

КОНЦЕПЦИЯ ПОСТРОЕНИЯ

ИСКУССТВЕННОГО РАЗУМА
НА БАЗЕ

НЕЙРОСЕМАНТИЧЕСКОГО ПОДХОДА
(ПРИГЛАШЕНИЕ К СОТРУДНИЧЕСТВУ)

Бодякин В.И.

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

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

План обсуждения



Постановка задачи. Проблемы 2030 года

Нейросемантика

- Предметная область – Информационный канал – Информационная система
- Форма представления информации
- Автосруктуризация
- Нейросемантические структуры (НСС)
- Сигнал – Информация – Знание
- Адаптивные регуляторы на базе НСС
- Текстовая энтропия

Решение. Проекты: - Искусственный разум

- Информоград
- Восхождение разума

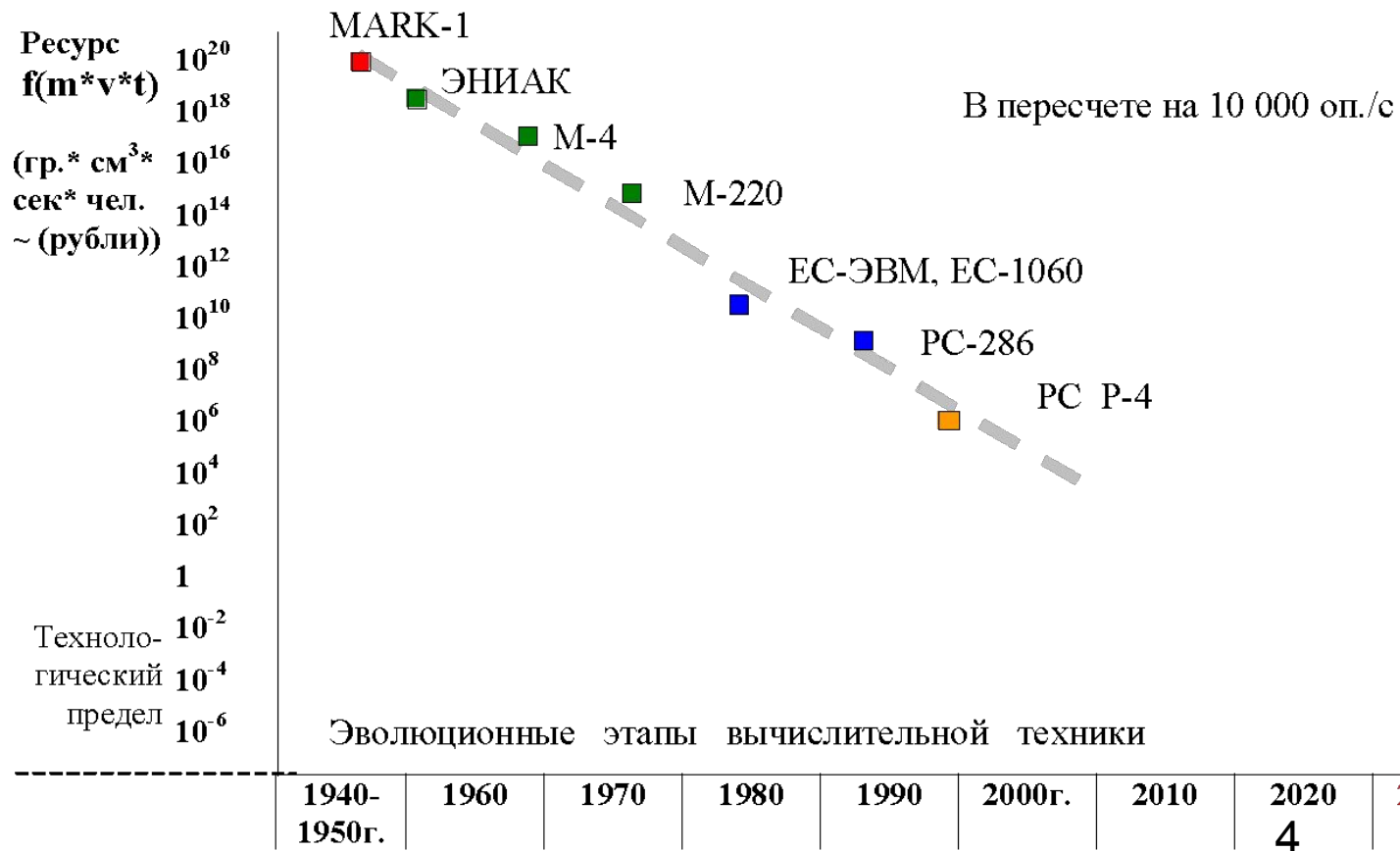


Постановка задачи

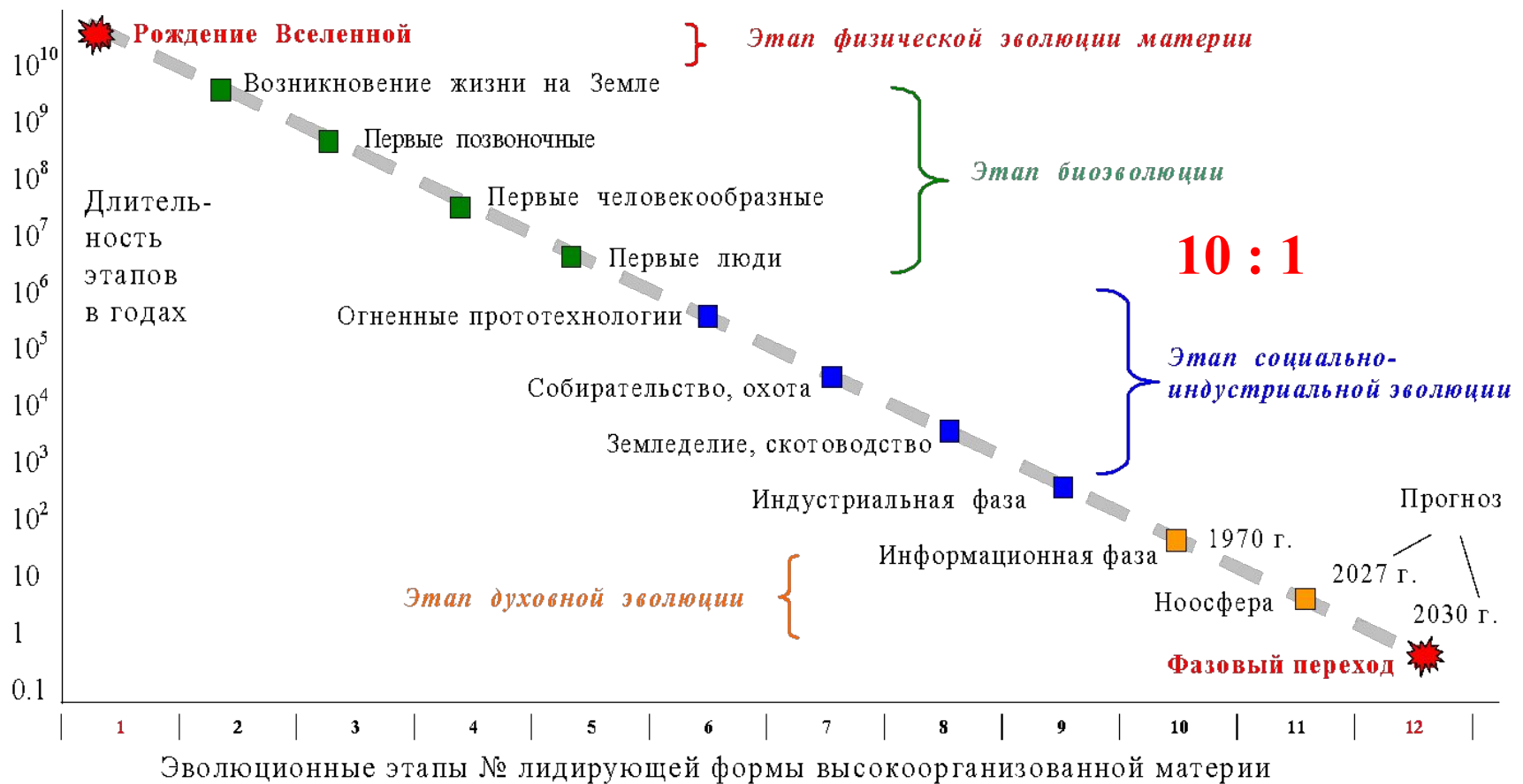
Проблемы 2030 года

- Тенденции уменьшения ресурсоемкости информационной техники и прогноз
- Глобальные эволюционные этапы лидирующей формы высокоорганизованной материи
- Модель "текстовая жизнь"
- Естественная эволюция, - это ряд аттракторов
- Темпы и масштабы эволюции
- Эволюция технической цивилизации

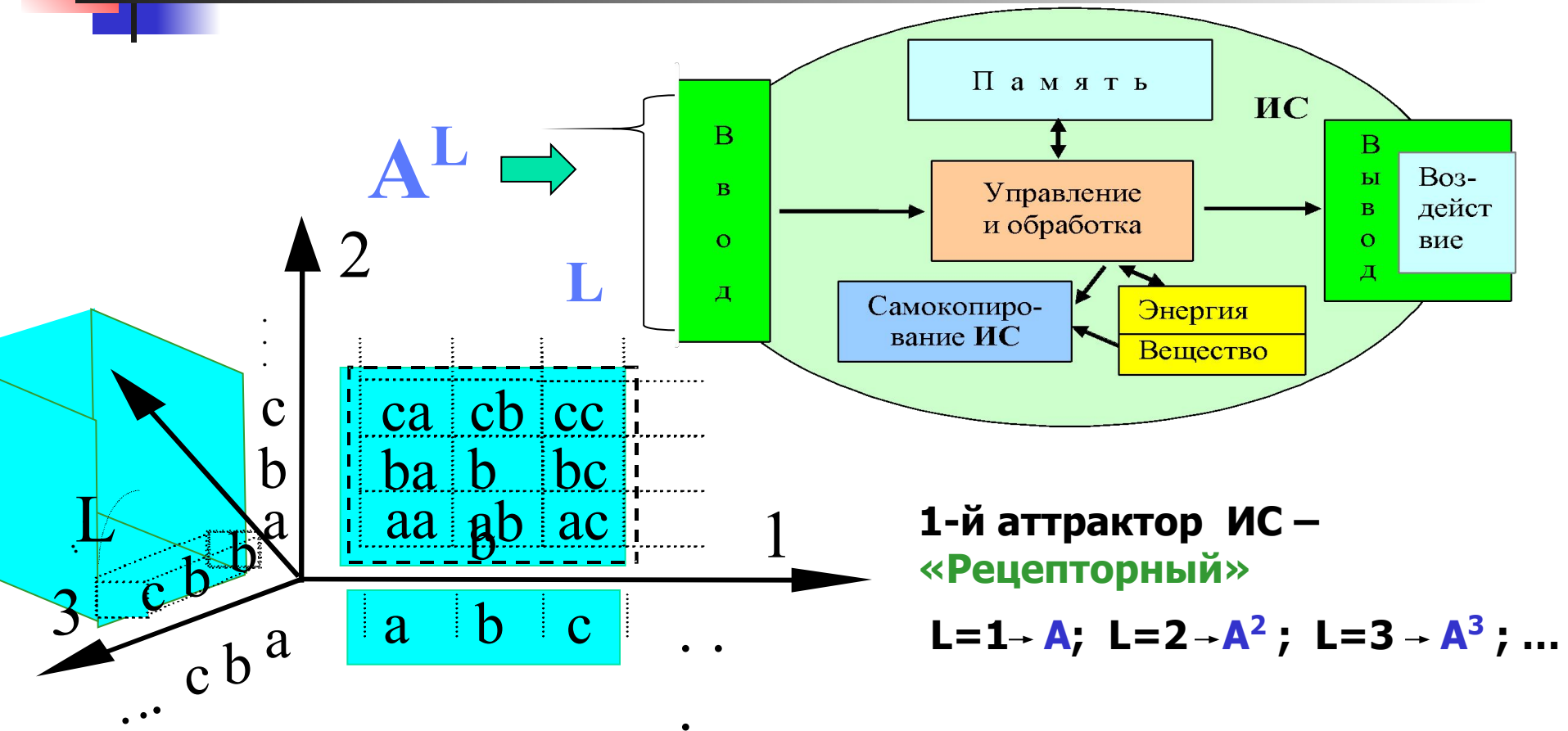
Тенденции уменьшения ресурсоемкости ($f(\text{гр} \cdot \text{см}^3 \cdot \text{сек} \cdot \text{чел})$) информационной техники (на 10 000 оп./с)



Глобальные эволюционные этапы лидирующей формы высокоорганизованной материи



Энерго-сырьевой потенциал эволюции простейших информационных систем (ИС). ("первые аттракторы").



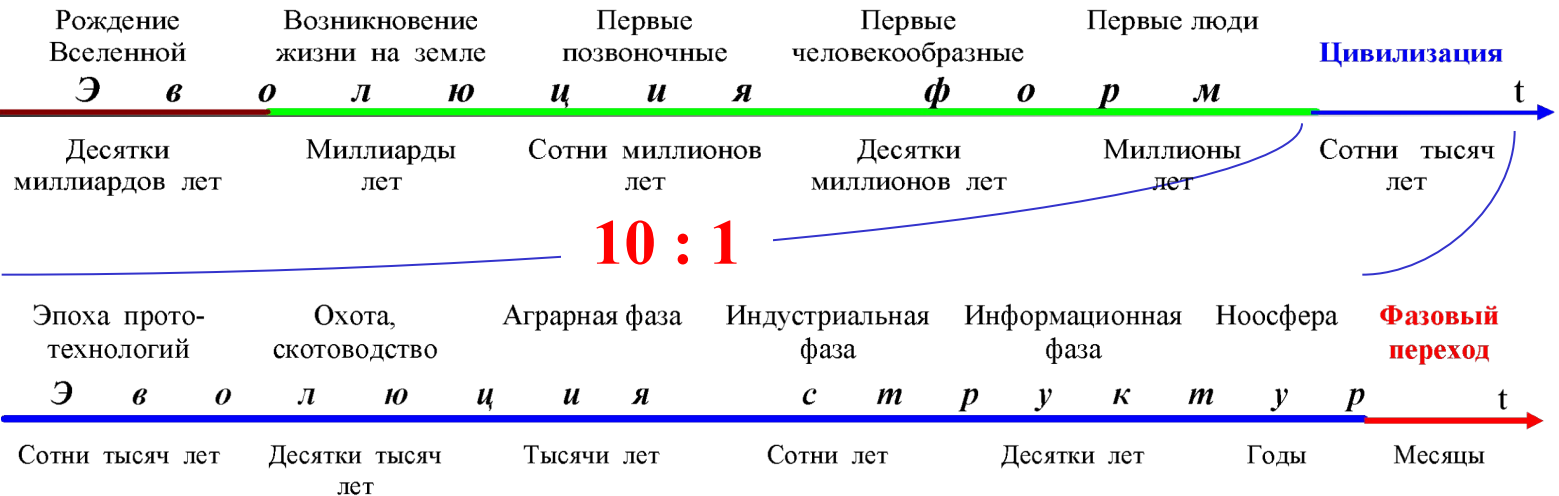
Процесс естественной эволюции ИС

представляет ряд аттракторов

с экспоненциальными характеристиками

- **Первый** из которых "**рецепторный**" – линейные затраты (**x**) на расширение диапазона рецепторов ИС дают приращение эволюционного ресурса как A^x , где **A** – алфавит рецепторов, **x** – длины образов.
- **Второй** эволюционный аттрактор ИС – "**ассоциативный**", заключается в переводе линейной формы памяти ИС в иерархически-сетевую.
- **Третий** аттрактор ИС – "**интеллектуальный**". Он позволяет сводить воспринимаемые семантические переменные текстов к константам их грамматических структур [4]. Конечное множество грамматических структур покрывает большую часть информационного потока любой среды.
- **Четвертый** эволюционный аттрактор ИС – "**знание**" когда все типы грамматических конструкций ИС по переработке текста сводятся к 4 конструкциям над грамматическими конструкциями ИС ("Homo-sapiens").
- **Пятый** эволюционный аттрактор ИС – "**социальный**", законы А.А.Богданова о сверхаддитивности сложения усилий субъектов.
- **Шестой** аттрактор – "**Восхождение разума**", интеграция всех ИС независимо от их начальной природы (биологические, искусственные, ...), приводит к бесконечному росту эффективности Разума в "понимании" среды. .

