

*В. А. Якубович*

## **МАШИНЫ, ОБУЧАЮЩИЕСЯ РАСПОЗНАВАНИЮ ОБРАЗОВ. I**

### **ВВЕДЕНИЕ**

Будем рассматривать машины, моделирующие в определенном смысле работу каких-либо органов чувств человека и, что особенно важно, работу той части коры головного мозга, которая руководит организацией впечатлений и вырабатывает простейшие понятия, связанные с процессом распознавания образов.

Принципам построения подобных машин уделялось много внимания в литературе последних лет. Позволим себе привести выдержку из речи академика В. А. Трапезникова [1], так как вряд ли лучше поясним постановку проблемы:

«Сейчас мы находимся на пороге создания удивительных машин для изучения управляемых объектов, машин, в которые заранее не изложены алгоритм или расчетная формула для обработки информации. Имеются в виду обучаемые и самообучающиеся устройства, которые в дальнейшем смогут определить состояние системы по ряду признаков, которые человек либо вовсе не способен оценить, либо оценивает лишь подсознательно. Это позволит вести процесс управления при неполной информации, упрощая тем самым ее сбор.

Один из наиболее интересных процессов обучения — это процесс обучения распознаванию зрительных образов. Образом называется восприятие, которое человек относит к определенной группе, умеет распознать его, отнести именно к этой группе, хотя может быть никогда раньше данный представитель группы ему не встречался. Так, например, буква «А» есть образ, потому что начертания этой буквы, написанные разными почерками, воспринимаются как буква «А». Другими примерами образов являются понятия: «мужской портрет», «цифра 1» и т. д.

Имитация процесса обучения является одной из важных проблем кибернетики, если иметь в виду дальнейшее приближение функций машин к функциям живого мозга.

В связи с этим за последние годы в мировой литературе появилось много работ, касающихся устройств для распознавания образов, так называемых «персентронов».

Как же происходит процесс обучения? Можно представить себе, что учитель, который уже умеет различать эти образы, продумывает систему признаков — своеобразную программу, при помощи которой можно распознать буквы или другие образы, и передает эту программу ученику. Такой подход имитируется во многих

разрабатываемых сейчас машинах для чтения печатных или письменных букв. В подобном процессе обучения дело сводится к выработке признаков для включения их в хорошо продуманную программу.

Существует, однако, другой подход к процессу обучения, значительно более интересный с точки зрения его имитации в машине. Учитель, не пытаясь выработать признаки и объяснять, как надо различать буквы, показывает ученикам по нескольку представителей букв и говорит, какие это буквы. Очень скоро ученики начинают уверенно отличать буквы, хотя признаки букв или программы для их распознавания ученикам не были переданы»<sup>1</sup>.

Автору известны статьи [1–7] по теме настоящей работы, от которых данная статья отличается большим уклоном в сторону математики, а также математическая работа [8]. Читатель обнаружит с самого начала некоторые различия в точках зрения.

Имея в виду читателя-нематематика, мы подробно проводили доказательства некоторых элементарных предложений.

## § 1. ПРИНЦИПИАЛЬНАЯ СХЕМА ПЕРСЕПТРОНА

Пусть  $X$  — некоторое множество, элементы которого будем называть образами; пусть  $S_1, S_2, \dots, S_k$  — какие-то его непересекающиеся подмножества (элементарные понятия), которым сопоставлены символы  $\xi_1, \xi_2, \dots, \xi_k$  («названия» понятий) из некоторого алфавита. Например, все возможные изображения на сетчатке глаза (мы отвлекаемся от различия изображений в двух глазах) образуют множество зрительных образов, а всевозможные изображения, возникающие, когда мы видим букву А, образуют элементарное понятие «изображение буквы А».

Будем называть персептроном машину, функционирующую следующим образом:

- 1) на вход машины подаются образы — элементы множества  $X$ ,
- 2) машина имеет конечное число  $k$  выходов («ответов»)  $\xi_j$ ,
- 3) учитель, которым может быть человек, природа или другая машина, подает на вход некоторую конечную последовательность («тренировочную последовательность») образов  $x_j$  ( $j = 1, 2, \dots, m$ ) из подмножества  $S_i$  ( $i = 1, 2, \dots, k$ ) и соответствующие ответы  $\xi_i \sim S_i \ni x_j$ ; тем самым учитель сообщает машине, какому подмножеству  $S_i$  принадлежит показанный элемент. Этот процесс будем называть процессом обучения машины<sup>2</sup>,
- 4) после окончания обучения машина должна на каждый вход  $x \in S_i$  выдавать ответ  $\xi_i$  или во всяком случае ошибаться достаточно редко,
- 5) подмножества  $S_i$  никак не связаны с конструкцией машины, они могут быть произвольными из некоторой достаточно широкой системы подмножеств.

<sup>1</sup> Отметим, что цитированная речь так же, как работа [2], стала известна автору, когда настоящая статья оформлялась к печати. Мы пользуемся в дальнейшем иной, может быть не очень удачной терминологией. Мы называем «образом» вход в машину (фиксированное изображение на сетчатке глаза и т. п.). Множество образов, объединенных некоторыми признаками («образ» в смысле акад. В. А. Трапезникова), мы называем элементарным понятием.

<sup>2</sup> Сопоставление  $S_i \sim \xi_i$  считается либо осуществленным заранее, либо его осуществляет учитель. Описанный способ обучения называется обучением с поощрением. Относительно обучения без поощрения, когда нет учителя и, следовательно, в машину не поступает информация о том, каков должен быть ответ, смотри [3], [4]. По-видимому, обучение без поощрения основывается на том, что показываемые образы в множестве  $X$ , которое естественно считать топологическим пространством, группируются в некоторые «кучи»  $S_i$ . Ниже рассматривается лишь обучение с поощрением.

