МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра информационной безопасности

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Локтионова
2017 г.

МОДЕЛИРОВАНИЕ РАБОТЫ ЭЛЕМЕНТОВ НЕЙРОКОМПЬЮТЕРНЫХ СИСТЕМ

Методические рекомендациипо выполнению
лабораторной работы №1для студентов специальности 10.03.01

Курск 2017

Лабораторная работа №1

Моделирование работы элементов нейрокомпьютерных систем

Цель работы: изучить структуру мажоритарного, порогового и нейроподобного элемента, по структурным схемам создать модели элементов нейрокомпьютерных систем.

Задача: По представленным структурам мажоритарного, порогового и нейроноподобного элемента разработать блок-схемы алгоритмов работы элементов и протестировать программы на языке высокого уровня

1. Теоретическая часть

Определение: элементами технических систем обработки информации называются их простейшие, структурно и функционально обособленные части способные принимать, преобразовать, сравнивать, хранить, передавать и производить другие операции с сигналами.

Производя операции с сигналами, элементы функционально связывают изменение параметров входных сигналов x с изменениями параметров входных сигналов y. По характеру связи y =f(x) элементы разбиваются на:

1. элементы непрерывного действия;
2. элементы дискретного действия.

У элементов первого типа зависимость y = f(x) линейная или близкая к линейной, т.е. при непрерывном изменении параметра входного сигнала x соответствующий параметр входного сигнала y меняется непрерывно и пропорционально изменению x.

У элементов же второго типа зависимость y=f(x) нелинейная, т.е. при непрерывном изменении x в установленных пределах y меняется скачкообразно.

Элементы дискретного действия по логике работы можно классифицировать следующим образом:

- элементы булевой логики (булевы элементы);
- элементы мажоритарной логики (мажоритарные элементы);
- элементы пороговой логики (пороговые элементы);
- элементы нейронной логики (нейронные элементы или формальные нейроны).

Элементы, реализующие простейшие логические функции конъюнкции (&), дизъюнкции (V), отрицания (-), будем называть простыми булевыми элементами (БЭ). Распространены и-не, или-не, и-или-не элементы.

Характерными свойствами типичного БЭ является однофункциональность и равноценность входов.

Мажоритарный элемент (МЭ)

МЭ является обобщением булева элемента. Характерной особенностью МЭ является обязательная нечетность числа входов.

Рис. 1

Θ –пороговое напряжение

$$\Theta=\frac{n+1}{2};n=2*k+1,\mathrm{k}=1,2,3,... \qquad (1)$$

Пороговый элемент (ПЭ)

Пороговый элемент (ПЭ) представляет собой устройство, имеющее л входов и один выход. Входы ПЭ характеризуются весами w_i (i = 1, 2, …, n) которые имеют смысл коэффициентов усиления по входам. На каждый вход поступает сигнал x_i, принимающий значения 0 или 1. Сигнал на выходе порогового элемента равен 1, если

$$\sum_{i=1}^{n} w_i x_i > T \qquad (2)$$

и равен 0, если

$$\sum_{i=1}^{n} w_i x_i \le T \qquad (3)$$

число Т называют порогом.

Условия функционирования ПЭ полностью определяются величинами весов *w*, и порога Т. Поэтому ПЭ задается совокупностью весов и порога и обозначается $[w_1,w_2,.., w_n;T]$, полагая, что на вход с весом w_i, подается переменная x_i .

Работа ПЭ $[w_1,w_2,.., w_n;T]$ может быть описана переключательной функцией $f(x_1,x_2,…, x_n)$ равной 1 на тех наборах аргументов

x_i, для которых выполняется неравенство (2), и равной 0 на тех наборах, для которых выполняется неравенство (3).

Переключательная функция, которая может быть реализована одним пороговым элементом, называется пороговой функцией. Переключательные функции однозначны, поэтому любому ПЭ соответствует единственная пороговая функция. В то же время, любой пороговой функции соответствует бесконечное множество пороговых элементов, реализующих эту функцию.

Формальный нейрон (нейроподобный элемент НЭ)

Большинство моделей основывается на схеме формального нейрона У.С.Мак-Каллока и У.Питтса (1943 год), согласно которой нейрон представляет собой пороговый элемент (Рис.2). На входах нейрона имеются возбуждающие и тормозящие синапсы, в нейроне определяется взвешенная сумма (с учетом весов синапсов) входных сигналов, при превышении этой суммой порога нейрона вырабатывается выходной сигнал.

Нейроподобный элемент имеет вид:

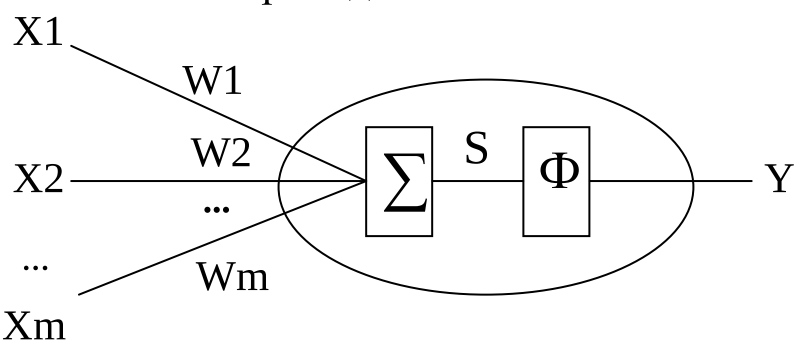

Рис. 2. Схема формального нейрона. X_i –входные сигналы, Y – выходной сигнал нейрона.

Работа формального нейрона (Рис.2) может быть описана уравнениями:

$$Yj = F(net_j\text{-}Kj), \quad (4)$$

$$net_j=\sum w_{ij}\, x_i, \quad (5)$$

где j - номер нейрона в сети, X_i - входные сигналы, У, - выходной сигнал нейрона, w_{ij} - веса синапсов, net_j - суммарное входное воздействие на нейрон, *Kj*- порог нейрона, *F* -активационная функция. Активационная функция характеризует реакцию нейрона на входное воздействие net_j, она может быть пороговой: или некоторой непрерывной, например линейной:

$$F(a) = ka \tag{6а}$$

или логистической:

$$F(a)=1/[1+exp(-a)]. \tag{6б}$$

В зависимости от реализуемого алгоритма на допустимые значения входов и выходов нейрона накладываются определенные ограничения: значения X_i, и Y_j могут бинарными (т.е. равными 0 или 1), бинарными биполярными (+1 или -1) принадлежащими интервалу (0,1), неотрицательными или действительными. Аналогичные ограничения накладываются на веса синапсов нейронов w_{ij}.

