

ФОРМАЛЬНОЕ ОПИСАНИЕ ТЕХНОЛОГИИ ИСПОЛЬЗОВАНИЯ НЕЙРОННЫХ СЕТЕЙ С ИСПОЛЬЗОВАНИЕМ АТРИБУТИВНЫХ ТРАНСЛИРУЮЩИХ ГРАММАТИК

Л.А. Жуков

Сибирский государственный технологический университет
г. Красноярск

В работе приведены особенности построения и использования формальных языков и грамматик для описания технологии использования нейронных сетей при решении прикладных задач. Обсуждается использование атрибутивных грамматик. Модели использовались при обучении дисциплинам «Нейрокомпьютерные системы» и «Формальные языки и грамматики» и для проведения исследований.

Цель работы – построение и использование моделей описания технологии использования нейронных сетей с использованием формальных грамматик.

Задачи:

- 1) построение формальных моделей использования нейронных сетей разных типов;
- 2) использование построенных формальных моделей для решения различных прикладных задач, в том числе для языкознания;
- 3) использование формальных моделей для повышения эффективности обучения студентов и школьников старших классов работе с нейронными сетями;
- 4) расширение формальных моделей в виде контекстно-свободных грамматик до атрибутивной формы.

Первая задача - разработка модели или схемы использования различных операций. Модель построена для определения конкретного порядка операций для решения задачи, что может подсказать исследователю, какие операции выполнять сначала, что следует делать после выполненных операций, осуществить контроль за действиями исследователя и проверить правильность выбранного им интуитивного порядка следования операций. Для построения формальных моделей на данном этапе работы использованы формальные языки и грамматики [1, 2].

Существующие формальные языки описывают особенности реализации архитектуры нейронных сетей или особенности описания структур сетей для переноса на другое программное обеспечение. Примером могут служить работы Гилева С.Е., Горбаня А.Н., Мир-

кеса Е.М., Дорогова А.Ю. и многих других отечественных и зарубежных исследователей [3][4][5]. Расширение большинства известных языков не является возможным, т.к. они описывают аппаратные нейросетевые модели, а не их реализацию в специализированных пакетах и не их использование конечными исследователями. Известные модели предлагают нижний уровень детализации, если использовать параллели с языками программирования – уровень ассемблера или Си.

Отсутствует универсальное и полное описание технологии использования нейронных сетей для решения прикладных задач, что безусловно является важным компонентом при проведении исследовательских работ. Многие исследователи самостоятельно выбирают и применяют различные нейросетевые методы. Начинаящие исследователи решают поставленную задачу методом проб и ошибок, сами определяют конкретную последовательность операций, зачастую не представляют последствия и результаты, полученные при этом. Особенно это характерно для студентов, следовательно, проблема становится актуальной и для ведущего занятия преподавателя. Можно сказать, что проблема многих исследований с использованием НКС – методика выполнения исследования, последовательность, интерпретация результатов, разнообразие подходов. У каждого исследователя – своя технология, чаще всего она не описана и всегда недостаточно формализована. Встречается большое число ошибок. Требуется формализованная технология работы и для пользователей, и для НКС. Пользователю она облегчит выполнение работы, для программных или аппаратных НКС позволит перейти к автоматизированному выполнению технологии. Язык описания формализованной технологии должен быть высокого уровня.

Предлагается семейство формализованных технологий использования нейронных сетей, ориентированных на конечного пользователя. Может быть уместно сравнение

ФОРМАЛЬНОЕ ОПИСАНИЕ ТЕХНОЛОГИИ ИСПОЛЬЗОВАНИЯ НЕЙРОННЫХ СЕТЕЙ С ИСПОЛЬЗОВАНИЕМ АТРИБУТИВНЫХ ТРАНСЛИРУЮЩИХ ГРАММАТИК

данной проблемы и предлагаемого решения с использованием формальных языков, в том числе БНФ, для формализации языков программирования. В определенном смысле можно ограничить решаемую задачу следующим образом. Моделируется последовательность выполнения операций, каждый символ терминального алфавита представляет собой отдельную операцию, слово представляет собой последовательность операций.

Результаты данной работы могут быть использованы для организации исследований пользователями нейросетевых программ и инструментов, и для оптимизации разработки аппаратных нейросистем и их имитаторов.

В данной работе представлен верхний уровень – описание использования готового программного и аппаратного обеспечения. Если привести некоторую аналогию с аппаратным обеспечением – известные модели, это как описание работы процессора, памяти и т.п., а данная модель – как описание операционной системы. Может быть уместно сравнение данной работы с использованием формальных языков, в том числе БНФ, для формализации языков программирования.

