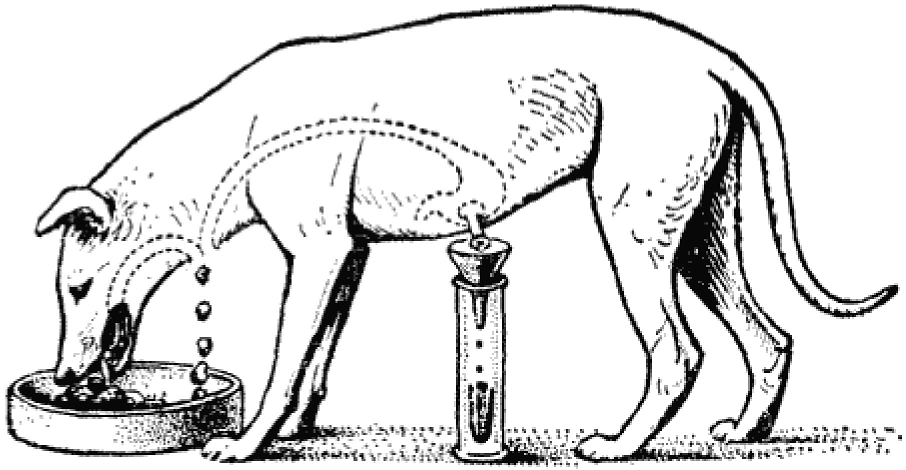
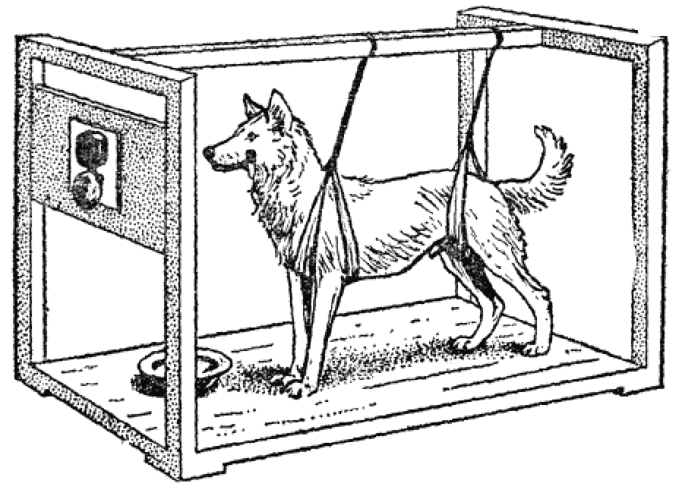
Карпов В.Э.

Модели поведения

Условно-рефлекторное поведение

Рефлексы





Безусловный стимул (БС) Пища → Безусловная реакция (БР) Слюноотделение

↕

Условный стимул (УС) Колокольчик → Условная реакция (УР) Слюноотделение

Формирование и угасание условного рефлекса

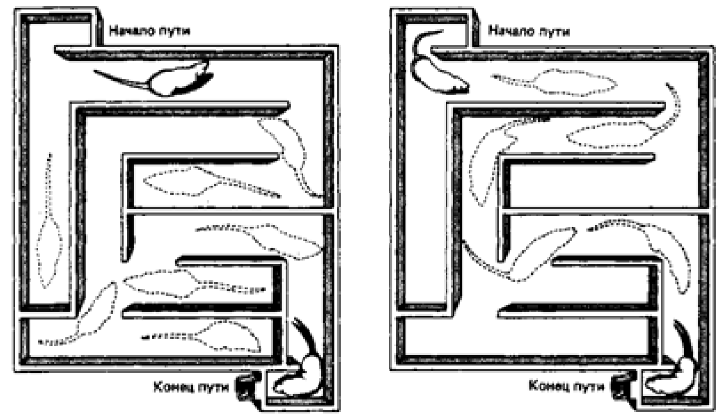
$$A_{n+1}^{i} = a_n \bigvee_{i=1}^{s} b_n^i P(N_n^i \geq N_0^i), \qquad i = \overline{1, s}$$

$$N_n^i = \sum_{k=1}^{n-1} \Delta N_k^i + \varepsilon_n^i \qquad \Delta N_n^i = \begin{cases} +\varphi_n^i, & a_{n-1} b_{n-1}^i = 1 \\ -\psi_n^i, & a_{n-1} \overline{b_{n-1}^i} = 1 \\ -\chi_n^i, & \overline{a_{n-1}} b_{n-1}^i = 1 \end{cases}$$

- i – номер нейтрального раздражителя
- P - предикат
- A_n – рефлекторный акт в момент времени n
- N_0^i – порог срабатывания i-го раздражителя
- φ_n^i – функция запоминания
- ψ_n^i, χ_n^i – функции забывания
- ε_n^i – случайная составляющая
- a_n – безусловный раздражитель
- b_n^i – нейтральный раздражитель

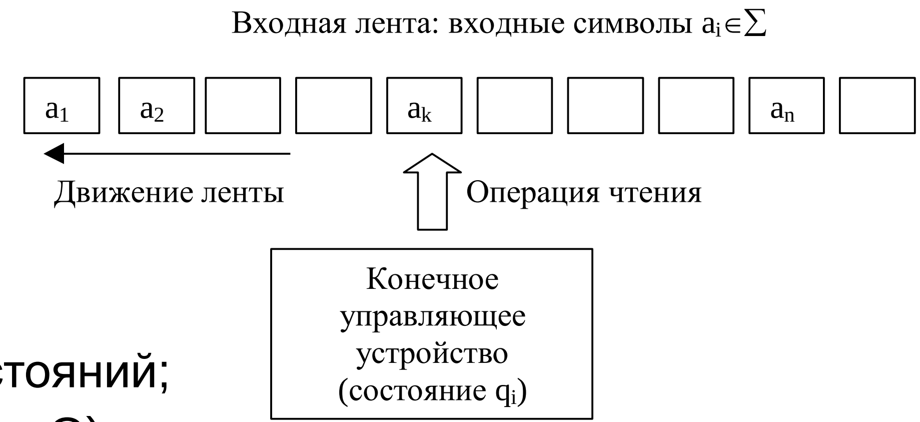
Автоматные модели

- 30-е гг. Т-образные лабиринты. Зоопсихология. Йеркс: опыты по формированию условных рефлексов у дождевых червей. Торндайк: крысы в лабиринте.
- М.Л.Цетлин. 60-е гг. Тезис: любое достаточно сложное поведение слагается из совокупности простых поведенческих актов.



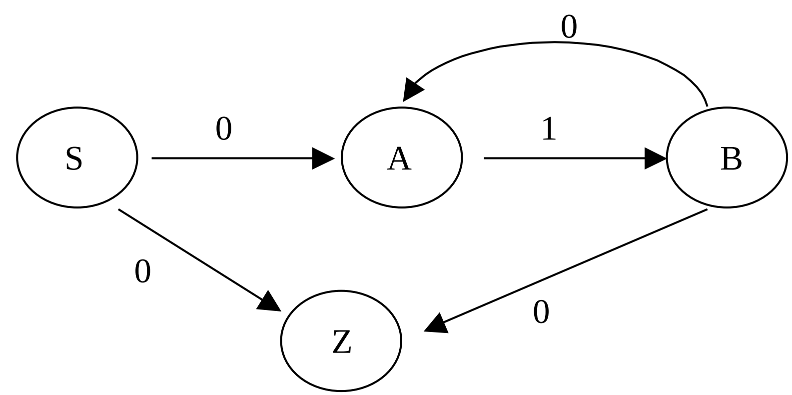
Цетлин М.Л. Исследования по теории автоматов и моделированию биологических систем. М.: Наука, 1969. – 316 с.

