CafeTran Handbook

Contents:

1. Fast start
2. Documents and Resources
3. User interface
4. Projects
5. Memories
6. Rendezvous Memory Server
7. Databases
8. Automatic translation, fuzzy and subsegment matching
9. Dictionaries and Glossaries
10. Internet resources
11. Image resources
12. Synchronization
13. Segmentation workflow
14. Image-Paper workflow
15. Review and translation of projects prepared in other translation tools
16. Clipboard workflow
17. Alignment workflow
18. Connection to OpenOffice (LibreOffice) suite
19. Spell checking, thesaurus and auto-correction
20. Quality Assurance (QA)
21. Frequent terms
22. Segmentation
23. Placeables
24. Useful functions and options
25. Formating tags
26. Splitting and joining segments
27. Conversions
28. Statistics and billing

29. Translation management

30. Keyboard shortcuts

Fast start

This introduction presents a few simple steps to start working with the program.

Creating a new translation project

Before the start you need to prepare a sample document in your source language for the trial translation. Preferably, a short document in the format of an office suite such as Ms Office (Ms Word) or OpenOffice is chosen for the test. In case of Ms Word, CafeTran makes it possible to translate the documents saved in the latest format with the *.docx extension. This is the default format starting with the 2007 Ms Office version. If you use older versions of Ms Office where the default extension is .doc, make sure you save your documents with the .xml extension which the older versions enable. So let us assume that our sample document called Test.docx is a Ms Word document in Spanish which has to be translated into English.

Launch the program by clicking at the Start.jar icon. After a while the Project window will pop up where you can choose your sample Test.docx for translation. Click the "Document" button. The window with files and folders will appear where you need to locate and choose the Test.docx. In the lower part of this window, choose the format of the document - Ms Word (.*docx/xml) from the drop-down list (Files of type). Next, accept the selection with the OK button and you will be back in "Project File" window.

Now set the source language code - "es" for Spanish in the "Source language" field and the target language code - "en" for English in the "Target language" field.

We will also create a new memory file for this project called ProjectTM.tmx which will be located in the same folder as your source document. Please, check in the field "New project memory".

If the source language of the document does not contain spaces between words (e.g. Japanese) you need to change the Matching Type option from "Fuzzy & Subsegment" to "Detailed Matching" in the "Memory settings" tab.

To keep things simple at the start, skip other memories that might be used leaving the fields "Load memory file" and "Load database memory" unchecked. You will check in any of these fields when loading existing segments either from the TMX memory file or the database.

Now we are ready to accept the project settings with the OK button and move on to the main translation panel. In the lower part of the panel you should see three tabs: a document tab (Test_en.docx), a project memory tab (ProjectTM) and machine translation tab of the Google Translate and Bing Translator online services. In the upper part there is an empty window for project segments and a split window for translation of the current segment. The top part of the split window will display source segments which are translated into target language in the bottom part of this window.

Before the start of the translation we may also create a memory for the terms which we will find important during the translation process. To create a new terms memory, choose from the menu Memory | New memory. The "New memory" panel will pop up where you need to check in the "For

terms" field and leave the "For segments" field unchecked. After accepting our choice with the OK button the new memory tab will appear in the translation panel. During the translation of segments we will be able to select and add any important terms or frequent phrases to this memory.

The tabs in the main translation panel can be joined together or docked to other windows in the panel. Joining or docking the tabs can be done by clicking at a tab with the right button of the mouse and choosing the "Synchronize with.." (the tabs joining command) or "Dock window at view 1, 2, or 3" (the tab docking command). After docking the tab to another view you will notice a new button on this view to let you undock the tab at any time.

a new button on this view to let you undock the tab at any time.

Start of the translation

1. Click the "Start" button on the toolbar of the target segment window.

2. The "Frequent terms" tab will open displaying any frequent terms from the current segment. These phrases can be added to the "Terms memory" during the translation process.

3. The first segment for translation will appear n the upper source segment window. The lower window is for the translation of the current segment. This is a target segment window.

Translation of segments

Using the buttons on the target window toolbar or keyboard shortcuts proceed as follows:

1. Based on the results provided by online machine translation services and any matches in the memory window, type the translation in the target segment window. You may use keyboard shortcuts (see menu Translation) or mouse selection to transfer any good phrases or matches to the target segment window.

2. Press the F2 key to see the matches from all the memories and glossaries being used. The F1 key displays the panel of automatic translation which is based on these matches. In both the panels F1 and F2, you can see numbers next to the results. Pressing the number in the keyboard will transfer the corresponding result to the target segment window. Obviously, the matches from the memory will start to show up only if your memory contains a number of segments and/or terms.

3. The powerful auto-completion function remembers not only longer words as you type but also exact matches in the memory or glossaries, and results from online machine translation Google and Bing.

4. When the translation of the current segment is ready, click the button "Add current segment to memory" on the target window toolbar to add it to the memory and take the next segment. If you do not wish to add the current segment to the memory, click the "Next segment" button on the target window toolbar.

5. If you want to join two segments together click the "Join segments" button on the target window toolbar. If you see any numbered tags in the source window, they need to be transferred to the target window using keyboard shortcuts F3 or Alt-9. The tags are also easily transferred after typing the tag number in the target window and pressing the Esc key.

6. We can go back to the previous segment too. Just click the "Previous segment" button with the left arrow on the target segment toolbar.

7. If the "For terms" memory is created it is very easy to add selected phrases to it. You need to select with a mouse a term in the source segment window, and then do the same with the term's translation in the target segment window. Next, click the "Add term to memory"

   button on the target window toolbar. Finally, the "New term" window appears where you click the the "OK" button to add the term to the "For terms" memory. Similarly, you can add any frequent words or phrases which appear in the "Frequent terms" tab. You need to click on the button next to the given phrase.

8. Continue the translation as described above until you reach the end of the document.

9. You can preview the partially translated document in the work flow. Click on the Translation | Preview document" menu. The file with your translation called Preview.docx will be created in the same folder which contains the project.

Save and continue the translation

It is recommended to save the project and working memories before the exit of the program. Click

the "Save" button in the tabbed window toolbar to save the project and memories. To exit the program, click Project | Exit (shortcut - CTRL+Q) in the Project menu. After the restart of the program, click the button which displays the name of the last project to open the project. Alternatively, choose from the "Recent projects" menu a project that you want to continue with. The project's memory (ProjectTM), if created at the start of the project, opens automatically. Other memories can be opened via the "Memory | Open Memory..." menu. Next, click the "Start" button in the target window toolbar to continue the translation from the last working segment.

Export of the translated segments

1. Choose "Export current document" from the Project menu or click the "End" button on the target window toolbar. The target document is created in the same folder that contains the source document.

2. Open the target language document in Ms Word and check the result of the translation.

Saving memories in the database

1. Create a new database table by the menu command "Database | New table...".

2. Select the memory tab from the tabbed window.

3. Choose "Database | Store memory in table" from the menu.

4. In the table selection window choose a table which will hold segments and/or terms from the memory.

5. Press the "OK" button.

6. Likewise, you can store other memories in a new or existing database table.

Memories can also be saved in TMX files if you want to send them to your clients or open in another translation program, for example. To save a memory in the TMX file, choose from the menu Memory | Save memory. Clicking on the "Save" button in the main toolbar also saves any working memories in the TMX files.

Documents and Resources

CafeTran handles the following formats:

1. Text.
2. HTML.
3. XML.
4. OpenOffice.
5. MS Office (2003 or later).
6. Adobe Indesign
7. Adobe Framemaker
8. AutoCAD
9. DTP Tagged Text
10. iWork.
11. Abiword.
12. KWord.
13. Image (jpg, gif, png).
14. XLIFF and TTX files.
15. TMX and TXT memories.
16. Uncleaned RTF files saved as .docx (Ms Word) or .odt (OpenOffice).

One of the main design goals was to use CafeTran for translation of the documents which are not available in the editable form such as image, scanned or just paper documents lying in front of you on the desk.

CafeTran offers a very flexible interface for the creation, access and search of various types of

resources such as:

1. Dictionaries.
2. Glossaries, which are a kind of specialized dictionaries (a glossary of state institutions, for example).
3. Internet (Google search engine or KudoZ glossaries, for example).
4. Images (various image references such as scanned dictionaries and other documents saved in an image format).
5. Translation memories.

As you will see later, each kind of a resource can be integrated in the translation workflow.

