

С.Ю. Семёнова (ИНИОН РАН, РГГУ, Москва)

Модельный пример в автоматическом семантическом словаре: к поиску прототипа употребления лексемы

Тема понимания в сообщении затрагивается в двух аспектах. С одной стороны, прикладной задачей исследования является создание компьютерного словаря, нацеленного на автоматическое понимание текста. С другой стороны, в процессе текстового иллюстрирования словаря возник вопрос о подборе таких модельных примеров, которые как можно ближе соответствовали бы прототипическому употреблению вокабулы, и в этом отношении речь идет о нацеленности на понимание того, каков может быть прототип употребления слова.

Более конкретно. В настоящее время продолжается разработка русского прикладного семантического словаря РУСЛАН, предназначенного для автоматического анализа текста. Словарные описания, отражающие семантические, сочетаемостные, тезаурусные, а также некоторые энциклопедические свойства слова, строятся на формальном метаязыке, предоставляющем лексикографу довольно разнообразные средства кодирования. Создателем метаязыка, как и всей концепции словаря, является Н.Н. Леонтьева. Первые относительно завершенные версии словаря были получены под ее руководством к середине 2000-х гг. Наряду с другими (формализованными) зонами и полями заполнялась свободная зона иллюстраций; в основном она содержала модельные примеры. С 2017 г. под руководством автора начались работы по модернизации словаря, и при этом существенное внимание уделено иллюстрированию. Зона иллюстраций теперь структурирована, она включает ряд новых полей для корпусных иллюстраций (подбираемых, главным образом, в НКРЯ); также в формируемой ныне версии сохранено поле модельных иллюстраций.

Иллюстративная работа в существенной мере нацелена на применение словарных описаний при автоматическом распознавании лексем полисемичных слов. Для модельных примеров (обычно кратких, часто минимальных) это связано со стремлением отразить прототип (или ряд прототипов) употребления лексемы-вокабулы.

Определенным ориентиром при осмыслении прототипа служат модельные примеры в МАС - одном из наиболее авторитетных русских толковых словарей. При этом не все стратегии иллюстрирования, применяемые в МАС, приемлемы для прикладного словаря РУСЛАН – в том числе, в силу более разветвленной структуры последнего. Так, ряд иллюстраций МАС «укладывается» в поля основной части словарной статьи РУСЛАН (таковы, в частности, контекстные лексические функции, а также термины и словосочетания, обозначающие гипонимы). Возникает своего рода конкуренция полей словарной статьи. В сообщении предполагается обсудить эти вопросы, провести определенный анализ иллюстраций в МАС и в РУСЛАНе (в аспекте отражения прототипа), затронув выбор лексического состава примеров.