Сети Хопфилда и Хэмминга

Конфигурации сетей с обратными связями

Рассмотренный нами ранее персептрон относится к классу сетей с направленным потоком распространения информации и не содержит обратных связей. На этапе функционирования каждый нейрон выполняет свою функцию — передачу возбуждения другим нейронам — ровно один раз. Динамика состояний нейронов является неитерационной.

Несколько более сложной является динамика в сети Кохонена. Конкурентное соревнование нейронов достигается путем итераций, в процессе которых информация многократно передается между нейронами.

В общем случае может быть рассмотрена нейронная сеть, содержащая произвольные обратные связи, по которым переданное возбуждение возвращается к данному нейрону, и он повторно выполняет свою функцию. Наблюдения за биологическими локальными нейросетями указывают на наличие множественных обратных связей. Нейродинамика в таких системах становится итерационной. Это свойство существенно расширяет множество типов нейросетевых архитектур, но одновременно приводит к появлению новых проблем.

Неитерационная динамика состояний нейронов является, очевидно, всегда устойчивой. Обратные связи могут приводить к возникновению неустойчивостей, подобных тем, которые возникают в усилительных радиотехнических системах при положительной обратной связи. В нейронных сетях неустойчивость проявляется в блуждающей смене состояний нейронов, не приводящей к возникновению стационарных состояний. В общем случае, ответ на вопрос об устойчивости динамики произвольной системы с обратными связями крайне сложен и до настоящего времени является открытым.

Остановимся на важном частном случае нейросетевой архитектуры, для которой свойства устойчивости подробно исследованы. На [рис. 8.1](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=1#image.8.1) показана сеть с обратными связями, состоящая из двух слоев. Способ представления несколько отличается от использованного в работе Хопфилда и других сходных, но эквивалентен им с функциональной точки зрения, а также хорошо связан с сетями, рассмотренными на предыдущих лекциях. Нулевой слой, как и на предыдущих рисунках, не выполняет вычислительной функции, а лишь распределяет выходы сети обратно на входы. Каждый нейрон первого слоя вычисляет взвешенную сумму своих входов, давая сигнал NET, который затем с помощью нелинейной функции F преобразуется в сигнал OUT. Эти операции сходны с нейронами других сетей.

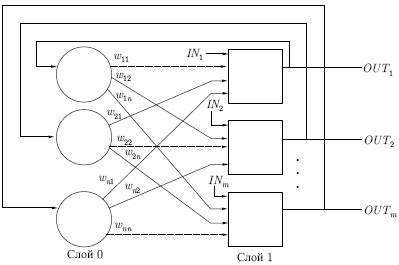


Рис. 8.1.

Бинарные системы

В первой работе Д.Хопфилда функция F была просто пороговой функцией. Выход такого нейрона равен единице, если взвешенная сумма выходов с других нейронов больше порога T_j, в противном случае она равна нулю. Порог вычисляется следующим образом:

\begin{gathered}
NET_j=\sum_{i\ne j}w_{ij} OUT_i+IN_j,\\
OUT_j=\left\{\begin{aligned}
1, \quad & \text{если } NET_j>T_j,\\
0, \quad & \text{если } NET_j<T_j,\\
\text{не меняется}, \quad & \text{если } NET_j=T_j.
\end{aligned}\right.
\end{gathered}

Состояние сети — это просто множество текущих значений сигналов OUT от всех нейронов. В первоначальной сети Хопфилда состояние каждого нейрона менялось в дискретные случайные моменты времени, в последующем - состояния нейронов могли меняться одновременно. Так как выходом бинарного нейрона может быть только ноль или единица (промежуточных уровней нет), то текущее состояние сети является двоичным числом, каждый бит которого является сигналом OUT некоторого нейрона.

Задачи, решаемые данной сетью, как правило, формулируются следующим образом. Известен некоторый набор двоичных сигналов (изображений, оцифровок звука, прочих данных, описывающих некие объекты или характеристики процессов), которые считаются образцовыми. Сеть должна уметь из произвольного неидеального сигнала, поданного на ее вход, выделить ("вспомнить" по частичной информации) соответствующий образец (если такой есть) или "дать заключение" о том, что входные данные не соответствуют ни одному из образцов. В общем случае, любой сигнал может быть описан вектором X=\{x_i\colon i=0\ldots
n-1\}, n — число нейронов в сети и размерность входных и выходных векторов. Каждый элемент x_i равен либо 1, либо 0. Обозначим вектор, описывающий k-й образец, через X^k, а его компоненты, соответственно, — x_i^k, k=0,\ldots,m-1, m — число компонентов. Когда сеть распознaет (или "вспомнит") какой-либо образец на основе предъявленных ей данных, ее выходы будут содержать именно его, то есть Y = X^k, где Y --вектор выходных значений сети: Y = \{ y_i\colon
i=0,\ldots,n-1\}. В противном случае, выходной вектор не совпадет ни с одним образцовым.

