The Effect of Context on Semantic Similarity Measurement

Carsten Keßler[1], Martin Raubal[2], and Krzysztof Janowicz[1]

[1] Institute for Geoinformatics, University of Münster, Germany
{carsten.kessler|janowicz}@uni-muenster.de
[2] Department of Geography, University of California at Santa Barbara, USA
raubal@geog.ucsb.edu

Abstract. Similarity measurement is currently being established as a method to explore content on the Semantic Web. Semantically annotated content requires formal concept specifications. Such concepts are dynamic and their semantics can change depending on the current context. The influence of context on similarity measurement is beyond dispute and reflected in recent similarity theories. However, the systematics of this influence has not been investigated so far. Intuitively, the results of similarity measurements should change depending on the impact of the current context. Particularly, such change should converge to 0 with a decreasing impact of the respective contexts. To hold up to this assertion, a quantification of the impact of context on similarity measurements is required. In this paper, we use a combination of the SIM-DL theory, which measures similarity between concepts represented using description logic, and a context model distinguishing between internal and external context to quantify this impact. The behavior of similarity measurements within an ontology specifying geospatial feature types is observed under varying contexts. The results are discussed with respect to the corresponding impact values.

1 Introduction and Motivation

Information integration and retrieval are central aspects of the Semantic Web. It has been argued that cognitively plausible methods for achieving such integration and retrieval require the employment of semantic similarity measurement [1]. The influence of context on semantic similarity measurement is a well-known phenomenon that has long been observed in psychological experiments [2]. As an example, consider an information retrieval scenario where a user is looking for buildings that are similar to *churches*. For a similarity-based ranking of the results, the best matches depend on the context: within a *religion* context these could be *cathedral*, *chapel* and *convent*, whereas for a *sightseeing* context they could be *palace*, *castle* and *museum*.

Although the results of human similarity ratings depend on the context, and this dependency is also reflected in recent similarity theories [3–5], the nature of this influence and its actual impact have not been subject to thorough research yet. Previous work looked primarily at how contextual information can be employed to adjust the results of a similarity measurement. In this paper, we present observations of similarity measurements under changing contexts using the description logic based SIM-DL theory [5]. The long-term objective of our research is a generic context model for similarity measurement, which represents the important characteristics of context. This envisioned model can then be used for the development of similarity-based applications, particularly for the assessment of the impact of different context parameters. Building such context model requires a better understanding of the interaction between context and similarity measurement.

The focus of this paper lies on observing and quantifying the change in the results of a similarity measurement under different contexts. We introduce and formalize the notion of *impact* for contexts: the more the context changes the involved concepts, the larger its impact. The results of a similarity measurement should directly depend on the impact the context has on the compared concepts. In other words, if the same similarity measurement is performed under a series of contexts with decreased impact, the change in the similarity values will converge to 0. To enable such behavior, an impact measure for contexts is required. In the following, we introduce such an impact measure, based on the SIM-DL theory and a context model consisting of an internal context and a set of external context rules. A scenario from geographic feature type lookup is used to test the model with varying concept pairs and for different contexts.

The remainder of this paper is organized as follows: we first present related work on semantic similarity measurement and context, followed by an introduction of the SIM-DL theory, and the formal specification of context impact. Section 3 describes the use case and its context formalization. In Section 4, the resulting similarity and impact values are evaluated, and the results discussed. Section 5 presents conclusions and gives directions for future research.

2 Similarity Measurement and Context

This section presents previous work on similarity measurement and context, and introduces the SIM-DL theory. The context representation and the according formalization of context impact are demonstrated.

2.1 Semantic Similarity Measurement

The notion of similarity was established in psychology to determine why and how entities are grouped to categories, and why some categories are comparable to each other while others are not [2, 6]. The main challenge with respect to semantic similarity measurement is the comparison of meanings. A language must be specified to represent the nature of entities and metrics are needed to determine how (conceptually) close these compared entities are. While entities can be expressed in terms of attributes, the representation of entity types is more complex [7]. Depending on the expressivity of the representation language, types are specified as sets of features, dimensions in a multidimensional space, or formal restrictions specified on sets using various kinds of description logic. Similarity is measured between entity types, which are representations of concepts in human minds, therefore the results depend on what is said (in terms of computational representation) about these types. This again depends on the employed representation language, and therefore most similarity measures cannot be compared. Besides the question of representation, context is another major challenge for similarity assessments. In general, meaningful notions of similarity cannot be determined without defining in respect to what similarity is measured [2, 8, 9].