Уточняя сказанное, будем считать данными:

- 1) множество образов  $X = \{x\}$ ,
- 2) некоторую систему  $\Sigma = \{S\}$  его подмножеств (элементарные понятия),
- 3) множество  $U = \{u\}$ , элементы которого будем называть параметры, и
- 4) конечное множество  $Z = \{\xi_1, \xi_2, \dots, \xi_k\}$  — алфавит «выходов» или «ответов».

С математической точки зрения персептрон описывается функцией  $\xi = f(x, u)$ , осуществляющей отображение множества  $X \times U$  на множество  $Z$ , и некоторым определенным ниже «алгоритмом обучения».

С технической точки зрения персептрон — машина, которая может работать в двух режимах — в режиме распознавания и в режиме обучения и состоит из двух блоков: блока распознавания (БР) и блока обучения (БО).

В режиме распознавания работает лишь блок распознавания; параметр  $u$  фиксирован, на вход персептрона подается образ  $x \in X$ , выходом является ответ  $\xi = f(x, u)$ .

В режиме обучения на вход персептрона подается тренировочная последовательность образов из  $\cup_{i=1}^m S_i$ . Эта последовательность подается на входы БО и БР (рис. 1). Кроме того, на вход БО подается информация о состоянии БР после подачи каждого из образов тренировочной последовательности и поступающая от «учителя» информация о том, какому из множеств  $S_i$  принадлежит каждый образ тренировочной последовательности. Эта информация перерабатывается в БО, после чего БО выдает значения параметра  $u$ , поступающие в БР.

Процесс повторяется некоторое число  $N$  раз, после чего обучение считается законченным. (Возможно, что  $N = 1$ ; тренировочная последовательность при повторениях может меняться.)

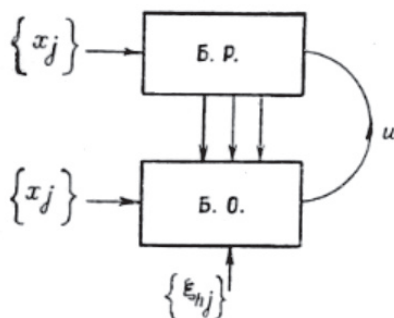


Рис. 1.

Алгоритм определения нового значения параметра  $u$  по старому значению и по последовательности  $\{x_1, \xi_{i_1}\}, \{x_2, \xi_{i_2}\}, \dots, \{x_m, \xi_{i_n}\}$ , в которой  $x_k \in S_k \sim \xi_{i_k}$  ( $k = 1, 2, \dots, n$ ), будем называть алгоритмом обучения.

Образы тренировочной последовательности выбираются из множеств  $S_i$  случайным образом согласно некоторому распределению вероятностей.

Схема персептрона должна быть таковой, что вероятность каждого из  $m$  равенств

$$f(x, u) = \xi_i \text{ при } x \in S_i \quad (i = 1, 2, \dots, m)$$

сколь угодно близка к единице, если параметр  $u$  выбран применением алгоритма обучения по достаточно длинной тренировочной последовательности и, возможно, при достаточно большом числе  $N$  повторений алгоритма обучения.

## § 2. СХЕМА БЛОКА РАСПОЗНАВАНИЯ. ЛИНЕЙНЫЙ ПЕРСЕПТРОН

Будем считать, что БР состоит из четырех частей: 1) ретины, 2) ассоциативной части, 3) суммирующей части и 4) логической части.

В обычных конструкциях [4–7] ретиной называется конечный набор элементов («рецепторные нейроны»), каждый из которых может находиться обычно в двух состояниях (состояние возбуждения и состояние торможения) или в нескольких состояниях. Состоянием ретины является набор состояний каждого рецепторного нейрона. Образ поступает на ретину и переводит ее в определенное состояние. Будем ниже отождествлять образы  $x \in X$  с состояниями ретины.

Вторая ступень персеプトрона — набор  $N_a$  ассоциативных элементов (а. э.). Каждый а. э. связан некоторым образом с ретиной. С математической точки зрения важно лишь, что каждому а. э. соответствует некоторая вещественная функция  $\phi_j(x)$  на  $X$ , называемая реакцией  $j$ -го а. э. на образ  $X$ .

Третья ступень персепторона — набор  $N_s$  суммирующих элементов (с. э.). Каждый с. э. связан, вообще говоря, с каждым а. э. Суммирующему элементу с номером  $h$  соответствует система  $N_a + 1$  чисел («весов»)  $\alpha_0^{(h)}, \alpha_j^{(h)}$  ( $j = 1, 2, \dots, N_a$ ). В ответ на образ  $x$  с. э. выдает значение

$$\psi_h(x) = \alpha_0^{(h)} + \sum_{j=1}^{N_a} \alpha_j^{(h)} \phi_j(x),$$

называемое реакцией с. э. на образ  $x$ . Набор весов  $\{\alpha_j^{(h)}\}$ ,  $h = 1, 2, \dots, N_s$ ,  $j = 0, 1, \dots, N_a$ , задается блоком обучения.

Сопоставим каждому с. э. логическую переменную  $\pi_j$ :

$$\pi_j = \begin{cases} 0, & \text{если } \psi_j < 0, \\ 1, & \text{если } \psi_j \geq 0. \end{cases} \quad (2.1)$$

Последняя, логическая ступень БР персепторона реализует  $k$  логических функций  $\xi_i = \xi_i(\pi_1, \pi_2, \dots, \pi_{N_s})$  ( $i = 1, 2, \dots, k$ ), где число  $t$  равно числу выходов БР. Эти логические функции не заданы заранее: они формируются блоком обучения.

