

## ProMT 98 - the current state

Svetlana Sokolova  
Project MT Ltd,  
199053, St.-Petersburg,  
Russia  
svetlana@prompt.spb.su  
[www.prompt.ru](http://www.prompt.ru)

ProMT 98 is a new name of the next version of the well-known MT Software - STYLUS - the most popular machine translation software for the Russian language. It can translate from Russian and into Russian for many European languages: English, French, German, and Italian. This version, ProMT 98, includes also non-Russian language pairs: German-French and English-French.

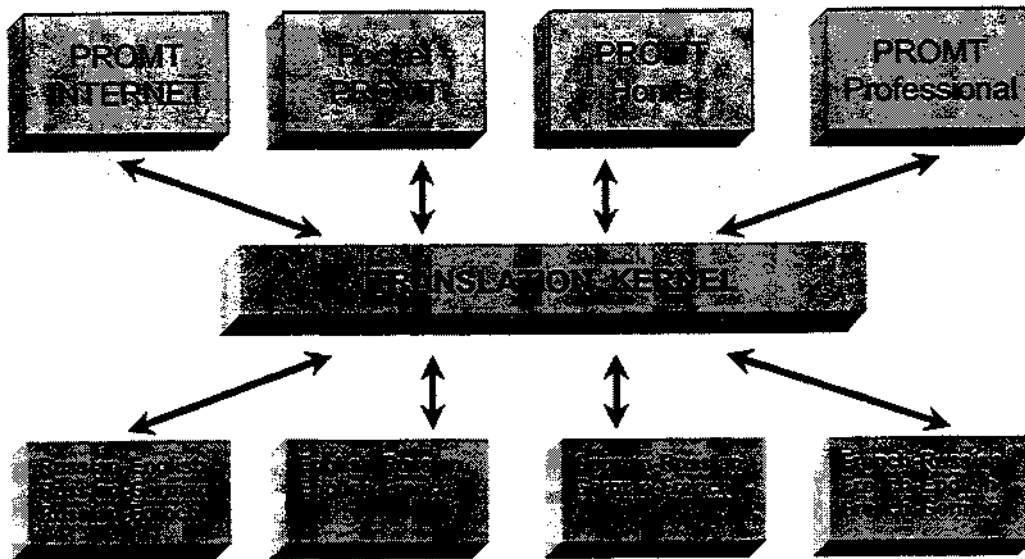


Fig.1 PROMT Family

ProMT 98 is a family of several applications with the same Machine Translation kernel inside. The interface solutions are intended for wide range of end users. There are ProMT Internet Kit, ProMT Home, Pocket ProMT, and ProMT Professional.

ProMT Internet Kit provides two different solutions to help surfing in foreign languages. The first one is realised in WTS application for very fast translation inside an Internet browser like Internet Explorer and Netscape Navigator. On any page during Websurfing a user can make a selection and drag it to the application bar. Translation results will appear immediately in a small moveable box. One more nice possibility inside WTS: translating a search query and then bringing a user to the search server at his discretion.

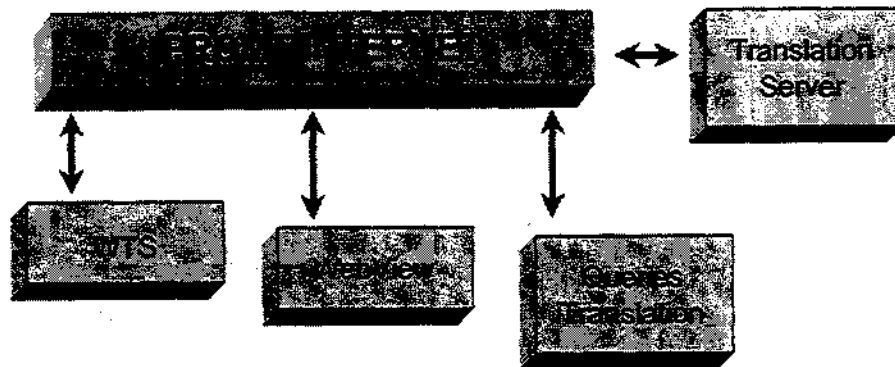


Fig.2 ProMT Internet Kit

Another Internet solution - WebView application. This is a browser-translator. It provides all the functions of Internet Explorer and additionally a possibility to translate the HTML-page with saving of HTML format in the target text. This browser-translator gives the user a unique opportunity for Web surfing in target language. The user does not need to care about source text links because the links will be saved in the target language page in the same places as in the source text. This translator browser solves this extremely difficult task for such different language pairs as German-French or Russian-English. There is also Internet translation server to provide size-limited on-line translation on the Web (for example: [www.translate.ru](http://www.translate.ru))

