The i-Disc - A Tool To Visualize and Explore Topic Maps

T. Hofmann, H. Wendler, and B. Froehlich

Faculty of Media, Bauhaus University Weimar, Germany

Abstract
We present the i-Disc, a tool to interactively visualize and explore medium sized topic maps. Topic maps contain two basic structures: the topic hierarchy and the associations between topics. Our system presents the topic hierarchy in a radial planar layout by encoding different hierarchy levels as separate rings. Associations are displayed on demand as three-dimensional arcs across the topic landscape. By separating these two topic map structures into different spatial dimensions, we untangle the often complex topic map graph.

A perspective rendering of our layout generates a natural focus and context display. Our elementary circular design allows quick perception of the overall topic map structure while interactive navigation and exploration provide access to details on demand. The tool is implemented as a client-server application and integrates seamlessly into existing web based environments.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [User Interfaces]: Graphical user interfaces

1. Introduction
We are working on an e-learning project, which develops tools for recording and presenting lectures and courses in a digital format. Even though a series of lectures is originally presented in a linear fashion, the exploration of the material may benefit significantly from non-linear access to the contained information. Lectures are dealing with lots of different topics and associations between these topics. Our idea was to capture this type of information in topic maps.

We designed and implemented the i-Disc, a tool for interactively exploring the content of medium sized topic maps, which contain some hundred to a few thousand topics. Our tool displays the two basic topic map structures - the topic hierarchy and the associations between topics - in the following way:

- **Topic hierarchies** are in most cases tree structures, which we render into a planar circular layout consisting of concentric rings (see Figure 1). Each ring is partitioned into slots and topics are assigned to these slots. Each hierarchy level of a topic map is encoded as a separate ring. Parent-child relationships between topics are displayed as arcs between rings. Our hierarchical layout of the topic hierarchy is similar to well-known radial space-filling techniques.
- **Associations** build a semantically orthogonal structure on top of the topic hierarchy. They are displayed geometrically orthogonal to the planar topic hierarchy layout as three-dimensional arcs across the map.

The topic map is perspectively rendered and allows interactive navigation and exploration of the topic landscape. Displaying both topic map structures simultaneously results in general in a complex graph visualization problem. We separate these structures into different spatial dimensions and display associations as well as further topic map information only on demand, thus avoiding clutter and unnecessary occlusion of the underlying topic hierarchy.

During our work on formatting and structuring course material, we found that much of that material is inherently organized hierarchically. One course is divided into a certain number of chapters each dealing with a separate area. Within these chapters, sections and subsections continue to build the hierarchy, which ends with index terms found in paragraphs. This leads to an initial tree structure with chapters and section titles serving as inner nodes of the hierarchy, and index terms serving as leaves. Initial associations can be found by parsing cross references from the text, and by associating text elements with identical index terms.

Since the topic maps for our course ware are still under development, we used a topic map throughout this paper, which contains information about concepts, tools, standards,
and organizations involved in the area of XML processing and development. The topic map was compiled by Lars Marius Garshol [Gar] and contains approximately six hundred topics and one thousand associations. This size matches our i-Disc design well.

The main contribution of this paper is the development of a new topic map visualization tool - the i-Disc. The circular design provides always an overview of the topic map structure, while the perspective rendering generates a natural focus and context display similar to the perspective wall [MRC91]. Interactive exploration allows responsive access to details on demand. Navigating the topic map involves rotation of the three-dimensional model, which introduces motion parallax as a cue for better understanding of the three-dimensional association link structure. Our implementation uses VRML as a front end, supports hardware accelerated browsing, and integrates seamlessly into existing web based environments. An informal user test indicates that users quickly acquire a basic mental model of the topic landscape while browsing and exploring our circular layout.

2. Background

The paradigm of topic maps has first been fully described in [BBN99], and is based on SGML and Hytime [ISO86] [ISO97]. With the increasing popularity of XML [BPSM98], serialized interchange formats for topic maps emerge in XML, such as [XTM01].

An application creates the topic map by reading the elements from such files. The resulting map basically consists of:

- Topics
- Associations
- Occurrences

Technically, a topic is a resource that acts as a proxy for some subject; it is the topic map system’s representation of that subject. This means that the topic in the topic map is an abstract representation of a “subject” (a person, an entity, a concept, really anything) we can think of.

