

ПЕРЕЧЕНЬ ССЫЛОК

1. Лацис А. Как построить и использовать суперкомпьютер. – М.: Бестселлер, 2003. – 240 с.
2. Букатов А. А., Дацюк В. Н., Жегуло А. И. Программирование многопроцессорных вычислительных систем. – Ростов-на-Дону. Издательство ООО «ЦВВР», 2003. – 208 с.
3. Таненбаум Э. Современные операционные системы. 2-е изд. – СПб.: Питер, 2002. – 1040 с.

Надійшла 25.09.07

Розроблена математична модель оптимального розподілу завдань для паралельної обчислювальної системи

з урахуванням завантаженості її обчислювальних вузлів. Представлені результати експериментальних досліджень моделі, які демонструють значний приріст продуктивності в порівнянні з рівномірним розподілом паралельного завдання по вузлах кластера.

The mathematical model of tasks optimum distribution for the parallel computing system is developed in view its computing nodes load. Results the model experimental researches which show a significant increase of productivity in comparison with uniform distribution of the parallel task on nodes of a cluster are submitted.

УДК 004.087

А. А. Егошина

ФОРМАЛЬНАЯ МОДЕЛЬ СЛОВООБРАЗОВАТЕЛЬНОЙ СЕМАНТИКИ

В статье рассмотрен вопрос формализации семантики естественного языка. Предложена формальная модель словообразовательной семантики, учитывающая специфику словообразовательного значения в сравнении с заданными грамматическими и лексическими значениями производного слова.

ВВЕДЕНИЕ

Семантический компонент уже достаточно давно признается необходимой частью полного описания языка. Семантика как раздел лингвистики отвечает на вопрос, каким образом человек, зная слова и грамматические правила какого-либо естественного языка, оказывается способным передать с их помощью самую разнообразную информацию о мире (в том числе и о собственном внутреннем мире), даже если он впервые сталкивается с такой задачей, и понимать, какую информацию о ми-

ре заключает в себе любое обращенное к нему высказывание, даже если он впервые слышит его.

Семантика как информация, передаваемая языком или какой-либо его единицей (словом, грамматической формой слова), представляет собой не жестко детерминированную систему. Ячейка семантики (полнозначное слово) организована по принципу «семантического треугольника» [1] и может быть схематично представлена следующим образом.

Схема, представленная на рис. 1, резюмирует семантические отношения, т. е. аналогичным образом семантика организована во всех единицах языка.

Свой вклад в формирование общих принципов семантического описания вносят разные теории языка. Семантика как наука начала развиваться еще во второй половине 19 века. Фундаментальные лингвистико-гносеологические концепции В. Фон Гумбольта, А. А. Потебни, В. Вудта и др. определили первый этап развития

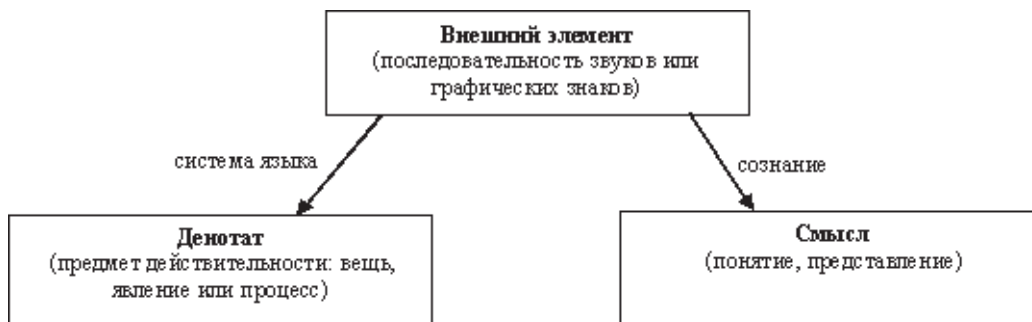


Рисунок 1 – Семантический треугольник

© Егошина А. А., 2007

семантики. Второй этап ознаменовался выделением семантики в особую часть языкознания (труды М. М. Покровского и других русских и немецких ученых). Третий этап развития семантики начинается приблизительно в двадцатых годах 20 века. Он характеризуется сближением семантики с логикой и ориентацией на синтаксис. К началу семидесятых годов прошлого века устанавливаются более гармоничный и комплексный подход к семантическим явлениям. В работах В. А. Звегинцева, Ю. Д. Апресяна и др. основной ориентацией становится анализ не абстрактного, изолированного предложения, а рассматривается предложение в реальной речи с учетом прагматики языка. А. В. Бондарко и Т. В. Бултина уже исследуют семантику морфологических форм.

