GTM*alpha* – Towards a Graphical Notation for Topic Maps

Hendrik Thomas[1], Tobias Redmann[2], Maik Pressler[2], and Bernd Markscheffel[2]

[1]Knowledge and Data Engineering Group,
School of Computer Science and Statistics, Trinity College Dublin,Ireland
hendriktho@googlemail.com

[2]Chair of Information and Knowledge Management,
Faculty of Economic Sciences, Ilmenau University of Technology, Germany
{tobias.redmann,bernd.markscheffel}@tu-ilmenau.de
{maik.pressler}@googlemail.com

Abstract. In the last years several drafts, recommendations and concepts for a graphical notation for Topic Maps have been published, but till today no graphical notation is generally approved and used in the Topic Maps community. In this paper we present GTM*alpha* as a conceptual new notation for a graphical representation of Topic Maps. Our objective is, to provide a practical usable notation, which allows a complete, consistent as well as easy to use graphical representation of any given topic map draft. GTM*alpha* provides a domain as well as a subject centric view and most important it considers the unique characteristics of the Topic Maps paradigm. This paper serves as a user oriented GTM*alpha* manual for ontology designers, domain experts as well as users.

1 Introduction

Modeling a Topic Maps ontology is generally a complex and time-consuming process which involves many different actors [1,2]. To support the necessary discussion and to demonstrate modeling options a graphical representation of a topic map draft can be helpful [3, 4]. Using a standardization graphical notation for Topic Maps (GTM) ensures that involved ontology engineers interpret a topic map graphic correctly and uniformly, which is especially important for a collaborative modeling process [5, 6].

In the last years several drafts, recommendations and concepts for a GTM have been published, but till today no graphical notation is general accepted and used in the Topic Map community[3-7]. Recent studies showed that the common trend to reuse and adapt existing graphical notations from the field of data [3]and knowledge modeling [6, 7] is not suitable for this task [8, 9]. Evidences could be identified which indicates that none of these existing notations [1] could be used for a graphical representation of Topic Maps without adjustments or extensions. However, reusing a existing graphical notation in a different manor as original intended, forces the user to relearn the notation elements and rules as well as significantly increases the risk of misinterpretations. Based on the conducted evaluation we concluded that the creation of a conceptual new graphic notation for Topic Maps is inevitable. A suitable GTM from a pragmatic as well as from a research point of view has to consider general requirements for modeling as well as the specific characteristics of the Topic Maps paradigm [13, 2]. Existing notations can't provide this, because they have been designed for a different purpose and domain.

As a consequence of this insight, we will present in this paper a new conceptual notation draft for the graphical representation of Topic Maps. We designate it as GTM^{alpha} to highlight our objective to provide an every-day usable notation, which allows a complete, consistent as well as easy to use graphical representation of any given topic map draft. This paper serves as an user oriented manual for ontology designers and users who need a graphical representation of a topic map.

In section two we will explain in detail how the GTM^{alpha} should be used and why specific design decisions had to be made to allow a complete and consistent representation of a topic map according to the Topic Maps standard 13250 [2]. In section three we will explain the two pre-defined views of GTM^{alpha}, domain view and the subject centric view. The paper concludes with a summary and an outlook.

2 Manual for GTM^{alpha}

From a scientific point of view a GTM has to allow a complete representation of a topic map. Furthermore the graphical representation should be as consistent as possible to ensure a unambiguous interpretation for users. As a result every graphical representation based on GTM must be transformable into a valid

---

[1]   During the conducted evaluation the notation of concept maps [10], frames [11] and entity relationshipc [12] model have been analyzed in order to determine their suitability to represent a topic map draft under the condition that all individual notation rules are followed by the book.

formal topic map (e.g. XTM, LTM [2]) without loosing any information or adding additional elements. A GTM should also provide different views on relevant aspects to simply and support understanding. Considering a cost-benefit-ration, the amount of effort which is necessary to create a graphical visualization and using the notation should be as low as possible. This has obviously strong impact on the amount of notation elements, the complexity of these elements and modeling rules. Considering the required principle of clearness of a model [8, 14], a graphical Topic Map representation the quality of layout should be high and therefore the GTM design should support the user in the creation of a clear and easy understandable presentation.

