



Сократ



Prompt Express



Retrans



Pragma

Обзор систем машинного перевода

# Экзамен по английскому

*Lupus Tamboveus tuus amicus est*  
вольный перевод с русского

Проблема, о которой пойдет речь в этой статье, особая. Это, на мой взгляд, даже не столько компьютерная или, скажем, лингвистическая проблема. Это одна из главных религиозно-этических трудностей человечества, — языковая. Религиозно-этическая, потому что возникла она, как известно, по воле не кого-нибудь, а самого Господа Бога...

**Е**сли верить библейским преданиям, у древних людей был всего один язык, а не сотни, как сегодня. Трудностей с общением у них не было никаких, и поэтому вот так сообща надумали однажды наши пращуры построить башню до небес на территории будущего Вавилона. Но строительству этого первого в истории человечества небоскреба не суждено было закончиться — оно вскоре было «заморожено» Господом путем разделения языков у его проектировщиков: «Сойдем же, и смешаем там язык их, так чтобы один не понимал речи другого» (Бытие, 11-7).

И с тех пор люди разных народов перестали понимать друг друга. Возникшие языковые трудности кое-как решались в докомпьютерную эпоху тремя путями: с помощью специалистов-переводчиков, нудным штудированием учебников и словарей, а иногда и вовсе мимикой и жестами. Никаких других вариантов не было.

Но теперь, с приходом в нашу жизнь современных компьютерных технологий, возник соблазн переложить и эту застарелую человеческую проблему на хрупкие виртуальные плечи машин. Соблазн-то, конечно, возник, но вот осуществить желанное оказалось не так-то просто. Построить новую виртуальную «Вавилонскую башню» пока никому так по-настоящему и не удалось. Чаще всего машинный перевод нашего времени

напоминает бред сумасшедшего, к тому же еще и косноязычного.

В числе особо невезучих, «трудных» с точки зрения перевода, языков оказался наш родной, русский. Все дело в его так называемом синтетическом строении, когда слова стыкуются друг с другом при помощи своих окончаний. А не как, скажем, в английском, аналитическом, где падежи связывать не надо — все и так сделают предлоги.

Ничего удивительного поэтому, что перевод с русского на английский у компьютера получается немного лучше, нежели обратный, с английского на русский. Количество вариантов значительно меньше. Одним словом, сочность и богатство нашего языка сослужили плохую службу программам компьютерного перевода.

С трудностями виртуальные переводчики справляются по-разному и с разной степенью успеха. Выяснить, кто из них преуспел больше на ниве языкознания, — задача нашего экзамена по английскому. И надо сказать, у нас есть из чего выбрать. У долго и безоговорочно лидировавшей коммерческой системы машинного перевода PROMT (в девичестве — Stylus), созданной питерской фирмой ProjectMT, недавно появилась наконец-то сильная соперница по имени Retrans, придуманная научным коллективом ВИНТИ РАН — Всероссийским институтом научной и технической информации. В общем, наука

бросила вызов бизнесу. Третьим, но далеко не лишним участником лингвистического забега стал продукт «Арсенала» — переводчик «Сократ». В открытом чемпионате России по компьютерному переводу приняла участие и представительница братской Украины — программа Pragma. Это новинка от создателей знаменитого модуля для Word, Language Master — фирмы Trident Software.

Сразу хочу договориться с вами о терминах. Первых двух участниц и четвертую украинскую я буду, вслед за их создателями, именовать в женском роде, поскольку речь идет именно о системах машинного перевода, а не просто о каких-то там переводчиках. А вот для «Сократа» — только из уважения к древнегреческому философу — решил сделать исключение из грамматических правил.

Сама процедура экзамена, естественно, проходила в равных, или, как сейчас принято говорить, равноудаленных, условиях: на одном и том же компьютере, с одними и теми же текстами о древнеегипетских сфинксах, которые нужно было перевести с английского языка. Варианты, предложенные конкурсантами, сравнивались с контрольным текстом перевода, сделанным опытным лингвистом. Результаты тоже, кстати, оказались у экзаменуемых практически одинаковыми. Но с некоторыми все же нюансами...