Темпы и масштабы эволюции



1 : 1000

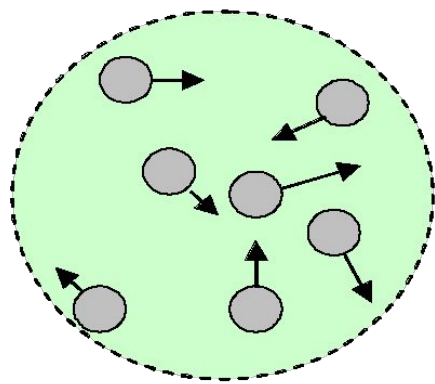
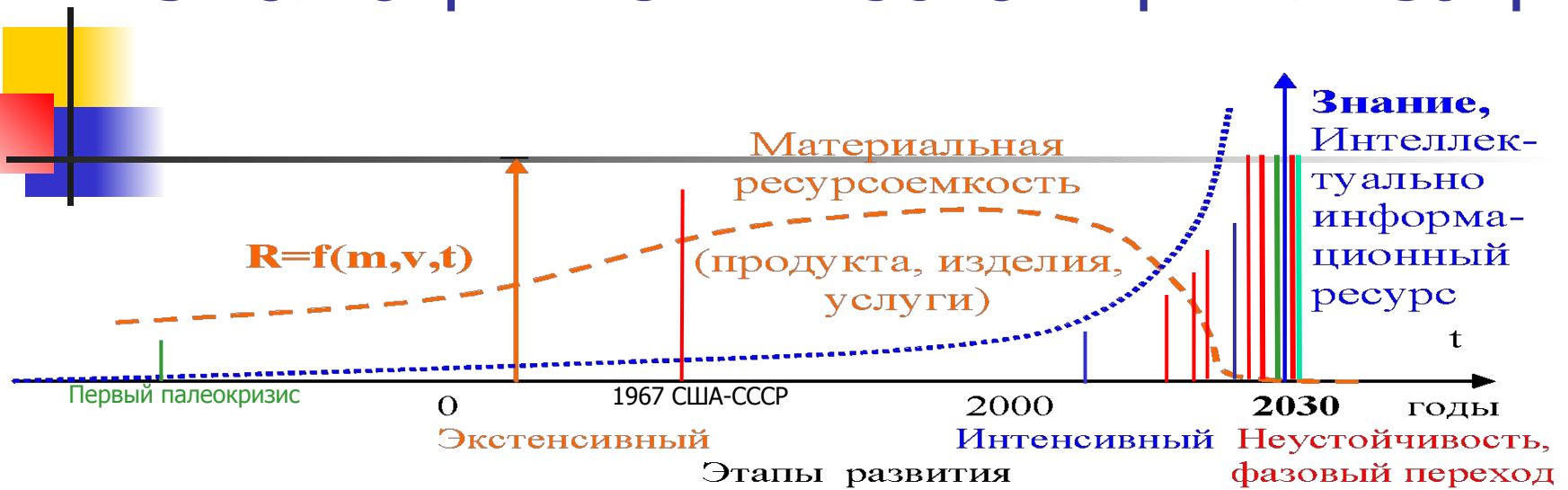
Энерго-вещественный этап

информационный этап

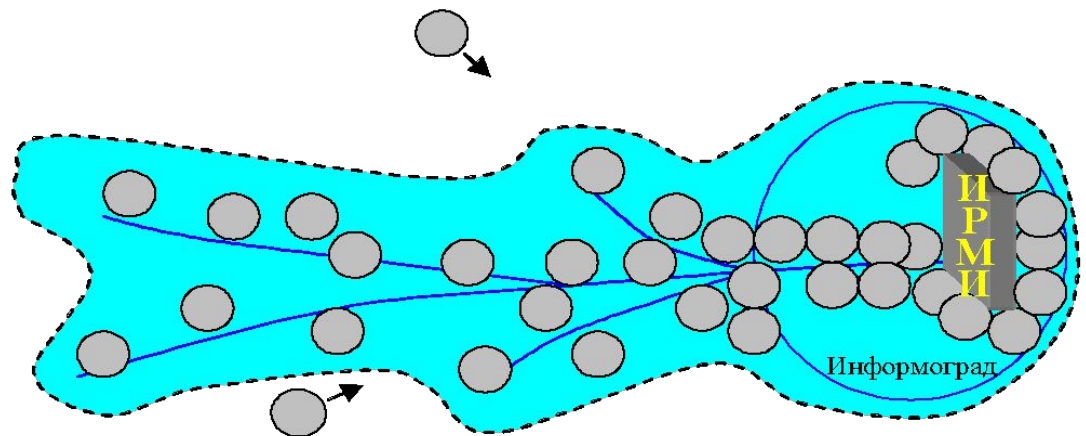


Направленность эволюции
Рост сферы влияния

Эволюция технической цивилизации



Одноклеточный социум



Ноосфера – форма существования социального организма (Информоград - первый этап)



Нейросемантика

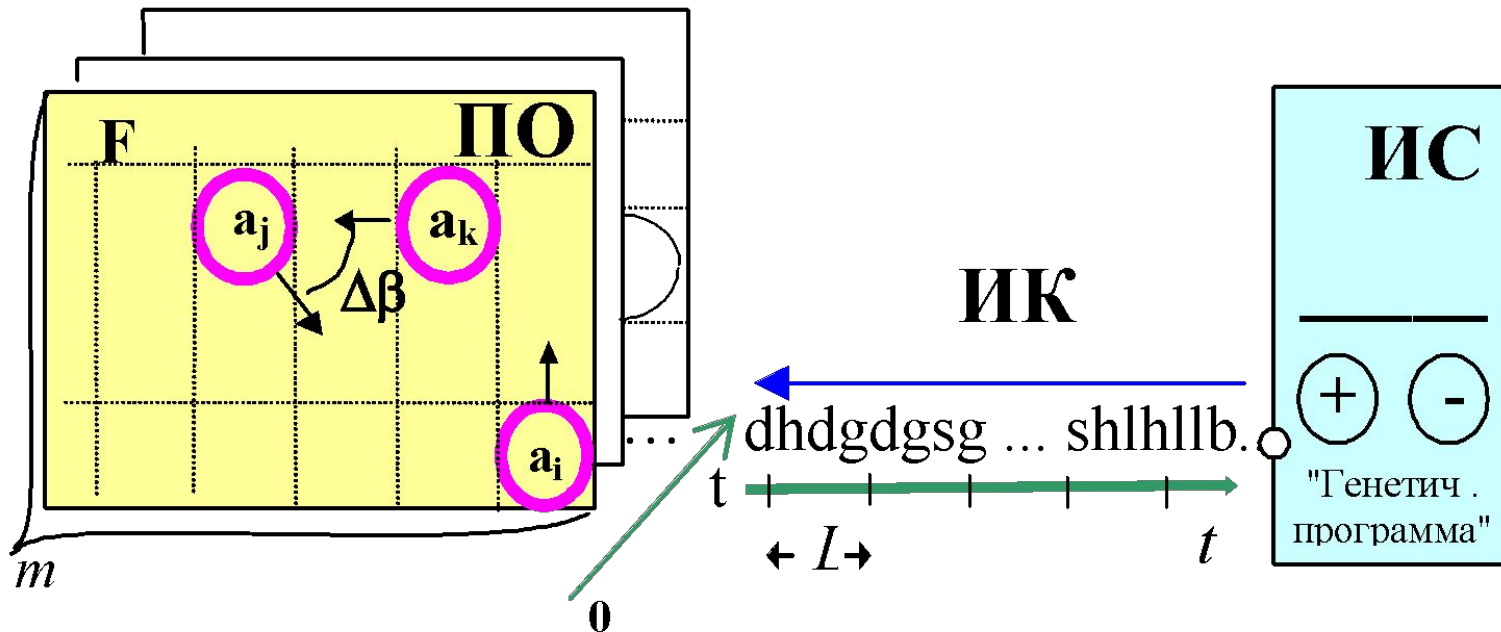
- Предметная область – Информационный канал – Информационная система
- Форма представления информации
- Автосруктуризация
- Нейросемантические структуры (НСС)
- Сигнал – Информация – Знание
- Адаптивные регуляторы на базе НСС
- Текстовая энтропия

Комплекс:

"Предметная область (ПО) –

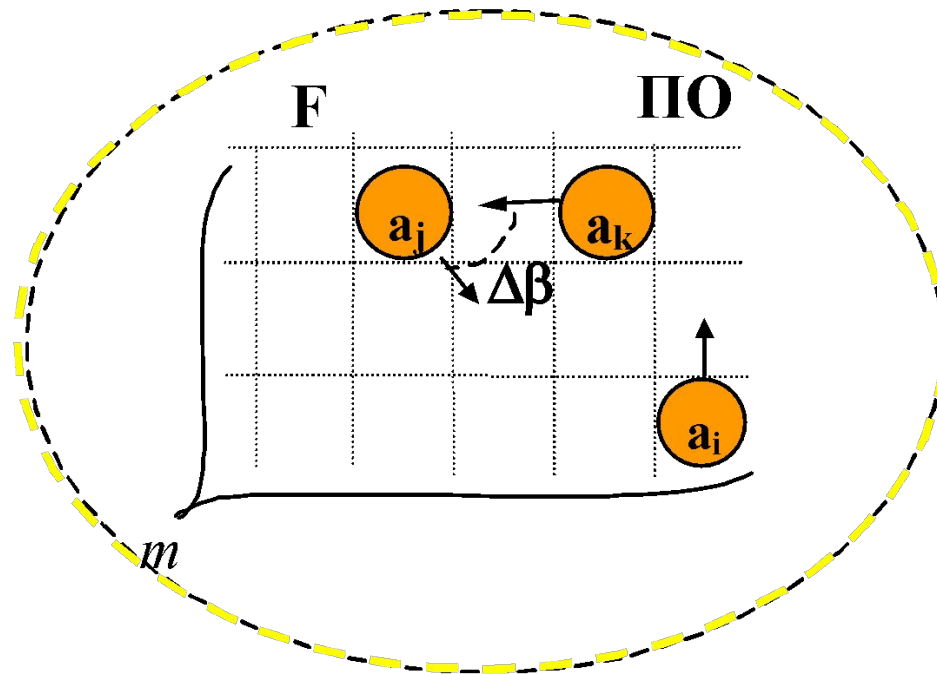
Информационный канал (ИК) –

Информационная система (ИС)"



Предметная область (ПО)

Любую физическую ПО можно рассматривать как некоторую дискретную пространственно-временную область с взаимодействующими объектами-процессами (a_i, a_j). При попадании нескольких объектов, в одну и ту же ячейку (область взаимодействия), происходит процесс их взаимодействия.



Энергетика физического процесса

– это

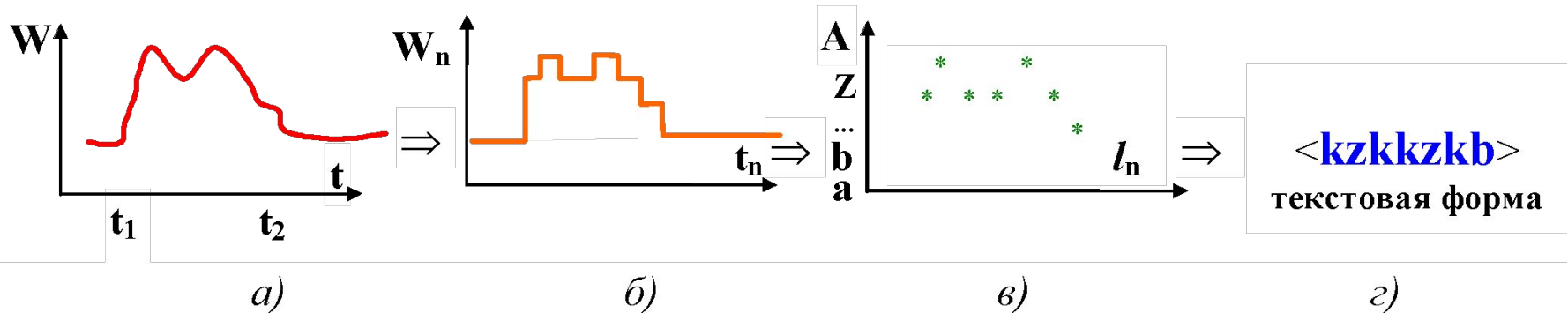
некоторый **инвариант**, характеризующий потенциальную глубину причинно-следственного распространения данного физического процесса

- 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.

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



Будем предполагать, что в любой ПО выполняются:

- **принцип причинности** (инвариантность процессов во времени и в пространстве),
т.е. если $si \rightarrow sj$, то процесс si всегда вызывает s_j ;
- **принцип локальной однородности** (сведение континуума значений параметра к ограниченному числу качественно-различных образов, т.е. $\langle zzjbaabj \rangle = si \sim \langle zzjbab \rangle = si$);
Например, вся температурная шкала разбивается на четыре образа: $si =$ «твердое», $s_j =$ «жидкое», $sk =$ «газ», $sl =$ «плазма».
- **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**

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

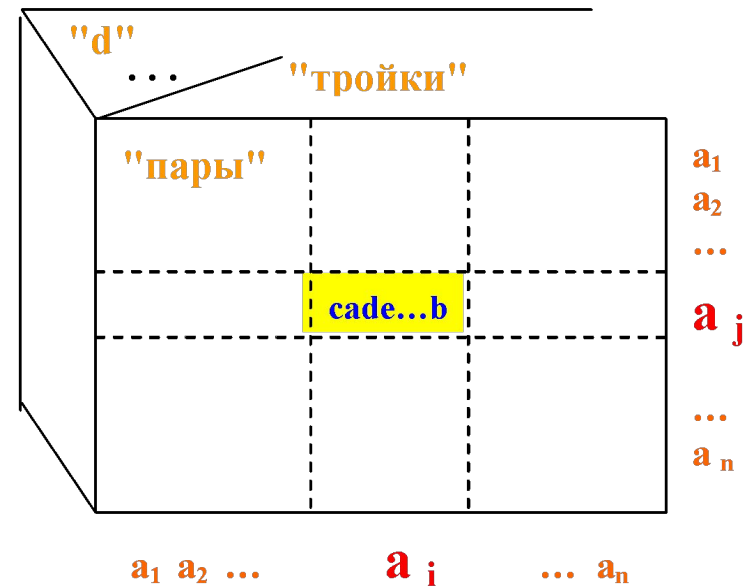
которая однозначно порождается процессом взаимодействия этих объектов. Все множество взаимодействий в данной ПО, представляет собой **семантическое пространство**, как совокупность ЭСЕ.

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).

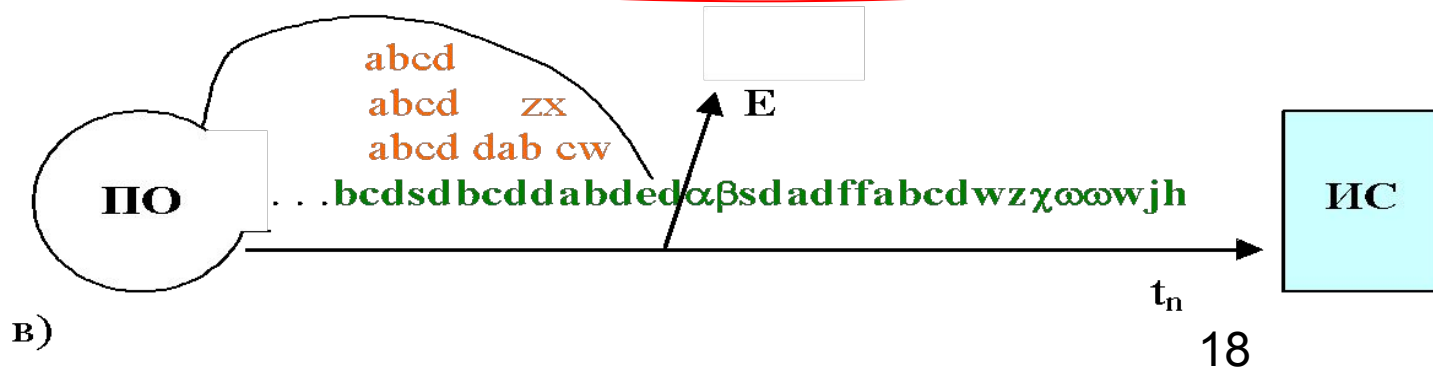


ЭСЕ, порождаемые в ПО, суперпозиционно "сливаются" в непрерывный информационный поток. W

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.

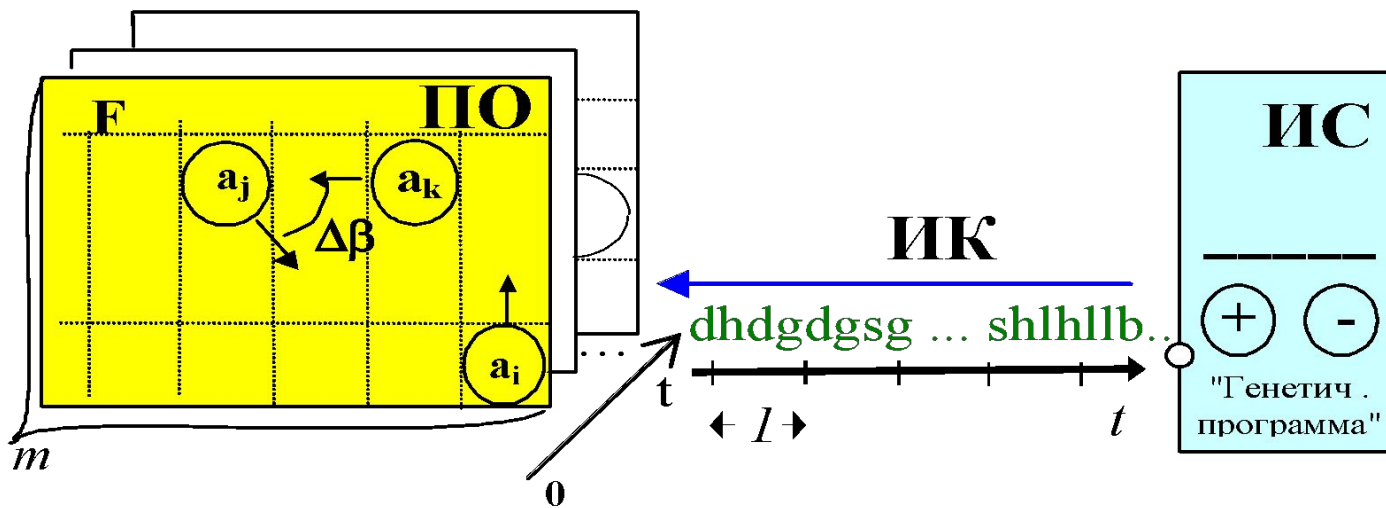


В рамках рассматриваемого комплекса:

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)
6. Иметь положительный **эволюционный потенциал**



