



СХЕМА КОНСТРУКЦИИ ИЗ НЕЙРОПРОЦЕССОРОВ, СПОСОБНОЙ РЕАЛИЗОВАТЬ ОСНОВНЫЕ ФУНКЦИИ МЫШЛЕНИЯ И НАУЧНОГО ТВОРЧЕСТВА

Д.С. Чернавский, В.П. Карп, А.П. Никитин, О.Д. Чернавская

Предлагается вариант построения конструкции из нейропроцессоров, которая потенциально способна решать задачи, традиционно относимые к творческим. Обсуждается роль условной информации, предлагается конструкция блока образования символа. Выделена подсистема, способная решать логические задачи. Продемонстрировано, что при распознавании процесса и его прогнозировании в символической подсистеме возможна интерполяция и экстраполяция, что приводит к понятию континуального времени. Показано, что решение творческих задач (при недостатке информации или противоречивости алгоритмов) в символической подсистеме невозможно. Обсуждаются понятия интуитивного и логического мышления применительно к нейрокомпьютерингу и их реализация в рассматриваемой схеме; предлагается концепция перехода от интуитивного к логическому.

Ключевые слова: Мышление, нейрокомпьютеринг, самоорганизация, динамическая теория информации, динамическая модель, символическая система, генерация информации, научное творчество.

Введение

В последнее время интерес к проблеме искусственного интеллекта вновь переживает несомненный подъем, на сей раз в духе синергетического подхода (см. [1–4] и ссылки *ibid*) с привлечением динамической теории информации [5,6]. На этом фоне появилось устойчивое словосочетание «пути моделирования мышления». Оно отражает тот факт, что единый подход к моделированию процесса мышления не выработан, *mainstream* в этой области пока не образовался. В данной работе проблема рассматривается с позиций представлений динамической теории информации, теории распознавания образов и нейрокомпьютеринга. Предлагается один из возможных вариантов конструкции, собранной из нейропроцессоров и потенциально способной решать определенные задачи научного мышления.

Вопрос «*Что такое мышление?*» исследовался авторами давно и активно [7–9]. В работе [9] было предложено определение мышления как *самоорганизующегося процесса записи, сохранения, обработки, обобщения и генерации новой информации*. Под *обработкой* информации при этом понимается распознавание образа (процесса), классификация, кодирование и прогноз развития событий, под ее *обобщением* – интерполяция и экстраполяция результатов обработки. Говоря о *генерации* новой информации, авторы подразумевают, прежде всего, творчество во всех его проявлениях. Наконец, определение «*самоорганизующийся*» делает акцент на том, что алгоритмы обработки и обобщения создаются в результате самоорганизации, а не привносятся кем-либо со стороны, извне.

В динамической теории информации ключевым является следующее определение. *Информация – запомненный выбор одного варианта из N возможных и равноправных* [10]. Таким образом, *генерация информации* предполагает *свободный* (случайный) выбор. Под *рецепцией информации* понимается выбор, *навязанный извне* (в результате прямого наблюдения или прямого указания). В этой связи большую роль играет *условная информация* – этот термин относится к выбору, возникшему в коллективе (обществе) в результате договоренности. Примеры условной информации (код, язык) обсуждались в [6].

Ценность информации зависит от цели, для которой она может быть использована. В рассматриваемом контексте базовая цель индивидуума – создание, сохранение и распространение «своей» условной информации. Конкретные, локальные цели, которые ставит перед собой человек, разнообразны, но подчинены базовой.

Важно, что задача сохранения (запоминания) информации и задача генерации информации *дуальны*, то есть дополнительные. В частности, в результате генерации новой информации практически всегда исчезает часть «старой» информации. Следовательно, как было показано в [8], для реализации этих задач принципиально необходимы *две разные подсистемы*, которые можно условно соотнести с левым и правым полушариями головного мозга.

Ниже мы будем использовать идеологию и термины теории распознавания образов, как то: *обучающее множество* – набор *известных системе* объектов (процессов); *целевое обучающее множество* – набор объектов, связанных с определенной целью; *пространство признаков* – многомерное пространство, в котором точка соответствует *образу* объекта.

Цель данной работы – предложить конкретную схему искусственной конструкции, способной решать определенные (прежде всего, научные) задачи мышления. Подчеркнем, что речь пойдет именно об искусственной конструкции, то есть на моделирование мыслительного процесса человека мы не претендуем, хотя и используем в своем рассмотрении ряд соответствующих аналогий.

1. Логическое и интуитивное мышление

Поясним, в каком смысле мы будем оперировать понятиями «логическое мышление» и «интуитивное мышление».

Термин *логика* означает искусство *убеждать словом*. Уже отсюда следует, что логическое мышление теряет смысл без владения общепринятым языком, то есть кодом, следовательно, оно обязано оперировать *символами* (но не образами).

К *логическому мышлению* следует относить решение задач на основе формализованных алгоритмов, выраженных на общепринятом языке. Под словом *алгоритм*

мы понимаем *способ вынесения суждения*. Алгоритм может быть оформлен в виде аксиомы, являющейся *обобщением опыта*; графа причинно-следственных связей; непрерывной или дискретной динамической математической модели и т.д. При этом причинно-следственные связи, лежащие в основе стандартных алгоритмов, *обще-признанны*, то есть имеют ранг *почти законов*, что и обеспечивает возможность убедить других в своих выводах. Таким образом, логические алгоритмы не индивидуальны и для подтверждения своего статуса требуют доказательства.

Отметим, что допускается потеря однозначности алгоритма при возникновении неустойчивости; в последнем случае требуется дополнительная информация, которая может быть получена путем *генерации* информации (чисто случайно или на основе индивидуальных привычек, эмоций или ассоциаций) или путем *рецепции* недостающей информации извне. Эта задача решается уже вне логики.

