

УДК 519.1

**А. А. Харламов**, д-р техн. наук., ст. науч. сотр.; e-mail: kharlamov@analyst.ru,  
Институт высшей нервной деятельности и нейрофизиологии РАН, г. Москва,  
**Т. В. Ермоленко**, канд. техн. наук, нач. отдела, e-mail: etv@iai.dn.ua,  
Институт проблем искусственного интеллекта, Украина, г. Донецк

## Нейросетевая среда (нейроморфная ассоциативная память) для преодоления информационной сложности. Поиск смысла в слабоструктурированных массивах информации. Часть II. Обработка информации в гиппокампе. Модель мира

*Ассоциативная память человека является средством для преодоления информационной сложности, возникшей и возникающей постоянно в процессе его жизнедеятельности. Рассмотрены архитектура, свойства и функциональность ассоциативной памяти в целях ее использования для моделирования способности преодоления информационной сложности. Описан процесс перенормировки весовых характеристик образов событий, хранящихся в иерархии колонок коры полушарий большого мозга человека, который происходит в ламелях гиппокампа. Даются представления о модели мира человека, формирующейся на основе иерархий словарей разных уровней разных модальностей в коре. Показан изоморфизм лингвистической модели мира и многомодальной модели мира человека.*

**Ключевые слова:** кортикоморфная ассоциативная память, структурная обработка информации, ассоциативная (однородная семантическая) сеть, переранжирование понятий, модель мира, классификация, обработка текстов

### Введение

В первой из двух частей работы, которая была опубликована ранее, освещены вопросы формирования когнитивных сетей в мозге человека и использование технологии их формирования для автоматического анализа текстов. В ней были представлены механизмы формирования ассоциативных (однородных семантических) сетей, в том числе формирования словарей образов событий различной частоты встречаемости, а также выявления связей между этими образами во входной информации. В качестве примера обработки информации с использованием этих механизмов была представлена обработка текстовой информации. Во второй части работы описываются механизмы переранжирования весовых характеристик, выявленных на предыдущем этапе образов событий. Описана в структурных терминах модель мира человека, состоящая из трех компонентов: лингвистического и двух многомодальных. Показан изоморфизм многомодальных и лингвистического компонентов модели мира. Представлены механизмы сравнения и классификации текстов, в том числе естественно-языковых текстов.

### 1. Обработка информации в гиппокампе

Кроме коры больших полушарий другой структурой мозга, существенно важной для формирования модели мира, является гиппокамп. Ламели гиппокампа (сечения, ортогональные к длинной оси гиппокампа) ответственны за хранение информации о связях образов событий, хранящихся в колонках коры, в рамках целых ситуаций. Пирамидные нейроны (теперь уже) поля CA<sub>3</sub> p-й ламели гиппокампа формируют искусственную нейронную сеть Хопфилда [1], весовые характеристики которой хранят информацию об объединении образов событий, хранящихся в колонках коры, относящихся к конкретной ситуации, в рамках этой ситуации:

$$N_p = \cup_i B_i. \quad (1)$$

Знак крышки ( $\wedge$ ) над соответствующими событиями в (1) отсутствует потому, что в ламелях гиппокампа в сеть объединяются не элементы словарей — фрагменты траекторий в многомерном пространстве, а их индексы.

Ламели гиппокампа получают информацию от колонок коры [2], причем и здесь работает ассоциативный принцип обращения к информации. Весь поток информации, поступающий через переключе-

ния из коры, приходит одновременно на все ламели гиппокампа. Но откликаются на этот поток только те ламели, которые содержат информацию о событиях, образы которых присутствуют во входном потоке. Отклик тем больше, чем сильнее ассоциация и чем больше вес образов событий в колонках коры.

На каждой итерации взаимодействия коры и гиппокампа поле  $CA_1$  гиппокампа (как конкурентная сеть) формирует отклик только одной ламели гиппокампа — наиболее близкой к входной ситуации.

Но на этом дело не кончается: в результате отклика текущей ламели в коре в колонке, инициировавшей процесс, происходит дообучение (в результате так называемой долговременной потенциации). И на следующей итерации ассоциативная проекция той же ситуации на ламели гиппокампа оказывается измененной из-за этого дообучения, и следующий отклик ламелей изменяется.