There are two general types of topics: typing topics and instance topics. Typing topics abstract instances and provide a grouping or classification similar to classes in object oriented programming languages. Topic instances are instances of typing topics. The “products” topic is such a typing topic, whereas the topic "expat" represents the real world software with the same name, which is of the type "product".

An association is a relationship between one or more topics, each of which plays a role as a member of the associa-
tion. In our example, an association might be "subcategory", where the topic plays the role of "category" and the children play the role of "subcategory". Another example is an association, which connects a specific product to its vendor.

An occurrence is any information that is relevant to a given subject. Such occurrences must be resources that are either addressable by reference using an URI or are capable of being placed in-line as character data. Here, most occurrences are URIs, linking to "real world" internet addresses the author of the topic map considered relevant for the topic.

Such a topic map can be seen as a semantic net because of the meaning of the associations and topics. A basic introduction to topic maps is given in [Pep00], an introduction with regard to their visualization can be obtained from [LS00] or [JS02a], and an overview of different aspects of topic maps can be found in [JS02b].

In general, topic maps are stored in "engines" or database backends using open or commercial tools [Bar02], [APW96], and their visualization is achieved by either rendering them as HTML pages (using hyperlinks to follow relations or occurrences) or different graph visualization techniques.

3. Related Work

Several graph and tree visualization concepts have been applied to topic maps, e.g. polar graphs, cone trees, and hyperbolic trees. These, and other applicable concepts, will be briefly discussed.

Auillans and Delahousse suggested the use of a polar graph in their implementation of the "Topic Navigator" [DA02]. In this visualization, the relations between topics are drawn as straight lines in the center of a circle, as well as arcs around the map. However, too many associations will draw a thick inarticulate circle around the map. Additionally, the approach depends on the planarity of the visualized graph.

Le Grand and Soto used the UNiversal Interactive Visualization Tool UNIVIT [LS00] to display topic maps. UNIVIT employs a cone tree concept to visualize graph data. Since cone trees cannot avoid to occlude nodes, there is no possible position from where all nodes can be seen. Additionally, the strength of cone trees lies in visualizing graphs that are structured hierarchically, but yet another layer of connections between nodes - such as associations - is difficult to integrate and leads to graph cluttering.

Hyperbolic trees hide the content at the outer areas of the screen to purposely reduce the amount of information displayed. The "k 42 TMV" visualization developed by empolis [JS02a], uses colors to differentiate the various connections to and from a topic (associations, occurrences, class-instance relations). However, such a concept leads to topics reappearing repeatedly at different locations with different colors defined by the current role. For this reason, the dynamic creation of a topic map visualization seems to be appropriate only for very small and concise maps where a user is able to detect these redundancies, or very large and complex maps where static algorithms fail by delivering cluttered views.

Munzner introduced the H3 layout technique to visualize large directed graphs [Mun97] in 3D hyperbolic space. This technique handles huge graphs with over 20,000 nodes. This work makes also use of specific domain knowledge to derive a spanning tree as the basic information structure. The hyperbolic navigation provides a focus and context display similar to our perspective rendering of the i-Disc.

Approaches to the visualization of hierarchies and graphs by means of radial space-filling techniques like InterRing [YWRP03] or Sunburst [SZ00] offer the ability to efficiently use the display space while effectively conveying the hierarchy structure. In both cases, different visualizations and navigation and interaction techniques are used to alternate between global and detailed views. Furthermore, techniques are described which enable the user to reconfigure hierarchies, distort objects, expose or hide objects, and pan, zoom and rotate the display. In our approach, we limit ourselves to a very basic navigation which does not require switching between different views.

The Harmony Information Landscape [APW96] visualizes hierarchical structures on a plane and connects documents located above and under the plane via visible links. Interaction with the visualization leads to the user’s viewpoint gently flying to the appropriate location. An unobstructed view of the whole scene seems very difficult to achieve.

Paley’s TextArc work [Pal02], allows the user to see the frequency and distribution of words of a text on a single screen. While this approach is an appealing visual, it does not make use of metadata and structure, even if such data were available, and seems therefore less suited for topic maps.

Duffett and Vernik’s netmap [DV97] places items around the perimeter of a circle and draws links between them. Our technique is a combination of this approach and radial drawings of a tree. Our contribution is in combining the two views and extending these concepts with a natural focus and context display, 3D interaction techniques, and domain specific content sorting and weighting.