Если, например, сигналы представляют собой некие изображения, то, отобразив в графическом виде данные с выхода сети, можно будет увидеть картинку, полностью совпадающую с одной из образцовых (в случае успеха) или же "вольную импровизацию" сети (в случае неудачи).

На стадии инициализации сети весовые коэффициенты синапсов устанавливаются следующим образом:

w_{ij}=\left\{
\begin{aligned}
\sum_{k=0}^{m-1}x_i^k x_j^k,\quad & \text{если } i\ne j,\\
0,& \text{если } i=j.
\end{aligned}
\right.

Здесь i и j — индексы, соответственно, предсинаптического и постсинаптического нейронов; x_i^k, x_j^k — i -й и j -й элементы вектора k -го образца.

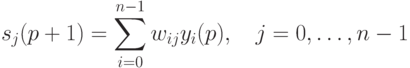
Алгоритм функционирования сети следующий ( p — номер итерации):

1. На входы сети подается неизвестный сигнал. Фактически его ввод осуществляется непосредственной установкой значений аксонов:

y_i(0)=x_i,\quad i=0,\ldots,n-1,

поэтому обозначение на схеме сети входных синапсов в явном виде носит чисто условный характер. Ноль в скобке справа от y_i означает нулевую итерацию в цикле работы сети.

1. Рассчитывается новое состояние нейронов:



и новые значения аксонов

y_j(p+1)=f\lfloor s_j(p+1)\rfloor.

где f — активационная функция в виде скачка.

1. Проверка, изменились ли выходные значения аксонов за последнюю итерацию. Если да — переход к пункту 2, иначе (если выходы стабилизировались) — конец процедуры. При этом выходной вектор представляет собой образец, наилучшим образом сочетающийся с входными данными.

Как говорилось выше, иногда сеть не может провести распознавание и выдает на выходе несуществующий образ. Это связано с проблемой ограниченности возможностей сети. Для сети Хопфилда число запоминаемых образов m не должно превышать величины, примерно равной 0,15n. Кроме того, если два образа А и Б имеют значительное сходство, они, возможно, будут вызывать у сети перекрестные ассоциации, то есть предъявление на входы сети вектора А приведет к появлению на ее выходах вектора Б и наоборот.

Когда нет необходимости, чтобы сеть выдавала образец в явном виде и достаточно, скажем, получать номер образца, ассоциативную память успешно реализует сеть Хэмминга. Данная сеть характеризуется, по сравнению с сетью Хопфилда, более экономным использованием памяти и меньшим объемом вычислений, что становится очевидным из ее структуры (см. [рис. 8.2](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=3#image.8.2)).

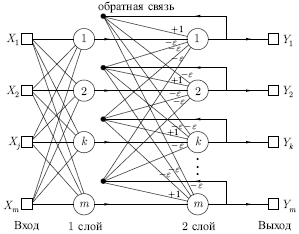


Рис. 8.2.

Сеть состоит из двух слоев. Первый и второй слои имеют по m нейронов, где m — число образцов. Нейроны первого слоя имеют по n синапсов, соединенных с входами сети (которые образуют фиктивный нулевой слой). Нейроны второго слоя связаны между собой ингибиторными (отрицательными обратными ) синаптическими связями. Единственный синапс с положительной обратной связью для каждого нейрона соединен с его же аксоном.

Идея работы сети состоит в нахождении расстояния Хэмминга от тестируемого образа до всех образцов. Расстоянием Хэмминга называется число отличающихся битов в двух бинарных векторах. Сеть должна выбрать образец с минимальным расстоянием Хэмминга до неизвестного входного сигнала, в результате чего будет активизирован только один выход сети, соответствующий именно этому образцу.

На стадии инициализации весовым коэффициентам первого слоя и порогу активационной функции присваиваются следующие значения:

\begin{gathered}
w_{ik}=\frac{x_i^k}{2},\quad i=0,\ldots, n-1,\; k=0,\ldots, m-1,\\
T_k=\frac n2,\quad k=0,\ldots,m-1.
\end{gathered}


Здесь x_i^k - i -й элемент k -го образца.

Весовые коэффициенты тормозящих синапсов во втором слое берут равными некоторой величине 0<\varepsilon<1/m. Синапс нейрона, связанный с его же аксоном, имеет вес +1.

