

МОДЕЛЬ «КОННЕКТОМА» ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ

УДК 519+61:681.3

ПРОКОПЧУК Юрий Александрович

к.ф.-м.н., доцент, с.н.с. отдела системного анализа и проблем управления Института технической механики НАНУ и ГКАУ,
зав. центром телемедицины Днепропетровского областного центра кардиологии и кардиохирургии

Научные интересы: интеллектуальные и когнитивные системы, базы знаний

E-mail: itk3@ukr.net

ВВЕДЕНИЕ

Создание биологически инспирированных интеллектуальных систем (когнитивных систем) предполагает моделирование и реализацию, в частности, следующих важнейших свойств биологической памяти (мозга): порождение сверхизбыточности и сверхразнообразия, авто/гетеро-ассоциативность, продуктивность, когерентность, самоорганизованную критичность, сверхпараллельность, грубость.

Во вступительной статье к переводу классической монографии Вертгеймера (Max Wertheimer) [1] В.П. Зинченко отмечает: «Из описаний продуктивного мышления Вертгеймера следует, что главным в этом процессе являются не столько операционально-технические процедуры, направленные на решение уже сформулированной задачи, сколько сама формулировка задачи, постановка проблемы. Именно на этой стороне мыслительного процесса должно быть сконцентрировано внимание исследователей».

Термин «коннектом» (connectome) родился в 2005 году. Им обозначили полное описание структуры связей в нервной системе человека (по аналогии с термином «геном») [1 - 4]. С тех пор коннектомикой называют компьютерный анализ строения естественных нейронных сетей, своего рода картографирование нейронных связей. Проект «Коннектом человека» (Human Connectome Project) стал самым масштабным американским проектом последних лет [5]. Анонсируя проект, Барак Обама заявил: «Настало время выйти на уровень научных исследований и разработок, невиданный с момента пика космической

гонки...». Аналогичный мегапроект реализует и Европейский союз. В Швейцарии неподалеку от Лозанны строится Нейрополис — аналог CERN и Кремниевой долины для исследования мозга. Нейрополис станет центром мегапроекта «Human Brain Project», в котором уже участвуют 120 команд ученых из 90 институтов 22 стран.

Коннектом объясняет, почему мозг обладает такой потрясающей автоассоциативностью. Многие когнитивные функции есть производные от его сетевой топологии. Концепция коннектома позволяет выявить ключевые свойства естественных нервных сетей, важные для создания искусственных когнитивных систем (ИКС), включая агентные и интеллектуальные среды. Возникает проблема инженерии искусственных «коннектомов» или «ИКС-коннектомов».

Исследуя механизмы памяти, невозможно обойти проблему консолидации фрагментов, зафиксированных в пространственно-временном континууме мозга. Такая консолидация в рамках целостного образа, по-видимому, обусловлена наличием функциональных контактов (ассоциативностью) фрагментов. Коммутационная концепция, объясняющая обучение и возникновение ассоциаций только за счет формирования (или трансформации) контактов между нейронными структурами, уступает место теории, основанной на системно-информационном подходе [6]. Соответственно, субстратом ассоциативной памяти служит не просто комплекс контактов нейронов, а система тех фрагментов, которые входят в состав ряда образов. Консолидация фрагментов памяти в процессе принятия решения происходит под контролем регулято-

ров, которые ответственны за целостность взаимодействия и ассоциативность образов [1 - 9].

Различные концептуальные наброски ассоциативной памяти можно строить в рамках парадигмы предельных обобщений (ППО) [7, 8]. ППО – это методология системной реконструкции концептуальных набросков когнитивной сферы на основе синергетической концепции критичности и ее применение для создания интеллектуальных приложений. ППО является конкретизацией парадигмы сложности применительно к ментальной сфере. ППО направлена на поиск возможной природы и фундаментальных закономерностей информационного синтеза в процессе познания, восприятия и опыта, а не на исследование частного вида ее реализации в специфических структурах головного мозга человека и животных. Этот частный вид реализации может быть интересен, но должен рассматриваться именно как один из возможных вариантов реализации некоторых более общих и фундаментальных механизмов переработки информации. Структура очень важна, она многое подсказывает, и в некоторых случаях ее надо заимствовать у природы, но гораздо важнее заимствовать "принцип биологической логики", т.е. то движущее начало, которое заставляет работать структуры и создает их.

