МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА

Механико-математический факультет

На правах рукописи

Половников Владимир Сергеевич

УДК 519.95

Об оптимизации структурной реализации нейронных сетей

01.01.09 — дискретная математика и математическая кибернетика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва — 2008

Работа выполнена на кафедре Математической теории интеллектуальных систем Механико-математического факультета Московского государственного университета имени М.В. Ломоносова.

Научный руководитель — кандидат физико-математических наук, доцент
Анатолий Александрович Часовских

Официальные оппоненты:
доктор физико-математических наук, профессор
Алексей Иванович Чуличков
кандидат физико-математических наук, доцент
Сергей Иванович Карташов

Ведущая организация — Московский Энергетический Институт
(Технический университет)

Защита диссертации состоится “18” апреля 2008 г. в 16 ч. 40 м. на заседании диссертационного совета Д.501.001.84 при Московском государственном университете имени М.В. Ломоносова по адресу: 119991, Российская Федерация, Москва, ГСП-1, Ленинские горы, Московский государственный университет имени М.В. Ломоносова, Механико-математический факультет, аудитория 14-08.

С диссертацией можно ознакомиться в библиотеке Механико-математического факультета МГУ имени М.В. Ломоносова (Главное здание, 14 этаж).

Автореферат разослан “18” марта 2008 г.

Ученый секретарь
диссертационного совета
Д.501.001.84 при МГУ
доктор физ.-мат. наук, профессор

А.О. Иванов

Общая характеристика работы

Актуальность темы

Многие практические проблемы могут быть решены с использованием нейронных сетей. Нейронные сети хорошо зарекомендовали себя как средство решения задач для которых практически невозможно использование статических методов (без обучения, адаптации). Это задачи классификации, распознавания образов, моделирования ассоциативной памяти и другие. Нейронные сети также применяются для нахождения приближенного решения NP-полных задач за приемлемое время. Однако, актуальна проблема синтеза нейронных сетей: как построить нейросеть способную решать поставленную задачу. Для начала необходимо представить исходные данные в виде, приемлемом для обработки нейронными сетями. В идеальном варианте после предварительной обработки исходных данных мы получим линейно разделимую задачу, что значительно упрощает построение нейросети-классификатора, которой в этом случае может быть представлена персептроном Розенблатта[1]. Хорошо известен алгоритм обучения такой нейросети, сходимость которого математически обоснована (Розенблатт[2], Новиков[3]). К сожалению, при решении реальных задач, как правило, такая предобработка данных невозможна. Минский и Пэперт [4] в своей работе "Персептроны" доказали, что простейшие однослойные нейронные сети способны решать только линейно разделимые задачи. Однако, это ограничение преодолимо при использовании многослойных нейронных сетей. Значит, необходимо переходить к более сложной архитектуре, чем у персептрона. В теории непрерывных нейронных сетей (с достаточно гладкой функцией активации) часто используется многослойный персептрон, а для его обучения — алгоритм обратного распространение ошибки[5], который, как известно, не применим для дискретных нейронных сетей, в том числе и для сетей модели Мак-Каллока-Питтса[1]. К сожалению, заранее не известно, какой сложности (размера) может потребоваться сеть для достаточно точной реализации требуемого отображения. К тому же может оказаться, что рассматриваемая нейронная сеть не обучаема для решения поставленной задачи.

Для построения математических моделей различных процессов удобно применить понятие функциональной системы введенное впервые Кудрявцевым[6]. В рамках понятия функциональной системы можно рассматривать функции алгебры

[1] Хайкин С. *Нейронные сети: полный курс, 2-е издание*, Вильямс, 2006.

[2] F. Rosenblatt *The perceptron: a probabilistic model for information storage and organization of the brain*, Psychological Review, 65:386-408, 1958.

[3] A. B. J. Novikoff *On convergence proofs on perceptrons*, Proceedings of the Symposium on the Mathematical Theory of Automata, XII:615-622, 1962

[4] Minsky M. L. and S. A. Papert *Perceptorns*, Cambridge, MA: MIT Press, 1969.