Finally, a GTM must reflect the unique characteristics of the Topic Maps paradigm [2, 13]. Essentially two features must be taken into account. First, we have to consider the fundamental rule of Topic Maps: one topic per subject. As a result in a graphical representation a subject should be modeled by exact on element. If this is not possible the notation must provide suitable indicators, which makes clear, that two elements represent the same subject [13, 2]. Furthermore in Topic Maps all kind of types (e.g. topic types, association types, association role types, name types and occurrence types) are represented by topics. As a result a topic can act as a class but at the same time as an instance. This quite unique feature must also be considered in the notation.

Beside these criteria, from a pure pragmatic point of view, a GTM is only suitable for the needs of the Topic Maps community, if it allows to draw a Topic Maps draft fast and easy – with a bad handwriting using a half-full pen on a dirty white board – and an foreign ontology expert is still able to grasp the structure and the elements of the topic map draft correctly and consistently. This is what we need to support the modeling process and communication. Turning to the actual design of a GTM, two observations can be made regarding to these requirements.

First, Topic Maps is clearly a special type of a semantic network consisting of topics and associations representing the relationships between subjects [13]. Consequently the design and main structural principle of the GTM should be network oriented to provide an adequate representation. Second, topics as well as associations are connected to multiple information elements, e.g. a value of a base name, a URI for the subject identity as well as topics acting as association types [2]. From the visualization point of view the resulting network consists of a wide variety of elements, which must be easily identifiable for a user [14].

Coloring the elements might be a good idea. However, typically only a limited number and quite different colors are available, if someone needs to draw something on a black or white board. Even more important is, that a lot of graphical representations of topic maps are made for research publications, like

this one. Those publications are traditionally limited to black and white prints, thus the usage of color is not recommendable for a GTM.

As an alternative we could use, like in previous GTM drafts [5], different node shapes to distinguish Topic Maps constructs. Generally those shapes should be limited to simple geometric forms like ellipses or rectangles to ensure that a topic map draft can be drawn fast and easily. However, the representation of any Topic Maps construct by an individual shape could lead to misinterpretations by the user and to inconsistencies regarding the TMDM[2]. For example, a topic could be symbolized as an ellipse and a scope as a rectangle. This would be suitable to distinguish the elements, but would indicate for a user that the nature of the elements is different.

In fact, both are topics and only one of them acts as a scope in order to define a specific valid context for a Topic Maps construct. Topics can be involved in different roles, e.g. acting as a topic type, scope, association role type, etc. We considered these circumstances in the GTM^{alpha} draft by representing topics with a unique shape and additional symbols indicating which role a topic is playing in the specific construct. Using these approach, a topic is represent by one shape but it can play more than one role, each clearly identified by the specific symbol.

Based on this thoughts the following subsection explains in detail the symbols and notation rules of GTM^{alpha}, which are used to represent the individual topic map constructs. In all further images the topic map elements are labeled. Keep in mind, that these labels are only for education purpose and are not needed in the final graphical visualization of Topic Maps.

2.1 How to represent topics and types relations?

Since the earliest days of Topic Maps [13] a topic has been described as a node in a network and was therefore often graphical represented as an ellipse. In GTM^{alpha} we will continue this good old tradition. Optional a text can be included in the ellipse, which is interpreted as the common name of the topic.[2] One of the most important constructs in Topic Maps are types, which are used for categorization. In GTM^{alpha} the topic type relation is represented by a arrow line connecting the topic acting as a type and the instance topic. By definition the arrow head points at the topic type element. This arrow line symbol is used in the whole GTM^{alpha} notation to indicate type relationship of all kinds, e.g. the assignment of an association type, role type, name type as well as occurrence types.

---

[2]   Please note, that the text is actual a shortcut for a base name of the topic without any scope or additional information. More information on this special case can be found in section 2.2. Topic Names.

