

УДК 681.513.6, 004.89:004.4

Г.А. Прокопович

НЕЙРОСЕТЕВОЙ БЛОК ПАМЯТИ ДЛЯ АДАПТИВНОЙ РАБОТЫ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ В ДИНАМИЧЕСКОЙ СРЕДЕ

Рассматриваются современные методы управления сложными техническими системами, которые предназначены для автономной работы в динамической среде с априори неизвестными свойствами. Для придания подобным системам свойств интеллектуального адаптивного поведения предлагается нейроподобный блок ассоциативной памяти, способный производить классификацию образов на основе имеющихся эталонов, а также добавлять в память новые.

Введение

Поведение таких сложных технических объектов управления, как беспилотные летательные аппараты и автономные мобильные роботы, трудно поддается формализации и математическому описанию. Процесс аналитического описания поведения этих объектов усложняется еще и тем, что зачастую не только внешняя среда, но и свойства самих объектов априори неизвестны или изменяются в процессе функционирования. Характерная задача для подобных автономных систем: в неизвестной местности за конечный промежуток времени при ограниченном запасе энергии найти и доставить в точку старта искомый объект или информацию о нем, причем объект описан неполными начальными данными. Несмотря на сложность, в естественной природе данная задача является типичной и имеется масса наглядных примеров ее успешного решения.

Неудивительно, что разработка и изучение подобных систем, названных в [1] интеллектуальными автономными системами (ИАС), ставятся большинством ученых во главу угла исследований в области искусственного интеллекта (ИИ). ИАС должны обладать собственными мотивациями и активно взаимодействовать с внешней средой. Целью создания ИАС является изучение фундаментальных принципов функционирования и механизмов эволюции естественных систем управления, определяющих поведение организмов, а не отдельных аспектов их проявления.

Среди методов построения и моделирования ИАС особого внимания заслуживают те, которые основаны на моделировании нервных систем живых организмов. Как правило, в подобных бионических методах заранее неизвестна аналитическая модель объекта управления, а необходимые знания добываются эмпирически в процессе взаимодействия системы управления со средой и самим объектом управления.

В отечественной науке первыми шагами в этом направлении были разработка транспортной тележки ТАИР (транспортного автономного интегрального робота), основанной на использовании М-автоматов Н.М. Амосова [2]; моделирование поведения в программе «Животное» М.М. Бонгарда [3]; моделирование индивидуального развития, обучения и эволюции организма с помощью теории функциональных систем П.К. Анохина [4]; создание модели целесообразного поведения «Гиромат» Д.А. Поспелова [5]. В настоящее время на основе указанных работ, а также последних нейрофизиологических данных активно разрабатываются схемы управления адаптивного поведения *аниматов* – моделей систем, поведение которых следует принципам поведения живых организмов [6, 7]. Параллельно в исследовательской группе под руководством А.А. Жданова на основе собственной концептуальной модели нервной системы разрабатывается универсальный метод автономного адаптивного управления, который уже сегодня находит практическое применение [8].

1. Постановка задачи

Согласно классической парадигме ИИ компьютер фон-неймановской архитектуры, являющийся практической реализацией универсальной машины Тьюринга, может рассматриваться как теоретическая модель человеческого мозга. Следовательно, его структура должна содержать отдельные блоки, выполняющие функции центрального процессора, который управля-

ет операциями и выполняет вычисления, и памяти, в которой хранятся команды и данные о текущей программе. Также необходим канал связи, через который они должны обмениваться информацией. Однако данная парадигма еще не привела к разработке технических устройств, которые в полной мере можно назвать интеллектуальными [9].

Поэтому для разработки архитектуры системы управления (СУ) ИАС, которая должна решать задачи перемещения в пространстве и воспринимать окружающую обстановку, необходимо, как уже говорилось выше, использовать бионический подход, основанный на выявлении и использовании аналогий в реализации поведенческих актов живыми организмами. Поведенческие акты свойственны не только человеку, но и животным [5]. Таким образом, по аналогии с живыми организмами ИАС должна представлять собой активную систему, которая воспринимает информацию из среды (внешней и внутренней) при помощи рецепторов, интерпретирует ее согласно текущим целям и накопленному опыту и соответствующим образом воздействует на среду посредством эффекторов (рис. 1).

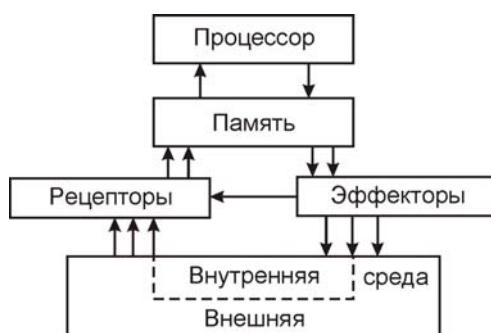


Рис. 1. Структурная схема системы «организм – среда»

В качестве примера рассмотрим функциональную схему СУ ИАС (рис. 2), предназначенную для оснащения комплексов летательных аппаратов и построенную на основе бионического подхода [10]. Данные, полученные от датчиков, проходят первичную обработку и классификацию в блоках распознавания ситуаций и формирования понятий. Блок памяти, представленный подблоками (базой данных и базой знаний), содержит знания о предметной области: описание объектов, явлений, фактов, а также отношений между ними. На основе результатов моделирования окружающей обстановки блок принятия решений производит выработку тех или иных управляющих воздействий. Несмотря на то что указанная функциональная схема содержит все блоки структурной схемы (см. рис. 1), в ней отсутствуют обратные связи между блоками процессора и памяти.

