

О МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ СЛОВООБРАЗОВАТЕЛЬНОЙ СЕМАНТИКИ ЦЕПОЧКИ СУФФИКСАЛЬНЫХ МОРФОВ

Задача математического описания смысла производного слова может быть решена после формализации деривационного анализа. Сложность решения этой задачи состоит в том, что смысл производной лексической единицы представляет собой в общем случае не простую сумму смыслов элементов деривата, а некоторую функцию, аргументами которой являются смыслы производящей основы и словообразовательного аффикса. В настоящей статье сделана попытка определения смысла многоаффиксных производных, и как первый шаг в этом направлении — определение словообразовательного значения, мотивируемого цепочкой суффиксов. Основу категориальной семантики производного слова составляет словообразовательное значение, мотивируемое аффиксом. Словообразовательная семантика, актуализируемая деривационной морфемой, обычно общая для некоторого класса производных, представляет собой некую семантическую константу. Зная конкретное словообразовательное значение аффикса и семантику мотивирующей основы, можно получить смысл производной лексической единицы. Для этого нужно конкретизировать словообразовательную семантику ссылкой на семантику производящей основы. В случае многоаффиксной производной единицы описанные действия производят несколько раз (для каждого морфа), причем определение словообразовательного значения цепочки аффиксов начинают с первого от основы морфа.

Основной единицей классификации словообразовательных значений является словообразовательный тип (СТ) [1]. Это некая схема построения слов, абстрагированная от конкретных лексических единиц, характеризующаяся общностью трех элементов: частью речи производящей основы; семантическим соотношением между производными и производящими; формальным соотношением между производными и производящими, а именно, тождественностью аффикса при аффиксальном словообразовании (здесь и в дальнейшем нас будет интересовать лишь лексическая деривация, несущая в себе полную трансформацию лексического значения производящего в отличие от синтаксической деривации). Общее значение, отличающее все мотивированные единицы данного типа от мотивирующих, называют словообразовательным значением (СЗ) производных слов этого типа. Содержательно интерпретировать СЗ удобно как сумму некоего инвариантного смысла суффикса и семантических добавлений [2].

Важнейшими характеристиками типа принято считать регулярность и продуктивность. Однако в ряде случаев можно говорить и о продуктивности суффикса, понимая под этим его способность

участвовать в различных словообразовательных моделях, характеризующихся сменой мотивирующих частей речи: существительное → существительное, глагол → существительное, прилагательное → существительное. Можно привести ряд морфем, способных образовывать дериваты от двух частей речи, встречаются также словообразовательные морфемы, способные участвовать в трех словообразовательных моделях: *утиль-щик, выдум-щик, паркетчик, потат-чик; греб-ец, краснодерев-ец, юн-ец*. Формальная модель морфемы <щик> (ее морфы-чик/-щик) записывается в виде уравнения алгебры конечных префикатов [3]:

$$S_1^- \bar{s}_{21}^m \bar{s}_{22}^m \bar{s}_{23}^m \bar{s}_{31}^n \bar{s}_{32}^k \bar{s}_{33}^- \bar{s}_4^- \vee S_1^- \bar{s}_{21}^- \bar{s}_{22}^n \bar{s}_{23}^- \bar{s}_{31}^n \bar{s}_{32}^k \bar{s}_{33}^- \bar{s}_4^- \vee \\ \vee s_{11}^n s_{12}^n s_{13}^- \bar{s}_{21}^- \bar{s}_{22}^m \bar{s}_{23}^- \bar{s}_{31}^n \bar{s}_{32}^k \bar{s}_{33}^- \bar{s}_4^- \sim M^l, \quad (1)$$

где s_{ij} — сегмент лингвистического регистра сегментированных суффиксов, i — номер сегмента регистра, j — номер позиции в сегменте, M^l — словообразовательная морфема с порядковым номером l . Таким образом можно говорить о морфеме, способной соединяться с основами, имеющими различную частеречную принадлежность.

При исследовании различных словообразовательных типов было замечено, что одним из решающих факторов, влияющих на изменение смысла семантического инварианта аффикса, является частеречная принадлежность мотивирующей основы, например: основа прилагательное — *дик-арь* — «лицо, мужской пол, характеризуется отношением к признаку»; основа глагол — *звон-арь* — «лицо, мужской пол, характеризуется отношением к действию»; основа существительное — *пушк-арь* — «лицо, мужской пол, характеризуется отношением к предмету». Внутри словообразовательной модели инвариант имеет некоторое постоянство. Таким образом, видно, что охарактеризовать суффикс полностью можно только в том случае, если охарактеризована его словообразовательная модель либо модели. Последнее возможно, «... если в описание суффикса включены все данные о сочетающихся с ним основах» [1]. Только тогда прояснится полная картина вариации словообразовательного значения, включающая как информацию о дополнительных элементах смысла, так и о семантическом инварианте аффикса.

