

На правах рукописи

Ионов Сергей Дмитриевич

**ПРОБЛЕМЫ ПОСТРОЕНИЯ
РАЗВИВАЮЩЕЙСЯ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ
С АССОЦИАТИВНОЙ ПАМЯТЬЮ**

Специальность 05.13.18 – математическое
моделирование, численные методы и комплексы программ

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Екатеринбург 2014

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте математики и механики им. Н.Н. Красовского Уральского отделения Российской академии наук.

Научный руководитель доктор физико-математических наук, профессор
Мазуров Владимир Данилович

Официальные оппоненты: **Жернаков Сергей Владимирович**,
доктор технических наук, профессор, ФГБОУ ВПО
«Уфимский государственный авиационный технический
университет», заведующий кафедрой электроники и
биомедицинских технологий;
Соколинский Леонид Борисович,
доктор физико-математических наук, профессор, ФГБОУ
ВПО «Южно-Уральский государственный университет»
(национальный исследовательский университет), проректор
по информатизации.

Ведущая организация Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Тюменский государственный университет»

Защита состоится «24» сентября 2014 г. в 14:00 часов на заседании диссертационного совета Д 212.285.25 на базе ФГАОУ ВПО «Уральский федеральный университет имени первого Президента России Б.Н.Ельцина» по адресу: 620000, г. Екатеринбург, пр. Ленина, д. 51, зал заседания диссертационных советов, к. 248.

С диссертацией можно ознакомиться в библиотеке и на сайте ФГАОУ ВПО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»,
<http://dissovet.science.urfu.ru/news2/>.

Автореферат разослан " ____ " _____ 2014 г.

Ученый секретарь диссертационного совета,
доктор физико-математических наук, профессор

Пименов Владимир Германович

Общая характеристика

Актуальность работы

Основные направления развития современных информационных технологий напрямую связаны с хранением и управлением знаниями и, следовательно, требуют автоматизации как сбора знаний, так и представления результатов их обработки либо пользователю, который должен принять решение, либо непосредственно устройствам управления тем или иным процессом. В любом случае сам ход принятия решений так же важен, как и результат. Не случайно в последнее время при построении систем принятия решений (СПР) отходят от искусственных нейронных сетей (ИНС), существующие реализации которых не могут предоставить пользователю информацию, убедительно обосновывающую полученный результат, из-за чего теряется важная часть процесса принятия решения – доказательство его допустимости и оптимальности. Однако отказ от нейронных сетей влечет и отказ от всех их преимуществ, в частности, распределенной параллельной природы функционирования и достаточно быстрой адаптации к новым обстоятельствам, что в современном постоянно меняющемся мире довольно актуально. Чтобы не отказываться от этих преимуществ, надо уметь выделять осмысленные части нейронной сети и обеспечивать возможность их контролируемого изменения.

Для реализации указанных возможностей в нейронных сетях наиболее простым подходом представляется такая модификация нейронных сетей, которая приближала бы их описания к описаниям процедур, реализующих тот или иной алгоритм. В этом случае уже на этапе проектирования выделяются осмысленные части, результаты которых могут быть предоставлены для анализа пользователем или даже самой нейронной сетью, что позволит контролировать изменения вплоть до возможности самоконтроля их допустимости.

Надо понимать, что управление знаниями и самоконтроль невозможны без наличия в нейронной сети той или иной разновидности памяти, которая позволила бы учитывать предыдущий опыт при появлении новых знаний. В нейросетевом подходе на этапе обучения знания чаще всего опосредованно сохраняются в виде значений весовых коэффициентов. Однако для наглядности при обосновании принятия того или иного решения или для контроля допустимости самомодификации в процессе переобучения необходимо явное хранение значений поступающих сигналов. Действительно, достаточно хорошим обоснованием принятого решения во многих случаях является демонстрация запомненного при обучении примера, максимально схожего с полученным во время работы.

Для построения новой модели нейронной сети с памятью, адекватной современным задачам построения систем принятия решений, следует обратиться к исследованиям физиологии мозга и, в частности, к теориям организации памяти. С течением времени познания в области нейрофизиологии развивались и уточнялись, и часть новых открытий была применена в построении различных подходов к моделированию нейронных сетей. Тем не менее, в этих подходах и получаемых моделях не было явной направленности на моделирование работы памяти. Из анализа существующих теорий и моделей можно выделить наиболее устоявшуюся концептуальную часть, основную идею которой положить в основу модели памяти для искусственной нейронной сети. Этот анализ должен быть достаточно тщательным, чтобы уровень детализации моделируемых процессов функционирования элементов нейронной сети соответствовал решаемым задачам.

Цельная ИНС, строящая свою структуру только посредством своих внутренних элементов и включающая память, подсети анализа сигналов и обучения с модификацией связей и порождением новых ансамблей нейронов с заранее заданной функциональностью, обладает рядом преимуществ. Так, представленные в единой форме различные функциональные блоки, касающиеся одной смысловой части нейронной сети, могут быть выделены в пакеты и включены в библиотеки для их повторного использования. Кроме того, возможно описание этих блоков с использованием одного специализированного языка, что представляется более последовательным решением, чем то, когда задание топологии сети и правил ее работы разрозненны и указываются в сложных взаимозависимых структурах одного или даже нескольких универсальных языков. Правда, специализированный язык требует создания, обучения использованию и развития среды исполнения, включая набор стандартных библиотек. Важно отметить, что этот язык должен быть способен выразить любую структуру потоковой нейронной сети, включающей классические примеры перцептрона или рекуррентных сетей.

