ISO/IEC JTC 1/SC34 N0394



ISO/IEC JTC 1/SC34

Information Technology --

Document Description and Processing Languages

TITLE: Simple Topic Maps (STM)

SOURCE: Michel Biezunski and Steven R. Newcomb

PROJECT: Topic Maps

PROJECT EDITORS: Michel Biezunski, Martin Bryan, Steven R. Newcomb

STATUS: Editor's Draft, Revision 1.5
ACTION: For review and comment

DATE: 27 April 2003

SUMMARY:

DISTRIBUTION: SC34 and Liaisons

REFER TO: SUPERCEDES:

REPLY TO: Dr. James David Mason

(ISO/IEC JTC1/SC34 Chairman) Y-12 National Security Complex Information Technology Services

Bldg. 9113 M.S. 8208

Oak Ridge, TN 37831-8208 U.S.A. Telephone: +1 865 574-6973 Facsimile: +1 865 574-1896 E-mail: mailto:mxm@y12.doe.gov

http://www.y12.doe.gov/sgml/sc34/sc34oldhome.htm

Ms. Sara Hafele Desautels, ISO/IEC JTC 1/SC 34 Secretariat

American National Standards Institute

25 West 43rd Street New York, NY 10036 Tel: +1 212 642-4937 Fax: +1 212 840-2298 E-mail: sdesaute@ansi.org

27 April 2003

JTC 1/SC 34 N 0394

Simple Topic Maps (STM)

Version 1.5, 2003/04/27

Go to http://www.isotopicmaps.org/TMMM/STM-latest.html to view or contribute to the current editors' working revision.

Table of Contents

0 Introduction

1 Scope

2 References

3 TM Application Definition for STM (incomplete)

3.1 Application Name

3.2 Included TM Application Definitions

3.3 Property Classes

3.4 Assertion Types

3.5 Built-in Topics and Assertions

4 Document Type

4.1 Comparison with of STM DTD with XTM DTD

4.2 STM Document Type Definition

5 STM Syntax Descrialization Definition (incomplete)

5.1 Topic demanders

5.2 Deserializing a <topicMap> element

5.3 Deserializing a <topic> element

5.4 Deserializing an <association> element

5.5 Deserializing a <mergeMap> element

0 Introduction

This document provides parts of an example of a "TM Application Definition" and a "Syntax Deserialization Definition", as those terms are defined in SC34/WG3 N0393. (It does not provide a complete TM Application Definition, nor a complete Syntax Deserialization Definition. However, it does provide a complete DTD.)

The best available introduction to TM Application Definitions and Syntax Deserialization Definitions is ISO/IEC JTC1 SC34/WG3 N0393.

The STM Application and its Syntax Deserialization Definition fulfill a commonplace subset of the larger set of user requirements that the XTM and HyTM syntaxes of IS13250 are designed to fulfill. The authors' experiences indicate that this subset is adequate in a significant number of commercial contexts. STM demonstrates that, because of the modularity offered by the TMM approach described in N0393, it is unnecessary for all topic map systems to implement all the features of the traditional interpretation of XTM or HyTM comprehensively. The modularity offered by the TMM approach allows implementers to avoid incurring development and operating overheads that are unnecessary in many commercial contexts. Similarly, STM also demonstrates that a modular approach to the specification of Topic Map Applications can allow the implementations of different Topic Map Applications to share modules, thus decreasing software development and maintainance costs, while increasing overall reliability.

1 Scope

This document specifies:

1. A portion of a TM Application Definition for Simple Topic Maps (STM), in conformance with ISO/IEC JTC1/SC34/WG3 N0393, the current editors' draft of the TMM (the March, 2003 version of the *Topic Maps Model*).

At the TM Application level (as opposed to the Syntax Deserialization level), the only difference between STM and most traditional interpretations of XTM is that, in STM, the concept of scoping, in the context of topic naming, does not include the concept of establishing a namespace. Instead, a distinct <space> element type is provided in the DTD, whose deserialization involves a distinct

1S13250-STM-1.1::AT_namespace assertion type, which is provided in the TM Application Definition for STM.