Конечный автомат

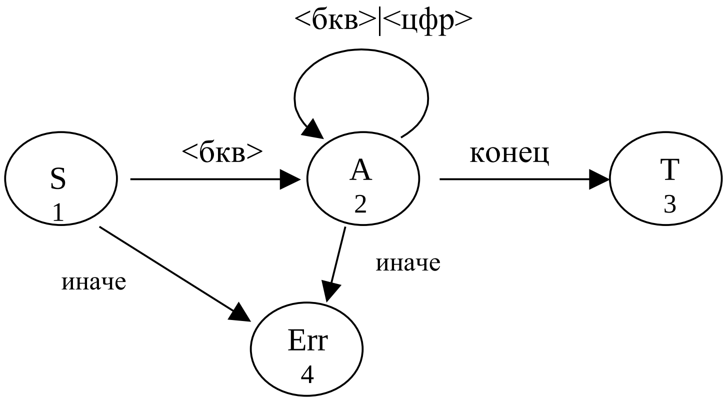


КА = (Σ, Q, q_0, T, P), где

- Σ - входной алфавит;
- Q - конечное множество состояний;
- q_0 - начальное состояние ($q_0 \in Q$);
- T - множество терминальных состояний, $T \subset Q$;
- P – подмножество отображения вида $Q \times \Sigma \rightarrow Q$, называемое функцией переходов. Элементы этого отображения называются *правилами:*
- $q_i a_k \rightarrow q_j$, где q_i и q_j - состояния, a_k - входной символ: $q_i, q_j \in Q$, $a_k \in \Sigma$.



Пример КА



P:

	1	2	3
<бкв>	**2**	**2**	**-**
<цфр>	**4**	**2**	**-**
<конец>	**4**	**3**	**-**
<иначе>	**4**	**-**	**-**

Автомат, распознающий идентификаторы

```
char c;              //текущий исходный символ
int q;               //номер состояния
int a;               //входной текущий символ для автомата
q=0;                 //начальное состояние автомата
while(1)             //бесконечный цикл
{  c = readchar();   //считывание входного символа
   a = gettype(c);   //распознавание   входного   символа   –   отнесение   его   к   одной   из
известных
                    // автомату категорий - <бкв>, <цфр>, <конец> или <иначе>
   q = P[a, q]; //Выполнение перехода
   //Обработка
   if (q==3) return 1;            //нормальный выход из программы
   if (q==4) return 0;            //выход по ошибке
}
```

Стационарные среды

Среда E = (E_1, E_2,…,E_n)

E_k характеризует реакцию среды (наказание или поощрение) объекта за действие k. В качестве E_k может фигурировать вероятность наказания P_k.

- При равновероятном выборе действия математическое ожидание штрафа определяется как

$$M^* = \sum_{i=1}^{n} \frac{1}{n} P_i$$

- Целесообразность поведения заключается в том, чтобы математическое ожидание наказания было меньше, чем M*.

Автоматы с линейной тактикой

- Автомат – объект, способный в каждый момент времени t=1,2,...воспринимать конечное число сигналов $S=(S_1,S_2,..,S_n)$ и в зависимости от них изменять свое состояние.

- Автомат может производить конечное число действий $f=(f_1, f_2,..,f_n)$ выбор, которого определяется внутренним состоянием автомата $\varphi=(\varphi_1, \varphi_2,..,\varphi_m)$, где m – емкость памяти автомата.

- Простейший случай: автомат воспринимает два сигнала S=0 (выигрыш) и S=1 (проигрыш).

Детерминированный и вероятностный КА

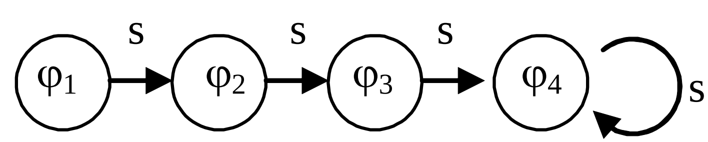
Детерминированный КА

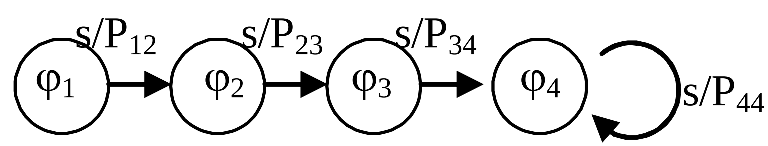
Канонические уравнения ДКА:

- $\varphi(t+1) = Q(\varphi(t))$ (1)
- $S(t+1) = F(\varphi(t))$ (2)

Матрица переходов Q: $[A_{ij}(S)]$

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$





Вероятностный КА

Матрицы состояний $[A_{ij}(S)]$, S=0,1 являются стохастическими

Стационарная случайная среда

Стационарная случайная среда $\mathbf{C=C(A_1,..,A_n)}$:

действие f_a, где a=1,..,n в момент t влечет за собой в момент t+1

S=1 (проигрыш) с вероятностью $p_a=(1-A_a)/2$

S=0 (выигрыш) с вероятностью $q_a=(1+A_a)/2$, $(A\leq 1)$.

Вероятность p_{ij} перехода автомата из состояния φ_i в состояние φ_j:

$$\mathbf{p_{ij} = p(a_i)\cdot A_{ij}(1) + q(a_i)\cdot A_{ij}(0);\ i, j = 1,.., m}$$

$p(a_i)$, $q(a_i)$ – вероятности наказания и поощрения за действие i.

Матрица $P=[p_{ij}]$ является стохастической

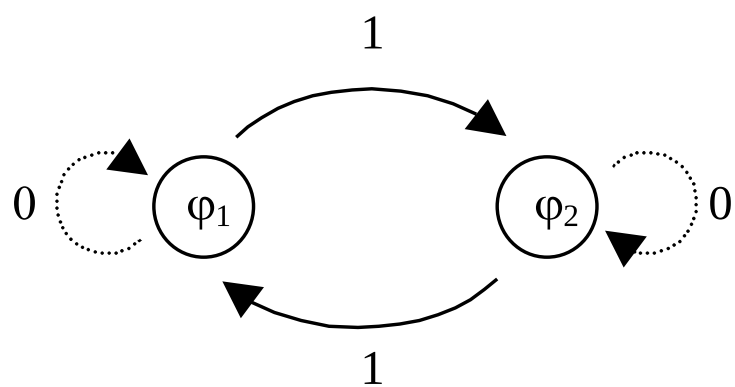
Пример автомата

L(2,2) - автомат двух состояний φ_1 и φ_2 двух действий $f_1=F(\varphi_1)$, $f_2=F(\varphi_2)$.

Автомат сохраняет свои состояния при выигрыше и изменяет при проигрыше.

$$\left[A_{ij}(1)\right]=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},$$

$$\left[A_{ij}(0)\right]=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$



Асимптотически оптимальные автоматы

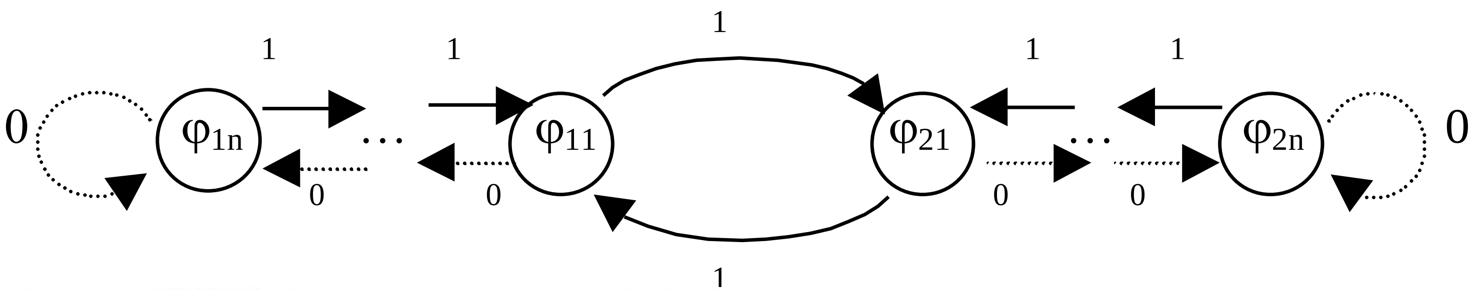
Последовательность автоматов $U_1,..,U_n$ *асимптотически оптимальна*, если в среде $C(A_1,..., A_x)$.

