



PROMT 9.0: для профессионалов и бизнесменов

Сегодня с системами машинного перевода (МП) так или иначе знакомы практически все пользователи. Соответствующие онлайн-службы встраиваются в самые разнообразные прикладные программы, хотя наибольший охват аудитории обеспечивают, конечно же, браузеры. Однако, несмотря на отсутствие революционных прорывов в области МП, развивается и традиционное настольное ПО, готовое стать верным помощником даже профессиональному переводчику.

Обычно обновляемая раз в два года система машинного перевода PROMT задержалась на уровне 8-й версии (включая релиз 8.5, ko.com.ua/42925) несколько дольше, зато разработчикам, похоже, удалось накопить кое-какой потенциал. Собственно ядро перевода PROMT 9 претерпело заметные изменения: помимо экстенсивных усовершенствований (улучшена обработка сложных семантических моделей, переработаны и дополнены лингвистические базы), в нем также применены корпусные и статистические методы (данное направление развивает, в частности, Google). Адекватно оценить объем таких новаций довольно трудно, хотя эффект в общем-то виден – перевод стал более гладким и связным, особенно когда речь идет об об-

работке конструкций, содержащих омонимы (представляющие разные части речи), такие неоднозначные служебные слова, как that, as, притяжательные местоимения, имена собственные, даты, время, цвета. С типовыми образцами непосредственных улучшений в переводе можно ознакомиться в официальном документе (www.promt.ru/images/ru/translation_software/pdf/compare_promt90_promt85.pdf).

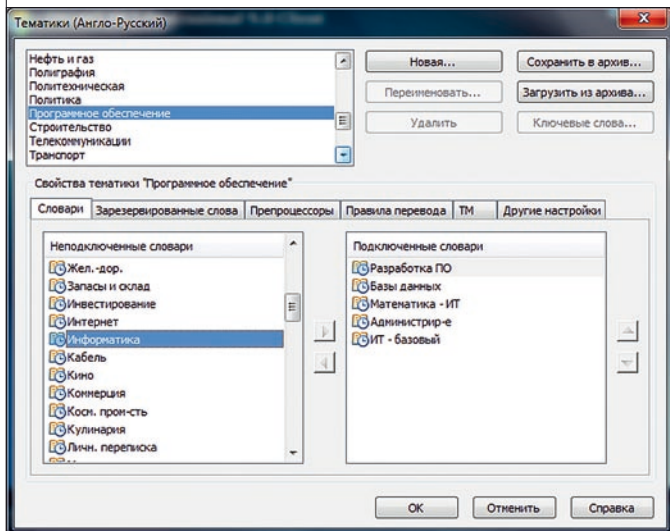
Также по общему впечатлению PROMT 9 быстрее предыдущей версии обрабатывает большие файлы, скажем, на перевод технической документации объемом 540 страниц ушло около получаса. Если при этом удастся обеспечить приемлемое качество «подстрочника», то экономия на чисто технических операциях становится заметной. Еще одно наблюдение: с английским языком программа работает и оперативнее, и лучше, чем с остальными направлениями (табл. 1). Когда же предметная область находится на стыке разделов знаний, определенно нужно позаботиться о подготовке узкоспециализированных словарей. В одном из наших экспериментов дополнение тематики «Финансы» всеми имеющимися в наличии математическими словарями не помогло добиться корректного перевода терминов из финансовой математики. В целом же правильный подбор тематик оказывает заметное влияние на качество перевода (табл. 2), и это одно из очевидных направлений развития

PROMT 9.0 – теперь пользователь может выбирать из пяти конфигураций с поддержкой шести языков (английский, немецкий, французский, итальянский, испанский, русский) и разными наборами специализированных словарей (до 123 тематик).

