ATop: A Software Tool for Topic Map Technology

Athanasios Hatzigaidas, Georgios Tryfon, Anastasia Papastergiou
Department of Electronics,
Technological Educational Institution of Thessaloniki,
P.O BOX 14561, GR - 541 01, Thessaloniki ,GREECE

Abstract: - Topic Map (TM) technology is a new semantic approach for information representation, management and navigation gaining increasing popularity. In order to realise and exploit the potentials of this new technology, the need for software development to manage TM is indispensable. In the context of this paper, we have attempted to outline some key factors, regarding the level of functionality that an integrated TM software application tool should provide. Consequently, keeping in touch with the main design principles, a software application tool (ATop) for efficient TM editing, browsing, navigation, management and querying has been designed, implemented and tested with encouraging results.

Key-Words: Topic map, software tool, edit, represent, manage, navigate, query

1 Introduction

Topic maps (TM) as an emerging semantic approach to information organization, describe knowledge structures, associating them with information resources [3,8,21]. TM were designed to support knowledge representation applied to information management from the perspective of humans and they provide a unifying framework for knowledge representation and information management [3,4].

One of the main applications of this new technology is enabling, in an efficient way, enhanced navigation and information retrieval in complex data sets, by adding semantics to these resources [12]. TM architecture enables also merging between topic maps. They can be applied in knowledge management systems, content management, web portal development, and enterprise application integration [5].

TM technology is gaining increasing popularity and consequently there is an indispensable need for developing effective software tools to support users create, edit, query and manage topic maps. Furthermore, the issues of TM navigation, representation and visualisation, as well as corresponding tools to support these functions, are also essential [9].

In the context of this paper we briefly present basic concepts of TM and discuss the main design requirements that should be taken under consideration when developing software tools for TM. Thereupon, based on the proposed design principles, we present the system architecture and basic functionalities of a software application tool (ATop) for efficient editing, browsing, navigating, managing and querying TM, which has been developed and tested with encouraging results.

2 TM basic concepts

The Topic Maps (TM) are a new semantic – ontological approach to knowledge representation and management. The first occurrence of TM was in 1993 when the basic concepts of topic maps where first described. In 2000 topic maps became an ISO/IEC standard 13250:2000 and in 2001, the standard was revised to include the XTM syntax of XML, TM language [3,8,21].

Although TM organise and represent very complex structures, the basic concepts of TM, Topics, Associations, and Occurrences, also known as the TAO of topic maps and the additional concepts of Identity, Facets and Scope (the IFS of topic maps), seem to be simple [16]. A topic is the basic concept in TM and represents a subject [16]. Topics can be categorized according to their types. Figure 1 represents a very small Topic Map, which contains 3 topics: Java, Papas and Athens. These topics are instances of other topics: Java is a book, Papas is an author and Athens is a city. A topic association establishes a relationship between two or more topics, and each topic that participates in an association plays an association role. An association adds semantics to data by expressing a relationship between several topics, such as Java written_by Papas, Papas work_in

Athens, etc. An association may have any number of roles and can easily express more complex relationships [16].
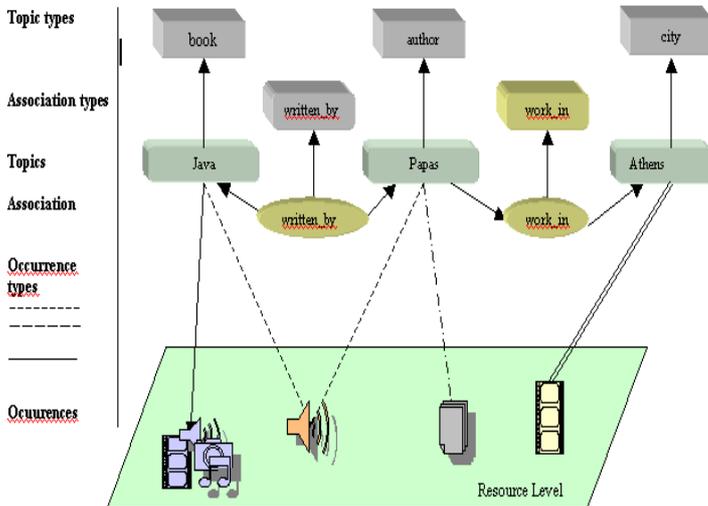

Fig. 1: Example of a Topic Map

A topic may be linked to one or more information resources e.g. web pages, video, images, which are relevant to the topic in some way. Such resources are called occurrences of the topic and they are also categorized according to their type [16]. In the TM represented on Figure 1, occurrences of the topic Athens may be URLs, pointing to some images.

