



OSTIS-2012

(Open Semantic Technologies for Intelligent Systems)

УДК 004.81:82:032.26

АВТОМАТИЧЕСКОЕ ПОСТРОЕНИЕ СЕМАНТИЧЕСКИХ МОДЕЛЕЙ ПРОИЗВОЛЬНЫХ ПРЕДМЕТНЫХ ОБЛАСТЕЙ НА БАЗЕ НЕЙРОСЕМАНТИЧЕСКОГО ПОДХОДА

Бодякин В.И.

Институт проблем управления им. В.А. Трапезникова РАН, Москва
body@ipu.ru

Рассматривается актуальность комплекса проблем, связанных с обработкой сверхбольших объемов неструктурированной информации произвольных предметных областей и автоматического построения их семантических моделей. Анализируются компоненты: нейроподобный N-элемент, иерархически-сетевая структура из N-элементов, а также базовые свойства нейросемантического подхода – семантическая автоструктуризация, которая достигается при минимизации ресурсных затрат в процессе отображения предметной области на структуру N-элементов.

Ключевые слова: автоструктуризация, физический процесс, образ, информационно-управляющие системы, нейроподобные среды.

ВВЕДЕНИЕ

Актуальность комплекса проблем, связанных с обработкой сверхбольших объемов неструктурированной информации, постоянно нарастает. Современный этап развития научно-технической цивилизации характеризуется небывалым инновационным ускорением по всему спектру жизнедеятельности человека. Более половины открытий, научных теорий, технических и технологических новшеств из многотысячелетней истории цивилизации приходятся на последние ее тридцать лет. В течение одного поколения мы удвоили цивилизационный потенциал и, одновременно, человечество не располагает средствами, чтобы оценить все возможные последствия нахлынувших на него инноваций.

Надежды на успешное применение методов искусственного интеллекта (ИИ) и нейронных сетей (НС) для автоматизации построения крупномасштабных информационно-управляющих систем (ИУС) принципиальным успехом не увенчались. Проектируемые сегодня ИУС качественно не отличаются от ИУС 60-х годов. Различие выражается лишь в быстродействии и информационных объемах, что является следствием технического развития электронной элементной базы. Тогда как, без качественного рывка, современные ИУС «захлебнутся» от экспоненциально нарастающих информационных потоков. Необходима ревизия, а также анализ методов и подходов в ИИ и НС, становление которых произошло более полувека назад, на базе экспериментальных данных психологии и нейрофизиологии того времени [Глушков, 1981, Бир, 1993].

Исследование и анализ когнитивных функций чрезвычайно сложно, в силу большого морфологического и функционального многообразия «мозговых» структур, в которых, одних только синаптических связей более триллиона. Поэтому, предлагается от исследования отдельного нейрона и мозга перейти к «теоретической когнитологии», когда выдвигается и моделируется на компьютере некоторая модель когнитивных процессов, результаты которой сравниваются со свойствами прототипа. По такому пути уже давно идут теоретическая физика, химия, фармакология, биология (благодаря генетике) и другие бурно развивающиеся области знаний.

Наша группа, идя по тому же пути «теоретизации», пришла к выводу, что создание глобальной ИУС для управления социально-экономическими и технологическими процессами в обществе и на производстве, характеризующимися все нарастающей сложностью, возможно только на пути реализации механизмов ее самообучения, функционально не уступающих механизмам человеческой психики.

Для этого мы разработали принципиально новую формальную модель нейроподобного N-элемента, структуру нейросети (НС) из N-элементов и рекурсивную архитектуру ИУС из сформированных НСС.

Это направление мы назвали «нейросемантикой» [Бодякин, 2006]. Обсуждение со специалистами нейрофизиологами показало большую корреляцию современных нейрофизиологических данных, с характеристиками предлагаемой нами теоретической НСС-модели, по сравнению с характеристиками «классических» НС [Бодякин, 2006, 1998, 2009] [Грекул, 2008].

Кратко перечислим базовые утверждения нейросемантики:

1. Сложность и устойчивость используемых природой процессов при эволюционном конструировании homo sapiens укладывается в набор арифметических и графовых операций, т.е. «природа не мудрее четвероклассника».
2. При минимизации ресурсных затрат (памяти) на отображение в нейрореподобной специализированной среде (НСС) причинно-следственных процессов исследуемой предметной области (ПО) в этой нейрореподобной среде автоматически формируется словарь образов-процессов этой предметной области.
3. Архитектура ИУС состоит из трех НСС и объединяет в триплет образы: <ситуация ПО> <реакция ИУС> <оценка>, что позволяет отображать условные рефлексы и, посредством (само)обучения, формировать из них прагматическую модель поведения ИУС в произвольной предметной области.
4. Если рассматривать информационные процессы в самой НСС с помощью другой ортогональной (рекурсивной) ей НСС, как некую новую виртуальную предметную область, то оказывается, что в ней можно найти физические процессы, эквивалентные различным абстрактным понятиям, порождаемым человеческой психикой, что в результате позволит ИУС перейти к оперированию уже множествами (классами) образов.