После 15...20 итераций образы событий, включенных в ситуацию, в колонке коры изменятся из-за дообучения (которое связано с моделями ситуаций, хранящихся в ламелях гиппокампа). Вообще говоря, и модели ситуаций в ламелях гиппокампа изменяются также. То есть этот итеративный процесс переупорядочивает информацию в коре о событиях текущей ситуации в соответствии с имеющимися моделями ситуаций, хранящимися в гиппокампе, а эти модели ситуаций принимают к сведению информацию о текущей ситуации, отображенной в виде образов событий в коре.

А поскольку поле  $CA_3$  гиппокампа работает в виде громадной автоассоциативной рекуррентной памяти по всей его длине и ширине, множество отдельных моделей ситуаций  $N_p$ , которые хранятся в ламелях  $p$ , совместно с образами событий, хранящихся в колонках коры, формируют единую семантическую сеть  $N$  на многомодальной модели мира, хранящейся в колонках коры:

$$N = \bigcup_p N_p. \quad (2)$$

При этом важна не дифференцировка хранимых образов, это обеспечивается в коре, а контекстуальные пространственно-временные связи образов в рамках целых ситуаций.

### 1.1. Семантическая сеть

Опишем более формально ассоциативную (однородную семантическую) сеть  $N$ , а также рассмотрим ее как некое подмножество метрического пространства.

**Определение 1.** Под семантической сетью  $N$  понимается множество несимметричных пар событий  $\{ \langle c_i, c_j \rangle \}$ , где  $c_i$  и  $c_j$  — события, связанные между собой отношением ассоциативности (совместной встречаемости в некоторой ситуации):

$$N \approx \{ \langle c_i, c_j \rangle \}.$$

В данном случае отношение ассоциативности несимметрично:

$$\langle c_i, c_j \rangle \neq \langle c_j, c_i \rangle.$$

**Определение 2.** Семантическая сеть, описанная таким образом, может быть представлена как множество так называемых звездочек  $\{ \langle c_i < c_j \rangle \}$ :

$$N \approx \{ z_i \} = \{ \langle c_i < c_j \rangle \}.$$

**Определение 3.** Под звездочкой  $\langle c_i < c_j \rangle$  понимается конструкция, включающая главное событие  $c_i$ , связанное с множеством событий-ассоциантов  $\langle c_j \rangle$ , которые являются семантическими признаками главного события, отстоящими от главного события на одну связь. Связи направлены от главного события к событиям-ассоциантам.

**Определение 4.** Звездочка с единичными значениями весов событий-ассоциантов называется единичной звездочкой (звездочкой-ортом).

**Определение 5.** Звездочкой-подпространством называется звездочка, полученная на единичной звездочке введением весов событий ( $w_j$  — веса событий):

$$Z \approx \langle c_i < w_j c_j \rangle.$$

**Определение 6.** Весом  $w_j$  образа события  $c_j$  в сети является значение счетчика повторных появлений событий во входном тексте.

Семантическая сеть в терминах этих определений представляет собой декартово произведение подпространств, порождаемых всеми звездочками, входящими в семантическую сеть, полученными на единичных звездочках за счет введения весовых характеристик понятий-ассоциантов:

$$N = Z_1 \times Z_2 \times \dots \times Z_l.$$

Для введения в метрическом пространстве, определенном на семантической сети операций вложения и пересечения, разобьем звездочки на пары событий  $\langle c_i, c_j \rangle$ . Здесь  $c_i$  — главное событие звездочки, а  $c_j \in \langle c_j \rangle$  — множество событий-ассоциантов события  $c_i$ . Связь направлена от  $c_i$  к  $c_j$ .

Все сказанное относится к когнитивной семантической сети, т. е. семантической сети, формируемой на нейронной сети как на субстрате [3]. Отчужденная от субстрата когнитивная семантическая сеть остается той же ассоциативной сетью, которая представлена в определениях 1–6. В этом случае множество событий  $\{ c_i \}$  — это индекс (квализлов А квазисловаря) квазитекста (в том числе и обычного текста), а сеть  $N$  — ассоциативная сеть, построенная на этом индексе.

### 1.2. Перенормировка семантических весов

При формировании сети на основе большого корпуса квазитекстов, каковыми являются все входные последовательности всех модальностей, получают некоторые весовые характеристики  $w_j$  образов событий  $c_j$ . При этом частота их встречаемости приближается к их смысловому весу. При анализе

малых по объему квазитекстов, каковыми являются отдельные ситуации, частота встречаемости уже не характеризует важности образа события.

