

ОПРЕДЕЛЕНИЕ ПОНЯТИЯ "ИНФОРМАЦИЯ" С ПОЗИЦИИ НЕЙРОСЕМАНТИКИ

Бодякин В.И.

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

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

Аннотация

В работе предлагается комплексное рассмотрение процесса взаимодействия системы и окружающей ее среды. Обосновываются базовые свойства среды для возможности естественной эволюции саморазвивающихся систем. Дается формально-объективное определение понятия "информации", как одной из фундаментальных проекций материи, равновеликой категориям: энергия, пространство и время.

Abstract

In work the process of interaction of system and environment is in complex considered. The properties of environment for an opportunity of natural evolution of developing systems are proved. The formal objective definition of concept of "information" is given, as one as of fundamental projections of a matter equal to categories: energy, space and time.

Москва 2006

Введение. Слово "*информация*", используемое как самостоятельно, так и в виде однокоренных производных, в широком понятийном ряду, сегодня, наверно, самое распространенное слово по широте охвата всех аспектов нашей жизни: науки, производства, культуры, быта. И влияние этого понятия на нас со временем только нарастает. На запрос "*информация*" в Интернете выдается самый широкий индексный список, включающий около полумиллиарда документов (Google, Yahoo!, Rambler, и др.).

Понятие "*информация*" известно с древнейших времен. Каждый коммуникативный акт между субъектами рассматривается как обмен информацией. Однако с середины XX века в результате социального прогресса и бурного развития науки и техники роль информации как основного цивилизационного ресурса неизмеримо возросла.

Начало этому было положено работой К.Шеннона "*Математическая теория связи*", изданной в 1946 г. Понятие "*информация*" стали рассматривать с позиции вероятности. Математическую теорию связи К.Шеннона положили в основание нового научного направления "*теории информации*". Затем из первоначально чисто статистического направления теории информации стали пытаться его расширять рассмотрением количества и качества информации в живых организмах. Появились исследования: топологического (Н.Рашевский и У.Р.Эшби 1960), алгоритмического (А.Н.Колмогоров 1965) направлений, а также исследовалась проблема измерения ценности информации (А.А.Харкевич, 1960) и содержания ее в семантическом подходе. Области приложения информационных подходов постоянно расширяются их использованием в лингвистике, биологии, психологии, социологии, педагогике и других предметных областях. Лавинообразное нарастание использования этого понятия в различных областях знаний и частота употреблений, получило название "*информационного взрыва*" [1,5,14].

Современное понятие "*информации*" было расширено и включило обмен сведениями не только между человеком и человеком, но также между человеком и автоматом, автоматом и автоматом; обмен сигналами в животном и растительном мире. Передачу признаков от клетки к клетке и от организма к организму также стали рассматривать как передачу информации.

В информатике известны различные подходы к определению понятия "*информация*", и каждый подход вносит свои акценты в это определение. В результате, на практике понятие "*информация*" очень широко распространилось, и такое словоупотребление закрепляется нормами естественного языка. Сегодня под "*информацией*" источники различных областей знаний и направлений называют:

- *любые сведения о каких-либо ранее неизвестных объектах;*
 - *содержательное описание объекта или явления;*
 - *результат выбора;*
 - *содержание сигнала, сообщения;*
 - *меру разнообразия, отраженное разнообразие;*
 - *сущность, сохраняющуюся при вычислительном изоморфизме;*
 - *уменьшаемую неопределенность;*
- (1)

- меру сложности структур, меру организации;
- результат отражения реальности в сознании человека;
- семантику и прагматику языка представления данных;
- атрибут материи и многое, многое другое

Естественно, что каждый из перечисленных вариантов определения информации в некоторой степени неполон.

С позиции различных философских концепций, которые занимаются изучением онтологического понятия информации как научного термина, определение "информации" еще более неопределенно. На одном полюсе утверждается, что весь мир и все его свойства созданы из информации как одного из основных субстанциональных свойств материи. Наиболее известные сторонники этого взгляда – английский нейрофизиолог У.Р.Эшби и французский физик Л.Бриллюэн. На другом, – вообще отрицается существование информации как действительности, и ее относят лишь к субъективной реальности, как знания, сведения, которые формируются в сознании человека. При этом утверждается, что информацию невозможно определить исчерпывающим образом через какие-то более простые понятия в силу их отсутствия, а ее значение может быть пояснено только на примерах или выявлено путем сопоставления с содержанием других понятий. Между этими крайностями существуют еще ряд промежуточных точек зрения: признание информации третьим атрибутом материи, равнозначным веществу и энергии; утверждение о независимости информации от какого бы то ни было материального носителя; трактовка ее как тонкоматериальной структуры с необычными для физического мира свойствами; признание в одном случае первичности информации и вторичности материи, а в другом – материальности информации и т.д.

Все многообразие окружающих нас терминов "информации" можно сгруппировать, например, в зависимости от области их возникновения. Так информацию, отражающую процессы и явления неодушевленной природы, называют элементарной, процессы животного и растительного мира – биологической, человеческого общества – социальной информацией.

В технологическом аспекте "информацию" также можно рассматривать по способу передачи и восприятия как: визуальную, передаваемую видимыми образами и символами; аудиальную – звуками; тактильную – ощущениями; органолентическую – запахами и вкусом; машинную, выдаваемую и воспринимаемую средствами вычислительной техники, и т.д.

В философском энциклопедическом словаре дается определение: *"Информация" (от лат. informatio - разъяснение, изложение), первоначально означало - сведения, передаваемые одними людьми другим людям устным, письменным или каким-либо другим способом (например, с помощью условных сигналов, с использованием технических средств и т.д.), а также сам процесс передачи или получения этих сведений. Однако в середине XX века, в результате бурного развития науки и техники, роль информации неизмеримо возросла. В связи с*

этим возникла потребность в научном подходе к информации, выявлении её наиболее характерных свойств, что привело к двум принципиальным изменениям в трактовке понятия информация. Во-первых, оно было расширено и включило обмен сведениями не только между человеком и человеком, но также между человеком и автоматом, автоматом и автоматом; обмен сигналами в животном и растительном мире. Во-вторых, была предложена количественная мера информации (работы К. Шеннона, А.Н. Колмогорова и др.), что привело к созданию "Теории информации". Сегодня развитие и исследование проблем, связанных с научным понятием информации, продолжается в следующих основных направлениях: - разработке математического аппарата применительно к кибернетике; - исследовании различных "ценностных" свойств информации; - использовании в лингвистике, биологии, психологии, социологии и др. [2,3]

Но что же такое в данном определении есть сообщение, или сведения? Эти понятия в свою очередь требуют определения. Такой подход, когда одно слово объясняется через другое, синонимичное, можно характеризовать всего лишь как обыденное толкование, а не научное дефинирование.

До сих пор не существует строгого научного определения "информации". Не очерчена и область применения этого понятия. Часто информацию рассматривают как форму отражения, но является ли она свойством всех материальных объектов, или только живых и самоуправляющихся, или присуща только сознательным существам, единого мнения также нет.

За последние десятилетия понятие "информация" и однокорневые его производные еще более расширили сферу употребления и стали применяться при описании таких социальных и экономических явлений, как: "экономическая информация", "информационный бизнес", "информационное право", "защита информации" и пр. Это говорит о том, что понятие "информация" сделалось не только важнейшей категорией современной науки, но и важнейшей характеристикой современной цивилизации [6].

В широком смысле "информация" – это общенаучная категория. В связи с этим возникла потребность в научном подходе к выявлению наиболее характерных свойств этой категории и определяющих ее о понятиях, а также исследования проблем, связанных с научным пониманием этого термина при разработке математического аппарата, отражающего основные количественные и качественные (семантические) аспекты этого понятия.

Научное исследование данной категории должно коснуться философских и методологических вопросов, связанных с соотношениями уже широко используемых понятий: "отражение" (сигнал) – "информация" – "знание" – "сознание" в процессе сбора, хранения и переработки информации, и их ролью, применительно к саморазвивающимся системам различной природы.

В эпоху информационных технологий ответы на эти вопросы становятся ещё более актуальными. И "как-то неудобно получается", активно оперировать понятием "информация", называть нашу эпоху веком информатики и не иметь ни четкого определения, ни точного описания основных ее свойств.

Так если понятие "бит" четко и ясно осознаваемо инженером в деталях ЭВМ, то понятие "информация", или "семантика" – уже не имеет понятий, через которые их можно определить. Поэтому и существуют весьма разработанные калькуляторы-компьютеры, а интеллектуальных систем – нет.

"Информация" – это такая же фундаментальная категория, как и энергия, которую до недавнего времени (до XVII века) также нельзя было эффективно использовать без ее теоретического осмысления в физике и, как следствие, в те времена отсутствовало ее однозначное определение для всех областей приложения.

Информатика еще совсем молодая фундаментальная область знания (ей менее полувека), и ей перед формированием стройного собственного научного здания, как и всем уже установившимся классическим естественнонаучным теориям, необходимо пройти через предварительный этап систематизации [6].

В основу предлагаемого нами подхода по формализации основных понятий информатики мы положили рассмотрение процесса взаимодействия "произвольной предметной области и биологической информационной системы (животного или человека)". В результате у нас сформировалась модель исследовательского лабораторного комплекса: *"предметная область (ПО) – информационно-управляющая система (ИУС)*, все компоненты которого мы попытаемся формализовать и выразить в физических единицах измерения: сантиметр (пространственное кодирование: номер-имя), бит (аналог грамма в Системе единиц СГС), секунда (порядок следования), – сокращенно *система СБС*.

В качестве отправного рабочего определения под "информацией" мы будем понимать:

обмен сигналами (процессами) между материальными образованиями (системами). (2а)

Всех, кто хотел бы более подробно проследить становление категории "информации" мы отсылаем к прекрасной монографии А.Д.Урсула [1], в которой на большом материале подробно рассмотрены все сложности этой проблемы. Наша же задача будет состоять в формулировании самого общего подхода к (корневому) определению понятия "информация" с позиции нашего нейросемантического подхода, основанного, в свою очередь, на достижениях нейрофизиологии и лингвистики [7,15,17,18].

Дать определение термину – это значит, подвести его под более общие понятия. В нашем случае это будут простые физические и математические понятия, измеряемые в сантиметрах, битах и секундах (система СБС). Так что трактовки определения понятия "информация" и родственных ему: "сигнал", "образ", "знание", "данные", "информационный ресурс" и др. будут иметь однозначное трактование (в СБС), и на них будут строиться определение аналогичное (2) и формулировки (1).

В понятие "*информационный ресурс*" мы будем вкладывать обобщение множества таких понятий как: "сигнал", "образ", "информация", "знание", "данные" и др. Под основной чертой "информационного ресурса" мы будем понимать нематериальную компоненту материи, т.е. ту, которая не является отдельно ни веществом, ни энергией, ни пространством, ни временем, но может заключаться в их *особом структурном взаимодействии*.

Часть I

Психофизическая ограниченность человека. Для начала мы попытаемся найти причину множественности столь неоднозначных определений понятия "информация" (1).

Как известно, *психофизические возможности человека ограничены*. Нейрофизиологи дают следующие оценки усредненным (техническим) характеристикам человека, как информационно-управляющей системы (ИУС). Так, *объем памяти среднего человека* оценивается в 10^6 образов.

Образ – это некоторая информационная сущность (фрагмент информации, знания), хранимая и используемая как единое целое. Образы имеют иерархическую природу. Сложные образы состоят из более простых образов. Самые простые – терминальные, находятся на границе ИУС с предметной областью.

Например, любое слово данной статьи – образ. Любая фраза из слов, – также образ. Слоги, из которых состоит слово – образы. Образы слогов, в свою очередь, состоят из терминальных образов – символов кириллического алфавита. Все терминальные образы атомарны, т.е. неделимы. Все сложные образы, с одной стороны, представляют некоторое цельное понятие (образ), которым можно оперировать как неделимым объектом, с другой стороны, тоже состоят из образов. Образами также могут быть как элементы какой-либо схемы, так и сама схема, акустические сигналы и целые симфонии. Все, что в понятийном аппарате человека имеет имя, – это есть пример образа в нашем понимании. Далее этому понятию, будет дано более формальное определение в системе СБС.

Другое психофизиологическое ограничение человека – это *быстродействие биологических нейронов* (как аналога электронной элементной базы компьютера), которое в нормальном функционировании центральной нервной системы человека оценивается в 10-30 Гц.

Скорость ввода-вывода осознаваемой человеком информации, оценивается в 5-10 образов/сек, или 10-20 байт/сек.

Число контролируемых сознанием человека параллельных и независимых каналов (зрительный, слуховой, тактильный и др.) одновременного ввода-вывода информации оценивается в 1-5.

Можно отметить, что количественные характеристики человека (как ИУС) по переработке информации достаточно скромные по сравнению с техническими характеристиками современной вычислительной техники. С другой стороны, человеческая цивилизация за десятки тысяч лет своего существования накопила поистине огромные объемы информации. Это и тексты, и графические изображения, и музыкальные произведения, различные телеметрические и экспериментальные данные, кардиограммы, звездные каталоги и т.д., и т.д. Вся эта информация, хранящаяся в различных библиотеках и архивах, а также в базах данных и файловых системах, оценивается в 10^{20} - 10^{24} байт или в 10^9 - 10^{12} образов [7].

Такой огромный объем накопленной информации не может охватить своим сознанием ни один человек. Человеку доступна лишь проекция реального мира, которая сегодня составляет уже менее десятой доли процента от необходимого знания о предмете его профессии.

Плеяда великих ученых XVII-XVIII веков, таких как: И.Ньютон, Г.В.Лейбниц и др., были, видимо, одними из последних ученых-энциклопедистов, которые могли своим сознанием охватить весь уровень научных данных того времени.

С тех пор, объемы порождаемой информации только нарастают, и затем началась все ускоряющаяся специализация по отдельным направлениям знаний. И чем дальше, тем все меньше и меньше специалист в одной области знания уже понимал узкоспециализированного специалиста из другой области знания. Изобретение Гуттенбергом печатного станка, появление ЭВМ и особенно персонального компьютера, только усугубило состояние дел с перенасыщенностью мира слабоструктурированной информацией. Информация становится все более избыточной, противоречивой, необозримой, причем, с каждой минутой эти проблемы только нарастают. И как следствие резко снижается эффективность работы всего научного сообщества, что, в свою очередь, переносит эти проблемы в современные информационные технологии, на которые надеются как на локомотив мирового экономического и технологического развития [6,12].

Таким образом, мы имеем, с одной стороны, и так уже запредельный, но еще и ежеминутно нарастающий информационный вал и, с другой стороны, человека, хотя и способного разобраться в этом информационном сырье ("втосырье"), но, к сожалению, психофизические возможности человека, как ИУС, *чрезвычайно ограничены*, по отношению к уже существующим "информационным эверестам". На сегодняшний день *не существует* достойных информационно-коммуникационных технологий, которые бы могли справиться с такими объемами и такой скоростью пополнения информационных ресурсов нашей цивилизации. Обещанного и долгожданного ИИ (в расширенном понимании) также до сих пор не создано.

