Л е к ц и я № 15

Искусственные нейронные сети прямого распространения

Разработал профессор П. М. Васильев
Кафедра фармакологии и биоинформатики

Для студентов, обучающихся по направлению 06.03.01 «Биология»
профили Биохимия, Генетика
при изучении дисциплины «Цифровые технологии в биологии»

План лекции

- **Нейронные сети прямого распространения**
- **Теорема Колмогорова**
- **Многослойный перцептрон**
- **Перцептрон Румельхарта**
- **Функции активации**
- **Обучение сети прямого распространения**
- **Примеры использования**

Нейронные сети прямого распространения

Feed-forward neural networks, FFNN

Нейронная сеть, в которой все сигналы от входных и скрытых нейронов передаются в одном направлении к выходным нейронам.

Полносвязные нейронные сети прямого распространения

Нейронная сеть прямого распространения, в которой все сигналы от нейронов предыдущего слоя передаются всем нейронам последующего слоя.

Теорема Колмогорова

$$f(\mathbf{x}) = f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right)$$

Доклады Академии наук СССР
1957. Том 114, № 5

МАТЕМАТИКА

Академик А. Н. КОЛМОГОРОВ

О ПРЕДСТАВЛЕНИИ НЕПРЕРЫВНЫХ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ В ВИДЕ СУПЕРПОЗИЦИЙ НЕПРЕРЫВНЫХ ФУНКЦИЙ ОДНОГО ПЕРЕМЕННОГО И СЛОЖЕНИЯ

Целью заметки является краткое изложение доказательства следующей теоремы:

Теорема. *При любом целом $n \geqslant 2$ существуют такие определенные на единичном отрезке $E^1 = [0; 1]$ непрерывные действительные функции $\psi^{pq}(x)$, что каждая определенная на n-мерном единичном кубе E^n непрерывная действительная функция $f(x_1, \ldots, x_n)$ представима в виде*

$$f(x_1, \ldots, x_n) = \sum_{q=1}^{q=2n+1} \chi_q \left[\sum_{p=1}^{n} \psi^{pq}(x_p) \right], \tag{1}$$

где функции $\chi_q(y)$ действительны и непрерывны.

При $n = 3$, положив

$$\varphi_q(x_1, x_2) = \psi^{1q}(x_1) + \psi^{2q}(x_2), \quad h_q(y, x_3) = \chi_q[y + \psi^{3q}(x_3)],$$

получаем из (1)

$$f(x_1, x_2, x_3) = \sum_{q=1}^{7} h_q[\varphi_q(x_1, x_2,), x_3], \tag{2}$$

что является небольшим усилением результата В. И. Арнольда ([2]), который показал, что любая непрерывная функция трех переменных представима в виде суммы девяти слагаемых того же вида, как слагаемые, входящие в формулу (2) в числе семи. Результаты моей заметки ([1]) не вытекают из сообщаемой сейчас новой теоремы в их точных формулировках, но принципиальное их содержание (в смысле возможности представления функций нескольких переменных суперпозициями функций меньшего числа переменных и их приближения суперпозициями фиксированного вида из многочленов от одного переменного и сложения) очевидным образом содержится в новой теореме. Метод доказательства новой теоремы элементарнее методов работ ([1,2]), сводясь к прямым конструкциям и подсчетам. Исчезла, в частности, необходимость употребления деревьев из компонент линий уровня. Фактически, однако, конструкции, употребленные в этой заметке, были найдены путем анализа конструкций, употреблявшихся в ([1,2]), и отбрасывания в них деталей, излишних для получения конечного результата.