User interface

After the program has opened you can see the following user interface elements:

1. The menubar
2. The resources and document window with its toolbar and search field
3. The project segments window with its toolbar
4. The current source segment window with its toolbar
5. The current target segment window with its toolbar

The layout, colors and fonts of the windows can be changed to suit your preferences.

The window layout can be changed in the View menu.

- Horizontal/Vertical layout of the segment windows
- Reverse or flip segment windows
- Merged segment windows for image or paper translation
- Detach/Attach main window
- Window layout 1
- Window layout 2
- Window layout 3

Colors and the font are set in the Edit menu.

- Background color
- Font color
- Selection color
- Segment color
- Subsegment match color
- Tag
- Font

To change a font or color of a window, click at this window, and then choose from the Edit menu respectively.

There are also a few predefined color themes such as bright, dark and texture-based theme. You can select the best colors depending on the time of day, lighting etc. Choose a color theme from menu View | Themes.

The background grid can be switched on/off in the menu Edit | Colors | Background grid.

Furthermore, you can change the overall Look and Feel of the program selecting it from the View | Look and Feel menu.

CafeTran also looks great with external Substance skins available at https://substance.dev.java.net/. You need to download two files - substance.jar and trident.jar and copy them to the lib folder in CafeTran's installation directory. Next, select your favorite skin in the View | Look and Feel | Substance menu.

Projects

Projects

CafeTran's project format follows the XLIFF specification, a widely accepted standard in the localization industry.

Creating a new project

1. Choose Project | New project from the menu.

Each project consists of one or more project documents for translation.

Adding a document to the project

1. Choose Project | Add document
2. In the Project info panel click the Document button and choose your source document and its type.
3. Fill in the source and target language abbreviated names in the language fields of the project dialog. You can switch the languages just by clicking at one of the two labels.
4. Press OK button and the source document will open ready for translation.

Since one project can comprise a few documents, you can put them all into one folder and add to the project. In the Project File dialog press the Document button and go to the folder with your files. When you set the folder path in the Document field, CafeTran will add all files in the folder to the project. Next, by choosing from Project | Documents menu you can open each file for translation.

Saving the project

1. Choose Project | Save or Save project as.
2. In the dialog window give a name to you project and click the Save button.
3. Project is also saved automatically after adding a few segments to it.

Opening the project

1. Choose Project | Open project... from the menu.
2. In the Project chooser window select a project and press the button Open.

Memories

Translation memory is a fast RAM computer memory for holding your current and previous translations in units consisting of source and target segments. The segment is usually a sentence or a paragraph. When the identical source language segment appears later during translation, its target language equivalent is automatically retrieved from the memory and presented to the translator as the exact match.

If such a match is not found in the memory, special algorithms try to find the closest match possible, which is often called a fuzzy match. Accuracy of fuzzy match search functions depends on the size of the translation memory. The bigger the memory is the more precise match you will get.

CafeTran stores memory segments in database or can save them to a file in the TMX format that follows the TMX specification, a widely accepted standard in the localization.

Memories in CafeTran can be used both for segments and terminology. Segment memories write in new segments in the workflow. Terms memories will write only phrases or terms which the translator selects from the current segment.

Both segment and terms memories are used by Autocompletion and Autotranslation functions.

Creating a new memory

1. Choose Memory | New memory from the menu

1. Choose Memory | New memory from the menu.
2. In the Memory dialog type in the language codes of your langauge pair.
3. Check in "Memory For Segments" field if it is the segments memory or/and "Memory For Terms" field if you create a terminology base.
4. Press OK.

Saving the memory

1. Make sure your memory tab is selected.
2. Choose Memory | Save memory or Save memory as.
3. In the Save dialog window give a name to your memory and click the Save button.

Opening the memory

1. Choose Memory | Open.
2. In the Memory dialog type in the language codes of your langauge pair.
3. Check in "Memory For Segments" box if it is a segment memory or/and "Memory For Terms" box if you load a terminology base.
4. Checking in "Read only" box will not let you write any new entries to this memory.
5. Press OK.

The Read Only mode reduces loading size of memories taking much less computer RAM memory. It is recommended to load huge TMX reference memories in this mode.

The Memory Type settings can be changed later in the workflow by clicking with a right mouse button at the memory window and checking in/out the appropriate field from the menu.

You can open as many memories as you like and hook them all into the workflow.

Project memories

CafeTran allows fast creation and access to the memory which is meant for project segments. As you create a new project, check in the New Project Memory field in the Memory Info tab of the Project File window. This memory opens automatically together with the project.

The Memory Info tab makes it possible to open a memory or memories which are set in the Memory field. Check in this field, click on the Memory button and choose a memory or a folder with memories to be loaded.

Memory integration with the workflow

Memories have three type of integration with the workflow, which you can set in the Memory opening dialog.

1. Automatic: up to 20000 - 30000 units.
2. Pretranslation: for large memories with more than 30000 units.
3. Manual: for translation of only chosen segments.

Generally, if you notice some look-up slow down due to the memory size you should switch to the Pretranslation mode where the response to your current segment query is instant.

For example, a typical workflow scenario would be to create a new memory for the project segments with Automatic integration and a large reference memory set in the Pretranslation mode for fast queries and Read Only access.

Matching type

There are some settings which have an important impact on the matching accuracy and the number of hits.

1. **Fuzzy and subsegment**

   This is the default setting appropriate for Latin based and other languages where there is a space separator between words. CafeTran analyzes source segments on a word basis and performs statistical analysis of smaller chunks called subsegments to determine their

translation.

2. **Fuzzy**

   To speed up the matching process you may wish to switch off the subsegment matching. The program shows only fuzzy matching results.

3. **Detailed matching**

   This option is suitable for languages without a word boundary (e.g. Japanese and Chinese). When this option is selected, the program analyzes source segments on a character basis. Please, select this setting if your source language does not have a word space boundary. Remember to switch back to the default Fuzzy & Subsegment matching when you translate from a Latin based language, for example.

4. **Prefix matching**

   When this option is selected, CafeTran will analyze the beginnings of words (here called prefixes) and discard any endings responsible for inflection of words. It is an option which increases significantly the number of hits for highly inflected languages. The length of prefixes is set by a percentage number. The bigger the percent number the longer the prefix of words which the program will analyze. The minimal prefix length option (menu Edit | Options | Memory | Minimal prefix length) lets you set the minimal allowed length of prefixes. The length can also be fixed, when the "fixed" option selected, instead of a set percentage length. It means that all the words will have the minimal prefix length, no matter their actual length.

5. **Custom prefixes**

   If the inflection of a word is too high for automatic prefix matching you can enter your terms to the memory determining the prefix of a word manually. This is done by inserting the pipe character | at the end of a prefix in a word. For example, the Polish phrase "piękny dzień" (a beautiful day) has a highly inflected word "dzień" occuring in a number of various cases (dnia, dni, dniom). If you insert the pipe characters at the following positions - "pięk|ny d|zień", CafeTran will also recognize other forms of the phrase (pięknego dnia, pięknych dni etc.). Note that inserting the pipe character at the first word in the phrase - "pięk|ny" is optional since its inflection is quite regular and CafeTran should recognize its prefix automatically.

Adding segments to the memory

CafeTran adds new segments to the segment memory automatically in the workflow as you accept currently translated segment with the "Workflow | Add segment to memory" menu command.

Adding terms to the memory

1. Select a term or double click a word in the source or target language window and press "Add term to memory" button in the target window toolbar.
2. In the pop-up term window type or correct a source language term in the Source text area and its translation in the Target text area.
3. If you hold Ctrl key while pressing "Add term to memory" button, the selected entries are added without popping up the term window.
4. Press OK button.

Filtering memories

1. Choose Memory | Open memory in the menu.
2. Choose a memory and press Filter button in the Memory dialog window.
3. Keep newer duplicates (default) filter - overwrites older identical source segments.
4. Keep older duplicates filter - ignores new identical source segments which are already present in the memory.
5. Keep all duplicates filter - both newer and older identical source segments are kept in the memory.

6. Properties and notes filter - loads only the translation units that have a particular property's name and value or a note.
7. Regular expression segment filter - loads only the translation units which contain a particular word, phrase or regular expression.
8. Languages filter - extracts only one language pair if your memory has units in more than two languages.
9. With Properties, Regular expressions or Languages filter set, remember to save the filtered memory under some new name. Otherwise, you will overwrite the original memory file while saving!

