

Temporal Qualification in Topic Maps

Lutz Maicher and Christoph Teichmann

Topic Maps Lab Leipzig

Abstract. This paper will introduce a number of predefined elements for the topic maps data model. The paper will start with a short introduction of the level and possible context of these elements. Then the problem the elements proposed here are supposed to solve will be described. This will be the problem of adding temporal qualifications to a topic map. A set of criteria for the quality of a possible solution will be given and possible solutions will be evaluated against these criteria. This will lead to the final proposal for adding temporal qualification to a topic map.

1 Predefined Constructions for Topic Maps

The topic of this paper can be associated with the larger goal of giving a set of predefined constructions for topic maps. These constructions could be used for modeling tasks that occur repeatedly in creating a topic map. With such predefined elements it would be possible to use their predefined semantics in order to allow for automatic reasoning with given topic maps. Many other generic interfaces are possible for topic maps.

An example for an already existing predefined construction would be the supertype-subtype relation. This also gives an idea of the relation to the topic maps data model these predefined constructions should have. It should be an additional level using the topic maps data model but not adding any elements to the model itself. This domain of predefined constructions will be referred to as *pred* in this paper for lack of an established domain.

Such constructions should be defined by giving the topic map constructs to be used and the PSIs to be used. With these elements given it is easy to identify predefined constructs. The PSIs ensure that when two topic maps which both use these constructions are merged a connection is established through these constructions.

One possible objection to this idea could be the fact that in a certain sense the semantics of these constructions are externalized. When and how they are used would be defined by a standard that establishes these constructions. However it would still be possible for any topic map creator to add any information that would be necessary for computations or other tasks. Another objection may be that such constructions may be overkill for certain tasks. They would need to be very general to apply to many tasks. So they may become complicated if they are applied to a special problem. This objection can be defused if interfaces for the automated inclusion of these constructions would be provided. In this case they could be added without much additional work to be done.

In order to create such a set of predefined constructions it will be necessary to identify different problems to be solved by these constructions and give the constructions which would best solve those problems. In order to establish which is the best solution to a problem it will also be necessary to develop criteria for the evaluation of different possible constructions for a problem.

This paper will introduce one such problem: temporal qualifications for topic map constructs. This will give a first example of how such a problem could be handled.

2 Introduction

First to clearly define the problem which will be discussed in this paper the following considerations will be helpful. It is an recurring task to model temporal relations within a topic map and then to connect the modeling of the temporal domain with other information in that topic map. Examples for this may be including the birth dates of people which are modeled in a topic map, adding dates to clarify the date at which some information included in the topic map was gathered, dating events or just stating when a something started to be the valid name of an organization and when it ended to be the valid name for the organization.

Each of these problems seems to have a natural solution in the topic maps framework and all of them are quite different in terms of the mechanisms used. Still this paper will argue for a general standard for modeling temporal information. This decision will be justified in the following chapters.

This paper will not discuss the way in which the dates should be modeled. This is not to say that there are no interesting and important questions concerning the modeling of the temporal domain, or that there should be no effort to create some standards for modeling this domain. But as both problems can be clearly separated this problem will not be addressed here and will be left for further discussion. It is assumed however that there is some solution for modeling the temporal domain in place.

3 Why Introduce a Standard?

Returning to the problems given in the introduction it will be helpful to discuss what would be the natural way of solving the problems given in the introduction. This will help in introducing the different types of solutions are possible and it will be a starting point for arguing that a standard for modeling time will be useful.

The first problem that was stated in the introduction was giving the birth dates of people modeled in a topic map. A very straight forward solution that comes to mind is creating a type 'born at' with the appropriate subject identifier and then create associations of this type. The next step is creating role types 'person born' and 'birth date'. Then associations of this type and with the introduced roles could be used to connect the people to their birth dates.