Отметим, что в основополагающей работе Мак-Каллока и Питтса входы и выходы нейронов предполагались бинарными, веса синапсов считались бинарными биполярными, а активационная функция - пороговой. Исследования нейросетей проводились с точки зрения анализа логических исчислений, которые могут быть построены на базе формальных нейронов. В частности было показано, что "для всякого логического выражения, удовлетворяющего некоторым условиям, можно найти сеть, имеющую описываемое этим выражением поведение".

Формальные нейроны до определенной степени отражают динамику передачи сигналов в реальных биологических нейронах. Живые нейроны состоят тела клетки, дендритов и аксона. Очень упрощая картину, работу нейрона можно описать следующим образом. Дендриты получают сигналы от других клеток через синапсы, эти сигналы поступают в тело клетки, где они суммируются с другими такими же сигналами. Если суммарный сигнал в течение короткого промежутка времени является достаточно большим, то клетка возбуждается, вырабатывая в аксоне импульс, который передается на следующие клетки. Не вдаваясь в подробности, подчеркнем, что формальные нейроны только очень грубо отражают работу биологических *живых* нервных клеток.

1.4. Перцептрон

История исследования нейронных сетей испытывала взлеты и падения. Первый всплеск энтузиазма был в 50-60-х годах. Его можно связать с работами Дж. фон Неймана по концептуальному сравнительному анализу работы биологических нейронных сетей и компьютеров и по разработке принципов построения надежных вычислительных систем из ненадежных компонент (фактически формальных нейронов) и с работами Ф. Розенблата по перцептронам. Работы по перцептронам - наиболее значимое направление исследований первого бионического бума.

Перцептрон состоит из элементов 3-х типов: *S* - элементов, *A* - элементов и *R* - элемента. *S* - элементы это - слой рецепторов. Эти рецепторы соединены с *A* - элементами, с помощью тормозных или возбуждающих связей. Каждый рецептор может находиться в одном из двух состояний - покоя или возбуждения. *A* - элементы представляют собой сумматоры с порогом (т.е. формальные нейроны). Это означает, что *A* - элемент возбуждается, если алгебраическая сумма возбуждений, приходящих к нему от рецепторов, превышает определенную величину - его порог. При этом сигнал от рецептора, приходящий по возбуждающей связи, считается положительным, а приходящий по тормозной связи - отрицательным. Сигналы от возбудившихся *A* – элементов передаются в сумматор *R*, причем сигнал i-го ассоциативного элемента передается с коэффициентом k_i.

Система связей между рецепторами *S и A* - элементами, так же как и пороги *A* - элементов выбираются некоторым случайным, но фиксированным образом, а обучение состоит лишь в изменении коэффициентов k_i. Считаем, что мы хотим научить перцептрон разделять два класса объектов, и потребуем, чтобы при предъявлении объектов первого класса выход перцептрона был положителен, а при предъявлении объектов второго класса - отрицательным. Начальные коэффициенты *k* полагаем равными нулю. Далее предъявляем обучающую выборку: объекты (например, круги либо квадраты) с указанием класса, к которым они принадлежат. Показываем перцетрону объект первого класса. При этом некоторые *A* - элементы возбудятся. Коэффициенты k_i, соответствующие этим возбужденным элементам, увеличиваем на 1. Затем предъявляем объект второго класса и коэффициенты k_i тех *A* - элементов, которые возбудятся при этом показе, уменьшаем на 1. Этот процесс продолжим для всей обучающей выборки. В результате обучения сформируются значения весов связей k_i.

После обучения перцептрон готов работать в режиме распознавания. В этом режиме перцептрону предъявляются "не знакомые" перцептрону объекты, и перцептрон должен установить, к какому классу они принадлежат. Работа перцептрона состоит в следующем: при предъявления объекта возбудившиеся A - элементы передают сигнал R - элементу, равный сумме соответствующих коэффициентов k_i. Если эта сумма положительна, то принимается решение, что данный объект принадлежит к первому классу, а если она отрицательна - то второму.

Исследования перцептронов показали, что перцептроны способны обучаться, хотя способности их обучения довольно ограничены. Справедлива *теорема о сходимости перцептрона,* согласно которой независимо от начальных значений коэффициентов и порядка показа образцов при обучении перцептрон за конечное число шагов научится различать два класса объектов, если только существуют такие значения.

Исследования также показали, что слабые стороны перцептрона (в частности большое время обучения) в значительной степени связаны со случайностью связей между его элементами. Однако эта конструктивная особенность обеспечивает перцептрону и положительное качество надежность: выход из строя заметного числа элементов перцептрона слабо сказывается на качестве его работы.

Первые успехи исследованиям перцептронов других нейросетей вызвал взрыв активности и энтузиазма. М. Минский, Ф. Розенблат, Б. Уидроу и другие разработали ряд искусственных нейронных сетей. В течение некоторого времени казалось, что ключ к интеллекту найден, и воспроизведение человеческого мозга является лишь, вопросом конструирования достаточно большой сети.

Возможности перцептронов оказались довольно ограниченными Серьезный математический анализ перцептронов был проведен М.Минским и С Пейпертом. Они, в частности, показали, что задачи, которые в принципе могут быть решены перцептроном могут потребовать нереально больших времен или нереально большой памяти. Например, для различения некоторых классов объектов коэффициенты части ассоциативных элементов должны быть столь велики, что для хранения их в вычислительной машине потребовался бы больший объем памяти, чем для того, чтобы просто запомнить все конкретные объекты этих двух классов.

Критика перцептронов М. Минским (а он - один из признанных авторитетов в теории искусственного интеллекта), а также сравнительно небольшой прогресс нейрокибернетики 50-60-х годов привели к тому, что период энтузиазма сменился периодом спада активности исследований искусственных нейронных сетей.

Только немногие кибернетики (Т. Кохонен, С. Гроссберг, Дж. Андерсон, Г.С. Бриндли, Д. Мар, В.Л.Дунин-Барковский, А.А.Фролов и др.) продолжали исследования нейросетей в 70-х годах.

Однако в середине 80-х годов снова возник нейросетевой бум. Причиной бума, по-видимому, послужил постоянный интерес человечества к изучению работы нервной системы и ряд новых интересных моделей, разработанных к этому времени. Одной из таких "стимулирующих" моделей стали работы Дж. Хопфилда, которые позволили привлечь методы теоретической физики к исследованию нейронных сетей.

