

**ВОЕННО-КОСМИЧЕСКАЯ АКАДЕМИЯ ИМЕНИ А.Ф. МОЖАЙСКОГО**

Кафедра Математического обеспечения

несекретно

Экз. № 1

УТВЕРЖДАЮ

Начальник 25 кафедры  
полковник \_\_\_\_\_ С.Петренко

« \_\_\_\_ » \_\_\_\_\_ 2007 г.

Автор: профессор 25 кафедры  
доктор технических наук  
профессор А. Хомоненко

Тема: Заключительная

по дисциплине: Системы искусственного интеллекта

Обсуждено и одобрено на заседании 25 кафедры

« \_\_\_\_ » \_\_\_\_\_ 2007 г.  
протокол № \_\_\_\_

Санкт-Петербург  
2007

## Содержание занятия и время

Введение.....	7 мин.
Учебные вопросы (основная часть)	
1. Устройство нейронной сети.....	30 мин.
2. Персептрон.....	15 мин.
3. Выбор структуры нейросети.....	35 мин.
Заключение.....	3 мин.

### Литература:

#### Основная:

1. Борисов В.В., Круглов В.В., Федулов А.С. Нечеткие модели и сети. - М.: Горячая линия-Телеком, 2007. – 284 с.
2. Рассел С, Норвиг П. Искусственный интеллект: современный подход, 2-е изд.: Пер. с англ.. – М.: Изд. дом «Вильямс», 2006. – 1408 с.
3. Конспект лекции.

#### Дополнительная:

1. \_\_\_\_\_  
(наименование издания, страницы)
2. \_\_\_\_\_  
(наименование издания, страницы)
3. \_\_\_\_\_  
(наименование издания, страницы)

#### Материально-техническое обеспечение:

1. Наглядные пособия (по данным учета кафедры): -
2. Технические средства обучения: проектор
3. Приложения (диафильмы, слайды): презентация «Заключительная лекция»

Организационно-методические указания: Во введении сформулировать тему лекции, цель и название изучаемых вопросов. Задать вопросы обучаемым по материалам предыдущих лекции:

1. Назовите основные формы представления знаний?
2. Перечислите проблемы, связанные с представлением знаний.

При изложении материалов обратить внимание обучаемых на соотношение продукционных систем и нейросетей.

Привести примеры использования нейросетей в прикладных экспертных системах. Отметить результаты ученых академии в развитии теории нейронных сетей.

В заключительной части обобщить изложенный материал и сформулировать задание на самостоятельную подготовку.

**Цель лекции:** Изложить общую характеристику нейронных сетей.

**Введение**

**Учебные вопросы:**

### 1. Устройство нейронных сетей

Несмотря на существенные различия, отдельные типы НС обладают несколькими общими чертами.

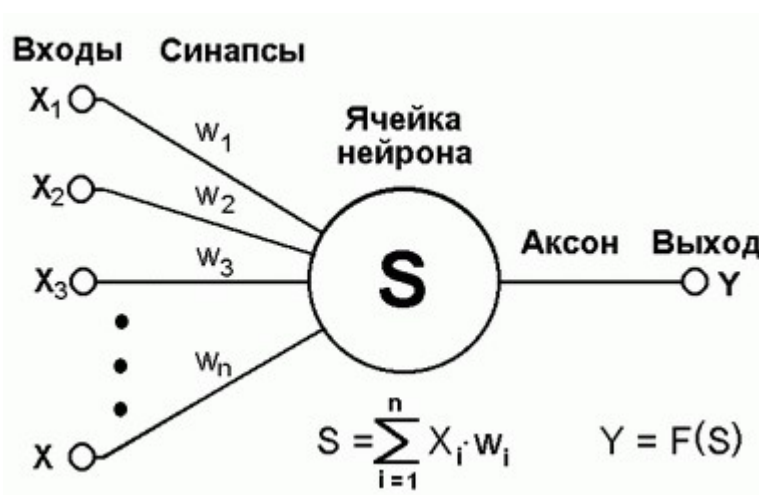


Рис. 6.1 - Искусственный нейрон

Во-первых, основу каждой НС составляют относительно простые, в большинстве случаев – однотипные, элементы (ячейки), имитирующие работу нейронов мозга. Далее под нейроном будет подразумеваться искусственный нейрон, то есть ячейка НС. Каждый нейрон характеризуется своим текущим состоянием по аналогии с нервными клетками головного мозга, которые могут быть возбуждены или заторможены. Он обладает группой синапсов – однонаправленных входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов. Общий вид нейрона приведен на рисунке 1. Каждый синапс характеризуется величиной синаптической связи или ее весом  $w_i$ , который по физическому смыслу эквивалентен электрической проводимости.