ПОСТАНОВКА ЗАДАЧИ

Общеизвестен тот интерес к семантике, который характерен для современного этапа развития лингвистики. Интеллектуальные системы с естественно-языковым интерфейсом, который включает активное словообразование, в этом смысле не составляют исключения. Описание модуля словообразования является важной составной частью описания языковой системы в целом. Словообразование – одно из основных средств пополнения словарного состава языка новыми словами. Оно тесно связано с лексикой и грамматикой. Иногда в современной лингвистике словообразование рассматривают как часть лексикологии. Однако более характерен для русской лингвистической традиции грамматический подход к словообразованию, когда оно включается в состав грамматики. Основанием для этого является тот факт, что словообразовательные средства языка – это морфемные средства, и поэтому механизм словообразования во многом сближается с механизмом морфологического уровня языка, в результате чего проблемы словообразования тесно переплетаются с проблемами морфологии.

В [2] предлагается формализация семантики путем сведения описания производных (мотивированных) слов к описанию мотивирующих, более простых по морфемному составу. В работе [3] полагается, что если целью грамматики является построение правил порождения текста на данном языке, то в такую грамматику войдут лишь семантически регулярные конструкции. Однако автор [4] утверждает, что словообразовательная семантика может быть предметом описания как грамматики, порождающей текст, так и аналитической грамматики, служащей различным целям – теоретико-познавательным, типологическим и др. Подходы, предложенные в рассмотренных работах, не являются эффективными из-за необходимости хранения полных значений словоформ мотивирующих слов.

В предыдущих работах автора решены задачи разработки структуры лингвистического процессора интеллектуальной системы с естественно-языковым интерфейсом на основе многокомпонентного словаря [5]. Применение подсловарей морфем позволяет существенно сократить затраты памяти и время поиска [6]. Целью настоящего исследования является формализация словообразовательной семантики в интеллектуальной системе с морфемы словарями.

ФОРМАЛИЗАЦИЯ СЛОВООБРАЗОВАТЕЛЬНОЙ СЕМАНТИКИ

Формализация особенностей организации семантических единиц словообразования позволяет автоматизировать процесс построения производных слов с учетом заданной специфики грамматического и лексического значений.

Одной из основных трудностей формализации словообразовательной семантики является создание правил порождения семантически нерегулярных конструкций. Однако для решения таких задач как построение автоматизированных систем перевода, поисковых систем и интеллектуальных систем обучения языку автором предлагается рассматривать правила порождения семантически нерегулярных конструкций не в их реальном значении, а в «структуральном» значении, т. е. в значении, равном сумме значений составляющих слово морфем.

Для решения данной задачи необходимо знать значение этих морфем, в том числе и инвариантных. Значение инвариантных морфем – это тот участок семантико-словообразовательного поля, в котором находятся и контекстные значения данной морфемы. Этот участок очерчен теми семантическими границами, в пределах которых возможны новообразования с помощью данной морфемы.

Формально определим семантику каждого слова X языка как функцию $f(X)$. Слово x представляет собой упорядоченное множество морфем, т. е.

$$X = (X_1, X_2, X_3), \quad (1)$$

где X_1 – множество приставок, входящих в слово. $X_1 = (x_1^1, x_1^2, \dots)$. Так как слово может не содержать приставку, то $|X_1| \geq 0$ – мощность множества X_1 ;

X_2 – множество корней, входящих в слово. $X_2 = (x_2^1, x_2^2, \dots)$. Так как слово обязательно содержит хотя бы один корень, $|X_2| > 0$;

X_3 – множество суффиксов, входящих в слово. $X_3 = (x_3^1, x_3^2, \dots)$. Так как слово может не содержать суффикс, то $|X_3| \geq 0$.

Исходным пунктом классификации слов обычно является деление на части речи. Производные слова определенной части речи связаны между собой многообразными отношениями словообразовательного характера. лишь некоторые префиксы и единичные суффиксы (например, *-оньк*) выступают одновременно в разных частях речи. Остальные аффиксы присущи словам только одной части речи. разного вида семантические взаимодействия (совмещение значений и их отдельных компонентов), взаимоотношения между производными словами разных способов словообразования осуществляются только в пределах одной части речи. Таким образом, семантическое значение аффикса есть функция двух переменных: аффикса и части речи, т. е. $f(x_j^i, z_k)$, где z_k – элемент множества всех частей речи Z .

