

В. И. БУЛКИН, Д. Э. СИТНИКОВ, Ю. П. ШАБАНОВ-КУШНАРЕНКО,
д-р техн. наук, Н. В. ШАРОНОВА, канд техн. наук

МАТЕМАТИЧЕСКИЕ МОДЕЛИ МЕЖМОРФЕМНЫХ СВЯЗЕЙ НА МНОЖЕСТВЕ ПОЛИСЕМАНТИЧЕСКИХ ПРОИЗВОДЯЩИХ ОСНОВ И СЛОВООБРАЗОВАТЕЛЬНЫХ СУФФИКСОВ

В процессе развития вычислительной техники и все более широкого внедрения ее в различные отрасли народного хозяйства особое значение приобретает удобство общения человека с ЭВМ. В настоящее время общение пользователя с ЭВМ, как известно, происходит с помощью программ, для написания которых используются специальные языки программирования. ЭВМ будущих поколений должны иметь такие аппаратные и программные средства, которые позволяют осуществлять общение с машиной на естественном языке. Для решения этой проблемы в автоматизированных информационных системах предусматривается создание человеко-машинного интеллектуального интерфейса, адекватно реагирующего на естественно-языковые запросы человека путем формирования осмысленных ответов на поступающие входные сообщения.

Одной из важных проблем при создании интеллектуального интерфейса является адекватная интерпретация смысла сообщений, решение которой невозможно без понимания смысла слова. Как известно, смысл производного слова (именно такие слова мы будем рассматривать) в общем случае нельзя описать как простую сумму смыслов составляющих его морфов. Для того чтобы лучше понять механизм формирования смысла деривата, необходимо рассмотреть и описать математически межморфемные семантические связи, возникающие в процессе словообразования между префиксальными и корневыми морфемами, корневыми и суффиксальными морфемами, а также между основами и окончаниями.

Рассмотрим математическое моделирование межморфемных семантических связей на морфемном шве между корневым и суффиксальным морфемами производного слова. Пусть имеется множество производных слов, каждое из которых может быть представлено в виде $P \nabla Q \nabla S \nabla R$, где P — префиксальная морфема; Q — корневая морфема; S — суффиксальная морфема; R — окончание. Знак ∇ показывает, что между этими морфемами существуют определенные семантические связи, которые необходимо описать математически. Согласно работе [2], если существует два множества семантических ролей S_1 и S_2 и два множества предикатов-морфем M_1 и M_2 , операция соединения двух морфем $P_1(t) \in M_1$ и $P_2(t_2) \in M_2$ характеризуется согласованием семантических ролей этих двух морфем. В результате соединения получим множество формализованных связей между семантическими ролями рядом стоящих морфем. Такое множество представ-

ляет собой некоторый бинарный предикат $P(t_1, t_2)$, причём $P(t_1, t_2) \rightarrow P_1(t_1) \cdot P_2(t_2)$ и, если предположить, что возможность установления семантических связей не зависит от вида морфем, то на декартовом произведении $S_1 \times S_2$ можно задать предикат $\lambda(t_1, t_2)$, принимающий значение 1, если между семантическими ролями t_1 и t_2 устанавливаются связи, и 0 — в противном случае. Логическое произведение $P_1(t_1) \cdot P_2(t_2)$ означает возможность согласования любой семантической роли морфемы $P_1(t_1)$ с любой семантической ролью морфемы $P_2(t_2)$. Однако в действительности часто некоторые роли рядом стоящих морфем не согласуются, поэтому операция соединения * морфемных семантических ролей запишется так:

$$P_1(t_1) * P_2(t_2) = \lambda(t_1, t_2) \cdot P_1(t_1) \cdot P_2(t_2), \quad (1)$$

где логическое произведение $P_1(t_1) \cdot P_2(t_2)$ описывает все связи между семантическими ролями морфем $P_1(t_1)$ и $P_2(t_2)$, а $\lambda(t_1, t_2)$ исключает те связи, которые при этом не реализуются.

