

# Об одном подходе к определению семантической информации

Н.В. Максимов, А.А. Лебедев

Национальный исследовательский ядерный университет «МИФИ»

[nv-maks@yandex.ru](mailto:nv-maks@yandex.ru), [lebedevalex@live.ru](mailto:lebedevalex@live.ru)

## Аннотация

Определяется понятие «семантическая информация», через определение понятия "смысл". Анализируется вопрос соотношения понятий «знак» – «смысл» – «значение». Принципиальным фактором для проявления семантики является наличие выделенного направления деятельности – цели, на которую производится проекция системы знаков (информационных объектов). Для семантической информации направленность возникает при взаимодействии исходного информационного объекта (сообщения, текста) и целевого (который может быть представлен, в частности, как аспектная онтология). Семантическая информация – это информация в контексте некоторого смысла, определяющего множество элементов и их возможных состояний.

**Ключевые слова:** семантика, смысл, семантическая информация

**Библиографическая ссылка:** Максимов Н.В., Лебедев А.А. Об одном подходе к определению семантической информации // Компьютерная лингвистика и вычислительные онтологии. Выпуск 6 (Труды XXV Международной объединенной научной конференции «Интернет и современное общество», IMS-2022, Санкт-Петербург, 23 – 24 июня 2022 г. Сборник научных статей). — СПб.: Университет ИТМО, 2022. С. 30-40. DOI: 10.17586/2541-9781-2022-6-30-40

## 1. Семантика, смысл, значение

Семантика, не в широком смысле слова, а в контексте информационном, понимается несколько специфично – как смысловое содержание текста, процесса и т.д. (более подробно см. в [1]).

Смысл в таком понимании - нечто, содержащееся в сообщении (тексте, сигнале), то, по чему мы можем судить об обозначаемом объекте («Что это», «Как устроено»). Смысл объекта – это некоторое *существо* (в пределе, архетип), образ, построенный субъектом применительно к этому объекту в соответствии с некоторой схемой (точкой зрения, шаблоном, целью), отражающей текущий интерес субъекта.

Понятие «смысл» связано с понятием «значение», содержание которого достаточно конструктивно определено в [2] следующим образом: значение можно понимать как нечто, возникающее из *взаимодействий* организма с его нишей в окружающей среде, воспринимаемой как субъективный (т.е. выстроенный в результате таких взаимодействий) мир организма. Значение есть отношение между организмом и средой, определяемое *ценностью*<sup>1</sup>, которую отдельные *аспекты* среды<sup>2</sup> (разделяемые на категории) имеют

<sup>1</sup> Очевидно, здесь «ценность» рассматривается как оценка степени зависимости организма от аспекта среды, т.е. это свойство взаимодействия (соотношения), представленное в форме знака-символа. И в этом контексте правильнее было бы говорить о значимости.

<sup>2</sup> В реальности — это физически величины, воспринимаемые как воздействия на органы чувств или измерители.

для этого организма. Оно не чисто субъективно ("в голове") и не объективно ("в мире"), но характеризуется взаимодействием между организмом и средой.

По сути, значение и смысл соотносятся как объект и функция, задающая этот объект, «сущностное свойство характеризуемого с его помощью объекта» [3], т.е. смысл – это то, что формирует в информационном взаимодействии [4] значение.

Смысл может иметь как содержательный, так и идентифицирующий характер. В первом случае – это представление содержания в некотором аспекте, во втором – обозначение, название этого аспекта (например, «Координаты в смысле географии»). При этом «смысл», как существо объекта, может быть представлено как «по-вертикали», так и «по-горизонтали». В первом случае – это обобщение – наиболее существенные составляющие, представленные понятиями относительно высокого уровня. Во втором – это некоторый аспектный срез, представляющий содержание с некоторой конкретной точки зрения, например, в рамках конкретной научной дисциплины или потребность субъекта (интереса, профиля деятельности и т.д.).

В семиотике [5] значением (предметным значением, денотатом) некоторого знака (в частности, имени) называется обозначаемый (называемый) им предмет или класс предметов (объем именуемого понятия – «Что»), а смыслом знака — содержание понятия (концепт денотата – «Как», «Какой»), т. е. то, понимание чего является условием адекватного восприятия данного знака. По Г. Фреге знак (как отдельный объект), во-первых, указывает на другой объект (значение знака), а во-вторых, на соответствующее означаемому объекту понятие (смысл знака).