Поэтому исследователь и вынужден довольствоваться лишь информационной проекцией мира, прикладывая ее на свою узкоспециализированную область знания. Но, как известно, любая проекция какого-либо явления всегда неполна относительно его самого. А выводы из различных проекций одного

и того же явления могут быть даже и противоречивы. Как в известной притче "о трех слепцах и слоне". Каждый из них, воспринимая только свою часть, утверждал, что: "слон – это веревка", "слон – это пузырь", "нет, слон – это большой столб". Хотя каждый объективно прав, но не относительно самого объекта, а только применительно к его проекции. Аналогичная ситуация сложилась и с понятием "информация", см. рис. 1.



Рис. 1.

Заметим, что хотя множество проекций одного объекта-явления интегрированное в некоторую целостность и может приближаться к адекватному пониманию его самого (например, из трех ортогональных проекций детали, можно получить ее полное представление как геометрического объекта), но, как мы уже отмечали, инструментария для синтеза существующих информационных объемов при их сверхчеловеческой многомерности, в единую мировоззренческую структуру знаний, у нас на сегодня – не разработано.

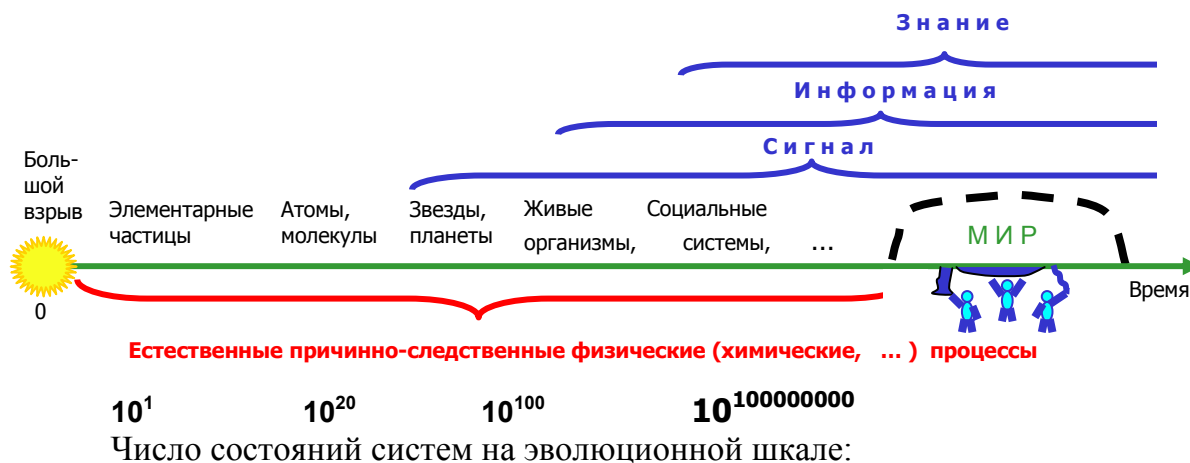
Генезис "информационной субстанции". Невольно возникают вопросы о соотношении "информации" с такими категориями как "*материя*" и "*сознание*": "Что первично, материя или сознание?", "Какая связующая роль информации между ними?", "Является ли информация неотъемлемым свойством материи, как, например, энергия?" и т.д.

Давайте примем наиболее общепринятую в научном мире точку зрения о происхождении нашей Вселенной в результате Большого Взрыва (БВ), как некоторой Сингулярности материи [10,12,13].

В результате БВ, на первых минутах существования Вселенной были порождены "элементарные" частицы: кварки → нуклоны → ядра. Остывая, они далее эмерджентно преобразовывались в: атомы → молекулы → химические соединения, образуя звездные и планетарные системы. Затем, через миллиарды лет, зародились простейшие формы жизни, которые начали био-

логическую эволюцию: клетки → организмы. Далее эволюция высокоорганизованной материи поднялась на следующую ступень – социальную и продолжилась в форме: семьи → племена → государства → человечества (цивилизации). Этот эволюционный процесс продолжается и в настоящее время, и у него есть будущее.

Подавляющее большинство согласится с тем, что информация, как мы ее сегодня понимаем (см. выше 2), в человеческом обществе существует. А на вопрос: "Существует ли информация в мире кварков и элементарных частиц?" – ответ уже затруднителен, см. рис. 2.



Вселенная состоит из $\approx 10^{80}$ частиц; возраст $\approx 13,7$ млрд. св. лет $\approx 10^{17}$ с; число реализованных квантовых состояний $\approx 10^{170}$ и $\approx 10^{10000}$ – потенциальных.

Рис.2.

Как мы уже говорили выше, в качестве отправного рабочего определения понятия "информация" мы будем понимать обмен сигналами (процессами) между материальными образованиями (см. 2а). Под результатом этого "обмена" мы будем рассматривать не просто детерминированный причинно-связанный физический процесс, а некоторую его возможную многовариантность развития. При этом мы должны сохранить причинно-следственную картину мира.

Так в "кварковом мире", самое сложное материальное образование – элементарная частица, имеет не более десятка состояний (см. рис. 2). За время существования Вселенной это материальное образование более чем $\approx 10^{100}$ (Гугол) раз побывало во всех возможных собственных состояниях. И до завершающей Сингулярности (модель пульсирующей Вселенной) еще не меньшее число раз их повторит.

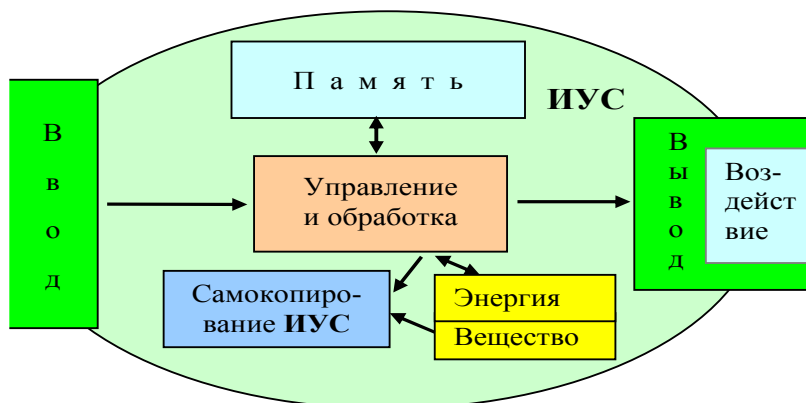
В результате эмерджентного процесса из элементарных частиц образуется более сложное материальное образование – атом. Число состояний атома, того же порядка что и число состояний элементарных частиц. Число состояний сложных белковых молекул, как результат дальнейшей эволюции мате-

рии, уже порядка $\approx 10^{20}$. Они также, за миллиарды лет, многократно будут повторять всевозможные собственные состояния.

Начиная с простейшего биологического организма, потенциальное число собственных его состояний, уже превышает число состояний физической "элементарно-частичной" Вселенной (если частицу Вселенной рассматривать как – "1", а ее пространственную ячейку как – "0", то получим $\approx 10^{170}$ состояний, см. рис. 2). Понятие детерминизма в случае отсутствия тождественной повторяемости состояний (материальных) объектов *теряет смысл*. Как в случае впервые происходящего события, можно точно предсказывать дальнейшее его развитие? Таким образом, лапласовский детерминизм, как инструмент теоретико-логического анализа систем, уже начиная с уровня сложности простых биологических организмов и далее, теряет свое гносеологическое значение. Точно предсказать дальнейшее развитие явления-процесса, которое происходит первый раз в сложной физической системе – становится *принципиально невозможным*, оно может иметь любое качественно новое (эмерджентное) продолжение.

Таким образом, уже с эволюционного уровня звездно-планетарных образований, появляется ниша для устойчивых материальных состояний, а точнее для *сигнала*, сохраняющегося (неизменного) за миллиарды лет существования Вселенной. След (сигнал) какого-либо процесса может оставаться уникальным во всей пространственно-временной Вселенной. Например, геофизики и палеонтологи заняты изучением различных сигналов из прошлого и их информационной интерпретацией.

И только на функциональном уровне сложности биологических организмов и кибернетических систем (и, возможно, других сложных инопланетных материальных образований), появляется возможность оперировать сигналом и информацией. Для этого необходима определенная устойчивая структура ИУС, включающая в себя блоки: ввода-вывода информационного ресурса, памяти, блока самокопирования + хранилища энергии и вещества, а также блока согласованного управления всеми процессами этой системы, см. рис 3.



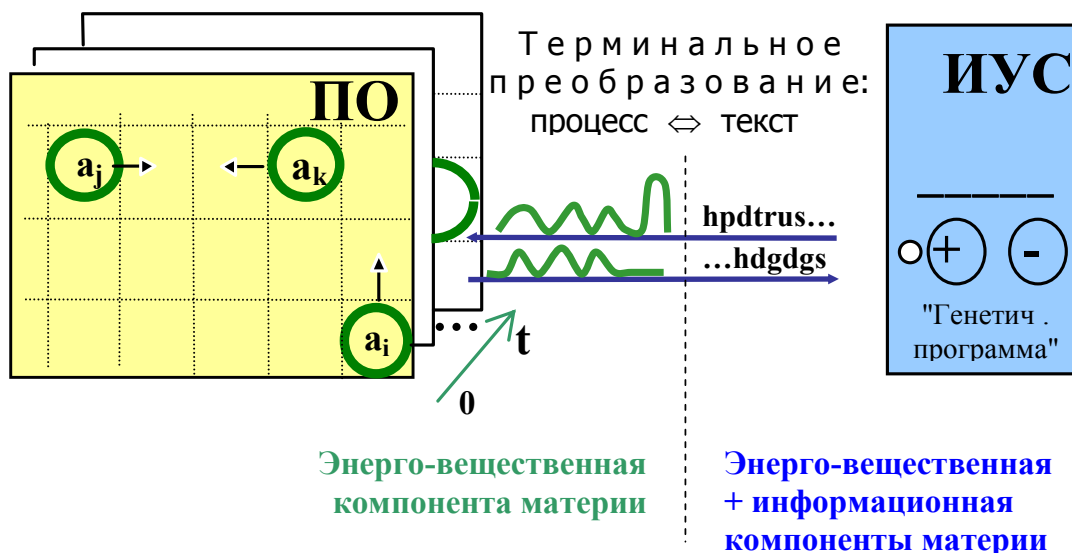
Структура простейшей информационно-управляющей системы

Понятно, что физические процессы предметной области, отображаемые и обрабатываемые в ИУС, должны иметь какую-то специфическую инвариантную форму для возможности организации межблоковых процессов в ИУС. Этой энерго-вещественно-информационной формой и является сигнал. Самая простейшая его форма – это появление/исчезновение какого-либо энерго-вещественного признака (W). *Количественная мера* простейшего сигнала – 1 бит.

Дальнейшая эволюция высокоорганизованной материи (см. рис. 2) приводит к образованию социальных систем, в которых, как мы априори считаем, уже появляются новые категории "информация" и "знание", в качестве следующей ступени развития информационного ресурса. Таким образом, на плазменном и газовом состоянии Вселенной информации, как категории, еще нет. Информация появляется только с появлением структур со множеством устойчивых состояний и с возможностью выбора одного из них, т.е. с появлением ИУС. Формальное определение понятий информационного ресурса: "сигнал", "образ", "информация" и "знание" будет изложено дальше.

Комплекс: "Предметная область – Информационно-управляющая система". Говоря об информации и различных информационных системах, мы редко принимаем во внимание предметную область (ПО), среду, в которых функционируют ИУС. Если же исходить из эволюционной модели, начинающейся с Большого Взрыва (БВ), см. рис. 2., то очевидно, что ПО и ИУС генетически связаны. ИУС фактически формируется под особенностями свойств ПО. Но если проанализировать литературу по информатике, то характеристики ПО не анализируются, обычно о них просто умалчивается, т.к. они предполагаются как самоочевидные для разработчиков ИУС.

Как и в любой молодой развивающейся науке, в информатике, помимо постоянного накопления данных, необходимы периодические попытки структурной и логической компактификации всего накопленного материала на базе новых конкурирующих гипотез. В качестве одной из таких гипотез нами предлагается формализованное описание свойств ПО с позиции ИУС, см. рис. 4.



Комплекс: "предметная область – информационно-управляющая система"

Рис. 4.

Базовые составляющие обобщенной предметной области. Под моделью предметной области (ПО), мы будем рассматривать причинно-связанную совокупность взаимодействующих материальных образований (физических объектов и процессов) на дискретном пространственно-временном поле, см. рис. 5.

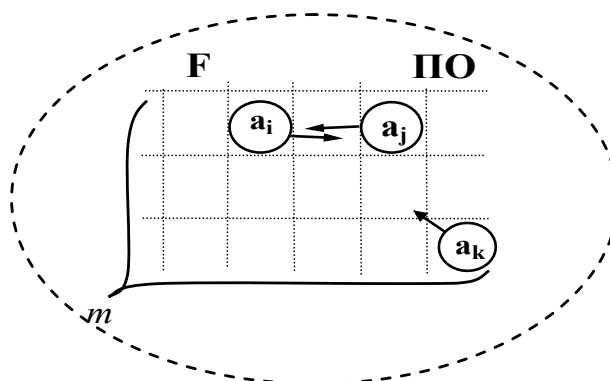


Рис. 5.

При попадании нескольких объектов-процессов ПО в область их взаимодействия (например, пары $a_i * a_j$) порождается характерный для них причинно-следственный физический процесс (W), при этом исходные объекты-процессы ($a_i * a_j$) изменяются (b_i, \dots, b_j).

$$a_i * a_j \rightarrow b_i, \dots, b_j, W \quad (3)$$

Для всех взаимодействий материальных образований в ПО выполняется закон энерго-вещественного (E) сохранения.

$$E(a_i, a_j) = E(b_i, \dots, b_j, W) \quad (4)$$

Если $a_i \approx b_i$, то a_i и b_i можно считать объектами ПО.

Если $a_i \neq b_i$, то a_i и b_i будем считать процессами ПО.

Объекты можно относить к вещественной (статичной) характеристике ПО, а процессы – к энергетической (динамичной). Хотя такое деление условно. В нашей эволюционирующей Вселенной нет абсолютно статичных объектов. Просто, все процессы имеют различные темпы изменений, с позиции ИУС. Так, камень, человек (в своем темпомире) относит к объекту, взрыв – к процессу, а растение – к промежуточному между ними состоянию. Поэтому, принимая сказанное мы все материальные образования во вселенной (ПО) будем считать процессами, но для удобства, в силу исторической привычки, "очень медленные" процессы будем иногда называть объектами. При этом метрику пространства и времени ПО мы будем считать Евклидовой, т.е. неизменной во всех ее точках.

Таким образом, под *процессом* (W) будем понимать независимое от времени и пространства ПО детерминированное изменение во времени некоторого вектора физических параметров. Например, два соприкасающихся бокала всегда порождают характерный процесс (W_i).

Нас процессы будут интересовать с энергетической точки зрения, т.е. возможности их преобразования в другую форму.

Гипотеза. Мы утверждаем, что:

Любой физический процесс может быть автоматизированно преобразован в текстовую форму (5)

Текстовая форма. Любой физический объект-процесс в природе, в конечном итоге, состоит из многих элементов и обладает энергией. Для любых объектов-процессов существуют энергетические пороги нечувствительности. Интервалы значения различных параметров (W), не выходящие за пределы таких порогов, качественно не влияют на результат. Следовательно, возможна *дискретизированная* форма представления квантово-континуальных объектов-процессов ПО в гораздо более компактной текстовой форме (ТФ), отражающей только то, что влияет на результат (причинно-следственное развитие процесса).