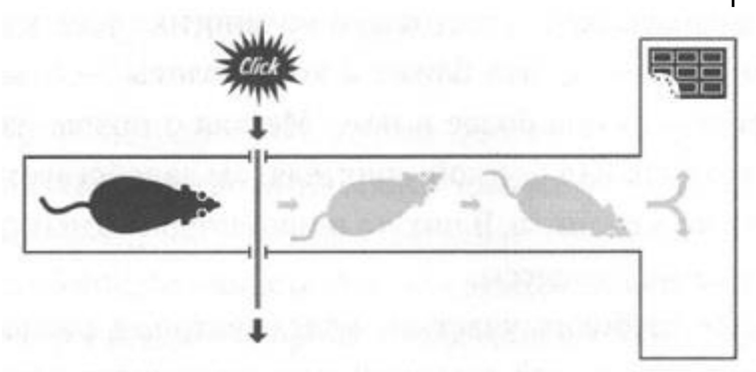
$$\lim_{n\to\infty} W(U_n, C) = A_{\max}$$

- Автомат, принадлежащий асимптотической последовательности, производит почти исключительно то действие, вероятность которого максимальна.

Примеры асимптотически оптимальных автоматов

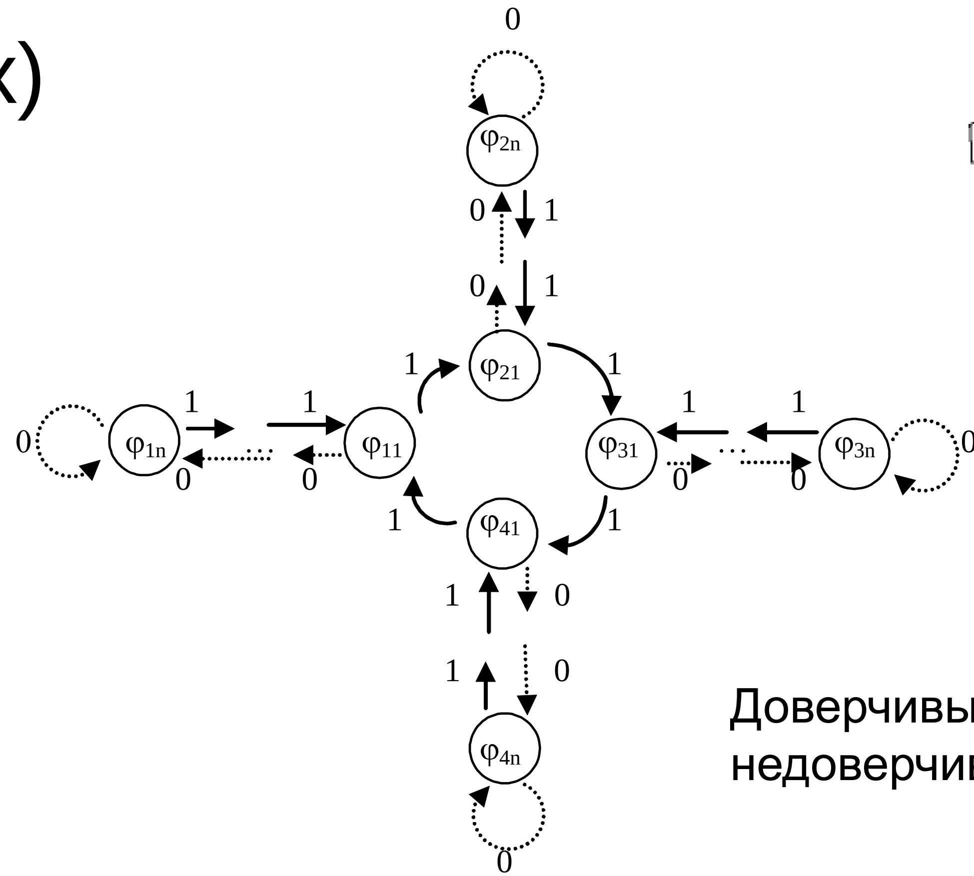
- Автомат с линейной тактикой L(2n,2)

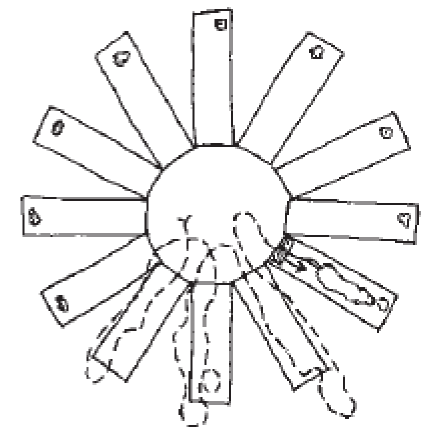




Автомат L(xn, x)

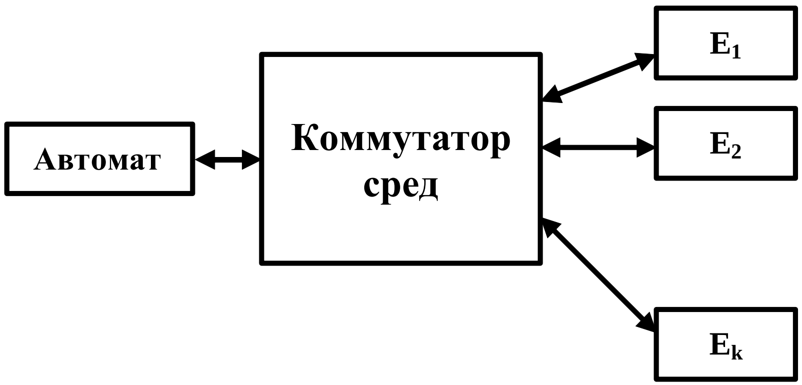
- L(xn, x)

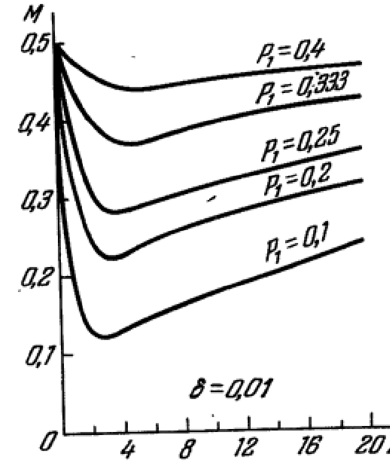




Доверчивые и недоверчивые автоматы

Автоматы с переменной структурой



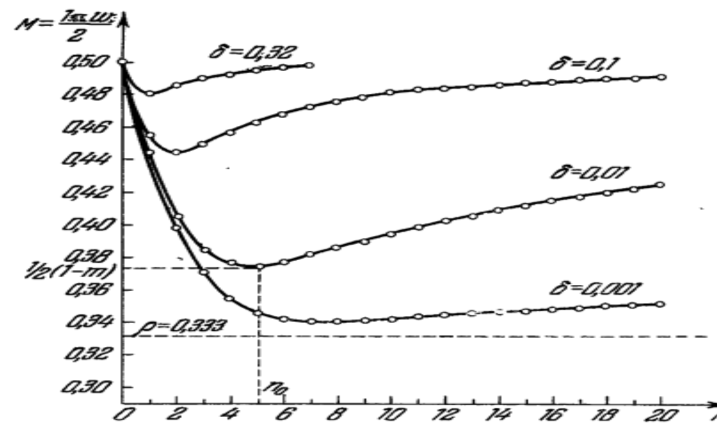


- $a_{ij}(t+1,s(t)) = a_{ij}(t,s(t))+(-1)^{s(t+1)}\cdot g\cdot a_{ij}(t,s(t))\cdot[1-a_{ij}(t,s(t))]$
- $a_{ik}(t+1,s(t)) = a_{ik}(t,s(t))-(-1)^{s(t+1)}\cdot g\cdot a_{ik}(t,s(t))\cdot a_{ij}(t,s(t))$

для $j\neq k$. $0\leq g\leq 1$

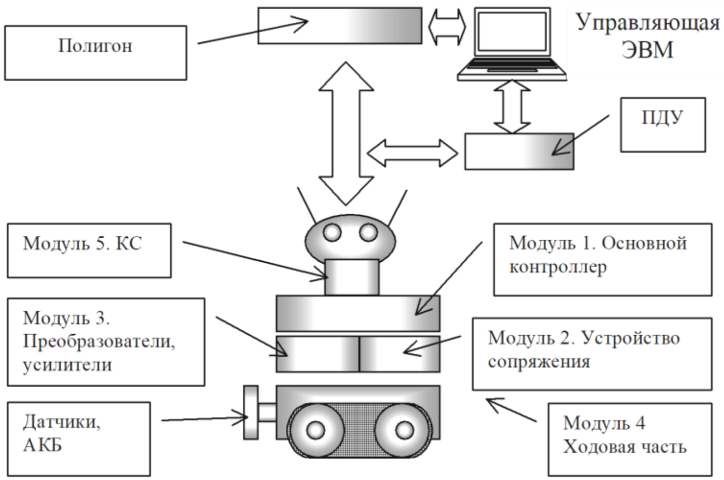
Вопрос оптимальности глубины памяти. «Городские» и «сельские» жители.