1. Нейросемантика

Нейросемантический подход, это принципиально новый, эволюционный взгляд на (само)формирование и (само)функционирование ИУС. Прежде всего, это непосредственная («генетическая») связь ИУС со свойствами, изначально ее порождающей ПО. Это, отражение всех основных характеристик физических процессов ПО в ИУС, как предтечу ее когнитивных функций. Свойства ИУС вторичны по отношению к физическим процессам ПО, т.к. ИУС формируется в самой ПО на уже устоявшихся ее свойствах, см. рис. 1.

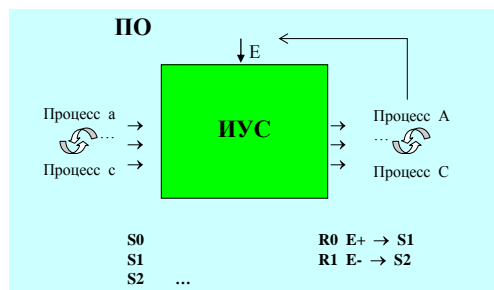


Рис. 1. Взаимоотношение ПО и ИУС

Отметим, что для физических процессов ПО характерны: детерминизм, целостность и причинно-следственность, иерархичность (фрактальность), энергетическая многомерность, параметрическая

количественно-качественная устойчивость, стремление к минимуму потенциальной энергии, функциональная эквивалентность аналоговой и дискретной форм представления и др.

Базовыми понятиями нейросемантического подхода являются:

- процесс;
- образ ИУС;
- взаимоотображение физического процесса и образа ИУС;
- текстовая форма (ТФ), как универсальный внутренний код ИУС и связующая форма взаимодействия ИУС и ПО.

Любой процесс рассматривается как некая целостность детерминированного изменения во времени вектора каких-либо физических параметров (характеристик). Процесс объединяет в себе как квази-устойчивую компоненту – параметры (данные), так и квази-динамическую компоненту – причинно-следственные детерминированные изменения параметров (функции). Такая модель процесса очень близка структуре объектно-ориентированных данных [Грекул, 2008].

Процесс также рассматривается нами как некая материальная субстанция, объединяющая в себе две разноплановые проекции: категорию вещество, обычно ассоциируемое с понятием – объект и категорию энергия, ассоциируемое с понятием – движение.

Объединение во вводимом понятии – «процесс», двух, воспринимаемых на сегодняшний день как разноразностных категорий, существенно упрощает концептуальную схему ИУС, т.к. пропадает необходимость в их обязательной «ручной» увязке для целостного функционирования в некой системе (напр., технической).

Любой общепринятый объект ПО, на более длительном временном интервале его наблюдения, всегда сводится к процессу. Например, придорожный камень на интервале миллионов лет, или человек от рождения и до его смерти – это все процессы, хотя каждый со своим темпом времени.

Все процессы переводимы в текстовую форму (ТФ), т.е. непрерывное, аналоговое изменение во времени компонент вектора физических параметров любого процесса может быть дискретизировано, или оцифровано («отекстовано») с любой требуемой степенью точности, как для отображения в форме образа, так и для последующего восстановления исходного физического процесса. Элементарные процессы отображаются знаками (символами) некоторого алфавита ТФ (например, A: <abazcw>).

Вообще, непрерывная и дискретная формы представления процесса (сигнала) принципиально различны. Непрерывная точна (до микроуровня квантовых процессов), но подвержена энтропии, т.е. деградации со временем. Напротив, дискретная форма сигнала точна лишь в на уровне шага дискретизации, но зато устойчива к действию энтропии в пределах этого же шага дискретизации.

Допустимость дискретизации физических процессов в ТФ следует из принципа

параметрической локальной однородности ПО (сведение континуума квантовых значений произвольного физического параметра к ограниченному числу качественно-различных образов. Например, вся температурная шкала от $0,03^\circ \text{K}$ до $10\,000,00^\circ \text{K}$ (10^6 значений), для любого вещества разбивается всего на четыре класса образов его состояния: «твердое», «жидкое», «газообразное», «плазма».

Из этого же принципа следует и возможность адаптивной подстройки в ИУС шага и равномерности дискретизации. В наиболее значимом для ИУС отрезке шкалы некоторого параметра происходит переход к более детальной дискретизации, тогда как на менее значимом интервале происходит загробление шага дискретизации. Тем самым происходит оптимизация начального представления ИУС информационного ресурса.

ТФ является универсальной формой для представимости разноуровневых процессов. Например, взаимодействия элементарных частиц – <ababsw> (микроуровень); процессов макроуровня – <«Привет, как дела?» – «Нормально!»>; и образования двойной звезды из двух астрообъектов – <zzjbaabj> (на гигауровне).