The second problem, giving the date at which information was gathered could be handled in a very different way. As the information that is gathered may be contained in names, occurrences and associations it would be natural to use reification for creating this relation to a date. This makes it possible to add information about the statement represented by these items. Then the date information could just be added as an occurrence of type 'gathering time' to the reifying topic.

Dating events could be handled in much the same way the birth date problem could be handled. The problem of dating the validity of some information over time lends itself to a very different solution. The topics representing dates could be used as scopes for the information that should be restricted in its validity. This could indicate that the context in which this statement is valid contains this date.

These considerations show a strength of topic maps that may be under appreciated. Topic maps are able to allow for very different ways of modeling the same domain and thereby allowing for very different perspectives on the same domain. The general problem of giving a temporal qualification to an element in the topic map was always the same but very different approaches seem reasonable for each of the special instances of the problem depending on how the temporal qualifications were supposed to be understood in the topic map. The considerations made here also show that different problems connected to temporal qualifications may best be solved by using different approaches to representing temporal qualifications.

So why should one want to introduce a standard for introducing temporal qualifications? The first reason that comes to mind is the use in general interfaces. These could be user interfaces for the generation of topic maps that provided predefined elements for topic maps or reasoning tools for topic maps. As introducing temporal qualifications is a recurring problem in creating topic maps advanced user interfaces should incorporate a way of easily integrating temporal information into a topic map. A predefined standard will help making this integration easier. Creating a standard for temporal qualifications would also be a step toward a general set of predefined constructions. Such constructions could also be used in information retrieval.

4 The Problems

In order to define a standard for integrating temporal qualifications it will be useful to identify some general classes of problems that should be solved using this standard. The first question in identifying these problems is how they should be characterized. In this paper the problems will be described by two different properties. The first property that will be used is the type of construct in a topic map or domain references which should be qualified with temporal information. This is of interest as a standard might have to use different approaches for different types of constructs or domain references. The second property that will be used to describe the types of problems is the type of information that is given

by way of a temporal qualification. This is of interest as it will be helpful to identify some basic types of temporal qualifications that can be used to give a general classification of the data that is added to the model.

The types of constructs will be separated into the classes *reifiable*, *scoped* and *topics*. These classes are chosen, as they include all topic maps constructs. They are also chosen because they are actual classes in the *TMAPI* [1]. They also represent a division in how additional information is added. Topics allow for occurrences, names and associations. Scoped and reifiable are classes that are even named by the way information can be added to these elements. Topics have to be chosen as a single class because they belong to no other general class in the topic maps standard. Note that these classes are not mutually exclusive. In fact the set of scoped constructs is included in the set of constructs that belong to the class of reifiable constructs.

The domain references which will be considered are subjects, statements and relations. As statements can be directly included in a topic map by using the constructs that fall in the class of scoped constructs. This further justifies adding the scoped class to the set of constructs discussed in this paper. Because statements can be included into a topic map this way they will not be regarded as an on class and the discussion will be limited to the class scoped. However domain references which will be discussed independently in this paper are the relations and subjects, as they can not be directly included into a topic map.

It should be noted that this choice of classes incorporates two important distinctions. First of all, the distinction between talking about the reference of a topic map construct and talking about the topic map constructs themselves. This distinction should be expressible in a topic map. This problem could be tackled by a standard for predefined constructs as these constructs could be used to designate elements of a topic map as meta information and it could be clarified how references to other elements of a topic map should be established. This may sound similar to the process of reification, but reification is originally designed to create topics for the elements the topic map constructs are actually *about* and not for the constructs themselves. The second distinction would be the distinction between qualifying information temporally and giving purely temporal information. An example for this could be the distinction between adding birth dates and telling when a certain relation was established. This distinction is represented by the distinction between topics and subjects on the one hand side and statements, relations, scoped elements and reifiable elements on the other hand.

