SEMANTIC REFERENCE SYSTEMS ACCOUNTING FOR UNCERTAINTY – A REQUIREMENTS ANALYSIS

Sven Schade
Institute for Geoinformatics, University of Münster, Germany – schades@uni-muenster.de

KEYWORDS: Semantic Reference Systems, Ontologies, Requirements, Uncertainty, Ambiguity, Measurement Error, Vagueness

ABSTRACT:

In analogy to spatial and temporal reference systems, semantic reference systems have been proposed to support interoperation between information communities. Such reference systems provide the structure for specifying intended interpretations of communities' vocabularies and the functionality of translating expressions used by one community to those used by another. First attempts towards designing semantic reference systems have been made and few prototypical implementations examined the notion for information communities in the geospatial domain. In addition to ambiguous, general and open-textured expressions that are applied to define data models, many geospatial concepts are vague in nature and geospatial information relies on error-prone measurements. Not all these aspects have yet been considered in reference systems. The work presented in this paper analyses the requirements, which semantic reference systems need to meet, in order to account for the varying facets of uncertainty. The facets are categorised and taxonomically structured. Design requirements are derived for each category. Possibilities of satisfying these requirements are outlined. The analysis defines a basis for developing semantic reference systems, which account for uncertainties underlying geospatial information.

1. INTRODUCTION

Attribute reference systems and, more generally, semantic reference systems have been proposed as a mean for reaching formal models of semantics (Chrisman, 2002; Kuhn, 2003). Such systems should be used to overcome semantic heterogeneities between information communities. An information community is a group of people, which agrees on a shared set of concepts (OGC, 2003). For the domain of geospatial information, these concepts are used to define feature names, attribute values, relations and operators of geospatial data models (Kuhn and Raubal, 2003). Concepts are expressed by either single symbols or by a combination of symbols. The expressions articulating the concrete concepts used by a distinct community are defined as their specialised vocabulary. Semantic reference systems contribute to semantic interoperability by providing methods to explain the meaning of vocabularies and to translate expressions from the context of one community to that of another, (Kuhn, 2003; Kuhn and Raubal, 2003; Kuhn, 2005).

Since geospatial information bases on measurements (Buyong et al., 1991; Goodchild, 2004) and many geospatial phenomena are vague in nature (Keefe and Smith, 1997; Bennett and Cristani, 2003), requirements that uncertainty modelling imposes on semantic reference systems for geospatial information, are investigated. The current efforts on designing semantic reference systems concentrate on dissolving ambiguities (Kuhn, 2005; Probst and Espeter, 2006). This work contributes to the design by discussing requirements for semantic reference systems to comprise a more general notion of uncertainty. Semantic referencing, i.e. referencing user vocabularies to semantic reference frames, entails *ambiguity*, *generality* and *open-texture*. Ontologies, which constitute these reference frames have to account for *data acquisition errors* and to provide descriptions of *vague concepts*. Both require grounding in a semantic datum. Semantic projection and semantic transformation, i.e. methods translating expressions between reference frames, have to include *propagation of uncertain information*. In addition, *uncertainty caused by transformations* between reference frames needs to be captured.

Based on these requirements, ideas to include uncertainty in semantic reference systems are presented. Measurements of physical unary moments* (Guizzardi et al., 2002), like weight, temperature, and length are considered as the initial step in establishing semantic reference systems for geospatial information (Probst, 2006). Possibilities of extending and combining existing approaches towards implementing semantic reference systems, which account for uncertainty related to such moments, are highlighted. The moment temperature is used as a running example.

The notion of a semantic reference system is reviewed in more detail in section 2. In section 3, relevant aspects of uncertainty are identified and categorised. These aspects are discussed in relation to the fundamentals provided by semantic reference systems in section 4. Section 5 outlines possibilities to account for the identified requirements. Conclusions and a discussion of future work follow.

2. SEMANTIC REFERENCE SYSTEMS

This section provides an overview of the conceptual components of semantic reference systems and selectively points to related wok. The structural elements (semantic reference frame and semantic datum) and the process of semantic referencing are explained. Based on these components, intended interpretations of vocabularies can be specified. Semantic projection and semantic transformation are introduced as means to support the translation of expressions.

* In order to avoid confusion with the term data quality, the notion *moment* as defined in the Basic Foundational Ontology (BFO) (Guizzardi et al., 2002) is used instead of the notion *quality* (Masolo et al., 2003). Both are considered to be interchangeable.