В процессе взаимодействия объектов в ПО запускается многомерный (многопараметрический) физический процесс и протекает в течение некоторого интервала времени t_2-t_1 . В течение этого интервала времени происходит изменение значений физических величин компонент вектора параметров (W), характеризующих этот процесс. Изменения значений физических величин по каждой из компонент данного вектора (например, W) могут фиксироваться в ИУС как с некоторым шагом ΔW по амплитуде изменения параметра, так и

по длительности времени – Δt , т.е. быть оцифрованы, или точнее "отекстованны" с преобразованием его в *текстовую форму* (ТФ), см. рис. 6.

Максимальное по амплитуде значение параметра, деленное на ΔW , определяет необходимый размер алфавита кодов $|A|$ ($A = \{a, b, c, \dots, z\}$), а число символов, отображающих процесс, определяется его длительностью $t_2 - t_1$ и дискретизацией Δt . Таким образом, любой исходный физический процесс преобразуется в ТФ.

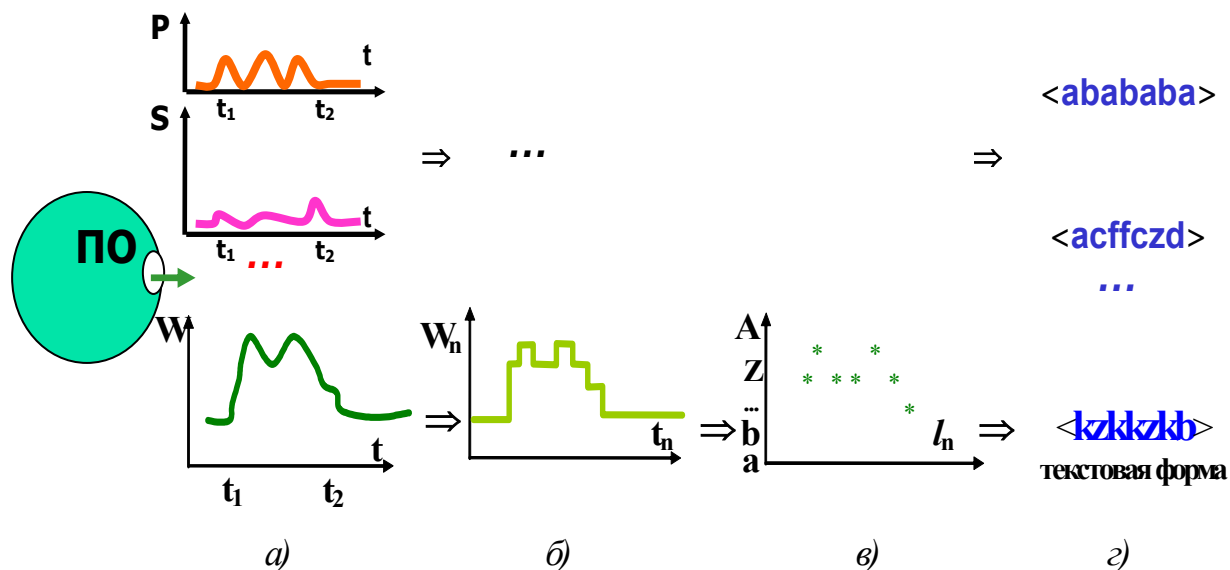


Рис.6.

Очевидно, что произвольный физический процесс любой мерности (графической, акустической, текстовой и т.д.), можно взаимнооднозначно отобразить в ТФ и обратно, с любой наперед заданной степенью точности соответствия (ΔW , Δt). ТФ – это "*цифровая*" форма представления аналогового сигнала и в литературе этот вопрос достаточно полно проработан [??].

Принципы построения алгоритмов оптимального для ИУС квантования (дискретизации) по ΔW и Δt , частично изложены в [7], но для проведения детального анализа планируется отдельная работа (статья).

Для нашего рассмотрения самые значимые преимущества, получаемые от преобразования в ТФ, заключаются:

- в *устойчивости к шумам* при хранении и передаче;
- *унифицированности формы представления* для процессов любой физической природы;
- *отсутствию (минимизации) потерь*, т.е. возможности обратного преобразования ТФ в физический процесс, что говорит об объективности представления в ТФ;
- *автоматизации* преобразования в ТФ (в каждом персональном компьютере есть звуковая карта с входом АЦП (аналого-цифрового преобразования) – фактическим переводом в ТФ);
- *отображению (представлению) в статической ТФ реальной динамики процессов ПО*, посредством повторения ее фрагментов;

- кодировании физического процесса последовательностью знаков (символов) из некоторого алфавита A , в результате чего появляется возможность операций над знаками (материальными объектами), характеризующими данный процесс.

Таким образом, при анализе и моделировании процессов, происходящих в ПО, для ИУС *достаточно* использования только ТФ.

Элементарная семантическая единица. Как уже отмечалось, каждый отдельный процесс может быть закодирован последовательностью знаков ТФ. Соответственно, каждому отдельному процессу соответствует знаковая последовательность, которую мы будем называть *элементарной семантической единицей* (ЭСЕ). "Элементарной единицей", потому что она соответствует одному процессу, а "семантической" потому, что эта ТФ еще и характеризует этот физический процесс (отметим, что вопрос целенаправленной однозначности отображения с позиции ИУС, т.е. анализ процесса дискретизации по ΔW и Δt будет рассмотрен отдельно). Так, например, процесс взаимодействия двух объектов $a_i * a_j - s_{\{a_i * a_j\}}$ отображается последовательностью знаков $\langle \text{cade...b} \rangle$, т.е. однозначно соответствующей ему ЭСЕ.

Допустим, что в ПО может одновременно взаимодействовать не более чем d объектов ("пар", "троек", ..., "d", см. рис. 7.). В результате мы получаем (фиксируем) d -мерный куб со стороной n (n равно числу типов объектов в ПО на данный момент времени). В d -мерном кубе всевозможные одномоментные взаимодействия объектов $\{a_1, a_2, \dots, a_n\}$ *детерминировано* отображаются множеством элементарных процессов (n^d). Число всевозможных ЭСЕ в ИУС (словарь $S = \{s_1, s_2, \dots, s_u\}$) равно u и не превышает числа процессов в ПО:

$$u \leq n^d. \quad (6)$$

Если некоторые участвующие в процессе взаимодействия объекты ПО не воспринимаются ИУС (например, как ультрафиолет для человека), т.е. находятся вне рамок диапазона воспринимаемого ее рецепторами-преобразователями (процесс-ТФ), то с позиции ИУС развитие одного и того же процесса описывается уже *вероятностно*. Соответственно, вероятностно будет следовать и причинно-следственная цепочка ЭСЕ при участии этого не проявляемого процесса. Обычно процесс связан многообразными связями с другими процессами, поэтому его научное познание возможно только при полной картине всех его отношений и связей.

Если в ПО удастся ввести дополнительный $n+1$ объект-прибор, который позволяет обнаружить в виде ТФ ранее не проявляемые ("скрытые параметры") в процессах взаимодействия объекты, то причинно-следственная цепочка ЭСЕ становится для ИУС *детерминированной*.

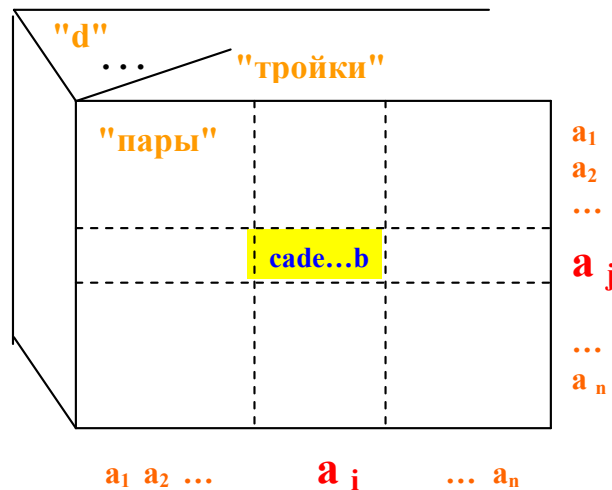


Рис. 7.

В соответствии с (3) в результате взаимодействия может происходить и синтез новых объектов-процессов. Когда в правой части (3) появляются новые ($n+1$, $n+\dots$) объекты-процессы, их свойства часто трудно или вообще не прогнозируемы, как, например свойство "солености" у соединения атомов Na и Cl. Это естественный эволюционный процесс материи, который мы назовем *эмерджентностью*. Таким образом, происходит расширение стороны d-мерного куба: $n+1$, $n+2$, ... (см. рис. 7).

При этом отметим, что в соответствии с (6) мощность числа состояний d-мерного куба существенно превосходит суммарное число реализуемых состояний в виде словаря ЭСЕ. Поэтому, в ИУС рассматривается и считаются корректными только детерминированные пары: *процесс – ЭСЕ* (ТФ).

Описание с помощью ЭСЕ универсально для всех разноуровневых ПО. Так, на микроуровне взаимодействие элементарных частиц может быть описано: <ababsw>. Взаимодействие макроуровневых объектов: <Привет, как дела? Нормально!> и на гигауровне, например, образование двойной звезды из двух астрообъектов как: <zzjbaabj>. Т.е. ТФ универсальна для описания различных ПО, она применима к предметным областям всего диапазона измерений от микро- до гигамира. Таким образом, процессы из произвольных ПО, сливаясь в непрерывный поток ТФ, поступают на вход ИУС.

Характеристики потоков ТФ. В качестве примера мы рассмотрим три основных типа потока в ТФ. Сначала допустим, что процессы в ПО достаточно редки, и тогда ТФ будет иметь вид, представленный на рис. 8а.



Рис. 8а.

Начало каждого причинно-связанного процесса (его фаза) хорошо различима по энергетической паузе между ненулевыми значениями параметра W . Соответственно, в ИУС легко построить словарь из ЭСЕ, отображающих процессы ПО.

Как уже упоминалась, процессы в ПО многопараметрические и различаются, в том числе, и по величине энергетической компоненты. Предположим, что менее энергоемкая компонента процесса ("сигнальная") распространяется с большей скоростью и раньше достигает ИУС. Так как все компоненты физического процесса имеют естественную корреляцию, то "сигнальная" компонента дает ИУС временной интервал в $\Delta\tau$ для подготовки к приходу основной энергетической компоненты.

В качестве следующего примера рассмотрим ПО, насыщенную взаимодействиями таким образом, что процессы идут встык один за другим, и формируемая ими ТФ будет непрерывна, т.е. без энергетических разрывов по параметру W , см. рис. 8б.



Рис. 8б.

Этот случай мы будем рассматривать в качестве основного и на нем будем проводить численный эксперимент. В качестве примера такого типа ТФ можно представить символьный поток на телетайпной ленте от какого-либо малоизвестного племени, или внеземной цивилизации.

Третий случай потока ТФ – это "восточный базар". Множество одновременных процессов накладываются друг на друга, интерферируют, таким образом, что в результате знаки ТФ выходят за диапазон воспринимаемого алфавита ИУС ($A = \{a, b, c, \dots, z\}$), см. рис. 8в. Такая ТФ не воспринимаема ИУС, и чтобы в ней выделить ЭСЕ какой-либо процесс, его необходимо выделить энергетически (по E), например, возвысив над другими. При одновре-

менном повышении порога чувствительности воспринимающих рецепторов ИУС, этот случай сводится к представленному на рис. 8а. Так часто поступают взрослые, повышая голос при обучении ребенка, чтобы быть воспринятыми.

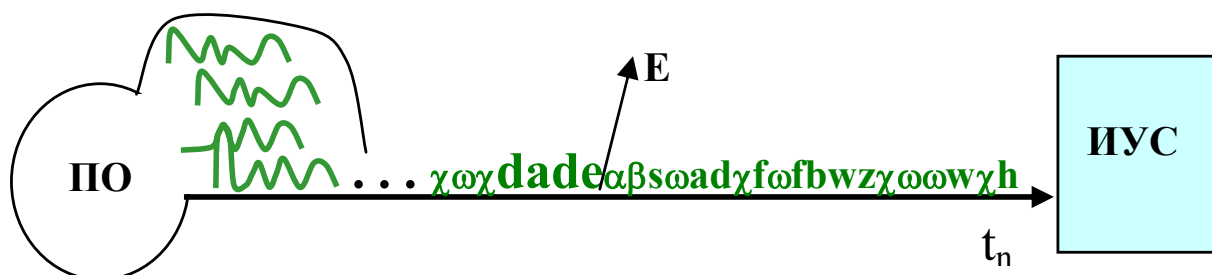


Рис. 8в.

Таким образом, мы рассмотрели все особенности происходящих в ПО процессов и поступающей ТФ на ИУС. Подытожим и уточним эти особенности.

Формальные характеристики ПО. Предположим, что в анализируемых ПО выполняются:

- *принцип причинности* (инвариантность процессов во времени и в пространстве), т.е. если процесс s_i всегда вызывает s_j , то это будет происходить всегда и везде в ПО;
- *принцип локальной однородности* (сведение континуума значений параметра к ограниченному числу качественно-различных образов, т.е. $\langle \text{zzjbaabj} \rangle = s_i \sim \langle \text{zzjbab} \rangle = s_i$); Так, например, вся температурная шкала, для какого-либо вещества, разбивается на четыре образа: $s_i = \langle \text{твердое} \rangle$, $s_j = \langle \text{жидкость} \rangle$, $s_k = \langle \text{газ} \rangle$, $s_l = \langle \text{плазма} \rangle$.
- *принцип субъектности*, т.е. каждый процесс ПО (s_i) имеет некоторую оценку полезности для ИУС. Физическая энергетика процесса (в Дж) входит в целевую функцию эволюционирующих ИУС.

Мы будем рассматривать ПО, в которых различны скорости распространения проекций-компонент физического процесса и при этом, малоэнергетические ("сигнальные") проекции, обычно являясь высокоскоростными, опережают основную энергетическую составляющую физического процесса (на $\Delta\tau$, см рис. 8а). Можно привести примеры таких ПО: перед грозой появляются ее предвестники – грозовые тучи, раскаты грома, отблески молний; перед восходом солнца – светлеет небо (кроме безатмосферных планет) и ночные виды биологических ИУС могут успеть скрыться от его ультрафиолетовых лучей.

В качестве примера не входящих в рассматриваемые нами ПО, можно привести области Вселенной, в которых часты взрывы Сверхновых звезд. Здесь $\Delta\tau = 0$ (как и при радиоактивном распаде элементарных частиц), и ни подготовится, ни обучиться существованию ИУС в таких ПО принципиально невозможно.

Информационно-управляющая система. В рамках рассматриваемого нами комплекса ПО – ИУС (см. рис. 4), задача ИУС (см. рис. 3) будет заключаться в:

- преобразовании физических процессов посредством рецепторов ИУС в ТФ;
- хранении ТФ и преобразовании ее в *информацию*;
- структурировании *информации* в наиболее эффективные формы для хранения и обработки (*знания*);
- прагматической оценки поступающей *информации* с целью выбора и запуска полезных для ИУС процессов;
- выдавать *информацию* в виде ТФ и преобразовывать ее через эффекторы ИУС в физические процессы;
- иметь положительный эволюционный потенциал.