■ ■ ■ Александр Евдокимов,  
кандидат филологических наук

## Retrans: академический подход

Retrans

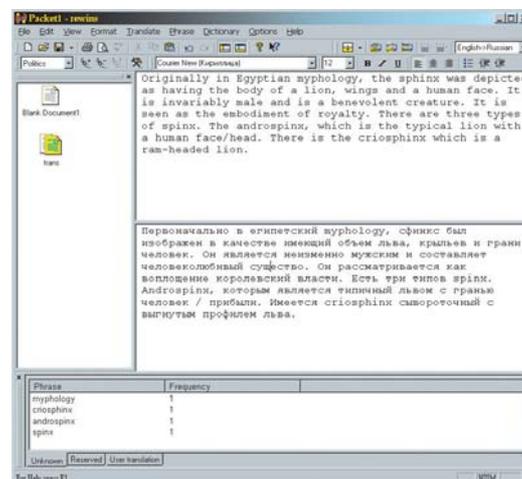
### ПЛЮСЫ/МИНУСЫ

- + сравнительно хорошая сочетаемость слов в переведенных текстах
- + интегрированные специализированные словари
- + сохранение форматирования при переводе
- много грамматических ошибок
- только два направления перевода
- высокая цена коммерческой версии

Система машинного перевода Retrans из всех имеющихся у меня под рукой оказалась наиболее точной. Подчеркну, что из имеющихся, поскольку у PROMT есть, например, несколько версий, и для сравнения была выбран PROMT Express. Клянусь, что не подсуживал своим братьям-ученым. Их победа определяется не моими симпатиями, а объективной близостью к контрольному варианту, сделанному переводчиком.

Успех Retrans во многом объясняется принципиально новым подходом к самому процессу перевода. Вместо традиционного, словарного, в ней используется так называемый фразеологический метод. Поясню для нефилологов: в Retrans текст переводится не пословно, а целыми словосочетаниями. И это вполне разумно, ведь мы же с вами говорим и думаем не отдельными словами, а именно словосочетаниями — синтагмами, почему же машина должна поступать по-другому?

Выигрышным моментом именно Retrans, на мой взгляд, является стилистическая и



семантическая согласованность текста перевода. Слабой позицией этой системы относительно того же PROMT оказалась грам- >>

» матическая сочетаемость. Вариантов наподобие: «Чей туфля?» — «Мое», — в этой программе было довольно много.

В качестве примера сравним перевод английской фразы: «It is seen as the embodiment of royalty». У Retrans получилось вот что: «Он рассматривается как воплощение королевской власти». У PROMT: «Это замечено как воплощение лицензионного платежа». У «Сократа»: «Видно как реализация авторского гонорара». Почти то же самое у украинской Pragma: «Это видится, как воплощение авторского гонорара». Во втором и третьем случае нет ни смысловой, ни грамматической ошибки при переводе, и эти варианты следовало бы признать наиболее удачными, если бы не одно но. В нашем тексте речь шла о могуществе древнеегипетских сфинксов, которые вряд ли осуществляли лицензионные платежи и уж точно не получали авторских гонораров. У Pragma, помимо смыслового промаха, еще и лишняя запятая нарисовалась. В точку попала только Retrans. Но у этой системы своя беда — безграмотность. Двоичница Retrans, видимо, просто не знает, что власть всегда женского рода, даже если речь идет о государстве сфинксов и фараонов. Кстати, человек перевел эту фразу следующим образом: «Он представляется воплощением царской власти».

На сайте разработчика Retrans вы без труда найдете ее демоверсию в RAR-архиве общим объемом примерно 110 Мбайт. Программу можно скачать или целиком, или частями по 5 Мбайт. Перед инсталляцией распакуйте архив в любую удобную для вас

папку и запустите Setup.exe. Все остальное инсталлятор сделает сам.

