

# Курс лекций «Введение в ИИ». Часть III.

## Нейрокомпьютеринг.

### Лекция 20. Когнитрон

О.Г. Чанышев

## Содержание

<b>1 Введение</b>	<b>1</b>
<b>2 Когнитрон</b>	<b>2</b>
2.1 Структура . . . . .	2
2.2 Обучение . . . . .	2
2.3 Возбуждающий нейрон . . . . .	3
2.4 Тормозящие нейроны . . . . .	4
2.5 Процедура обучения . . . . .	4
2.6 Латеральное торможение . . . . .	5
2.7 Рецептивная область . . . . .	6
2.8 Результаты моделирования . . . . .	7

## 1 Введение

Сходство и различия образов, являющиеся очевидными для человека, пока ставят в тупик даже наиболее сложные компьютерные системы распознавания. Таким образом, бесчисленное количество важных приложений, в которых компьютеры могут заменить людей в опасных, скучных или неприятных работах, остаются за пределами их текущих возможностей.

Кажется вполне логичным вернуться к биологическим моделям и попытаться определить, каким образом они функционируют так хорошо. Очевидно, что это трудно сделать по нескольким причинам. Во-первых, сверхвысокая сложность человеческого мозга затрудняет понимание принципов его функционирования. Трудно понять общие принципы функционирования и взаимодействия его приблизительно  $10^{11}$  нейронов и  $10^{14}$  синаптических связей. Кроме того, существует множество проблем при проведении экспериментальных исследований.

Несмотря на ограничения, многое было изучено благодаря блестяще задуманным экспериментам. Например, в [1] описан эксперимент, в котором котята выращивались в визуальном окружении, состоящем только из горизонтальных черных и белых полос. Известно, что определенные области коры чувствительны к углу ориентации, поэтому у этих котят не развились нейроны, чувствительные к вертикальным полосам. Этот результат наводит на мысль, что мозг млекопитающих не является полностью <предустановленным> даже на примитивном уровне распознавания ориентации линий. Напротив, он постоянно самоорганизуется, основываясь на опыте.

Рис. 1: Пресинаптические и постсинаптические нейроны

Рис. 2: Область связей нейрона

На микроскопическом уровне обнаружено, что нейроны обладают как возбуждающими, так и тормозящими синапсами. Первые стремятся к возбуждению нейрона; последние подавляют его возбуждение (см. приложение А). Это наводит на мысль, что мозг адаптируется либо изменением воздействия этих синапсов, либо созданием или разрушением синапсов в результате воздействия окружающей среды. Данное предположение остается пока гипотезой с ограниченным физиологическим подтверждением. Однако исследования, проведенные в рамках этой гипотезы, привели к созданию цифровых моделей, некоторые из которых показывают замечательные способности к адаптивному распознаванию образов.

## 2 Когнитрон

Основываясь на текущих знаниях анатомии и физиологии мозга, в работе [2] разработан когнитрон, гипотетическая модель системы восприятия человека. Компьютерные модели, исследованные в [2], продемонстрировали впечатляющие способности адаптивного распознавания образов, побуждая физиологов исследовать соответствующие механизмы мозга. Это взаимно усиливающее взаимодействие между искусственными нейронными сетями, физиологией и психологией может оказаться средством, посредством которого будет со временем достигнуто понимание механизмов мозга.

### 2.1 Структура

Когнитрон конструируется в виде слоев нейронов, соединенных синапсами. Как показано на рис. 1, пресинаптический нейрон в одном слое связан с постсинаптическим нейроном в следующем слое. Имеются два типа нейронов: возбуждающие узлы, которые стремятся вызвать возбуждение постсинаптического узла, и тормозящие узлы, которые тормозят это возбуждение. Возбуждение нейрона определяется взвешенной суммой его возбуждающих и тормозящих входов, однако в действительности механизм является более сложным, чем простое суммирование.

На рис. 2 показано, что каждый нейрон связан только с нейронами в соседней области, называемой областью связи. Это ограничение области связи согласуется с анатомией зрительной коры, в которой редко соединяются между собой нейроны, располагающиеся друг от друга на расстоянии более одного миллиметра. В рассматриваемой модели нейроны упорядочены в виде слоев со связями от одного слоя к следующему. Это также аналогично послойной структуре зрительной коры и других частей головного мозга.

### 2.2 Обучение

Так как когнитрон реализован в виде многослойной сети, возникают сложные проблемы обучения, связанные с выбранной структурой. Автор отверг управляемое обучение, как биологически неправдоподобное, используя взамен этого обучение без учителя. Получая обучающий набор входных образов, сеть самоорганизуется посредством изменения силы синаптических связей.

Рис. 3: Область связи с областью конкуренции

При этом отсутствуют предварительно определенные выходные образы, представляющие требуемую реакцию сети, однако сеть самонастраивается с целью распознавания входных образов с замечательной точностью.

