространению волн со специфической внутренней структурой, которые позволяют распространять информацию, возникающую в одной миниколонке по всему пространству зоны коры или проецировать ее на другие зоны [9].

Пространство миниколонки – это плотно упакованная система пересечений между аксонами и дендритами нейронов. Контакты между аксонами и дендритами называются синапсами.

Сигнал, представленный нейронным кодом в пределах одной миниколонки, вызывает распространение спайков по аксонам нейронов, участвующих в формировании этого кода. В результате в химических синапсах этих нейронов происходит выброс нейромедиаторов и нейромодуляторов.

За счет плотной упаковки, а также в результате работы плазматических астроцитов нейромедиаторы влияют не только на свой синапс, но и затрагивают ближайшее окружение. Мозг использует достаточно большое число нейромедиаторов и нейромодуляторов. Это позволяет создать в каждом месте миниколонки свой специфичный коктейль нейромедиаторов, уникальный по отношению к определенному фрагменту того кодового сигнала, который его породил. На мембране нейронов находятся рецепторы, которые способны образовывать кластеры. Такие кластеры могут фиксировать информацию о том, что определенные нейроны сработали совместно и впоследствии реагировать на повторение такого события.

В [10] показано, как такую конструкцию можно успешно использовать для реализации обучения с учителем. Смоделирован механизм обучения и обобщения, построенный на создании при предъявлении примеров множества рецептивных кластеров, как гипотез о том, что определенное сочетание нейромедиаторов соответствует некой закономерности. Последующий опыт позволяет достаточно быстро проверить, действительно ли созданный кластер является закономерностью или он следствие случайного сочетания.

Способность кластеров мембранных рецепторов фиксировать информационные картины дала возможность описать модель памяти, основанную на интерференции пространственных паттернов активности [3][8]. Один паттерн несет в себе ключ, другой кодирует значение. Показана роль гиппокампа, как генератора ключей для формирования памяти [8].

В [8] показано, что одна миниколонка способна хранить сотни мегабайт семантической информации, что вполне достаточно для реализации контекстного подхода.

Реализация Q-критика в пространстве контекстов

Поведение животных и человека хорошо объясняется принципами обучения с подкреплением. Состояния «хорошо – плохо», возникающие в моменты голода, жажды, боли, насыщения, секса и т. п. хорошо сопоставляются с сигналами подкрепления. Эмоциональные оценки, то есть те состояния «хорошо – плохо», которые соответствуют переживаниям, могут быть сопоставлены с оценкой качества ситуации в понимании обучения с подкреплением. В [11] показано, что все эмоциональные оценки можно свести либо к страху, либо к предвкушению и что для всех эмоциональных оценок можно показать путь их формирования в онтогенезе. Это позволяет утверждать, что эмоциональные оценки не являются врожденными и не наследуются генетически, а формируются уже при жизни.

Однако, обучение с подкреплением сталкивается с определенной сложностью, когда в практических задачах оказывается, что пространство возможных состояний агента и число возможных в этих состояниях поступков столь велико, что обучение оказывается затруднительным. Выход из этой ситуации был предложен в концепции адаптивных V-критиков [12] (рис. 4).

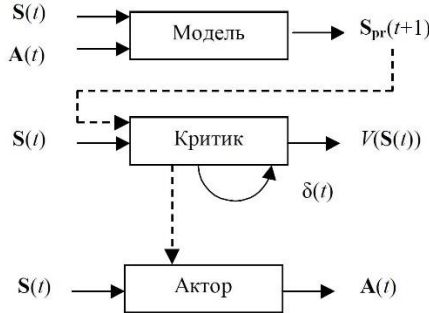


Рис. 4. Адаптивный V-критик

V-критик содержит модуль Актора, модуль Критика и модуль Модели. Актор - это тот, кто выполняет действие, а Критик это тот, кто выдает оценку качества ситуации. Модуль Модели способен строить прогнозы, то есть, имея ситуацию и имея предполагаемое действие, он может выдавать картину той ситуации, в которой мы окажемся, по его мнению, после совершения этого действия.

Преимущество, которое несет V-критик по сравнению, например, с Q-критиком - это то, что Критик выдает не оценку качества действия в какой-либо ситуации, а оценку качества самой ситуации, независимо от того, какое действие к этой ситуации привело. Соответственно, в такой конструкции мы можем пытаться перебирать доступные нам в текущей ситуации действия, проверять, какое из этих действий ведет к наилучшему прогнозу оценки качества ситуации, и выбирать его.

