


# НЕЙРОСЕМАНТИЧЕСКАЯ ФОРМА ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ



**Бодякин В.И.**

Институт проблем управления РАН  
им. В.А. Трапезникова, Москва

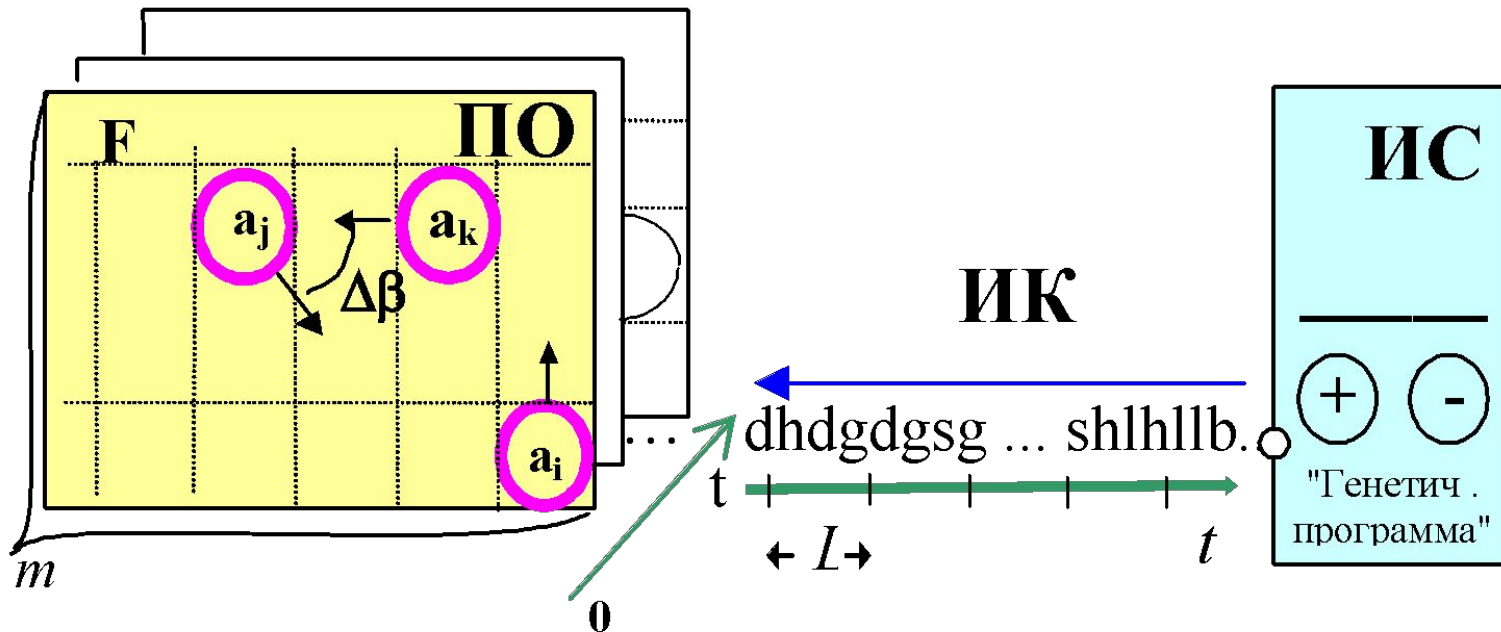
E-mail: [body@ipu.ru](mailto:body@ipu.ru) ,  
<http://www.informograd.narod.ru> ,  
служ.тел.:334-92-39

**Комплекс:**

**"Предметная область –**

**Информационный канал –**

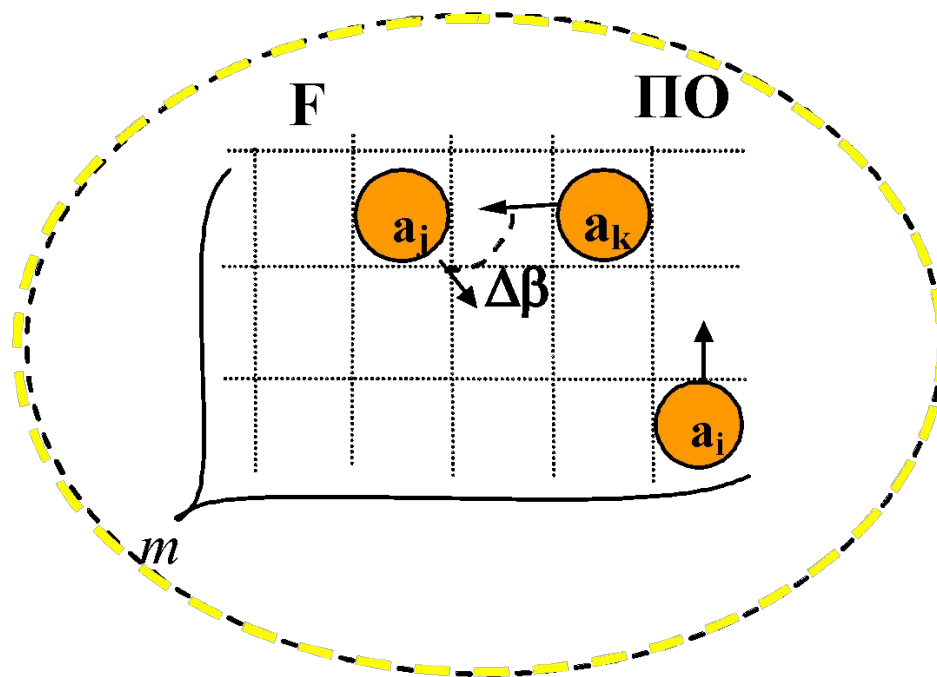
**Информационная система"**



## Предметная область (ПО)

Любую физическую ПО можно рассматривать как некоторую дискретную пространственно-временную область с взаимодействующими объектами ( $a_i, a_j$ ).

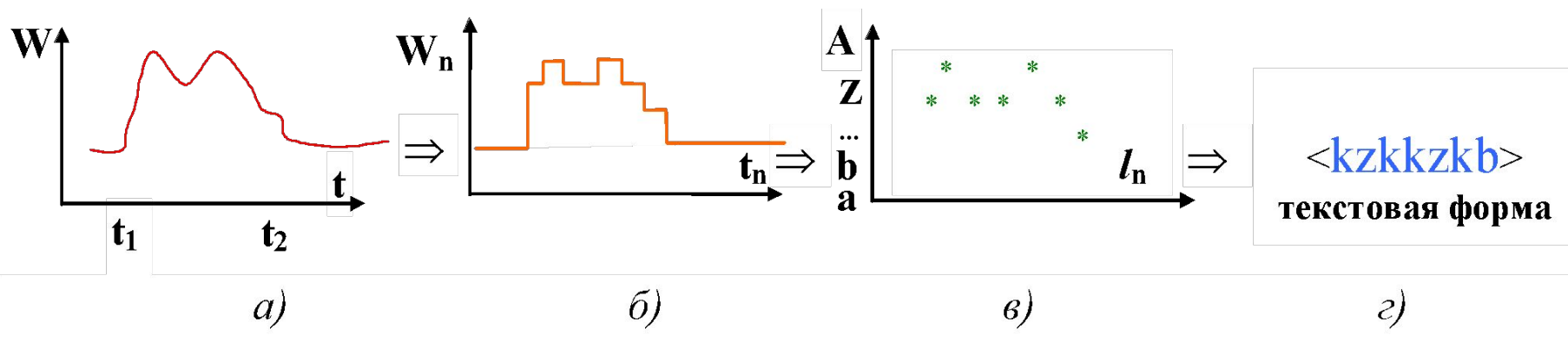
При попадании нескольких объектов, в одну и ту же ячейку (область взаимодействия), происходит процесс их взаимодействия.



**Процесс взаимодействия** объектов осуществляется в течение нескольких тактов времени  $t_2-t_1$ .

**Результат взаимодействия** объектов выражается в изменении величины одного или нескольких параметров  $W$ , характеризующих их взаимодействие.

Любой  $W(t)$  может быть преобразован в **текстовую форму**.



**многомерный физический процесс <-> текстовая форма**

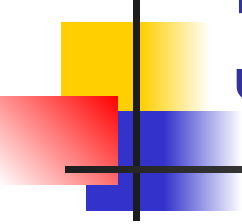
# Энергетика физического процесса

— это

некоторый **инвариант**, характеризующий потенциальную глубину причинно-следственного распространения данного физического процесса

---

- Предположим, что **скорости распространения** разнообразных физических процессов **различны**.
- Мы будем рассматривать только такие ПО, в которых существуют **малоэнергетические сопутствующие процессы с большими скоростями распространения**, относительно основного физического процесса.



Будем предполагать,  
что в любой ПО выполняются:

---

- **принцип причинности** (инвариантность во времени и в пространстве), т.е. если  $si \rightarrow sj$ , то процесс  $si$  всегда вызывает  $s_j$ ;
- **принцип глобальной дискретности** (гладкости) процессов взаимодействия, т.е. если  $si (<zzjbaabj>) \sim sk (<zzjbab>)$ , и  $si \rightarrow sj$ , то и, скорее всего,  $sk \rightarrow sj$ ;
- **принцип субъектности**, т.е. каждый процесс ПО ( $si$ ) имеет оценку полезности для ИС.

При взаимодействии объектов (например,  $\{a_i * a_j\}$ ), в соответствии с  $F(\text{ПО})$ , порождается **ЭСЕ - элементарная семантическая единица**

$$(s\{a_i * a_j\} = \langle \text{cade...b} \rangle),$$

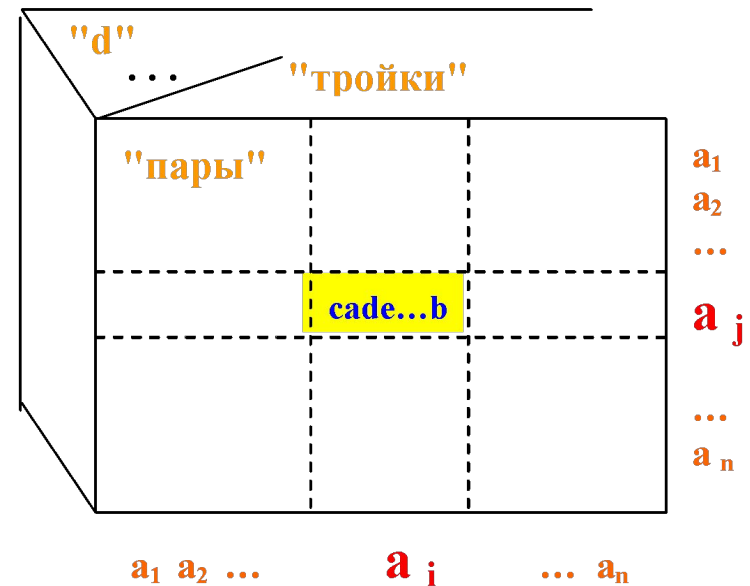
которая однозначно порождается процессом взаимодействия этих объектов. Все множество взаимодействий в данной ПО, представляет собой **семантическое пространство**, как совокупность ЭСЕ.

В качестве примеров ЭСЕ различных ПО можно привести:

взаимодействие элементарных частиц –  $\langle \text{ababcw} \rangle$ ;

$\langle \text{Привет Петров, как дела? Нормально!} \rangle$  – взаимодействие объектов макроуровневых ПО;

образование двойной звезды из двух астрообъектов –  $\langle \text{zzjbaabj} \rangle$  (на гигауровне).

