

УДК 519.768:004.738.5

Г. Г. Белоногов, Р. С. Гиляревский, В. С. Егоров,  
Ал-др А. Хорошилов, А. Н. Шогин

## Системы фразеологического машинного перевода RETRANS и ERTRANS в сети Интернет

*Сообщается об установке систем фразеологического машинного перевода RETRANS и ERTRANS на сетевой сервер ВИНТИ и о возможности поиска информации в русскоязычных базах данных по запросам на английском языке с выдачей результатов поиска также на английском языке.*

Научные исследования и разработки по проблеме машинного перевода велись в ВИНТИ в течение нескольких десятков лет. Сначала они носили преимущественно экспериментальный характер [1,2]. Затем, с середины 80-х гг., был взят курс на создание промышленных систем машинного перевода. В настоящее время созданы две системы фразеологического машинного перевода: 1) система русско-английского перевода (RETRANS) и 2) система англо-русского перевода (ERTRANS). Обе системы построены на одних и тех же принципах и имеют аналогичные характеристики [3,4]. Поэтому мы приведем краткие сведения только об одной системе — системе RETRANS.

Система RETRANS предназначена для автоматизированного перевода научно-технических текстов с русского языка на английский. Отличительной чертой этой системы является то, что в ней в качестве основных единиц смысла используются не отдельные слова, а словосочетания и фразы, выражающие научно-технические понятия. Это позволяет переводить тексты более точно. Русско-английский политематический машинный словарь системы содержит терминологию по естественным и техническим наукам, экономике, бизнесу, политике, законодательству и военному делу. В частности, он содержит термины и фразеологические единицы по следующим тематическим областям: Машиностроение, Электротехника, Энергетика, Транспорт, Аэронавтика, Космонавтика, Робототехника, Автоматика и Радиоэлектроника, Вычислительная Техника, Связь, Строительство и Архитектура, Математика, Физика, Химия, Биология, Медицина, Экология, Сельское Хозяйство, Астрономия, Геология, География, Геофизика, Горное дело, Металлургия, Нефтегазовая промышленность, Легкая промышленность, Внешняя Торговля, Банковское Дело. Перевод текстов может осуществляться в автоматическом и в диалоговом режимах.

Основные технические характеристики системы RETRANS:

• **ОБЪЕМ ПОЛИТЕМАТИЧЕСКОГО МА-**

**ШИННОГО СЛОВАРЯ:** более 1 млн 500 тыс. словарных статей; 80% из них составляют словосочетания длиной от двух до семнадцати слов. Объем дополнительных машинных словарей (для настройки системы на различные тематические области) — более 200 тыс. словарных статей.

• **ТИП ПЭВМ:** IBM PC/AT, PENTIUM.

• **ОПЕРАЦИОННЫЕ СИСТЕМЫ:** MS-DOS 6.0 и выше, WINDOWS 95, WINDOWS 98, WINDOWS NT, UNIX.

• **СКОРОСТЬ ПЕРЕВОДА:** 100–150 слов/сек. (зависит от мощности ПЭВМ).

При создании систем фразеологического машинного перевода их разработчики руководствовались следующими принципами:

1. Основными единицами языка и речи, которые прежде всего следует включать в машинный словарь, должны быть фразеологические единицы (словосочетания, фразы). Отдельные слова также могут включаться в словарь, но они должны использоваться только в тех случаях, когда не удается осуществить перевод, опираясь только на фразеологические единицы.

2. Наряду с основным политематическим словарем большого объема, в системах фразеологического машинного перевода целесообразно использовать также набор небольших по объему дополнительных тематических словарей. Дополнительные словари должны содержать только ту информацию, которая отсутствует в основном словаре (например, информацию о приоритетных переводных эквивалентах словосочетаний и слов для различных предметных областей, если эти эквиваленты не совпадают с приоритетными переводными эквивалентами основного словаря).

3. Основным средством разрешения полисемии (многозначности) слов в системах фразеологического перевода является их использование в составе фразеологических словосочетаний. Дополнительным — аппарат дополнительных тематических словарей, где для каждого многозначного слова или словосочетания указывается его приоритетный переводной эквивалент, специфичный для рассматриваемой предметной области.