INT. J. GEOGRAPHICAL INFORMATION SCIENCE,
VOL. 17, NO. 5, JULY–AUGUST 2003, 405–409

Guest Editorial

Semantic reference systems

WERNER KUHN
Institute for Geoinformatics, University of Münster, Robert-Koch-Str 26–28, 48149 Münster, Germany
e-mail: kuhn@ifgi.uni-muenster.de

(*Received 12 November 2002; accepted 5 February 2003*)

Four centuries after René Descartes watched a fly walk across his ceiling and wondered how to capture its position (Gribbin 2002), we use Cartesian coordinates routinely to describe locations. We identify the positions of entities in the real world, transform their GIS representations from one coordinate system to another, and integrate spatially referenced data across multiple coordinate systems. A theory of *spatial reference systems* standardises the notions of geodetic datum, map projections, and coordinate transformations (ISO 2002). Similarly, our temporal data refer unambiguously to temporal reference systems, such as calendars, and can be transformed from one to another.

Geographical information systems contain spatial, temporal, and thematic data. The first two kinds are firmly tied to reference system theories and tools. We now need to produce the third component—*semantic* reference systems. Descartes might wonder today how to establish common frames of reference for, say, a geneticist and an entomologist to talk about that fly. They would need methods to explain the meaning of their specialized vocabularies to each other, to detect synonyms and homonyms, and to translate expressions. Accordingly, a theory of semantic reference systems will enable producers and users of geographical information to explain the meaning of thematic data, to translate this meaning from one information community to another (OGC 1998), and to integrate data across differing semantics.

Reaching this goal may not take centuries again, but a decade or two could pass until we see the kind of on-the-fly (so to speak) integration that is now possible across different spatial reference systems in GIS (Lutz 2003). There is no need for universal geographical information semantics, as long as we can provide the means for any pair of information communities to define their concepts and translate between them. All the same, work on semantic primitives and universals (Wierzbicka 1996) as well as on top-level ontologies (Sowa 2000) strongly suggests that a common core of concepts exists and can be defined.

Formalizing the semantics of geographical information communities is, in any case, much simpler than defining natural language terms. The reason is that every information community agrees, per definition, on a shared set of concepts, expressed in conceptual models, feature-attribute catalogues, legal documents, work practices,

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online
http://www.tandf.co.uk/journals
DOI: 10.1080/1365881031000114116

regulations, and the like. These agreements are subject to inspection, analysis, and revision (Kuhn 2001). Given this pragmatic and encouraging foundation of information system semantics, where are we today on the path to semantic reference systems for geographical information? What is required to reach the goal?

At first sight, semantic reference systems may sound like yet another fancy name for what the GIS community has just learned to call *ontologies*. Indeed, ontologies constitute a core component of them. They describe the conceptualisations of the world to which the data in information systems refer (Gruber 1993). In other words, they provide a frame of reference for the vocabulary used in a database. For example, an ontology might specify what the term 'forest' means in one or more vegetation databases. Current ontology work focuses on making this notion practically useful through languages for designing and exposing ontologies (Welty and Smith 2001).

The spatial reference system analogy suggests something more powerful than today's ontology languages can offer. Producers and users of geographical information need tools for *transformations* among semantic spaces and *projections* to subspaces. A transformation may occur within or between information communities and involves a change to the reference system (for example, adding a new axiom to an ontology). A projection occurs typically within a community and reduces the complexity of a semantic space (for example, by generalizing two entity classes to a super-class).

The notions of semantic transformation and projection are suggested here in *analogy* to their geometric counterparts. They do not assume any metric for semantic spaces. All that is required are axiomatised concepts allowing for mappings between them, i.e., a formal version of the mental spaces in Fauconnier (1994). Nevertheless, a geometric structure for semantic spaces, such as the one proposed by Gärdenfors (2000), would have obvious advantages for computational support and for dealing with the core semantic issue of similarity (Rodríguez and Egenhofer, 2003).