Эволюционный потенциал (ЭП) характеризует эволюционную перспективность ИУС при ее функционировании в ПО. Эволюционный потенциал – двухкомпонентная характеристика. Она состоит из материальной компоненты (вещества и энергии), образующей тело и обеспечивающей двигательную активность ИУС, а также нематериальной компоненты (информации, знания), позволяющей прогнозировать развитие процессов. Материальную компоненту будем измерять в энергетических единицах $E+$, а нематериальную компоненту в Δt числом прогнозируемых знаков ТФ. Общую величину эволюционного потенциала можно оценивать как функцию от величин $E+$ и Δt , например, как их сумму. Подробному рассмотрению эволюционного потенциала будет посвящена отдельная статья (работа). Нам же сейчас это понятие будет необходимо для проведения численного эксперимента по выделению информационной компоненты из ТФ, в комплексе "ПО – ИУС".

Зарождение простейших ИУС. Исходя из принятой нами текстовой (лингвистической) метафоры, в качестве модели среды зарождения *простейших ИУС* сделаем предположение, что произошел первоначальный "Большой символный взрыв" и породил множество символов "а" алфавита $A \{a,b,c,d,\dots,w,x,y - \text{аналоги химических элементов и } z - \text{аналог пространства}\}$. Далее, символы под действием собственной потенциальной энергией ($a>b>c>d>\dots>w>x>y>z=0$) и в соответствии со своим "полупериодом синтеза" ($a \rightarrow bbz$; $b \rightarrow ccz$; $c \rightarrow ddz$; ...) начинают эволюционировать в различные тексты ($bb \rightarrow bccz$; $bb \rightarrow ccccz+z$; $ccb \rightarrow ddcbz$; ...). В результате порождается "текстовой бульон" (см. рис. 9).

ше, т.е. $\approx 10^3$. Таким образом, строя заданную фразу "Быть или не быть ...", уже из реальных слогов, мы увеличиваем вероятность ее формирования на семь порядков. На словах, словосочетаниях и фразах мы таким же способом выигрываем еще 39 порядков, итого – 46. И вероятность случайного формирования фразы: "Быть или не быть, вот в чем вопрос?", с изначально практически невозможной – $P \approx 10^{-80}$, становится уже практически вероятной – $P \approx 10^{-34}$. Таким образом, из этого примера следует, что иерархизация эволюционного процесса делает вполне реальным вероятностное формирование простейших ИУС, тем более, что при этом создается и окружающая их "биота" (ПО). Более детальный анализ иерархизации эволюционного процесса будет изложен в отдельной работе.

Следовательно, ресурсов реальной Вселенной вполне достаточно, чтобы в ней случайно сформировалась "текст-ДНК" простейших самодублирующихся ИУС, которые за счет дальнейших мутаций могут начать *эволюционный процесс* в "символьной Вселенной" ("текстовой жизни"). При этом, символьные тексты среды (ПО) рассматриваются одновременно: как строительный материал, как носитель энергетического ресурса и как источник упреждающей информации для ИУС.

Отметим, что направленность эволюционного процесса не случайна, а имеет аттракторный характер, т.е. пространство эволюционного развития неравномерно и имеет эволюционный градиент в предпочтительные области "символьной Вселенной". Таких предпочтительных областей мы насчитали шесть, от простейших информационных автоматов и до Ноосферы, иерархически входящих друг в друга. При этом, эволюционные траектории, проходящие через них, имеют тенденцию экспоненциального возрастания эволюционного потенциала при вхождении в эволюционный аттрактор "символьной Вселенной". Более подробно аттракторный характер эволюции мы также рассмотрим в отдельной работе.

Фактически всем предыдущим материалом этой статьи мы рассмотрели основные свойства ПО, комплекса "ПО – ИУС", свойства и механизмы возможного формирования простейшей ИУС. Теперь у нас есть все исходные данные, чтобы формально рассмотреть и проанализировать появление "*информации*" в ИУС как некоторой категории, равновеликой категориям: энергия, пространство и время. Исследовать свойства "информации" и выразить их в системе СБС. А также проанализировать другие категории информационного ресурса "*сигнала*" и "*знания*" и соотнести их в единую классификационную систему информационного ресурса.

Часть II

Автоструктуризация (автокластеризация). Начнем с общеизвестного факта, – все биологические организмы (ИУС) "единогласны" в кластеризации окружающего мира на отдельные образы, см. рис. 10.



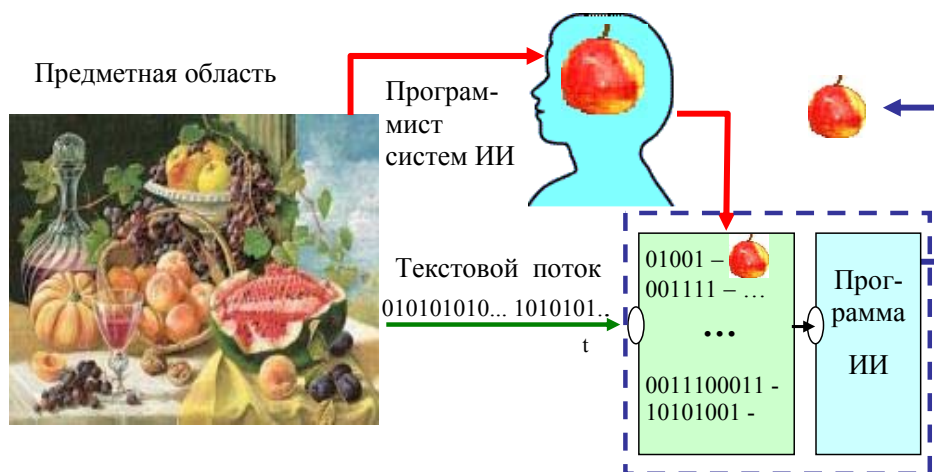
Рис. 10.

Выделение из непрерывного потока сигнала отдельных образов одно из основных свойств человеческой психики, которое мы воспринимаем как должное и фактически не замечаем. Но если вдуматься, ведь каждый предмет или процесс, которые у нас получают собственное имя, никогда не встречаются абсолютно изолированными в окружающем нас мире. Когда ребенку говорят "яблоко", то перед его глазами демонстрируется рука матери с яблоком и все это на сложном фоне. Почему же впоследствии накопления жизненного опыта, подавляющее большинство под "яблоком" понимают именно яблоко, а не "какую-то часть руки + часть фона + часть яблока", как они всегда наблюдали это в действительности?

Аналогичную задачу мы попытаемся решить, заставляя информационную систему разбивать (кластеризировать) непрерывный символьный поток на составляющие его образы. Естественно, допуская при этом предположение, что ни язык, ни грамматика этого символьного потока ИУС не известны, в отличие от современных методов разработки технических информационных систем (ИС).

При разработке традиционных ИС на разработчика-аналитика (программиста) ложится основная функция по структуризации ПО. Исходя из исторически сложившегося понимания некоторой конкретной ПО, а также решаемой в ней задачи, программист устанавливает в соответствие признаки

процессов-объектов ПО и их семантические значения. Например, в нашем случае, "круглое" и "красное" (0111001), которые иерархически могут распадаться на еще более детальные признаки, ставится в соответствие семантическому понятию-денотату – "яблоко". И далее с этим понятием будет оперировать разрабатываемая ИС, которая будет считаться "интеллектуальной" ИС, см. рис. 11.



Традиционный ("контрабандный") способ структуризации ПО в системах ИИ

Рис. 11.

Очевидно, что если при эксплуатации данной ИС изменится характеристики ПО, например, цвет или форма яблок, то в программу необходимо будет внести соответствующие изменения, или же вынуждено ограничить возможности ИС. Если же ПО открытая, или априорно неопределенная, то традиционный способ становится вообще неприемлемым (Это первый "грех" традиционного направления разработчиков ИИ-направления). Существующая технология построения современных ИУС является основным камнем преткновения на пути автоматизации обработки больших и сверхбольших неструктурированных (слабоструктурированных) информационных потоков. Хотя у нас перед глазами пример биологических ИУС, которые вполне успешно справляются с задачей кластеризации любых априорно неопределенных ПО. Встает вопрос: "Могут ли современные инженеры повторить результат естественно эволюционировавших ИУС?"

Рассмотрим *демонстрационный пример* комплекса "ПО–ИУС". Допустим, что в ПО всего три объекта: $\{a_i, a_j, a_k\}$. Далее предположим, что при их попарном взаимодействии: $(a_i * a_j)$, $(a_i * a_k)$, $(a_j * a_k)$, порождаются процессы, которые в текстовой форме имеют вид: <МАК>, <ЛЕН>, <РИС>. Когда взаимодействуют все три объекта $(a_i * a_j * a_k)$, то порождается процесс – <ДУБ>. Предположим, что эти четыре равновероятных процесса отображающихся в текстовых формах, сливаются в непрерывный символьный поток (ТФ), см. рис. 12.

Этот демонстрационный пример полностью укладывается в предлагаемый в статье формальный аппарат и показывает достаточность ТФ для адекватного отображения в ИУС любых физических процессов ПО. В качестве некоторого бытового аналога данного примера можно предположить, что: объекты – это шары, текстовая форма – это звук процессов их соударения, символьный поток – это непрерывный гул, порождаемый соударяющимися шарами.

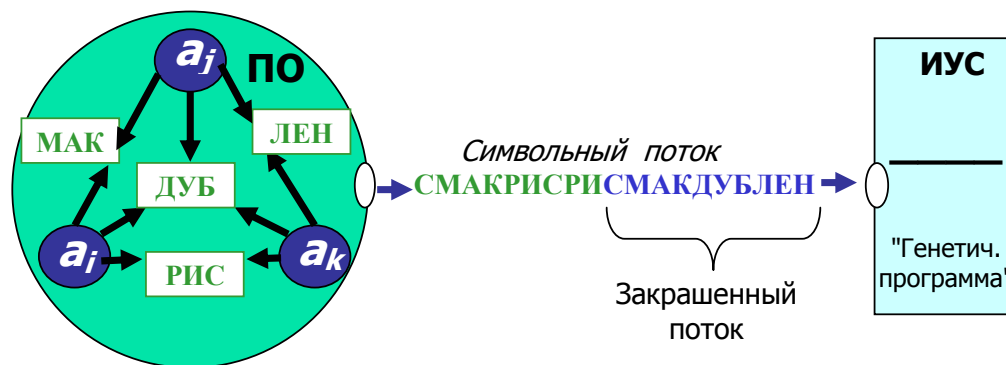


Рис. 12.

Минимально необходимая задача любой ИУС – сформировать "толковый" словарь кодовых последовательностей (структур) ПО и соответствующих им семантических образов ИУС, однозначно отображающих процессы ПО. Восприятие символьного потока (ТФ) должно быть без каких-либо пропусков, т.е. полностью покрывающим ("закрашивающим") весь символьный поток. Напомним, что в нашем случае, ИУС должна выполнить эту задачу самостоятельно на априорно неизвестной ПО.

Для приближения к реальности взаимодействия ИУС (биологических организмов) с ПО (внешней средой) и для формализации ряда определений информатики дополнительно введем некоторую "биологизацию" этого взаимодействия, которая будет заключаться в "энергетической" оптимизации процессов ИУС.

Для этого будем оценивать "энергетическое" состояние ИУС усреднено на 12-ти символьном (тактовом) интервале текстового потока (ТФ). Примем, что:

- время обработки одного символа ТФ в ИУС – один такт T ;
- "энергетические" затраты ($E(O)$) при распознавании на обработку одного образа равны $-1E$;
- каждые 12 тактов в ИУС на "энергетическое" самообеспечение ($E(N)$) каждого образа словаря потребляется по $-0,5E$;
- за каждый точно прогнозируемый символ текстового потока ИУС получает "энергетическое" подкрепление в $+1E$.

Энергетический баланс ИУС на каждые 12 тактов определяется как:

$$E_+ = -12(\text{обработка}) - 1/2N(\text{хранение}) + f(T)(\text{прогнозирование}) \quad (7)$$

где: N – число образов словаря ИУС;
 T – число прогнозируемых ИУС символов в ТФ

Таким образом, задача ИУС заключается в том, чтобы в непрерывном потоке, с произвольных символов ТФ (t_{01}, t_{02}, \dots) сформировать словарь в N -образов, которые бы без разрывов закрашивали текстовый поток и, в соответствии с (7), максимизировали величину E_+ , за счет прогнозирования будущих символов ТФ, см. рис. 13. Фактически, E_+ – эта величина характеризует эволюционный потенциал ИУС (см. выше).

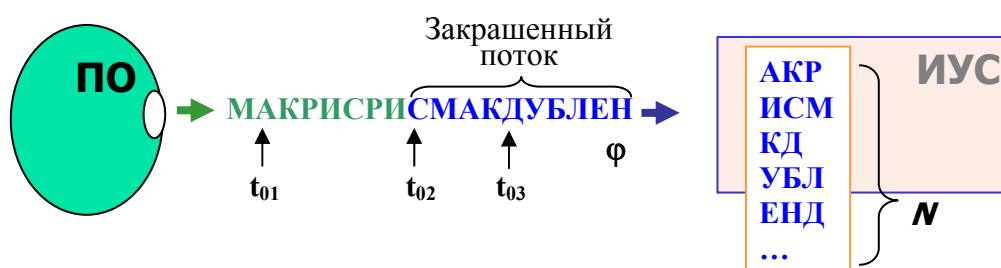


Рис. 13.

Результаты экспериментов. Первая структуризация словаря ИУС. Начнем с самой простейшей ИУС. Процедура формирования словаря предельно упрощена, и ИУС строит односимвольные образы словаря ($L=1$), т.е. размер образа не превышает одного символа: $\langle M \rangle$, $\langle A \rangle$, $\langle K \rangle$, $\langle P \rangle$, $\langle I \rangle$, $\langle C \rangle$, $\langle L \rangle$, $\langle E \rangle$, $\langle H \rangle$, $\langle D \rangle$, $\langle Y \rangle$, $\langle B \rangle$.

При этом "энергетический" баланс данной ИУС на каждом символьном интервале потока, по 12 символов, будет следующим:

- затраты на распознавание 12 символьного интервала потока
 $E(O) = -12E$;
- затраты на обслуживание словаря ИУС каждые 12 тактов
 $E(N) = 12 * (0,5E) = -6E$;
- потенциальное преимущество от прогнозирования $E(T) = 0E$ (т.к. образы состоят только из одного символа);

Итог: для односимвольного словаря $E(O) + E(N) + E(T) = -18E$.

Вторая структуризация словаря ИУС. Усложним ИУС. Теперь образы словаря не должны превышать двух символов ($L \leq 2$). Отметим, что при этом, в зависимости от начала разбиения символьного потока (t_{01}, t_{02}, t_{03} , см. рис.13) на образы, ресурсные затраты ИУС могут существенно различаться. Среди них можно выбрать как минимальные, так и максимальные словари.

