A Structuralistic Semantics for Ontology Alignments

Christian Schäufler[1], Clemens Beckstein[1,2], and Stefan Artmann[2]

[1] Artificial Intelligence Group, University of Jena, Germany
[2] Frege Centre for Structural Sciences, University of Jena, Germany

1 Introduction

The definition of a formalism for ontology alignments is straightforward. Problems start when one attempts to define what they mean. The existing semantics [3] all depend on certain preconditions for the alignments to make sense — e.g., that the ontology domains are empirically given, that the involved ontologies are compatible, and that either the domains are identical or, there is a practicable way for specifying a domain relation. Those preconditions, of course, are not always empirically justifiable. With scientific structuralism [1] another approach for interpreting inter-theoretical relations has been put forward. With the help of this framework it is possible to give an exact description of the formal context in which the distributed alignment semantics works.

2 Structuralistic Interpretation of Alignments

A structuralistic interpretation of an ontology O (see [2]) need not consist of just one undifferentiated domain D, but may consist of several domains $D_1, \ldots, D_n$. Structuralists call the factors of such a domain structure *(domain) types*. The domain terms $D_1, \ldots, D_n$ are given by a set of disjunct primitive concepts that, according to the ontological axioms of O, are just below the top concept $\top$ of the ontology. In analogy to the distinction between reduced and reducing theory we assign roles to the ontologies that are involved in an alignment. Alignments in our view always assume a specific flow of information — from a *foreign* knowledge base W over O to the initial inquirer with a commitment to O'. Domain inclusions relate the domains of the ontologies to be aligned. Despite the domain relation r in contextualized distributed semantics, a *domain inclusion* sets whole (echelons of) domains in relation. An echelon set on some base sets is a set resulting from arbitrary product or power-set-operations of the base sets or echelon sets (of the base sets).

Let m and m' be two models that satisfy a correspondence $e\,R\,e'$. Because the domains $D_1, \ldots, D_n$ match the top-level primitive concepts of ontology O, the extension of an ontology element e in O is a subset of exactly one domain D_i (or a pair of domains if e is a role) and e' that of exactly one D'_j. For an interpretation of the correspondence, both ontology elements have to be interpreted in the same domain. W.l.o.g. O' is the querying ontology, so the interpretation of both e and

e' should take place in D'_j. A domain inclusion between D_i and D'_j is either a domain inclusion $D_i \subseteq S$ where S is an echelon-set that actually uses the base set D'_j or $D'_j \subseteq S'$ where S' is an echelon-set that actually uses the base set D_i. The following disjoint cases can be distinguished:

1. *$D_i \subseteq D'_j$:* The extension of e is a subset of D'_j. e can be interpreted in the same domain D'_j as e'.

2. *$D'_j \subseteq D_i$:* To assure an interpretation in simple distributed semantics, it is necessary that in addition the domain inclusion $e^m \subseteq D'_j$ holds.

3. *$D_i \subseteq S$ with $S \neq D'_j$:* An interpretation in domain D'_j is possible if the elements of D_i can be ontological projected to D'_j: $p(D_i) \subseteq D'_j$, where p is that mapping that projects S to D'_j.

4. *$D'_j \subseteq S'$ with $S \neq D_i$:* The extension of e' consists of complex individuals from the point of view of the queried ontology O. Elements of e are missing some properties in order to be interpretable as elements of D'_j. An alignment containing such a domain inclusion does not have a model.

5. *Otherwise:*D_i and D'_j are ontologically incompatible. Even if some individuals do appear in both domains, there is no general rule holding for every individual. Even by introducing additional properties, neither D'_j can be reconstructed from D_i nor the other way around. For incompatible domains an interpretation of the correspondence is therefore only possible wrt. the intersection $D_i \cap D'_j$, i.e. if the domain inclusions $e^m \subseteq D'_j$ and $e^{m'} \subseteq D_i$ hold.

Each of the cases specifies a (maybe empty or unsatisfiable) additional precondition in the form of domain inclusions that have to be fulfilled for the alignment to have a model in the structuralistic sense. And whenever these *alignment specific domain inclusions* hold, the corresponding case can be reduced to an interpretation like that of the simple distributed semantics.

3 Conclusion

Structuralism points the way to a new semantics for ontology alignments that rests on structured domains, a distinguished direction of the alignment, and compatibility constraints in the form of term-by-term domain inclusions. Domain inclusions are much simpler to specify in practice than domain relations as they are used in the contextualized distributed semantics: they are expressed wrt. a given set of domain terms and not wrt. individuals of an empirical domain and many of them result as a byproduct of the structuralistic alignment process.

References

1. Balzer, W., Moulines, C.U., Sneed, J.D.: An architectonic for science - the structuralist program. Reidel, Dordrecht (1987)
2. Schäufler, C., Artmann, S., Beckstein, C.: A structuralistic approach to ontologies. In: KI 2009: Advances in Artificial Intelligence. Lecture Notes in Computer Science, vol. 5803, pp. 363–370. Springer, Berlin / Heidelberg (2009)
3. Zimmermann, A., Euzenat, J.: Three semantics for distributed systems and their relations with alignment composition. In: The Semantic Web - ISWC 2006. Lecture Notes in Computer Science, vol. 4273, pp. 16–29. Springer Berlin / Heidelberg (2006)