[5] Уосермен Ф. *Нейрокомпьютерная техника.* М.: Мир. 1992. – 240 с.

[6] Кудрявцев В. Б. *Функциональные системы.* М., Изд-во МГУ, 1982.

логики с операцией суперпозиции, конечные автоматы[7] с операциями суперпозиции или композиции, однородные структуры[8], схемы из функциональных элементов[9] с операцией суперпозиции и т.п. Унифицированный подход позволил обобщить свойства и методы исследования для различных функциональных систем.

Так в теории нейронных сетей, в частности при решении задачи синтеза, могут быть использованы идеи, возникающие при изучении других функциональных систем.

Цель работы

1. Рассмотреть нейронные сети с точки зрения теории функциональных систем, формализовать понятие обучения нейросети. Разработать аппарат исследования многослойных дискретных нейронных сетей.

2. Разработать аппарат исследования дискретных нейронных сетей. С помощью полученного аппарата исследовать сложностные характеристики оптимальных нейронных схем.

3. Исследовать проблему полноты для функциональной системы нейронных схем с операцией суперпозиции при естественных ограничениях.

4. Построить канонические формы нейронных схем, реализующих функции их определенных классов.

Структура и объем диссертации

Диссертационная работа изложена на 108 страницах и состоит из введения и четырех глав, разделенных на параграфы. Библиография включает 23 наименования.

Научная новизна

1. Доказано, что любая нейронная схема реализует кусочно-линейную функцию и наоборот, любая кусочно-линейная функция может быть реализована нейронной схемой нелинейной глубины не более двух.

2. Доказано, что любая нейронная схема Мак-Каллока-Питтса реализует кусочно-параллельную функцию и наоборот, любая кусочно-пралллельная функция

[7] Кудрявцев В. Б., Алешин С. В.,Подколзин А. С. *Введение в теорию автоматов.* М.:Наука.Гл.ред.физ.-мат.лит., 1985.
[8] Болотов А. А., Кудрявцев В. Б., Подколзин А. С. *Основы теории однородных структур.* М.: Наука, 1990.
[9] Лупанов О. Б. *Асимптотические оценки сложности управляющих систем.* М., МГУ, 1984.

может быть реализована нейронной схемой Мак-Каллока-Питтса нелинейной глубины не более двух.

3. Получен критерий реализуемости кусочно-линейных функций нейронной схемой нелинейной глубины не более 1.

4. Для функциональной системы нейронных схем Мак-Каллока-Питтса найден эффективный критерий полноты множеств функций, содержащих в замыкании все линейные функции.

5. Получены оценки функции Шеннона для сложности схем Мак-Каллока-Питтса и сложности нейронных схем, имеющих единичную нелинейную глубину I рода, которые приводят к асимптотике при фиксированной размерности признакового пространства.

6. С помощью системы канонических уравнений описано функционирование нейронных схем во времени. Приведен канонический вид нейронной схемы с памятью: доказано, что для любой нейронной схемы можно получить эквивалентную ей однократным применением операции обратной связи на последнем этапе ее построения.

Основные методы исследования

В диссертации использованы методы теории автоматов, теории функциональных систем, комбинаторики, алгебры, математического анализа.

Теоретическая и практическая ценность работы

Работа носит теоретический характер и может быть полезна специалистам, работающим в области теории нейросетей. Результаты могут быть использованы при решении практических задач методом нейросетевого моделирования для предварительной оценки сложностных параметров сети.

Апробация работы

Результаты диссертации докладывались на конференциях и семинарах Механико-Математического факультета МГУ им. М.В.Ломоносова: на семинаре "Теория автоматов" под руководством академика, профессора В.Б. Кудрявцева (неоднократно в 2005-2007 гг.), на семинаре "Математические вопросы кибернетики" под руководством профессора О.М. Касим-Заде (ноябрь 2007 г.), на семинаре "Нейронные сети” под руководством доцента А.А. Часовских (неоднократно 2002-2007 гг.), на семинаре “Теория дискретных функций и приложения” под руководством профессора Д.Н. Бабина (сентябрь 2002 г.), а также на конференции ”Ломносовские чтения” в 2006 и 2007 гг. Москва, МГУ им. М.В. Ломоносова; на IX международной конференции “Интеллектуальные системы и компьютерные науки” Москва, МГУ им. М.В. Ломоносова 2006 г.