- а) Минимальная форма словаря: $\langle MA \rangle$, $\langle K \rangle$, $\langle RI \rangle$, $\langle C \rangle$, $\langle LE \rangle$, $\langle H \rangle$, $\langle DY \rangle$, $\langle B \rangle$:
- $E(O) = -8E$;
 - $E(N) = 8 * (-0,5E) = -4E$;

- прогнозирование $E(T) = ((\text{двухсимвольный образ} = 1E) + (\text{односимвольный образ} = 0E)) * 4 = +4E$, или по $+0,5E$ на образ ИУС.

Итог а₂): $-4E - 8E + 4E = -8E$.

б) Максимальная форма (без перекрытия): <МА>, <КЛ>, <КР>, <КД>, <КМ>, <ЛЕ>, ..., <БД>:

- $E(O) = -20E$;

- $E(N) = 20 * (-0,5E) = -10E$;

- прогнозирование $E(T) = +4E$.

Итог б₂): $-20E - 10E + 4E = -26E$.

В результате, для двухсимвольной структуризации словаря качество ИУС, а точнее, ее эволюционный потенциал лежит в интервале от $-8E$ до $-26E$ (и до бесконечности при дублировании или перекрытии образов).

При третьей структуризации словаря образы словаря не должны превышать трех символов ($L \leq 3$).

а) Минимальная форма словаря: <МАК>, <РИС>, <ЛЕН>, <ДУБ>:

- $E(O) = -4E$;

- $E(N) = 4 * (-0,5E) = -2E$;

- прогнозирование $E(T) = (\text{трехсимвольный образ} = 2E) * 4 = +8E$, или по $+2E$ на образ ИУС.

Итог а₃): $-4E - 2E + 8E = +2E$.

б) Максимальная форма (без перекрытия = 36 образов): <АКР>, <ИСП>, <ИСМ>, <АКД>, <УБЛ>, <ЕНД>, ... <АКМ>:

- $E(O) = -36E$;

- $E(N) = 36 * (-0,5E) = -18E$;

- прогнозирование $E(T) = +4E$.

Итог б₃): $-36E - 18E + 4E = -50E$.

В итоге, опять имеем множество различных словарей, с введенной нами "энергетической" мерой от $+2E$ до $-50E$ (и далее). Естественно, что минимальная форма словаря: <МАК>, <РИС>, <ЛЕН>, <ДУБ> наиболее эффективна для функционирования ИУС. Так как у ИУС всего четыре образа памяти, то затраты на их обслуживание при распознавании символьного потока *минимальны*. Минимальны как по вещественным ресурсам (всего четыре образа), так и по "энергетике" обслуживания и по времени обработки входных символов. При этом, по своей функциональности, минимальная форма словаря не уступает максимальной форме.

Четвертая структуризация словаря (размер образа не превышает четырех символов, т.е. $L \leq 4$).

а) Минимальная форма словаря: <МАК>, <РИС>, <ЛЕН>, <ДУБ>.

Итог а₄): $-4E - 2E + 8E = +2E$.

б) Максимальная форма (без полного перекрытия = 48 образов): <АКРИ>, <ИСРИ>, <ИСМА>, <АКДУ>, <ЕНДУ>, ..., <АКМА>.

Итог б₄) = $-78E$.

Пятая и другие структуризации словаря в нашей модели дают следующие значения эволюционного потенциала:

$$a_{5,6,7\dots}): = +2E.$$

$$b_{5,6,7\dots}): = -96E, -128E, \dots -\infty.$$

Общий итог структуризации:

а) лучший итоговый эволюционный потенциал = +2E ;

б) худший = $-\infty E$!

Рассмотренные нами результаты эксперимента показывают, что эволюционный потенциал ИУС *обратно пропорционален* размеру словаря. Это и очевидно, исходя из логики эволюции биологических организмов и развития технических систем. Минимальная форма словаря наиболее эффективна для функционирования ИУС как по времени, так и по всем другим значимым для ИУС ресурсам.

Анализ результатов эксперимента указывает, что размеры минимальных a_j и максимальных b_j словарей ИУС, отображающих символьный поток из ПО, соотносятся как минимум, как величины в первой и в третьей степенях:

$$o(N) \text{ и } o(N^3) \quad (8)$$

где: N – размер минимального словаря ИУС
и исходного числа процессов в ПО

Соотношение (8) позволяет строить множество несложных "вычислительных" процедур в ИУС, находящих минимальный словарь. Более точные оценки малого a_j и большого b_j словарей, в зависимости от длин ТФ процессов (ϵ), их числа (N) и длин образов (L) приведены на рис. 14.

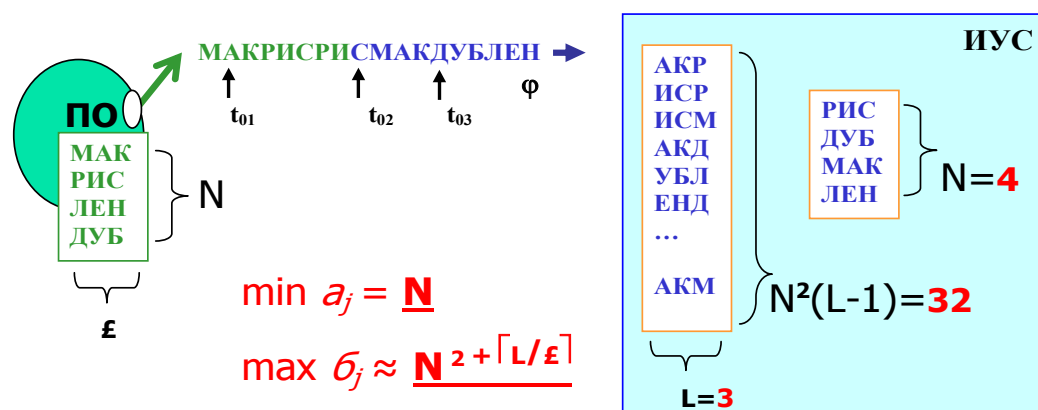


Рис.14

Из анализа проведенных экспериментов следует, что минимальная форма словаря ИУС может достигаться только при его *гоморфности* структуре исходных процессов в ПО. Сформулируем достаточно очевидную теорему.

Теорема 1 (анализа): минимальная форма словаря ИУС может достигаться только при ее *гомоморфности* исходной структуре процессов в ПО.

Доказательство: Представим равномерно случайное формирование ТФ двумя посимвольно непересекающимися процессами А (t_1, t_2) и Б (t_1+1, t_3), см. рис. 15.

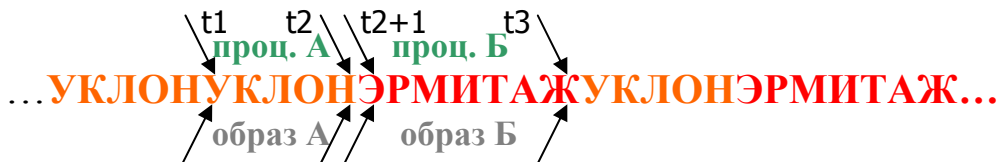


Рис.15

Предположим, что мы строим образы словаря на отрезках $[t_1, t_2]$ и $[t_2+1, t_3]$. Два эти образа в ИУС имеют общую длину $t_3 - t_1 + 1$, и так как символы образов различны, то данная простая линейная структура образов словаря, полностью закрашивающая поток ТФ, будет минимальна. Отметим, что при этом структура словаря ИУС гомоморфна структуре процессов ПО (см. рис. 14).

Если же границы образов словаря будут сдвинуты хотя бы на 1 от t_1 , t_2 , t_2+1 или t_3 , то возрастет количество образов словаря и его размер, и при этом, существенно уменьшится эволюционный потенциал данной ИУС (см. выше результаты эксперимента). При этом, естественно, что структура словаря ИУС уже будет отлична от структуры процессов в ПО. Таким образом, только словарь, гомоморфный структуре процессов ПО, полностью закрашивает поток ТФ, являясь одновременно минимальным и с максимально возможным эволюционным потенциалом, в соответствии с (7).

Утверждение теоремы 1 верно и для отображения сложных иерархических процессов ПО. Для установления данного утверждения заменим процессы-образы вышеприведенного доказательства иероглифами (номераами), переводя их в ТФ следующего уровня, и сводим задачу к уже рассмотренному доказательству. Тем самым доказывая справедливость утверждения теоремы 1 и для иерархических сложных процессов ПО, см. рис. 16.

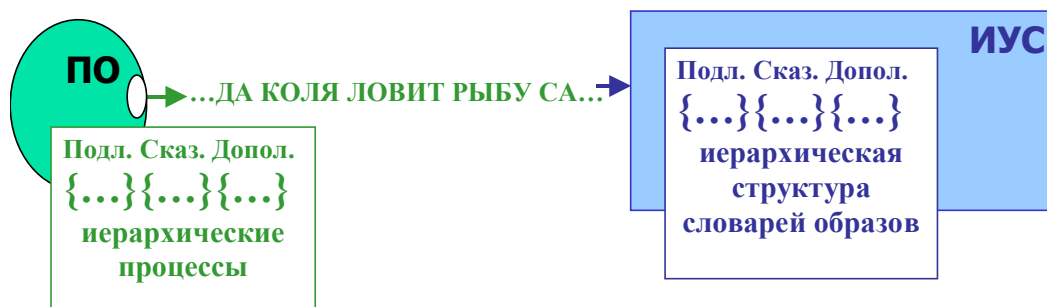


Рис. 16.

Для построения гомоморфного отображения (словаря ИУС) в соответствии (7) и условием (8) можно использовать достаточно широкий набор разнообразных методов и вычислительных алгоритмов. Результаты многочис-

ленных экспериментов подтверждают этот вывод. Так, строя минимальный словарь на символьных потоках, например, на конкретных ТФ: <RISMAK DUBLENLENDUBMAKMAKLENRISRISLENMAKRISDUBRIS>, легко находятся исходные слова-процессы: <МАК>, <РИС>, <МАК>, <ДУБ>; – на потоке <ДОМЗЕБРЫСКИТНАДОМДОМВНАДОМВСКИТВНАСКИТВВЗЕБРЫ>, правильно выделяются слова-процессы (семантические единицы) ПО: <ЗЕБРЫ>, <СКИТ>, <ДОМ>, <НА>, <В>, из которых и был сформирован этот поток ТФ. Отметим, что из ТФ, полученной из предыдущей ТФ, сдвигом текста ее алфавита А в кодах ASCII на +1 <ЕПНИЖВСЬТЛЙУОБЕПНЕПНГ ОБЕПНГТЛЙУГОБТЛЙУГЖВСЬОБГОБ>, были выделены те же самые семантические единицы, но также сдвинутые на +1 в новой нотации (<ИЖВСЬ>, ... , <Г>.).

Из теоремы 1 следует и объяснение явления "единогласия" биологических ИУС по кластеризации окружающего мира на отдельные образы, см. рис. 10. Минимальные структуры образов в различных биологических ИУС(НСС), построенные на одной и той же ПО, оказываются топологически эквивалентными на уровне причинно-следственных подструктур. Поэтому они и "понимают" друг друга в кластеризации окружающего мира, т.к. это объективный процесс. К тому же, минимальный словарь дает ИУС эволюционные преимущества. Вообще, принципы минимизации затрачиваемых материальных ресурсов (вещество, время, энергия), при максимизации прогнозистического отображения внешней среды (ПО) – это основной принцип направленности эволюции ИУС. Который совпадает по направленности с эволюционным потенциалом (см. выше).

Естественно, что автокластеризация открывает большие перспективы и для перехода разработки технических информационных систем с обманчиво легкого "контрабандного" пути (см. рис. 11), на проверенный биологической эволюцией, – *нейросемантический* (см. рис. 10).

Под технической информационной системой мы будем понимать систему, у которой раз и навсегда заданна целевая функция "поведения" (функционационирования). Для ИС – целевая функция постоянна и не зависит от времени и пространства. Например, ИС – замок и ключ. В отличие от ИС у ИУС имеется возможность модификации целевой функции в зависимости от последовательности ситуаций (процессов) и их контекста (траектории) при взаимодействии ИУС с ПО. При этом, полное число возможных вариантов состояния ИУС не может быть опробовано (пройдено) за время существования Вселенной, т.е. ИУС не детерминированная строго система. В качестве примера ИУС можно привести адаптивную систему распознавания образов.

Свойство гомоморфности, вытекающее из теоремы 1, открывает широкие перспективы по автоматизации построения баз данных и АСУ для сложноструктурированных крупномасштабных и априорно неизвестных ПО. А это один из трудоемких этапов в процессе существующих технологий построения автоматизированных систем управления ($\approx 20-30\%$ от стоимости АСУ). Т.е., эксплуатируя соотношения (7) и (8), можно быстро и дешево про-

водить автоматический анализ структуры ПО, для разработки и постройки автоматизированных систем управления, см. рис. 17.

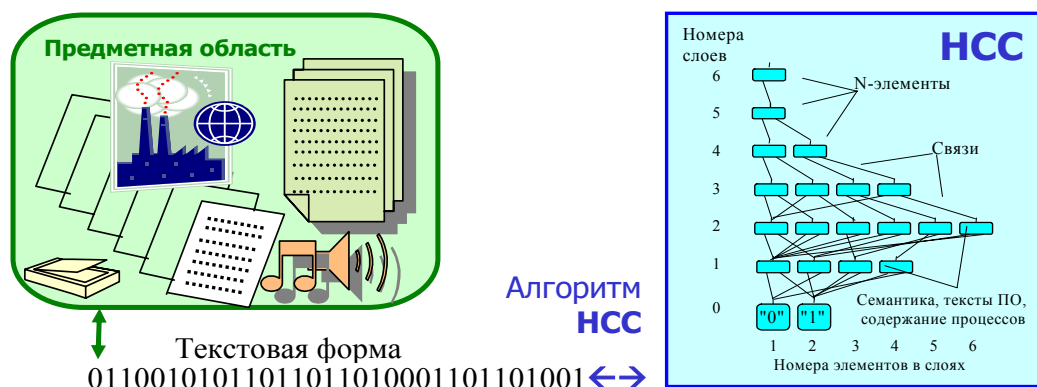


Рис. 17.

Нейросемантическая форма представления информации (HCC). Как уже было показано, на ТФ можно построить минимальный покрывающий ее словарь. Если же закрашенные участки ТФ заменить индексами (номерами-именами), соответствующих им образов словаря, то получим новую, более короткую ТФ₂. Строя аналогичным образом новый словарь на ТФ₂ и, заменяя ее индексами, получим ТФ₃, и т.д., пока длина ТФ_m не будет равна 1. В результате, исходная ТФ превращается в структуру иерархически взаимосвязанных словарей, которую мы назвали *нейросемантической структурой* (HCC). Исходная ТФ и HCC взаимопереводимы, но HCC обладает большей информативностью [7].

Как пример, на рис. 17 показана ТФ и соответствующая ей HCC. Если представить, что все возможные процессы изучаемого объекта преобразуются в ТФ и затем переходят в HCC, то в HCC мы получаем препарированную причинно-следственную семантическую модель произвольного исследуемого объекта. Допустим, что исследуемый объект – это некое предприятие. Тогда в многодольном иерархическом графе (HCC), мы находим подписанные в алфавите A ($A = \{0,1\}$) объекты-процессы с понятным для нас смыслом.

