



Technical Issues on Topic Maps

Hans Holger Rath

STEP Electronic Publishing Solutions GmbH

Germany – Hungary – Norway – Poland – UK
Web: www.step.de – eMail: consulting@step.de

Some exercises

- ❖ Have you read the standard or visited the topic map tutorial?
- ❖ Have you done first experiments?
- ❖ Have you already designed a map?
- ❖ Do you understand the concepts?
- ❖ Do you **not** understand the concepts?

Overview

- ❖ **Topic map model**
 - ◆ A very brief summary
- ❖ **Topic map templates**
 - ◆ Purpose, concept, application
- ❖ **Consistency checking**
 - ◆ Rule-based constraints
- ❖ **Automatic generation**
 - ◆ Using existing information resources
- ❖ **Conclusions**



Topic map model

Very brief summary

- ❖ **Topic**
 - ◆ Name, type
- ❖ **Occurrence**
 - ◆ Role, role type
- ❖ **Association**
 - ◆ Type, role, role type
- ❖ **Scope for topic characteristics**
 - ◆ Theme
- ❖ **Public subject**
- ❖ **Facet**

Where STEP's topic map experience comes from

- ❖ Standards work: Steve Pepper and H.H. Rath are members of ISO topic map working group
- ❖ Fruitful discussions with leading reference works (encyclopedias, dictionaries), legal and journal publishers
- ❖ Cooperations with
 - ◆ AI company selling expert systems
 - ◆ Scientists from University of Würzburg



Topic Map Templates

Topic map templates
Missing pieces

- ❖ **Mostly everything is a topic**
- ❖ **Even the “objects” declaring a topic map are topics (themes and all types)**
- ❖ **There is no “DTD” concept that separates the schema from the instances**

Topic map templates
Shortcomings

- ❖ **Mix of “declaring” and “regular” topics**
- ❖ **Topic map design, creation, and maintenance are hard to separate**
- ❖ **No clean concepts for**
 - ◆ Modularization
 - ◆ Re-use
 - ◆ Application profiles

