On Compositionality of ISO 25964 Hierarchical Relationships (BTG, BTP, BTI)

Vladimir Alexiev	Antoine Isaac	Jutta Lindenthal
Ontotext Corp	Europeana and VU Amsterdam	Independent consultant
Sofia, Bulgaria	The Hague, Netherlands	Lübeck, Germany
vladimir.alexiev@ontotext.com	aisaac@few.vu.nl	jutta.lindenthal@gmail.com

Knowledge Organization Systems have used different hierarchical relations for a long time. The ISO thesaurus standards (ISO 2788, now superseded by ISO 25964) distinguish between three kinds of hierarchical relationships: the generic, hierarchical whole-part and instance relationship, marked by the tags BTG, BTP, BTI. The latest ISO standard on thesauri (ISO 25964, 2011) includes a field *HierarchicalRelationship.role* in its data model. The corresponding OWL ontology expresses these relationships as properties *broaderGeneric*, *broaderPartitive*, and *broaderInstantial* respectively in the http://purl.org/iso25964/skos-thes namespace ('iso:' for brevity). Similar relations are used in actual data: some vocabularies hosted by *digiCULT-Verbund eG*, the German *Gemeinsame Normdatei*, the FinnONTO SKOS Extensions, and most recently, the *Getty Art & Architecture Thesaurus* (AAT).

Hierarchical relations are one of the most important relations in Knowledge Organization Systems (KOS). In view of the Linked Data scenario and vocabulary mappings, the logical grounding of relationships becomes increasingly important to ensure true interoperability. Considerable research is done in the field of subsumption and mereology in general, yet the compositionality of hierarchical relationships has not been investigated systematically so far. Compositionality matters with respect to transitive closure and inference in information retrieval vocabularies. It is a prerequisite for sensible search expansion, more specifically, search explosion over hierarchy chains. The exploration of compositionality of BTG, BTP, BTI, is a first step towards this end.

Our analysis builds on examples from the Getty vocabularies (AAT, TGN, ULAN), considering all nine possible compositions of the three kinds of relationships. We use property chains (denoted by "/") and analyze appropriate inferences ("→") by case. From the nine compositions, we found the following five to yield valid inferences as outlined above, provided that the semantic rules given in the ISO standard are observed.

On the left side BT*x means "BT* or BT*E":

BTGx/BTGx→BTGE
gvp:broaderGenericExtended owl:propertyChainAxiom
(gvp:broaderGenericExtended gvp:broaderGenericExtended).
If X is a *kind of* Y and Y is a *kind of* Z then X is a *kind of Z*.
BTGx/BTPx→BTPE
gvp:broaderPartitiveExtended owl:propertyChainAxiom
(gvp:broaderGenericExtended gvp:broaderPartitiveExtended).
If X is a *kind of* Y, which is *part of* Z then X is *part of* Z since X can play the role of Y.
BTPx/BTGx→BTPE
gvp:broaderPartitiveExtended owl:propertyChainAxiom
(gvp:broaderPartitiveExtended gvp:broaderGenericExtended).
If X is *part of* Y, which is a *kind of* Z then X is *part of* Z, since Z can play the role of Y.
BTPx/BTPx→BTPE
gvp:broaderPartitiveExtended owl:propertyChainAxiom

(gvp:broaderPartitiveExtended gvp:broaderPartitiveExtended).

If X is *part of* Y and Y is *part of* Z then X is *part of* Z. Note that transitivity for whole-part relationships only holds under certain conditions.

BTIx/BTGx→BTIE

gvp:broaderInstantialExtended owl:propertyChainAxiom

(gvp:broaderInstantialExtended gvp:broaderGenericExtended).

If X is an *instance of* Y and Y is a *kind of* Z then X is *an instance of* Z.

The 'iso:' properties have not seen wide deployment yet. As far as we know, there are very few datasets that use the ISO hierarchical relationship roles at all. One reason is the relative novelty of this ontology (created 2013-12-09), but we see two other reasons:

1. Improper mixing of different kinds of hierarchical relationships: The relations all are sub-properties of (contribute to) skos:broader, which unconditionally contributes to skos:broaderTransitive. But compositionality is not always appropriate, e.g. it does not apply for mixed paths of BTG and BTP. This means that skos:broaderTransitive makes less sense in many cases of qualified hierarchical relationships.

2. Inflexibility of the 'iso:' definition of hierarchical properties, inherited from SKOS: ISO 25964 restricts the hierarchical properties to skos:Concept, but important thesauri like the AAT need to use them also for items that the publishers have chosen to model as non-concepts (Facet, Hierarchy Name, and Guide Term). That is why non-standard properties, e.g. gvp:broaderPartitive ('gvp:' stands for *Getty Vocabulary Program Ontology*) had to be introduced for the AAT case.