Алгоритм функционирования сети Хэмминга следующий:

1. На входы сети подается неизвестный вектор

X=\{x_i|i=0,\ldots,n\},

исходя из которого рассчитываются состояния нейронов первого слоя (верхний индекс в скобках указывает номер слоя):

y_j^{(1)}=s_j^{(1)}=\sum_{i=0}^{n-1}w_{ij}x_i+T_j,\quad
j=0,\ldots, m-1.

После этого полученными значениями инициализируются значения аксонов второго слоя:

y_j^{(2)}=y_j^{(1)},\quad j=0,\ldots, m-1.

1. Вычисляются новые состояния нейронов второго слоя:

s_j^{(2)}(p+1)=y_j(p)-\varepsilon\sum_{k=0}^{m-1}y_k^{(2)}(p),\quad
k\ne j,\quad j=0,\ldots, m-1
\vspace{-1mm}

и значения их аксонов:

y_j^{(2)}(p+1)=f\lfloor s_j^{(2)}(p+1)\rfloor,\quad j=0,\ldots, m-1.

Активационная функция f имеет вид порога, причем величина F должна быть достаточно большой, чтобы любые возможные значения аргумента не приводили к насыщению.

1. Проверить, изменились ли выходы нейронов второго слоя за последнюю итерацию. Если да — перейти к шагу 2. Иначе — конец процедуры.

Из оценки алгоритма видно, что роль первого слоя весьма условна: воспользовавшись один раз на шаге 1 значениями его весовых коэффициентов, сеть больше не обращается к нему, поэтому первый слой может быть вообще исключен из сети.

Устойчивость

Как и в других сетях, веса между слоями в этой сети могут рассматриваться в виде матрицы W. Сеть с обратными связями является устойчивой, если ее матрица симметрична и имеет нули на главной диагонали, т. е. если w_{ij} = w_{ji} и w_{ii} = 0 для всех i.

Устойчивость такой сети может быть доказана с помощью элегантного математического метода. Допустим, что найдена функция, которая всегда убывает при изменении состояния сети. В конце концов, эта функция должна достичь минимума и прекратить изменение, гарантируя тем самым устойчивость сети. Такая функция, называемая функцией Ляпунова, для рассматриваемых сетей с обратными связями может быть введена следующим образом:

E=-\frac12\sum_i\sum_j w_{ij} OUT_i OUT_j-\sum_j I_j OUT_j+\sum_j
T_j OUT_j,

где E — искусственная энергия сети; w_{ij} — вес от выхода нейрона i к входу нейрона j ; OUT_i — выход нейрона j ; I_j — внешний вход нейрона j ; T_j — порог нейрона j.

Изменение энергии E, вызванное изменением состояния j -нейрона, есть

\delta E=\left[\sum_{i\ne j}(w_{ij}OUT_i)+I_j-T_j\right]\delta OUT_j=
-[NET_j-T_j]\delta OUT_j,

где \delta OUT_j — изменение выхода j -го нейрона.

Допустим, что величина NET нейрона j больше порога. Тогда выражение в скобках будет положительным, а из данных уравнений следует, что выход нейрона j должен измениться в положительную сторону (или остаться без изменения). Это значит, что \delta OUT_j может быть только положительным или нулем и \delta E должно быть отрицательным. Следовательно, энергия сети должна либо уменьшиться, либо остаться без изменения.

Далее, допустим, что величина NET меньше порога. Тогда величина \delta OUT_i может быть только отрицательной или нулем. Следовательно, опять энергия должна уменьшиться или остаться без изменения.

И окончательно, если величина NET равна порогу, \delta_i равна нулю и энергия остается без изменения.

Мы показали, что любое изменение состояния нейрона либо уменьшит энергию, либо оставит ее без изменения. Благодаря такому непрерывному стремлению к уменьшению энергия, в конце концов, должна достигнуть минимума и прекратить изменение. По определению такая сеть является устойчивой.

Симметрия сети является достаточным, но не необходимым условием для устойчивости системы. Имеется много устойчивых систем (например, все сети прямого действия), которые ему не удовлетворяют. Можно продемонстрировать примеры, в которых незначительное отклонение от симметрии будет приводить к непрерывным осцилляциям. Однако приближенной симметрии обычно достаточно для устойчивости систем.

Ассоциативность памяти и задача распознавания образов

Динамический процесс последовательной смены состояний нейронной сети Хопфилда завершается в некотором стационарном состоянии, являющимся локальным минимумом энергетической функции E(S). Невозрастание энергии в процессе динамики приводит к выбору такого локального минимума S, в бассейн притяжения которого попадает начальное состояние (исходный, предъявляемый сети образ) S_0. В этом случае также говорят, что состояние S_0 находится в чаше минимума S.