If we compare the SAM model proposed in N0396 with the TM Application proposed for STM, there are additional differences: STM reifies some subjects that N0396 does not reify, including: URIs, the streams that result from resolving URIs, the relationships between URIs and their streams (HTTP GET operations), and the relationships between streams regarded as subject indicators and the subjects that they indicate. The theory behind this is that, while the syntax of STM should be very simple, the information expressed in STM should be ready to be managed at a very high power level by information aggregators, without compromising its integrity, and without having to convert it to another TM Application. In other words, original authors of STM topic maps should be able to keep maintaining them, while their aggregators can continue to add value to them, and to aggregate them with other STM topic maps, losslessly, in real time.

2. A definition of the STM Document Type 1.5, in the form of an XML DTD. The STM DTD bears an obvious resemblance to the XTM DTD, but some names are different, and it is somewhat simpler, except for the support it provides for the added distinction between scopes and namespaces.

A key simplifying feature of STM is that its DTD sacrifices some generality in order to avoid always having to be syntactically explicit about whether a referenced piece of information is a subject constituter or a subject indicator.

3. A portion of an STM Syntax Description 1.5 for instances of the STM document type.

2 References

ISO/IEC JTC1 SC34/WG3 N0393

- **3 TM Application Definition for STM (incomplete)**
- 3.1 Application Name

The name of the TM Application defined in this document is IS13250-STM-1.5

3.2 Included TM Application Definitions

Note 1: The TM Applications listed below are included, by this reference, in their entirety, in this IS13250-STM-1.5 TM Application. See <u>ISO/IEC JTC1 SC34/WG3 N0393</u>.

IS13250-TMASD-1.2

IS13250-HTTPGET-1.1

IS13250-SETS-1.1

IS13250-CLASS-1.1

IS13250-SCOPE-1.1

IS13250-OCCURRENCE-1.1

3.3 Property Classes

3.3.1 IS13250-STM-1.5::PR_nameText

Value type: string

Semantic: If a value has been assigned to this property, the subject of the topic is the string that is value of the property, considered as a topic name.

Constraints on values: The value cannot be less than one character in length.

Consistency constraints: (None.)

SIDP or OP? SIDP

Topic Merging Rule: Whenever two or more topics both exhibit values for their IS13250-STM-1.5::PR_nameText properties, the topics must be merged if their values are the same.

Effect of merging on values: Whenever two or more topics that exhibit values for their IS13250-STM-1.5::PR_nameText properties are merged, one of the two values must become the value of the same property of the merged topic. If one of the two values is a built-in value, the built-in value becomes the value of the same property of the merged topic.

3.4 Assertion Types

3.4.1 IS13250-STM-1.5::AT_name

Semantic: Each instance asserts that a subject has a name (and, from the other perspective, that a name is the name of a subject).

3.4.1.1 Role: IS13250-STM-1.5::RL_name

Semantic: The name that is asserted to be the name of the subject that plays the other role (IS13250-STM-1.5::RL_namedSubject).

Constraints on player: Must be a name.

Value conferred on the IS13250-STM-1.5::PR_namedThings{ } property of the role player: The topic that plays the IS13250-STM-1.5::RL_namedSubject role of the assertion is added as a value component.

3.4.1.2 Role: IS13250-STM-1.5::RL_namedSubject

Semantic: The subject that is being asserted to have the name that plays the other role (IS13250-STM-1.5::RL_name).

Constraints on player: (None.)

Value conferred on the IS13250-STM-1.5::PR_names{ } property of the role player: The value of the IS13250-STM-1.5::PR_httpAddress property of the player of the other role (IS13250-STM-1.5::RL_address) is assigned.

3.4.2 IS13250-STM-1.5::AT_namespace

Semantic: Each instance asserts that a subject has one of its names within a specific namespace.

3.4.2.1 Role: IS13250-STM-1.5::RL_nameAssertion

Semantic: The name assertion whose namespace is being asserted. (IS13250-STM-1.5::RL_namedSubject).

Constraints on player: Must be an instance of a IS13250-STM-1.5::AT name assertion.