По аналогии с перевзвешиванием образов событий в коре под воздействием результатов обработки информации в гиппокампе разумно каким-то образом менять весовые характеристики образов событий ассоциативной сети с учетом их связей с соседними вершинами сети. При такой перенормировке на каждой итерации образы событий, связанные с образами, имеющими большой вес, должны свой вес увеличивать. Другие события должны равномерно терять свой вес.

Сформированное первоначально статистическое представление квазитекста — сеть образов событий с их связями (в виде частоты встречаемости событий и связей между ними) — перенормируется, что позволяет перейти от частотного портрета квазитекста (сети с частотами) к ассоциативной сети ключевых образов событий квазитекста (сети с рангами этих событий — их смысловыми весами):

$$w_i(t+1) = \left( \sum_{i \neq j} w_i(t)w_{ij} \right) \sigma(\bar{E}),$$

здесь  $w_i(0) = z_i$ ;  $w_{ij} = z_{ij}/z_j$  и  $\sigma(\bar{E}) = 1/(1 + e^{-k\bar{E}})$  — функция, нормирующая на среднее значение энергии всех вершин сети  $\bar{E}$ ;  $z_i$  — частота встречаемости  $i$ -го слова в тексте;  $z_{ij}$  — частота совместной встречаемости  $i$ -го и  $j$ -го слов в фрагментах текста;  $t$  — номер итерации. Полученная числовая характеристика слов — их смысловой вес — характеризует степень их важности в тексте.

## 2. Автоматическое структурирование больших объемов разнородной информации

Используя приведенный выше формализм обработки информации, представленной в виде внутренне структурированных последовательностей (квазитекстов и естественных языковых текстов в том числе), мы можем структурировать произвольные потоки информации различных модальностей. Мы можем формировать словари  $\{B_i\}$  образов событий разных модальностей разной частоты встречаемости (разной сложности). Мы можем формировать из входных последовательностей  $\{A\}$  последовательности  $\{C\}$  связей слов словарей. Таким образом, мы реализуем структурный анализ информации. Далее мы можем формировать весовые характеристики  $w_i$  слов этих словарей в зависимости от их смыслового (структурного) веса в конкретных квазитекстах (в естественных языковых текстах в том числе). Другими словами, мы можем формировать смысловой портрет квазитекста (в том числе и естественного языкового текста) в виде семантической сети со взвешенными (ранжированными) вершинами.

Пользуясь этими механизмами, мы можем перейти к пониманию структуры модели мира чело-

века, которая включает в себя лингвистический и многомодальные компоненты. И далее, мы можем сравнивать эти тексты по смыслу, вычисляя степень пересечения их семантических сетей. И, следовательно, классифицировать квазитексты, сравнивая их смысловые портреты со смысловыми портретами классов — моделей предметных областей.

### 2.1. Модель мира — иерархическая структура словарей образов событий разной сложности и разных модальностей

Рассмотрим формирование с помощью описанных механизмов представления о мире в терминах одной модальности. Рассмотрим иерархическую структуру словарей образов событий мира, имеющую вид, представленный на рис. 1, где на каждом уровне имеется множество словарей образов событий, связанных с образами событий следующего уровня по типу каждый с каждым. В каждом подсловаре  $\{\hat{B}_i\}_{jkm}$   $i$  — слово в подсловаре,  $j$  — номер подсловаря на уровне,  $k$  — номер уровня, а  $m$  — номер модальности.

События внешнего (и внутреннего) мира, отображающиеся сенсорами в иерархическую структуру данной ( $m$ -й) сенсорной модальности, формируют в этой иерархической структуре иерархию словарей, слова которых оказываются связанными между собой синтаксическими последовательностями. При этом синтаксические последовательности, сформированные в процессе обработки предыдущего уровня, являются исходными для формирования словарей следующего уровня. Эта иерархия словарей является моделью мира в терминах  $m$ -й модальности.