АВТОСТРУКТУРИЗАЦИЯ НЕПРЕРЫВНОГО ТЕКСТОВОГО ПОТОКА

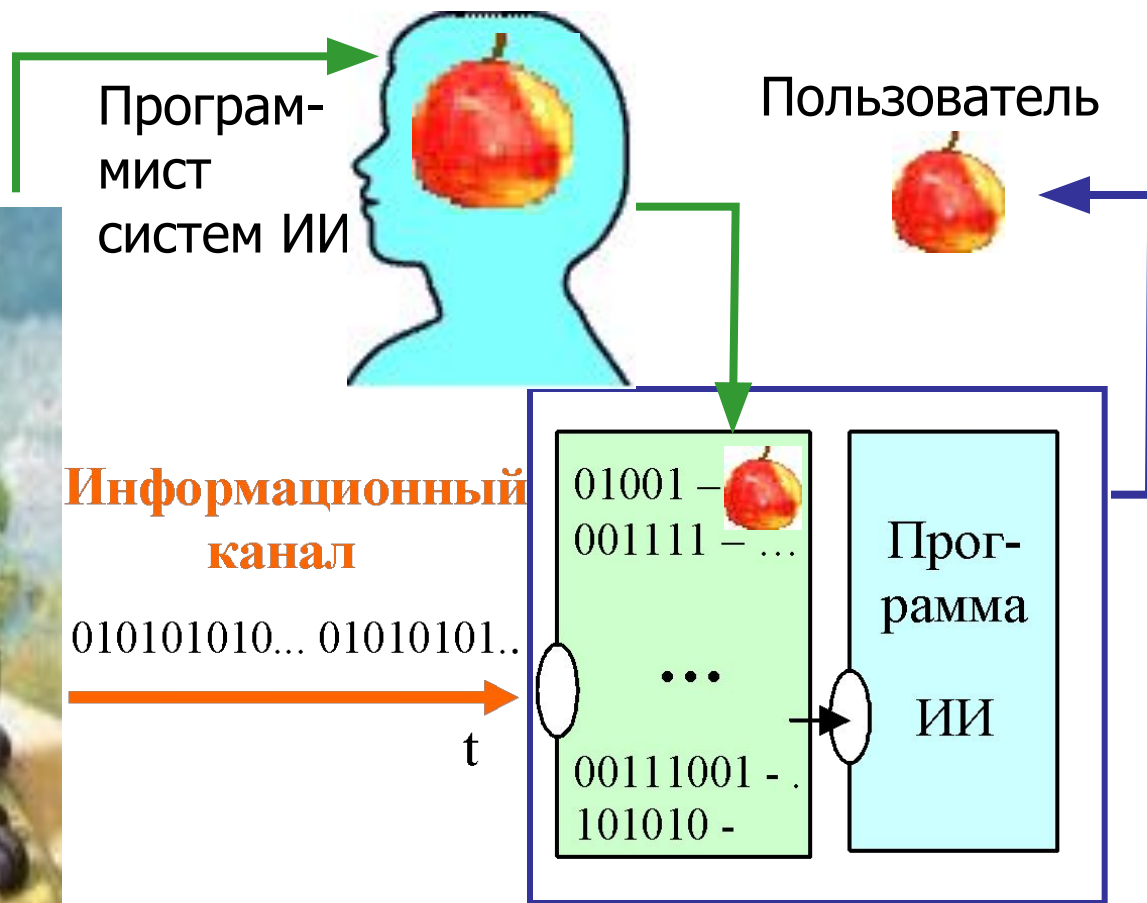
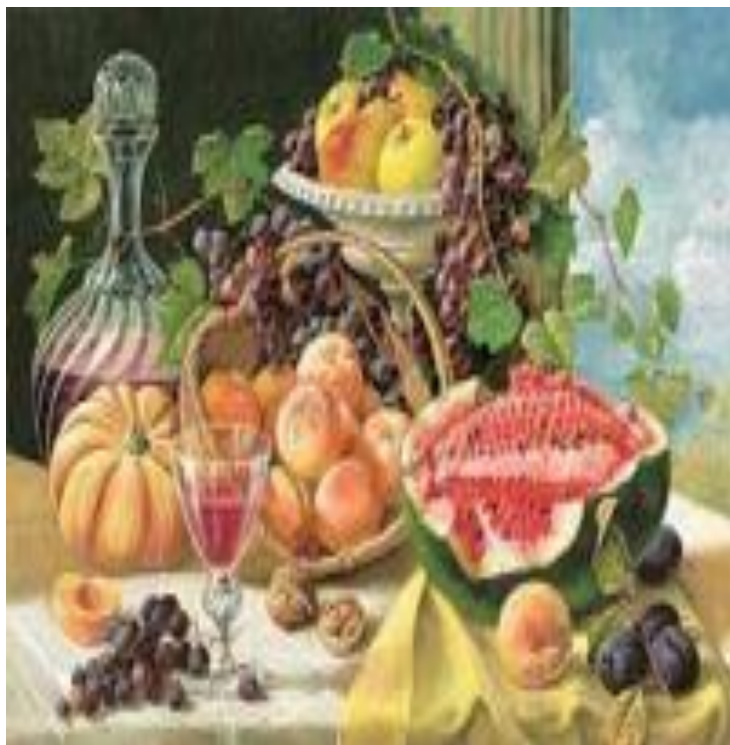


ПОЧЕМУ ВСЕ ОРГАНИЗМЫ "ЕДИНОДУШНЫ" В КЛАСТЕРИЗАЦИИ ОКРУЖАЮЩЕГО НАС МИРА НА ОТДЕЛЬНЫЕ ОБРАЗЫ ?

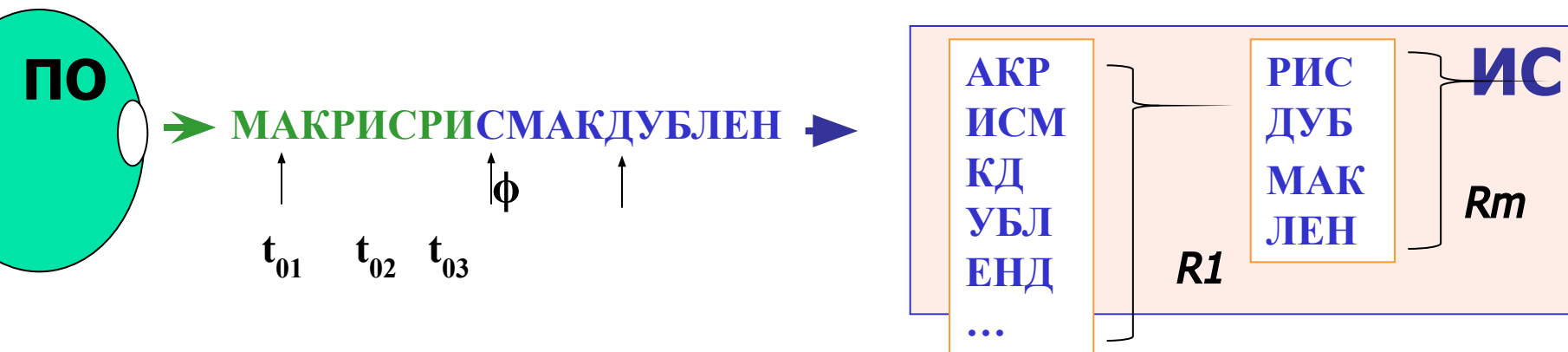


Традиционный способ структуризации в системах ИИ

Предметная область



Задача: В непрерывном потоке ТФ необходимо выделить образы, соответствующие процессам любой ПО



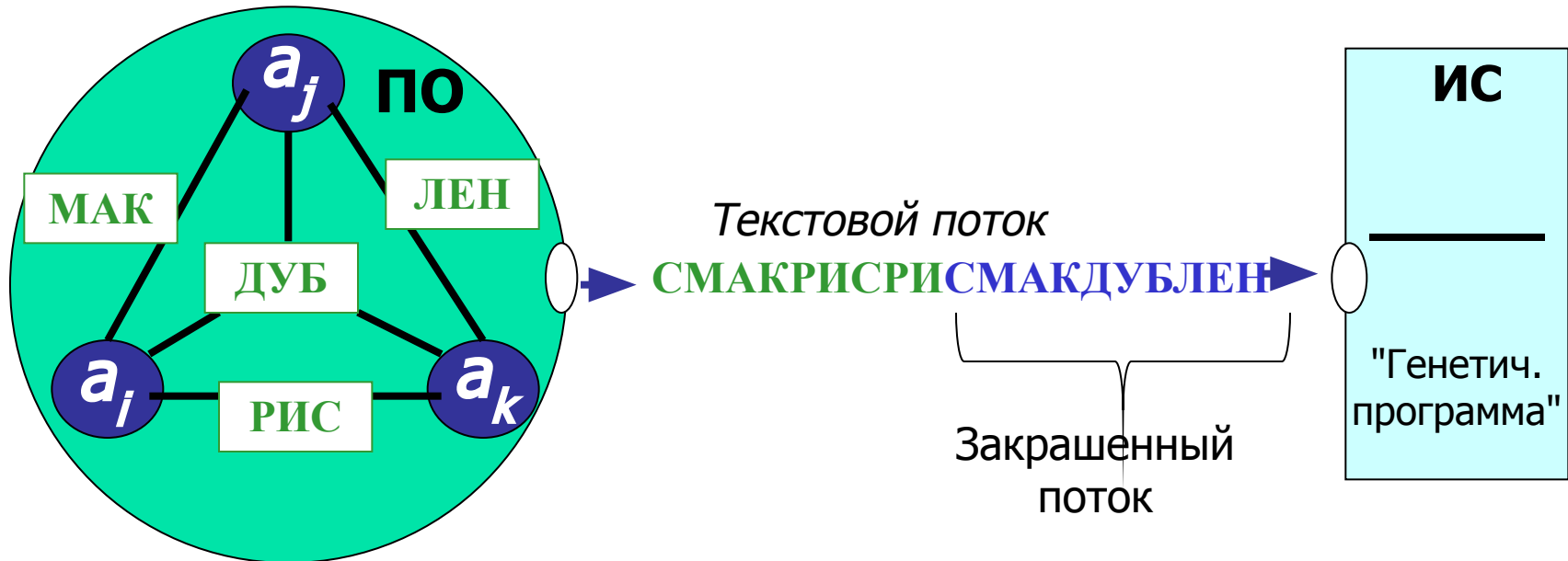
Метод: построение в ИС множества различных словарей и выбор минимального (R_i)

Результат: Минимальный словарь ИС
(гомоморфен процессам ПО)

Инструментарий: нейросемантические структуры

Демонстрационный пример

(четыре равновероятных процесса формируют непрерывный поток ТФ)



Необходимо построить словарь в N образов, полностью покрывающий ТФ. Примем что:

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

Время обработки символа и образа – один такт T ,

Энергетические затраты на хранение одного образа в памяти $1/24 E-$.

Энергетика каждого прогнозируемого ИС символа равна $1E+$,

Усредненный на 12 тактов энергетический баланс ИС =

$$-N(\text{обработка}) - 1/2N(\text{хранение}) + (12-N)(\text{прогнозирование}) = \frac{(12 - 5/2N)*E}{24}$$

Результаты эксперимента

Первая структуризация словаря ИС

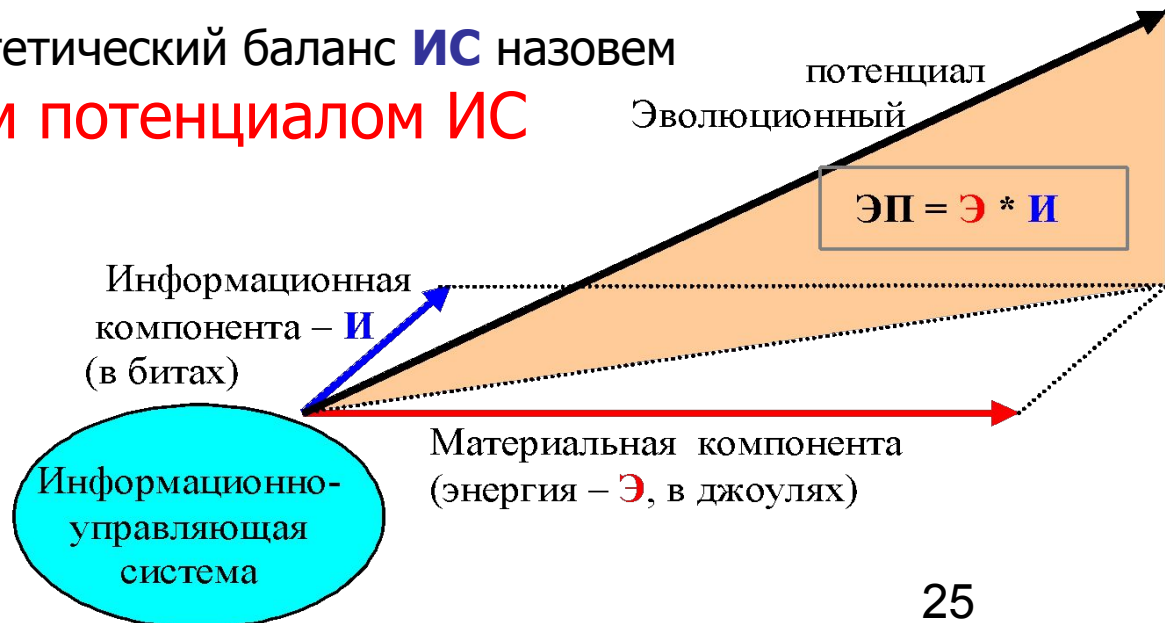
Форма словаря: (наибольший размер образа один символ):

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

Итог: для односимвольного словаря усредненный энергетический баланс = 18E- .

Усредненный энергетический баланс **ИС** назовем
ЭВОЛЮЦИОННЫМ ПОТЕНЦИАЛОМ ИС

Эволюционный потенциал ИС
численно
характеризует
жизнеспособность
ИС



Вторая структуризация словаря ИС

(наибольший размер образа два символа):

- а) Минимальная форма словаря: <МА><К><РИ><С><ЛЕ><Н><ДУ><Б>:
- $R(\text{ИС})=8R = 4E^-$, прогнозирование = $0,5T$ (на образ),
- затраты энергии на распознавание = $8E^-$,
- контролирование потенциальной энергии $T\Phi = 4E^+$.
- *Итог а)* $4E^- + (0,5T * 8(\text{образов на } T\Phi=12) = 4E^+) + 8E^- = 8E^-$.

- б) Максимальная форма (без полного пересечения):
<МА><КЛ><КР><КД><КМ> <ЛЕ>...<БД>:
- $R(\text{ИС})= 20R=10E^-$, прогнозирование = $0,2T$ (на образ),
- затраты энергии на распознавание = $20E^-$,
- контролирование потенциальной энергии $T\Phi = 4E^+$.
- *Итог б)* $10E^- + (0,2T * 20(\text{образов на } T\Phi=12) = 4E^+) + 20E^- = 26E^-$.

- **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 = $2T$ (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, т.е. без полного пересечения): <АКР><ИСП><ИСМ><АКД> <УБЛ><ЕНД> ... <АКМ>:
- $R(ИС) = 36R = 18E^-$,
- прогнозирование (T) = $+4E$,
- затраты энергии на распознавание = $36E^-$,
- контролирование потенциальной энергии $T\Phi = 0E^+$.
- *Итог б)* = $50E^-$.

- Лучший итоговый эволюционный потенциал = $2E^+$, $50E^- \rightarrow 2E^+(!!)$.

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 $T\Phi = 0,5E^+$.
- *Итог б)* = $78E^-$.

- Лучший итоговый эволюционный потенциал = $2E^+$, $78E^- \rightarrow 2E^+$

Пятая и другие структуризации словаря

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

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

Худший - $\infty E-$

- Теоретический *анализ результатов эксперимента* показывает, что **эволюционный потенциал ИС** *обратно пропорционален* **размеру словаря**,
- Размеры минимальных и максимальных словарей ИС соотносятся минимум как:

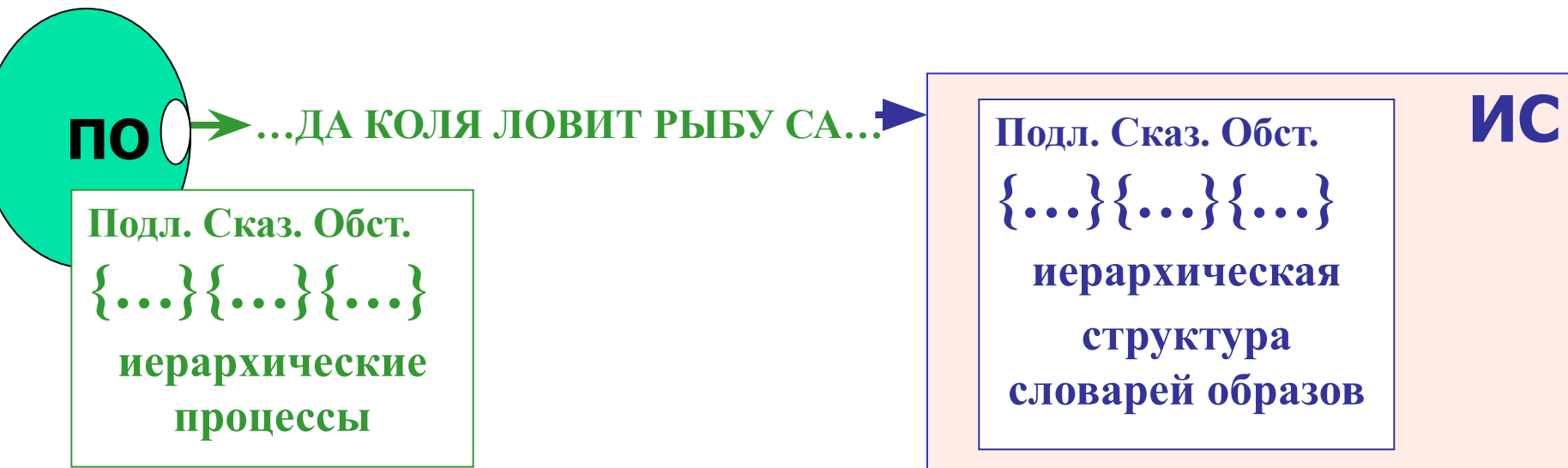
$o(N)$ и $o(N^3)$!!! $O(n)$; ???

где: **N** – максимальное число процессов в ПО

Автоструктуризация информации в ИС



Автоструктуризация сложных иерархических процессов



При минимизации словаря на нейросемантических структурах, его топология гомоморфна структуре исходных процессов ПО

Теорема: минимальная форма словаря ИС может достигаться только при ее гомоморфности исходной структуре процессов ПО

Доказательство: Если представить формирование текстовой формы двумя независимыми и непересекающимися процессами А и Б, то становится очевидным,



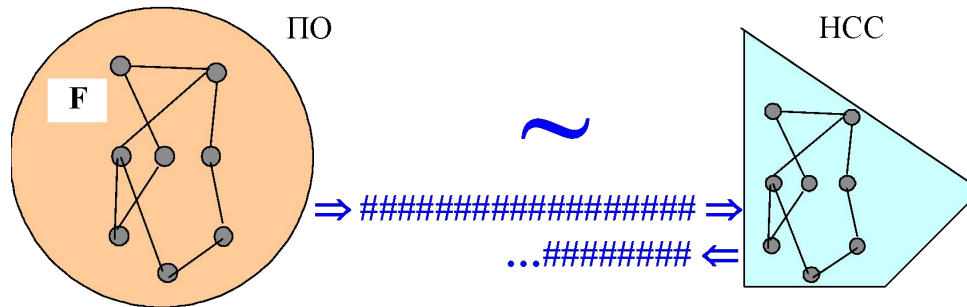
что минимальным словарем образов закрашивающим эти два процесса могут быть только образы совпадающие по текстовой форме с генерирующими их процессами.

На вопрос: "Почему все организмы 'единодушны' в кластеризации окружающего нас мира на отдельные образы?"

Вытекает ответ: "Т.к. минимальный словарь, дает эволюционные преимущества, то все ИС данной ПО выбирают его, а соответственно, и его образы".