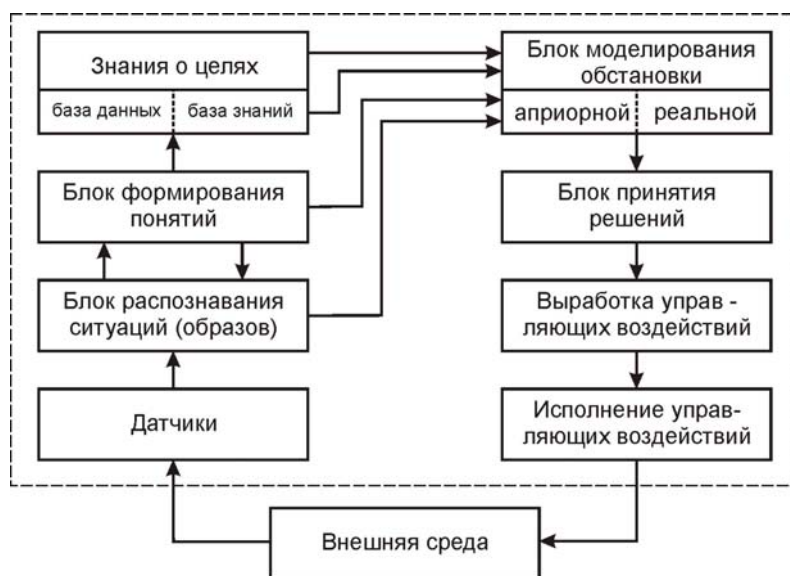


Рис. 2. Функциональная схема СУ ИАС

Большинство исследователей указывают на то, что биологическая память тесно связана с процессами мышления [11]. Так, в [5] на основе наблюдений и исследований реальных форм поведения было сделано заключение, что «богатство, разнообразие и вариабельность (форм поведения) существенно зависят от сложности и объема нервной системы, т. е. в первую очередь от свойств и возможностей памяти и в меньшей степени от возможностей оперирования с ее содержимым, от возможности преобразования хранящейся в ней информации». В следующей работе [12] приводится пример того, что в повседневной жизни поведение человека в некоторых ситуациях основано не столько на алгоритмах, сколько на интуиции, которая является хранилищем неформального опыта. Следовательно, можно сделать предположение, что проявление сложных форм поведения представляет собой организацию работы механизмов памяти. Тогда проявление любой формы поведения ИАС, последовавшей за изменениями внешней среды, должно восприниматься как реакция памяти на это изменение, выраженная специфическим действием или цепочкой действий эффекторов. Другими словами, память, находящаяся в гетероассоциативном режиме, на введенный специфический ключевой образ должна выдавать соответствующие отклики, представленные управляющими сигналами.

В контексте нейробиологии под памятью понимается относительно продолжительная во времени деформация структуры нейронов, вызванная влиянием на организм внешней среды. Без такой деформации память не существует [13]. Способность нервной системы изменяться под действием сенсорных синапсов, приспособляясь к свойствам внешней среды, обусловлена ее пластичностью, которая возникает в результате изменения эффективности и числа связей между нейронами [14, 15]. Чтобы память была полезной, она должна быть доступной для нервной системы. Только тогда она может влиять на будущее поведение организма. Однако для этого в памяти предварительно должны быть накоплены определенные модели поведения, что осуществляется с помощью процесса обучения.

Искусственные нейронные сети (ИНС), которые активно используются для построения нейрокомпьютеров и интеллектуальных СУ, как и их живые аналоги, способны обучаться в двух режимах. Рассмотрим указанные режимы обучения ИНС на примере управления ИАС.

Первый режим представляет собой процесс обучения с учителем, с которым ИАС должен находиться в состоянии отрицательной обратной связи. Учителю как эксперту следует заранее записать в память ИАС начальные знания, необходимые для его ориентации в пространстве и самосохранения, что соответствует созданию безусловных рефлексов. Также ИАС должна быть обеспечена набором знаний, необходимых для достижения поставленной экспертом цели.

Во втором режиме реализуется обучение без учителя, когда на основе полученных в процессе работы данных и фактов в памяти ИАС самостоятельно будут вырабатываться необходимые знания. Указанные самоорганизующиеся процессы, протекающие в системе управления ИАС, позволяют сгенерировать соответствующие условные рефлексы, которые помогут наилучшим образом приспособиться к среде даже в случае непредвиденных изменений собственной структуры (поломки).

Для реализации перечисленных свойств блок памяти ИАС должен обладать следующими свойствами: восстанавливать входные образы и сравнивать их с уже имеющимися эталонами, а также добавлять новые. Другими словами, по аналогии с биологической памятью разрабатываемый блок даже по произвольной части ключевого образа должен его классифицировать и определять соответствующий отклик либо по результатам его воздействия на ИАС создавать новую ассоциирующую пару.