Условием применимости знака служат признаки, образующие в своей совокупности его интенционал (или, в другой терминологии – сигнификат). Это не сам класс обозначаемых объектов, а те признаки (свойства), на основании которых эти объекты/ситуации объединены в данный класс и противопоставлены членам других классов. Т.е. знак отсылает к объекту (денотату) не непосредственно, а опосредованно, через понятие о данном классе объектов, имеющееся в сознании человека. Это позволяет выделять (различать) некоторый объект среди других не только по наличию «метки» (как еще одного дополнительного элемента, в общем то инородного по отношению этому объекту), но и по некоторому «существо» этого объекта, представленному в соответствии с некоторым общим для данной предметной области (Про) порядком вещей и свойствами, выбираемыми сообразно цели деятельности. Что, собственно, и составляет понимание.

Подобно тому, как знаки (слова) именуют объекты (в широком смысле, т.е. включая вещи, понятия, свойства, соотношения, действия и т.д.), так же и предложения именуют то, что можно назвать «ситуацией» объектов – взаимодействия, состояния, взаимосвязи, обстоятельства и т.д., которые могут быть как реально существующими, так и воображаемыми, в т. ч. реально невозможными и даже бессмысленными.

Здесь важно отметить, что смысловое содержание слова (его семантика) определяется обычно предопределенной ранее парадигматикой (вне или в начале текста) или контекстом – лингвистической информацией или статистикой совместного словоупотребления, извлекаемой при разборе текста и позволяющей уточнить парадигму термина. Семантика же предложения или полного текста определяется «ситуативно»: текст описывает некоторые, обычно нетривиальные результативные процессы или объекты, для которых характерна завершенность и, в частности, свойство новизны. То есть для семантики текста характерна целостность (а лингвистике для текста – сверхфразовое единство), которая динамически и, в общем случае – вариантно, формируется в процессе использования содержания текста субъектом при решении его практической задачи. В автоматизированных ИПС для этого может использоваться прагматика ситуации – например, указанные пользователем категории, поисковая история, спецификация задачи.

Интенционалу знака в некотором смысле можно поставить в соответствие понятие семантического поля<sup>3</sup> – контекста, который, собственно, и определяет выбор конкретного значения (что можно отождествлять со смыслом) объекта из множества существующих.

В общем случае, можно сказать, что поле<sup>4</sup> – это *набор* функционально взаимосвязанных объектов, а точнее – между этими объектами есть определенного типа *взаимодействия* (например, силовые в физике, функциональные группы в лингвистике), которые, собственно, и реализуют поля. В этом определении, которое будем рассматривать применительно к семантике текста с ее нечеткими лингвистическими переменными, ключевыми являются два понятия: «набор» и «взаимодействие».

«Набор» предполагает существование механизма формирования – функции включения объекта в этот набор. Наиболее очевидным является критерий наличия у объекта характеристического свойства (в физике и математике – это функция), общего для всех объектов этого, по существу, класса. Однако, нет строгого правила, определяющего, что является первичным – свойство (дифференциальный признак класса) или набор: такой же набор мог бы быть сформирован и с использованием какого-либо другого признака, либо этот признак мог бы быть свойственен некоторому объекту другого класса. Кроме того, поскольку свойство само является производным значением (отображением объекта на предмет, результатом взаимодействия с измерительным инструментом), набор свойств, используемых в качестве классификационных признаков, тоже может быть разным. И отсюда следует возможная нечеткость класса.

«Взаимодействие» также надо рассматривать шире, чем как функцию, задающую поле. То есть существует некоторая среда, которая и создаёт поле, причем в ней есть взаимодействия и с объектами, не обладающими дифференциальным признаком класса, но без наличия таких взаимодействий не появится соответствующий объект поля. Причем и само множество взаимодействий между объектами поля определяется некоторым правилом более высокого порядка.

Для понятия «семантическое поле» более широким является понятие «ассоциативное поле» – набор объектов, связанных по какой-либо ассоциации, например, сходству, смежности, контрасту. Более узким можно считать понятие «контекстное поле», задающее, например, при документальном поиске семантику отбираемых документов через перечисление характеристических терминов.

Таким образом, можно сказать, что смысловым содержанием сообщения (документа) будет такое, которое соотносимо (сопрягаемо) с некоторым другим содержанием в заданном контекстном поле (наборе концептов, специфицирующих соответствующий смысл). Здесь уместно привести схему восприятия, приведенную в [3]: «Восприятие (узнавание) объекта активизирует концепт – сложную репрезентацию (ментальную структуру опыта в памяти). И в зависимости от того, какие элементы концепта и с какой степенью интенсивности активизируются при восприятии, итоговая конфигурация может включать в себя образ языкового объекта (например, слова, имени воспринимаемого объекта). Такой характер процесса восприятия, когда неязыковой объект вступает в знаковое отношение с языковым объектом, приводит к росту числа возможных интерпретаций знаковых отношений (семиотическое умножение мира), а потому особую значимость приобретает способность "извлечения" из знаков знания непротиворечивым образом».

