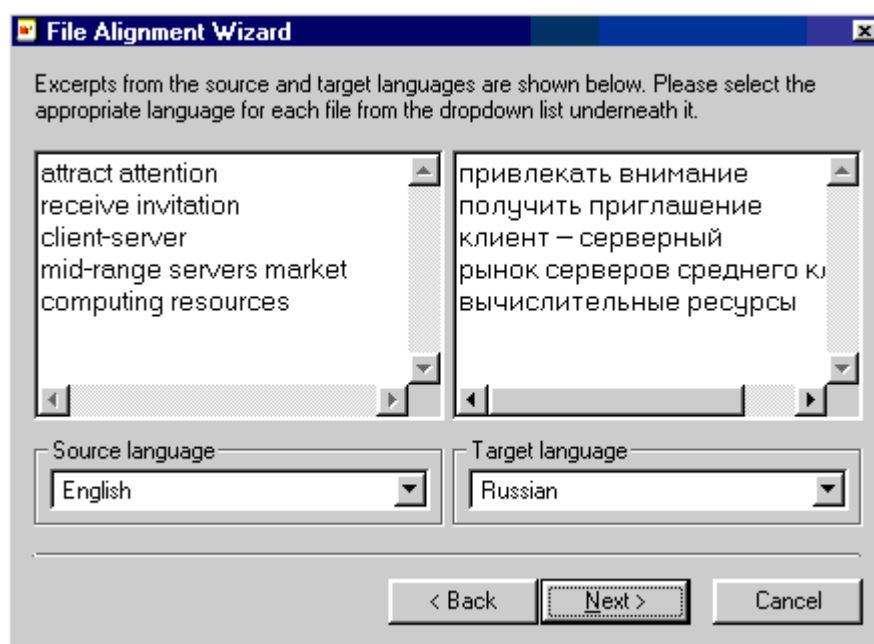


Программы, помогающие переводчику

Александр Силонов

НЕДАВНО (см. PC Week/RE, № 13/2000, с. 45) была опубликована статья Н. Шаховой, посвященная ТМ-программам, или “накопителям переводчика” (возможно, более удачным является вариант “накопители переводов”), — теме, очень редко затрагиваемой на страницах компьютерных изданий. Продолжая эту тему, предлагаем вам обзор ТМ-программ, доступных в настоящее время на рынке. Напомним, что речь идет об интерактивных инструментах, позволяющих переводчику накапливать в специальной базе данных (БД) эквивалентные текстовые фрагменты на двух языках, чтобы в дальнейшем быстро находить образцы для перевода новых текстов. В качестве фрагментов могут выступать слова, словосочетания и целые фразы. При работе над текстами, близкими по жанру и тематике, такие инструменты по мере пополнения БД все больше упрощают и ускоряют перевод.



Мастер сопоставления параллельных текстов в DeJaVu

Итак, основные функции ТМ-программы — пополнение БД парных фрагментов и анализ новых текстов для их поблочного перевода. Программа сегментирует переводимый текст (выделяет фразы или обособленные обороты) и сличает полученные сочетания с элементами БД, в случае совпадения предлагая переводчику подставить готовый перевод фрагмента в конечный текст (или же подставляя его автоматически). Во многих программах предусмотрен “нечеткий” (fuzzy) поиск.

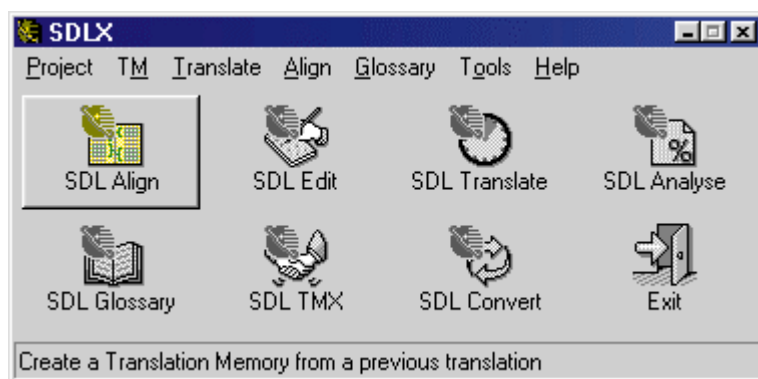
Стандартной для этого класса продуктов стала возможность вести словари терминов и непереводимых имен, что очень важно для работы с текстами из узкой предметной области. Обычно ассортимент поддерживаемых ТМ-системами языковых пар довольно обширен; в последнее время все большее число разработчиков декларируют поддержку “двунаправленных языков” (bi-directional languages). Термин этот

представляется не совсем корректным. Правильнее было бы в данном случае говорить не о языках, а скорее о текстах, содержащих фрагменты на языках с разным направлением письма (например, о смеси английского с арабским или ивритом).