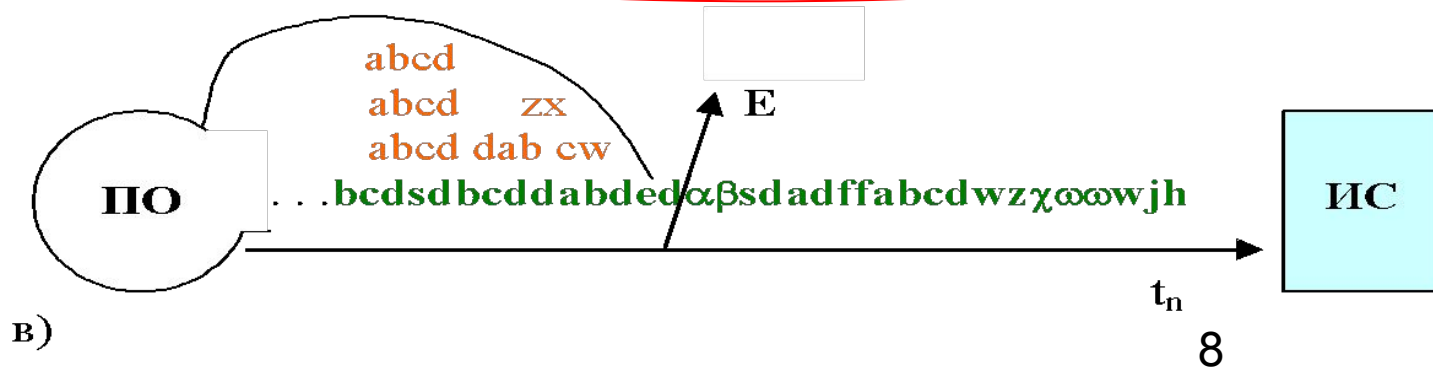


ЭСЕ, порождаемые в ПО, суперпозиционно "сливаются" в непрерывный информационный поток.

## Информационный канал (ИК).

В ИК происходит процесс формирования информационного ресурса и его транспортировка к ИС.

Частота взаимодействия объектов в ПО определяет плотность информационного потока в ИК.



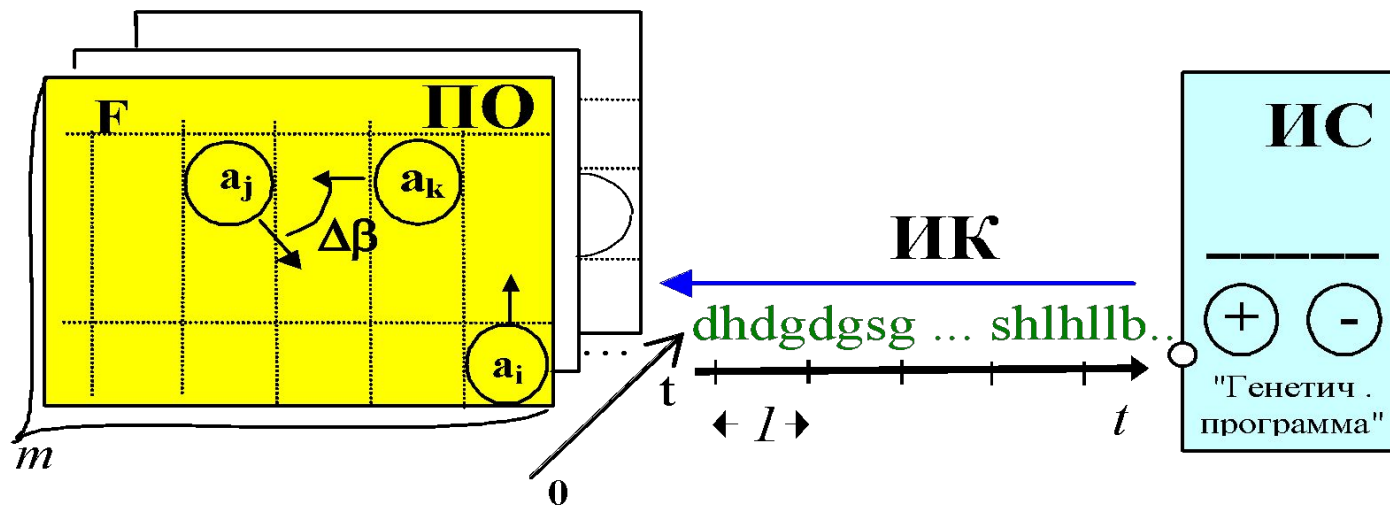


В рамках рассматриваемого комплекса:

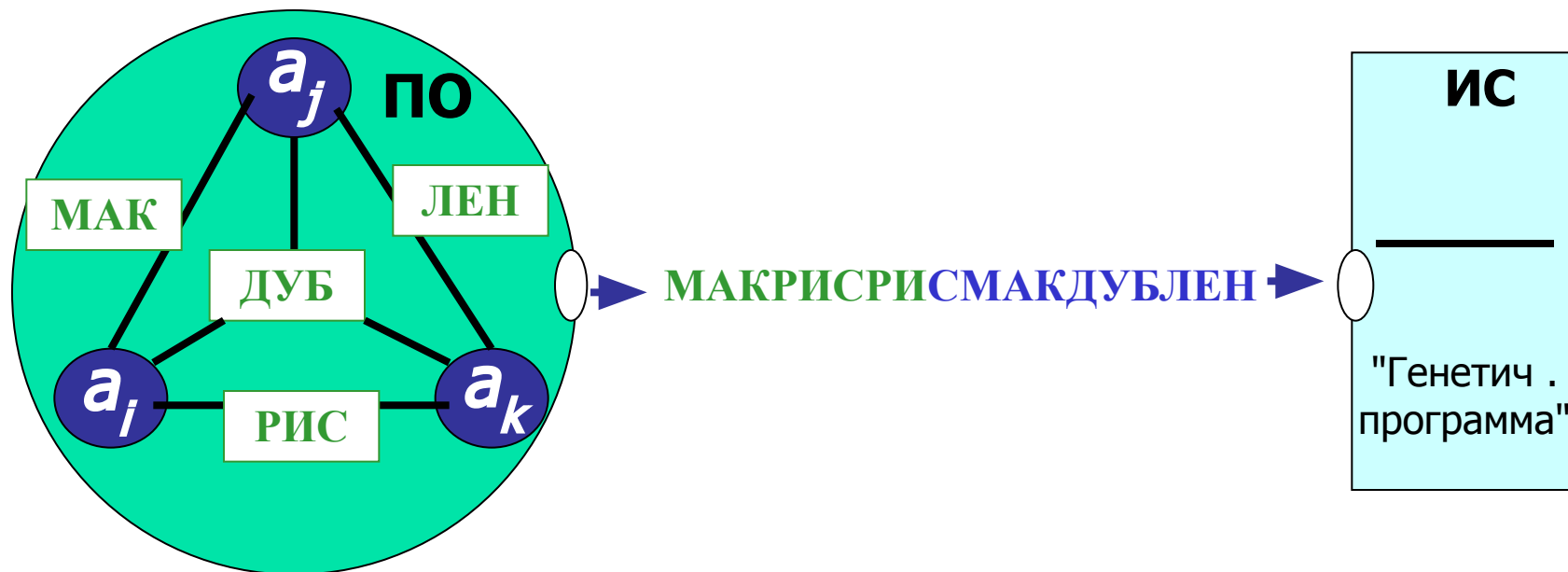
ПО → ИК → ИС,

задача ИС заключается в:

1. Воспринимать текстовую форму
2. Хранить информацию
3. Преобразовывать (качественно) информацию
4. Оценивать информацию
5. Выдавать информацию (текстовую форму)



# Демонстрационный пример



Энергетика каждого прогнозируемого символа равна  $1E+$ ,

Время обработки символа – **один такт**,

Энергетические затраты на обработку одного образа в ИС равна  $1E-$ .

Необходимо построить словарь образов, **полностью покрывающий ТФ**.

Каждые **12 тактов** все образы потребляют на самообеспечение по  $0,5E-$

# Результаты эксперимента

## Первая структуризация словаря ИС

---

Форма словаря: (наибольший размер образа один символ):

- $\langle M \rangle \langle A \rangle \langle K \rangle \langle P \rangle \langle I \rangle \langle C \rangle \langle L \rangle \langle E \rangle \langle H \rangle \langle D \rangle \langle Y \rangle \langle B \rangle$ ,
- $R(ИС) = 12R * (0,5E-/R) = 6E-$ ,
- прогнозирование (Т) = 0Т (т.к. у образа только один символ),
- затраты энергии на распознавание = 12E-
- контролирование потенциальной энергии ТФ = 0E+.
- **Итог: для односимвольного словаря = 18E- .**



## Вторая структуризация словаря ИС

---

- а) Минимальная форма словаря: <МА><К><РИ><С><ЛЕ><Н><ДУ><Б>:
- $R(ИС) = 8R = 4E^-$ , прогнозирование =  $0,5T$  (на образ),
- затраты энергии на распознавание =  $8E^-$ , контролирование потенциальной энергии ТФ =  $4E^+$ .
- *Итог а)*  $4E^- + (0,5T * 8(\text{образов на ТФ}=12) = 4E^+) + 8E^- = 8E^-$ .
  
- б) Максимальная форма (без пересечения): <МА><КЛ><КР><КД><КМ><ЛЕ>...<БД>:
- $R(ИС) = 20R = 10E^-$ , прогнозирование =  $0,2T$  (на образ),
- затраты энергии на распознавание =  $20E^-$ ,
- контролирование потенциальной энергии ТФ =  $4E^+$ .
- *Итог б)*  $10E^- + (0,2T * 20(\text{образов на ТФ}=12) = 4E^+) + 20E^- = 26E^-$ .
  
- **Итоговый лучший эволюционный потенциал =  $8E^-$ .**

## Третья структуризация словаря

(наибольший размер образа в три символа).

- а) Минимальная форма словаря: <МАК><РИС><ЛЕН><ДУБ>:
  - $R(ИС)=4R = 2E^-$ , прогнозирование = 2Т (на образ),
  - затраты энергии на распознавание =  $4E^-$ ,
  - контролирование потенциальной энергии ТФ =  $8E^+$ .
  - *Итог а)  $2E^- + (2Т * 4(образов на ТФ=12) = 8E^+) + 4E^- = 2E^+$ .*
- б) Максимальная форма (без дублирования, т.е. без полного пересечения): <АКР><ИСП><ИСМ><АКД> <УБЛ><ЕНД> ... <АКМ>:
  - $R(ИС)= 36R=18E^-$ ,
  - прогнозирование (Т) = 0Т (на образ),
  - затраты энергии на распознавание =  $36E^-$ ,
  - контролирование потенциальной энергии ТФ =  $0E^+$ .
  - *Итог б) =  $54E^-$ .*
- Лучший итоговый эволюционный потенциал =  $2E^+$  ,  $54E^- \rightarrow 2E^+(!!)$ .

## Четвертая структуризация словаря

(наибольший размер образа в четыре символа)

- а) Минимальная форма словаря: <МАК><РИС><ЛЕН><ДУБ>:
- $R(ИС)=4R = 2E^-$ , прогнозирование =  $2T$  (на образ),
- затраты энергии на распознавание =  $4E^-$ ,
- контролирование потенциальной энергии  $T\Phi = 8E^+$ .
- *Итог а)*  $2E^- + (2T * 4(\text{образов на } T\Phi=12) = 8E^+) + 4E^- = 2E^+$ .
  
- б) Максимальная форма (без дублирования):  
<АКРИ><ИСРИ><ИСМА> <АКДУ><УБЛЕ><ЕНДУ> ... <АКМА>:
- $R(ИС)= 48R=24E^-$ , прогнозирование ( $T$ ) =  $0T$  (на образ),
- затраты энергии на распознавание =  $48E^-$ ,
- контролирование потенциальной энергии  $T\Phi = 0E^+$ .
- *Итог б)* =  $96E^-$ .
  