Таким образом, из вышеизложенного следует, что понятие процесс является универсально-базовым и самодостаточным в нейросемантическом подходе.

ТФ для ИУС является источником ее информационного ресурса. Компонентами информационного ресурса являются: сигнал, образ, информация и знание. Сигнал – произвольная цепочка знаков ТФ. Образ – определенная знаковая последовательность ТФ, тождественная целому числу причинно-связанных процессов.

1.1. Нейроподобный N-элемент

Базовым функциональным элементом нейроподобной специализированной среды (НСС) является конечный автомат с памятью – N-элемент. Нейроподобный N-элемент имеет – k упорядоченных входов и один условно-рефлекторный выход и несколько рекурсивных выходов, отражающих его внутреннее состояние. N-элемент осуществляет функцию идентификации собственного образа с сигналами в потоке из предметной области (ПО), по некоторому активированному функционалу сходства. В случае положительной идентификации собственного образа в потоке сигнала, N-элемент переходит в активное состояние, подавая на условно-рефлекторный выход $U(t)=1$, в противном случае открывается доступ к рекурсивным выходам, отражающим его внутреннее состояние. Собственный образ N-элемента определяется в момент его формирования из неидентифицированного ни одним из уже существующих N-элементов символического участка информационного потока из ПО. Таким образом,

данный механизм формирования слоя N-элементов позволяет отображать в НСС весь поток из ПО без пропусков (без потерь) и, в тоже время, избегать дублирования среди собственных образов N-элементов, что дает возможность строить ассоциативную память.

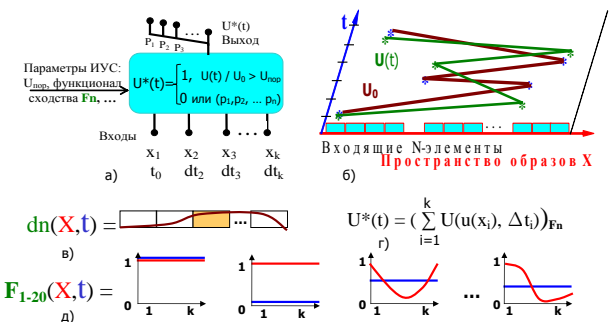


Рис. 1. а) Схематический вид N-элемента. Переход в активное состояние $U^*(t)=1$, при $F_i(U(t), U_0) > U_{оп}$, в остальных случаях доступны параметры внутренних состояний: P_1, P_2, \dots, P_n . б) Пространственно-временной образ N-элемента – U_0 и текущий входной сигнал – $U(t)$. в) Динамика проводимости входов в процессе обработки сигнала. г) Вычисление (свертка) $U^*(t)$ в процессе обработки сигнала. д) Пример набора функционалов сходства, отображающих максимальную проводимость с 1-го по k входы для X и t.

N-элемент принципиально отличается от общепринятого классического формального нейрона Маккалока и Питтса, разработанного ими в 1943 году [Бодякин, 2006]. Математический нейрон Маккалока и Питтса обычно представляют как некоторую пороговую или нелинейную функцию (функцией активации) от алгебраической суммы всех входных сигналов. Если суммы превышает некоторый параметрический порог, то на единственный выход нейрона подается «1», в остальных случаях – «0».

В N-элементе происходит постоянное сравнение его пространственно-временного образа, сформированного в момент его формирования в сети, с текущим входным пространственно-временным сигналом. Если в результате вычисления сходства по $U^*(t) = (\sum_{i=1}^k U(u(x_i), \Delta t_i)) F_n$, при текущем функционале F_n , $U^*(t) > U_{оп}$, то на выходе N-элемента будет «1», а в остальных случаях «0». Но при этом, в случае «0» становятся доступны текущие значения «микровыходов» (параметров внутреннего состояния N-элемента) – P_1, P_2, \dots, P_n , для других N-элементов специализированной рекурсивной «ортогональной» нейросети (НСС_{1...5}).

Таким образом, функционально N-элемент становится включенным в два информационно-вычислительных процесса нейросемантической ИУС. Если на выходе будет «1», то N-элемент включается в простейший условно-рефлекторный

контур ИУС. При этом, N-элементы ИУС обрабатывают только образы, отражающие процессы внешней среды. Т.е. в этом случае, вся семантика вычислительных процессов ИУС структурно разлагается на совокупность признаков («язык») процессов предметной области и, например, программисту при написании какой-либо программы достаточно использования данной совокупности признаков.

Если же на выходе не будет ожидаемой «1», то N-элемент включается уже в рекурсивный контур ИУС. При этом, N-элементы «ортогональной» нейросети (НСС_{1...5}) обрабатывают помимо образов, отражающих процессы внешней среды еще и образы внутренних состояний N-элементов (P_1, P_2, \dots, P_n) условно-рефлекторного контура, которые являются следствиями внутренних процессов в N-лементах при первичной обработке внешней среды. В этом случае, N-элементы обрабатывают в том числе и состояния самих N-элементов.

