SAM

Light at the end of the tunnel?

Status

- **SAM has been worked on since May 2001**
  - we hope to resolve all remaining issues at this meeting
- **Proposals for all issues provided here**
- **Text currently not in accordance with ISO requirements**
  - next draft will be rewritten to ISO style and have issue resolutions edited in
  - the goal is to send the next draft to ballot as Committee Draft
- **Relevant documents**
  - N0396: SAM
  - N0397: Proposed resolutions to open SAM issues

Plan for this session

- **Do the issues one by one**
  - first introduce the issue with background
  - then present proposed resolution from the authors
  - committee discussion
  - finally: resolve issue or put it aside to return to it later
- **Note that the resolutions we propose are for the most part edited into the text of N0396 already**
- **We don't have many hours, so we should aim for efficiency...**
  - **should probably continue until done, and if necessary skip the CXTM discussion entirely**

locator-notation-support

- ***"What locator notations, if any, must be supported?"***
- **That is, what are we going to require implementations to support? URIs? HyTime? Both?**
- **Considerations:**
  - each syntax will use specific locator syntaxes and thus require particular locator notations
- **Proposal:**
  - leave the requirements for the syntaxes
  - if you want to support XTM, you must support URIs
  - if you want to support HyTM, you must support HyTime locators
  - and so on...

prop-value

- ***“Regarding topic name.[value]; is [label] a better name?”***
- **Considerations:**
  - the string is a label of the topic, but the value of the base name
  - the base name is already a label
  - we could call it [base name] but topic name.[base name] does not really have the right ring to it
- **Proposal:**
  - leave it as [value]

prop-variant-scope-superset

- ***"ISO 13250:2000 allows a display name to have a scope that is a subset of that of the corresponding base name. This apparent contradiction needs to be resolved."***
- **HyTM syntax example:**
  - <topname scope="A">
  - <basename scope="B">Foo</basename>
  - <sortname scope="C">foo</sortname>
- **The view of the authors is that XTM improved HyTM in this respect**
- **Proposal:**
  - leave the SAM as is

names-as-subjects

- ***“Should base name items be merged, so that assertions made about one base name will also apply to all other base names that have the same identity? (This also applies to occurrences.)”***
- **Was raised by the RM group, but no longer seems to be important**
- **Does anyone still care about this?**
- **Proposal:**
  - forget about it

strings-as-subjects

- ***"Should it be possible to create topics that represent strings, and for it to be formally clear that these topics do represent particular strings? If so, how?"***
- **Considerations:**
  - data URLs can be used with subject addresses to achieve this
  - it is not an important use case in practice
- **Proposal:**
  - do not add any additional features to support this

psi-subclassing-loops

- ***“Are subtype loops allowed?”***
- **Considerations:**
  - detecting such loops is costly for implementations
  - the subtype relation can be seen as a 'strict order' or an 'order'
    - the first does not allow loops, the second does
  - in other words: there is not necessarily any logical contradiction involved
  - disallowing them allows optimizations during logical inferencing
  - whichever we choose, anyone who wants the other kind can define a PSI for it
- **Proposal:**
  - allow subclass loops; make supertype-subtype an 'order'

prop-subj-address-class

- ***"Are topics representing information resources allowed as types?"***
- **Considerations:**
  - we could put it in as a SAM constraint
  - increases implementation cost
  - we already agreed SAM shouldn't be bound by implicit XTM restrictions
  - we have the same issue with scope **(prop-subj-address-scope)**
  - can we be certain that there is absolutely no circumstance under which this might make sense?
- **Proposal:**
  - allow it

prop-subj-address-values

- ***"Should topic.[subject address] only accept a single value?"***
- **Considerations:**
  - is it possible for different locators to reference the same resource?
  - different URIs? different HyTime locators?
  - do we want to enable users to claim that different locators refer to the same resource without SAM being able to verify it?
  - is there a value to having a restriction?
  - OASIS XMLvoc expressed a desire for it to accept multiple values
- **Proposal:**
  - change property name to [subject addresses] and make it a set

constr-single-subj-address

- ***"What happens when the single subject address constraint is violated?"***
- **When different topics are merged they may be found to have different subject addresses**
  - is this an error? does this mean the TM is invalid?
- **Alternatives:**
  - [subject addresses]: no problem
  - [subject address]: it must be an error, and the TM declared invalid
- **Proposal:**
  - use [subject addresses]

psi-set-psi

- ***"How does one uniquely identify the set of published subjects defined in SAM? Is there a need to do so? Is a published subject for these published subjects needed? (Does it include itself?)"***
- **Considerations:**
  - the PubSubj TC recommends that the top page of the PSI set be used as the identifier for the whole set
- **Proposal:**
  - use the top page as the identifier (for the set independent of version)
  - there should also be one PSI for each version

psi-identification

- ***"How does one determine which subjects are published subjects and which are not? Is it necessary for the SAM model to provide a mechanism for this at all?"***
- **Considerations:**
  - this is information about the subject indicator, not the subject
  - it does not seem to be very important to applications in general
  - anyone who wants to identify this can create a PSI used to type inforamation resources that are published subject indicators
  - this appears to be out of scope for SAM, but may be in scope for the OASIS PubSubj TC
- **Proposal:**
  - don't attempt to do this in SAM

subject-identity-establish

- ***“ISO 13250 states that subject identity may be 'inferred from the topic's characteristics.' Does SAM need words to the same effect?”***
- **Considerations:**
  - this is not a processing constraint
  - it is really information to the users that they are free to do this
  - it does seem like useful information that should be provided
- **Proposal:**
  - Add to last para of 3.4: “Applications and users are therefore free to merge topics as they see fit. Most commonly this will be done by inferring the subject of the topics from their characteristics.”