ProMT Home is a multimedia translator with animated agent that accepts speech commands, pronounces texts in source and target languages and also allows Internet surfing. This application is designed in game style and is a very good assistant for children.

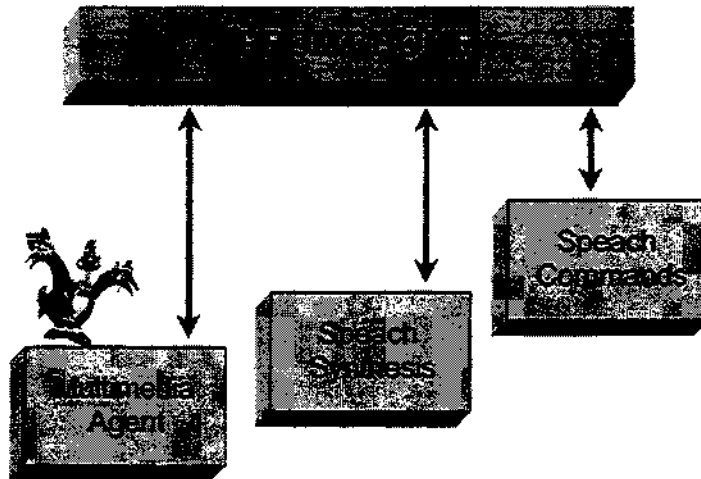


Fig.3 ProMT Home

Pocket ProMT is an application for Windows CE for hand-held computers. It takes about 4 MB RAM for one language pair.

ProMT Professional is intended for skilled users who need customisation for their domain. This solution includes the main application that creates the translation environment to translate texts in different modes and provides possibilities to translate and customise the system directly from MS Word. Translation Environment can open the different format files to translate, to choose a language pair, to manipulate with dictionaries by the dictionary manager, to attach dictionaries needed and to change priority of looking up in these dictionaries. ProMT translation environment includes a unique text editor allowing simultaneous scrolling of source and target texts after translation. It is very helpful for professional using the system. Speed of translation is rather high.

File Translator is an application included in the system to provide professional users with batch translation of file queue by giving various options. For corporate users there is a network solution based on client - server architecture that provides administrative and monitoring tasks for translation process.

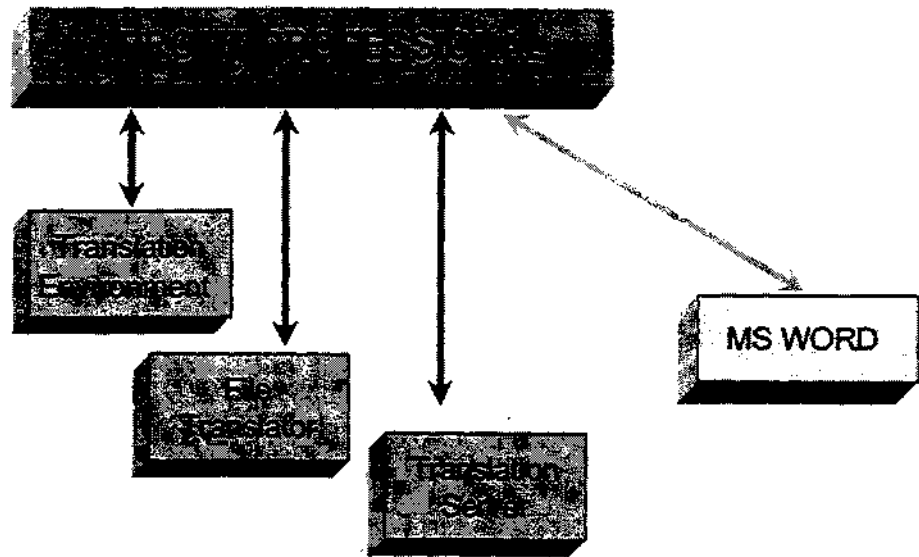


Fig .4 ProMT Professional

The external office applications such as MS Word can be integrated with Prompt Professional. In MS Word, the ProMT installer builds new features to translate, to call Dictionary Editor, to mark words not to be translated and so on. Using the usual office environment the users can get all these features in the same way as they could do in ProMT Translation Environment

All the applications use the same translation kernel that includes translation engines for the above mentioned language pairs. Each translation engine combines Translator to provide translation process, Dictionary Editor to provide editing the user dictionaries, and Linguistic Database.