Алгоритм обучения когнитрона является концептуально привлекательным. В заданной области слоя обучается только наиболее сильно возбужденный нейрон. Автор сравнивает это с <элитным обучением>, при котором обучаются только <умные> элементы. Те нейроны, которые уже хорошо обучены, что выражается силой их возбуждения, получают приращение силы своих синапсов с целью дальнейшего усиления своего возбуждения.

На рис. 3 показано, что области связи соседних узлов значительно перекрываются. Это расточительное дублирование функций оправдывается взаимной конкуренцией между ближайшими узлами. Даже если узлы в начальный момент имеют абсолютно идентичный выход, небольшие отклонения всегда имеют место; один из узлов всегда будет иметь более сильную реакцию на входной образ, чем соседние. Его сильное возбуждение будет оказывать сдерживающее воздействие на возбуждение соседних узлов, и только его синапсы будут усиливаться; синапсы соседних узлов останутся неизменными.

### 2.3 Возбуждающий нейрон

Грубо говоря, выход возбуждающего нейрона в когнитроне определяется отношением его возбуждающих входов к тормозящим входам. Эта необычная функция имеет важные преимущества, как практические, так и теоретические.

Суммарный возбуждающий вход в нейрон является взвешенной суммой входов от возбуждающих предшествующем слое. Аналогично суммарный тормозящий вход является взвешенной суммой входов от всех тормозящих нейронов. В символьном виде

$$E = \sum_i a_i u_i, \quad I = \sum_j b_j \nu_j$$

где  $a_i$  - вес  $i$ -го возбуждающего синапса,  $u_i$  - выход  $i$ -го возбуждающего нейрона,  $b_j$  - вес  $j$ -го тормозящего синапса,  $\nu_j$  - выход  $j$ -го тормозящего нейрона. Заметим, что веса имеют только положительные значения. Выход нейрона затем вычисляется следующим образом:

$$NET = \frac{1 + E}{1 + I} - 1$$

$$OUT = NET, \quad NET \geq 0,$$

$$OUT = 0, \quad NET < 0.$$

Предполагая, что NET имеет положительное значение, это можно записать следующим образом:

$$OUT = \frac{E - I}{1 + I}$$

Когда тормозящий вход мал ( $I \ll 1$ ), OUT может быть аппроксимировано как

$$OUT \approx E - I$$

Рис. 4: Слои когнитрона

что соответствует выражению для обычного линейного порогового элемента (с нулевым порогом).

Алгоритм обучения когнитрона позволяет весам синапсов возрастать без ограничений. Благодаря отсутствию механизма уменьшения весов они просто возрастают в процессе обучения. В обычных линейных пороговых элементах это привело бы к произвольно большому выходу элемента. В когнитроне большие возбуждающие и тормозящие входы результируются в ограничивающей формуле вида:

$$OUT \approx \frac{E}{I} - 1 \text{ если } E \gg 1 \text{ и } I \gg 1.$$

В данном случае OUT определяется отношением возбуждающих входов к тормозящим входам, а не их разностью. Таким образом, величина OUT ограничивается, если оба входа возрастают в одном и том же диапазоне  $X$ . Предположив, что это так,  $E$  и  $I$  можно выразить следующим образом:

$$E = px, I = qx, x \in X, p, q - \text{ константы,}$$

и после некоторых преобразований

$$OUT = \frac{p-q}{2q} \left[ 1 + th\left(\frac{\log(pq)}{2}\right) \right].$$

Эта функция возрастает по закону Вебера-Фехнера, который часто используется в нейрофизиологии для аппроксимации нелинейных соотношений входа/выхода сенсорных нейронов. При использовании этого соотношения нейрон когнитрона в точности эмулирует реакцию биологических нейронов. Это делает его как мощным вычислительным элементом, так и точной моделью для физиологического моделирования.

## 2.4 Тормозящие нейроны

В когнитроне слой состоит из возбуждающих и тормозящих узлов. Как показано на рис. 4, нейрон слоя 2 имеет область связи, для которой он имеет синаптические соединения с набором выходов нейронов в слое 1. Аналогично в слое 1 существует тормозящий нейрон, имеющий ту же область связи. **Синаптические веса тормозящих узлов не изменяются в процессе обучения; их веса заранее установлены таким образом, что сумма весов в любом из тормозящих нейронов равна единице.** В соответствии с этими ограничениями, выход тормозящего узла INHIB является взвешенной суммой его входов, которые в данном случае представляют собой среднее арифметическое выходов возбуждающих нейронов, к которым он подсоединен. Таким образом,

$$INHIB = \sum_i c_i OUT_i,$$

где  $\sum_i c_i = 1$ ,  $c_i$  - возбуждающий вес  $i$ .