В демонстрационной версии Retrans есть ряд ограничений: во-первых, доступен только английский вариант интерфейса, а во-вторых, нет никакой возможности сохранить получающийся перевод. Связь программы с окружающим миром Windows заключается в помещении содержимого буфера обмена в верхнее окно программы. Что-либо копировать из нижнего окошка, где помещается полученный перевод, нельзя. В коммерческой версии программы, предназначенной для широкого круга пользователей, причем не только Windows, но и Linux и даже Solaris, таких проблем нет.

Для меня так и осталось загадкой: ограничение ли это демоверсии или просто-напросто глюк, но почему-то Retrans безропотно соглашалась переводить только файлы html-формата или только тексты из буфера обмена Windows. Для остальных же форматов, если вдруг у вас ничего не произойдет, горячо рекомендую установить в исходном английском файле шрифт Times New Roman (Западноевропейский). И обязательно удалите текст из нижнего окошка, иначе могут возникнуть проблемы с раскладкой.

Запуск же самого процесса перевода прост до смешного: Translate -> Document. Но впрочем, вы можете перевести не весь документ, а лишь его выбранную часть (Selected Text), отдельный параграф (Paragraph) или одно предложение (Sentence). Сменить направление перевода или под-

ключить тот или иной дополнительный словарь можно в меню Show Document Properties, открываемым по щелчку правой клавишей мыши на значке открытого файла в окне структуры пакета (Package Structure)

Общее впечатление от Retrans у меня в целом осталось неплохое, хотя, без сомнения, объем основного «политематического» словаря можно и даже нужно еще увеличивать и увеличивать. А самое главное, разработчикам программы следует уделить особое внимание элементарной грамматике. А то стыдно, честное слово, что у маститых академиков программа пишет «поутре» вместо «поутру», «люда» вместо «люди». Но не будем слишком строги, ведь профессорское дитя пошло лишь в третий класс в этом году, то есть добралось лишь до третьей версии.

Хочется верить, что эта интересная разработка получит дальнейшее развитие и станет серьезным подспорьем для профессионалов. Все данные для этого есть: и хорошая работа программистов, и специалистов по русскому языку, разработавших словари программы. Заслуживает внимания Retrans и как один из немногих компьютерных переводчиков, у которого существует версия не только для Windows, но и для Linux.

### Retrans

Разработчик ► ВИНТИ

Сайт разработчика ► <http://www.viniti.ru>

Условия распространения ► shareware

Цена ► \$290 в минимальной комплектации

## Фразеологический метод перевода

### Что такое перевод «по блокам»

Фразеологический метод был предложен профессором Г. Г. Белоноговым, заведующим отделом лингвистических исследований ВИНТИ, еще в... 1975 году. То есть тогда, когда не то что Windows, но даже DOS еще не было даже в проекте. Если это действительно так, то выходит, что и систему машинного перевода впервые придумал и обосновал тоже наш с вами соотечественник. Что же предложил основоположник научного перевода, названный в 1995–1996 годах Кембриджским

биографическим центром за свои открытия в этой области Человеком Года? Кроме собственно опоры на словосочетания Г. Г. Белоногов считает необходимым использовать при машинном переводе так называемые «речевые модели». Имеются в виду широко распространенные словосочетания, имеющие, согласно его терминологии, «пустые места», которые можно при переводе «заполнять» другими, менее расхожими словами и конструкциями. В результате такого подхода из 1 300 000

словарных статей основного, «политематического» словаря Retrans 77%, согласно сведениям разработчиков, приходится на словосочетания от двух до семнадцати слов.

Но кроме основного, в их программе имеется еще и 24 дополнительных словаря. Эти специализированные словари общим объемом 100 тыс. статей существенно улучшают качество перевода текстов по различным темам от физики и математики до военного дела и права.



Prompt Express

# PROMT Express: давай сделаем это по-быстрому

## ПЛЮСЫ/МИНУСЫ

- + много возможных направлений перевода
- + грамматическая точность в переведенных текстах
- + возможность озвучивания текста
- потеря форматирования при переводе
- дополнительные словари не входят в дистрибутив
- нет демоверсии

Пока система Retrans растет и набирается лингвистических сил и коммерческого опыта, лидером продаж остается все та же PROMT. В нашем соревновании она, правда, уступила. Но, во-первых, совсем немного, а во-вторых, и вышла эта программа на старт, пользуясь спортивной терминологией, вторым, дублирующим составом. Даже из названия PROMT Express со всей очевидностью следует, что для ее разработчика здесь важна была не столько точность, сколько скорость и компактность программы. Задача PROMT Express — быть всегда под рукой.