It is common that a topic type has more than one instances. However, the fundamental rule of Topic Maps demands, that a subject is represented by exact one topic, can make the creation of a suitable layout quite difficult. Because every instance has to be connected with the one topic representing the type in order to be consistent to this rule. As a solution we propose the following approach. Topics acting as types can be drawn more than once, depending on the best position in the graphic for a high quality layout. To clearly indicate for a user, that these multiple topics represent the same subject, inside the type topic a small rectangle containing an ID must be added. All topics representing the same subject must share the same ID. In the formal representation this ID can be preserved as an item identifier, which forces an application to merge the multiple topics automatically [2].
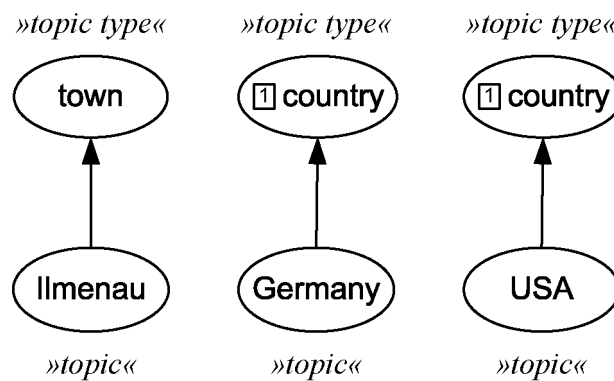

Fig. 1. Topic with topic types

Fig. 1 shows three topics with their according topic types, e.g. Ilmenau is an instance of the class town. As you can see the topic type is represented as topics and only the arrow lines indicates that the topic is used as a type. Germany and USA are both instances of the topic country. The topic type country is drawn twice, but because they share the same ID "1" it is clearly visible for a user that both topic types elements represent the same subject.

2.2 How to represent Topic Names?

As one can see in Fig. 2 the square symbol is used to model topic names. A topic name consists of a base name and optional multiple variant names with specific data values [2, 13]. For a consistent visualization values must be strictly distinguished from topics representing a subjects [8]. Therefore in GTM*alpha* all

values in a topic map are graphically represented by a rectangle containing the specific value.

A filled square symbol indicates that the value inside a rectangle represents a base name, e.g. the string "Ilmenau" is a suitable name for the topic Ilmenau. Origin from such a base name any number of variant names can be attached, e.g. "IK" and "ILM" as abbreviation for the town name. An empty squares symbol is used to indicate that the value represents a variant name.
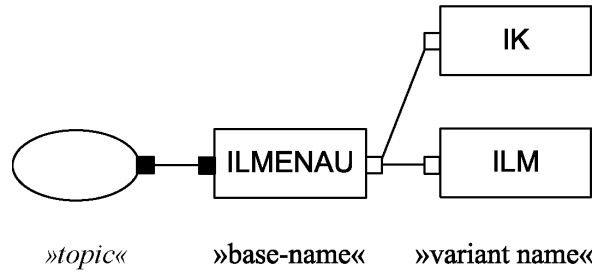

Fig. 2. Base name and variant names

To assign one or more scopes to a topic name, simply attach the topic representing the valid context to the connecting line. A crosshairs was chosen as symbol according to the character of a scope which pin points a valid context. The way to assign a topic name type is similar: simply draw an arrow line to the topic representing the topic type from the connecting line.

Fig. 3 shows how these descriptive elements for a topic name can be modeled. In this example the scope German is assigned to the base name "Ilmenau" as well as the topic name type "official name".
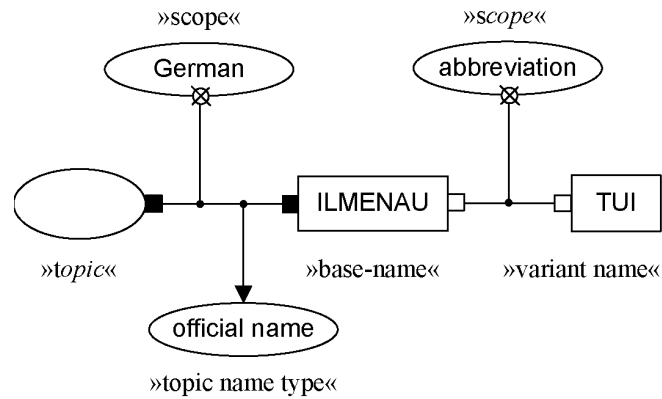
Fig. 3. Topic names with scope, topic name type and data type

