

# Когнитивная наука

2007/2008

[900igr.net](http://900igr.net)

Материалы к курсу

М.В. Фаликман:

<http://virtualcoglab.cs.msu.su>



# Что это такое?

область междисциплинарных исследований  
познания, понимаемого как совокупность  
процессов приобретения, хранения,  
преобразования и использования знаний  
живыми и искусственными системами



# ОСНОВНЫЕ ДИСЦИПЛИНЫ

Экспериментальная психология познания

Лингвистика

Компьютерные науки, кибернетика, искусственный  
интеллект

Нейробиология

Философия познания (Гносеология)

Антропология



# ПРОБЛЕМЫ МЕЖДИСЦИПЛИНАРНОГО ВЗАИМОДЕЙСТВИЯ

- единый («общепринятый») язык;

## Что мешает договориться?

От «научной омонимии»

*(«Депрессия альфа-ритма, говорите? А Вы антидепрессанты  
вводить не пробовали? »)*

до концептуальных разногласий

*(Активность мозга **определяет** психические процессы или  
**обеспечивает** их протекание?)*



# Общие допущения:

Познание = «обработка информации» =  
(1) *представление* знаний + (2) вычислительные  
операции по их *преобразованию*

Мозг - вычислительное устройство  
(«суперкомпьютер»), осуществляющее операции  
по преобразованию структур, посредством  
которых представлены знания



# Методология и методы КОГНИТИВНОЙ науки

Методология -- «обратная инженерия»  
(Дэниэл Деннетт).

Методы:

1. Частные

2. Междисциплинарные

- компьютерное моделирование
- функциональное картирование мозга (?)



# «Слабые звенья»

- мотивационно-эмоциональная регуляция познания
- социальная природа человеческого познания
- познание и телесность
- мозг как вычислительное устройство

...

Зоны роста когнитивной науки в XXI  
веке?



# РОЖДЕНИЕ КОГНИТИВНОЙ НАУКИ

«Когнитивная  
**контр**революция»  
в США

«Три кита» в Европе:

- Фредерик Чарлз Бартлетт (1886-1969)
- Жан Пиаже (1896-1980)
- Александр Романович Лурия (1902-1977)

# РОЖДЕНИЕ КОГНИТИВНОЙ НАУКИ

MIT (Кембридж, Массачусетс), 11 сентября 1956 года

- Ноэм Хомский «Три модели языка»
- Джордж Миллер «Магическое число  $7 \pm 2$ »
- Аллен Ньюэлл, Герберт Саймон «Логик-теоретик»



# РОЖДЕНИЕ КОГНИТИВНОЙ НАУКИ

Джордж Миллер: “ ...Я уходил с Симпозиума с твердой уверенностью, скорее интуитивной, чем рациональной, в том, что экспериментальная психология человека, теоретическая лингвистика и компьютерное моделирование познавательных процессов – части еще большего целого, и в будущем мы увидим последовательную разработку и координацию их общих дел... Я двигался навстречу когнитивной науке в течение двадцати лет, прежде чем узнал, как она называется...” (см. Миллер Дж. «Когнитивная революция с исторической точки зрения» // Вопросы психологии, 2005, №6, с.104-109)



# Продолжение следует...

1957 -- группа искусственного интеллекта в MIT  
(Марвин Минский, Джон Маккарти)

1960 -- Центр когнитивных исследований в Гарварде  
(Джером Брунер, Джордж Миллер)

1976/77 -- журнал «Когнитивная наука»

1979 -- Общество когнитивной науки (Cognitive Science Society, Inc.), Массачусетс: Д. Норман, Р. Шенк и др.

1979 -- Первая конференция по когнитивной науке, Ла Хойя, Калифорния

1981 -- широкомасштабное финансирование в США  
(Sloan Foundation), университетские программы

# А у нас?

психология познавательных процессов

нейронаука ([www.neuroscience.ru](http://www.neuroscience.ru))

искусственный интеллект ([www.raii.org](http://www.raii.org))

прикладная и компьютерная лингвистика

...

2002 -- Московский семинар по когнитивной науке  
(очередная встреча -- 9 марта 2006 г., 18:30)

2003 -- Первая российская Интернет-конференция по  
когнитивной науке ([Auditorium.ru](http://Auditorium.ru)),  
10 февраля - 10 апреля



# Продолжение следует...

Октябрь 2004, Казанский университет -- Первая  
российская конференция по когнитивной науке

Борис Митрофанович Величковский  
(Москва-Дрезден-Москва)



Валерий Дмитриевич Соловьев  
(Казань)



# Продолжение следует...

Июнь 2006, Санкт-Петербургский университет -- Вторая  
российская конференция по когнитивной науке

Подробнее см. <http://www.cogsci.ru> --  
сайт Российской Ассоциации  
Когнитивных Исследований  
(создана в 2004 г.)



# Компьютерная метафора познания

- Специалист подобен флюсу
- Человеческий мозг подобен компьютеру





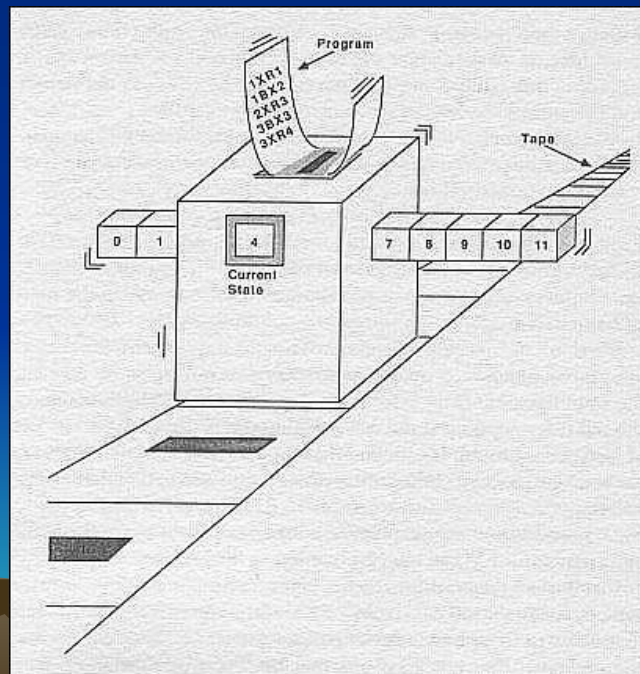
# ОСНОВНЫЕ ПОДХОДЫ В КОГНИТИВНОЙ НАУКЕ

## Часть 1. Символьный подход



# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

Машина Тьюринга:  
принципы обработки информации



# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

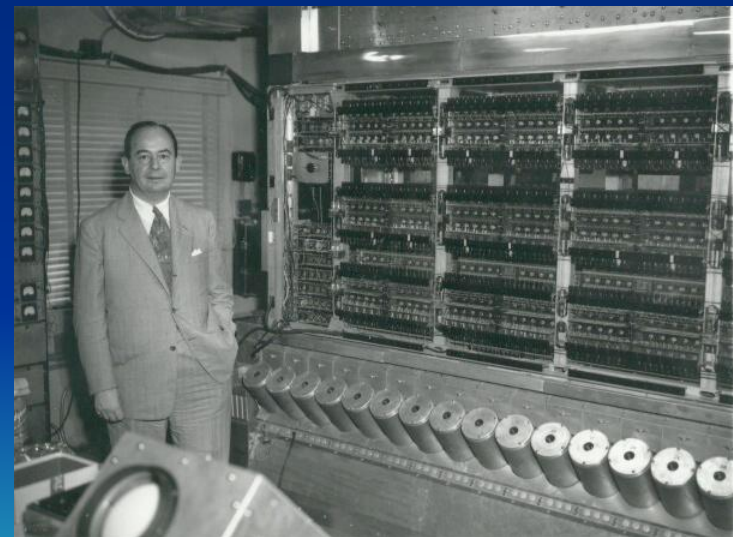
Основные принципы архитектуры компьютера:  
Джон/Янош фон Нейман (1903-1957)



# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

Основные принципы архитектуры компьютера:

- Периферические устройства ввода-вывода;
- центральный процессор;
- оперативное запоминающее устройство;
- постоянное запоминающее устройство.