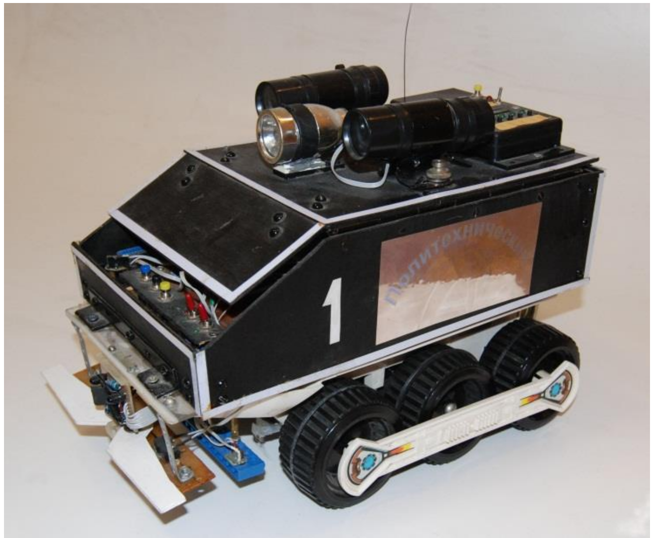
M=(1-W)/2



Формирование условного рефлекса

Робот АМУР-1. Лаборатория робототехники и искусственного интеллекта Политехнического музея.





ATMega162 - 7 МГц, флэш-память для хранения программного кода - 16 К, ОЗУ - 512 байт.

Структура автомата

$$y(t+1) = F(x(t), q(t), P(t))$$

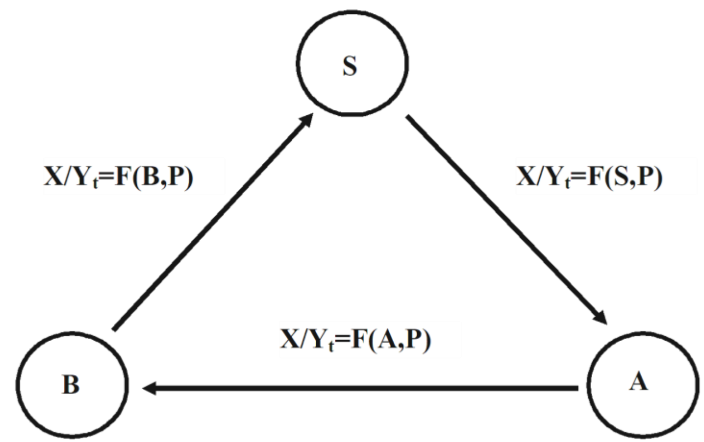
$$q(t+1) = Q(x(t), q(t))$$

Действия Y

Входные сигналы X

p_{11}	p_{12}	p_{13}	…	p_{1n}
p_{21}	p_{22}	p_{23}	…	p_{2n}
p_{31}	p_{32}	p_{33}	…	p_{3n}
…				
p_{m1}	p_{m2}	p_{m3}	…	p_{mn}

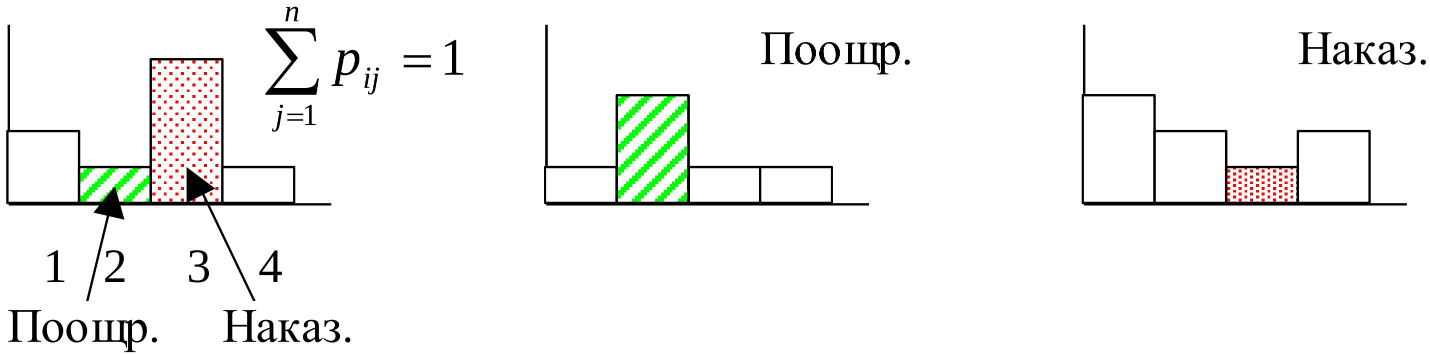
$$\sum_{j=1}^{n} p_{ij} = 1$$



Поощрение и наказание

$$p_{ij}(t+1,s(t)) = p_{ij}(t,s(t))+(-1)^{s(t+1)}\cdot g\cdot p_{ij}(t,s(t))\cdot[1-p_{ij}(t,s(t))]$$

$$p_{ik}(t+1,s(t)) = p_{ik}(t,s(t))-(-1)^{s(t+1)}\cdot g\cdot p_{ik}(t,s(t))\cdot p_{ij}(t,s(t)) \text{ для } k\neq j.\ 0\leq g\leq 1$$



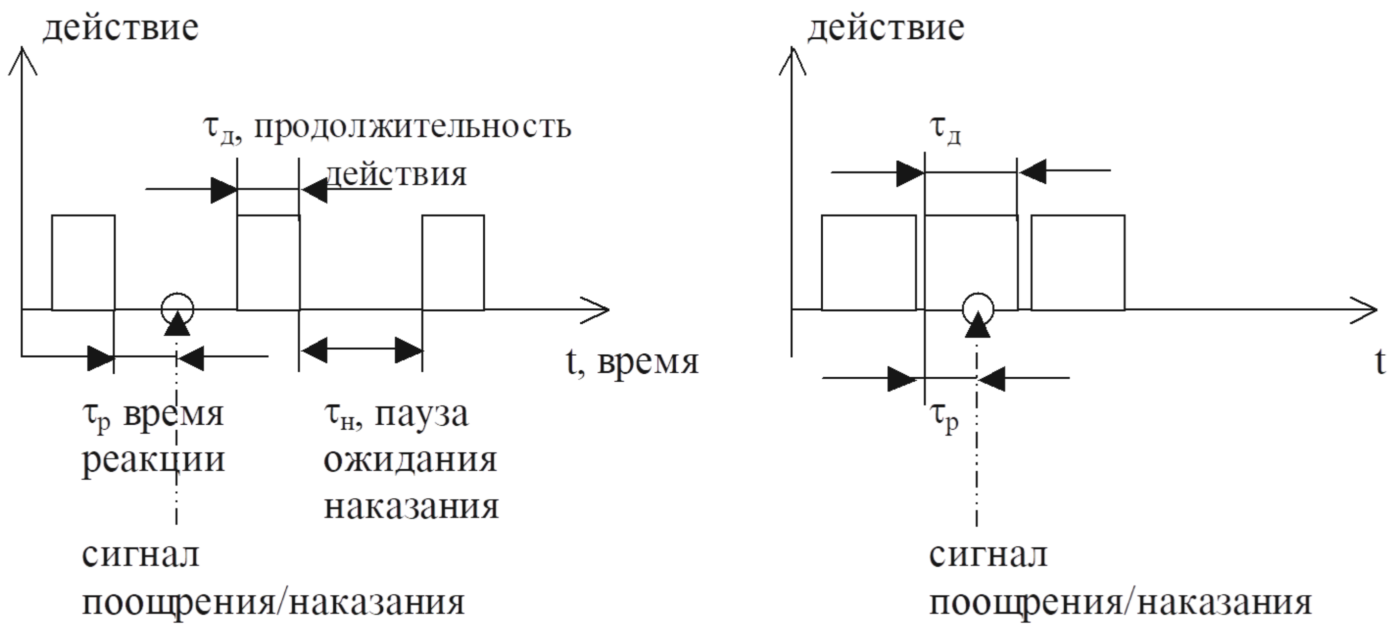
- При выработке условных рефлексов можно обойтись одними **наказаниями**
- «Отсутствие наказания может и должно рассматриваться как поощрение»