Рассмотрим пример взаимодействия корневых и суффиксальных морфем, в результате которого образуются производные отглагольные имена существительные. В частности, взаимодействие корневых морфем, входящих в состав полисемантических производящих основ, с суффиксальными морфемами, семантические роли которых частично совпадают. Для нашего случая формула (1) имеет вид

$$P_2(t_2) * P_3(t_3) = \lambda(t_2, t_3) \cdot P_2(t_2) \cdot P_3(t_3), \quad (2)$$

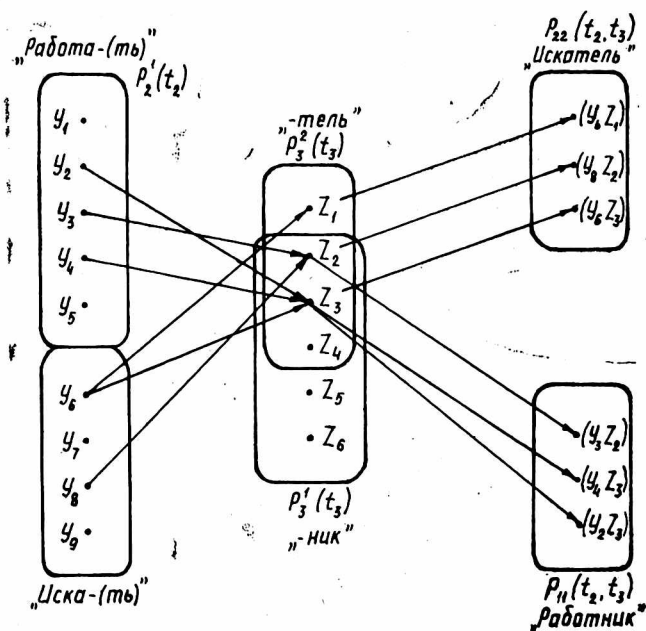
где $P_2(t_2)$ — предикат, описывающий семантические роли корневой морфемы из множества M_2 корневых морфем; $P_3(t_3)$ — предикат, описывающий семантические роли суффиксальной морфемы из множества M_3 суффиксальных морфем; $\lambda(t_2, t_3)$ — предикат, равный 1, если семантические роли t_2 и t_3 согласуются и 0 в противном случае; t_2 и t_3 — переменные семантические роли корневых и суффиксальных морфем соответственно.

Рассмотрим конкретный пример образования производных многозначных слов «работник» и «искатель», полученных путем соединения полисемантических производящих глагольных основ «работа-ть» и «иска-ть» и суффиксальных морфем *-ник* и *-тель*, семантические роли которых частично совпадают, образуя своеобразное пересечение семантических ролей — компонентов «семантических облаков» этих морфем.

Основы «работа-ть» и «иска-ть» имеют семантические роли $Y_1 \div Y_5$ и $Y_6 \div Y_9$ соответственно [3]: Y_1 — находиться в действии (завод работает), Y_2 — заниматься чем-нибудь, применяя свой труд (работать в поле), Y_3 — иметь занятие, должность, служить (работать на заводе), Y_4 — обслуживать кого-нибудь своим трудом (работать на семью), Y_5 — приводить в действие, управлять чем-нибудь (работать веслом), Y_6 — стараться найти или обнаружить (искать книгу), Y_7 — стараться получить (искать защиты, искать работы), Y_8 — стремиться к чему-нибудь новому (ищущий художник), Y_9 — предъявлять иск к кому-нибудь (спец).

Суффиксальная морфема *-тель* характеризуется следующими семантическими ролями: Z_1 — предмет — субъект действия, Z_2 — лицо — субъект действия, профессия, Z_3 — лицо — субъект действия, не профессия, Z_4 — предмет — орудие для выполнения действия.

Суффиксальная морфема *-ник* имеет следующие семантические роли; Z_5 — животное, субъект действия, Z_6 — место, помещение, а также семантические роли Z_2 , Z_3 , Z_4 , которые являются общими для суффиксальных морфем *-тель* и *-ник*.