<sup>3</sup>Примеры семантических полей: поле времени, поле химии, поле имен, поле глаголов движения, поле предлогов, поле суффиксов и т.п., для каждого из которых характерны некоторый дифференциальный признак и функция.

<sup>4</sup>Поле — это функция нескольких переменных: говорят, что поле задано на пространстве, которому принадлежат точки  $M$ , если задано правило, по которому каждой точке из  $M$  ставится в соответствии некоторая величина  $K$ , которая может быть и числом, и вектором, и тензором.

Отсюда следует существенное уточнение в терминологии: смысл приписывается знаку (т.е. реализуется некоторая функция поля) при его употреблении или формируется при восприятии, в результате чего знак «обретает» конкретное значение. Это – динамика, что противоположно статике и детерминированности, которые предполагает формулировка «знак имеет смысл».

## 2. Аналогии между физическими моделями и семантикой

Согласно [6] информация не тождественна информационному объекту (ИнфОб). Если ИнфОб ассоциируется с формой существования, то информация – с содержанием, т.е. действием, потенциально производимым этим объектом. С точки зрения физики процесса информация – это информационный объект во взаимосвязи с теми объектами, с которыми он может «информационно» взаимодействовать, или иначе – информационный объект и некоторое predetermined направление. То есть информация в фазе стабильного состояния (хранимый или передаваемый информационный объект) обладает свойствами макрообъекта, а в фазе взаимодействия – волновыми. До взаимодействия – это некоторый целостный объект, во время взаимодействия – это макрообразование «квантов». Причем эти связанные микрообъекты<sup>5</sup>, в тоже время могут использоваться в качестве элементов образов для других, существующих или будущих, объектов-оригиналов.

Отсюда следует, что информация может быть представлена как *суперпозиция возможных состояний информационного объекта* в  $n$  предметных областях, характеризуемая функцией:

$$A = c_1E_1 + c_2E_2 + \dots + c_nE_n,$$

где  $E_i$  – функция распределения<sup>6</sup> ИнфОб для  $i$ -ой предметной области;

$c_i$  – вероятностный<sup>7</sup> показатель реализации  $i$ -ой предметной области;

$A$  – характеристика актуального и потенциального разнообразия существования ИнфОб в рамках выделенных предметных областей. Поскольку не всегда возможно объективно оценить, насколько ИнфОб соответствует некоторой предметной области, то его функция распределения должна отражать и уже зафиксированные знания (актуальная, действительная часть распределения) и содержащиеся в неявной форме (потенциальная, мнимая часть).

Тогда *информационное взаимодействие* можно характеризовать функцией:

$$\Omega = c_1E_1\theta_1^T + c_2E_2\theta_2^T + \dots + c_nE_n\theta_n^T,$$

где  $\theta_i$  – функция распределения для информационного объекта, с которым взаимодействует исходный ИнфОб, для  $i$ -ой предметной области.

Последнее выражение отражает то, что информационное взаимодействие несимметрично (неассоциативно) и имеет «направленность», задаваемую для *воздействующего* информационного объекта вектором  $C$ . Именно при взаимодействии возникают неравновесные комбинации, создавая предпосылки для появления таких качеств, как ценность и новизна, которые станут характеристиками нового ИнфОб – результата взаимодействия. Направленность, возникающая во взаимодействии, будет создавать смысл, в котором будет использоваться ИнфОб.

<sup>5</sup> Для знаковой (текстовой) формы информации такими квантами являются понятия, обозначаемые в тексте терминами языка, которые могут использоваться в любых других текстах.

<sup>6</sup> Для текстовой (точнее, дескриптивной) формы представления информации, реализующих координатный метод идентификации содержания, пространство (первичные координаты) задается на дискретном множестве терминов, представляемых целостными понятийными конструкциями слов и их комбинаций (словосочетаниями, фразами, полными текстами). Функция распределения в этом случае может быть задана с помощью матриц типа «термин-документ» для соответствующих ПрО.

<sup>7</sup> В общем случае  $c_i$  – комплексное число, т.к. предметные области представлены в не одномерном пространстве и их выделение определяется известным знанием, т.е. необходимо предполагать, что существуют и другие «части» ПрО, «мнимые» для текущего состояния.

