Kinds of Contexts and their Impact on Semantic Similarity Measurement

Krzysztof Janowicz
Institute for Geoinformatics, University of Muenster, Germany
janowicz@uni-muenster.de

***Abstract*—Semantic similarity measurement gained attention over the last years as a non-standard inference service for various kinds of knowledge representations including description logics. Most existing similarity measures compute an undirected overall similarity, i.e., they do not take the context of the similarity query into account. If they do, the notion of context is usually reduced to the selection of particular concepts for comparison (instead of comparing all concepts within an examined ontology). The importance of context in deriving meaningful similarity judgments is beyond question and has been examined within recent research. This paper argues that there are several kinds of contexts. Each of them has its own impact on the resulting similarity values, but also on their interpretation. To support this view, the paper introduces definitions for the examined contexts and illustrates their influence by example.**

I. Introduction and Motivation

Semantic similarity measurement plays an increasing role in information retrieval and organization. Beside classical knowledge organization systems such as gazetteer services, similarity is also used within mobile decision support systems such as the pedestrian navigation service Utopian[1]. The benefit of similarity lies in delivering a ranked list of alternatives for the user's query if no exact match is available. This includes geographic feature types in case of gazetteer services as well as alternative routes or activity recommendations for decision support systems. One major shortcoming of computer based similarity judgments is that the results do not necessarily fulfill all user requirements. This is mostly caused by a lack of context information. To compute similarity judgments comparable to those of human users, a machine has to be able to adjust its knowledge representation and weightings to the user's task. Existing similarity theories either ignore the influence of context information or reduce the notion of context to a restriction of the domain of discourse.

In this work, we argue that there are several context types which have to be addressed during similarity measurement. These contexts have impact on both the measurement process and the later interpretation of similarity values. While some contexts can be inferred [1] or explicitly stated by the user, other kinds are difficult to capture. The resulting classification is adoptable to most similarity theories, however we focus on those developed for description logics here. The context types proposed in this paper are relevant for similarity measurement, for further classifications from other application and research domains see [2]–[4].

[1]Software & documentation available at www.utopian-online.de

The paper is structured as follows: first, a brief overview about similarity and context research is given. Second, the proposed kinds of contexts are introduced in detail. Next, their influence on similarity is demonstrated using a simplified example from hydrology[2]. An extended version of the SIM-DL plug-in for the Protégé ontology editor is introduced, to illustrate how the proposed contexts can be integrated within an existing reasoning infrastructure. Finally, conclusions and directions for further work are presented.

II. Related Work

This section gives a brief overview about semantic similarity measurement and points to some recent research on context related to similarity and description logics.

A. Semantic Similarity

The notion of similarity originated in psychology and has been established to determine why and how entities are grouped into categories, and why some categories are comparable to each other while others are not [5], [6]. The main challenge with respect to *semantic* similarity measurement is the comparison of meanings as opposed to purely structural comparison. A language has to be specified to express the nature of entities and functions are needed to determine how (conceptually) close the compared entities are. While entities can be expressed in terms of attributes, the representation of entity types is more complex. 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 logics. While some representation languages have an underlying formal semantics, the grounding of several representation languages remains on the level of an informal description.

Because similarity is measured between entity types which are representations of concepts in human minds, similarity depends on what is said (in terms of computational representation) about these types. This again is connected to the chosen representation language, leading to the fact that most similarity measures cannot be compared. Beside the question of representation, context is another major challenge for similarity assessments. In many cases meaningful notions

[2]A more sophisticated example based on a geographic feature type ontology can be downloaded together with the SIM-DL similarity server and plug-in at http://sim-dl.sourceforge.net.

of similarity cannot be determined without defining in respect to what similarity is measured [6]–[8].

Several measures [9]–[11] were developed to close the gap between ontologies described by various kinds of description logics, and similarity theories that had not been able to handle the expressivity of such languages.

By studying several similarity theories from information science and their application areas, we found generic patterns which jointly form a framework for measuring similarity between concepts (see also [11], [12]). The framework consists of the following five steps. Their concrete realization depends on the semantic similarity theory and the underlying representation language. Consequently, while some of these steps are important for a particular theory they may play a marginal role for another theory.

1) Selection of search concept and target concept
2) Transformation of concepts to canonical form
3) Definition of an alignment matrix to match concept descriptors[3]
4) Application of constructor specific similarity functions
5) Determination of normalized overall similarity

Each of these steps is affected by context in a different way.