Similarity measurements have been investigated as a method for information retrieval in the semantic web over the last years [1, 10]. Stroulia and Wang [11] developed a context-free similarity measure for Web services based on Web Service Description Language specifications. Based on Tversky's feature model [12], Rodríguez and Egenhofer [3] built an extended model called Matching Distance Similarity Measure (MDSM) that supports a basic context theory, automatically determined weights, and asymmetry. Raubal and Schwering [13, 14] used conceptual spaces [15] to implement models based on distance measures within geometric space. Several measures [5, 16, 17] were developed to close the gap between ontologies described by various kinds of description logic, and similarity theories that had not been able to handle the expressivity of such languages.

2.2 Similarity Theory: SIM-DL

SIM-DL [5, 18, 19] is an asymmetric and context aware similarity measurement theory used for information retrieval within an ontology (or several ontologies using the same shared vocabulary [20]). An early implementation of SIM-DL as DIG-compliant [21] server is available at http://sim-dl.sf.net. The latest version supports comparison between concepts specified using the expressive description logic $\mathcal{ALCHQ}$ [19].

In SIM-DL, similarity between concepts in canonical form [5, 22] is measured by comparing their definitions for overlap, where a high level of overlap indicates high similarity and vice versa. In description logic (complex) concepts are specified based on primitive concepts and roles using language constructors such as intersection, union, and existential quantification. Hence, similarity is defined as a polymorphous, binary, and real-valued function X $\times$ Y $\rightarrow$ R[0,1] providing implementations for all language constructs offered by the used description logic. The overall similarity between concepts is the normalized (and weighted) sum of the single similarities calculated for all parts (i.e., superconcepts) of the concept definitions. A similarity value of 1 indicates that the compared concepts cannot be differentiated, whereas 0 implies total dissimilarity. As most feature and geometric approaches, SIM-DL is a asymmetric measure, i.e. the similarity $sim(C_s, C_t)$ is not necessarily equal to $sim(C_t, C_s)$. The comparison of two concepts depends therefore not only on their descriptors but also on the direction in which both are compared. Further details on SIM-DL and the involved similarity functions are given in [5, 19].

2.3 Previous Work on Context

Context has been investigated from the perspectives of different research areas such as ubiquitous computing, interoperability, automatic metadata generation and web search. Accordingly, any definition of context largely depends on the field of application[1]. Concerning research on context for similarity measurement [3–5], existing context definitions are often tailored to specific similarity theories, and the context is mostly used to select the domain of application, i.e. a set of concepts that is taken into account for the similarity measurement. Moreover, context is used to assign weights to (parts of) the different concepts or instances within the domain of application. To clarify what we refer to as a similarity measurement's context, we use the following definition from [9]:

Definition 1. *A similarity measurement's context is any information that helps to specify the similarity of two entities more precisely concerning the current situation. This information must be represented in the same way as the knowledge base under consideration, and it must be capturable at maintainable cost.*

While capturability and cost are not relevant for this research, it is important that context only refers to such information that has an impact on the similarity measurement. Furthermore, it must be represented in the same way as the given knowledge base. For reasons of simplicity, we assume here that the context only refers to parts of the knowledge base at hand, as described in the following section.

2.4 Context and Impact Specification

The domain of application alone is not always sufficient to reflect the context of a similarity measurement; for example, a domain of application, such as transportation, cannot be used to specify that, by law, it is not allowed for trucks to drive on the German *Autobahn* on Sundays—hence, an external rule is required that removes the superconcept *NavigableByTrucks* from the *Autobahn* concept on Sundays. Accordingly, we define a context K as a combination of *internal* and *external* context (eq. 1): The internal context c_{int} is a concept which specifies the domain of application, i.e. all concepts subsumed by c_{int}, such as *NavigableByTrucks*. The external context is a set of rules R that allows for the modification of the concepts selected via c_{int}. Every rule consists of a condition that specifies the circumstances under which the rule is activated[2], a number of modifying concepts c_m, and the affected concepts c_a to which these modifications apply (eq. 2). Every modification either adds $(+)$ a superconcept to the affected concepts by intersection, or removes it $(-)$. In the special case where a negated superconcept is present, and the same (non-negated) superconcept is added via

[1] See [23] for an overview of different research areas investigating context, and the according definitions.

[2] A mechanism for the automatic selection of the appropriate rules R based on the conditions is required for an implementation, but out of scope for this paper.