## 2.5 Процедура обучения

Как объяснялось ранее, веса возбуждающих нейронов изменяются только тогда, когда нейрон возбужден сильнее, чем любой из узлов в области конкуренции. Если это так, изменение в процессе обучения любого из его весов может быть определено следующим образом:

$$\delta a_i = q c_j u_j,$$

где  $c_j$  - тормозящий вес связи нейрона  $j$  в слое 1 с тормозящим нейроном  $i$ , и  $u_j$  - выход нейрона  $j$  в слое 1,  $a_i$  - возбуждающий вес  $i$ ,  $q$  - нормирующий коэффициент обучения.

Изменение тормозящих весов нейрона  $i$  в слое 2 пропорционально отношению взвешенной суммы возбуждающих входов к удвоенному тормозящему входу. Вычисления проводятся по формуле

$$\delta b_i = \frac{q \sum_j a_j u_j}{2 \times INHIB_i}.$$

Когда возбужденных нейронов в области конкуренции нет, для изменения весов используются другие выражения. Это необходимо, поскольку процесс обучения начинается с нулевыми значениями весов; поэтому первоначально нет возбужденных нейронов ни в одной области конкуренции, и обучение производиться не может. Во всех случаях, когда победителя в области конкуренции нейронов нет, изменение весов нейронов вычисляется следующим образом:

$$\delta a_i = q' c_j u_j, \delta b_i = q' INHIB_i,$$

где  $q'$  - положительный обучающий коэффициент меньший, чем  $q$ .

Приведенная стратегия настройки гарантирует, что узлы с большой реакцией заставляют возбуждающие синапсы, которыми они управляют, увеличиваться сильнее, чем тормозящие синапсы. И наоборот, узлы, имеющие малую реакцию, вызывают малое возрастание возбуждающих синапсов, но большее возрастание тормозящих синапсов. Таким образом, если узел 1 в слое 1 имеет большой выход, синапс  $a_1$  возрастет больше, чем синапс  $b_1$ . И наоборот, узлы, имеющие малый выход, обеспечат малую величину для приращения  $a_i$ . Однако другие узлы в области связи будут возбуждаться, тем самым увеличивая сигнал  $INHIB$  и значения  $b_i$ .

В процессе обучения веса каждого узла в слое 2 настраиваются таким образом, что вместе они составляют шаблон, соответствующий образам, которые часто предъявляются в процессе обучения. При предъявлении сходного образа шаблон соответствует ему и узел вырабатывает большой выходной сигнал. Сильно отличающийся образ вырабатывает малый выход и обычно подавляется конкуренцией.

## 2.6 Латеральное торможение

На рис. 10.4 показано, что каждый нейрон слоя 2 получает латеральное торможение от нейронов, расположенных в его области конкуренции. Тормозящий нейрон суммирует входы от всех нейронов в области конкуренции и вырабатывает сигнал, стремящийся к торможению целевого нейрона. Этот метод является эффективным, но с вычислительной точки зрения медленным. Он охватывает большую систему с обратной связью, включающую каждый нейрон в слое; для его стабилизации может потребоваться большое количество вычислительных итераций.

Для ускорения вычислений в работе [2] используется остроумный метод ускоренного латерального торможения (рис. 5). Здесь дополнительный узел латерального торможения обрабатывает выход каждого возбуждающего узла для моделирования требуемого латерального торможения. Сначала он определяет сигнал, равный суммарному тормозящему влиянию в области конкуренции:

Рис. 5: Ускоренное торможение

$$LAT\_INHIB = \sum_i g_i OUT_i,$$

где  $OUT_i$  - выход  $i$ -го нейрона в области конкуренции,  $g_i$  - вес связи от этого нейрона к латерально-тормозящему нейрону;  $g_i$  выбраны таким образом, что

$$\sum_i g_i = 1.$$

Выход тормозящего нейрона  $OUT'$  затем вычисляется следующим образом:

$$OUT' = \frac{1 + OUT_i}{1 + LAT\_INHIB} - 1$$

Благодаря тому что все вычисления, связанные с таким типом латерального торможения, являются нерекурсивными, они могут быть проведены за один проход для слоя, тем самым определяя эффект в виде большой экономии в вычислениях.

Этот метод латерального торможения решает и другую сложную проблему. Предположим, что узел в слое 2 возбуждается сильно, но возбуждение соседних узлов уменьшается постепенно с увеличением расстояния. При использовании обычного латерального торможения будет обучаться только центральный узел. Другие узлы определяют, что центральный узел в их области конкуренции имеет более высокий выход. С предлагаемой системой латерального торможения такой ситуации случиться не может. Множество узлов может обучаться одновременно и процесс обучения является более достоверным.

