

1. Tovek Topic Mapper

Tovek Topic Mapper application has been developed by Tovek to support curriculum experts in mapping their curricula. The application is written entirely in Java so that it is platform independent. Tovek Topic Mapper is compatible with the ISO standard Topic Maps (ISO 132 50) and uses the standardized XTM data format for storing and exchanging data.

The application offers two main functionalities. It allows users to identify concepts in a text and to associate them with concepts contained in several vocabularies. This process is called mapping of the text. But it also supports users in building and enhancing these vocabularies by adding relations between existing terms or adding missing terms that have been identified in a text. Tovek Topic Mapper has been designed to make the mapping process as comfortable as possible so that the user can simply map the document by several mouse clicks or just by dragging and dropping vocabulary terms to text paragraphs. The application supports multilanguage vocabularies and also the user interface can be easily translated.

The current version of the Tovek Topic Mapper can be downloaded at <http://www.tovek.cz/produkty/topicmapper.html>

1.1 Application data and its structure

In the current stage of development the application stores all its data locally in XTM files. In general there are three types of objects that are stored in these files:

1. **Vocabularies** – each vocabulary is stored in a separated XTM file. By default each vocabulary contains three predefined term associations: BT-NT, RT, UF, but users can extend the semantics of relations between vocabulary terms by defining and using their custom term associations. All introduced term associations can be later used to navigate through the vocabularies within the application. To make the navigation even easier a vocabulary can be divided at the top level into vocabulary sections. Vocabulary terms must be defined in the main vocabulary language and can be translated to other languages supported by the vocabulary. In addition each term can have any number of definitions and examples in any by the vocabulary supported languages.
2. **Documents** – every imported document is stored as a Topic Map in a separate file. The document Topic Map contains information about the structure of the document and about its content.
3. **Mapping** – Mapping information is stored separately from documents and vocabularies in so called link files. This way documents remain untouched during the mapping process so that they can be remapped against other vocabularies just by replacing their link files.

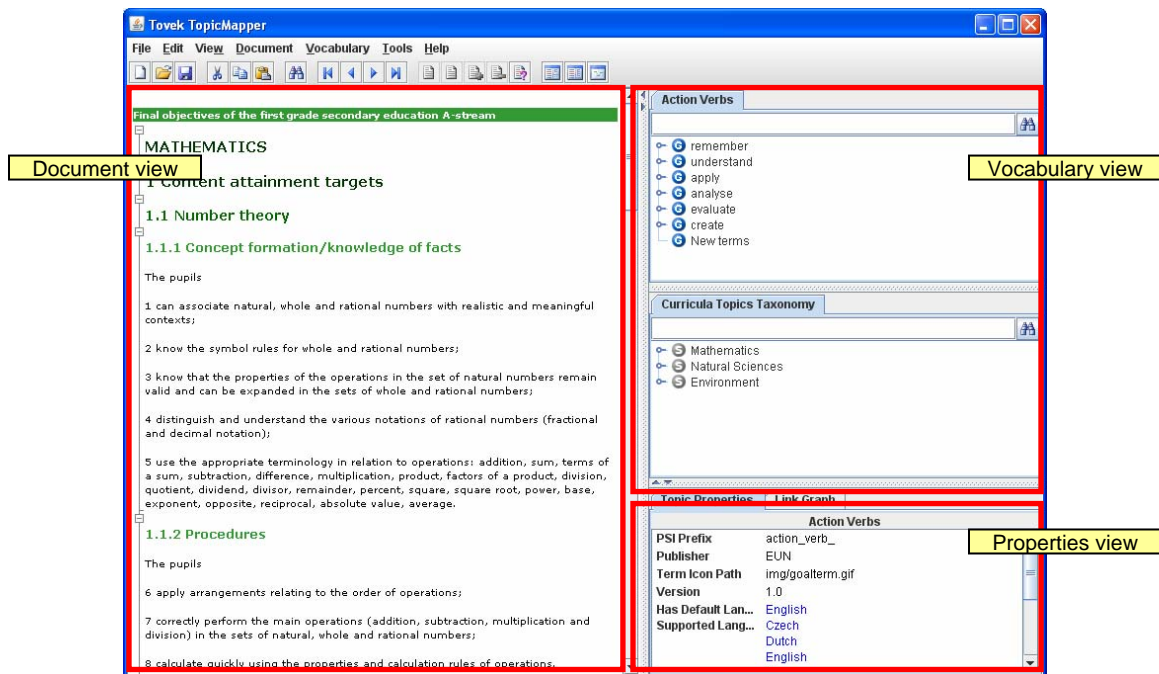
All Topic Map files and data are handled with the help of the open source Topic Map engine TM4J¹.