In contrast to previous graphics the topic node is drawn empty. This results from a special rule of GTM^{alpha}. The text inside of a topic ellipse node is interpreted as a base name of the topic without any scope or additional information. This is suitable because mostly it is not necessary to draw the complex topic name construct. In many situations it is more helpful to show topics with a standard label. As a results the two graphical representations in Fig. 4. are identical.
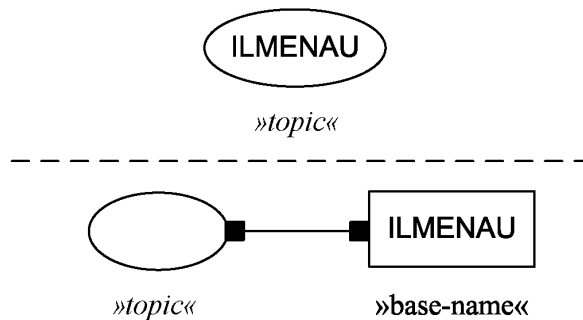

Fig. 4. Topic names shortcut

2.3 How to represent a subject identity?

To symbolize the subject identity of a topic in a graphical representation, we choose a symbol in the shape of the number eight in dependence to the

mathematically symbol for infinity. This should highlight the inescapable bound between a topic and a subject. In addition this symbol is easy to draw and provides an unambiguous indicator. This subject identity symbol can only be used on connections between a topic and a value box containing URI's or in some cases URL's. To indicate that the text inside a value box is a URI, simply underline it. In case of a subject locator, where the URL points to the digital resource the topic is representing, a filed eight symbol is used.
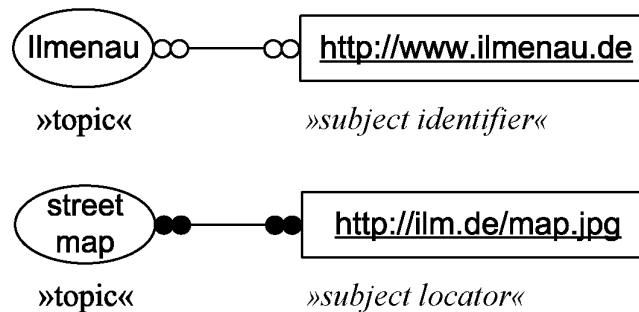

Fig. 5. Subject locator and subject identifier

If a subject identifier needs to be modeled, in terms of that the referred resource acts as an subject indicator, an empty eight symbol is used. This choice was made, because, based on the experiences of the authors, the majority of topics are not direct addressable and therefore subject identifier are more common and an empty eight symbol is a little bit easier to draw that a filled one. The Fig. 5 demonstrates how this notation element should be used, e.g. the URL *http://www.ilmenau.de* is assigned as subject identifier to the Topic Ilmenau and the URL *http://ilm.de/map.jpg* as a subject locator to the topic street map.

2.4 How to represent occurrences?

To represent occurrences in GTM^{alpha} we choose an empty circle as an equivalent to a two dimensional version of the common database symbol. This symbol can be used on connections between a topic and a suitable value element only. An internal occurrence us represented by a rectangle representing the piece of relevant data. For an external occurrence the rectangle must contain a valid URI and must therefore be underlined. The Fig. 6 shows how occurrences are visualized in GTM^{alpha}. In this example the number "98693" is assigned to the

topic Ilmenau as some piece of relevant data. Additional the web page *http://leipzig.de* is assigned as an external occurrence to the topic Leipzig.
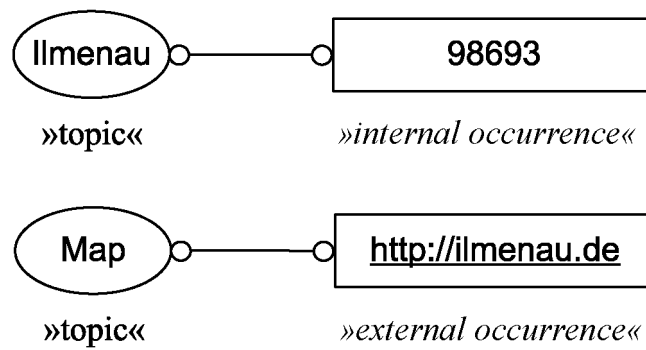