Merging memories

1. Move all the memories to merge in one folder on your system.
2. Choose Memory | Open memory in the menu.
3. Select the folder in the File Chooser and press Open. In the memory settings window that follows you may set your filtering preferences for all the memories.
4. After loading is complete choose Memory | Save in the menubar to store units from all the memories in one memory file.

Editing the translation memory

1. Select memory window.
2. From the Memory menu choose Show contents.
3. Press Forward or Back button in the menubar to browse through the translation units.
4. Click on a translation unit that you want to edit.
5. In the translation unit window you can edit content, notes and attributes.
6. Press Memory | Save or Save as to store any changes that you made to the chosen translation memory.

Much more convenient edition of the memory can be done through the Project interface. Go to Project | Open Project menu and choose your TMX memory file. It will open for edition just like any regular Project file and its segments are accessed by choosing the Next or Previous segment button in the target window toolbar.

Searching the translation memory

1. Check in Find | Memory Source or Target in the main menu.
2. Make sure your memory window is selected.
3. In the search field, type a word, phrase or regular expression to find its occurrences in memory source or target segments.
4. Click the "Search" button and the list of translation units containing your search criteria will appear.
5. Hover the mouse over the links to see the content of the unit segment or click to edit them.
6. If you hold the Ctrl button while pressing the "Search" button, you can replace the search results with the content of the target language window or its selection.

Adding memories to the Library menu

1. Go to Library | New resource.
2. Select Memory.
3. Fill in the required fields in the following way:
    - Name:
      Give the name of the new memory (if you are accessing an existing memory leave the field empty).

    - Path:
      By clicking on the button, choose a directory of the new memory or the full path (along with the file) to the existing one.

    - Source language:
      Give the short name (for example: pl) of the source language in the memory.

- Target language:
  Give the short name (for example: en) of the target language in the memory.

- Author:
  Give the name of the memory creator.

- Encoding:
  This field allows to specify an encoding for your memory. Unicode (UTF-8) is the default and generally it should not be changed unless your existing memory was created in a different encoding.

4. Press OK button and save the memory configuration file under some name in the Memories directory.

   You can also make subdirectories of the parent directories and save the file there organizing your memories in a tree-like structure.

Opening memories from the Library

1. Select Library | Memories | Your memory.
2. In the memory settings window you can switch the languages direction.
3. Press OK button to load the memory.

Editing the Library configuration file

1. Select your memory tab.
2. Choose Library | Edit resource info from the menu.

Memory Import

You can also move segments from the project and entries from glossaries to the memory. Open an existing project or glossary, select a memory tab and click menu Memory | Conversions | Import from project or glossary. Next, save the memory with new segments.

Joining Terms Memory

A terminology base in the TMX format can be joined with the main memory so that all matches are shown in the common window. Go to Memory | Join terms memory and choose your base to be joined. Make sure that language codes of the two memories match.

Rendezvous Memory Server

CafeTran can work in a server mode, which offers connection to its memories from other computers over network. The typical use of the server mode can be a team translation where two or more translators share their memories while working on a project. Of course, the computer which works as a server and wants be available to other computers over the Internet must have a public IP address. Another example of the server mode is connection and running the program on a server computer which has huge RAM memory and processing power for fast queries of very big resources.

Starting the Rendezvous memory server

1. Go to Options | Memory | Rendezvous server port and check in the box. Default server port is 4242 and can be changed if your computer has this port reserved.

2. Check in the Read Only box if you do not want to give write access to your memories.

3. Press OK to accept the settings and CafeTran will start waiting for search queries from other users.

4. When a new user will try to connect to your memories a request window pops up asking for acceptance.

5. The list of all remote guests to your server is visible in the Memory | Remote guests menu. You can terminate the connection with a guest by clicking at the list item.

6. Check out the box in Options | Memory | Rendezvous server port to stop the server.

Connection to the Rendezvous memory server

1. Click the menu Memory | Connect to remote memory...

2. When the Connection window pops up, type the Host name or IP address of the server in the Host field, a server port in the Port field, if different from default, and press OK.

3. To change the Remote memory integration and type of access, select accordingly in the Connection window.

4. Now, if everything goes well, you will be able to access memories on other computers as your own.

Databases

CafeTran offers database functionality for access, storage and searching of translation segments as well as terminology. It is a perfect solution to organize your language data based on various categories.

The default database engine is an efficient H2 database developed by Thomas Mueller. However, it is also possible to connect to other preferred databases.

Database tables

CafeTran stores any linguistic data in multilingual database tables. You can define various table columns such as Field, Notes, Reference and Languages in Edit | Options | Database options. If you leave any of the fields empty, the column for the field will not be created. Make sure that Source and Target language column names match the codes of the languages in the Project and Memory.

Creating a table

1. Go to the menu Database | New table.
2. Type the name of the table and press OK. The name of the table must be unique and it cannot contain interpunction characters.

Opening a table

1. Go to Database | Tables menu and choose a table.
2. A tabbed window with the table will open.

You can browse through table contents by choosing Back and Forward buttons in the toolbar. Searching the table is done after pressing the Search button in the toolbar. The default search setting looks up the source language column. You can change this by selecting the Find | Database Target check box.

Adding an entry

1. Choose the Database | Add entry command or "Add term to resource" command in the Workflow menu. The latter method will show the window to type in source, target text and other properties such as a note and a field.
2. Save and refresh the table after finishing adding entries.

Editing an entry

1. Select the row with the entry and type in the text areas below table rows.
2. Save and refresh the table after finishing.

Deleting an entry

1. Select the row with the entry and press Database | Delete entry in the menu.

Deleting a table

1. Go to Database | Show table info and press the Delete button.

Storing memory in database

Both segments and terms memories can be stored in database tables. It is usually done at the end of the project after reviewing the project memory segments and completing any terms lists.

1. Select a memory tab.
2. Go to Database | Store memory in table menu.
3. Choose a database table for segments or terms storage.
4. Press OK.

Loading database to memory

Segments or terms that are stored in the database can be loaded into memory for automatic translation of documents or sending to TMX files. To load a database table to memory:

1. Go to Database | Load database to memory.
2. In the Database Memory window choose a table to load.
3. Set any filtering options if necessary. For example, if you want to filter segments based on the subject, go to the Memory Filter window and type in a name of the subject column in the Name field and a subject in the Value field of the Segment Properties section.
4. Press OK.

If you are loading table contents to Memory only for autotranslation make sure to check in the Read Only box. This ensures faster matching and requires much less computer RAM memory.

Use the above steps to send a database table to a TMX file. After loading finishes go to Memory | Save memory to save in the TMX file format. Of course, the Read only box should be checked out to store data in a TMX file.

Working with references

The advanced function of a database enables setting and viewing a reference to a given unit of linguistic data. A typical reference example can be an image, a document file or a web address. Creation of a database table with a reference field is simple:

1. Go to Edit | Option | Database and type a name of the reference column in the field called "Reference column name". This field is empty by default because it is up to the user what reference type she works with. The name can be general such as Reference, Picture or Address. You can also give the name in your language.
2. Go to the menu Database | New table.
3. Type the name of the table and press OK. The name of the table must be unique and it cannot contain interpunction characters.

The created table will contain one additional field holding a reference address and two buttons for setting and viewing the reference respectively. The "Add entry to Database" window will also contain an extra field to let you set a reference for the new entry easily.

Database backup

Like with all important data it is recommended to perform regular backups of the database files. By default, the files are located in the resources/databases folder of CafeTran's installation. You should save periodically the databases folder with its contents to your prefered backup medium.

External databases

CafeTran has been tested to work with other popular databases such as H2, MySQL, Oracle 10g, HSQLDB 2.0 and Derby (Java DB). This gives you the unique opportunity to connect to terminology bases which are on some other machines, for instance. If you are interested in connecting your database which is not listed below, please contact the users support for more information.

When you have installed and created your database by using its adminstration tools, follow these steps to connect it to CafeTran:

1. Install a java driver of your database.

   This is a special .jar file which should be copied to CafeTran's lib folder. The file comes along with your database or can be downloaded separately. The following .jar files are java drivers for a specific database:

   - MySQL - mysql-connector-java-5.1.13.jar (rename the file to mysql.jar)
   - HSQLBD - hsqldb.jar
   - Oracle 10g - ojdbc14.jar (rename the file to ojdbc.jar)

   As you can see some of the .jar driver files should be renamed before copying them to CafeTran's lib folder. Renaming lets you install any future versions of the drivers without updating the program.