2. Выбор прототипа

В качестве прототипа для разработки блока памяти для ИАС с заданными выше свойствами предлагается адаптивный нейросетевой классификатор, описанный в [16]. Данный классификатор является гибридом двух нейроподобных сетевых архитектур, каждая из которых проявляет различные внешние свойства биологической памяти: АРТ-1 (рис. 3) позволяет безопасно добавлять в долговременную память новые образы, не искажая старые, а двунаправленная ассоциативная память (ДАП) позволяет в параллельном режиме быстро производить поиск заданных образов. Предложенную нейроподобную гибридную сеть далее будем условно назы-

вать АРТ-ДАП. Дадим краткое обоснование выбора нейроразличного классификатора в качестве прототипа для разработки блока памяти.

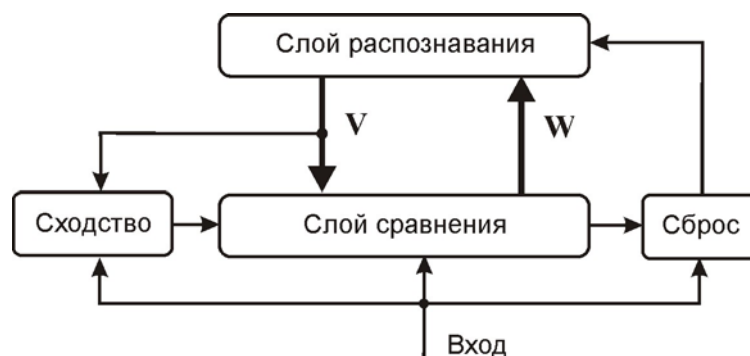


Рис. 3. Блок-схема сети АРТ-1

Проблема *стабильно-пластичного* запоминания образов в ИНС является одной из самых сложных и труднорешаемых задач при построении искусственных систем, моделирующих распознавание образов и восприятие окружающего мира [17, 18]. Она заключается в определении баланса между процессами запоминания нейронами новых образов и стабильного хранения образов, уже существующих в памяти.

В некоторых случаях процесс «затирания» старой информации является несущественным. Если набор обучающих векторов фиксирован, то, например, процесс обучения многослойного персептрона по методу обратного распространения ошибки будет заключаться в многократном циклическом предъявлении обучающей выборки. В итоге он запомнит весь пакет обучающей информации. Однако при обучении персептрона новому образу произойдет модификация синаптических связей с неконтролируемым разрушением структуры памяти о предыдущих образах. Таким образом, персептрон не способен к запоминанию новой информации без полного переобучения сети [18].

Аналогичная ситуация имеет место и в сетях Кохонена и Хэмминга [13, 17], обучающихся на основе принципов самоорганизации. Данные ИНС при классификации всегда выдают положительный результат. Поэтому они не способны отделить новые образы от искаженных или зашумленных версий старых образов. В реальной ситуации сеть подвергается постоянно изменяющимся воздействиям и может никогда не увидеть один и тот же обучающий вектор. При таких обстоятельствах сеть, скорее всего, не будет обучаться, так как будет непрерывно изменять свои веса, не достигая удовлетворительных результатов.

Между тем мозг человека ежесекундно способен выполнять трудную задачу обработки непрерывного потока сенсорной информации, получаемой из окружающего мира. Из огромного объема тривиальных данных ему приходится выделять и обрабатывать жизненно важную информацию и, возможно, регистрировать в долговременной памяти. Если некоторый образ для человека является новой информацией, реакция на него происходит поисково-познавательная, с сохранением этого образа в памяти, если же этот образ является вариантом уже знакомой информации, реакция организма соответствует ранее накопленному опыту и специальное запоминание образа не требуется. Таким образом, восприятие человека одновременно пластично (адаптировано к записи новой информации) и при этом стабильно, т. е. не «затирает» в памяти ранее сохраненные образы.

По данной аналогии работает нейросетевой векторный классификатор АРТ-1 (см. рис. 3), разработанный на основе теории адаптивного резонанса С. Гроссберга [15, 17], в которой постулируется необходимость самоадаптации алгоритма сравнения новых образов с уже существующими в памяти. В основе этой теории лежит использование слоя распознавания «снизу вверх» (матрица синаптических связей W), который предсказывает тип наиболее подходящего образа, и слоя сравнения «сверху вниз» (V), который проверяет предположения первого слоя. Если входной образ совпадает с одним из хранящихся эталонов, возникает состояние адаптивного резонанса (усиление и продление нейронной активности).

Входной вектор классифицируется в зависимости от того, на какой из хранящихся в долговременной памяти эталонов он похож. Результат классификации выражается в форме возбуждения только одного нейрона-победителя из распознающего слоя при условии, что восстановленный в слое сравнения входной вектор окажется достаточно близким по расстоянию Хэмминга к своему первоначальному значению. В противном случае выбирается новый нейрон-победитель. Если входной вектор не соответствует ни одному из хранящихся образов, то в сети создается новый нейрон, содержащий образ этого вектора. В процессе сравнения происходит модификация не только входного вектора, но и синаптической матрицы. Если определено, что входной вектор похож на один из ранее запомненных образов с точки зрения определенного критерия сходства, то под его действием запомненный в долговременной памяти образ будет изменяться (адаптироваться) таким образом, чтобы в процессе обучения стать более похожим на входной вектор [15].