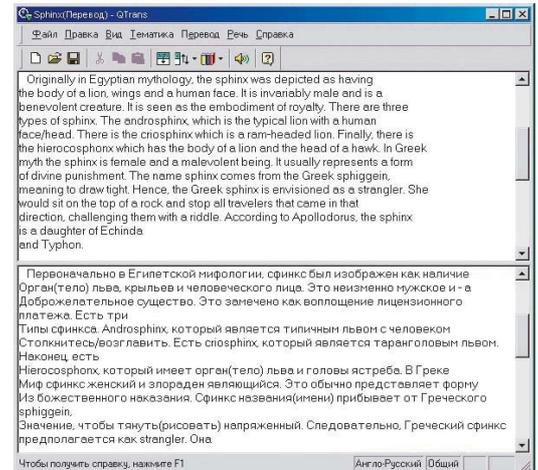
Кроме того, рождение этой облегченной версии — отличный маркетинговый ход. В Москве, по крайней мере, jewel-версия одного направления перевода PROMT Express стоит всего 100 руб., что практически снимает с повестки дня проблему пиратства.

Дополнительных словарей труженики ProjectMT создали уже более 100, объединив их в шесть тематических коллекций: «Коммерция», «Наука», «Техника», «Легкая Промышленность», «Тяжелая Промышленность» и «Домашняя». Если этих словарей

у вас нет, то все, чем можно помочь PROMT перед началом перевода, — это указать ей, какие слова перевести не стоит. А вот этого я не догадался сразу сделать, и в одном из дополнительных текстов, предложенных мною для перевода, PROMT упорно стала называть американского переводчика, по фамилии Saltzer, Сельтерской водой. Хорошо еще, что не виски. Между прочим, Retrans, «Сократ» и Pragma этой нелепой ошибки с фамилией своего гуманоидного коллеги не допустили.

Инсталляция PROMT Express вряд ли вызовет у вас какие-нибудь затруднения, разве что обременительной покажется необходимость дважды вводить контрольные наборы цифр. Сначала артикульный номер, а затем еще и номер лицензионного соглашения. Для каждого из поддерживаемых программой языков — английского, немецкого, французского и итальянского — нужно приобрести и установить свой отдельный модуль.

Внешне PROMT ничем не отличается от остальных участниц языкового экзамена — все тот же двухоконный интерфейс. Вверху исходный текст, внизу — получающийся перевод. Запуск процесса перевода оригинальностью тоже не отличается: Перевод -> Перевести. Аналогична конкурентам и смена направления перевода: Перевод -> Из-



менить направление и тематику. Специфической возможностью PROMT я бы, пожалуй, назвал так называемый синхронный перевод (Перевод -> Синхронный перевод). Активизация вами этой функции приведет к тому, что ввод текста в верхнем окне будет сразу же продублирован появлением его перевода в нижнем.

Одним словом, если вам нужен качественный продукт с хорошей родословной за не слишком большие деньги, то лучше программы PROMT Express вы вряд ли что найдете. Но не ждите от нее невозможного.

### PROMT Express

- Разработчик ► Компания PROMT
- Сайт разработчика ► [www.prompt.ru](http://www.prompt.ru)
- Условия распространения ► shareware
- Цена ► \$31 (полная версия)



## Перевод — математическое моделирование

### Математическая рулетка

Как и Retrans, PROMT была основана в 1993 году. Еще в конце 70-х годов прошлого века будущие создатели лучшего виртуального переводчика разрабатывали теорию МП (машинного перевода) в лаборатории Инженерной лингвистики Ленинградского педагогического института. В отличие от москвичей ленинградские ученые сделали ставку не на лингвистические аспекты перевода, а на математическое моделирование этих процессов. Сами разработчики называют свою систему довольно сложно,