Текущее состояние нейрона определяется, как взвешенная сумма его входов:

$$s = \sum_{i=1}^n x_i \cdot w_i \quad (6.1)$$

Выход нейрона есть функция его состояния:

$$y = f(s). \quad (6.2)$$

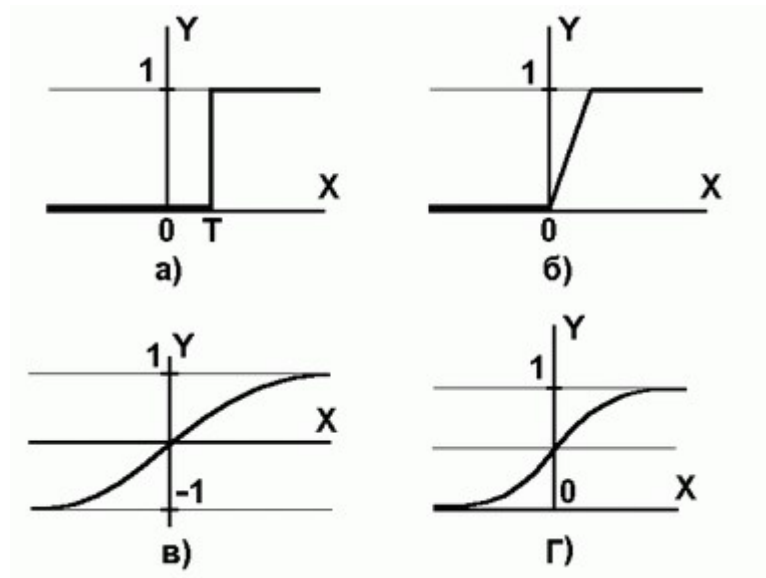


Рис. 6.2 - а) функция единичного скачка;  
 б) линейный порог (гистерезис);  
 в) сигмоид – гиперболический тангенс;  
 г) сигмоид – формула (3)

Нелинейная функция  $f$  называется активационной и может иметь различный вид, как показано на рисунке 6.2. Одной из наиболее распространенных является нелинейная функция с насыщением, так называемая логистическая функция или сигмоид (т.е. функция S-образного вида)[2]:

$$f(x) = 1 / (1 + e^{-\alpha x}). \quad (6.3)$$

При уменьшении  $\alpha$  сигмоид становится более пологим, в пределе при  $\alpha = 0$  вырождаясь в горизонтальную линию на уровне 0.5, при увеличении  $\alpha$  сигмоид приближается по внешнему виду к функции единичного скачка с порогом  $T$  в точке  $x=0$ . Из выражения для сигмоида очевидно, что выходное значение нейрона лежит в диапазоне  $[0, 1]$ . Одно из ценных свойств сигмоидной функции – простое выражение для ее производной, применение которого будет рассмотрено в дальнейшем.

$$f'(x) = \alpha * f(x) * (1 - f(x)). \quad (6.4)$$

Следует отметить, что сигмоидная функция дифференцируема на всей оси абсцисс, что используется в некоторых алгоритмах обучения. Кроме того она обладает свойством усиливать слабые сигналы лучше, чем большие, и предотвращает насыщение от больших сигналов, так как они соответствуют областям аргументов, где сигмоид имеет пологий наклон.

Возвращаясь к общим чертам, присущим всем НС, отметим, во-вторых, принцип параллельной обработки сигналов, который достигается путем объединения большого числа нейронов в так называемые слои и соединения определенным образом нейронов различных слоев, а также, в некоторых конфигурациях, и нейронов одного слоя между собой, причем обработка взаимодействия всех нейронов ведется послойно.