Появление у ИнфОб смысла позволяет говорить не просто о связанной с ним информации, но об информации семантической. Таким образом, семантическая информация – это информация в контексте некоторого смысла, определяющего множество элементов и их возможных состояний.

Как отмечалось ранее, смысл – это существо (архетип) – образ, построенный (на сообщении) в соответствии с некоторой схемой (точкой зрения, шаблоном), отражающей интересы приемника (его «направленностью»). Но не совсем корректно говорить «сообщение имеет смысл»: за редким исключением сообщение (если оно не тривиальное) «имеет» много смыслов. Значение слова есть вероятностная характеристика процесса и результата соотнесения его знакового образа с некоторым явлением окружающей действительности, и задается некоторой функцией распределения в многомерном пространстве.

### 3. Графы онтологий как инструменты отображения смысла

Так же, как и осмысленный текст, граф онтологии, построенный по этому тексту, может быть рассмотрен, как набор фактов, в совокупности выражающий некоторый смысл. Онтология, как семиотически целостное образование, определяется в [7] как совокупность трех взаимосвязанных систем

$$O = \langle S_f, S_c, S_t \rangle, \text{ где}$$

$S_f$  – функциональная система (объекты и связи действительности), определенная как  $S_f = \langle M_f, A_f, R_f, Z_f \rangle$ , где  $M_f$  – множество объектов (сущностей),  $A_f$  – множество характеристических свойств,  $R_f$  – множество функциональных отношений, представленных типизированными ситуативными связями, характерными для ПрО,  $Z_f$  – закон композиции, т.е. правила и схемы упорядочения объектов (например, мерономияПрО);

$S_c$  – понятийная система, определенная как  $S_c = \langle M_c, A_c, R_c, Z_c \rangle$ , где  $M_c$  – множество понятий ПрО,  $A_c$  – множество признаков систематизации понятий (таксономия ПрО),  $R_c$  – множество отношений (прежде всего, парадигматических),  $Z_c$  – закон композиции (схема представления);

$S_t$  – терминологическая система, определенная как  $S_t = \langle M_t, A_t, R_t, Z_t \rangle$ , где  $M_t$  – множество терминов,  $A_t$  – множество свойств,  $R_t$  – множество отношений эквивалентности и включения, а также лингвистических отношений,  $Z_t$  – закон композиции (грамматика);

$\equiv$  – операция сопоставления элементов различных систем на уровне знаков, обеспечивающая их тождество в функциональной, понятийной и терминологической системах.

В качестве модели данных функциональной системы онтологии используется помеченный (для вершин и дуг которого определены свойства  $A_f$ ) ориентированный граф  $G(V, E) = \langle V, E \rangle$ , где  $V$  – множество вершин, а  $E$  – множество дуг. Множество вершин и множество дуг в совокупности соответствуют множеству элементарных фактов.

На множествах  $V$  и  $E$  определяются (могут быть динамически построены):

1. Метаграф, формально задаваемый как  $MG = \langle V, MV, E, ME \rangle$ , где  $V$  – множество вершин,  $MV$  – множество метавершин,  $E$  – множество дуг,  $ME$  – множество метадугов. Каждая метавершина соответствует метасущности ситуативного факта и представляет собой граф  $mv_i = \langle V_i, E_i \rangle$ , где  $V_i \subset V$ ,  $E_i \subset E$ , а метадуга – метаотношению. Метаграф также обладает свойством мультиграфа, так как ситуативные факты могут, например, различаться только метаотношениями.

Для графовых форм, отображающих семантику текстов (и познания), наличие метавершин вполне конструктивно и естественно. Это, например, разрешение анафорических ссылок, использование определений понятий, именование агрегатов и т.п. Метавершина соответствует по своему имени сущности (понятию, узлу, композиции) и выступает в качестве атомарного семантического эквивалента смысла, определяемого неатомарной конструкцией (текстовым выражением и, соответственно, графом).

2. Гиперграф, когда на множестве вершин  $V \cup MV$  формируется множество гиперребер  $W$ , при этом в основе правил задания гиперребра лежат множества  $A_f$  и  $Z_f$ :

$$W = \{w_1, w_2, \dots, w_n\}, w_i = V_i \cup MV_i, \text{ где } V_i \subset V, MV_i \subset MV; V_i \neq \emptyset \forall MV_i \neq \emptyset.$$

3. Т.к. одна и та же пара атомарных сущностей может участвовать в нескольких элементарных фактах, в графе  $G(V, E)$  для двух вершин может существовать более одной дуги, то  $G(V, E)$  обладает свойством мультиграфа. Наличие свойства мультиграфа обусловлено тем, что онтология, как форма информации, обладает свойством эмерджентности и межотраслевым свойством.