Например, вершина №1 слоя №1 <01>, а вершина №1 слоя №2 <0110> и этот образ-процесс входит в вершины (образы-процессы) №1, №2 и №4 слоя №2 HCC. Допустим, что для специалиста данной ПО это значит, что "продукция" <0110> поступает в "цеха" №1, №2 и №4, и т.д. Т.е. все, что было отображено в ТФ, для специалиста переходит в графовую модель HCC. И эта большая часть работ по автоматизации, причем она выполняется *оперативно и автоматически*. На HCC, как гипертекстовом графе [16], пользователь-специалист может надстроить топологическую копию, но уже с интересующими его операциями (например, сбор в список образов-процессов слоя №2 – будет "номенклатурным списком продукции данного предприятия", а сложение значений определенных образов-процессов, – "суммарный выпуск продукции данным предприятием" и т.д.) и АСУ данного предприятия готова к эксплуатации.

НСС – это пример формального преобразования статистических (количественных) характеристик ТФ, отражающей процессы в ПО, в качественно новую форму – причинно-следственную структуру образов-процессов ИУС. Алгоритм этого преобразования эффективно выполняется на нейроподобных N-элементах [7,9,10]. В результате его работы физические процессы ПО через ТФ целостно переходят в НСС, в которой каждый процесс, в зависимости от своего уровня сложности, отображается соответствующей вершиной (N-элементом, см. рис. 17) определенного иерархического уровня НСС. Т.е. формируется взаимнооднозначное соответствие:

Процесс ПО (ЭСЕ) ↔ N-элемент (образ НСС) (9)

И так как физический процесс – это семантическое понятие (если он значим для ИУС, то он обязательно имеет имя), и он отображается в нейроподобных N-элементах НСС [7], то эта структура и называется *нейросемантической формой представления информации*.

Хотя в рамках поставленной перед нами задачи, нас будет интересовать только свойство (9) – взаимно-однозначная локализация семантических единиц в физических объектах, но это свойство лежит в основе всего направления, названного *нейросемантикой*.

Вообще, свойство (9) позволяет существенно упростить или даже отказаться от многих понятий информатики и вычислительной техники. Так в ИУС на базе НСС можно вообще отказаться от файловой системы, заменив ее автоматической локализацией смысловых (семантических) единиц на собственных ("личных") областях памяти, а необходимость адресного обращения через каталоги, заменяется ассоциативной связью по содержанию. При этом, N-элемент фактически выполняет функции распознавания, прогнозирования и моделирования причинно-следственных связей в ПО, а небольшие добавления к структуре НСС, и у пользователя уже готовая база данных (знаний). С другой стороны, глядя на структуру нейросемантического графа, отмечаем статистику сразу всех процессов данной ПО: вычисляем значения текстовой энтропии вершин НСС (процессов — N-элементов), определяем точки роста НСС и причинную-взаимосвязанность в ПО [7].

Следует также отметить, что все технические характеристики ИУС(НСС), такие как: *время доступа, коэффициент компрессии-сжатия, надежность-пластичность хранения информации* и др., имеют тенденцию к улучшению, как в среднем, так и в абсолютных значениях по мере роста объема вводимой ТФ.

Программистам и разработчикам ИС очевидно, что свойство (9) открывает широкие возможности, как по упрощению и ускорению построения АСУ, так и по открытию пути эволюционного самопрограммирования ИУС. Нейросемантика – это новое многогранное направление в информатике XXI века, и ему будет посвящена отдельная статья, но в рамках сейчас поставленной перед нами задачи, нас будет интересовать только свойство (9).

Одно из следствий (9) приводит к тому, что *компрессия* (P) отображения ТФ в НСС монотонно возрастает с ростом объемов вводимого символического потока:

$$P = \frac{\text{ТФ}_{\text{ИУС}} = \text{объем текстовой формы в ИУС в битах}}{\text{НСС}_{\text{ИУС}} = f(\text{число N-элементов, число связей}) \text{ в битах}} \quad (10)$$

и обратно, пытаясь максимизировать величину компрессии на конкретной ТФ, мы добиваемся выполнения условия (9). Т.е. условие (10) – это еще один метод автоматического формирования НСС.

Понятно, что не на любой ТФ можно добиться абсолютного выполнения условия (9). Для построения НСС необходимо выполнение условия достаточного представительства в ТФ всех причинно-следственных отношений данной ПО. В вышеприведенных примерах автоструктуризации показаны предельно минимальные размеры конкретных ТФ(ПО): <RISM... RIS> и <ДО... РЫ>, когда правильно выделяются все семантические единицы (ЭСЕ).

Под автоструктуризацией мы будем понимать процесс формирования структуры словарей ИУС, гомоморфной структуре процессов ПО. Механизмов автоструктуризации достаточно много: от вариационных методов, до используемого нами нейросемантического метода [2].

Критерии достаточности ТФ(ПО) следующие:

- в ТФ отображено все пространство всевозможных причинно-следственных процессов ПО (так называемый "гиперпериод" [10], равный длине $L(\text{ТФ}) = N^{\aleph}$, где N – число различных процессов в ПО, \aleph – длина причинно-связанной цепочки процессов);
- если человек может правильно структурировать данный текстовый материал в непривычной для него нотации (например, в сдвинутом алфавите A кодов ASCII);
- наличие характерных особенностей динамического процесса при автоструктуризации (например, начало лавинообразного процесса (9) – выделение ЭСЕ, см. рис. 18);
- если каждый процесс будет иметь более одного сочетания с другими последующими процессами, то минимальная длина ТФ может не превышать удвоенной суммы всех длин процессов (образов) ПО (Это условие будет подробно рассмотрено в дальнейших работах).

Качество самой НСС отображается отношением числа корректных пар (процесс ПО (ЭСЕ) \leftrightarrow N-элемент (образ НСС)) на общее число N-элементов в НСС. Естественно, что в процессе автоматической подстройки ИУС под конкретную, априорно неизвестную ПО, на начальном этапе ввода текстового потока "*НСС-качество*" будет невелико ($\ll 1$), а далее "*НСС-качество*"

полностью зависит от алгоритмов формирования НСС. Качественный график динамики автоструктуризации (НСС) имеет следующий характерный вид, см. рис. 18.



Рис. 18.

Скорость автоструктуризации можно существенно улучшить, выделив и введя уже известные экспертам данной ПО "семантические заправки" (ЭСЕ). Такой процесс автоструктуризации принято называть обучением "с учителем" (см. рис. 18). Эффективность обучения "с учителем" хорошо известна для биологических ИУС.

В работе [7] исследовано и показано, что высокое качество автоструктуризации (процесс ПО (ЭСЕ) \leftrightarrow N-элемент (образ НСС) или свойство гомоморфности НСС к ПО) открывает возможность естественного эволюционного саморазвития ИУС.

Основной вопрос философии (первичность материи или сознания). При рассмотрении комплекса "ПО – ИУС" (см. рис. 4) обнаруживается интересная особенность, которая заключается в том, что, обладая ограниченным числом собственных процессов, ПО порождает неограниченный поток ТФ. Также как и ИУС – ограниченным словарем образов может закрасить (покрыть) неограниченную ТФ. Т.е. ПО и ИУС как бы зеркально взаимоотобразимы относительно друг-друга в комплексе "ПО – ИУС". И более того, с позиции НСС каждый участник комплекса содержит в себе другого: ПО ИУС и ИУС моделирует ПО. Таким образом, в модели пульсирующей Вселенной вопрос: "первичности материи или сознания" *становится неактуальным*. Если когда-то ПО и породила ИУС, то при последующих Больших Взрывах (сингулярностях), ИУС вполне могла участвовать в процессе формирования ПО, выбором набора Космологических Констант [7].

Продолжая рассуждения об обратимости потока ТФ, но уже в количественной форме, можно представить, что ИУС на базе НСС как бы генерирует этот поток, а ПО его теперь воспринимает. Тогда для разносимвольных (нет посимвольного пересечения, см. рис. 15) отображений процессов-образов доказательство гомоморфности ПО к ИУС (НСС) становится следствием теоремы 1.

Из гомоморфности ИУС(НСС) к ПО и гомоморфности ПО к ИУС(НСС) следует *изоморфность* ПО и ИУС(НСС), т.е.:

Процессы ПО ↔ ИУС(НСС) (12)

Теорема 2 (творения): минимальная форма словаря ИУС(НСС) и структура процессов ПО *изоморфны*.

Доказательство: будем считать, что ПО и ИУС подобны, – ограниченные физические структуры порождают / принимают неограниченный поток ТФ. Обратим направленность потока из ИУС в ПО и рассмотрим два предельных случая:

1-й – сформированный на минимальном словаре ИУС(НСС), см. рис.19 а;

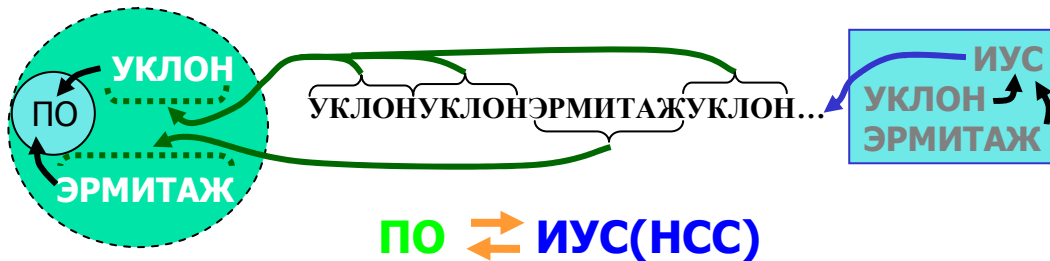


Рис. 19 а.

2-й – сформированный на произвольном словаре ИУС (не минимальном) см. рис. 19 б.



Рис. 19 б.

Различие между случаями на рис 19 а и 19 б очевидно. В случае 19 б обратимость не выполняется. Следовательно, только минимальная форма словаря ИУС(НСС) *изоморфна* структуре процессов ПО.

Изоморфность открывает возможность ИУС строить объективные модели ПО, и, как следствие, подниматься ИУС по эволюционной лестнице, все более и более активно влияя на переустройство процессов ПО в направлении увеличения эволюционного потенциала. Вообще, многочисленные результаты вычислительных экспериментов с моделями ИУС(НСС) и их теоретический анализ позволил сформулировать основные постулаты для эволюцион-

но-перспективных ИУС, но их изложение и анализ предполагается издать отдельной работой.

Изоморфность между материальным (ПО) и идеальным (ИУС(НСС)) мирами связывает их в одно неразрывное гармоничное целое, и у естествоиспытателя появляется чувство не только исследователя, но и творца мира, а, соответственно, и глобальной осмысленности, как существования Вселенной, так и предназначения Человека.

Часть III

Определения основных компонент информационного ресурса. Фактически, в изложенном материале предыдущих частей статьи у нас есть все необходимое для вынесения на обсуждение формализованных определений различных компонент информационного ресурса (*ТФ-сигнал, информация, знание*) в физических единицах измерения системы СБС (см., бит, сек.).

Текстовая форма (ТФ). Это первый *качественный переход* от физических квантово-континуальных объектов-процессов ПО к их возможным дискретным модельным замещениям (образам) в ИУС, дополнительно имеющим еще и нематериальную характеристику информационного ресурса – семантическую (см. рис. 4).

Качество (число разбиений) дискретизации ТФ по амплитуде и по длительности определяется через обратную связь ИУС – ПО: ТФ₁ (входная, состояние ПО) – ТФ₂ (выходная, действие ИУС) – ТФ₃ (оценка "полезности" внутреннего состояния ИУС, как результата реакции ТФ₂ на ТФ₁), см. далее – информация.

Как уже отмечалось, мы предполагаем, что процессы в ПО детерминированы (см. рис. 7). Таким образом, тройка: ТФ₁ – ТФ₂ – ТФ₃, также должна быть детерминирована. При этом, если встречаются ситуации:

$$\langle ab \rangle - \langle dbg \rangle - \langle + \rangle,$$

$$\langle ab \rangle - \langle dbg \rangle - \langle - \rangle,$$

то следует, что исходные ситуации $\langle T\Phi_1 \rangle$ и/или реакции $\langle T\Phi_2 \rangle$ *различны*, и их необходимо различать, посредством увеличения в ИУС числа градаций по амплитуде и/или длительности.

Если же встречаются ситуации:

$$\langle abc \rangle - \langle dbg \rangle - \langle e1 \rangle,$$

$$\langle abd \rangle - \langle dbg \rangle - \langle e1 \rangle,$$

то следует, что исходные ситуации $\langle T\Phi_1 \rangle$, без ущерба для ИУС, можно *загрубить*, уменьшив число градаций по амплитуде и/или длительности.

Для ситуаций:

$$\langle abc \rangle - \langle dklm \rangle - \langle e2 \rangle,$$

$$\langle abc \rangle - \langle dm \rangle - \langle e2 \rangle,$$

$$\langle abc \rangle - \langle dddkkk \rangle - \langle e3 \rangle,$$

$$\langle abc \rangle - \langle dk \rangle - \langle e3 \rangle,$$

следует, что процессы реакций $\langle T\Phi_2 \rangle$, также можно загрузить как по амплитуде, так и по частоте.

Можно привести и другие конструкции троек, но основное, что процесс адаптивной подстройки механизмов дискретизации ИУС при преобразовании процессов ПО в ТФ и обратно ТФ в процессы ПО, определяется лишь двумя правилами: в тройках $(T\Phi_1 - T\Phi_2 - T\Phi_3)$ для ИУС всегда *должен выполнять принцип причинности*, и при этом, суммарный *объем ТФ должен быть минимален* (стремится к минимуму).

Поэтому ТФ, по отношению к исходным объектам-процессам ПО, всегда *менее ресурсоемка*, при потенциальном сохранении всех значимых для ИУС причинно-следственных характеристик исходных объектов-процессов ПО.

Очевидно, что попытки представления информации (ТФ) о каких-либо процессах не в дискретной, а в *континуальной форме* потребуют "бесконечных" ресурсов для отображения (каждый отсчет – 20-40 десятизначным числом, 10^{20-40} рецепторов и эффекторов ИУС), что практически *невыполнимо* в рамках выбранной нами модели Вселенной.

ТФ не существует вне материального носителя, являющегося частью (телом) ИУС. Вне ИУС ТФ становится объектом-процессом (действием, книгой, ...). ТФ инварианта по отношению к физическим (физико-химическим) свойствам своего носителя. Принцип инвариантности определяет возможность многообразного кодирования одних и тех же объектов-процессов ПО. Что позволяет осуществлять коммуникацию ИУС даже различной природы (конструкций). Например, если одна ИУС может воспринимать графическую форму и выдавать информацию только в акустической форме, а другая ИУС наоборот, воспринимает только в акустической форме и может выдавать информацию только в графической форме, то такие ИУС вполне могут вести диалог.