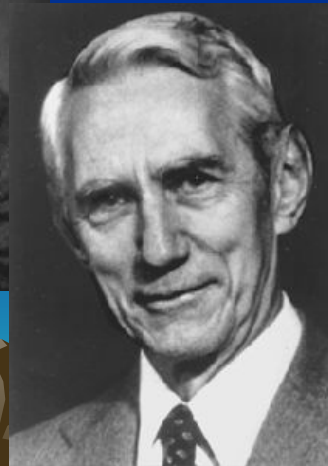
# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

Принципиальная архитектура познания:

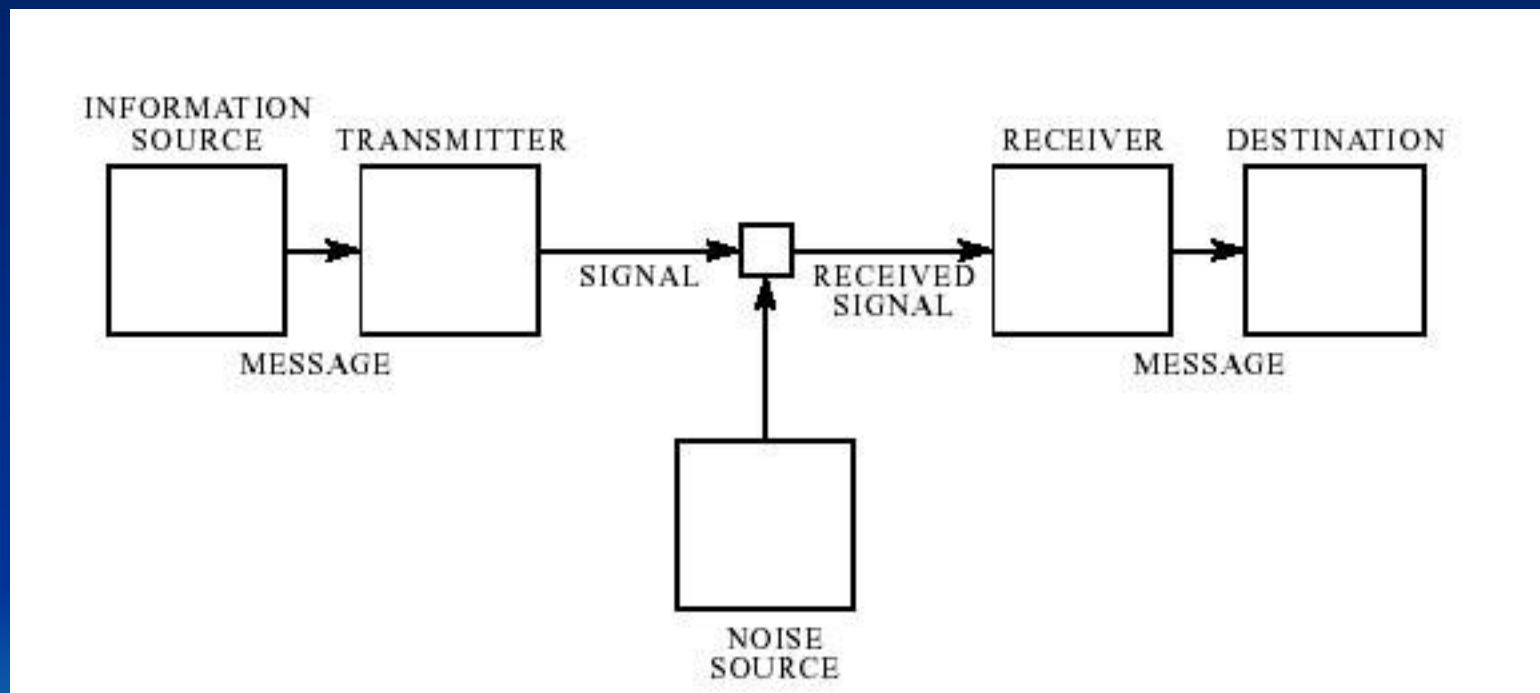
- Периферические устройства ввода-вывода;
- центральный процессор;
- оперативное запоминающее устройство;
- постоянное запоминающее устройство.
- Сенсорные и моторные системы;
- «центральный процессор»;
- кратковременная (рабочая) память;
- долговременная память.

# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

Теория информации и теория коммуникации:  
Клод Элвуд Шеннон (1916-2001)

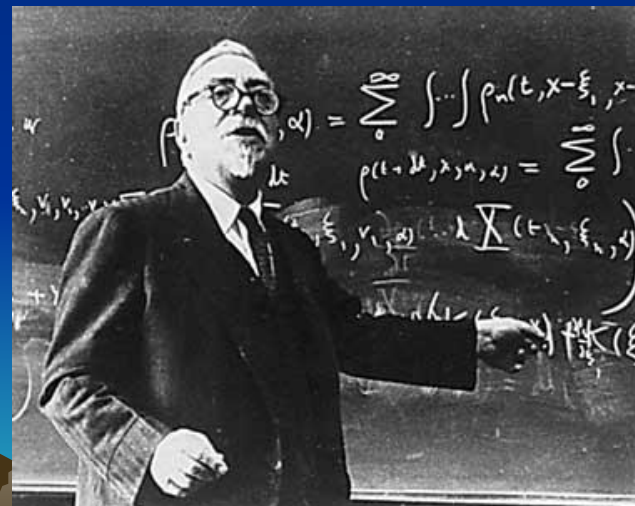


# Модель передачи информации: Клод Элвуд Шеннон



# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

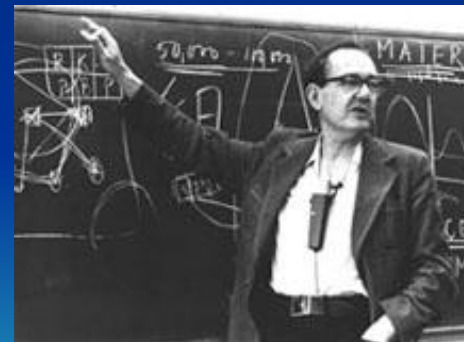
Кибернетика, или теория управления:  
Норберт Винер (1894-1964)





# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

Аллен Ньюэлл, Герберт Саймон  
«Логик-теоретик»  
«Универсальный решатель задач»



# ОБЩИЙ РЕШАТЕЛЬ ЗАДАЧ

Мышление -- преобразование символов и  
символьных систем по определенным *правилам*.

Алгоритм

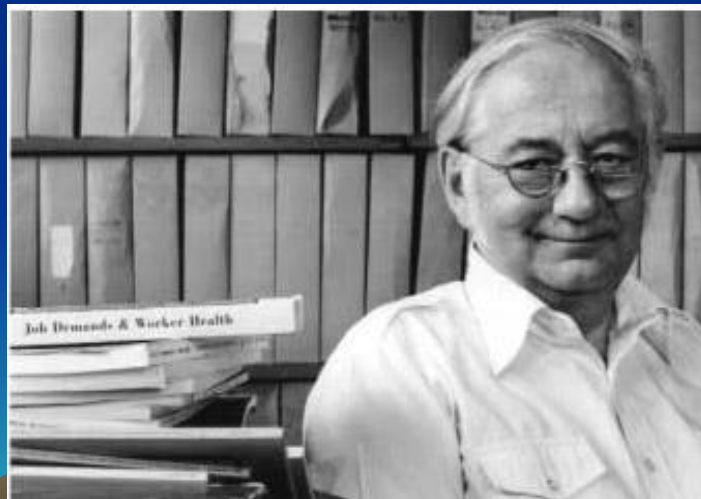


Эвристика

А. Ньюэлл, Дж. Шоу, Г. Саймон «Моделирование мышления  
человека с помощью электронно-вычислительной машины» //  
Хрестоматия по психологии мышления. М.: 1981. С. 305-327.