Следует подчеркнуть, что такое «строгое» понимание логики серьезно отличается от того, что в обыденной жизни считается «логичным». Более «мягкое» понимание логики применительно к нейрокомпьютерингу обсуждалось в работе [8].

Под *интуитивным мышлением* будем понимать вынесение суждений, *не основанных* на общепринятых причинно-следственных связях. Интуитивное мышление базируется на обобщении опыта человека, точнее, на *всем* индивидуальном опыте. Этот тип мышления манипулирует *образами*, а не символами объектов, и оперирует *ассоциативными* связями между образами. Интуитивное мышление строго индивидуально – его результат формируется внутри индивидуума и не апеллирует к «доказательству». Оба типа мышления играют важную роль и часто рассматриваются как независимые и дополняющие друг друга [11]. Тем не менее возможен переход от интуитивного мышления к логическому.

2. Аппарат мышления как конструкция из блоков нейропроцессоров

Обсудим общие свойства аппарата мышления (далее – АМ). Согласно изложенному, он не должен предполагать непосредственное привнесение в него алгоритмов извне. Тем не менее внешние воздействия необходимы и сводятся к следующему:

- внесение первичной информации из рецепторных систем;
- постановка цели самим организмом или обществом;
- внедрение общепризнанного языка (кода), поскольку без обмена информацией с членами общества мышление невозможно.

АМ должен решать задачи распознавания образа, кластеризации и классификации, прогнозирования, принятия решения, творчества. Подчеркнем, их реализации в необученном (аналог мышления ребенка) и в обученном (мышление взрослого человека) АМ различны. Предлагаемая конструкция не претендует на описание задач взаимодействия с обществом и взаимодействия с организмом (влияние организма выступает как внешнее по отношению к АМ, обратное влияние не учитывается, хотя и имеет место). Мы не претендуем также на анализ *эмоций*, роль которых в интуитивном мышлении велика, но трудно формализуема.

Обсудим кратко свойства нейропроцессоров, из которых строится АМ.

2.1. Нейропроцессор Хопфилда [12] представляет собой пластину, на которой расположены бистабильные элементы – нейроны, соединенные связями. Число связей N больше числа нейронов n , но меньше максимального числа комбинаций $n < N < n(n - 1)$.

Согласно идеологии Дж. Хопфилда, каждый нейрон стремится «переключить» те нейроны, с которыми он связан, в «свое» состояние. Переключение происходит при достижении алгебраической суммы импульсов порогового значения (то есть нейрон – сумматор импульсов). Кроме того, имеются *внешние* по отношению к данной пластине связи.

Мы будем использовать следующую математическую модель:

$$\frac{dH_i(t)}{dt} = \frac{1}{\tau_i} \{ H_i + \beta_i (H_i^2 - 1) - H_i^3 \} + \sum_{j \neq i}^N \Omega_{ij} H_j + Z(t) \xi(t) + S(t), \quad (1)$$

где H_i – переменная нейрона со стационарными состояниями $H_i = +1$ (возбужденное), $H_i = 0$ (неустойчивое седло), $H_i = -1$ (невозбужденное); τ_i – характерное время; β_i – порог активации; связи между нейронами симметричны $\Omega_{i,j} = \Omega_{j,i}$; $Z(t)\xi(t)$ – шум (случайное хаотическое воздействие), амплитуда Z которого зависит от времени; $S(t)$ – предъявление образа, то есть возбуждение определенных нейронов по внешним связям.

Предъявление образа можно рассматривать как начальные условия, заданные на конечном отрезке времени – интервале обучения. Возбужденные извне нейроны являются признаками образа. Пространство признаков образует многомерный куб, координаты которого – состояния нейронов.

В процессоре Хопфилда ключевую роль играет обучение связей. Согласно оригинальной работе [12], обучение происходит следующим образом: при обработке образа возникают импульсы только между нейронами в противоположных состояниях. Связи исходно *сильные*, но при прохождении импульса слабеют, в чем и состоит механизм обучения. В этом случае уравнение для обучения имеет вид

$$\Omega_{ij}(t) = \Omega_0 \left\{ 1 - \frac{1}{2T} \int_0^t [1 - H_i(t') H_j(t')] \zeta(t') dt' \right\}, \quad (2)$$

$$\int_0^\infty \zeta(t') dt' \leq 1,$$

Ω_0 – исходная сила связи, T – интервал обучения.

Подобный метод обучения является по сути *фильтром*, способом «отсечения ненужного». Процессор, обученный таким образом, служит инструментом распознавания образа.

Можно предложить иной вариант обучения, обеспечивающий принцип *усиления* связей между нейронами при активации данного образа (другие связи, исходно *слабые*, при этом не затрагиваются). В этом случае математическая модель обучения имеет вид

$$\frac{d\Omega_{ij}(t)}{dt} \sim \frac{\Omega_0}{T_0} (H_i(t) + 1)(H_j(t) + 1)\zeta(t), \quad (3)$$

где $\zeta(t)$ – монотонная интегрируемая функция, обладающая эффектом насыщения. Ниже мы будем называть этот вариант обучения *альтернативным*. Такой процессор естественно использовать для *записи* информации.

В обоих случаях связи между нейронами, входящими в образ, в результате обучения оказываются *сильнее* связей между ними и нейронами, не входящими в образ.

Благодаря симметричности связей $\Omega_{ij} = \Omega_{ji}$ уравнения обученного процессора Хопфилда квазипотенциальны, то есть аналогичны уравнениям движения легкого шарика в вязкой жидкости в потенциальном поле по ландшафту, который содержит особенности, описываемые понятиями *лунки, водоразделы, сепаратрисы, русла, каналы* и т.п.