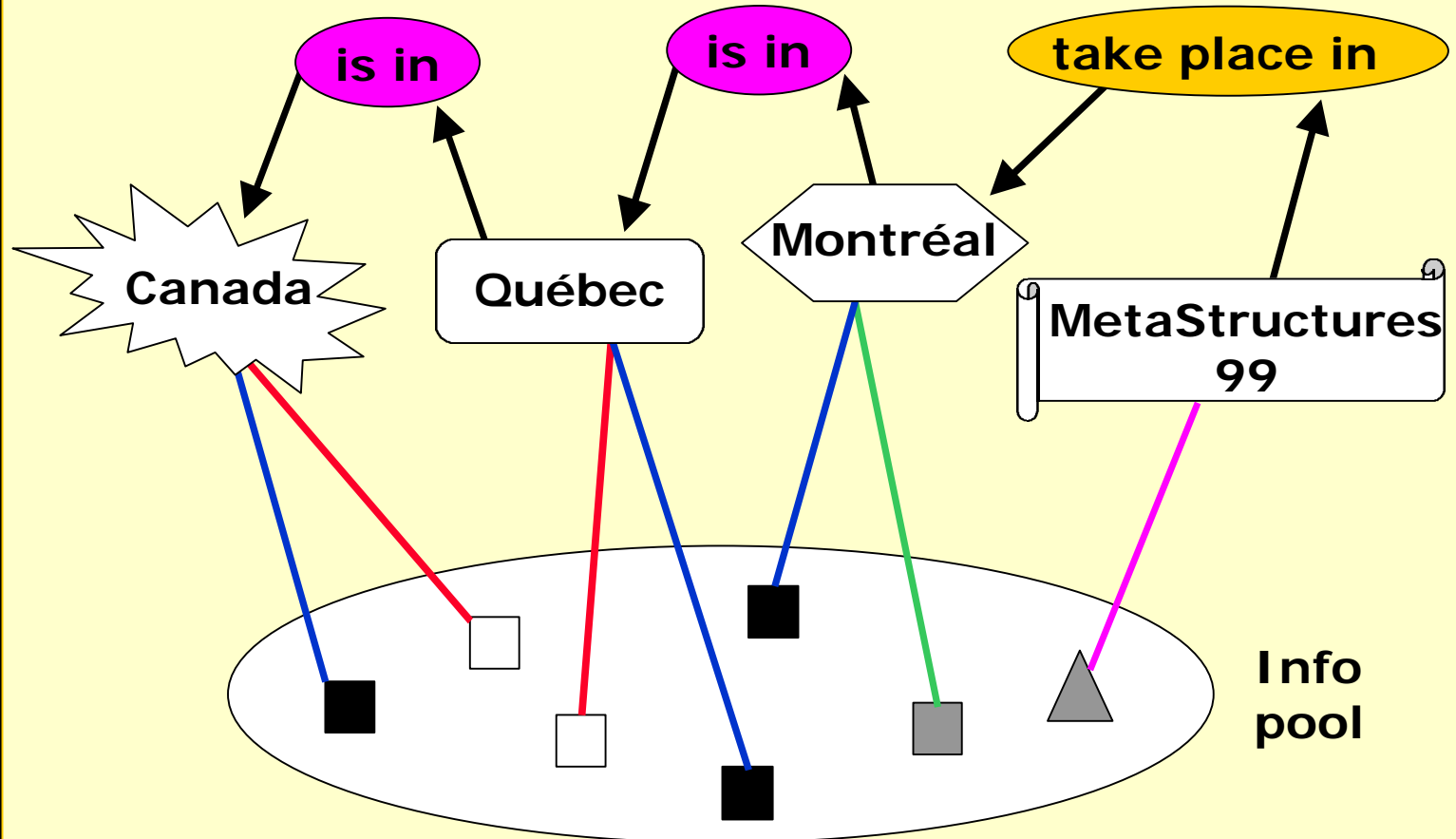
Purpose of templates

- ❖ Invented by ISO working group
- ❖ Logical container for the “declaration” part of a map
- ❖ Distinguish between map design and map creation
- ❖ Modularization
- ❖ Re-use (copy, reference) in other maps as kind of profile
- ❖ Templates might be standardized for application areas

Concept

- ❖ A topic map template is a topic map
- ❖ Consists of all constructs which have a **declarative meaning** for a map
 - ◆ topics which are candidates for themes and types (topics, occurrence roles, associations, association roles)
 - ◆ consistency constraints

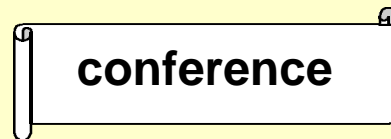
Topic map templates "Real" map



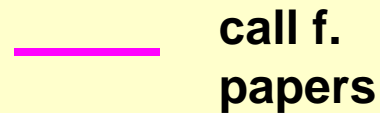
Topic map templates

"Template" map

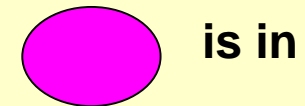
Topic types



Occurrence role types



Association types



Others:
Assoc. role types,
themes

Topic map templates
Concept cont'd

- ❖ **Topics “say” for which type they shall be used for**
- ❖ **Topics in template are distinguishable from “real” topics**
- ❖ **Template becomes manageable entity**
 - ◆ Identifier
 - ◆ Owner
 - ◆ Version



Consistency Checking

Consistency checking

- ❖ **Why is this an issue?**
 - ◆ Manual checking of large maps is impossible
 - ◆ Standard does not cover it
- ❖ **Should be part of topic map software supporting design and creation**
 - ◆ Permanently or on demand
 - ◆ Like structure validation in SGML/XML editors
 - ◆ Checks template also

Consistency checking **Constraints**

- ❖ **Content models of topic map elements are not expressive enough**
- ❖ **Separate schema needed: "constraints"**
- ❖ **Constraints control validation process**
- ❖ **Constraints are expressed by rules which are modeled with topic map constructs**