# НА ЗАРЕ КОГНИТИВНОЙ НАУКИ: СИМВОЛЬНЫЙ ПОДХОД

Дональд Эрик Бродбент (1926-1993)  
модель переработки информации



# ОСНОВНЫЕ ДОПУЩЕНИЯ:

Познание -- переработка информации

Линейный характер переработки:  
последовательный ряд блоков  
от входа до выхода

Блок/канал с ограниченной  
пропускной способностью



# ОСНОВНЫЕ ДОПУЩЕНИЯ:

В системе переработки информации  
должен быть защитный фильтр --  
механизм ВНИМАНИЯ:



Теории внимания как *отбора*:

Э.М. Трейсмэн  
Д. и Дж.Э. Дойч  
Д. Норман

...

# Модели языка:

- Вероятностная (стохастическая) модель Дж. Миллера
- Теория трансформационных грамматик Н. Хомского

## Общее допущение:

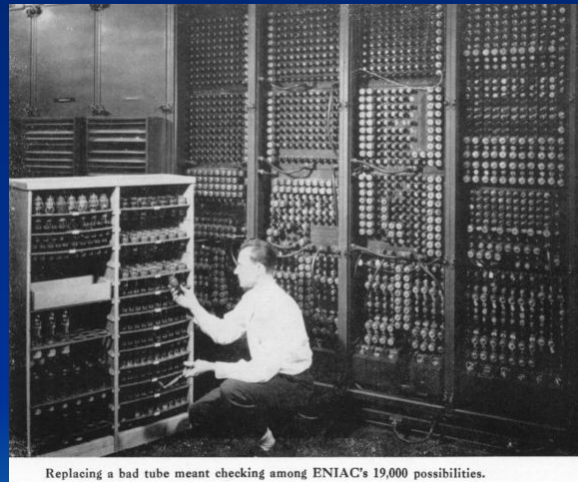
понимание и порождение речи как преобразование символов и их систем (словарных единиц и грамматических конструкций) по определенным правилам

# Модели памяти:

- Кратковременная память:  $7 \pm 2$  ячейки
- Теория двойственности памяти («постоянное запоминающее устройство» и «оперативное запоминающее устройство»)
- Трехкомпонентная теория памяти (сенсорный регистр -- «буфер» Дональда Бродбента -- и те же системы).

# СИМВОЛЬНЫЙ ПОДХОД К ПОЗНАНИЮ

Познавательные процессы ~ переработка  
символьной информации компьютером



Психика -- «универсальное перерабатывающее  
устройство»



# СИМВОЛЬНЫЙ ПОДХОД К ПОЗНАНИЮ

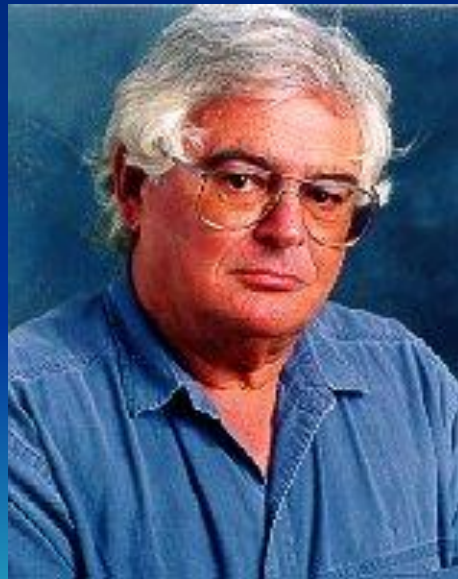
Развитие вычислительной техники :  
от «вычислений вообще» к частным задачам

Появление специализированных  
*«микропроцессоров»*  
в пределах одной архитектуры  
(видеокарта, звуковая карта, управление  
внешними устройствами и т.д.)



# РОЖДЕНИЕ МОДУЛЬНОГО ПОДХОДА К ПОЗНАНИЮ

1983 -- Джерри Фодор, «Модульность психики»  
(*The Modularity of Mind*)

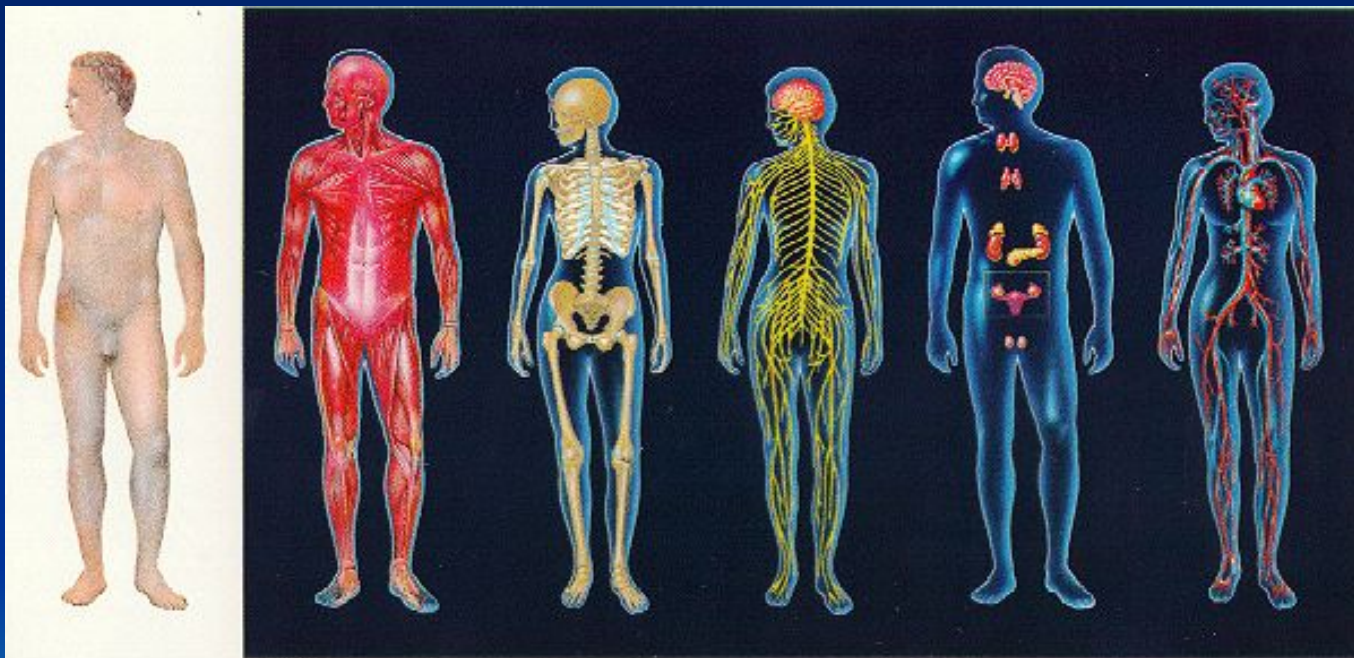


# МОДУЛЬНЫЙ ПОДХОД К ПОЗНАНИЮ



(с) Леда Космидес, Джон Туби

# МОДУЛЬНЫЙ ПОДХОД К ПОЗНАНИЮ

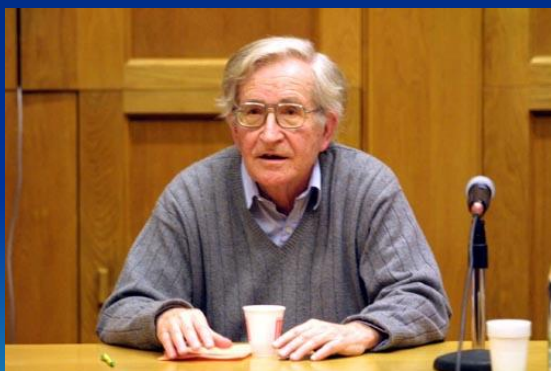


СИСТЕМЫ БИОЛОГИЧЕСКОГО ОРГАНИЗМА

# ОСНОВЫ МОДУЛЬНОГО ПОДХОДА К ПОЗНАНИЮ

**Нейропсихология XIX века:**

речь может нарушаться при сохранности прочих функций (зона Брока, зона Вернике)



**Нозм Хомский (1988):**

врожденность языковой способности и ее независимость от других способностей -- язык как отдельный «умственный орган»