+, the negation is overridden. Note that these changes are only temporary and revoked after the similarity measurement.

$$K = \langle c_{int}, \{R_1, ..., R_n\}\rangle \tag{1}$$

$$R : condition \longrightarrow \langle\{\pm c_{m_1}, ..., \pm c_{m_n}\}, \{c_{a_1}, ..., c_{a_n}\}\rangle \tag{2}$$

To allow statements about the impact of a context on a similarity measurement, we introduce a formal measure that quantifies this impact (eq. 3): A context's impact on a similarity measurement is defined as the overall change the corresponding context rules cause to the search- and target concepts c_s, c_t. For every modifying concept c_m, the absolute change is quantified in terms of how many of the superconcepts are changed (added or removed) in c_s or c_t. Beyond this, the kind of change is reflected: if the application of c_m makes c_s and c_t more similar, the absolute value is counted positive; if it makes them less similar, the value is counted negative. Whether a single c_m is positive or negative is determined according to Table 1, which lists all possible combinations for adding or removing superconcepts, depending on whether they are already part of c_s or c_t (or both). The combined measure (eq. 3) takes into account that a rule which makes two concepts more similar, and one which makes them less similar, may compensate for each other if both appear in the same context. The outcoming impact measures range from 0, where the concepts under consideration are not changed, to 1, where all of the original superconcepts are removed.

$$Imp(K, c_s, c_t) = \sum \frac{\pm c_m}{|\{c_a | c_a \sqsupseteq c_s \sqcup c_t\}|} \tag{3}$$

Table 1. Possible combinations of adding or removing superconcepts from search concepts, and target concepts respectively. The contents of the table show the development of similarity under the preconditions given in the header; an increase in similarity is marked +, decrease is marked −, no change by 0.

	$c_m \sqsupset c_s, c_m \sqsupset c_t$	$c_m \sqsupset c_s, c_m \not\sqsupset c_t$	$c_m \not\sqsupset c_s, c_m \sqsupset c_t$	$c_m \not\sqsupset c_s, c_m \not\sqsupset c_t$
$+c_s$	0	0	+	−
$+c_t$	0	+	0	0
$-c_s$	−	+	0	0
$-c_t$	−	0	0	0
$+c_s, +c_t$	0	+	+	+
$+c_s, -c_t$	−	0	−	−
$-c_s, +c_t$	−	+	0	0
$-c_s, -c_t$	−	+	0	0

As mentioned in the introduction, the results of a similarity measurement should intuitively depend on the impact the context has on the compared concepts: when the context's impact increases, the change in the measurement's result should also increase. In other words, if the same similarity measurement is performed under a series of contexts with decreased impact, the change in the similarity values should converge to 0. The specification of this behavior (eq. 4) requires a standard context K_{std}, consisting of totality (i.e. the whole ontology, $\top$) as the internal context and without external context (i.e. without any rules): $\langle \top, \{\emptyset\} \rangle$.

$$\lim_{Imp(R_K) \to 0} sim_K(a, b) = sim_{K_{std}}(a, b) \tag{4}$$

3 Use Case: Geographic Feature Type Lookup

In this section, we present an application scenario involving a feature type ontology. Different geographic feature types, as represented in the Web Ontology Language (OWL), are compared under different contexts.

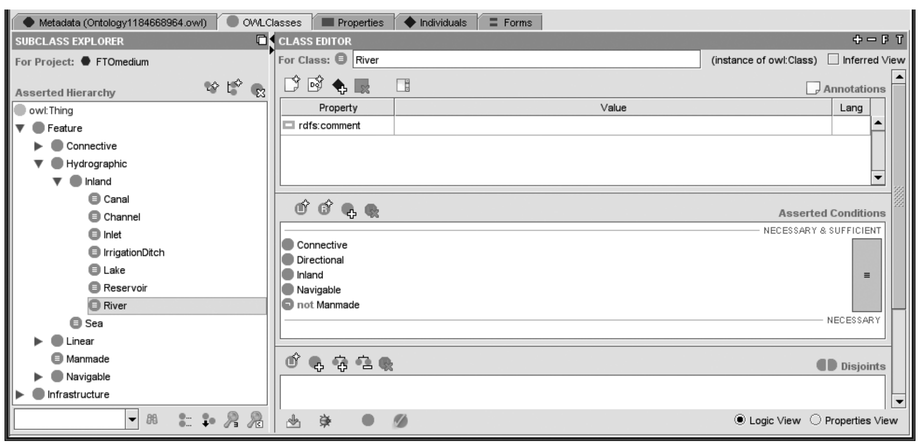

Fig. 1. Snapshot of the feature type ontology (see http://sim-dl.sf.net/downloads/).

