

УДК 510.62

Н. В. РЯБОВА, Н. В. ШАРОНОВА

О МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ ОТСУБСТАНТИВНЫХ ИМЕН СУЩЕСТВИТЕЛЬНЫХ РУССКОГО ЯЗЫКА

Развитие вычислительной техники достигло того уровня, когда ЭВМ все больше используется для решения сложных, плохо поддающихся формализации, интеллектуальных задач. К ним относятся распознавание на ЭВМ смысла документа, машинный перевод с иноязычного текста, человеко-машинный диалог на естественном языке. Проблема формализованного описания деривационных процессов, и в частности деривационной номинации,— одна из самых актуальных. Это обусловлено тем, что деривационная номинация — основное средство обогащения лексического состава языка. Напомним, что под номинацией понимается «назечение предметов и ситуаций с помощью языковых средств, закрепление за определенным референтом того или иного специального знака» [1]. Распространенные способы номинации — словообразовательные способы. Среди них выделяется прежде всего словопроизводство — построение слов с помощью аффиксов. Самую распространенную систему аффиксов имеет имя существительное, в то же время оно — «наиболее номинативная часть речи, поскольку уже само выделение существительного как части речи происходит на основании функции называния» [2]. Отсюда следует вывод об актуальности формализации номинативного словообразования имен существительных, особенно тех, которые наделены «предметным признаком», т. е. полученных в результате суффиксальной деривации при субстантивной производящей основе.

Отсубстантивные имена существительные представлены наименованием лиц и предметов по предмету, как-то с ними связанному. В их образовании чаще всего используются суффиксы *-ник* (с его вариантами *-овник*, *-арник*, *-атник*), *-чик/-щик*, *-ец*. Например: *сырник*, *ельник*, *шиповник*, *волчатник*, *ватник*, *тральщик*, *датчик*, *малинник*, *печник*, *лесник*, *терновник* и т. д.

Рассмотрим словообразовательную номинацию предметов. Для конкретизации и полноты построения математической модели возьмем только суффиксальные производные, полученные от субстантивных основ. Самый высокоупотребимый деривационный суффикс данного класса производных — суффикс *-ник*. При построении математической модели нами было выделено 19 «семантических ролей» этого суффикса. Отметим, что выделяемые «семантические роли», необходимые для формализации описания, зачастую являются гораздо более дробными, чем словообразовательные значения, приводимые, например, в работе [3]. При этом мы берем словообразовательное значение как «обобщенное структурно-лексическое значение, его инвариант».

В качестве формального аппарата описания используется алгебра конечных предикатов [4], которая позволяет записывать отношения на множестве деривационных морфов и ранее успешно применялась при описании ряда словоизменительных и словообразовательных процессов. Согласно [5] мы понимаем словообразовательный суффикс как фрагмент текста, обозначаемый X . Смысл фрагмента текста, вводимый при замене действия отброшенной части контекста, понимается как переменный вектор $Y = (y_1, y_2, \dots, y_n)$, где n — число компонентов смысла. Между X и Y существует связь, записываемая в виде отношения $L(X, Y)$. Поскольку в роли X выступает часть слова (деривационный суффикс), рассматриваем это отношение как морфологическое. При описании будем придерживаться принципа однозначности, согласно которому текст всегда однозначно зависит от его смысла, интерпретируем этот принцип как требование полноты набора признаков — компонентов смысла. С целью ограничения и конкретизации задачи вводим следующие фиксированные значения признаков: $\xi_1 = c$ (тип грамматических преобразований со значением «словообразование»); $\xi_2 = c$ (способ словообразования со значением «суффиксальный»); $\xi_3 = c$ (часть речи производящей основы со значением «существительное»); $\xi_4 = c$ (часть речи производного со значением «существительное»); $\xi_5 = n$ (признак одушевленности со значением «неодушевленный»). Введем также переменные признаки: x_1 — «абстрактности» со значениями a — абстрактный, k — конкретный; x_2 — «рукотворности» со значениями p — рукотворный, n — нерукотворный; x_3 — «единичности» со значениями e — единичный, n — неединичный; x_4 — «целый или часть предмета» со значениями c — целый, $ч$ — часть предмета; x_5 — «совокупности» со значениями c — совокупность, n — нет; x_6 — «содержания» со значениями x_6^c — «содержит в себе то, что названо в основе», x_6^n — «нет»; x_7 — «предназначения» со