Сама по себе идеология накопителей переводов не подразумевает использования развитых лингвистических технологий, однако в наиболее совершенных ТМ-программах имеются встроенные морфологические модули и средства проверки орфографии для нескольких языков, а также относительно интеллектуальные средства сопоставления (aligning) параллельных текстов с целью автоматического формирования парных фрагментов на двух языках. Процедуры сопоставления в разных программах различны, но, как правило, они включают элементы диалога и иногда требуют модификации одного из параллельных текстов (особенно если имеет место вольный перевод).

Большинство программ совместимы с популярными текстовыми процессорами и издательскими пакетами, с распространенными форматами типа RTF, PDF и HTML. Почти все ТМ-системы позволяют для удобства группировать рабочие файлы, БД и терминологические словари в отдельные “проекты”.

Накопители переводов — дорогое профессиональное средство. Как правило, разработчики предлагают как корпоративные, сетевые, так и индивидуальные (freelance) версии своих ТМ-систем. В первом случае (особенно при поставке многофункциональных инструментальных комплексов) цена может превышать 1000 долл., во втором она составляет несколько сот долларов. Программы нередко защищаются с помощью аппаратных ключей.



Панель SDLX обеспечивает удобный доступ ко всем функциям

Trados 3.0

Одним из пионеров и лидеров рынка ТМ-программ является многонациональная компания Trados (www.trados.com), предлагающая пакет Trados Translation Solution — линейку инструментов автоматизации перевода. Пакет поставляется в корпоративном (Team Edition) и однопользовательском (Freelance Edition, цена около 1 тыс. долл.) вариантах. Основной модуль пакета, Translator’s Workbench, предназначен для работы с базой данных ТМ, а также обеспечивает функции преобразования дат и единиц измерения, а при работе со сложными документами — еще и манипуляции со сносками, указателями и т. п. Модуль MultiTerm служит для ведения многоязычной БД терминов. Модуль WinAlign позволяет в полуавтоматическом режиме пополнять БД накопителя путем анализа параллельных текстов на двух языках. Кроме того, в пакет входит средство для перевода документов, содержащих разметку SGML и HTML, — TagEditor.

Версия Trados 3.0 совместима с Windows 2000, Office 2000 (Word 2000, PowerPoint 2000), издательскими пакетами FrameMaker, Interleaf, PageMaker, QuarkXPress и Ventura, языками XML, SGML и HTML. При инсталляции Trados встраивает свой инструментарий в панель инструментов и меню Word. Тесная интеграция с Office позволяет при переводе текстов в Word автоматически пополнять базу данных ТМ.

Программа автоматически анализирует тексты, разбивая их на сегменты, которые можно затем экспортировать в базу данных ТМ. Ограничителями сегментов по умолчанию служат табуляторы, маркеры абзацев и некоторые знаки препинания. Сегменты, не имеющие аналогов в БД, можно либо включить туда, снабдив переводом, либо поместить в словарь терминов. После сегментации текста функция автоматизированного перевода предлагает пользователю фрагменты для поблочного формирования перевода.

Инструментарий DeJaVu

Компания Atril (www.atril.com) выпускает набор мощных средств класса ТМ, главным из которых является программа DeJaVu Interactive. Ее функция Pretranslate сегментирует исходный текст и автоматически подставляет в конечный переводы сегментов, найденных в БД. Для фрагментов, лишь частично соответствующих имеющимся образцам, функция fuzzy match в диалоговом режиме предлагает вариант перевода. Более интеллектуальные функции Assemble и Propagation позволяют, как утверждается, повысить степень автоматизации перевода и пополнения базы данных ТМ.

Как и в других ТМ-системах, в DeJaVu есть средство работы с терминами, кроме того, добавлена возможность запрещать перевод отдельных фраз в исходных текстах.

Программа совместима с многочисленными текстовыми редакторами и верстальными программами, в том числе Word, PowerPoint, FrameMaker, PageMaker, QuarkXPress, Interleaf, а также ТМ-программами IBM и Trados; поддерживаются форматы RTF, Help Contents, Java Properties files, HTML (включая ASP), HTML Help, SGML, RC, C/C++.



Панель управления WordAlign

DeJaVu позволяет в переводческих проектах выбирать языковые пары из обширного списка языков мира; имеется встроенное средство проверки орфографии для десятка европейских языков. В комплект DeJaVu входит также модуль TermWatch, который обеспечивает доступ к терминологическому словарю из среды приложений, работающих с иноязычными текстами. Программа снабжена удобным справочником.

Все модули DeJaVu используют общий OLE-сервер, на его базе можно разрабатывать приложения, содержащие вызовы функций API.

Цена DeJaVu — от 790 до 990 долл. в зависимости от числа приобретаемых лицензий. На Web-узле разработчика можно загрузить бесплатную 30-дневную демоверсию программы.