B. Context

As the presented paper discusses context types relevant for similarity measurement, at first we will define context as *any kind of additional information which has impact on similarity judgments at execution time*.

In terms of similarity, the role of context has been examined by Rodríguez and Egenhofer [13] for so-called feature based measures and by Janowicz, Keßler and colleagues [8], [11], [12], [14] for similarity measures based on various description logics. Turhan et al. [15] introduced a framework for processing context information based on modeling context as concepts using the Web Ontology Language OWL-DL. While this approach focuses on modeling the application (domain), other approaches (such as [1], [16]) try to come up with a generic (top-level) context ontology. Jurisica [17] proposed a context-based similarity theory for information retrieval using the Similarity Query Language (SimQL).

III. Kinds of Contexts

In the following, six kinds of contexts will be introduced and their impact on semantic similarity measurement will be discussed in detail. The contexts will be presented in the order they usually appear during the measurement process. While some are best represented as sets, other kinds of contexts are rather functions. Finally, the contexts will be applied to a simplified measurement example to demonstrate their impact.

A. User Context

The first kind of context underlying every information retrieval task is the user context ($\mathcal{C}_u$). It describes the user's cognitive capabilities and cultural background ($CCCB$), the current environment (ENV), and the user's motivation (M) for using an information retrieval system [3], [18]. There are strong clues from cognitive science that similarity judgments depend on previous knowledge as well as age [19]. For instance, children tend to a perception driven similarity while adults tend toward so-called theory driven similarity. Recent studies from Mark and colleagues point out that similarity also depends on cultural background and language [20]. The influence of the user's (comparison) environment has been examined by Goldstone et al. [21]. Finally, one clearly needs to distinguish between the user's motivation and the query typed into an information retrieval system. If a user is searching for rivers or similar entities, this does neither answer the question of why nor how the data will be used. While the user's capabilities, cultural background, environment, and motivation influence similarity, their impact is difficult to measure (at least from a computer science point of view).

For this reason a definition of the user context is out of scope for this paper. Nevertheless, these aspects are crucial for understanding the difficulties of measuring similarity arising from context. Consequently, while we define user context as a triple *CCCB*, *ENV*, *M* here, we do not state anything about the interaction between these components or their formal characteristics.

$$\mathcal{C}_u = \langle CCCB, ENV, M \rangle \tag{1}$$

B. Noise and Intended Context

Based on the definition of context as additional information influencing similarity, one has to distinguish between intended and undesired context. We assume that noise, i.e., undesired context, is the part of the user context that is not formally represented within an context-aware similarity measure. The term noise is chosen here, because this kind of context has impact on human similarity judgments while it is not accessible for computational similarity theories. This results in (apparently random) deviations between human and computer similarity judgments.

The problem of noise is especially important in case of human participants tests. For instance, while comparing pictures, human subjects do not only use the depicted entities (*intended stimuli*) for comparison, but also aspects such as the size of the picture or background, e.g., a cloudy sky.

In contrast, intended context is what we are trying to take into account while developing similarity theories and reasoning services (independent of whether we are able to catch all this information). In the following, we assume that intended context ($\mathcal{C}_{int}$) and noise ($\mathcal{C}_n$) form a partition, such that:

$$\mathcal{C}_{int} = \mathcal{C}_u \setminus \mathcal{C}_n \tag{2}$$

C. Application Context

Measuring semantic similarity does not end in itself, but is used to solve a given task. As argued by Goodman [7] and Medin et al. [6] there is no global law stating how similarity measurement works and what it measures. In implementing

[3]The term concept *descriptor* is used here as placeholder for feature, dimension, superconcept etc., which are used to describe the nature of a particular concept.

specific similarity functions, each application defines the semantics of similarity (values) with respect to its application area. We define the application context ($\mathcal{C}_a$) as additional parameters the user can pass to the application to influence the way similarity is measured.

For instance, the asymmetric Matching Distance Similarity Measure (MDSM) [13] allows the user to chose between a commonality or variability weighting to elevate the role of specific concept descriptors. In contrast, the SIM-DL theory [11], [12] distinguishes between average and maximum similarity and additionally allows the user to decide whether the measure should be symmetric or not. It is also possible to define a threshold as minimum similarity of interest. In case of (mobile) recommendation systems, one may also think of K.O. criteria such as a price limit or duration specified by the user.