В большинстве моделей запоминание информации в нейронной сети (обучение) происходит в результате формирования весов синапсов нейронов. Во многих случаях это интерпретируется как формализация гипотезы Хебба, в соответствии с которой изменение состояния произвольного синапса определяется его текущим состоянием и активностью пре- и постсинаптических нейронов.

Предполагается, что определенные практические задачи должны решаться нейрокомпьютерами и нейрочипами - искусственными нейроподобными сетями, созданными на основе микроэлектронных вычислительных систем. Спектр задач для нейрокомпьютеров достаточно широк: распознавание зрительных и звуковых образов, создание экспертных систем и их аналогов, управление роботами, создание нейропротезов для людей, потерявших слух или зрение. Достоинства нейрокомпьютеров - параллельная обработка информации и обучаемость.

Простой персептрон. В середине 50-х годов была предложена одна из основных моделей нейронных сетей, которая вызвала большой интерес из-за своей способности обучаться распознаванию простых образов.

Эта модель - персептрон - состоит из бинарных нейроподобных элементов, имеет простую топологию, что позволило достаточно полно проанализировать ее работу и создать многочисленные физические реализации. Типичный персептрон состоит из трех основных компонент:

матрицы бинарных входов $r_1..r_n$ (сенсорных нейронов или «сетчатки», куда подаются входные образы);

набора бинарных нейроподобных элементов $x_i \dots x_m$ (или предикатов в наиболее общем случае) с фиксированными связями к

подмножествам сетчатки («детекторы признаков»);
бинарного нейроподобного элемента с модифицируемыми связями к этим предикатам («решающий элемент»).

На самом деле число решающих элементов выбирают равным количеству классов, на которое необходимо разбить предъявляемые персептрону образы.

Таким образом, модель персептрона характеризуется наличием только прямых связей, один из слоев которых является модифицируемым. В простейшем случае, когда $n = m$ и $x_i = r_i$ детекторы признаков могут рассматриваться как входной слой. Тогда персептрон становится одним бинарным нейроподобным элементом. Это классическая модель М - входового нейрона или простой персептрон Розенблата. В общем случае каждый элемент x, может рассматриваться как булева функция, зависящая от некоторого фиксированного подмножества сетчатки. Тогда величина выходных сигналов этих обрабатывающих элементов является значением функции x_i, которое равно 0 или 1. Устройство реагирует на входной вектор генерацией выходного сигнала *y* решающего элемента по формуле (4). Таким образом, персептрон формирует гиперплоскость, которая делит многомерное пространство $x_i \ldots x_m$ на две части и определяет, в какой из них находится входной образ, выполняя, таким образом, его классификацию. Возникает вопрос, как определить значения весов, чтобы обеспечить решение персептроном конкретной задачи. Это достигается в процессе обучения. Предложены различные правила обучения персептрона. Один из алгоритмов называется процедурой сходимости персептрона Розенблата и является вариантом хеббовского правила изменения весов связей с учителем. Алгоритм работает следующим образом. Вектор весов Wi устанавливают в произвольное состояние.

На сетчатку поочередно подают образы из обучающей выборки, которые трансформируются в выходной сигнал *y* решающего элемента. При правильном отклике ничего не изменяют. При неправильном отклике $y = 0$ веса всех связей от активных элементов сетчатки увеличивают, а при неправильном отклике $y = 1$ - уменьшают. Величина изменения связи определяет степень адаптации.

Если решение существует, оно будет достигнуто при циклической по образов обучающей выборки за конечное число шагов при любом начальном выборе связей.

Таким образом, если два класса образов могут быть раздел гиперплоскостью, то при достаточно долгом обучении персептрон бv различать их правильно. Однако линейная разделяющая поверхности упрощающая анализ персептрона, ограничивает решаемый им крут задач. Этот вопрос тщательно исследовали Минский и Пейперт, показав, какие задачи принципе не может решить персептрон с одним слоем обучаемых связей. Одним, из таких примеров является выполнение логической операции «исключающее ИЛИ».

Многослойный персептрон. Как отмечалось выше, простой персептрон с одним слоем обучаемых связей формирует границы областей решений в виде гиперплоскостей. Двухслойный персептрон может выполнять операцию логического «И» над полупространствами, образованными гиперплоскостями первого слоя весов. Это позволяет формировать любые, возможно неограниченные, выпуклые области в пространстве входных сигналов. С помощью трехслойного персептрона, комбинируя логическими «ИЛИ» нужные выпуклые области, можно получить уже области решений произвольной формы и сложности, в том числе невыпуклые и несвязные. То, что многослойные персептроны с достаточным множеством внутренних нейроподобных элементов и соответствующей матрицей связей в принципе способны осуществлять любое отображение вход-выход, отмечали еще Минский и Пейперт, однако они сомневались в том, что можно открыть для них мощный аналог процедуры обучения простого персептрона. В настоящее время в результате возрождения интереса к многослойным сетям предложено несколько таких процедур. Алгоритм обратного распространения ошибки. Этот алгоритм является обобщением одной из процедур обучения простого персептрона, известной, как правило, Уидроу - Хоффа (или дельта-правило), и требует представления обучающей выборки. Выборка состоит из набора пар образов, между которыми надо установить соответствие, и может рассматриваться как обширное задание векторной функции, область определения которой - набор входных образов, а множество значений - набор выходов.

Многослойная нейроподобная сеть с прямыми связями. Входные элементы (блоки) образуют нижний слой сети, выходные - верхний. Между ними может быть много слоев скрытых блоков. Каждый блок может быть соединен модифицируемой связью с любым блоком соседних слоев, но между блоками одного слоя связей нет. Каждый блок может посылать выходной сигнал только р вышележащие слои и принимать входные сигналы только от нижележащих слоев. Входной вектор подается на нижний слой, а выходной вектор определяется путем поочередного вычисления уровней активности элементов каждого слоя (снизу - вверх) с использованием уже известных значений активности элементов предшествующих слоев.