2.1 Specifying Intended Interpretations

Structural elements provide the backbone of semantic reference systems. They constitute the conceptual structure that is required to reduce doubt in possible interpretations of expressions. Guarino's definition of an ontology, as "an engineering artefact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words" (Guarino, 1998) is applied. Upper-level, also called top-level ontologies are used as *semantic reference frames*. The vocabulary used in an upper-level ontology serves as a formally defined, application-independent conceptual structure. It is characterised as upper-level, because it is more general than the vocabulary of a specific information community.

Considering the measurement of air temperature, temperature may be seen as a dimension characterizing an intrinsic unary moment of the concept air. Pressure and humidity are further examples. This dimensional view for structuring reference frames corresponds to Gärdenfor's notion of conceptual spaces (Gärdenfors, 2000).

As each ontology, the reference frame constitutes of building blocks. The basic building blocks are called primitives. Ensuring the correct interpretation of the conceptual structure provided by the reference frame requires a clear definition of the meaning of these primitives, i.e. grounding. The meaning of primitives applied in the semantic reference frame, is grounded in the *semantic datum*. The semantic datum is part of the semantic reference system, but not of the ontologies constituting the frame.

Interpretation rules for vocabularies used by information communities are defined by establishing a reference between the vocabulary and the semantic reference frame. The process of establishing these links is called *semantic referencing*.

2.2 Translating Expressions

Once a vocabulary is semantically referenced, *semantic projection* may be used to simplify the underlying formal model. Semantic projection is a dimensional simplification or reduction of the reference frame. If a second vocabulary is semantically referenced, a *semantic transformation* may be used to switch from one vocabulary to the other. In the simplest case a semantic transformation is used to modify the units of a dimension. For example, air temperature described in Fahrenheit can be translated to a description using Celsius as the unit of measure.

So far, the translation rules required to implement semantic transformations have to be built manually. In an overall mathematical structure based on category theory (Barr and Wells, 1990; Asperti and Longo, 1991), these rules can be precisely defined (Raubal and Kuhn, 2003). A future scenario envisions the derivation of translation rules as an empirical procedure, akin to the inference of transformation parameters in similarity transformation between spatial reference systems (Iliffe, 2000). Parameters may, for example, be discovered using text mining techniques (Manning and Schutze, 2001) for ontology alignment.

2.3 Related Work

An attempt of capturing the meaning of unary moments, their function as dimensions and their measurements is taken by Probst and Espeter (Probst, 2006; Probst and Espeter, 2006). Primitives are *grounded* in formal ontology (Probst, 2006). Following Probst, *semantic referencing* can be established by annotating expressions with elements of the ontology. This method is also known as registration mapping (Bowers and Ludäscher, 2004).

Opposed to Probst and Espeter who focus on dimensions as characteristics of concepts Frank uses dimensions for defining so called distinctions (Frank, 2006). Instead of facilitating dimensions directly, he intends using distinctions as language-independent conceptual units for building *reference frames*. At the current stage, distinctions can be used to maintain and integrate taxonomies. An extension towards ontologies is planned. *Grounding* of distinctions has not been discussed yet. If using a lattice of distinctions as a reference frame (Frank, 2006), *semantic referencing* may be established by labelling notes of the lattice with expressions of the vocabulary.

3. FACETS OF UNCERTAINTY

Before uncertainty in geospatial information can be accounted for, possible facets need to be defined. For this purpose uncertainty in the use of concepts and of expressions are separated. The use of expressions within vocabularies is exploited. Subsequently, requirements for the specification of intended interpretations of these expressions and for the translation of expression between communities can be derived.

The notion of uncertainty is widely used, but in distinct ways. According to Fisher and others, uncertainty can be understood as a general concept that can be separated into varying kinds (Fisher et al., 2006). In this section uncertainty is defined and categorized with the goal to derive specific requirements for the design and implementation of semantic reference systems. The definitions revise the categorization proposed by Fisher (figure 1). The gray boxes in the figure denote categories of uncertainty, while the white boxes point to existing approaches that account for the connected category. These approaches are focussed on in section 5.