2. Connect CafeTran to the database of your choice.

   - Go to Edit | Options | Database.
   - Choose New database... from the drop down list.
   - Fill in the following fields:

     Driver file: leave this field empty.

     Driver class:

     - MySQL database - com.mysql.jdbc.Driver
     - HSQLBD database -org.hsqldb.jdbcDriver
     - Oracle 10g database - oracle.jdbc.driver.OracleDriver

     User name: type your user name here.

     Password: type your password here.

     Connection URL:

     - MySQL database - jdbc:mysql://localhost:3306/your_database_name?useUnicode=yes&characterEncoding=UTF-8
     - HSQLBD database - jdbc:hsqldb:file:./your_database_name;shutdown=true
     - Oracle 10g database - jdbc:oracle:thin:@localhost:1521:xe

       This field depends on your database settings. The above examples are default values.

   - Press OK and save the configuration file under some name.

- Now select the file from the drop down list and press OK closing the Options window. The database tables, if any present, should be visible in the Database | Tables menu, and you are ready to work with the connected database.

Automatic translation, fuzzy and subsegment matching

These functions make up the core of CafeTran's translation engine where the current segment is matched against the segment and terms memories.

Automatic translation (Automatic assemby function)

The function is highly dependent on the size of the memory. The bigger the memory is the more accurate results you can get. Glossaries can also be included by the function's algoritms. First, they need to be converted to the Terms memory as shown in the Conversions section.

Fuzzy matches

The program gives three best fuzzy matches for the current segment. The default minimum threshold is 33% of accuracy and can be changed by setting the Fuzzy Match Threshold field in the Edit | Options | Memory dialog.

Fuzzy AutoCorrector

Fuzzy matches can be corrected automatically by applying exact subsegment matches to their missing (incorrect) parts. For example, you have in the memory the following English-Spanish translation unit:

```
I like the red car = Me gusta el coche rojo
```

Later, when you translate the similar English sentence such as:

```
I like the blue car (Me gusta el coche azul)
```

CafeTran retrieves from the memory the fuzzy match: "Me gusta el coche rojo" and with the Fuzzy AutoCorrector function on, the program can substitue the missing word "azul" for the word "rojo" provided that:

- it finds the translation for "blue" (blue = azul) in the memories, and also if
- it finds the translation for "red" (red = rojo) in the memories.

The Fuzzy AutoCorrector can be turned off by unchecking the "Fuzzy AutoCorrector" box in the Memory menu.

Exact subsegment matches

They are 100% matches for the subsegments of the current segment.

Subsegment matches

They are the approximate matches for the subsegments of the current segment.

Subsegment matches are shown with the examples in the full context. Move the mouse over the numbers of the hits to display the subsegment in the context.

Virtual matches

These are the best subsegment matches of all the subsegment matches selected by the program . The Subsegment to Virtual threshold can be set in the Edit | Options | Memory dialog.

All the results can be transfered to the target segment window either by clicking on the match header or mouse selection. The F2 keyboard shortcut is a convenient way of listing and transfering matches from all memories. The transfer of the best matches from all memories and autotranslation can be done via the Translation menu and its shortcuts, e.g. F1, too.

autotranslation can be done via the Translation menu and its shortcuts, e.g. F1, too.

Dictionaries and Glossaries

CafeTran offers a flexible interface to access and update your dictionaries in the workflow.

Glossaries are more specialized dictionaries such as terminology lists. They enable an automatic and fast look-up for any specific terminology that should be used in the translation.

This distinction between glossaries and dictionaries in CafeTran only affects the resource integration in the workflow. The lookup in glossaries is done "on the fly" each time you take a new segment, whereas the dictionary check happens when you click on the Search button in the main toolbar.

Creation and access to dictionaries

1. Select Library | New resource.
2. Select Dictionary or Glossary.
3. Fill in the fields in the following way:

    - Name:
      Give the name of the new dictionary (if you access an existing dictionary leave the field empty).

    - Path:
      After clicking on the button, give the directory of a new dictionary or the full path (along with the file name) to the existing one.

    - Entries separator:
      You can specify the separator between consecutive entries in the dictionary. If you leave the field empty it will use your operating system line separator.

    - Entry word separator:
      Here you may put in a separator between a word and its definition or translation. The default is "=".

    - Text delimiter:
      Some existing dictionaries have entries which are between some text delimiter such as inverted commas, for example.

    - Encoding:
      This field allows to specify the encoding for your language. Unicode is the default and generally it should not be changed unless your existing dictionary was created with a different encoding.

    - Description:
      If you wish, write down a short description. It will appear at the start of your dictionary.

4. Press OK button and save the resource configuration file under some name in the Dictionaries or Glossaries directory. You can also make subdirectories of the parent directories and save the file there organizing your resources in a tree-like structure.

Opening dictionaries

1. Select Library | Dictionaries or Glossaries | Your dictionary.

   In the tabbed view a dictionary tab will appear.

Adding entries to dictionaries

1. Select your dictionary window.
2. Press "Add term to glossary" button in the target window toolbar.
3. In the pop-up term window type a new entry in the Source area and its definition or translation in the Target area.
4. Press OK button.

Extracting terminology

CafeTran offers fast selection of words and expressions from both source and target segments and their quick addition to the chosen dictionary.

1. Choose your dictionary window.
2. Select a phrase with a mouse or double click a word in the source or target window and press "Add term to glossary" button in the target window toolbar.
3. Press OK button to accept the word and its translation in the dictionary update window that pops up.
4. If you hold Ctrl key while pressing "Add term to glossary" button, the selected entries are added automatically without the term window popping up.

Saving dictionary

After a number of new entries has been added, you have to save the dictionary to make the changes permanent. Otherwise, they will not have any effect when you exit the program.

1. Make sure your dictionary window is selected.
2. Press Save button in the toolbar.

Searching dictionaries

1. Type a word in a search field or select it with a mouse from a source window.
2. Press the Enter key or the Search button in the toolbar. If the entry exists it will show up in the dictionary window.

Searching options

CafeTran provides a few convenient search options shown below based on the example of the English word "cat".

1. cat - search for the word cat in both entries and their translations or definitions.
2. cat= - search for the word cat only in the entries.
3. =cat - search for the word cat only in the translations or definitions.
4. cat* - search for the words beginning with the 'cat' sequence (cataract).
5. *cat - search for the words ending with the 'cat' sequence (tomcat).
6. *cat* - search for the words including the 'cat' sequence (delicate).

Furthermore, you can make a regular expression search by enclosing the regular expression with " symbol.

Editing the Library configuration file

1. From the tabbed view select the dictionary window.
2. Choose Library | Edit resource info from the menu.

Internet resources

CafeTran lets you access and query your favorite web resources.

Accessing Internet resources

KudoZ terminology bases are a good example of the web resource that translators use and share. It requires a few steps to access it.

1. Select Library | New resource.
2. Select Internet.
3. Fill in the fields in the following way:

    - Address:

      Leave it empty if you intend to set Address start and end fields. This field is useful if you simply want to access a particular web page without querying it.

    - Address start and end:

      A lot of Internet sites provide services such as searching terminology bases. They present the form to fill in and the button to press. Next, your browser generates the URL query that is sent to the resource provider. Based on this query you get the response.

      Usually, the URL query that is generated and sent consists of three main parts: prefix, searched word and suffix. Sometimes, you may see this query in the browser address field, and thus determine the parts.

      For example, accessing KudoZ Polish-English terminology base the browser generates the following URL query:

      http://www.proz.com/?sp=ksearch&submit=1&term=
      your_searched_word
      &from=pol&to=eng&whole_words=y&edited=y&unedited=y&glossaries=y&glosspost=y

      As you can see the prefix part is:
      http://www.proz.com/?sp=ksearch&submit=1&term=

      and the suffix part:
      &from=pol&to=eng&whole_words=y&edited=y&unedited=y&glossaries=y&glosspost=y

      Now, fill in the Address start (prefix) and end (suffix) fields with those information. For other languages replace pol and eng in the suffix part with your language pair.

    - Page start and end:

      These two fields are optional. When you load a web page from the Internet, it often contains a lot of unnecessary information. It may be a good idea to filter the page and get only what you really need. In the start field you should provide the start of relevant information on the HTML page and in the end field the end of it.

      Based on it, CafeTran will narrow the page down to what you really want to see. Occasionally, the web page authors change its content so that the page start and end fields might need to be updated too.