term-subject-indicator-def

- ***"If the subject identifier is a locator that does not refer to an information resource, what is the subject indicator then? This also applies to the subject address."***
- **Considerations:**
  - clearly this is bad practice
  - equally clearly it cannot be reliably detected by software
- **Proposal:**
  - add a note that warns against doing this to make it clear that it is not the right way to use this feature

merge-srcloc-vs-subjid

- ***"What happens when the same locator appears as a source locator for one topic and as a subject identifier for another?"***
- **Considerations:**
  - SAM does not allow the same locator be a source locator and subject identifier for the same topic
  - the Baltimore meeting decided the locator should in this cases become a source locator
  - later practical experience has shown that this is problematic because a <topicRef/> to a subject indicator will cause the subject identifier to disappear
  - removing the constraint does not appear to cause any difficulties
- **Proposal:**
  - allow the same locator to be both a source locator and a subject identifier for the same topic

occurrence-variants

- ***"Should occurrences be allowed to have variants in the same way that topic names are?"***
- **Considerations:**
  - this is not a bug fix, but an extension of the existing standard
  - the use case is to support renditions of occurrences appropriate in different contexts
  - also, it was claimed that it would make occurrences and topic names more symmetric
- **Proposal:**
  - we feel that whatever the merits or demerits of this proposal it goes beyond "restatement of ISO 13250", and so we should not do this

reification-effects

- ***"If you reify a topic name, does that affect your allowed type? If you reify an association, must you inherit its type?"***
- **Examples:**
  - must a topic reifying a topic name be of type 'topic name'?
  - must a topic reifying the association representing 'the marriage of Winston and Clementine Churchill' be of type 'marriage'?
- **Considerations:**
  - such checks add costs for implementors
  - it is not clear that these semantics are always appropriate
- **Proposal:**
  - leave this alone
  - define a set of PSIs for core topic map contructs

term-scope-def

- **When an occurrence of type start-date has scope A and B, what can an application safely conclude?**
  - that the start date is valid when A applies and when B applies (any subjects)?
  - that the start date is valid when A and B both apply (all subjects)?
  - or is the interpretation of scope undefined, so that both are possible?

Any subjects

- **In this view, the two examples below are obviously equivalent**
  - {rema-1000, opening-hours, [[0800-2200]]} / mon
    {rema-1000, opening-hours, [[0800-2200]]} / tue
  - {rema-1000, opening-hours, [[0800-2200]]} / mon tue
- **What that means is that**
  - if two topic characteristics are equal, but have different scopes, they can always be merged and their scopes unioned
  - in other words, in this view scope does not affect merging
  - to add another topic to the scope of a topic characteristic widens its applicability
- **This would be broader than the examples above**
  - {rema-1000, opening-hours, [[0800-2200]]} / mon tue wed
- **In other words, in this view the unconstrained scope must be the universal set**
- **TMCL and TMQL can use a single scope matching operator**

All subjects

- **Here, semantics are different**
  - as topics are added to the scope it becomes increasingly restricted
  - {tnc, opinion, [[TNC is a bug.]]} / lmg tuesday
  - the unconstrained scope must therefore be the empty set
  - topic characteristics can be merged when they are equal and the scope of one is a subset of the scope of the other
- **Expressing X is valid when Y and when Z becomes awkward**
  - PSIs should be provided to enable users to construct the set of {Y,Z} as this would solve the problem
- **TMCL and TMQL can use a single scope matching operator**

Undefined

- **In this case the unconstrained scope must be a special value**
- **Merging requires equal scopes**
- **Different applications can use scope differently**
- **TMQL and TMCL must provide multiple scope matchers**
  - A is a superset of B
  - A is a subset of B
  - A == B
  - (A is a superset of B) OR (A is a subset of B)
  - $(A \cap B) \neq \emptyset$
- **Pepper and Grønmo concluded that none of these algorithms yielded “consistently intuitive and useful results”**

Proposal

- **Define scope to be 'all subjects'**
- **Define the unconstrained scope as the empty set**
- **psi-type-instance-scope could be explained with an example**
  - type-instance(a : instance, b : type) / Y X
  - super-sub(b : sub, c : super) / Y Z
  - here, a will be an instance of b when Y and X apply
  - it will be an instance of c when Y, X, and Z apply
- **It seems clear that any-subjects is untenable**
- **It also seems clear that the 'undefined' alternative is less desirable**
  - scope matching becomes complex and non-untuitive
  - the interaction of type with scope becomes unclear

mnemonic-representation

- ***"Should HyTM mnemonics be represented in the model using special properties, or should they be represented using already existing constructs?"***
- **Considerations:**
  - mnemonics have the same semantics as types
  - they are less structured, but can be mapped to topics
  - their use is strongly discouraged
  - removing them would not be backwards-compatible
- **Proposal:**
  - map them to topics and associations
  - discourage their use in the specification

sam-conformance

- ***"Should the SAM have a conformance section of its own? If so, what does it mean to conform to the SAM?"***
- **Considerations:**
  - conformance requirements are intended to provide interoperability
  - SAM on its own cannot really provide any interoperability
  - interoperability must be on the API, syntax, query, or constraint level
  - other specifications will still reference the SAM as normative text, however, so SAM will still be normative
- **Proposal:**
  - remove the conformance section and leave it to the other specifications to define what conformance to them mean
- **This issue is essentially the same as 'scope-extension'**

The next step

- **After this it seems that we should**
  1. edit in all resolutions not already in the document
  2. edit it to conform to proper ISO style
  3. submit it for ballot as a Committee Draft