It must be pointed out, that TM technology enables the clear separation between the description of the information structure and the resource level [5].

3 TM tool design requirements

To delve into the world of Topic Maps, a number of operational factors should be taken under consideration, namely: editing, representation and navigation, merging, visualization, querying, storage data format and system platform. A TM administrator needs an integrated software tool that supports these development and management steps.

3.1 Editing Topic Maps

A Topic Map designer should have an explicit understanding of XTM syntax and specifications [11] in order to write XTM code. The current situation is discouraging for new comers in this new emerging technology simply because working with XTM code is tedious. Hence, an absolute necessity lies in the design and development of a visual topic map editor that releases users from directly working on the XML coding.

A main feature of the intended software tool should be to provide an easy-to use GUI (Graphical User Interface) .An integrated TM editor should enable import, and manipulation and export of existing TM or creation and export of new TM. Moreover, to support easy editing and updating of the TM structures and content. A list of the topics, associations, occurrences, scopes, etc could be presented and features like add, remove, rename and delete would be available for manipulating such characteristics of the topic map.

3.2 Representation and Navigation of Topic Maps

Using topic maps for navigating in large information sets, requires an "easy-to-use and an easy-to understand user interface". User interaction, different ways of output and navigation, and rapid searches retrieval, make navigating processes easier [7,9].

For an efficient navigation, users should be able to focus on any part of the Topic Map and observe precise information concerning a specific topic object. Providing different forms of representation requires different levels of detail, which are displayed in several windows simultaneously, offering the opportunity to the user to select the most appropriate representation at a given level of detail [10]. An innovative approach would encompass the possibility to edit and update the topic map objects in any of these forms, and automatically refresh all the other representatives' forms.

3.3 Topic Map merging

Merging of topic maps is another essential function that must be provided to the TM administrator. User should be able to control merging of TM, thus a TM software tool must provide a user interface to allow the parameterization of merging process.

3.4 Visualization of Topic Maps

A tedious problem that TM application developers are facing has to do with the fact that structures of topic maps are not easy to display visually. It is essential to create a software tool that encompasses a visual topic editor that releases users from directly working on discouraging XML coding. Additionally it is of outmost importance for a developer to be able to visualize the overall topic map structure tree hierarchy simultaneously with different levels of granularity

concerning associations, scopes, occurrences, instances e.t.c.

In an efficient representation and consequently navigation of a topic map, special techniques are needed, to provide besides the navigation part, a conceptual image of the stored data [10,11]. A Topic Map can be displayed as a simple graph consisting of nodes and arcs, while different shapes and colours may be used to symbolize various dimensions of nodes and arcs of the graph [10,11]. Each technique is suitable for a specific level of detail.

3.5 Querying Topic Maps

In the same way as relational databases can be queried, a structured retrieval of Topic Map data, is offered to the TM author via TOLOG language [6]. This language can query topic maps for topics of specific types, which participate in certain combinations of associations, and supports inference rules. In order to add a predefined rule, it is necessary to specify the relationship between the parameters of the rule [2]. An integrated software application tool for TM should enable intelligent retrieval of information through the use of rules and queries for even quicker access to the information. Various rules and queries should be taken under consideration for better filtering of information in representations and visualisations of topic map structures.

3.6 Storage data format

Another key issue that should be taken under consideration is the format that is going to be used for data storage. XTM is a standard format based on XML [3], is highly portable and is an acceptable solution for small Topic Maps. Ozone [15] is a fully featured, object-oriented database management system, completely implemented in Java and distributed under an open source license.

3.7 System platform

A topic map tool addressing end users, it would be better to be cross platform portable and compatible. A feasible solution would be to provide standalone windows executable tool and not a minimalist java applet, offering an advanced and friendly graphical user interface.