Пусть M – множество всех слов языка, тогда справедливо следующее

$$\forall x \in M f(x) = f(x_1^i, z_k) \cup f(x_2^j, z_m) \cup f(x_3^h, z_n). \quad (2)$$

В качестве примера рассмотрим семантику слова *белеть* (корень слова *бел-*, *-e* – суффикс, означающий абсолютное становление признака, заданного корнем, *-ть* – глагольный суффикс, указывающий на то, что выполняется действие).

X_1 – пустое множество, так как слово не имеет приставки;

$$X_2 = \{x_2^1\};$$

$$x_2^1 = \text{бел};$$

$$X_3 = \{x_3^1, x_3^2\};$$

$$x_3^1 = e;$$

$$x_3^2 = \text{ть};$$

$$Z = \{z_1\};$$

$$z_1 = \text{глагол};$$

$$f(x) = f(x_2^1, z_1) \cup f(x_3^1, z_1) \cup f(x_3^2, z_1);$$

$$f(x) = \text{«становится абсолютно белого цвета»}.$$

Учитывая изложенные выше позиции, можно определить причинно-следственную связь между семантической и морфологической формами слова. Определим процесс словообразования как

$$G = (I, O, \varphi), \quad (3)$$

где I – множество входных воздействий; O – множество мотивированных слов, образующих словообразовательное гнездо; φ – функция перехода (закон словообразования), задающая отображение $\varphi : I \rightarrow O$.

Множество входных воздействий определяется следующим образом:

$$I = (C, z, X), \quad (4)$$

где C – множество ключевых слов, определяющих семантику производного (мотивированного) слова;

z – часть речи;

X – исходное (мотивирующее) слово.

При выборе аффикса, необходимого для данного закона словообразования, учитывается его семантическое значение, ключевые слова которого должны соответствовать множеству C , т. е., если x^i – некоторый аффикс, то $f(x^i, z) \cap C \neq \emptyset$.

Рассмотрим предложенный подход на примере. Пусть задано следующее множество входных воздействий I :

$C = \{\text{лицо, выполняющее действие}\};$

$z = \{\text{существительное}\};$

$X = \{\text{преподавать}\};$

φ : «сохранение конечной гласной инфинитивной основы с добавлением суффикса»;

x^i – суффикс *-тель*;

$f(x^i, z) = \{\text{лицо или предмет, выполняющее или производящее действие}\};$

$$f(x^i, z) \cap C \neq \emptyset;$$

$$\varphi : I \rightarrow O = \text{преподаватель}.$$

ЗАКЛЮЧЕНИЕ

Таким образом, в настоящей работе построена формальная модель словообразовательной семантики, учитывающая специфику словообразовательного значения в сравнении с заданными грамматическими и лексическими значениями производного слова.

В дальнейшем планируется модификация модели с целью детализации словообразовательных категорий и учета синонимии и омонимии словообразовательных средств. Полученные результаты могут применяться при автоматизированном переводе и при разработке интеллектуальных информационно-поисковых и обучающих систем.

ПЕРЕЧЕНЬ ССЫЛОК

1. Лингвистический энциклопедический словарь / Гл. ред. В. Н. Ярцева. – М.: Сов. энциклопедия, 1990. – 685 с.: ил.
2. Тузов В. А. Математическая модель языка. – Л.: Изд-во Ленингр. ун-та, 1984. – 176 с.
3. R. Grzegorzczkova, J. Puzinina. Gramatyka opisowa a slowotworstwo // Biuletyn polskiego towarzystwa jezykownawczego. – XXX. – С. 82.
4. Улханов И. С. Словообразовательная семантика в русском языке. – М.: Наука, 1977. – 255 с.
5. Егошина А. А. Языковые и алгоритмические аспекты построения морфологических процессоров для интеллектуального поиска в полнотекстовых базах данных // VI международная конференция «Интеллектуальный анализ информации ИАИ-2006»: Киев, 16–19 мая 2006 г.: Сб. тр. / Рос. ассоц. искусств. интеллекта и др.; Под ред. Т. А. Таран. – К.: Просвіта, 2006. – 334 с.: ил.
6. Егошина А. А., Об одном способе построения статического словаря морфологического процессора // Материалы Седьмой Международной научно-техни-

ческой конференции „Искусственный интеллект. Интеллектуальные и многопроцессорные системы – 2006». – Т.2. – Таганрог: Изд-во ТРТУ, 2006. – 404 с.