Value conferred on the IS13250-STM-1.5::PR_namespaces { } property of the role player: The topic that plays the IS13250-STM-1.5::RL_namespace role of the assertion is added as a value component.

3.4.2.2 Role: IS13250-STM-1.5::RL_namespace

Semantic: The namespace of the IS13250-STM-1.5::AT_name assertion that plays the IS13250-STM-1.5::RL_nameAssertion role.

Constraints on player: (None.)

(No property values are conferred.)

3.5 Built-in Topics and Assertions

Note 2: This section is not written yet.

4 Document Type

4.1 Comparison with of STM DTD with XTM DTD

The following is a list of differences between the STM DTD and the XTM DTD:

- 1. STM has no name variants, and no parameters for name variations.
- 2. STM uses <name> instead of <baseName>, and <text> instead of <baseNameString>. (In the absence of the notion of variant names, the notion of base names is not very meaningful.)
- 3. STM does not use <xlink> to refer to pieces of information. Instead, it uses <a>, on the theory that more people are comfortable with syntax than are comfortable with <xlink>.
- 4. STM has neither <subjectIndicatorRef> nor <resourceRef>. The vital distinction between subject indicators and subject constituters is preserved, however; the interpretation of the <a> elements, as to whether their referents are to be understood as subject constituters or subject indicators, is determined by the contexts of the <a> elements, as follows:
 - An <a> always references a subject indicator when it appears in the context of:
 - an <instanceOf>,
 - a <subjectIdentity>,
 - a <member>,
 - a <roleSpec>,
 - a <scope>, or
 - a <space>.
 - An <a> always references a subject constituter when it appears in the context of:
 - an <occurrence> or
 - a <mergeMap>.
- 5. STM has no <topicRef>. If an <a> refers to a <topic>, and the context of the <a> requires it to be understood as referencing a subject indicator, then the referent is the subject of the referenced <topic>. Otherwise, the <topic> is understood as a piece of information -- a subject constituter.
- 6. STM provides a <space> element, in addition to a <scope> element, for qualifying the relationships between named subjects and their names (see 1).
- 7. STM does not provide as many element types with id attributes as XTM does. There are two explanations for this:
 - a. It may be true that some of the id attributes provided by the XTM

DTD were added in order to support "lazy" reification via references to the syntactic instance. However, in order to preserve the integrity of topic maps when they are merged with other topic maps, the TMM requires pre-emptive reification, so this requirement is considered moot.

- b. STM is not proposed as a full-featured syntax for Topic Maps. Instead, it is proposed as a lightweight, easily implementable alternative to XTM. As such, its SDD declares very few "topic demanders" (see the TMM for more information).
- 9. STM requires that topics that have the same names in the same namespaces (declared via <space>) are always merged. (In other words, the name-based merging rule is always "on", but, on the other hand, in STM, namespacing and scoping are different things, whereas in XTM, they are the same.)
- 10. STM requires that <association>s declare at least two roles. This makes the descrialization of STM syntax considerably simpler than the descrialization of XTM syntax, since one-role <association>s in XTM syntax must be descrialized in a special way that creates a class-instance assertion.
- 11. STM requires all <association> roles to be explicitly invoked; this obviates the need for special defaulting rules in the Syntax Deserialization Definition, and/or special default assertion types and/or roles to be supplied by the governing TM Application.