Несмотря на то что теория адаптивного резонанса предлагает эффективное решение проблемы стабильно-пластичной организации памяти в ИНС, ее реализация в сети АРТ-1 не лишена недостатков. Биологически неправдоподобным является то, что информация о каждом хранимом образе находится в своем отдельном нейроне. Отсюда вытекает один из главных недостатков АРТ-1 – последовательная организация поиска образов в долговременной памяти. При большом объеме накопленной информации это отрицательным образом сказывается на общем быстродействии сети [13, 17]. Также локальное распределение информации в сети отрицательно сказывается на ее безопасности и надежности: при потере одного нейрона теряется информация об одном эталоне. С другой стороны, нейронные сети адаптивного резонанса характеризуются процессами самоорганизации и обучаются без учителя. Поэтому эксперт не в силах заставить сеть объединить те или иные образы (даже близкие с точки зрения эксперта) в один класс, как требуется для решения конкретной задачи.

Для устранения перечисленных выше недостатков в [16] было предложено заменить в архитектуре сети АРТ-1 слой распознавания и сравнения двухслойной гетероассоциативной сетью ДАП. Сеть ДАП (рис. 4), как и широко известная сеть Хопфилда, способна к обобщению, вырабатывая правильные реакции даже на искаженные входы. Однако последняя является автоассоциативной. Это означает, что входной образ может быть завершен или исправлен, но не может быть ассоциирован с другим образом. Данный факт является результатом одноуровневой структуры ассоциативной памяти, в которой выходной вектор генерируется на тех же нейронах, на которые поступает входной вектор.

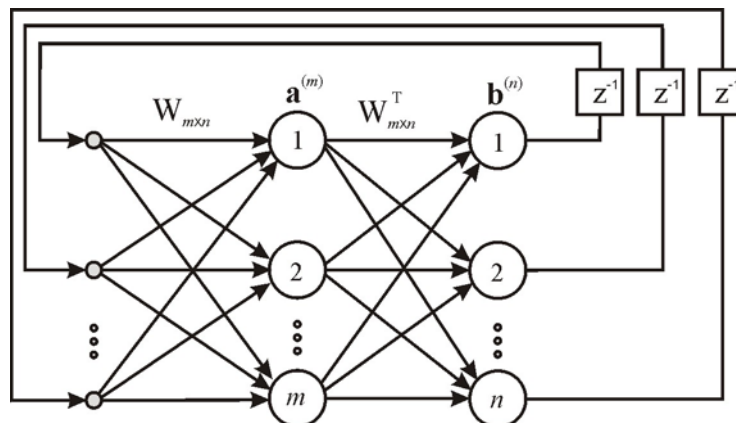


Рис. 4. Обобщенная структура сети ДАП

В двухслойной сети ДАП входной вектор поступает на один набор, а соответствующий выходной вектор вырабатывается на другом наборе нейронов. Таким образом реализуется гетероассоциативная память. Поэтому обученная сеть на поступающие двоичные ключевые векторы, обладающие общими характерными признаками, вырабатывает на выходе один и тот же вектор-отклик, т. е. наблюдается процесс адресации памяти по ее содержанию.

Процесс обучения сети ДАП реализуется в форме суммирования внешних (матричных) произведений всех векторных пар обучающего набора согласно правилу Хебба [17]:

$$W_{m \times n} = \sum_{i=1}^k \mathbf{a}_i^{(m)} \otimes \mathbf{b}_i^{(n)} = \sum_{i=1}^k \begin{bmatrix} a_1^i b_1^i & a_1^i b_2^i & \dots & a_1^i b_n^i \\ a_2^i b_1^i & a_2^i b_2^i & \dots & a_2^i b_n^i \\ \vdots & \vdots & \dots & \vdots \\ a_m^i b_1^i & a_m^i b_2^i & \dots & a_m^i b_n^i \end{bmatrix}, \quad (1)$$

где $W_{m \times n}$ – весовая матрица; m и n – количество нейронных элементов во входном и выходном слоях соответственно; $\mathbf{a}_i^{(m)}$ и $\mathbf{b}_i^{(n)}$ – обучающие пары векторов; k – число обучающих пар; \otimes – знак тензорного произведения [16], который необходим для корректной и однотипной записи формулы вычисления весовой матрицы при использовании одно- и многомерных векторов.

Из формулы (1) вытекает ряд свойств, которые качественно характеризуют ИНС, обучаемые с учителем. Одним из таких свойств является то, что при необходимости эксперт может обучить сеть ДАП извлекать (ассоциировать) искомый отклик даже при вводе неодинаковых ключей, описывающих различные образы.

Нейроны в сети ДАП функционируют наподобие других нейросетевых парадигм: сначала суммируются взвешенные входы, а затем вычисляется значение функции активации $f(x)$. Формула получения отклика сети в векторной форме выглядит следующим образом:

$$\mathbf{b}^{(n)} = f\left(W_{m \times n}^T \mathbf{a}^{(m)} - \mathbf{T}^{(n)}\right), \quad (2)$$

где $W_{m \times n}^T$ – транспонированная матрица весов $W_{m \times n}$; $\mathbf{T}^{(n)}$ – динамический вектор внешнего смещения [16].