# ОСНОВЫ МОДУЛЬНОГО ПОДХОДА К ПОЗНАНИЮ

**Идея модульности познания  
-- Дэвид Марр (1945-1980):**

«Любой большой массив вычислений  
должен быть разбит и реализован как  
набор частей, независимых друг от  
друга настолько, насколько это  
допускает общая задача...» (1976)



# РОЖДЕНИЕ МОДУЛЬНОГО ПОДХОДА К ПОЗНАНИЮ

Джерри Фодор (1983):

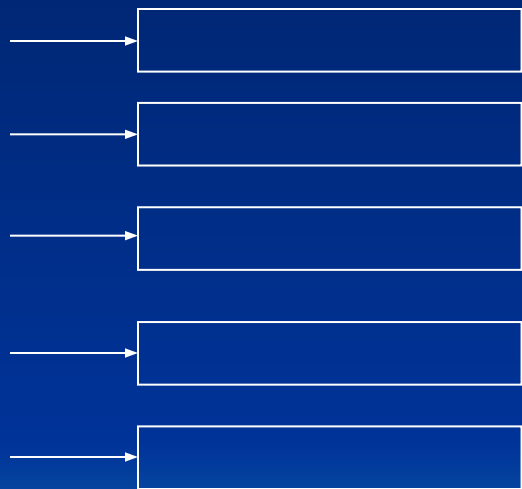
общая концепция «модульности»:  
познание как мозаика  
специализированных *модулей*

Насколько этот принцип универсален?



# МОДУЛЬНЫЙ ПОДХОД К ПОЗНАНИЮ

Когнитивная архитектура:



Модульные  
системы ввода



Центральные системы:  
планирование,  
принятие решения



# КРИТЕРИИ ВЫДЕЛЕНИЯ МОДУЛЕЙ

## 1. Особая сфера влияния, или специализация (*domain specificity*):

каждый модуль компетентен в обработке одного из видов информации или в решении одного из классов познавательных задач и не участвует в решении других классов задач



# ЕЩЕ КРИТЕРИИ ВЫДЕЛЕНИЯ МОДУЛЕЙ

7. Закономерное разворачивание в онтогенезе: ряд последовательных ступеней (собственная «история развития»)
8. Локализация в мозге: специфические нервные механизмы
9. Избирательное нарушение: выпадение модуля не сказывается на работе других модулей (пример: лицевая агнозия)

Прямое следствие --  
**УЗКАЯ СПЕЦИАЛИЗАЦИЯ МОДУЛЯ**

# Насколько речь модульна?

Элизабет Бейтс  
(1947-2003)

- Недавнее появление в филогенезе
- Пластичность поведенческих проявлений
- Пластичность нервных механизмов
- Произвольность связей между обозначением и обозначаемым



Синдром Уильямса,  
SLI, афазии у  
взрослых:  
неоднозначность  
проявлений!

# ГИПОТЕЗА ВСЕОБЩЕЙ МОДУЛЬНОСТИ



Дэн Спербер:

познание полностью модульно -- так же,  
как биологический организм.  
Неспециализированных систем  
переработки информации,  
использующих обобщенный  
«умственный лексикон», НЕТ.

# КРИТИКА МОДУЛЬНОГО ПОДХОДА

## 1. Теоретическая:

- проблема обучения и пластичности познания;
- влияние культуры на «модульные» процессы (иллюзия Мюллера-Лайера в «круглом» мире);
- проблема нисходящей регуляции решения познавательных задач.

*За пределами рассмотрения: взаимодействие модулей!*

Примеры: две системы зрительного восприятия; феномен отчуждения руки, etc.

# КРИТИКА МОДУЛЬНОГО ПОДХОДА

## 2. Эмпирическая:

- двойные диссоциации внутри двойных диссоциаций -- Аннет Кармилофф-Смит и др. (Оксфорд);
- развитие речи и ранние локальные поражения головного мозга: возможности компенсации (Элизабет Бейтс и др.);
- «ген грамматики» FоxP2: исследования экспрессии гена у человека и животных.



**ВЫВОД:** познавательные процессы не обусловлены наследственностью настолько, как того хотелось бы представителям модульного подхода...

*Адекватная модель?*

Неспециализированная обучаемая система!

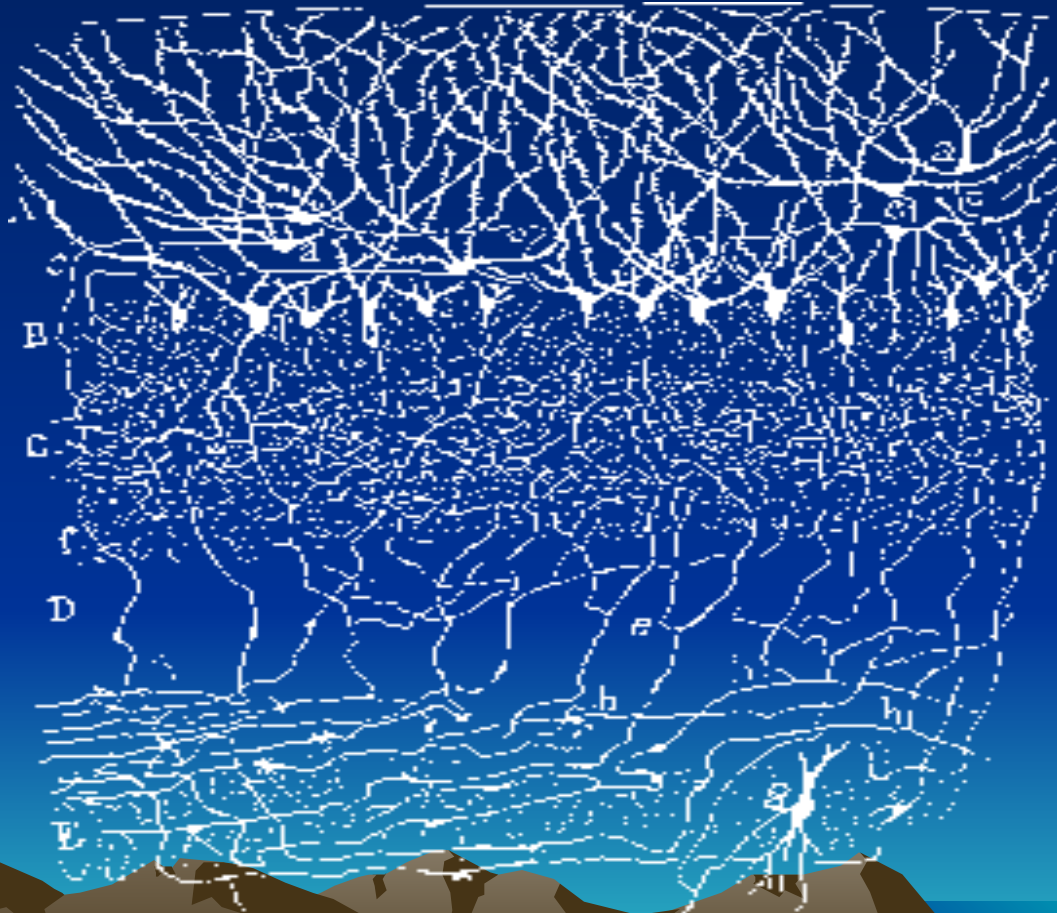


Представление и  
приобретение знаний:  
есть ли альтернатива  
компьютерной метафоре?





# Нейронные сети: основные положения



# Мозг человека: преимущества перед компьютером

- $10^{11}$  нейронов,  $10^{14}$ - $10^{15}$  связей между нейронами.
- Частота импульсации --  $10^2$  Гц (современные персональные компьютеры -- до  $10^9$  Гц).

*NB!* Медлительность и ненадежность отдельных нейронов компенсируется их количеством.

- Параллельная переработка информации (в компьютерах -- преимущественно последовательная).
- «Переход количества в качество»: богатство поведения.
- Нельзя сказать, что мозг исходно «готов к использованию»: велика роль обучения.