Модель мира данной модальности — представляет собой объединение слов словарей  $\hat{M}_m$ . На верхнем ( $K$ -м) уровне в нем объединяются все слова  $\hat{B}_{ijkm}$  словаря (подсловарей) верхнего уровня, а на всех более низких уровнях в него включаются (в соответствующие места — по ассоциативному принципу) слова  $\hat{B}_{ijkm}$  словарей нижних уровней:

$$\hat{M}_m = \bigcup_{ij} \bigvee_{ijk < K} \hat{B}_{ijkm}. \quad (3)$$

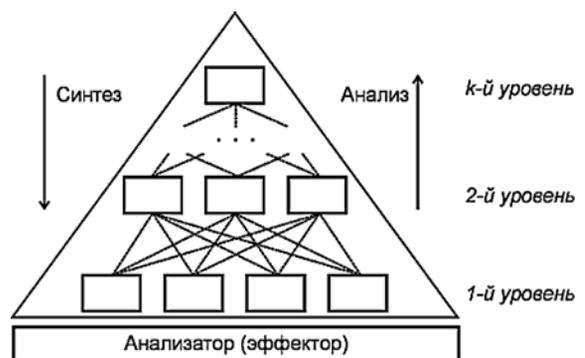


Рис. 1. Многоуровневая иерархическая структура из словарей образов событий одной модальности

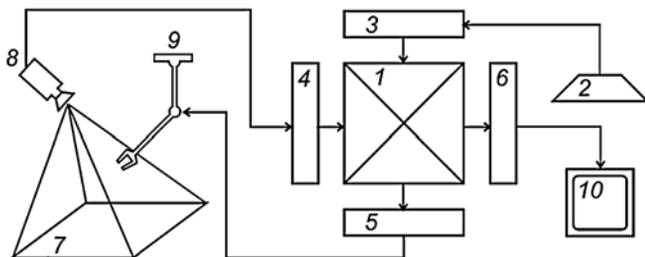


Рис. 2. Объединение иерархий для многомодального представления знаний в антураже задачи "Ханойская башня" (интегральный робот):

1 — многомодальная иерархия из словарей разных уровней разных модальностей; 2 — клавиатура; 3 — сенсорная подсистема ввода текстовой информации; 4 — сенсорная подсистема ввода зрительной информации, предварительно обработанной и развернутой, полученной с модели сцены 7, наблюдаемой светочувствительной матрицей 8; 5 — эффекторная подсистема управления манипулятором 9; 6 — эффекторная подсистема синтеза текста с выдачей на экран видеотерминала 10

Здесь  $\cup$  — операция объединения, а операция  $\vee$  обозначает включение на свое место в слова словаря более высокого уровня. Это включение аналогично логическому сложению последовательности, соответствующей слову словаря нижнего уровня, с синтаксической последовательностью (вложению слова словаря в соответствующую купюру синтаксической последовательности).

Объединив несколько иерархических структур, соответствующих разным сенсорным модальностям (рис. 2) со сформированными на них модальными моделями мира  $\hat{M}_m$ , мы получим многомодальную модель мира:

$$\hat{M} = \cup_m \hat{M}_m = \cup_{ijm} \hat{B}_{ijkm} \vee_{ijk < Km} \hat{B}_{ijkm}. \quad (4)$$

## 2.2. Формирование многомодального многоуровневого представления

Любое многомодальное представление можно рассматривать как квазитекстовое, поскольку оно имеет многоуровневую внутреннюю структуру. Поэтому рассмотрим формирование такого представления у человека на примере обработки текстовой информации. Язык, представленный в виде множества текстов, с помощью описанного выше формализма подвергается структурному анализу, в результате которого выявляются его словарные компоненты разных уровней.

При обработке текстов обычно рассматриваются несколько уровней обработки информации: морфологический, лексический, синтаксический и семантический. На каждом из уровней возможно формирование нескольких типов словарей. Мы рассмотрим только некоторые из них: на морфологическом уровне — словарь флективных морфем; на лексическом уровне — словарь корневых основ; на синтаксическом уровне — словарь синтаксиса, представляющих собой флективную структуру синтаксических групп с выколотыми корневыми ос-

новами; на семантическом уровне — словарь попарной сочетаемости корневых основ [4].

Рассмотрим последовательные этапы формирования вышеперечисленных словарей. Сначала формируется словарь флективных морфем  $\{\hat{B}\}_1$ , так как они являются наиболее часто встречающимися языковыми единицами.

На следующем этапе формируется словарь корневых основ слов. После того как сформировался словарь флективных морфем, фильтрация словарем флективных морфем множества текстов приводит к формированию словаря корневых основ  $\{\hat{B}\}_2$ , так как в результате взаимодействия множества текстов со словарем флективных морфем возникает множество синтаксических последовательностей с купюрами вместо флексий — множество последовательностей корневых основ.