значениями x_7^n — «предназначен для того, что названо в основе», x_7^H — «нет»; X_8 — «стилевой характеристики» со значениями с — специальный, терминологический, р — разговорный, о — областной, п — просторечье. Кроме введенных семантических признаков нам необходимо также использовать ряд признаков лексико-семантического свойства. Это связано со спецификой внутрикатегориальной деривации, когда ряд аффиксов, например *-ник*, несет на себе только общее категориальное значение «предмета, связанного с предметом, названным в основе». Конкретное же семантическое значение этот аффикс приобретает только при соединении с конкретной основой, причем это значение может не входить ни в значение основы, ни в значение аффикса. Для выявления всех семантических ролей деривационных аффиксов введем следующие значения лексико-семантических признаков: t_1 — «участок земной поверхности, территории, пространство» со значениями t_1^n — «участок поверхности, покрытый тем, что названо в основе»; t_1^p — «участок поверхности, на котором растет то, что названо в основе»; t_2 — «части тела, организма человека или животного» со значениями t_2^d — «да», t_2^H — «нет»; t_3 — «продукты питания» со значениями t_3^d — «да», t_3^H — «нет»; t_4 — «предметы быта, обихода» со значениями t_4^n — «посуда, утварь», t_4^o — «одежда», t_4^c — «санитария, гигиена»; t_5 — «продукты интеллектуальной деятельности» (в данном случае «книги») со значениями t_5^n — «с обязательным порядком перечисления того, что названо в основе», t_5^H — «нет»; t_6 — «сооружения, постройки, помещения» со значениями t_6^d — «да», t_6^H — «нет». Следует отметить, что при разработке системы признаков были использованы данные традиционной грамматики [3], а также последние разработки в области лингвистики [6].

Семантический признак мы рассматривали как переменную, заданную на некотором множестве смысловых оттенков. При таком подходе, чтобы сформировать семантический признак, достаточно указать для него множество смысловых оттенков — значений признака. Разрабатывая семантические оттенки, мы использовали прием, частично употребляемый в описательной форме в традиционной лингвистике. Он состоит в попытке представить «значение производного слова дефиницией, содержащей исходное или мотивирующее слово или слова» [1]. Мы разработали конструкцию, называемую «пробный текст», в которой устанавливается однозначная зависимость между производящей основой Z , деривационным суффиксом Y и производным словом ZY . Например, текст «шиповник обладает шипами» представляем в виде конструкции « ZY обладает Z -ами». Выделяя множество текстов с такими конструкциями, мы получаем оттенок «обладания». Рассмотрим другой пример: «чайник предназначен для чая» соответствует конструкции « ZY предназначен для Z », таким приемом выделяется оттенок «предназначения». Данный подход позволяет четко разграничивать

деривационные суффиксы в их конкретных семантических ролях, поскольку один и тот же суффикс в разных производных словах может характеризоваться различными наборами смысловых оттенков. С помощью введенной системы признаков устанавливается однозначное соответствие между деривационным суффиксом, с одной стороны, и конкретным набором семантических признаков, с другой. Покажем это на примерах. Обозначим буквы суффикса через y_1, y_2, y_3, y_4, y_5 . Таким образом, имеем набор переменных y_i ($i = \overline{1,5}$) в соответствующих областях определения, задаваемыми уравнениями

$$\begin{aligned} y_1^0 \vee y_1^a \vee y_1^{\square} \vee y_1^{\text{я}} \vee y_1^{\text{е}} &= 1; y_2^b \vee y_2^{\text{м}} \vee y_2^{\square} = 1; \\ y_3^{\text{н}} \vee y_3^{\square} &= 1; y_4^{\text{н}} \vee y_4^{\text{а}} \vee y_4^{\text{я}} &= 1; y_5^{\text{а}} &= 1. \end{aligned} \quad (1)$$

Суффикс *-ник* (с его вариантами *-овник*, *-арник*, *-атник*) как наиболее продуктивный встречается в различных словообразовательных типах. Покажем, как выражается его словообразовательная семантика с помощью введенной системы признаков. Пусть заданы значения признаков: $X_2^{\text{н}}$ — нерукотворный, $X_3^{\text{е}}$ — единичный, $X_4^{\text{а}}$ — целый. В указанном случае производное слово имеет семантику «единичное растение». Уравнение имеет вид

$$x_2^{\text{н}} x_3^{\text{е}} x_4^{\text{а}} = (y_1^0 \vee y_1^a \vee y_1^{\square}) (y_2^{\square} \vee y_2^{\text{м}} \vee y_2^{\text{б}} \vee y_2^{\text{я}}) y_3^{\text{н}} y_4^{\text{н}} y_5^{\text{а}}. \quad (2)$$