Результат распознавания в такой механической аналогии – достижение шариком определенной лунки. Координаты шарика в данный момент соответствуют положению изображающей точки в многомерном фазовом пространстве модели (3)

$$\frac{dH_i}{dt} = \frac{\partial \Xi(H_i, H_j \dots H_N)}{\partial H_i},$$

$$\Xi(H_i, \dots H_N) = - \sum_i \left\{ \frac{1}{2} H_i^2 - \frac{1}{4} H_i^4 + \beta \left(\frac{1}{3} H_i^3 - H_i \right) + \sum_j \Omega_{ij} H_i H_j \right\}. \quad (4)$$

Потенциальный ландшафт создается в результате обучения, причем лунки становятся четкими при достаточно продолжительном или многократном предъявлении образа.

Обученный процессор Хопфилда способен выполнять следующие операции.

- Запоминать предъявленные образы и отличать их друг от друга, то есть *распознавать*. Возможен отказ от распознавания, если образы слишком близки. При этом процессор заикливаясь, происходит так называемая «фрустрация».
- Классифицировать предъявляемые образы.
- Регулировать «внимание» как к определенному образу, так и к определенной части пластины. Для этого необходимо допустить действие неких *внешних связей*, не изменяющих состояние нейронов, но влияющих на порог переключения – параметр β .

Ниже мы будем использовать два типа процессоров Хопфилда, отличающихся характером обучающего множества. В первом типе обучающее множество построено не целесообразно и случайно. Такое *размытое* или «полное», множество должно быть велико и плотно: разные образы имеют общие возбужденные нейроны. Во втором типе обучающее множество невелико и построено целесообразно, путем многократного предъявления объектов, соответствующих данной цели – такое множество можно назвать «*четким*». Подобные множества формируются из размытого на базе определенного способа самоорганизации.

Оба типа играют важную роль в АМ, причем в «размытом» процессоре используется альтернативный способ обучения связей, а в «четком» – способ обучения «по Хопфилду».

2.2. Нейропроцессор Гроссберга [13] отличается от процессора Хопфилда функционально, иными свойствами характеризуются нейроны и их связи. Назначение процессора Гроссберга – локализация образа, то есть преобразование образа из нескольких возбужденных нейронов в один, который можно считать символом образа.

В пластине Гроссберга нейроны – бистабильные элементы, которые могут находиться в возбужденном (1) и пассивном (0) состояниях. Каждый нейрон способен к

самовозбуждению (порог переключения низок). Связи в общем случае несимметричны и обучаемы. Сигналы по связям могут посылать только возбужденные нейроны, стремящиеся при этом подавить все прочие. Число связей N больше n нейронов, но меньше максимального $n(n - 1)$.

Математическая модель процессора подробно рассматривалась в [14] и имеет вид

$$\frac{dG_i(t)}{dt} = \frac{1}{\tau_i} \{ -(\alpha - 1)G_i + \alpha G_i^2 - G_i^3 \} - \sum_{j \neq i} \Gamma_{ij}(t) G_i G_j, \quad (5)$$

где G_i – переменная нейрона, α – параметр, Γ_{ij} – подавляющие (благодаря отрицательному знаку) связи внутри пластины.

В необученной пластине Гроссберга при предъявлении образа посредством возбуждения нескольких нейронов происходит конкуренция нейронов за сохранение своего состояния. В результате в возбужденном состоянии остается лишь один нейрон-«победитель», который и уместно именовать «символом». Символом может стать как один из нейронов – представителей образа, так и нейрон, не входивший в предъявленный образ. Выбор по сути происходит случайно в силу неустойчивости начального состояния, а вероятность выбора зависит от связей; в области скопления нейронов предъявляемого образа она выше.

Уравнение обучения связей имеет вид (T_c – характерное время обучения):

$$\frac{d\Gamma_{ij}(t)}{dt} = \frac{1}{T_c} \{ G_i G_j (G_i - G_j) \}. \quad (6)$$

После первичного обучения символ образуется однозначно в том же месте, где и в первом акте. Однако при наличии шума и не слишком глубокого первичного обучения при повторении процесса возможно «переобучение» и смена символа. При образовании символа другого образа (отличного от предыдущего) процесс протекает аналогично, но нейрон – символ первого образа – уже не должен стать «победителем».

Фазовое пространство модели (5) – то же, что и модели (1), и совпадает с пространством признаков. Несмотря на то, что модель (5) не является квазипотенциальной и механическая аналогия здесь, строго говоря, неприменима, все же для фазового пространства имеют смысл понятия *аттракторы (лунки)*, *сепаратрисы (водоразделы)* и т.д. Плоскость пластины Гроссберга представляет собой свертку многомерного признакового пространства в двумерное. Координаты символа на этой плоскости можно считать «кодом» этого символа.

3. Блоки конструкции аппарата мышления

Конструкция АМ состоит из блоков, включающих упомянутые выше процессоры. Обсудим структуру и функции этих блоков.

3.1. Блок формирования целевого множества. Рассмотрим один из возможных вариантов процесса самоорганизации имеющейся информации. На рис. 1 представлены схемы блоков из двух процессоров. Слева – до формирования целевого множества, справа – после.