Consistency checking
Constraints cont'd

- ❖ **Constraining occurrences**
 - ◆ Topic type →
 - ◆ Occurrence role type →
 - ◆ Info resource "type" (if available)
- ❖ **Example:**
 - ◆ *person*
 - ◆ *biography, portrait*
 - ◆ SGML/XML instance with public identifier "-//STEP//DTD biography//EN"
 - ◆ Object types TIFF, GIF, JPEG

Consistency checking
Constraints cont'd

❖ **Constraining associations**

- ◆ Association type →
- ◆ Association role type(s) →
- ◆ Topic types of associated topics

❖ **Example:**

- ◆ *is in*
- ◆ 1 *containee* 1 *container*
- ◆ city country, state, county
 county country, state
 state country

Consistency checking
Constraints cont'd

- ❖ **Constraining scopes**
 - ◆ Topic type →
Scope for topics, names, occurrences
 - ◆ Association type →
Scope for associations and associated topics

Consistency checking
Constraints cont'd

- ❖ **Example of scope constraint:**
 - ◆ **Theme:** *before Einstein's theory of relativity*
 - ◆ **Topic types:** *physical law, mathematical axiom*
 - ◆ **Occurrence role type:** *definition*
 - ◆ **Constraint:** The scope *before Einstein's theory of relativity* might be used for occurrences with role *definition* for topics of type *physical law*; but it must not be used for *definitions of mathematical axioms*.