Increasing popularity of Topic map technology resulted to a proportional indispensable need for developing effective tools to help users manage their topic maps. Consequently a number of tools that support TM technology have been developed in recent years to support TM administrators. Nevertheless there

is still place for a powerful, free and easy to use software tool for TM, and this was the motivation key for this work. Available tools to the potentials developers/users who want to exploit TM technology, have been briefly explored and knowing their shortcomings, we haven't been aware of the existence of a complete free software tool for the creation, manipulation, navigation, visualization, querying and merging of topic maps [7].

4 Basic Architecture and functionalities of TM software tool (ATop)

A software application tool that conforms to the design requirements discussed in Section 3, is ATop. It is a Java-based application that can be used in standalone mode and has been implemented as a module under NetBeans 3.6, in the JSDK programming environment of SUN publication 1.5 [18], due to portability advantages that ensure the platform independence of the developed application. Netbeans 3.6 [14], is a programming environment of open source code in Java, that enables programmers to manufacture and add extra "modules". The use of Netbeans enables the expandability and scalability of the developed software and allows future modular extension of its functionalities.

Moreover, the publication 3.6 has been selected due to its advanced features for creating and managing windowing systems. Thus, the developed application provides a facile graphical user interfaces (GUI), enriched with a lot of enhanced operations. As the GUI is a windowing environment and the user is familiar to work in such environments, he can easy learn and explore all provided possibilities.
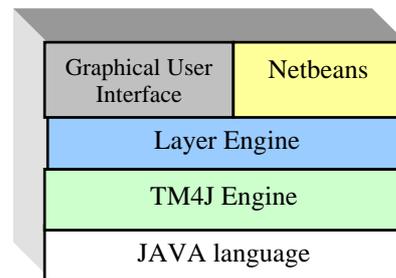

Fig. 2: System components

Low-level management of Topic Maps is achieved by using an open source code library, TM4J and more specifically publication 0.9 [19]. TM4J engine is a software package supporting the XTM-standard, easy to use and provides a rich set of operations (as Java-APIs) enabling import, export and manipulation of Topic Maps. For higher-level management of topic

maps, a new Layer Engine (Figure 2) was designed and implemented in order to manage TM4J engine and organize and represent the topic map information in a user-oriented structure. This way, ATop enables user to view, create and manage TM information in a more human comprehensive form as it will be explaining in the following.

4.1 Representation and Navigation issues

The main application window of ATop is separated in two basic areas: the "TM explorer" area that is located at the left side of the application environment, and the "Properties panel" at the right (Figure 3).
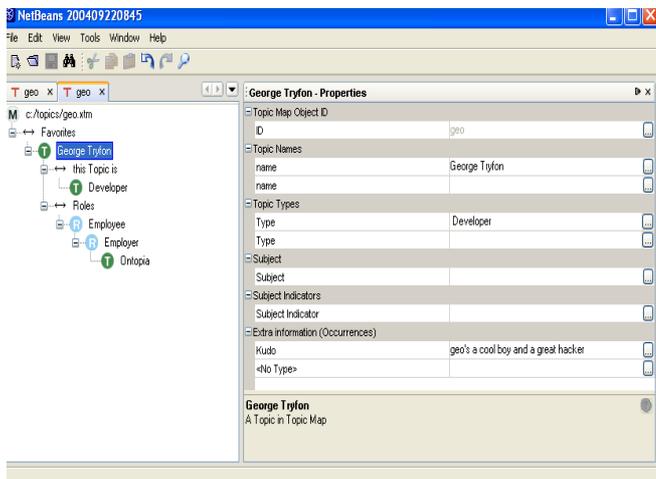
Fig 3. The main application window

When opening an XTM file, the name of the file appears in the "TM explorer" area. The information of XTM file is represented as tree with nodes and subnodes. This tree is collapsible and expandable. For a clear view of the knowledge domain, various icons are used to spot different kind of topic map objects. These icons follow the design of icons that TMNAV program uses [20], as the user may be already familiar to them. Layer Engine organizes information in a user-oriented structure, and at a higher level of XTM structure. This way, a specific knowledge in human readable way can be illustrated and the underlying knowledge can be made explicit to the user by extracting information from the XTM file and representing it in a more comprehensible form.