В русском языке имеется ряд суффиксов, способных в одном случае выступать неким носителем категориального значения, актуализируя при этом и дополнительную семантику, в другом — выступать в качестве элемента, модифицирующего семантику производящей основы минимально, привнося лишь один оттенок смысла «женскость», «фамильярность» и т. д. Например: *ззнай-ка, побирущ-ка* — «лицо, общий род, субъект действия, просто производит действие»; *смугл-янка, резв-ушка* — «лицо, женский пол, характеризуется отношением к признаку, внешний признак»; *шпионка, цыган-ка* — «женскость». Сказанное выше свидетельствует о необходимости серьезного учета информации об основе на каж-

дом этапе анализа деривата. Допустим, что удалось организовать классы основ по критерию принадлежности к одному словообразовательному типу. При этом учитываются словообразовательные свойства мотивирующих единиц и их способность сочетаться с производной лексической единицы необходимо произвести следующие шаги. 1. Расчленив дериват на словообразующий аффикс и мотивирующую основу. 2. Определить морфемную принадлежность аффикса. 3. Зная морфему, найти группу семантических классов основ, с которыми эта морфема может соединяться. 4. Пользуясь текстом основы, найти ее семантический класс. 5. Определить словообразовательное значение аффикса и смысл производного слова, пользуясь семантикой основы и текстом суффиксальной морфемы.

Исследуем производящие основы одного семантического класса с точки зрения их производности. Оказывается, что в одном случае в качестве мотивирующих могут выступать немотивированные лексические единицы, в другом случае — суффиксальные: *цыган — цыганка, грузин — грузинка*. Таким образом, в одном случае мотивирующую основу можно найти в машинном словаре корневых морфем (и модифицировать ее, например, семой «женскость», как в нашем примере), в другом случае основа производная и, следовательно, возникает проблема определения ее смысла. Напомним, что производные основы «физически» не хранятся в памяти ЭВМ. Они находятся там лишь формально, т. е. могут быть получены в процессе словообразования посредством некоторой формальной операции. Поэтому их объединение в один семантический класс с корневыми морфемами носит тоже чисто формальный характер.

Итак, при анализе многослойного (многосуффиксного в данном случае) деривата возникает проблема определения семантики производной основы и проблема определения ее словообразующих свойств: задача определения ее принадлежности к тому или иному семантическому словообразовательному классу. Как показано выше, производные и производные лексические единицы могут обладать одинаковыми словообразовательными свойствами. Объединим их в один семантический словообразовательный класс, причем производные лексические единицы будут реальными членами класса, а производные — формальными. Словообразовательные свойства производящей основы, т. е. принадлежность словообразовательному классу, будем определять одновременно с ее семантикой. Исходным материалом в этой задаче считаем очередной словообразующий аффикс и мотивирующую основу, получаемые на очередном шаге препарации. На языке алгебры конечных предикатов этот процесс описывается следующим образом [3]:

$$M^l k_l^n \sim k_l^n, \quad (2)$$

$$M^l k_l^n \sim x_{11}^{\sigma_1} x_{22}^{\sigma_2} \dots x_{nn}^{\sigma_n} : x_{11}^{\sigma_1} \cdot x_{22}^{\sigma_2} \dots x_{nn}^{\sigma_n} = r^n, \quad (3)$$

где M^l — морфема, имеющая условный номер l ; k_l^n — семантический класс n внутри группы основ, соединяющихся с морфемой l ;

k_p^q — семантический класс q внутри группы основ, соединяющихся с морфемой p ; x_i^i — семантические оттенки, σ_i — значения оттенков смысла; r^n — n -я семантическая роль суффиксальной морфемы.

