

АВТОМАТИЧЕСКОЕ РЕФЕРИРОВАНИЕ ТЕКСТОВОЙ ИНФОРМАЦИИ: ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ ПРИ ПОДГОТОВКЕ КАДРОВ

Л.В. СТЕПУРА

Одним из главных условий высокого качества подготовки кадров является возможность обрабатывать большие объемы исходного материала, выделяя из них наиболее значимую информацию с последующим более тщательным ее изучением. Данная задача сводится к анализу и реферированию текстовой информации. Основной функцией системы реферирования является выявление в реферируемом тексте кортежа информативных предложений и синтез реферата. В состав разрабатываемого модуля реферирования входят три основных структурных компонента. Во-первых, база знаний, включающая систему словарей и совокупность тематических корпусов текстов. Во-вторых, программы поиска информативных предложений, которые выявляют в тексте информативные предложения и "высвечивают" их. Количество таких предложений регулируется путем задания пользователем порогового уровня информативности. При этом возможно предъявление пользователю (по его требованию) контекста каждого информативного предложения.

В-третьих, программы синтеза рефератов. Реферат строится из информативных предложений путем поиска релевантной информации в специальной системе словарей и последующего синтеза выходного текста. При решении задачи реферирования используется подход, основанный на математическом моделировании известных в лингвистике семантических отношений, которые обеспечивают обработку документов на различных входных языках при наличии в базе знаний соответствующих двуязычных словарей синтагм. В докладе представлен алгоритм реферирования.