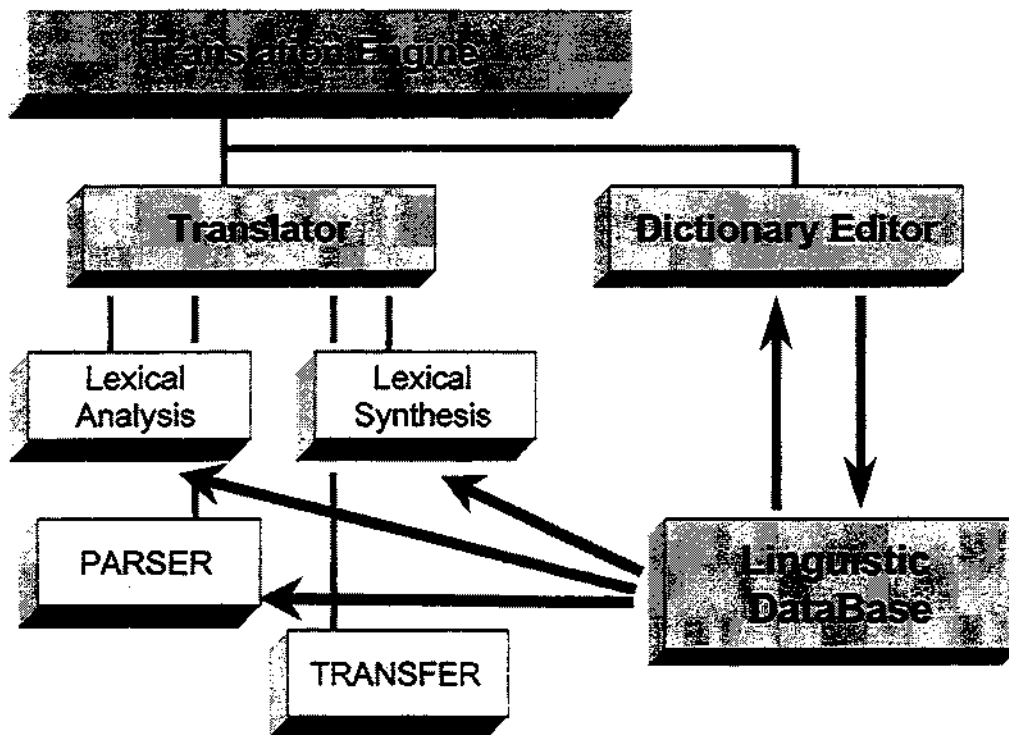


Fig 6. Structure of ProMT Translation Engine

Linguistic Database consists of Flexion tables for source and target languages, Source Morphology tree, and Augmented Transition networks adapted to translation and Dictionaries. A source morphology tree combines a set of source endings in a tree to fast find a word dividing into two parts: a stem and an ending. This structure is rather convenient for morphological analysis. Analysis starts with the final letter of the word and moves down the tree as far as possible. When it is impossible to continue moving, potential dividing is ready.

Source morphology and Target morphology of the language (for example Russian source morphology in Russian-English and Russian target morphology in English-Russian) can be different because of difference of analysis and synthesis tasks. However, source morphology of the language (for example French source morphology in French-English or in French-Russian) is the same, as well as Target Morphology (for example French target morphology in German-French or Russian-French).

Another part of Linguistic Database is the set of transition networks. We adopted this formalism to describe translation processes for Noun and Verb groups. Parsing by ATN moves along a source chain of words from one node to another according to the marks on the arcs and generates a target chain along with meanings of some characteristics inherited from the upper level of

analysis. The most part of ATN sets is common for each source language in different language pairs.