Надійшла 13.11.07

В статті розглянуте питання формалізації природної мови. Надана формальна модель словотворчої семантики, яка ураховує специфіку словотворчого значення в порів-

нянні з заданими граматичними та лексичними значеннями похідного слова.

In the article the formalization of semantics of natural language is considered. The formal model of semantics of creation of words is offered which takes into account word-formation specificity of meaning in comparison with the given grammatic and lexical meanings of a derivative word.

УДК 681.391

В. Н. Журавлев, Е. А. Архипова

АНАЛИЗ МЕТОДА РАСЧЕТА ПАРАМЕТРА ЭФФЕКТИВНОСТИ МАСКИРОВАНИЯ РЕЧИ В ТЕХНИЧЕСКИХ КАНАЛАХ УТЕЧКИ

В статье проведен анализ метода цифровой корреляционной обработки контрольного фрагмента речи, который позволяет на основе расчета коэффициента корреляции обосновать аналитическую оценку параметра эффективности аддитивного маскирования речевых сигналов.

ВВЕДЕНИЕ. ПОСТАНОВКА ЗАДАЧИ

Среди разнообразных средств спецтехники в последние годы интенсивно развивается и совершенствуется рынок устройств, предназначенных для несанкционированного доступа (НСД) к конфиденциальной информации, содержащейся в речевом сигнале (РС), который является первичным общедоступным выражением результата мыслительного процесса человека.

При проектировании систем технической защиты речевой информации в выделенных помещениях (ВП) от утечки по каналам несанкционированного доступа (НСД) технической разведки противника (ТРП) основным параметром, определяющим, как показатель эффективности, так и категорию защиты, служит параметр словесной разборчивости $W(t)$. Данный параметр предлагается рассчитывать инструментально-расчетным методом [1–3], предполагающим следующие основные допущения.

1. Тестовый сигнал (модель РС) $si(\Delta\omega, t)$, маскирующий сигнал (МС) $sn(\Delta\omega, t)$ и сигнал канала утечки $su(\Delta\omega, t) = si(\Delta\omega, t) + sn(\Delta\omega, t)$ представляют собой случайные процессы с нормальным законом распределения плотности вероятности, при этом

$$\Delta\omega \times t \geq 10, \quad t \in [t_1, t_1 + T_a], \quad (1)$$

где $\Delta\omega$ – эффективная полоса частот РС, t_1 – время начала процесса анализа, T_a – интервал анализа эф-

фективности маскирования, который не превышает среднего интервала времени между ложными тревогами и выбирается из ряда фиксированных значений $T_a = (1, 5, 10, 30)$ секунд.

2. Тестовый и маскирующий сигналы представляют собой сигналы с равномерным распределением спектральной плотности мощности $N(\Delta\omega, t)$, т. е. со спектром «белого шума».

3. Точка канала НСД, в которой установлены средства ТРП, расположена на расстоянии 1 метр от источника тестового сигнала $si(\Delta\omega, t)$, при этом акустическое поле сигнала $su(\Delta\omega, t)$ полагается плоским с полностью активным акустическим сопротивлением среды распространения.

4. На основании предыдущего допущения, в части плоского акустического поля, а так же постоянства АЧХ рецептора средства ТРП на интервале времени T_a , отношение сигнал – помеха $SN(\Delta\omega, t)$ рассчитывается не как классическое [4] отношение эффективных энергий сигнала и помехи, поглощающихся рецептором средства ТРП за время интервала анализа:

$$SN(\Delta\omega, t) = 10 \lg \frac{\int_{t_1}^{T_a} p_{si}(\Delta\omega, t) dt}{\int_t^{T_a} p_{sn}(\Delta\omega, t) dt}, \quad t \in [t_1, t_1 + T_a], \quad (2)$$

где $p_{si}(\Delta\omega, t)$ и $p_{sn}(\Delta\omega, t)$ – мгновенные мощности сигнала и помехи, соответственно, а как логарифм отношения усредненных на интервале времени T_a акустических давлений, что возможно только для моногармонических анализируемых сигналов при активном входном сопротивлении рецептора давления.