Уравнение (2) содержательно интерпретируется таким образом: если основа, находящаяся в семантическом классе n , принадлежащем морфеме l , соединится в процессе словообразования с одним из морфов морфемы M^l , то образуется дериват, который потенциально может соединиться с морфемой, имеющей номер p , попадая при этом в словообразовательный класс q морфемы M^p . Уравнения (3) описывают семантическую роль r^n в виде набора оттенков смысла [2], которая актуализируется, если основа семантического класса n соединится с одним из морфов морфемы с номером l . Заметим, что в семантический словообразовательный класс q (k_p^q) морфемы p могут потенциально попасть дериваты, образованные различными путями, т. е. посредством соединения основ с морфами разных морфем и в результате соединения с морфами одной морфемы из разных ее словообразовательных классов.

Анализ деривата необходимо начинать с препарации последнего словообразовательного морфа. Процесс препарации продолжается до тех пор, пока последняя мотивирующая основа будет ничем иным, как корнем. Корни машинного словаря снабжены наборами помет, которые объединяют их по семантическим словообразовательным признакам в классы. Поскольку любая производная единица в результате препарации может быть расчленена на корень и аффиксы, то, зная текст первого словообразовательного аффикса и семантический класс корневого морфа, можно определить словообразовательное значение и словообразовательные свойства первой мотивирующей основы, полученной на этом деривационном шаге. Процесс моделирования словообразования будет повторяться до тех пор, пока не присоединится последний словообразовательный аффикс.

Таким образом, определение словообразовательной семантики сводится к последовательной препарации и затем моделированию процесса словообразования с определением семантических словообразовательных классов, к которым принадлежат основы, участвующие в каждом очередном шаге суффиксации. Осуществив морфную препарацию деривата, последовательно восстанавливаем семантическую предысторию всех деривационных шагов (путем последовательного наращивания семантики). Последнее обстоятельство позволяет пошагово определять словообразовательный смысл очередного морфа и словообразовательные свойства основы. Полная словообразовательная семантика аффиксальной цепочки многослойного деривата определяется, как видно, на последнем шаге.

В семантическом словообразовательном классе могут содержаться производные и непроизводные лексические единицы. С учетом этого обстоятельства сформулируем алгоритм определения

семантики многослойного суффиксального деривата. 1. Отсекаем первый словообразующий аффикс. 2. Определяем морфему, к которой принадлежит этот аффикс. 3. Устанавливаем группу классов основ по частеречной принадлежности, и класс, с основами которого может соединяться эта морфема (номер класса l совпадает с порядковым номером морфемы M^l). 4. Пользуясь текстом основы, находим номер n семантического класса k_n^l , к которому она принадлежит; если класс не найден, выполняем п. 5, в противном случае — п. 6. 5. Отсекаем очередной словообразующий аффикс и переходим к выполнению п. 2; если отсечение аффикса невозможно, считаем, что исследуемое слово «потенциально», переходим на специальную ветку алгоритма (определению смысла потенциальных слов будет посвящена отдельная статья). 6. Определяем словообразовательное значение, мотивируемое первым словообразующим суффиксом, и семантический класс, в который потенциально может перейти основа, если к ней будет присоединен следующий словообразовательный морф. 7. Соединяем очередной аффикс с основой, повторяем выполнение п. 6; если аффикс последний, процесс моделирования семантики закончен. Таким образом, в памяти ЭВМ могут храниться производные лексические единицы и их семантика. Вся языковая информация записывается в виде уравнений алгебры конечных предикатов, с которыми и оперирует вычислительная машина. Приведем примеры реальных машинных уравнений, описывающих переход производных из одного семантического класса в другой. Рассмотрим морфему -к(а) (M^{23}). Объединение морфов в морфему M^{23} записывается с помощью следующего уравнения алгебры конечных предикатов:

$$\begin{aligned} S_1^- s_{31}^- s_{32}^k s_{33}^- S_4^- (s_{21}^0 s_{22}^b s_{23}^- \vee s_{21}^a s_{22}^h s_{23}^- \vee s_{21}^n s_{22}^c s_{23}^- \vee \\ \vee s_{21}^a s_{22}^c s_{23}^- \vee s_{21}^v s_{22}^m s_{23}^- \vee s_{21}^a s_{22}^m s_{23}^- \vee s_{21}^n s_{22}^m s_{23}^- \vee \\ \vee s_{21}^n s_{22}^h s_{23}^- \vee s_{21}^n s_{22}^c s_{23}^-) = M^{23}. \end{aligned} \quad (4)$$