Публикации по теме диссертации

Основные результаты диссертации опубликованы в пяти статьях, список которых приведен в конце автореферата. [1-4]

Краткое содержание работы

Во введении содержится краткая история развития теории нейронных сетей и некоторых функциональных систем, приводятся формулировки необходимых понятий и фактов, ссылки на них.

Глава 1

В первой главе изучаются схемы из функциональных элементов полученных из элементных базисов с помощью операций суперпозиции. В главе рассматриваются два базиса, исследуются пространства функций выразимых схемами над данными базисами. Вводится понятие одной из сложностных характеристик нейронных схем — нелинейной глубины, соответствующей количеству слоев в нейросети, и доказывается, что для представления любой выразимой функции достаточно нейронных схем нелинейной глубины два. Подробно исследуются класс функций, представимых схемами нелинейной глубины один.

В параграфе 1.1 даются основные определения. Вводятся следующие функции:

1) постоянная функция $g_c \equiv c$, $c \in \mathbb{R}$;
2) сумматор $\Sigma_n(x_1, \ldots, x_n) = x_1 + \ldots + x_n$, где $x_i \in \mathbb{R}$
при $i = 1, 2, \ldots, n$;
3) усилитель $f_\gamma(x) = \gamma x_1$, где $\gamma, x \in \mathbb{R}$;
4) функция $\theta(x) = \begin{cases} 1, & \text{если } x \geq 0 \\ 0, & \text{если } x < 0, \end{cases}$ где $z \in \mathbb{R}$;
5) функция $F(x_1, x_2) = \begin{cases} x_1, & \text{если } x_2 > 0 \\ 0, & \text{если } x_2 \leq 0, \end{cases}$ где $x_1, x_2 \in \mathbb{R}$.

Множество всех таких функций обозначается через Δ'.

Далее, по аналогии с автоматными схемами определяются схемы из функциональных элементов Δ', называемые нейронными схемами без памяти.

Вводятся понятия нелинейной глубины и сложности таких схем.

Путем в нейронной схеме называется последовательность функциональных элементов $F_1, \ldots, F_k$, где $F_i \in \Delta'$. Причем один из входов F_1 является входом нейронной схемы, выход F_k является выходом нейронной схемы, а выход F_i — один из входов F_{i+1}, $i = 1, \ldots, k-1$. Рассматриваются все пути в нейронной схеме без памяти. Длиной пути называется число нелинейных элементов (F или θ), содержащихся в нем. Нелинейной глубиной нейронной схемы называется длина самого длинного пути.

Число нелинейных элементов в схеме называется нелинейной сложностью нейронной схемы.

Множество всех линейных функций обозначается через L, далее определяются кусочно-линейные функции многих переменных.

Пусть l_i — гиперплоскость, задаваемая уравнением $\bar{x} \cdot \bar{a}_i + c_i = 0$, $\bar{a}_i \in \mathbb{R}^n \backslash \{0\}$, $c_i \in \mathbb{R}$, $i = 1, \ldots, k$. Для каждой точки $\bar{x} \in \mathbb{R}^n$ рассмотрим вектор $\sigma(\bar{x}) = (\sigma_1, \ldots, \sigma_k)$ с компонентами из множества $\{-1, 0, 1\}$, $\sigma_i = sgn(\bar{x} \cdot \bar{a}_i + c_i)$, где

$$sgn(b) = \begin{cases} -1, & b < 0 \\ 0, & b = 0 \\ 1, & b > 0. \end{cases}$$

Две точки $\bar{x}, \bar{y} \in \mathbb{R}^n$ эквивалентны относительно гиперплоскостей $l_1, \ldots, l_k$ тогда и только тогда, когда $\sigma(\bar{x}) = \sigma(\bar{y})$, что обозначается $\bar{x} \sim \bar{y}$.