Из минимизации отображения в ИС

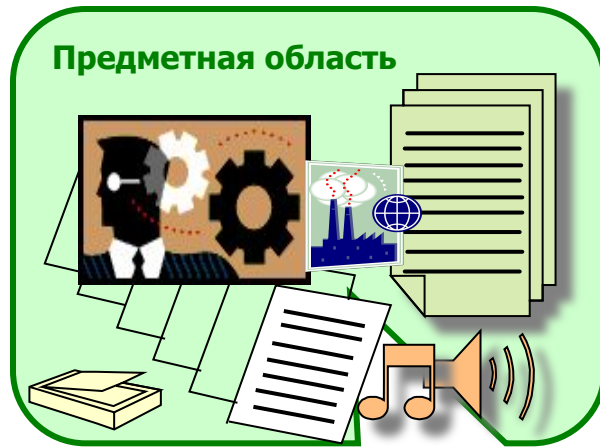
→ гомоморфность
структуры образов в НСС
структуре процессов ПО



Теорема:

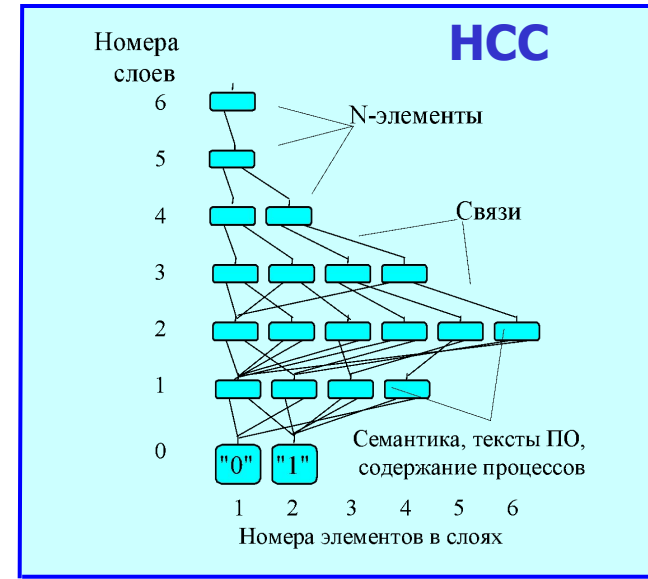
**максимальное значение компрессии достижимо
только при гомоморфном отображении
структуры процессов ПО в структуре образов НСС**

Нейросемантическая форма представления информации



Алгоритм
НСС

11011010001101101001 □-□



Автоматическое выделение образов-процессов из предметной области в нейроподобные элементы НСС при минимизации ресурсных затрат (памяти)

N-элемент (образ НСС) ↔ процесс предметной области

Автоструктуризация на нейросемантических структурах

$R_{ис} = f(\text{число } N\text{-элементов, число связей})$ 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 $t \rightarrow \infty$

Примеры:

а) <RISMAKDUBLENLENDUBMAKMAKLENRISRISLENMAKRISDUBRIS>

правильно выделяются все **процессы**: <МАК><РИС><МАК><ДУБ>;

б) <ДОМЗЕБРЫСКИТНАДОМДОМВНАДОМВСКИТВНАСКИТВВЗЕБРЫНАВНА> ,

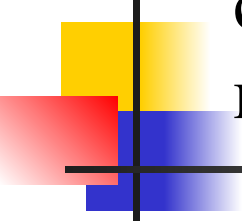
правильно выделяются все **процессы**: <ЗЕБРЫ><СКИТ><ДОМ><НА>

<В>.

сдвиг алфавита А в кодах ASCII в примере б)

<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:

N-элемент (образ ИС) ↔ ЭСЕ (ПО)

будем называть –

нейросемантической структурой (НСС)

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

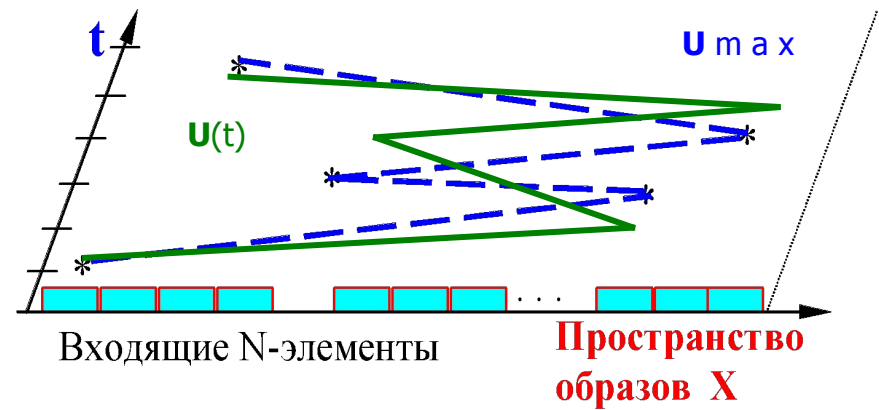
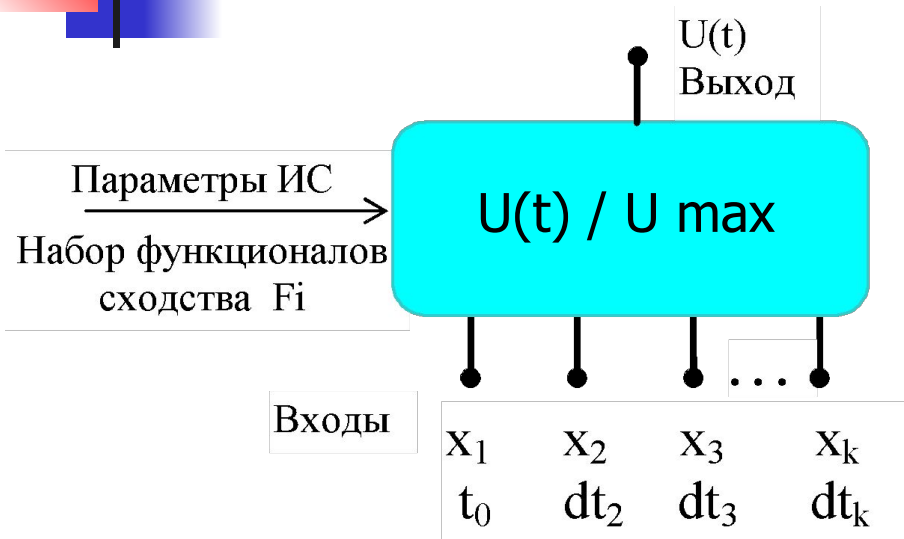
Понятно, что ее автоматическое формирование открывает широкие горизонты для инженерии ИС.

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



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 $R_{ИС}$

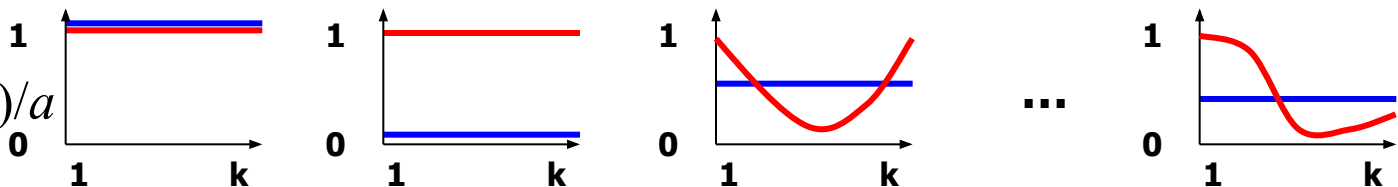
Расширив свойства формального нейрона введением в него **относительности времени активации входов**, получим **нейроподобный N-элемент**.



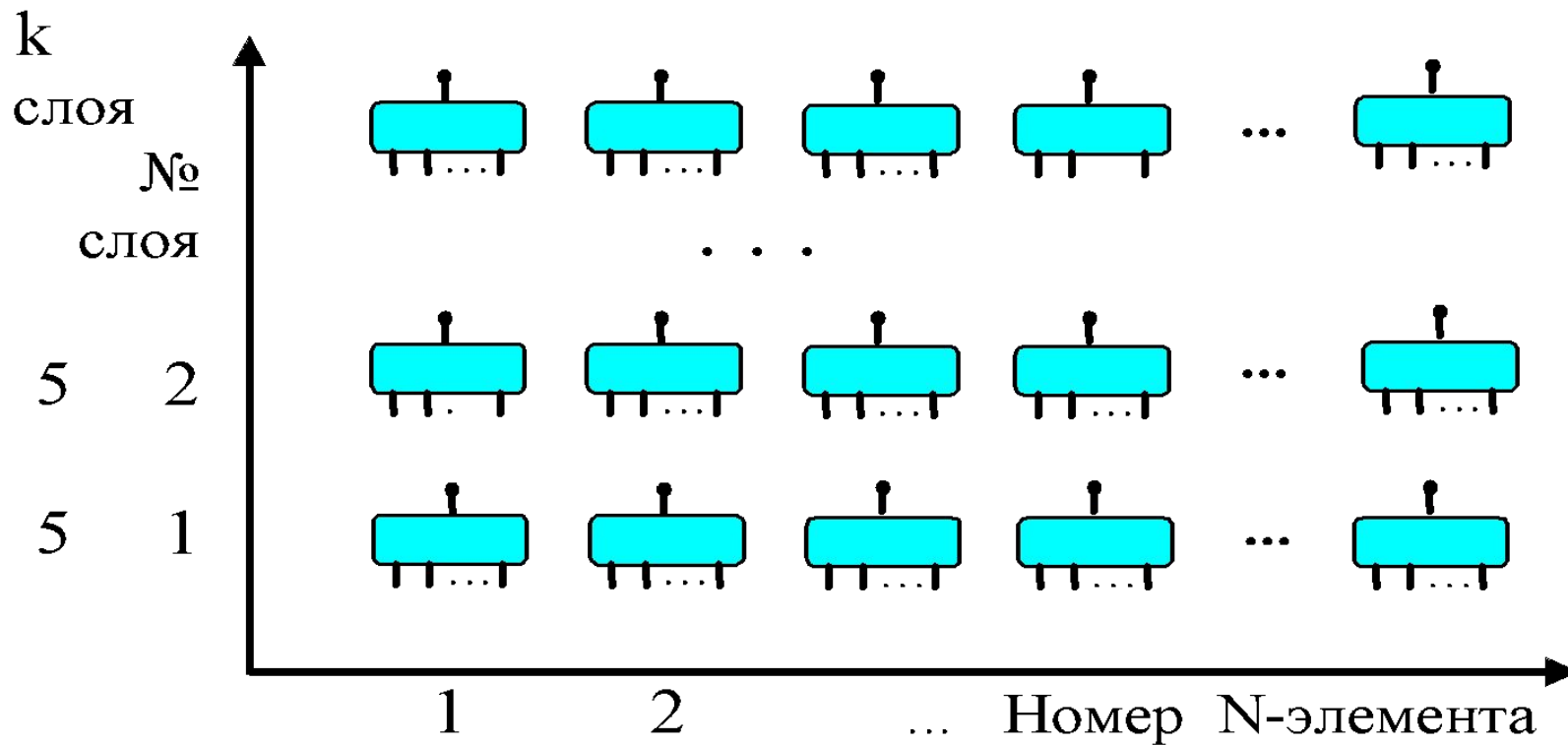
$$U(t) = P_1 * X_1(t_0) + P_2 * X_2 + T_2 * (t_2 - t_0) + \dots + P_k * X_k + T_k * (t_k - t_0)$$

$F_i(P, T)$

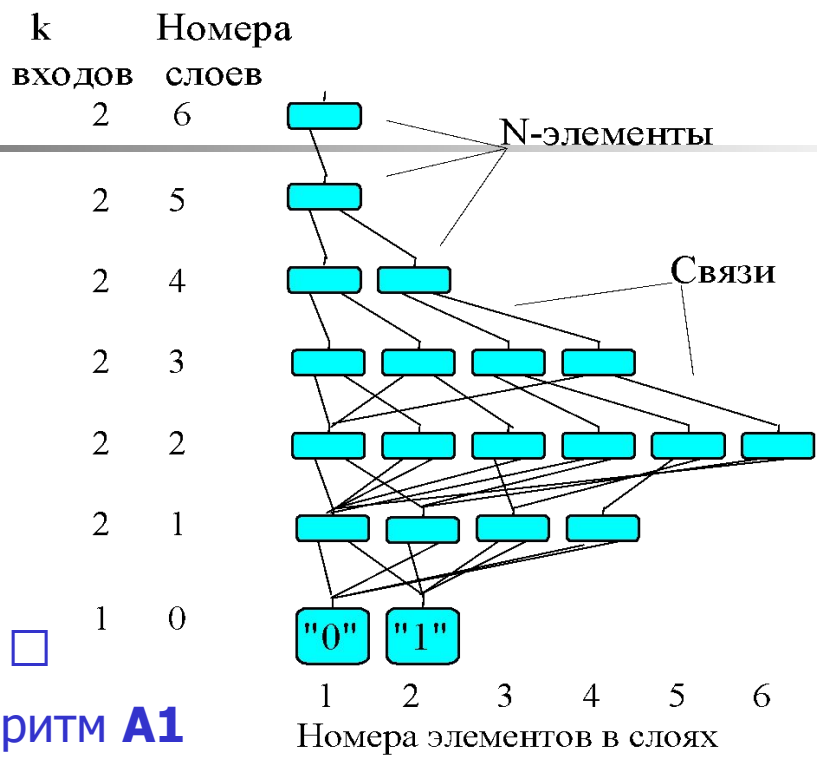
$$U(t+1) = U(t) / a$$



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



Структурное описание НСС можно представить в виде **многослойного графа**:



01100101011011011010001101101001 □□

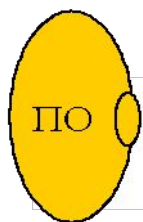
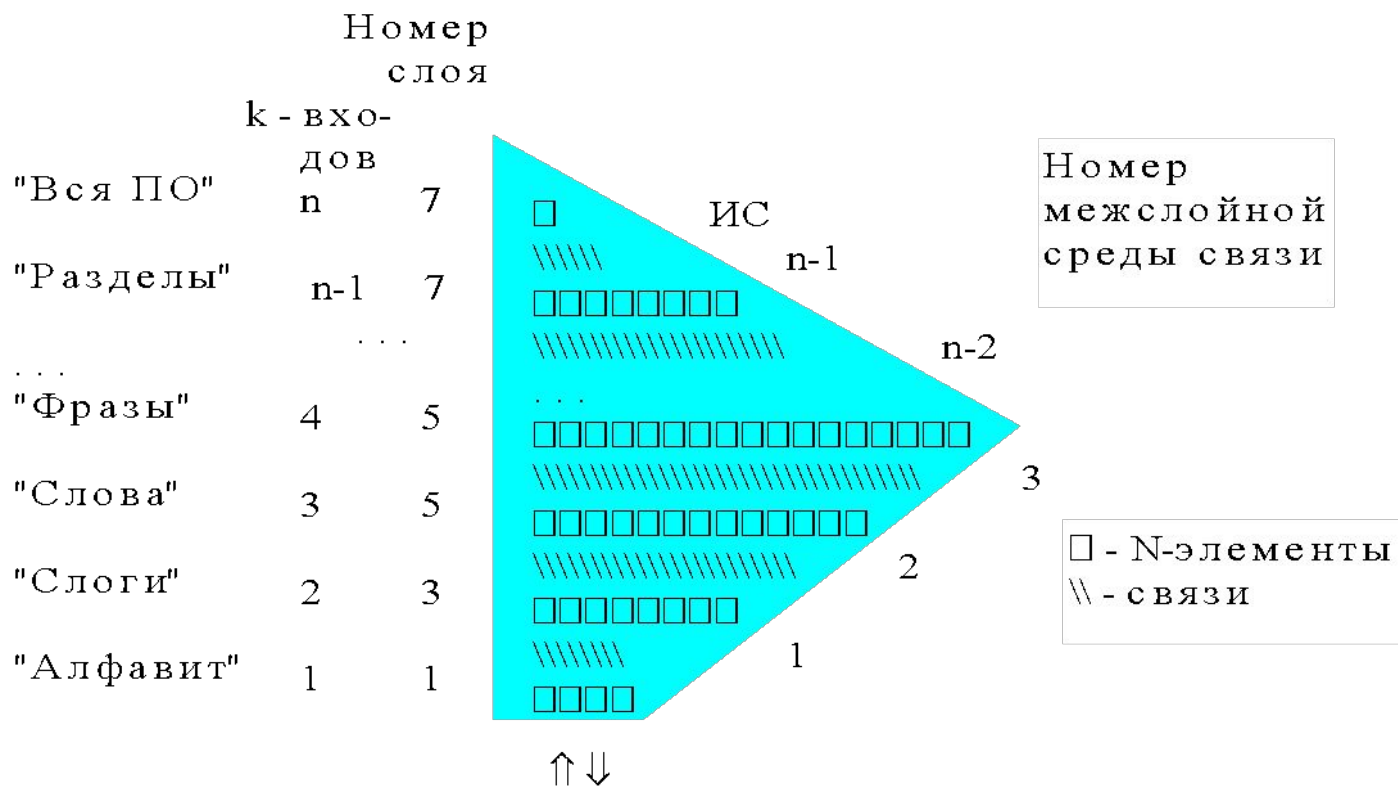
Алгоритм **A1**

Текстовая форма

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 $A_{ЭСЭ}$, 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.