SDLX 3.0

Компания SDL International (www.sdlintl.com) предлагает ТМ-инструмент SDLX, позволяющий создавать и редактировать базу данных ТМ, пополнять ее путем сопоставления параллельных текстов, производить анализ, сегментацию и автоматизированный перевод текстов, вести пользовательский многоязычный глоссарий. Предусмотрен “нечеткий” поиск фрагментов в ТМ. Поддерживается перевод “двунаправленных” (неоднородных по направлению письма) текстов.

SDLX совместима с форматами документов HTML, RTF, MIF и TXT; предусмотрен импорт и экспорт баз данных ТМ различных форматов, включая Trados. Важным достоинством программы является возможность обрабатывать (анализировать и переводить) файлы в пакетном режиме.

На узле компании SDL можно загрузить демоверсию SDLX с ограниченной функциональностью. Полная корпоративная версия программы стоит 1000 долл. (она

защищается аппаратным ключом), версия Light — 399 долл., версия для одиночных пользователей (без функций анализа текста, автоматического перевода и фильтров для SGML и XML) — 350 долл.

Transit и TermStar

Швейцарская компания STAR (www.star-transit.com) предлагает серию программ для автоматизации деятельности переводчика.

ТМ-программа Transit Professional использует технологию “нечеткого индекса” (fuzzy index): при поиске фрагментов в базе данных накопленных переводов ищутся сочетания, приблизительно соответствующие выделенному фрагменту.

Transit работает практически со всеми европейскими языками (включая русский), турецким, арабским, китайским и японским, поддерживаются также “двунаправленные” тексты.



Благодаря наличию фильтров импорта-экспорта программа совместима с множеством текстовых процессоров, издательских и прочих пакетов (Excel, PowerPoint, Word, WordPerfect, AmiPro, Interleaf, FrameMaker, PageMaker, Xyvision, Ventura, QuarkXPress), а также с распространенными форматами (HTML, SGML, XML, RTF).

Вместе с Transit распространяется программа TermStar — многоязычный менеджер терминологических гипертекстовых словарей, взаимодействующий с ТМ-средой и пригодный для многопользовательской работы.

ТМ-инструменты фирмы STAR поставляются в нескольких вариантах комплектации. Вариант Transit Professional 3.0, включающий TermStar, стоит 1245 евро; вариант Workstation, не имеющий функций импорта и экспорта, — 695 евро. Корпоративный пакет Translation Center Suite стоит около 4 тыс. евро, а его вариант с сетевыми возможностями Corporate Suite 3.0 — около 8 тыс. евро.

WordFisher 4

Независимый переводчик из Венгрии Корней Тибор создал недорогую (всего 30 долл.) ТМ-программу WordFisher. Программа, реализованная в виде макроса, интегрируется с Microsoft Word (версии 6 и выше) и обладает многими функциями, свойственными профессиональным пакетам этого класса: автоматически создает папку для хранения файлов переводческого проекта, формирует таблицу параллельных фрагментов и позволяет искать в исходном документе текстовые элементы из накопителя ТМ. Модуль управления глоссариями служит для работы со словарем терминов. Для автоматического построения БД накопителя переводов на основе сопоставления параллельных текстов служит дополнительный макромодуль WordAlign, более простой и быстрый, чем аналогичные средства в составе пакетов Trados, DejaVu и Star.

Полнофункциональную ознакомительную версию WordFisher можно бесплатно загрузить по адресу: www.wordfisher.com/wf4.htm.

IBM TranslationManager 2.0

Программа TranslationManager корпорации IBM предоставляет стандартные для ТМ-систем возможности работы с БД накопителя переводов, группировки рабочих файлов в проекты, управления словарем терминов, автоматического анализа и сегментации текстов. Предусмотрен API для встраивания функций TranslationManager в приложения.

TranslationManager работает в среде Windows 9x, NT и OS/2 и обеспечивает совместимость с форматами RTF, HTML, WinWord, Word Pro, WordPerfect, PageMaker и QuarkXPress. Программа поддерживает перевод для любых пар языков из списка, включающего английский, немецкий, французский, испанский, итальянский, португальский, нидерландский, русский, шведский, греческий, иврит, японский (полная поддержка гарантируется для 34 языков). В комплект поставки входят базовые словари терминов (для полноценной работы необходимо расширять их или приобретать более совершенные), средство управления базами данных ТМ и сами данные.

На странице www-4.ibm.com/software/ad/translat/tm/tmdown.html доступна демоверсия программы, действующая в течение 60 дней. К сожалению, судя по информации на узле IBM, корпорация не очень активно развивает направление своей деятельности, связанное с ТМ (демоверсия датирована 1996 г.), однако и имеющихся функций достаточно, чтобы существенно облегчить работу переводчика.