Fig. 6. Internal and external occurrence

Similar to the approach used for topic names additional information can be modeled around the occurrences. Fig. 7 demonstrate how an occurrence type can be assigned by drawing an arrow line to the topic representing the occurrence type, e.g. zip code. Also a scope topic is assigned to the occurrence by using the scope symbol, e.g. limiting the zip code to the context Germany.
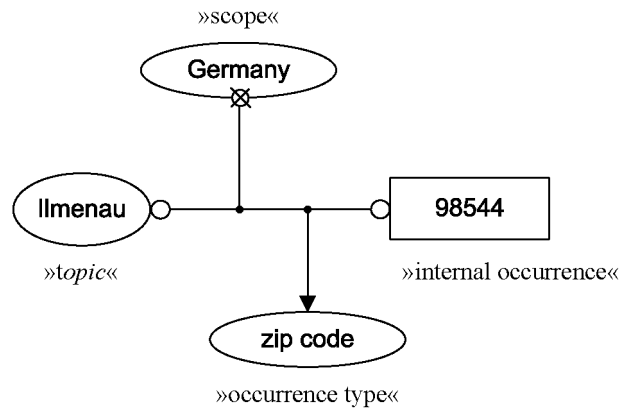
»scope«

Germany

Ilmenau   98544

»topic«   »internal occurrence«

zip code

»occurrence type«

Fig. 7. Occurrence with type and scope

2.5 How to represent an association?

The association is the second fundamental element of Topic Maps, it can be represented in GTM^{alpha} by drawing a line between the associated topics. The additional descriptive elements of the association can be attached directly to the connecting line, which represents the association. In the center of this line a arrow line can points to the association type. The same concept is used to point to the role types for the topic players. The arrow line should origin close to the corresponding topic player. For a better distinction the connecting node for role types should be smaller than the node for the association type. Additional one or more scope elements can be assigned to the association by drawing a line from center to the scope topic combined with the scope symbol. Fig. 8 demonstrate these recommendations for drawing associations, e.g. the topic Ilmenau plays the role "part" and the topic Thuringia plays the role "whole" in the association "is-part-of".

During the modeling process an association do not necessary posses all of these information. Especially in the early modeling stages a modeler often needs only to represent the fact, that there is some kind of association between two topics. In later modeling stages the association is refined and role types as well as the association types are incremental added. Therefore in the GTM^{alpha} it is allowed to draw an association without role types or without role types and association types.
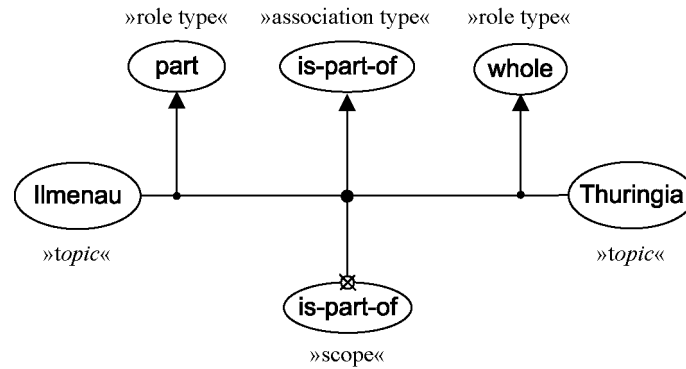
Fig. 8. Association

2.6 How to represent reification?

Sometimes it is necessary to model a statement about other Topic Maps constructs. Such meta-knowledge about the ontology itself can be represented by reification. In terms that a topic represents another topic map construct, like association, a base name, etc. In GTM*alpha* this can be easily represented by drawing a dotted rectangle around the construct which should be reified. Additionally a dotted line from this rectangle must be drawn to the topic which shall represent the specific construct. As you can see in Fig. 9 the filled eight was chosen as symbol, to indicate that the subject represents the specific topic map construct.
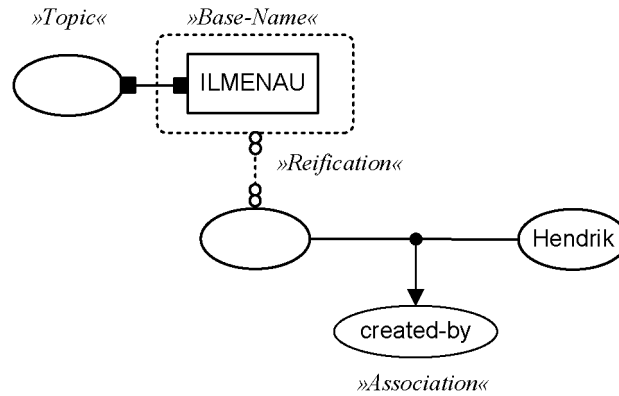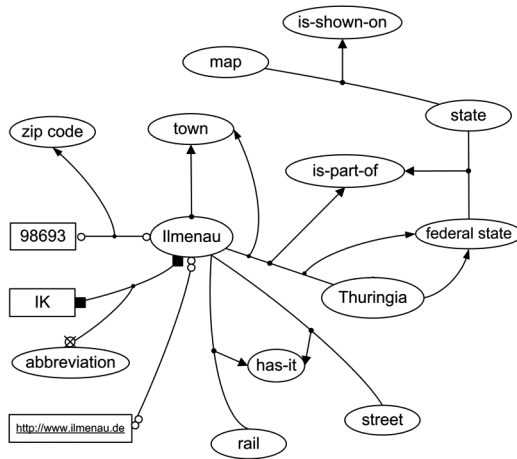
Fig. 9. Reification