Преобразование процессов ПО в ТФ осуществляется *рецепторами* (сенсорами) ИУС. Обратное преобразование ТФ в процессы (действие) ПО осуществляется *эффекторами* ИУС. Таким образом, ИУС и ПО представляют взаимодействующий комплекс (см. рис. 4), при этом, ИУС может *управлять* процессами в ПО для повышения собственного эволюционного потенциала, выбирая среди "равновероятных" реакций $\langle T\Phi_2 \rangle$, которая будет максимизировать $\langle T\Phi_3 \rangle$ на текущую ситуацию ПО $\langle T\Phi_1 \rangle$.

Дискретную форму, выражаемую последовательностью кодов алфавита ($A = \{a, b, c, \dots, z, \dots\}$) и отображающую с любой необходимой степенью точности изменение физического параметра во времени, будем называть *текстовой формой*. (13)

Понятно, что ТФ характеризуется размером $|A|$ алфавита отображения A , частотой отсчетов (рецепторов и эффекторов) ω и ее длительностью $t = (t_2 -$

t_1+1). Длина ТФ в символах $L = f(t) = (t_2 - t_1 + 1) * \omega$. Объем ТФ измеряется ТФ в битах:

$$\begin{aligned} \text{объем ТФ в битах} &= \lceil \log_2 |A| \rceil * L = \lceil \log_2 |A| \rceil * t * \omega, \\ \text{где: знак } \lceil \rceil &\text{ обозначает – ближайшее большее целое.} \end{aligned} \quad (13a)$$

Таким образом, единицами измерения ТФ являются: *см.* (место от начала символа в алфавите A , – его имя), *сек* (длительность ТФ) и сек^{-1} (частота отсчетов (символов) за одну единицу времени, $\omega = \omega(\text{см} * \text{сек}^{-1})$), и в *битах* измеряется объем ТФ.

В реальных ИУС частота отсчетов может быть неравномерна во времени, также и символам алфавита могут соответствовать различные (не равные) диапазоны амплитуды изменения параметра. Все определяется качеством адаптационных возможностей ИУС.

Сигнал. Сигналом является произвольный участок ТФ по длине соизмеримый с размером алфавита A , т.е. $\forall t_1, t_2: t_2 - t_1 \approx |A|$. Сигнал характеризует особенности ТФ. Например, размер, цвет знаков алфавита A . Сигнал может характеризовать ПО, принадлежность алфавита, но не касается никаких целостных семантических аспектов, касающихся процессов ПО.

Сигнал – произвольный отрезок текстовой формы (14)

Единицами измерения сигнала являются единицы измерения ТФ: *см.*, *сек.* и *бит*. Дополнительными относительными характеристиками сигнала является: длительность ($\text{сек} \sim \text{см.}$) и безразмерная величина текстовой энтропии [7] (от 1 и до 0, 1 – полный хаос, "белый шум" и 0 – полный детерминизм)

Образ. Семантическая единица ИУС, замещающая процесс-объект ПО. Хранится и используется ИУС как неделимое целое (смысловой квант). Самые простые образы (терминальные) – коды алфавита A текстовой формы. Образы имеют иерархическую природу. Сложные образы состоят из более простых образов. Пример образов: буквы, слоги, слова, фразы и т.д.

Образы: k, g, m – это множество пар: $t_1^k t_2^k; t_1^g t_2^g; \dots, t_1^m t_2^m$; где верхний индекс – имя образа, а нижний – время (индекс символа) начала и завершения образа в конкретной ТФ. Если $\exists t_1^k = t_1^g$, или $t_2^m = t_2^d$, то образы: k, g, m, d вложенные, или с перекрытиями. Если $\exists t_1^k > t_1^g$ и $t_2^k < t_2^g$, при $\forall g$, то образ k является уникальным.

Помимо обычных характеристик в единицах измерения ТФ: *см.*, *сек.* и *бит*, у образа появляется "имя", которое измеряется в координатах множества пар и может быть выражено в пространственной координате – в см^* шкалы.

Образ – вычлененная ИУС (НСС) знаковая последовательность ТФ, соответствующая целому числу причинно-связанных процессов ПО (15)

Информация. Общепринято, что информация всегда несет какую-либо законченную смысловую нагрузку. Например, фраза: "идет снег" – информация (т.к. она вызывает мысль, например, "пойти на лыжах" и, как следствие, получить от этой прогулки удовольствие), тогда как: "идет" и "снег", – это образы (т.к. чтобы эти сообщения вызвали у вас реакцию, вам необходимо получить дополнительные сообщения – ответы на ряд вопросов: "кто или что идет?", "причем здесь – снег?", и ряд других, если конечно "идет" и "снег" не являются ответом на предыдущие ваши вопросы).

Информация является прагматической единицей (квантом) ИУС. Информация всегда имеет определенную грамматико-прагматическую структуру: *образы состояния ПО; образы действия ИУС; E_+* , т.е. в тройках (ТФ₁–ТФ₂–ТФ₃, E_+), см. выше ТФ. ИУС всегда настраивается на восприятие ТФ через эту прагматическую структуру (E_+), которая, в свою очередь, может быть представлена как внутренний образ ИУС (ТФ₃). Таким образом, из ТФ выделяется только прагматически окрашенная для ИУС информация.

Особенность конструкции информации заключается в том, что она объединяет себе дискретную часть: ТФ₁, ТФ₂ и аналоговую ТФ₃ или E_+ . E_+ – это своеобразный "эмоциональный" уровень удовлетворенности собственной реакции ТФ₂ на ТФ₁. Эта "эмоциональность" пронзает все процессы ИУС, хоть как-то связанные с полученным результатом. Аналоговая форма E_+ позволяет безадресно (глобально, масштабно) управлять настройкой ИУС на получение ею положительных "эмоций" – E_+ .

Ценность информации всегда оценивается по величине E_+ в данной структуре. Информация содержит образ управления ПО. При отсутствии в информации образа E_+ она впоследствии для ИУС (генетически) игнорируется.

Если есть информация, то должна быть и "не информация". Так если, ТФ не укладывается ни в одну из структур восприятия ИУС, то ТФ не воспринимается (не запоминается) как информация, и относится ИУС либо к образу, либо к сигналу (шуму).

Информация – прагматически причинно-связанная последовательность образов, в виде тройки: <ситуация ПО> <реакция ИУС><оценка>, ориентированная на достижение ИУС(НСС) какого-либо ее целевого образа ПО (например, E_+) (16)

У информации, дополнительно к образу, появляется еще две дополнительные характеристики: структура тройки (ТФ₁–ТФ₂–ТФ₃), для которой нужна аппаратная поддержка НСС в виде N-элементов, и измеряемая величина "прагматичности", которая может определяться как образами лингвисти-

ческой переменной: "хорошо", "средне", ... , так и просто положением на "шкале прагматичности" ИУС – единицами см* шкалы.

Знание. В литературе определения понятия "знания" рассматривается как продукт общественной материальной и духовной деятельности людей; идеальное выражение в знаковой форме объективных свойств и связей мира, природного и человеческого. В знании кристаллизуется, преемственно накапливается и объективизируется общественная сила человека. В знании осуществляется перевод разрозненных представлений в теоретически систематизированную, общезначимую форму, удержание того, что может быть сохранено, передано, преемственно развито в качестве устойчивой опоры последующей человеческой деятельности [2]. Знание – это отражение реального мира в сознании. Знания в информатике – это вид информации, отражающей опыт специалиста (эксперта) в определенной предметной области, его понимание множества текущих ситуаций и способы перехода от одного описания объекта к другому [3]. Как мы видим, данные определения понятия "знания", понятные нам в контексте нашей культуры, но формализовать и сделать их понятными инженеру и доступными компьютеру, т.е. формализовать, очень затруднительно. Тем более, что видов "знаний" как и традиционных видов "информации" (см. (1)) великое множество. Это и "декларативные", "процедурные", "научные", "эвристические", "метазнания", "знания о предметной области" и др.

Знание – эта категория информационного ресурса, требует дополнительных инструментальных особенностей ИУС, которые в данной статье не рассматриваются и которые вполне могут сформироваться у ИУС естественным эволюционным путем. Знание – значительно более сложная категория, чем информация. Так, если в формировании информации достаточно присутствия образов, относящихся только к ПО, то *в структуру знания* необходимо включить еще и *образы состояния самой ИУС*, возникающие в процессе переработки образов этого текущего состояния ПО.

Текстовые последовательности, как конкретные примеры, относящиеся к одному типу (классу) причинно-следственных процессов ("знанию"), вызывают характерные и однотипные процессы (образы) состояния самой НСС. Формирование (связывание) в единую функциональную подструктуру образов (N-элементов), относящихся непосредственно к ПО и образов состояния НСС, вызванных обработкой ТФ из внешней среды (ПО) – это и есть физическое представление "образа знания". Таким образом, одноранговое включение в НСС *функции рекурсии*, позволяет получить N-элемент, выполняющий "функцию знания", а именно, реагирующий на весь класс примеров, относящихся к какому либо правилу причинно-следственных процессов ПО для различных объектов, т.е. вычислять семантические закономерности ПО.

Знание – структура ИУС, включающая в себя, причинно-связанное объединение информации (образов) о ПО и текущей информации (образов) о внутреннем состоянии самой ИУС(НСС), вызванном анализируемыми процессами ПО. (17)

В знании, ко всем физическим характеристикам информации (см. 16), добавляются еще образы, состояния самой НСС, вызываемые обработкой текущей информации. Для вовлечения этих образов в процессы обработки НСС одноранговой рекурсии (интроскопии), возникает необходимость в дополнительных структурах, которые также реализуются N-элементами.

Знание более *контекстно* (зависимо от пространственно временного окружения процессов) в рамках ПО и ИУС в сравнении с информацией и механизм его настройки более сложен. Более подробно эти вопросы будут рассмотрены отдельно, здесь мы лишь приводим определение.

Таким образом, мы дали формализованные определения понятиям информационного ресурса: "текстовая форма", "сигнал", "образ", "информация" и "знания" с позиции нейросемантики. Другие понятия информационного ресурса, такие как сообщение, данные и др., могут быть легко выведены из вышеприведенных "корневых" определений.

Помимо четкого формального различия в данных нами определениях, компоненты информационного ресурса имеют качественные и функциональные физические различия.

Так, если представить, что у нас есть три типа ПО: случайная ("белый шум" – БШ), периодическая (колебательные физические процессы, телесериал), реальная (излагающая систематизированное знание), и также три типа ИУС: автомат ("техническая ИС"), животное ("адаптивная ИУС"), человек ("разумная ИУС").

Теперь представим, что сначала процессами в ПО являются "белым шумом" (БШ). Так как это случайные процессы, то ни "техническая ИС", ни "адаптивная ИУС", ни "разумная ИУС", ничего более чем случайного сигнала (образов), при формировании структуры словарей (см. рис. 17) из этого потока выявить не смогут, см. рис. 20.

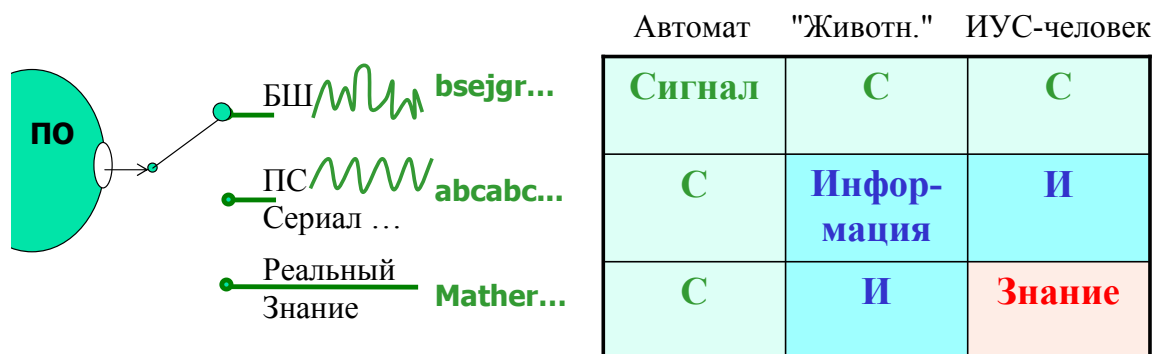


Рис. 20.

Если же представить, что в ПО будут происходить некоторые априорно неизвестные периодические процессы, то "техническая ИС" (например, бухгалтерская программа "С1"), не настроенная на них, будет воспринимать их как сигнал. "Адаптивная ИУС" – найдет некоторую регулярность или период в этом потоке ТФ, и образ этой регулярности будет рассматриваться как информация из данной ТФ. "Разумная ИУС" из этого потока также более чем регулярности, т.е. образа или информации, выявить не сможет.

Теперь представим, что процессы в ПО – это интеллектуальный процесс, например, передача "В мире животных". "Техническая ИС" более чем сигнала из данной ПО не почерпнет. "Адаптивная ИУС" выявит только иерархию различных образов (информацию), а "разумная ИУС" из этого потока выявит семантические закономерности (например, различные классы, отряды животных и т.д.), т.е. выявит знание.

Выводы из этого примера (уже наши с вами семантические закономерности) заключаются в том, что ПО и ИУС тесно взаимосвязаны. При этом, *ИУС не может выделить из ТФ более того, что в ней есть, и ИУС не может выявить из ТФ более того, что позволяет ее конструкция (информационная архитектура)*, см. рис. 20.

Касательно определенных нами компонент информационного ресурса, рассмотрим еще один параметр физических отличий, который заключается в существенно различных объемах памяти, необходимых для полного отображения в ИУС (НСС) постоянного потока ТФ ($TФ(t)=const$), попеременно состоящего из: сигнала, информации, или знания.

Будем предполагать ТФ равномерным потоком, отображающим процессы ПО: шумовые, информационные и интеллектуальные. При этом под ИУС будем рассматривать как одну из "разумных ИУС", см. рис. 20.

Рассмотрим *первый случай*. Понятно, что отобразить шумовую ТФ без потерь можно только полной ее записью (например, на магнитную ленту). Структуризация априорно неопределенного шумовой ТФ невозможна никакой ИУС. Теоретически и экспериментально показано, что соотношение (10) в этом случае не зависит от длины ТФ и не превышает 1/3. Таким образом, если затраты ИУС (НСС) необходимой памяти $R_{ИУС}$ под ТФ будут характеризоваться линейной функцией: $R_{ИУС} = k*t + b_1$, где, k и b_1 некоторые константы, то, при удовлетворении критерия достаточности ТФ (11) будет следовать, что ТФ является принципиально неструктурируемой, т.е. "белым шумом", см. рис. 21а.

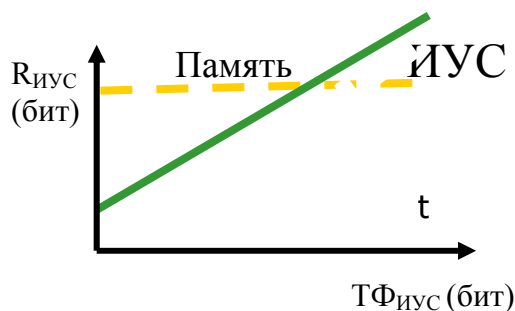


Рис. 21а

Пример ТФ в форме "белого шума" есть, конечно, теоретическая идеализация. В реальной Вселенной причинно-связанных процессов линейный рост затрат на память ИУС(НСС) может говорить: – либо о недостаточности объема введенной ТФ (невыполнение условия 11); – либо о недостаточности в точности дискретизации по ΔW или Δt (см. рис. 6); – либо вообще об упущении или не рассмотрении некоторых из проекций физических процессов; – либо о временной конструктивной (эволюционной) неготовности ИУС(НСС) к взаимодействию с данными ПО.