При кратковременном предъявлении входного вектора на первом слое через транспонированную матрицу весов W^T на втором слое вырабатывается ассоциирующий вектор-отклик. Затем полученный вектор воздействует через матрицу весов W на первый слой, восстанавливая значения исходного вектора:

$$\mathbf{a}^{(m)} = f\left(W_{m \times n} \mathbf{b}^{(n)} - \mathbf{T}^{(n)}\right). \quad (3)$$

Каждый следующий цикл вызывает уточнение выходных векторов первого и второго слоев до тех пор, пока не будет достигнута точка стабильности в сети, при которой ни первый, ни второй векторы не будут изменяться. Эта точка в фазовом пространстве может быть рассмотрена как резонансная, так как вектор передается между слоями сети в прямом и обратном направлениях, всегда вырабатывая, но не изменяя их текущие выходные значения [16].

Рассмотрим принцип работы сети АРТ-ДАП (рис. 5). Видно, что в предложенном классификаторе имеются основные функциональные блоки сети АРТ-1. Однако благодаря наличию ассоциативного способа извлечения информации из сети ДАП и предложенного в [16] механизма сравнения входного образа с хранящимися в долговременной памяти эталонами процесс классификации стал значительно быстрее (сравнительная характеристика работы двух нейросетевых классификаторов будет представлена ниже).

В блоке «Сходство» каждый k -й новый вектор-отклик кодируется k -м ортом в пространстве R^q , где $q = 1, 2, \dots, k$ – максимальное количество записываемых образов. По принятому в [16] правилу входной вектор считается распознанным, если бинарный вектор-отклик содержит только одну единицу и номер его орта не превышает число k (количества хранимых образов). Во всех других случаях вектор считается нераспознанным и, следовательно, необходимо генерировать новый класс.

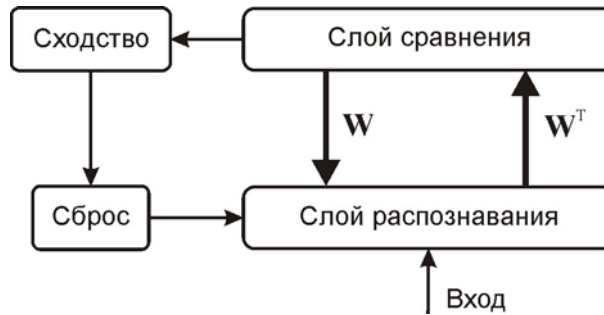


Рис. 5. Блок-схема сети АРТ-ДАП

Блок «Сброс» останавливает циркуляцию информации между слоями нейронов в одном из двух случаев: при утвердительном решении классификатора (распознавании входного вектора), при превышении наперед заданного числа произведенных итераций.

3. Принцип работы блока памяти

Предлагаемый блок памяти будет эффективным средством для разработки СУ ИАС, основанной на системе классификаторов [19]. Данный подход может быть рассмотрен как специальный случай теории обучения с подкреплением [20]. Системы классификаторов работают на основе множества правил (классификаторов), имеющих вид

Если $\langle \text{условие}_1 \rangle, \langle \text{условие}_2 \rangle, \dots, \langle \text{условие}_n \rangle$, то $\langle \text{действие} \rangle$,

где условие – правила, описывающие ситуацию, к которой применим данный классификатор, а заключение – действие, которое необходимо реализовать. При этом каждый классификатор может иметь свой вес (силу классификатора), показывающий полезность его системы. При совершенствовании действий предпочтение отдается тем классификаторам, которые имеют больший вес. Другими словами, классификатор выберет тот класс действий, которому больше всего соответствуют начальные данные. Структура принятия решений показана на рис. 6, где $\mathbf{A} = \{a_1, a_2, \dots, a_m\}$ – вектор условий, $\mathbf{B} = \{b_1, b_2, \dots, b_n\}$ – класс ситуации.



Рис. 6. Система классификаторов на основе разработанного блока памяти

Для своего успешного функционирования блок памяти ИАС должен хранить информацию о различных объектах и явлениях окружающего мира. Так как в большинстве случаев информация в указанном блоке представлена большими бинарными векторами и не поддается зрительному восприятию, для примера его работы воспользуемся процессом классификации монохромных изображений букв латинского алфавита. На рис. 7 показана классификация обу-

ченной сетью АРТ-ДАП входных образов, представленных различными вариантами написания буквы «А» при заданном числе элементов в векторе-отклике $q = 8$ и количестве записанных в долговременную память образов $k = 4$, буквам «А», «В», «С» и «D» соответствуют номера 1, 2, 3 и 4. В результате входной образ распознан блоком памяти как класс № 1.

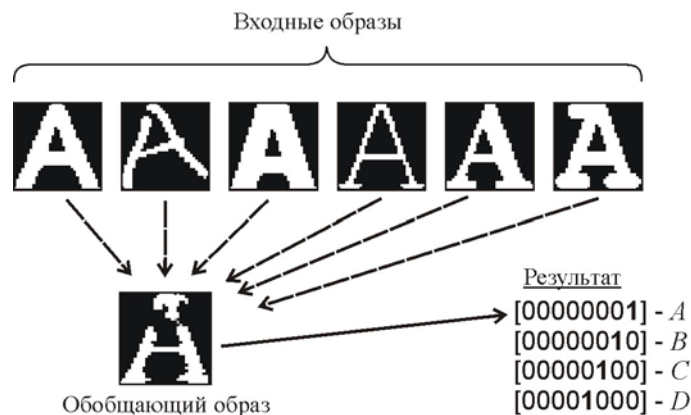


Рис. 7. Процесс ассоциации графических образов с одномерным бинарным вектором

Подобным образом блок памяти для ИАС способен корректно классифицировать все входные образы. Благодаря оригинальному методу представления векторы-отклики содержат в себе не только номер одного из известных классов, но и информацию о новизне.