3 Domain and Subject Centric Views

The comprehensibility of a graphic model depends on its complexity and its volume. Especially helpful are different views in which only selected aspects of the complex model are visualized. This can reduce the cognitive workload of a user. In the GTM^{alpha} two views are predefined. First, we have the so called *domain view,* which provides an overview of the whole or selected fragments of the modeled domain for a user. A graph oriented structural layout was chosen, in order to highly especially the relationships between the topics in order to allow a user to grasp the big picture. All elements in the domain view are drawn according to the presented GTM^{alpha} notation rules. In Fig. 10 a topic map draft is displayed in LTM and the corresponding GTM^{alpha} domain view.

**domain view of the topic map draft in GTM^{alpha}:**   **topic map draft in LTM:**


```
/*topics*/
[Ilmenau:town = ``Ilmenau''; ; /German
            = ``IK''/abbreviation)
 @``http://www.ilmenau.de'']
[Thuringia: federal state]
[street: traffic connection]
[rail: traffic connection]
[state] [map]
/*associations*/
is-part-of (Ilmenau:town,
Thuringia:federal state)
is-part-of (federal state, state)
is-shown-on (state, map)
has-it (Ilmenau, street)
has-it (Ilmenau, rail)
has-status (Ilmenau, county seat)
/*occurrence*/
{Ilmenau, zip code, [[98693]]}
```

Fig. 10. Domain view of GTM

In the topic map developing process very often the discussion is limited to a specific topic rather than to the whole complex topic map network. To provide users a detailed view on the modeled knowledge on a relevant subject in GTM*alpha* the so-called *subject-centric view* was pre-defined. In this view all topic map elements are drawn according to the presented GTM*alpha* notation rules, which ensures a cross-model consistent representation and visualization. The layout of the subject-centric view is tree oriented in contrast to the network orientation of the domain view. This enables a well-structured and easy-to-grasp presentation. In the subject centric view the topic of interests is placed at the top and all modeled information are arranged in an explore tree beneath it. We recommend to start with the subject identity, followed by all topic types, topic names, occurrences and at the bottom all association the topic is involved in. Fig. 11 shows the subject-centric view for the central topic Ilmenau of the previous topic map draft. In the end in both views the same information are displayed, only different default layouts are used in order to highlight a specific aspect, e.g. overview or detailed information. Other views are possible.

subject centric view of the topic map draft in GTM^{alpha} :
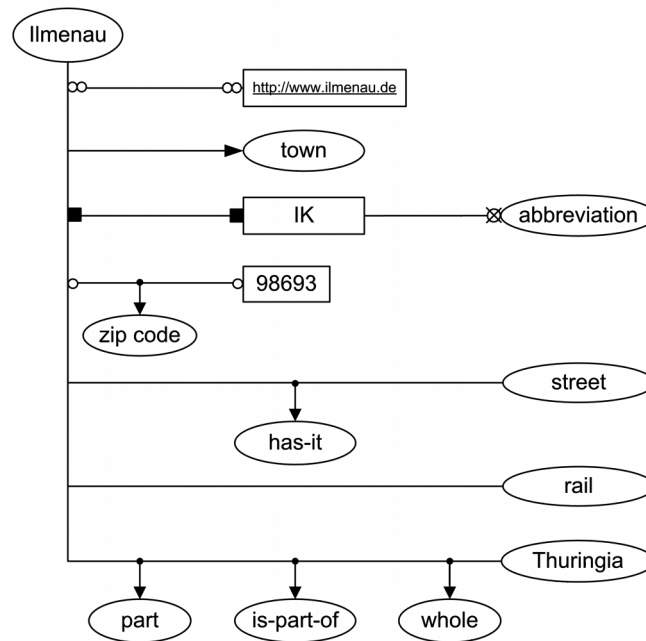
Fig. 11. Subject centric view of GTM