С точки зрения познавания образов входной вектор соответствует набору признаков, а одной - классу образов. Скрытый слой используется для представления части знаний. Перед началом обучения связям присваиваются небольшие случайные значения. Каждая итерация процедуры состоит из двух фаз. Во время первой фазы на сеть подается входной вектор путем установки в нужное состояние входных элементов. Затем входные сигналы распространяются по сети элементов предшествующих слоев, порождая некоторый выходной вектор. Для работы алгоритма требуется, чтобы характеристика вход-выход нейроподобных элементов была неубывающей и имела ограниченную производную. Обычно для этого используют сигмоидную нелинейность вида (6б). Полученный выходной вектор сравнивается с требуемым. Если они совпадают, обучение не происходит. В противном случае вычисляется разница между фактическими и требуемыми выходными значениями, которая передается последовательно от выходного слоя к входному.

При обучении веса связей перестраиваются таким образом, чтобы минимизировать частоту смены активности каждого блока. Таким образом, обученная сеть имеет стабильные состояния и может функционировать в режиме ассоциативной памяти. В настоящее время многослойные персептроны являются наиболее популярной моделью нейронных сетей. Это в значительной степени объясняется тем, что с их помощью удалось продемонстрировать решение ряда задач, в том числе классической для персептронов задачи «исключающего ИЛИ», задачи синтеза речи по тексту, а также задач, требующих принятия экспертных решений. Возможно, что подобные многослойным персептронам нейронные структуры используются мозгом для предварительной обработки сенсорной информации, например, при выделении признаков.

Ансамблевые нейронные сети. Минский и Пейперт отмечали, что недостатки простых персептронов можно преодолеть как с помощью многослойных сетей, так и введением в сеть обратных связей, допускающих циркуляцию сигналов по замкнутым контурам. Использовать свойства такого рода сетей для моделирования функций мозга еще в 1949 г. предложил Хебб. Согласно взглядам Хебба нервные клетки мозга соединены друг с другом большим количеством прямых и обратных возбуждающих связей и образуют нейронную сеть. Каждый нейрон осуществляет пространственно-временную суммацию приходящих к нему сигналов от возбужденных нейронов, определяя потенциал на своей мембране. Когда потенциал на мембране превышает пороговое значение, нейрон возбуждается. Нейрон обладает рефрактерностью и усталостью. Эффективность связей может изменяться в процессе функционирования сети, повышаясь между одновременно возбужденными нейронами. Это приводит к объединению нейронов в клеточные ансамбли – группы клеток, которые чаще всего возбуждались вместе, и к обособлению друг от друга. При возбуждении достаточной части ансамбля он возбуждается целиком. Различные ансамбли могут пересекаться; один и тот же нейрон может входить и разные ансамбли. Электрическая активность мозг обусловлена последовательным возбуждением отдельных ансамблей. Идеи Хебба оказали большое воздействие на представления о работе мозга, послужили основой для создания нейронных моделей долговременной памяти. Действительно, ансамблевую нейронную сеть можно рассматривать как структуру, реализующую функции распределенной ассоциативной памяти. Формирование ансамблей в такой сети соответствует запоминанию образов (признаков, объектов, событий, понятий), закодированных паттерном активности нейронов, а сформированные ансамбли являются их внутренним представлением. Процесс возбуждения всего ансамбля при активации части его нейронов можно интерпретировать как извлечение запомненной информации по ее части - ключу памяти. Модель памяти на основе ансамблевой нейронной сети обладает некоторыми свойствами, присущими биологической памяти, такими, как ассоциативность, распределенность, параллельность, устойчивость к шуму и сбоям, надежность. Проводятся также структурные аналогии между ансамблевыми моделями нейронных сетей и строением коры головного мозга. Имеются экспериментальные данные о синаптической пластичности, постулированной Хеббом. Модель ансамблевой сети состоит из большого количества нейроподобных элементов, каждый из которых обычно соединен со всеми другими элементами сети. Входной образ подается на сеть путем активации нужных нейроподобных элементов. В отличие от персептрона ансамблевая сеть может обучаться как с учителем, так и без него. Обучение производится по правилу Хебба или одному из его вариантов. Значение коэффициента *a* при обучении с учителем определяет величину подкрепления, а при обучении без учителя эта величина может быть установлена, например, постоянной. Отметим, что правило Хебба формирует симметричную матрицу связей. В процессе обучения при подаче на сеть набора входных образов - представителей среды - в сети формируются ансамбли, которые могут иметь сложную структуру, отражающую свойства среды.

Сеть Хопфилда. Хотя многочисленные результаты моделирования демонстрировали стабильность ансамблевых сетей с обратными связями и хеббовским правилом обучения (эволюцию сети к устойчивому состоянию), отсутствие математического обоснования такого введения препятствовало их популярности. Положение изменилось с появлением работ, где было определено подмножество нейронных сетей с обратными связями, которые гарантированно достигают устойчивого состояния. В 1982 г. американский биофизик Джон Хопфилд опубликовал статью, где на основании аналогии между нейронными сетями и особым классом физических систем - спиновыми стеклами - ему удалось привлечь к анализу нейросетевых моделей мощный математический аппарат статистической физики. Это стимулировало вторжение в область моделирования нейронных сетей большого отряда ученых-физиков, которыми в настоящее время получено много интересных аналитических результатов. Сам Хопфилд в упомянутой статье рассмотрел поведение модели полносвязной сети бинарных нейроподобных элементов с симметричными связями ($w_{ij}=w_{ji}$). Элементы функционировали в асинхронном режиме, т. е. каждый нейрон в случайные моменты времени с некоторой средней частотой определял свое состояние в соответствии с правилом (6а).

Это позволило описать поведение сети как релаксационный процесс, при котором минимизируется энергетическая функция E (функция Ляпунова, гамильтониан) модели: где сеть состоит из w_{ij} - матрицы связей; y и q -состояние и порог модельного нейрона. Действительно, изменение E при изменении состояния нейрона (учитывая симметрию w_{ij} и полагая q=0). Так как знак Dyi совпадает со знаком, ясно, что *E* по мере срабатывания нейронов будет монотонно убывать, а так как *E* ограничена, будет достигнуто состояние ее минимума. Таким образом, эволюция сети из любого начального состояния приводит к состоянию, соответствующему локальному минимуму *E.* Можно провести аналогию поведения сети с движением легкой частицы по некоторому вязкому рельефу под действием силы тяжести.