То есть данные, представляющие онтологию, всегда интерпретируются в конкретном контексте, в итоге формируя конкретное знание. Но, поскольку содержание текста может использоваться для решения более чем одной задачи и, возможно, разными путями, это означает, что по одному тексту может быть построено несколько онтологий, соответствующих отдельным решениям (контекстам, аспектам и т.п.). Можно говорить, что на общем множестве сущностей и связей, выделяемых из одного текста, будет построен мультиграф – множество семантически самостоятельных графов, каждый из которых представляет свою целостную картину (онтологию) конкретной задачи.

Наличие в функциональной системе онтологии, помимо множеств сущностей и функциональных отношений, множества характеристических свойств и закона композиции позволяет группировать сущности не только в динамике, например, по принципу соответствия синтезируемой цепочки фактов, но и в статике – например, по принципу обладания общим свойством, по лексикографическому включению и т.п.

#### 4. Задание смысла на графах онтологий

Проиллюстрируем введенные определения для понятия «смысл» на примере применения операции семантического масштабирования – обобщения, через приведение к более высокому уровню общности, что соответствует трактовке понятия «смысл» как основного понятийного *существа* рассматриваемого объекта.

Для этого рассмотрим фрагмент текста, в котором описываются назначение и состав главного циркуляционного насосного агрегата:

«Главный циркуляционный насосный агрегат (ГЦНА) предназначен для создания циркуляции теплоносителя в первом контуре и отвода тепла из активной зоны реактора. ГЦНА имеет дополнительную функцию для обеспечения циркуляции теплоносителя на выбеге при различных авариях с обесточиванием. Дополнительная функция позволяет осуществлять плавный выход на режим естественной циркуляции.

Насос состоит из корпуса и выемной части. Выемная часть для обеспечения герметичности уплотняется медной прокладкой трапециевидального сечения. Корпус из теплоустойчивой стали марки 48ТС защищен изнутри нержавеющей наплавкой. К корпусу приварены опорные лапы. Опорные лапы опираются на фундаментную раму.

Выемная часть состоит из верхнего радиально-осевого подшипника, вала насоса, рабочего колеса, направляющего аппарата, станины и крышки с горловиной, сваренной из поковок, стали 48ТС. В горловине расположены торцевое уплотнение вала и ГСП. Вал насоса – цельнокованный и выполнен из стали 20Х13. Рабочее колесо с лопатками двойной кривизны сварено из двух частей: диска с лопатками и диска покрывного. Рабочее колесо и направляющий аппарат выполнены из нержавеющей стали 10Х18Н9ТЛ».

Граф, построенный по данному тексту, представлен на рис. 1. Для удобства восприятия некоторые вершины перемещены вручную.

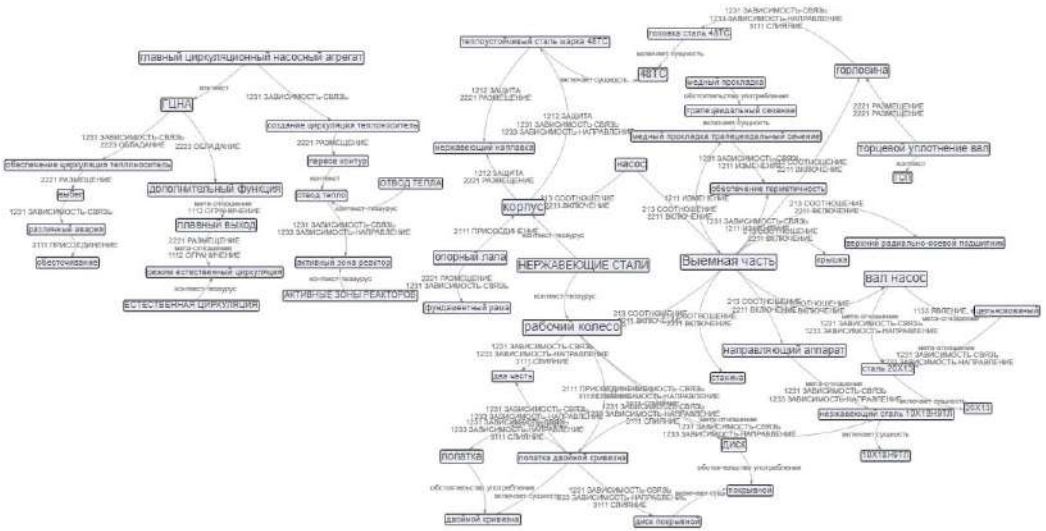


Рис. 1. Граф, построенный по тексту описания конструкции главного циркуляционного насоса

Механизм семантического обобщения смыслового содержания предполагает замену контекстно определенных имен сущностей понятиями более высокого уровня общности (например, из тезауруса предметной области).