Такая своеобразная рекурсивная обработка N-элементов существенно (эмерджентно) расширяет пространство признаков в ИУС о процессах в предметной области. Так, ни в одной реальной (физической) предметной области не существует процесса (признаков) характеризующих абстрактные понятия, например, число «два». Хотя могут быть «два яблока», «два кубика», но «два» в физической природе нет. И ни один программист (без использования нейросемантического подхода) не может написать программу, которая бы узнавала два совершенно новых процесса (объекта). Для решения поставленной задачи программисту нужно знать характерные признаки этих, априорно неизвестных процессов. Кстати, высшие биологические ИУС, в частности, человек, легко справляются с решением этой задачи.

Легко с этой задачей справляется и нейросемантическая модель рекурсивной ИУС. Благодаря автоструктуризации, происходит взаимоотображение структуры физических процессов предметной области (внешней среды) и структуры нейросемантических образов в ИУС [Бодякин, 2011]. Каждому новому процессу в нейроподобной среде выделяется новый N-элемент, в том числе и для априорно неизвестного процесса (объекта).

Конструктивно N-элементы устроены так, что внутренний параметр (P_1) величины выходного сигнала после его активации ($U^*(t)=1$) постепенно спадает по экспоненте. Если этот же N-элемент активируется повторно, то величина P_1 будет уже больше «1» и например, параметр «температура» поднимется до $P_1=1.5$ [Бодякин, 2011]. Значение этого параметра и будет характерным признаком состояния N-элементов для случаев графических изображений: «два яблока» и «два кубика».

Если же теперь на вход нейросемантической ИУС поступит пара совершенно новых для нее графических объектов-процессов, например, «два

ёжика», то контур условно-рефлекторной модели ИУС здесь «промолчит» ($U^*(t)=0$). Не получив реакции, эта задача автоматически поднимется на рекурсивный уровень контур обработки в НСС_{1...5}, где будет найдена ассоциация, что у новообразованного N-элемента сигнал $P_1=1.5$, так же как и у тех N-элементов, которые получили положительную оценку, выдав реакцию «два» на парные графические образы-процессы, и эта их ассоциация активируется. Таким образом, в ИУС автоматически сформируется новый абстрактный класс – «два», который объединяет все парные графические образы-процессы с внутренним параметром N-элемента «температура» $P_1=1.5$ и теперь, любая пара совершенно новых процессов будет определяться ИУС как – «два».

Этот пример показывает, как отсутствие рефлекторной реакции на N-элементах ($U^*(t)=0$) у нейросемантической ИУС включает ее механизмы «подсознания», для поиска новых ассоциаций уже в расширенном внутренними параметрами N-элементов (P_1, P_2, \dots, P_n) пространстве признаков. В этом расширенном пространстве признаков на параметрах P_1, P_2, \dots, P_n появляется возможность выявления таких принципиально новых признаков, как различные «отношения» между реальными пространственно-временными процессами предметной области. Параметры P_1, P_2, \dots, P_n N-элементов полностью задаются конструктором или природой.

Из вышеизложенного можно выдвинуть гипотезу, что для любого когнитивного отношения всегда можно сконструировать такой физический процесс на N-элементах, который будет отражать требуемое отношение-абстракцию. Например, абстрактное понятие «число» может отражаться в виде параметра некой обобщенной «температуры» образа, вызванного неоднократной активацией его N-элемента.

Также возможно формирование и других базисных образов-абстракций для ИУС: больше-меньше; фон; симметрии; повороты на угол; обороты; верх-низ-лево-право; большой-маленький; легкий-тяжелый; образы естественного языка: прилагательные, наречия, глаголы, ... и т.д. В человеческой культуре известно всего несколько десятков тысяч образов-абстракций, что соответствует мощности естественных языков.

Таким образом, снимая рабочие процессы в N-элементах многочисленными датчиками и структурируя их с помощью НСС, мы можем объективно определять формируемые абстрактные понятия.

Абстрактные понятия-образы позволяют сгладить-заполнить признаковое пространство между конкретными примерами, сразу покрывая весь класс (множество элементов) данного отношения.

Из этого факта и модели нейросемантической рекурсивной ИУС следует гипотеза, что

нейроморфология мозга высших животных должна эволюционно изменяться относительно нарастанием числа межнейронных контактов типа: «синапс – тело нейрона», к «синапс – аксон».

1.2. Нейросемантическая структура

Еще одна отличительная черта нейросемантического подхода заключается в том, что если в классических нейронных сетях изначально устанавливают связи между нейронами («каждый с каждым») со случайными весами и затем в процессе обучения (тренинга) подстраивают веса этих связей, таким образом, чтобы вся нейросеть на входной вектор сигнала выдавала требуемый вектор реакции, то в нейросемантической ИУС процессы запоминания входного информационного потока и обучения объединены в единый процесс.