Наличие такого языка позволит нейросетевому подходу не уступать доминирующим в настоящее время подходам при построении систем принятия решений: экспертные системы, системы логического вывода, вопросно-ответные системы, – а на ряде задач превзойти их.

Степень разработанности проблемы

Основы теории биологических нейронных сетей были предложены А. Бэйном в 1873 году и У. Джеймсом в 1890. В 1898 г. Ч. Шеррингтон провел эксперименты для проверки теории Джеймса и способствовал разработке теории привыкания. В 1900 г. В. М. Бехтерев опубликовал «Проводящие пути спинного и головного мозга». Значительных

успехов в области исследования нервных систем в XX веке достигли Г. Бишоп, Дж. Эрлангер, Г. Гассер, И. С. Беритон, Дж. Экклс, Л. Проссер, Ф. Браун, Ч. Шеррингтон, Н. П. Бехтерева, Э. Кендел.

Теория искусственных нейронных сетей была создана в 1943 году Мак-Каллоком и Питтсом как множество математических алгоритмов, ориентированных на изучение биологических процессов головного мозга и на создание нейросетевого искусственного интеллекта. Дальнейшее развитие включало разработку однослойного перцептрона Ф. Розенблаттом в 1957 году и многослойного в 1965 году. Для многослойного перцептрона метод обучения впервые был описан в 1974 г. А. И. Галушкиным, а также независимо и одновременно П. Дж. Вербосом. Впоследствии в направлении искусственных нейронных сетей были предложены различные модели Т. Кохоненом, Б. Видроу, Т. Хоффом, С. Гроссбергом, Г. Карпенгером, Дж. Андерсоном, Б. В. Хакимовым, К. Фукусимой, Дж. Хопфилдом, Г. Н. Борисюком, Х. Симизду и Ж.-Ж. Слотеном, Дж. Дином, М. Веллингем, Р. Салахутдиновым. Из последних достижений стоит отметить созданные в 2007 году Дж. Хинтоном алгоритмы глубокого обучения многослойных нейронных сетей.

Портретный подход в моделировании ИНС представлен в работах Г. Маркрама, Ш. Хилла, В. А. Филиппова, Ю. И. Александрова, К. В. Анохина, С. Гранта, В. Герстнера.

Структурно-иерархический подход используется в трудах Д. Хокинса, Д. Джорджа, П. А. Лалетина, Ю. П. Ланкина.

Функциональный подход рассматривается в концепциях Дж. Хопфилда, Г. Йу, Ж.-Ж. Слотена, Д. Хьюбеля, Т. Н. Визеля.

Хотя сдвиг в сторону выделения сохранения полученной информации в последнее время прослеживается, но явно блоков хранения исходной информации в ИНС не выделяется. Противопоставление нейросетевого и процедурного подходов к решению задач во многом не соответствует биологическим нейронным сетям и препятствует решению сложных задач с применением ИНС.

Цели и задачи работы

Цель работы – построение нового класса нейронных сетей, предоставляющих в своей структуре возможности реализации алгоритмов и сохраняющих способности нейронных сетей к изменчивости и обучению.

Для достижения этой цели необходимо решить ряд задач, относящихся к прикладным задачам современного математического моделирования.

1. Упрощение использования математической модели ИНС за счет усложнения минимального элемента сети – нейрона. Для решения этой задачи необходимо выбрать минимальный достаточный функционал, который бы позволял строить логику произвольной сложности и оставаться в терминах нейронов и связей между ними. Кроме создания самого элемента, также требуется обоснование его структуры с точки зрения теории нейронных сетей и, в частности, приводимость одного вида нейронных сетей к другому.

2. Внесение в нейронную сеть возможностей обучения, указанных в терминах самой нейронной сети, так как именно в нейронной сети должны задаваться алгоритмы ее работы. При этом должна быть продемонстрирована возможность динамического изменения весов связей между нейронами в соответствии со стандартными алгоритмами обучения перцептрона.

3. Добавление в ИНС средств для изменения структуры сети в зависимости от проходящих сигналов. Решение этой задачи должно позволить в нейронной сети создать память, по потребности расширяемую за счет порождения новых нейронов. Таким образом, например, в рассматриваемой математической модели должна быть возможность создания теоретически бесконечной ленты машины Тьюринга.

Также в диссертации ставятся задачи разработки численных методов и комплекса программ для практического применения новой модели ИНС.