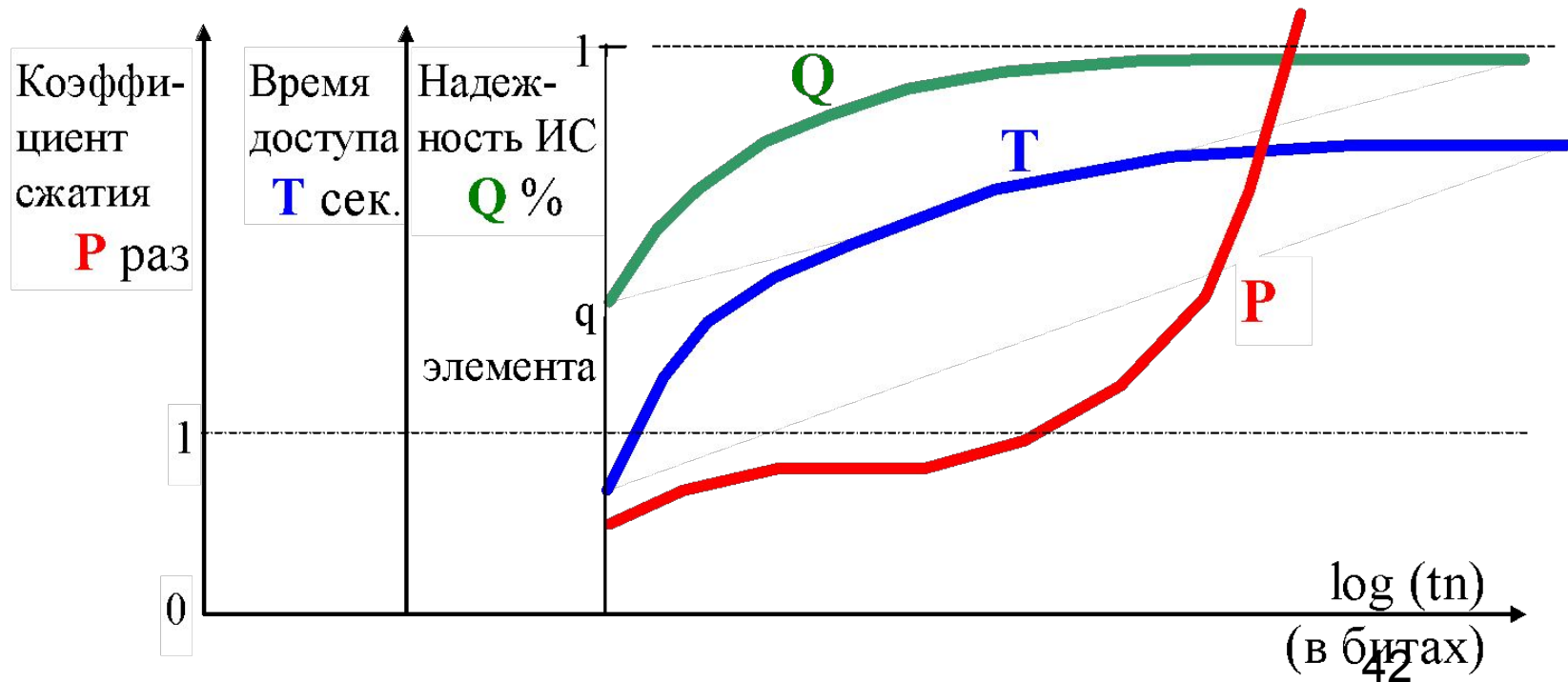


→ abcklltfdbababafgkdkediejddjchhc
Текстовой поток

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

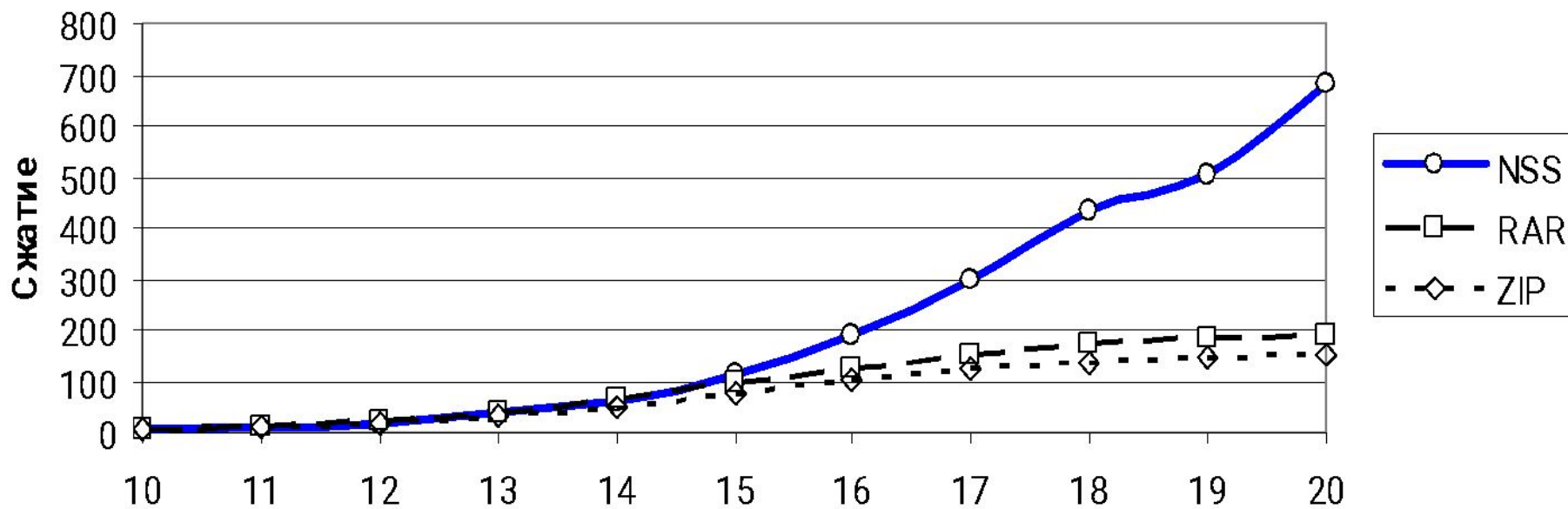
- *время доступа*,
- *коэффициент компрессии-сжатия*,
- *надёжность хранения информации* и др.)

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



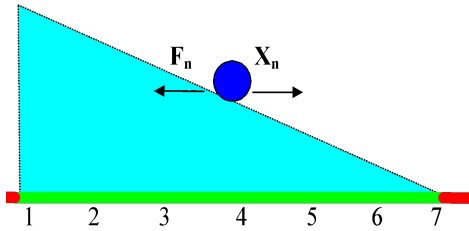
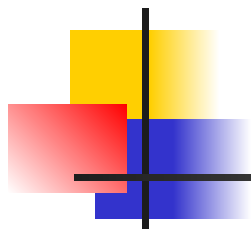
Величина компрессии отражает потенциальную интеллектуальность ИС. Псевдофрактальные файлы.

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

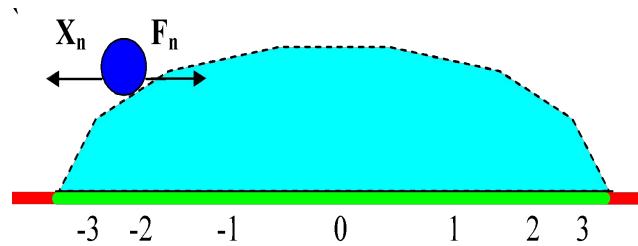


Адаптивные регуляторы на базе НСС

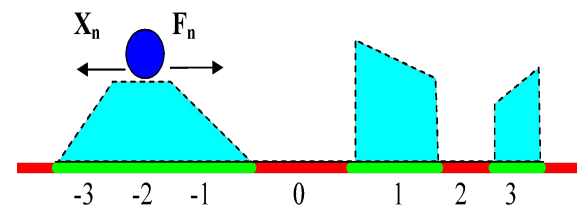
Примеры объектов управления



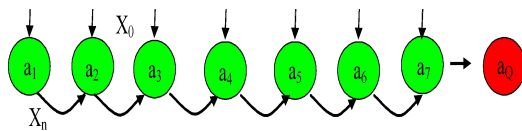
а) "Наклонная плоскость"



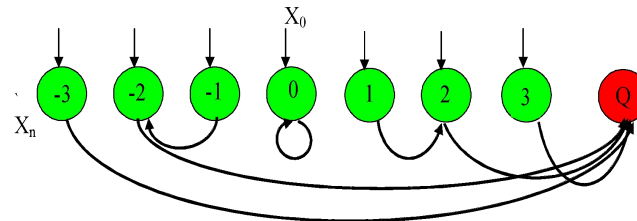
б) "Обратный маятник"



с) "Неизвестный объект"



$$X_{n+1} = 1 + X_n + F_n + Z_n$$



$$X_{n+1} = 2 * X_n + F_n + Z_n$$

$$X_{n+1} = f(X_n) + F_n + Z_n$$

Таблица пространства состояний

$f(X_n)$ и F_n

| | | F_n | | | | | | |
|-------|-------|--------|--------|--------|--------|--------|--------|--------|
| | | $-a_1$ | $-a_2$ | $-a_3$ | $-a_4$ | $-a_5$ | $-a_6$ | $-a_7$ |
| X_n | a_1 | a_1 | Q | Q | Q | Q | Q | Q |
| | a_2 | a_2 | a_1 | Q | Q | Q | Q | Q |
| | a_3 | a_3 | a_2 | a_1 | Q | Q | Q | Q |
| | a_4 | a_4 | a_3 | a_2 | a_1 | Q | Q | Q |
| | a_5 | a_5 | a_4 | a_3 | a_2 | a_1 | Q | Q |
| | a_6 | a_6 | a_5 | a_4 | a_3 | a_2 | a_1 | Q |
| | a_7 | a_7 | a_6 | a_5 | a_4 | a_3 | a_2 | a_1 |
| | | | | S | | | | |

Форма задания объекта

$N = 7$ - число состояний объекта.

$(1\ 2)(2\ 3)(3\ 4)(4\ 5)(5\ 6)(6\ 7)(7\ 8)$

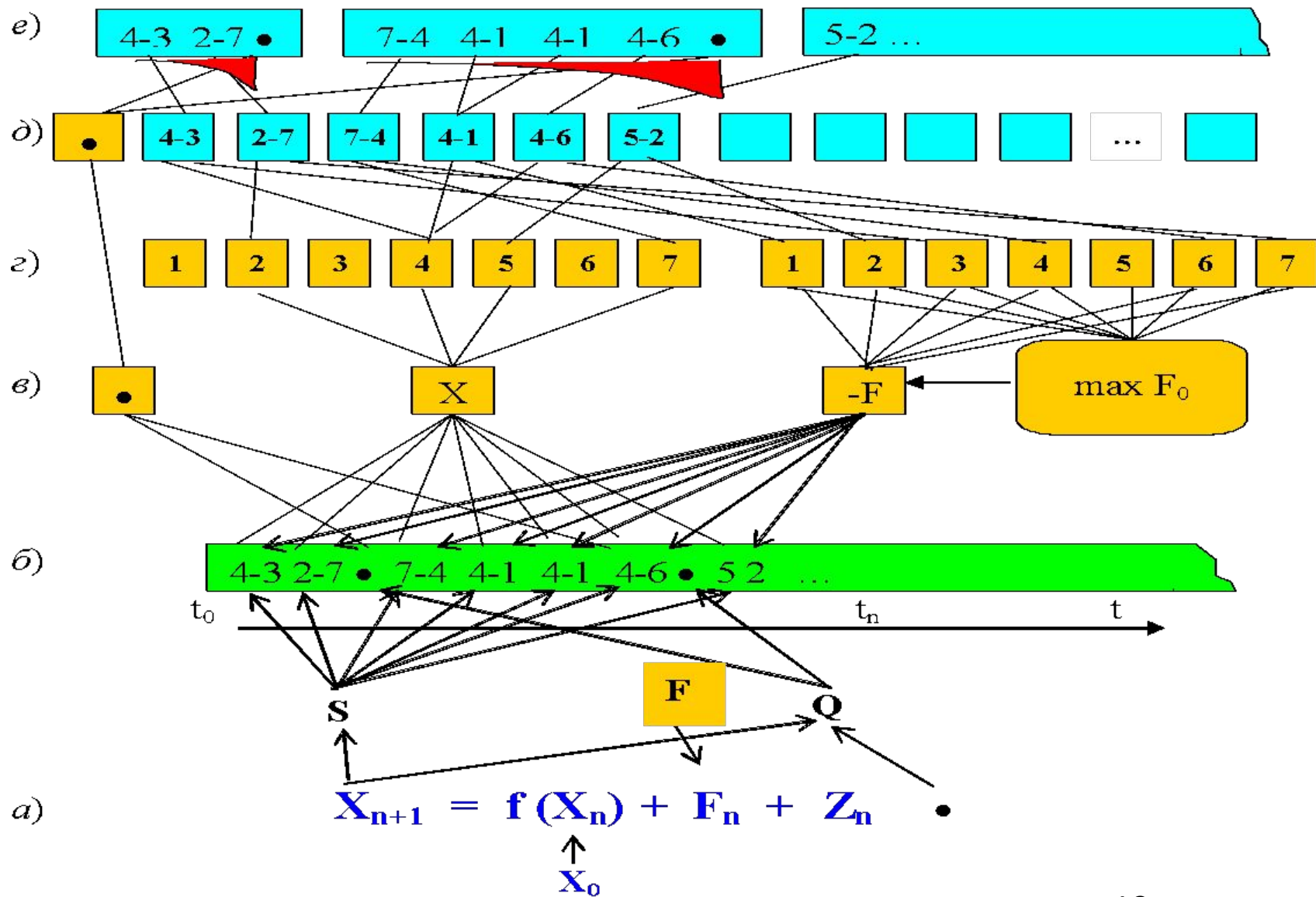
- пары $(X_n\ X_{n+1})$ отражают $f(X_n)$ перехода состояний объекта
 Область нормальных состояний объекта **S** – первый элемент пары $(X_n\ X_{n+1}) = \{1, 2, 3, 4, 5, 6, 7\}$.

$F = -1, -2, -3, -4, -5, -6, -7,$

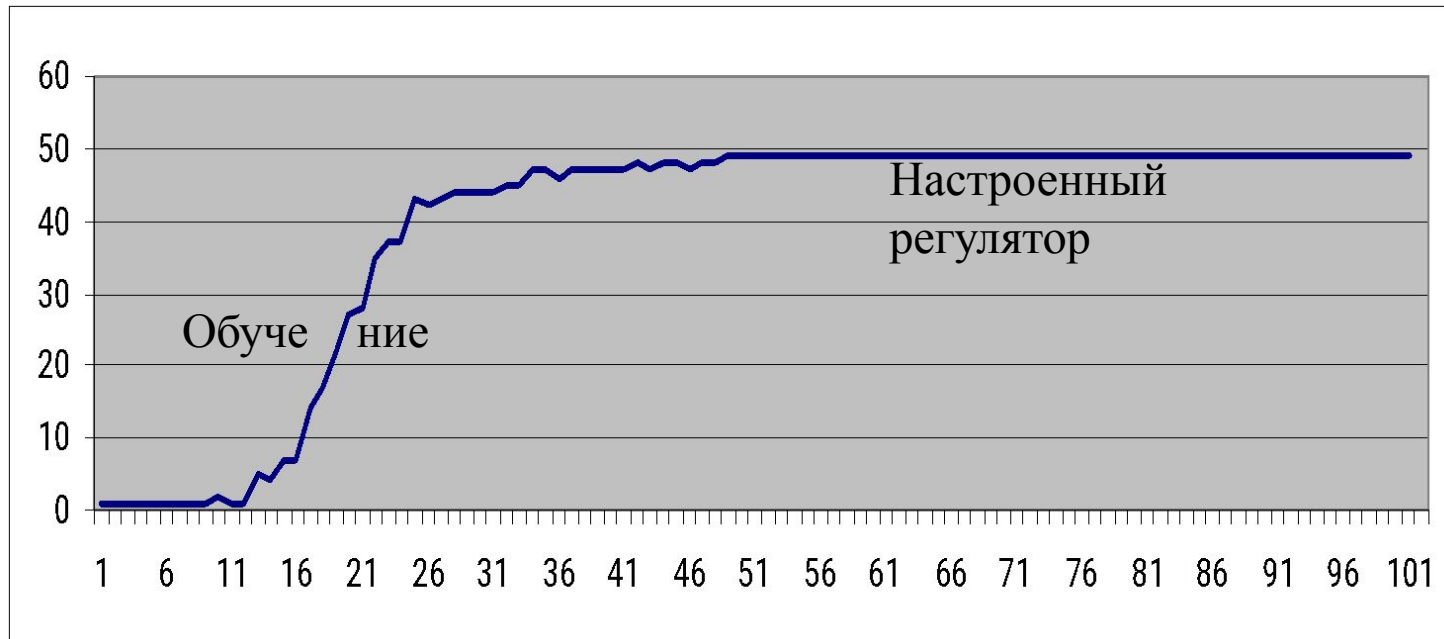
- возможные управляющие воздействия регулятора

$$X_{n+1} = f(X_n) + F_n + Z_n$$

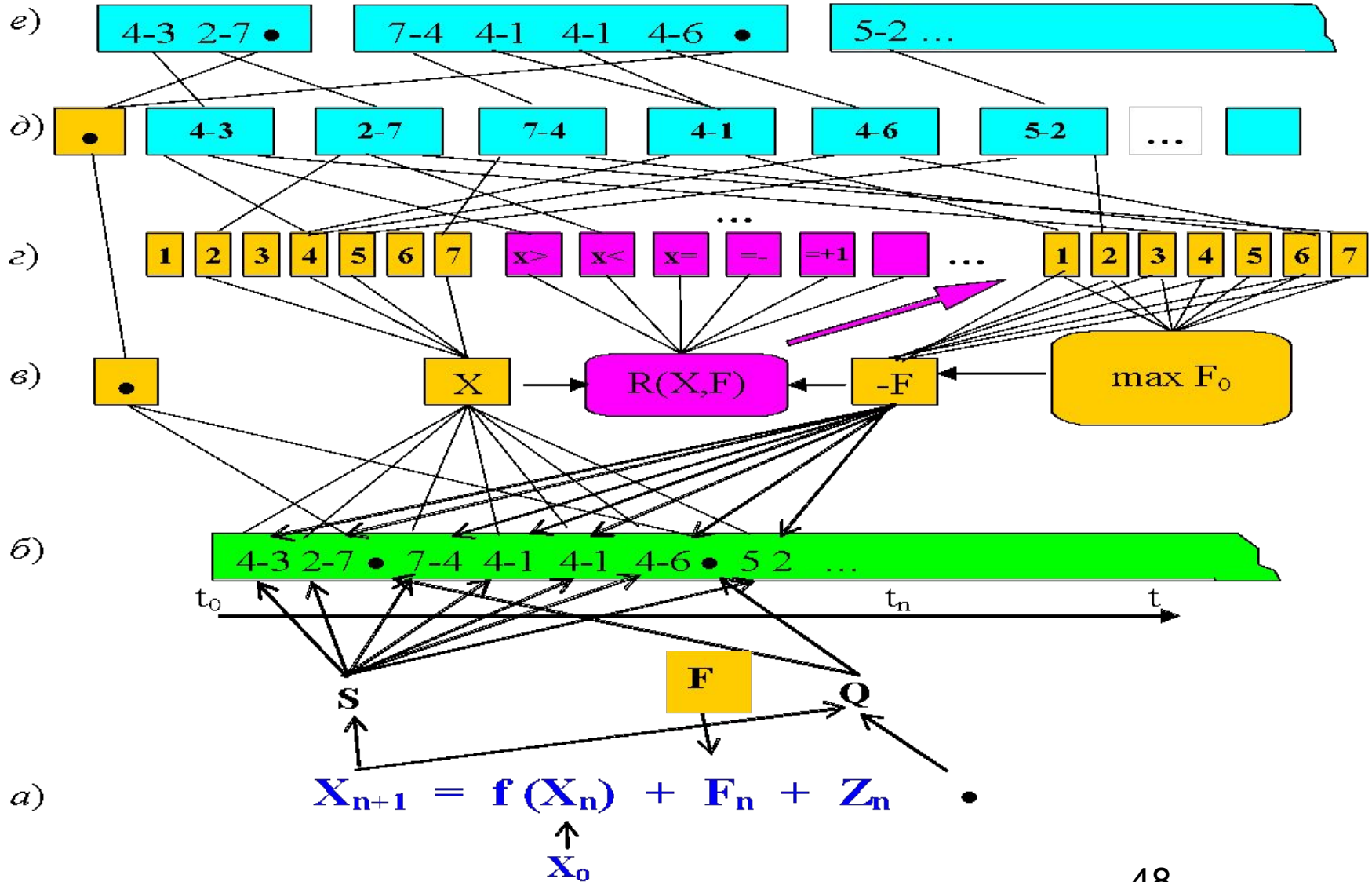
Адаптивный регулятор №1 на базе НСС.