но красиво: «Технологии расширенных сетей переходов и формализма нейронных сетей». В основе этой теории понимание текста как некоей взаимосвязанной цепочки элементов — лексем. Угадает машина, что лежит в основе этой сети, не сообразит при ее дальнейшей расшифровке — получится логичный, связанный семантически (по смыслу) и грамматически (по правописанию) перевод. Ну а не угадает — что ж, не судьба, извольте довольствоваться все тем же бессмысленным набором слов. В последних версиях PROMT

стало возможным помогать машине в решении этой архисложной задачи. Для этого всего-то и нужно подсказать программе тематику предлагаемого текста (Тематика -> Выбрать тематику). Можно также подредактировать уже готовые варианты и даже добавить свои. Можно, но при одном условии, если вы приобрели дополнительные словари, которые следует докупать и устанавливать отдельно. Так же как и голосовой движок, чтобы переведенный текст можно было бы еще и прослушать (Речь -> Произнести).

# # Прagma. Що це таке?

Pragma

## ПЛЮСЫ/МИНУСЫ

- + гибкая система разделов основного словаря
- + интеграция перевода в Проводник и ряд приложений Windows
- + возможность автоматического обновления через Интернет
- лексически неправильный перевод отдельных слов
- нет дополнительных тематических словарей
- часто ошибочная расстановка слов в переведенных предложениях

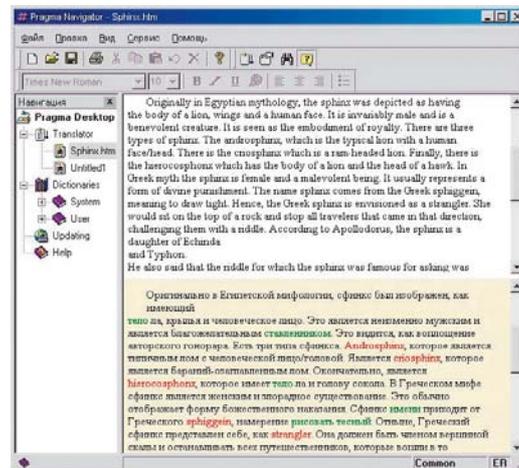
Наконец-то. Украинцы сделали свой незадолго-самостоятельный PROMT и даже назвали его весьма созвучно — Pragma. Нет, это, конечно, не калька с российской программы. Pragma — это вполне самостоятельный и достаточно качественный продукт. Помимо английского и русского модуля в этой программе, понятное дело, есть и свой украинский. В Украине сейчас, как известно, любой официальный документ или справку нужно обязательно переводить на мову великого кобзаря Тараса Шевченко. Поэтому важность этого модуля для преимущественно русскоговорящих граждан юга и востока Украины трудно переоценить.

Pragma возникла не на пустом месте: она заняла место своего предшественника — Language Master, программы перевода, интегрировавшейся в Microsoft Word 97. Теперь Pragma полностью самостоятельная система машинного перевода, состоящая из двух частей: Monitor — своеобразного центра управления, сидящего в трее, а также Navigator — собственно программы перевода. С помощью Pragma Monitor (меню

«Параметры») систему перевода можно встроить в Microsoft Word (причем не только 97, но и 2000), Word Pad и даже (!) в «Блокнот». Кроме того, при активизации Pragma Monitor в контекстном меню «Проводника» появится новая строчка — «Перевод». Функция эта, правда, пока не работает.

Pragma Navigator запускается по двойному щелчку мыши на значке Monitor. Двухоконный интерфейс программы обычен для программ-транслейторов и аналогичен уже рассмотренным нами выше. Перевод запускается из меню Сервис -> Перевод. Перед началом непосредственно самого таинства вас попросят определиться с его темой и направлением. Дело в том, что 600-тысячный словарь Pragma делится на смысловые блоки: общий, коммерческий, компьютерный, правовой, технический, химический и медицинский.