Consequently, the application context influences steps 3, 4 and 5 of the similarity framework introduced above. Asymmetry is usually achieved by changing the alignment matrix, i.e., by specifying which concept descriptors from the search and the target concept are compared to each other (step 3). Weightings and the distinction between averaged and maximum similarity are defined based on particular similarity functions (step 4). The threshold is used to decide whether a particular target concept will be presented to the user (within the result set) and hence is part of step 5.

Beside such explicitly stated information, parts of the application context can be inferred from the user's behavior or spatio-temporal constraints. Daytime and opening hours are classical examples, but user profiles would allow for additional information. One has to keep in mind that the limiting factor is not how much context information can be collected about the user's behavior and motivation[4], but whether it can be incorporated into the similarity measure (e.g., through weights) and whether it plays a significant role (i.e., has a clear impact the resulting similarity judgements).

The application context is the part of the intended context which is captured by the application. A particular similarity service such as Utopian may take spatio-temporal aspects (and their influence; see section III-E) into account, but fail to support other aspects such as legal restrictions. One can argue that computed similarity judgments correlate the better with human judgments the better $\mathcal{C}_a$ approximates $\mathcal{C}_{int}$ (see also [8]).

$$\mathcal{C}_a \subseteq \mathcal{C}_{int} \tag{3}$$

D. Discourse Context

In a typical information retrieval scenario, the user only defines the search concept, while the compared to (target) concepts depend on the domain of discourse, e.g., the examined ontology. The discourse context defines which concepts are compared to the search concept (step 1 of the framework). Along with similarity measures such as MDSM or SIM-DL, we assume that the user is able to restrict the search to a set of concepts by defining a context concept (C_c). This context concept is either part of the ontology or phrased using an (graphical) interface such as the SIM-DL extension [12] for the Protégé ontology editor (see Figure 1). After reclassification, the discourse context is the set of those target concepts which have C_c as their least common subsumer (lcs); see [22] for more details. Finally, the search concept is compared to each target concept (C_t) out of the set $\mathcal{C}_d$.

$$\mathcal{C}_d = \{C_t | C_t \sqsubseteq C_c\} \tag{4}$$

The discourse context does not only define which concepts are selected, but also influences similarity (step 4). In case of the SIM-DL theory, all descriptors (i.e., superconcepts) defining C_c are not taken into account to compare search and target concept (as they appear in all C_t). This is comparable to the variability weighting proposed by Rodríguez and Egenhofer for MDSM [13]. Up to now, SIM-DL only allows for primitives (and their intersections) as context concept. The usage of arbitrary concepts would require more complex substitution operations on DL concepts as proposed by Teege [23] (see also [24]).

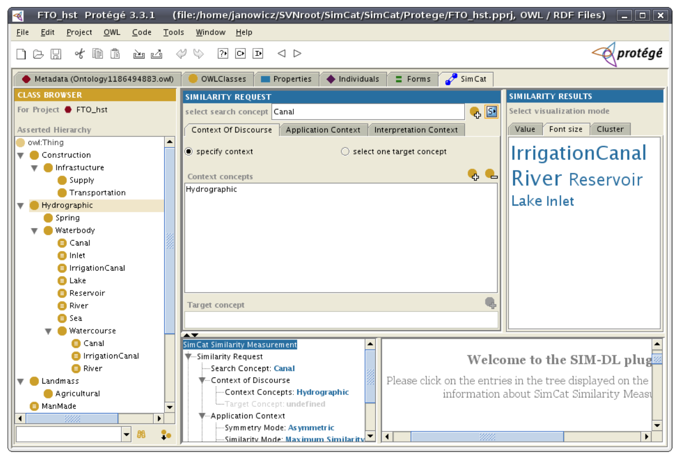

Fig. 1. Selection of search concept and discourse context (via C_c) using an extended version the SIM-DL Protégé plug-in (compare to [12]).