- Лучший итоговый эволюционный потенциал =  $2E^+$  ,  $96E^- \rightarrow 2E^+$



## Пятая и другие структуризации словаря

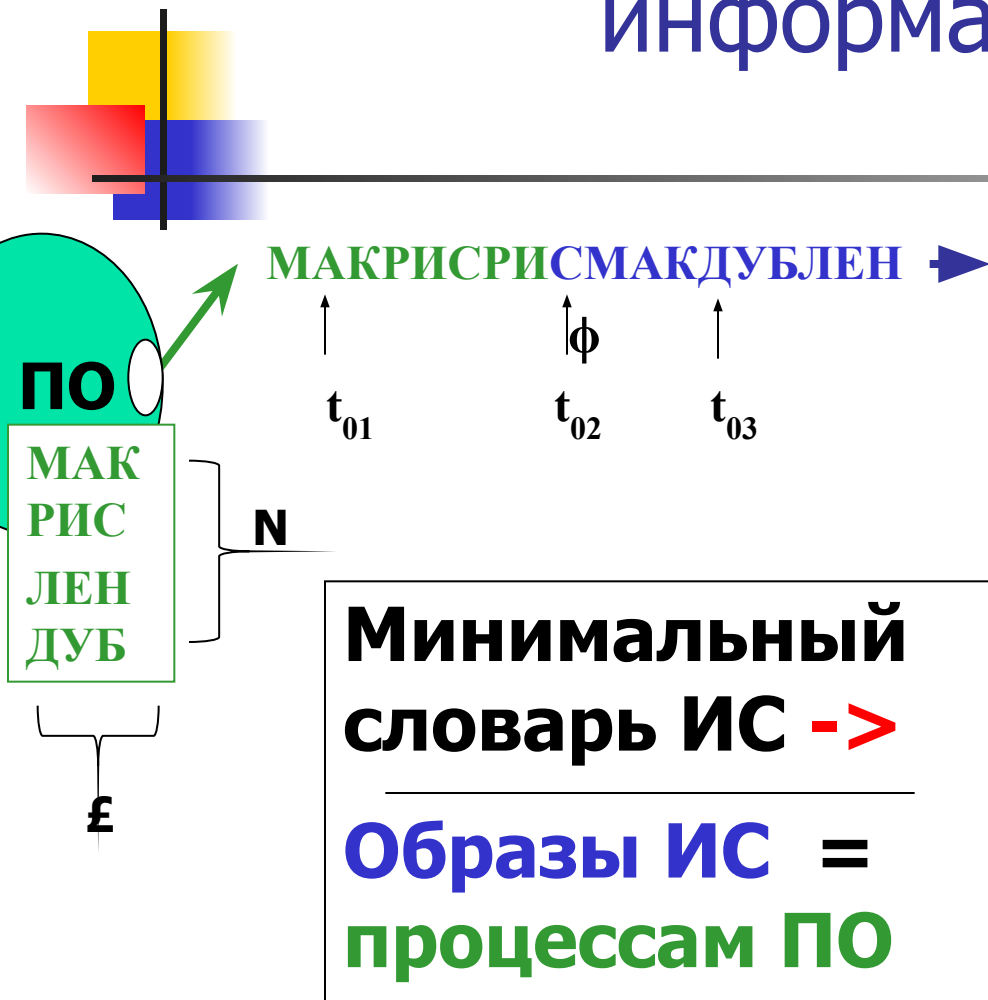
---

Лучший итоговый эволюционный потенциал  
 $2E+$  ,  $204E-$  →  $2E+$

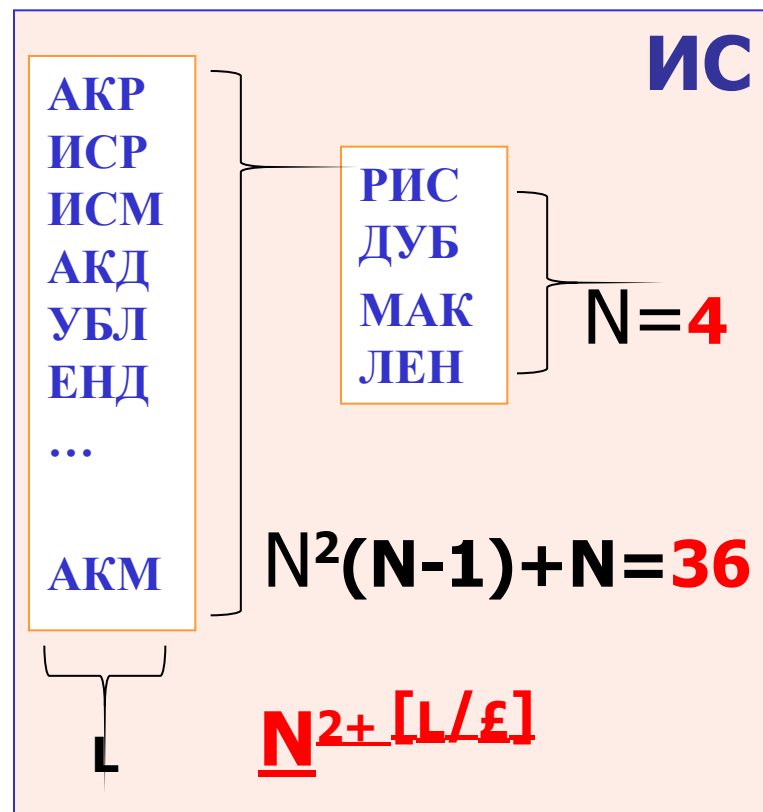
Шестая структуризация словаря, седьмая ... и т.д. →  $2E+$  !!!

- Теоретический *анализ результатов эксперимента* показывает, что эволюционный потенциал ИС *обратно пропорционален размеру словаря*,
- Размеры минимальных и максимальных словарей соотносятся как:
- $o(N)$  и  $o(N^3)$  !!!
- где:  $N$  – максимальный размер образа словаря

# Автоструктуризация информации в ИС



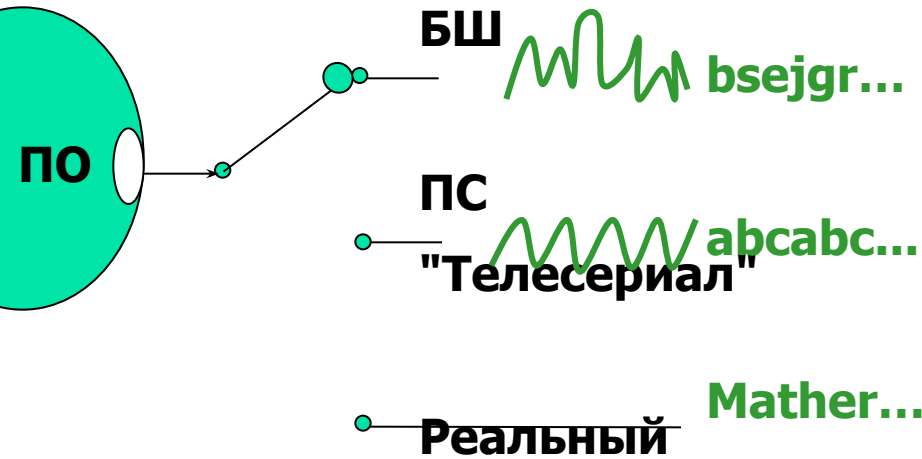
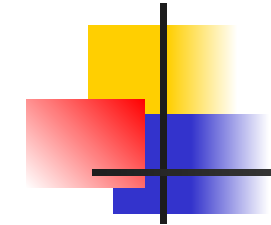
**ИНФОРМАЦИЯ ПО – это есть образы ИС**



**$o(N^3) \gg o(N) !!!$**



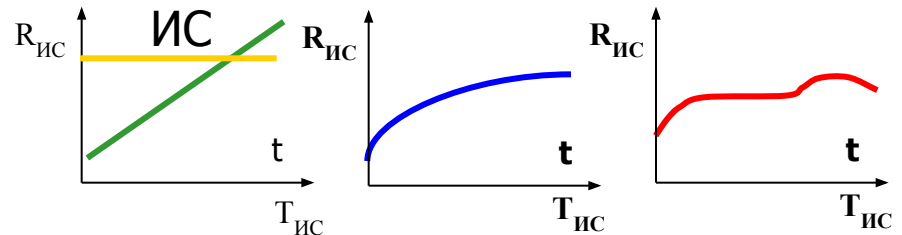
# Отображение информационного ресурса тремя классами ИС



Автомат "животн." ИС-человек

Сигнал	С	С
С	Информация	И
С	И	Знание

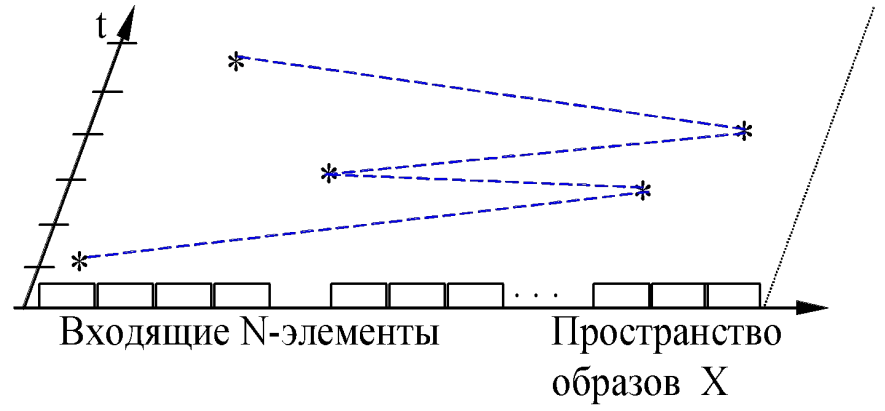
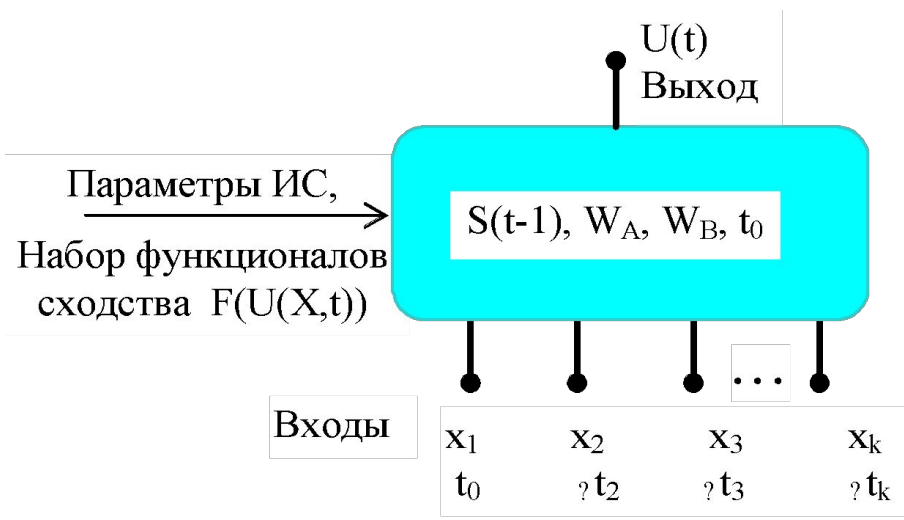
При  $T_{ИС} = \text{const} (t)$



$f = (R_{ИС} (t)) :$  линейный; логарифмический; const;