Себастьян Сеунг (Sebastian Seung) в работе [3] предложил научную метафору: «I Am My Connectome». Перефразируя ее для ИКС, можно сказать, что «ИКС определяется своим коннектомом». Такая трактовка фундамента ИКС обосновывает актуальность настоящей работы.

ПОСТАНОВКА ЗАДАЧ ИССЛЕДОВАНИЯ

Основными задачами исследования являются разработка концепции «коннектома» интеллектуальной систе-

мы (Ag-connectome) и его программная реализация на основе парадигмы предельных обобщений.

ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

В рамках ППО предполагается, что произвольную ситуацию действительности (прецедент, систему, процесс) можно описать с помощью множества элементарных тестов (квалиа), представимых в виде «тест = значение». Результаты любого теста τ могут выбираться из разных доменов T (множеств значений), которые образуют оргграф доменов $G(\tau) = \{T \rightarrow T\}_\tau$. В совокупности оргграфы формируют *Банк тестов* $\{G(\tau)\}$. Для решения той или иной когнитивной задачи (Z -задачи) формируется множество прецедентов с известными исходами $\Omega = \{\alpha(\{\tau/T, z/Z\})$, где $Z = \{1, \dots, N\}$ – множество заключений (различий, диагнозов, прогнозов, управлений); $\{\tau/T\}$ – множество значений тестов. Множество Z является одним из доменов оргграфа $G(z)$. Z -задачи вместе с другими тестами формализуют акты различения (дифференциации). *Акт различения* – это системоквант «мыследействия» когнитивной системы, базовая функция наблюдателя [7, 8].

Оргграфы доменов тестов $G(\tau)$ являются специальным представлением оргграфов значений тестов $G(\tau)$. Последние реализуют «физику» ИКС – первичные и вторичные каналы передачи информации (рис. 1). «Физика» ИКС в виде $\{G(\tau)\}$ является аналогом биологических нейронных сетей. Назначение первичных сетей – передача сигналов от сенсориума к верхним отделам обработки информации («мозгу» ИКС). Вторичные сети моделируют квалиа.

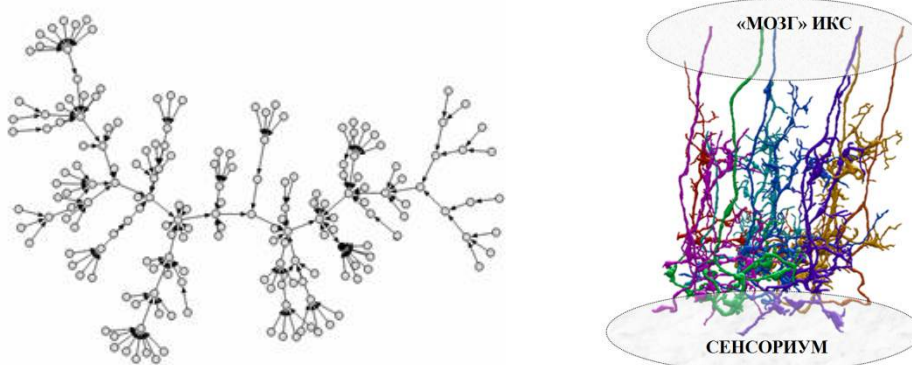
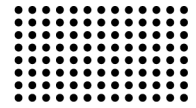
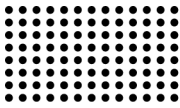


Рис. 1. – «Физика» ИКС: оргграфы значений тестов



С помощью орграфов значений/доменов формируются оргграфы набросков образов $Gs(W) = \{P \rightarrow_e P'\}_{W}$, где W – образ, P, P' – наброски. В частности, строятся наброски ситуаций из базы прецедентов $\Omega = \{Gs(\alpha)\}$. Оргграфы набросков, в свою очередь, порождают оргграфы значений, что является ничем иным, как *возникновением нового теста, qualia* (структурогенез вторичного теста).

Решением любой Z-задачи является базис собственных форм (предельных моделей знаний) $\{\{S^*\}_{Min}\}_{Full}$. Каждая форма $\{S^*\}_{Min}$ состоит из предельных синдромов $\{S^*\}$ – неизбыточных и максимально обобщенных идеальных закономерностей, позволяющих однозначно установить заключение, в частности, для всех ситуаций из базы прецедентов Ω . Собственные формы являются инвариантами когнитивной эволюции (внутренними кодами) [7]. Воплощение собственных форм приводит к возникновению функциональных систем и критических путей (КП) на их основе. Связка «собственная форма (СФ) – функциональная система (ФС)» является инвариантом когнитивно-моторной категоризации. Другими словами, связки «СФ – ФС» являются *элементарными единицами опыта*, кодирующими соотношение целого организма (или ИКС) с теми или иными аспектами мира.