There are two types of nodes represented: "Topic Map Object nodes", and "Arc nodes". "Topic Map Object nodes" represent some "Topic Map Object" while "Arc nodes" represents relationships between "Topic Map Objects". Arcs can't be deleted, they can only be hidden On the other hand generally, "Topic Map Object nodes" can be deleted and the meaning of

deletion depends on the Arc that this node is child of. For example, if the user wishes to delete a "Topic node" under the "All Topics" node, ATop will check if this Topic is used in the topic map, i.e. it is a "Topic Type" for another Topic, and consequently will permit or not the deletion procedure or it will appear message describing the reason for not deleting it.

Different representatives forms are also supported, providing so various levels of detail in multiple windows that are displayed simultaneously, and offering the opportunity to the user to select the most appropriate view at a given level of detail of TM information. By right clicking on the name of the file, one may choose the preferred view to be appeared in the same window and either expand the existing tree (Expand tree) or replace previous tree or view the information in a new window (New panel). ATop provides the ability to select the detail of information that is going to appear i.e. Topics, Topics Types, Occurrences Types, Member Types, Association Types, Association or All Themes. Furthermore user can edit, alter the topic map in any of these views, and automatically refresh and update TM information in all forms of representation.

Fig 4. A snapshot of menu

User may right click on each node, and select to see all the information concerning the selected node. For example if the selected node is a topic, information about what this topic is, and its role and members are available in "TM explorer" area (Figure 4). All the properties of the selected component are represented in an editable form in the "Panel Properties" at the right frame (Figure 3). This way, a simple interactive navigation and search within all the available "topic" components is enabled.

4.2 Editing issues

Enhanced features, for efficient editing and management of topic map structures, are also provided. With a simple mouse move, users can employ the editor to create, alter and update topics, associations, occurrences, scopes, resources, instances, and the rest of topic aspects, and consequently convert them to their equivalent XTM form. Layer engine

enables the TM administrator to work in a higher editing level, and simplify the editing process. The means that Layer Engine uses to achieve its goal include organisation and representation of information in user-oriented structure without following the strict XTM structure. Additionally, ATop provides a "Panel Properties" frame and a "Choose a topic" dialog box, which help user to create new "topic map object" in a more intuitive way.

4.2.1 Create a new topic
In order to add a new topic user may right click on "All Topics" and choose the "Add new Topic" command. The user specifies an ID name and then a new topic is created and opened in an editable form in the "Panel Properties" at the right frame (Figure 3). In this panel, user may add all the properties concerning the new topic. For example if he wants to add a " topic type" he can click on the button in the line topic type and a new dialog box will appear (Figure 5).
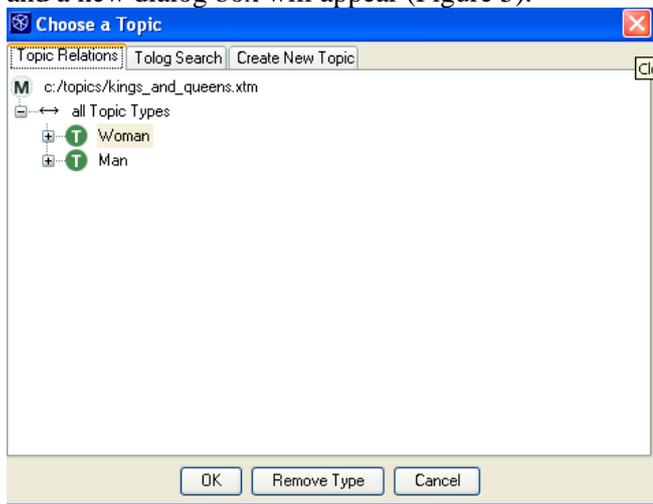

Fig 5. The "Choose a Topic" Dialog Box

The "Choose a Topic" dialog box provides 3 tabs: "Topic Relations', "Create New Topic" and "TOLOG Search", aiming to facilitate the user to choose the topic we wants to add. "Topic Relation" provides to the user a pool of predefined topics. Alternatively, the user can find the topic he is looking for using TOLOG queries, for more sophisticated information retrieval. If the user adds as a topic type, a topic that is not yet defined as a topic type, ATop considers it further on as a new topic type and adds it to the list.