Легко проверить, что отношение "$\sim$" действительно является отношением эквивалентности. Таким образом, пространство $\mathbb{R}^n$ разбивается на классы эквивалентности $R_1, \ldots, R_s$.

Сигнатурой класса R называется вектор $\sigma(R) = \sigma(\bar{x})$, где $\bar{x}$ — точка класса R.

Пусть $R_1, \ldots, R_s$ — все классы эквивалентности на которые гиперплоскости $l_1, \ldots, l_k$ разбивают $\mathbb{R}^n$.

Функция $f : \mathbb{R}^n \to \mathbb{R}$ называется *кусочно-линейной*, если $\forall j \in \{1, \ldots, s\}$ найдутся $\bar{b}_j \in \mathbb{R}^n$ и $d_j \in \mathbb{R}$, что для всех $\bar{x} \in R_j$ выполняется $f(\bar{x}) = \bar{x} \cdot \bar{b}_j + d_j$. Множество всех кусочно-линейных функций обозначаются PL.

Функция $f : \mathbb{R}^n \to \mathbb{R}$ называется *кусочно-постоянной*, если $\forall j \in \{1, \ldots, s\}$ найдутся $d_j \in \mathbb{R}$, что для всех $\bar{x} \in R_j$ выполняется $f(\bar{x}) = d_j$. Множество всех кусочно-постоянных функций обозначаются PC.

В параграфе 1.2 доказывается теорема о нелинейной глубине нейронных схем без памяти и показывается, что схем нелинейной глубины один не достаточно для реализации всех кусочно-линейных функций.

Теорема 1 *Для любой нейронной схемы без памяти существует эквивалентная ей схема с нелинейной глубиной не больше 2.*

Доказательство теоремы является конструктивным.

Утверждение. Существует нейронная схема, для которой нет эквивалентной схемы нелинейной глубины меньше двух.

Для доказательства утверждения рассмотрена нейронная схема, реализующая следующую кусочно-линейную функцию:

$$\delta(x,y) = \begin{cases} 1, & x = 0 \text{ и } y = 0 \\ 0, & \text{иначе.} \end{cases}$$

В параграфе 1.3 подробно изучены схемы нелинейной глубины один: вводятся необходимые определения и доказывается основная теорема параграфа.
Пусть кусочно-линейная функция f задана гиперплоскостями $l_1, \ldots, l_k$. Любая тройка классов эквивалентности (R_+, R_0, R_-), $R_+, R_0, R_- \in \{R_1, \ldots, R_s\}$, таких, что элементы векторов сигнатур этих классов удовлетворяют условиям: $\sigma_+^t = \sigma_0^t = \sigma_-^t$ при $t = 1, \ldots, i-1, i+1, \ldots, n$, а $\sigma_+^i = 1$, $\sigma_0^i = 0$ и $\sigma_-^i = -1$, называется смежной тройкой к гиперплоскости l_i, где $(\sigma_+^1, \ldots, \sigma_+^n) = \sigma(R_+)$, $(\sigma_0^1, \ldots, \sigma_0^n) = \sigma(R_0)$ и $(\sigma_-^1, \ldots, \sigma_-^n) = \sigma(R_-)$.

Если для гиперплоскости l_i функции f найдутся линейные функции $l_{+;i}(\bar{x}) = \bar{a}_{+;i}\bar{x} + c_{+;i}$ и $l_{-;i}(\bar{x}) = \bar{a}_{-;i}\bar{x} + c_{-;i}$, такие что для всех троек (R_+, R_0, R_-) смежных с l_i выполняются равенства $f_{R_+} - f_{R_0} = l_{+;i}$ и $f_{R_-} - f_{R_0} = l_{-;i}$, то тройка $(l_i, l_{+;i}, l_{-;i})$ называется *переходом* f через гиперплоскость l_i. В противном случае переход через гиперплоскость l_i не существует. Если гиперплоскость l не совпадает ни с одной из гиперплоскостей $l_1, \ldots, l_k$ задающих f, то тройка $(l, 0, 0)$ также является переходом для f через l. Переход называется нетривиальным, если хотя бы одна из двух его последних компонент отлична от нуля.