Условия потенциального перехода производных, образованных с помощью словообразовательных морфем $M^1, M^7, M^{15}, M^{16}, M^{20}$ в класс R_{23}^{80} морфемы M^{23} , запишутся в виде:

$$\begin{aligned} M^1 (k_1^3 \vee k_1^{23} \vee k_1^{24} \vee k_1^{25} \vee k_1^{26} \vee k_1^{27} \vee k_1^{28} \vee k_1^{51} \vee k_1^{53} \vee \\ \vee k_1^{55} \vee k_1^{56} \vee k_1^{57} \vee k_1^{58} \vee k_1^{59} \vee k_1^{60} \vee k_1^{61} \vee k_1^{64} \vee k_1^{68} \vee k_1^{70} \vee k_1^{71}) \sim k_{23}^{80}, \end{aligned} \quad (5)$$

$$M^7 (k_7^1 \vee k_7^{33} \vee k_7^{35} \vee k_7^{59} \vee k_7^{64} \vee k_7^{66}) \sim k_{23}^{80}, \quad (6)$$

$$M^{15} (k_{15}^{22} \vee k_{15}^{23} \vee k_{15}^{63} \vee k_{15}^{66}) \sim k_{23}^{80}, \quad (7)$$

$$M^{16} (k_{16}^1 \vee k_{16}^8 \vee k_{16}^{34} \vee k_{16}^{59} \vee k_{16}^{60} \vee k_{16}^{64} \vee k_{16}^{71} \vee k_{16}^{72} \vee k_{16}^{73}) \sim k_{23}^{80}, \quad (8)$$

$$M^{20} (k_{20}^8 \vee k_{20}^{60} \vee k_{20}^{64} \vee k_{20}^{69} \vee k_{20}^{70}) \sim k_{23}^{80}. \quad (9)$$

Если любой дериват из класса k_{23}^{80} соединится с морфемой $M^{23-к}(а)$, то всякая семантическая роль, актуализируемая морфемами M^1 , M^7 , M^{15} , M^{16} , M^{20} , в этом классе будет модифицирована семей «женскость»

$$M^{23} k_{23}^{80} \sim x_5^ж, \quad (10)$$

где $x_5^ж$ — оттенок «женскость». Рассмотрим пример. Пусть в деривате *мотоцикл-ист-ка* проведены морфемные швы. Корень *мотоцикл-* относится к семантическому классу k_1^3 морфемы M' (*-ист*). В результате соединения этого корня с морфемой M^1 полученный дериват перейдет в словообразовательный класс k_{23}^{80} (см. уравнение (5)). Одновременно актуализируется семантическая роль r^3 : $M^1 k_1^3 \sim r^3$ и $M^1 k_1^3 \sim k_{23}^{80}$. При соединении основы *мотоциклист-* с морфом *-к(а)* семантическая роль r^3 модифицируется семей «женскость»: $M^{23} k_{23}^{80} \sim x_5^ж$.

Список литературы: 1. Принципы и методы семантических исследований. М., 1976. 266 с. 2. Левицкий А. С., Шаронова Н. В., Рябова Н. В. О математическом моделировании деривационного анализа суффиксальных существительных со значением лица//К., 1988. 15 с. Деп. в УкрНИИИТИ. 15.03.88. № 1313-Укр88. 3. Шабанов-Кущинаренко Ю. П. Теория интеллекта. Математические средства. К., 1984. 144 с.

Поступила в редколлегию 11.04.88.

УДК 652.22

С. Н. ГЕРАСИН, Е. Б. КАЙКОВА, В. Я. ТЕРЗИЯН, канд. техн. наук

ИНТЕРВАЛЬНЫЙ ПОДХОД К ОБРАБОТКЕ ЭКСПЕРТНЫХ МНЕНИЙ. МЕТОД ПОСЛЕДОВАТЕЛЬНЫХ ОТСЕЧЕНИЙ

В условиях неопределенности для принятия решений широко используется метод экспериментального оценивания. Обычно решались две задачи: сбор экспертной информации (в этом плане разрабатывается метод Дерби [1]) и обработка оценок экспертов с целью получения из субъективной информации какого-либо объективного результата. Основой для решения второй задачи служит метод статистического точечного оценивания [2]. По-видимому, такой подход к обработке экспертных мнений не единственен. Действительно, предположим, что эксперт оценивает некоторый числовой параметр системы, причем вместо точечной оценки он определяет интервал, в который, по его мнению, попадает данный параметр. Сам эксперт, т. е. любой носитель субъективного мнения, обладает некоторой суммой знаний о параметре, поэтому очевидно, что чем выше квалификация эксперта, тем уже интервал мнений.

Рассмотрим следующую задачу. Требуется определить значение некоторого параметра X . Предположим, что изменение точечного