Так как значения  $\pi_j$  определяются образом  $x$ , то и  $\psi_i$  является функцией  $x$ . Считаем, что на образ  $x$  персептрон выдает ответ  $\xi_i$ , если значение соответствующей функции  $\xi_i = 1$ . Будем говорить, что персептрон разделяет множества  $S_i$ , если 1) каждому  $S_i$  сопоставлена переменная  $x_i$  и 2) из  $x \in S_h$  следует, что

$$\xi_h(x) = 1, \quad \xi_i(x) = 0 \text{ при } i \neq h. \quad (2.2)$$

Однозначность ответа означает, что функции  $\xi_i$  должны удовлетворять условию:

$$\text{если } \xi_h = 1, \text{ то } \xi_i = 0 \text{ при } i \neq h.$$

Однако, удобно расширяя определение § 1, можно отказаться от условия (2.2) и интерпретировать ответы следующим образом:

$$\xi_i = 0 \text{ означает, что } x \notin S_i.$$

Тогда, например, ответ

$$\xi_1 = 1, \quad \xi_2 = 0, \quad \xi_3 = 1, \quad \xi_4 = 1, \quad \dots, \quad \xi_k = 0$$

будет означать, что  $x \in S_1 \cup S_3 \cup S_4$ . Поскольку множества  $S_1$  по условию не пересекаются, подобные ответы свидетельствуют, что перцептрон недостаточно хорошо «обучен».

Из изложенного следует, что параметрами перцептрона, которые может менять блок обучения, являются веса а. э. и параметры, задающие логическую функцию  $\xi_i$ .

Для дальнейшего важную роль играет понятие линейного перцептрона.

**Определение 1.** Перцептрон называется *линейным*, если

1) множество образов — евклидово или вещественное гильбертово пространство,

2) реакции ассоциативных элементов на образ — линейные функционалы.

Условие 1 не является серьезным ограничением. Зрительные образы можно трактовать как функции двух переменных, звуки — как функции одной переменной. В обоих случаях это элементы подходящего гильбертова пространства или, имея в виду ограниченную точность, элементы евклидова пространства. (На самом деле при техническом осуществлении это конечное множество, некоторая  $\epsilon$ -сеть ограниченной части евклидова пространства.)

По-видимому, бесспорно, что человек имеет представление о близости воспринимаемых образов. Таким образом, множество образов  $X$  — во всяком случае топологическое пространство. Естественно рассмотреть вначале предположение, что  $X$  — гильбертово или евклидово пространство.

Условие 2 совсем не обязательно. В перцептронах Розенблатта [4, 5], Джозефа [8] а. э. являются пороговые элементы, т. е. условие 2 не выполнено. Автору, однако, неясны преимущества нелинейных перцептронов. Последовательной теории нелинейных перцептронов нет. Мы убедимся сейчас, что линейность делает понятными основные принципы работы перцептрона (см. также [3] с очень близкой трактовкой).

Для линейного перцептрона каждому а. э. соответствует вектор  $a_j \in X$  так, что  $\phi_j(x) = (x, a_j)$ ,  $j = 1, 2, \dots, N_a$ . Каждому с. э. соответствует плоскость  $\psi_h(x) = 0$  в пространстве образов.

Естественно предполагать, что множества  $S_i$  не «патологические», именно что их можно разделить конечной системой плоскостей. Более точно:

**Определение 2.** Непересекающиеся множества  $S_i \subset X$  ( $i = 1, 2, \dots, k$ ) будем называть *допустимыми*, если каждое из  $S_i$  компактно и существует набор плоскостей  $\psi_h(x) = \alpha_0^{(h)} + (a^{(h)}, x) = 0$  ( $h = 1, 2, \dots, N_s$ ) и набор  $k$  логических функций  $\xi_i(\pi_1, \pi_2, \dots, \pi_{N_s})$  ( $i = 1, 2, \dots, m$ ) таких, что из  $x \in S_h$  следует  $\xi_h = 1$ ,  $\xi_i = 0$  ( $i \neq h$ ), причем логические переменные  $\pi_1, \pi_2, \dots, \pi_{N_s}$  определяются формулами (2.1).

Рассмотрим вначале случай, когда  $X$  является  $n$ -мерным евклидовым пространством. Плоскости

$$\psi_h(x) = \alpha_0^{(h)} + \sum_{j=1}^{N_i} \alpha_j^{(h)}(a_j, x) = 0$$

смогут принимать произвольное положение за счет выбора весов  $\alpha_j^{(h)}$ , если выполнено следующее условие:

(А) Число ассоциативных элементов  $N_a \geq n$ , где  $n$  — размерность пространства образов  $X$ ; среди  $N_a$  функций  $\phi_j(x) = (a_j, x)$  найдутся  $n$  линейно независимых.

При выполнении сформулированного условия для всякой допустимой системы множеств  $S_i$  можно, очевидно, подобрать веса а. э. и логические функции так, чтобы персептрон разделял эти множества.

В случае, когда  $X$  — гильбертово пространство, разделение множеств можно получить лишь при достаточно большом  $N_a$  и при условии, что векторы  $a_j \in X$ , соответствующие а. э., образуют в пределе (при  $N_a \rightarrow \infty$ ) полную систему. В практических реализациях  $N_a$  конечно и определяется необходимой точностью восприятия.

Таким образом, для линейного персептрона теорема существования тривиальна и обеспечивается основными предположениями<sup>3</sup>.