Далее формируется словарь синтаксического уровня  $\{\hat{B}\}_3$ . Этот словарь формируется фильтрацией фрагментов текста, соответствующих по длине синтаксической группе, через словарь корневых основ слов. Полученные при этом цепочки флективных морфем кластеризуются на подклассы по их графемной структуре. Эти подклассы являются частями синтаксических классов, соответствующих основным синтаксическим группам.

После того как сформирован словарь синтаксиса, формируется словарь попарной (смысловой) сочетаемости слов [4]  $\{\hat{B}\}_4$ , которая определяет семантику. Для простоты будем считать, что именно слова словарей этого уровня объединяются в семантическую сеть  $N$ :

$$N = \cup_i B_i,$$

здесь слова естественного языка  $B_i$  соответствуют словам-траекториям  $\hat{B}_i$  в многомерном пространстве словарей  $\{\hat{B}_i\}_4$ .

## 2.3. Модель мира.

### Языковой и многомодальные компоненты

Модель мира человека представляет собой совокупность модальных компонентов, которые являются множеством образов ситуаций, хранящихся в ламелях гиппокампа.

Однако существует более точно очерченная функциональность многомодальных компонентов модели мира, которая связана с латерализацией полушарий, зависящей от представительства в коре речевых функций. В левом полушарии под воздействием социума формируется лингвистическая модель мира  $\hat{M}_L$  (3) в терминах естественного языка.

Таким образом, в субдоминантном (правом у правой) полушарии локализован многомодальный образный компонент модели мира  $\hat{M}_R$  (4), а в доминантном (соответственно, левом у правой) полушарии — локализованы многомодальный схематический  $\hat{M}_L$  (4) и лингвистический  $\hat{M}_L$  (3) компоненты модели. Такое распределение не случайно:

в связи с наличием в левом полушарии лингвистической модели мира  $\hat{M}_L$  социум через язык как сегментирующую функцию влияет на формирование левополушарного многомодального компонента  $\hat{M}_L$ . А образный компонент правого полушария  $\hat{M}_R$  формируется под воздействием индивидуальной истории развития конкретного индивидуума (конечно, используя знания левополушарного компонента). При этом схематические образы событий левополушарной многомодальной схематической модели  $\hat{M}_L$  наполняются содержанием — соответствующими образами событий правополушарной образной многомодальной модели  $\hat{M}_R$ .

Все три компонента имеют поуровневые связи и потому работают как единое целое:

$$\hat{M} = \hat{M}_L \cup \hat{M}_L \cup \hat{M}_R.$$

Ключевые понятия лингвистической модели мира соответствуют ключевым образам событий многомодальной модели, т. е. лингвистическая модель изоморфна многомодальной модели мира.

Таким образом, многомодальная модель, представленная в виде объединенной семантической сети  $\hat{M}_L \cup \hat{M}_R$ , имеет левополушарный языковой эквивалент (также в виде сети)  $\hat{M}_L$ : многомодальным образом событий мира в их взаимосвязях в ней сопоставляются понятия (слова, словосочетания) с их соответствующими связями. Поэтому операции на языковой модели можно считать отражением операций на многомодальной модели (мы можем оценить операции над многомодальной моделью, моделируя их на лингвистической модели). Для этого сопоставим многомодальной модели (4) соответствующую сеть (2). Сеть же для языковой модели (3) сформируем аналогично (1) и (2), но только для лингвистической модальности  $m = l$ :

$$N_{pl} = \bigcup_{ijklp} B_{ijklp}.$$

Здесь индекс  $l$  означает лингвистическую модальность:

$$N_l = \bigcup_p N_{pl}.$$

#### 2.4. Сравнение семантических сетей. Классификация

Поскольку далее нам придется выявлять смысл текстов их соотношением с моделями предметных областей, рассмотрим вопросы сравнения текстов по смыслу и, далее, классификацию текстов. Все, что будет сказано ниже, в равной степени относится как к семантическим сетям и прочим упоминавшимся выше конструкциям в текстах (и квазитекстах), так и к когнитивным представлениям в многомерном (когнитивном) пространстве (которое моделируется колонками коры и ламелями гиппокампа).

**Определение 7.** Введем скалярное произведение на векторах  $c_i, c_j$ , где угол между векторами по-

нятий  $c_i, c_j$ ;  $w_{ij}$  пропорционален весу связи от  $c_i$  к  $c_j$ ;  $w_{ij} \in (0...90^\circ)$ .