However, tools for transforming and projecting between semantic spaces need stronger formal foundations than most current ontology languages provide. They require sound typing, parameterised polymorphism, multiple inheritance of behaviour, and higher order reasoning capabilities (Frank and Kuhn 1999, Kuhn 2002). Furthermore, in order to relate the meaning of terms to GIS applications, they need to place entities and relationships in the context of human activities (Kuhn 2001). An example of a formalized semantic reference system, applying this approach to the domain of vehicle navigation, can be found at http://musil.uni-muenster.de. It contains a complete axiomatisation and computational solution for a simple transformation (among two interpretations of a navigation data model) and projection (simplifying a data model), and is intended as a proof of concept for the ideas presented here.

The mapping of terms from one context to another and the construction of common ontologies require powerful mathematical instruments of the kind found in category theory (Bench-Capon and Malcolm, 1999). A semantic reference system consists of ontologies that specify concepts as well as *mappings* between them, embedded in a formalism that supports the computation of these mappings. These requirements go beyond the formal apparatus underlying coordinate projections and transformations. Most of today's ontology formalisms, however, are even weaker than that. How would you describe in your favourite ontology language, which semantic aspects of a house and of a boat are projected into boathouses and houseboats, respectively, and how they get combined (Kuhn 2002)?

The formal requirements for semantic reference systems are more than a theoretical goal. They represent a practical necessity to achieve *semantic interoperability*, i.e. the capacity of information systems or services to work together without the need for human intervention (Harvey *et al.* 1999, OGC 1998). Focusing on the case of (web) services, I see three core problems that need to be solved on the way to semantic interoperability. Each of them identifies a special kind of *matchmaking* (Sycara *et al.* 1999) between information providers and requesters. And each of them calls for semantic reference systems to support that matchmaking:

- Service providers need to be able to determine whether a *data source* offers useful semantics for a planned service (e.g. can a runoff model use the road widths provided by a cadastral database? can the number of building levels in that database be used to calculate building heights for noise propagation?);
- Client services need to be able to determine whether a given *service* offers useful semantics as input to their processing (e.g. is an elevation model with a certain resolution and linear interpolation sufficient for a particular visibility calculation?);
- *Human* users need to be able to determine whether a service provides useful semantics to answer a question (e.g. should the question 'which parcels touch this road?' be posed to the database or to the GIS?).

In all three problem statements, the term 'useful' indicates that the answers will typically not be a simple yes or no, and that a process of enabling actual information use will follow. The interesting and challenging cases are those where an information source contains useful content that needs to be supplemented with additional information or transformed to other contexts. These cases need the kind of projection and transformation support that the idea of semantic reference systems suggests.

The three semantic interoperability problems raise two practical questions, which also show the way forward to address them:

1. What needs to be stated about an information source, so that a requester (human or machine) can assess and exploit its semantic value?
2. How does this assessment and exploitation work?

The answer to the first question clearly depends on the answer to the second: how can we reasonably decide on the contents, let alone the representations, before specifying their use? While most current work on information system ontologies starts with the first, focusing on metadata and mark-up languages, the second question raises the issues posed by the reference system analogy: What conceptualisations occur in an application? Where and how are they specified? How can they be projected to other conceptualisations? How can they be merged? How can a common ontology be constructed for them? How can expressions be translated from one application context to another?

Finally, in case you wondered, *spatial* reference systems can be seen as a special kind of semantic reference systems: they explain the meaning of coordinates and how it shifts between contexts. This restriction of the difficult general semantics problem has allowed for highly successful solutions to the spatial case. The best strategy for progress towards general semantic reference systems is, thus, to solve the next easiest special cases. This could be gazetteers, which translate geographical names to coordinates. They map the semantic problem of geographic identifiers to the geometric problem already solved by spatial reference systems. Or it could be cases where semantics rests primarily on geometric and topological properties, such as those of navigation services. Ongoing work on these and other applications confirms the utility and fruitfulness of the idea of semantic reference systems. The semantic interoperability discussion in science (Meersman and Tari 2002) and industry (OGC 1998) documents the need for them.