На выбор конкретного деривационного морфа оказывает влияние производящая основа («ближний текст»). Правило таково: если производящая основа односложная, мужского рода, неодушевленная, то к ней присоединяется морф *-овник*, в остальных случаях — морф *-ник*. Обозначим это правило через A . Тогда правило выбора морфа *-овник* запишется в виде уравнения

$$A x_1^{\text{н}} x_2^{\text{н}} x_3^{\text{е}} x_4^{\text{а}} = y_1^0 y_2^{\text{б}} y_3^{\text{н}} y_4^{\text{н}} y_5^{\text{а}}. \quad (3)$$

Например: *шиповник*, *крестовник*, но *лимонник*, *морковник*.

Пусть теперь заданы следующие значения признаков: $x_1^{\text{к}}$ — «конкретный», $x_2^{\text{н}}$ — «нерукотворный», $x_3^{\text{е}}$ — «неединичный», $x_5^{\text{е}}$ — «совокупность». На таком наборе признаков имеем два вида производных: первый — названия совокупностей деревьев, кустарников, растений (*березник*, *ельник*, *малинник*), второй — части тела, организма человека или животного (*костяк*, *кишечник*). Зададим значение признака $t_2^{\text{а}}$. Тогда будем иметь уравнение

$$x_1^{\text{к}} x_2^{\text{н}} x_3^{\text{н}} x_5^{\text{е}} t_2^{\text{а}} = y_1^{\square} y_2^{\square} (y_3^{\text{н}} \vee y_3^{\square}) (y_4^{\text{н}} \vee y_4^{\text{я}}) y_5^{\text{а}}. \quad (4)$$

Чтобы снять неоднозначность в третьей и четвертой буквах суффикса, введем значение переменной X_8 — «тилевая характеристика со значением «просторечный»». Тогда условие выбора суффикса *-як* запишется

$$x_1^{\text{к}} x_2^{\text{н}} x_3^{\text{н}} x_5^{\text{е}} t_2^{\text{а}} = y_1^{\square} y_2^{\square} y_3^{\square} y_4^{\text{я}} y_5^{\text{а}}. \quad (5)$$

Если признак X_2 принимает значение «рукотворный», то с помощью суффикса *-ник* и его вариантов образуются названия конкретных предметов (в широком смысле), изготовленных человеком в процессе его жизнедеятельности. Задав значения признаков: x_2^p — «рукотворный», x_6^c — «содержит в себе то, что названо в основе» (обладает признаком «содержания»), t_3^n — «продукты питания», имеем следующую группу производных: *сырник, творожник, морковник* и т. п. Уравнение таково:

$$x_1^k x_2^p x_6^c t_3^n = y_1^{\cup} y_2^{\cup} y_3^n y_4^y y_5^k. \quad (6)$$

При комбинации наборов признаков «рукотворный» (обладает признаком «содержания»), t_4^o — «одежда»), имеем уравнение

$$x_1^k x_2^p x_6^c t_4^o = y_1^{\cup} y_2^{\cup} y_3^n y_4^y y_5^k. \quad (7)$$

Производное из этой группы — *ватник*.

Таким образом, на примерах мы убедились, что между конкретными наборами смысловых переменных и деривационными аффиксами в их конкретных семантических ролях устанавливается однозначное соответствие. Для более полного описания необходимо учитывать влияние «ближнего текста», т. е. производящей основы.

Список литературы: 1. Кубрякова Е. С. Части речи в ономаσιологическом освещении. — М.: Наука, 1978. — 113 с. 2. Земская Е. А. Русская разговорная речь. — М.: Наука, 1981. — 275 с. 3. Русская грамматика. — М.: Наука, 1980. — 709 с. 4. Шабанов-Кушнарченко Ю. П. Об алгебре конечных предикатов. — Пробл. бионики, 1979, вып. 22, с. 12—15. 5. Бондаренко М. Ф., Шабанов-Кушнарченко Ю. П. О математическом описании естественного языка. — Пробл. бионики, 1981, вып. 27, с. 9—13. 6. Кубрякова Е. С. Типы языковых значений. — М.: Наука, 1981. — 199 с.

Поступила в редколлегию 23.12.83.