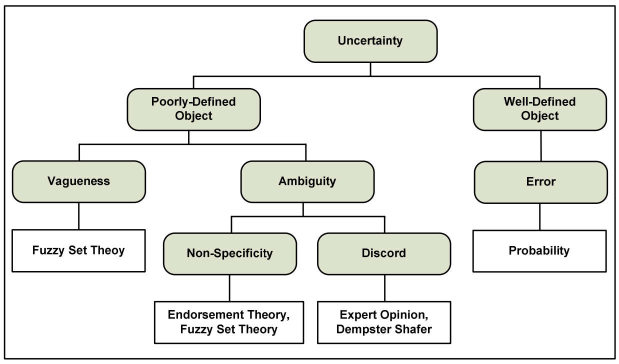

Figure 1. Fisher's categories of uncertainty (adapted from (Fisher et al., 2006; Fisher, 1999; Klir and Yuan, 1995)).

At the highest level of abstraction, Fisher distinguishes poorly-defined and well-defined objects. Opposed to this, the work presented in this paper elaborates the uncertainty of expressions that are used to define object models. In the case of well-defined expressions, the distinct meaning (i.e. the intended interpretation) of the expressions, is the only one possible. This criterion is clarified by changing the label for the two upper-level categories into "poor-defined interpretation" and "well-defined interpretation". In the following, the lower level categories are defined. Further changes to Fisher's proposal are considered if appropriate.

3.1 Poorly-Defined Interpretation

Given an expression, it may be unclear to which concept it denotes to. Depending on the nature of concept or characteristics of the expression used, varying facets can be separated.

3.1.1 Ambiguity: Ambiguity arises if two expressions x and y have the same outward appearance, but one denotes something, the other does not (Scheffler, 1979). It is not a characteristic of concepts, but of interpreting symbols and symbol combinations. Accordingly, expressions used in information communities might be ambiguous. For example, the attribute "temperature" considered by itself might be used to express the measurement of air temperature at a point in space and time, with the unit degree Celsius and resulting of a one time measurement with a quicksilver thermometer. But, it might as well express the monthly average temperature.

In the presented taxonomy, ambiguity is not specified any further. A detailed separation of ambiguities is provided by Scheffler (1979). Other authors suggest adding discord as a sub-concept of ambiguity (figure 1). Discord arises if expressions are clearly specified, but are instantiated differently according to varying interpretations. In this work discord is conceptualised as a situation that may be caused by ambiguity instead of being a specific kind of it.

3.1.2 Generality: General concepts conceptualise multiple phenomena, no matter how dissimilar these may be or what similarity criteria are used (Scheffler, 1979). The concept of being *cold*, for example, can be used to conceptualise air temperature at a certain location in space-time, but can also be used to characterise the personality of somebody. In the same way, the concept of *temperature* can be used in one context to conceptualise the temperature of air, in another the one of water.

Generality of concepts is frequently mixed up with vague concepts. Considering general concepts, membership is always "true" or "false"; there are no borderline cases. Thus, generality differs from vagueness (see also section 3.2.2).

General expressions are symbols expressing many things, no matter how dissimilar these may be and what similarity criteria are used (Scheffler, 1979). They reflect generality at the concept level. Considering attribute values for example, "cold" is a general expression.

This kind of uncertainty is frequently mixed up with ambiguity. For example in (Fisher, 1999; Fisher et al., 2006; Klir and Yuan, 1995), generality appears as "non-specificity" and is considered a specific case of ambiguity. In case of ambiguous expressions, membership may shift from "true" to "false" when changing the interpretation. This does not hold for general expressions. Being "cold", for example, is always evaluated "true" or "false", once the value is assigned. This is independent of whether the specific use of the attribute value is to characterize air temperature at a specific location in space-time, the character of a certain person or anything else.

3.1.3 Open texture: Open texture targets membership assignments to concepts (Scheffler, 1979). It is present, if the situations causing membership to a specific concept cannot be defined distinctively. Thus, it can be seen as an under-specification on purpose. In case of open textured concepts, a human being is required to judge for the specific instantiation. Similarly, open texture might be defined on expressions. Here the expressions reflect open texture at the concept level.

Open textured concepts (respectively expressions) do not occur in relation to the air temperature example. Nevertheless, they frequently arise in the geospatial domain as soon as the observed moment becomes more complex. It is common practice that domain experts, like landscape ecologists assign attribute values based on personal experiences. Consider, for example, *land use* as a dimension of the concept *earth's surface*. Domain experts use their practical know-how to assign class membership for land use based on individual expertise. It was shown previously that such expert opinions may be very heterogeneous (Comber, 2002).