Площадь треугольника  $s_j$ , построенного на векторах  $c_i, c_j$ , развернутых на угол  $w_{ij}$  относительно друг друга, будет использована для вычисления степени пересечения сначала звездочек, а потом семантических сетей как совокупностей звездочек.

**Определение 8.** Под пересечением двух звездочек понимается сумма по всем событиям-ассоциантам данного главного события звездочки пересечений площадей двух треугольников, построенных в плоскости векторов  $c_i, c_j$ , один из которых построен на векторах, развернутых на угол, пропорциональный связи  $(w_{ij})_1$  между событиями в одной звездочке, а другой — на угол, пропорциональный связи  $(w_{ij})_2$  между теми же событиями в другой, сравниваемой с первой, звездочке. В случае если в одной из звездочек пары, для которой считается пересечение

$$s_{12} = \langle c_{i_1} \langle c_{j_1} \rangle \rangle \cap \langle c_{i_2} \langle c_{j_2} \rangle \rangle = \sum_{j=1}^{\max(N_1, N_2)} (s_{j_1} \cap s_{j_2}),$$

не нашлось соответствующего события-ассоцианта, пересечение считается равным 0. Здесь  $N_1, N_2$  — число ассоциантов в звездочках соответственно 1 и 2.

**Определение 9.** Под пересечением семантических сетей понимается сумма пересечений звездочек, включенных в эти сети (считая по главным понятиям):

$$S_{12} = \langle c_{i_1} c_{j_1} \rangle \cap \langle c_{i_2} c_{j_2} \rangle = \sum_{k=1}^{\max(M_1, M_2)} (s_{k_1} \cap s_{k_2}),$$

$M_1, M_2$  — число звездочек, входящих соответственно в семантические сети 1 и 2.

**Определение 10.** Под классификацией входной ситуации можно понимать отнесение семантической сети ситуации  $N$  к сети  $N_n$  (где  $n = 1, \dots, N$  — число предметных областей) одной из предметных областей модели мира. Здесь объединение сетей  $\bigcup_n N_n$  соответствует модели мира  $\hat{M}_m$   $m$ -й модальности. В идеальном случае семантическая сеть ситуации вкладывается в сеть соответствующей предметной области.

Используя операцию пересечения сетей  $N_1$  и  $N_2$ , определенную выше, мы можем оценивать степень подобия двух сетей  $N_1 \cap N_2$  (рис. 3) и, тем самым, сравнивать по смыслу (по структуре) ситуации (их модели). Имея модели предметных областей в виде ассоциативных семантических сетей, мы можем классифицировать входные ситуации (описывающие их модели) вычислением степени совпадения (вложения) сети входной ситуации и сетей предметных областей (рис. 4), относя входную ситуацию к той предметной области, у которой степень

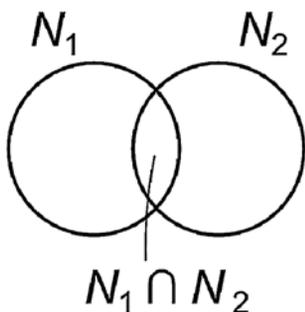


Рис. 3. Пересечение  $N_1 \cap N_2$  двух сетей  $N_1$  и  $N_2$ , характеризующее степень их смыслового подобия

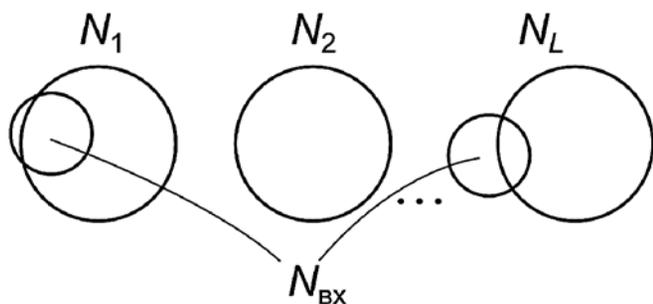


Рис. 4. Классификация входного текста путем выявления степени вложенности его семантической сети  $N_{вх}$  в одну или несколько семантических сетей классов-рубрик — предметных областей  $N_1, N_2, \dots, N_L$

совпадения сети входной ситуации с сетью предметной области окажется выше.