Клиентские приложения PROMT 9 могут встраиваться в Microsoft Outlook и основные приложения Microsoft Office (версия 2010 пока не поддерживается), OpenOffice.org, Adobe Acrobat/Reader, ICQ, Skype, QIP, MSN Messenger. Однако при работе с документами различных форматов необходимо учитывать особенности их формирования. Например, качество перевода PDF во многом зависит от того, каким образом расставлены переносы («жесткий» дефис программой не воспринимается) и как документ разбит на абзацы (перевод строки программа интерпретирует как конец предложения, соответственно, количество ошибок и неточностей заметно увеличивается). К сожалению, подобные нюансы все еще не редкость, и разработчики могли бы учесть наиболее частые из них. Впрочем, также понятно, что нельзя объять необъятное, поэтому перед отправкой в перевод большого документа нелишне вначале поэкспериментировать с каким-то его фрагментом.

Всю продуктовую линейку PROMT можно разбить на три группы: для индивидуальных пользователей (от домашних до профессионалов), для перевод-

Набор словарей – основа формирования тематик, более тонкая настройка осуществляется за счет определения зарезервированных терминов и подключения правил перевода



ческих бюро, для обычных компаний, в которых перевод не является основной деятельностью. Формирование последней отражает тенденции глобализации и интеграции современного бизнеса. При этом задачи, стоящие перед компаниями, выходящими на мировой рынок или взаимодействующими с зарубежными партнерами, можно условно разделить на два класса: оперативное ознакомление с документом на иностранном языке (когда поток такой деловой корреспонденции достаточно велик) и точный профессиональный перевод в формате исходного документа для последующего использования или распространения.

Для решения задач первого типа предлагается PROMT Translation Server (PTS) Intranet Edition 9.0 – эта система развертывается и работает централизованно на сервере, а доступ к инструментам перевода осуществляется через веб-интерфейс (поддерживаются браузеры Internet Explorer, Mozilla Firefox, Opera) и внутри локальной сети, и (при соответствующей настройке)

Таблица 1. Результаты перевода в пакетном режиме с использованием тематик «Информатика и вчт» для двух языков

Язык	Исходный текст	Перевод
Английский	<p>Typographical conventions</p> <p>Menu commands and dialog box components are identified with bold type throughout this manual. For example, File > New refers to the command New in the File menu.</p> <p>File names are given as follows: C:\IMAGES\BEACH.JPG</p> <p>The Enter key is represented by (↵), the arrow keys by (←), (→) etc. To enter a key combination such as (Alt)(F), press the (F) key while holding down the (Alt) key.</p> <p>Important: Especially important notes or useful tips are highlighted like this.</p>	<p>Типографские соглашения</p> <p>Команды меню и компоненты диалогового окна идентифицируются с жирным шрифтом всюду по этому руководству. Например, Файл> Новый обращается к команде New в меню File.</p> <p>Имена файлов даются следующим образом: C:\IMAGES\BEACH.JPG</p> <p>Клавиша Enter представляется (↵), клавишами со стрелками (←), (→) и т.д. Ввести а сочетание клавиш, такое как (Alt) (F), нажмите (F) клавишу во время как удержание (Alt) ключ.</p> <p>Важный: Особенно важные примечания или полезные подсказки выделяются как это.</p>
Немецкий	<p>Schreibweisen</p> <p>Menubefehle werden in diesem Handbuch durch Fettdruck gekennzeichnet. Datei > Neu bezeichnet beispielsweise den Befehl Neu im Menu Datei.</p> <p>Bestandteile eines Dialogfensters sind ebenfalls fettgedruckt. Zum Beispiel: „Wählen Sie in der Liste Name einen Dateinamen und klicken Sie auf die Schaltfläche OK.“</p> <p>Dateinamen werden folgendermaßen dargestellt: C:\BILDER\SCHULZ.JPG</p> <p>Die Eingabetaste wird als (↵) dargestellt, die Richtungstasten als (←), (→) etc.</p> <p>Um eine Tastenkombinationen wie beispielsweise (Alt)(F) einzugeben, betätigen Sie die Taste (F) bei gedruckter Alt-Taste.</p>	<p>Стили</p> <p>Команды меню характеризуются в этом руководстве жирной печатью. Файл > По-новому обозначает, к примеру, команду По-новому в меню Файл.</p> <p>Составные части окна диалога также напечатаны жирным шрифтом. Например: „Выбирать Они в списке имя имя файла и Вы нажимаете кнопку ОК.“</p> <p>Имена файлов представляются следующим образом: C:\BILDER\SCHULZ.JPG</p> <p>Клавиша ввода представляется как (↵), клавиши направления как (←), (→) и т. д.</p> <p>Чтобы давать комбинации клавиш как (Alt)(F), приводят в действие Они клавишу (F) при нажатой клавише Alt.</p>