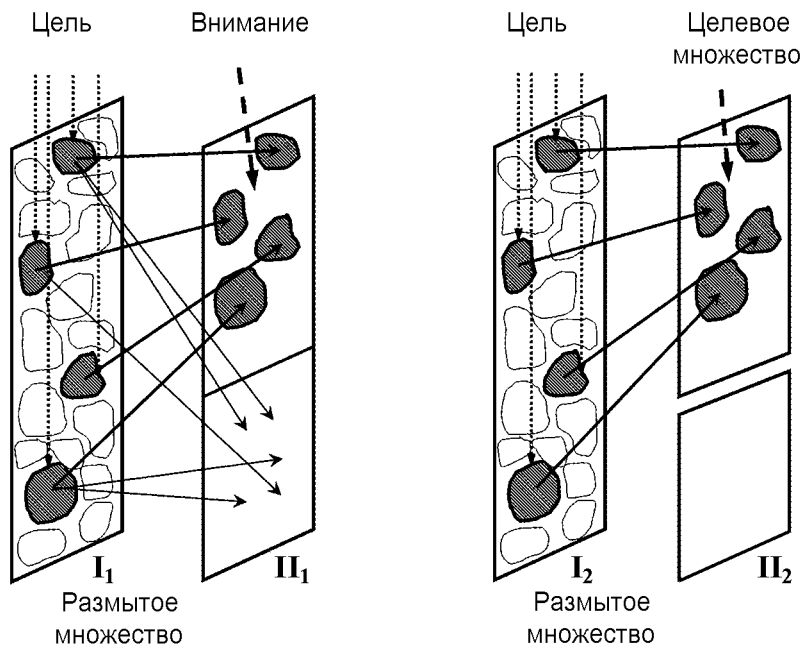


Рис. 1. Схема формирования целевого множества

На схеме слева пластина I_1 – процессор Хопфилда первого типа, содержащий *размытое* множество образов. В нем связи обучены альтернативным способом согласно (3). Пластина II_1 – пластина Хопфилда второго типа, где связи обучены «по Хопфилду» согласно (2). Межпластинные связи обильны, не обучены, но способны к обучению аналогично (3).

Постановка цели означает, что извне, путем многократного предъявления, активируются образы конкретных объектов, соответствующих данной цели. Связи между нейронами, входящими в размытые образы, усиливаются. На языке механической аналогии: лунки целевых образов углубляются. При достижении некоего критического значения образ передается на пластину II_1 по межпластинным связям. До их обучения образ может быть подан в произвольное место пластины II_1 , и даже не в одно, как показано на схеме слева. Во избежание этого, одновременно с постановкой цели повышается «внимание» в определенной области пластины II_1 (на рис. 1 – верхняя часть II_1). В силу этого все соответствующие цели образы локализуются в области повышенного внимания. При повторении процедуры межпластинные связи обучаются, так что связи, ведущие в иные области, ослабевают или даже исчезают.

Справа на рис. 1 представлена схема обученного блока из двух пластин с аналогичными свойствами. Межпластинные связи немногочисленны, благодаря тому, что из них оставлены только ведущие в ту область пластины II_2 , которая соответствует целесообразному множеству. В II_2 по мере поступления образов происходит «очищение», в результате которого в них остаются только нейроны, возбуждаемые наиболее часто – образуется *типичный образ*. Процедура очищения не однозначна и при ее повторении типичный образ может несколько измениться. В результате эта область пластины II_2 содержит набор четко различимых типичных образов, соответствующих поставленной цели.

При постановке другой цели внимание повышается уже в другой области пластины II_1 и процедура повторяется. При этом один и тот же образ размытого множества может войти в несколько разных целевых множеств.

Представленная схема является первым шагом целесообразной классификации накопленной информации. В ней число целей счетно и невелико. Число конкретных индивидуальных целей существенно больше, но они иерархически структурированы. Схема, изображенная на рис. 1, соответствует нижнему уровню иерархии. В ней каждый типичный образ содержит достаточно много признаков.

При формировании типичного образа на каждом уровне иерархии часть информации предыдущего уровня теряется. Если впоследствии она окажется важной, то это послужит поводом к повторению процедуры, включая обращение к размытому множеству.

3.2. Блок преобразования образа в символ. Сам процесс формирования символа из типичного образа подробно обсуждается в [15,16]. Здесь мы представляем его конкретную реализацию в конструкции АМ.

Блок схематически представлен на рис. 2. Он состоит из трех частей: верхней, нижней и соединяющей их пластины Integr. Верхняя и нижняя части идентичны и состоят из 4-х процессоров каждая.

Рассмотрим сначала нижнюю часть рис. 2. Символами I,H обозначена пластина Хопфилда, на которой представлен типичный образ. Он по прямым связям передается на пластину Хопфилда II,H. С нее образ по прямым связям передается на пластину Гроссберга III,Gr, где образ локализуется и выбирается первый случайный символ.

Между пластинами III,Gr и II,H существуют разветвленные обратные связи. Каждый нейрон пластины III,Gr связан с несколькими (но не со всеми) нейронами пластины II,H. По этим связям из символа идут сигналы в обратном направлении, и на пластине II,H возникает *инверсный образ*. Эта процедура является декомпозицией символа в инверсный образ, сравниваемый с исходным. Если отличие мало, то образ в силу свойств пластины Хопфилда восстанавливается до исходного, проверка считается успешной и символ закрепляется.

Если исходный образ сильно отличается от инверсного (а это – общий случай), то возникает фибриляция (шум), которая является сигналом к повторению процедуры. При этом выбирается новое положение символа – происходит отбор наилучшего. Проверка продолжается, пока отличие исходного образа от инверсного не