4.2 STM Document Type Definition

```
<!-- topicMap: root element -->
<!ELEMENT topicMap ( topic | association | mergeMap )* >
<!ATTLIST topicMap
                             #TMPLTED
                  CDATA #FIXED 'http://www.isotopicmaps.org/TMMM/STM-1.5/
  xmlns
<!-- topic: referenceable syntactic surrogate of a subject -->
<!ELEMENT topic
   ( instanceOf*, subjectIdentity?, ( name | occurrence )* )
<!ATTLIST topic
  id
                  ID
                             #REQUIRED
<!-- instanceOf: points to a <topic> representing a class -->
<!ELEMENT instanceOf (a) >
<!-- subjectIdentity: pointers to subject indicators -->
<!ELEMENT subjectIdentity ( a* ) >
<!-- a: wrapper for an -href- attribute -->
<!ELEMENT a EMPTY >
<!ATTLIST a
                CDATA #REQUIRED ID #IMPLIED
  href
  id
<!-- name: a name of the subject of the containing <topic> -->
<!ELEMENT name ( text, space?, scope?) >
```

```
<!-- NOTE: The default topic name space is the empty
                     set (of topics). (Every topic-name
                     relationship exists within at least one
                     namespace.)
          <!-- NOTE: The default scope in which topics have
                    their names is the empty set (of topics).
<!-- text: the name in #PCDATA form -->
<!ELEMENT text ( #PCDATA ) >
<!-- space: pointer to a <topic> or subject indicator whose
           subject is the name space within which the topic
           has the name -->
<!ELEMENT space (a) >
<!-- scope: pointers to <topic>s or other indicators of
           subjects that comprise the scope -->
<!ELEMENT scope (a)+ >
<!-- occurrence: relationship between a subject and relevant
                information -->
<!ELEMENT occurrence ( instanceOf?, scope?, a ) >
         <!-- NOTE: <instanceOf> points at an assertion type that
                     is either the topic-occurrence assertion type
                     (the default), or one of its subclasses. -->
<!-- association: relationship between 2 or more subjects, each
                 playing one or more distinct roles -->
<!ELEMENT association ( instanceOf, scope?, member, member+ ) >
<!ATTLIST association
  id
                             #TMPLTED
<!-- member: a casting of a role player in a role in a
            relationship -->
<!ELEMENT member ( roleSpec, ( a )* ) >
<!-- roleSpec: pointer to a topic whose subject is the role -->
<!ELEMENT roleSpec (a)
<!-- mergeMap: Include (merge with) other topic maps -->
<!ELEMENT mergeMap ( a )+ >
```

5 STM Syntax Deserialization Definition (incomplete)

Note 3: In this clause (5), a typographical convention is used to distinguish topic elements (syntactic constructs) from the more generic TMM usage of the word "topic", in which topics are reified subjects. Herein, all element instances and element types are set in monospace font, surrounded with angle brackets. For example, the typography "<topic" indicates an element or element type whose generic identifier is "topic". By contrast, the more generic notion of "topic" as "reified subject" does not receive any special typographical treatment.

5.1 Topic demanders

Instances of the following element types are topic demanders. In other words, references to them, by means of the -href- attributes of <a> elements, are considered references to subjects, as follows:

5.1.1 <a>

If the referencing <a> is in the context of an <occurrence>, then the referent is not considered a topic demander. The referent is the <a> element, considered as a string.

If the referencing <a> is not in the context of an <occurrence>, then the referent is one of the topics demanded by the <a>, as follows:

- If the referenced <a> is in the context of an <instanceOf>, the referent is the IS13250-CLASS-1.1::AT_classInstance assertion demanded in accordance with 5.3.2.6.
- If the referenced <a> is in the context of a <subjectIdentity>, the referent is the IS13250-HTTPGET-1.1::AT_subjectIndication assertion demanded in accordance with 5.3.3.
- If the referenced <a> is in the context of a <member>, the referent is the c-topic that casts the role player topic demanded in accordance with 5.4 in the role.
- If the referenced <a> is in the context of a <rolespec>, the referent is the role topic demanded in accordance with 5.4.
- If the referenced <a> is in the context of a <scope>, the referent is the IS13250-SETS-1.1::AT_set-member assertion demanded in accordance with 5.3.6, 5.4, or 5.3.5.2.5 (depending on whether the <scope> is in the context of an <association>, an <occurrence>, or a <name>).
- If the referenced <a> is in the context of a <space>, the referent is the IS13250-STM-1.1::AT_namespace assertion demanded in accordance with 5.3.5.3.3.
- If the referenced <a> is in the context of an <occurrence>, the referent is the IS13250-OCCURRENCE-1.1::AT_occurrence assertion demanded in accordance with 5.3.6.

5.1.2 <association>

If the referencing <a> is in the context of an <occurrence>, then the referent is not considered a topic demander. The referent is the <association> element, considered as a string.