    - Cache:

      This field lets you provide a number of web pages that are kept in the cache memory for fast browsing using Back/Forward buttons in the main toolbar. The default number is 10.

    - Encoding:

      In case you have problems viewing national characters in the page, provide the language encoding for it (for example ISO-8559-2 for Central European languages). Since this value is often provided in the page HTML code you may leave this field empty.

    - Style:

      In this optional field you can provide the path to your preferred stylesheet to change

In this optional field you can provide the path to your preferred stylesheet to change the colors or fonts of the internal browser. Choosing System browser in this field will let you view the search results in the system browser such as Firefox or Internet Explorer.

- Description:

  Here you can type any short description of the resource.

4. Press OK button and save the Internet configuration file under some name in the Internet directory. You can also make subdirectories of the parent directories and save the file there organizing your Web resources in a tree-like structure.

Opening internet resources

1. Select Library | Internet | Your resource. A new window will appear to access the Internet site.

2. Type a word in a CafeTran search field or select it with a mouse from the source language window.

3. Press the Enter key or the Search button in the toolbar. The results of the query should appear in either CafeTran's internal or system browser. The Back and Forward buttons in the main toolbar let you navigate through the search results.

External browser

It is possible to open all your Internet resources in the system web browser such as Internet Explorer, Opera or Firefox. In the Edit | Options | Browser field choose System. Then, the search query will be redirected to the system web browser. If you wish to set only particular Web resources to view in the system browser, choose "System browser" in the Style field of their configuration dialog.

Editing the Library configuration file

1. Select your Internet resource tab.
2. Choose Library | Edit resource info from the main menu.

Image resources

Even today in the digital era there are still great dictionaries available only on paper. You can scan them to the computer and save in some image format such as jpeg or png.

In CafeTran, not only is it possible to open scanned images but also search for their pages much more rapidly than manual browsing of paper books.

Creation of image resources

1. Scan your paper resource to the computer.
2. Save in a directory your scanned pages each under a unique name in the jpeg, gif or png image format. The name of an image file might be the first entry on the page.

Access to image resources

1. Select Library | New resource.
2. Select Image.
3. Fill in the fields in the following way:

   - Path:

     Set the complete path to one image file or a directory containing all saved image files.

- Zoom:

  You can set the default size of images opened in CafeTran. For example, if you set zoom to 10, the size of the images will be increased by 10 percent. Setting the zoom value with a minus sign, for example -10, will decrease all the images by 10 percent. The default value is 0.

- Description:

  Here you can type any short information about the resource.

4. Press OK button and save the image configuration file under some name in the Images directory. You can also make subdirectories of the parent directories and save the file there organizing your images in a tree-like structure.

Opening and searching images

1. Select Library | Images | Your image. In the tabbed view a new window appears.
2. Now you can browse through images by clicking the Back and Forward buttons in the toolbar.
3. You can decrease and increase the size of images by clicking the Zoom+ and Zoom- buttons in the toolbar.
4. Invert the colors of the image by choosing the Invert colors in the window context menu which pops up if you click the window with the right button of the mouse.
5. You can search for image files by typing a name in the Search field and then pressing Enter key or clicking the Search button. The image will appear in the window.

   If the file with this name does not exist the alphabetically closest file is loaded. Press the Back/Forward buttons in the main toolbar to go to the next or previous image.

Editing the configuration file

1. Select the image window.
2. Choose Library | Edit resource info from the menu.

Synchronization

CafeTran can join resources in one view so that any query is directed to them at the same time. You may look up a phrase in a glossary and online resources simultaneously seeing the results in one joined window. Synchronization also helps to organize the whole workplace on the screen.

1. Open two resources to work with in adjacent tabs.
2. Select one of the tabs.
3. Choose View | Synchronized views from the main menu.
4. If you wish to add another resource, open it in the adjacent tab and synchronize with the group as above.

Alternatively, click with a right mouse button at the tab title of the resource and in the pop-up menu choose Synchronize with... | Some other resource to join them in one view. The tab can be moved by clicking the "Dock window at..." menu.

Segmentation workflow

The Segmentation workflow enables the translation of all the documents which are supported by CafeTran's editor. Use buttons or keyboard shortcuts instead of described menu actions.

1. Create a new project and add the source document as a new project file. Click the Document button in the Project info panel. When the Open file dialog appears, choose the source file along with its type.

2. Select Translation Mode: Progressive segmentation - from the list of modes in the Project info panel.

3. Open as many translation memories as you wish to work with. You may also create a new memory for the project segments.

4. If needed, open other resources from the Library such as a glossary, online or image references.

5. Arrange the tabs. Click with the right button of the mouse at the tab's title and synchronize or dock as you like.

6. Choose the source document tab and start the workflow - Translation | Start translation menu.

7. Use Translation | Next/Add/Join segment actions to proceed with the workflow. The Next segment action takes a new segment for translation. The Add segment action does the same and adds the current segment to the memory. The Join segment action joins two segments. The source segments are visible in the source window.

8. Check for any matches in the memory, glossary and Google Translate tabs. Use them by clicking "Exact/Hits/Fuzzy match" headers in the memory window or by selecting the match with a mouse.

9. You can also apply the memory matches from the menu Memory | Insert Fuzzy match/Autotranslation/Subsegment matches. Use keyboard shortcuts.

10. Memory can be searched manually too. Press Translation | Translate in the menu.

11. Use other opened resources to search for unknown words or phrases. Select a resource tab, type a word in a search field or select it with a mouse in the source window. Next, press the "Search" button.

12. Edit the final translation of the current segment in the target window.

13. Transfer any placeables by the menu command Translation | Transfer/Next/Previous/ placeables. Use keyboard shortcuts.

14. You have to transfer any source tags, which appear by joining segments, to the target segment to keep formating. Use the menu Translation | Transfer tags/current tag or keyboard shortcuts.

15. You might wish to provide some additional information to the current segment such as a note or alternative translation. Use the menu Translation | Alternative translation/Add segment note.

16. If you wish, add a new entry to the glossary or memory manually. Select it with the mouse in both the source and target window and click the appropriate button in the target window toolbar: Add term to glossary or Add term to memory. The whole segments are added to the segment memory automatically.

17. Keep pressing Translation | Next/Join/Add segment menu until you finish the translation. Use target toolbar buttons or keyboard shortcuts.

18. As the segments are added to the project they can be seen in the Project window. This window has two view modes. You can switch between the modes by clicking at the eye icon in the project toolbar. Use arrow buttons in this toolbar to browse through the segments.

19. If you want to go back to some previous segment, use Translation | Previous segment menu or the Previous segment button in the target toolbar. Press Translation | Next/Add/Join segment (or target window buttons) to go forward. To come back fast to the current segment, click Start translation button in the target toolbar.

20. Anytime you may preview the target document. Choose Translation | Preview document menu. The preview document will be created in the project directory.

21. Use the "Transfer to target window" button in the target window if you wish to move the source segment to the target window.

22. When finished, export target segments to the target document. Use the menu Project | Export translation or the End - "Export target segments" button in the target window toolbar.

23. Save any updated memories and glossaries. Choose a memory/glossary tab and click the "Save" button in the main toolbar.

24. Open or print the document by its native application installed in your computer. Select the document tab and choose the Edit | Print/View the document menu.

Automatic segmentation

In the above described Progressive segmentation workflow the source document is segmented and translated at the same time, segment after segment. It is also possible to segment the whole document first, and then translate the segments.

1. Choose the Automatic segmentation mode from the list of modes in the Project info panel.
2. Start translation as desribed above.
3. The whole source document wil be segmented and the first segment will appear in the source window.
4. There are two important options here, which you can check in/out as needed in the Edit | Options dialog. The Automate with source transfer option lets you transfer the source segments to target ones. The Translation in Review option in the Options | Memory tab activates/desactivates the automatic matching in memories and glossaries.

Image-Paper workflow

The Image-Paper workflow lets you translate the image or paper documents. Use buttons or keyboard shortcuts instead of described menu actions.

1. Create a new project as described above and choose an image document for translation in the Documents field. In case of paper documents just set a project name in the Documents field of the Project info panel.

2. Select Translation Mode: Image-Paper from the list of modes in the Project info panel.

3. Decrease or increase its size clicking Zoom buttons in the main toolbar. You can invert the image's colors by using the image window context menu.