В определенном смысле можно ограничить решаемую задачу следующим образом. Моделируется последовательность выполнения операций, каждый символ терминального алфавита представляет собой отдельную операцию, если хотите нажатие кнопки пользователем, слово представляет собой последовательность операций.

Следует заметить, что первые работы в этом направлении были сделаны для формализации нейронных сетей с учителем (supervised neural network), которые чаще называют двойственными сетями, сетями обратного распространения, сетями BackProp и т.д. (Нейронные сети с учителем – это сети обратного распространения или двойственные нейронные сети; учитель выполняет функцию обучения сети, с точки зрения математики происходит минимизация функции оценки). К сожалению, распространенные названия такого типа не включают явно сети, обучаемые случайными методами, например, с использованием генетических алгоритмов. Термин «нейронные сети с учителем», будучи менее используемым, включает другие виды методов и алгоритмов обучения и потому более точно определяет область выполненных работ [6]. Формальные модели для этого вида сетей подробно рассмотрены в указанных работах, в данной работе не при-

водятся. Кроме того, была выполнена работа по формализации сетей Хопфилда [7].

Для построения формально-языковой модели нейронных сетей без учителя использовалось текстовое описание технологии в контексте применения нейронных сетей без учителя пользователем для решения прикладной задачи [6][7]. Модель предполагает операции пользователя атомарными, т.е. неделимыми и уже реализованными в специализированных программах-нейроимитаторах, например STATISTICA с пакетом Neural Networks, MDN [8] и других. Технология представлена в виде формального описания.

Следует отметить, что на первых двух уровнях детализации не обнаружено существенного различия алфавита языка (семейства языков) от аналогичных моделей для нейронных сетей с учителем. Отличия алфавита появляются на следующем уровне. Одно из самых существенных отличий – нет почти всех видов контрастирования, а контрастирование (минимизация) входных сигналов применимо только в отдельных случаях, специальными методами, почти выводящими систему из типа «сети без учителя». Кроме того, существенно сокращается использование процедур бинаризации.

Предложенные модели и методы использованы при проведении занятий с аспирантами, студентами и школьниками в СибГТУ и КГТУ, в Зимних Политехнических Школе по Нейроинформатике, проходивших в 2000–2007 годах.

Формальная методика использована для решения нескольких задач. В задаче классификации на лексико-грамматические классы слов английского языка исследовалась возможность определения принадлежности слов английского языка к части речи (существительное, глагол, прилагательное и т.п.). Для проведения экспериментов использовалось текстовое описание методики применения нейронных сетей с учителем [8]. Результаты данной работы могут быть использованы при изучении английского языка, при построении переводчиков и т.д. Выделение «трудно распознаваемых» классов слов в некоторых случаях может быть рекомендовано в качестве списка слов для заучивания. «Легко распознаваемые» слова (по части речи) могут быть определены по структуре и месту расположения в контексте.

Для задачи определения продолжительности сушки зерна входными данными были использованы температура и влажность зерна и воздуха и некоторые другие параметры.

При сушке зерна важно получить зерно с влажностью в ограниченном диапазоне. Формальная методика работы позволила эффективно использовать знания специалиста в предметной области, поставить задачу, организовать данные и процесс вычислений, получить результаты прогноза.

Формальное описание нейронных сетей различных типов и технологии работы с использованием бесконтекстных формальных грамматик позволяет удовлетворительно решить задачу описания. Например, в книге Миркеса Е.М. основной метаязык – модифицированные БНФ [5], которые также являются подвидом контекстно-свободных грамматик. Для данной работы при описании технологии ранее также использовались контекстно-свободные грамматики.

При этом возникали некоторые трудности при описании передачи параметров. Предлагается расширить грамматики [6] до атрибутивных контекстно-свободных грамматик, в которых атрибуты будут использованы для передачи параметров. Атрибутивные грамматики, которые являются в основе своей бесконтекстными, но позволяют связывать с символами формальной грамматики не только понятие «класс», но и «значение». Наследуемые и синтезируемые атрибуты, используемые для описания технологии, позволяют учесть полученные на предшествующих этапах результаты, для повышения эффективности технологии.

В ходе исследования был выявлен минимальный набор атрибутов: имя файла данных, имя/имена обученных сетей, количество сетей, счетчик сетей (номер используемой сети), число нейронов, число слоев, число нейронов в слоях, количество входных параметров, количество выходных параметров, качество обучения, значимость входных параметров, количество параметров и другие. Можно особо выделить наиболее важные атрибуты: текущая оценка, оценка на предыдущем шаге функционирования, количество тактов обучения, количество классов и др.