If the referencing <a> is *not* in the context of an <occurrence>, then the referent is the assertion demanded by the <association>, in accordance with 5.4.

5.1.3 <topic>

If the referencing <a> is in the context of an <occurrence>, then the referent is not considered a topic demander. The referent is the <topic> element, considered as a string.

If the referencing <a> is *not* in the context of an <occurrence>, then the referent is the topic demanded by the <topic>, in accordance with 5.3.1.

5.2 Deserializing a <topicMap> element

<topicMap> has no effect. The topic map itself is not automatically reified.

5.3 Deserializing a <topic> element

5.3.1 Deserializing the <topic> itself

Each <topic> demands a corresponding topic.

5.3.2 Deserializing an <instanceOf> within a <topic>

The following topics are demanded by each <instanceOf> in the content of the <topic>:

5.3.2.1 addressing expression

A topic is demanded whose subject is the addressing expression that is the value of the _href- attribute of the <a> in the content of the <instanceOf>. The value of the _href- attribute is built-in as the value of the IS13250-HTTPGET-1.1::PR_httpAddress property.

5.3.2.2 addressed stream

A topic is demanded whose subject is the stream that is the result of resolving the subject of the topic demanded in accordance with 5.3.2.1.

5.3.2.3 IS13250-HTTPGET-1.1::AT_httpGet assertion

An IS13250-HTTPGET-1.1::AT_httpGet assertion is demanded in which the IS13250-HTTPGET-1.1::RL_address role is played by the topic demanded in accordance with 5.3.2.1, and in which the IS13250-HTTPGET-1.1::RL_stream role is played by the topic demanded in accordance with 5.3.2.2.

5.3.2.4 class topic

A topic is demanded whose subject is the class of which the topic demanded in accordance with 5.3.2.1 is an instance.

5.3.2.5 IS13250-HTTPGET-1.1::AT_subjectIndication assertion

An IS13250-HTTPGET-1.1::AT_subjectIndication assertion is demanded in which the IS13250-HTTPGET-1.1::RL_subjectIndicator role is played by the topic demanded in accordance with <u>5.3.2.2</u>, and in which the IS13250-HTTPGET-1.1::RL_indicatedSubject role is played by the topic demanded in accordance with <u>5.3.2.4</u>.

5.3.2.6 IS13250-CLASS-1.1::AT_classInstance assertion

An IS13250-CLASS-1.1::AT_classInstance assertion is demanded in which the IS13250-CLASS-1.1::RL_class role is played by the topic demanded in accordance with 5.3.2.4, and in which the IS13250-CLASS-1.1::RL_instance role is played by the topic demanded in accordance with 5.3.1.

5.3.3 Descrializing a <subjectIdentity> element within a <topic>

5.3.4 Deserializing a <name> element within a <topic>

5.3.4.1 Deserializing a <text> within a <name>

5.3.4.1.1 name text

A topic is demanded whose subject is the name text. The content of the <text> is built-in as the value of its IS13250-STM-1.5::PR_nameText property.

5.3.4.2 Deserializing a <space> within a <name>

5.3.4.2.1 name space

A topic is demanded whose subject is the name space.

5.3.4.2.2 addressing expression

A topic is demanded whose subject is the addressing expression that is the value of the -href- attribute of the <a> in the content of the <space>. The value of the -href- attribute is built-in as the value of the IS13250-HTTPGET-1.1::PR_httpAddress property.

5.3.4.2.3 addressed stream

A topic is demanded whose subject is the stream that is the result of resolving the subject of the topic demanded in accordance with 5.3.4.2.2.

5.3.4.2.4 IS13250-HTTPGET-1.1::AT_httpGet assertion

An ${\tt IS13250-HTTPGET-1.1::AT_httpGet}$ assertion is demanded in which the ${\tt IS13250-HTTPGET-1.1::RL_address}$ role is played by the topic demanded in accordance with ${\tt 5.3.4.2.2}$, and in which the ${\tt IS13250-HTTPGET-1.1::RL_stream}$ role is played by the topic demanded in accordance with ${\tt 5.3.4.2.3}$.