## 2. Персептрон

В качестве примера простейшей НС рассмотрим трехнейронный персептрон (рис.6.3), то есть такую сеть, нейроны которой имеют активационную функцию в виде единичного скачка. На  $n$  входов поступают сигналы, проходящие по синапсам на 3 нейрона, образующие единственный слой этой НС и выдающие три выходных сигнала:

$$y_j = f \left[ \sum_{i=1}^n x_i \cdot w_{ij} \right], j=1...3. \quad (6.5)$$

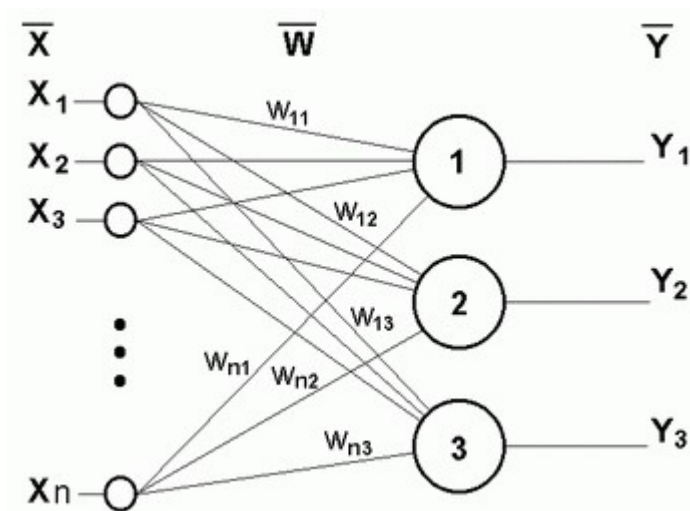


Рис. 6.3 - Однослойный персептрон

Все весовые коэффициенты синапсов одного слоя нейронов можно свести в матрицу  $W$ , в которой каждый элемент  $w_{ij}$  задает величину  $i$ -ой синаптической связи  $j$ -ого нейрона. Таким образом, процесс, происходящий в НС, может быть записан в матричной форме:

$$Y = F(XW). \quad (6.6)$$

где  $X$  и  $Y$  – соответственно входной и выходной сигнальные векторы,  $F(V)$  – активационная функция, применяемая поэлементно к компонентам вектора  $V$ .

Теоретически число слоев и число нейронов в каждом слое может быть произвольным, однако фактически оно ограничено ресурсами компьютера или специализированной микросхемы, на которых обычно реализуется НС. Чем сложнее НС, тем масштабнее задачи, подвластные ей.

### 3. Выбор структуры нейросети

Выбор структуры НС осуществляется в соответствии с особенностями и сложностью задачи. Для решения некоторых отдельных типов задач уже существуют оптимальные, на сегодняшний день, конфигурации. Если же задача не может быть сведена ни к одному из известных типов, разработчику приходится решать сложную проблему синтеза новой конфигурации. При этом он руководствуется несколькими основополагающими принципами: возможности сети возрастают с увеличением числа ячеек сети, плотности связей между ними и числом выделенных слоев; введение обратных связей наряду с увеличением возможностей сети поднимает вопрос о динамической устойчивости сети; сложность алгоритмов функционирования сети (в том числе, например, введение нескольких типов синапсов – возбуждающих, тормозящих и др.) также способствует усилению мощи НС. Вопрос о необходимых и достаточных свойствах сети для решения того или иного рода задач представляет собой целое направление нейрокомпьютерной науки. Так как проблема синтеза НС сильно зависит от решаемой задачи, дать общие подробные рекомендации затруднительно. В большинстве случаев оптимальный вариант получается на основе интуитивного подбора.

Очевидно, что процесс функционирования НС, то есть сущность действий, которые она способна выполнять, зависит от величин синаптических связей,

поэтому, задавшись определенной структурой НС, отвечающей какой-либо задаче, разработчик сети должен найти оптимальные значения всех переменных весовых коэффициентов (некоторые синаптические связи могут быть постоянными). Этот этап называется обучением НС, и от того, насколько качественно он будет выполнен, зависит способность сети решать поставленные перед ней проблемы во время эксплуатации. На этапе обучения кроме параметра качества подбора весов важную роль играет время обучения. Как правило, эти два параметра связаны обратной зависимостью и их приходится выбирать на основе компромисса. Развивая дальше вопрос о возможной классификации НС, важно отметить существование бинарных и аналоговых сетей. Первые из них оперируют с двоичными сигналами, и выход каждого нейрона может принимать только два значения: логический ноль ("заторможенное" состояние) и логическая единица ("возбужденное" состояние). К этому классу сетей относится и рассмотренный выше перцептрон, так как выходы его нейронов, формируемые функцией единичного скачка, равны либо 0, либо 1. В аналоговых сетях выходные значения нейронов способны принимать непрерывные значения, что могло бы иметь место после замены активационной функции нейронов перцептрона на сигмоид.