При последовательной динамике в качестве стационарного состояния будет выбран такой образ S, который потребует минимального числа изменений состояний отдельных нейронов. Поскольку для двух двоичных векторов минимальное число изменений компонент, переводящее один вектор в другой, является расстоянием Хемминга \rho_H(S,S_0), то можно заключить, что динамика сети заканчивается в ближайшем по Хеммингу локальном минимуме энергии.

Пусть состояние S соответствует некоторому идеальному образу памяти. Тогда эволюцию от состояния S_0 к состоянию S можно сравнить с процедурой постепенного восстановления идеального образа S по его искаженной (зашумленной или неполной) копии S_0. Память с такими свойствами процесса считывания информации является ассоциативной. При поиске искаженные части целого восстанавливаются по имеющимся неискаженным частям на основе ассоциативных связей между ними.

Ассоциативный характер памяти сети Хопфилда качественно отличает ее от обычной, адресной, компьютерной памяти. В последней извлечение необходимой информации происходит по адресу ее начальной точки (ячейки памяти). Потеря адреса (или даже одного бита адреса) приводит к потере доступа ко всему информационному фрагменту. При использовании же ассоциативной памяти доступ к информации производится непосредственно по ее содержанию, т.е. по частично известным искаженным фрагментам. Потеря части информации или ее зашумление не приводит к катастрофическому ограничению доступа, если оставшейся информации достаточно для извлечения идеального образа.

Поиск идеального образа по имеющейся неполной или зашумленной его версии называется задачей распознавания образов. В нашей лекции особенности решения этой задачи нейронной сетью Хопфилда будут продемонстрированы на примерах, которые получены с использованием модели сети на персональной ЭВМ.

В рассматриваемой модели сеть содержала 100 нейронов, упорядоченных в матрицу 10\times 10. Сеть обучалась по правилу Хебба на трех идеальных образах — шрифтовых начертаниях латинских букв M, A и G (см. [рис. 8.3](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.3)). После обучения нейросети в качестве начальных состояний нейронов предъявлялись различные искаженные версии образов, которые в дальнейшем эволюционировали с последовательной динамикой к стационарным состояниям.

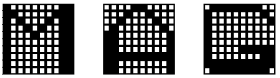


Рис. 8.3.

Для каждой пары изображений на [рисунке 8.4](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4), левый образ является начальным состоянием, а правый — результатом работы сети, достигнутым стационарным состоянием.

Образ на [рис. 8.4(А)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) был выбран для тестирования адекватности поведения на идеальной задаче, когда предъявленное изображение точно соответствует информации в памяти. В этом случае за один шаг было достигнуто стационарное состояние. Образ на [рис. 8.4(Б)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) характерен для задач распознавания текста независимо от типа шрифта. Начальное и конечное изображения безусловно похожи, но попробуйте это объяснить машине!

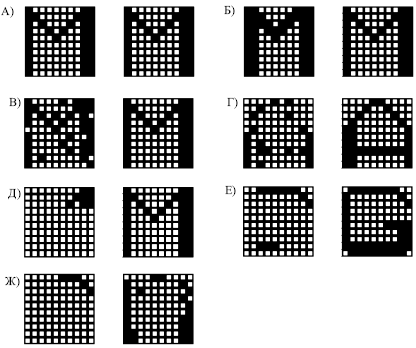


Рис. 8.4.

Задания на [рис. 8.4(В, Г)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) характерны для практических приложений. Нейросетевая система способна распознавать практически полностью зашумленные образы. Задачи, соответствующие рисункам [8.4(Д, Е)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4), демонстрируют замечательное свойство сети Хопфилда: она способна ассоциативно узнавать образ по его небольшому фрагменту. Важнейшей особенностью работы сети является генерация ложных образов. Пример ассоциации к ложному образу показан на [рис. 8.4(Ж)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4). Ложный образ является устойчивым локальным экстремумом энергии, но не соответствует никакому идеальному образу. Он является в некотором смысле собирательным образом, наследующим черты идеальных собратьев. Ситуация с ложным образом эквивалентна нашему "Где-то я уже это видел".

В данной простейшей задаче ложный образ является "неверным" решением и поэтому вреден. Однако можно надеяться, что такая склонность сети к обобщениям может быть как-то использована. Характерно, что при увеличении объема полезной информации (сравните [рис. 8.4 (Е) и (Ж))](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) исходное состояние попадает в область притяжения требуемого стационарного состояния, и образ распознается.

* Просмотр
* [Редактировать](https://moodle.bgpu.ru/mod/lesson/edit.php?id=71565)
* [Отчеты](https://moodle.bgpu.ru/mod/lesson/report.php?id=71565)
* [Оценить эссе](https://moodle.bgpu.ru/mod/lesson/essay.php?id=71565)

×В лекции используются переходы «Непросмотренный вопрос в кластере» или «Страница непросмотренного вопроса из раздела». Сейчас вместо них будет использован переход «Следующая страница». Зайдите как студент, чтобы протестировать эти переходы.