5.3.4.2.5 IS13250-HTTPGET-1.1::AT_subjectIndication assertion

An IS13250-HTTPGET-1.1::AT_subjectIndication assertion is demanded in which the IS13250-HTTPGET-1.1::RL_subjectIndicator role is played by the topic demanded in accordance with 5.3.4.2.3, and in which the IS13250-HTTPGET-1.1::RL_indicatedSubject role is played by the topic demanded in accordance with 5.3.4.2.1.

5.3.5 Deserializing a <scope> within a <name>

5.3.5.1 scope

A topic is demanded whose subject is the set of topics that comprises the scope.

5.3.5.2 For each value of an -href- attribute of each <a> in the content of the scope:

5.3.5.2.1 addressing expression

A topic is demanded whose subject is the addressing expression that is the value of the -href- attribute of an <a> in the content of the <scope>. The value of the -href- attribute is built-in as the value of the IS13250-HTTPGET-1.1::PR_httpAddress property.

5.3.5.2.2 addressed stream

A topic is demanded whose subject is the stream that is the result of resolving the subject of the topic demanded in accordance with 5.3.5.2.1.

5.3.5.2.3 IS13250-HTTPGET-1.1::AT_httpGet assertion

An IS13250-HTTPGET-1.1::AT_httpGet assertion is demanded in which the IS13250-HTTPGET-1.1::RL_address role is played by the topic demanded in accordance with 5.3.5.2.1, and in which the IS13250-HTTPGET-1.1::RL_stream role is played by the topic demanded in accordance with 5.3.5.2.2.

5.3.5.2.4 set member topic

A topic is demanded whose subject is a member of the set of topics which is the subject of the topic demanded in accordance with 5.3.5.1.

5.3.5.2.5 IS13250-SETS-1.1::AT set-member assertion

An IS13250-SETS-1.1::AT_set-member assertion is demanded in which the IS13250-SETS-1.1::RL_set role is played by the topic demanded in accordance with 5.3.5.1, and in which the IS13250-SETS-1.1::RL_member role is played by the topic demanded in accordance with 5.3.5.2.4.

Note 4: This assertion confers a value component on the set topic's IS13250-SETS-1.1::PR_members{} SIDP. The component is the topic demanded in accordance with <u>5.3.5.2.4</u>.

5.3.5.3 Deserializing the <name> itself

5.3.5.3.1 IS13250-STM-1.1::AT_name assertion

An IS13250-STM-1.1::AT_name assertion is demanded in which the IS13250-STM-1.5::RL_name role is played by the topic demanded in accordance with 5.3.4.1.1, and the IS13250-STM-1.5::RL_namedSubject role is played by the topic demanded in accordance with 5.3.1.

5.3.5.3.2 IS13250-SCOPE-1.1::AT_scope assertion

An IS13250-SCOPE-1.1::AT_scope assertion is demanded in which the IS13250-SCOPE-1.1::RL_scope role is played by the topic demanded in accordance with 5.3.5.1, and the IS13250-SCOPE-1.1::RL_scopedAssertion role is played by the assertion demanded in accordance with 5.3.5.2.4.

Note 5: This assertion confers, as a value component of the scoped assertion's IS13250-SCOPE-1.1::PR_scopes{ } property, the topic that plays the IS13250-SCOPE-1.1::RL_scope role.

5.3.5.3.3 IS13250-STM-1.1::AT_namespace assertion

An IS13250-STM-1.1::AT_namespace assertion is demanded in which the IS13250-STM-1.5::RL_namespace role is played by the topic demanded in accordance with 5.3.4.2.1, and the IS13250-STM-1.5::RL_nameAssertion role is played by the assertion demanded in accordance with 5.3.5.2.4.

- 5.3.6 Deserializing an cocurrence> element within a <topic>
 - 5.4 Deserializing an <association> element
- 5.4.1 Deserializing an <instanceOf> element within an <association>
- 5.4.2 Deserializing a <scope> element within an <association>
- 5.4.3 Deserializing a <member> element within an <association>
 - 5.5 Deserializing a <mergeMap> element