4. Обеспечение эффективного распределенного параллельного запуска получаемых ИНС путем применения соответствующих численных методов. В задаче требуется разместить граф нейронной сети на узлах графа физической сети, выполнив ряд требований, обеспечивающих эффективное использование ресурсов кластера вычислительных узлов. Так как при некоторых упрощениях задача может быть сформулирована как квадратичная задача размещения, то требуется построить приближенный численный алгоритм. Примененный к кластеру физических узлов этот алгоритм в результате должен обеспечивать как можно большую утилизацию вычислительных ресурсов с малой задержкой работы, вызванной передачей сигнала по сети между узлами. Кроме самого размещения, в задачу также включается вычисление настроек синхронизации работы отдельных нейронов, расположенных на удаленных физических узлах. Синхронизация осуществляется определением задержки получения сигналов нейроном, зависящей от производительности как отдельных узлов, так и связей между ними.

5. Разработка языка описания алгоритмов для ИНС и комплекса всех соответствующих инструментов: среды разработки, компилятора, среды отладки и

тестирования, системы запуска. Решение должно включать в себя описание грамматики языка, рекомендаций по построению компилятора и пример компилятора. На основе выдаваемых компилятором кодов определяется поведение остальных инструментов, принимающих эти коды в качестве входных данных для исполнения.

Основная практическая цель работы – создание системы, которая базируется на новой модели нейронной сети и осуществляет обработку сигналов с датчиков окружающей среды и подачу управляющих сигналов на контроллеры для приведения окружающей среды к некоторому заданному оптимальному состоянию.

Объект исследования

Объектом исследования является СПР на основе искусственных нейронных сетей, обеспечивающих хранение и применение правил для вывода рекомендаций.

Предмет исследования

Предметом исследования являются модели ИНС, методы их описания и наглядного представления, а также алгоритмы распределенного запуска.

Методы исследования

В ходе диссертационного исследования применялась методология системного анализа; использовались методы теории графов в отношении нейронных сетей и математической логики в представлении арифметических операций в виде нейронов, статистические методы обработки результатов обучения полученной СПР, методы объектно-ориентированного и прототипно-ориентированного проектирования и программирования.

Область исследования

Содержание работы соответствует Паспорту специальностей ВАК РФ в области исследования п.1 – Разработка новых математических методов моделирования объектов и явлений; п.4 – Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента; п.8 – Разработка систем компьютерного и имитационного моделирования.

Научная новизна

1. Разработана новая модель искусственных нейронных сетей, позволяющая описывать алгоритмы обработки сигналов в терминах элементов и связей между ними. Созданная модель является расширением классических искусственных нейронных сетей на основе нейронов МакКаллока-Питтса.

2. Обоснована применимость новой модели в построении алгоритмических решений и определены ее выразительные возможности. Доказательство основывается на определении полноты рассматриваемой модели по Тьюрингу.

3. Разработан декларативный язык описания структуры графа ИНС, учитывающий ее параллельную и распределенную природу. Для языка показана его полнота по Тьюрингу, а также раскрыты возможности задания растущих массивных нейронных сетей.

4. Разработаны численные методы размещения нейронной сети и алгоритмы вычисления задержек для настройки распределенного запуска.

Теоретическая значимость

В работе описана новая математическая модель искусственных нейронных сетей, являющаяся расширением классической модели; конструктивно доказана ее связь с существующей теорией нейронных сетей; показаны ее выразительные возможности в решении алгоритмических задач, а также основания для применения методов обучения нейронных сетей. Для созданной модели приводится соответствующий язык ее описания, применимый для задания обработки потока сигналов в распределенной параллельной системе.

Кроме того, разработаны численные методы размещения нейронной сети на множестве вычислительных узлов и определения общей задержки работы вычислителей для настройки поведения нейронов.

Практическая ценность результатов

Представленные в работе примеры отдельных блоков нейронных сетей, а также созданная для языка библиотека стандартных блоков-ансамблей искусственных нейронов (БАИН) могут быть применены для решения множества типовых задач построения нейронных сетей. Кроме того, приводятся примеры сетей, вычисляющие различные математические функции.

Разработанный специализированный язык позволяет описывать любой сложности нейронные сети в предложенной модели и может быть расширен для задания произвольной параллельной распределенной системы. Разработанная грамматика языка занимает 188 строк.

Созданный в процессе работы комплекс программ может использоваться для описания, компиляции, визуализации, отладки и запуска нейронных сетей в рамках новой модели. При этом разработанные алгоритмы и архитектура позволяют реализовать и применять и другие модели нейронных сетей. Исходные тексты компилятора состоят из

более чем 9000 строк и 60 классов на языке Python. Написанная на том же языке среда разработки занимает более 7500 строк и 40 классов. Система отладки и тестирования, созданная на языке Perl, состоит из более чем 3000 строк. В исходных текстах системы распределенного запуска находится более 18000 строк на языке C++.

В ходе достижения практической цели работы была построена СПР, способная как к точному применению полученных во время обучения правил, так и к выведению возможных вариантов поведения на основе состояний в изученных ранее схожих правилах. В качестве подтверждения результатов СПР применена для поддержки систем автоматизированного управления.

Положения, выносимые на защиту

1. Математическая модель искусственной нейронной сети, обеспечивающая полноту описания алгоритмов с использованием ее структурных элементов, и язык описания ИНС, позволяющий задавать структуру графа сети, выделять отдельные блоки-ансамбли искусственных нейронов для их повторного использования и учитывающий параллельную распределенную природу исследуемой модели.
2. Система принятия решений на основе ассоциативной памяти, построенной из элементов исследуемой модели ИНС и опирающейся на хранение известных и выведение новых правил поведения.
3. Комплекс программ разработки ИНС, включающий компилятор языка описания ИНС, среду разработки для хранения исходных текстов и визуализации графа полученной сети, систему отладки и тестирования, а также систему распределенного запуска и, в частности, численные методы размещения графа нейронной сети на графе физических узлов и настройки работы нейронов.