4. Design Intentions

With the design of the i-Disc we wanted to provide a few expressive structuring elements for the topic map display and in addition details to browse. The most prominent elements are the rings composed of detail proxies representing the topics. This results in a very calm and concise startup state.

The semantic net represented by the topic map has to be
turned into knowledge in the users’ minds. This means that
the users have to be supplied with a clear, significant, and
quickly perceivable representation facilitating fast memo-
ration. We separated these general requirements into the
following individual issues:

- **elementary structure**
  To offer a simple appearance yet hide none of
  the nodes, the large constituting structures are
  composed of single entities.

- **immediate browsing**
  All nodes are immediately accessible in one in-
  terface area. Details needed for further investi-
  gation are covered within the nodes and show
  up at once.

- **immediate navigation**
  Navigation is done within the interface. No
  scrolling or change of the interface metaphor is
  needed.

- **continuous orientation**
  A persistent overview position is fixed and the
  user gets never disoriented. In case of mutually
  occluding information, the map can be viewed
  from different directions by using the rotation
  handle. The user has no need to move around.

- **spatial coherency**
  User actions may change the appearance of ob-
  jects in the map, but they do not change the
  structure of the map. All relations inside the
  map stay where previously observed and brows-
  ing can be resumed easily.

These design intentions were developed while we exper-
imented with prototypes of various components of the sys-
tem.

5. The i-Disc in detail

In this section we will explain the various elements of our
topic map visualization using the already mentioned XML
topic map. The XML topic map’s ontology avoids multiple
inheritance for topics like almost all other ontologies, even
though multiple inheritance is supported by the XTM stan-
dard. We do not expect the need to use multiple inheritance
for our course ware topic maps, which reduces some of the
complexity of displaying these specific semantic link maps.

5.1. Topics and hierarchies

The concentric rings of our topic map visualization contain
topics according to their hierarchy level. The inner ring rep-
resents the top level of the topic hierarchy. Due to the struc-
ture and organization of topic maps, the top level contains
almost always typing topics of a generic type, like "topic
type" or "association role". In the middle rings, the next lev-
els of the hierarchy are displayed. Here, instanced and typing
topics are located, forming the grouping structure. For our
example, there is just one middle ring. However, we found
that there are typically only three to five hierarchy levels for
topic maps currently publicly available, resulting in one to
three inner rings. In the outermost ring, we find the leaf top-
ics as endpoints of the topics’ hierarchy.

Our topic map display conveys a lot of meta information:

- the size of the map
- the number of hierarchy levels
- the approximate number of topics on various levels
- the dominating concepts.

An experienced user is able to quickly spot the main aspects
by taking a closer look at the middle ring. In the given exam-
ple, two types, “product” and “vendor”, cover three quarters
of the circumference (see Figure 2). Only a few small sec-
tions are contained in the fourth quarter. It is easy and correct
to interpret this visual information as a representation of the
main subjects of the map: the topic map deals mainly with
XML software and the corresponding distributors.

![Figure 2: The two typing topics "Products" and "Vendors"
on the inner ring in an activated state. The degree of pro-
trusion of an activated topic is determined by the number of
associations it shares.](image)

Touching the topics with the cursor displays details like
associations, parent-child relationships, base names, and oc-
currences. This action is implemented in a very responsive
way, to let users quickly gain an impression of the underly-
ing relations by sweeping over the rings.

Selection of a topic is achieved by simply clicking on it.
This behaves similarly to the touching action, but changes
the color of the selected segment and makes the highlighting
of the topic and its children permanent, as well as the pre-
sentation of its name. This allows the user to mark the topic
for later investigation.

© The Eurographics Association 2005.
5.2. Associations

The outermost topics are extended according to their "relevance". Relevance here is derived from the number of a topic's associations. It is assumed that a topic which is associated twenty times is more important within the map's ontology compared to a topic which may have only ten associations. Therefore, it is visualized "longer" when it is activated. Topics can be locked by a mouse click. In this state, yellow lines connect the associated topics to static association type labels. These labels appear as a pair at the left and right hand side at the bottom of the screen, displaying the association type and association role (Figure 3). The yellow lines are updated interactively during map rotation. When too many associations clutter the screen, the labels can be folded, and the yellow traces disappear temporarily. Since the number of associations determines a topic's size, it is important to switch off uninteresting association types while exploring the map.