### § 3. ПЕРСЕПТРОН С ОДНИМ СУММИРУЮЩИМ ЭЛЕМЕНТОМ

Для разделения двух множеств может служить персептрон с одним с. э. Логическая часть в этом случае отсутствует; два возможных ответа соответствуют двум знакам реакций с. э.:

$$\psi(x) = \alpha_0 + \sum_{j=1}^{N_a} \alpha_j (a_j, x).$$

Очевидно, не всякие два множества могут быть разделены. К числу множеств, которые могут быть разделены персептроном с одним с. э., относятся выпуклые множества.

**Лемма 1.** *Предположим, что  $S_1$  и  $S_2$  — выпуклые, компактные множества в гильбертовом или евклидовом пространстве  $X$ , замыкания которых не пересекаются. Тогда*

1) *существует неоднородный линейный функционал*

$$\psi(x) = \alpha + (a, x)$$

*такой, что*

$$\begin{aligned} \psi(x) &> 0 \text{ при } x \in S_1, \\ \psi(x) &< 0 \text{ при } x \in S_2; \end{aligned} \tag{3.1}$$

<sup>3</sup> Для нелинейного персептрона, например типа «Розенблатта» [4, 5], является совсем неясной даже принципиальная возможность разделения. Теорема Джозефа [8] о разделении для персептрона Розенблатта доказана в предположении, что число  $N_a$  достаточно велико, а образы подчинены некоторому условию «положительности», проверить которое невозможно и которое, по существу, содержит утверждение теоремы. Если рассматривать персептрон Розенблатта как «испорченный» линейный персептрон, то делается понятной принципиальная возможность разделения некоторых простейших, но далеко не всяких типов образов. С этой точки зрения делается понятной также случайность связей а. э. с ретиной в персептроне Розенблатта — эта случайность обеспечивает линейную независимость векторов  $a_j$ . Отметим еще, что в отличие от теоремы Джозефа, которая утверждает, что вероятность разделения стремится к единице при  $N_a \rightarrow \infty$ , для линейного персептрона увеличение  $N_a$  сверх некоторого, определенного типом ретины значения, бесполезно.

2) существует число  $\sigma > 0$  такое, что для любых  $a', \alpha'$  таких, что  $|a - a'| < \sigma, |\alpha - \alpha'| < \sigma$ , соотношения (3.1) будут выполнены также для  $\psi(x) \equiv \psi'(x) = \alpha' + (a', x)$ .

ДОКАЗАТЕЛЬСТВО. Из выпуклости множества  $S = S_j$  ( $j = 1, 2$ ) следует выпуклость замыкания  $\bar{S}$ . Действительно, пусть  $x' \in \bar{S}, x'' \in \bar{S}$ . Найдутся последовательности  $x'_n \in S, x''_n \in S$  такие, что  $x'_n \rightarrow x', x''_n \rightarrow x''$  при  $n \rightarrow \infty$ . Поэтому для любого  $t$  ( $0 \leq t \leq 1$ )

$$x_n = tx'_n + (1-t)x''_n \rightarrow x = tx' + (1-t)x''.$$

Так как  $x_n \in S$ , то  $x \in \bar{S}$ , т. е.  $\bar{S}$  выпукло.

Поскольку  $\bar{S}_1$  и  $\bar{S}_2$  являются компактными, компактом будет и множество  $\bar{S}_1 \times \bar{S}_2$ . Из непрерывности функции

$$\rho(x_1, x_2) = |x_1 - x_2|$$

на  $\bar{S}_1 \times \bar{S}_2$  следует существование

$$x' \in \bar{S}_1, \quad x'' \in \bar{S}_2$$

таких, что

$$\text{Inf}_{x_1 \in \bar{S}_1, x_2 \in \bar{S}_2} |x_1 - x_2| = |x' - x''|. \quad (3.2)$$

Положим<sup>4</sup>

$$\psi(x) = \left( x - \frac{x' + x''}{2}, x' - x'' \right). \quad (3.3)$$

Мы имеем

$$\psi(x') = \frac{1}{2}|x' - x''|^2 > 0, \quad \psi(x'') = -\frac{1}{2}|x' - x''|^2 < 0.$$

Покажем, что

$$\begin{aligned} \psi(x) &> 0 \text{ при } x \in \bar{S}_1, \\ \psi(x) &< 0 \text{ при } x \in \bar{S}_2, \end{aligned} \quad (3.4)$$

откуда будет следовать выполнение (3.1).

Предположим, что нарушается первое неравенство (3.4). Тогда существует образ  $x \in \bar{S}_1$  такой, что  $\psi(x) \leq 0$ . Покажем, что тогда нарушается (3.2).

Из выпуклости  $\bar{S}_1$  следует, что

$$x(t) = tx' + (1-t)x_0 \in \bar{S}_1 \text{ при } 0 \leq t \leq 1.$$

Определим  $t_0$  из условия

$$(x(t_0) - x'', x_0 - x') = 0,$$

т. е.

$$t_0 = \frac{(x_0 - x', x_0 - x'')}{|x_0 - x'|^2}.$$

Если  $t_0 < 0$ , то из соотношения

$$|x' - x''|^2 = |x' - x_0|^2 - 2t_0|x' - x_0|^2 + |x_0 - x''|^2$$

<sup>4</sup> Уравнение  $\psi(x) = 0$  является уравнением гиперплоскости, ортогональной к вектору  $x' - x''$  и проходящей через середину отрезка, соединяющего  $x'$  и  $x''$ .

выводим, что  $|x_0 - x''| < |x' - x''|$ . Так как теперь  $x(t_0) \in \overline{S_1}$ , мы снова приходим к противоречию с (3.2).