В своей работе Хопфилд исследовал сеть с нейроподобными элементами, имеющими сигмоидную характеристику. Состояния нейронов такой сети изменяются одновременно и непрерывно, и сеть описывается системой дифференциальных уравнений. Хопфилд доказал сходимость и такой сети к стабильным энергетическим минимумам и нашел соответствие между ее устойчивыми состояниями и устойчивыми состояниями сети с бинарными элементами. Это послужило основой для построения аппаратных моделей, где сеть реализуется как аналоговая электронная схема, состоящая из операционных усилителей, моделирующих нейроны, соединенных резисторами с проводимостями *Wij,* и со смещениями входными токами *q.* Понятно, что если сделать минимумами энергии заданный набор паттернов нейронной активности (образов), оба варианта сети Хопфилда смогут выполнять функции ассоциативной памяти, «скатываясь» к тому образу, в чей «бассейн притяжения» попадает начальный паттерн активности нейронов сети.

Рассмотрим сеть, состоящую из 40 нейронов, которые расположены в виде матрицы 5x8, запомненного в сети изображения буквы *«E».* Активному нейрону соответствует измененный элемент изображения. Из поданного на сеть зашумленного изображения восстанавливается правильное изображение. Как и следовало ожидать, одним из способов получения нужной энергетической функции является формирование матрицы связей в соответствии с вариантом хеббовского правила: где z^p- образы, которые надо запомнить в сети; *L* - их количество. Это правило, как и правило, предложенное Хеббом, обеспечивает формирование симметричной матрицы связей, однако постулирует увеличение веса связей между не только одновременно активными, но и одновременно неактивными нейронами, а также его уменьшение между нейронами, находящимися в разном состоянии. Такое правило допускает существование тормозящих модифицируемых связей между нейронами и даже переход возбуждающих связей в тормозящие. Оно позволяет сети автоматически саморегулировать уровень активности и работать с нулевыми порогами нейронов. Однако при этом значительно снижается емкость памяти сети количество случайных образов, которое можно записать в сеть с возможностью восстановления, не превышает 0.14 количества нейронов. Кроме того в дополнение к энергетическим минимумам, соответствующим запомненным образам, возникают ложные минимумы функции *E*. Положение еще более осложняется для скоррелированных образов, которые после запоминания не становятся минимумами *E*. В настоящее время ведется интенсивная работа по улучшению характеристик модели Хопфилда, предлагаются ее интересные расширения и обобщения Предпринимаются попытки создать алгоритмы обучения, позволяющие работать со скоррелированньми образами.

Машина Больцмана. Машина Больцмана представляет собой стохастический вариант сети Хопфилда. Бинарные нейроподобные элементы (блоки) трактуются здесь как представители элементарных гипотез, а веса - как слабые парные взаимоограничения между ними. Положительный вес связи указывает, что две гипотезы стремятся поддерживать друг друга, а отрицательный - на их несовместимость. Симметрия связей позволяет проанализировать поведение сети с использованием энергетической функции.

Энергию определенного паттерна активности можно интерпретировать как степень нарушения ограничений, присутствующих в проблемной области, со стороны конкретной комбинации гипотез или как стоимостную функцию, которая должна быть минимизирована для решения оптимизационной задачи. Если зафиксировать состояния некоторых блоков, подав, таким образом, на сеть входное воздействие, остальные блоки начнут изменять свое состояние так, чтобы минимизировать энергию *E*. Поступающая на каждый блок взвешенная сумма сигналов от активных блоков из-за симметрии связей совпадает с величиной разности между значениями, энергетической функции, зависящей от его собственного состояния: Поэтому алгоритм изменения состояния бинарных нейроподобных элементов автоматически приводит к минимизации энергии *E*.

Машина Больцмана может использоваться для классификации образов. При этом в ней, как и в многослойном персептроне, выделяют входные, выходные и внутренние (скрытые) блоки, однако для каждой прямой связи между блоками существует равная ей по величине обратная (симметричная) связь. Процесс распознавания состоит из следующих шагов: на входных блоках фиксируют входной образ; рандомизируют состояние скрытых и выходных блоков, а затем медленно понижают температуру; наблюдают за состоянием сети при конечной низкой температуре и собирают статистику состояний выходных блоков. На основании этой статистики делают вывод о входном образе.

Для машины Больцмана существует алгоритм обучения, который, как и для многослойного персептрона, создает путем модификации связей внутреннюю модель среды, позволяющую достаточно хорошо классифицировать входные образы. Для работы алгоритма требуется обучающая выборка, состоящая из пар вход - выход, которые должна научиться ассоциировать сеть. Если после обучения зафиксировать на входных блоках один из входных образов, на выходных блоках должен появиться соответствующий выходной образ. Если подать на вход неизвестный образ, система на основе выявленных в обучающей выборке закономерностей должна провести правильное обобщение. Каждый цикл обучения состоит из трех шагов.

1. Фаза тренировки. Для каждой пары образов фиксируются состояния входных и выходных блоков, а остальная часть сети подвергается отжигу к низкой температуре. Затем для каждой связи собирается статистика, какую часть времени p^+_{ij} были одновременно активны соединяемые ею блоки.

2. Фаза проверки. Вычисляется аналогичная величина p^-_{ij} однако теперь выходные блоки не зафиксированы и свободно меняют состояние.

3. Изменение связей. В хорошо обученной сети ее поведение будет идентично для обеих фаз. Если p^- и p^+ не совпадают для конкретной связи, ее изменяют: где e масштабирует размер изменения. Каждый цикл необходимо повторить много раз, пока матрица связей не стабилизируется в достаточной степени.

К сожалению, алгоритм обучения машины Больцмана имеет типичные недостатки, присущие процедурам градиентного спуска в многопараметрических пространствах. Прежде всего это неточность вычисления градиента, обусловленная неполным достижением теплового равновесия и ограниченным временем сбора статистик. Из-за своей стохастичности алгоритм требует гораздо больших временных затрат по сравнению даже с алгоритмом обучения многослойного персептрона методом обратного распространения ошибки. Имеющаяся аппаратная реализация, однако, смягчает этот недостаток по крайней мере для небольших сетей. Известны гримеры применения машины Больцмана для решения классических персептронных задач, таких, как задача «исключающего ИЛИ», обнаружение симметрии во входном образе и т. д., а также для распознавания речи.

2. Практические схемы на элементах нейрокомпьютерных систем.

