

УДК 510.62

М. Ф. БОНДАРЕНКО, канд. техн. наук, С. И. МАЛЕНЧЕНКО

МЕТОД ФОРМИРОВАНИЯ ГРАММАТИЧЕСКИХ КАТЕГОРИЙ ПО СМЫСЛОВЫМ ОТТЕНКАМ МОРФЕМ

В грамматике русского словоизменения большое развитие получило учение о словоизменительных категориях (род, число, падеж, время и др.) [1, с. 465—483, 583—646]. Эти категории необходимы при математическом моделировании словоизменительных процессов [2]. Совершенно иное положение сложилось в области грамматики русского словообразования. Систематическое учение о словообразовательных категориях отсутствует, о них имеются лишь разрозненные и отрывочные данные [1, с. 133—142]. Чтобы получить возможность беспрепятственного математического моделирования словообразовательных процессов, необходимо прежде сформировать словообразовательные категории. В этой статье описывается метод формирования грамматических категорий (в том числе словообразовательных).

Любая грамматическая категория представляет собой переменную, заданную на некотором множестве смысловых оттенков. Например, категория рода имеет три смысловых оттенка: мужской, женский и средний. Поэтому, чтобы сформировать грамматическую категорию, достаточно указать для нее множество смысловых оттенков, на которое она распространяется. Множество это не может быть выбрано произвольным образом, оно должно удовлетворять определенным требованиям. Четкая формулировка этих требований позволит указать метод формирования грамматических категорий.

Смысловые оттенки относятся к отдельным морфемам, при этом одни и те же морфемы в различных словах (правильнее было бы сказать — текстах) могут характеризоваться различными наборами смысловых оттенков. Так, окончание *ой* словоформы *дорогой* в тексте *дорогой брат* характеризуется набором смысловых оттенков <именительный, мужской>, а в тексте *дорогой сестре* — набором <дательный, женский>. Текст, окружающий морфему и однозначно характеризующий набор относящихся к ней смысловых оттенков, назовем *текстовой ситуацией*. В нашем примере текстовыми ситуациями будут служить тексты *дорог... брат*, *дорог... сестре*. Набор смысловых оттенков, однозначно определяющий соответствующую ему морфему в данной текстовой ситуации, назовем *полным*. Две текстовые ситуации, которые характеризуются одинаковыми полными наборами смысловых оттенков, будем считать тождественными.

Введем множество $A = \{a_1, a_2, \dots, a_k\}$ всех морфем, для которых требуется сформировать грамматические категории. Введем также множество $B = \{b_1, b_2, \dots, b_l\}$ всех интересующих нас текстовых ситуаций с морфемами из множества A . Наконец, введем множество $C = \{c_1, c_2, \dots, c_m\}$ всех смысловых оттенков, применимых к морфемам из множества A , включенным в текстовые ситуации, взятые из множества B . Рассмотрим предикат $P(x, y)$, где $x \in C$, $y \in B$, равный единице, когда оттенок x присущ морфеме, находящейся в текстовой ситуации y , и равный нулю в противном случае.

Образуем предикат $Q(x_1, x_2)$, где $x_1, x_2 \in C$, равный

$$Q(x_1, x_2) = D(x_1, x_2) \vee \neg \exists y (P(x_1, y) \wedge P(x_2, y)). \quad (1)$$

Здесь D — предикат равенства на C^2 , определяемый следующими условиями:

$$D(x_1, x_2) = \begin{cases} 1, & \text{если } x_1 = x_2, \\ 0, & \text{если } x_1 \neq x_2. \end{cases} \quad (2)$$

Предикат $Q(x_1, x_2)$ обращается в единицу, если оттенки x_1 и x_2 совпадают или же не существует ни одной текстовой ситуации, которой были бы одновременно присущи смысловые оттенки x_1 и x_2 . В противном случае $Q(x_1, x_2) = 0$.

Очевидно, что два различных значения одной и той же грамматической категории не могут быть одновременно присущи какой бы то ни было ситуации. Например, в тексте *дорогой сестре* окончанию *ой* присущ дательный падеж, а винительный — не присущ. Поэтому если существует хотя бы одна текстовая ситуация y , которой одновременно присущи два различных оттенка x_1 и x_2 , то эти оттенки принадлежат различным категориям. Мы видим, что если $Q(x_1, x_2) = 1$, то смысловые оттенки x_1, x_2 принадлежат одной грамматической категории, если же $Q(x_1, x_2) = 0$, то оттенки x_1, x_2 должны быть отнесены

к различным категориям. Таким образом, предикат Q в скрытом виде содержит искомые грамматические категории.