# Нейросетевой подход: основные положения

- Процессы познания -- результат взаимодействия большого числа простых перерабатывающих элементов, связанных друг с другом и организованных в слои («модули»). «Переработка информации» -- определенный ответ элемента на воздействия извне.
- Знания, управляющие процессом переработки, хранятся в форме *весовых коэффициентов связей* между элементами сети. Главное -- не элементы, а связи между ними («*субсимвольный подход*»).
- Обучение -- процесс изменения весовых коэффициентов связей между элементами сети (приспособления их к решению определенной задачи).

# Классы задач, решаемых современными нейросетями:

- **Классификация:** распознавание образов, распознавание голосов, верификация подписей, постановка диагноза, анализ экспериментальных данных и т.д.
- **Моделирование:** поведение системы, поставленной в определенные условия.
- **Прогноз:** погода, ситуация на рынке ценных бумаг, бега, выборы и т.д.

## Комплексные задачи:

- управление
- принятие решений

«Центральные системы»  
модульного подхода

# Нейронные сети: рождение идеи (1943)

Уоррен Маккаллох  
(1898-1969)



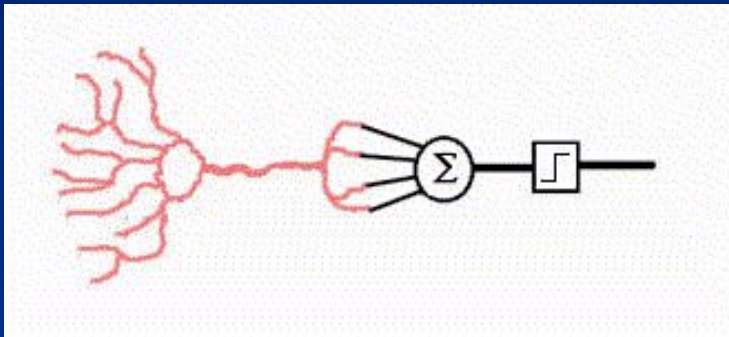
Уолтер Питтс  
(1923-1969)



«Логическое исчисление присуще нейронной активности» (1943)

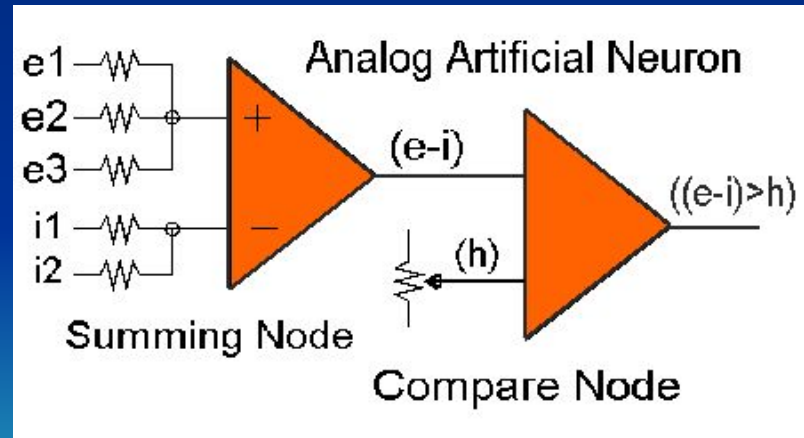
# Нейронные сети

## ФОРМАЛЬНЫЙ НЕЙРОН



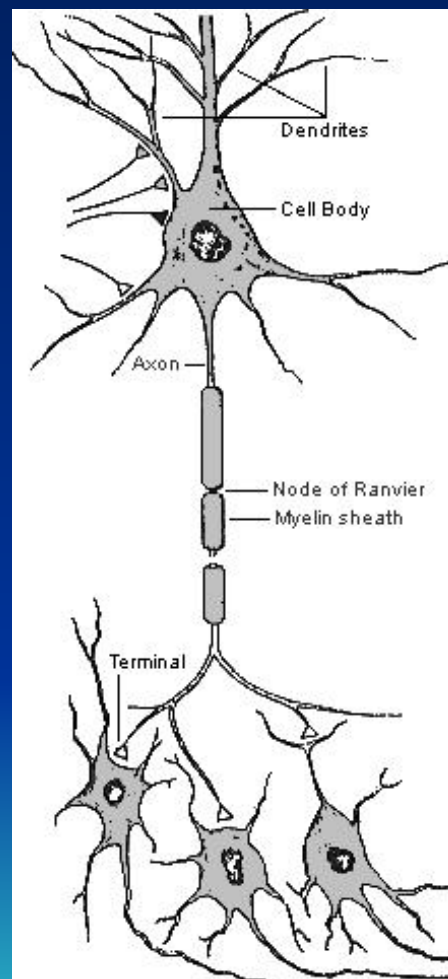
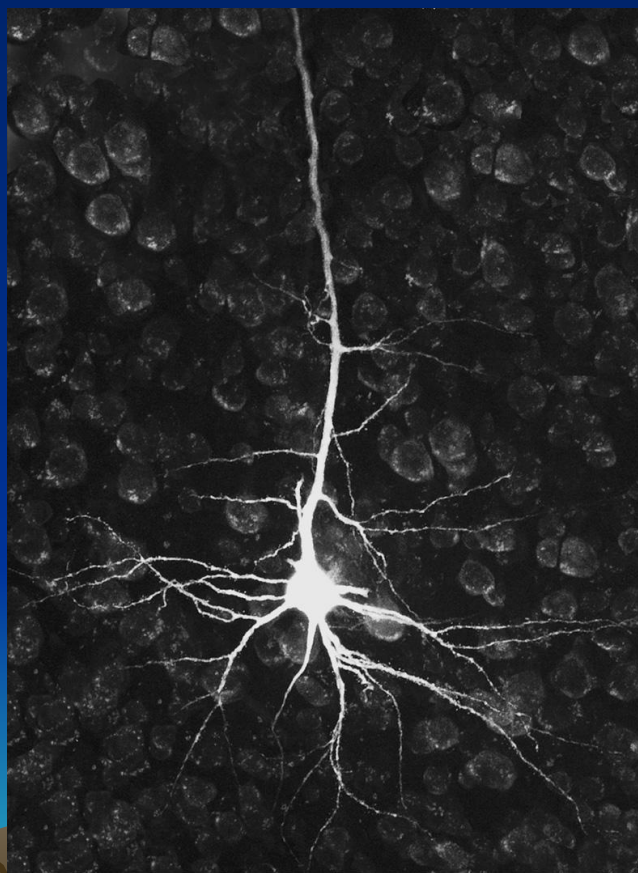
Элемент с пороговой логикой (TLU):

преодоление порога -- 1,  
иначе -- 0.



# Нейронные сети

## РЕАЛЬНЫЙ НЕЙРОН





# Теоретическая концепция искусственной сети Маккаллоха и Питтса

Три типа нейронов:

- входные (рецепторы) -- активируются извне;
- внутренние (центральные) -- активируются входными и прочими нейронами и активируют входные и прочие нейроны;
- выходные (эффекторы) -- получают импульсы от центральных и входных нейронов и отвечают за выполнение действия.





# Теоретическая концепция искусственной сети Маккаллоха и Питтса

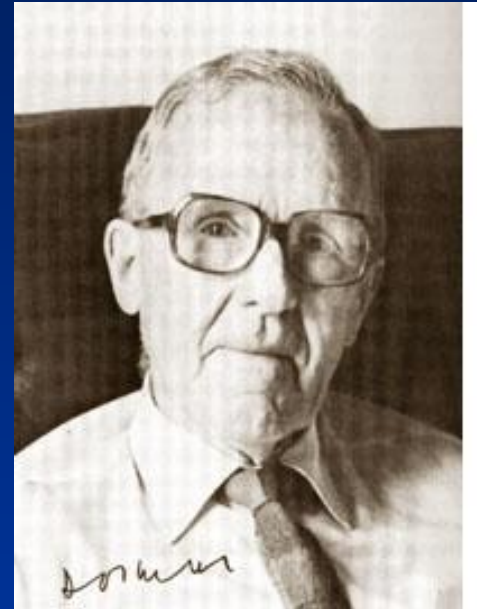
Правила функционирования сети:

- задержки в распространении активации одинаковы для всех нейронов сети;
- нейроны импульсируют не постоянно, а только в определенные моменты;
- каждый выходной синапс одного нейрона соответствует только одному входному синапсу следующего нейрона;
- на любом нейроне может сходиться несколько синапсов;
- входные синапсы вносят вклад в преодоление порога активации, при переходе через который (и только в этом случае) нейрон начинает передавать импульс.

# ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

Дональд Олдинг Хебб  
(1904-1985)

Правило Хебба (1949):  
между одновременно  
активированными  
нейронами сети пороги  
синаптической связи  
снижаются.

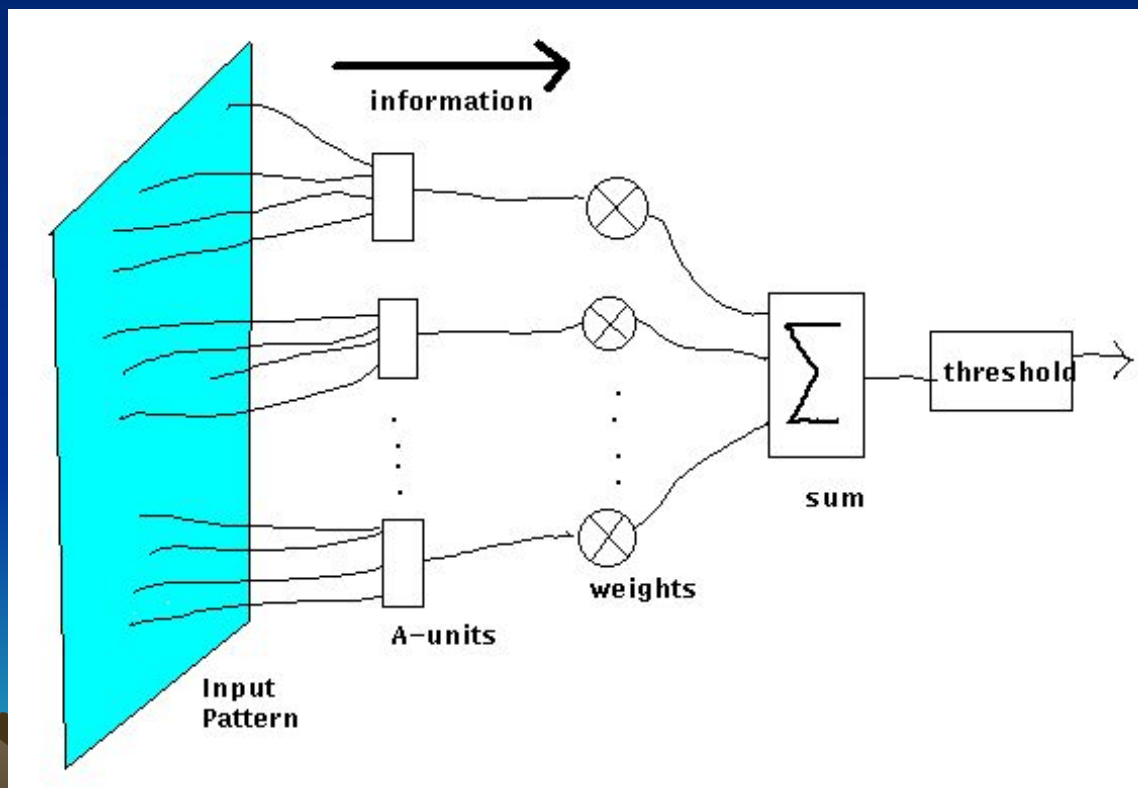


↓

Итог -- образование *«нейронного ансамбля»*, который все быстрее активируется при каждом очередном повторении входа.

# Развитие нейронных сетей

Фрэнк Розенблатт (1928-1969),  
Корнельский университет, США --  
перцептрон (1958)



# Развитие нейронных сетей

Фрэнк Розенблатт (1928-1969),  
Корнельский университет, США

1962 -- «Принципы нейродинамики:  
перцептроны и теория мозговых  
механизмов»:

интеграция данных компьютерного  
моделирования (включая перцептрон),  
нейрохирургии, регистрации  
активности отдельных нейронов и т.д.



# УПАДОК КОННЕКЦИОНИЗМА



1969 -- Марвин Минский,  
Сеймур Пейперт  
«Перцептроны»:  
приговор нейронным сетям?



Критика перцептронов: математическое обоснование их неэффективности в решении задач распознавания образов (в ходе поэлементного анализа связанных и несвязанных изображений теряется информация о связанности, которую невозможно задать линейно).

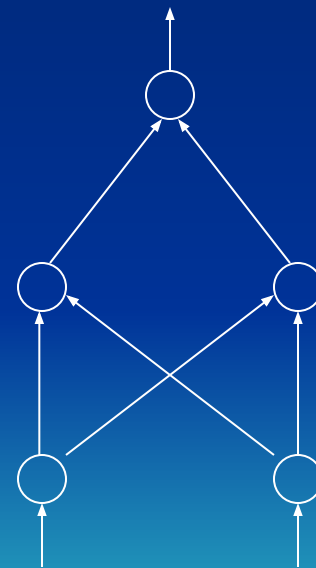
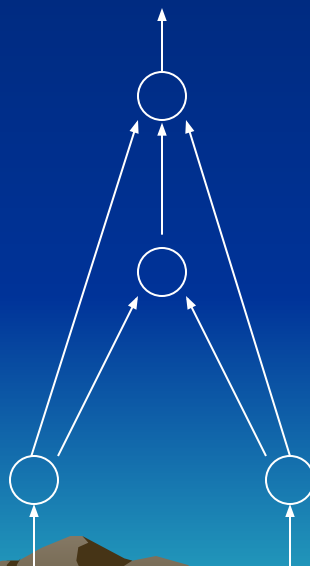
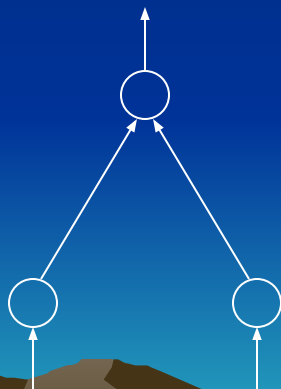


# УПАДОК КОННЕКЦИОНИЗМА

Проблема «исключающего ИЛИ» (*XOR*):