Работа мажоритарного элемента аналитически описывается следующим выражением:

$$F = sign[\sum_{i=1} x_i - \frac{n+1}{2}], n = 3,5,7,\ldots$$

Функция *sign y* определяется так:

$$sign\_y = \begin{cases} 0, при\_y < 0 \\ 1, при\_y >= 0 \end{cases}$$

Работа мажоритарного элемента на три входа описывается в следующем виде:

X	**У**	**Z**	F
0	**0**	**0**	**0**
0	**0**	**1**	**0**
0	**1**	**0**	**0**
0	**1**	**1**	**1**
1	**0**	**0**	**0**
1	**0**	**1**	**1**
1	**1**	**0**	**1**
1	**I**	**1**	**1**

$F(x, y, z) = xzVxyVyz$

Эта функция сохраняет нуль, единицу, является самодвойственной и монотонной.

Чтобы образовать функционально полную систему для синтеза логических схем на мажоритарных элементах к функции F добавляют инверсию и константы:

F(x,y,0)=xy - схема И

$$F(x,y,1)=xVy$$ - схема ИЛИ

$$\overline{F(x,y,0)}=\overline{xy}=\overline{x}V\overline{y}=x/y$$ - Шеффера

$$\overline{F(x,y,1)}=\overline{xVy}=\overline{x}*\overline{y}=x\downarrow y$$ - Пирса

Мажоритарный элемент с пятью входами образует функцию:

$$F(x_1,x_2,x_3,x_4,x_5)=x_1x_2x_3Vx_1x_2x_4Vx_1x_2x_5Vx_1x_3x_4Vx_1x_3x_5Vx_1x_4x_5Vx_2x_3x_4Vx_2x_4x_5Vx_3x_4x_5$$

Эта функция сохраняет нуль, единицу, является самодвойственной, монотонной и нелинейной.

С помощью функции с пятью входами можно построить схемы И, ИЛИ, Пирса, Шеффера на три входа.

$$F(x_1,x_2,x_3,0,0)=x_1x_2x_3$$

$$F(x_1,x_2,x_3,1,1)=x_1Vx_2Vx_3$$

$$F(\overline{x_1,x_2,x_3,0,0)}=\overline{x_1Vx_2Vx_3}=\overline{x_1}\setminus\overline{x_2}\setminus\overline{x_3}$$

$$F(\overline{x_1,x_2,x_3,1,1)}=\overline{x_1x_2x_3}=x_1\downarrow x_2\downarrow x_3$$

Схема сумматора-вычитателя на мажоритарных элементах.

a	**b**	**P_H**	**S**	**Pi**
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

B / P_H **a**	00	01	11	10
0	0	1	0	1
1	1	0	1	0

$$S=a\overline{b}\overline{P}_{i-1}V\overline{a}\overline{b}P_{i-1}VabP_{i-1}V\overline{a}bP_{i-1}=a\oplus b\oplus P_{i-1}$$

$$Y=x\overline{y}V\overline{x}y=x\oplus y=(x\overline{y},\overline{x}y,1)$$

$$(x,\overline{y},0)=x\overline{y}$$

$$(\overline{x},y,0)=\overline{x}y$$

$Y = (x\bar{y}, \bar{x}y, 1) = x\bar{y} V \bar{x}y$

$S = (a, \bar{b}, \bar{P}_{i-1}, 0,0)(\bar{a}, \bar{b}, P_{i-1}, 0,0)(a, b, P_{i-1}, 0,0)(\bar{a}, b, P_{i-1}, 0,0)$

$S = (a, \bar{b}, \bar{P}_{i-1}, 0,0, \bar{a}, \bar{b}, P_{i-1}, 0,0,1)$

Найти пороговую функцию, реализуемую ПЭ [1,-1,2; 1]. Используя неравенство (1) имеем:

$$x_1 - x_2 + 2x_3 > 1$$

X_1	X_2	X_3	F
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

$$F = \bar{x}_1 \bar{x}_2 x_3 V x_1 \bar{x}_2 x_3 V x_1 x_2 x_3$$

После карты Карно имеем: $F = x_1 x_3 V \bar{x}_2 x_3$

Переключательные функции однозначны, поэтому любому ПЭ соответствует единственная пороговая функция. В то же время, любой пороговой функции соответствует бесконечное множество пороговых элементов, реализующих эту функцию.

Рассмотрим вопрос о функциональной полноте системы пороговых функций.

1. Функция Шеффера n аргументов (n>2) является пороговой и одна образует функционально полную систему пороговых функций: [w_1 =-1, w_2 =-1.. w_n =-1;T = -n].

Для двух входов n =2 имеем, [w_1 = -l,w_2 =-1;T = -2] $-x_1-x_2>-2$

- штрих Шеффера (таблица истинности)

x_1	x_2	F
0	0	1
0	1	1
1	0	1
1	1	0

2. Функция Пирса является пороговой. [w_1 =-1,w_2 =-1..w_n =-1;T = -1]

Для двух переменных имеем: $-x_1 - x_2 > -1$

- стрелка Пирса (таблица истинности)

x_1	x_2	F
0	0	1
0	1	0
1	0	0
1	1	0

3. Дизъюнкция [$w_1 = 1$, $w_2 = 1$.. $w_n = 1$; $T = 0$]
4. Конъюнкция [$w_1 = 1$, $w_2 = 1$.. $w_n = 1$; $T = n-1$]
5. Отрицание [$w_i = -1$, $T = -1$]

Нейронный элемент (НЭ)

Нейронный элемент (НЭ) является обобщением порогового элемента (или формальный нейрон (ФН)). ФН представляет собой сочетание порогового элемента с булевыми элементами И, ИЛИ, НЕ. Булевы элементы (БЭ), входящие в его структуру, являются его неотъемлемой частью, как бы встроены в нейрон и не могут быть структурно выделены как самостоятельные элементы.

Следует отметить, что для произвольно заданной булевой функции всегда существует ФН, реализующий эту функцию. Это свойство ФН и называется Н-полнотой (от слова нейрон).

На рисунке показана схема обобщенного ФН.

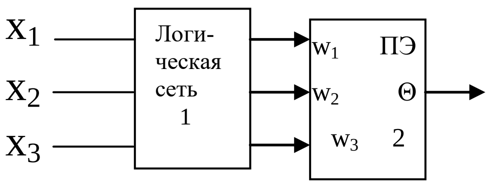

Рис. 3.

Обозначения, используемые в рис.3.

1 - блок взаимодействия входных волокон нейрона, представляющий собой б х И логическую сеть булевых элементов.

2 - пороговый элемент нейрона.