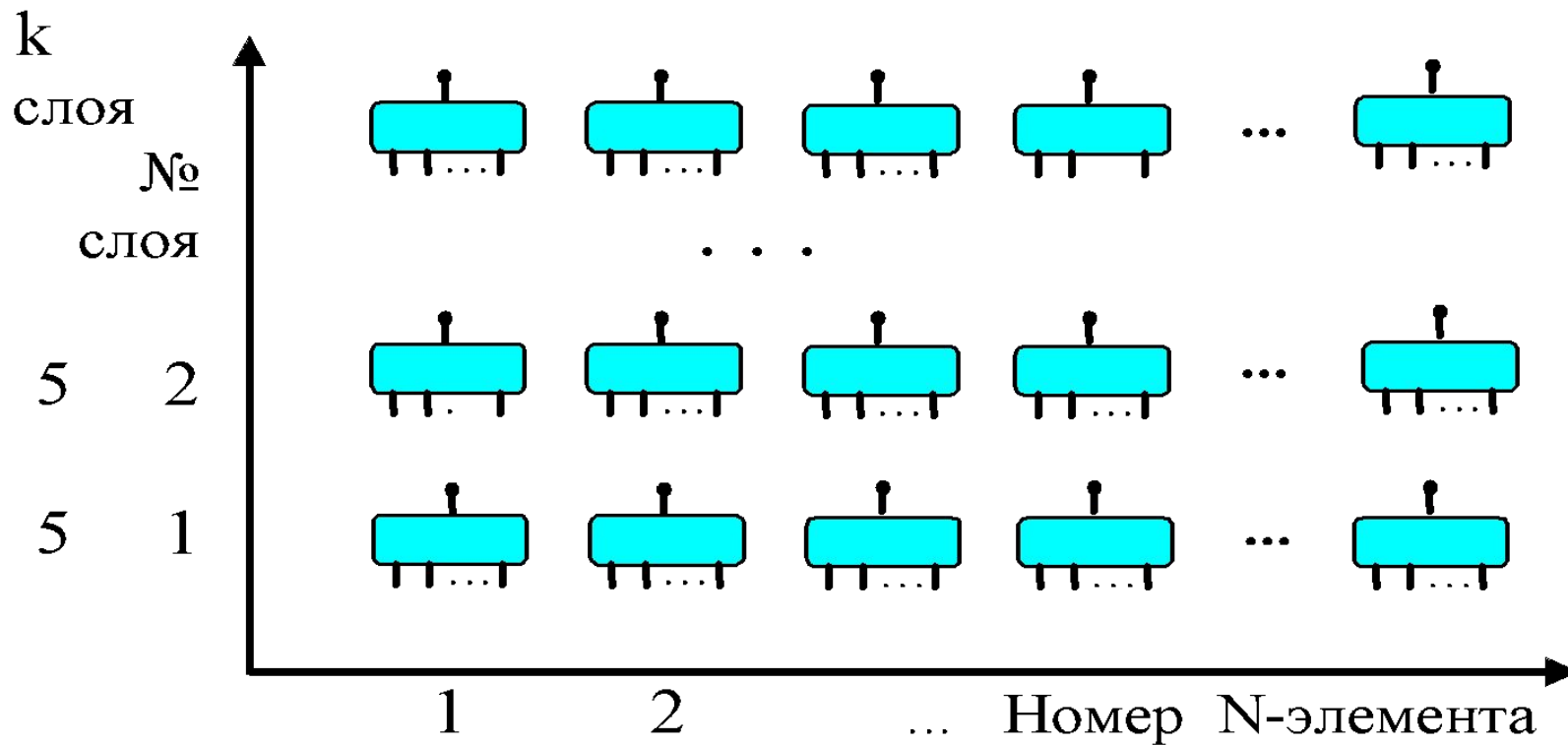
Решением проблемы **автоструктуризации** стало расширение понятия формального нейрона (МакКаллока-Питтса от 1943г.), вводом в него **относительности времени активации входов**, что позволило получить **нейроподобный N-элемент**.

что



$$U(t) = F_i(U(X,t)), \quad F_i(t) = F_{ИС}(U(t-1))$$

Объединив  $N$ -элементы в потенциальный многодольный иерархический граф, удалось получить структуру аналогичную естественно-языковым.



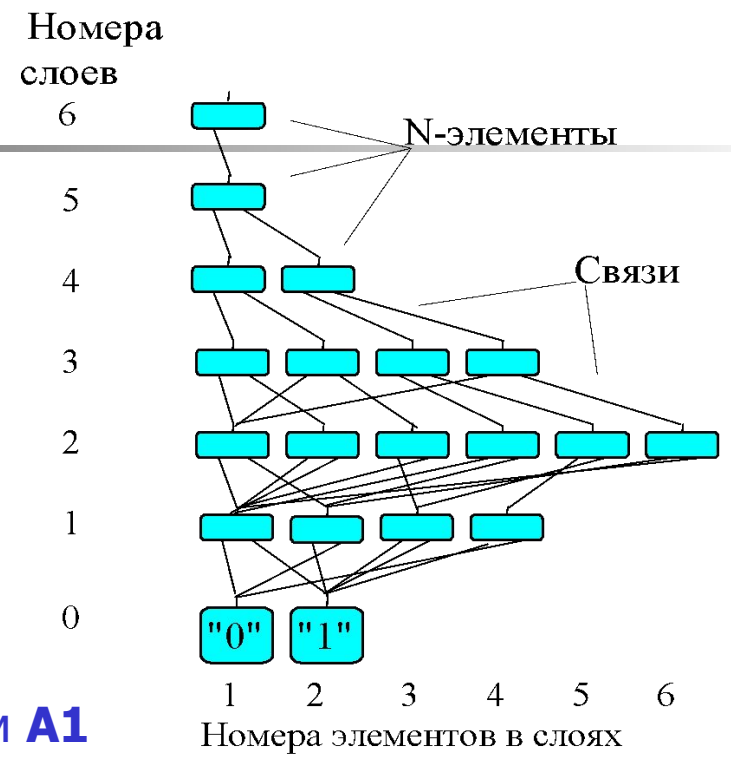
# Структурное описание НСС можно представить в виде многодольного графа:



01100101011011011010001101101001    □ □

Текстовая форма

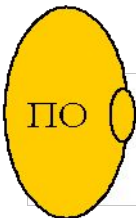
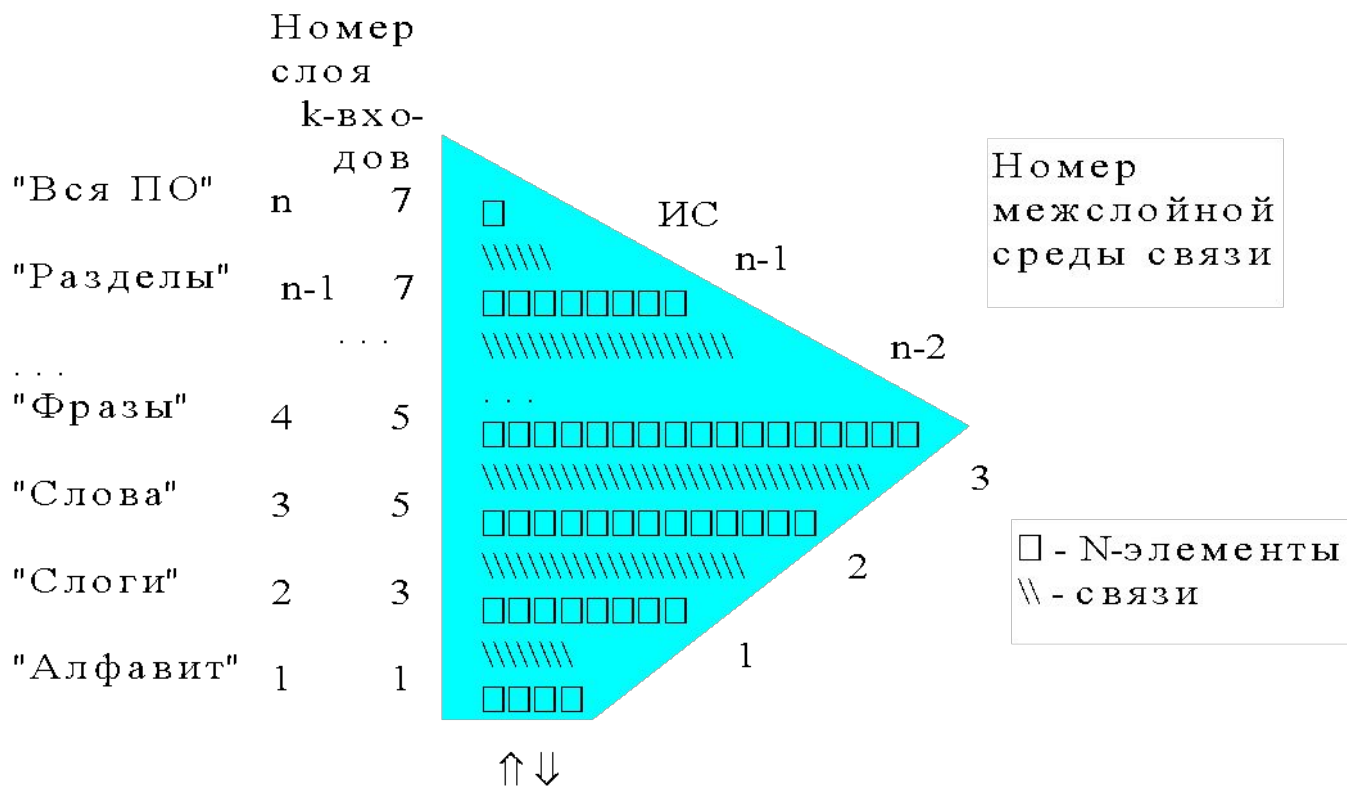
Алгоритм А1



Форма НСС

**Алгоритм обратного преобразования НСС в текстовую форму** осуществляется уже за меньшее число операций и идет "сверху-вниз".

Первый слой (доля графа) N-элементов – терминальный, фактически отображает **алфавит A ЭСЕ**, второй слой – "**псевдослоги**" и строится на пространственно-временных ссылках на предыдущий (терминальный) слой - **информационное содержание N-элемента**, слой "**псевдослов**" – ссылается на "псевдослоги" и т.д., до самого верхнего N-элемента, отображающего в себе через связи всю ПО.



→ **abcklltfdbababafgkdkediejddjchhc**  
 Текстовой поток

# Автоструктуризация

$R_{ИС} = f(\text{число } N\text{-элементов, число связей})$  в битах

$1/P$  (компрессия) = -----  $\rightarrow 0$   
при  $t \rightarrow \infty$   $T\Phi_{ИС} = \text{объем текстовой информации в ИС}$  в битах

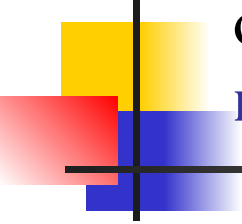
или  $\Delta R_{ИС} / \Delta T\Phi_{ИС} \rightarrow 0$  и  $\Delta R_{ИС} \rightarrow \text{const}$ , при  $t \rightarrow M$   
при  $t \rightarrow \infty$

Пример:

<ДОМЗЕБРЫСКИТНАДОМДОМВНАДОМВСКИТВНАСКИТВВЗЕБРЫНАВНА>, правильно выделяются все **ЭСЕ**: <ЗЕБРЫ> <СКИТ> <ДОМ> <НА> <В>.

сдвиг алфавита А в кодах ASCII

<ЕПНИЖВСЬТЛЙУОБЕПНЕПНГОБЕПНГТЛЙУГОБТЛЙУГЖВСЬОБГОБ> на +1  
<?IGB@<KVLECMH;?IG?IG=H;?IG=LECM =H;LECM==B<KVH;=H;{> на -133.



Структуру памяти ИС, в которой выполняется свойство гомоморфного отображения ЭСЕ ПО и их структуры в образы ИС и обратно, будем называть **нейросемантической структурой (НСС)**

---

**N-элемент (образ ИС) ↔ ЭСЕ (ПО)**

Назовем процесс **автоматического** выделения семантических единиц в НСС –

**автоструктуризацией.**

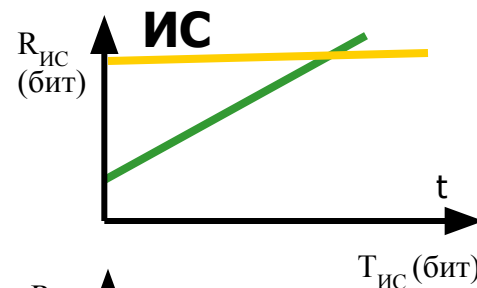
НСС – это готовая **структура данных** (процессов и объектов) произвольной ПО для любой ИС.

Понятно, что ее автоматическое формирование открывает широкие горизонты для инженерии ИС.