Реализация схемы наказания/поощрения

Основная проблема: как, когда, кого и за что наказывать

Синхронный и *асинхронный* способы подачи оценивающих воздействий

Асинхронный вариант



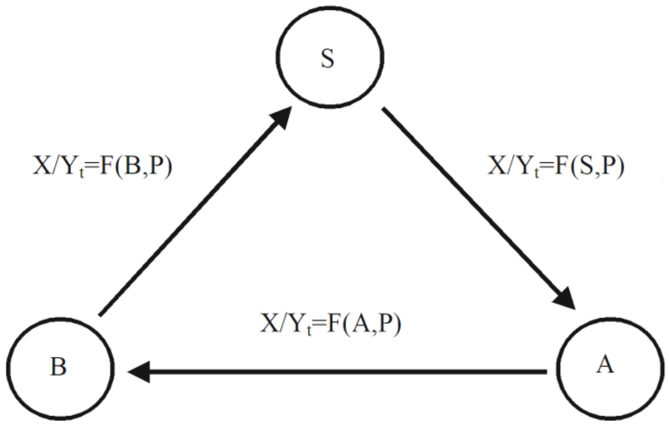
Режимы «периодического» и «непрерывного» функционирования

Синхронный метод

- После совершения действия автомат выдает **сигнал готовности** к приему оценки и ждет в течение некоторого времени. По окончании времени ожидания автомат выдает сигнал неготовности к приему

- Нейробиологические основы

Обучение

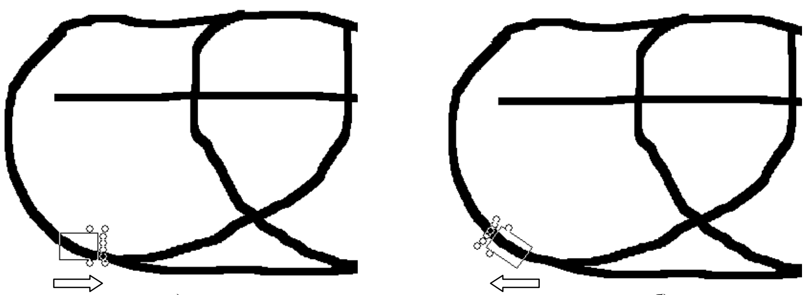


Действия Y

Входные сигналы X

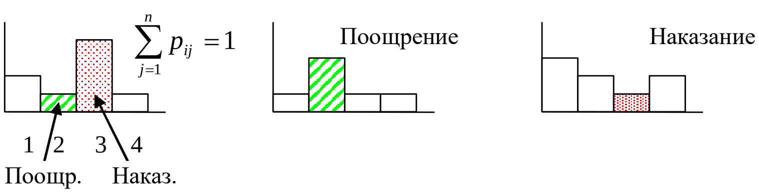
p_{11}	p_{12}	p_{13}	…	p_{1n}
p_{21}	p_{22}	p_{23}	…	p_{2n}
p_{31}	p_{32}	p_{33}	…	p_{3n}
…				
p_{m1}	p_{m2}	p_{m3}	…	p_{mn}

$$\sum_{j=1}^{n} p_{ij} = 1$$



а) б)

Б-обучение (± - обучение). Учитель ограничивается сигналами поощрения и наказания;
Формирование вектора управляющих воздействий.



Псевдовзаимодействие роботов

- Пусть робот A_1 функционирует в некоторой среде S_1, а робот A_2 - в среде S_2.
- Среды S_1 и S_2 являются частями некой единой среды S.
- Задача – реализовать схему взаимного обмена навыками, полученными роботами с тем, чтобы робот, успешно живущий в S_1, мог бы функционировать и в «чужой» среде S_2.

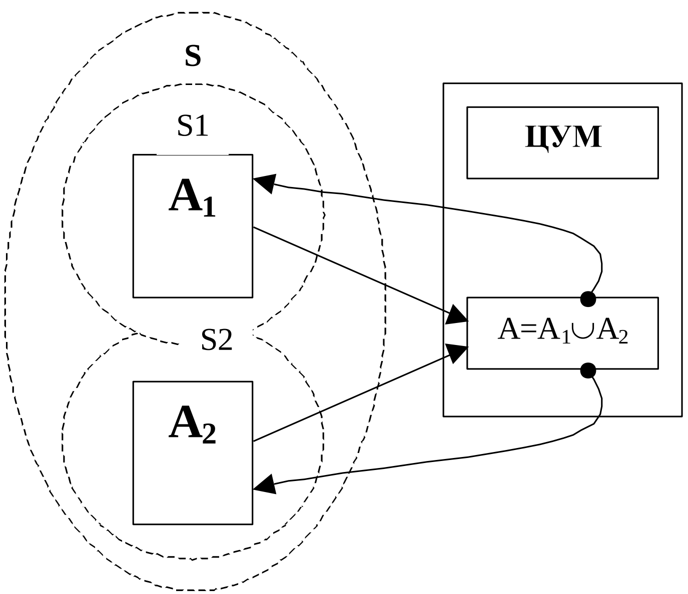


Схема «1ГХ2Л»

Модель 1ГХ2Л-А («Автоматы»)

A_1

P_i^1

p_{11}	p_{12}	…	p_{1n}
p_{21}	p_{22}	…	p_{2n}
p_{i1}	p_{i2}	…	p_{in}
…			
p_{m1}	p_{m2}	…	p_{mn}

$d_1 = D(\mathbf{P}^1)$

A_2

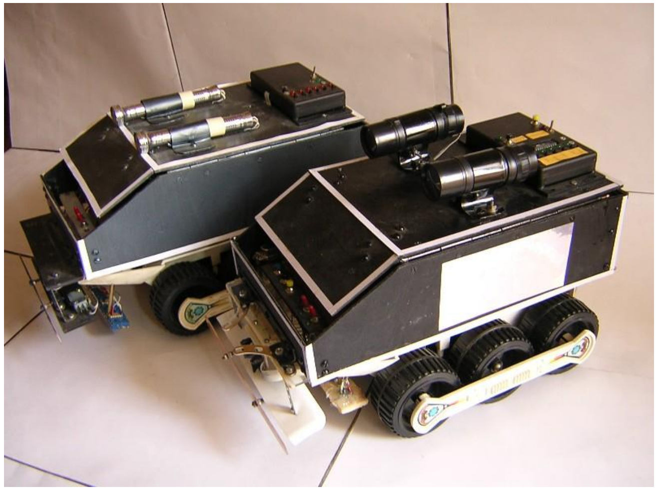
P_i^2

p_{11}	p_{12}	…	p_{1n}
p_{21}	p_{22}	…	p_{2n}
p_{i1}	p_{i2}	…	p_{in}
…			
p_{m1}	p_{m2}	…	p_{mn}

$d_2 = D(\mathbf{P}^2)$

$$\overline{P_i} = \begin{cases} P_i^1, d_1 \geq R, d_2 < R \\ P_i^2, d_2 \geq R, d_1 < R \\ P_i^1 \oplus P_i^2 \, \textit{иначе} \end{cases}$$

R – порог (СКО, δ и т.п.)



- Роботы, управляемые вероятностными автоматами.
- Процедура объединения матриц вероятностей действия.
- Сопоставление некоторых характеристик d векторов каждой матрицы. d определяет степень неоднородности элементов (дисперсии, среднего квадратичного отклонения и т.д.)