1.2 Tovek Topic Mapper user interface

The application window is divided into three main views (See picture 1).

The left part is called document view. It shows the text of the currently opened curriculum document with highlighted parts that have been mapped so far. In addition the view shows the document structure and allows the user to collapse or expand selected document sections so that the handling of long curricula is more comfortable. Tooltips along this view show overview information about the mapping of the text piece under the mouse cursor. The application supports the user in navigating through the text by automatic recognition of text paragraphs. The user can jump to next or previous paragraphs just by clicking the appropriate toolbar buttons.

¹ TM4J Project homepage: <http://www.tm4j.org/>



Picture 1: Default window layout

The right top view contains all opened vocabularies. In general the user can open as many vocabularies as desired while he can choose whether a vocabulary appears in the top or bottom half of the vocabulary view. Each vocabulary is displayed in form of a tree according to the predefined broader-narrower term relation and each term is shown with a vocabulary depended icon. Terms of multilanguage vocabularies are displayed either in the current user language if their translation exists or in the default vocabulary language. In the later case the terms are displayed with a special yellow icon indicating missing translation.

If the fulltext search is enabled for the Tovek Topic Mapper each vocabulary view contains a field for entering queries. The simple search runs only through the concrete one vocabulary that is contained in the focused vocabulary view and only the best match is highlighted. The advanced search involves all opened vocabularies and a dialog with the search result pops out. The open source fulltext search engine Lucene² carries out both searches and it searches through all term titles, definitions and examples.

The third view is located at the bottom right of the application window. It is a property view and it displays all available context information of the currently selected items. The property view itself consists of two windows. The first one shows the context information in a textual form, the second one in a graphical way. The concrete displayed data depends on the kind of the selected item. When a vocabulary term is selected, the text part of the vocabulary view shows its name, all its definitions and examples in the current user language. It also shows all other terms that are associated to the selected one and a description of all text parts which are tagged by the current term. All the displayed items are clickable so that the user can use them to navigate though the data. For a vocabulary term the graphical part of this view draws a graph showing the selected term and its associations to other terms. The context information of a selected document part is a description its mapping.

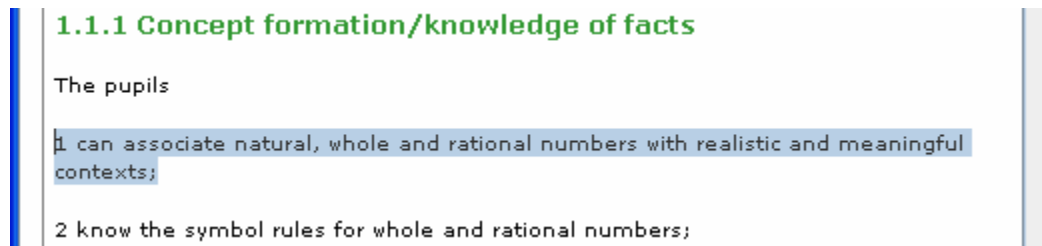
1.3 Mapping of a document

² The Apache project Lucene Java homepage: <http://lucene.apache.org/>

Before the mapping process of a document can start, the document must be imported into the Tovek Topic Mapper. The current application version supports importing only of HTML documents. Importing itself can be done using the menu item *Document->Import Document...* Once a document is imported its content is displayed in the document view. At the beginning the document does not have any structure. It is recommended to define the document structure before the actual mapping starts. It can be done by selecting the menu item *Document->Document Structure...* which opens a dialog that shows all already defined document sections, allows adding new sections or removing existing ones.