При создании семейства атрибутивных грамматик выявилась необходимость дополнительно выделить несколько видов атрибутов. Фактический атрибут – атрибут с фиксированной семантикой, соответствующей определенному значению. Формальный атрибут – атрибут с фиксированным синтаксисом, использование которого в правилах грамматики происходит с изменением буквы, т.е. его семантика может меняться при необходимости. Абстрактный атрибут – атри-

бут с недоопределенным синтаксисом или семантикой, другими словами – информационная модель, которая может быть полезна в будущем, но пока в явном виде представлена быть не может и не использована.

После модификации грамматик в атрибутивные технология была использована для прогнозирования оценки успеваемости студентов четвертого курса математического факультета КГУ (СФУ) 2007 года обучения по предмету СППО. Обучение нейронных сетей выполнялось по результатам обучения студентов четвертого курса в 2006 году. Количество правильных ответов по тестовой выборке составляло от 70 до 95 процентов, в зависимости от сети и изменяемых условий обучения.

Выводы.

1. Одна из задач, решаемых в ходе моделирования – устранение некорректного порядка следования операций. Следует заметить, что ответом может служить несколько пригодных вариантов действий, т.к. некоторые операции допускают перестановку. Задавать вопрос «Какой из технологических процессов правильный?» неправильно, т.к. пока еще не существует общепризнанных критериев определения оптимального решения для нейронных сетей.

2. Одна из основных возможностей использования – для начинающего пользователя и пользователя среднего уровня. Начинающий пользователь иногда не может определить какое решение выбрать и даже с чего начать, в процессах может присутствовать заведомо неправильный порядок действий и операций, он может сделать неправильный вывод, основываясь на неверных данных. В этом случае важным является отсеять некорректные последовательности. Например, при определении значимостей параметров необходимо учитывать результаты тестирования, если этого не было сделано, выводы будут неверными.

3. Использование формальной модели из атрибутивных контекстно-свободных грамматик не показало существенных преимуществ применения для исследователей. Разумеется, для оптимизации разработки аппаратных нейросистем и их имитаторов использование формальных моделей должно быть проверено отдельно.

4. С использованием описанных технологий проведены многочисленные эксперименты для решения некоторых лингвистических и других задач.

ФОРМАЛЬНОЕ ОПИСАНИЕ ТЕХНОЛОГИИ ИСПОЛЬЗОВАНИЯ НЕЙРОННЫХ СЕТЕЙ С ИСПОЛЬЗОВАНИЕМ АТРИБУТИВНЫХ ТРАНСЛИРУЮЩИХ ГРАММАТИК

Список литературы

1. Ахо А., Теория синтаксического анализа, перевода и компиляции: Синтаксический анализ / А. Ахо, Дж. Ульман – М.: Мир, – 1978. – Т.1. – 487 с.
2. Молчанов А.Ю. Системное программное обеспечение / А.Ю. Молчанов – СПб.: Питер, 2003. – 396 с.
3. Горбань А.Н. Обучение нейронных сетей / А.Н. Горбань – М.: СП «Paragraph», 1990. – 160 с.
4. Дорогов А.Ю. Быстрые нейронные сети / А.Ю. Дорогов М.: Изд-во Санкт-Петербургского университета, 2002. – 77 с.
5. Миркес Е.М. Нейрокомпьютер. Проект стандарта. – Новосибирск: Наука, 1999. – 337 с.
6. Жуков Л.А. Формализация технологии применения нейронных сетей с учителем и особенности их использования для решения прикладных задач / Л.А. Жуков, Н.В. Решетникова. – Красноярск: ИПЦ КГТУ, 2005. – 168 с.
7. Жуков Л.А., Решетникова Н.В., Корчевская О.В. О формализации нейросетевой технологии решения прикладных задач на примере сетей с учителем и сетей Хопфилда // Научная сессия МИФИ-2005. VII Всероссийская научно-техническая конференция «Нейроинформатика-2005»: Сб. науч. тр. В 2-х частях. Ч.1. – М.: МИФИ, 2005. – С.68 – 75.
8. Жуков Л.А. Формально-языковая модель технологии использования нейронных сетей с учителем и без учителя // Лесной комплекс – проблемы и решения: Всеросс.конф. – Красноярск, СибГТУ, 2005.