4. Процесс дообучения

Как уже говорилось выше, в сети АРТ-1 память организована таким образом, что информация о каждом образе хранится обособленно в соответствующем нейроне. Поэтому добавление в долговременную память нового образа путем ввода в сеть нового нейрона не затронет обученные синаптические веса других нейронов. Однако это является и недостатком сети АРТ-1: она не может моделировать такую важную характеристику функций мозга, как распределенная память [17].

Сеть ДАП, напротив, демонстрирует преимущества распределенной памяти. На основании формулы обучения (1) можно сделать вывод, что добавление новой пары обучающей выборки в сеть АРТ-ДАП, используемой в блоке памяти ИАС, не требует полного пересчета матрицы весов $W_{m \times n}$, так как входные и выходные векторы в процессе обучения взаимодействуют только со своими парами. Поэтому процесс дообучения как безопасное добавление новых образов в долговременную память можно описать следующей формулой:

$$W_{m \times n}^{k+l} = W_{m \times n}^k + \sum_{i=1}^l \mathbf{a}_i^{(m)} \otimes \mathbf{b}_i^{(n)}, \quad (4)$$

где k – число записанных пар векторов; l – число новых пар обучающей выборки, которые необходимо добавить в долговременную память.

Как и сети Хопфилда, ДАП имеет ограничения на максимальное количество ассоциаций, которые она может точно воспроизвести. Если это число превышено, то сеть может выработать неверный выходной вектор, воспроизводя ассоциации, которым не обучена. Такие векторы обычно называются химерами или ложной (паразитной) памятью. Появление этих непредвиденных образов свидетельствует о том, что сеть ДАП способна не только ассоциировать выборку запомненной информации, но и активно преобразовывать исходную информацию, т. е. не является пассивным хранилищем информации. Примером может служить обобщающий образ буквы «А» (см. рис. 7).

Объем записанной в весовую матрицу информации зависит от соотношения числа входных и выходных нейронов, а также от корреляции между хранимыми образами. Так, в [17] приводится формула, по которой сеть ДАП может хранить только p случайных бинарных векторов:

$$p_{\max} = \frac{n}{2 \ln n}, \quad (5)$$

где n – количество нейронов в наименьшем слое. При наличии корреляции между векторами объем памяти (p) резко уменьшается.

Одним из возможных путей для решения поставленной задачи увеличения объема разрабатываемой памяти ИАС является использование алгоритма отображения распознаваемых бинарных образов в некие образы в пространстве иной размерности, при котором происходит декорреляция образов [21]. Данный алгоритм используется в так называемых векторных нейросетях, которые функционируют по аналогии с обучающим правилом Хэбба. Функционирование подобных сетей адекватно поведению системы q -мерных спинов. В [22] приведено описание векторной ДАП, в которой максимально возможное число хранимых образов достигает

$$p_{\max} = 2 + \frac{\ln n}{\ln q}, \quad (6)$$

где n – эффективное число выходных нейронов; q – мерность пространства, в котором хранятся преобразованные образы.

Следует указать, что с ростом q -го пространства в q^2 раз растет и число синаптических связей, так как величина межсвязи W_{ij} между i -м и j -м нейронами – не скаляр, как в обычной модели Хопфилда, а матрица $q \times q$.

5. Свойства разработанного блока памяти для ИАС

Синергетический эффект, полученный в результате проведенной выше замены, позволяет говорить о том, что свойства блока памяти ИАС, основанного на использовании сети АРТ-ДАП, можно считать приближенными к биологическим, так как они являются характерными для головного мозга человека.

Для более глубокого понимания работы блока памяти ИАС подробно рассмотрим свойства блока памяти ИАС и одновременно проведем сравнение принципов работы сетей АРТ-1 и АРТ-ДАП.

1. *Распределенное хранение информации.* В головном мозге человека информация хранится в виде своеобразной коллективной структуры. Так, в биологической памяти, основанной на распределенном хранении информации, каждый запоминающий элемент или участок запоминающей среды содержит следы многих образов. С другой стороны, каждый элемент записываемой информации распределяется по большой области запоминающей среды [13]. Данное явление наблюдается и в сети АРТ-ДАП, когда в процессе обучения с помощью корреляционной функции (1) элементы входного образа «размазываются» по индуцированным локальным полям всех выходных нейронов. Аналогично в процессе восстановления выходного образа по формуле (2) каждое значение отдельного элемента формируется путем опроса (суммирования) синаптических весов соответствующего нейрона.

2. *Адресация памяти по содержанию.* Как уже отмечалось в разд. 4, ассоциативное кодирование образов представляет собой либо автоассоциативный, либо гетероассоциативный процесс. Поиск образа происходит не по его адресу в памяти, а по содержанию (или даже по произвольной части) при условии, что эта часть достаточна для того, чтобы отличить данный образ от других, записанных в эту ассоциативную память. В отличие от АРТ-1, где для нахождения номера наиболее близкого образа необходимо произвести определенный последовательный перебор хранящихся в памяти образов, в АРТ-ДАП результат достигается путем единовремен-

ной ассоциативной выборки по формуле (2), требующей лишь незначительного числа итераций для уточнения значений входных и выходных нейронов.