Сочетание семантических ролей

Сочетание семантических ролей (установление связей между семантическими ролями корневых морфов производящих основ «иска-ть», «работа-ть» и суффиксальных морфов *-тель*, *-ник*) удобно представить, как показано на рисунке.

В процессе словообразовательного акта при взаимодействии производящей основы «работать» и суффикса *-ник* образуется производное слово «работник», имеющее несколько лексических значений. Как видно, каждому значению этого слова соответствует определенная связь между семантическими ролями корневой и суффиксальной морфем. При образовании производного слова «работник» актуализируются три семантические связи, соответствующие трем смысловым значениям этого слова: $Y_3 Z_2$ — человек, работающий в какой-нибудь сфере трудовой деятельности (научный работник), $Y_2 Z_3$ — тот, кто работает, трудится (отличный

работник), Y_4Z_3 — наемный сельскохозяйственный рабочий (хозяин и работник).

В случае же, когда в результате взаимодействия производящей основы «иска-ть» и суффиксальной морфемы *-тель* образуется производное слово «искатель», между семантическими ролями корневого и суффиксального морфов устанавливаются также три семантические связи, которые соответствуют трем смысловым значениям деривата «искатель»: Y_6Z_1 — приспособление в различных приборах, облегчающее нахождение наблюдаемого предмета, какого-нибудь объекта искатель в фотоаппарате, Y_8Z_2 — человек, стремящийся к новому, увлеченный исканиями, Y_6Z_3 — тот, кто занят поисками, добыванием чего-нибудь, искатель жемчуга, искатель приключений.

Для того чтобы описать семантические связи, образующиеся между корневыми и суффиксальными морфами в нашем примере, воспользуемся формулой (2), где $\lambda(t_2, t_3)$ для нашего примера имеет вид

$$\lambda(t_2, t_3) = t_{22}^y t_{33}^z \vee t_{23}^y t_{32}^z \vee t_{24}^y t_{33}^z \vee t_{26}^y t_{31}^z \vee t_{26}^y t_{33}^z \vee t_{28}^y t_{32}^z. \quad (3)$$

Множества семантических ролей производящих основ «работа-ть» и «иска-ть» будем описывать с помощью предикатов $P_2^1(t_2)$ и $P_2^2(t_2)$ соответственно:

$$P_2^1(t_2) = t_{21}^y \vee t_{22}^y \vee t_{23}^y \vee t_{24}^y \vee t_{25}^y; \quad (4)$$

$$P_2^2(t_2) = t_{26}^y \vee t_{27}^y \vee t_{28}^y \vee t_{29}^y.$$

Множества семантических ролей суффиксов можно описать с помощью предикатов $P_3^1(t_3)$ и $P_3^2(t_3)$ соответственно:

$$P_3^1(t_3) = t_{32}^z \vee t_{33}^z \vee t_{34}^z \vee t_{35}^z \vee t_{36}^z; \quad (5)$$

$$P_3^2(t_3) = t_{31}^z \vee t_{32}^z \vee t_{33}^z \vee t_{34}^z.$$

В процессе словообразования при соединении корневого морфа производящей основы «раба-ть» и суффикса *-ник* образуется производное слово «работник». Семантические связи, возникающие при этом, могут быть описаны с помощью предиката

$$P_{11}(t_2, t_3) = \lambda(t_2, t_3) \cdot P_2^1(t_2) \cdot P_3^1(t_3) = (t_{22}^y t_{33}^z \vee t_{23}^y t_{32}^z \vee t_{24}^y t_{33}^z \vee t_{26}^y t_{31}^z \vee t_{26}^y t_{33}^z \vee t_{28}^y t_{32}^z) (t_{21}^y \vee t_{22}^y \vee t_{23}^y \vee t_{24}^y \vee t_{25}^y) (t_{32}^z \vee t_{33}^z \vee t_{34}^z \vee t_{35}^z \vee t_{36}^z) = t_{22}^y t_{33}^z \vee t_{23}^y t_{32}^z \vee t_{24}^y t_{33}^z. \quad (6)$$

Полученный предикат $P_{11}(t_2, t_3)$ является математической моделью межморфемных связей корневого и суффиксального морфов производной словоформы «работник».