Наконец, если  $t_0 > 1$ , то

$$(x_0 - x', x_0 - x'') > (x_0 - x', x_0 - x')$$

или

$$(x_0 - x', x' - x'') > 0,$$

т. е.

$$\psi(x_0) = \psi(x') + (x_0 - x', x' - x'').$$

Последнее предположение невозможно по определению.

Таким образом, нарушение первого неравенства (3.4) ведет к противоречию. Аналогично проверяется выполнение второго неравенства (3.4).

Переходим к доказательству утверждения 2 теоремы. Предполагая противное, получим, что существует  $a_n, \alpha_n$ ,

$$a_n \rightarrow a, \quad \alpha_n \rightarrow \alpha \text{ при } n \rightarrow \infty$$

такие, что для

$$\psi_n(x) = \alpha_n + (a_n, x)$$

нарушается (3.1). По крайней мере одно из неравенств (3.1), например первое, нарушается для бесконечного числа значений  $n$ :  $\psi_n(x_n) \leq 0$  для  $x_n \in S_1$ . Выделяя, в силу компактности, сходящуюся подпоследовательность, получим

$$x_{n_k} \rightarrow x^* \in \overline{S_1}, \quad \psi_{n_k}(x_{n_k}) \rightarrow \psi(x^*).$$

Следовательно,  $\psi(x^*) \leq 0$ , что противоречит (3.4). Тем самым мы доказали справедливость утверждения 2.

**Замечание.** Утверждение 2 несправедливо для некомпактных множеств  $S_1, S_2$ . Чтобы убедиться в этом, достаточно, например, в качестве  $S_1$  взять выпуклое множество, ограниченное на плоскости  $\xi, \eta$  ветвь гиперболы  $\xi\eta = 1$ ,  $\xi > 0$  и в качестве  $S_2$  — область, симметричную  $S_1$  относительно оси  $\xi$ .

**Определение 2.** Будем называть выпуклым телом в гильбертовом или евклидовом пространстве выпуклое множество, содержащее хотя бы одну внутреннюю точку.

**Теорема 1.** Пусть  $S_1, S_2$  — непересекающиеся компактные выпуклые тела в гильбертовом или евклидовом пространстве  $X$ .

Предположим, что из множества  $S_1$  выбираются случайным и независимым образом элементы  $x'_1, x'_2, \dots, x'_{m_1}$  и из множества  $S_2$  — элементы  $x''_1, x''_2, \dots, x''_{m_2}$ . Предположим, что какова бы ни была внутренняя точка  $x' \in S_1$  и какова бы ни была ее окрестность, целиком содержащаяся в  $S_1$ , вероятность того, что точка  $x'_j$  будет выбрана в этой окрестности, не зависит от  $j$  и  $> 0$ . Предположим также, что аналогичное утверждение справедливо для множества  $S_2$ .



Пусть

$$\min_{j,h} |x'_j - x''_h| = |x'_r - x''_s| \quad (3.5)$$

и  $P_{m_1, m_2}$  означает вероятность выполнения соотношения (3.1) для

$$\psi(x) = \left( x - \frac{x'_r + x''_s}{2}, x'_r - x''_s \right). \quad (3.6)$$

Тогда  $P_{m_1, m_2} \rightarrow 1$  при  $\min(m_1, m_2) \rightarrow +\infty$ .

Эта теорема показывает, что персептрон с одним с. э. будет разделять множества  $S_1, S_2$ , если: 1)  $S_1$  и  $S_2$  являются выпуклыми телами, 2) веса а. э. выбираются согласно соотношениям

$$\alpha_0 = -\frac{1}{2}(x'_r + x''_s, x'_r - x''_s), \quad (3.7)$$

$$\sum_{j=1}^{N_a} \alpha_j a_j = x'_r - x''_s$$

и 3) тренировочная последовательность достаточно велика. Соотношения (3.5), (3.7) задают простейший алгоритм обучения.

**Лемма 2.** Пусть  $M$  — выпуклое тело,  $x_0$  — внутренняя и  $z$  — граничная точка  $M$ . Тогда отрезок  $x(t) = tz + (1-t)x_0$  при  $0 \leq t < 1$  состоит из внутренних точек  $M$ .

ДОКАЗАТЕЛЬСТВО ЛЕММЫ 2. Пусть  $x(t_0)$  ( $0 < t_0 < 1$ ) не является внутренней точкой  $M$ . Тогда существуют последовательности  $z_n \in M, x_n \notin M, z_n \rightarrow z, x_n \rightarrow x(t_0)$  при  $n \rightarrow \infty$ .

Очевидно

$$x'_n = \frac{x_n - t_0 z_n}{1 - t_0} \notin M,$$

так как в противном случае мы получили бы, что  $x_n = t_0 z + (1 - t_0)x'_n \in M$ .

Поскольку

$$x'_n \rightarrow \frac{x(t_0) - t_0 z}{1 - t_0} = x_0,$$

мы получаем, что  $x_0$  не является внутренней точкой. Лемма доказана.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1. Пусть  $x', x''$  — образы, удовлетворяющие соотношению (3.2), и  $\psi(x)$  определяется равенством (3.3). Тогда, как мы видели выше, выполнено (3.1). По пункту 2 теоремы 1 существует  $\epsilon > 0$  такое, что для любых  $x'_0, x''_0$ , удовлетворяющих неравенствам

$$|x' - x'_0| < \epsilon, \quad |x'' - x''_0| < \epsilon, \quad (3.8)$$

соотношение (3.1) выполнено также для функции

$$\psi_0(x) = \left( x - \frac{x'_0 + x''_0}{2}, x'_0 - x''_0 \right).$$

Пусть  $x'_* \in S_1$  — какая-либо внутренняя точка в  $S_1$  и  $x'_0 \neq x'$  — какая-либо точка, лежащая на отрезке, соединяющем  $x'$  и  $x'_*$ , и внутри  $\epsilon$ -окрестности точки  $x'$ . Так как  $x'$ , очевидно, — граничная точка множества  $S_1$ , то  $x'_0$  будет по лемме 1 внутренней

точкой множества  $S_1$ . Найдем аналогичным образом точку  $x''_0 \in S_2$ , внутреннюю во множестве  $S_2$  и удовлетворяющую второму неравенству (3.8).