Еще одна классификация делит НС на синхронные и асинхронные. В первом случае в каждый момент времени свое состояние меняет лишь один нейрон. Во втором – состояние меняется сразу у целой группы нейронов, как правило, у всего слоя. Алгоритмически ход времени в НС задается итерационным выполнением однотипных действий над нейронами. Далее будут рассматриваться только синхронные НС.

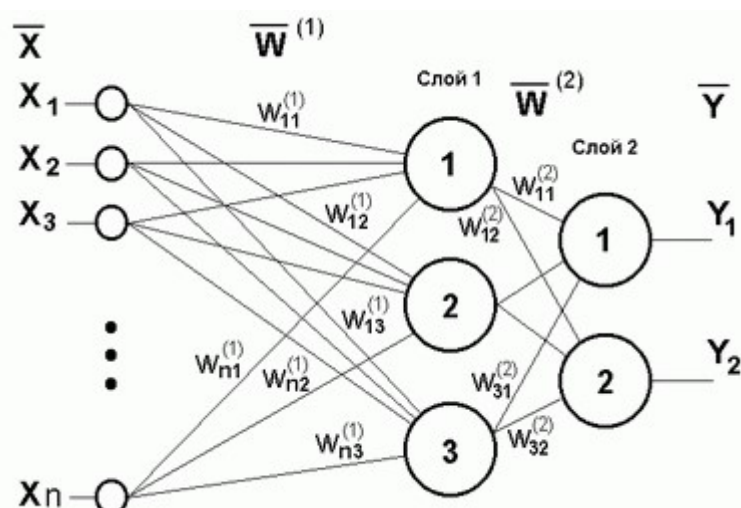




Рис. 6.4 - Двухслойный перцептрон

Сети также можно классифицировать по числу слоев. На рисунке 6.4 представлен двухслойный перцептрон, полученный из перцептрона с рисунка 6.3 путем добавления второго слоя, состоящего из двух нейронов. Здесь уместно отметить важную роль нелинейности активационной функции, так как, если бы она не обладала данным свойством или не входила в алгоритм работы каждого нейрона, результат функционирования любой  $p$ -слойной НС с весовыми матрицами  $W^{(i)}$ ,  $i=1,2,...,p$  для каждого слоя  $i$  сводился бы к перемножению входного вектора сигналов  $X$  на матрицу

$$W^{(\Sigma)} = W^{(1)} * W^{(2)} * ... * W^{(p)}. \quad (6.7)$$

то есть фактически такая  $p$ -слойная НС эквивалентна однослойной НС с весовой матрицей единственного слоя  $W^{(\Sigma)}$ :

$$Y = XW^{(\Sigma)}. \quad (6.7)$$

Продолжая разговор о нелинейности, можно отметить, что она иногда вводится и в синаптические связи. Большинство известных на сегодняшний день НС используют для нахождения взвешенной суммы входов нейрона формулу (6.1), однако в некоторых приложениях НС полезно ввести другую запись, например:

$$s = \sum_{i=1}^n x_i^2 \cdot w_i. \quad (6.9)$$

или даже

$$s = \sum_{i=1}^n x_i \cdot x_{((i+1) \bmod n)} \cdot w_i. \quad (6.10)$$

Вопрос в том, чтобы разработчик НС четко понимал, для чего он это делает, какими ценными свойствами он тем самым дополнительно наделяет нейрон, и каких лишает. Введение такого рода нелинейности, вообще говоря, увеличивает вычислительную мощь сети, то есть позволяет из меньшего числа нейронов с "нелинейными" синапсами сконструировать НС, выполняющую работу обычной НС с большим числом стандартных нейронов и более сложной конфигурации.

Основными видами нейронных сетей для решения практических задач являются:

- многослойные персептроны (сети с прямой передачей сигнала);
- сети на радиальных базисных функциях;
- самоорганизующиеся карты Кохонена;
- вероятностные (байесовские) нейронные сети;
- обобщенно-регрессионные нейронные сети;
- сети главных компонент;
- сети для кластеризации;
- линейные сети.

### **Заключение:**

Обратить внимание обучаемых на актуальность вопросов, связанных с использованием нейронных сетей.

На самостоятельной подготовке прочитать материалы из рекомендуемой литературы.

ДОКТОР ТЕХНИЧЕСКИХ НАУК ПРОФЕССОР

\_\_\_\_\_ А. ХОМОНЕНКО

«\_\_\_\_\_» \_\_\_\_\_ 200\_\_ г.