5.3. Occurrences

On the top right-hand side of the visualization area, the occurrences are displayed in a blue box (see Figure 4). Usually, such occurrences are URIs. Clicking on them opens a browser window and directly leads the user to the information linked to, allowing to directly discover the content for which the topic is only a proxy.

5.4. Search and Logs

Sometimes users do not want to explore the topic map directly, but they are interested in the location of a specific topic. For these cases we support regular expression queries entered in the input field at the bottom of the i-Disc window. As a result of these queries the corresponding topics with matching base names are highlighted. These results, as well as a history, are displayed in the logging and history area above the input field.

6. Implementation

Here we describe the technologies and architecture used in the i-Disc environment, as well as our approach for computing the geometric representation from the topic map data.

6.1. Software Architecture

The i-Disc is implemented using a VRML plug-in embedded in a HTML page. Using this platform, full hardware 3D acceleration can be achieved while still keeping open standards. Two components work on the server side: a Perl script generates an initial VRML file based on the current topic map, which is kept as long as the topic map does not change. This VRML file is transferred to the client machine only once at the beginning of a session. Then, an XMLRPC server handles online queries. This combination was chosen to keep the system as responsive and lightweight as possible: rendering an initial state once reduces startup time to a minimum. Querying for occurrence details or base names on the server side avoids distributing the complete xml topic map to the client.

6.2. Structuring the Data

Our Perl script computes an i-Disc based on the following steps:

1. The number of hierarchy levels is calculated from the XTM file.
2. The number of topics in the outermost circle is determined.
3. All circles are geometrically divided by the number of topics. Typing topics are positioned and scaled on the inner rings to cover approximately the same circle segment as all of their outer children. Topics on inner rings without children are displayed with a certain minimal width, which results in a slight shift of some children’s segments with respect to their parent segments. We chose this approach to make full use of the total circumference of each ring instead of leaving some segments unoccupied in the outer rings.
4. The script calculates geometries for the topics in their active and inactive state, depending on the number of associations. The position and the extension of activated topics contains redundant information: the number of the topic’s associations. This redundancy was added on purpose to keep the structure of the map simple, easily readable and highly characteristic. Sorting the topics in any other order has led to a confusing contour of the rings.

Our topic maps had neither associations between topics in the inner circles nor associations between topics on different hierarchy levels. Such associations encode generic association concepts of the underlying ontology.

7. User Observations

We observed people using the i-Disc while giving demos and during a planned informal user observation session. In observing users, we hoped to assess whether our approach and the system as a whole were natural and intuitive for people. Our sixteen subjects were colleagues, administration staff, and students from our department, who are all more or less experienced computer users.

We attempted to answer the following questions during these observations:

- Do users accept this kind of presentation?
- Do users accept the interaction model?

In summary, it can be said that users accepted the presentation without any hesitation. The visualization of hierarchies as concentric rings as well as the rotation of the model was reported as being "natural". Likewise, the interaction with the map by hovering over the topics to browse through the map and activating/locking topics for further investigation seemed logical to the users.

During the begin of the session, users were asked to complete several interaction and selection tasks, such as rotating the map, activate topics, describe and trace relations. The next task was to do a text query for the topic "product", an action that highlights a typing topic named "product" in the middle ring. The users were requested to estimate the number of depending topics as well as to name the most "important" one. This request implies that the user activates the topic, rotates the map, checks for the most extended topic, hovers over it with the mouse, and reads the base name. After a short hesitation period, most users solved the whole task without any further instruction.

After working with the topic map model for about ten minutes, the users were asked to close the application and sketch the i-Disc. When the sketch was done, they were asked to point at the topic "standard". Generally, the users had no problem to sketch the map, and mention some structures they remembered. Most of the users were able to point at the topic "standard" in their own sketch. This indicates that they quickly acquired a model of the coarse spatial layout of the map and the high level content. Our experiments also revealed that users mainly remember the three subdivided rings, but they are not consciously aware of the six hundred items they are effectively looking at.

The arrangement of the association labels was less successful: even though the users had no problem with the concept of associations as arcs between topics, the connection to the association labels at the bottom left and right side of the screen was perceived as another association. In addition, the fact that occurrences were placed unrelated to the topics at the upper right side of the screen was not always directly memorized.