Все N-элементы объединены в слои и образуют многодольный иерархически-сетевой граф, который будем называть – нейросемантической структурой (НСС). Вершины графа (N-элементы) представляют образы отображаемых физических процессов текущей задачи и предметной области, а порядок входящих в вершину дуг отражает причинную связанность этих процессов. Межслойные связи в НСС формируются в режиме реального времени восходящей активностью ($U(t)=1$) N-элементов, вызванной обработкой ими символического потока из предметной области. Число слоев сформированной НСС определяется как $\log_k(\text{ТФ})$ от объема обрабатываемого потока. Размер каждого слоя НСС определяется разнообразием образов данного иерархического уровня, отображающих процессы в предметной области, см. рис. 3.

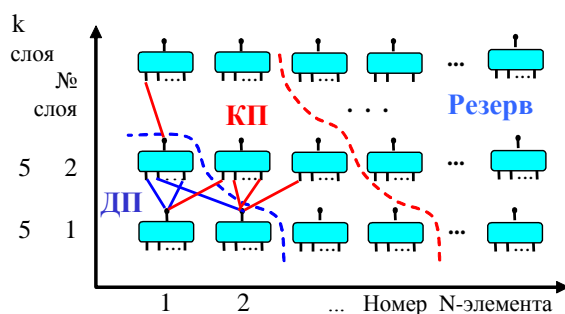


Рис. 3. Динамика роста нейросемантической структуры.

Процесс заполнения связями (памяти) НСС идет слева направо и снизу вверх (долговременная память – КП, промежуточная память – КП и резерв). В слое N-элементы вычислительно независимы, следовательно, возможна максимальная распараллеливаемость их обработки на ЭВМ.

Автоструктуризация (взаимоотображение физического процесса и образа в ИУС). Одним из значимых свойств НСС, является возможность автоматического формирования в ней графа, гомоморфного причинно-следственной структуре отображаемых физических процессов задачи для любой предметной области. Для этого необходимо

лишь минимизировать суммарные ресурсные затраты памяти, расходуемые на формирование вершин (N-элементов) и дуг (связей) в НСС при отображении в ней неограниченного символического информационного потока ТФ из произвольной ПО [Бодякин, 2006].

$$\text{Если} \quad \lim_{t \rightarrow \infty} \frac{\text{Объем НСС (в битах)}}{\text{Объем ТФ (в битах)}} \rightarrow 0, \quad (1)$$

$$\text{то НСС будет гомоморфна} \quad (2)$$

структуре процессов ПО и будет представлять собой информационную (семантическую) модель данной ПО. Семантическая модель ПО – это совокупность ее причинно-связанных процессов.

Можно предложить множество несложных вычислительных процедур по минимизации ресурсных затрат НСС на базе вариационных методов, например, метода «частотной конкуренции образов словаря».

Рассмотрим данный метод на упрощенном примере. Допустим, что ПО состоит из трех «объектов» – трех шаров: «чугунного», «алюминиевого» и «деревянного». При попарных и тройном соударениях они порождают отекстованные образы четырех (n) процессов: «МАК», «ЛЕН», «РИС», «ДУБ», которые с равной вероятностью ($1/n$) без перекрытий, пауз и каких-либо разделителей поступают в поток ТФ, см. рис. 4. Задача ИУС заключается в том чтобы, анализируя только данный текстовый поток, построить информационную модель ПО. Для начала ИУС следует построить словарь «образов-процессов», в котором для упрощения демонстрации предположим, что все формируемые образы словаря имеют длину, соответствующую длинам процессов.

Метод «частотной конкуренции образов» базируется на частотном различии активации образов словаря принадлежащих двум множествам: «образов-процессов» и «образов-кентавров». «Образы-процессы» состоят из последовательности символов одного причинно-связанного процесса ПО, тогда как «образы-кентавры» состоят из двух причинно-связанных процессов: отображающих последовательность символов окончания одного процесса и последовательность символов начала следующего процесса, см. рис. 4.



$$\begin{aligned} \text{Образ-процесс } p_p &= 1/n. & \omega_p &= 1/p_p. \\ \text{Образ-кентавр } (1/n * 1/n) & p_k = 1/n^2 & \omega_k &= 1/p_k. & \omega_p/\omega_k &= n \\ \text{Число образов-кентавров / число образов-процессов} & & & & &= (L-1)*n \end{aligned}$$

Рис. 4. Образы-процессы и образы-кентавры при формировании словаря.