Сети Хопфилда и Хэмминга

Конфигурации сетей с обратными связями

Рассмотренный нами ранее персептрон относится к классу сетей с направленным потоком распространения информации и не содержит обратных связей. На этапе функционирования каждый нейрон выполняет свою функцию — передачу возбуждения другим нейронам — ровно один раз. Динамика состояний нейронов является неитерационной.

Несколько более сложной является динамика в сети Кохонена. Конкурентное соревнование нейронов достигается путем итераций, в процессе которых информация многократно передается между нейронами.

В общем случае может быть рассмотрена нейронная сеть, содержащая произвольные обратные связи, по которым переданное возбуждение возвращается к данному нейрону, и он повторно выполняет свою функцию. Наблюдения за биологическими локальными нейросетями указывают на наличие множественных обратных связей. Нейродинамика в таких системах становится итерационной. Это свойство существенно расширяет множество типов нейросетевых архитектур, но одновременно приводит к появлению новых проблем.

Неитерационная динамика состояний нейронов является, очевидно, всегда устойчивой. Обратные связи могут приводить к возникновению неустойчивостей, подобных тем, которые возникают в усилительных радиотехнических системах при положительной обратной связи. В нейронных сетях неустойчивость проявляется в блуждающей смене состояний нейронов, не приводящей к возникновению стационарных состояний. В общем случае, ответ на вопрос об устойчивости динамики произвольной системы с обратными связями крайне сложен и до настоящего времени является открытым.

Остановимся на важном частном случае нейросетевой архитектуры, для которой свойства устойчивости подробно исследованы. На [рис. 8.1](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=1#image.8.1) показана сеть с обратными связями, состоящая из двух слоев. Способ представления несколько отличается от использованного в работе Хопфилда и других сходных, но эквивалентен им с функциональной точки зрения, а также хорошо связан с сетями, рассмотренными на предыдущих лекциях. Нулевой слой, как и на предыдущих рисунках, не выполняет вычислительной функции, а лишь распределяет выходы сети обратно на входы. Каждый нейрон первого слоя вычисляет взвешенную сумму своих входов, давая сигнал NET, который затем с помощью нелинейной функции F преобразуется в сигнал OUT. Эти операции сходны с нейронами других сетей.

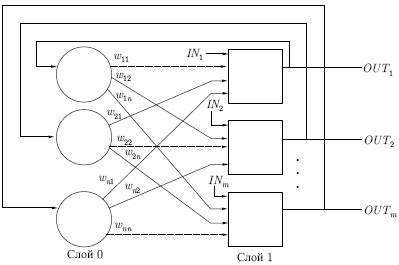


Рис. 8.1.

Бинарные системы

В первой работе Д.Хопфилда функция F была просто пороговой функцией. Выход такого нейрона равен единице, если взвешенная сумма выходов с других нейронов больше порога T_j, в противном случае она равна нулю. Порог вычисляется следующим образом:

\begin{gathered}
NET_j=\sum_{i\ne j}w_{ij} OUT_i+IN_j,\\
OUT_j=\left\{\begin{aligned}
1, \quad & \text{если } NET_j>T_j,\\
0, \quad & \text{если } NET_j<T_j,\\
\text{не меняется}, \quad & \text{если } NET_j=T_j.
\end{aligned}\right.
\end{gathered}

Состояние сети — это просто множество текущих значений сигналов OUT от всех нейронов. В первоначальной сети Хопфилда состояние каждого нейрона менялось в дискретные случайные моменты времени, в последующем - состояния нейронов могли меняться одновременно. Так как выходом бинарного нейрона может быть только ноль или единица (промежуточных уровней нет), то текущее состояние сети является двоичным числом, каждый бит которого является сигналом OUT некоторого нейрона.

Задачи, решаемые данной сетью, как правило, формулируются следующим образом. Известен некоторый набор двоичных сигналов (изображений, оцифровок звука, прочих данных, описывающих некие объекты или характеристики процессов), которые считаются образцовыми. Сеть должна уметь из произвольного неидеального сигнала, поданного на ее вход, выделить ("вспомнить" по частичной информации) соответствующий образец (если такой есть) или "дать заключение" о том, что входные данные не соответствуют ни одному из образцов. В общем случае, любой сигнал может быть описан вектором X=\{x_i\colon i=0\ldots
n-1\}, n — число нейронов в сети и размерность входных и выходных векторов. Каждый элемент x_i равен либо 1, либо 0. Обозначим вектор, описывающий k-й образец, через X^k, а его компоненты, соответственно, — x_i^k, k=0,\ldots,m-1, m — число компонентов. Когда сеть распознaет (или "вспомнит") какой-либо образец на основе предъявленных ей данных, ее выходы будут содержать именно его, то есть Y = X^k, где Y --вектор выходных значений сети: Y = \{ y_i\colon
i=0,\ldots,n-1\}. В противном случае, выходной вектор не совпадет ни с одним образцовым.