3.2 Well-Defined Interpretation

The discussion in this section assumes that the expressions used in the vocabulary are defined clearly. Each facet of uncertainty discussed here is orthogonal to any other category introduced in section 3. Every facet may occur in conjunction with one or more of the others.

3.2.1 Incomplete Information: Error is the most prominent occurrence of incomplete information in the geospatial domain. Fisher targets this facet in his categorisation (Fisher et al., 2006). Anyhow, incomplete information covers more than just errors. It acknowledges the fact, that the complete truth of a situation cannot be accessed in many cases. This view corresponds to what is called uncertainty in the Artificial Intelligence (AI) community (Russell and Norvig, 2003). In order to reason under incomplete information, data models may explicitly represent dependencies between the values of two or more different attributes (Russell and Norvig, 2003).

Error as one aspect of incomplete information does not apply to expressions and concepts, but to value assignments. Following Heuvelink and others (Heuvelink, 1998; Goodchild, 2004), error is conceptualized as the difference between the "true value" and the represented one. Errors caused by transportation of information over varying channels (Shannon, 1948) are out of the scope of semantic reference systems. Instead, accuracy due to measurement and subsequent errors during operation (Heuvelink, 1998) is focussed on. Error is always present when relying on empirical data. Due to the nature of observation, it can never be eliminated entirely. In the example, measurement error occurs when defining spatio-temporal location of an observation, as well as during the measurement of temperature values.

3.2.2 Vagueness: A concept is vague if instantiations may cause “borderline cases” (Tomai and Kavouras, 2004). A borderline case occurs if the membership of a certain instance to the concept cannot be defined with “true” or “false”. It may not give clear criteria for being a member. For example, consider the concepts *cold*, *warm* and *hot* used as attribute values for temperature. An air temperature of plus 20 degrees Celsius might be conceptualised as *warm* and one of minus 5 degrees Celsius as *cold*. But, how to conceptualise 14 degrees Celsius? How 13.5?

The use of the example concept *cold* in the paragraph above and the use of a similar example in section 3.1.2 illustrates that vagueness is in fact orthogonal to generality.

Expressions articulating vague concepts reflect their characteristics. Vague expressions denote vague concepts in human minds. Scheffler defines vagueness of expressions as indeterminacy and concomitant interpretations and as ambivalence in deciding the applicability of an expression to an entity (Scheffler, 1979). Take “cold”, “warm” and “hot” as used in every day language as an example.

Unlike Fisher’s suggestion, vagueness is not categorized as “poorly-defined”. Whereas Fisher targets objects as subject to poor definitions, here expressions used to define object models are in focus. Vague expressions represent the vague nature of a concept in an object model, not the possibility of misinterpretation. This is what interferes with the definition of “poorly-defined interpretation”. Consequently, vagueness is a facet of uncertainty that relates to well-defined interpretation.

The revised categorization of uncertainties is summarized in figure 2. The connections indicate subcategory relations.

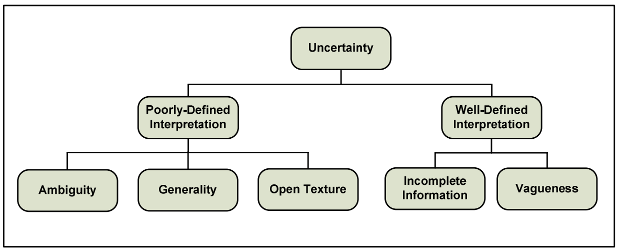

Figure 2. Fisher’s categories of uncertainty revised to accentuate semantic aspects.

In contrast to Fisher’s initial proposal, vagueness is not categorised as poorly-defined. The notion of ambiguity is more specific and separated from other facets of poorly-defined interpretations. Effects caused by ambiguity have been rejected, because they are not particular categories of uncertainty.

4. UNCERTAINTY MEETS SEMANTIC REFERENCE SYSTEMS

The categories of uncertainty are used to derive the requirements for their consideration within semantic reference systems.

4.1 Requirements to Clarify Intended Interpretations

The reduction of possible interpretations to the intended ones is topic to the reference frame, the datum and to the process of semantic referencing. Intended interpretations need to face ambiguity, generality and open texture.

4.1.1 Requirements Posed by Ambiguity: Natural language is subject to ambiguity, as are semi-formal concept representation approaches. Requirements are:

A1. Usage of formal approaches, in order to reduce possible interpretations of the vocabulary to the one intended by a specific community. Since taxonomic relationships between concepts alone proved to be too weak for the required specifications, axioms to express further relationships have to be provided (Guarino, 1998; Kuhn, 2005).