Теорема 2 *Пусть f — кусочно-линейная функция, заданная гиперплоскостями $l_1, \ldots, l_k$. Она представима нейронной схемой нелинейной глубины один с некоторой нелинейной сложностью p тогда и только тогда, когда для $i = 1, \ldots, k$ у f существуют переход $(l_i, l_{+;i}, l_{-;i})$ через гиперплоскость l_i, причем в последовательности $l_{+;1}, l_{-;1}, \ldots, l_{+;k}, l_{-;k}$ не более p ненулевых компонент, т.е. выполнено соотношение*

$$2k - \sum_{i=1}^{k} \left(\chi(l_{+;i}, 0) + \chi(l_{-;i}, 0)\right) \le p.$$

Где функция $\chi(a,b)$ определяется равенством

$$\chi(a,b) = \begin{cases} 1, & \text{при } a = b \\ 0, & \text{иначе.} \end{cases}$$

Следствие. Кусочно-линейную функцию f можно реализовать нейронной схемой нелинейной глубины один тогда и только тогда, когда существует натуральное

k, что для некоторых $\bar{a}_{+;i}, \bar{a}_{-;i}, \bar{b} \in \mathbb{R}^n$ и $c_{+;i}, c_{-;i}, d \in \mathbb{R}$, $i = 1, \ldots, k$, f представима в виде

$$f(\bar{x}) = \sum_{i=1}^{k} (F(\bar{a}_{+;i}\bar{x} + c_{+;i}, \bar{a}_i \cdot \bar{x} + c_i) + F(\bar{a}_{-;i}\bar{x} + c_{-;i}, -\bar{a}_i \cdot \bar{x} - c_i)) + \bar{b} \cdot \bar{x} + d,$$

В параграфе 1.4 подробно изучаются схемы Мак-Каллока-Питтса, получаемые из элементов базиса Δ' без использования элемента "F". Определяется множество кусочно-параллельных функций:

$$PP = \{f | f = f_c + f_l, f_c \in PC, f_l \in L\}.$$

Для этого подкласса нейронных схем доказывается теорема, аналогичная теореме 1.

Теорема 3 *Множество функций, реализуемых нейронными схемами Мак-Каллока-Питтса, совпадает с множеством всех кусочно-параллельных функций. И для любой кусочно-параллельной функции существует нейронная схема Мак-Каллока-Питтса нелинейной глубины не более 2.*

Глава 2

В главе 2 рассматривается типичная для функциональных систем задача полноты. Для начала вводятся необходимые определения.

Функция одной переменной $f(t)$ из PC называется *C-финитной функцией*, если $\exists T_f, c_f \in \mathbb{R}$, $T_f > 0$ такие, что при $t \in (-\infty, -T_f) \cup (T_f, +\infty)$ выполнено $f(t) = c_f$. Обозначим класс всех C-финитных функций от 1 переменной $\Phi(1)$.

Для кусочно-постоянной функции $f : \mathbb{R}^n \to \mathbb{R}$ и прямой $\mathfrak{N}$, заданной параметрически: $x_1 = a_1 t + b_1, \ldots, x_n = a_n t + b_n$, функция $g(t) = f(a_1 t + b_1, \ldots, a_n t + b)$ называется сечением функции f прямой $\mathfrak{N}$.

Функция $f \in PC$ называется C-финитной, если сечение f любой прямой является C-финитной функцией от 1 переменной. Обозначим класс всех C-финитных функций через Φ.

Вводится класс C-финитно-линейных функций:

$$\Phi L = \{f | f = \phi + l, \phi \in \Phi, l \in L\}.$$

Доказывается утверждение о замкнутости рассмотренного класса.

Лемма. Класс ΦL замкнут относительно операций суперпозиции.

Далее приводится доказательство основной теоремы главы.