§ 1. Построение функций ψ^{pq}. Индексы p, q, k, i всюду далее пробегают целые значения

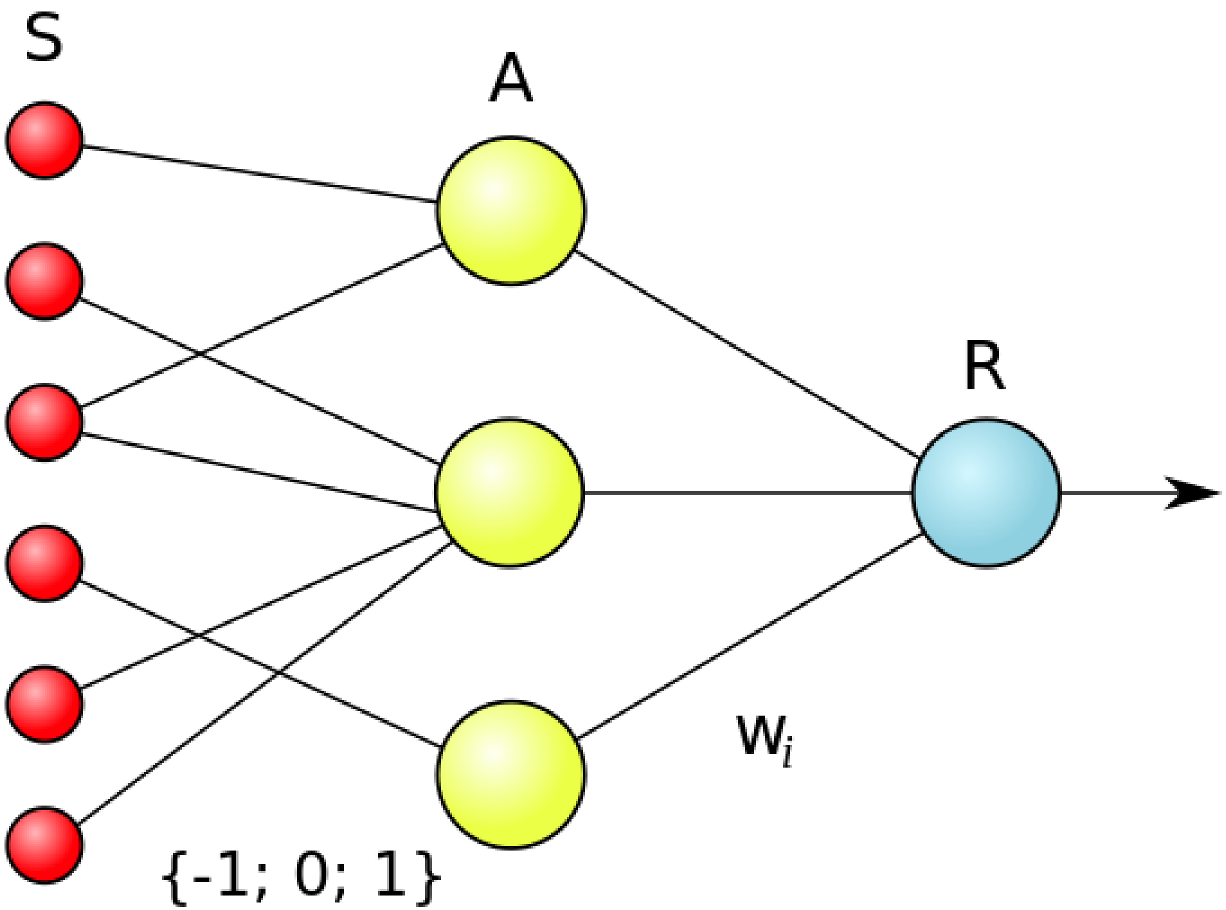
$$1 \leqslant p \leqslant n, \quad 1 \leqslant q \leqslant 2n+1, \quad k = 1, 2, \ldots, \quad 1 \leqslant i \leqslant m_k = (9n)^k + 1.$$

При суммировании и перемножении в этих пределах пределы не обозначаются.