Достоверность

Достоверность научных положений и выводов обеспечивается корректным использованием формальных доказательств в рамках математических моделей и методов, отсутствием противоречий с известными теоретическими положениями, а также подтверждается успешным решением конкретных прикладных задач на основе полученных результатов.

Апробация результатов работы

Полученные результаты докладывались и обсуждались на следующих научных

конференциях, симпозиумах и семинарах:

- Межвузовская научная конференция «СПИСОК-2009», апрель 2009 г.
- 42-ая Всероссийская молодежная школа-конференция «Современные проблемы математики », февраль 2011 г.
- Международная суперкомпьютерная конференция «Научный сервис в сети Интернет: поиск новых решений», сентябрь 2012 г.
- 44-ая Всероссийская молодежная школа-конференция «Современные проблемы математики », февраль 2013 г.
- XI Всероссийская научная конференция «Нейрокомпьютеры и их применение», март 2013 г.
- Семинар ИММ УрО РАН, 10 октября 2013 г.
- XII Всероссийская научная конференция «Нейрокомпьютеры и их применение», март 2014 г.

Личный вклад автора

Автором лично осуществлялась разработка элементов модели и языка их представления; обоснование положений о выразительных возможностях модели и языка; разработка блоков искусственных нейронных сетей, представленных в работе; реализация комплекса программ, включающего компилятор, среду разработки, систему тестирования и систему распределенного запуска; реализация интеграции с САУ «Умный дом». Проработка отдельных фундаментальных оснований модели и структуры языка проводилась при содействии С.В. Шарфа.

Публикации

По теме диссертации опубликовано 6 печатных работ, в том числе 2 статьи – в журналах, рекомендованных ВАК РФ.

Структура работы

Диссертация состоит из введения, 6 глав, заключения, списка использованной литературы и приложений. Общий объем работы составляет 153 стр., 5 табл., 45 рис., 7 приложений, список литературы – 63 наименований.

Краткое содержание диссертации

Введение посвящено постановке научных и технических задач диссертации, дана краткая аннотация работы.

Первая глава представляет обзор различных моделей памяти. С точки зрения философии рассматривается основополагающее понятие памяти и его роль в структуре сознания. Далее изучаются различные виды памяти, выстраивается их типология согласно научным трудам в психологии. На основе детального рассмотрения модели работы памяти с точки зрения нейробиологии показывается более материальная и четкая структура памяти как элемента нервной системы организма, причем основной акцент делается на принципах развития памяти как инструмента хранения информации. При обобщении знаний, полученных в ходе исследований в других областях, изучаются различные математические подходы к моделированию памяти и построению систем принятия решений в целом. А именно, рассмотрены подходы:

- классический, включающий перцептрон Розеблатта, Адалайн Видроу и Хоффа, а также теорию адаптивного резонанса Гроссберга;
- портретный, максимально точно воспроизводящий процессы, протекающие в нервных клетках, и показанный на примере проектов IBM Blue Brain Project и Тюменского Ассоциативных Систем Объединения (ТАСО);
- структурно-иерархический, обладающий важной способностью к наращиванию структуры сети и распределению процессов обучения между уровнями иерархии, что продемонстрировано на примере иерархических самоадаптирующихся нейронных сетей, разработанных в Красноярской школе нейроинформатики;
- функциональный, моделирующий поведение нейронной сети путем усложнения алгоритмов ее работы, что показано на примере осциллирующей нейронной сети и автоассоциативной памяти на основе сети Хопфилда;
- алгоритмический, включающий нелинейную логику, системы черной доски, методы опорных векторов, эволюционные алгоритмы, обладающий свойствами предсказуемости результата и структурности с выделением смысловых блоков, которых не имеют существующие нейронные сети в силу традиционного противопоставления подходов к решению задач.

Производится анализ соответствующих подходов и определяется дальнейшее направление развития, связанное с интеграцией выявленных положительных черт в новой основывающейся на функциональном подходе модели. Исходя из рассмотренных теорий,

строятся предположения о возможных расширениях модели нейрона и структурном представлении нейронной сети в целом для построения алгоритмов хранения знаний и применения их в процессе принятия решений.

Вторая глава описывает особый тип нейронной сети, созданный в ходе исследования для достижения цели работы и являющийся одним из основных результатов. В основе предлагаемой математической модели воспроизводящихся искусственных нейронных сетей (ВИНС) лежит элемент, описываемый формулой

$$E(f, b, C): \begin{cases} a_t(f) = f(a_{t-1}(f), \sum_i w_i I_i), \\ D(f, b) = a_t(f) \cdot (b + \sum_i v_i I_i), \\ a_0 = C, \\ t \in N, \end{cases}$$

где a_t – внутреннее активационное состояние элемента, изменяющееся на каждом новом такте t работы сети, f – функция вычисления нового значения активации элемента, I – вектор входных сигналов, w_i, v_i – значения соответствующих весовых векторов, собранных для всех нейронов ВИНС в матрицы W, V , b – константа отклонения вычисленного значения элемента, C – начальное активационное состояние элемента. Полученное в D итоговое значение элемента далее передается во входные вектора остальных связанных с ним элементов.