3.1 Feature Type Ontology

Gazetteers are place name directories that make use of type lookup functionality to determine a geographic feature's type, such as *Road*, *Country* or *City*. The feature types are often classified in semi-formal feature type thesauri. However, similarity measurement among different feature types requires a formal description of the types. In the following, we refer to a feature type ontology for hydrographic features (figure 1) that was created based on the Alexandria Digital Library Feature Type Thesaurus[3]. Once complete, such an ontology allows

[3] http://www.alexandria.ucsb.edu/gazetteer/FeatureTypes/ver070302/index.htm.

for similarity-based queries to a semantic geo-webservice [19]. Since any complex similarity tasks build upon the comparison of concept pairs, we use SIM-DL and the context definition presented in section 2.4 to measure similarities between *River* and *Lake*, *Lake* and *Reservoir*, *Canal* and *River*, and between *Reservoir* and *Canal*.

3.2 Contexts within the Scenario

The internal context c_{int} is set fixed to *Hydrographic* for the use case. We introduce a set of context rules that can be combined to different contexts[4]; for example, R_1 states that seas and lakes are not navigable during a storm, whereas R_4 states that it is allowed to navigate in reservoirs in the case of an emergency:

R_1: Storm $\longrightarrow \langle\{-Navigable\}, \{Sea, Lake\}\rangle$
R_2: Flooding $\longrightarrow \langle\{-Linear\}, \{River, Channel, IrrigationDitch\}\rangle$
R_3: Night-time $\longrightarrow \langle\{-Navigable\}, \{Canal\}\rangle$
R_4: Emergency $\longrightarrow \langle\{+Navigable\}, \{Reservoir\}\rangle$

Based on R_1-R_4, we define the following scenarios referring to different situations such as emergency at night (K_1), storm at night (K_2) and stormtide (K_3):

$K_1 = \langle Hydrographic, \{R_3, R_4\}\rangle$
$K_2 = \langle Hydrographic, \{R_1, R_3\}\rangle$
$K_3 = \langle Hydrographic, \{R_1, R_2\}\rangle$

4 Calculation of Context Effects and Discussion

In this section, we calculate and discuss the similarity and impact values for the example search and target concepts as well as context scenarios presented above.

4.1 Calculation with SIM-DL and External Context

Similarity in SIM-DL is calculated as the number of superconcepts the target concept c_t shares with the search concept c_s, divided by the number of superconcepts of c_s for standardization. For example, the similarity of River (c_s) to Lake (c_t) is $\frac{3}{6}$ (0.5, see first line of results in table 2), as River is represented by 6 superconcepts, of which Lake shares 3. This calculation is valid under K_{std}; if the context changes, the similarity may also change. For example, K_2 removes the superconcept Navigable from Lake, so that there are only two common superconcepts left and similarity changes to $\frac{2}{6}$ (0.33). On the other hand, K_1 does not affect the similarity of River to Lake because these concepts are not part of R_3 and R_4.

All similarity results in table 2 are also annotated with the impact values as calculated according to eq. 3. For example, the impact of K_2 on the similarity

[4] For reasons of readability, we only use combinations of two context rules in this paper; however, an arbitrary number of rules can be combined in principle.

of River to Lake is $-\frac{1}{6}$ (-0.17), since one of the 6 superconcepts of River $\sqcup$ Lake is removed. The impact is negative, because a superconcept is removed from c_t which is a superconcept of both c_s and c_t, resulting in a negative impact on the overall similarity (see first column, fourth line in table 1). Table 2 gives an overview of all similarity and impact values for the concept pairs under consideration.

Table 2. Similarity results for the three different external contexts. The internal context was set fixed to Hydrographic.