Respectively, the button at the right of the occurrence activates also the "Choose a Topic" dialog box, and user may choose the "Occurrence type" which is also a topic, and consequently fill in the information of the occurrence type. In order to add name, subject

indicator or subject, by clicking the corresponding button a text box appears, where the necessary information can be added.

4.2.2 Create a new association
Associations between topics are created as instances of the ASSOCIATION class. The MEMBER class is used to define each member role of the association and the topics, which play that role. Layer Engine, organizes and manage information in a user-oriented structure, thus enables TM administrator to create associations, by simply specifying for each topic its role, members and role player. Then it manages TM4J Engine, converts user input and creates all the necessary structure for the XTM file in order to establish the new association.

For every new role the user may define its association type (topics) and one or more members (topics) along with the roles they play in the relationship. For example in order to add a new role to the topic "George", by right clicking on Roles and selecting "Add new Role", the "Choose a topic" dialog box appears, providing to the user three options, as explained above. This makes existing experience transparent to the user enabling him to see the existing topics, make a TOLOG search to locate what he is looking for or create a new topic. By this way the user subsequently searches for components and adds them to the current configuration.
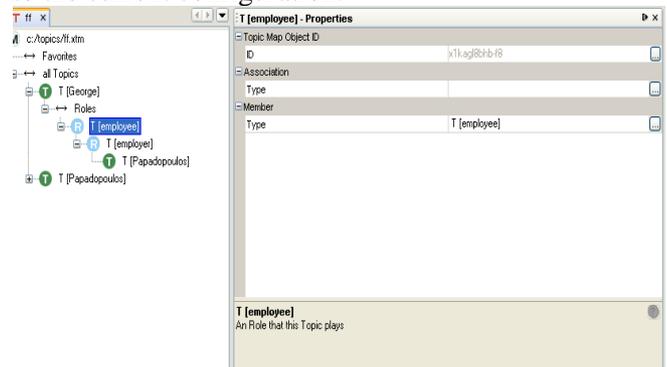
Fig 6.Association Panel Properties

For example, we create a new topic "employee". The editable form of the new Role appears in the "Panel Properties" frame (Figure 6), where the user respectively can add an Association Type, using the "Choose a topic" box (Figure 5). User may enter a predefined Association Type or introduce a new one to be added to the "Association Type" list for future use. Afterwards, by right click on the role "employee", ATop guides the user to import a Role (for other Players) and following the same procedure a new role

is added i.e. "employer". Finally by right clicking on "employer", user introduces the Player for this Role, i.e. Papadopoulos, and the relationship is established. If the user wants to find the Roles that "Papadopoulos" topic plays, is apparent that ATop has also updated the information of this topic too, and represents "Papadopoulos" as an employer, with employee George.

4.3 Rules and queries

ATop supports the use of rule and queries for a more sophisticate information retrieval. In order to successfully launch a query, in a topic map and explore knowledge, there is a need to synchronize various different files that keep different contents. Firstly the topics map itself, which resides in one or more XTM files. Secondly the rules that resides in one or more text files. Rules are written in "TOLOG" language and are kept in files with the extension ".tolog". Finally queries must be defined. Assuming that we would like to keep many queries in a repository, we understand that we need a file to store the queries. Atop's Layer engine loads all the involved XTM and TOLOG files, and it is ready to process the query. For this reason, a new file format, "ATop project" file is created that organizes all the above.

4.4 Storage format

Another key issue kept in mind during the development of ATop, was the format that is going to be used for data storage. TM4J engine supports two fully-featured backends - one based on storing the processed XTM file in memory and one as persistent storage in an Ozone OODBMS back-end.

Two options are provided to ATop user for data storage format:

- XTM file: A standard format based on XML which is highly portable and is proposed for Topic Maps
- ATop project: ATop specific format that can handle many different Topic Map sources simultaneously. It supports merging and OZONE database. It can store custom queries and also supports User rights.

ATop is hosted on Sourceforge, available for free download, so that researchers that are interested, are encouraged to use it and explore all the features it offers. Potentials user's opinions and evaluations will provide valuable input on how ATop should be refined according to each user needs.