Задача индуктивной классификации

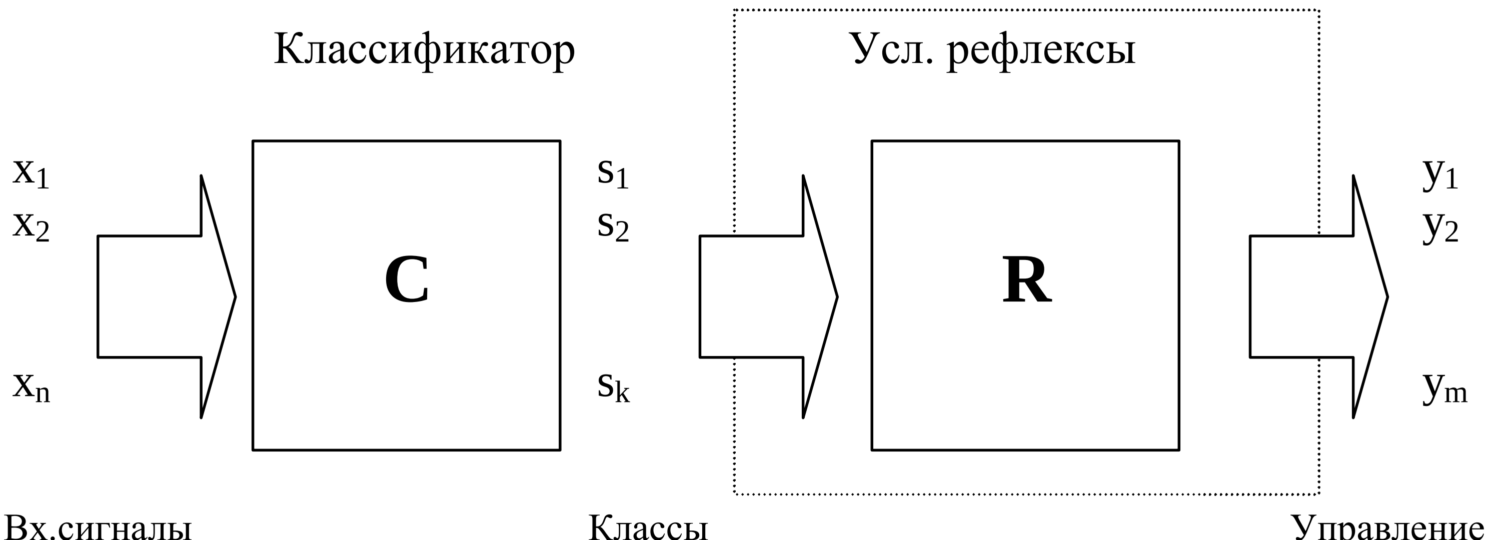
Действия Y

Входные сигналы X

p_{11}	p_{12}	p_{13}	…	p_{1n}
p_{21}	p_{22}	p_{23}	…	p_{2n}
p_{31}	p_{32}	p_{33}	…	p_{3n}
…				
p_{m1}	p_{m2}	p_{m3}	…	p_{mn}

$$\sum_{j=1}^{n} p_{ij} = 1$$

dim $X=2^R$

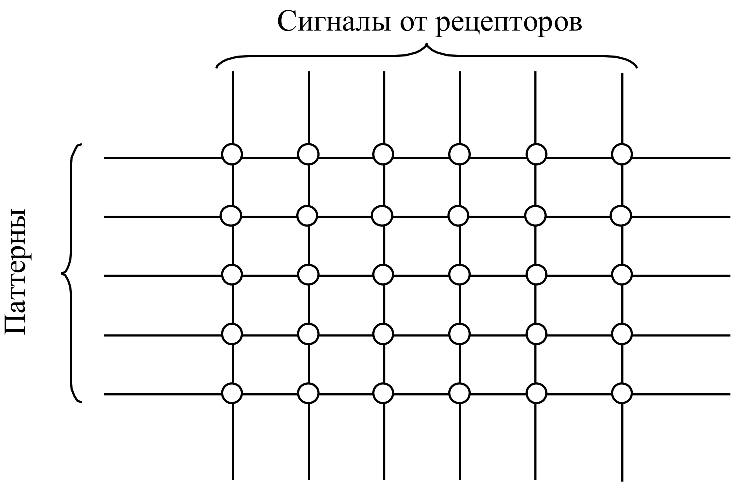


$\mathbf{Y} = \mathrm{R}(\mathbf{S}),\ \mathbf{S} = \mathrm{C}(\mathbf{X})$

$\mathbf{Y} = \mathrm{RC}\ \mathbf{X}$

Сложные рефлексы

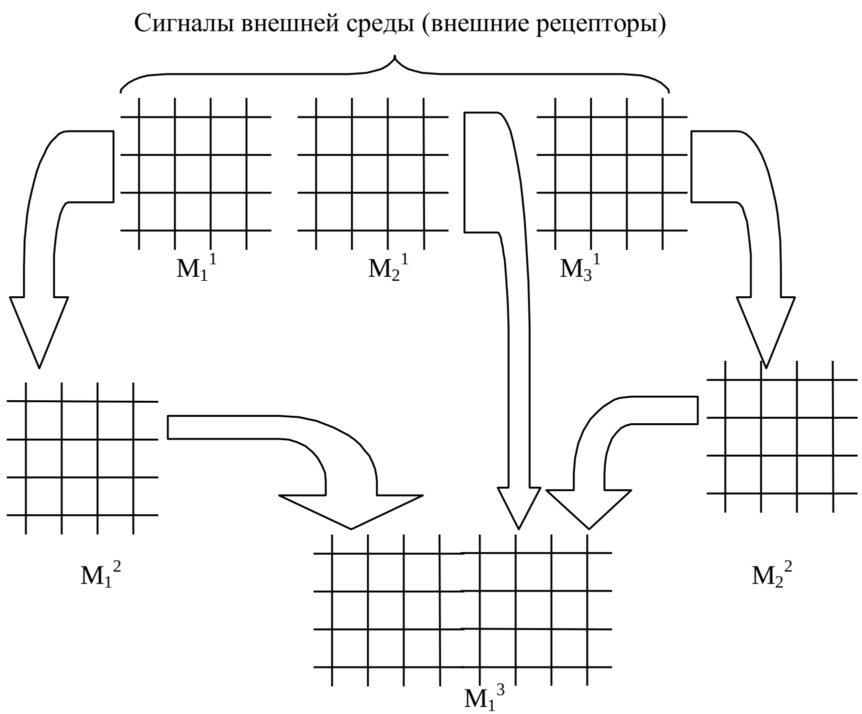
- К.Штайнбух. Обучаемые матрицы.



Условные рефлексы (связи в местах пересечений) образуются на этапе обучения при многократном повторении одновременного возбуждения пересекающихся шин. Сложные рефлексы образуются путем соединения нескольких обучающихся матриц.

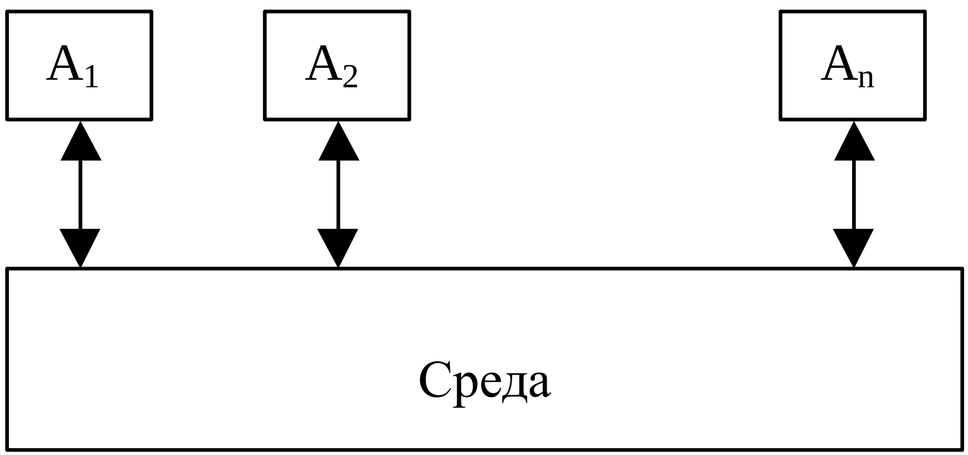
Соединение матриц

- Моделирование процесса абстрагирования.
- Если позволить распространяться сигналам в противоположном направлении – от паттернов к рецепторам, то получим эффект **конкретизации образа** в виде чувственного представления. Если вместо рецепторов использовать эффекторы, то мы получим процесс **управления**.