Обозначим  $O'$  и  $O''$  окрестности точек  $x'_0$  и  $x''_0$  соответственно, содержащиеся в  $\epsilon$ -окрестности этих точек и такие, что  $O' \subset S_1, O'' \subset S_2$ .

Пусть  $p_1 > 0, p_2 > 0$  — вероятности попадания случайной точки в окрестности  $O'$  и  $O''$ . Вероятность того, что хотя бы одна из точек  $x'_j$  попала в  $O'$  и хотя бы одна из точек  $x''_h$  попала в  $O''$ , есть

$$(1 - q_1^{m_1})(1 - q_2^{m_2}),$$

где  $q_j = 1 - p_j$  ( $j = 1, 2$ ). Поскольку при этом будет выполнено (3.1) для функции (3.5), то

$$P_{m_1, m_2} \geq (1 - q_1^{m_1})(1 - q_2^{m_2})$$

и утверждение теоремы доказано.

#### § 4. ПОНЯТИЯ, КОТОРЫЕ МОЖЕТ РАЗДЕЛЯТЬ ПЕРСЕПТРОН С ОДНИМ СУММИРУЮЩИМ ЭЛЕМЕНТОМ

Для того чтобы понятия разделялись персептроном с одним суммирующим элементом, достаточно, чтобы при введении подходящего пространства образов соответствующие множества были выпуклыми. Необходимое и достаточное условие состоит в том, чтобы не пересекались выпуклые оболочки соответствующих множеств.

Гласные фонемы (а возможно, и согласные) являются понятиями, удовлетворяющими первому условию. Мы исходим здесь из интуитивного представления о том, что звук « $O$ » остается звуком « $O$ » при увеличении или уменьшении громкости звучания и что при одновременном произношении звука « $O$ » полученный звук также будет звуком « $O$ ». В переводе на математический язык наше утверждение выглядит так. Образом (звуком) длительности  $T$  называется функция из  $L_2(0, T)$

$$\phi(t) = \int_{-\infty}^{+\infty} \xi(\lambda) e^{i\lambda t} d\lambda \quad (\xi(\lambda) > 0) \quad (4.1)$$

(представляющая запись в каких-либо единицах этого звука на ленте идеального магнитофона).

Звукам, отличающимся лишь громкостью, будут соответствовать функции  $\phi(t)$  и  $K\phi(t)$ , отличающиеся лишь мультипликативной постоянной. Если произносятся одновременно два звука  $\phi_1(t)$  и  $\phi_2(t)$ , то воспринимается  $\phi_1(t) + \phi_2(t)$ . Мы утверждали выше, что звуки  $\phi_1(t)$  и  $\phi_2(t)$  принадлежат одной фонеме, той же фонеме принадлежит звук  $\alpha\phi_1(t) + \beta\phi_2(t)$ , где  $\alpha > 0, \beta > 0, \alpha + \beta = 1$ .

Принадлежность звука определенной фонеме удобнее определять в терминах спектральной плотности  $\xi(\lambda)$ . Характер спектральных плотностей для различных фонем определен в работе [10].

Поскольку преобразование (4.1) является линейным, в гильбертовом пространстве  $H = \xi(\lambda)$  гласным фонемам будут также соответствовать выпуклые множества. Таким образом, определению гласных фонем может быть обучен персептрон с одним с. э., на выход которого вместо звуков подаются соответствующие спектральные плотности. Это же справедливо в отношении звучаний слов (элементарным понятием является множество звуков, соответствующее фиксированному слову). Подчеркнем, что мы имеем в виду здесь «нормальное» звучание, начинающееся в момент

$t = 0$  и заканчивающееся в момент  $t = T$  и без нарочитого растяжения отдельных слогов. (Это условие может, конечно, выполняться приближенно.) Для разделения двух множеств, соответствующих словам, когда в каждое множество включаются звуки различной длительности, нужен перцептрон со многими с. э.

Аппроксимируя функции  $\phi(t) \in L_2(0, T)$  тригонометрическими полиномами

$$\psi(t) = \sum_{j=1}^N \xi_j \cos \lambda_j t \quad (\xi_j > 0),$$

получим  $N$ -мерное пространство образов (частоты  $\lambda_j$  предполагаются фиксированными). Число  $N$  определяется требуемой точностью воспроизведения звука. Таким образом, для того чтобы перцептрон разделял фонемы не хуже, чем ухо человека, число а. э. перцептрона, которое, как было показано выше, также равно  $N$ , должно быть равно числу гармоник, которое может различать ухо человека.

Переходя к обонянию, предположим, что каждому запаху соответствует некоторая, может быть, векторная функция на поле рецепторных нейронов. Предположим далее, что одновременно воспринимаемым запахам соответствуют суммы функций, а запахам, отличающимся лишь «разной силой», соответствуют функции, отличающиеся численным множителем, пропорциональным «силе запаха». Если эти предположения справедливы, то понятиям, объединяющим однотипные запахи (таким как «запах ландыша», «запах резины»), соответствуют выпуклые множества и, следовательно, разделению подобных запахов может быть обучен перцептрон с одним с. э.

