Notes on Modeling of Places / Localizations

Martin Scholz

These notes try to summarize the discussions on the modeling of places and place names at the ICOM CIDOC Transdisciplinary Approaches in Documentation workshop in Shanghai, 2010 and earlier meetings of the working group.

Fuzzy Localizations

While today most politically motivated place appellations (e.g. Germany, England, EU, Berlin) have clear boundaries, many places have no clear boundaries or they are at least arbitrary chosen. There are several reasons for this:

- fuzzy geogr. or political boundaries: the Alps, Sahara, transition North Sea ↔ Atlantic, Sun Belt, Scandinavia
- imprecise knowledge: boundaries of China during the Tang Dynasty, finding place near XYZ, Vinland
- negligent or colloquial usage: Berlin ↔ Berlin metropolitan area

A fuzzyness-based localization mechanism similar to that for temporal positioning with the introduction of a place primitive would be a logical solution. Thus, each place would be linked with its maximal and minimal extent. Fuzzy localizations may also be modeled using the existing schema: The fuzzy place P is linked with its maximal extent P_{max} via *P89 falls within* and with its minimal extent P_{min} via *P89 contains*.

The latter modeling has the disadvantage that the specific semantics is not expressed and that the modeling of fuzzy places and the modeling of fuzzy times follows different approaches (one with and one without using a primitive), despite the clearly exposed analogies.

Modeling of Place Names with Temporal Positioning

Starting point:

- Places mostly are not labeled directly, but through entities present at the place at a particular time. The place is unambiguously identified through the tuple (appellation of the entity, time appellation). Examples: Poland 1920, Germany 1950-1989, Pangaea in the Upper Triassic.
- The entities can be of different nature, e.g. political (Germany), geographical (Alps), cultural (Franconia). However, they are never time appellations, epochs or the like. This means the tuple always has exactly one time appellation.

The following suggestions for modeling have been made so far. The suggestions do not exclude

each other.

Flat Appellation

The appellation of the entity and the time appellation are added to the place appellation without semantic itemization.

Advantages:
- Simplest way of place appellation
- Sufficient for scenarios without direct semantic reference to the time appellation or the entity
- Maximum freedom in appellation

Disadvantages:
- The underlying concept/pattern is ignored completely
- Reference to both components impossible. Modeling in retrospect is probably expensive and maybe only feasible manually.

Modeling of the Entity

Here, the weight is put on the entity standing behind the place appellation. The intuitive idea that "Poland 1920" and "Poland 2010" indeed do not describe the same place but both places are connected semantically ("Poland") is taken into account. The entity X standing behind the appellation is modeled accordingly. An event E (E5 Event) assigns X the expansion P (E53 Place) at the respective time T (E52 Time-Span). Suitable events are those who express a placement, for example E9 Move, E12 Production, etc. The appellation of the place AP itself is calculated based on the event: One creates the tuple (AX, AT) from the appellation for X and T. A further localization P1 at the point in time T1 != T may be added through a new Event E1. The place appellation AP1 is analogously (AX, AT1).

Advantages:
- "Clean", detailed modeling of the concept/pattern standing behind the location
- All localizations are connected through X
- Automatic calculation of alternate place appellations for all possible timespans between two events. Recognition of incorrect appellations before/after creation/destruction

Disadvantages:
- Modeling of the entity may be difficult and arbitrarily detailed. (Is "Germany" just E74 Group or also a state entity? How to model the entity "Sahara"?)
- Entity mostly not important for mere localization
- Algorithm necessary (for alternative locations only)

Combined place appellations

This approach leaves the entity itself unmodeled, just the two elements of the place appellation are modeled as independent appellations. While the appellation of the entity stays unspecific, the time appellation is modeled as such, whereby both can be accessed unambiguously at all times. An E53 Place P is being identified through E44 Place Appellation A_P. A_P is connected

through *P106 is composed of* with E41 Appellation A_X and E49 Time Appellation A_T, which are appellations for the entity and the timespan. Since A_X can never be an E49 Time Appellation, unambiguous assignment of appellations is assured. The wording of A_P can well differ from the concatenation of A_X and A_T: "Poland (1920)" (A_P) consists among others of "Poland" (A_E) and "1920" (A_T).

There is neither provision for the modeling of the entity nor of the timespan to avoid the above problems. Yet a modeling is entirely imaginable and maybe indeed purposeful for at least the timespans.

Advantages:
- The composition of the place appellation is reflected semantically.
- Minor modeling overhead.
- Easy sorting by appellation parts.

Disadvantages:
- If T is not modeled, no real temporal localization can take place, just a numeric-lexical sorting.
- The special tuple-semantics are not reflected through E44 Place Appellation.
- P106 is composed of reflects the semantics insufficiently. Through this property no differentiation between time and entity is made.

Special classes and properties for entity-time-locations

This approach takes into account the special semantics of place appellations through entity-time tuples. It extends the approach of the combined place appellations by introducing a new subclass of E44 Place Appellation: "Tuple Place Appellation". For this, two properties are introduced which both have the cardinality "exactly one": "*P106e refers to entity*" is a subproperty of *P106 is composed of* and has E41 Appellation as its range. "*P106t refers to time-span*" is a subproperty of *P106* as well and has E49 Time Appellation as its range.

See advantages and disadvantages above.