Acknowledgments

Our work on geographic information semantics is supported by the European Commission, in the ACE-GIS (IST-2002-37724) and BRIDGE-IT (IST-2001-34386) projects, by the German ministry of education and science (BMBF) in the Geotechnologies Research Program. The ideas presented have been shaped by discussions with members of the semantic interoperability research group at Münster (http://musil.uni-muenster.de) and of the Meaning and Computation Laboratory at the University of California at San Diego (http://www.ese.uesd.edu/users/goguen/). Comments from anonymous reviewers helped to improve the presentation.

References

Bench-Capon, T. and Malcolm, G., 1999, Formalising Ontologies and their Relations, DEXA '99, edited by G. Soda and A. M. Tjoa (Berlin: Springer-Verlag), Lecture Notes in Computer Science 1677, pp. 250–259.

Fauconnier, G., 1994, *Mental Spaces* (Cambridge: Cambridge University Press).

Frank, A. U., and Kuhn, W., 1999, A specification language for interoperable GIS. In: *Interoperating Geographic Information Systems*, edited by M. F. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman (Dordrecht: Kluwer), pp. 123–132.

Gärdenfors, P., 2000, *Conceptual Spaces* (London: MIT Press).

Goguen, J., 1991, Types as theories. In: *Topology and Category Theory in Computer Science*, edited by G. M. Reed, A. W. Roscoe, and R. F. Wachter (Oxford: Oxford University Press), pp. 357–390.

Gribbon, J., 2002, *Science: A History 1543–2001* (London: Penguin/Allen Lane).

Gruber, T. R., 1993, Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human and Computer Studies*, **43**, 907–928.

Harvey, F., Kuhn, W., Pundt, H., Bishr, Y., and Riedemann, C., 1999, Semantic interoperability: a central issue for sharing geographic information. *Annals of Regional Science*, **33**, Geo-spatial data sharing and standardization, 213–232.

ISO, 2002, Geographic information—Spatial referencing by coordinates. ISO TC211, Final Draft International Standard.

Kuhn, W., 2001, Ontologies in support of activities in geographical space. *International Journal of Geographical Information Science*, **15**, 613–631.

Kuhn, W., 2002, Modelling the semantics of geographic categories through conceptual integration. In: *Geographic Information Science—Second International Conference, GIScience 2002 Boulder, Colorado*, eidted by M. J. Egenhofer and D. M. Marks (Berlin: Springer-Verlag), Lecture Notes in Computer Science 2478, pp. 108–118.

Lutz, D., 2003, GeoMedia and FME: The 'Universal Information Integrator' Meets the 'Universal Translator'. White Paper, Safe Software Inc.: http://www.safe.com/solutions/whitepapers/geomedia_fme.htm.

Meersman, R., and Tari, Z., editors, 2002, *On the Move to Meaningful Internet Systems 2002-DOA/CoopIS/ODBASE 2002* (Berlin: Springer-Verlag), Lecture Notes in Computer Science 2519.

OGC, 1998, *The OpenGIS® Guide* (Wayland, MA Cheel: The Open GIS Consortium).

Rodríguez, A., and Egenhofer, M., 2003, Determining semantic similarity among entity classes from different ontologies. *IEEE Transactions on Knowledge and Data Engineering*, **15(2)**, 442–456.

Sycara, K., Klusch, M., Widoff, S., and Lu, J., 1999, Dynamic service matchmaking among agents in open information environments. *SIGMOD Record*, **28**, 47–53.

SOWA, J. F., 2000, *Knowledge Representation—Logical, Philosophical and Computational Foundations* (Pacific Grove, CA: Brooks/Cole).

WELTY, C., and SMITH, B., editors, 2001, Formal ontology in information systems. In *Proceedings of the International Conference on Formal Ontology in Information Systems* (London: ACM Press).

WIERZBICKA, S., 1996, *Semantics—Primes and Universals* (Oxford: Oxford University Press).