Если, например, сигналы представляют собой некие изображения, то, отобразив в графическом виде данные с выхода сети, можно будет увидеть картинку, полностью совпадающую с одной из образцовых (в случае успеха) или же "вольную импровизацию" сети (в случае неудачи).

На стадии инициализации сети весовые коэффициенты синапсов устанавливаются следующим образом:

w_{ij}=\left\{
\begin{aligned}
\sum_{k=0}^{m-1}x_i^k x_j^k,\quad & \text{если } i\ne j,\\
0,& \text{если } i=j.
\end{aligned}
\right.

Здесь i и j — индексы, соответственно, предсинаптического и постсинаптического нейронов; x_i^k, x_j^k — i -й и j -й элементы вектора k -го образца.

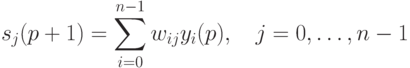
Алгоритм функционирования сети следующий ( p — номер итерации):

1. На входы сети подается неизвестный сигнал. Фактически его ввод осуществляется непосредственной установкой значений аксонов:

y_i(0)=x_i,\quad i=0,\ldots,n-1,

поэтому обозначение на схеме сети входных синапсов в явном виде носит чисто условный характер. Ноль в скобке справа от y_i означает нулевую итерацию в цикле работы сети.

1. Рассчитывается новое состояние нейронов:



и новые значения аксонов

y_j(p+1)=f\lfloor s_j(p+1)\rfloor.

где f — активационная функция в виде скачка.

1. Проверка, изменились ли выходные значения аксонов за последнюю итерацию. Если да — переход к пункту 2, иначе (если выходы стабилизировались) — конец процедуры. При этом выходной вектор представляет собой образец, наилучшим образом сочетающийся с входными данными.

Как говорилось выше, иногда сеть не может провести распознавание и выдает на выходе несуществующий образ. Это связано с проблемой ограниченности возможностей сети. Для сети Хопфилда число запоминаемых образов m не должно превышать величины, примерно равной 0,15n. Кроме того, если два образа А и Б имеют значительное сходство, они, возможно, будут вызывать у сети перекрестные ассоциации, то есть предъявление на входы сети вектора А приведет к появлению на ее выходах вектора Б и наоборот.

Когда нет необходимости, чтобы сеть выдавала образец в явном виде и достаточно, скажем, получать номер образца, ассоциативную память успешно реализует сеть Хэмминга. Данная сеть характеризуется, по сравнению с сетью Хопфилда, более экономным использованием памяти и меньшим объемом вычислений, что становится очевидным из ее структуры (см. [рис. 8.2](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=3#image.8.2)).

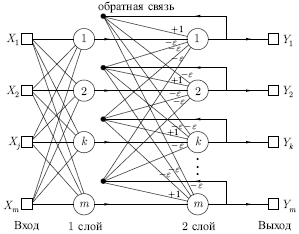


Рис. 8.2.

Сеть состоит из двух слоев. Первый и второй слои имеют по m нейронов, где m — число образцов. Нейроны первого слоя имеют по n синапсов, соединенных с входами сети (которые образуют фиктивный нулевой слой). Нейроны второго слоя связаны между собой ингибиторными (отрицательными обратными ) синаптическими связями. Единственный синапс с положительной обратной связью для каждого нейрона соединен с его же аксоном.

Идея работы сети состоит в нахождении расстояния Хэмминга от тестируемого образа до всех образцов. Расстоянием Хэмминга называется число отличающихся битов в двух бинарных векторах. Сеть должна выбрать образец с минимальным расстоянием Хэмминга до неизвестного входного сигнала, в результате чего будет активизирован только один выход сети, соответствующий именно этому образцу.

На стадии инициализации весовым коэффициентам первого слоя и порогу активационной функции присваиваются следующие значения:

\begin{gathered}
w_{ik}=\frac{x_i^k}{2},\quad i=0,\ldots, n-1,\; k=0,\ldots, m-1,\\
T_k=\frac n2,\quad k=0,\ldots,m-1.
\end{gathered}


Здесь x_i^k - i -й элемент k -го образца.

Весовые коэффициенты тормозящих синапсов во втором слое берут равными некоторой величине 0<\varepsilon<1/m. Синапс нейрона, связанный с его же аксоном, имеет вес +1.