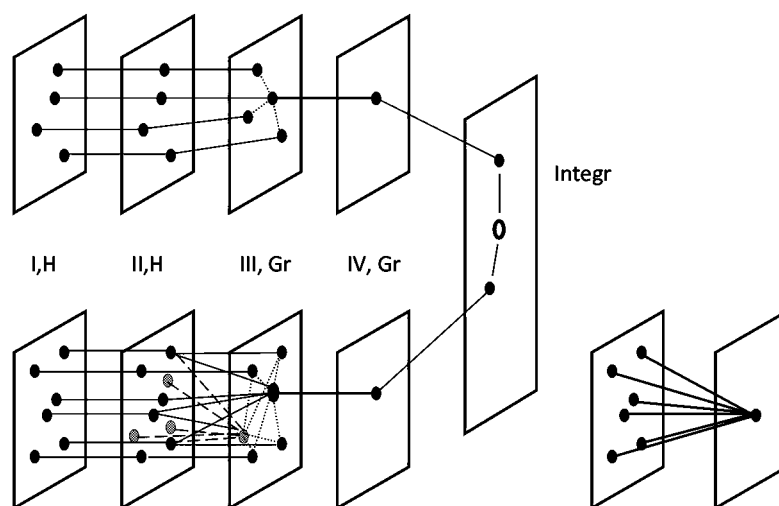


Рис. 2. Блок преобразования образа в символ

станет достаточно малым. Если проверка закончилась успешно, то символ передается по прямым связям в пластину IV,Gr. Таким образом, проверка осуществляется в процессорах II,Н и III,Gr. Именно эта пара и составляет блок *back propagation* (далее – ВР).

Итерационная процедура проверки может и не сойтись. Это означает, что блок закичивается – возникает «фрустрация», что является сигналом о том, что типичный образ нуждается в коррекции. Тогда процедура формирования типичного образа повторяется.

Когда процедуры проверки и коррекции как символов, так и образов отработаны, необходимость в ВР отпадает. Тогда образ сразу, без проверки передается по прямым межпластинным связям с I,Н на IV,Gr. Этот блок изображен на рис. 2 справа.

В случае успешной проверки возникает однозначный алгоритм преобразования образа в символ, и задача становится решаемой логически. Однако такой алгоритм индивидуален, поскольку при его построении присутствует элемент запомненной случайности (генерации информации – творчества). Код символа (его координаты на пластине IV,Gr) тоже индивидуален. Это означает, что в другом, по сути аналогичном, АМ код символа в общем случае будет другим. Поэтому отработанный в данном АМ однозначный алгоритм, строго говоря, нельзя считать логическим, так как он выражен на внутреннем, индивидуальном коде, а не на общепринятом языке. Его можно назвать невербализованной логикой – аналогично понятию «невербализованное мышление» в [1].

В верхней части рис. 2 представлен идентичный блок из четырех процессоров. В нем решается задача формирования символа образа того же объекта, но в другом пространстве признаков. Например, в нижнем блоке образ построен на признаках формы цветка, а в верхнем – на признаках запаха. Образы одного и того же объекта на верхних и нижних пластинах I,Н и II,Н различны, равно как и символы на пластинах IV,Gr. Для формирования единого символа, учитывающего все признаки, предназначена пластина Integr гроссберговского типа, на которую поступают символы с верхней и нижней пластин IV,Gr. Заметим попутно, что схема такого блока была впервые построена для описания функций аутодиагностической системы человека. Структура блока была сопоставлена с реальным органом (пластины Рекседа спинного мозга) и было обнаружено удивительное сходство [7].

3.3. Блок обобщения – интеграция символов. Выше был рассмотрен простейший случай интеграции символов. В действительности интеграция информации важна для выполнения еще двух функций.

Первая – возникновение общепринятого языка. Для этого необходимо рассмотреть несколько (хотя бы два) однотипных АМ, в каждом из которых имеется блок, изображенный на рис. 2, а в пластинах IV,Gr символы одного и того же объекта отображаются в одном и том же или схожем пространстве признаков. Как уже говорилось, в силу индивидуальности, сочетания образы↔символы в разных АМ в общем случае различны. Для формирования единого символа необходимо допустить наличие в каждом АМ интегрирующей пластины Integr. Необходимо также наличие связей (не нервных!) между этими пластинами из разных АМ. Имея иную природу, они должны передавать информацию о сочетаниях образ↔символ из одного АМ в другой АМ. В результате «конкуренции» символов в пластинах Integr может происходить коррекция индивидуальных символов. В целом эта проблема выходит за границы текущего рассмотрения и может быть решена в рамках модели «борьбы условных информаций» [6]. В ней главную роль играют не связи внутри АМ, а связи

между АМ – между «индивидуумами». Подчеркнем, что сказанное имеет отношение только к искусственным системам.

Вторая дополнительная функция интеграции информации – обобщение результатов, полученных в разных целевых множествах. Очевидно, что задача обобщения иерархична и имеет несколько уровней. Процедура интеграции символов из разных целевых множеств фактически соответствует схеме на рис. 2. Каждый акт интеграции проходит проверку и коррекцию (блок ВР) вплоть до обращения к размытому множеству. Если процедура успешна, то возникает однозначный алгоритм и необходимость в блоке ВР опять-таки пропадает.

3.4. Блок распознавания процесса и прогноза. В теории распознавания прогноз строится на основе прецедентов, то есть уже наблюдавшихся процессов. В АМ процесс наблюдается «покадрово», в дискретные моменты времени t_0, t_1, \dots, t_n распознается образ, строится его символ, положение символа проверяется блоком ВР. Далее полагаем, что, если интервал между кадрами (*скважность*) мал, то образы меняются мало, а их символы близки.

На рис. 3 приведена схема покадрового распознавания образов и формирования их символов в дискретные моменты времени t_0, t_1, \dots, t_n . На схеме слева отражены блоки проверки ВР. Они существенно замедляют процесс формирования символов. Как следствие, скважность велика, и символы могут быть расположены далеко друг от друга.

На схеме показана финальная пластина Fin, на которую последовательно передаются и где запоминаются символы. Внутрипластинные связи в ней практически отсутствуют, так как основное ее назначение – запечатлеть весь процесс с начала и до конца. Результат запоминания можно схематично представить в виде ломаной линии.