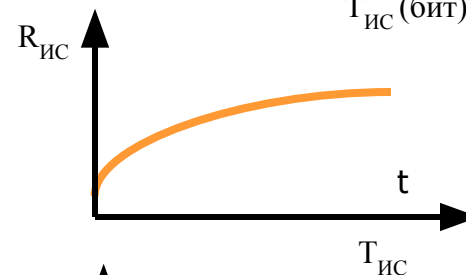
Информационный ресурс в ИС можно представить как:

- "сигнал" или **текстовая форма** – простая суперпозиция ЭСЕ ПО;
- "информация" – сигнал, структурированный на иерархию ЭСЕ ПО;
- "знание" – иерархия НСС в ИС.

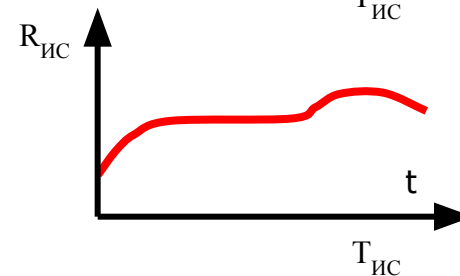
"линейный – **сигнал**",



"логарифмический  
– **информация**"



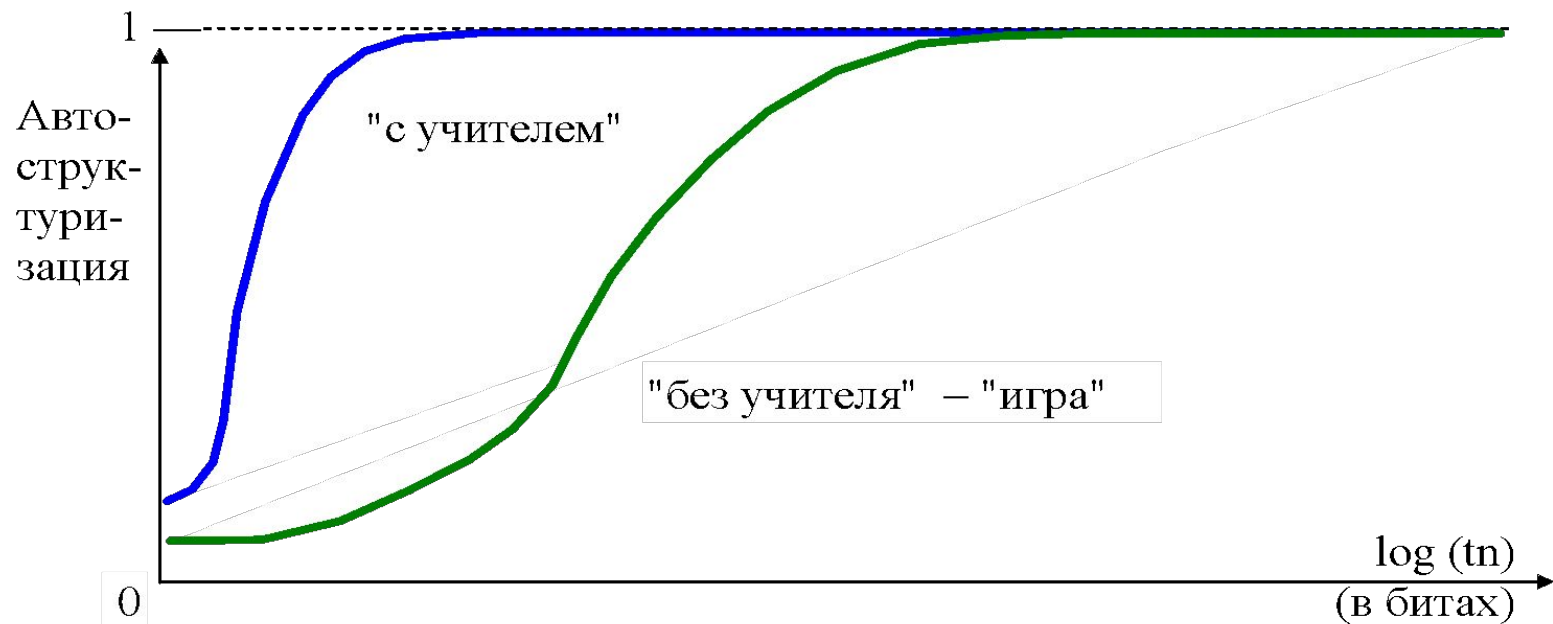
"const - **знание**".



$$\text{При } T_{ИС} = \text{const } (t), \quad R_{ИС} = f(S_{ИС}(t))$$



НСС – это пример **1-го формального преобразования** количественной текстовой формы представления информации в качественно новую форму – структуру ЭСЕ

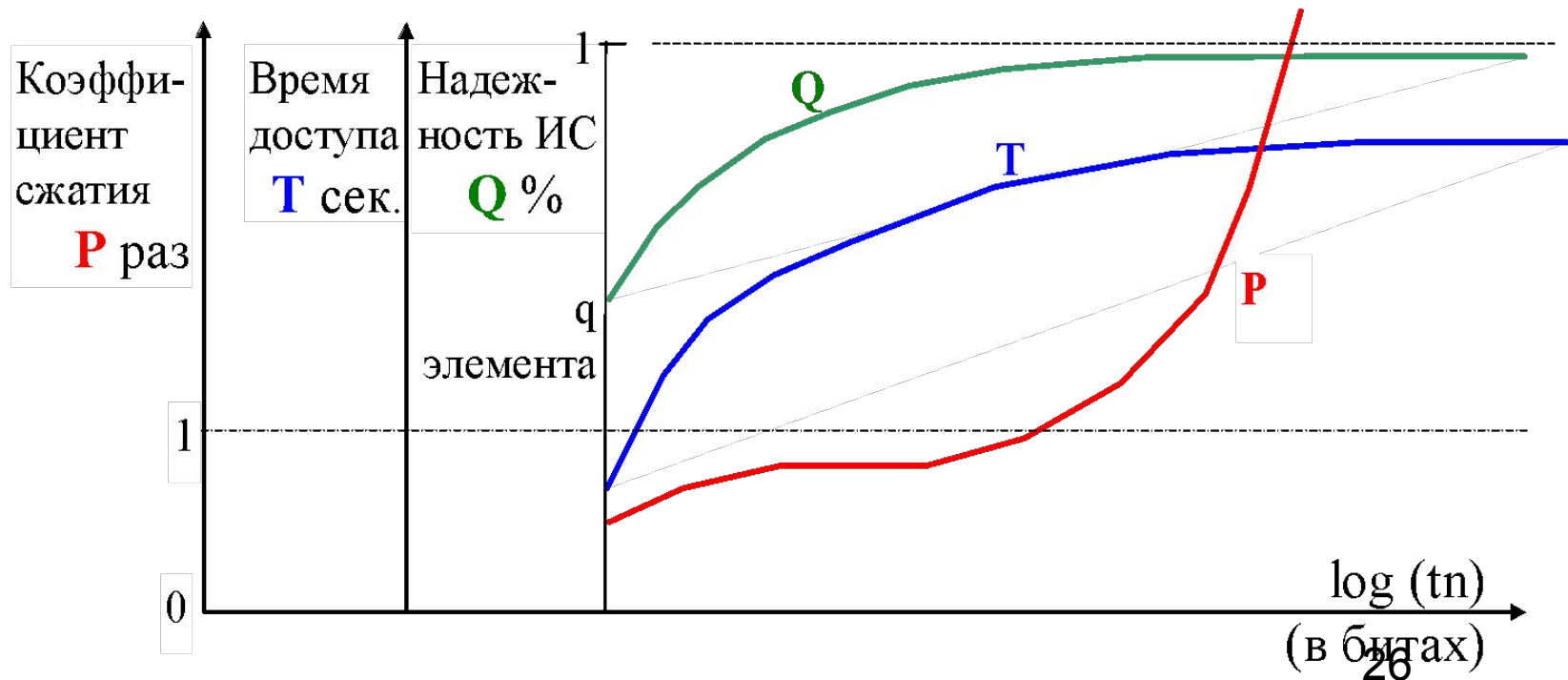


**Критерии достаточности:** а) все пространство состояний; б) если человек может правильно структурировать данный текстовой материал в непривычной, но взаимнооднозначной нотации, в) наличие характерных особенностей динамического процесса при минимизации ресурса  $R_{ИС}$

Следует также отметить, что все технические характеристики ассоциативной памяти на базе НСС:

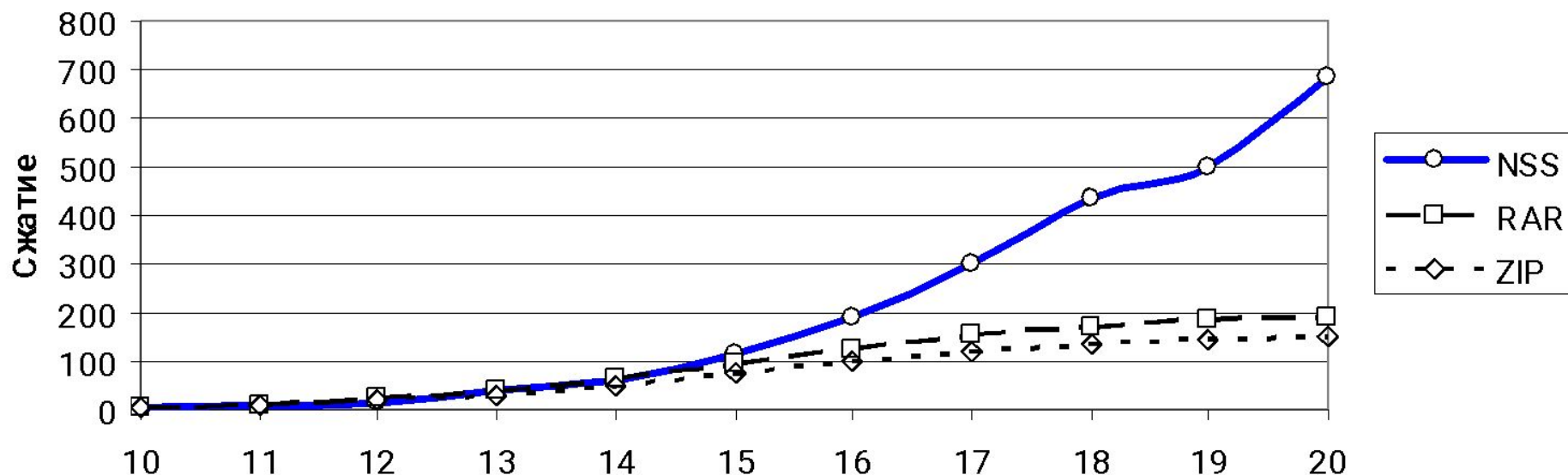
- *время доступа*,
- *коэффициент компрессии-сжатия*,
- *надежность хранения информации* и др.)

имеют тенденцию **к улучшению**, как в среднем, так и в абсолютных значениях, **по мере роста объема вводимой информации** из ПО.