В работе [16] приведено сравнение эффективности работы процессов обучения и распознавания образов сетями АРТ-1 и АРТ-ДАП, программно реализованных в пакете моделирования MatLab, на примере набора монохромных изображений (рис. 8). На рисунке видно, что предложенной сети АРТ-ДАП потребовалось значительно меньшее число итераций по сравнению с АРТ-1: 17 и 15 против 29 и 22 соответственно.

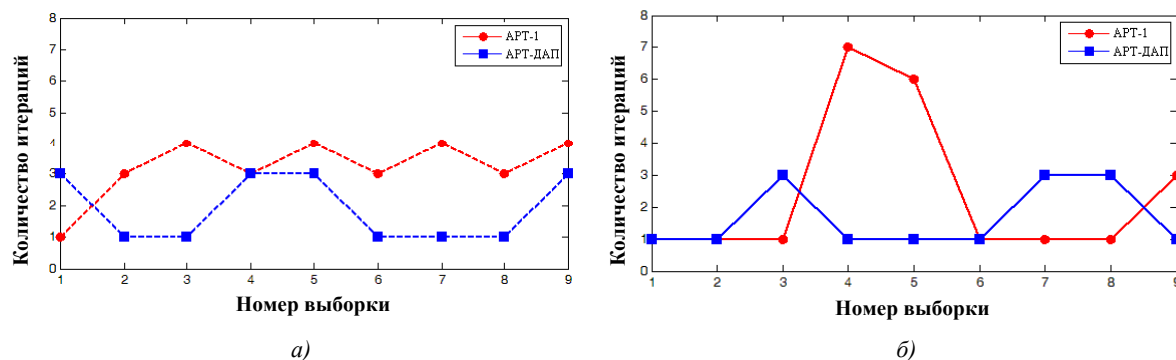


Рис. 8. Сравнение режимов работы сетей АРТ-1 и АРТ-ДАП:
а) процесс обучения; б) процесс распознавания

3. *Гетероассоциативный механизм.* В режиме гетероассоциативной выборки выходной образ, как правило, не соответствует ни одному из ключевых элементов и формируется как ответ на специфический ключевой образ. Подобным образом в сети АРТ-ДАП входные векторы, содержащие сенсорную информацию о состояниях внешней и внутренней среды, являются стимулами к ассоциативному поиску соответствующих выходных векторов, являющихся ответной реакцией ИАС.

4. *Адаптивное обучение.* Под реализацией адаптивных свойств понимается безопасное добавление в долговременную память сети АРТ-ДАП новых образов. Благодаря правилу дообучения, которое описывается формулой (4), старые образы не уничтожаются. Это позволит блоку памяти ИАС приобретать новые знания, а также генерировать и запоминать новые типы поведения, необходимые для приспособления к новым изменениям внешней среды.

5. *Обучение с учителем.* Как уже отмечалось выше, сеть АРТ-ДАП обучается с учителем. На основании формулы (1) в процессе обучения ИАС эксперт имеет возможность определенные ключевые образы принудительно ассоциировать в блоке памяти с конкретными выходными управляющими образами, позволяющими ИАС адекватно реагировать на изменения внутренней и внешней среды.

6. *Безытерационное обучение.* Итерационный способ обучения, который присутствует в других нейросетевых архитектурах, требует большого объема машинного времени и вычислений. Сеть АРТ-ДАП лишена этого недостатка, поэтому способна быстро пополнять свои знания и полученный опыт. Однако не следует путать итерационный способ обучения с процессом повторения заучиваемой информации, так как на примере биологической памяти последний приносит явную пользу. Такие же положительные результаты можно достичь и в рассматриваемой сети. В этом случае будет наблюдаться усиление пространственного распределения значений коэффициентов матрицы синаптических весов, характерных для определенного образа.

7. *Обобщение образов.* Поскольку выходные образы в системах распределенной памяти формируются путем опроса очень большого числа отдельных элементов, локальные искажения ключевых образов оказываются статистически сглаженными [13], т. е. выходные образы оптимально корректируются относительно объема всей информации, содержащейся в ключевом образе. Поэтому при работе сети АРТ-ДАП наблюдаются свойства фильтрации шума и обобщения входных образов.

Заключение

Предложен блок памяти для интеллектуальных автономных систем – сложных технических устройств, самостоятельно работающих в режиме реального времени. Используемые подходы, основанные на возможности самообучения, позволяют разрабатывать системы управления, которые способны самостоятельно формировать свое поведение, обучаясь на опыте взаимодействия с внешней средой.

Благодаря своим антропоморфным свойствам описанный блок памяти для ИАС на основе имеющихся в памяти эталонов производит классификацию поступающих образов, а также, при необходимости, добавляет в память новые. Гетероассоциативный механизм сети АРТ-ДАП, используемой в предложенном блоке памяти, позволяет быстро извлекать и обрабатывать информацию, а также реализует фильтрацию шумов во входных и выходных образах. Данная разработка дает возможность ИАС адаптироваться к изменяющимся условиям внешней среды и оптимально контролировать свои внутренние параметры.