Search concept C_s	**Target concept** C_t	K_1	K_2	K_3	K_{std}
River	Lake	0.5	0.33	0.4	0.5
Impact		0	-0.17	-0.03	0
Lake	River	0.6	0.5	0.5	0.6
Impact		0	-0.17	-0.17	0
Lake	Reservoir	0.8	0.75	0.75	0.6
Impact		+0.17	+0.17	+0.17	0
Reservoir	Lake	0.8	0.6	0.6	0.6
Impact		+0.17	0	0	0
Canal	River	0.75	0.75	0.6	0.8
Impact		-0.14	-0.14	-0.14	0
River	Canal	0.5	0.5	0.6	0.66
Impact		-0.14	-0.14	-0.17	0
Canal	Reservoir	0.38	0.38	0.3	0.3
Impact		0	+0.17	0	0
Reservoir	Canal	0.5	0.3	0.3	0.3
Impact		+0.17	0	0	0

4.2 Discussion of Results

The similarity values produced by SIM-DL using the combined context model introduced in section 2.4 generally appear plausible. For example, $sim(Lake, Reservoir)$ increases when the Reservoir becomes Navigable (K_1), or if the Lake is not Navigable (K_1, K_2). Note that K_2 and K_3 do not affect the similarity in the inverse direction, since the according rule R_1 only makes the Lake not navigable, i.e. the rule removes a superconcept from c_s that did also not match c_t before the rule was applied, therefore there is no change. It must be pointed out that leaps in the similarity values under different contexts, such as in the comparison of Reservoir and Lake, are for the most part due to the small number of concepts contained in the ontology: since every concept is only described by a small number of superconcepts, every modification to this set of superconcepts causes comparably large changes in the similarity values. The observations made in this paper need to be verified for more complex ontologies in the future.

Concerning the calculation of the impact values, the behavior as described in eq. 4 cannot be observed. While the tendency for the impact is generally correct, the actual impact value does not correlate with the corresponding change in similarity. For example, K_1 has an impact of $+0.17$ on $sim(Reservoir, Canal)$, as shown in the last line of table 2. The according similarity value increases by 0.2 with respect to the same comparison under K_{std}. At the same time, $sim(River, Lake)$ decreases by 0.1 under K_3, where the corresponding impact value is as low as -0.03. These results point to the fact that the impact model introduced in section eq. 3 may be too loosely coupled to the similarity theory.

5 Conclusions and Future Work

This paper has introduced a method, which quantifies the impact of context on the results of semantic similarity measurements. SIM-DL produces plausible results, even for ontologies that use only a small number of primitives for concept description. The separation between internal and external context allows for the addition of conditional rules that cannot be expressed within the ontology. The notion of context impact presented in this paper, however, does not fully correspond to the expected behaviour. Although the general tendency of the impact values corresponds to the trend of the change in similarity values, the quantification of this change is not yet reflected in the impact value. To solve this issue, further research on different strategies for impact specification is required. For example, knowledge about individuals (via assertions) [16] could be used to make more precise statements about the impact of a given context.

Beyond this, the limitation of the current model to the methods for adding and removing concepts needs to be enhanced by more sophisticated ways of assigning weights to the superconcepts. This would allow for rules such as "Glaze $\longrightarrow \langle\{Navigable : 0.3\}, \{Road\}\rangle$". This weighting would also allow for resolution of contradicting rules in the same context—a case that cannot be handled so far. To allow such rules, methods for the determination of the weights and the corresponding application rules are needed. Moreover, future work should focus on how the developed strategies can be generalized and transferred to other concept representations and similarity theories. The behavior observed for concept pairs in this paper also needs to be compared to similarity rankings and tested for cognitive plausibility in a human subjects test.

Acknowledgements

This research is partly funded by the German Research Foundation (DFG) under the project title "Semantic Similarity Measurement for Role-Governed Geospatial Categories", see http://sim-dl.sf.net.