Если алгоритм обработки каждого кадра отработан и в этом контексте однозначен, то блоки ВР не нужны. Справа представлена упрощенная схема без блоков ВР, в ней покадровое распознавание ускорено и скважность существенно меньше. В результате на финальной пластине Fin процесс выглядит не как ломаная линия, а

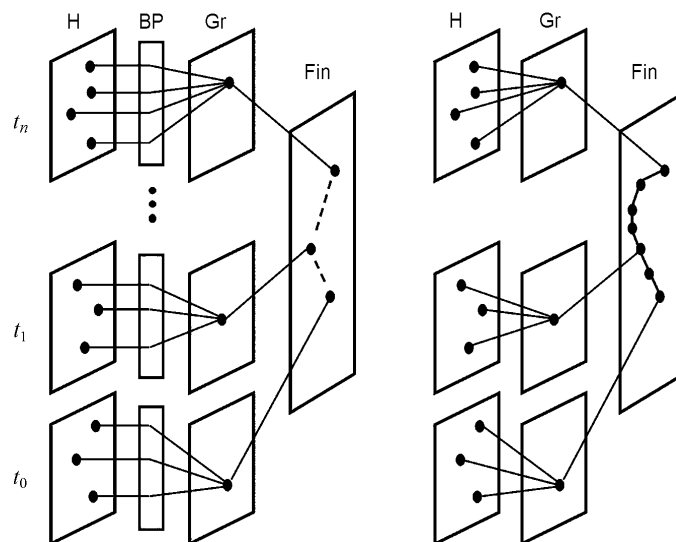


Рис. 3. Схема покадрового распознавания образов

почти как непрерывная кривая. Заметим, что эта кривая является аналогом траектории изображающей точки в математической модели процесса.

Таким образом, появляется возможность ввести понятие континуального (непрерывного) времени. Это позволяет проводить интерполяцию и экстраполяцию процесса во времени, то есть осуществлять прогнозирование. Отметим, что прогноз может оказаться ошибочным, если в какой-то момент извне поступает информация, противоречащая алгоритму. В левой части схемы на рис. 3 время воспринимается как дискретное, а при большой скважности теряются понятия «до» и «после».

4. Архитектура АМ – комбинация блоков

На рис. 4 приведена схема АМ, на которой объединены упомянутые выше блоки.

Слева представлено размытое множество образов. За ним следуют целевые множества, обсуждавшиеся в подразделе 3.1. Непосредственно далее идет блок распознавания, формирования символов и их интеграции, а также блок проверки и коррекции ВР. Именно в этой совокупности блоков «U» (обведена пунктиром) возникает *однозначный алгоритм* перехода образ→символ, который согласно определению в разделе 1 является актом *логического мышления*, выраженного на внутреннем коде («невербализованная» логика). Переход к вербализованной логике, то есть к логике в строгом понимании, является социальной задачей и на схеме не показан.

Блок ВР по своему построению содержит элемент *случайности и неоднозначности*, именно в нем происходит *генерация* информации (творческий процесс). Согласно идеологии, сформулированной в разделе 1, его следует отнести к инструменту *интуитивного мышления*. На основании всего сказанного можно предложить алго-

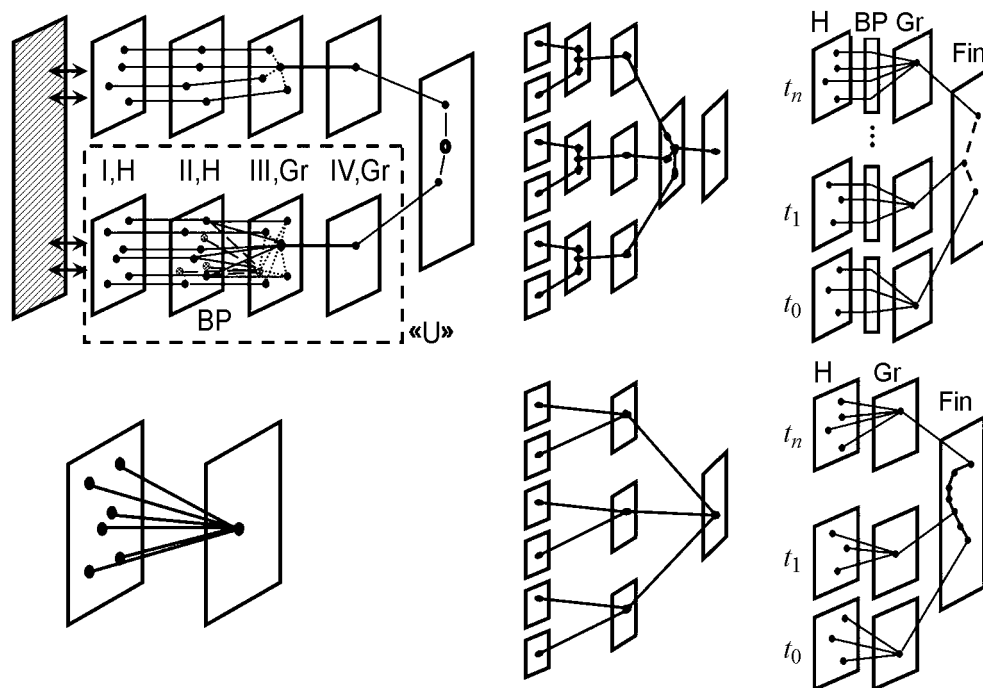


Рис. 4. Архитектура аппарата мышления

ритм выделения логической и интуитивной подсистем АМ: присутствие блока ВР в подсистеме служит индикатором ее «интуитивности», отсутствие – индикатором «логичности».