A2. The formal specifications, i.e. ontologies, need to be free of cyclic references.

A3. The primitives used in ontologies need to be grounded outside the ontologies.

4.1.2 Requirements Posed by Generality: General expressions are loosely defined. Their specific meaning is only given by a particular context. The same expression may appear in varying contexts, but with distinct specialized meaning. Accounting for generality requires:

G1. The possibility to use the same concept specification in varying contexts with different meanings.

G2. The exact definition of the context of use related to a specific vocabulary.

G3. A distinction between contexts.

G4. The assurance that general expressions which appear in vocabularies are put into the intended context.

4.1.3 Requirements Posed by Open Texture: The use of open textured expressions is highly subjective, because it depends on personal interpretation of facts given in a distinct situation. One requirement is:

O1. Identifying the agent which assigned a membership, i.e. which interpreted the available expressions and chose a specific assignment.

Open texture does not apply to the current example. The nature of open texture requires further research.

4.2 Requirements if Intended Interpretations are clarified

Incomplete information and vagueness constrain the formalisms and languages, which are able to implement semantic reference systems. Models for measurement error need to be included in the semantic frame and the propagation of errors relates to semantic projections and transformations.

4.2.1 Requirements Posed by Incomplete Information: Considering complex application scenarios, usually only partial information is available (Russell and Norvig, 2003). The state of the model cannot be generated completely. Since, data models may explicitly represent dependencies between data values, semantic reference systems require:

II1. A possibility to express models that capture information dependencies.

Errors (or the inverse of error: accuracy (Veregin, 1999)) affects spatial, temporal and thematic information. Expressions capturing error or accuracy information could be semantically referenced to ontologies provided in semantic reference systems. Considering measurement errors **all** of the following requirements need to be concerned:

E1. Ontologies constituting reference frames for information items have to account for data acquisition, i.e. measurement errors. Accordingly, models for attribute values need to be extended. The required extensions include a grounding in the semantic datum.
E2. Methods for specifying the error propagation mechanisms, which are used within data models, have to be defined.
E3. The propagation of error values needs to be enabled for all functionalities of semantic reference systems, i.e. for projections of and transformations between semantic reference frames.

In addition to these rather technical requirements, an ontological requirement needs to be considered (Navratil and Frank, 2006). In view of comparison operators on temperature values as example, measurement error poses the following question:

- Is air with a temperature of 26.8 degrees Celsius and a standard deviation of 0.5 degrees Celsius *colder* then air with a temperature of 27 degrees Celsius?

4.2.2 Requirements Posed by Vagueness: Vague concepts need to be described appropriately in the semantic reference system. Appropriately means explicitly ignoring the vague characteristics or using a formalisation and according representation language accounting for vagueness.

If vagueness is ignored, concept definitions remain crisp. Membership is restricted to “true” and “false”. Only crisp links from community vocabularies, i.e. expressions used in data models, to the reference frame are allowed.

Experiences in AI showed that ignoring vagueness leads to systems without extensive practical use (Russell and Norvig, 2003). Accounting for vagueness requires **at least one** of the following points:

V1. The semantic reference frame offers the possibility to describe vague concepts and the required primitives are grounded in a semantic datum.
V2. Instantiations within the semantic reference frame allow the assignment of vague membership values.
V3. The semantic reference frame stays crisp, but possibilities for vague semantic referencing are provided.

5. SEMANTIC REFERENCE SYSTEMS ACCOUNTING FOR UNCERTAINTY

Existing work targeting the identified requirements are put into the context of the semantic reference system approach. The measurement of physical unary moments, such as temperature, and their description constitute the basis for designing and implementing semantic reference systems. Following the idea of measurement-based GIS (Buyong et al., 1991; Goodchild, 2002; Goodchild 2004) these considerations form the foundation for defining reference systems for further geospatial information. Measurement-based GIS retain basic observations (corresponding to physical unary moments) as a foundation for all geospatial information. Derivation functions are applied to basic observations to obtain more complex information.

It turns out, that semantic reference systems, as currently envisioned, target only parts of the “poorly-defined interpretation” branch of the diagram (figure 2). References to approaches accounting for the remaining parts are identified.