Многослойный перцептрон

Фрэнк Розенблатт, 1958

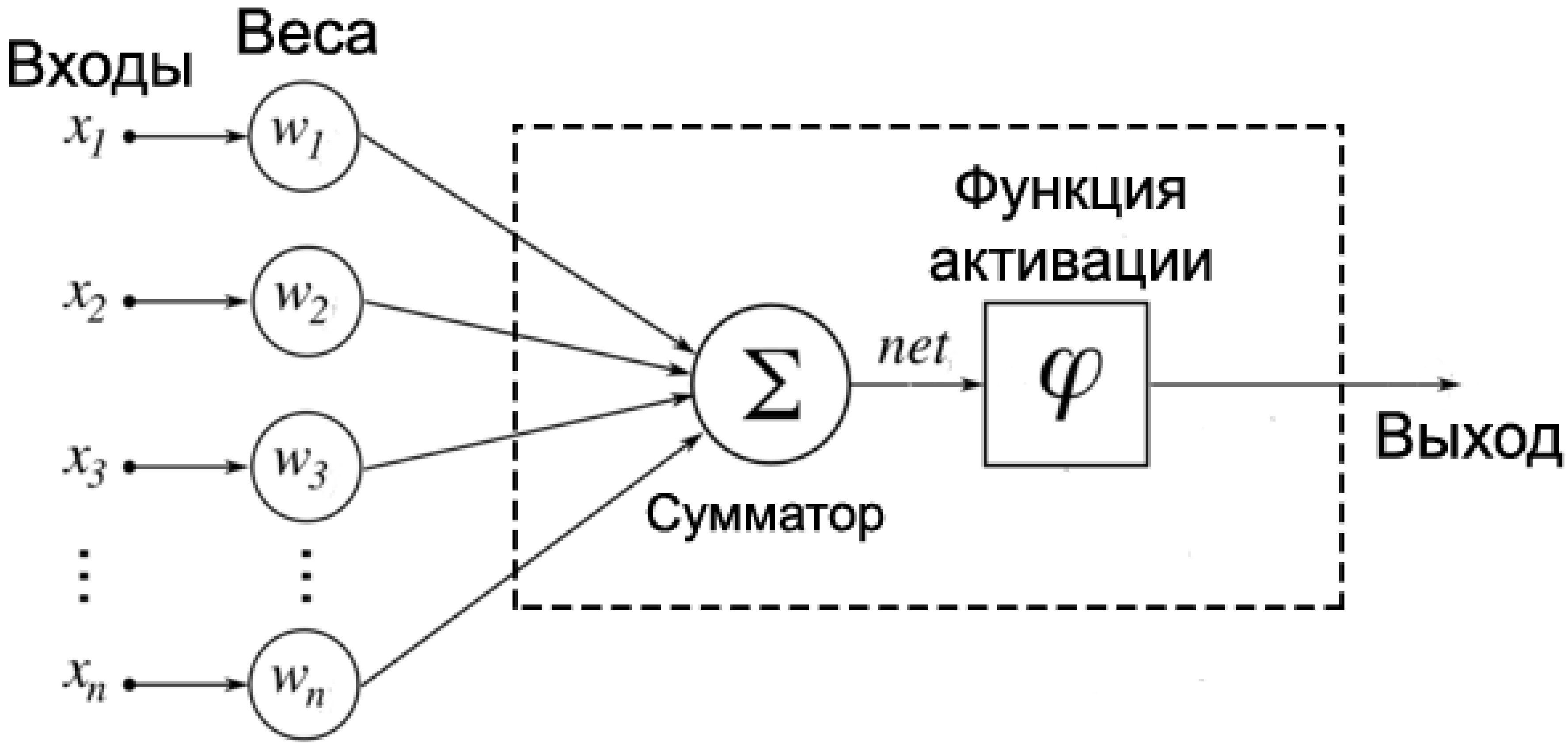
Дэвид Румельхарт, 1986



Перцептрон Румельхарта

- Используется нелинейная функция активации
- Число обучаемых слоев больше одного
- Сигналы, поступающие на вход и получаемые с выхода, не бинарные
- Допускается произвольная архитектура связей (в том числе, и полносвязные сети)
- Ошибка сети вычисляется как некоторая статистическая мера невязки между нужным и получаемым значением
- Обучение проводится до стабилизации весовых коэффициентов

Искусственный нейрон



Уровень активации нейрона

Представляет собой простую линейную функцию входов и определяется как взвешенная сумма входов нейрона плюс пороговое значение:

$$x = bias + weight_1 \cdot input_1 + weight_2 \cdot input_2 + \ldots$$

Эта активация затем преобразуется с помощью передаточной функции (функции активации).