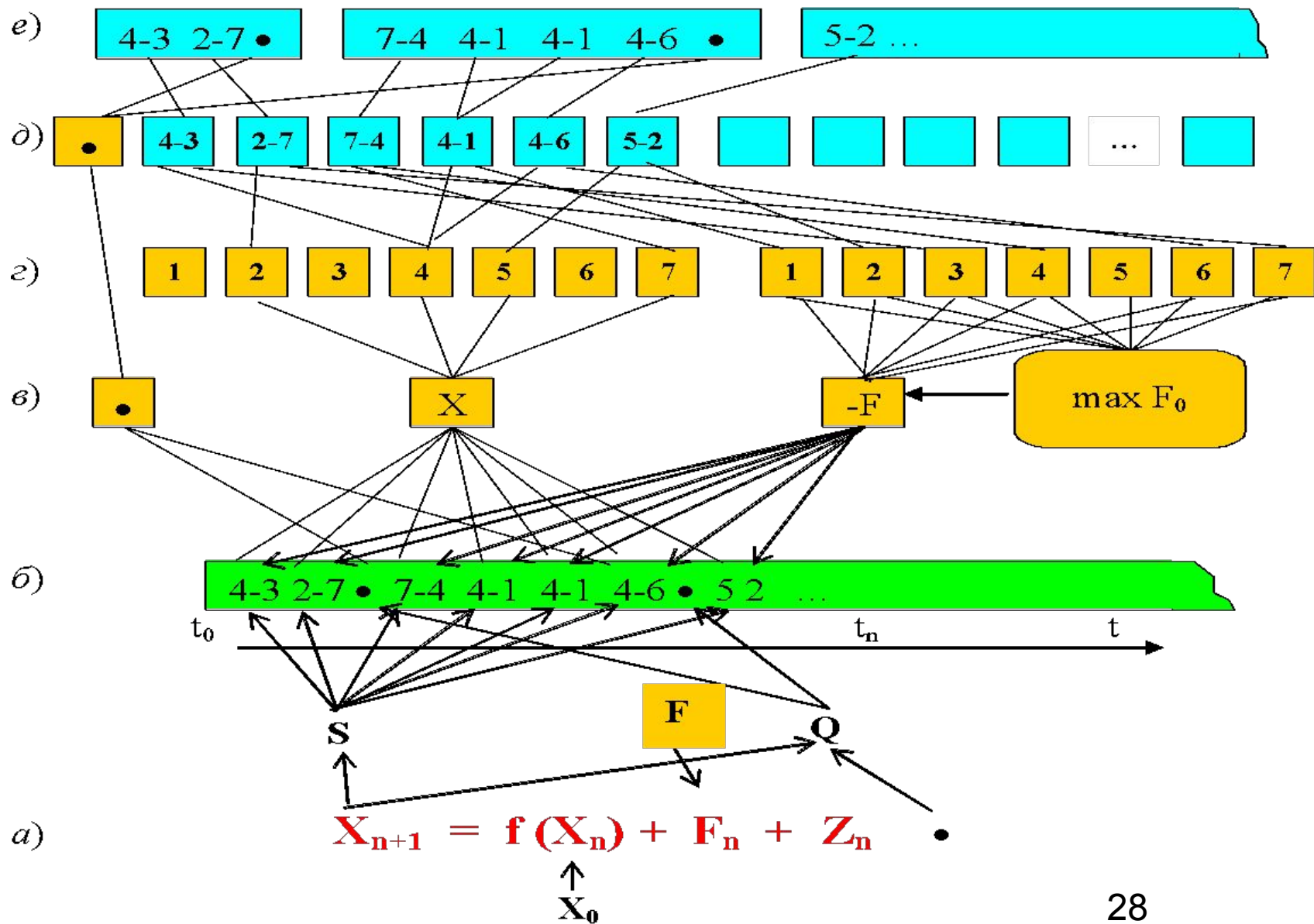


# Величина компрессии отражает **потенциальную интеллектуальность ИС.** Псевдофрактальные файлы.

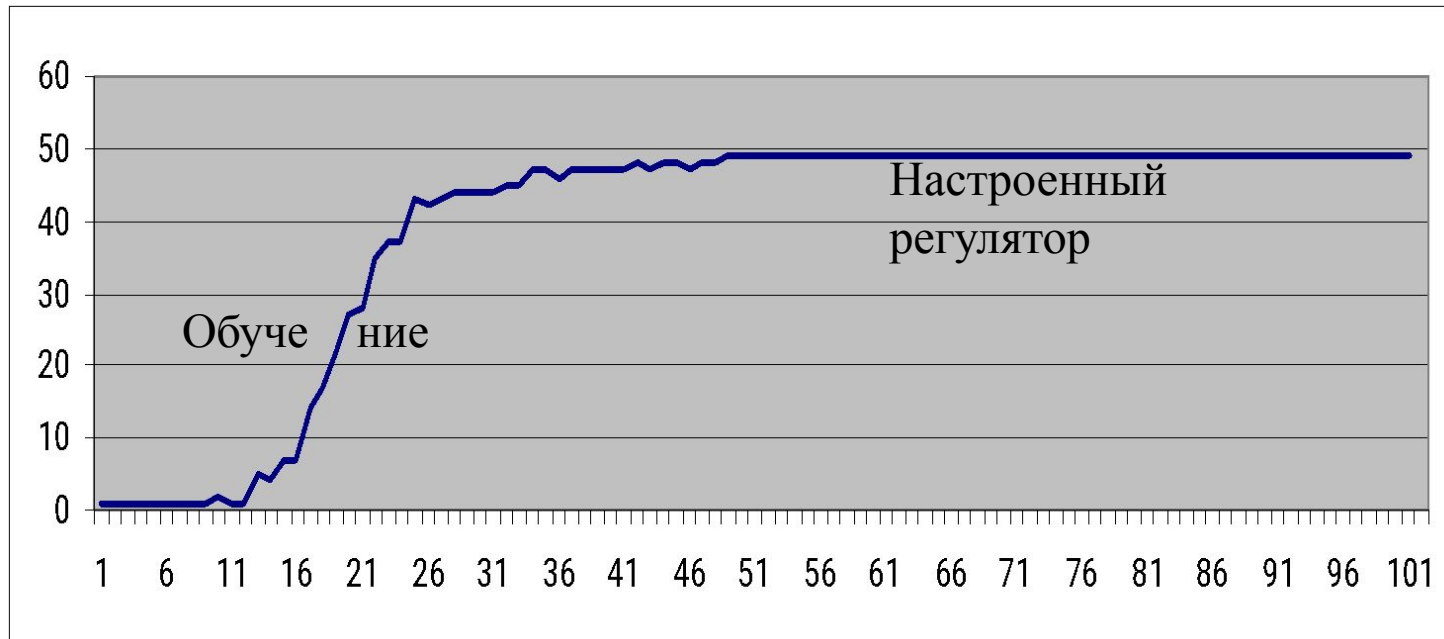
Зависимость степени сжатия от размера файла, как  $\text{Log}_2(X)$  байт



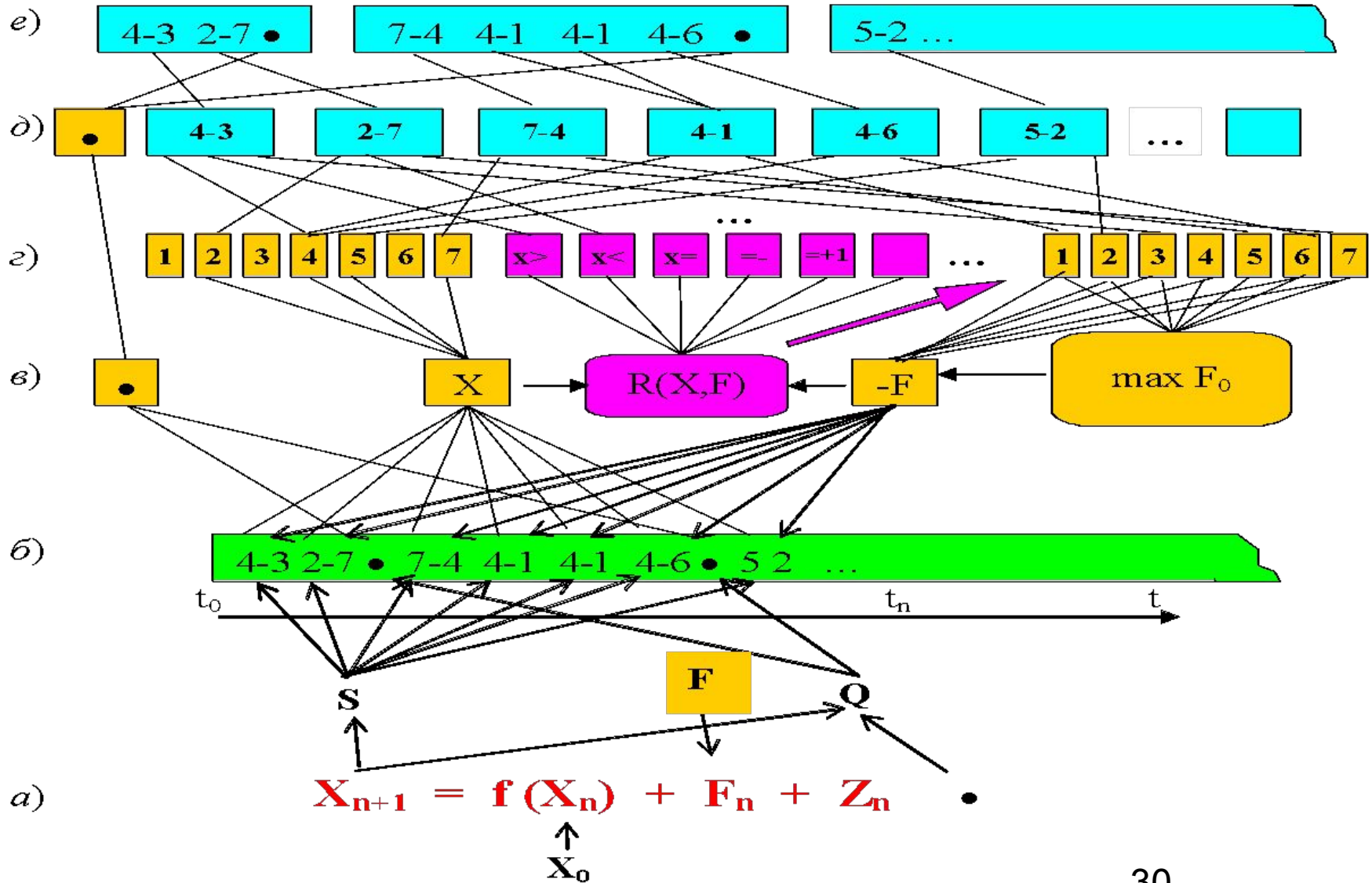
# Адаптивный регулятор №1 на базе НСС.



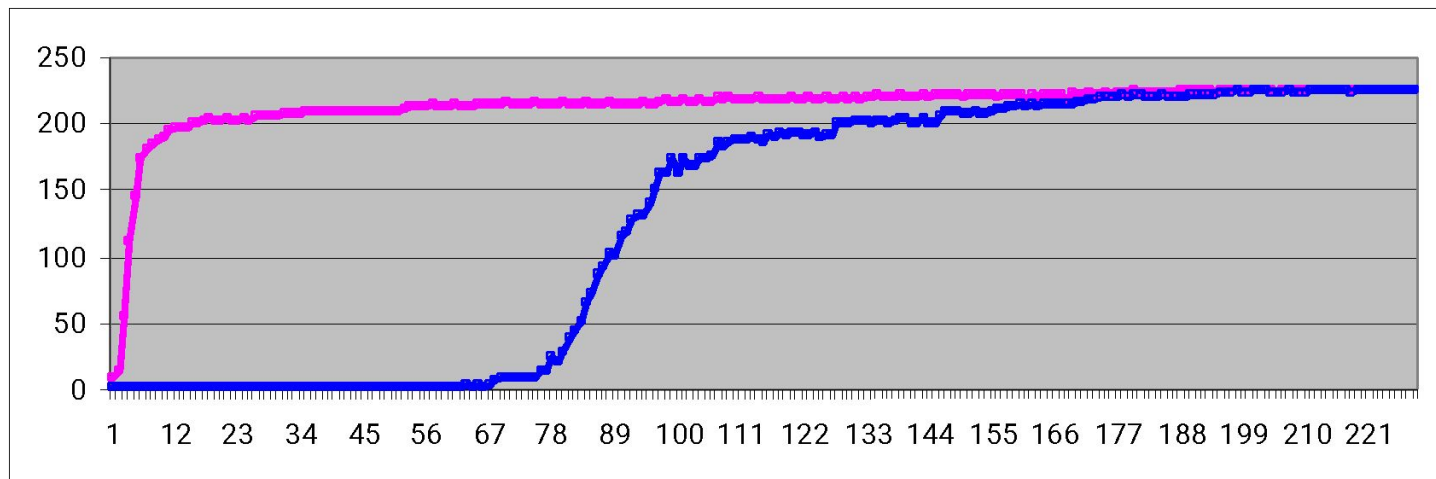
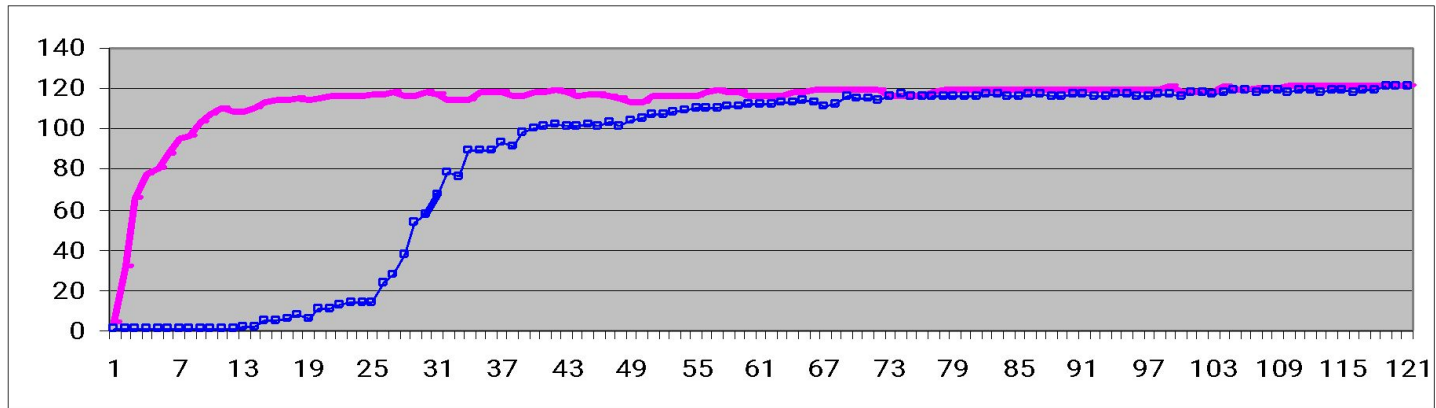
Запоминая пары  $\langle X \rangle \langle Y \rangle$  и их оценку  $\langle E \rangle$  для любого априорно неизвестного объекта управления, регулятор №1, фактически перебором, заполняет все возможное пространство его состояний ( $N \cdot M$ ).