5.1 Accounting for Ambiguity

Accounting for ambiguity is at the heart of each semantic reference system. The whole notion developed in order to meet requirements A1 and A3. Ontology engineers are responsible for eliminating cyclic references (A2) during the ontology development process.

Explicitly for measurements, a reference frame has been put into place (Probst and Espeter, 2006). This approach applies and refines conceptual spaces (Gärdenfors, 2000) to specify observable moments of entities. Along each dimension, reference units are established to specify the meaning of measurement results (A1). Semantic grounding (A3) is reached by aligning the measurement ontology to the foundational ontology DOLCE (Masolo et al., 2003). The provided frame can be used directly to dissolve ambiguities in vocabularies relating to measurement of unary moments.

5.2 Accounting for Generality

General concepts can be put into different contexts (G1), by using the construct of thematic roles (Sowa, 1999). In this way, *cold* may play the role of an attribute for temperature in one context and of an attribute value to characterize a certain person in another.

Defining the exact context of use within a vocabulary (G2) is in the responsibility of the person or machine implementing the semantic referencing. After the distinction between contexts (G3) has been clarified by the ontology engineers defining semantic reference frames, the references need to be established correctly. Contexts can be distinguished from each other, because the formal model promotes dissolving ambiguities. The identification of general expressions (G4) remains with the person defining the semantic reference between the vocabulary and the frame.

Before context can be targeted in more detail, the notion of context itself requires a precise specification. As a recent survey of 150 definition of context illustrates, the notion is frequently ill-defined (Bazire and Brézillon, 2005).

5.3 Accounting for Open Texture

Although open texture is not relevant for basing semantic reference systems in measurements, it applies frequently in geospatial information data models, where pure factual information is not available. The report on the expert which interpreted the open textured expression (O1) is a pragmatic way to account for open texture. This is mainly a question of metadata recording.

Regarding further investigations, research results from the domain of AI might be of value. Specific work exists concerning open textured expressions in laws (Bench-Capon, 1999). In terms of existing theories Dempster-Shafer’s theory of evidence (Dempster 1967, Shafer, 1976) might prove useful for including models of reliability between expert opinions into semantic reference frames. Dempster-Shafer Theory can be regarded as a generalised Bayesian approach (Dempster, 1968), which is introduced in the following subsection. Expert judgements can also be targeted using Cohen’s Endorsement Theory (Cohen, 1985). Approaches of combining both theories are already carried out (Wadsworth and Hall, 2007).

5.4 Accounting for Incomplete Information

Concerning the inclusion of incomplete information (II1) into semantic reference frames, probabilistic models seem most appropriate (Russell and Norvig, 2003). Bayesian networks are the standard means to implement such models. Approaches to include Bayesian networks in ontologies have already been proposed (Gu et al., 2004; Costa and Laskey, 2006). They target either the inclusion of an additional ontology in order to translate instances into models that can be accessed by Bayesian network algorithms or extend the language used to represent the ontologies as such.

For measurements with ratio scaled units, the standard deviation with root mean square error may be used as a first error model (Heuvelink, 1998). If more then one observable is targeted, for example if air temperature and pressure are measured, correlations can be captured in variance-covariance matrixes (E1). If operators are offered as part of the data model, those include error propagation laws (E2). Available reference frames need to be extended accordingly.

Besides the extensions in the reference frame and possibly in the datum, transformations need to account for error propagation (E3). In the simplest case of a unit transformation, only the unit of the standard deviation attached to the measured value has to be changed. Concerning more complex transformations, a first order Taylor approximation might be used (Goodchild, 2004). Since this method only approximates the error for all non-linear transformations, improvement might be required in the long term.

5.5 Accounting for Vagueness

Accounting for vagueness in semantic reference systems requires analyzing classical approaches to vagueness, model-theoretical and of the epistemic kind (Sorensen, 1988). Fuzzy logic (Zadeh, 1965) and supervaluationism (Fine, 1975; Kamp, 1981) are the two common examples for the former and have both been widely applied to geospatial information (Bennett, 2001; Bennett and Cristani, 2003; Bittner and Smith, 2001). Epistemic vagueness (Vardi, 1986) is (up to the author's knowledge) hardly applied within the domain of GI.

Although the use of fuzzy membership functions has often been criticized (Keefe and Smith, 1997), information communities may use either of the approaches. In case that two communities with different approaches desire to communicate, meta-level mappings between the approaches need to be defined. This might lead to ignorance of vagueness in practice.