4. Open a memory and other resources such as a glossary, online or image references, Google Translate etc.

5. Choose the image document tab if you loaded the image and start the workflow from the Translation | Start translation menu.

6. Arrange the tabs. Click with the right button of the mouse at the tab's title and synchronize or dock as you like.

7. Translate any part of the source image or paper document in the target window.

8. Use Translation | Next segment action to divide your translation into target segments.

9. Use resources to search for unknown words or phrases. Select a resource tab, type a word in the search field and press the "Search" button.

10. Memory can be searched manually too. Press Translation | Translate in the menu.

11. You might wish to provide some additional information to the current segment such as a note or alternative translation. Use the menu Translation | Alternative translation/Add segment note.

12. If you wish, add a new entry to the glossary or memory. Type a term in the source window and select its translation in the target window. Next, click the appropriate button in the target window toolbar - Add term to glossary or Add term to memory.

13. Keep pressing the Next segment button until you finish the translation. Use the target toolbar button or a keyboard shortcut.

14. The translated segments are visible in the project window. Use project window toolbar to navigate or switch the view mode of the translation units (the eye icon).

15. If you want to go back to some previous segment, use Translation | Previous segment menu or the Previous segment button in the target toolbar. Press Translation | Next segment (or a target window button) to go forward. To come back fast to the current segment, click Start translation button in the target toolbar.

16. When finished, export the target segments to create a target text document. Use the menu Project | Export translation or the End - "Export target segments" button in the target window toolbar.

17. The exported text will appear in a new tab. The text can be saved (Project | Save document as) or copied to open or paste in a word processor such as MS Word or OpenOffice Writer for further formating.

18. If you save the exported translation with the .doc extension, click Edit | View the documen to open it in MS Word.

19. Save any updated memories and glossaries. Choose a memory/glossary tab and click the "Save" button in the main toolbar.

Review and translation of projects prepared in other translation tools

In CafeTran you can check and translate any supported bilingual files such as XLIFF, TMX and TTX.

1. Click the Project button in the Project info panel and choose a Xliff, TMX or TTX file to open it in CafeTran. These file types may be also opened through the Project | Open project menu.

2. Open as many translation memories as you wish to work with. You may create a new memory for the project segments.

3. If needed, open other resources from the Library such as a glossary, online or image references, Google Translate etc.

4. By default, the automatic look-up in this workflow is off. You can activate it checking in Translation in Review box in the Project info panel.

5. Start the translation through the Translation | Next segment menu. You may also start it by clicking on a segment number in the project window.

6. After the start, the process is the same as in the Segmentation workflow described above.

7. Use the buttons in the target window or keyboard shortcuts to go to the next or previous segment.

8. When finished, save the project and any updated memories and glossaries. Choose a memory/glossary tab and click the "Save" button in the main toolbar.

Clipboard workflow

This workflow makes it possible to translate segments from external applications whose documents cannot be easily converted to some editable form. The best example are PDF files which can only be edited by special software. CafeTran can communicate with the other software by means of clipboard.

1. Create a new project and give it a name in the Document field of the project dialog.

2. Select Translation Mode: Clipboard - from the list of modes in the New Project dialog.

3. Open as many translation memories as you wish to work with. You may also create a new memory for the project segments.

4. If needed, open other resources from the Library such as a glossary, online or image references, Google Translate etc.

5. Arrange the tabs. Click with the right button of the mouse at the tab's title and synchronize or dock as you like.

6. Start the workflow from the Translation | Start translation menu.

7. Copy a part of text such as a sentence or a paragraph in the external application. There is no need to paste. The the first segment of the text will appear in the source window and the queries will be made automatically.

8. Use other resources the same way as in the Segmentation workflow.

9. Edit the final translation of the current segment in the target window.

10. Use Translation | Next/Add segment actions to proceed with the workflow. The Next segment action sends a target segment back to clipboard to paste in the external application. The Add segment action does the same and adds the current segment to the memory.

11. Carry on with this Copy-Translate-Send back cycle until you finish. The rest of the process looks like in the Segmentation workflow.

12. Use the buttons in the target window or keyboard shortcuts to go to the next or previous segment.

13. When finished, save the project and any updated memories and glossaries. Choose a memory/glossary tab and click the "Save" button in the main toolbar.

Alignment workflow

If you have a source document and its translation, this workflow lets you add their segments to the memory. You can also extract some words and phrases from the source and target segments to the glossary or just compare the two documents. The process can be automatic or manual. Before the automatic conversion, the two documents need to be prepared so that their contents would be equally separated by paragraphs or sentences. In the manual mode you have the full control over the alignment of the segments.

1. Create a new project and set the source document in the Document field of the New project dialog.

2. Select the Segmentation translation mode from the list of modes.

3. Check in the Documents Alignment box.

4. Press OK and then choose the target document in the File dialog that pops up.

5. Open as many translation memories as you wish to work with. You may also create a new memory for the project segments.

6. If needed, open other resources from the Library such as a glossary, online or image references, Google Translate etc.

7. Select the tab with the two documents and start the workflow from the Translation | Start translation menu.

8. The Alignment menu and toolbar will appear up in the target segment window.

9. The first three Next/Add/Join source segment buttons in the alignment toolbar take a segment from the source document.

10. The last three Next/Add/Join target segment buttons in the alignment toolbar take a segment from the target document.

11. The Next/Add/Join segment buttons in the target window down toolbar take segments from both documents at the same time.

12. Use keyboard shortcuts ALT - F1,F2,F3 for the source document segments and ALT - F10,11,12 for the target document segments.

13. The process of working with memories, glossaries and other resources is the same as in the Segmentation workflow.

14. When finished, save the project through the Project | Save menu. Save any updated memories and glossaries, too. Choose a memory/glossary tab and click the "Save" button in the main toolbar.

You can press Translation | Automate in the menu to automate the process. Next, send the created project segments to the memory through the Memory | Conversions | Import from project menu command. However, it requires the contents of the two documents to be equally segmented prior to the automatic alignment workflow.

Connection to OpenOffice (LibreOffice) suite

CafeTran is a standalone application and can translate OpenOffice documents in its own editor. You have the option to connect to OpenOffice to use its functions such as spell checker and thesaurus. Such a connection also lets you translate directly from OpenOffice applications.

- In the Edit | Options | "Connection to:" drop-down list choose OpenOffice, and then click the "Connection to:" button.
- In the Connection dialog click the Classpath button and choose directory of your OpenOffice or LibreOffice installation.
- After a few seconds the Classpath field will be set automatically.
- Save the setting and restart the program. The new menu item "Spell checker" will appear in the Edit menu. The language of a spell checker and thesaurus is chosen based on the target language in the Project.
- The spellchecker and thesaurus will be available for CafeTran's internal editor, as well.
- You can choose the editor for the workflow in the Tool-id field of the Project file dialog.

If CafeTran fails to set the Classpath field in the Connection dialog automatically, you will need type the full path to the four jar files and one directory manually. Based on the below examples for Windows and Mac OS, set the Classpath field appropriately. All the paths should be set in the Classpath field in one line, with no break.

The example of the Classpath field for Mac OS:

/Applications/OpenOffice.org.app/Contents/basis-link/ure-link/share/java/ridl.jar:
/Applications/OpenOffice.org.app/Contents/basis-link/ure-link/share/java/jurt.jar:
/Applications/OpenOffice.org.app/Contents/basis-link/ure-link/share/java/juh.jar:
/Applications/OpenOffice.org.app/Contents/basis-link/program/classes/unoil.jar:
/Applications/OpenOffice.org.app/Contents/MacOS

The example of the Classpath field for Windows:

C:\Program Files\OpenOffice.org 3\program;
C:\Program Files\OpenOffice.org 3\URE\java\juh.jar;
C:\Program Files\OpenOffice.org 3\URE\java\jurt.jar;
C:\Program Files\OpenOffice.org 3\URE\java\ridl.jar;
C:\Program Files\OpenOffice.org 3\Basis\program\classes\unoil.jar

OpenOffice (LibreOffice) as an external editor

1. Create a new project and click the Document button in the project dialog. When the Open File dialog appears, choose the OpenOffice source file and select OpenOffice in the Files of type field.

2. Switch to the OpenOffice Writer editor. This is done by choosing CafeTran-OpenOffice in the Tool-id field of the Project dialog.

3. Press OK in the New Project dialog. In the OpenOffice File Chooser select the OpenOffice file which has the target language extention in its name.