Поэтому, получив экспериментально линейный рост потребляемого ресурса памяти ($R_{ИУС}$) в реальных физических ПО, нужно проанализировать возможные вышеназванные причины и, только в случае невозможности устранения причин, считать ТФ "шумовой".

Разберем *второй случай*. Будем считать, что ТФ – это радиопередача на неизвестном языке некоторого сериала типа: "про дона Педро ...", в котором встречаются одинаковые фрагменты ТФ, но между ними нет устойчивых причинно-следственных связей. Это типичная информационная ТФ. За счет того, что ссылки (в битах) на повторяющиеся образы в НСС существенно меньше, чем сами образы (в битах), то величина компрессии (10) для такой ТФ будет иметь *монотонный рост, близкий к логарифмической функции*: $R_{ИУС} = \log(t) + b_2$, где, b_2 некоторая константа, см. рис. 21б.

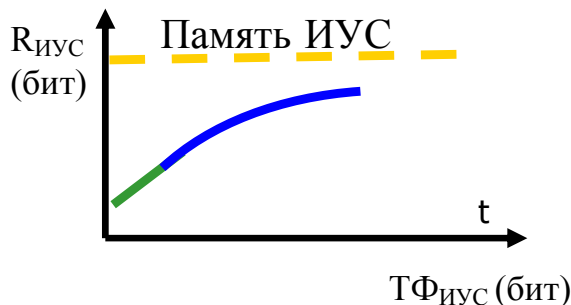
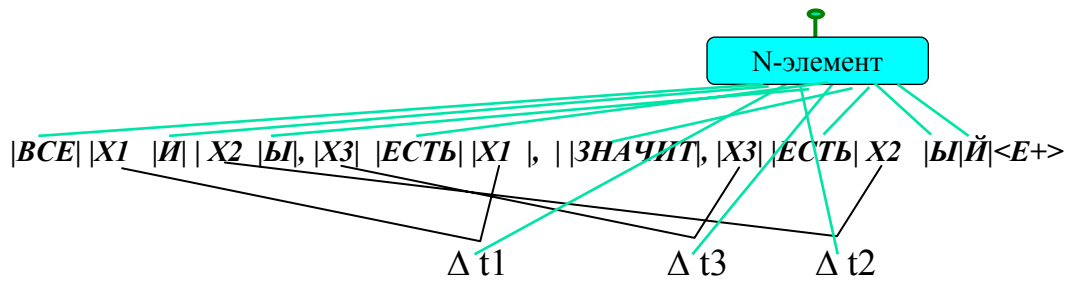


Рис. 21б

Если в результате эксперимента мы получим монотонный рост компрессии в соответствии с (10), то можно сделать вывод о том, что ТФ является потоком образов или информации. Если же нам изначально известно, что ТФ точно является информационным потоком, но мы не получаем компрессии, то естественен вывод о "некачественности" алгоритмов формирования НСС.

В третьем случае будем преобразовывать в НСС текстовый поток, сформированный взаимоувязанными причинно-связанными процессами ПО. В этом случае у ИУС (НСС) определенной архитектуры появляется возможность не обременять память конкретными фактами (образами), а только один раз построить и запомнить правило, причинно связывающее образы. Например, на основании двух информационных отрезков ТФ: $\langle |ВСЕ||ВОЛК|И||СЕРЫ| |А||ЕСТЬ||ВОЛК|,||ЗНАЧИТ|,|А||ЕСТЬ|СЕРЫЙ| \ E+\rangle$ $\langle |ВСЕ||ПЕТУХ|И||КРАСИВЫ|В||ЕСТЬ$

||ПЕТУХ|,||ЗНАЧИТ|,|В||ЕСТЬ|КРАСИВ|ЫЙ| E+>, логически формируется завершенная подструктура, см. рис. 21в1:



где: Δt_i – "ритмические" образы внутреннего состояния ИУС(НСС₍₂₎),
 X_i – "повторно" активирующиеся образы ПО,
 знак "|" - обозначает найденные ИУС границы между образами.

Рис. 21в1

Если же затем на вход ИУС поступает неполная фраза, например, <ВСЕ ПЛЮКИ КАНЫ, С ЕСТЬ ПЛЮК, ЗНАЧИТ, С ЕСТЬ >, то не найдя ассоциативно близкого N-элемента (см. рис. 17) в информационной области НСС, ИУС включается на поиск ассоциаций в области семантических закономерностей НСС₍₂₎. Здесь ассоциативно активизируется N-элемент, отображающий структуру, приведенную на рис. 21в1. В результате, на выходе ИУС будет сформирована ТФ – <КАНЬЙ>, как итог активации образов ИУС и ПО: $\Delta t_2 + Ы|Й$ [7,9,10].

Понятно, что если формируются N-элементы, аналогичные показанному на рис. 21в1, то у ИУС отпадает дальнейшая необходимость в записи в НСС информационных отрезков ТФ, которые перекрываются образами семантических закономерностей в НСС₍₂₎. Таким образом, информационные отрезки ТФ, на которых сформированы семантические закономерности, в память ИУС(НСС) не записываются. И затраты ИУС(НСС) необходимой памяти $R_{ИУС}$ под такие ТФ будут нулевые, а функция $R_{ИУС} = \text{const}(t)$, см. рис. 21в2.

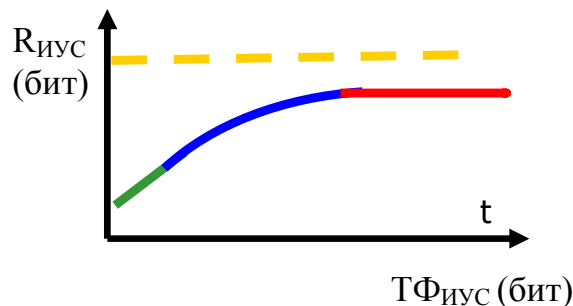


Рис. 21в2.

Таким образом, по экспериментальным данным, соответствующим функции $R_{ИУС} = \text{const}(t)$, можно утверждать, что данный текстовый поток из ПО содержит информационные ТФ, на которых строятся в ИУС подструктуры знания, состоящие из образов ПО и образов ИУС(НСС).

Таким образом, судя по характеру функции $R_{ИУС}(t)$ можно объективно классифицировать процессы ПО на: "шумовые", "информационные" и "интеллектуальные" ("разумные").

Отметим, что с точки зрения "операционной системы" ИУС(НСС) оперирование различными элементами информационного ресурса "сигналом", "образом", "информацией" и "знанием" осуществляется различными типами образов-операторов НСС (N-элементов). Принцип действия всех этих образов-операторов подобен, различие лишь в том, что они включаются на разных этапах обработки ТФ. Так, первоначально ТФ обрабатывается "образом распознавания", далее включается "образ оценки-действия" и затем "образ оператора знания".

Информационный ресурс всегда связан с материальным носителем (бумага, камень, акустическая и электромагнитная волны, состояния молекул и различных полей и т. д.). В данной работе мы рассмотрели только статические особенности информационного ресурса, т.е. наличие или отсутствия той или иной категории информационного ресурса (сигнала, информации или знания). С другой стороны, проявление всех информационных категорий это всегда переходный процесс, имеющий сигнальную форму – "есть" или "нет" в данном отрезке ТФ "сигнал" или "информация", т.е. категория "информация" имеет двойственную характеристику: потенциальное ее наличие или отсутствие определяется в статической форме ее носителя, а само проявление "информации" возможно только в процессе взаимодействия ее носителя с ИУС.

Еще одна особенность, которую не хотелось бы упустить, заключается в том, что одна единица "знания" соответствует целому множеству конкретных единиц "информации", т.е. целому классу.

Количественные характеристики компонент информационного ресурса в системе СБС. Во всех определениях компонент информационного ресурса (ТФ, сигнал, образ, информация, знание) мы использовали как физические характеристики, см., бит, сек., так и строили специальные шкалы ("прагматичности" и др.) на которых пространственным кодированием, например, в см. По нашему мнению, их вполне достаточно, чтобы объективно и формально отображать все процессы, связанные с функционированием НСС и точной верификаций различных компонент информационного ресурса.

Стоит отметить, что уровень "математического образования" Природы (Вселенной), до появления homo-sapiens, это не более чем арифметические операции: сложение, вычитание, умножение и деление, а также простая интеграция, как суммирующее накопление и простое дифференцирование величин параметров процессов в НСС. Но арсенала этих простых операций, при их множественности и автономности в функционировании элементной базе НСС (N-элементы), вполне достаточно чтобы породить знание и, соответственно, разум.

Измерение объема информации можно точно проводить при вводе ее в виде ТФ. В НСС и знании НСС₍₂₎ количественное измерение информации зависит от ее контекста. Допустим, что у нас существует сформированная

висит от ее контекста. Допустим, что у нас существует сформированная НСС на некоторой ПО. Теперь если в нее ввести ТФ содержащую определенную информацию, то потребуется определенный ресурс ($R_{иус}$) памяти НСС (в битах) для этого, и количество это ресурса будет существенно зависеть от контекста водимой ТФ и ТФ описывающей ПО. Поэтому, точно измерять можно в битах только сигнал (ТФ), а не информацию (или же информацию в определенных состояниях, например, "гиперпериода", когда в ТФ отображено все пространство всевозможных причинно-следственных процессов ПО, при этом достигается особые состояния "самоописания НСС" [7,10].

Заключение. Формализовать активно используемые современной цивилизацией такие понятия информационного ресурса как: "сигнал", "информация", "знание" жизненно необходимо, особенно в "информационную эпоху".

Это необходимо для формирования основ научного мировоззрения. Прежде всего, формирования представлений об информационных процессах как одних из основополагающих понятий науки: пространства-времени, вещества, энергии и информации, на основе которых строится современная научная картина мира. Также необходимо рассмотрение единых информационных принципов строения и функционирования саморазвивающихся ИУС различной природы. И, в конечном счете, в формировании общенаучных и общекультурных навыков работы с информацией, хотя бы на уровне, как мы уже работаем с категориями времени, пространства и энергии.

Первоначально физическая категория "времени" способствовала порождению понятия причинно-связанных процессов у простейших биологических информационно-управляющих систем. Активное использование категории "пространство" первобытными людьми позволило открыть им Землю. Инженерное использование категории "энергии" дало человеку пропуск во Вселенную. Осмысленное и нравственное применение категории "информация" откроет Человеку путь в духовный мир (Мироздание) [8,9].

На каждом из этих четырех эволюционных этапов высокоорганизованные формы материи (ИУС) многократно увеличивали динамику своего эволюционного потенциала. И если на предыдущем, "энергетическом" этапе человек научился направлять на свои нужды, как великое достижение, несколько долей процента потенциальной энергии материи (например, при термоядерном синтезе), то на "информационном" этапе фактически вся материя Вселенной становится подвластна Разуму человека.

Из этого следует вся важность и ответственность в скорейшем наведении порядка нашего научного понимания сути информационных процессов. И сложности на этом пути имеют вполне объективный характер. Это и психофизическая ограниченность человека (см. рис. 1), а также и отсутствие приборного арсенала, измеряющего "величины информационности", кроме как человеческого сознания, которое до сих пор для науки также остается Великой тайной. Эту сложность хорошо иллюстрируют слова А. Эйнштейна: *"В нашем стремлении понять реальность мы подобны человеку, который хочет понять механизм закрытых часов. Он видит циферблат и движущиеся"*

ся стрелки, даже слышит тиканье, но не имеет средств открыть их. Если он остроумен, он может нарисовать себе картину механизма, которая отвечала бы всему, что он наблюдает, но он никогда не может быть вполне уверен в том, что его картина единственная, которая могла бы объяснить его наблюдения. Он никогда не будет в состоянии сравнить свою картину с реальным механизмом, и он не может даже представить себе возможность и смысл такого сравнения” [11].

Изначальная цель данной статьи, заключалась в том, чтобы через аксиоматически простые понятия дать формализованные определения основным компонентам информационного ресурса, но необходимость введения понятия автоструктуризации и обоснования аксиоматической базы, а также и планируемое расширение дальнейшей формализации основ информатики и управления (на базе нейросемантического подхода) вынудили существенно расширить изложение материала. Фактически данный материал постоянно просится в формат монографии, причем не обычной линейной формы, а в форме гипертекстовых книг на машинных носителях [19]. За все утверждения данной статьи полностью отвечает автор, все о чем заявлено, будет изложено в следующих статьях и монографии, возможно, в гипертекстовой форме. Я приглашаю всех к участию в обсуждении сущности информационного ресурса и основ информатики с позиции эволюции высокоорганизованной материи во Вселенной.

Список использованной литературы

1. Урсул А.Д. Природа информации. Философские очерки // М.: Политиздат, 1968. 288с.
2. Философский энциклопедический словарь. – М.: Советская энциклопедия, 1983, стр. 217-218
3. Философский словарь /под ред. И.Т.Фролова – 4-е изд. -М.: Политиздат, 1981,-445 с.
4. Шеннон К. Работы по теории информации и кибернетике. – М.: 1963
5. Эшби У.Р. Введение в кибернетику. М., ИЛ, 1959.
6. Ракитов А.И. Философия компьютерной революции / М.: Политиздат, 1991. 287 с.
7. Бодякин В.И. Куда идешь, человек? (Основы эволюциологии. Информационный подход). – М. СИНТЕГ, 1998, 332с. <http://www.ipu.ru/stran/bod/monograf.htm>
8. Бодякин В.И. Направленность эволюции информационных систем. //Всероссийский симпозиум "Современные проблемы неравновесной термодинамики и эволюции сложных систем", тезисы, 15-16 с, М., МАКС Пресс, 2004,124 с.
9. Бодякин В.И. Проект "Искусственный разум", <http://www.iformograd.narod.ru>
10. Бодякин В.И. "Исследование структурных моделей открытых динамических систем", специальность: 05.13.01 (управление в технических системах), автореферат и диссертация на соискание ученой степени к.ф.-м.н., М. 1999 г.
11. Эйнштейн А., Инфельд Л. Эволюция физики / Пер. с англ. М., 1966. С. 30.
12. Эбелинг В., Энгель А., Файстель Р. Физика процессов эволюции. – М.: Эдиториал УРСС, 2001
13. Вайнберг С. Первые три минуты. М., 1981
14. Урсул А.Д. Проблема информации в современной науке. Философские очерки / М.: Наука. 1975. 288с
15. Уосермен Ф. Нейрокомпьютерная техника. М.: Мир. 1992. – 240 с.
16. Нейман Дж. Теория самовоспроизводящихся автоматов. – М.: Мир, 1971.

17. Редько В.Г. Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики: Кибернетика. – М: КомКнига, 2006. 224
18. Редько В.Г. Эволюционная кибернетика. – М.: Наука, 2001.
19. Эпштейн В.Л. Гипертекст – новая парадигма информатики. ИПУ РАН им. В.А.Трапезникова, М.2006г.

Замечания