Automatic Generation

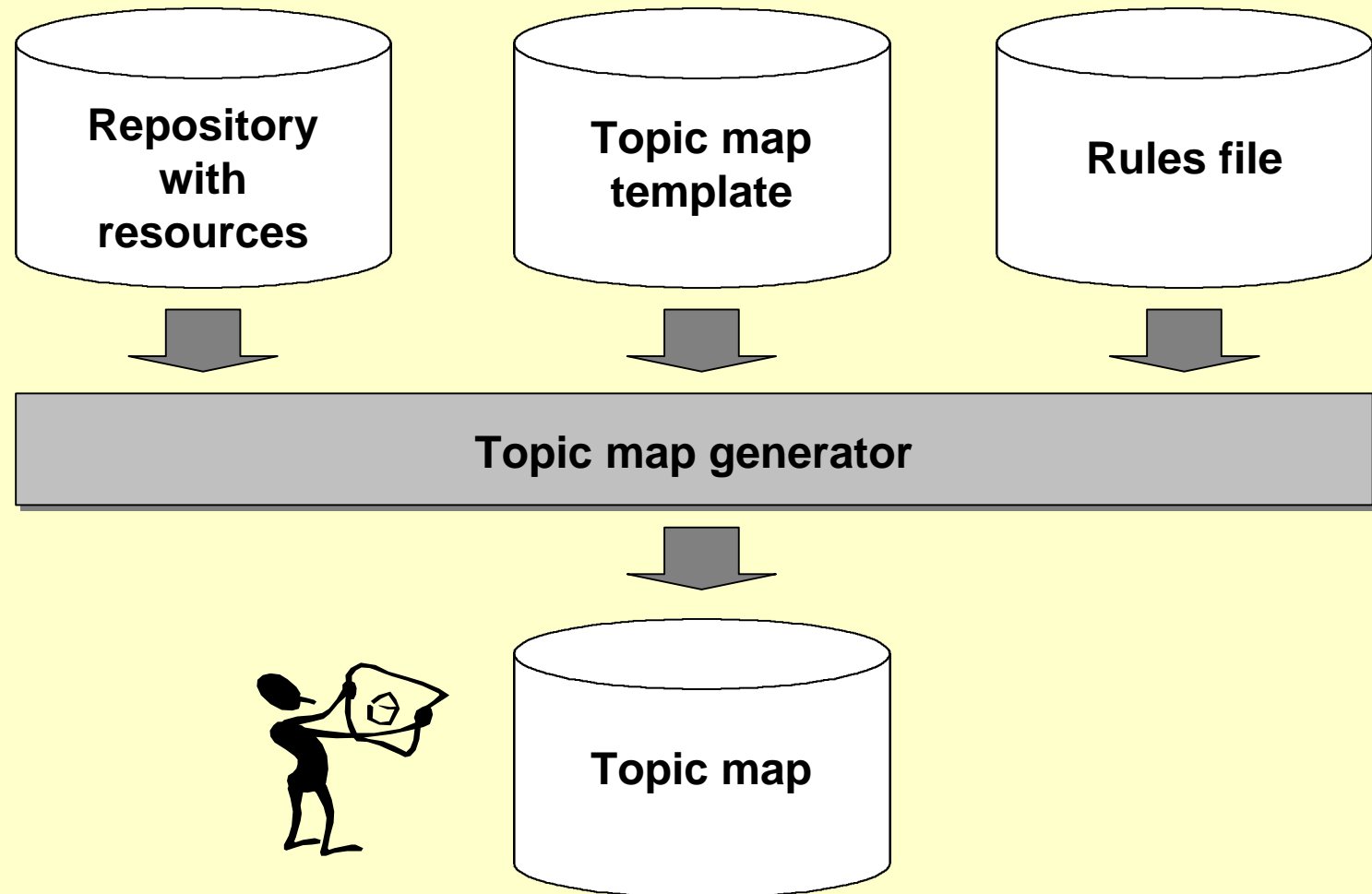
Automatic generation

- ❖ Applying a topic map to existing information resources
- ❖ Like *up translation* from legacy data to SGML/XML instances
- ❖ Uses existing information as hints for the initial creation of the map

Automatic generation
Basic setting

- ❖ **Topic map template**
- ❖ **info resources in repository (file system, database, Web)**
- ❖ **Metadata about info resources (e.g. name, format, classification code)**
- ❖ **Structure information in resources (SGML/XML elements)**
- ❖ **Script with generation rules**

Automatic generation Topic map generator



Automatic generation

Topics and occurrences

If resource **fulfills** metadata
<condition> **and/or**
contains <element> in <context>
containing <content>
then **create topic** of <type> **with name**
derived from metadata <field> **or**
from <element> in <context> **and**
create occurrence with <role> to
resource **or** <element> in
<context> in resource.

Automatic generation
Scopes

- ❖ **Scopes can be assigned to topics, names, occurrences**
- ❖ **Info for scopes can be extracted from resources and/or their metadata**
- ❖ **Result: Topic map with topics, names, occurrences, scopes**

Automatic generation
Associations

- ❖ **Very complex task because we have to build knowledge structures (AI)**
- ❖ **Pragmatic approach: use topic types and consistency constraints**
- ❖ **List of topics which are candidates for associations**
- ❖ **Creation of associations is done manually using this list**



Conclusions

Conclusions

- ❖ **ISO/IEC 13250 Topic maps**
 - ◆ Provides concepts to model knowledge structures
 - ◆ Needs further improvements / extensions
- ❖ **Topic map templates**
 - ◆ Logical container for the “declaration” part of a map
 - ◆ Modularization and re-use
 - ◆ Standardization of templates

Conclusions cont'd

❖ **Consistency checking with constraints**

- ◆ Manual checking is impossible
- ◆ Rule-based constraints control validation process

❖ **Automatic generation**

- ◆ Existing info resources are used to build the map by an automatic process
- ◆ Easy: topics, names, occurrences, scopes
- ◆ Very complex: associations



WordNet

WordNet

Lexical database for English developed by
Princeton University

- ❖ WordNet® is an on-line lexical reference system whose design is inspired by current psycholinguistic theories of human lexical memory.
- ❖ English nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexical concept.
- ❖ Different relations link the synonym sets.

WordNet Cont'd

- ❖ Thus, WordNet is a topic map (even it is not aware of it)
- ❖ WordNet demo ...
 - ◆ tree, question, shark, topic
- ❖ Reference:
<http://www.cogsci.princeton.edu/~wn>



<end>

**Thank you for
your attention!**

Questions?

</end>