Шоста науково-технічна конференція-виставка

DOCFLOW УКРАЇНА 2010



Управління електронними та паперовими документами

21 вересня 2010, Київ

Ресстрація на www.DOCFLOW.ua

Для зареєстрованих слухачів відвідування БЕЗКОШТОВНЕ

Оргкомітет: тел. (44) 490-9999, docflow@abbyu.ua

<p>Організатор:</p> 	<p>Генеральний спонсор:</p> 	<p>Спонсори:</p>  	<p>Бізнес-медіа партнер:</p> 
<p>ІТ-медіа партнер:</p>  <p>www.ko-online.com.ua</p>		<p>Інформаційні спонсори:</p>   	



Таблица 2. Перевод технической документации, демонстрирующий влияние нерелевантной тематики на его качество в части, не относящейся к терминологии

Оригинал на английском	Оригинал на русском	Перевод PROMT, тематика «Информатика и вчт»	Перевод PROMT, тематика «Бизнес»
<p>TEST "VIRUS" EICAR AND ITS MODIFICATIONS</p> <p>This test «virus» was specially developed by (The European Institute for Computer Antivirus Research) for the testing of anti-virus products.</p>	<p>ТЕСТОВЫЙ «ВИРУС» EICAR И ЕГО МОДИФИКАЦИИ</p> <p>Тестовый «вирус» был специально разработан организацией (The European Institute for Computer Antivirus Research) для проверки работы антивирусных продуктов.</p>	<p>ПРОТЕСТИРУЙТЕ «ВИРУС» EICAR И ЕГО МОДИФИКАЦИИ</p> <p>Этот тест «вирус» был особенно разработан (Европейский Институт Компьютерного Исследования Антивируса) для тестирования антивирусных продуктов.</p>	<p>ПРОВЕРЬТЕ «ВИРУС» МОДИФИКАЦИИ ITS И EICAR</p> <p>Этот тест «вирус» был особенно развит (Европейский Институт Компьютерного Антивирусного Исследования) для тестирования антивирусных продуктов.</p>
<p>Never use real viruses for testing the operation of an anti-virus product!</p>	<p>Никогда не используйте в качестве проверки работоспособности антивирусного продукта настоящие вирусы!</p>	<p>Никогда не используйте реальные вирусы для того, чтобы протестировать работу антивирусного продукта!</p>	<p>Никогда не используйте реальные вирусы для того, чтобы проверить работу антивирусного продукта!</p>
<p>The application identifies the file downloaded from the EICAR site as an infected object containing a virus that cannot be disinfected and performs the actions specified for this type of object.</p>	<p>Программа идентифицирует файл, загруженный с сайта компании EICAR, как зараженный объект, содержащий не подлежащий лечению вирус, и выполняет действие, установленное для такого объекта.</p>	<p>Приложение идентифицирует файл, загруженный с сайта EICAR как зараженный объект, содержащий вирус, который не может быть дезинфицирован и выполняет действия, определенные для этого типа объекта.</p>	<p>Применение идентифицирует файл, загруженный с сайта EICAR как зараженный объект, содержащий вирус, который не может быть дезинфицирован и выполняет действия, указанные для этого типа объекта.</p>