Функции активации (передаточные функции)

$$f(x) = \frac{1}{1+e^{-x}}, \quad 0 \ldots 1$$

Логистическая (сигмоидная)

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}, \quad -1 \ldots +1$$

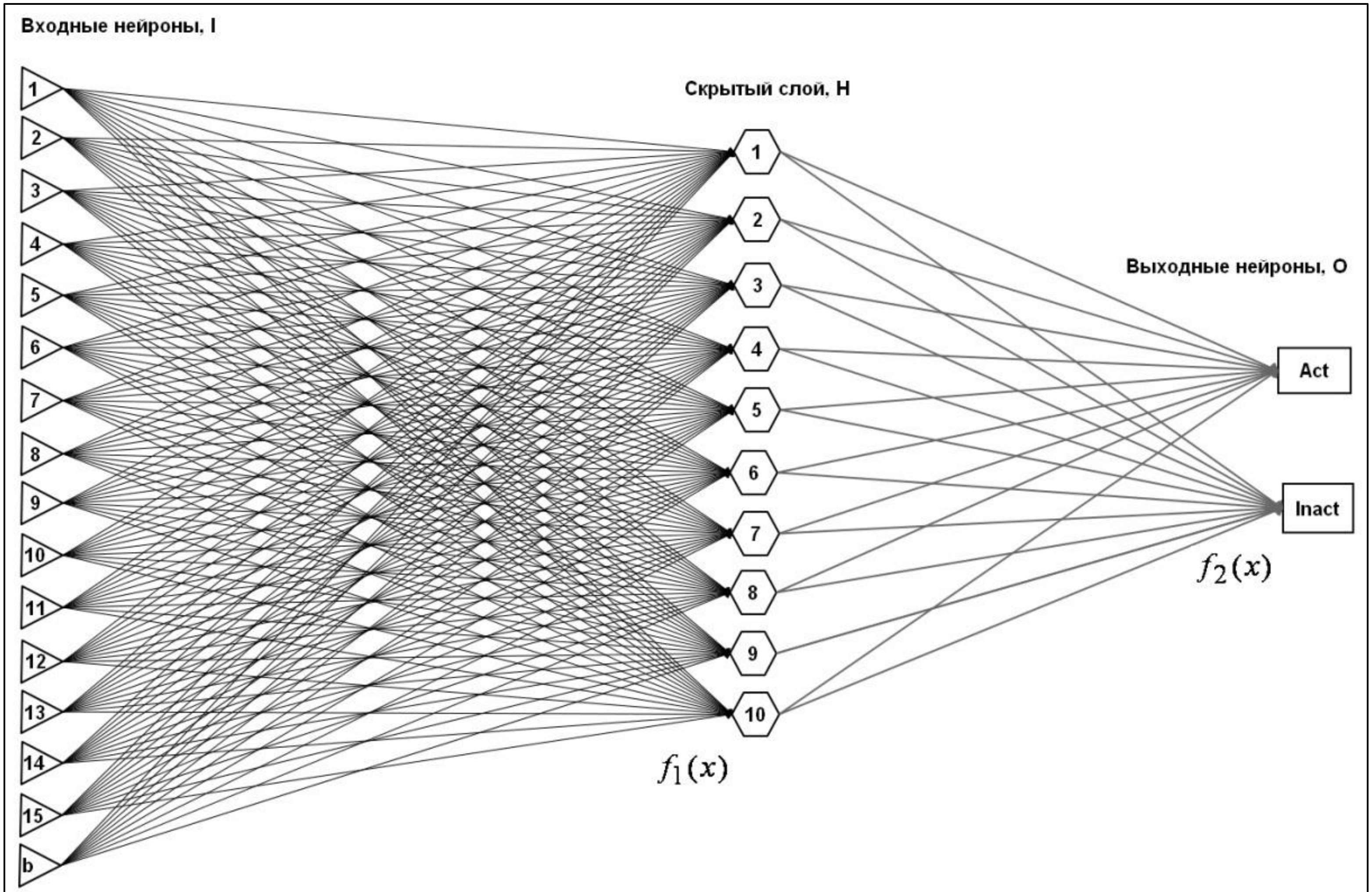
Гиперболический тангенс

$$f(x) = e^{-x}, \quad 0 \ldots +\infty$$

Обратная экспонента

Многослойный перцептрон

Multi-Layer Perceptron, MLP




To be continued ...