

Semantic Similarity Measurement and Geographic Applications

GEOG 288MR, Spring 08, 2-4 credits

Instructor: Dr. Martin Raubal, raubal@geog.ucsb.edu

Office (hours): Wednesday 2-4pm, EH 5713

Phone: 893-4839

Class meets: Thursday 1-3pm, EH 5824

Course description:

Semantic similarity measurement is central for enabling semantic interoperability between geographic information systems and web services. It is also the basis for information retrieval and integration. In this seminar we cover several theoretical and practical aspects of semantic similarity measurement regarding geographic concepts. Topics include psychological models to explain human similarity judgment, representation of geographic semantics, theories for measuring semantic similarity, context-awareness, and the use of similarity measurement in geographic applications.

The seminar will mainly consist of readings, student presentations, and discussions, but students will also get some hands-on experience using a semantic similarity server. Each student will be responsible for one session including presentation of the weekly material (and possibly additional material) and lead of discussion. Students are responsible for preparing questions to be discussed in the seminar and writing short commentaries on the weekly readings.

Required reading: Literature for the course is supplied online by the instructor (through electronic reserve). Weekly readings are listed below. Please come to class prepared to discuss the readings, which are listed for that day.

Grades: Grades will be determined by your presentation and lead of discussion (35%), the written assignments (35%), and your general class participation (30%). Class participation is mandatory.

Written assignments: The written assignments consist of brief critical commentaries (not longer than 1 page each) and your questions connected to the readings. You should have prepared at least 3 questions (possibly related to your own field of study) that can be discussed in the weekly sessions. You must hand in a printed version to the instructor at the beginning of each session. Late assignments will not be accepted and count as failed.

Weekly schedule and readings

Week 1, 3 April 2008: Introduction – Why do we need semantic similarity measurement?

Week 2, 10 April 2008: Similarity approaches

- R. Goldstone and J. Son (2005) Similarity. in: K. Holyoak and R. Morrison (Eds.), *Cambridge Handbook of Thinking and Reasoning*. Cambridge University Press, Cambridge.
- Schwering, A., *Approaches to Semantic Similarity Measurement for Geo-Spatial Data: A Survey*. Transactions in GIS, 2008. **12**(1): p. 5-29.

- A. Rissland (2006) AI and Similarity. *IEEE Intelligent Systems* (May-June 06): 39-49.

Week 3, 17 April 2008: Representation (AAG Meeting, self-study; please submit your assignment by email!)

- S. Palmer (1978) Fundamental aspects of cognitive representation. in: E. Rosch and B. Lloyd (Ed.), *Cognition and categorization*. pp. 259-303, Lawrence Erlbaum, Hillsdale, NJ.

Week 4, 24 April 2008: Semantic interoperability

- W. Kuhn (2005) Geospatial Semantics: Why, of What, and How? *Journal on Data Semantics (Special Issue on Semantic-based Geographical Information Systems, LNCS 3534)* Spring 2005: 1-24.
- M. Lutz, C. Riedemann, and F. Probst (2003) A Classification Framework for Approaches to Achieving Semantic Interoperability between GI Web Services. in: W. Kuhn, M. Worboys, and S. Timpf (Ed.), *Spatial Information Theory - Foundations of Geographic Information Science, International Conference COSIT 2003. Lecture Notes in Computer Science 2825*, pp. 186-203, Springer, Berlin.

Week 5, 1 May 2008: Feature-based models

- A. Tversky (1977) Features of Similarity. *Psychological Review* 84(4): 327-352.
- A. Rodríguez and M. Egenhofer (2003) Determining Semantic Similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering* 15(2): 442-456.
- C.-C. Feng and D. Flewelling (2004) Assessment of semantic similarity between land use/land cover classification systems. *Computers, Environment and Urban Systems* 28 (2004): 229-246.

Week 6, 8 May 2008: Geometric models

- P. Gärdenfors (2004) How to Make the Semantic Web More Semantic. in: A. Varzi and L. Lieu (Ed.), *Formal Ontology In Information Systems*. pp. 17-34, IOS Press, Amsterdam.
- M. Raubal (2004) Formalizing Conceptual Spaces. in: A. Varzi and L. Vieu (Ed.), *Formal Ontology in Information Systems, Proceedings of the Third International Conference (FOIS 2004)*. *Frontiers in Artificial Intelligence and Applications* 114, pp. 153-164, IOS Press, Amsterdam, NL.
- A. Schwering and M. Raubal (2005) Spatial Relations for Semantic Similarity Measurement. in: J. Akoka, S. Liddle, I.-Y. Song, M. Bertolotto, I. Comyn-Wattiau, W.-J. vanden Heuvel, M. Kolp, J. Trujillo, C. Kop, and H. Mayr (Ed.), *Perspectives in Conceptual Modeling: ER 2005 Workshops CAOIS, BP-UML, CoMoGIS, eCOMO, and QoIS, Klagenfurt, Austria, October 24-28, 2005*. *Lecture Notes in Computer Science* 3770, pp. 259-269, Springer, Berlin.

Week 7, 15 May 2008: Context

- A. Dey and G. Abowd (2000) Towards a Better Understanding of Context and Context-Awareness. *CHI 2000 Workshop on the What, Who, Where, When, Why and How of Context-Awareness*. *The Hague, The Netherlands*.

- C. Keßler (2007) Similarity Measurement in Context. in: B. Kokinov, D. Richardson, T. Roth-Berghofer, and L. Vieu (Ed.), *Modeling and Using Context. 6th International and Interdisciplinary Conference, CONTEXT 2007, Proceedings. Roskilde, Denmark, August 20-24, 2007*. 4635, pp. 277-290, Springer, Berlin.
- A. Rodríguez and M. Egenhofer (2004) Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure. *International Journal of Geographical Information Science* 18(3): 229-256.

Week 8, 22 May 2008: Semantic similarity server - Protege & Sim-DL

- <http://www.co-ode.org/resources/tutorials/ProtegeOWLTutorial.pdf> (Pizza tutorial)
- K. Janowicz, C. Keßler, I. Panov, M. Wilkes, M. Espeter, and M. Schwarz (2008) A Study on the Cognitive Plausibility of SIM-DL Similarity Rankings for Geographic Feature Types. *11th AGILE International Conference on Geographic Information Science (AGILE 2008). Girona, Spain, 5-8 May 2008*.

Week 9, 29 May 2008: Semantic relatedness and Wikipedia

- A. Budanitsky and G. Hirst (2006) Evaluating WordNet-based Measures of Lexical Semantic Relatedness. *Computational Linguistics* 32(1): 13-47. (Sections 1-4)
- B. Hecht and M. Raubal (in press) GeoSR: Geographically Explore Semantic Relations in World Knowledge. in: L. Bernard, H. Pundt, and C. Friis-Christensen (Ed.), *Proceedings of the 11th AGILE International Conference on Geographic Information Science in 2008, Girona, Spain. Springer Lecture Notes Series on Geo-Information* Springer, Berlin.

Week 10, 5 June 2008: Future issues and wrap-up

Further references:

- Y. Bishr (1998) Overcoming the Semantic and Other Barriers to GIS Interoperability. *IJGIS* 12(4): 299-314.
- F. Harvey, W. Kuhn, H. Pundt, Y. Bishr, and C. Riedemann (1999) Semantic Interoperability: A Central Issue for Sharing Geographic Information. *Annals of Regional Science* 33 (2) (Geo-spatial data sharing and standardization): 213-232.

Links:

- (Semantic) similarity blog: <http://www.similarity-blog.de/>