Алгоритм функционирования сети Хэмминга следующий:

1. На входы сети подается неизвестный вектор

X=\{x_i|i=0,\ldots,n\},

исходя из которого рассчитываются состояния нейронов первого слоя (верхний индекс в скобках указывает номер слоя):

y_j^{(1)}=s_j^{(1)}=\sum_{i=0}^{n-1}w_{ij}x_i+T_j,\quad
j=0,\ldots, m-1.

После этого полученными значениями инициализируются значения аксонов второго слоя:

y_j^{(2)}=y_j^{(1)},\quad j=0,\ldots, m-1.

1. Вычисляются новые состояния нейронов второго слоя:

s_j^{(2)}(p+1)=y_j(p)-\varepsilon\sum_{k=0}^{m-1}y_k^{(2)}(p),\quad
k\ne j,\quad j=0,\ldots, m-1
\vspace{-1mm}

и значения их аксонов:

y_j^{(2)}(p+1)=f\lfloor s_j^{(2)}(p+1)\rfloor,\quad j=0,\ldots, m-1.

Активационная функция f имеет вид порога, причем величина F должна быть достаточно большой, чтобы любые возможные значения аргумента не приводили к насыщению.

1. Проверить, изменились ли выходы нейронов второго слоя за последнюю итерацию. Если да — перейти к шагу 2. Иначе — конец процедуры.

Из оценки алгоритма видно, что роль первого слоя весьма условна: воспользовавшись один раз на шаге 1 значениями его весовых коэффициентов, сеть больше не обращается к нему, поэтому первый слой может быть вообще исключен из сети.

Устойчивость

Как и в других сетях, веса между слоями в этой сети могут рассматриваться в виде матрицы W. Сеть с обратными связями является устойчивой, если ее матрица симметрична и имеет нули на главной диагонали, т. е. если w_{ij} = w_{ji} и w_{ii} = 0 для всех i.

Устойчивость такой сети может быть доказана с помощью элегантного математического метода. Допустим, что найдена функция, которая всегда убывает при изменении состояния сети. В конце концов, эта функция должна достичь минимума и прекратить изменение, гарантируя тем самым устойчивость сети. Такая функция, называемая функцией Ляпунова, для рассматриваемых сетей с обратными связями может быть введена следующим образом:

E=-\frac12\sum_i\sum_j w_{ij} OUT_i OUT_j-\sum_j I_j OUT_j+\sum_j
T_j OUT_j,

где E — искусственная энергия сети; w_{ij} — вес от выхода нейрона i к входу нейрона j ; OUT_i — выход нейрона j ; I_j — внешний вход нейрона j ; T_j — порог нейрона j.

Изменение энергии E, вызванное изменением состояния j -нейрона, есть

\delta E=\left[\sum_{i\ne j}(w_{ij}OUT_i)+I_j-T_j\right]\delta OUT_j=
-[NET_j-T_j]\delta OUT_j,

где \delta OUT_j — изменение выхода j -го нейрона.

Допустим, что величина NET нейрона j больше порога. Тогда выражение в скобках будет положительным, а из данных уравнений следует, что выход нейрона j должен измениться в положительную сторону (или остаться без изменения). Это значит, что \delta OUT_j может быть только положительным или нулем и \delta E должно быть отрицательным. Следовательно, энергия сети должна либо уменьшиться, либо остаться без изменения.

Далее, допустим, что величина NET меньше порога. Тогда величина \delta OUT_i может быть только отрицательной или нулем. Следовательно, опять энергия должна уменьшиться или остаться без изменения.

И окончательно, если величина NET равна порогу, \delta_i равна нулю и энергия остается без изменения.

Мы показали, что любое изменение состояния нейрона либо уменьшит энергию, либо оставит ее без изменения. Благодаря такому непрерывному стремлению к уменьшению энергия, в конце концов, должна достигнуть минимума и прекратить изменение. По определению такая сеть является устойчивой.

Симметрия сети является достаточным, но не необходимым условием для устойчивости системы. Имеется много устойчивых систем (например, все сети прямого действия), которые ему не удовлетворяют. Можно продемонстрировать примеры, в которых незначительное отклонение от симметрии будет приводить к непрерывным осцилляциям. Однако приближенной симметрии обычно достаточно для устойчивости систем.

Ассоциативность памяти и [задача](https://moodle.bgpu.ru/mod/assign/view.php?id=109703) распознавания образов

Динамический процесс последовательной смены состояний нейронной сети Хопфилда завершается в некотором стационарном состоянии, являющимся локальным минимумом энергетической функции E(S). Невозрастание энергии в процессе динамики приводит к выбору такого локального минимума S, в бассейн притяжения которого попадает начальное состояние (исходный, предъявляемый сети образ) S_0. В этом случае также говорят, что состояние S_0 находится в чаше минимума S.