E. Representation Context

While the discourse context defines which concepts are compared, the representation context ($\mathcal{C}_r$) modifies their descriptors in dependence of $\mathcal{C}_a$ (which has influence on step 4). This is comparable to the focus change (*dressing*) introduced by Brézillon [25]. Keßler and colleagues [14] describe the representation context as a set of rules (R). Each rule maps from an activation condition ($cond$) to a set of concept modifiers (C_m) and affected (modified) concepts (C_a).

$$\mathcal{C}_r = \{R_1, ..., R_n\} \tag{5}$$

$$R_i : cond \rightarrow \langle\{\pm C_{m_1}, ..., \pm C_{m_n}\}, \{C_{a_1}, ..., C_{a_n}\}\rangle \tag{6}$$

If the condition for a particular rule from the representation context is true, the rule gets activated. Every affected concept (C_{a_i}) is modified temporarily by either adding or removing the concept descriptors C_{m_i}. These descriptors are concepts themselves; $+$ means that they are added to the definition of all C_a by intersection, $-$ indicates that they are removed

[4]Which also raises all kinds of privacy issues.

from the definitions of all C_a (see [14] for further details). In dependence of the application, a condition may be a checkbox in an user interface, a FOL axiom or information extracted from the user's query (e.g., the user's location).

While it is easy to see that modifying the concepts changes their similarity, a quantification of this change turns out to be difficult (see [14]). To measure the impact of the representation context on similarity is interesting, as it would allow to infer which parts of the context are of major importance and which could be left aside.

F. Interpretation Context

Similarity maps compared concepts to a real number, without pointing out in respect to which descriptors these concepts differ. A similarity value (e.g., 0.67) computed between two concepts hides most of the important information. It does neither answer the question whether there are more or less similar target concepts in the examined ontology. It is not enough to know that possible similarity values range from 0 to 1 as long as their distribution is unclear. Imagine an ontology where the least similar target concept has a value of 0.6 (compared to the source concept), while the comparison to the most similar concept yields 0.9. In such a case, a similarity value of 0.67 is not high at all. Beside these interpretation problems, isolated comparison puts too much stress on the concrete similarity value. It is hard to argue that (and why) the result is plausible without other values as reference [17].

Consequently, measures such as SIM-DL focus on similarity rankings. The search concept is compared to all target concepts $\in \mathcal{C}_d$. The result is an ordered list with descending similarity values. In general, one would not argue that certain similarity values are cognitively plausible, but that the computed order correlates with human ranking judgment. In this paper, we argue that such a rating puts a single similarity value in context - namely into the context established by the order of similarity values. This context is called the interpretation context ($\mathcal{C}_i$) here and influences step 5 of the framework.

$$\mathcal{C}_i : (C_s, C_t, simV) \in \Delta_{sim} \times \mathcal{C}_a \rightarrow \Psi(C_s, C_t) \in \Delta_\Psi \quad (7)$$

The interpretation context maps the triple search concept (C_s), target concept (C_t), similarity value ($simV$) from the set of measured similarities[5] (Δ_{sim}) and the restrictions specified by the application context ($\mathcal{C}_a$) to an interpretation value ($\Psi(C_s, C_t)$) from the domain of interpretations (Δ_Ψ).

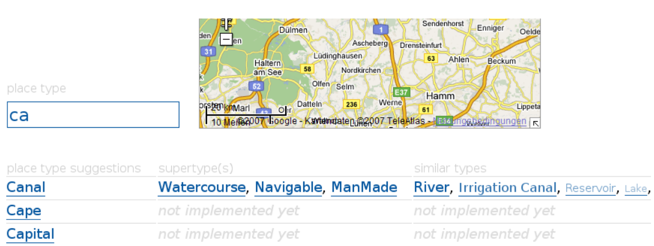

Fig. 2. Partial view on the web gazetteer interface.

[5]between the search concept and those target concepts $\in \mathcal{C}_d$.

In its simplest case, the domain of interpretation is formed by $\Delta_\Psi = \{t, f\}$. Depending on the application area and the remaining pairs of compared concepts from Δ_{sim}, each triple is either mapped to $true$ or $false$. In such a case, the question of whether search and target concepts are similar is answered by yes or no. For the web gazetteer example proposed in [12], similarity values are mapped to font sizes (for visualization) using a logarithmic tag cloud algorithm (see figure 2).

One has to keep in mind, that $\mathcal{C}_i$ does not map an isolated similarity value to another domain, but depends on Δ_{sim}. For instance, the maximum font size is always assigned to the target concept with the highest similarity (to the search concept), independent of a particular value.