Коллективное поведение автоматов

- Однородные структуры
- Оценке подлежит совокупное воздействие всего коллектива автоматов $D=(d^1_{i1},d^2_{i2},\ldots,d^k_{ik})$, где d^m_k – k-е действие m-го автомата.

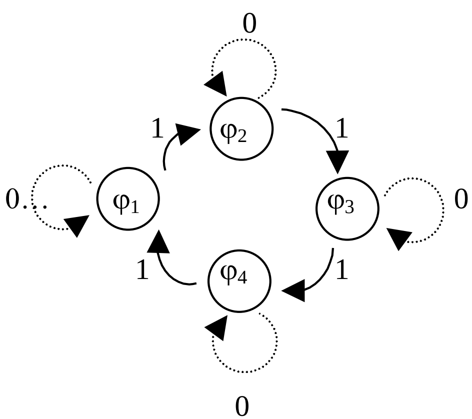


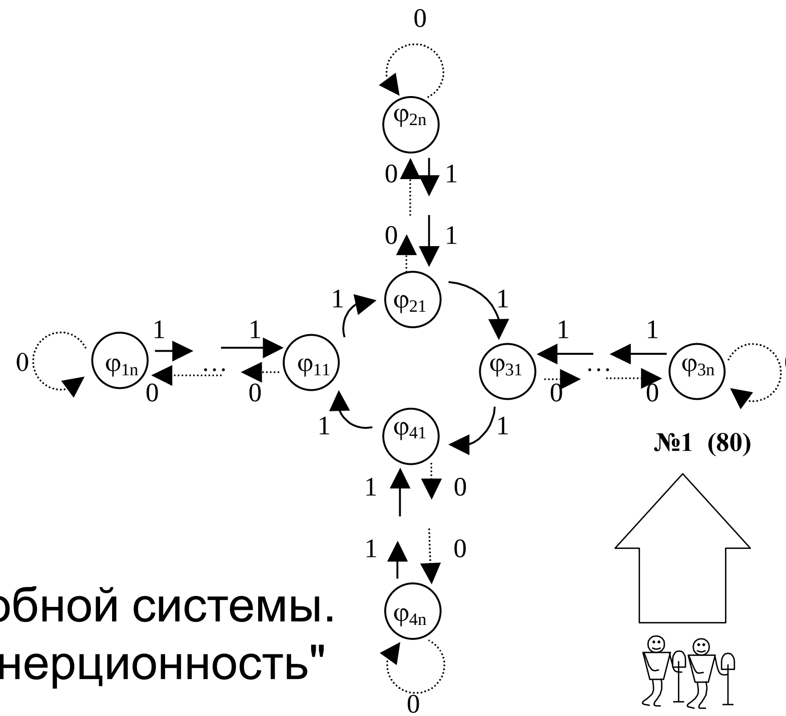
Игры автоматов

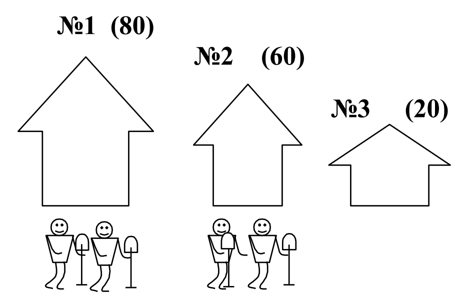
- Равновесие Нэша. Набор стратегий в игре для двух и более игроков, в котором ни один участник не может увеличить выигрыш, изменив свою стратегию, если другие участники своих стратегий не меняют.
- Устойчивая по Нэшу партия максимальной цены называется *точкой Мора* или *партией Мора.*

Задача 1. N мест и M претендентов, N>M

- Пусть N=3, a_1 = 80, a_2 = 60, a_3 = 20. M=4.
- В отсутствии информации о значениях a_i и динамике размещения по местам работы можно добиться такого положения, при котором каждый индивид максимизирует свой выигрыш (система выйдет на точку Мора).
- Для n рабочих мест







Неустойчивость работы подобной системы. Колебания. Можно ввести "инерционность" принятия решения.

Задача 2. Игра с общей кассой

Необходимо максимизировать суммарную зарплату, получаемую всем коллективом (схема с общей кассой).

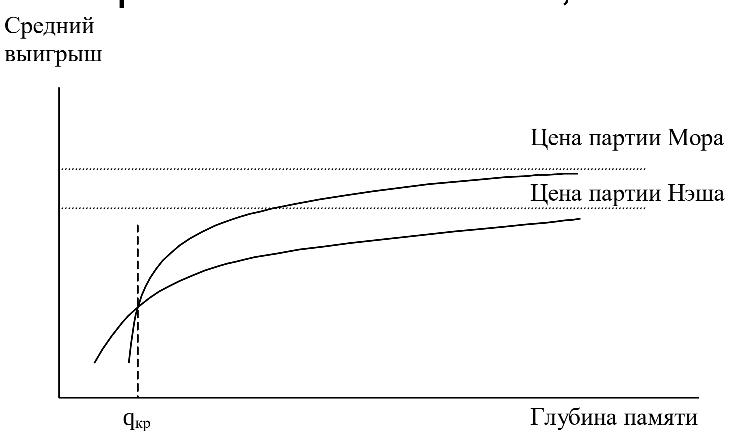
- $\Sigma a_i=80+60+20=160$, $x^{cp}_{max}=\Sigma a_i/M=160/4=40$

Если окажется, что $x^{cp}< x^{cp}_{max}$, то будем наказывать тех игроков, чей выигрыш оказался меньше x^{cp}.

При показанном выше распределении $x^{cp}=(40+40+35+35)/4=35$. Будут наказаны игроки, зарабатывающие по 30 рублей. В этом случае кто-нибудь из них поменяет место работы.

Наказания прекратятся тогда, когда 2 будут работать на предприятии №1, 1– на предприятии №2 и 1 – на предприятии №3.

Выигрыш автоматов в игре с общей кассой при глубине памяти меньше критической меньше, чем в игре без общей кассы



Введение процедуры общей кассы делает партию максимальной цены устойчивой по Нэшу.

"Вред уравниловки при низкой сознательности"

Задача Майхилла (задача о цепи стрелков)

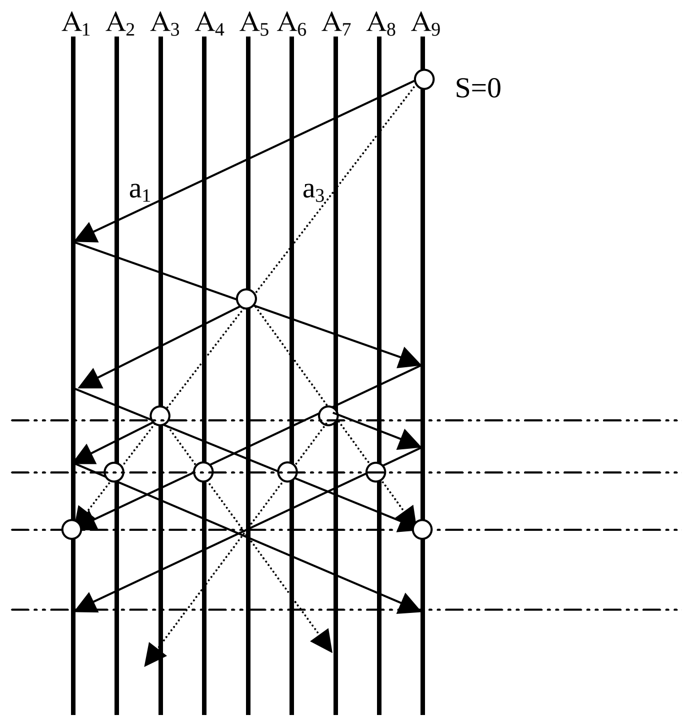
Существуют ли правила поведения стрелков, обеспечивающих синхронизацию, если количество слов, которыми могут обмениваться стрелки и объем внутренней памяти каждого из них ограничены и не зависят от длины цепи.

Минимально возможное время решения составляет 2N-2 тактов (N-количество стрелков).

- Э.Гото, 1962. Конечный автомат с несколькими тысячами состояний.
- В.И.Левенштейн, 1965. 9 внутренних состояний. Можно и с восемью состояниями.