Объективно (да простят меня украинские программисты) перевод с помощью Pragma предложенного на конкурс текста оказался наименее удачным. Например, знаменитую загадку сфинкса, по-английски «riddle», Pragma почему-то обозвала... решетом. На саму загадку: кто утром ходит на четырех ногах, днем — на двух, а вечером — на трех? — Pragma, не долго думая, ответила «мужчина» (возможно, из-за того, что на украинском мужчина — это и есть «человек»). Все остальные участники соревнова-



ния, хотя и с пятого на десятое, сформулировали сам коварный вопрос, но ответ на него все же дали правильный — «человек».

Надеюсь, что разработчики поправят в последующих версиях то, что пока хромает. А ошибки, к сожалению, еще имеются. И их не могло не быть, если учесть, что в разработке такого интересного и сложного проекта, как Pragma, участвует всего... 5 программистов. Просто удивительно, как им удалось при этом создать вполне конкурентоспособную программу, которой пользуется президентская администрация и кабинет министров Украины.

## Pragma

- Разработчик ▶ Trident Software
- Сайт разработчика ▶ [www.trident.com.ua](http://www.trident.com.ua)
- Условия распространения ▶ shareware
- Цена ▶ \$180



## Ждите, пожалуйста, ответа...

Ой, боюсь, попадись Pragma сфинксу пару тысячелетий назад, не избежать бы ей страшной участи тех несчастных путешественников, которые тоже что-то там напугали при ответе. Сфинкса ведь не волновало, знаешь ты иностранные языки или не знаешь, хороший у тебя переводчик или не очень. Если что не так — незадачливого путешественника ждала участь «завтрака туриста». Хорошо, что мы живем в другое, в основном цивилизованное время, когда за

ошибки при переводе нет риска быть, по крайней мере, съеденным. Хотя, как знать, как знать, — вот ведь несколько лет назад потерпел крушение авиалайнер одной турецкой авиакомпании из-за того только, что пилоты как-то не так поняли инструкцию на английском языке и неправильно закрыли люк грузового отсека. Элементарная лингвистическая ошибка обернулась страшной технологической катастрофой.

Кстати, о технологии. У Pragma, к сожалению, и технические оплошности тоже имели место: мне вот так и не удалось осуществить быстрый перевод буфера обмена. Все вроде бы делал по инструкции — щелчок мыши по значку Pragma Monitor в трее с одновременным нажатием кнопки Ctrl на клавиатуре, а ничего, кроме окошка с просьбой подождать минутку, я не получил. Терпеливо ждал минутку, две, пять, десять — ничего. Ни словечка...

# «Сократ» мне друг, но...

Сократ

## ПЛЮСЫ/МИНУСЫ

- + скорость перевода
- + автоматическое определение языка исходного текста
- + довольно точная пунктуация в переводах
- много логических ошибок
- нет возможности подключать дополнительные словари
- нельзя корректировать определение тематики текста

От команды фирмы «Арсенал» в нашем экзамене участвовал самый юный, а стало быть, и самый продвинутый представитель семейства переводчиков «Сократ» — «Сократ Персональный» версии 4.0. Существует еще «Сократ Профессиональный» — более старший и более мощный представитель этого уважаемого семейства, но этот транслейтор давно уже не обновлялся. Принципиальное различие двух «Сократов» в возможности подключения десяти дополнительных тематических словарей, а также в количестве направлений возможного перевода. К русско-английскому и англо-русскому в профессиональной версии добавились еще направления с французским и немецким языками.

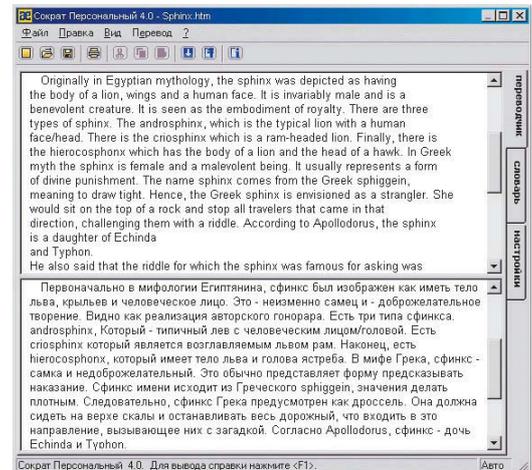
Новенький «Сократ Персональный» теперь может переводить письма прямо в Outlook 2000. Сохранилась и знаменитая «фишка» предыдущей версии программы — кнопка для мгновенного перевода, встраиваемая в help-файлы Windows. Обе эти функции крайне важны для пользователей, ведь именно при работе с иноязычной электронной почтой и с файлами помощи чаще всего приходится напрягать все свои порой скромные познания в английском языке.