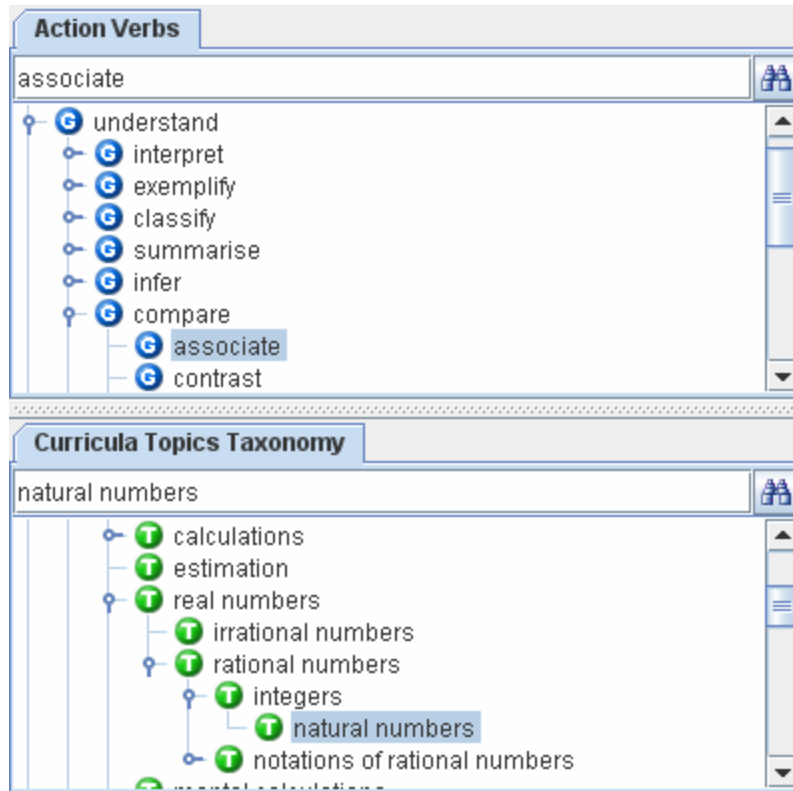
When the document is ready for mapping the user can open vocabularies that (s)he is going to use. Picture 1 shows the application with a mathematics curriculum with already defined structure and two vocabularies – topic vocabulary and action/goal vocabulary. Now the actual mapping process can start. It involves following three steps:

1. **Selecting part of the document** – in general the user may select any part of the document text but for the purpose of curricula mapping it is strongly recommended to select complete document paragraphs to achieve the desired granularity level. Working with paragraphs is fully supported by the application so that the user has a possibility to navigate paragraph by paragraph through the whole document text just by using toolbar navigation buttons or he can select a paragraph by holding down the Ctrl-key and pressing left mouse button. (See picture 2)



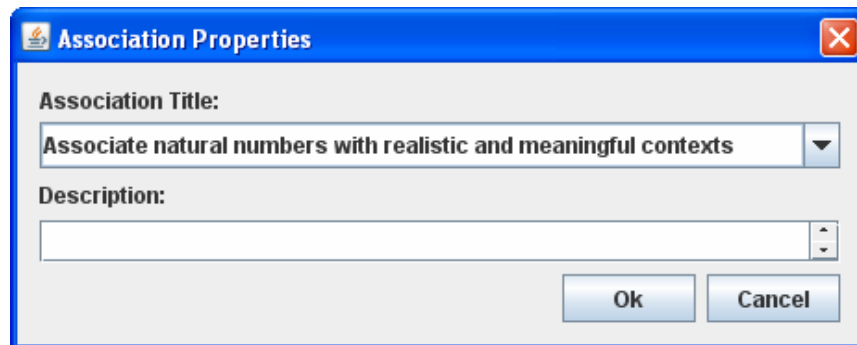
Picture 2: Selected document paragraph

2. **Selecting terms** – in general any number of terms can be selected and used for mapping. In case of curricula mapping the application should be started in so called "tuple mode". In this mode the user has to select one action verb and at least one topic. All the terms can be selected using mouse in combination with pressing Ctrl-key. (See picture 3)



Picture3: Selecting goal term “associate” and topic term “natural numbers”

3. **Mapping** – the mapping is simply done by pressing the *Tag Selection* toolbar button or pressing *Ctrl+T*. If the application runs in the “tuple mode” a dialog for entering the association title appears automatically. The user can enter a title and a description of the current mapping. If the combination of the terms has already been used for tagging the previously typed title is suggested to the user. It is recommended to use only one title for each term combination. (See picture 4)



Picture4: Enter title for the current mapping

To complete the mapping of the whole document the user has to repeat the steps 1 to 3 until all important paragraphs of the document have been processed.
 After a paragraph has been mapped it appears underlined within the document view (see picture 5) so that it is easy to identify all remaining document pieces that still need to be mapped.

1.1.1 Concept formation/knowledge of facts

The pupils

1 can associate natural, whole and rational numbers with realistic and meaningful contexts;

- Associate natural numbers with realistic and meaningful contexts (natural numbers, asso

Picture5: Mapped document paragraph is underlined, tooltip shows information about mapping

Once the whole document has been mapped Tovek Topic Mapper can export it in the ZThes format, which can be directly processed and displayed by the portal.