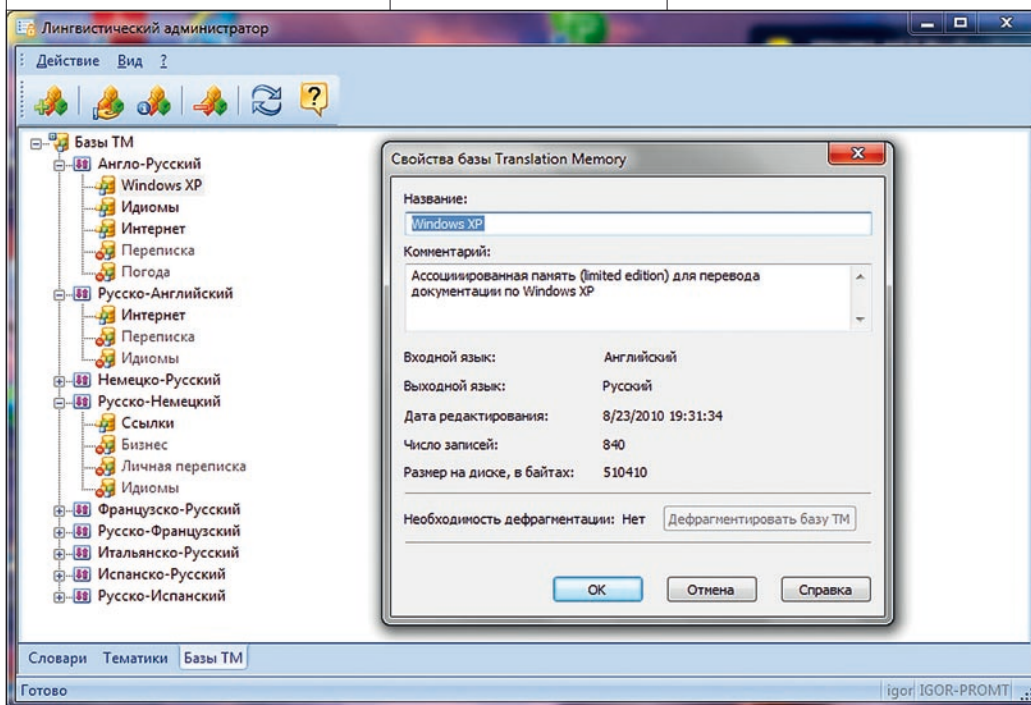
извне. По сути, схема такая же, как у бесплатного сервиса translate.ru: с рабочего места пользователя на сервер с PTS отправляется запрос на перевод, который в зависимости от загрузки обрабатывается либо помещается в очередь на обслуживание. Однако на PTS-сервере применяются новейшее ядро, специализированные словари (в том числе пользовательские), единая корпоративная база переводов.

Управление базой знаний переводчика и доступом к ним сосредоточено в специальном приложении для лингвистического администратора

На качественную подготовку документов в рамках компании рассчитан PROMT NET Professional 9.0. Он пригодится как профессиональным переводчикам, так и квалифицированным специалистам, способным сформировать терминологическую базу, обеспечить сходную стилистику и внутренние стандарты. При этом речь идет вовсе не о замене персонала на системы MT, а о минимизации времени на первич-

ную обработку текста, обращение к словарям, поддержку используемых форматов, т. е. о повышении производительности труда. Затраты на окончательное редактирование перевода также могут снизиться, особенно при активном задействовании технологии Translation Memory, однако надо иметь в виду, что последнее предполагает определенный период накопления базы эталонов и обучения пользователей. В целом, по оценкам самих разработчиков, применение PROMT 9 сокращает время на профессиональный перевод документов примерно на 25%. На самом деле можно добиться даже большей экономии, прибегнув к пакетной обработке, выполняемой в фоновом режиме, причем для каждого исходника можно задать собственные параметры – от направления и тематики перевода до формата результирующего документа и пр.