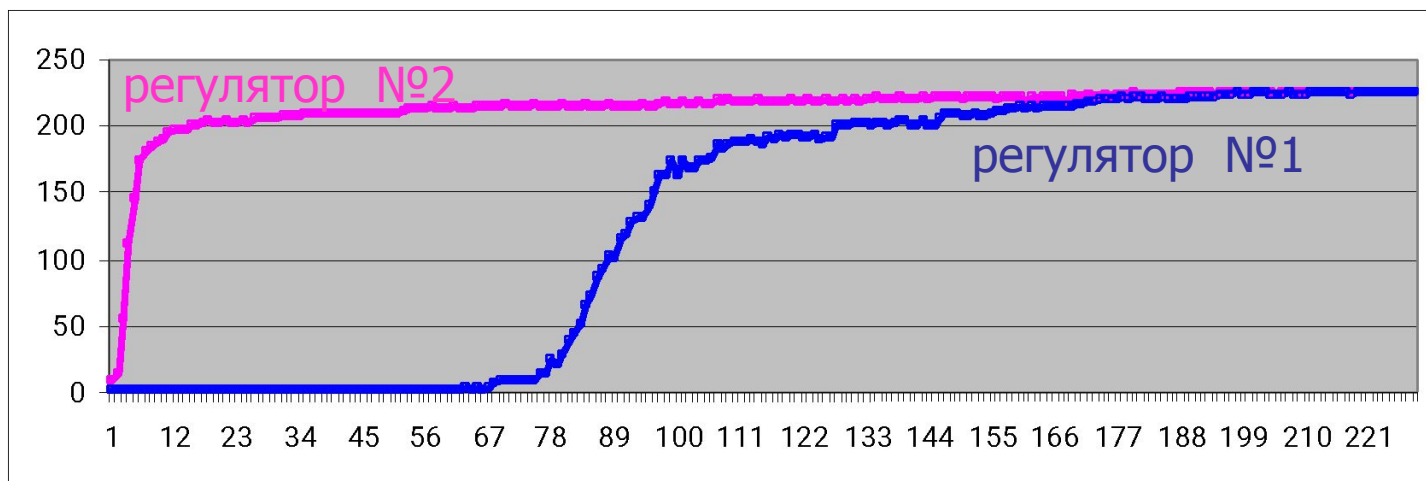
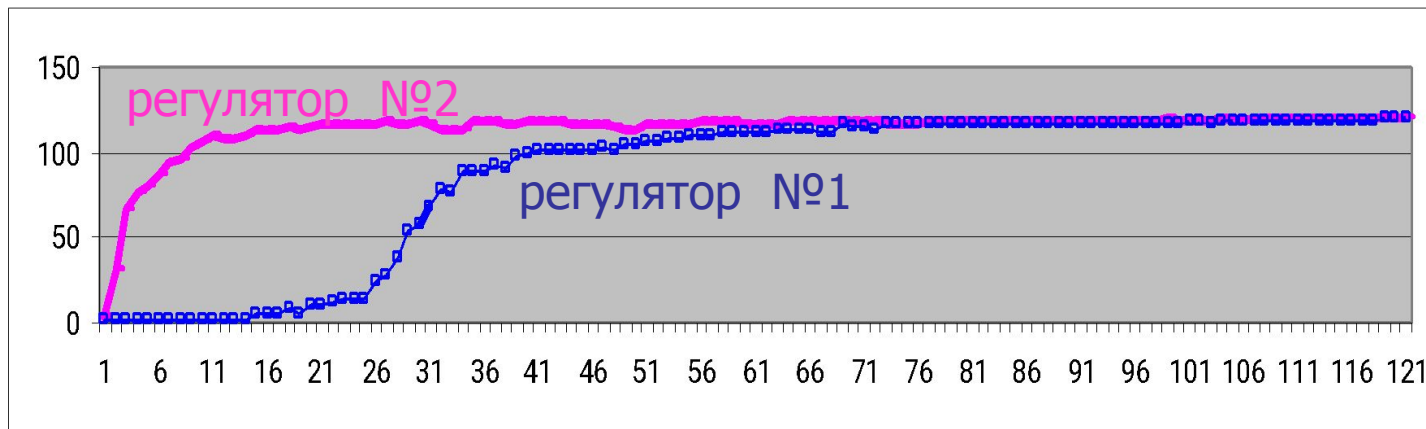
Запоминая пары $\langle X \rangle \langle Y \rangle$ и их оценку $\langle E \rangle$ для любого априорно неизвестного объекта управления ($f(X_n)$), регулятор №1, фактически полным перебором, проходит все возможное пространство его состояний ($N \cdot M$).



Адаптивный регулятор №2



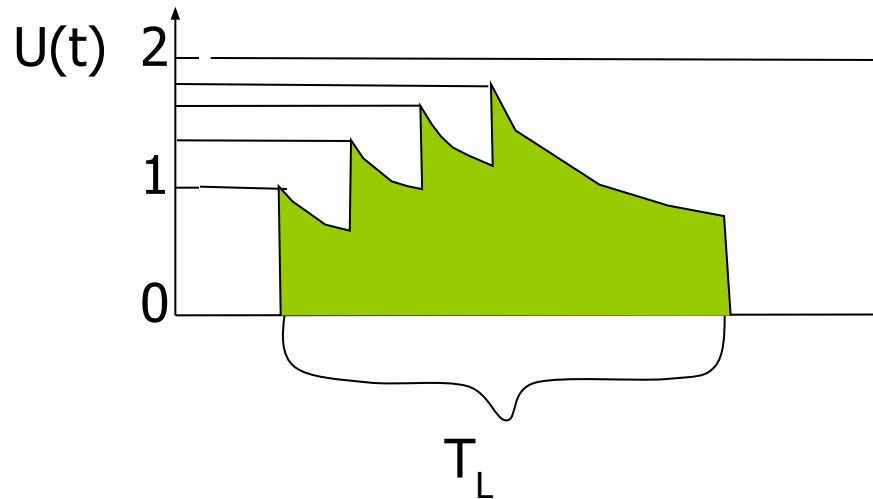
Скорость обучения регулятора №2 почти **не зависит от размерностей N и M** (N – число состояний объекта, M – число состояний возможного управления, в эксперименте для простоты принималось $N = M = 7, 15, 17, 19, 21$), Т.е., пример регулятора №2 демонстрирует возможность практического **преодоления "проклятия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



Формирование понятия абстрактного числа в ИС

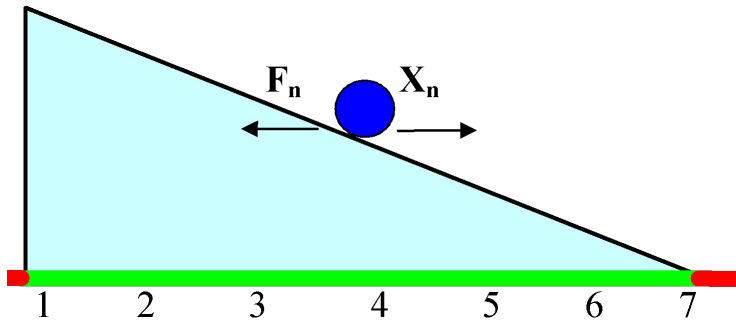
(задача неразрешимая для классического ИИ, и легко решаемая биологическими ИС, как высшими животными, так и человеком).

Одно из физических свойств N-элемента состоит в экспоненциальном падении величины $U(t+1) = (U(t) + 1^*) / \tau$ после его активации (напр., при $\tau=2$)

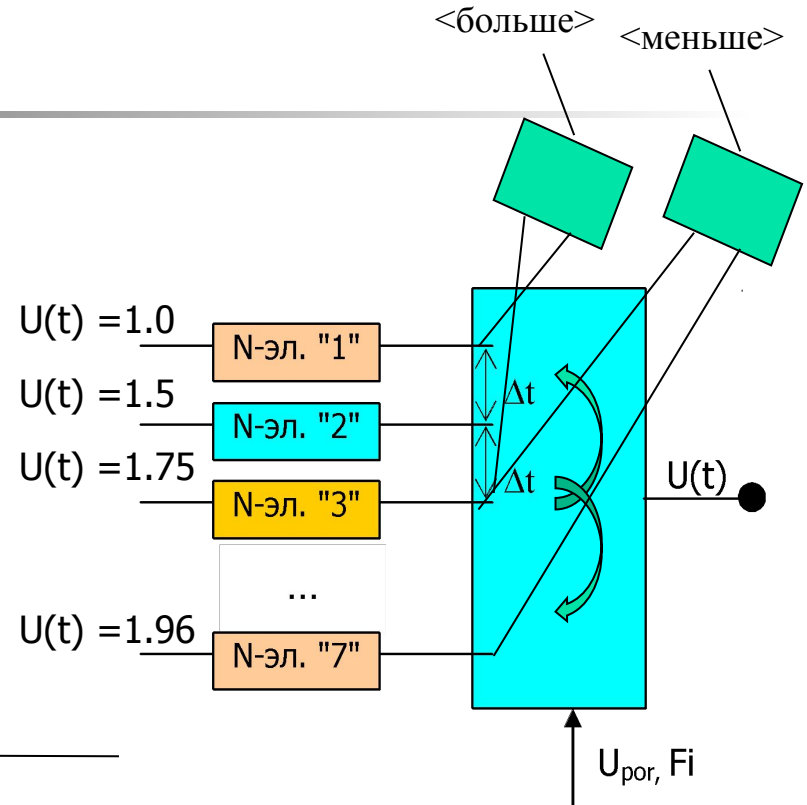
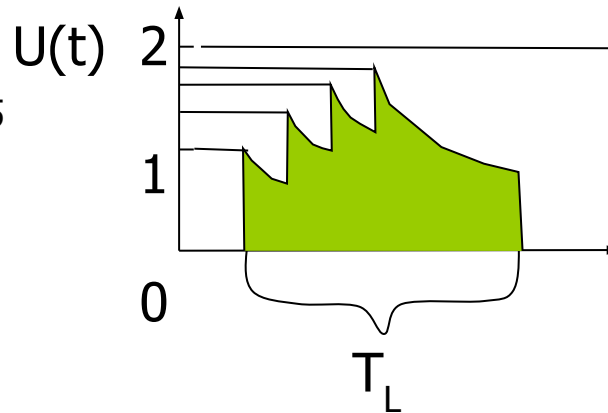
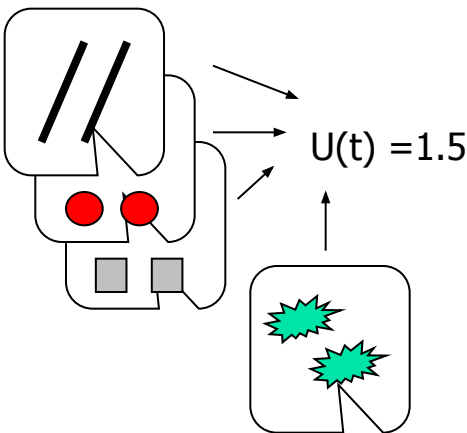


Это пример сведения семантических характеристик ПО к физическим параметрам элементной базы ИС, т.е. сводимым к: См.-Бит-Сек – Система СБС₅₀

Адаптивный регулятор №3



|; ||; |||; ||||; |||||; ... |||||;



<3; 4><число><+Δt><больше>
 <1; 4><число><+2Δt><больше>
 ...
 <3; 2><число><-Δt><меньше>
 <4; 1><число><-3Δt><меньше>

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 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 НСС-регуляторах можно продемонстрировать:

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

Пространство №1 >> №2 >> №3 : **полного перебора >> пространство отношений >> энтропийное пространство**

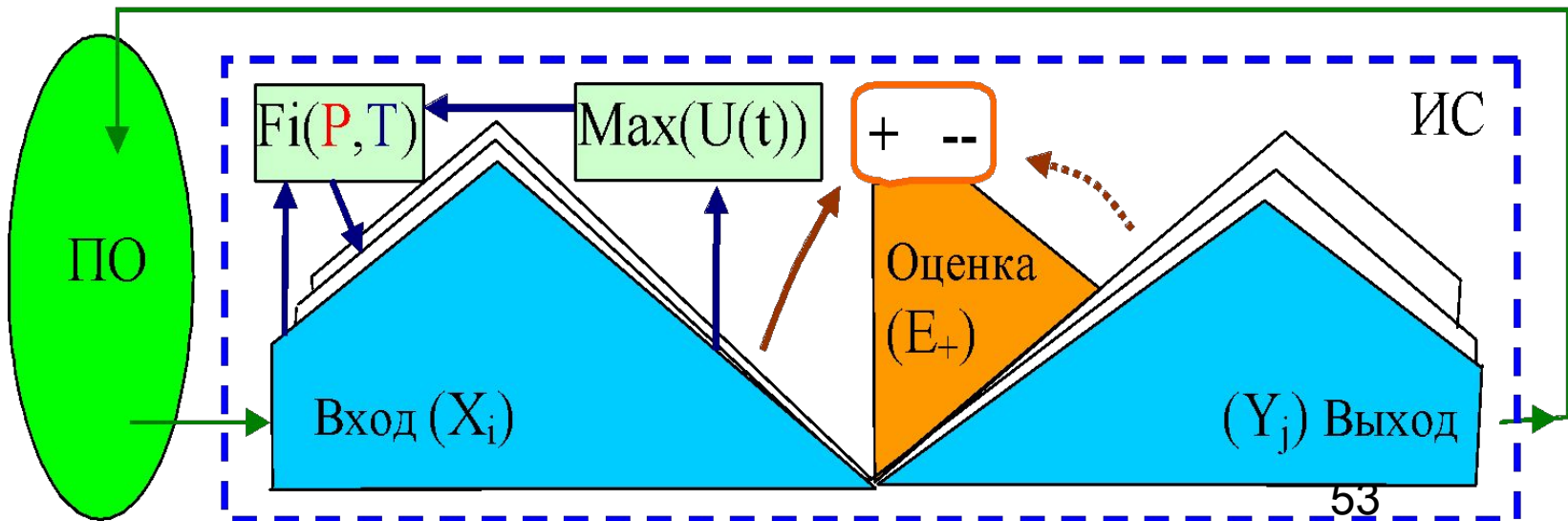
Функционирование примитивных ИС

В ИС отображается только **объективная информация**.

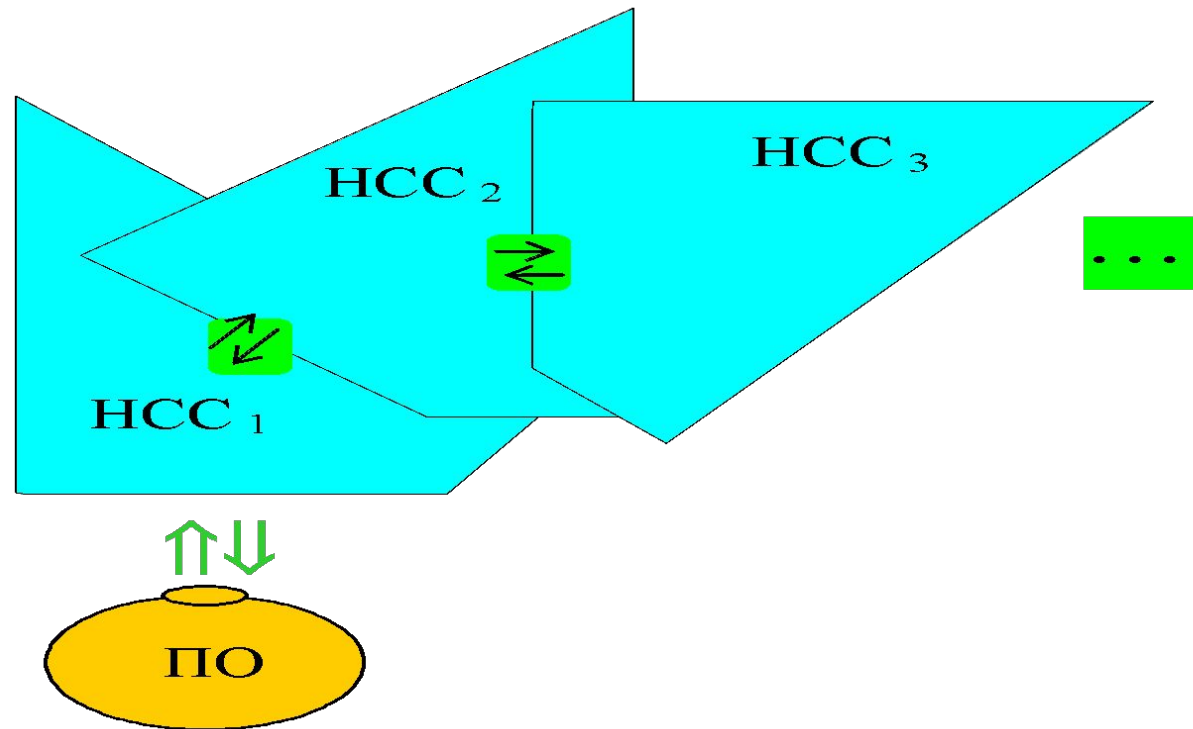
При этом, на каждое **конкретное раздражение (X_i)**

– **конкретная реакция (Y_j)**,

с "эмоциональной" оценкой ее полезности для ИС.



Память "сложных" ИС строится как иерархическая структура из НСС. На фактографическую НСС₁, настраивается НСС₂ ("ортогональная"), которая функционально воспринимает НСС₁, также как НСС₁ воспринимает ПО.