IV. Example

The following simplified example is intended to demonstrate the impact of the contexts described before. While SIM-DL and other theories are able to compute similarity between concepts specified using expressive description logics, we restrict this example to concepts formed by intersection (of primitives). Five concepts from the domain of hydrography are examined with respect to their similarity. The concepts are specified as follows:

Concepts:
$Waterbody \sqsubseteq HydrographicFeature$
$Spring \sqsubseteq HydrographicFeature$
$Canal \sqsubseteq Waterbody \sqcap Linear \sqcap Navigable \sqcap ManMade$
$River \sqsubseteq Waterbody \sqcap Linear \sqcap Navigable \sqcap \neg ManMade$
$Lake \sqsubseteq Waterbody \sqcap \neg Linear \sqcap Navigable \sqcap \neg ManMade$

One has to keep in mind that similarity is computed based on representations. There are more differences and commonalities between the introduced concepts which are not specified here.

Discourse Context:
$\mathcal{C}_d = Waterbody$

The discourse context is set to $Waterbody$. On the one hand this restricts the comparisons to $Canal$, $River$ and $Lake$ (because $Spring$ is a $HydrographicFeature$ but not a $Waterbody$); on the other hand the concept $Waterbody$ has no further influence on the computed similarities.

Representation Context:
$\mathcal{C}_r = \{R_f\}$
$R_f = Flooding \rightarrow (+\neg Linear, \{River, Canal\})$

The representation context consists of a rule that adds $\neg Linear$ to $River$ and $Canal$ in case of flooding. Using $-Linear$ would only remove the primitive $Linear$ from the concept definitions, without explicitly stating that flooded rivers and canals are no longer linear. As the representation context overwrites existing definitions, adding $\neg Linear$ does not lead to a contradiction (i.e., an unsatisfiable concept).

Interpretation Context:
$\mathcal{C}_i : \{t, f\}; sim(C_s, C_t) \geq 0.75 \rightarrow t$

In our example the interpretation context maps similarity values to true, i.e., similar with respect to the user's goal, or false (dissimilar). For simplification, the mapping only depends on a threshold (0.75) and every similarity value is considered separately.

TABLE I
SIMILARITY RESULTS WITH AND WITHOUT CONTEXT.

(C_s, C_t)	sim_{dc}	$sim_{\mathcal{C}_d}$	$sim_{\mathcal{C}_r}$	$sim_{\mathcal{C}_d+\mathcal{C}_r}$
(Lake, Canal)	0.5	0.33	0.75_t	0.66
(Lake, River)	0.75_t	0.66	1.0_t	1.0_t
(Canal, River)	0.75_t	0.66	0.75_t	0.66

Table I shows the similarities computed for three pairs of concepts. The first similarity value is calculated without any context information, i.e., decontextualized (sim_{dc}). The second similarity takes the discourse context into account ($sim_{\mathcal{C}_d}$). Similarity decreases, because the ratio between common and distinctive descriptors changes ($Waterbody$ is not used for comparison). The third value is computed taking the flooding scenario into account. As $Canal$ and $River$ are not longer linear, their similarity to $Lake$ is increased. In the second case, the linearity was the only distinctive concept and hence similarity is 1 during flooding (with respect to this representation). The last column is computed using both discourse and representation context. Whether a target concept will be displayed as alternative to the user depends on $\mathcal{C}_i$ and is indicated by a small t (true) in table I.

V. IMPLEMENTATION

Based on the introduced kinds of contexts, we have extended the Protégé plug-in for the SIM-DL similarity server [12]. As depicted in figure 1, in the first step the user selects a search (query) concept and either specifies a context of discourse or a certain target concept. In the following tab, named *Application Context*, the user can chose between a symmetric and a asymmetric version of the SIM-DL measure, as well as define a threshold. Additionally, the user can decide how to compute similarity if concept descriptions contain logical disjunction. Using the *Interpretation Context* tab, three outputs can be selected. The results can be either represented as a descending list of similarity values (between 0 and 1), using font size scaling or as clusters. In the last case, the user can define the number of categories. Short descriptions for each tab support the users by explaining the possible settings.

The integration of the representation context into the SIM-DL plug-in is still under development and requires further work on how to create the necessary rules. The plug-in is developed for ontology engineers familiar with the used vocabulary and description logics. The integration into an end-user centric interface such as the gazetteer web interface is out of scope for this work.

The plug-in and SIM-DL similarity server can be downloaded at http://sim-dl.sourceforge.net.

VI. CONCLUSIONS AND FURTHER WORK

The presented work distinguishes six types of context which influence similarity judgments and the interpretation of single similarity values. As for other domains, the *context gap* [2] between the user context and its computational representation is also relevant for similarity measurement. While this mismatch cannot be solved, other contexts can be used to improve the correlation between human and machine similarity judgments. The context information used to adjust these judgments can be either inferred or explicitly stated by the user. Context is more than the domain of discourse. A better understanding of the different types of context and their influence allows to improve the accuracy of machine based similarity ratings and make them situation-aware. In addition, as not all context information can be modeled, one can still examine which information is most relevant and which could be left aside (e.g., using context impact measures [14]). Doing so would also help to differentiate between noise and intended context.