Упрощенно этапы субъективной категоризации, возникновения и воплощения в рамках произвольной Z-задачи изобразим в виде следующей схемы (внутрисистемная консолидация памяти) [7]:

$$\begin{aligned} \Omega(\{ \tau/T_0, Z \}, \{G(\tau)\} \vee \{Gv(\tau)\} \rightarrow \{Gs(\alpha)\} \rightarrow \{R\}_{Full}, \\ \{V\}_{Full} \rightarrow \{S\}_{Full} \rightarrow \{R^*\}_{Full}, \{S^*\}_{Full} \rightarrow \\ \rightarrow \{\{S^*\}_{Min}\}_{Full} \rightarrow \{\{S^*\}_{Min}\}^* \rightarrow \{ФС\} \rightarrow \{КП\}, \quad E_Z \geq 0. \end{aligned} \quad (1)$$

где $\{R\}, \{R^*\}$ – вероятностные закономерности или предвестники; $\{V\}$ – идеальные закономерности; $\{S\}$ – синдромы (неизбыточные идеальные закономерности); $\{\{S^*\}_{Min}\}^*$ – критический путь в рамках базиса СФ (наиболее востребованные СФ). Схема обязана первично-воплощенным инвариантам «собственные функции И-среды», а также информационному/моторному напряжению E_Z (параметру порядка эволюции). Она напрямую связывает память с задачей о минимизации энергии функционирования. Если на любом этапе эволюции информационного/моторного напряжения недостаточно ($E_Z = 0$), то дальнейшая эволюция останавливается и может даже начаться регресс (забывание).

Переход $\{\{S^*\}_{Min}\}^* \rightarrow \{ФС\}$ символизирует *логику исполнения* или *воплощение* знаниевых структур. Воплотиться могут только стабильные и максимально компактные модели знаний. Такими моделями и являются собственные формы (это обстоятельство рассматривается как определение собственных форм). Воплощенные модели знаний (функциональные системы) активируются рефлексивно, что резко снижает затраты энергии. Именно стадия воплощения характеризует наивысший профессиональный опыт (приобретение навыков). Если модели знаний не воплощены, то их использование (активация) требует значительной «психической» энергии (ресурсов), которой в нужный момент может и не быть, что препятствует *пониманию* ситуации (выделению параметров порядка). Кроме того, такие модели знаний быстро деградируют (забываются, наступает амнезия).

Схема (1) показывает, как функциональные системы возникают в эволюции, развитии и при обучении. Она описывает системогенез как общую закономерность эволюционного процесса. Взаимопревращение между оргграфами набросков и оргграфами значений приводит к *интеллектуальной эволюции* или спиральной когнитивной метадинамике [7, 9].

Покажем, как введенные выше базовые структуры и процессы формируют «коннектом» ИКС (интеллектуальной системы).

Поскольку любой дискретный домен любого теста образует Z-задачу различения, для которой имплицитно формируется базис собственных форм (моделей знаний), то оргграф доменов теста τ можно представить в виде:

$$G(\tau) = \{T\{\{V\}_\tau\} \rightarrow_e T' \{\{V\}_{T'}\}\}_\tau \quad (2)$$

где $\{\{V\}_\tau\}$ – инвариантный базис моделей знаний T -задачи, основанный на всех тестах (квалиа) банка $\{G(\tau)\}$ и базе прецедентов $\Omega(\alpha\{\tau, T\})$. Примеры лингвистических оргграфов доменов (из медицины):

Уровень гемоглобина \wedge Hb (Пол?) {

D4 {норма \wedge 0; отклонение \wedge 1 4} $\{\{S\}_{D4}\}$

D3 {анемия \wedge 1 2 3; норма \wedge 0; повышен \wedge 4 5 6} $\{\{S\}_{D3}\}$

D2 {Резко выраж. анемия \wedge 3 М [80; 110] Ж [80; 102];