Из рисунка 4 легко понять, что образ-процесс имеет частоту активации (w_p), соответствующую частотам процессов ($1/n$) на интервале ТФ в 12 символов, т.е. $w_p = 1/n$. Соответственно, так как образ-кентавр построен из двух процессов ПО, то его частота активации (w_k) на интервале ТФ в 12 символов, определяется произведением вероятностей составляющих его процессов ($1/n * 1/n$), т.е. в нашем случае, $w_k = 1/n^2$.

Этот частотный признак и служит для предварительной маркировки формируемых N-элементов слоя как образов словаря в промежуточной памяти ИУС, см. рис.3. Большая частота активации N-элемента соответствует большей вероятности, что его образ принадлежит к классу образов-процессов. В итерационном алгоритме формирования словаря-слоя, N-элементам с большей частотой активации дается приоритетное право закрашивать ТФ и они постепенно вытесняют N-элементы с образами-кентаврами из формируемого словаря.

В соответствии (1), еще одной индикативной характеристикой при формировании информационной модели, гомоморфной причинно-следственной структуре процессов ПО, является размер формируемого словаря. Слои-словарь НСС, построенный только из образов-кентавров будет, как минимум, в n-раз больше слоя-словаря образов-процессов (3). Что при больших n алгоритмически (программно) легко определяется и служит управляющим фактором для корректного расширения словаря новыми образами.

$$\frac{\text{Размеры случайных словарей}}{\text{Размеры минимальных словарей}} > n \quad (3)$$

Таким образом, только при совпадении начал фаз образов словаря НСС и процессов ПО (см. рис. 4) и соответствии длин образов словаря и длин процессов получается минимальный словарь образов-процессов, который полностью соответствует исходным причинно-следственным процессам ПО и мы автоматически получаем информационную модель произвольной, априорно неизвестной ПО.

Формирование информационных моделей иерархических ПО. Очевидна справедливость механизма минимизации при формировании информационной модели ИИС и для иерархически сложных процессов ПО, см. рис. 5. В этом случае, сначала формируется словарь образов слов («Коля», «Саша», «ловит», ...), по методам рассмотренным выше. Затем, вместо образов словаря («Коля», «Саша», «ловит», ...) в ТФ ставятся их коды (номера) в словаре 1-го уровня НСС. В результате, мы получаем ТФ индексов и сводим данную задачу к предыдущей.

Проводим процедуру минимизации словаря уже 2-го уровня с кодами-индексами образов 1-го уровня словаря («134256 ...») и получаем двухуровневую структуру словаря («134», «256», ...), топологически изоморфную структуре

исходных процессов ПО. На втором уровне словаря будут сформированы образы псевдо-фраз, эквивалентные фразам-процессам ПО, см. пример на рис. 5.

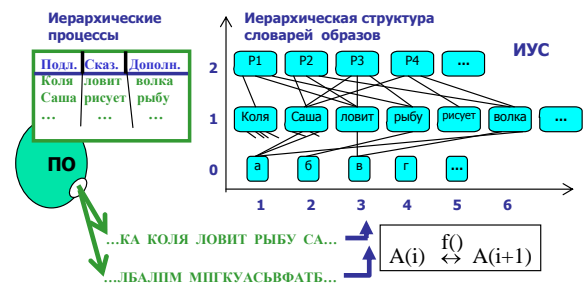


Рис. 5. Формирование иерархических словарей НСС.

На рис. 5 показано, что для взаимоднозначных алфавитов различных ТФ, порожденных одним процессом, информационные модели получаются топологически тождественными. Из этого факта следует, что у различных ИУС может быть сформирована объективная картина мира, независимо от алфавита и даже языка которые они используют.

Процесс формирования структуры словарей ИУС, гомоморфной структуре процессов ПО, мы назвали – автоструктуризацией. Механизмов автоструктуризации достаточно много, от простейших вариационных методов до рассмотренного нами частотного метода [Бодякин, 2006, 2008, 2009]. Саму процедуру мы назвали – нейросемантической потому как благодаря автоструктуризации в каждом отдельном нейроразнообразном N-элементе формируется («стягивается») образ семантики ПО.

1.3. Автоструктуризация

Явление автоструктуризации, как взаимоотношение структуры физических процессов ПО и структуры образов в НСС, первоначально было выявлено при моделировании на ЭВМ. Затем теоретически доказано, что для автоматического формирования в НСС графа, гомоморфного причинно-следственной структуре отображаемых физических процессов задачи любой предметной области, необходимо лишь минимизировать суммарные ресурсные затраты памяти, расходуемые на формирование вершин (N-элементов) и дуг (связей) в НСС, при отображении в ней символического информационного потока из произвольной ПО, см. рис. 4 и рис. 5.