## 2.7 Рецептивная область

Анализ, проводимый до этого момента, был упрощен рассмотрением только одномерных слоев. В действительности когнитрон конструировался как каскад двумерных слоев, причем в данном слое каждый нейрон получает входы от набора нейронов на части двумерного плана, составляющей его область связи в предыдущем слое.

С этой точки зрения когнитрон организован подобно зрительной коре человека, представляющей собой трехмерную структуру, состоящую из нескольких различных слоев. Оказывается, что каждый слой мозга реализует различные уровни обобщения; входной слой чувствителен к простым образам, таким, как линии, и их ориентации в определенных областях визуальной области, в то время как реакция других слоев является более сложной, абстрактной и независимой от позиции образа.

Аналогичные функции реализованы в когнитроне путем моделирования организации зрительной коры. На рис. 6 показано, что нейроны когнитрона в слое 2 реагируют на определенную небольшую область входного слоя 1. Нейрон в слое 3 связан с набором нейронов слоя 2, тем самым реагируя косвенно на более широкий набор нейронов слоя 1. Подобным образом нейроны в последующих слоях чувствительны к более широким областям входного образа до тех пор, пока в выходном слое каждый нейрон не станет реагировать на все входное поле.

Если область связи нейронов имеет постоянный размер во всех слоях, требуется большое количество слоев для перекрытия всего входного поля выходными нейронами. Количество слоев может быть уменьшено путем расширения области связи в последующих слоях. К сожалению,

Рис. 6: Области связей когнитрона

Рис. 7: Результаты экспериментов с когнитроном

результатом этого может явиться настолько большое перекрытие областей связи, что нейроны выходного слоя будут иметь одинаковую реакцию. Для решения этой проблемы может быть использовано расширение области конкуренции. Так как в данной области конкуренции может возбудиться только один узел, влияние малой разницы в реакциях нейронов выходного слоя усиливается.

В альтернативном варианте связи с предыдущим слоем могут быть распределены вероятно с большинством синаптических связей в ограниченной области и с более длинными соединениями, встречающимися намного реже. Это отражает вероятностное распределение нейронов, обнаруженное в мозге. В когнитроне это позволяет каждому нейрону выходного слоя реагировать на полное входное поле при наличии ограниченного количества слоев.

## 2.8 Результаты моделирования

В [3] описываются результаты компьютерного моделирования четырехслойного когнитрона, предназначенного для целей распознавания образов. Каждый слой состоит из массива  $12 \times 12$  возбуждающих нейронов и такого же количества тормозящих нейронов. Область связи представляет собой квадрат, включающий  $5 \times 5$  нейронов. Область конкуренции имеет форму ромба высотой и шириной в пять нейронов. Латеральное торможение охватывает область  $7 \times 7$  нейронов. Нормирующие параметры обучения установлены таким образом, что  $q = 16,0$  и  $q' = 2,0$ . Веса синапсов проинициализированы в 0.

Сеть обучалась путем предъявления пяти стимулирующих образов, представляющих собой изображения арабских цифр от 0 до 4, на входном слое. Веса сети настраивались после предъявления каждой цифры, входной набор подавался на вход сети циклически до тех пор, пока каждый образ не был предъявлен суммарно 20 раз.

Эффективность процесса обучения оценивалась путем запуска сети в реверсивном режиме; выходные образы, являющиеся реакцией сети, подавались на выходные нейроны и распространялись обратно к входному слою. Образы, полученные во входном слое, затем сравнивались с исходным входным образом. Чтобы сделать это, обычные однонаправленные связи принимались проводящими в обратном направлении и латеральное торможение отключалось. На рис. 7 показаны типичные результаты тестирования. В столбце 2 показаны образы, произведенные каждой цифрой на выходе сети. Эти образы возвращались обратно, вырабатывая на входе сети образ, близкий к точной копии исходного входного образа. Для столбца 4 на выход сети подавался только выход нейрона, имеющего максимальное возбуждение. Результирующие образы в точности те же, что и в случае подачи полного выходного образа, за исключением цифры 0, для которой узел с максимальным выходом располагался на периферии и не покрывал полностью входного поля.

## Список литературы

- [1] Blakemore C., Cooper G. F. 1970. Development of the brain depends on the visual environment. Nature 228(5270):477-78.



- [2] Fukushima K. 1975. Cognitron: A self-organizing multilayered neural network. *Biological Cybernetics* 20:121-36.
- [3] Fukushima K. 1981. Cognitron: A self-organizing multilayer neural network model. NHK Technical Monograph No. 30, pp. 1-25. Available from Nippon Hoso Kyokai (Japanese Broadcasting Corp.), Technical Research Labs, Tokio, Japan.

## **Следующая лекция**

*Лекция 21. От информационного пространства к пространству знаний. Онтологии в Интернет.*