$(0;0) (1;1) \rightarrow 0$

$(0;1) (1;0) \rightarrow 1$



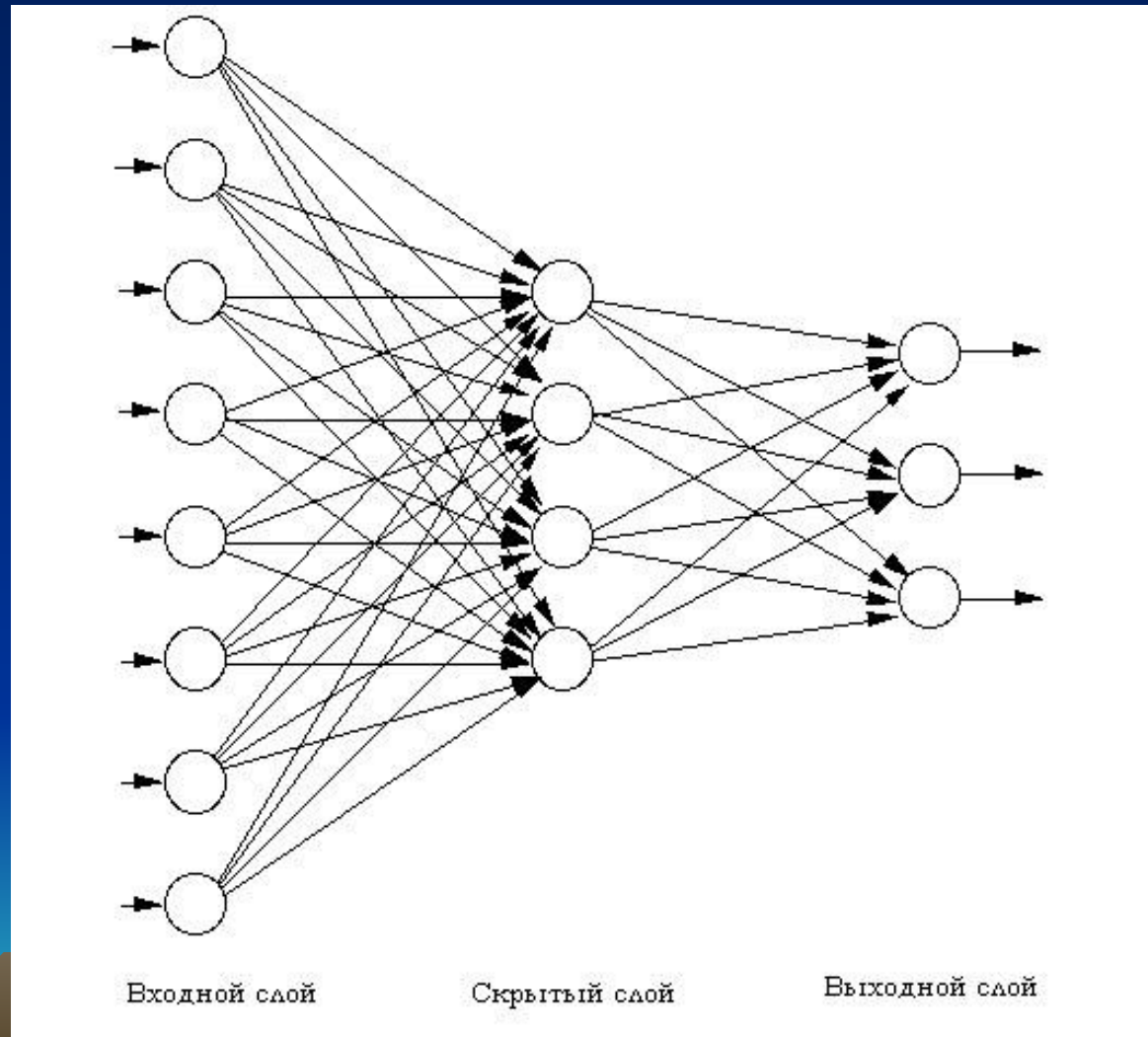
# РЕНЕССАНС КОННЕКЦИОНИЗМА

1986 -- Дэвид Румельхарт (Стэнфорд),  
Джеймс Макклелланд (Карнеги-Меллон)

«Параллельно-распределенная переработка» (*PDP*)



# АРХИТЕКТУРА НЕЙРОННОЙ СЕТИ





## Основные понятия:

«Нейрон» (*unit, node*) -- элемент сети, который суммирует входные сигналы и, в случае превышения *порога* его активации, выдает выходной сигнал (1 или 0) , выполняющий функцию активации или торможения в соответствии с *весовым коэффициентом* связи между ним и последующими нейронами.

Функция связи между элементами сети («синапса») -- умножение *сигнала* на *весовой коэффициент*.

*Порог* -- весовой коэффициент, связанный с постоянным входным сигналом, равным 1.

# ВИДЫ АРХИТЕКТУР:

- Сеть прямого распространения
- Сеть обратного распространения (рекуррентная)




# ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ:

- «Обучение с наставником»: задачи *распознавания* (заранее известен правильный ответ -> сеть настраивается на выдачу ответов, максимально близких к нему).

Алгоритм:

обратное распространение ошибки (*backpropagation*)

«Психологический механизм»:

«Предвосхищение» (результат работы сети)  «Истинное положение дел» (эталон)

# ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ:

- «Обучение с наставником»: задачи *распознавания* (заранее известен правильный ответ -> сеть настраивается на выдачу ответов, максимально близких к нему).
- «Обучение без наставника»: задачи *классификации* (правильный ответ неизвестен, но набор параметров относительно устойчив -> раскрытие внутренней структуры данных или связей между образцами).
- Смешанные формы обучения.



# ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ:

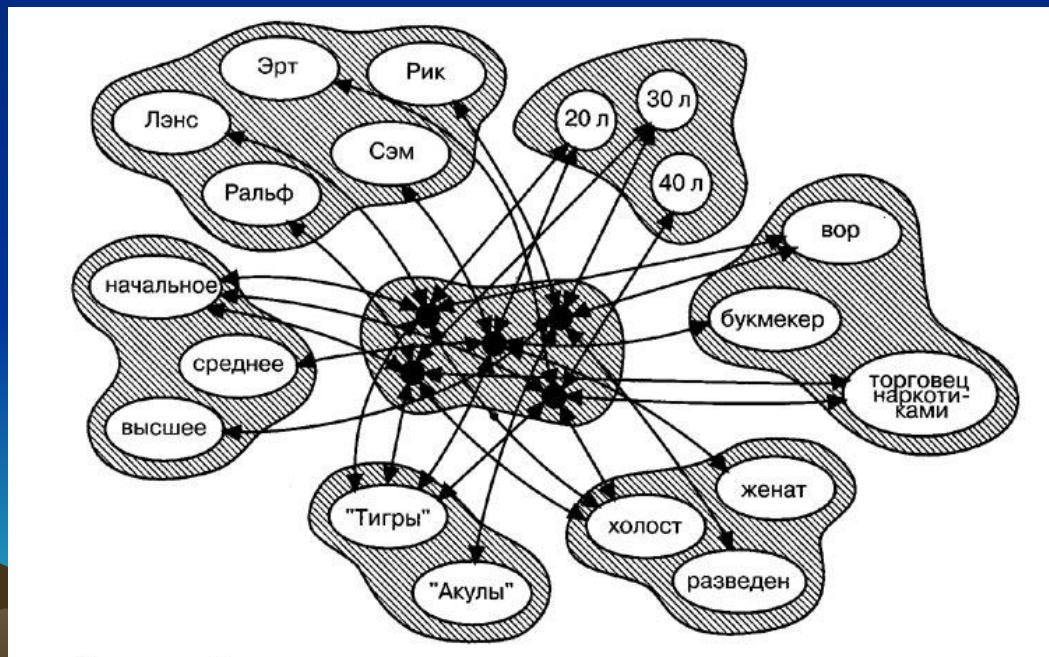
- Проблема устойчивости обучения: система обучения устойчива, если ни один из примеров обучающей выборки не изменит своей принадлежности к установленной категории после определенного числа итераций (повторных предъявлений).
- Феномен «переобученности» сети: хорошее функционирование на примерах обучающей выборки и плохое -- на сходных, но не идентичных тестовых примерах.

*СРАВНИМ: стадия дифференциации при выработке условного рефлекса  
(по данным лаборатории И.П. Павлова).*

# «БИБЛИЯ КОННЕКЦИОНИЗМА»

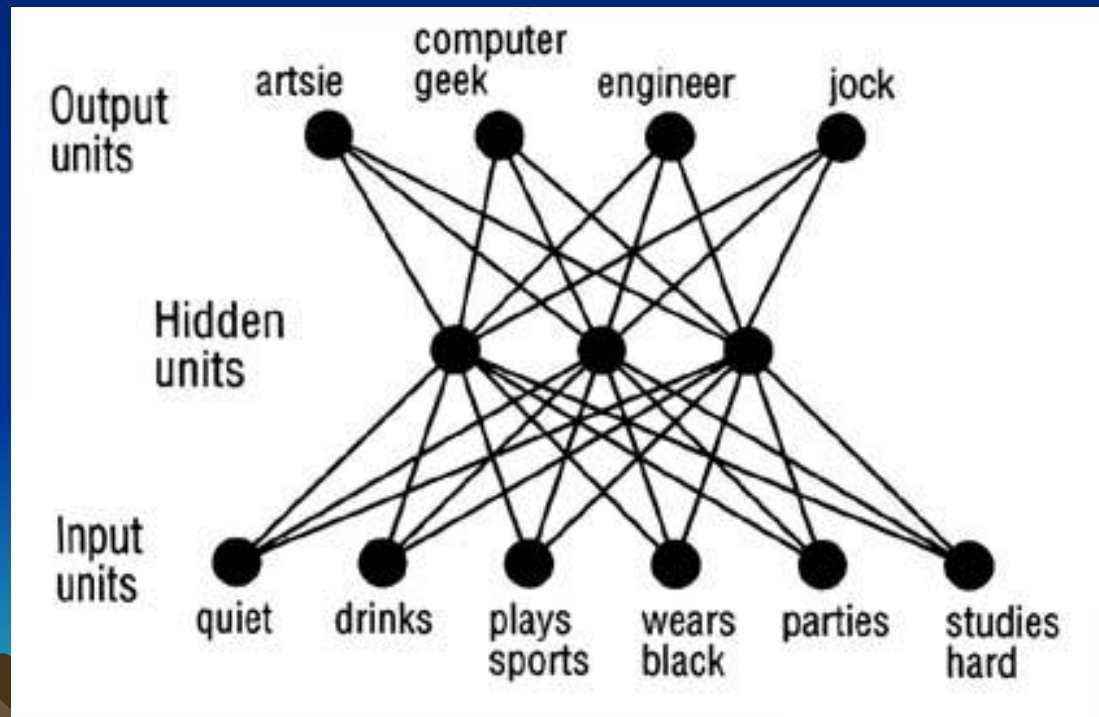
Организация памяти (Макклелланд, 1981):