1.4. Эволюционный потенциал ИУС

Известно, что для всех живых организмов энерго-вещественная и прогностическая составляющая их поведения, являются определяющими в характеристики их жизнеспособности. Отсутствие или недостаточность энергии, через некоторое время, приводит к гибели организма. Также практически нежизнеспособен биологический организм, если он

не может предвидеть развитие событий в окружающей его среде. Т.е. энергетическая и информационная составляющие, в своей совокупности, характеризуют потенциальную жизнеспособность организма. Объединение энергетической и информационной составляющих в некоторую комплексную характеристику назовем эволюционным потенциалом. Материальной компонентой эволюционного потенциала является энергия и вещество (Э), измеряемые в джоулях и нематериальная, т.е. информационная компонента (И), измеряемая в битах или образах, см. рис 6.

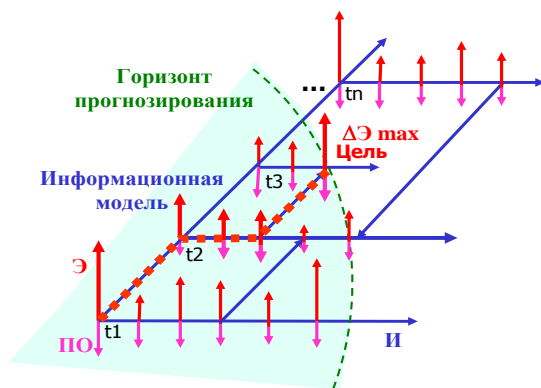


Рис. 6. Задача ИУС – максимизация материальной компоненты (Э) эволюционного потенциала до горизонта прогнозирования (И) информационной модели.

Если первая компонента эволюционного потенциала – это свободная энергия (Э), как необходимый ресурс для преобразования окружающей среды, то вторая компонента – нематериальная (И), содержит в себе информацию или знания о том, как в окружающей среде можно добыть и использовать эту свободную энергию, см. рис. 6.

Эволюционный потенциал характеризует жизнеспособность ИУС, как биологических организмов, так и технических систем. Если в контекстных условиях среды происходит столкновение интересов двух и более ИУС, то победителем выходит ИУС, имеющая больший эволюционный потенциал. Накопление эволюционного потенциала является целевой функций всех ИУС. Чем больше эволюционный потенциал, тем лучше ИУС может противостоять разрушительным действиям энтропии.

Но эффективность использования этой свободной энергии определяется когнитивными характеристиками ИУС. Соответственно, формально вычисляемая величина эволюционного потенциала ИУС, позволяет делать прогнозы относительно ее будущего.

1.5. Условно-рефлекторная и рекурсивная модели ИУС.

ИУС «генетически» настроена на стремление к накоплению свободной энергии среды, как необходимого ресурса для противодействия разрушающим действиям физической энтропии.

Соответственно, ИУС из всевозможных реакций на текущую ситуацию должна выбрать ту, которая дает ей наибольшее приращение свободной энергии. При варианте «с учителем», целевой функцией ИУС будет стремление к получению положительных оценок «от учителя» за выбор реакций на текущие ситуации ПО. Из такой постановки вытекает условно-рефлекторная модель ИУС.

Если расширить схему условно-рефлекторной ИУС, добавлением в нее внутреннего канала обратной связи, переводящего образы реакции системы на ее вход (своеобразный аналог мышления) и введением дополнительной структуры из иерархических НСС, то мы получим схему рекурсивной ИУС, которая будет способна автоматически формировать абстрактные понятия и алгоритмы-эвристики для обработки процессов ПО, т.е. самообучаться, см. рис 7.

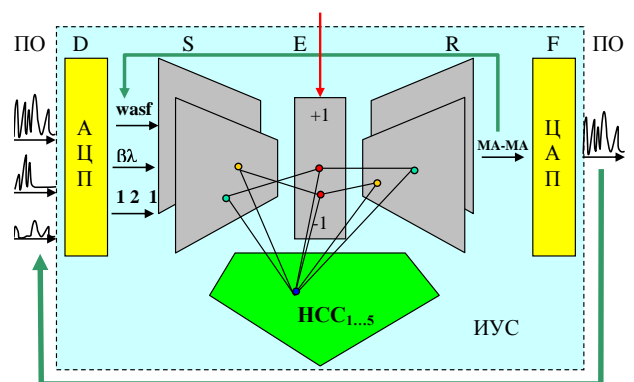


Рис. 7 Архитектура «рекурсивной» ИУС состоит из шести типов блоков: детекторно-рецепторного (D), распознающего (S), двигательного (R), прагматического (E), эффекторного (F) и рекурсивного (HCC1...5).

ЗАКЛЮЧЕНИЕ

Из тезисного изложения материала следует, что на базе нейросемантического подхода можно построить ИУС для работы с неструктурированными и априорно неизвестными информационными потоками. При этом, нейросемантический подход позволяет решить основную проблему современного программирования – автоматизацию формирования информационной (семантической) модели из априорно неизвестных информационных потоков, порожденных из процессов исследуемой предметной области, а также, формирование эвристических правил ИУС не программированием, а посредством ее самообучения. Что позволяет полностью автоматизировать весь цикл создания крупномасштабных ИУС.