Значительно сложнее обстоит дело со зрительными образами. Отвлекаясь от различия изображения в различных глазах и считая для простоты, что воспринимаются лишь черно-белые изображения, нарисованные на экране, получим, что зрительным образом можно считать вещественную функцию, заданную на экране. При такой естественной трактовке в соответствующем гильбертовом пространстве образов таким простым понятиям, как «изображение буквы А», соответствуют, очевидно, невыпуклые множества. В этом пространстве выпуклые множества соответствуют лишь совсем простым понятиям, таким как «светлое пятно», «темное пятно», «темное пятно в левой стороне экрана» и т. п.<sup>5</sup>

Рассмотрим теперь случай, когда образом является контурный рисунок. Заменяя контур системой точек, будем считать, что образом является упорядоченная система  $k$  точек на плоскости<sup>6</sup>. Вводя сложение и умножение на числа по компонентам, превратим пространство образов  $X$  в  $2k$ -мерное евклидово пространство. Пусть  $M = |\xi|$  — какое-либо множество  $M + \eta = \xi + \eta$ , т. е. множество точек, полученное из  $M$  сдвигом на вектор  $\eta$ . При этом  $M = M(0)$ . Следующая теорема доставляет широкий класс выпуклых множеств в пространстве  $X$ .

**Теорема 2.** *Предположим, что  $M_1, M_2, \dots$  — конечная система выпуклых множеств на плоскости. Тогда множество  $S \subset X$  образов  $x = (\xi_1, \xi_2, \dots, \xi_k)$ , опреде-*

<sup>5</sup> Эти пятна притом не должны быть слишком яркими — соответствующие интенсивности окраски должны быть выпуклыми или вогнутыми функциями.

<sup>6</sup> Недостатком такой трактовки является аксиоматизация упорядоченности точек образа. Без упорядоченности, однако, пространство не будет линейным. Впрочем, для широкого класса образов (профиль человека) упорядочение точек может быть просто произведено.

ленные соотношениями

$$\begin{aligned} \xi_1 &\in M_1(0), \\ \xi_2 &\in M_2(0) \cap M_3(\xi_1), \\ \xi_3 &\in M_4(0) \cap M_5(\xi_1) \cap M_6(\xi_2), \\ &\dots, \end{aligned} \tag{4.2}$$

является выпуклым. Если, кроме того, каждое из множеств  $M_j$  содержит хотя бы одну внутреннюю точку, то множество  $S$  является выпуклым телом.

ДОКАЗАТЕЛЬСТВО. Второе утверждение теоремы очевидно. Пусть  $x^{(j)} = (\xi_1^{(j)}, \dots, \xi_k^{(j)}) \in S$  ( $j = 1, 2$ ). Согласно (4.2), имеем

$$\begin{aligned} \xi_1^{(j)} &\in M_1(0), \quad \xi_2^{(j)} \in M_2(0), \quad \xi_2^{(j)} - \xi_1^{(j)} \in M_3(0), \\ \xi_3^{(j)} &\in M_4(0), \quad \xi_3^{(j)} - \xi_1^{(j)} \in M_5(0), \quad \xi_3^{(j)} - \xi_2^{(j)} \in M_6(0), \dots \end{aligned}$$

Используя выпуклость множеств  $M_1(0), M_2(0), \dots$ , получим для любых  $\alpha \geq 0, \beta \geq 0, \alpha + \beta = 1$ :

$$\begin{aligned} \xi_1 &= \alpha\xi_1^{(1)} + \beta\xi_1^{(2)} \in M_1(0), \quad \xi_2 = \alpha\xi_2^{(1)} + \beta\xi_2^{(2)} \in M_2(0), \\ \xi_2 - \xi_1 &\in M_3(0), \quad \xi_3 = \alpha\xi_3^{(1)} + \beta\xi_3^{(2)} \in M_4(0), \quad \xi_3 - \xi_1 \in M_5(0), \dots \end{aligned}$$

Это означает, что для чисел  $\xi_h = \alpha\xi_h^{(1)} + \beta\xi_h^{(2)}$  выполнены соотношения (4.2). Таким образом,  $x = \alpha x^{(1)} + \beta x^{(2)} \in S$ , что и требовалось доказать.

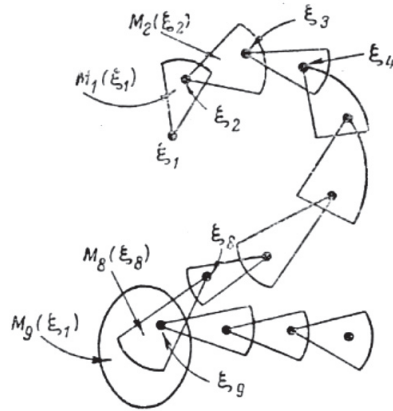


Рис. 2.

Отметим, что некоторые из множеств  $M_j$  могут совпадать с плоскостью, тогда в формуле (4.2) эти множества просто не будут фигурировать. Такие понятия, как «изображение цифры 2», «профиль лица человека», «профиль мужского лица с толстым носом», можно задать формулой (4.2), подбирая некоторым образом выпуклые множества  $M_j$ . Рис. 2 поясняет это на примере понятия «изображение цифры 2». Множества  $M_j$  имеют указанный на рис. 2 вид, а цепочка включений (4.2) в других обозначениях является следующей:

$$\begin{aligned} \xi_1 &\in M_0, \\ \xi_2 &\in M_1(\xi_1) \cap M_0, \end{aligned}$$

$$\begin{aligned}
&\xi_3 \in M_2(\xi_2) \cap M_0, \\
&\dots \\
&\xi_6 \in M_7(\xi_7) \cap M_0, \\
&\xi_9 \in M_8(\xi_8) \cap M_9(\xi_1) \cap M_0, \\
&\xi_{10} \in M_{10}(\xi_9) \cap M_0, \\
&\dots \\
&\xi_{12} \in M_{12}(\xi_{11}) \cap M_0.
\end{aligned}$$

Здесь  $M_0$  является экраном. Теорема показывает, что подобным понятиям отвечают выпуклые множества<sup>7</sup>.