В этом блоке происходит алгебраическое суммирование взвешенных входных сигналов, т.е. сумма равна

$\sigma = \sum_{i=1}^{n} w_i a_i$, где w_i е {±1,±2,±3,...} - весовые коэффициенты ПЭ, *а,* е {0,1} - логическая переменная, показывающая состояние входа (сигнала) в данный момент времени, a_i = 1 означает, что i-й сигнал возбужден, a_i = 0 - не возбужден. Пороговое устройство (дискриминатор) срабатывает и на выходе элемента появляется сигнал только, если $\sigma > \Theta$ где Θ - порог нейрона. Различные типы ФН отличаются только по структуре блока 1, а структура блока 2 у всех ФН одинакова.

На нейроподобный элемент поступает набор входных сигналов $x_1, x_2, ..., x_m$ (или входной вектор X). Входной вектор представляет собой выходные сигналы других нейроподобных элементов. Этот входной вектор соответствует сигналам, поступающим в синапсы биологических нейронов. Каждый входной сигнал умножается на соответствующий вес связи $\omega_1, \omega_2, \ldots, \omega_m$ - аналог эффективности синапса. Вес связи является скалярной величиной, положительной для возбуждающих и отрицательной для тормозящих связей.

Обычно используются простейшие нелинейные функции:

1. Бинарная

$$Y = \begin{cases} 1 & \text{при } S > 0 \\ 0 & \text{при } S \le 0 \end{cases}$$

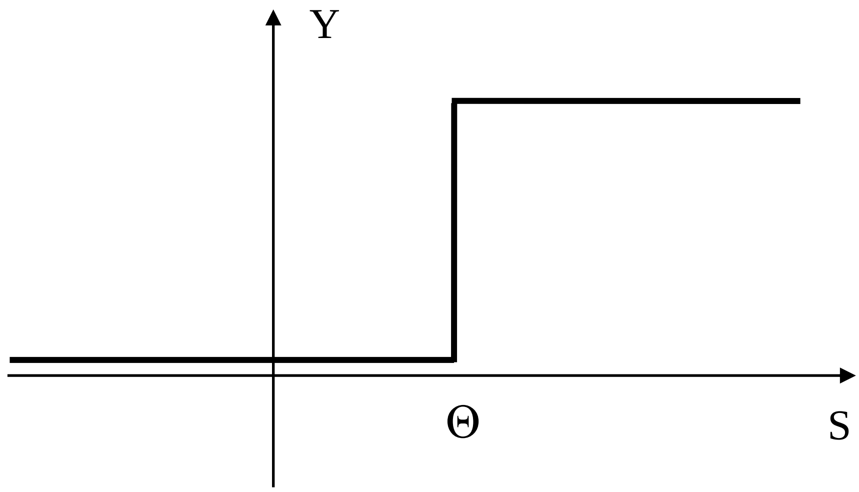

Рис.4

2. Сигмоидная

$$Y = \frac{1}{1 - e^{-(S\,-\,\Theta)}}$$

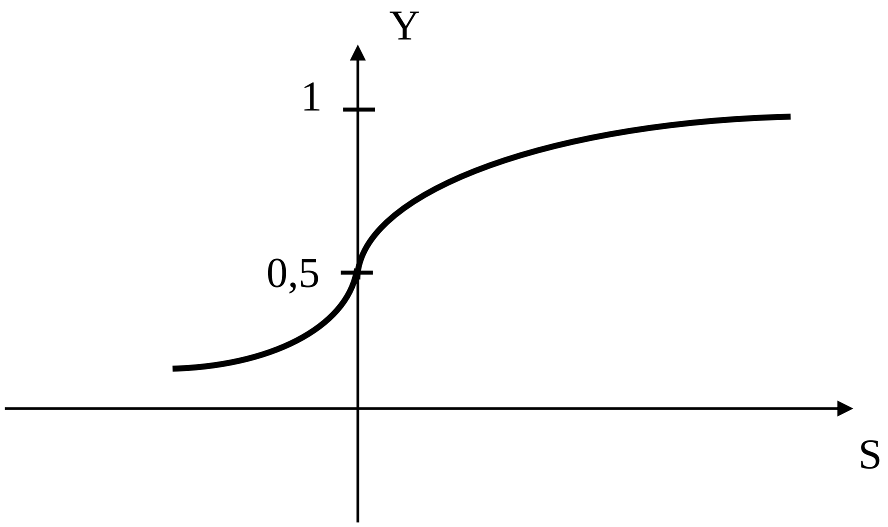

Рис.5

Взвешенные весами связей входные сигналы поступают на блок суммации, соответствующий телу клетки, где осуществляется их алгебраическая суммация и определяется уровень S возбуждения нейроподобного элемента S(y)

$$S = \sum\nolimits_{j=1}^{M} \omega_i x_i$$

Выходной сигнал нейрона y определяется путем пропускания уровня возбуждения S через нелинейную функцию φ.

$$y = \varphi(S - \Theta)$$

где Θ - некоторое постоянное смещение. Построенные на основе таких простых нейроподобных элементов системы демонстрируют ассоциативные свойства, напоминающие свойства биологических систем.

Логические устройства на нейронных элементах

Построение триггеров на ФН

Функцию асинхронного RS-триггера аналитически можно описать следующим образом: $Q(t+1) = \bar{R}[S \vee Q(t)] \vee pRSQ(t)$

где p e {0,1}: p=l, если R&S=1, и p=0, если R&S=0.

допустим, что в рассматриваемом триггере R=l, S=l, является запрещенной, т.е. R&S=0. Тогда обозначая R=x_1 S=x_2, Q(t)=x_3, Q(t+Δ)=F, получим

$$F=\bar{x}(x \vee x_3)$$

получаем схему триггера RS на ФН:
при R=l или S=l состояние нейрона изменяется.

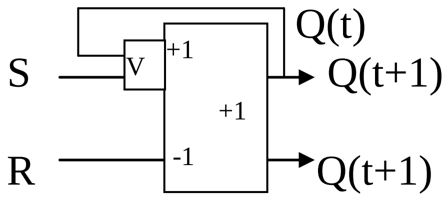

Рис. 6

3. Задание

1. Составить блок - схемы алгоритмов работы трех и пяти входового мажоритарного элемента МЭ. Входными данными для трех входового мажоритарного элемента являются переменные - x_l, x_2, x_3, для пяти входового - a_l, a_2, a_3, a_4, a_5. Выходная функция Y =1 определяется по принципу большинства, т.е. число возбужденных входов составляет абсолютное большинство от общего числа входов. Пороговая функция 0 вычисляется по формуле (1).

2. По разработанной блок-схеме алгоритма составить программу на языке высокого уровня, моделирующую работу мажоритарного элемента.