Связи между элементами ВИНС задаются тремя матрицами: W – веса активационных связей, V – веса значимых связей, G – веса генерационных связей.

В работе используется два типа элементов ВИНС: аддитивный и мультипликативный нейроны.

Аддитивный нейрон задается формулой

$$A(p, b) \equiv E(\phi_p, b, 0), \phi_p(a, x) = \begin{cases} 0, & x < p, \\ 1, & x \geq p, \end{cases}$$

где p – значение порога, указываемое для каждого нейрона отдельно. Таким образом, сигналы по активационным связям складываются и определяют результат вычисления пороговой функции активации, пропускающей или не пропускающей сигналы значимых связей на выход элемента.

Мультипликативный нейрон использует активационное значение в качестве константы для умножения входящего значимого сигнала, что создает возможность изменения весов связей ВИНС изнутри во время ее работы. Мультипликатор задается такой функцией:

$$M(C, b) \equiv E(\psi, b, C), \psi(a, x) = a + x,$$

где C – начальное значение константы умножения, ψ – функция изменения этой константы. Важно заметить, что данный элемент использует свое внутреннее состояние, равное предыдущему значению активационной функции, благодаря чему открывается возможность инкрементального изменения константы умножения.

Для элементов ВИНС предусмотрено два режима функционирования: рабочий и шаблонный. Рассмотренные ранее модели поведения применяются в рабочем режиме. В шаблонном режиме каждый элемент игнорирует все сигналы активационных и значимых связей. Вместо этого он ожидает значения индексов по генерационным связям и выполняет воспроизведение шаблонных элементов. Далее к новым порожденным элементам применяется принцип растущих связей, гарантирующий восстановление связности нейронной сети согласно шаблону. На рис. 1 представлен вид нейронной сети до и после прихода из входного нейрона I последовательности сигналов 7, 4, 7.

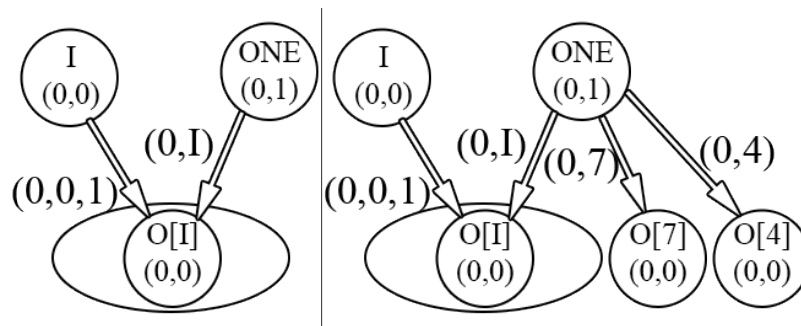


Рис. 1. Принцип работы ВИНС в шаблонном режиме

Можно заметить, что в указанной ВИНС реализуется сохранение уникальных значений полученных сигналов в виде весов связей. Уникальность гарантируется действующим принципом динамической генерации, по которому повторной генерации элемента с тем же индексом не происходит. Связь каждого нового выходного нейрона с нейроном ONE, постоянно генерирующим сигнал со значением 1, обеспечивается реализацией принципа растущих связей.

На основе предложенной математической модели показано построение алгоритмических решений задач. В качестве обоснования конструктивно доказывается Теорема о полноте ВИНС по Тьюрингу через построение эмулятора машины Тьюринга.

Теорема

Пусть дана ВИНС с ограничением на передаваемый в сети сигнал $S \in R$. Тогда ВИНС полна по Тьюрингу.

Для обоснования применимости существующей теории нейронных сетей для модели ВИНС проводится построение элементов с использованием классических нейронов МакКаллока-Питтса с пороговой функцией активации. Каждый элемент выполняет операции сложения, умножения и сравнения, которые могут быть представлены в двоичной логике и выражены в виде нейронов. Построения выполнены в работе, а их корректность и оценка размеров получаемых нейронных сетей обоснованы в соответствующих теоремах.

Помимо алгоритмических возможностей модели ВИНС в данной главе показано их применение для решения классических для нейронных сетей задач и, в частности, приводится построение перцептрона.

В третьей главе описывается соответствующий модели ВИНС язык задания алгоритмов SAINNL, учитывающий ее природу и возможности. Приводится сравнительный анализ существующих языков задания нейронных сетей: универсальные языки с соответствующими нейросетевыми библиотеками, специализированные языки NeuronC, NSL, NeuroML. Каждый из рассмотренных языков не в полной мере удовлетворяет набору требований:

- 1) язык должен учитывать предметную область, т. е. нейронные сети;
- 2) язык должен позволять строить абстракции более высокого уровня, нежели отдельные элементы;
- 3) язык должен отражать реализованную в сети логику обработки сигнала, а не структуру графа;
- 4) язык должен быть декларативным или функциональным, чтобы отразить параллельную природу нейронных сетей.