# Адаптивный регулятор №2



Скорость обучения регулятора №2 почти **не зависит от размерностей N и M** (N – число состояний объекта, M – число состояний возможного управления, в эксперименте для простоты принималось  $N = M = 7, 15, 17, 19, 21$ ), Т.е., пример регулятора №2 демонстрирует возможность практического **преодоления "проклятия размерности"**.



# Адаптивный регулятор №3

На базе нейросемантического регулятора №3 формально показано, как на основе вышеописанных физических свойств N-элементов возможно естественное **самоформирование R-отношений**, представленных в регуляторе №2. Тем самым, подведено теоретическое основание для инженерного построения ИР, т.е.:

**теоретически решить (закрыть) проблему построения ИИ.**

Как и вычислительная техника начала развиваться с теоретических моделей "машин Тьюринга и Поста", так и анализ работы нейросемантического регулятора будет способствовать формированию широкого фронта научных работ по разработке ИР.

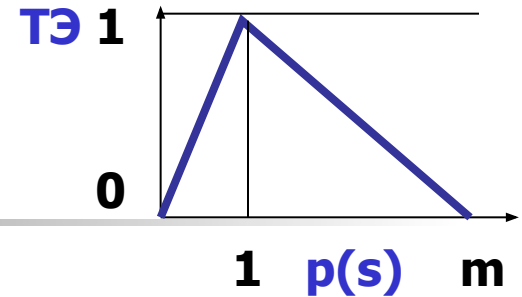
Таким образом, на нейросемантических регуляторах можно продемонстрировать:

- №1 – "**адапционность**", как возможность адаптироваться в любой ПО;
- №2 – "**интеллектуальность**", как возможность существенного сокращения перебора;
- №3 – "**разумность**", как целенаправленное порождение нового знания.



# Текстовая энтропия

$$TЭ(s) \begin{cases} = p(s), & \text{при } p(s) \leq 1 \\ = 1 - (p(s) - 1) / (m-1), & \text{при } p(s) > 1 \end{cases}$$



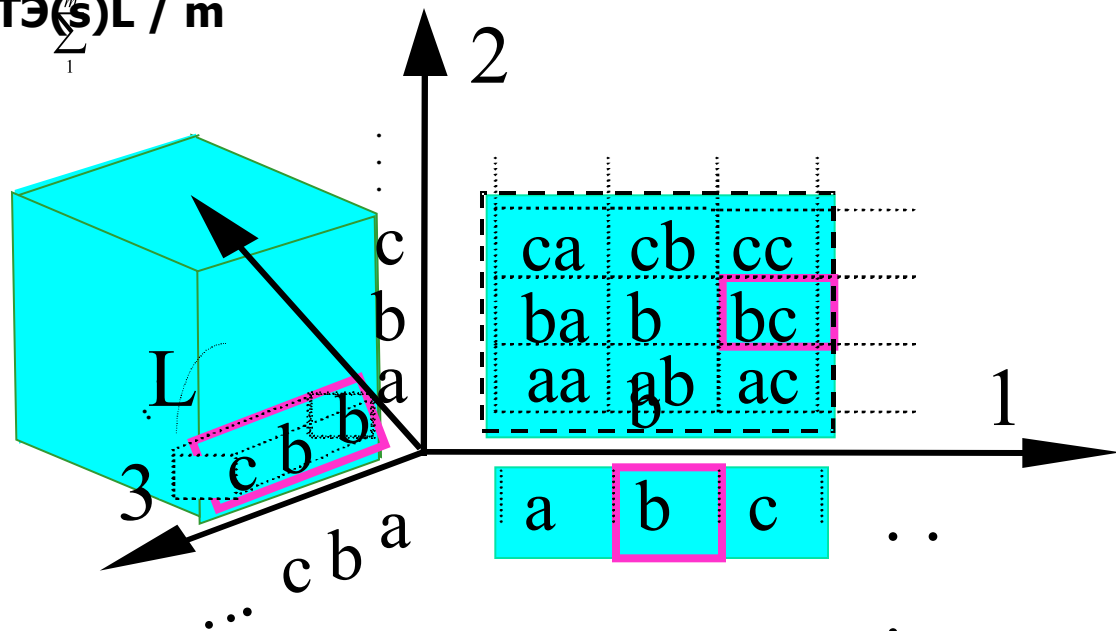
$$0 \leq p(s) \leq m;$$

$$\sum_1^m p(s)_L / m = 1 \quad (\text{условие нормировки})$$

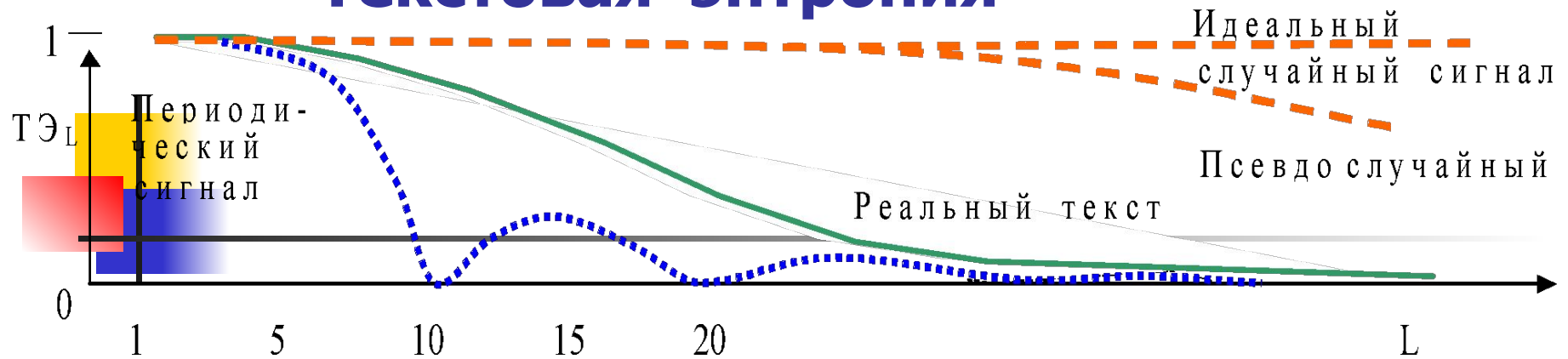
s - некоторое слово длиной в L символов;

m - размер потенциального S-словаря в  $m=A^L$  S-слов  
информационное пространство в  $L*m = L*A^L$  СИМВОЛОВ