In case of mobile applications, certain context information might be available only at a given time or at a given location. This leads to the question on how to update similarity judgments on-line, which relates to AI planning. How to proceed in the absence of this information and how to interpolate or infer it?

While first human participants tests show a significant correlation between human and SIM-DL similarity rankings, the same kind of testing is necessary for the proposed contexts. This is especially important in case of the interpretation context. Different applications (such as mobile decision support systems) may require their own visualization and interpretation of the results.

Finally, the relationship between context and similarity is reciprocal. While this paper describes how to improve the accuracy of machine based similarity judgments using context, one could also infer context information out of similarity judgments. An user of a mobile, location based service who rates biking to be more similar to walking than to taking a taxi, is probably not in a hurry (or has experienced rush hours in cities such as New York; which leads back to the question of user context and noise) [26].

VII. ACKNOWLEDGMENT

Comments from the members of the SimCat project as well as from the participants of the Semantic Similarity and Geospatial Applications Workshop at COSIT 2007 provided suggestions to improve the content and clarity of the paper. The author is also grateful for the comments provided by the four anonymous reviewers. This work is funded by the "Semantic Similarity Measurement for Role-Governed Geospatial Categories (SimCat)" project granted by the German Research Foundation (DFG Ra1062/2-1).

REFERENCES

[1] X. H. Wang, D. Q. Zhang, T. Gu, and H. K. Pung, "Ontology based context modeling and reasoning using owl." in *Proceedings of Workshop on Context Modeling and Reasoning (CoMoRea '04), in conjunction with the 2nd IEEE International Conference on Pervasive Computing and Communications*. IEEE Computer Society, 2004, pp. 18–22.

[2] L. Barkhuus, “The context gap: An essential challenge to context-aware computing,” PhD Thesis, IT University of Copenhagen, December 1st 2004.

[3] K. Henricksen and J. Indulska, “Modelling and using imperfect context information,” in *Proc. of the Second IEEE International Conference on Pervasive Computing and Communications. Workshop on Context Modelling and Reasoning (CoMoRea’04)*. Orlando, Florida, USA: IEEE Computer Society, 2004, pp. 33–37.

[4] M. Bazire and P. Brézillon, “Understanding context before using it.” in *Modeling and Using Context, 5th International and Interdisciplinary Conference (CONTEXT 2005)*, ser. Lecture Notes in Computer Science, A. K. Dey, B. N. Kokinov, D. B. Leake, and R. M. Turner, Eds., vol. 3554. Springer, 2005, pp. 29–40.

[5] R. Goldstone and J. Son, “Similarity,” in *Cambridge Handbook of Thinking and Reasoning*, K. Holyoak and R. Morrison, Eds. Cambridge: Cambridge University Press, 2005.

[6] D. Medin, R. Goldstone, and D. Gentner, “Respects for similarity,” *Psychological Review*, vol. 100, no. 2, pp. 254–278, 1993.

[7] N. Goodman, “Seven strictures on similarity,” in *Problems and projects*. New York: Bobbs-Merrill, 1972, pp. 437–447.

[8] C. Keßler, “Similarity measurement in context,” in *6th International and Interdisciplinary Conference, CONTEXT 2007*, B. Kokinov, D. C. Richardson, T. R. Roth-Berghofer, and L. Vieu, Eds. Roskilde, Denmark: Springer-Verlag Berlin Heidelberg, 2007, vol. 4635, pp. 277–290.

[9] C. d’Amato, N. Fanizzi, and F. Esposito, “A dissimilarity measure for $\mathcal{ALC}$ concept descriptions.” in *Proceedings of the 2006 ACM Symposium on Applied Computing (SAC)*, Dijon, France, 2006, pp. 1695–1699.

[10] A. Borgida, T. Walsh, and H. Hirsh, “Towards measuring similarity in description logics.” in *Proceedings of the 2005 International Workshop on Description Logics (DL2005)*, ser. CEUR Workshop Proceedings. Edinburgh, Scotland, UK: CEUR, 2005, vol. 147.