Выраж. анемия \wedge 2 М [110; 120] Ж [102; 110]; Умеренная анемия \wedge 1 М [120; 133] Ж [110; 114]; Норм.уровень Hb \wedge 0 М [133; 185] Ж [114; 168]; Уровень Hb умеренно пов. \wedge 4 М [185; 190] Ж [168; 170]; Повыш.уровень Hb \wedge 5 М [190;

200] Ж (170; 172]; Резко повыш. уровень Hb ^6 M (200; 250] Ж (172; 250] $\{\{S\}_{D2}\}$

D1 $\{\{80; 250\}\}$;

$G(Hb) = \{D1 \rightarrow D2 \rightarrow D3 \rightarrow D4\}$;

Степень выраженности воспалительного процесса ^{CBVP}

D2 {в.п. отсутствует ^1; в.п. имеется ^2 3 4} $\{\{V\}_{D2}\}$

D1 {в.п. отсутствует ^1; незначительная ^2; средняя ^3; выраженная ^4} $\{\{S\}_{D1}\}$;

$G(CBVP) = \{D1 \rightarrow D2\}$.

Тест «CBVP» определяется на основе анализа крови, включающего тест «Hb», и является вторичным тестом. Инвариантные модели знаний $\{\{S\}_{D2}\}$, $\{\{S\}_{D3}\}$, $\{\{S\}_{D4}\}$ теста «Hb» и $\{\{V\}_{D2}\}$ и $\{\{S\}_{D1}\}$ теста «CBVP» формируются автоматически и имплицитно на основе процессов авто/гетероассоциативности в рамках банка тестов $\{G(\tau)\}$ и базы прецедентов Ω .

Инженерия орграфов доменов тестов включает в себя построение структурно-завершенных орграфов доменов $G^{++}(\tau)$ [8]. Орграфы $G^{++}(\tau)$ образуются из орграфов $G(\tau)$ путем выполнения операции расщепления вершин: для любой дискретной вершины $D = \{a_1, a_2, \dots, a_n\}$ такой, что $n > 2$, порождаются n терминальных вершин (листьев) вида $\{a_j; не a_j\} \equiv \{a_j; \neg a_j\} \equiv \{a_j; D \setminus a_j\}$, где $j = 1, \dots, n$. Все совпадающие вершины-листья объединяются. Структурно-завершенные орграфы иллюстрируют автоматическую генерацию новых Z-задач-доменов, т.е. продуктивность и креативность когнитивной системы [1]. На рис. 2 показаны структурно-завершенные орграфы доменов для первичного теста «Hb» и вторичного теста «CBVP» (нижняя вершина – базовый домен).

Фиксация орграфа доменов $G(\tau)$ означает автоматическую постановку $|G(\tau)|$ задач (кроме непрерывного домена). Решение каждой задачи строится в соответствии со схемой (1). Вместе с автоассоциативностью банк тестов $\{G(\tau)\}$ формирует транзитивно-индукторное пространство или основу ИКС-коннектама.

Полный ИКС-коннектом формируется с учетом орграфов набросков образов $G_S(W)$ или в развернутом виде:

$$G_S(W) = \{P \{\{V\}_P\} \rightarrow_e P' \{\{V\}_{P'}\}\}_W, \quad (3)$$

где P, P' – наброски образа W , $\{\{V\}_P\}$ – инвариантные модели знаний (базисы собственных форм), а также иерархий-гетерархий собственных форм.

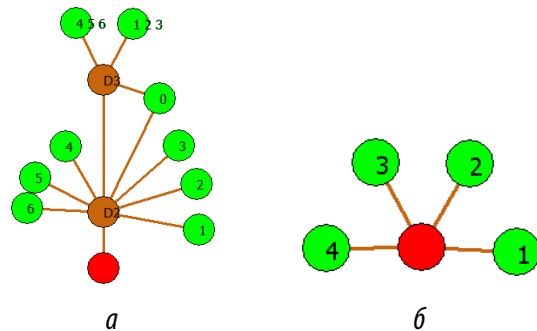


Рис.2. – G^{++} – Структурно-завершенные орграфы доменов а - теста «Hb»; б – теста «CBVP»

Опишем процесс построения фрагмента «коннектама» на основе банка тестов $\{G(\tau)\}$. Пусть имеется кадр данных $\{\tau/T_0\}$ и банк тестов $\{G(\tau)\}$. Без потери общности примем, что каждый тест входит в кадр не более одного раза. Поток данных представляет собой текущую совокупность всех кадров $\cup_i \{\tau/T_0\}_i$. Пример потока данных в рамках госпитальной системы (ГИС) показан на рис. 3.