Далее глава 3 содержит описание языковых единиц языка SAINNL: блоков, выражений, элементов – и их применения в грамматике языка для получения корректных описаний нейронных сетей. Указаны представления элементов ВИНС в терминах языка

$\text{ADDITIVE} = \# w_i * I_i \geq p, t_i * I_i + b$ $\text{MULTIPLICATIVE} = \& w_i * I_i, t_i * I_i + b, C$

Также приводится способ задания шаблонного режима элементов ВИНС в виде динамического генератора

$[B] A[B] = \# X - 2 * Y[B] + B * Z[B] \geq 0, I$ $[B] M[B] = \& X, Y, B$

Для определения соответствия языка и модели построены отображения их друг на друга в виде следующих двух теорем, а также показаны свойства инъективности и сюръективности этих отображений.

Теорема

Пусть дана программа на языке SAINNL в виде множества глобальных констант и блоков с объявлениями внутри, причем один из блоков имеет имя «network». Тогда эту программу можно представить в виде графа ВИНС.

Теорема

Пусть дан граф ВИНС. Тогда этот граф можно представить в виде исходного текста на языке SAINNL.

Более того, представлено доказательство теоремы о полноте языка по Тьюрингу, а также показаны другие выразительные возможности самого языка.

В главе приводятся подробные описания по обработке исходных текстов на языке SAINNL. Указывается требуемое поведение при преобразовании сложных выражений в подсети. А именно:

$$O1 = A + B * C + (D + E) * F$$

эквивалентно

$$B_C = B * C$$

$$D_E = D + E$$

$$D_E_F = D_E * F$$

$$O1 = A + B_C + D_E_F$$

Также описывается основной принцип подстановки подсетей при обращении к блокам на примерах простых сетей.

Отдельное внимание уделено функции статической генерации, позволяющей задавать массивные нейронные сети, используя шаблонные описания.

Далее выставляются требования к функциям компилятора, которые позволят реализовать корректное преобразование исходного текста на языке в описание графа ВИНС.

В четвертой главе решается задача построения системы принятия решений с использованием предложенной модели нейронной сети. В основе решения лежит применение ассоциативной памяти, определяемой в теории памяти как рабочая или кратковременная память. В первую очередь вводится понятие ассоциативной памяти и определяются основные требования к ее поведению. Приводятся подробности реализации в виде отдельных логических блоков. Определяется единица хранения – ячейка памяти, представляющая собой блок нейронов с циклической передачей сохраненного сигнала. На ее основе строится слот памяти, способный сохранить кортеж сигналов. Его структура показана на рис. 2.

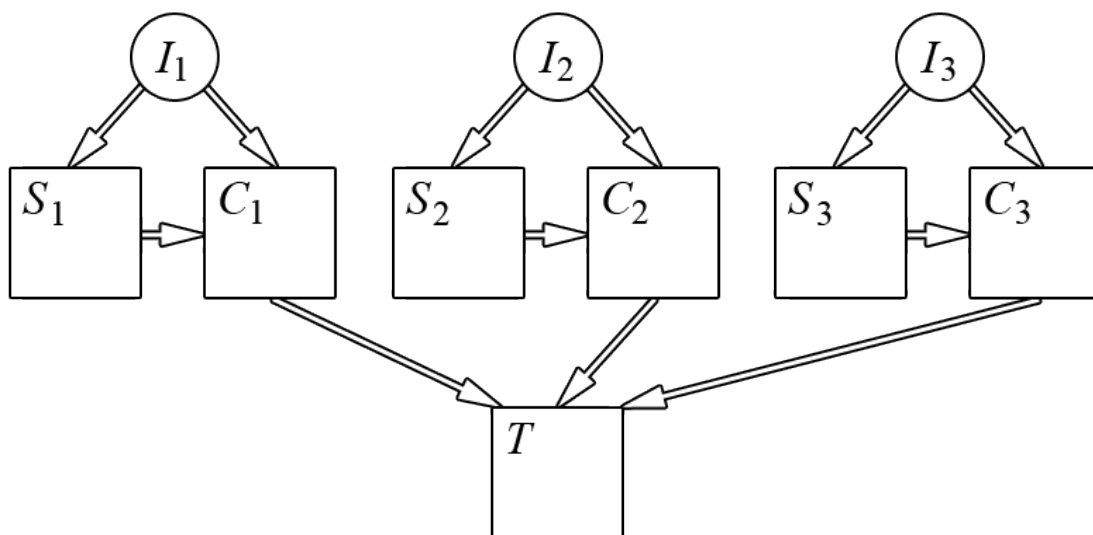


Рис. 2. Слот памяти для кортежа из трех сигналов. Блоки: I_i – сигнал входного кортежа, S_i – ячейка памяти для i -ого сигнала, C_i – блок сравнения сохраненного сигнала с входящим, T – отвечает за обнаружение полного соответствия сохраненного кортежа входящему

Из множества слотов образуется блок памяти, хранящий сигналы от одного из источников. Для каждого источника строится отдельный блок памяти, и между ними задается блок ассоциативных связей, позволяющий объединять в ассоциации кортежи сигналов на основе времени их получения. В предложенной системе ассоциированными считаются одновременно полученные сигналы.