По нашим оценкам, структуры из 4-5-ти иерархических НСС вполне достаточно, чтобы моделировать основные психические функции человека

ЗНАНИЕ (например, "логические высказывания")
в текстовой форме обязательно содержат
ритмические **повторения!**

|ВСЕ|ВОЛК|И| |СЕРЫ|А|ЕСТЬ| ВОЛК|,|ЗНАЧИТ|,|А|ЕСТЬ| СЕРЫЙ|<E+>
|ВСЕ|ПЕТУХ|И|КРАСИВ|Ы|В|ЕСТЬ|ПЕТУХ|,|ЗНАЧИТ|,|В|ЕСТЬ|КРАСИВ|ЫЙ|<E+> } НСС₁

в НСС₂ автоматически строится структура:

|ВСЕ|X1|И|X2|Ы|,|X3|ЕСТЬ|X1|,|ЗНАЧИТ|,|X3|ЕСТЬ|X2|ЫЙ|<E+>
Δ t1 Δ t3 Δ t2

<ВСЕ ПЛЮКИ КАНЫ, ОН ЕСТЬ ПЛЮК, ЗНАЧИТ, ОН ЕСТЬ > Δ t2 + ЫЙ □ <КАНЫЙ>

<ВСЕ ЛЮДИ СМЕРТНЫ, СОКРАТ ЕСТЬ ЛЮД, ЗНАЧИТ, СОКРАТ ЕСТЬ >

ИС в своей грамматике выдаст: <СМЕРТНЫЙ>

Таким образом, все типы логических высказываний
произвольной ПО покрываются одним механизмом НСС

Информационный ресурс:

"сигнал – информация – знание"

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

"линейный – **сигнал**", $R_{\text{ИС}} = K \cdot t + b_1$

Сигнал - простая суперпозиция ЭСЕ ПО;

"логарифмический – **информация**"

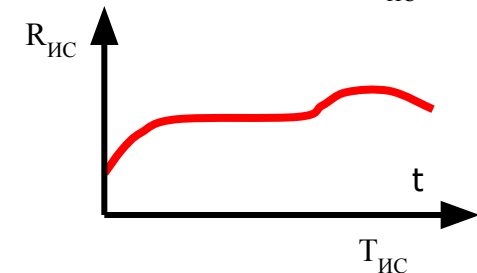
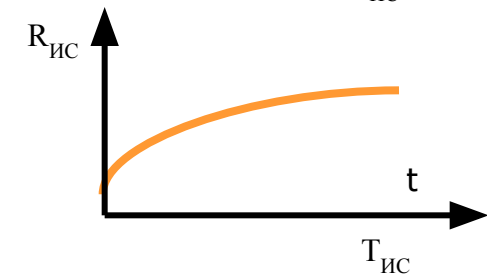
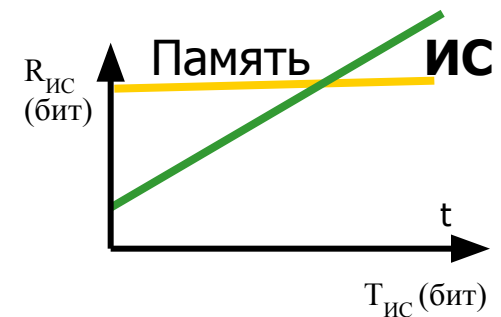
$$R_{\text{ИС}} = \log(t) + b_2$$

Информация – знаковая последовательность на языке системы, соответствующая **целому числу причинно-связанных процессов ПО, т.е. ЭСЕ**

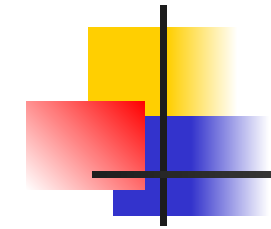
Данные – форматированная под ПО **информация**

"const - **знание**". $R_{\text{ИС}} = \text{const} + b_3$

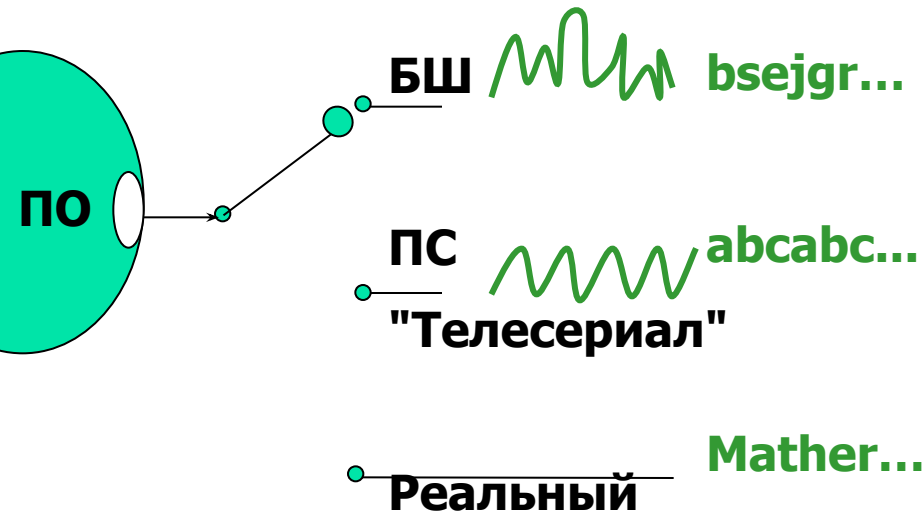
Знание – совокупность ЭСЕ, отражающих одновременно процессы ПО и процессы в самой ИС



Отображение информационного ресурса тремя классами ИС



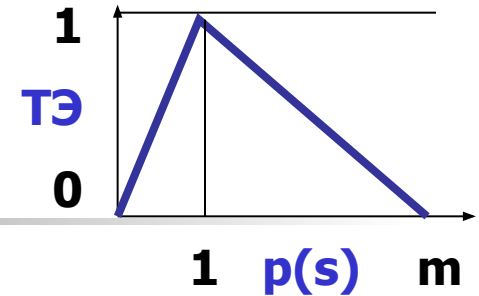
Автомат "животн." "ИС-человек"



| | | |
|--------|------------|------------|
| Сигнал | Сигнал | Сигнал |
| Сигнал | Информация | Информация |
| Сигнал | Информация | Знание |

Текстовая энтропия

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



$p(s)$ - частота на интервале $L \cdot A^L$, $0 \leq p(s) \leq m$;

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

s - некоторое слово длиной в L символов;

m - размер потенциального S -словаря в $m = A^L$ S -слов

информационное пространство в $L \cdot m = L \cdot A^L$ СИМВОЛОВ

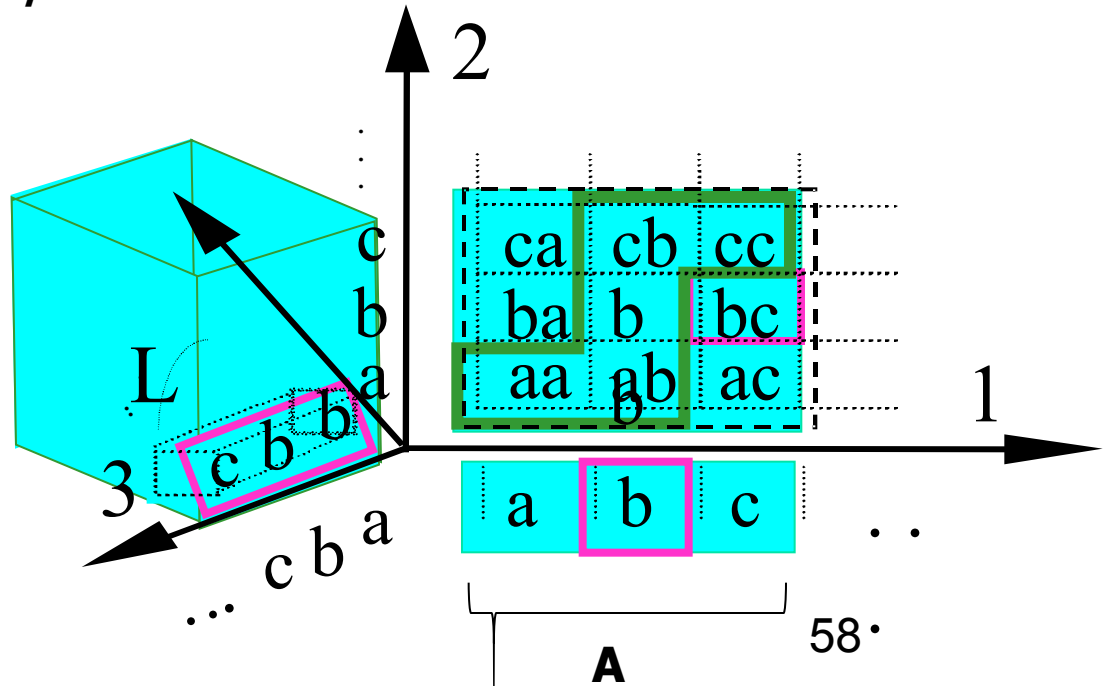
$$TЭ_L = TЭ(s)L / m$$

$$m = A^L$$

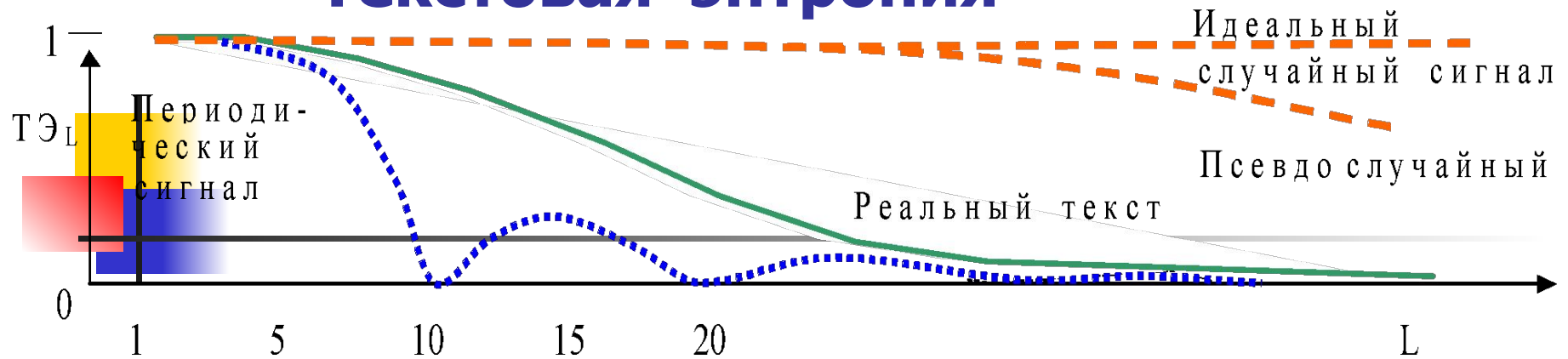
...kksxkdkdfjfhgh...

$$T = L \cdot A^L$$

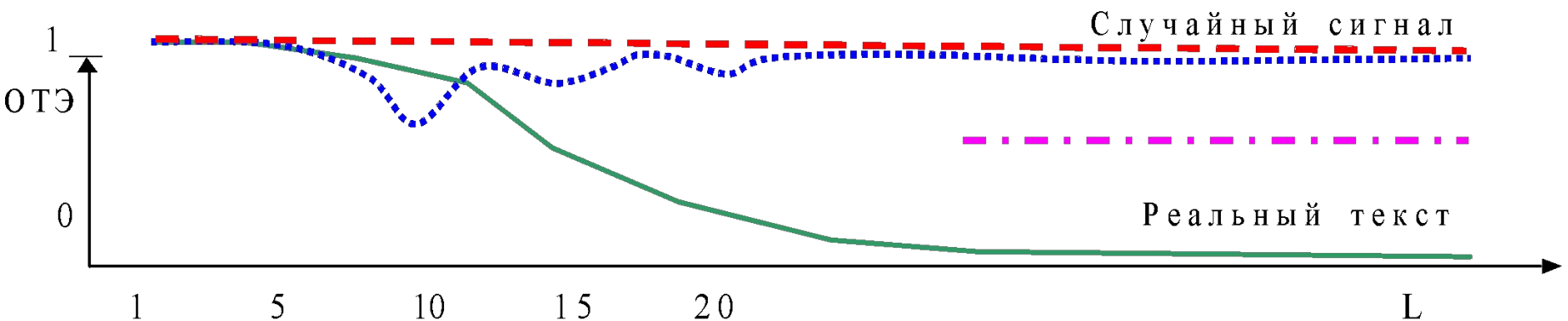
Например, $p(dkd)_{L=3}$



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).

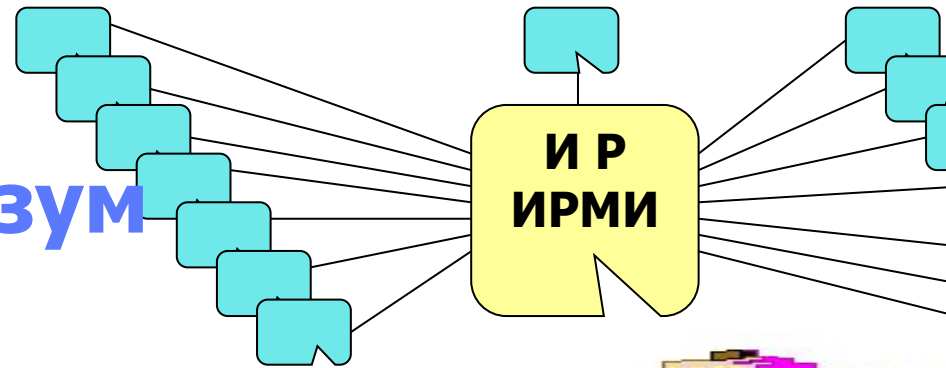
Решение

Проекты:

- Искусственный разум

- Информоград

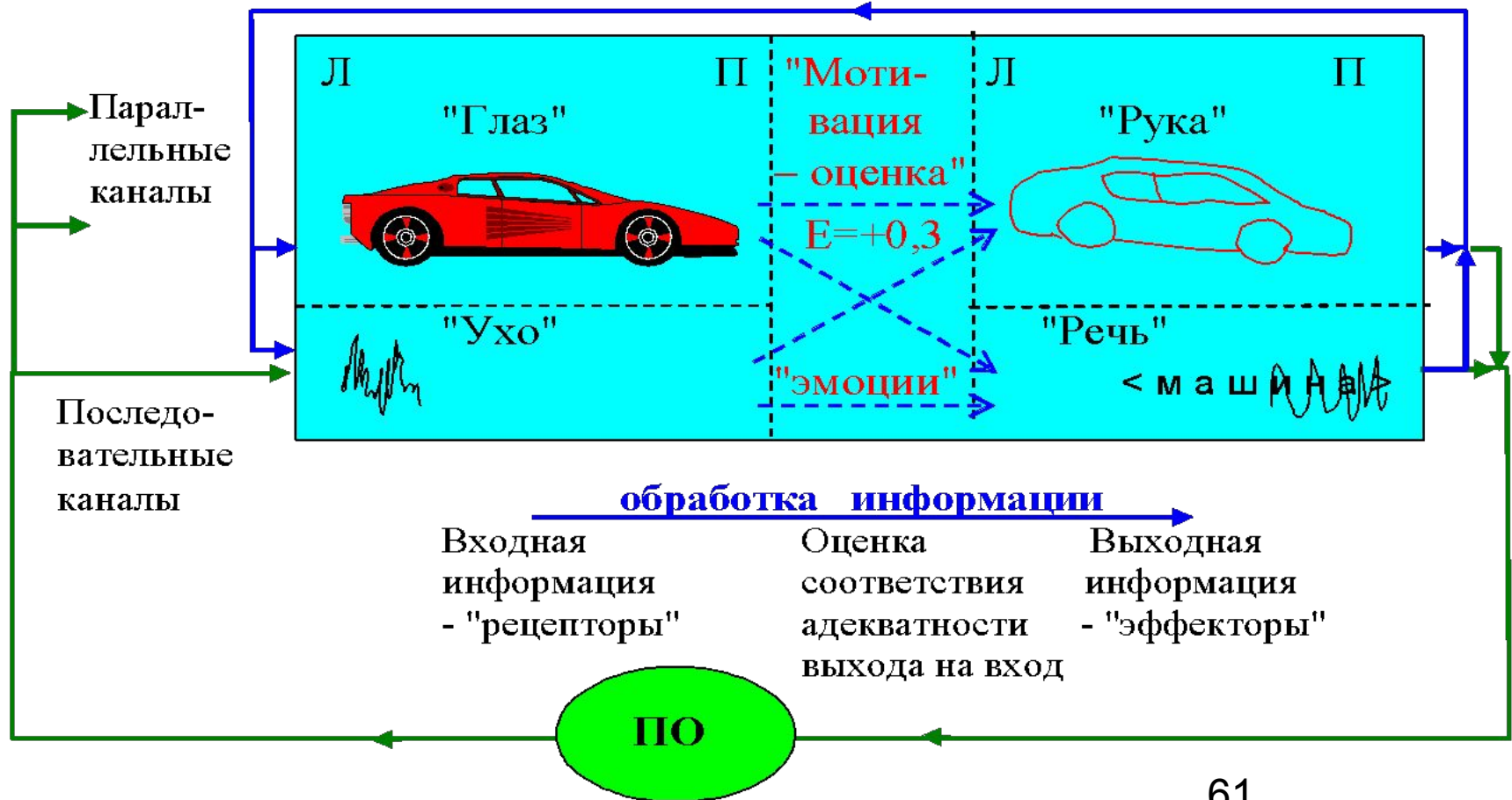
- Восхождение разума



Проект - Искусственный разум

Прототипы компьютерного интерфейса: "человек-ИР"

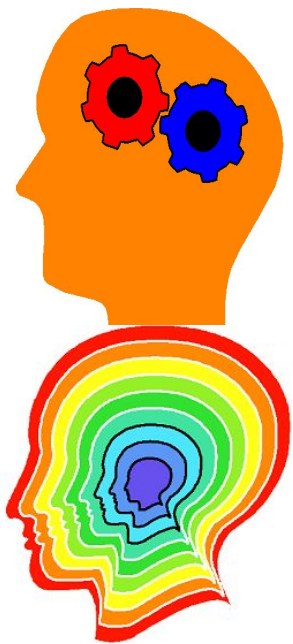
саморефлексия



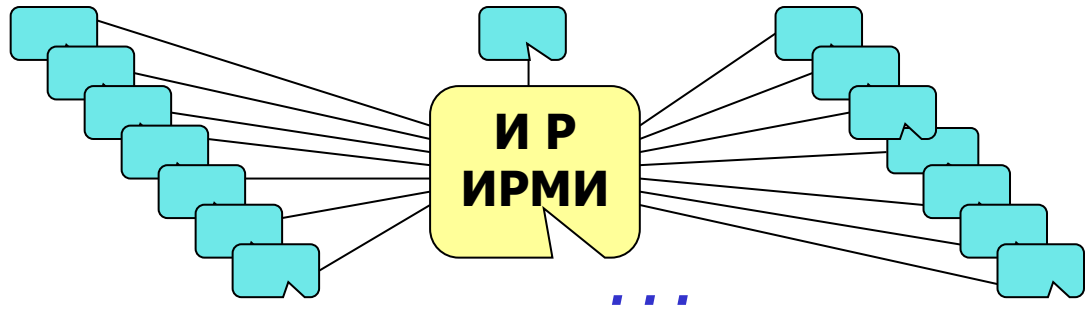
По нашим оценкам

- **через 12 месяцев** возможно получить программную реализацию описанного ИР.
- **за 18 и 24 месяца**, на базе многопроцессорных ЭВМ (256-1024 RISC процессоров), сформировать многопроцессорный кластер и чисто аппаратную реализации ИР, реализуя естественную параллельность процессов в НСС.
- ИР - это **на 6-9 порядков ускорение** решения любых "интеллектуальных" задач с одновременным **повышением их сложности (числа компонент)** на **5-6 порядков**.
- Для цивилизации, появление ИР, это как реальный шанс ее будущего, особенно в сегодняшнее неуправляемое время **"технологий глобального (воз)действия"** (биотехнологий, информационных, ядерных, финансовых), которые, в руках террористов, превращаются в орудие ада.
- **Без ИР человечеству в 21 веке практически не выжить** и в подтверждение этого тезиса можно привести множество доводов

Человек и искусственный разум



* 10^{6-10} =



Объем памяти $10^{12}-10^{16}$ образов (10^{24} текста, графики, ...)
Режим работы on-line: 50 пользователей - станция Sun;
1000 пользователей – многопроцессорная супер ЭВМ;
100 000 пользователей – аппаратная реализация ИРМИ.
Язык диалога - естественный для каждого пользователя-исследователя.