Теорема 4 *Пусть замыкание множества B кусочно-параллельных функций содержит все линейные функции. Тогда замыкание B по операции суперпозиции совпадает с PP в точности тогда и только тогда, когда B целиком не содержится в ΦL.*

Доказательство теоремы конструктивное и позволяет с помощью схемы из линейных элементов и любой не C-финитно-линейной функции выразить θ-функцию.

Глава 3

В главе 3 приведены оценки сложности минимальных нейронных схем. Через $Z(S)$ обозначена нелинейная сложность схемы S. Для произвольной кусочно-параллельной функции $f(x_1, \ldots, x_n)$ в $\mathbb{R}^n$, заданной k гиперплоскостями вводится сложность ее реализации схемами Мак-Каллока-Питтса

$$Z_{MP}(f) = \min_{S \text{ реализует} f} Z(S),$$

где минимум берется по всем нейронным схемам Мак-Каллока-Питтса реализующим f.

Вводится функция Шеннона $Z_{MP}(k) = \max\limits_{f \in PP: f_{\text{задана } k \text{ гиперплоскостями}}} Z_{MP}(f).$

При этом уточняется понятие нелинейной глубины. Пусть для пути $F_1, \ldots, F_p$ схемы S выполнены следующие условия:

1. Для всех $i = 1, \ldots, p-1$ выполнено, что если F_{i+1} является элементом "F", то выход F_i соединен лишь с первым входом F_{i+1};
2. Для всех $i = 1, \ldots, p$ выполнено $F_i \neq \theta$.

Тогда рассматриваемый путь называется путем I рода. Максимальное число нелинейных элементов "F", в таких путях схемы S называется нелинейной глубиной I рода для нейронной схемы S. Рассматриваются простые нейронные схемы без памяти, нелинейная глубина I рода которых не превосходит 1.

Для простых схем определяются аналогичные введенным ранее функции:

$$Z(f) = \min_{S \text{ реализует} f} Z(S), \text{ где минимум берется по всем}$$

простым нейронным схемам реализующим f.

Функция Шеннона $Z(k) = \max\limits_{f \in PL: f_{\text{задана } k \text{ гиперплоскостями}}} Z(f).$

Пусть $H(n, k)$ — максимальное число классов эквивалентности, на которые k гиперплоскостей могут разбить $\mathbb{R}^n$. Известна теорема Шлефли, позволяет подсчитать число n-мерных открытых многогранных конусов, которые получаются

в результате разбиения k гиперплоскостями пространства $\mathbb{R}^n$. В работе использованы следующие вспомогательные утверждения:
Лемма. Для всех натуральных n и k верно $H(1,k)=2k+1, H(n,1)=3$. А для $n\geq 2$ и $k\geq 2$ выполнено неравенство

$$H(n,k)\leq H(n,k-1)+2H(n-1,k-1).$$

Лемма. При $k>n$ выполнено

$$H(n,k)=3\sum_{i=0}^{n-1}C_{k-1}^{i}2^i+2^n\sum_{j=n-1}^{k-2}C_j^{n-1}.$$

Далее в диссертации приводятся оценки для функций $Z_{MP}(k), Z(k)$ и их асимптотики.

Теорема 5 *Для функции Шеннона $Z_{MP}(k)$ верно:*
1) при $k\leq n$:

$$3^k-1\leq Z_{MP}(k)\leq 3^k+2k.$$

2) при $k>n$:

$$3\sum_{i=0}^{n-1}C_{k-1}^{i}2^i+2^n\sum_{j=n-1}^{k-2}C_j^{n-1}-1\leq Z_{MP}(k)\leq 3\sum_{i=0}^{n-1}C_{k-1}^{i}2^i+2^n\sum_{j=n-1}^{k-2}C_j^{n-1}+2k.$$

Следствие. Для функции Шеннона $Z_{MP}(k)$ при $n=const$, $n>1$ верна асимптотика

$$Z_{MP}(k)\sim\frac{2^n}{n!}k^n,\ k\to\infty.$$

Теорема 6 *Для функции Шеннона $Z(k)$ верно:*
1) при $k\leq n$:

$$3^k-1\leq Z(k)\leq 3^k+2k.$$

2) при $k>n$:

$$3\sum_{i=0}^{n-1}C_{k-1}^{i}2^i+2^n\sum_{j=n-1}^{k-2}C_j^{n-1}-1\ \leq\ Z(k)\ \leq\ 3\sum_{i=0}^{n-1}C_{k-1}^{i}2^i+2^n\sum_{j=n-1}^{k-2}C_j^{n-1}+2k.$$