В соответствии с [8] информационно-поисковые тезаурысы используются при формировании по тексту триплетов, положенных в основу графа, для построения структурно-лингвистических отношений «контекст-тезаурус». Такие отношения можно наблюдать, в том числе, и на рис. 1. Особенностью построения отношений является добавление в качестве вершин только тех дескрипторов тезауруса, длина которых (в словах) больше 1.

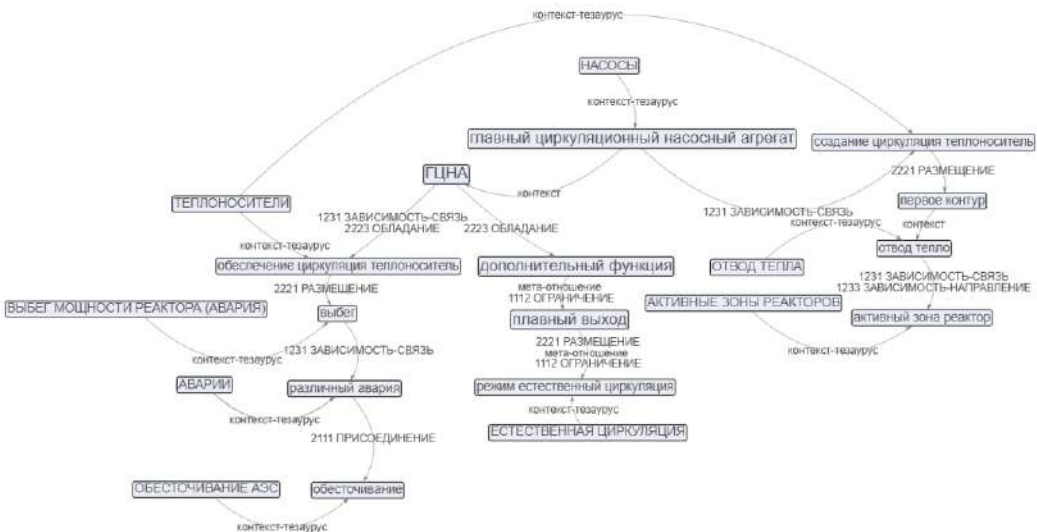


Рис. 2. Граф с присоединёнными терминами тезауруса. Исходный граф представлен на рис. 1, слева. Термины тезауруса приведены верхнем регистре (кроме вершины «ГЦНА»)

При семантическом обобщении дополнительно происходит отбор дескрипторов тезауруса, имеющих минимальное пересечение (но не менее одного слова) с терминами,

расположенными в вершинах графа. Все отобранные таким образом дескрипторы присоединяются к соответствующим вершинам через отношение «контекст – тезаурус». В случае наличия в тезаурусе аскрипторов и, соответственно, их отбора, происходит замена аскрипторов соответствующими им дескрипторами (через отношение *USE*).

Следующий шаг – поиск корней у ветвей иерархических деревьев тезауруса, начиная с вершин-дескрипторов, и замена всей ветви (включая исходную вершину отношения «контекст – тезаурус») корневым дескриптором.

Для иллюстрации семантического обобщения в качестве исходного графа рассмотрим левый подграф на рис. 1.

На первом шаге в результате присоединения дескрипторов тезауруса (использован тезаурус ИНИС МАГАТЭ [9]) мы получили граф, представленный на рис. 2.

Далее, вершина «главный циркуляционный насосный агрегат» будет заменена вершиной «ОБОРУДОВАНИЕ» (поскольку в тезаурусе есть связь «НАСОСЫ» – *BT* – «ОБОРУДОВАНИЕ»). Вершины «создание циркуляция теплоноситель» и «обеспечение циркуляция теплоноситель» заменяются вершиной «ТЕПЛОНОСИТЕЛИ» (вышестоящие дескрипторы отсутствуют). Вершина «отвод тепло» будет заменена вершиной «ОТВОД ТЕПЛА» (вышестоящие дескрипторы отсутствуют). Вершина «активный зона реактор» будет заменена на вершину «УЗЛЫ РЕАКТОРОВ» (в тезаурусе есть связь «АКТИВНЫЕ ЗОНЫ РЕАКТОРОВ» – *BT* – «УЗЛЫ РЕАКТОРОВ»). Вершины «выбег», «различный авария», «обесточивание» будут заменены одной вершиной «АВАРИИ» («ВЫБЕГ МОЩНОСТИ РЕАКТОРА (АВАРИИ)» – *BT* – «АВАРИИ», «ОБЕСТОЧИВАНИЕ АЭС» – *BT* – «АВАРИИ»). Вершина «режим естественный циркуляция» будет заменена вершиной «ПЕРЕДАЧА ЭНЕРГИИ» («ЕСТЕСТВЕННАЯ ЦИРКУЛЯЦИЯ» – *USE* – «ЕСТЕСТВЕННАЯ КОНВЕКЦИЯ» – *BT* – «ПЕРЕДАЧА ЭНЕРГИИ»).