Дата	Время	Документ	Составитель
04.02.15	15:46	План обследования	Калыц О.В.
05.02.15	09:00	Общий анализ мочи	Орлановский
	10:28	Дневник	Калыц О.В.
	13:26	Клиническое	Врангелер
	14:44	Печеночный комплекс	Врангелер
	15:50	Печеночный комплекс	Врангелер
	15:50	Липидный комплекс	Врангелер
	15:50	Электролиты и микроэлементы	Врангелер
	12:00	Глюкоза крови	Врангелер
	14:23	Общий анализ крови	Врангелер
06.02.15	09:52	Дневник	Яременко Е.В.
	11:08	ЭИТ	Карпова Г.А.
	16:50	Дневник	Калыц О.В.
	19:45	Дневник дежурного врача	Мисюк-Степан Д.
27.02.15	15:25	Продолжительный интервал	Скороходов А.С.
	16:50	Дневник	Калыц О.В.
02.03.15	10:41	Дневник	Скороходов Е.Е.
	13:31	Коронарография	Скороходов А.С.
	14:46	Ангиография почечных артерий	Скороходов А.С.
	15:50	Дневник	Калыц О.В.
	20:53	Дневник дежурного врача	Щенячкин П.И.
03.03.15	09:00	Клинический	Калыц О.В.
	13:27	Дневник	Калыц О.В.
04.03.15	11:27	Продолжительный интервал	Карникова Д.В.
	12:04	Дневник	Калыц О.В.
09.03.15	09:30	Дневник	Яременко Е.В.
	11:20	Стендирование коронарных артерий	Карникова Д.В.
	11:42	Стендирование почечных артерий	Карникова Д.В.

Рис. 3. – База данных ГИС: поток первичных данных

Имплицитное обучение без учителя на основе потока данных с кадром $\{\tau/T_0\}$ означает следующее:

- 1) фиксируется текущее множество данных;
- 2) фиксируется произвольный тест $z \in \{\tau\}$;
- 3) фиксируется произвольный домен Z из $G(z)$;
- 4) все данные z-теста в потоке $\cup_i \{\tau/T_0\}_i$ преобразуются к домену Z , следовательно, возникает Z-задача с базой прецедентов $\Omega = \cup_i \{\tau/T_0\}_i$;
- 5) для возникшей Z-задачи формируются базисы моделей знаний по схеме (1).

Процедура повторяется для всех $z \in \{z\}$ и всех доменов Z из $G(z)$. После добавления нового кадра весь процесс повторяется заново. В результате обновляются все авто/гетеро-ассоциативные модели знаний «ИКС-коннектома». На рис. 4 показано сетевое отображение связности «мозга» агента или когнитивного ядра интеллектуальной среды.

Возбуждение части тестов через тысячи связей с другими тестами ИКС-коннектома способно почти моментально вызвать богатейшую сеть ассоциаций (решений задач). Задачи «вшиты» в «коннектом», поэтому решения возникают автоматически (аналог «подсознательной» активности мозга). ИКС-коннектом, как и биологический

прообраз, иллюстрирует авто/гетеро-ассоциативность, распределенность, робастность и активность памяти. ИКС-коннектом – это фракталоподобная, постоянно растущая и изменяющаяся структура. Он задает первичную структуризацию ПрО (семантическую сеть предметной области). Можно говорить о глобальном ИКС-коннектоме интеллектуальной среды (инструмент: web based connectom viewer), ПрО-коннектомах, Z-коннектомах (обеспечивают гетеро-ассоциативность). В совокупности они образуют масштабируемую систему опыта конкретной ИКС (когнитивного агента, интеллектуальной среды) [7].