The types of information will be divided into the classes introduced by the examples given so far. The birth date problem introduced temporal qualifications that model a certain aspect of the domain. This gives the first general class of qualifications the *domain information*. This type of information can only be given about the elements of the domain. Therefore this information can only be given for the classes *scoped*, *subject* and *relation*. The question how to model the information when data was collected introduced *meta information*. This kind of information should be clearly separated from the domain information. This

information will be given about topic maps constructs. Stating when a certain company name was valid adds *validity information*. This type of information is similar to scope. It models the context in which a certain construct or domain element is valid. However it should not be confused with scope insofar as there may also be validity information about a topic map construct. To give an example one could want to state that a certain association was only confirmed for a few days and afterwards it was replaced by a different association which could be relevant for a processing tool. Topics may also have such validity information associated to them, for a subject such validity information does not seem to make much sense, therefore this type of information is not considered for them. As a result of this discussion the following geometry can be used. Note that from this point on this paper will only refer to the class *scoped*, which is taken as a placeholder for statements as well as *scoped* topic maps constructs as they are considered to be the same in the topic maps data model [2].

	domain information	meta information	validity information
scoped	?	?	?
reifiable	-	?	?
topic	-	?	?
relation	?	-	?
subject	?	-	-

The standardization approach proposed here will try to give one solution for each of the valid fields in this table. '-' designates an invalid element entry in the table, because there is no problem for this combination of information and class. '?' is written in a field of the table if there is an actual problem for this combination of this kind of information and this class.

5 What to Expect

Before different standardization approaches are discussed it should be there needs to be a way of deciding what would constitute a good standardization approach as opposed to a bad one. Therefore a list of desirable qualities for a standardization approach will be given and justified.

1. The solution should be minimal.
 - The spirit of the topic maps data model includes the attempt to allow the for giving models of a domain that recreate an ontology with a minimal amount of of constructions while still keeping the created model human readable. This means that if there is a way to express what should be expressed by a construction in two different ways and one way introduces less new PSIs and/or demands the creation of less topic map constructs, then this approach should be preferred to the other one if both are equally good considering the other criteria. If two approaches differ greatly in the numbers of PSIs and constructs they use than the one using the smaller number should be preferred.

2. The different dates should be represented as topics and the temporal information should be connected to these topics.
 - This is a criterion of quality because one aim of using topic maps is the ability to merge different domain models in any easy way and this should result in more informative topic maps. To this end dates should be introduced as topics so the temporal qualifications from topic maps can be merged simply by merging the date topics. This paper will use the word date throughout and while the authors would themselves chose to have discreet time stamps and combine them to speak about intervals this paper itself should not be taken to favor any modeling of the temporal domain. So the proposal given here is intended to work with intervals as well as with dates.
3. It should be possible to add temporal information to any topic map construct and it should be possible to add temporal information about each domain reference.
 - This simply amounts to saying that each field in the table given in the last chapter should be filled. This is important as the standard should solve all general problems connected to introducing temporal qualifications in topic maps.
4. It should be possible to add further information to each temporal qualification.
 - When a temporal qualification is given it might be necessary to add some information to these qualifications. E. g. when saying when some information was checked it might also be useful to add the information who checked it and how.
5. It should be possible to add multiple temporal qualifications to a construct.
 - This requirement is needed simply because complex domains may include many kinds of temporal qualifications and it should not be necessary to use different approaches for each kind of temporal data.
6. No extension should be required for the topic maps data model.
 - The proposal is supposed to contribute to a set of predefined constructions for topic maps and these constructions should be an addition on top of the topic maps data model and not an addition to the topic maps data model.

With combination of those criteria and the table of temporal problems given it will be easy to evaluate different approaches to modeling temporal qualifications in topic maps. Approaches will be presented and it will be asked which fields in the table they can fill and how well they do in terms of the different criteria. Then the best approach according to this evaluation can be chosen. Some of these criteria could also be seen as criteria for the evaluation of predefined constructions in general.

6 Possible Solutions

This section will introduce four approaches to including temporal information into topic maps. All of them are a very straight forward use of possibilities of the topic maps data model. The goal here is not to introduce new and inventive ways of representing dates but rather to evaluate the straight forward approaches and then decide which ones would be best for a standardization. This is a necessary step on the way to a standardization proposal.