Дальнейшие исследования будут направлены на увеличение объема памяти сети АРТ-ДАП, что позволит повысить эффективность работы как самого блока памяти, так и всей ИАС. Также необходимо уделить внимание разработке универсальных методов для ввода, хранения и обработки сенсорной информации, получаемой от разнообразных типов датчиков (графические изображения, звуки, запахи, касания и т. д.), которые выдают различные по форме и содержанию данные.

Список литературы

1. Тюменцев, Ю.В. Интеллектуальные автономные системы – вызов информационным технологиям / Ю.В. Тюменцев // Сб. тр. Восьмой Национальной конференции по искусственному интеллекту (КИИ'2002). Коломна, 7-12 октября 2002 г. – Т. 2. – М. : Физматлит, 2002. – С. 827–836.
2. Автоматы и разумное поведение / Н.М. Амосов [и др.]. – Киев : Наукова думка, 1973. – 375 с.
3. Бонгард, М.М. Проект модели организации поведения «Животное» / М.М. Бонгард, И.С. Лосев, М.С. Смирнов // Моделирование обучения и поведения. – М. : Наука, 1975. – С. 152–171.
4. Анохин, П.К. Очерки по физиологии функциональных систем / П.К. Анохин. – М. : Медицина, 1975. – 477 с.
5. Гаазе-Рапопорт, М.Г. От амебы до робота: модели поведения / М.Г. Гаазе-Рапопорт, Д.А. Поспелов. – М. : Наука, 1987. – 288 с.
6. From animals to animats // Proceedings of the First International Conference on Simulation of Adaptive Behavior ; Eds. J.-A. Meyer, S. Wilson. W. – Cambridge, Massachusetts, London, England : The MIT Press, 1990. – 551 p.
7. Непомнящих, В.А. Поиск общих принципов адаптивного поведения живых организмов и аниматов / В.А. Непомнящих // Новости искусственного интеллекта. – 2002. – № 2. – С. 48–53.
8. Жданов, А.А. Метод автономного адаптивного управления / А.А. Жданов // Известия академии наук. Теория и системы управления. – 1999. – № 5. – С. 127–134.
9. Кузнецов, О.П. Неклассические парадигмы искусственного интеллекта / О.П. Кузнецов // Теория и системы управления. – 1995. – № 5. – С. 3–23.
10. Стариков, А.И. Бортовые интеллектуальные системы управления комплексами оснащения летательных аппаратов : учеб. пособие / А.И. Стариков, А.С. Саркисов. – М. : Изд-во МАИ, 1994. – 72 с.
11. Кохонен, Т. Ассоциативные запоминающие устройства / Т. Кохонен. – М. : Мир, 1982. – 384 с.
12. Хакен, Г. Принципы работы головного мозга. Синергетический подход к активности мозга, поведению и когнитивной деятельности ; пер. с англ. / Г. Хакен. – М. : ПЕР СЭ, 2001. – 351 с.
13. Хайкин, С. Нейронные сети: полный курс / С. Хайкин. – 2-е изд. – М. : Издательский дом «Вильямс», 2006. – 1104 с.
14. Фролов, А.А. Нейронные модели ассоциативной памяти / А.А. Фролов, И.П. Муравьев. – М. : Наука, 1987. – 160 с.

15. Головкин, В.А. Нейронные сети: обучение, организация и применение. Кн. 4 : учеб. пособие для вузов ; под общ. ред. Л.И. Галушкина. – М. : ИПРЖР, 2001. – 256 с.
16. Прокопович, Г.А. Адаптивный нейросетевой классификатор / Г.А. Прокопович // Информатика. – 2009. – № 3 (23). – С. 68–81.
17. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика / Ф. Уоссермен. – М. : Мир, 1990. – 240 с.
18. Рассел, С. Искусственный интеллект: современный подход. / С. Рассел, П. Норвиг. – 2-е изд. – М. : Издательский дом «Вильямс», 2006. – 1408 с.
19. Induction: Processes of Inference, Learning, and Discovery / J.H. Holland [et al.]. – Cambridge : MIT Press, 1986. – 416 p.
20. Sutton, R. Reinforcement Learning: An Introduction / R. Sutton, A. Barto. – Cambridge : MIT Press, 1998. – 342 p.
21. Крыжановский, Б.В. О распознающей способности нейросети на нейронах с параметрическим преобразованием частот / Б.В. Крыжановский, А.Л. Микаэлян // Доклады РАН. – 2002. – Т. 383, № 3. – С. 318–321.
22. Крыжановский, Б.В. Быстрая система распознавания и принятия решения на основе векторной нейросети / Б.В. Крыжановский, В.М. Крыжановский // Искусственный интеллект. – 2004. – № 3. – С. 534–541.

Поступила 25.01.10

*Объединенный институт проблем
информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: bigznich@tut.by*

R.A. Prakapovich

NEURONET MEMORY UNIT FOR ADAPTIVE COMPLEX TECHNICAL SYSTEMS FUNCTIONING IN DYNAMIC ENVIRONMENT

New methods of complicated technical systems control for autonomous operation in a dynamic environment are considered. To impact the property of intelligent adaptive behavior to similar systems, the block of neural associative memory capable to pattern classification on the basis of available and additional standards is proposed.