4. The OpenOffice document will load into the OO Writer. Set the cursor there at the start of the source text and, back in CafeTran, start the translation.

5. A new tab with OO segments will appear. The rest of the process is the same as in the Segmentation workflow.

6. As you go along the current OO segment will be highlighted in the OO Writer.

7. When finished, click Project | Export translation in the menu. The translation will be sent to OO Writer.

As you can see, the OpenOffice (LibreOffice) documents can be translated both ways: through CafeTran's internal editor or directly through OpenOffice suite. It is up to you which one you choose but it's recommended that longer OO documents be translated in the internal editor while short forms with tables, invoices etc. is more convenient to translate straight from the OpenOffice Writer.

Remember to switch back to CafeTran in the Tool-id field of the Project dialog to return to the internal editor.

Spell checking, thesaurus and auto-correction

CafeTran can use OpenOffice (LibreOffice) writing aids in its interface. It uses the installed OpenOffice dictionaries and underlines any misspelled words in the target language window. For instance, if you set "en-US" as a target-language in the project, the spell check function will use the American-English dictionary out of the OpenOffice available dictionaries.

1. Connect to the OpenOffice as described in the previous section.
2. The language of the spell-checker and thesaurus is chosen based on the target language of the project.
3. The mispelled words are underlined in the target language window.
4. Click with the right mouse button on a misspelled word to list alternatives for replacement.

The keyboard shortcut for this action is CTRL+ALT+Space.
5. Automatic correction switches on when holding the CTRL key and pressing the Space key. Then, the misspelled word will be replaced with the alternative. If there are more alternatives, keep pressing CTRL+Space to choose the best.
6. The OpenOffice thesaurus is accessed similar to the spell checker. Hold the CTRL key while clicking with the right mouse button on a word. The list of meanings and their synonyms for the word will pop up to check or replace. The keyboard shortcut for this action is CTRL+SHIFT+Space.

Quality Assurance (QA)

The Quality Assurance system preforms important checks to ensure the target segments do not contain some frequently occurring errors. It can also check for the consistent usage of terms in your translation.

Terms consistency check

1. In the memory start-up panel check the "Terms consistency check" box. You can also set this option after the memory is loaded. Just click with the right mouse button at the memory panel and activate the "Terms consistency check" box.
2. At the end of your translation choose menu QA | Terms consistency check.
   If there are any inconsistencies in the terminology usage, segments containing them will be filtered out and presented in the workflow for correction.
3. After the correction is made, return to the normal workflow by unchecking the "QA filter" box in the menu QA.

Similarly, you can perform the other following QA checks at the end of the translation:

1. Tags check.
2. Spaces check.
3. Numbers check.
4. Empty segments check.
5. End punctuation check.
6. Check all - perform all the above checks in one process.

After the correction of the filtered segments is made, you can either re-check the segments or return to the normal workflow by unchecking the "QA filter" box in the menu QA.

Frequent terms

This feature extracts frequent phrases out of the source document to overview used terminology and select good candidates for new terms to the glossary or terms memory.

1. Make sure that the Extract frequent terms option is checked in the Edit | Options dialog.

2. As you take next segments, phrases from these segments together with a number of occurrences will be displayed in "Frequent terms" window. This gives you a hint which terms will be often repeated and may be added to the glossary or memory for automatic translation.

3. For convenience, join the Frequent terms tab either with the source document or the glossary tab.

Segmentation

Segmentation divides the source text into small parts called segments. There are five predefined types of segmentation which you can select from Edit | Options | Segmentation list:

- sentence (default)

- paragraph
- tag
- word
- document (the whole text is treated as one segment)

Segmentation rules

CafeTran also supports the SRX standard for setting segmentation rules for a given language. It means that you can define exceptions to the segmentation in your language. The default srx file Rules.srx is located in the rules/segmentation directory. It contains the default rules and examples of some exceptions in a few languages.

To edit or add your own exceptions, you need to go to Edit | Options | Segmentation. Next, select a srx file from the list and click the Segmentation button.

The segmentation editor has two menus: languagerules and maprules. In maprules menu you may edit or add a new language map based on the provided examples. Once a new language map has been created, the language for segmentation will be chosen automatically as you make a new Project.

In the languagerules menu, we can create, add and edit the rules for segmentation in a given language. The Default rules take precedence over any other rules and should not be changed.

To add a new language to the languagerules, click "New languagerule" and provide a name of your language.

There are two kinds of rules we can define. Yes rules which determine the segmentation breaks and No rules that define the exceptions to the breaks. Again, look at the given examples to set your own Yes and No rules. In most cases you will need to create only No rules, that is, exceptions to the default breaks since the Default Yes rules are sufficient. The rules must be in the form of regular expressions. The SRX specification available in the Internet gives many examples and can serve as a general guide to setting your own rules.

The rule menu has two submenus for the beforebreak rule and afterbreak rule. For example, if we want to define the exception to segmentation for the English word Mr., first we must create the No break rule with the field set to "no". Then, create the beforebreak rule: \sMr\. (\s stands for a white space and dot must be preceded by the backslash symbol). Next, create the afterbreak rule \s for a white space.

You can also prepare your own srx file in any other xml editor and place it in the rules/segmentation directory. The srx files are selected from the list in Edit | Options | Segmentation.

Placeables

These are source segment words that need no translation and should be quickly placed into the target segment. By default, CafeTran treats numbers and double spaces as placeables. This setting can be customized by providing a regular expresion in the Placeables field in Edit | Options dialog.

When a placeable word occurs in the source segment, it is highlighted. Clicking a "Transfer placeable" menu item copies the current placeable to the target segment. You can easily navigate through placeables and transfer them using keyboard shortcuts shown in the Translation menu - Next, Previous and Transfer placeable. The F4 keyboard shortcut shows all the placeables for transfer.

Useful functions and options

Printing, viewing document and sending e-mail

- You can print the current document choosing from menu Edit | Print.

- The Edit | View document menu lets you see the document in a native application such as the browser or word processor.

- The Edit | Send e-mail menu opens your mail application with a message edited in the target language window. The message reference can be the selection of text in the source language window and the subject of the message is edited in the search field.

Auto-completion

This is a very useful function that completes words automatically while typing. The Auto-completion can also be set to work with the glossary and memory, which means it can autocomplete typing any good matches. The word length and the starting index of words to be completed automatically are set in the Edit | Options | Auto-completion. Check out the "Auto-completion" box of the Edit | Options dialog to disable this function.

Word selection history

This function extracts words from the source segment and remembers them together with their context. Each remembered word is highlighted at its next occurrence in the source segment.

Switch on/off the function clicking the "Word selection history" button in the source window toolbar. Next, select a word with the mouse and press the F5 key to add this selection to the list of selected words.

If there is no selection, pressing the F5 key will pop up a list of all the selected words in their full context. To remove a word from the list, just click at the list item.

Segment Timer (Edit | Options | Segment Timer)

The segment timer has two functions:

1. Automating the process of checking translation in the Review mode. In Edit | Options check in the Segment Timer box and choose the following:
    - Timer Mode: Review
    - Timer Action: Next segment
    - Timer Delay: change the delay time to accelerate or slow down the timer action.
2. Building segments in heavy formatted documents. In Edit | Options check in Segment Timer box and choose the following:
    - Timer Mode: Translation or Review depending how you organize you workflow.
    - Timer Action: Join segments
    - Timer Delay: change the delay time to accelerate or slow down the timer action.

To stop the timer action, press Esc key or a clock button in the source window toolbar. Take a next segment to restart it.

To switch on/off the timer temporarily, check in/out the Timer box in the source window toolbar.

Autosave (Edit | Options | Workflow | Autosave)

The project and memory are saved automatically after every n-th segment added, where n is your chosen number in the field. If you set 0 here, then the autosave function is off.

Clipboard sensitive target (Edit | Options | Workflow | Clipboard sensitive target)

After enabling this function, CafeTran captures any clipboard content, sets it in its search field and looks it up in a current resource. This is especially useful when you work with other applications and need to search in CafeTran's resources. For example, copying a phrase in a web browser will send it automatically to CafeTran for look up. Check it in and select Look up as Clipboard sensitive target in Options to enable this function.

You can also redirect the clipboard content to the target window instead of the search field. For instance, this will let you fast copy any selection from a web browser to the target segment

window. Check it in and select Segment as Clipboard sensitive target in Options.

When this function is on, CafeTran "steals" any clipboard content, so please check it out in Options for normal use of copy, cut and paste functions.