6.1 Scope

In order to model temporal information by using scope all that needs to be done is create date topics and use them as scopes. This approach seems especially natural for the upper, right corner of the table that was introduced earlier on. 'x' is used to mark problems that could be solved effectively by using a certain approach.

	domain information	meta information	validity information
scoped	?	x	x
reifiable	-	x	x
topic	-	?	?
relation	?	-	x
subject	?	-	-

The domain field is not filled, because all temporal qualifications would have to be added into one set of scopes. This would not allow for the necessary differentiations necessary to model a domain. It is also unclear how temporal qualifications could be declared to belong to a special type of qualification. This could be solved by creating special topics connected to the original dates, that express the kind of information given by adding this topic to the scope. This would overpopulate any topic map and make the approach all but minimal.

6.2 Occurrences

Another way of solving the temporal qualification problem is to introduce occurrences of a special type and let them have the dates as their values. With this approach it is possible to fill the second and third line of the table and because all scoped elements are reifiable thereby the whole table.

	domain information	meta information	validity information
scoped	x	x	x
reifiable	-	x	x
topic	-	x	x
relation	x	-	x
topic	x	-	-

To model the different types of temporal information given different types of occurrences could be used. These types would be defined by the standard

by giving the appropriate PSIs All constructs that do not allow for occurrences themselves could have the temporal qualifications introduced as occurrences of their reifiers. Additional information could be given by reification of the occurrences for temporal qualifications themselves. It is also no problem to add multiple temporal qualifications as it is no problem to have many separate occurrences if this is necessary. So this approach seems to be very well suited for a standardization. However the only criteria for which this approach fails outweighs the other criteria here. Even if topics were used as the data stored in those occurrences there would be no link between those topics and the the qualified constructs in the way such a link is usually created in the topic maps standard and the only way to navigate from the dates to the topics qualified with those dates would be to search through all the occurrences of appropriate type. Therefore this approach will not be used in this paper for it's inability to make correct use of the topics representing the dates.

6.3 Associations

If it is decided that dates should be represented as topics a natural way to connect them to other topics is by using associations. By using different types of associations it would be possible to do represent differences in the type of information given by the association. So it would be possible to fill the third line of the table.

	domain information	meta information	validity information
scoped	?	?	?
reifiable	-	?	?
topic	-	x	x
relation	?	-	?
subject	x	-	-

This approach allows the dates to be topics. Additional information can be added by reifying the associations and multiple association can be used when different information needs to be attached to one topic. The use of association demands that association types and role types are given for a complete modeling therefore this approach is not as minimalistic as the solutions mentioned before. However to greatest deficit of this approach is the inability to handle the all the lines in the table. Associations can only contains roles which themselves can only contain topics as their players. Therefore this approach is not applicable to most entries in the table. Therefore the last approach presented will be an extension of this one intended to fix exactly this problem.

6.4 Reifier

In order to include all topic map constructs the mechanism of reification is used. Anything that is not a topic will be augmented with temporal information by creating an association of appropriate type between it's reifier and a date. This solves the problem of lacking generality of the association approach but it also

makes the approach even less minimal. Still this seems to be the best approach of all those discussed here. Therefore the rest of this paper will deal with developing an standard for the addition of temporal qualifications to topic maps based on this association and reification model.

7 The Standard Proposal

With the decision on how to model temporal qualifications made this section will focus on how to standardize this approach. No standardization is needed for reification except to determine that this in connection with the associations presented will be used to model the temporal information for all constructs other than topics. To standardize associations the following things need to be defined.

1. the association types
2. the role types
3. the possible players for each role type
4. the number of roles for each association

There will be no restriction on the number of roles in any of the associations defined here. Therefore it is possible to apply temporal qualifications not only to single constructs but to groups of constructs as well and it will be possible to use multiple temporal roles to signify complex temporal information.