[11] K. Janowicz, “Sim-dl: Towards a semantic similarity measurement theory for the description logic $\mathcal{ALCNR}$ in geographic information retrieval,” in *SeBGIS 2006, OTM Workshops 2006*, ser. Lecture Notes in Computer Science, R. Meersman, Z. Tari, P. Herrero, and e. al., Eds. Berlin: Springer, 2006, vol. 4278, pp. 1681 – 1692.

[12] K. Janowicz, C. Keßler, M. Schwarz, M. Wilkes, I. Panov, M. Espeter, and B. Baeumer, “Algorithm, implementation and application of the sim-dl similarity server,” in *Second International Conference on GeoSpatial Semantics (GeoS 2007)*, ser. Lecture Notes in Computer Science. Springer, 2007.

[13] A. Rodriguez and M. Egenhofer, “Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure,” *International Journal of Geographical Information Science*, vol. 18, no. 3, pp. 229–256, 2004.

[14] C. Keßler, M. Raubal, and K. Janowicz, “The effect of context on semantic similarity measurement,” in *Proceedings of the 3rd International IFIP Workshop On Semantic Web & Web Semantics (SWWS ’07)*, ser. Lecture Notes in Computer Science. Vilamoura, Portuga: Springer Verlag, November 29-30 2007; forthcoming.

[15] A.-Y. Turhan, T. Springer, and M. Berger, “Pushing doors for modeling contexts with owl dl a case study.” in *Proc. of the Fourth IEEE International Conference on Pervasive Computing and Communications. Workshop on Context Modelling and Reasoning (CoMoRea’06)*. IEEE Computer Society, 2006, pp. 13–17.

[16] H. Chen, T. Finin, and A. Joshi, “An ontology for context-aware pervasive computing environments,” *Knowledge Engineering Review*, vol. 18, no. 3, pp. 197–207, 2003.

[17] I. Jurisica, “Technical report dkbs-tr-94-5: Context-based similarity applied to retrieval of relevant cases,” University of Toronto, Department of Computer Science, Toronto, Tech. Rep., 1994.

[18] A. Schmidt, “Ontology-based user context management: The challenges of dynamics and imperfection,” in *On the Move to Meaningful Internet Systems 2006: CoopIS, DOA, GADA, and ODBASE. Part I.*, ser. Lecture Notes in Computer Science, R. Meersman and Z. Tahiri, Eds., vol. 4275. Springer, 2006, pp. 995–1011.

[19] L. Kotovsky and D. Gentner, “Comparison and categorization in the development of relational similarity,” *Child Development*, vol. 67, pp. 2797–2822, 1996.

[20] D. Mark, A. Turk, and D. Stea, “Does the semantic similarity of geospatial entity types vary across languages and cultures?” in *Workshop on Semantic Similarity Measurement and Geospatial Applications (at COSIT 2007)*, Melbourne, Australia, 19th September 2007.

[21] R. L. Goldstone, D. L. Medin, and J. Halberstadt, “Similarity in context,” *Memory & Cognition*, vol. 25, pp. 237–255, 1997.

[22] R. Küsters, *Non-Standard Inferences in Description Logics*, ser. Lecture Notes in Artificial Intelligence. Springer-Verlag, 2001, vol. 2100.

[23] G. Teege, “Making the difference: A subtraction operation for description logics,” in *4th International Conference on Principles of Knowledge Representation and Reasoning (KR’94)*, J. Doyle, E. Sandewall, and P. Torasso, Eds. Morgan Kaufmann, 1994, pp. 540–550.

[24] S. Brandt, R. Küsters, and A.-Y. Turhan, “Approximation and difference in description logics,” in *Proceedings of the Eighth International Conference on Principles of Knowledge Representation and Reasoning (KR2002)*, D. Fensel, F. Giunchiglia, D. McGuiness, and M.-A. Williams, Eds. San Francisco, CA: Morgan Kaufman, 2002, pp. 203–214.

[25] J. Brézillon and P. Brézillon, “Context modeling: Context as a dressing of a focus,” in *Modeling and Using Context, 6th International and Interdisciplinary Conference (CONTEXT 2007)*, ser. Lecture Notes in Computer Science, B. N. Kokinov, D. C. Richardson, T. Roth-Berghofer, and L. Vieu, Eds., vol. 4635. Springer, 2007, pp. 136–149.

[26] A. Tversky, “Features of similarity,” *Psychological Review*, vol. 84, no. 4, pp. 327–352, 1977.