5 Conclusion

Topic Map (TM) technology is an emerging semantic approach, which is gaining increasing popularity nowadays. As a result, there is proportional increasing need for effective tools to manage topic maps. A considerable number of tools that support TM technology have been developed in recent years to support TM administrators. However, there is still place for a functional, easy to use and free software tool for TM, and that was the motivation for this research work.

This paper presents our attempt in developing an integrated software application tool for Topic Maps. Firstly we have illustrated the core issues that should be taken under consideration in order to form an efficient software tool for TM. Thereupon, architecture and basic functionalities of ATop were introduced, and the enhanced features that offers to TM administrator for Topic Map editing, browsing, management, querying and navigation were outlined. ATop is a standalone windows executable tool, providing a friendly windowing GUI and its main advantage is that incorporates "Layer Engine" which has been developed for organising information in a user-oriented architecture. Hence, ATop represents knowledge in a human scrutable form and consequently enables ATop's user to create, update, view, manage and query topic map easier and efficiently.

Further work needs is currently under progress in order to build a complete suite. The next step of this work is to intergrade the process of rule and queries construction, in a more perceivable and automatic form, and to experiment with visualization techniques for a more sophisticated representation of TM information.

References:
1. Ahmed, Kal (2000). Topic Maps for repositories, XML Europe 2000 Proceedings
2. Ann Wrightson and Hans Holger Rath, *Requirements for a Topic Map Query Language (TMQL)*, version 0.6, in the tmql-wg archive
3. Biezunsky, M., Bryan, M., Newcomb, S., (1999). ISO/IEC FCD 13250:1999 – Topic Maps, www.ornl.gov.
4. Biezunsky, M., (1999). Topic Maps at a Glance, in: Proc. of XML Europe 99.
5. Garshol, L.M., (2001) Topic maps, RDF, DAML, OIL, A comparison. XML Europe 2001, http:// www.ontopia.net/topicmaps/materials/tmrdfoild

6. Garshol, Lars Marius, (2001) TOLOG, a topic map query language, http://www.ontopia.net/ [Moore, Graham, RDF and TopicMaps: An Exercise in Convergence, In *XML Europe 2001*, Berlin]

7. Hatzigaidas, A., Papastergiou, A., Tryfon, G. & Zaharias, Z. (2004) "Topic Map Existing Tools: A Brief Review", ICTAMI 2004 (International Conference on Theory and Applications of Mathematics and Informatics), Greece

8. ISO 13250, (2000)International Organization for Standardization, *ISO/IEC 13250, Information technology — SGML Applications — Topic Maps* (ISO, Geneva 2000)

9. Le Grand, B., Soto, M., (2000). Information management – Topic Maps visualization, *XML Europe 2000, Paris, France.*

10. Le Grand, B., Soto, M., (2001). *XML Topic Maps and Semantic Web Mining*, Semantic Web Mining Workshop,ECML/PKDD 2001 conference, Freiburg,Germany.

11. Lin, X. and Qin. J., (2002). Building a topic map repository. In: Knowledge Technologies Conference, March 11-13, 2002, Seattle, Washington.

12. Moore G. (2000). Topic Map technology - the state of the art. Proc. of XML Europe, Paris http://www.infoloom.com/gcaconfs/WEB/paris2000/S22-04.HTM, June 2000

13. Munzner, T., (1997). *H3: Laying Out Large Directed Graphs in 3D Hyperbolic Space*, IEEE Symposium on Information Visualization.

14. Netbeans IDE Organisation, http:www.netbeans.org

15. Ozone database homepage www.ozone-db.org/

16. Pepper, S. (2000). The TAO of Topic Maps. Proc. of XML Europe, Paris

17. Rath, H.H. (1999) Technical Issues on Topic Maps, in: Proc. Of Metastructures 99 Conference, GCA, Alexandria, VA

18. Sun Microsystems, Inc.: Java programming language. An interpreted language, http://java.sun.com/

19. TM4J: Topic Maps for Java, http://www.tm4j.org/

20. TMNav http://tm4j.org/tmnav.html

21. XTM 1.0 ,Pepper, S. and Moore, G (2001): *XML Topic Maps (XTM) 1.0* TopicMaps.Org, http://www.topicmaps.org/xtm/1.0/