Семантические связи, возникающие при сочетании производящей основы «иска-ть» и суффикса *-тель* в процессе словообразовательного акта, опишем с помощью предиката

$$P_{22}(t_2, t_3) = \lambda(t_2, t_3) \cdot P_2^2(t_2) \cdot P_3^2(t_3) = (t_{22}^y t_{33}^z \vee t_{23}^y t_{32}^z \vee t_{24}^y t_{33}^z \vee t_{25}^y t_{31}^z \vee t_{26}^y t_{33}^z \vee t_{28}^y t_{32}^z) \cdot (t_{26}^y \vee t_{27}^y \vee t_{28}^y \vee t_{29}^y) \cdot (t_{31}^z \vee t_{32}^z \vee t_{33}^z \vee t_{34}^z) = t_{26}^y t_{31}^z \vee t_{26}^y t_{33}^z \vee t_{28}^y t_{32}^z. \quad (7)$$

Предикат $P_{22}(t_2, t_3)$ является математической моделью межморфемных семантических связей производного слова «искатель» на морфемном шве между корневым морфем производящей основы и словообразовательным суффиксом.

Если попытаться соединить производящую основу «работа-ть» и суффикс *-тель*, семантические роли которых описываются предикатами $P_2^1(t_2)$ и $P_3^2(t_3)$, то можно убедиться, что при этом «правильная» словоформа не образуется, так как нормы русского языка не допускают образования производного слова «работатель», хотя семантические связи вполне позволяют сформировать на выходе системы такую «неправильную» словоформу. Формирование связей, возникающих в этом случае, можно описать с помощью формулы

$$P_{12}(t_2, t_3) = \lambda(t_2, t_3) \cdot P_2^1(t_2) \cdot P_3^2(t_3) = (t_{22}^y t_{33}^z \vee t_{23}^y t_{32}^z \vee t_{24}^y t_{33}^z \vee t_{26}^y t_{31}^z \vee t_{26}^y t_{33}^z \vee t_{28}^y t_{32}^z) \cdot (t_{21}^y \vee t_{22}^y \vee t_{23}^y \vee t_{24}^y \vee t_{25}^y) \cdot (t_{31}^z \vee t_{32}^z \vee t_{33}^z \vee t_{34}^z) = t_{22}^y t_{33}^z \vee t_{23}^y t_{32}^z \vee t_{24}^y t_{33}^z. \quad (8)$$

Таким образом, мы получили для «неправильной» с формальной точки зрения словоформы «работатель» те же семантические связи и те же смысловые значения слова, что и для словоформы «работник».

Учитывая изложенное, можно сделать вывод, что без привлечения дополнительных признаков, характеризующих формальную сочетаемость производящих основ и словообразовательных суффиксов, невозможно однозначно сформировать словоформу на выходе системы, имея на входе только рассмотренные выше множества семантических ролей морфем. Для устранения возникающей неоднозначности необходимо ввести признак класса основ, который характеризует способность основ данного класса сочетаться с определенной группой суффиксов и не сочетаться со всеми остальными. Введение этого признака и использование его при формализации семантической сочетаемости морфем дает возможность избавиться от неоднозначности.

Полученные в данной работе уравнения описывают межморфемные отношения на семантическом уровне между полисемантическими производящими основами и деривационными суффиксами и являются частью математической модели межморфемных связей.

Список литературы: 1. Шаронова Н. В. Математические модели суффиксального словообразования и их использование для автоматической обработки отглагольных имен существительных в текстах русского языка: канд. дис. Х., 1984. 219 с. 2. Ожегов С. И. Словарь русского языка. М., 1988. 749 с. 3. Цыганенко Г. П. Состав слова и словообразование в русском языке. К., 1978. 151 с.

Поступила в редколлегию 10.05.90