В архитектурном плане PROMT NET Professional радикально отличается от PTS. В этом случае сервер выполняет лишь второстепенные задачи, к которым относятся управление лицензиями и доступом, хранение словарных баз и персональных настроек. А на рабочие станции устанавливается клиентское приложение, функционально соответствующее персональному PROMT NET Professional 9.0. Авторизация

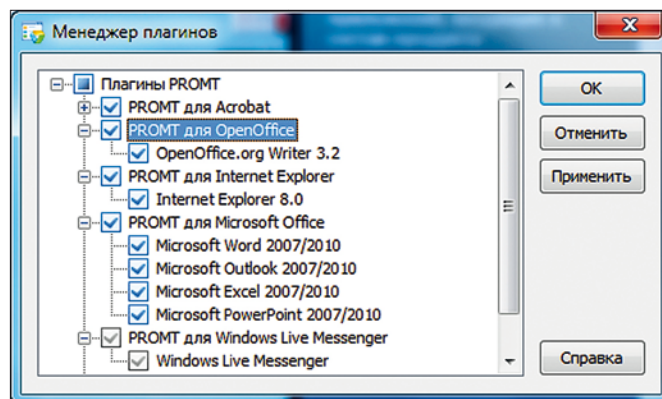


пользователей в такой распределенной системе производится средствами Windows, внутри PROMT учетным записям назначаются три стандартные роли: администратор, лингвистический администратор и пользователь (права последнего ограничены доступом исключительно к клиентской части PROMT). Как обычно, обеспечивается достаточно гибкое централизованное управление лицензиями, в том числе параллельными, что позволяет подобрать оптимальную схему и застраховаться от определенных ситуаций (когда, к примеру, пользователю англо-русского ПО вдруг понадобится перевод с немецкого).

Пожалуй, наиболее существенным архитектурным ограничением PROMT NET Professional является необходимость использования Windows Server для серверного компонента – потенциальным заказчиком, вероятно, имеет смысл заранее задуматься о применении для подобных целей виртуальных машин в рамках соответствующих лицензий Microsoft. При этом, однако, доменная ар-

хитектура не требуется, документация содержит необходимые сведения для развертывания системы в рамках одноранговой сети.

В PROMT NET Professional 9.0 предложен новый механизм настройки перевода на основе выбора тематики переводимого текста, управление которым полностью ложится на плечи лингвистического администратора. Именно он определяет, какой перечень тематик и словарей может применять каждый пользователь, а также права этих пользователей по пополнению имеющихся баз данных и специализированных словарей. В ведении лингвистического администратора также находится формирование новых тематик на основе общих и специализированных словарей и загрузка их в/из архива. Естественно, создать словарь для персонального применения вправе любой пользователь PROMT, но лингвистический администратор посредством соответствующих инструментов сравнения и слияния может на основе разрозненных источников сформировать общий словарь узкоспециализи-



Удобство использования PROMT обеспечивают подключаемые модули для популярных приложений

рованной лексики либо с помощью программы извлечения терминологии из текста (с сохранением контекста, в котором она применяется) заранее позаботиться о наличии единого глоссария для конкретного проекта.

Для мобильных пользователей, не подключенных к сети и не имеющих возможности постоянного доступа к серверу, предусмотрено специальное приложение для синхронизации данных – этот компонент приобретается отдельно и выполняет перенос выбранных словарей, тематик и баз между локальным ПК и сервером.

ITC PUBLISHING
технологии для человека

ДОПОЛНИТЕЛЬНАЯ ВОЗМОЖНОСТЬ ДЛЯ ПОДПИСЧИКОВ

Теперь вы можете получать журнал
«Компьютерное Обозрение» и в
PDF-формате
дополнительно к печатной версии

Для этого вам необходимо прислать заявку и копию документа, подтверждающего подписку, на факс **(044) 496-34-69** или e-mail **sveta@itc.ua, oksana@itc.ua**.

В заявке укажите срок подписки на печатную версию и адрес электронной почты.

Ссылка на PDF-версию будет отправляться на указанный вами e-mail начиная с выхода ближайшего номера после подачи заявки и до окончания срока подписки.