Recall that the domain of predefined constructions is to be referred to as *pred*. As long as no other name for this domain is determined it will be used as a prefix for all the PSIs used in defining topics used as types.

7.1 Predefined Association Types

The topic with the subject identifier *pred/temporal-qualification* should be the supertype of all types of associations that represent temporal qualifications. All associations typed by a subtype of this topic should represent temporal qualifications.

In order to discriminate different types of temporal qualifications the following topics are introduced. Each is supposed to be a subtype of the *temporal-qualification* type.

1. The subtypes of the topic that has *pred/meta-temporal-qualification* as it's subject identifier should be the types of those associations that represent temporal meta information.
2. The subtypes of the topic that has *pred/domain-temporal-qualification* as it's subject identifier should be the types of those associations that represent domain related temporal information.
3. The subtypes of the topic that has *pred/validity-temporal-qualification* as it's subject identifier should be the types of those associations that represent temporal information about validity. It's subtype *pred/validity-temporal-qualification-domain* and it's subtypes should be used for validity information about the domain and the subtype *pred/validity-temporal-qualification-construct* and it's subtypes to talk about validity information for constructs.

7.2 Predefined Roles

The last element that is needed for a standardization is the definition of the roles of the associations described so far and a restriction on the players of these role.

1. roles of some subtype of the topic that has *pred/Timed* as it's subject identifier should be the roles of topics that represent elements to which a temporal qualification should be added. There is no restriction to the type of players for these roles.
2. roles of some subtype of the topic that has *pred/Time* as it's subject identifier should be the roles of topics that represent the date which is used in a temporal qualification. These roles should only be played by topics that represent dates.

8 Something Date Specific

An Example how the standard described here could be used in an actual topic map is the following. If one has decided to use dates in order to model time and wants to express that something denoted by a certain topic existed in a certain time interval. To this end a subtype of *pred/Time* to declare something as the begin of an interval. It could be called *domainname/Begin*. Then another subtype *domainname/end* could be created to denote the end of the interval. The type of the association could be subtype of *pred/domain-temporal-qualification* of called *domainname/existence-qualification*. Then associations could be created of this type with roles for the start of an interval, it's end and the element that is qualified.

9 Solving the Example Problems

How to would the example problems from the second section of this paper be solved with the standard defined in this paper? Birth dates could be modeled by creating a subtype of the *pred/domain-temporal-qualification* type, with the identifier *domainname/birth-qualification*. No new roles are needed, as *pred/time* and *pred/timed* seem to be perfectly reasonable for this task. To clarify when certain data was collected, the according statements may be reified and the reifiers can be used in an association of a subtype of *pred/meta-temporal-qualification*, for example *domainname/collected* to signify that this association determines the collection date for the data element that is reified. The problem of validity for a name could be solved by using the *begin/end* approach from the last section, reification and the *pred/validity-temporal-qualification-domain* type.

10 Conclusion

The idea for a set of predefined constructions for topic maps was introduced and justified. Then an example for the kind of problem to be solved by such an construction was introduced.

This paper gave a discussion of different possibilities of introducing temporal information into topic maps. One approach was chosen as a candidate for a possible standardization of the handling of temporal qualifications in topic maps. Then it was described how this approach could be standardized.

This approach seems to be very versatile and simple. It would even be possible to say that a set of topics shares a certain temporal property by including them all in an association.

Related to this topic is the question of how to represent the domain of time. Further research may include a proposal for a standardization of the representation of the time domain. Then tools could be introduced that use both standardizations in order to allow for an elegant handling of temporal qualifications in topic maps.

References

1. Heuer, L., Schmidt, J.: Tmap 2.0. In: TMRA 2008, University of Leipzig (2008)
2. for Standardization. Geneva. Switzerland, I.A.: Iso/iec is 13250-2:2006: Information technology – document description and processing languages – topic maps – data model
3. Ahmed, K.: Topic map design patterns for information architecture