Диаграмма распространения сигналов

- В точках пересечения сигналов a_1 и a_3 автомат переходит в состояние S_1 и сам начинает генерировать сигналы a_1 и a_3. Переход в синхронизирующее состояние S осуществляется тогда, когда и сам автомат и оба его соседа находится в состоянии S_1.



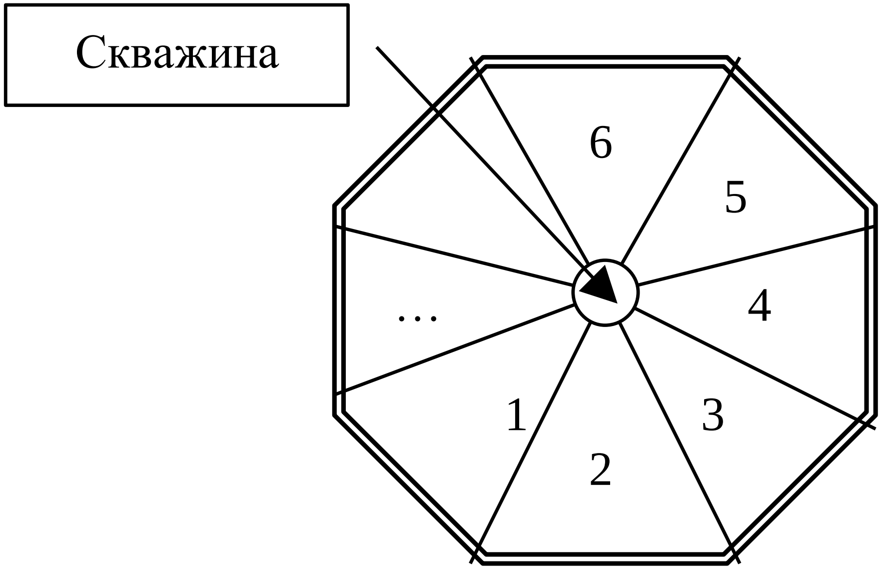
Ранг рефлексии

Определение

- Индивид имеет нулевой ранг рефлексии, если при выборе своего действия он никак не учитывает наличия других участников коллектива.
- Выбор действия при нулевом ранге рефлексии определяется только той информацией, которая поступает на вход решающего устройства из внешней среды.
- Индивид имеет первый ранг рефлексии, если он считает, что остальные участники имеют нулевой ранг рефлексии и сам он может выбирать действия за них.

Задача о поливе садовых участков

- Скважина.
- Кольцевой коллектор.
- Критерии: (1) экономия электроэнергии, (2) надо поливать.



Решение

- Кольцо из N автоматов. Каждый из них может находиться в одном из двух состояний – 0 и 1.
- Плохо, когда: 1) не экономится электроэнергия; 2) когда все засыхает.

Состояние			Вероятность наказания
Левый сосед	**Собственное**	Правый сосед	
0	0	0	1
0	0	1	0.5
0	1	0	0
0	1	1	0.5
1	0	0	0.5
1	0	1	0
1	1	0	0.5
1	1	1	1

Схема поведения автомата с 0 РР

Автомату с первым РР лучше сохранить свое текущее состояние.

Автомату с первым РР необходимо знать не только соседей, но и соседей соседей. Автомату со вторым РР – аналогично. Чем выше РР, тем о большем количестве соседей необходимо иметь информацию.

	Левые соседи		**Основной автомат**	Правые соседи		
…	1	0	**1**	1	0	…

Михаил Львович Цетлин (1924 – 1966)

Вадим Львович Стефанюк и Лотфи А. Заде
4-я Международная конференция по мягким вычисления, май 2014 года, Беркли, США.
http://raai.org/about/persons/stefanuk/

Список некоторых работ В.Л. Стефанюка по коллективному поведению и играм автоматов

- Стефанюк В.Л. Пример задачи на коллективное поведение двух автоматов//Автоматика и телемеханика. - 1963. - Т.24. - N.6. - С.781-784
- Стефанюк В.Л., Цетлин М.Л. **О регулировке мощности в коллективе радиостанций** //Проблемы передачи информации. - 1967. - Т.3. - N.4. - С.59-67.
- Стефанюк В.Л. **Некоторые локальные критерии устойчивой регулировки мощности в коллективе радиостанций** //Проблемы передачи информации. - 1968. - Т.4. - N.1. - С.90-91.
- Стефанюк В.Л., Бутрименко А.В. Игры автоматов как модель группового поведения// Тезисы статей 3-го Симпозиума по человеко-машинным проблемам: групповая активность в малых коллективах (проведен в Баку).- Научный совет по кибернетике. - Москва, 1968, С.21-23.
- Стефанюк В.Л. Коллективное поведение автоматов и задача устойчивого локального управления системой связи//Кандидатская диссертация, - М:ИПУ. - 1968.- 115 с.
- Стефанюк В.Л. Локальное управление мощностью в большой системе связи// 3-й Международный симпозиум "Новости в радиоэлектронике", Варна, 1970. - Ч.2. - С.1-7.
- Микийчук А.М.,Стефанюк В.Л. Об одном типе взаимодействия , гарантирующем глобальную устойчивость локального управления//2-е всесоюзное совещание по теории релейных устройств и конечных автоматов. Тезисы докладов - Рига. - 1971. - С.100-101.
- Стефанюк В.Л. Об описании игр ε-автоматов //Автоматика и телемеханика. - 1971. - N.4. - С.83-88.
- Стефанюк В.Л. О "взаимопомощи" в коллективе радиостанций// Проблемы передачи информации. - 1971. -Т.7. - N.3. - С.103-107.
- Stefanuk V.L. Collective Behaviour of Automata and the Problems of Stable Local Control of a Large Scale System//Second International Joint Conference on Artificial Intelligence (IJCAI), London, pp. 51-56, 1971.
- Котляр С.Б., Стефанюк В.Л. Об одной упрощенной модели взаимодействия в коллективе автоматов//Problems of Control and Information Theory. - 1972. - V.1(3-4). - С.297-305.
- Стефанюк В.Л. О взаимодействии при локальном управлении// Автоматика и телемеханика. - 1973. - N.6. - С.48-56.
- Стефанюк В.Л. Анализ целесообразности локально-организованных систем через потоки вероятности// Модели в системах обработки данных . - М.: Наука, 1989 . - С.33-45.
- Стефанюк В.Л. Равновесие в дробно-линейной системе взаимодействия при локальных данных//Модели в системах обработки данных . - М.: Наука, 1989 . - С.45-54.
- Стефанюк В.Л. Консультирующая экспертная система с локальной организацией// Всесоюзная конференция "Проблемы разработки и внедрения экспертных систем". - М.:ВНИИНС, 1989. - С.33-34.
- Стефанюк В.Л. Локальная организация целесообразного поведения технических систем.- М:МИЭМ. - Докторская диссертация. - 1990. - 423с.
- Стефанюк В.Л. От многоагентных систем к коллективному поведению. Труды международного рабочего совещания "Распределенный искусственный интеллект и многоагентные системы" (DAIMAS'97), 1997, С. Петербург, С. 327-338
- Stefanuk V.L. From Multi-Agent Systems to Collective Behaviour. In Proc. of the Workshop "Distributed Artificial Intelligence and Multi-Agent Systems (DAIMAS'97), June 15-18, 1997, St.Petersburg, Russia, p. 223 (in English).
- Стефанюк В.Л. Коллективное поведение систем с переспросом. Научная сессия МИФИ-2006. Сборник научных трудов, Т3: Интеллектуальные системы и технологии, Министерство образования и науки Российской федерации, Москва: МИФИ, 2006, с. 44-45
- Стефанюк В.Л. Локальная организация интеллектуальных систем. Модели и приложения. М.: Физматлит, 2004 , 328 с.
- Стефанюк В.Л. Коллективное поведение в противовес перемешиванию мнений. 3-я международная конференция "Системный анализ и информационные технологии", САИТ-2009: Труды конференции, С. 211-218, ИСА РАН, 2009
- Vadim Stefanuk. Reaching Collective Opinion. International Journal of Computational Intelligence. Theory and Practice, Vol. 5, No. 1, June 2010. pp. 31-35.