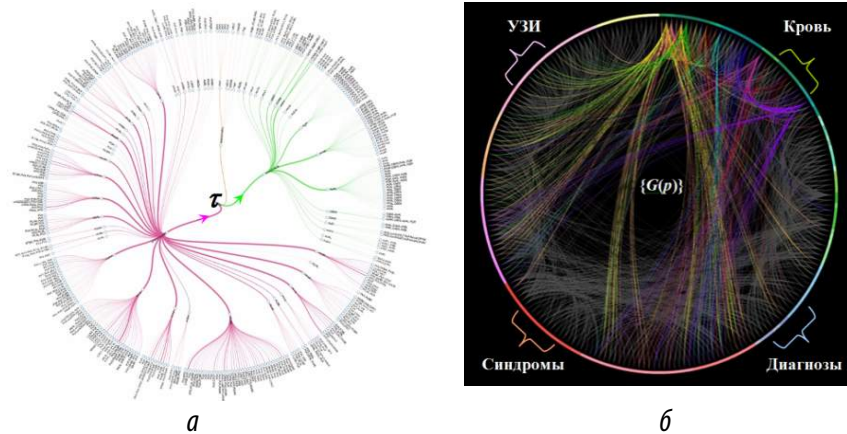


Рис. 4 – Иллюстрации ИКС-коннектома (Relationship Viewer):
 а - конус причинности и конус эффекта произвольного теста z ;
 б - «Коннектом» интеллектуальной госпитальной системы

Разработано программное обеспечение, позволяющее формировать разномасштабные «коннектомы» интеллектуальных систем. На рис. 5 показаны примеры

СФ-коннектомов задач прогнозирования исходов заболеваний и послеоперационных осложнений в рамках госпитальных и телемедицинских систем/сред [8].

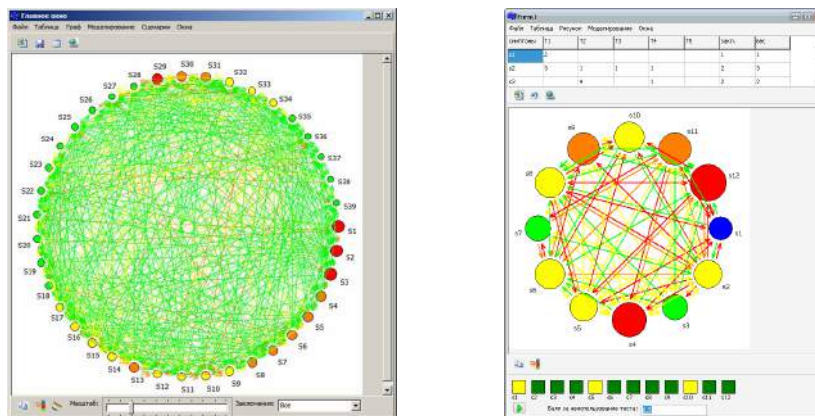


Рис. 5 – Примеры СФ-коннектомов ИКС

На рис. 6 показана гетеро-асоциативность базисов СФ/ФС и гетеро-асоциативность СФ внутри базисов. Гетерархия и гетеро-асоциативность возникают, например, в

рамках орграфа задач $G(z)$. Показано также распределение активности радикалов-синдромов внутри СФ (кадр активности).

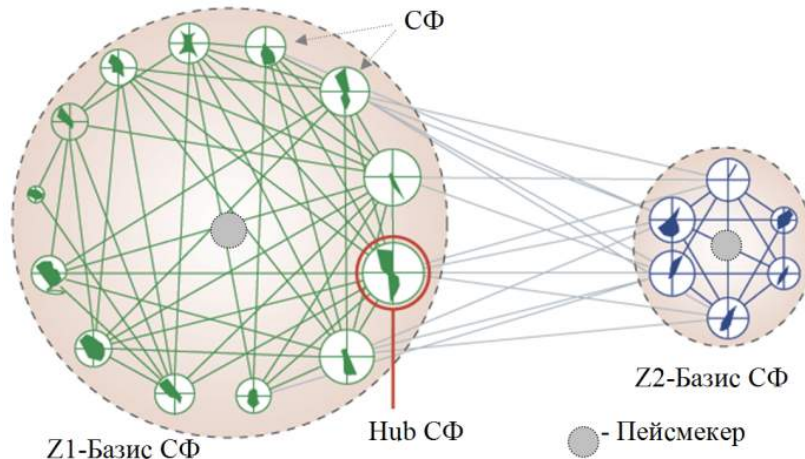


Рис. 6 – Гетеро-асоциативность базисов СФ и гетеро-асоциативность СФ внутри базисов

Приведем некоторые оценки. Если $|\{\tau\}| > 100$ и $|\{G(\tau)\}| > 10$, то потенциальное количество Z-задач и, соответственно, Z-базисов моделей знаний превышает 10^{100} , т.е. гугол (а также гугол процессов самоорганизации). При этом каждый базис содержит большое количество конкурирующих собственных форм, а каждая СФ содержит большое количество конкурирующих собственных значений (предельных синдромов, радикалов). Реальные процессы категоризации осуществляются для тех Z-задач, для которых $E_z > 0$. Возможно, отчасти в этом кроется причина непостижимости работы мозга, его продуктивности и креативности.