Кроме явно заложенных в ходе обучения правил, выполнение которых происходит в точности согласно сохраненным значениям, система выводит новые возможные правила, используя блок анализа различий, выявляющий общие и частные компоненты кортежей и задающий новые преобразованные правила. Включение этого блока происходит в случае срабатывания для слота одного источника нескольких слотов другого, что означает сходство хранящихся в этих слотах значений и требует соответствующего анализа.

На основе хранящихся в памяти ассоциаций определяется действие с наибольшим весом связи в текущем контексте – это действие считается рекомендуемым решением, которое должно быть применено системой автоматизированного управления.

Пятая глава посвящена созданному в рамках работы практическому инструментарию, необходимому для апробации теоретических результатов. Приводится подробное описание всех основных систем программного комплекса.

Описаны особенности реализации компилятора для предложенного в третьей главе языка. Также в главе ставится задачи размещения графа нейронной сети на плоскости для

наилучшего его наглядного представления в последующих инструментах. Приводится решение задачи с применением алгоритмов укладки графа на основе спектральных методов, используя собственные вектора Лапласиана графа, и силовых алгоритмов Фрюхтермана-Рейнгольда.

Далее показана архитектура среды разработки, особенности реализации версионного хранения данных, разграничения прав и организации проектов. Отдельно указывается интеграция с различными редакторами и компиляторами, которые могут разрабатывать и подключать сторонние разработчики.

В работе особое внимание уделяется системам тестового и распределенного запуска нейронной сети. Для системы тестового запуска дается исчерпывающее описание ее возможностей и набора инструментов для максимально эффективной работы.

Приводится подробное описание архитектуры системы, осуществляющей распределенный запуск нейронных сетей, позволяющей увеличить производительность и мощность за счет использования кластерных систем. Также решается задача размещения графа нейронной сети на граф физических узлов, и доказывается ρ -оптимальность построенного приближенного алгоритма. Кроме того, приводится численный метод решения важной дополнительной задачи по определению времени задержек, в течение которых каждый нейрон должен ожидать получения сигналов. Только по истечении настроенной задержки нейрон считает сигнал по соответствующей связи потерянным и равным нулю.

В шестой главе представлено практическое применение построенной системы принятия решений в интеграции с промышленной системой автоматизированного контроля датчиков и управления множеством устройств (САУ). Дается представление об архитектуре шины, протоколах внутреннего взаимодействия компонентов САУ и принципах событийного управления. В качестве точки интеграции с системой принятия решения используется специальный комплекс виртуальных контроллеров, собирающих события в шине и передающих их в качестве входных сигналов в СПР. Также эти контроллеры отвечают за передачу рекомендованных воздействий на физические устройства в шину, где уже другие компоненты определяют необходимость и метод выполнения указанного воздействия.

В качестве итогов приводятся результаты обучения и использования новой модели нейронных сетей, подтверждающие применимость построенной модели и языка. В частности, приводится методология построения обучающей и тестовой выборки, гарантирующая наилучший показатель обучения, определяемый как число корректных рекомендаций изменений среды к числу входящих состояний датчиков.

Заключение

Согласно поставленным в настоящей работе целям и задачам, была разработана новая модель, названная ВИНС, расширяющая класс решаемых нейронными сетями задач, доказаны ее теоретические основания и связь с классическими моделями ИНС. Кроме того, созданы методы, алгоритмы и инструменты, позволяющие эффективно применять предложенную модель на практике.

В ходе работы была построена математическая модель ВИНС, и доказана достаточность представленных в ней элементов для построения логики произвольной сложности. Для сохранения всех свойств классических нейронных сетей и, в частности, параллельной распределенной природы показана представимость элементов ВИНС в виде простейших нейронов с пороговой функцией активации. Также показана возможность представления классической нейронной сети с использованием элементов ВИНС, чем доказывається, что построенная математическая модель является расширением существующей. Таким образом, была решена задача упрощения использования нейронных сетей, за счет усложнения их минимальных элементов.

Для решения задачи обучения ВИНС был введен второй минимальный элемент – мультипликативный нейрон. В работе продемонстрирована его применимость на основе построения классического примера самообучающейся системы – перцептрона. При этом акцент в построенных ВИНС сделан на то, что обучение производится средствами и алгоритмами самой сети, описанными в ее структуре и с использованием ее терминов. Данный механизм реализован через изменение внутреннего состояния мультипликативного нейрона, эмулирующего изменение веса связи.

Кроме управления весами связей в процессе обучения, предложенная модель дает возможность изменять структуру нейронной сети за счет порождения новых элементов по заранее заданным шаблонам и алгоритмам. В конструктивном доказательстве Теоремы о полноте ВИНС по Тьюрингу приведена методология построения теоретически бесконечной памяти с использованием динамической генерации ВИНС. Эта методология решает задачу расширения модели нейронной сети возможностями изменения структуры в процессе работы.