Анализ нейросемантического подхода показал большую корреляцию по различным характеристикам (работа с пространственно-временным сигналом, многопараметрическим выходом нейрона, «прорастиваемостью»

межнейронных связей в режиме реального времени, иерархичностью нейросетевых структур и др.) с современными нейрофизиологическими данными.

Интересны рассмотренные, вытекающие из нейросемантики естественно-научные определения компонент информационного ресурса:

- Сигнал – произвольный отрезок оцифрованного процесса (отражает характеристику свойств потока текстовой формы);
- Образ – вычлененная информационной системой знаковая последовательность текстовой формы, тождественная целому числу причинно-связанных процессов ПО (семантическая единица ПО);
- Информация – прагматически причинно-связанная последовательность образов, ориентированная на достижение ИУС какого-либо ее целевого образа ПО (прагматическая единица ИУС);
- Знание – причинно-связанное объединение информации о ПО с информацией о состоянии ИУС (семантические закономерности).

А так же, определения степени когнитивности различных классов ИУС:

- Автоматная – с жестко «запаянным» алгоритмом (характерно «зацикливание» при ошибках и необучаемость);
- Адаптивная – возможность самообучения по управлению любым априорно неизвестным объектом за счет перебора всего пространства его состояний (аналог универсальной адаптируемости живого);
- Интеллектуальная – свойство практического преодоления «проклятия размерности» в «гладких» пространствах состояний управляемых объектов за счет набора эвристик;
- Разумная – возможность самоформирования эвристик (отношений).

Анализ приведенных определений указывает на фундаментальность нейросемантического подхода.

Данная работа выполнена при частичной поддержке РФФИ, проект № 11-07-00007-а.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- [Бодякин, 2006] Бодякин В.И. Определение понятия «информация» с позиций нейросемантики. М.: ИПУ РАН. 2006, 48 с. http://www.mtas.ru/search_results.php?short_view=0&publication_id=3033
- [Бодякин, 1998] Бодякин В.И. Куда идешь, человек? (Основы эволютологии. Информационный подход). - М. СИНТЕГ, 1998, 332с. <http://www.ipu.ru/stran/bod/monograf.htm>
- [Бодякин, 2009] Бодякин В.И. Разработка инструментария информационно-управляющих систем на базе нейросемантического подхода для построения крупномасштабных информационно-управляющих систем // XI Всероссийская научно-техническая конференция "Нейроинформатика-2009": М.: МИФИ, 2009. Ч.2. С. 27-38,

[Грекул, 2008] Грекул В.И., Денищенко Г.Н., Коровкина Н.Л. Проектирование информационных систем. Интернет-университет информационных технологий ИНТУИТ.ру, 2008 г., 304 стр.

[Бодякин, 2011] Бодякин В.И. «Механизм автоматического формирования информационной модели в информационно-управляющей системе, построенной на базе нейросемантической парадигмы» // Вторая Всероссийская конференция "Нелинейная динамика в когнитивных исследованиях", г. Нижний Новгород, 2011 г. 16-19с. - 260 с.

[Бодякин, 2011] Бодякин В.И. «Анализ формирования когнитивных функций, построенной на базе нейросемантической парадигмы» // Вторая Всероссийская конференция "Нелинейная динамика в когнитивных исследованиях", г. Нижний Новгород, 2011 г. 20-23с. - 260 с.

[Глушков, 1981] Глушков В.М., Валах В.Я. Что такое ОГАС? – М.: Наука, 1981 – 160 с

[Бир, 1993] Бир С. Мозг фирмы: Пер. с англ. – М.: Радио и связь, 1993. 600с

AUTOMATIC CONSTRUCTION OF SEMANTIC MODELS OF ARBITRARY SUBJECT DOMAINS ON BASE OF NEURONSEMANTIC APPROACH

V.I. Bodyakin

V.A. Trapeznokov Institute of Control Sciences, 65 Profsoyuznaya, Moscow 117997, Russia,

body@ipu.ru

The method of automatic exposure of semantic component is in-process investigated from the continuous text stream of the real subject domain. The basic idea of method consists in minimization of reflection of text stream on the specialized neuronsimilar environments. In theory it is well-proven and confirmed by the results of computer experiments on a hypothetical subject domain, that during minimization of reflection an informative model is automatically formed in a neuronsemantic environment, homomorphous to the arbitrary subject domain.

Keywords: neurosimilar structure, neurosemantic, intellectual systems.

CONCLUSION

The computer construction of dictionaries of minimum form showed that both for simple subject domains and for the hierarchical processes of arbitrary complication property of homomorphism is executed. Thus, the examined mechanism opens practical possibility of the automatic forming of informative (semantic) models of arbitrary subject domains, that will allow to give a new impulse to works on self-training.