References

1. Gärdenfors, P.: How to make the semantic web more semantic. In Varzi, A.C., Vieu, L., eds.: Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004). IOS Press, Amsterdam (2004) 17–34
2. Medin, D., Goldstone, R., Gentner, D.: Respects for similarity. Psychological Review **100**(2) (1993) 254–278
3. Rodríguez, A., Egenhofer, M.: Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure. International Journal of Geographical Information Science **18**(3) (2004) 229–256
4. Albertoni, R., Martino, M.D.: Semantic similarity of ontology instances tailored on the application context. In Dey, A.K., Kokinov, B., Leake, D., Turner, R., eds.: ODBASE - OTM Conferences. Volume 4275 of Lecture Notes in Computer Science. Springer (2006) 1020–1038
5. Janowicz, K.: SIM-DL: Towards a semantic similarity measurement theory for the description logic $\mathcal{ALCNR}$ in geographic information retrieval. In Meersman, R., Tari, Z., Herrero, P., eds.: SeBGIS 2006, OTM Workshops 2006. Volume 4278 of Lecture Notes in Computer Science. Springer, Berlin (2006) 1681–1692
6. Goldstone, R., Son, J.: Similarity. In Holyoak, K., Morrison, R., eds.: Cambridge Handbook of Thinking and Reasoning. Cambridge University Press (2005) 13–36
7. Donnelly, M., Bittner, T.: Spatial relations between classes of individuals. In Cohn, A., Mark, D., eds.: Spatial Information Theory - International Conference, COSIT 2005, Ellicottville, NY, USA. Volume 3693 of Lecture Notes in Computer Science. Springer, Berlin (2005) 182–199
8. Goodman, N.: Seven strictures on similarity. In Goodman, N., ed.: Problems and projects. Bobbs-Merrill, New York (1972) 437–447
9. Keßler, C.: Similarity measurement in context. In Kokinov, B., Richardson, D., Roth-Berghofer, T., Vieu, L., eds.: Sixth International and Interdisciplinary Conference on Modeling and Using Context. Volume 4635 of Lecture Notes in Artificial Intelligence. Springer-Verlag Berlin Heidelberg (2007) 277–290
10. Rissland, E.: Ai and similarity. IEEE Intelligent Systems **21**(3) (2006) 39–49
11. Stroulia, E., Wang, Y.: Structural and semantic matching for assessing web-service similarity. International Journal of Cooperative Information Systems **14** 407–437
12. Tversky, A.: Features of similarity. Psychological Review **84**(4) (1977) 327–352
13. Raubal, M.: Formalizing conceptual spaces. In Varzi, A., Vieu, L., eds.: Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004). Volume 114 of Frontiers in Artificial Intelligence and Applications. IOS Press, Amsterdam, NL (2004) 153–164
14. Schwering, A., Raubal, M.: Spatial relations for semantic similarity measurement. In Akoka, J., Liddle, S., Song, I.Y., Bertolotto, M., Comyn-Wattiau, I., vanden Heuvel, W.J., Kolp, M., Trujillo, J., Kop, C., Mayr, H., eds.: Perspectives in Conceptual Modeling: ER 2005 CoMoGIS Workshop, Klagenfurt, Austria. Volume 3770 of Lecture Notes in Computer Science. Springer, Berlin (2005) 259–269
15. Gärdenfors, P.: Conceptual Spaces - The Geometry of Thought. Bradford Books, MIT Press, Cambridge, MA (2000)
16. d'Amato, C., Fanizzi, N., Esposito, F.: A dissimilarity measure for $\mathcal{ALC}$ concept descriptions. In: Proceedings of the 2006 ACM Symposium on Applied Computing (SAC), Dijon, France (2006) 1695–1699

17. Borgida, A., Walsh, T., Hirsh, H.: Towards measuring similarity in description logics. In: Proceedings of the 2005 International Workshop on Description Logics (DL2005). Volume 147 of CEUR Workshop Proceedings. CEUR, Edinburgh, Scotland, UK (2005)
18. Janowicz, K.: Similarity-based retrieval for geospatial semantic web services specified using the web service modeling language (wsml-core). In Scharl, A., Tochtermann, K., eds.: The Geospatial Web - How Geo-Browsers, Social Software and the Web 2.0 are Shaping the Network Society. Springer, London (2007) 235–246
19. Janowicz, K., Keß ler, C., Schwarz, M., Wilkes, M., Panov, Espeter, M., Bäumer, B.: Algorithm, implementation and application of the sim-dl similarity server. In Fonseca, F., Rodríguez-Tatsets, A., eds.: Second International Conference on GeoSpatial Semantics (GeoS 2007), Mexico City, Mexico, Springer (forthcoming 2007)
20. Lutz, M., Klien, E.: Ontology-based retrieval of geographic information. International Journal of Geographical Information Science **20**(3) (March 2006) 233–260
21. Bechhofer, S.: The dig description logic interface: Dig/1.1. In: DL2003 Workshop, Rome (2003)
22. Horrocks, I.: Implementation and optimization techniques. In Baader, F., ed.: The description logic handbook: theory, implementation, and applications. Cambridge University Press, New York, NY, USA (2003) 306–346
23. Bazire, M., Brézillon, P.: Understanding context before using it. In Dey, A.K., Kokinov, B., Leake, D., Turner, R., eds.: Modeling and Using Context - 5thInternational and Interdisciplinary Conference CONTEXT 2005. Volume 3554 of Lecture Notes in Computer Science. Springer, Paris, France (2005) 29–40