However, we would like to assert an ISO relation when two AAT concepts are connected by an appropriate chain of BT* properties, interleaved by non-concepts. By defining Extended properties (BTGE, BTPE, BTIE), we remove this obstacle. The Extended relations capture these "appropriate chains", so we can fulfill this requirement. Analyzing compositions of the immediate "step" properties and the Extended properties, we determine which of these are appropriate. We call these properties "Extended" instead of Transitive, because not all of them are transitive.

In the case of property chains with interleaving concepts and non-concepts, ISO properties can be inferred from the Extended properties. We have to limit only to pairs of concepts, bypassing interleaving non-concept(s), and use the fact that we have already asserted skos:broader between such concepts.

In summary, we have obtained a set of rules by which ISO relationships can be inferred from the GVP ones.

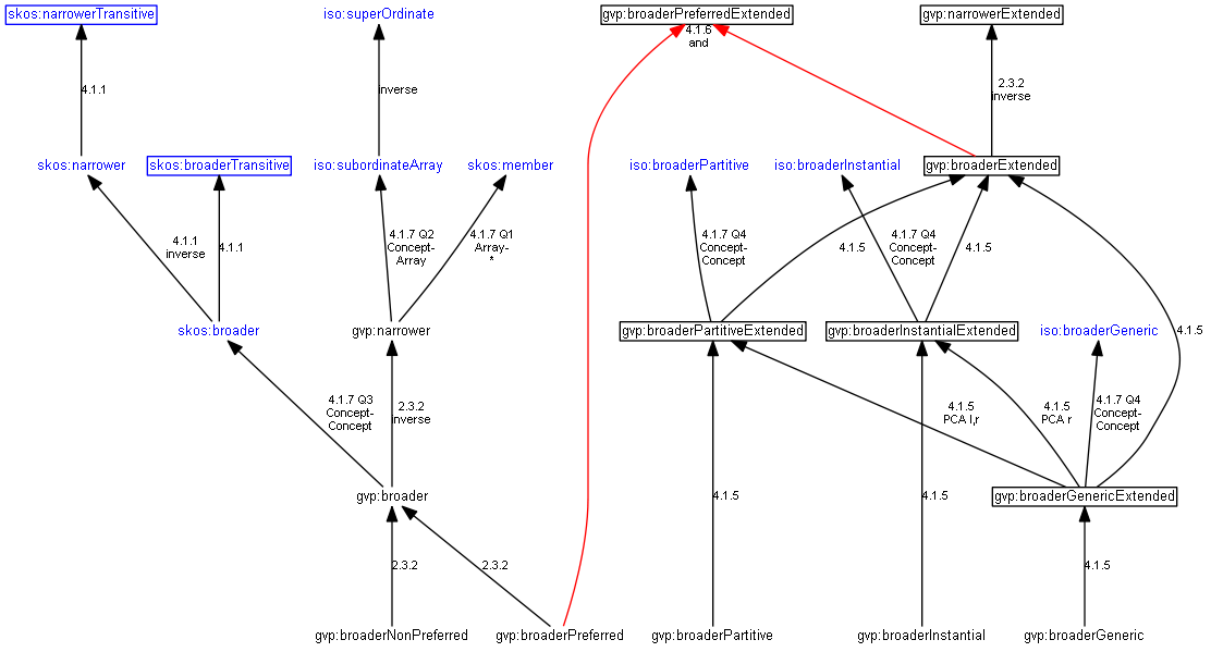

Figure 1: Inference of hierarchical relationships using Extended properties.
(The numbers refer to the Linked Open Data specification for the Getty Vocabularies).

The main purpose of a proper broader relation is to enable query expansion in information retrieval. Exploiting transitivity of hierarchical relationships is a prerequisite for this and other advanced search features. For example:

- If *Sofia* BTP *Bulgaria* BTP *Europe* then *Sofia* BTPE *Europe*. This enables a search for places in *Europe* to also find *Sofia*.
- If *Mt Athos* BTI *orthodox religious centers* BTG *Christian religious centers* BTG *religious centers* then *Mt Athos* BTIE *religious centers*. This enables a search for *religious centers* to also find *Mt Athos*.

Real-world implementations of such query expansion would allow the user to limit the length of the inference chain and to opt out of including instances, such as in the second example.

Besides addressing the issue of interoperability, our proposed Extensions to hierarchical thesaurus relationships may also open up new avenues for faceted retrieval which, in spite of its popularity, still rarely exploits semantic relationships within concept hierarchies.