Далее мы можем абстрагироваться от входных ситуаций. Будем рассматривать их языковые модели. Пойдем еще дальше: перейдем к анализу текстов, на основе которых эти языковые модели формируются. Как и в случае классификации входных ситуаций, используя операцию пересечения сетей, мы можем оценивать степень подобия двух сетей и, тем самым, сравнивать по смыслу уже тексты. Имея модели предметных областей  $N_n$  в виде ассоциативных семантических сетей соответствующих тематических текстовых выборов, мы можем классифицировать входные тексты вычислением степени совпадения (пересечения/вложения) сети  $N$  входного текста и сетей предметных областей  $N_n$ , относя входной текст к той предметной области, у которой степень совпадения его сети с сетью предметной области окажется выше.

Покажем далее как, используя все выше перечисленные механизмы, подойти к понятию сложности чего-либо: некоторого явления (например, ситуации) в виде последовательности взаимосвязанных многомодальных образов, т. е. квазитекстов.

### 3. Сложность квазитекста как отражения некоторой ситуации

Попробуем оценить сложность входной информации, учитывая число модальностей  $m$  (отдельно

выделим лингвистическое представление  $l$ ), представленных в ней, число формируемых уровней  $k$  словарей, число словарей  $j$  на уровнях, число слов  $i$  в словарях. Поскольку эта информация структурируется в процессе ее обработки в виде семантических сетей, удобно найти способ оценки ее сложности с учетом ключевых понятий сети, имеющих ранг выше заданного. Можно, например, определять число существенно важных тем в семантической сети, а также — число ассоциантов понятий в сети.

**Определение 11.** Под тематической структурой текста будем понимать минимальный древовидный подграф  $T = \langle K, C, R, D, W, V \rangle$ , извлеченный из сети  $N$ , где в качестве корневой вершины берется вершина с максимальным рангом, а остальные вершины ранжируются в соответствии с их смысловыми весами. Здесь  $K$  — корневая вершина,  $C$  — множество вершин в сети,  $R$  — множество связей этих вершин в семантической сети,  $N, D$  — конкретный способ связей вершин в конкретной сети,  $N, W$  и  $V$  — весовые характеристики вершин и связей соответственно. Минимальный древовидный подграф получается удалением слабых связей  $V$  (с весом ниже заданного порога  $h_p$ ) и слабых вершин (также ниже заданного порога  $h_c$ ).

Оперируя числом тем, подтем и подподтем в тематических структурах текстов, а также числом ассоциантов вершин семантической сети, можно оценивать сложность текстов и сравнивать тексты по сложности. Поскольку любая вершина в сети имеет ближайшие соседи-ассоцианты, которые являются ее семантическими признаками, по числу ассоциантов можно определять сложность текста в шири-

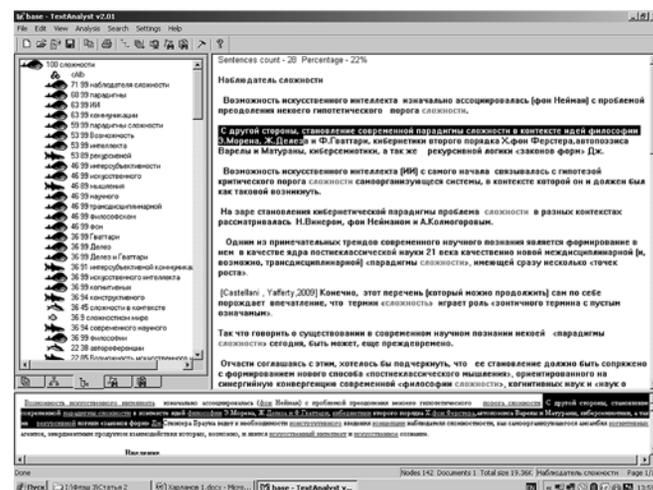


Рис. 5. Ближайшие ассоцианты конкретной вершины семантической сети, определяющие сложность текста в ширину (в левом окне) в тексте статьи В. И. Аршинова "Наблюдатель сложности как модель искусственного интеллекта" на графическом интерфейсе программы TextAnalyst. Вверху — корневая вершина, ниже справа — ассоцианты