The authors have been using the GTM*alpha* draft successfully for some time. Based on these experiences we identified some best practice recommendations. First the size of the topic element is not pre-defined and can be changed. It is suitable to increase the size of a topic to draw attention to it, e.g. in a discussion or in documentation. The connection line can be strait or curved depending on the layout but crossings of lines should be avoided. The lines for association should be drawn thicker than lines between Topic Maps constructs. Be aware that all symbol and elements can be rotated, without loosing their ability to identify an element unambiguously.

4 Summary and Outlook

To support discussion and documentation of the ontology modeling process, we presented in this paper a new draft for a graphical notation for Topic Maps. We showed how GTM*alpha* should be used to enable a complete and consistent

graphical representation of any given topic map draft according to the TMDM [2]. The amount of effort to learn and use the notation can be considered as moderate. Only few and simple shapes as well as only few notation rule have to be considered.

As the name GTM indicates, the focus was on the design of a practical usable notation. However, it is still a draft and only the broad usage of GTM in the Topic Maps community can lead to a final answer of the question: is GTM really suitable for representing topic map drafts. Overall a GTM has been missing for so long and with this proposal we hope to start a fruitfully discussion, which will finally lead to a official standardized GTM.

References

1. Garshol, L. M.: "Towards a Methodology for Developing Topic Maps Ontologies", in Maicher, L, Siegel, A, Garshol, L. M. (eds.): *Leveraging the Semantics of Topic Maps – Second International Conference on Topic Map Research and Applications, TMRA 2006, Leipzig, Germany, October 11-12, 2006*, Berlin Heidelberg New York, Springer (2007) 20–31

2. Garshol, L. M., Moore, G.: *ISO/IEC JTC1/SC34, Information Technology*, `http://www.isotopicmaps.org/sam/ sam-model/` (2006)

3. Gulbrandsen, A. D.: *ORM vs UML for Topic Maps*, `http://www.informatik.uni-leipzig.de/~tmra05/PRES/AG.pdf` (2005)

4. Lee, J.: *Graphical Notation for Topic Maps – Presentation*, `http://www.jtc1sc34.org/repository/0704.pdf`, Seoul, Korea (2005)

5. Henriksen, I.: *Graphical Notation for Topic Maps, Draft 1.2*, `http://cafe.teria.no/ingeh/files/6/13/GTM.pdf` (2006)

6. JTC 1/SC 34/WG 3 ISO/IEC: *N0883 - Draft GTM 13250-7 – Requirements*. `http://www.jtc1sc34.org/repository/0883.htm` (2007)

7. *ISO/IEC ISO 13250-7: GTM (Graphical Notation)* `http://isotopicmaps.org/gtm/` (2008)

8. Becker, J., Rosemann, N., Schütte, M.; "Grundsätze ordnungsgemäßer Modellierung", in *Wirtschaftsinformatik*, 5 (1995) pp. 435–445

9. Schütte, R.: Die neuen Grundsätze ordnungsmäßiger Modellierung: Paper zum Forschungsforum 97, Müunster (1997)

10. Noak J. D., Cañas A. J.: *The Theory Underlying Concept Maps and How to Construct Them*, Technical Report IHMC CmapTools 2006-01 Rev 01-2008, Florida Institute for Human and Machine Cognition, `http://cmap.ihmc.us/Publications/TheoryUnderlyingConceptMaps.pdf` (2008)

11. Reimer, U.: *Einführung in die Wissensrepräsentation: Netzartige und Schema-basierte Repräsentationsformate*, BG Teubner (1991)

12. Chen, P. P.: "The Entity-Relationship Model – Toward a Unified View of Data." in: *ACM Transactions on Database Systems*, 1 (1), (1976) pp. 9–36

13. Pepper, S.: *The TAO of Topic Maps - Finding the Way in the Age of Infoglut*, `http://www.ontopia.net/topicmaps/materials/tao.html`, Norway (2002)

14. Bennett, W. S.: Visualizing Software, a Graphical Notation for Analysis, Design and Discussion: A Graphical Notation for Analysis, Synthesis and Discussion, Marcel Dekker Ltd, Binghampton New York (1992)