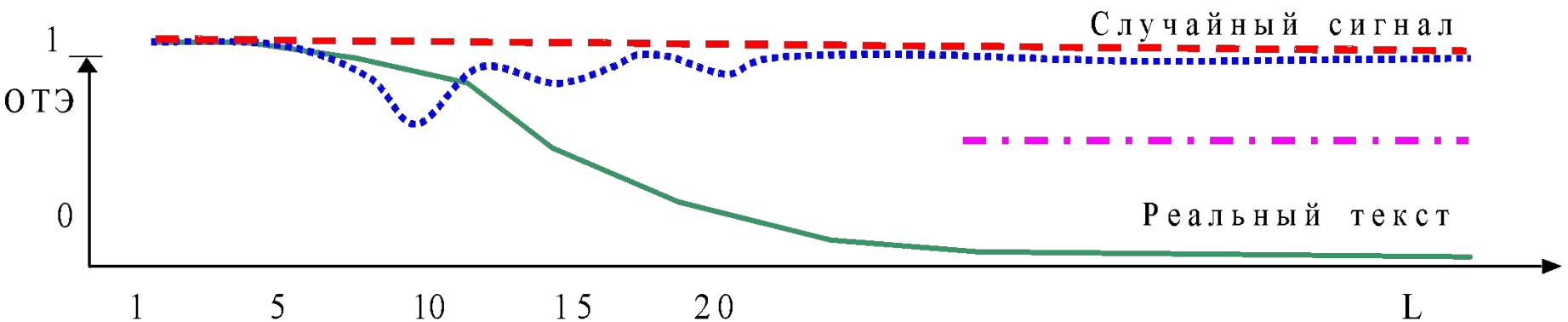
$$TЭ_L = TЭ(s)L / m$$



# Текстовая энтропия

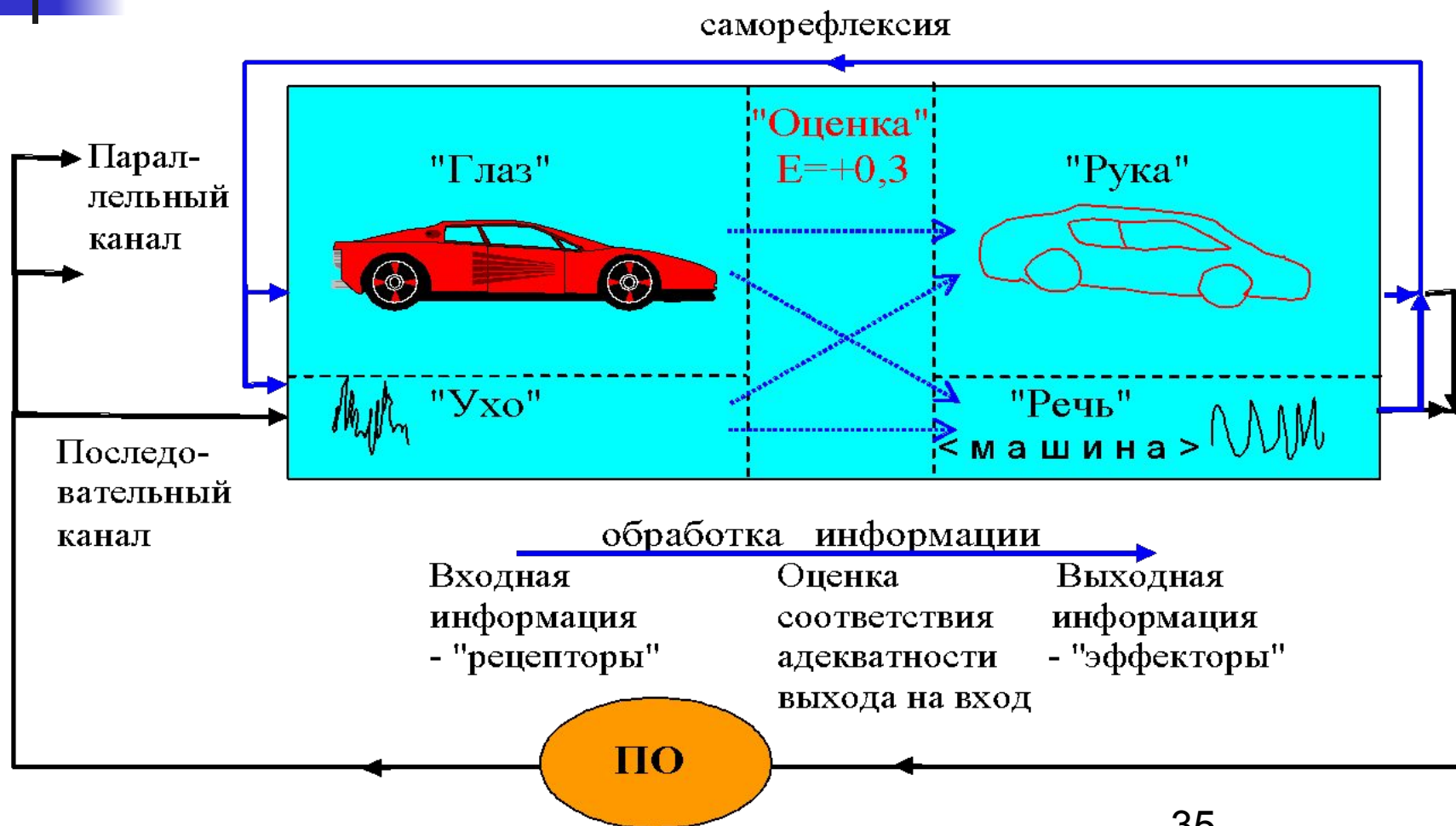


# Относительная текстовая энтропия



Таким образом, числовые значения ТЭ и ОТЭ являются эффективными *параметрами-индикаторами*, которые характеризуют возможность **семантического анализа** конкретной ПО (например, при поиске сигналов от внеземных цивилизаций).

# Прототипы компьютерного интерфейса: " человек-ИР "





# *СПАСИБО ЗА ВНИМАНИЕ*

---

По всем вопросам Проекта обращаться  
к **Бодякину Владимиру**,

E-mail: [body@ipu.ru](mailto:body@ipu.ru)

[www.informograd.narod.ru/](http://www.informograd.narod.ru/)

тел. (095)334-92-39,

в ИПУ тел 1-346

# Человек и Машина (ИР)

- Страхи что ИР поработит человечество – типичный пример широко распространившегося шаманизма от киноиндустрии. Для человека более опасен другой человек, т.к. у них одна ниша потребления. Генетически же **ИР нацелен на космос** – именно там широкое поле для его деятельности. Неограниченное количество любых материальных ресурсов вселенной не дает даже теоретических основ для конфликта между ИР с земным человеком.
- С точки же зрения сотрудничества, ИР для человека представляется **идеальным партнером**, т.к. у них одна область производства легко тиражируемого результата. Где в итоге, **каждый получает весь конечный продукт – новое знание** (пример ИРМИ). Так что, никаких естественных **оснований для конфликта между человеком и ИР нет**. Сотрудничество же человека с ИР станет мощным стимулом для заключительного экспоненциального этапа научно-технического прогресса нашей цивилизации.
- Работы над созданием ИР должны проводиться только в рамках **нового гуманистического мировоззрения** и специализированной международной академической инфраструктуры. Это необходимо, чтобы такой фактор, как ИР не стал "информационной дубиной" в руках какой-либо эгоистической группировки. В качестве такой начальной социально-экономической структуры по разработке ИР предлагается **проект "Информоград"**.

# Алгоритм (A1) преобразования текстовой формы в иерархическую структуру словарей (НСС)

$L = \{01100101011011011010001101101001\}$ .

Исходный текст

1-й шаг  $k_1=2$

|    |                                  |                                      |      |      |      |      |    |
|----|----------------------------------|--------------------------------------|------|------|------|------|----|
| /0 | 01100101011011011010001101101001 | Номера цепочек<br>в словаре          |      |      |      |      |    |
|    |                                  | Последовательность индексов (ссылок) | 1    | 2    | 3    | 4    |    |
| /1 | 1 2 1 1 1 2 3 1 2 2 4 3 1 2 2 1  | +                                    | "01" | "10" | "11" | "00" | L1 |



---

2-й шаг  $k_2=2$

|    |                  |    |    |    |    |      |      |      |                           |
|----|------------------|----|----|----|----|------|------|------|---------------------------|
| /1 | 1211123122431221 |    |    |    |    |      |      |      | Номера цепочек в словарях |
|    |                  | 1  | 2  | 3  | 4  | 5    | 6    |      |                           |
|    |                  | 12 | 11 | 31 | 22 | 43   | 21   | L2   |                           |
| /2 | 1 2 1 3 4 5 1 6  |    |    | +  |    | "01" | "10" | "11" | "00"                      |
|    |                  |    |    |    |    |      |      |      | L1                        |



---

3-й шаг  $k_3=2$

|           |          |    |    |    |      |      |      |      |                           |
|-----------|----------|----|----|----|------|------|------|------|---------------------------|
| $\bar{D}$ | 12134516 |    |    |    |      |      |      |      | Номера цепочек в словарях |
|           |          | 1  | 2  | 3  | 4    | 5    | 6    |      |                           |
|           |          | 12 | 13 | 45 | 16   |      |      |      | L3                        |
|           |          | 12 | 11 | 31 | 22   | 43   | 21   |      | L2                        |
| $\bar{B}$ | 1 2 3 4  |    |    | +  | "01" | "10" | "11" | "00" | L1                        |





---

4-й шаг  $k_4=2$

Номера цепочек в словарях

|       |      |    |    |    |    |      |      |      |      |    |
|-------|------|----|----|----|----|------|------|------|------|----|
| $I_3$ | 1234 | 1  | 2  | 3  | 4  | 5    | 6    |      |      |    |
|       |      | 12 | 34 | L4 |    |      |      |      |      |    |
|       |      | 12 | 13 | 45 | 16 |      | L3   |      |      |    |
|       |      | 12 | 11 | 31 | 22 | 43   | 21   | L2   |      |    |
| $I_4$ | 1 2  |    |    |    | +  | "01" | "10" | "11" | "00" | L1 |

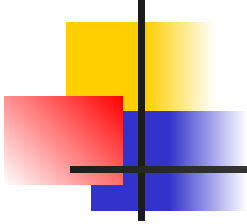


---

5-й шаг  $k_5=2$

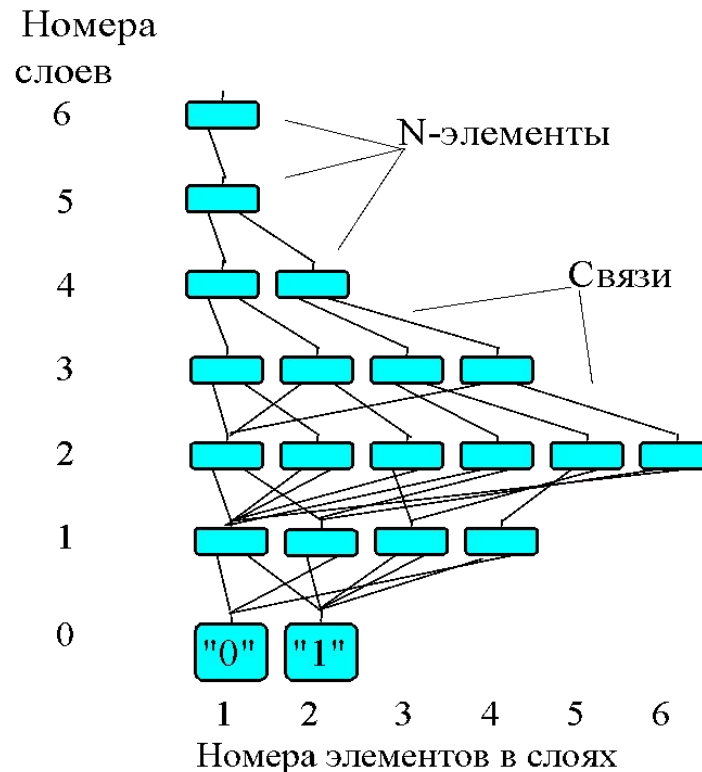
и исходная последовательность символов полностью переходит в НСС.

| Номера словарей (слоев) | Ссылки        |    |                            |    |    |    |                                   |   |
|-------------------------|---------------|----|----------------------------|----|----|----|-----------------------------------|---|
|                         | 6             | 1  |                            |    |    |    | на предыдущие словари и "алфавит" |   |
|                         | 5             | 12 |                            |    |    |    |                                   |   |
| 4                       | 12            | 34 |                            |    |    |    |                                   |   |
| 3                       | 12            | 13 | 45                         | 16 |    |    |                                   |   |
| 2                       | 12            | 11 | 31                         | 22 | 43 | 21 |                                   |   |
| 1                       | "01"10"11"00" |    |                            |    |    |    |                                   |   |
|                         |               |    | 1                          | 2  | 3  | 4  | 5                                 | 6 |
|                         |               |    | Номера элементов в словаре |    |    |    |                                   |   |



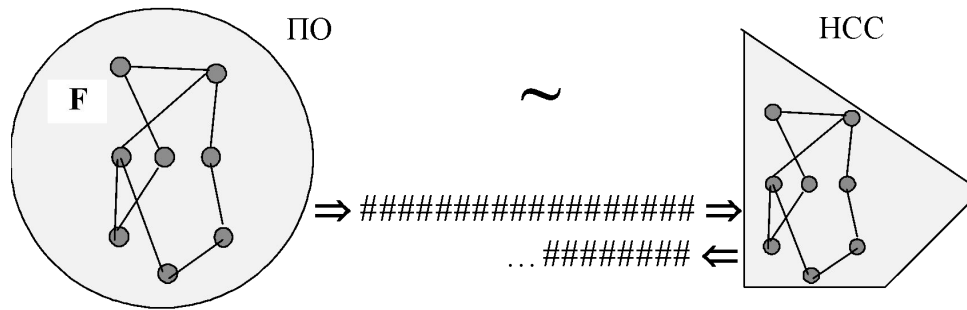
|                               |   |                            |     |        |    |    |    |   |    |
|-------------------------------|---|----------------------------|-----|--------|----|----|----|---|----|
| Номера<br>словарей<br>(слоев) |   | 6                          | 6   | Ссылки |    |    |    | на предыдущие<br>словари<br>и "алфавит" |    |
|                               | 5 | 5                          | 12  |        |    |    |    |   |    |
|                               | 4 | 4                          | 12  | 34     |    |    |    |   |    |
|                               | 3 | 3                          | 12  | 13     | 45 | 16 |    |   |    |
|                               | 2 | 2                          | 12  | 11     | 31 | 22 | 43 |   | 21 |
|                               | 1 | 1                          | 12  | 21     | 22 | 11 |    |   |    |
| 0                             | A | "0"                        | "1" |        |    |    |    |   |    |
|                               |   | 1                          | 2   | 3      | 4  | 5  | 6  |   |    |
|                               |   | Номера элементов в словаре |     |        |    |    |    |   |    |

# Структурное описание НСС можно представить в виде многоуровневого графа:



**Алгоритм обратного преобразования НСС в текстовую форму** осуществляется уже за меньшее число операций и идет "сверху-вниз".

# Изоморфность структур процессов в ПО и НСС



## Теорема:

**максимальное значение компрессии достижимо только при изоморфном отображении структуры процессов ПО в структуре образов НСС**