Google Translate (Edit | Options | Workflow | Google Translate)

By selecting this option, CafeTran adds the Google Translate online service to its workflow. Then, each next source segment is also translated automatically by the service. You can also translate the source window content or selection by clicking with the right mouse button at the Google Translate window. The result is shown in the window and used by the auto-completion function. The keyboard shortcut or mouse selection transfer the translation to the target segment window. The source and target languages are chosen from the list in the Edit | Options dialog.

Memory options (Edit | Options | Memory)

- Memory matching type - the default setting for the matching type.
- Memory integration - the default setting for the integration of memories with the workflow.
-
- Do not match - the regular expression which determines the characters skipped during matching for the speed reason. Punctuation characters are default setting.
- Fuzzy match threshold - the minimum accuracy in the percent number for the fuzzy matches.
- Function words threshold - the percent number which determines the statistical threshold for function words.
- Subsegment to Auto threshold - the statistical number for subsegment matches to be included in automatic translation.
- Subsegment to Virtual threshold - the statistical number for subsegment matches to be placed in the virtual memory for fast access.
- Auto-translation usage accuracy - the minimum accuracy in the percent number for automatic translation to be transfered to the target window.
- Subsegment look-up limit - the maximum number of subsegment fuzzy matches to search for in the current segment.
- Transfer numbers to matches - switches on/off automatic transfering of numbers in the source segment to matches.
- Translation in Review - switches on/off automatic integration of memories and glossaries in the Review workflow.

Formating tags

Most of the documents to be translated are in some kind of the XML format, which means dealing with tags to preserve formating. CafeTran offers a very clear way of transferring tags to the target segment. The menu Translation | Transfer tags lists all the source tags to be transferred to the target segment after the user clicks on a tag in the list. Choosing "Transfer current tag" in the Translation menu transfers the current tag at the top of the tag list.

The tags can also be transfered by typing the tag number and pressing the Esc key.

By joining two segments together you will see a tag which has to be transfered to the target segment window. If you do not do this, the creation of the target document may not be correct. CafeTran performs checking if the source and target tags match and issues a warning message otherwise.

Warning: Make sure that any tags visible in the source segment window are transfered to the target segment window. The export may not work properly if the source and target segment tags do not match!

Splitting and joining segments

Splitting segments

1. Set the cursor at the splitting point in the source segment.
2. Set the cursor at the splitting point in the target segment.
3. Go to Translation | Split segment menu or press the Split segment button in the source window toolbar.

Joining segments

1. Go to Translation | Join segments menu or press the Join segment button in the target window toolbar.
2. The current segment will be joined to the next one.

Use keyboard shortcuts to speed up the above segment operations.

Conversions

Given the XLIFF project format and TMX memory format, CafeTran's conversion is a process which takes place in the Segmentation workflow. So by translation of a document you end up in a valid XLIFF and/or TMX file.

Automatic document conversion

1. Start a new project.
2. Open a document.
3. Press the "Start translation button" in the target window toolbar.
4. Click Translation | Automate in the menu.
5. Save the project.

The document will be segmented into the XLIFF project file and ready to translate in the Review workflow.

Conversion of TMX memories

A memory can be easily converted into the project for translation or edition purposes.

1. Start a new project.
2. Open a memory to be converted.
3. Click Memory | Conversions | Export to project.
4. Edit or translate the imported segments in the Review workflow.
5. After completion you may send the segments back to the memory through the Memory | Conversions | Import from project menu.
6. Save the project and the memory.

Glossary conversion

To convert a glossary into the memory:

1. Open a glossary from the Library.
2. Open or create a new memory.

3. Select the memory tab.
4. In the menu choose Memory | Conversions | Import glossary.
5. Save the memory.

Statistics and billing

During the translation you can see all statistical information for project segments, memory matches as well as billing calculation, duration and pace of translation in the Project | Statistics menu.

At first, in the field Edit | Options | Billing you need to provide basic information about the billing that you use and the currency. You can set three sorts of billing calculations:

- based on the number of words
- based on the number of characters (white characters such as space, tab. line separator are included)
- based on the number of characters without whites

Press the Billing button in Options and fill in the fields providing a number of words/characters/all characters, price you charge for this number and the currency. Just type your currency in the currency field if there is no selection for your currency.

Here are some examples:

1. 1500 characters 40 USD (1500 characters are worth 40 dollars).
2. 1 word 0.50 EUR (one word is worth 50 euro cents).
3. 1500 all characters 40 USD (1500 characters with whites are worth 40 dollars).

Translation management

For each project it is possible to set billing specific information regarding your order, job, document and payment. Order details can be added to the chosen Excel or OpenOffice Calc spreadsheet. Invoice functionality includes generation of invoices in your language for printing either from OpenOffice Writer or Ms Word.

Billing and invoicing

1. Choose Project | Project info | Billing.
2. Fill in the fields.
3. Press the Invoice button to create an invoice in your language.

   In the Invoice dialog, pressing the total or sub-total buttons calculates the sum based on the given tax. The possible invoice templates are in the Ms Word (english.docx or english.xml) and OpenOffice Writer (english.odt) formats. They are selected after clicking the Invoice template button in the Invoice dialog window.

   It is also very easy to adjust the invoice templte to your language. Just edit the appropriate template file located in layouts/invoice/ in the CafeTran installation directory and change the English language information to your language leaving only the entries starting with the set prefix. Then, save the new file under some name (i.e german.odt) in the same directory. Next, click the Invoice template button in the Invoice dialog and select your new template file. Finally, click print and you will edit the invoice ready for printing in your favorite editor.

4. Click the OK button to accept the information.

Order registration

1. Choose Project | Project info | Billing.

2. Press the Save button.

   The Order registration field in the Register section will change from new to registered. If you have not set the file path to your register (File button), the new register will be created for you automatically. CafeTran gives you the option to add the information to the spreadsheet file if you use either MS Excel or OpenOffice Calc. Just choose the relevant template by pressing the Register template button and selecting another template (english.ods for Calc format or english.xlsx if you use MS Excel).

   As it was the case with the invoice template, the register may be adapted to your own language too. You only need to edit the default English template located in the layouts/register directory of CafeTran installation to your language replacing the names of the fields with your language equivalents. After saving the template file under a new name (for example spanish.xlsx) set your new template file in the register section by clicking the Register template button and choosing this file.

   The first time you press the Save button, it adds the new entry to the register. Any subsequent pressing of this button opens the register.

Keyboard shortcuts

The keyboard acceleration that CafeTran offers is visible next to a particular menu item in the main menu. The default key modifier may be different depending on operating system (ALT - Windows and Linux, CTRL+ALT - Mac OS) and can be changed in the menu Edit | Keyboard shortcuts modifiers. Below are shown the keyboard shortcuts where ALT is a default key modifier.

- F1 - list and transfer automatic translation and best fuzzy matches.
- F2 - list and transfer subsegment matches.
- F3 - list all tags for transfer.
- F4 - list and transfer current placeables.
- F5 - show all selected words.
- F6 - automate the next segment step.
- Alt-9 - transfer the tag at the top of the tag list.
- Alt-Equals - next placeable.
- Alt-Minus - previous placeable.
- Alt-0 - transfer placeable to target window.
- Alt-P - add selection to placeables
- Ctrl- , (comma key) and Ctrl- . (dot key) - tabbed windows navigation.
- Ctrl-Shift- [1-9] - numeric tabbed windows navigation.
- Ctrl-Enter - selects the word at the caret and looks it up in the chosen resource.
- Ctrl- / (slash key) - search box focus.
- Ctrl- [- source window focus.
- Ctrl-] - target window focus.
- Alt-Home - start the translation button.
- Alt-Right - next source segment button.
- Alt-Up - Join segments button.
- Alt-Down - add units to memory button.
- Alt-Enter - translate button.
- Alt-Left - previous segment button.
- Alt-D - preview current translation in document application.
- Alt-End - export segments button.
- Alt-S - split segment.
- Alt-G - add selected phrase to a glossary.
- Alt-M - add selected phrase to a memory.
- Alt-K - transfer Google translation to the target window.

- Alt-K - transfer Google translation to the target window.
- Alt-J - transfer Bing translation to the target window.
- Alt-I or Ins - transfer a source segment to the target window.
- Alt-F7 - look up in a selected resource.
- Ctrl-W - close the current tab.
- Ctrl-Q - exit the program.