- адресация по содержанию
- возможность «восстановления» информации:  
правило «щадящего разрушения» (*graceful degradation*)



# «БИБЛИЯ КОННЕКЦИОНИЗМА»

Последующие разработки:  
формирование у нейронной сети  
«СОЦИАЛЬНЫХ СТЕРЕОТИПОВ»



# «БИБЛИЯ КОННЕКЦИОНИЗМА»

1986 -- Дэвид Румельхарт, Джеймс Макклелланд

Освоение языка -- ряд стадий, характерных для развития ребенка, в том числе *стадия сверхобобщения* (4-5 лет):

to play -- played

to help -- helped

to kiss -- kissed

to go -- went

to jump -- jumped

to shout -- shouted

to go ... wented!

goed!

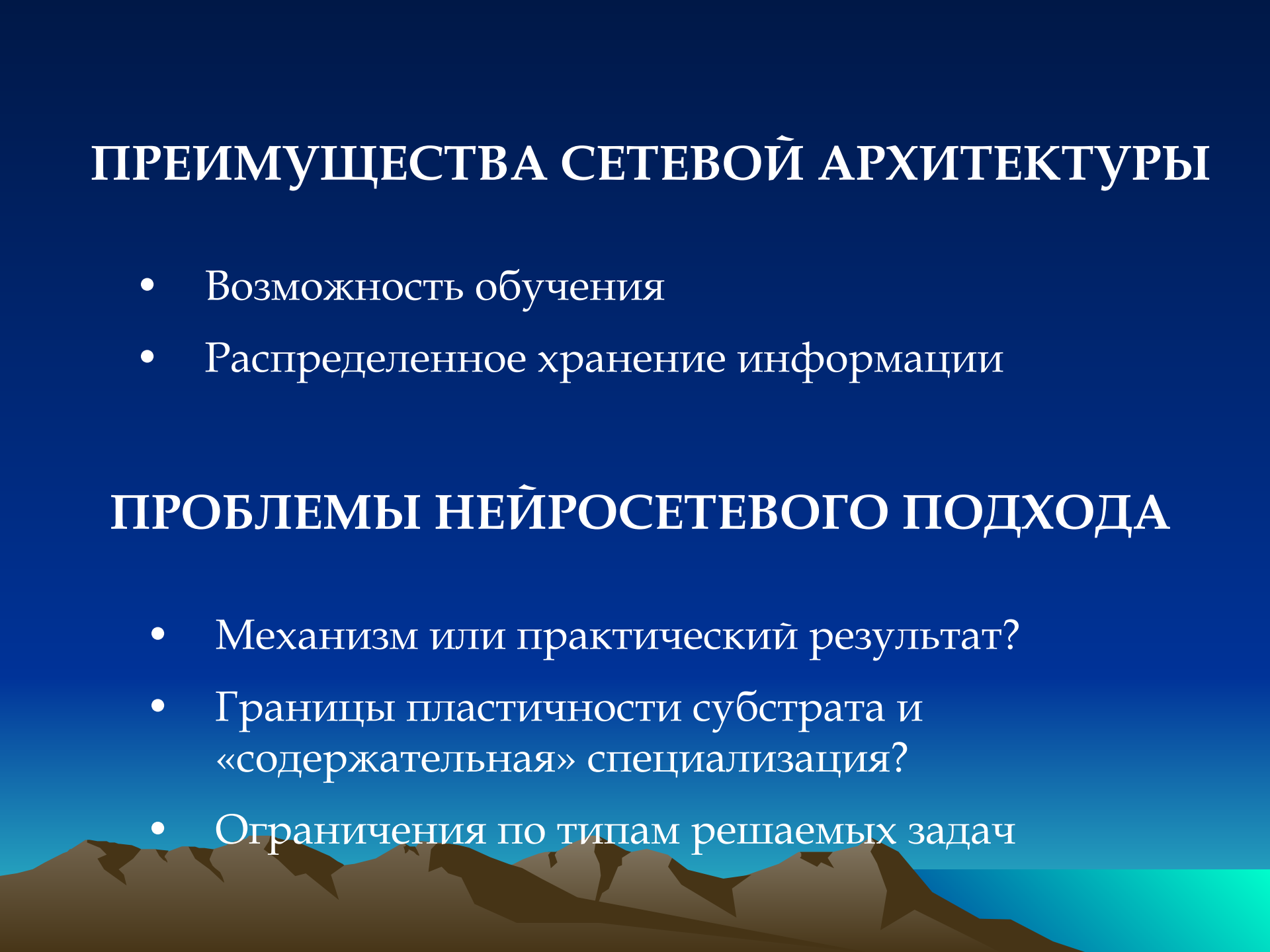




# ПРЕИМУЩЕСТВА СЕТЕВОЙ АРХИТЕКТУРЫ

- Возможность обучения
- Распределенное хранение информации

# ПРОБЛЕМЫ НЕЙРОСЕТЕВОГО ПОДХОДА

- Механизм или практический результат?
  - Границы пластичности субстрата и «содержательная» специализация?
  - Ограничения по типам решаемых задач
- 

# СИМВОЛЬНЫЕ И НЕЙРОСЕТЕВЫЕ МОДЕЛИ: «СФЕРЫ ВЛИЯНИЯ»

## Нейронные сети

неявные правила,  
«интуитивные» задачи  
(индивидуальные знания):  
умозаключение по  
анalogии, выделение  
фигуры на фоне и т.п.

Задачи, требующие  
обучения.

## Символьные модели

явные правила,  
формализуемые задачи  
(культурно-обусловленные  
общедоступные знания):  
например, логические и  
математические задачи.

Задачи, требующие  
конечного набора знаний.

# СИМВОЛЬНЫЕ И НЕЙРОСЕТЕВЫЕ МОДЕЛИ: «СФЕРЫ ВЛИЯНИЯ»

Нейронные сети

Символьные модели

«Холистическая»  
стратегия правого  
полушария



«Аналитическая»  
стратегия левого  
полушария

# ВОЗМОЖНОСТИ ИНТЕГРАЦИИ НЕЙРОСЕТЕВОГО И СИМВОЛЬНОГО ПОДХОДОВ: ЭКСПЕРТНЫЕ СИСТЕМЫ

Нейронная сеть



Экспертная система

распознавание  
образов, быстрые  
ответы на запросы  
сложной  
окружающей среды

принятие решений,  
логическая проверка  
выводов с учетом  
дополнительной  
информации



# ВОЗМОЖНОСТИ ИНТЕГРАЦИИ НЕЙРОСЕТЕВОГО И МОДУЛЬНОГО ПОДХОДОВ: ПРОБЛЕМА ВРОЖДЕННОГО И ПРИОБРЕТЕННОГО В ПОЗНАНИИ

«Наследственность» нейронной сети:

- количество элементов
- количество слоев
- правила и параметры распространения активации и изменения весов в разных слоях

*Достаточно ли этого для развития форм познания,  
характерных для человека?*