Результат семантического обобщения представлен на рис. 3.



Рис. 3. Граф, построенный с применением механизма преобразования с помощью операции семантического обобщения. Исходный граф представлен на рис. 1, слева. Термины тезауруса приведены в верхнем регистре (кроме вершины «ГЦНА»)

Таким образом, применяя операцию обобщения, мы переходим к понятиям более высокого уровня абстракции (оборудование, передача энергии, узлы реакторов и т.д.) – формируем «обобщенный смысл» – существо объекта. Тем самым появляется возможность выявить содержательный смысл документа.



Для иллюстрации понятия «смысл», как предметного *существа* – содержания в некотором аспекте, используем механизм аспектной проекции<sup>8</sup> [10]. Рассмотрим подграф, расположенный в правой части рис. 1, где, например, конструкционные особенности главного циркуляционного насосного агрегата могут быть выявлены через применение аспектной проекции «Композиция (Отдельное – Целое)» (аспект задан отношениями, код которых начинается с цифры 2). Результат работы механизма аспектной проекции представлен на рис. 4.



Рис. 4. Граф, построенный с применением механизма преобразования с помощью операции аспектной проекции. Исходный граф представлен на рис. 1, слева

Использование проекции позволяет сократить число отображаемых вершин, облегчая восприятие графа без потери смысла описания насоса как совокупности элементов, его составляющих. Таким образом, проявляется идентифицирующий характер смысла («в смысле состава»).

## 5. Заключение

Семантическая информация – это информация в контексте некоторого смысла, определяющего множество элементов и их возможных состояний.

Смысл – это существо (архетип) – образ, построенный (на сообщении) в соответствии с некоторой схемой (точкой зрения, шаблоном), отражающей интересы приемника (его «направленность»). Смысл может иметь как содержательный, так и идентифицирующий характер.

Смысл документа появляется тогда, когда появляется возможность *сопоставления* элементов исходного множества знаков (документа) элементам системы онтологии знаний, т.е. если заданы соответствующие система онтологии знаний и правила сопоставления.

Смысл не существует сам по себе, он субъективен и не материален. Материализуется смысл в виде артефакта, но артефакт — это не смысл, это только его производная.

Работа выполнена при поддержке Министерства науки и высшего образования РФ (проект государственного задания № 0723-2020-0036).

<sup>8</sup> Задание аспекта сводится к формированию функциональной системы  $S_f = \langle M_f^i, A_f^i, R_f^i, Z_f^i \rangle$  с непустым множеством  $R_f^i$  и возможно непустыми множествами  $M_f^i, A_f^i$  и/или  $Z_f^i$ , где  $R_f^i$  – множество отношений,  $M_f^i$  – множество опорных понятий,  $A_f^i$  – множество свойств и  $Z_f^i$  – множество правил построения цепей на графе онтологии. Множество опорных понятий  $M_f^i$  и множество отношений  $R_f^i$  задаются в нормальной форме, что не исключает наличия условий при формировании цепи на графе онтологии, связанных с формой, склонением и спряжением опорных понятий и отношений.