3. Составить блок - схему алгоритма работы порогового элемента ПЭ. Входными переменными X_i является вектор, принадлежащий множеству чисел: X_i е {0,1}. Количество переменных X_i задается. Определяется массив весовых коэффициентов ПЭ, в виде матрицы W_{ij}. Значения весовых коэффициентов это натуральные числа с учетом знака.

Например:

$$W_{ij}=\begin{vmatrix} +1 & -2 & -1 \\ +2 & 0 & -1 \\ -1 & +1 & 0 \\ -2 & +2 & +1 \end{vmatrix}$$

Для задания порогового напряжения формируется одномерный массив, элементы массива - натуральные числа со свом знаком. Количество элементов массива задается. Для примера возьмем:

$$T=\{+2,-3,+1,0,-1\}$$

Количество строк в матрице коэффициентов связей равно количеству входных переменных X_i порогового элемента. В нашем случае 4 г.е. на вход ПЭ поступают x_l, x_2, x_3, x_4. Для работы ПЭ необходимо из массива пороговых напряжений выбрать одно значение. Выходное значение ПЭ определяется по формулам (2) и (3). Для получения одной точки на графике нужно получить сумму произведений значений вектора X_i на значения коэффициентов связей в столбце матрицы.

$$\sum_{j=1}^{4} x_i * W_{ij}$$

Затем нужно сравнить полученную сумму со значением порогового напряжения Т. Для получения остальных точек необходимо перейти к следующим значениям столбцов матрицы, изменяя параметр j.

$$\sum_{j=1}^{3} \sum_{j=1}^{4} x_i * W_{ij}$$

Для каждого порогового значения из массива Т строится график функции Количество графиков соответствует количеству значений пороговых напряжений. Количество точек на графике соответствует количеству столбцов в матрице связей W_{ij}.

4. По разработанной блок-схеме алгоритма составить программу на языке высокого уровня, моделирующую работу порогового элемента. По заданным значениям входных величин: X_i, W_{ij}, Т построить графики функций.

5. Составить блок - схему алгоритма работы нейроподобного элемента НЭ (Рис.3). Для построение логической части НЭ необходимо задать булевы функции: И, ИЛИ, НЕ, а также количество входов каждой секции логических микросхем. Для определения порогового элемента необходимо выполнить задание пункта 3. Задается множество булевых функций: V={Н, ИЛИ, НЕ}, количество входов в каждую секцию Н={2,3,4,8}, количество секций р={ 1,2,3,4). Для примера сформируем логическую часть нейрона. Выберем значения из множеств V, Н, Р. Из множества V выберем логическую функцию И, количество входов из множества Н - 2, количество

секций из множества Р -1. Из множества V выберем ИЛИ, количество входов сначала - 2, затем - 3, количество секций - 2 в обоих случаях. Из множества V выбираем НЕ, количество входов - 1, количество секций - 1. Количество выходов из логической части соответствует количеству строк матрицы коэффициентов связи, в нашем случае - 4. Количество столбцов матрицы задается. Пороговое напряжение Т выбирается из массива пороговых напряжений. На основании заданных величин формируется нейрон НЭ (рис.7).

6. По разработанной блок-схеме алгоритма составить программу на языке высокого уровня, моделирующую работу нейрона. По заданным значениям входных величин: X_i, W_{ij}, T, V, H, P разработать структуру нейрона.

7. Промоделировать (тестировать) работу элементов нейрокомпьютерных систем на ПЭВМ.

8. Построить график зависимости выходной функции У от значений входного вектора X_i.

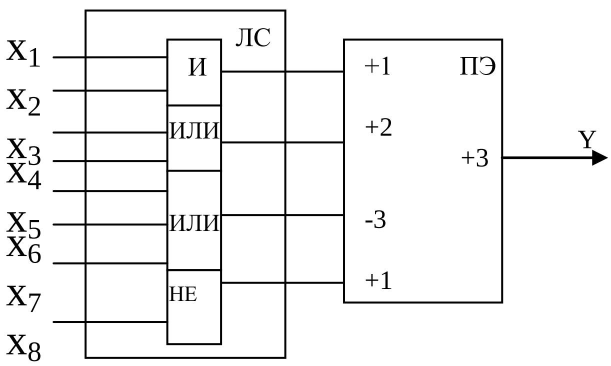

Рис. 7

4. Содержание отчета.

Отчет должен содержать:

- титульный лист;
- задание;
- алгоритм;
- текст программы;
- структуры элементов нейрокомпьютерных систем;
- результаты работы программы, графики зависимости выходной функции Y от входных значений вектора X_i;

Контрольные вопросы

1. Какую структуру имеет мажоритарный элемент МЭ.
2. Как он описывается с помощью логических функций.
3. Приведите примеры построение структурных схем на трех, пяти входовых мажоритарных элементах.
4. По какой формуле вычисляется пороговое напряжение мажоритарного элемента.
5. Дайте определение пороговому элементу ПЭ. Какие функции он выполняет.
6. Какие исходные данные формируют структуру порогового элемента.
7. Как определяется пороговое напряжение порогового элемента ПЭ.
8. Как реализуются основные булевы функции на пороговых элементах.
9. Что такое нейроподобный элемент. Его основные функции.
10. Как строятся логические устройства на нейронных элементах.
11. Как формируется выходной сигнал нейрона. Укажите формулу.

Библиографический список

1. Нейрокомпьютеры и интеллектуальные роботы. Под ред. Н.М. Амосова, Киев, Наукова думка, 1991.
2. В.Д. Цыганков. Нейрокомпьютер и его применение. М., СолСистем, 1993.
3. Е.Н. Соколов, Г.Г. Вайткявичус. Нейроинтеллект от нейрона к нейрокомпьютеру. М., Наука, 1989.
4. Отчет о НИР «Принципы представления знаний в системах управления роботов». Под руководством Н.М. Амосова, Киев, 1985.
5. Ю.В. Чернухин. Микропроцессорное и нейрокомпьютерное управление адаптивными мобильными роботами. Таганрог, 1993.
6. Ю.В. Чернухин. Нейропроцессоры. Таганрог, 1994.
7. Труды третьего международного симпозиума «Интеллектуальные системы», Псков, 1998.
8. В.Ф.Венда. Системы гибридного интеллекта. М., Машиностроение, 1990
9. Л.И. Волгин. Комплиментарная алгебра нейросетей. Таллин, АО KLTK, 1993.