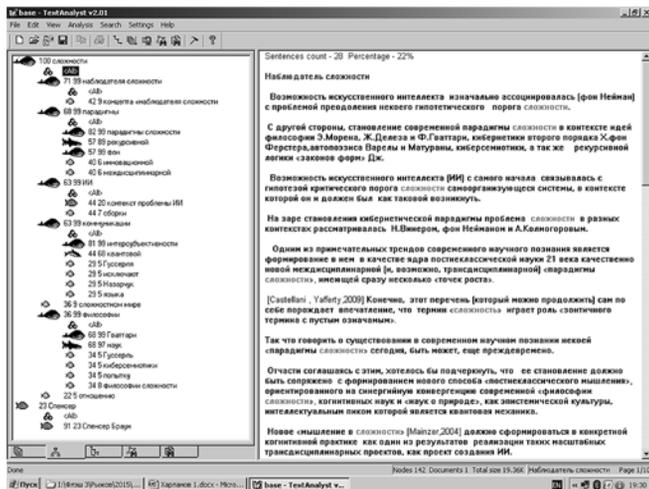


Рис. 6. Тематическая структура текста, определяющая сложность текста в глубину (в левом окне) в тексте статьи В. И. Аршинова "Наблюдатель сложности как модель искусственного интеллекта" на графическом интерфейсе программы TextAnalyst. Тематическое дерево имеет две корневые вершины — "Сложность" и "Спенсер". В первом дереве — семь подтем: "Наблюдатель сложности", "Парадигмы" и т. п. Наличие второй корневой вершины говорит о том, что в тексте есть отдельная тема, посвященная Спенсеру

ну (рис. 5), а по числу уровней подтем у главной темы — определять сложность текста в глубину (рис. 6).

### Заключение

Представленные во второй части работы механизмы обработки информации в мозге человека,

в том числе в ламелях гиппокампа, связанные с переранжированием весовых характеристик образов событий, хранящихся в памяти, позволяют выявлять наиболее важные ключевые образы, что является наиболее существенным моментом при анализе информации. Изоморфизм лингвистического и много-модальных компонентов модели мира человека, которая строится на основе описанных в первой и второй частях работы механизмах, позволяет интерпретировать картину мира на основе анализа текстовых представлений. Описанные механизмы сравнения и классификации текстов являются основой для такой интерпретации. В работе также тезисно представлены соображения об анализе сложности текстового материала, которые будут подробно изложены в следующей публикации.

Работа выполнена при поддержке РГНФ (проект № 15-03-00860 "Методология управления сетевыми структурами в контексте парадигмы сложности").

### Список литературы

1. Hopfield J. J. Neural networks and physical systems with emergent collective computational abilities // Proc. Natl. Acad. Sci. 1982, Vol. 79. P. 2554—2558.
2. Brown T. H., Zador A. M. Hippocampus / G. M. Shepherd (Ed.), The synaptic organisation of the brain. New York, Oxford: Oxford University Press, 1990. P. 346—388.
3. DARPA SyNAPSE Program. URL: <http://www.artificial-brains.com/darpa-synapse-program>.
4. Рахилина Е. В. Когнитивный анализ предметных имен: семантика и сочетаемость. М.: Русские словари, 2000.

A. A. Kharlamov, Senior Researcher, e-mail: kharlamov@analyst.ru,  
 Institute of Higher Nervous Activity and Neurophysiology of Russian Academy of Sciences, Moscow,  
 T. V. Yermolenco, Head of Department, e-mail: etv@iai.dn.ua,  
 Institute of Artificial Intelligence Problems, Ukraine, Donetsk

## Neuronetwork Environment (Neuromorphic Associative Memory) for Negotiation of Information Complexity. Meaning Identification in Semystructured Information Corpuses. Part II. Information Processing in Hippocampus. World Model

*Human associative memory is a tool for negotiation of information complexity which is a result of its living. In the paper shown an architecture, properties and functions of associative memory for its using in negotiation of information complexity. A rearranging process of weight characteristics of input event images stored in columns hierarchy of hemisphere of human cortex which is fulfilled in hippocampus lamellas are shown too. A representation of human world model in cortex which are formed on base of several levels and several modalities vocabularies hierarchy is shown. Isomorphism of linguistic and multimodal human world models is represented too.*

**Keywords:** korticomorphic associative memory, structural information processing, associative (homogeneous semantic) network, concept rearranging, world model, classification, text processing

### References

1. Hopfield J. J. Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci., 1982, vol. 79, pp. 2554—2558.
2. Brown T. H., Zador A. M. Hippocampus. G. M. Shepherd (Ed.), The synaptic organisation of the brain. New York, Oxford: Oxford University Press, 1990, pp. 346—388.
3. DARPA SyNAPSE Program. URL: <http://www.artificial-brains.com/darpa-synapse-program>.
4. Rakhilina E. V. Kognitivnyi analiz predmetnikh imen: semantika i sochetayemost, M.: Russkiye slovari, 2000.