## Литература

- [1] Максимов Н. В., Голицына О. Л. От семантического к когнитивному информационному поиску. Основные положения и модели глубинного семантического поиска // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2022. № 6. С. 1-16.
- [2] Zlatev J. Meaning = Life (+ Culture): An outline of a unified biocultural theory of meaning // Evolution of Communication. 2000. Vol. 4. № 2. P. 253-296.
- [3] Кравченко А. В. Методологические основания когнитивного анализа значения // Когнитивный анализ слова. Иркутск: Изд-во ИГЭА, 2000. С. 8-32.
- [4] Максимов Н.В., Лебедев А.А. О природе и определениях информации: физика и семантика // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2020. № 7. С. 1-12.
- [5] Философский энциклопедический словарь: Советская энциклопедия / Л. Ф. Ильичёв, П. Н. Федосеев, С. М. Ковалёв, В. Г. Панов. М.: Советская энциклопедия, 1983. 840 с.
- [6] Максимов Н.В. Информация и знания: природа, концептуальная модель // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2010. № 7. С. 1-10.
- [7] Голицына О.Л., Максимов Н.В., Окропишина О.В., Строгонов В.И. Онтологический подход к идентификации информации в задачах документального поиска // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2012. № 5. С. 1-9.
- [8] Голицына О. Л., Гаврилкина А. С. Об одном подходе к выделению имён сущностей и связей в задаче построения семантического поискового образа // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2021. № 3. С. 17-26.
- [9] INIS Multilingual Thesaurus. – URL: <https://inis.iaea.org/search/thesaurus.aspx> (дата обращения 06.04.2022).
- [10] Максимов Н.В., Голицына О.Л., Монанков К.В., Лебедев А.А., Баль Н.А., Кюрчева С.Г. Средства семантического поиска, основанные на онтологических представлениях документальной информации // Научно-техническая информация. Серия 2: Информационные процессы и системы. 2019. №7. С. 8-19.

### On One Approach to Semantic Information Definition

N.V. Maksimov, A.A. Lebedev

National Research Nuclear University MEPhI

The "semantic information" concept is defined through the "sense" concept definition. The question of "sign" – "sense" – "meaning" concepts correlation is analyzed. The fundamental factor for semantics manifestation is the presence of distinguished activity direction – a goal, to which the signs system (informational objects) projection is made. For the semantic information, directionality arises in interaction of the original informational object (message, text) and the target one (which can be represented, in particular, as an aspect ontology). The semantic information is information in context of some sense, defining an elements set and their possible states.

**Keywords:** semantics, sense, semantic information

**Reference for citation:** Maksimov N.V., Lebedev A.A. On One Approach to Semantic Information Definition // Computer Linguistics and Computing Ontologies. Vol. 6

(Proceedings of the XXV International Joint Scientific Conference «Internet and Modern Society», IMS-2022, St. Petersburg, June 23-24, 2022). - St. Petersburg: ITMO University, 2022. P. 30 – 40. DOI: 10.17586/2541-9781-2022-6-30-40

## Reference

- [1] Maksimov N. V., Golitsina O. L. Ot semanticheskogo k kognitivnomy informacionnomy poisky. Osnovnye pologeniya i modeli glybinno semanticheskogo poiska // Nauchno-technicheskaya informaciya. Seriya 2: Informacionnye processy i sistemy. 2022. № 6. P.1-16.
- [2] Zlatev J. Meaning = Life (+ Culture): An outline of a unified biocultural theory of meaning // Evolution of Communication. 2000. Vol. 4. № 2. P. 253-296.
- [3] Kravchenko A. V. Metodologicheskieosnovaniyakognitivnogoanalizaznacheniya // Kognitivnyjanalizslova. Irkutsk: Izd-vo IGEA, 2000. P. 8-32.
- [4] Maksimov N.V., Lebedev A.A. On the nature and definitions of information: physics and semantics // Automatic Documentation and Mathematical Linguistics. 2020. Vol. 54. P. 185–195.
- [5] Filosofskijenciklopedicheskij slovar': Sovetskayaenciklopediya / L. F. Il'ichyov, P. N. Fedoseev, S. M. Kovalyov, V. G. Panov. M.: Sovetskayaenciklopediya, 1983. 840 p.
- [6] Maksimov N.V. Information and knowledge: nature and the conceptual model // Automatic Documentation and Mathematical Linguistics. 2010. Vol. 44. P. 177–186.
- [7] Golitsyna O.L., Maksimov N.V., Okropishina O.V., Strogonov V.I. The ontological approach to the identification of information in tasks of document retrieval // Automatic Documentation and Mathematical Linguistics. 2012. Vol. 46. P. 125–132.
- [8] Golitsyna O.L., Gavrilkina A.S. On one approach to the extraction of entity and relationships names in the task of building a semantic search image // Automatic Documentation and Mathematical Linguistics. 2021. Vol. 55, № 2. P. 54-62.
- [9] INIS Multilingual Thesaurus. – URL: <https://inis.iaea.org/search/thesaurus.aspx> (access date 06.04.2022).
- [10] Maksimov N. V., Golitsina O. L., Monankov K. V., Lebedev A. A., Bal N. A., Kyurcheva S. G. Semantic search tools based on ontological representations of documentary information// Automatic Documentation and Mathematical Linguistics. 2019. Vol.53. P. 167-178.