Интересно отметить, имеются некоторые данные [9], что зрительная система человека производит декорреляцию изображения, выделяя на уровне ганглиозных клеток основные контуры из полного изображения.

В заключение этого параграфа отметим, что различные математические идеализации для «реально» одних и тех же образов приводят к совершенно различным, с инженерной точки зрения, принципам построения перцептрона.

Предположим теперь, что понятие  $S$  допустимой системы инвариантно относительно некоторых преобразований. Например, каждое понятие из системы «изображения буквы и цифры» инвариантно относительно плоско-параллельного сдвига изображения и относительно подобного увеличения или уменьшения изображения.

В этом случае естественно заложить в конструкцию перцептрона требование, чтобы реакция каждого ассоциативного элемента, а значит и всего перцептрона на произвольный образ, не изменялась бы, когда образ подвергается заданным преобразованиям.

Мы рассмотрим случай, когда в пространстве образов имеется подпространство  $Z_0$ , конечномерное и для простоты такое, что для любого понятия  $S$  допустимой системы и произвольного  $z \in Z$  из  $x \in X$  следует  $x + z \in S$ . (Понятия  $S$  инвариантны относительно сдвигов вдоль подпространства  $Z_0$ .)

Пусть  $z_1, \dots, z_p$  — ортогональный базис в  $Z_0$  и  $Y$  — ортогональное дополнение подпространства  $Z_0$ .

Если  $\psi(x) = \alpha + (a, x)$  разделяет  $S_1$  и  $S_2$ , т. е. выполнено (3.1), то, как легко проверить,  $a \in Y$ . Поэтому ассоциативные элементы  $a_j$  без ограничения возможностей перцептрона можно брать из  $Y$ . Если  $n$  — размерность пространства образов, то для числа ассоциативных элементов мы получаем неравенство

$$N_a \geq n - p. \quad (4.3)$$

Для произвольного  $x$  получим

$$\tilde{x} = x - \sum_{j=1}^p \rho_j z_j, \quad (4.4)$$

<sup>7</sup> Выделение соответствующих множеств по заданному понятию может быть непосильной задачей. Однако правдоподобной является гипотеза, что эти множества существуют даже для таких сложных понятий, как «профиль лица данного человека». Разумеется, нам важен лишь факт существования этих множеств.

где

$$\rho_j = \frac{(x, z_j)}{(z_j, z_j)}.$$

(Вектор  $\tilde{x}$  является проекцией в  $Y$  вектора  $x$ .) Мы имеем  $\tilde{x} \perp Z$ , и из  $x \in S$  следует  $\tilde{x} \in S$ . Пусть

$$x'_j \in S_1 \quad (j = 1, \dots, m_1), \quad x''_h \in S_2 \quad (h = 1, \dots, m_2)$$

— заданные тренировочные последовательности. Алгоритм, вырабатывающий вектор  $a$  по последовательности  $x'_j, x''_h$ , обычно не выводит из подпространства, натянутого на векторы  $x'_j, x''_h$ . Этим свойством обладают, в частности, простейший алгоритм §1 (см. (3.7)), а также алгоритмы, описанные в следующем параграфе. Если выполнено это условие, то, заменяя векторы  $x'_j, x''_h$  на векторы  $\tilde{x}'_j, \tilde{x}''_h$ , которые получаются из  $x'_j, x''_h$  по формуле (4.4), получим  $a \in Y$ . В этом случае реакция перцептрона на образы  $x$  и  $x + z$  для любого  $z \in Z$  будет одинакова.

#### ЛИТЕРАТУРА<sup>8</sup>

1. Трапезников В. А. Кибернетика и автоматическое управление // Автоматика и телемеханика. Т. XXIII, № 3. 1962.
2. Браверман Э. М. Опыты по обучению машины распознаванию зрительных образов // Автоматика и телемеханика. Т. XXIII, № 3. 1962.
3. Браверман Э. М. Некоторые вопросы построения машин, классифицирующих объекты по не заданному заранее признаку // Автоматика и телемеханика. Т. XXI, № 10. 1960.
4. Rosenblatt F. Two Theorems of Statistical Separability in the Perceptron // Symposium of the Mechanisation of Thought Processes. England, November, 1958.
5. Rosenblatt F. Perceptron Simulation Experiments // Proc. IRE. Vol. 48, № 3. 1960.
6. Selfridge O. G. Pandemonium: a Paradigm for Learning // Symposium of the Mechanisation of Thought Processes. England, November, 1958.
7. Харкевич А. А. Опознание образов // Радиотехника. Т. 14, № 5. 1959.
8. Joseph R. D. On Predicting Perceptron Performance // IRE Intern. Convent. Rec. Vol. 8, № 2. 1960.
9. Глезер В. Д., Цукерман И. И. Информация и зрение. М.: Изд. АН СССР, 1961.
10. Варшавский Л. А., Чистович Л. А. Средние спектры русских гласных фонем // Проблемы физиологической акустики. Т. IV. 1959.

---

<sup>8</sup> Примечание при корректуре: см. также недавно опубликованные работы М. А. Айзермана, М. М. Бонгарда и других авторов в сборнике «Биологические аспекты кибернетики» (Изд-во АН СССР, 1962) и книгу М. Глушкова «Введение в теорию самосовершенствующихся систем» (Киев, 1962).