Вся верхняя часть схемы содержит блоки ВР и расположена в определенной области конструкции АМ, которую мы будем условно называть *пред-логической подсистемой* (ПП). Очевидно, что этот термин апеллирует к ассоциациям с правым полушарием человека (и был уже использован в работе [8] именно в этом качестве). Однако настаивать здесь на этой аналогии мы не можем, поскольку соответствие АМ реальному мышлению человека не обсуждается, и такая претензия была бы слишком смелой.

В нижней части схемы приведены те же блоки, но принято, что однозначные алгоритмы уже выработаны, и поэтому размытое множество и блоки ВР отсутствуют. Эта часть расположена в другой области АМ, которую уместно назвать *логической подсистемой* (ЛП) с апелляцией к левому полушарию человека (со всеми оговорками, приведенными выше). В ней присутствуют типичные образы, символы и только логические алгоритмы.

Отметим, что концепция ЛП и ПП, предложенная в [8], и рассматриваемая здесь конструкция АМ не противоречат друг другу в главном: *элемент случайности* должен отсутствовать в ЛП и присутствовать в ПП. В остальном отличия связаны с тем, что в [8] структура не конкретизировалась. В данной работе, напротив, основной целью является построение конкретного варианта самоорганизующейся конструкции, нацеленной на определенные (но далеко не все) задачи мышления. Иными словами, мы не отрицаем возможность АМ, построенного на базе иных способов самоорганизации, однако считаем, что *принцип разделения* «логической» и «интуитивной» подсистем всегда должен быть связан с *фактором случайности* (неоднозначности).

Важно, чтобы подсистемы ЛП и ПП могли работать координированно. В предлагаемой схеме это возможно, так как символы одинаковых образов в ЛП и ПП имеют одинаковый код, поскольку обученная структура первого получена копированием второго. Полагаем, что одинаковые символы связаны друг с другом так, что возбуждение одного из них в ЛП может приводить к возбуждению его аналога в ПП и наоборот.

Обсудим, как функционирует АМ, а для этого рассмотрим возможные механизмы решения типичных задач из сферы научного мышления.

Во-первых, существуют задачи, в которых однозначный алгоритм решения существует и имеется необходимая и достаточная информация для его реализации. Примером может служить решение уравнений с заданными начальными и граничными условиями. Последние задаются в АМ извне в виде набора символов – возбуждением соответствующих нейронов в ЛП. Далее задача решается в ЛП без участия ПП.

Во-вторых, распространены задачи, в которых информации для однозначного ответа достаточно, но алгоритм его получения отсутствует (например, задача Эйнштейна, головоломки типа sudoku и т.д.). Задача ставится извне в символическом виде и ответ должен быть получен в той же форме. Решение может быть получено на числовом компьютере методом полного перебора, но число вариантов при этом комбинаторно велико. В АМ такая задача может быть решена в конструкции «U», но без коррекции образа и без обращения к размытому множеству. Речь идет о последовательности актов выбора наилучшего символа.

В-третьих, есть задачи, где алгоритм решения существует, но информации для его реализации недостаточно. В АМ такая задача ставится предъявлением нетипичного образа (ряд признаков «типичного» в нем отсутствует). Задача решается в конструкции «U» в ПП.

Если нехватка информации незначительна, то в пластине Хопфилда образ самодостраивается до типичного, что и является решением. Если недостает большого объема информации, то в пластине возникает фрустрация, что служит сигналом для обращения к размытому множеству. В нем, по ассоциации с предъявленным, выбирается один из типичных образов, и задача решается. Такой выбор не однозначен и потому является актом творчества.

Обсудим, какое отношение к сказанному имеет разрешение научных парадоксов. Задача ставится изнутри, в символьном блоке на высоком уровне иерархии, при невозможности интеграции символов предыдущего уровня. Задача становится актуальной при появлении объекта, признаки которого в алгоритмах ЛП несовместимы, а решение логическим путем невозможно. В АМ проблема проявляется в том, что при попытке интеграции алгоритмов наступает фрустрация. Она и является идущим изнутри сигналом к постановке задачи.

Решается такая задача в ПП путем декомпозиции символов вплоть до образов размытого множества. В нем ищется (по ассоциациям) новый типичный образ, обладающий признаком, который ранее не был учтен. Этот образ перемещается в целевое множество. Далее происходит коррекция целевых множеств и алгоритмов обработки информации, как в ПП, так и в ЛП. После этого интеграция алгоритмов становится возможной.

Показательным примером может служить проблема обратимости. В алгоритмах динамики Ньютона все процессы обратимы во времени, в статистической физике и термодинамике процессы необратимы – энтропия может только возрастать. Рассмотрим бильярд Больцмана – набор упруго сталкивающихся шаров. С одной стороны, шары должны подчиняться алгоритмам динамики, с другой стороны, эта система – типичный пример статистического ансамбля и в ней должны действовать алгоритмы термодинамики.

Разрешение парадокса было дано в работах Н.С. Крылова [17] и Я.Г. Синая [18]: динамический процесс может стать необратимым, если он неустойчив. Процессы в бильярде Больцмана именно таковы. Новый признак, который не учитывался при построении алгоритмов динамики, – неустойчивость. Чтобы найти этот признак, необходимо было обратиться к размытому множеству. В нем устойчивые и неустойчивые динамические процессы ассоциативно связаны. Кроме того, ассоциативно связаны хаотические необратимые процессы и неустойчивость. Таким образом, признак «неустойчивость» породил новый типичный образ – неустойчивый динамический процесс и, далее, новый класс процессов – динамический хаос. Этот результат следует считать одним из главных достижений науки XX века.

Заключение

Предложен один из возможных вариантов самоорганизующейся системы (набора нейропроцессоров), реализующей определенные функции мышления. Эта искусственная конструкция – аппарат мышления – способна решать ряд задач мышления, в частности, разрешать логические научные парадоксы. В ее основе лежат положения динамической теории информации, теории распознавания и теории ней-

росетей. Вопрос, каково соответствие между предложенным АМ и реальным процессом мышления человека, не обсуждается. Совершенно ясно, что здесь необходим следующий шаг по пути интеграции наук, подключающий нейрофизиологию и нейробиологию. Тем не менее полагаем, что некоторые важные черты процесса мышления конструкция АМ описывает адекватно.