Такой подход хорошо ложится на сделанное нами предположение о строении коры. Миниколонки соответствуют различным контекстам. Если контексту соответствует некое действие, то его правила трактовки – это правила изменения описания таким образом, чтобы смоделировать результат применения этого действия. То есть, трактовка в данном случае – это

описание того, что будет, если это действие будет совершено. По сути, контекст действия, осуществляя трактовку, выполняет функции модуля Модели.

Каждый контекст обладает собственной копией памяти. Причем память может содержать не только описание запомненного события, но и оценку качества ситуации (оценку переживаний), которая этому событию соответствовала. Такая память позволяет, найдя все «похожие» события, сформировать прогноз оценки для той трактовки, что присутствует в контексте. То есть выполнить функцию модуля Критика. За счет того, что каждый контекст имеет собственную копию памяти, оказывается возможным не только сделать одновременное параллельное моделирование возможных поступков, но и дать каждому из них оценку качества той ситуации, к которой они приведут.

В таком подходе может быть реализовано многошаговое прогнозирование.

При формировании прогноза оценки качества ситуации может быть получена не только сама оценка качества, но и оценка достоверности прогноза. Аналогично, при определении смысла ситуации, может быть получена оценка уверенности в найденной аналогии. Оценка уверенности может использоваться в дальнейшем. Так, например, при формировании поступка исследовательское поведение может быть получено при выборе контекста с высоким прогнозом оценки при достаточно малой ее достоверности.

Описания, сохраняющие основной смысл

Вычислительные эксперименты показали, что масштабы миниколонок позволяют успешно работать с относительно короткими описаниями, состоящими из 5-7 понятий. Это означает, что для реализации описанной схемы требуется использовать описания, избавленные от несущественных подробностей, и содержащие “главную суть”. Такие описания позволяют, например, представить результаты действий в лаконичных формулировках по принципу: (дверь к сокровищу закрыта) в контексте «открыть» приводит к (дверь к сокровищу открыта).

Формат таких лаконичных описаний, то есть, грамматика внутреннего языка мозга, и механизм их получения – предмет изучения. Можно предположить, что не существует жесткого алгоритма, позволяющего получать описания, определяющие суть ситуации. И можно предположить, что такие

описания сами могут быть получены как результат обучения с подкреплением. То есть, мыслительная деятельность подчиняется тем же законам и принципам, что и процессы формирования поведения.

Выводы

Сделано предположение о возможной схеме, позволяющей мозгу использовать механизм обучения с подкреплением. Высказано предположение о природе формирования описаний, сохраняющих главный смысл. Есть надежда, что развитие контекстно-смыслового подхода применительно к обучению с подкреплением позволит получить результаты, сопоставимые с возможностями биологического мозга.

Список литературы

1. <https://habr.com/post/308268/>
2. Rudolf Wille Restructuring lattice theory: An approach based on hierarchies of concepts // Nato Science Series C. 83. Springer Netherlands. 1982 pp. 445–470.
3. Redozubov A. Holographic Memory: A Novel Model of Information Processing by Neuronal Microcircuits // The Physics of the Mind and Brain Disorders, Springer 2017
4. LeCun Y., Bengio Y. Convolutional networks for images, speech, and time-series // MIT Press, Cambridge. 1995
5. Gerald M. Edelman, Vernon B. Mountcastle The Mindful Brain: Cortical Organization and the Group-Selective Theory of Higher Brain Function // MIT Press 1978
6. Peters, A., and Yilmaz, E. Neuronal organization of area 17 of cat visual cortex // Cerebral Cortex 3, 1993 pp. 49–68.
7. Sara Sabour, Nicholas Frosst, Geoffrey E. Hinton Dynamic Routing Between Capsules // 2017, <https://arxiv.org/pdf/1710.09829.pdf>
8. <https://habr.com/post/309366/>
9. <https://habr.com/post/308878/>
10. <https://habr.com/post/326334/>
11. Редозубов А. Логика эмоций // СПб.: Амфора, 2011
12. Prokhorov D., Wunsch D. Adaptive critic designs // IEEE Trans. on Neural Networks. 1997. Vol. 8. N.5. P.997-1007.