Следствие. Для функции Шеннона $Z(k)$ при $n=const$, $n>1$ верна асимптотика

$$Z(k)\sim\frac{2^n}{n!}k^n,\ k\to\infty.$$

Глава 4

В четвертой главе изучаются нейронные схемы с памятью, описано функционирование таких схем во времени с помощью системы канонических уравнений. Задача минимизации количества применений операции обратной связи для произвольной схемы сводится к однократному применению этой операции на последнем этапе построения схемы эквивалентной исходной.

В параграфе 4.1 рассматриваются нейронные схемы с памятью. Через $\mathfrak{Z}_{a_0}$ обозначим детерминированную функцию, реализующую единичную задержку с начальным состоянием a_0, $a_0 \in \mathbb{R}$, а через Δ — новый базис $\Delta = \Delta' \cup \{\mathfrak{Z}_{a_0} | a_0 \in \mathbb{R}\}$.

Схема, полученная из элементов базиса Δ с использованием операции суперпозиции и обратной связи, называется нейронной схемой (с памятью).

Детерминированная функция $f : (\mathbb{R}^\infty)^n \to \mathbb{R}^\infty : \tilde{z} = f(\tilde{x}_1, \ldots, \tilde{x}_n)$, для которой существуют числа $k \in \mathbb{N}$, $a_0^1, \ldots, a_0^k \in \mathbb{R}$ и кусочно-линейные функции $Z, \Phi_1, \ldots, \Phi_k : \mathbb{R}^n \to \mathbb{R}$, что выполнены канонические уравнения

$$\begin{cases} \bar{d}(0) = (a_0^1, \ldots, a_0^k) \\ \bar{d}(t+1) = \bar{\Phi}(\bar{d}(t), \bar{x}(t+1)) = (\Phi_1(\bar{d}(t), \bar{x}(t+1)), \ldots, \\ \qquad\qquad\qquad\qquad\qquad\qquad \Phi_k(\bar{d}(t), \bar{x}(t+1))) \\ z(t+1) = Z(\bar{d}(t), \bar{x}(t+1)), \end{cases}$$

называется нейронной функцией.

Далее доказывается следующая лемма:

Лемма. 1) Любая нейронная схема реализует некоторую нейронную функцию. 2) Любая нейронная функция реализуется некоторой нейронной схемой.

В параграфе 4.2 приводится доказательство основной теоремы главы

Теорема 7 *Пусть нейронная схема S реализует нейронную функцию f. Тогда из элементов Δ, используя лишь операцию суперпозиции, можно построить схему, однократное применение операции обратной связи к которой приводит к нейронной схеме, реализующей f.*

S — *нейронная схема без обучения*, если задает нейронную функцию f, удовлетворяющую каноническим уравнениям:

$$\begin{cases} \bar{d}(0) = (a_0^1, \ldots, a_0^k) \\ \bar{d}(t+1) = \bar{d}(t) \\ z(t+1) = Z(\bar{d}(t), \bar{x}(t+1)). \end{cases}$$

Пусть схема S задает нейронную функцию $f(\tilde{x}_1, \ldots, \tilde{x}_n, \tilde{y})$

S — *нейронная схема с обучением* тогда и только тогда, когда для всех $p \in \mathbb{N}$ и любых слов $\tilde{\alpha} \in \mathbb{R}^p$, $\tilde{\alpha} = (\alpha_1, \ldots, \alpha_p)$, существует нейронная схема $S'_{\tilde{\alpha}}$ без обучения, задающая функцию

$$f'_{\tilde{\alpha}}(\tilde{x}_1, \ldots, \tilde{x}_n) = f(\tilde{x}_1, \ldots, \tilde{x}_n, \tilde{\alpha}\tilde{0}), \text{ где } \tilde{\alpha}\tilde{0} = (\alpha_1, \ldots, \alpha_p, 0, 0, \ldots) \in \mathbb{R}^\infty.$$

Следствие. Для любой нейронной схемы с обучением можно получить эквивалентную ей схему, которая строится из элементов Δ с помощью операций суперпозиции и не более чем однократного применения операции обратной связи. При этом операция обратной связи, в случае необходимости, применяется последней.