При последовательной динамике в качестве стационарного состояния будет выбран такой образ S, который потребует минимального числа изменений состояний отдельных нейронов. Поскольку для двух двоичных векторов минимальное число изменений компонент, переводящее один вектор в другой, является расстоянием Хемминга \rho_H(S,S_0), то можно заключить, что динамика сети заканчивается в ближайшем по Хеммингу локальном минимуме энергии.

Пусть состояние S соответствует некоторому идеальному образу памяти. Тогда эволюцию от состояния S_0 к состоянию S можно сравнить с процедурой постепенного восстановления идеального образа S по его искаженной (зашумленной или неполной) копии S_0. Память с такими свойствами процесса считывания информации является ассоциативной. При поиске искаженные части целого восстанавливаются по имеющимся неискаженным частям на основе ассоциативных связей между ними.

Ассоциативный характер памяти сети Хопфилда качественно отличает ее от обычной, адресной, компьютерной памяти. В последней извлечение необходимой информации происходит по адресу ее начальной точки (ячейки памяти). Потеря адреса (или даже одного бита адреса) приводит к потере доступа ко всему информационному фрагменту. При использовании же ассоциативной памяти доступ к информации производится непосредственно по ее содержанию, т.е. по частично известным искаженным фрагментам. Потеря части информации или ее зашумление не приводит к катастрофическому ограничению доступа, если оставшейся информации достаточно для извлечения идеального образа.

Поиск идеального образа по имеющейся неполной или зашумленной его версии называется задачей распознавания образов. В нашей лекции особенности решения этой задачи нейронной сетью Хопфилда будут продемонстрированы на примерах, которые получены с использованием модели сети на персональной ЭВМ.

В рассматриваемой модели сеть содержала 100 нейронов, упорядоченных в матрицу 10\times 10. Сеть обучалась по правилу Хебба на трех идеальных образах — шрифтовых начертаниях латинских букв M, A и G (см. [рис. 8.3](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.3)). После обучения нейросети в качестве начальных состояний нейронов предъявлялись различные искаженные версии образов, которые в дальнейшем эволюционировали с последовательной динамикой к стационарным состояниям.

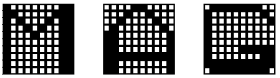


Рис. 8.3.

Для каждой пары изображений на [рисунке 8.4](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4), левый образ является начальным состоянием, а правый — результатом работы сети, достигнутым стационарным состоянием.

Образ на [рис. 8.4(А)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) был выбран для тестирования адекватности поведения на идеальной задаче, когда предъявленное изображение точно соответствует информации в памяти. В этом случае за один шаг было достигнуто стационарное состояние. Образ на [рис. 8.4(Б)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) характерен для задач распознавания текста независимо от типа шрифта. Начальное и конечное изображения безусловно похожи, но попробуйте это объяснить машине!

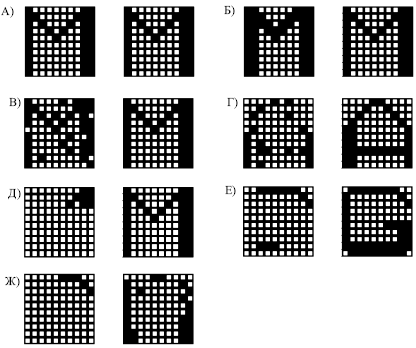


Рис. 8.4.

Задания на [рис. 8.4(В, Г)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) характерны для практических приложений. Нейросетевая система способна распознавать практически полностью зашумленные образы. Задачи, соответствующие рисункам [8.4(Д, Е)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4), демонстрируют замечательное свойство сети Хопфилда: она способна ассоциативно узнавать образ по его небольшому фрагменту. Важнейшей особенностью работы сети является генерация ложных образов. Пример ассоциации к ложному образу показан на [рис. 8.4(Ж)](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4). Ложный образ является устойчивым локальным экстремумом энергии, но не соответствует никакому идеальному образу. Он является в некотором смысле собирательным образом, наследующим черты идеальных собратьев. Ситуация с ложным образом эквивалентна нашему "Где-то я уже это видел".

В данной простейшей задаче ложный образ является "неверным" решением и поэтому вреден. Однако можно надеяться, что такая склонность сети к обобщениям может быть как-то использована. Характерно, что при увеличении объема полезной информации (сравните [рис. 8.4 (Е) и (Ж))](https://www.intuit.ru/studies/courses/88/88/lecture/20541?page=5#image.8.4) исходное состояние попадает в область притяжения требуемого стационарного состояния, и образ распознается.