ИКС-коннектом позволяет создавать когнитивные надстройки над СУБД, включая: семантический язык запросов, интерактивный транслятор с естественного языка в SQL/XML (семантический и лексический процессоры [8]); встроенные индуктивные и дедуктивные системы. Подобный инструментарий позволяет преодолеть семантический разрыв между моделями создаваемыми специалистом предметной области и программистом. Примеры запросов: «СВВП/D2?» «Дать прогноз послеоперационных осложнений».

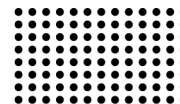
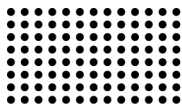
ЗАКЛЮЧЕНИЕ

ИКС-коннектом является важнейшей компонентой когнитивного ядра интеллектуальной среды, реализуя модельные и сервисные функции, включая когнитивные

вычисления. Разработаны прототипы информационных приложений, которые подтверждают универсальный характер предложенных конструкций.

В рамках концепции искусственного «коннектома» следует отметить такие аспекты:

- продуктивность – принципиальная возможность автоматического создания бесконечного множества иерархий новых задач и, соответственно, процессов ассоциирования на основе единых алгоритмов; возможность и легкость новаций;
- чрезвычайная роль контекста (банка тестов, базы прецедентов, уровня общности, эмоций, информационного напряжения и т.д.), а значит — возможность множественных трактовок сообщения и событий вообще;
- неожиданность сопоставляемых объектов или процедур: чем дальше отстоят друг от друга объекты, чем неожиданнее процедура сопоставления, тем эффективнее может оказаться творческий процесс;
- способность мыслить по аналогии, то есть искать сходство с уже известными явлениями, что, в свою очередь, позволяет усваивать огромные массивы информации;
- все процессы основываются на работе ассоциативной памяти, и мы имеем дело с постоянной сложной перестройкой всей «нейронной» сети, происходящей по определенным правилам;



- избыточность и возможность различных путей для поиска одного и того же: дает человеку несопоставимо больше степени свободы выбора алгоритмов как фиксации, так и считывания информации (базисы форм), что на порядки увеличивает уровень сложности; избыточность и в нахождении — попутно — того, что не искали (множество побочных решений).
- характерна размытость, неточность, приближенность описаний, не снижающая эффективность поиска в памяти и построения алгоритма поведения (грубость).

ЛИТЕРАТУРА

1. Max Wertheimer. Productive thinking. - Harper Torchbooks, 1982. – 302 p.
2. Sporns O, Tononi G, Kötter R (2005) The Human Connectome: A Structural Description of the Human Brain. *PLoS Comput Biol* 1(4): e42.
3. Seung S. Connectome: How the Brain's Wiring Makes Us Who We Are. – Houghton Mifflin Harcourt. — 2012. — 360 p.
4. Anokhin K.V. Cognitome: In search of a general theory for cognitive science // The Sixth International Conference on Cognitive Science. Abstracts (Kaliningrad, Russia, June 23-27, 2014). – Kaliningrad: Immanuel Kant Baltic Federal University, 2014. – PP. 26 - 28
5. Human Connectome Project <http://www.humanconnectomeproject.org/>
6. Lavrov V.V., Rudinskiy A.V. Neurophysiological rationale for the model of management of processing and registering of visual information // *Modern high technologies* №1, 2013. – pp. 49 – 54 (in Russian).
7. Prokopchuk Y.A. Paradigm of Limiting Generalizations: models of cognitive architectures and processes. - Saarbrücken, Deutschland: LAP LAMBERT Academic Publishing, 2014. – 204 p. (in Russian)
8. Prokopchuk Y.A. Principle of Limiting Generalizations: Methodology, Problems, and Applications. Monograph. Dnepropetrovsk, Institute of Technical Mechanics of the NAS and the State Space Agency of Ukraine Publ., 2012. 384 p. (in Russian)
9. Prokopchuk Y.A. Spiral Cognitive Metadynamics: Ways of Implementation in Information Technologies. *Problems of Information Technology*, 2013, no. 01 (013), pp. 6–17. (in Russian)