Теорема доказывается конструктивно и позволяет по заданной нейронной функции (или схеме ее реализующей) построить схему вида описанного в теореме (с однократным применением операции обратной связи на последнем шаге).

В параграфе 4.3 к множеству $\Delta(\Delta')$ добавляется элемент "умножение", реализующий функцию $M(x, y) = x \cdot y$, и рассматриваются схемы с умножением с памятью (без памяти). Вводится определение кусочно-полиномиальной функции. Далее доказывается лемма о характере функций схем без памяти, реализуемых над базисом с умножением (все такие функции обозначены через $\mathfrak{P}$):

Лемма.Множество кусочно-полиномиальных функций совпадает с множеством всех функций, реализуемых схемами над базисом $\Delta'_x = \Delta' \cup \{M(x, y)\}$.

Нейронную схему с памятью с умножением (над базисом $\Delta_x = \Delta \cup \{M(x, y)\}$) можно задать системой канонических уравнений:

$$\begin{cases} \bar{d}(0) = (a_0^1, \ldots, a_0^k) \\ \bar{d}(t+1) = \bar{\Phi}(\bar{d}(t), \bar{x}(t+1)) = (\Phi_1(\bar{d}(t), \bar{x}(t+1)), \ldots, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \Phi_k(\bar{d}(t), \bar{x}(t+1))) \\ z(t+1) = Z(\bar{d}(t), \bar{x}(t+1)), \end{cases}$$

где Z, $\Phi_1, \ldots, \Phi_k$ – кусочно-полиномиальные функции.

Лемма. 1) Любая нейронная схема с умножением задает нейронную функцию с умножением.
2) Любая нейронная функция с умножением задается некоторой нейронной схемой с умножением.

Наконец, формулируется теорема, являющаяся обобщением теоремы 7 на схемы над Δ_x.

Теорема 8 *Пусть нейронная схема с умножением S реализует нейронную функцию с умножением f. Тогда из элементов Δ_x, используя лишь операцию суперпозиции, можно построить схему, однократное применение операции обратной связи к которой приводит к нейронной схеме, реализующей функцию f.*

Автор благодарит своего научного руководителя кандидата физико-математических наук, доцента А.А. Часовских за интересную тему, постановку задачи и постоянное внимание к работе. Спасибо профессору О.М. Касим-Заде за ценные обсуждения. Автор выражает глубокую благодарность заведующему кафедрой академику, профессору В.Б. Кудрявцеву и всему коллективу кафедры математической теории интеллектуальных систем за творческую атмосферу и доброжелательность, способствующую работе.

Работы по теме диссертации

[1] Половников В.С. *О некоторых характеристиках нейронных схем.* М., Вестн. Моск. Ун-та. Сер.1, Математика. Механика. 2004. №5, стр. 65-67.

[2] Половников В.С. *О некоторых характеристиках нейронных схем.* М., Интеллектуальные системы том 8, выпуск 1-4, 2004. Стр. 121-145.

[3] Половников В.С. *Критерий нелинейной однослойности нейронных схем.* М., Вестн. Моск. Ун-та. Сер.1, Математика. Механика. 2006. №6, стр. 3-5.

[4] Половников В.С. *О нелинейной сложности нейронных схем Мак-Каллока-Питтса.* М., Интеллектуальные системы том 11, 2007. Стр. 261-276.

Издательство ЦПИ при механико-математическом факультете
МГУ им. М.В. Ломоносова.
Подписано в печать

Формат 60 x 90 1/16 Усл. печ. л.
Тираж 100 экз. Заказ