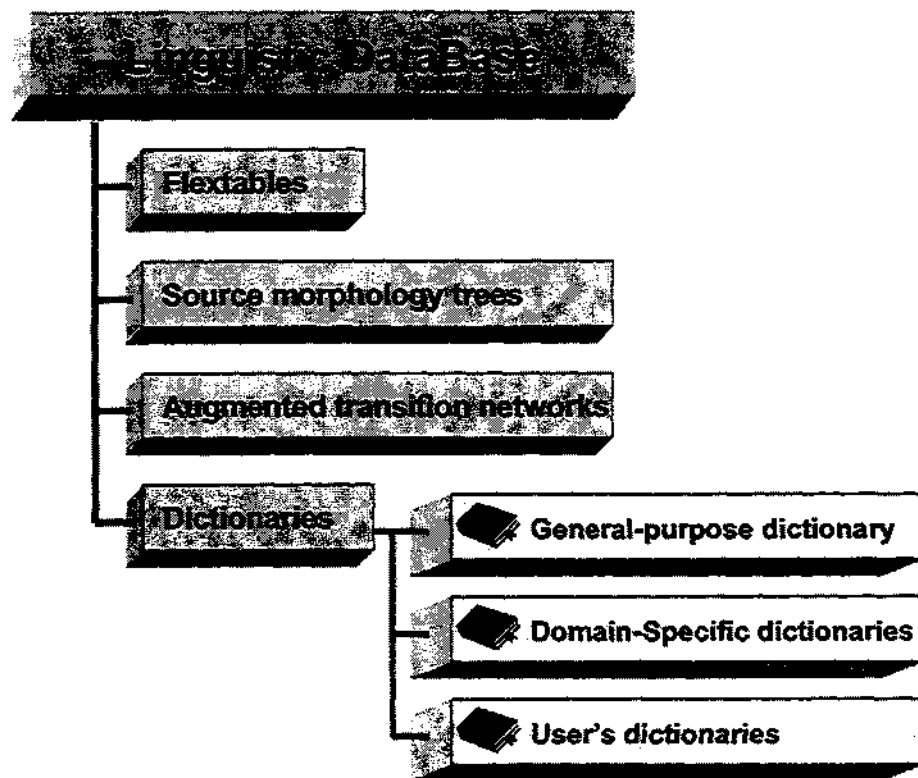


Fig 6. Structure of Linguistic Database

The largest part of Linguistic Database is Dictionaries. System Dictionaries have a three level structure: General-purpose dictionary, domain-specific dictionaries and User dictionaries created by the user with the identical logical and internal structure. This three level structure gives a possibility to customise a translation process to the different texts.

The general-purpose dictionary for one language pair, for example German-French, contains, as a rule, about 100000 entries and appears as the basic dictionary of the system during translation. It contains entries for the most frequent words and phrases of the source language. These entries have a functionally sophisticated collection of semantic and syntactic tags that are used for the translation algorithms.

The domain-specific dictionary contains terms specific to the corresponding domain and also general-purpose words, if they have some specific functions in the domain. For example, to translate software documentation in the correct way the words such as "program", "application" need a specific information. The volume of a domain-specific dictionary varies from 10000 to 40000 stems.

The user dictionary is created by a user himself. There is no limitation in the system on the number of user dictionaries used in translation process. As a

rule, if the user translates texts in different domains, he distributes his own terminology into different user dictionaries and then uses the suitable user dictionaries together with the general-purpose dictionary and domain-specific dictionaries if necessary.

Using Dictionary Editor a user can add new words and phrases, change meanings for the entries existing in the general-purpose dictionary and for the entries existing in a domain-specific dictionary. All these changes will be saved in the user dictionaries, because domain-specific dictionaries as well as the general-purpose dictionary are closed for a user's adaptation. These can contain a user-hidden information, and, to prevent basic linguistic algorithms that employ this hidden information, the system saves all user-customised information in the user dictionaries.

Our translation system first appeared on the Russian market seven years ago. Now we changed the name of the system and together with our French partner Softissimo we started to distribute our new language pairs without Russian under the trade name ProMT/ REVERSO. There are more than 50000 users of the system world-wide and feed back from the users gives us the positive feeling that MT is a real tool.