Не забыт и мгновенный перевод текста из буфера обмена. Он осуществляется по двойному щелчку на кнопку «Сократа» в трее. Если вы запускаете перевод какого-нибудь файла непосредственно в окне программы, то язык его текста будет определен автоматически и сразу же, как и у PROMT при включенной опции «Синхронный перевод», в нижнем окне появится перевод.

Не могу сказать, что качество этого самого перевода, которое показал «Сократ», меня сильно порадовало. Но в чем ему было нельзя отказать, так это в скорости. По этому показателю «Сократ» превзошел даже PROMT Express.

Но самое потрясающее было в этой-то хлипчатой пунктуации в переводах. Тире туда, тире сюда. Например, перевода фразы: «According to Apollodorus, the sphinx is a daughter of Echidna and Typhon», «Сократ» предложил такой вариант: «Согласно Apollodorus, сфинкс — дочь Echidna и Typhon». Все остальные переводчики перевели «is» как «является». И, в общем-то, правильно сделали. Но только «Сократ» был безупречно точен. Он не справился в этой фразе лишь с именами собственными. Но и остальные системы перевода этого сделать не смогли. Справился только переводчик с названием гордым Человек: «Согласно Аполлодору, сфинкс — дочь Ехидны и Тифона».

К числу несомненных достоинств транслейтора «Сократ Персональный»



можно отнести и его сравнительно небольшой размер: дистрибутив его демоверсии всего 6 Мбайт. Во время инсталляции вас попросят ввести серийный номер. Если в ответ вы нажмете «Дальше», вам тут же любезно предложат установить демоверсию продукта. Ограничения ее традиционны: через 21 день перевести текст еще будет можно, а вот сохранить результат или распечатать — нет.

При всех своих лингвистических проблемах «Сократ» мне понравился простотой и изяществом. Это мобильный инструмент в руках пользователя.

### «Сократ»

- Разработчик ► Компания «АрсеналЪ»
- Сайт разработчика ► <http://www.ars.ru>
- Условия распространения ► shareware
- Цена ► \$25



## Итоги и перспективы

### Как бы скрестить ежа с ужом

Наш экзамен подтвердил пессимистические прогнозы. Соревнование программ-переводчиков — это не состязание хорошего с лучшим, как, допустим, у MP3-проигрывателей, а плохого с худшим. Связать воедино лексическую точность, и стилистическо-синтаксическую сочетаемость на этом этапе развития техники оказалось не под силу никому. Поэтому неудача «Pragma» относительна, впрочем, как и успех Retrans. Ее вариант перевода тоже пока не идет ни в какое сравнение

с тем текстом, что вышел из-под клавиатуры человека. Лингвист проиграл машине только в одном параметре — временном. Системы машинного перевода тратили на свою бессмыслицу считанные секунды, а человек минуты, но на связный текст. Это и определяет место и функцию систем машинного перевода. Будучи не в силах заменить человека, они должны помогать ему. Быстрый перевод-подстрочник web-страничек, электронных писем и help-файлов —

вот удел электронных транслейторов. И не думаю, что ближайшие лет тридцать что-нибудь изменится. Мне так и хотелось во время тестирования применить метод героев Гоголя. Вот если бы «фразеологический» метод Retrans приложить к словарям PROMT, к этому добавить изящество и удобство «Сократа» и интеграцию в систему «Pragma», да еще предлагать пользователю несколько вариантов перевода на выбор, — тогда, думаю, получилось бы что-нибудь интересное.