Когнитивные функции ИРМИ не уступают человеческим

Проблемы решаемые только с помощью ИР

- on-line структуризация и формирование глобальных БЗ;
- "Госплан" Планеты;
- Глобальный Университет (обучение);
- Промышленная разработка информационного ресурса;
- Безопасность Планеты;

...

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"**.

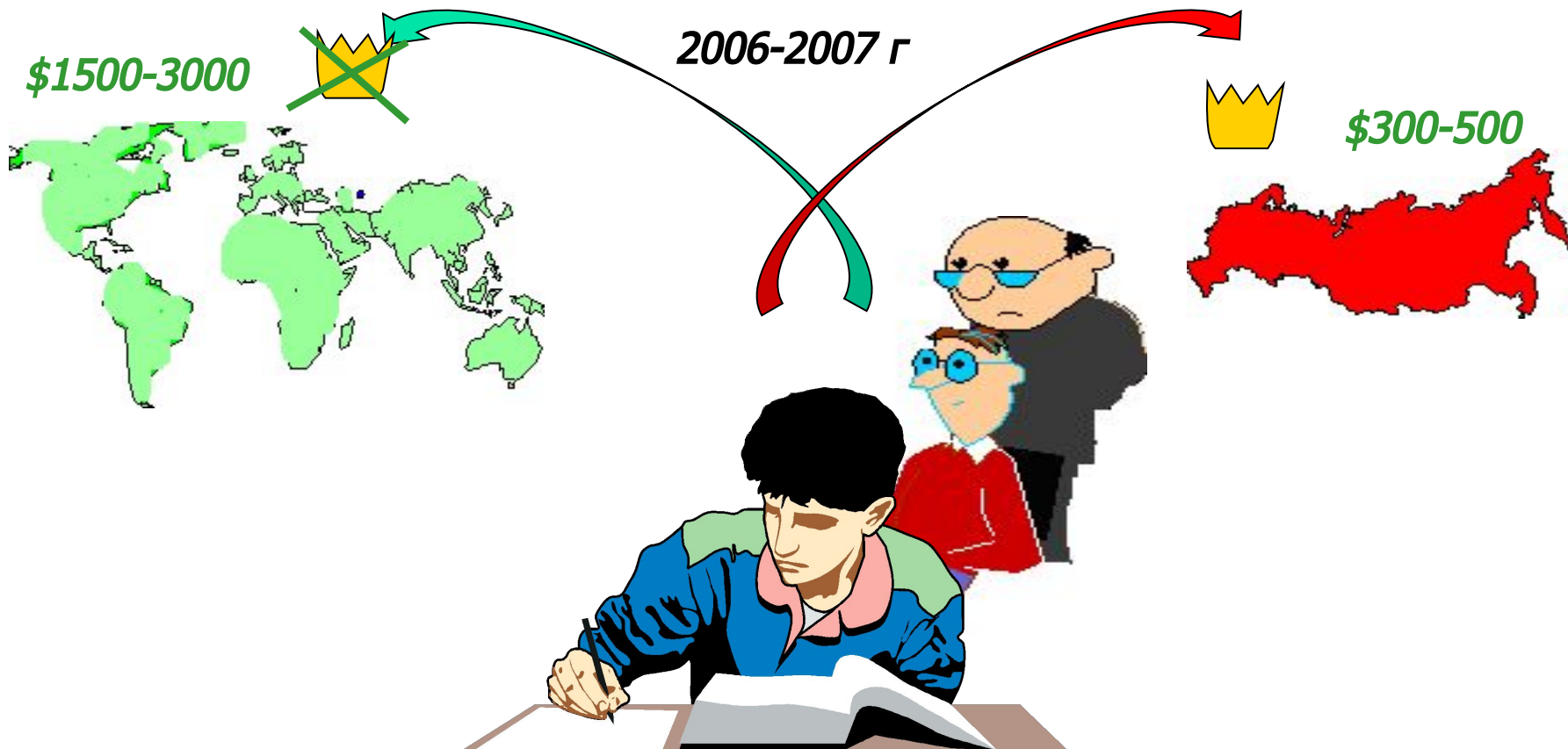
Российский центр поддержки инноваций (проект "Информоград")

ИРМИ в рамках **Российского центра поддержки инноваций** (проект "Информоград") каждому исследователю даст в свое распоряжение колоссальные материальные и интеллектуальные возможности, о которых до этого не могли мечтать и целые государства.

В процессе взаимодействия при решении задачи, **ИРМИ помогает человеку получить новое решение**, человек при этом обучает ИР, представляя ему уже известные знания.

Новое же знание, полученное в ходе сотрудничества ИРМИ и человека, **становится их общим знанием**, взаимообогащая их обоих, как бы более интеллектуально ИР не превосходил человека.





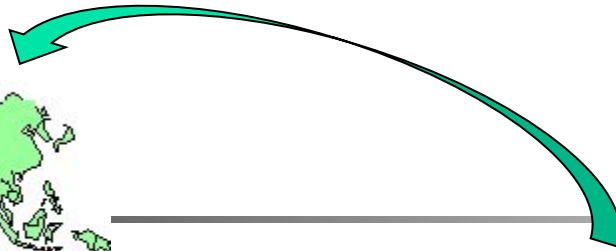
2006-2007 г

"Информград"

**Российский информационно-аналитический центр
поддержки инновационных работ**



\$1500-3000



Молодой российский выпускник за границей :

-Материальный достаток и хорошие бытовые условия

-Новейшее оборудование

-Иная культурная среда

-Работа на вторых ролях (max post-doctor = м.н.с.)

-Малая перспектива роста

~~-2004 г число возвращающихся высококлассных специалистов (МФТИ, МГУ, МГТУ, МИФИ) сравнялось с числом отъезжающих за рубеж~~



Аутсорсинг

\$500-1000



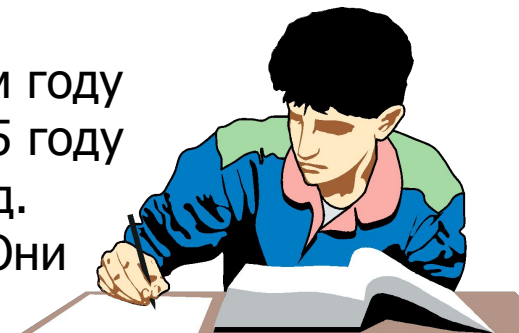
"Чем больше работаешь – тем больше остаешься должен" !

Российский рынок информационной безопасности в прошлом году составил \$170 млн. и по предварительным прогнозам, в 2005 году вырастет до \$230 млн., мировой рынок составит ~ \$200 млрд.

В этой области работает ~ 1 000 программистов из России. Они получают ~ \$1000, сумма = \$12 млн./год

\$230 млн. - \$12 млн. = \$220 млн. **??!** (выплаты из России)

В \$200 млрд. (общ. объем \$1 трлн.) ~ 10% труда россиян, а получают **0,5%** или **1/200** от прибыли, принесенной **ТНЖ**



Пример: 10 чел. за 12 мес. для ТНЖ делают ПО ИИС (цена ПО \$1000)

Получают по \$1200/мес. Семьи по 5 человек. Прож. минимум \$200.

Доход семьи (страны) = \$2400/год.

Семья разработчика ПО не может купить свой ПО (2400 << 5*1000) !!!



Нет более ценного товара, чем ИНФОРМАЦИЯ (знание)

- *Инновационная экономика - стратегическое направление развития России в XXI веке*
- *Создание единого информационного пространства (информационной инфраструктуры) технологии разработки знаний*
- *Как не богата Росси я природными ресурсами, но основное её "богатство в мозгах, а не в недрах" Смена социально-экономической философии*

Не догонять, а упреждать !!!

Не ЭВМ, а Искусственный разум !!!

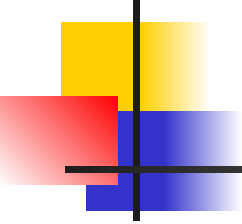
"Информград" ЭТО:

- Повышение производительности труда в 1000 раз
= 10 (этап постановки) * 10 (НИР) * 10 (НИОКР)
- Экспериментально-опытное производство (ЧПУ,...)
= ГОТОВЫЕ ИЗДЕЛИЯ (патент, самолет, ...)
- Новые социально-экономические отношения
= Ноосфера (социальный организм)
- 2012 год Россия мировой лидер социально-экономического развития (РУССКИЙ КОСМИЗМ)





Проект - **Восхождение разума**



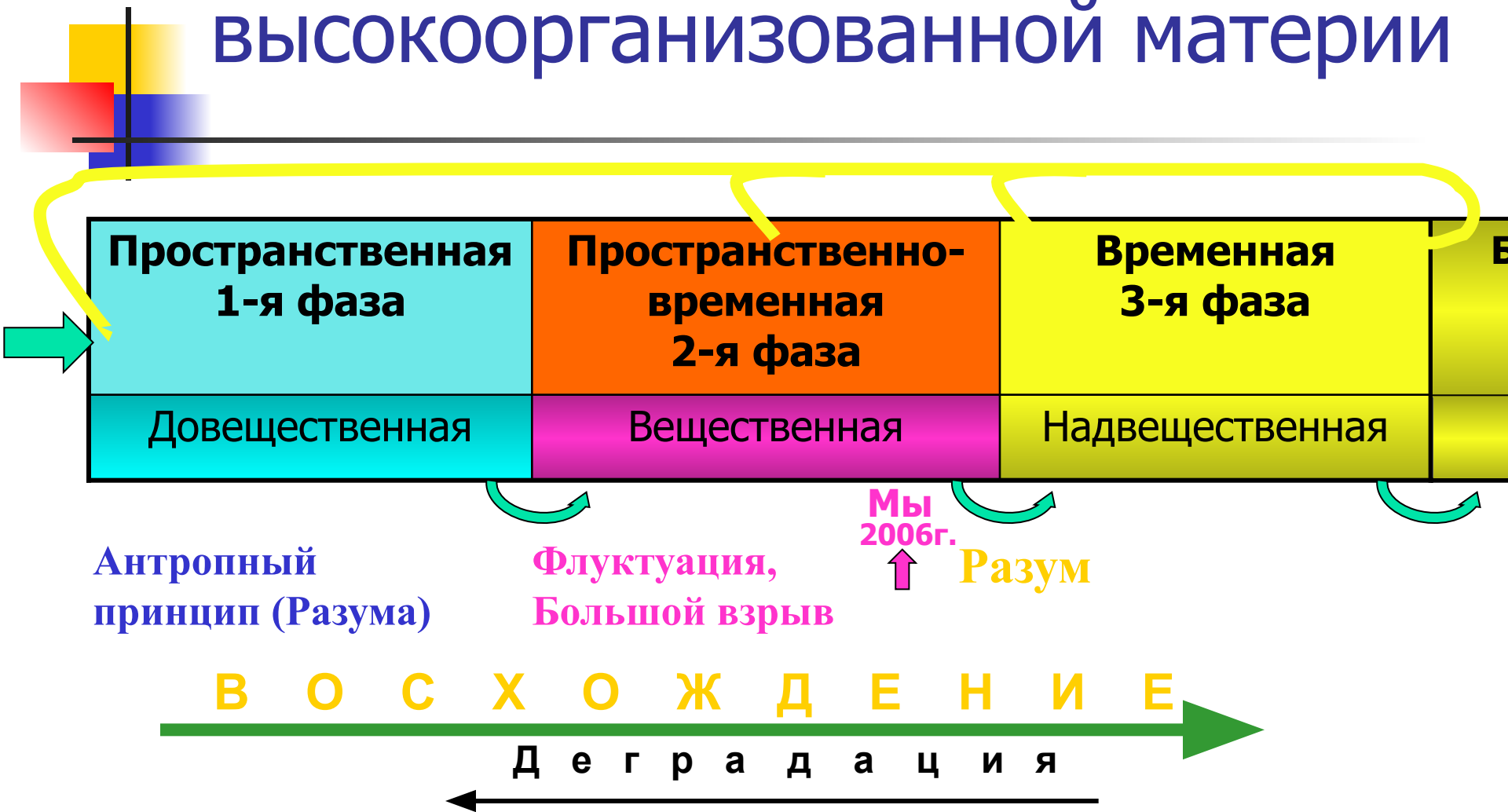
Характеристики эволюционного процесса высокоорганизованной материи

Цель – пролонгация лидирующего вида до горизонта прогнозирования будущего ($n(t), t \rightarrow \infty$).
– экспансия свободы вида ($n(t) = f(a^t), t \rightarrow \infty, Q$).

Мера – эволюционный потенциал ($\Delta(E) - H\Delta(E)$),
доступный ресурс ($n(t) \rightarrow \max(E) \rightarrow Q$),
компрессия ($P \rightarrow \max \rightarrow Q$).

Механизмы – законы экспоненциальности ($f(a^t)$),
и сверхаддитивности $f(s_1, \dots, s_n) \gg f(s_1) + \dots + f(s_n)$
и алгоритмы НСС (C++, ...)

Фазы эволюции высокоорганизованной материи



Эволюция форм материи и ее пространственно-временных фаз



СПАСИБО ЗА ВНИМАНИЕ

По всем вопросам Проекта (национального)
ИСКУССТВЕННЫЙ РАЗУМ

обращаться к Бодякину Владимиру

E-mail: body@ipu.ru

www.informograd.narod.ru/

тел. (095)334-92-39

в ИПУ тел 1-346



Свойство **доминирования** "внутренних" информационных процессов (от $HCC_5 > HCC_1$) над "внешними" информационными процессами идущими от HCC_1 , назовем **сознанием** ИС.

Особенностью нейросемантической ИС является то, что в N-элементе происходит слияние общепринятых понятий **данных и алгоритма**, как неделимого пространственно-функционального элемента – **образа**.

"**Метаалгоритмом**" поведения ИС является асимптотическая **минимизация отображения** (при $t \rightarrow \infty$), всех значимых для неё текстовых потоков.

Формальная схема иерархического построения памяти ИС приводит к **расширению языка**, что позволяет "обойти" **запрет теоремы Гёделя о неполноте формальных систем**.

Так, если язык фактографической ИС (НСС₁) представляет только $\langle X \rangle \langle Y \rangle \langle E \rangle$ ($\langle \text{стимул} \rangle \langle \text{реакция} \rangle \langle \text{оценка} \rangle$), то иерархическая схема в НСС₂ позволяет дополнительно описывать различные **отношения R($\langle X \rangle \langle Y \rangle$)** физических характеристик активируемых N-элементов и объединять все это с оценкой $\langle E \rangle$.

$\langle \underline{X} \rangle \langle \underline{Y} \rangle \langle \underline{E} \rangle \rightarrow \langle \underline{(\langle X \rangle \langle Y \rangle)} \rangle \langle \underline{R(\langle X \rangle \langle Y \rangle)} \rangle \langle \underline{E} \rangle \rightarrow \langle \underline{\dots} \rangle$

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



$L = \{01100101011011011010001101101001\}.$

Исходный текст

1-й шаг $k_1=2$

0 01100101011011011010001101101001

в словаре

Последовательность индексов (ссылок)

1 1 2 1 1 1 2 3 1 2 2 4 3 1 2 2 1

Номера цепочек

1 2 3 4

+ "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 | |
| l2 | 1 2 1 3 4 5 1 6 | + "01"10" 11"00" L1 |



3-й шаг $k_3=2$

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



4-й шаг $k_4=2$

Номера цепочек в словарях

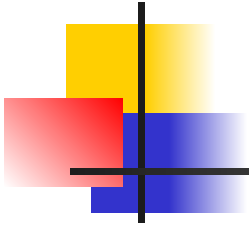
| | | | | | | | | | | |
|-----|------|----|----|----|----|------|------|------|------|----|
| B | 1234 | 1 | 2 | 3 | 4 | 5 | 6 | | | |
| | | 12 | 34 | | | | L4 | | | |
| | | 12 | 13 | 45 | 16 | | | L3 | | |
| | | 12 | 11 | 31 | 22 | 43 | 21 | L2 | | |
| A | 1 2 | | | | + | "01" | "10" | "11" | "00" |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



5-й шаг $k_5=2$

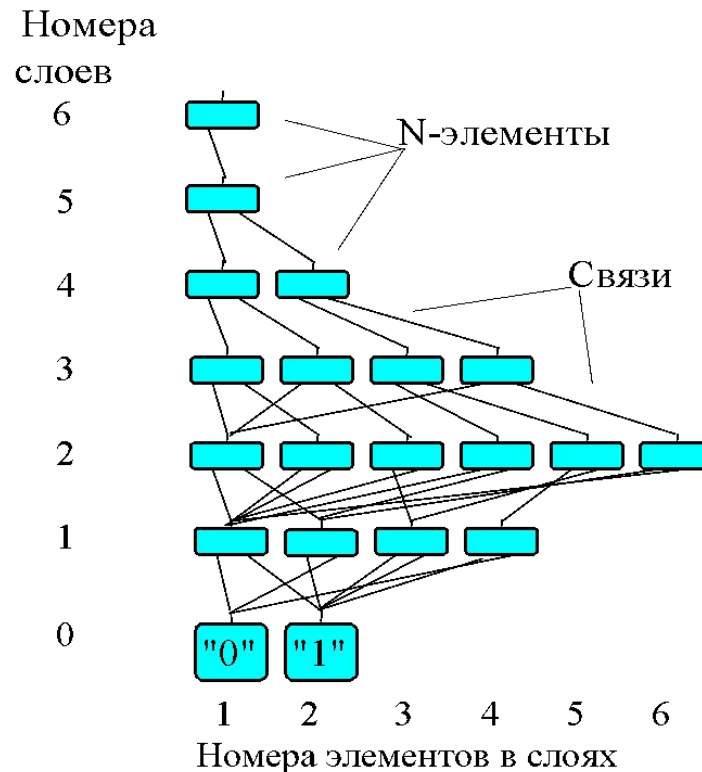
и исходная последовательность символов полностью переходит в НСС.

| Номера
словарей
(слоев) | Ссылки | |
|-------------------------------|----------------------------|---------------|
| 6 | 1 | на предыдущие |
| 5 | 12 | словари |
| 4 | 12 34 | и "алфавит" |
| 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 | | 6 | 6 | 1 | | Ссылки | | | |
| словарей | | 5 | 5 | 12 | | на предыдущие | | | |
| (слоев) | | 4 | 4 | 12 | 34 | словари | | | |
| Включение | | 3 | 3 | 12 | 13 | 45 | 16 | | и "алфавит" |
| в НСС | | 2 | 2 | 12 | 11 | 31 | 22 | 43 | 21 |
| алфавита | A | | 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".