8. Discussion

We have chosen a three-dimensional layout to visually separate the semantically orthogonal information contained in the association link structure from the basic tree-structured layout of the topic landscape. The navigation of the scenario by rotating the map provides motion parallax, which facilitates the understanding of the three-dimensional structure. In addition perspective rendering of the map provides a natural focus and context display, which is easy to understand and interpret.

In a first approach, the i-Disk was designed with free spatial interaction. A user could freely zoom, pan, pitch, and yaw. This navigation soon turned out to be exhausting. The
development of the prototype continued with offering different viewpoints located next to each topic. From a design point of view, this offered the interesting aspect of “viewing the topic map from different perspectives”, but in experiment this led to serious loss of orientation. In a next step, the transition was animated to exploit inertia effects as suggested in the concept of cone trees [RMC91]. This appeared as an enhancement but was still not convincing. The best navigation interface turned out to be a handle for manually rotating the map.

From our experience we found our visualization to work particularly well if an overview of the topic map landscape is required or desirable at all times, and only some of the associations need to be seen simultaneously. A clear limitation of our current approach is that it does not scale well beyond a some hundred nodes and possibly up to a few thousand nodes. Our XML topic map contains 600 nodes distributed over three hierarchy levels. This seems to be a good match for our approach. It is conceivable that topic maps with four or may be even five hierarchy levels could be represented in a similar way in particular considering the trend to higher resolution screens. We do not think that our approach will work well for general graphs. In our work we exploited the specific domain knowledge about topic map graphs by separating topics and associations into different spatial dimensions, which untangles the often complex graph structure.

Based on our recent initial experience with the construction of topic maps from well structured and marked up course material, we expect the i-Disc approach to allow students to quickly navigate and access the represented information. The initial presentation of the structure of the course material and the domain specific terms and concepts, serves to get an orientation within the domain. The subsequent exploration of the map, using both, the orthogonal layer of associations as well as the text search, should support the construction of a mental model of the domain.

9. Conclusions and Future Work

We presented a new visualization technique for medium sized topic maps that integrates seamlessly into existing web based environments. The elementary circular design allows quick perception while interactive exploration allows access to details on demand. The precalculated three dimensional model is optimized for comprehensibility and responsiveness. The topic hierarchy and the association links, which are semantically orthogonal structures, are separated into orthogonal spatial dimensions resulting in an uncluttered visual representation. An informal user study confirmed the natural and intuitive access for users.

As a whole, the i-Disc follows our design intentions, yet in some aspects there are improvements possible. For instance, the restriction of navigation and movement limits the number of topics that can be visualized in practice, especially when the display resolution or pointing device resolution (e.g. touch pads) is a concern. The spatial coherency in combination with the desired elementary structure forces precalculation of the map, which in turn makes dynamic changes by e.g. editing very difficult. In addition, the i-Disc does not deal with multiple inheritance of typing topics, which seems to be a minor issue.

However, there are still a number of remaining challenges for future work. A full evaluation of the introduced technique needs to be performed to determine the effectiveness in understanding information in comparison to other topic map visualization systems. We also need to explore the applicability of our system to the real-life course ware content during the daily use of our e-learning platform. We are also interested in exploring our tool for different application domains.

The display of relations by using arcs has potential for improvement. An idea might be the application of a linkmap concept [BEW95] to reduce occlusion where the arcs set off. Another idea is to reduce the number of items displayed by e.g. simplifying many association arcs into a single texture and displaying the arcs themselves on demand only. Employing a texture would offer many opportunities to sum up trends without cluttering the screen. In particular appropriate geometric and semantic low pass filtering can be employed to avoid geometric or semantic aliasing.

A larger step is the introduction of an interactive editing functionality, for both a topic map designer who develops the semantic structures and for topic map editors who may add, delete, and edit topics, associations, and occurrences.

10. Acknowledgments

This work has been funded and is based on work funded by the BMBF (Bundesministerium fuer Bildung und Forschung), in the program ”NMB” (Neue Medien in der Bildung - new media in education), under the project ”Medienquadrat - Medien fuer Medien” at the Bauhaus University Weimar.

References


[Bar02] BARTA R.: Xtm::base is a [perl] topic map engine that supports xtm 1.0., 2002.