Опишем алгоритм формирования категорий по известному предикату $Q(x_1, x_2)$. Алгоритм будем сопровождать примером. Предикат $Q(x_1, x_2)$ задан так. Для каждого значения x_1 находим класс всех x_2 таких, что $Q(x_1, x_2) = 1$. Имеем: для $x_1 = 1$ — класс $\{1, 3\}$, для $x_1 = 2$ — класс $\{2, 4, 5\}$, для $x_1 = 3$ — класс $\{1, 3\}$, для $x_1 = 4$ — класс $\{2, 4\}$, для $x_1 = 5$ — класс $\{2, 5\}$. Предикат Q симметричен, поэтому его левые и правые классы совпадают. Полученные классы графически представлены на рис. 1.

$x_2 \backslash x_1$	1	2	3	4	5
1	1	0	1	0	0
2	0	1	0	1	1
3	1	0	1	0	0
4	0	1	0	1	0
5	0	1	0	0	1

Исключая «лишние» оттенки, добиваемся превращения предиката Q , который есть предикат толерантности [3, с. 78] (как рефлексивный и симметричный), в предикат эквивалентности. Для этого нужно исключить минимальное число оттенков, при котором система множеств превращается в разбиение. Исключение «лишних» оттенков можно осуществить многими различными способами, при этом получаются различные разбиения множества оттенков. В нашем примере получаются два варианта разбиения множества оттенков. Одно из них (рис. 2, а) получается в результате исключения оттенка 5, второе (рис. 2, б) — в результате исключения оттенка 4.

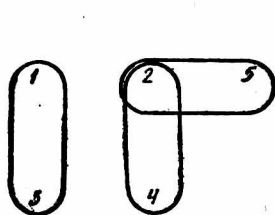


Рис. 1

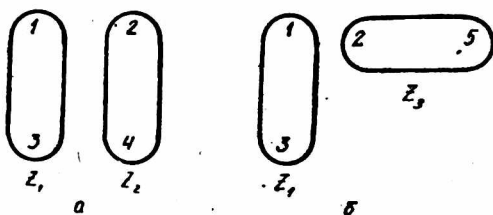


Рис. 2

В соответствии с полученным результатом получаем два набора категорий $\langle z_1, z_2 \rangle$ и $\langle z_1, z_3 \rangle$. Категория z_1 имеет значения 1, 3, категория z_2 — значения 2, 4, категория z_3 — значения 2, 5. В нашем примере наборы категорий получились совершенно равноценными, поэтому можно с равным успехом принять любой из них. Однако в общем случае наборы категорий могут получиться неравноценными, например: они могут отличаться общим числом различных оттенков, числом категорий, числом оттенков в категориях, более или менее равномерным распределением оттенков в смысле их числа по категори-

ям. В этом случае может возникнуть вопрос о выборе оптимального набора категорий. Появление «лишних» оттенков связано с тем, что первоначальный выбор оттенков случаен и вполне может обладать избыточностью. Можно также поставить вопрос о «недостающих» оттенках, введение которых позволило бы отыскать еще более оптимальный вариант набора категорий или же расширить множество тестовых ситуаций, охватываемых грамматическими категориями. Нельзя формировать категории, содержащие по одному оттенку, поскольку переменная, представляющая такую категорию, превращается в константу и по самому смыслу понятия категории не может являться ею.

В более сложных случаях, чем тот, который был рассмотрен в примере, необходим специальный метод последовательного исключения оттенков. Один из возможных методов таков. Исключаем любой оттенок, который охватывается более чем одним множеством, однако нежелательно производить такие исключения, которые приводят к одноэлементным множествам. Так, в нашем примере следует воздержаться от исключения оттенка 2, поскольку при этом появляются множества {4} и {5}. Если же появление одноэлементных множеств неизбежно, то на следующем шаге исключаем все элементы, входящие в одноэлементные множества. Процесс повторяется до тех пор, пока не останется ни одного элемента, охватываемого более чем одним множеством. В результате получаем разбиение множества всех оставшихся оттенков на классы эквивалентности. Исключая оттенки во всевозможных различных порядках, получаем все разбиения множества оттенков. Им соответствуют всевозможные наборы грамматических категорий.

Список литературы: 1. Русская грамматика, т. 1. — М.: Наука, 1980. — 816 с. 2. Шабанов-Кушнаренко Ю. П. Применение метода нуля-органа в лингвистике. — Проблемы бионики, 1978, вып. 21, с. 3—15. 3. Шрейдер Ю. А. Равенство, сходства, порядок. — М.: Наука, 1971. — 76 с.

Поступила в редколлегию 17.03.81.