В работе решена задача по построению численных методов и алгоритмов размещения графа ВИНС на граф физической сети, используемой в системе распределенного запуска. Для обоснования разработанного алгоритма приведена постановка задачи в терминах квадратичной задачи о назначениях, а также доказана корректность примененного алгоритма поиска произвольного минимального разреза и p -оптимальность полученного

приближенного алгоритма. Также представлен алгоритм определения задержек, описывающих периодичность работы нейронов и гарантирующих время накопления нейронами их входящих сигналов. Предложенные численные методы размещения и настройки нейронной сети успешно применены в системе распределенного запуска и дают оптимальный результат в случае его применения в гомогенной кластерной сети физических узлов.

Для возможности практического применения модели ВИНС были созданы язык описания алгоритмов и комплекс программ, а именно: среда разработки, компилятор, среда отладки и тестирования и система распределенного запуска. Среда разработки позволяет не только наглядно оценить промежуточный результат построения графа нейронной сети, но и представить полученный граф на плоскости в нужном виде. Кроме того, среда разработки способна хранить все результаты работы в надежном версионном репозитории. Среда отладки и тестового запуска предоставляет пошаговый интерфейс исполнения ВИНС, выдающий подробную информацию о состоянии каждого элемента и выполняемого им действия в процессе вычисления отдельного такта сети. Система распределенного запуска в противоположность среде отладки обеспечивает максимальную производительность в распределенной кластерной сети, но не предоставляет какой-либо отладочной информации. Каждая система требует определенный формат входных данных и настроек. Был создан специализированный язык описания SAINNL и компилятор, преобразующий исходные тексты к необходимым форматам систем. Таким образом, создан специализированный комплекс программ для работы с ВИНС, решающий одну из задач настоящей работы.

Главным итогом работы является создание нового класса нейронных сетей. Их отличительная черта заключается в возможности представления произвольных алгоритмов в своей структуре, сохраняя основные способности нейронных сетей: изменчивость и обучение. В работе представлены особенности применения новой модели, а также набор инструментов для практического использования. Следует отметить, что при описании алгоритмов в терминах ВИНС они автоматически наследуют свойства нейронных сетей и, в частности, возможность параллельного распределенного запуска.

В качестве практических примеров в работе приведены функциональные блоки, демонстрирующие базовые возможности ВИНС. Также построен ряд фундаментальных примеров, таких как реализация машины Тьюринга и перцептрон, на основании которых делаются выводы о выразительности предлагаемой модели. В качестве примера промышленного использования показано практическое внедрение системы принятия решений на основе ассоциативной памяти в САУ «Умный дом». В работе предложена

методология обучения построенной системы, а также приведены данные о результатах тестовых запусков. Благодаря алгоритмам, заложенным в нейронную сеть системы принятия решений, удалось не только получить итоговый результат, выдаваемый ассоциативной памятью, но и проследить весь процесс определения решения на основе текущего состояния наблюдаемой окружающей среды.

В итоге проделанной теоретической и практической работы можно утверждать, что полученная модель ВИНС отличается от существующих и предоставляет новые возможности применения нейронных сетей в решении задач. Так, данная модель может использоваться в различных промышленных отраслях, где окружающая среда является частью технологического процесса, требующего прогнозируемого и обоснованного поведения. Возможно, система может быть применена и в других областях, например, при создании медицинской системы рекомендаций. Рассмотренная СПР за счет постоянного обучения, надежного хранения знаний, возможности работы при неточных или неполных данных и выдачи обоснований для итоговой рекомендации сможет существенно помочь эксперту при принятии решения.

Дальнейшее развитие модели предполагает как расширение инструментальной базы для практического применения, так и пополнение алгоритмической базы библиотеки с готовыми решениями наиболее часто встречающихся задач.

Список основных работ по теме диссертации

Статьи в рецензируемых научных журналах и изданиях, определенных ВАК РФ:

1. Ионов С.Д. Распределенный запуск нейронных сетей на множестве вычислительных узлов / С.Д. Ионов // Вестник УГАТУ, Т. 17 № 2. – 2013. – С. 85-91.
2. Ионов С.Д. Построение системы ассоциативной памяти с использованием воспроизводимых искусственных нейронных сетей (ВИНС) / С.Д. Ионов, С.В. Шарф // Нейрокомпьютеры: разработка, применение, №9. – 2013. – С. 32-44.

Публикации в других изданиях:

3. Ионов С.Д. Распределенная нейронная сеть: принципы работы и протоколы взаимодействия / С.Д. Ионов // Материалы межвузовской научной конференции по проблемам информатики «СПИСОК-2009». – 2009. – С. 26-29.
4. Ионов С.Д. Распределенная потоковая нейронная сеть / С.Д. Ионов // Тезисы 42-й Всероссийской молодежной школы-конференции «Современные проблемы математики». – 2011. – С. 288-290.

5. Ионов С.Д. Многовариантность применения воспроизводящихся искусственных нейронных сетей в задачах распознавания / С.Д. Ионов // Тезисы докладов XI Всероссийской научной конференции «Нейрокомпьютеры и их применение». – 2013. – С. 43-45.

6. Ионов С.Д. Программирование нейросетевого эмулятора машины Тьюринга / С.Д. Ионов // Тезисы докладов XII Всероссийской научной конференции «Нейрокомпьютеры и их применение». – 2014. – С. 44-45.