Перечислим наиболее важные результаты моделирования функционирования АМ.

- Выделено размытое обучающее множество, которое содержит всю накопленную информацию. Оно плотно, ассоциативно, но не структурировано.
- Описано формирование *целевых* образных множеств, то есть структуры памяти.
- Предложена схема блока нейропроцессоров Хопфилда и Гроссберга, который обеспечивает процедуру формирования символа образа. Структура блока фактически повторяет структуру пластин Рекседа спинного мозга.
- Выделена символично-образная подсистема, способная решать *логические* задачи и формулировать научные парадоксы.
- Показано, что решение *творческих* задач (при недостатке информации или противоречивости алгоритмов) в логической подсистеме невозможно. Они решаются путем обращения к размытому множеству, что естественно ассоциировать с интуитивным мышлением. В рамках обсуждаемой схемы предложена концепция трансформации интуитивного процесса мышления в логический.
- При распознавании и прогнозировании процесса в *символьной подсистеме* возможна интерполяция и экстраполяция, при этом возникает понятие *континуального времени* (в отличие от кадрового представления в образной подсистеме).

Изложенное в работе можно считать попыткой построить в духе синергетики «мост» между естественными и гуманитарными науками. Авторы уверены, что полный ответ на вопрос «Как человек мыслит?» возможен только на пути интеграции наук.

Работа выполнена при поддержке проекта 09-07-12135 ОФИ-М.

Библиографический список

1. Penrose R. Shadows of the Mind. NY–Oxford: Oxford University Press, 2005.
2. Шамис А.С. Пути моделирования мышления. М.: КомКнига, 2006.
3. Яхно В.Г. Проблемы на пути конструирования симулятора живых систем // Труды конференции «Нелинейная динамика в когнитивных исследованиях–2011». С. 246.
4. Голдберг Э. Управляющий мозг; Парадокс мудрости. М.: Поколение, 2007.
5. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам. М.: УРСС, 2005.
6. Чернавский Д.С. Синергетика и информация: Динамическая теория информации. М.: УРСС, 2004.
7. Чернавский Д.С., Карп В.П., Родитат И.В., Никитин А.П., Чернавская Н.М. Распознавание. Аутодиагностика. Мышление. М.: Радиотехника, 2003.
8. Чернавская О.Д., Никитин А.П., Чернавский Д.С. Концепция интуитивного и логического в нейрокомпьютеринге // Биофизика. 2009. Т. 54, № 6. С. 1103.

9. Чернавская О.Д., Чернавский Д.С., Карп В.П., Никитин А.П., Рожило Я.А. Процесс мышления в контексте динамической теории информации / Препринт ФИАН, 2011. № 10. 20с.
10. Quastler H. The emergence of biological organization. New Heaven, London: Yale University Press, 1964; Кастлер Г. Возникновение биологической организации. М.: Мир, 1967.
11. Фейнберг Е.Л. Две культуры. Интуиция и логика в искусстве и науке. М.: Наука, 1992.
12. Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities // PNAS. 1982. Vol. 79. P. 2554.
13. Grossberg S. Studies of Mind and Brain. Boston, Riedel, 1982; Nonlinear neural networks: Principles, mechanisms and architecture// Neural networks. 1988. Vol. 1. P. 17.
14. Чернавский Д.С., Карп В.П., Васильев А.Н., Чернавская О.Д. Математическая модель процессора локализации образа / Препринт ФИАН, 2011. №9. 19 с.
15. Чернавская О.Д., Чернавский Д.С., Карп В.П., Никитин А.П. О возможном механизме перехода «образ↔символ» в нейропроцессорах // Труды конф. «Нелинейная динамика в когнитивных исследованиях–2011». С. 224.
16. Чернавская О.Д., Чернавский Д.С., Карп В.П., Никитин А.П. О роли понятий «образ» и «символ» в моделировании процесса мышления средствами нейрокомпьютинга // Изв. вузов. ПНД. 2011. Т. 19, № 6. С. 5.
17. Крылов Н.С. Работы по обоснованию статистической физики. М.: УРСС, 2003.
18. Синай Я.Г. К обоснованию эргодической гипотезы для одной динамической системы статистической механики // ДАН СССР. 1963. Т. 158. С. 1261.

*Физический институт
им П.Н. Лебедева РАН*

Поступила в редакцию 13.07.2011

THE CONSTRUCTION SCHEME OF NEUROPROCESSORS ABLE TO REALIZE THE BASIC FUNCTIONS OF THINKING AND SCIENTIFIC CREATIVITY

D.S. Chernavskii, V.P. Karp, A.P. Nikitin, O.D. Chernavskaya

We propose a version of the neuroprocessor construction scheme, which in principle capable to solve problems commonly treated as creative work. The role of conventional information is discussed, and the specific block for the symbol formation is suggested. The system capable to solve logic problems has been pointed out. It is shown that the symbol (logical) subsystem is able to interpolate and extrapolate in the process of pattern recognition and prognosis. It is shown that creative problems (connected with the lack of information or the algorithm internal conflicts) could not be solved within the symbol (logical) subsystem. The notions of intuitive and logic thinking as applied to neurocomputing, and their realization in the given scheme is discussed. The concept of transformation of intuitive into logical thinking is presented.

Keywords: Thinking, neurocomputing, self-organization, dynamic theory of information, dynamic model, symbol system, generation of information, scientific creativity.

Авторские сведения см. на стр. 19–20.