On the implementation level, accounting for vague concepts and grounding relevant primitives (V1) can be done by defining an ontology for vagueness that captures the conceptualization underlying varying approaches. This ontology then becomes a basic element of the semantic reference frame. An alternative is the extension of the formalism used to encode the reference frames.

Vague instantiation of attribute values to attributes (V2) can be implemented in the semantic reference frame by assigning vague memberships. As before, this requirement can be met by extending the ontology or the formalism as such.

If the semantic reference frame should remain crisp (V3), either of the approaches accounting for vagueness may be implemented during the semantic referencing. In this case, the link between the user vocabulary and the reference frame is vague.

5.6 May Parts of Semantic Reference Systems Cause Uncertainties?

Considering manual definitions of projections and transformations, uncertainties are bound to human error. Apart from that drawback, mathematically defined mapping functions do not introduce error.

Considering the vision that semantic reference systems allow for automated similarity transformations, uncertainties may be caused by semantic transformations. Here, probabilistic approaches as introduced in section 5.4 might aid in quantifying the arising uncertainty.

6. CONCLUSIONS AND FUTURE WORK

The notion of semantic reference systems has been related to categories of uncertainty. The developed categorization revises an earlier attempt by Fisher and others (Fisher, 1999; Fisher et al., 2006) accentuating semantic issues. Ambiguity, generality or open texture may cause an expression to require further specification. Once the intended interpretations are defined, incomplete information and vagueness require additional consideration.

All five categories of uncertainty pose distinct requirements towards the design of semantic reference systems for geospatial information, but indicate the power of this approach to deal with uncertainty in a general sense. One possibility to account for ambiguity concerning physical unary moments is offered by Probst and Espeter (2006). This work provides the basis for extending semantic reference system to account for uncertainty. Concepts from measurement-based GIS (Buyong et al., 1991; Goodchild, 2002) can be used to establish semantic reference systems for more complex data models of geospatial information in the future.

Surveying the possibilities to account for generality, incomplete information and vagueness in semantic reference systems is part of ongoing work. The current status has been reported in this paper. The combination of the remaining facets of uncertainty with semantic reference systems for basic physical measurements is the next step. In includes the elaboration of extending the approach followed by Probst and Espeter (Probst 2006; Probst and Espeter, 2006), as well as, analysing alternatives that are based on the distinctions framework suggested by Frank (2006). Open textured concepts do not need to be considered in the current attempts. They will be focused when upgrading to more complex measurements.

Practically, experimentation regarding models of air temperature will be focussed. On the theoretical level, the suggested categorisation of uncertainties will be completed by adding metaphor (already occurred as an example at the beginning of section 3.1.2) to the branch of "poorly-defined interpretation", and by separating incomplete information from error. It is planned to add one additional level to the taxonomy, where especially ambiguity will be examined in more detail. This work will be based on achievements by Scheffler (1979).

ACKNOWLEDGEMENTS

The work was carried out while the author was funded by the European Commission under the *SWING* project (FP6-26514, http://www.swing-project.org). I thank various members of the MUSIL group (http://musil.uni-muenster.de) for their valuable input.

REFERENCES

Asperti, A. and G. Longo, 1991. *Categories, Types and Structures - An Introduction to Category Theory for the Working Computer Scientist*. MIT Press, Cambridge, Massachusetts.

Barr, M. and C. Wells, 1990: *Category Theory for Computer Science*. Prentice Hall, London.

Bazire, M. and P. Brézillon, 2005. Understanding Context Before Using It. In: *Proceedings of CONTEXT 2005*, Paris, France, pp. 29-40.

Bennett, B., 2001. What is a Forest? On the Vagueness of Certain Geographic Concepts. *Topoi*, 20(2), pp. 189-201.

Bennett, B. and M. Cristani, 2003. Special issue on spatial vagueness, uncertainty and granularity. *Spatial Cognition and Computation*, 3(2,3), pp. 93-255.

Bench-Capon, T. 1999. Knowledge Based Systems in the Legal Domain. Article for: *Encyclopaedia of Computer Science and Technology*, Marcel Dekker, Inc, New York, pp. 163-186.

Bittner, T. and B. Smith, 2001. A unified theory of granularity, vagueness and approximation. In: *Proc. of the 1st Workshop on Spatial Vagueness, Uncertainty, and Granularity (SVUG01)*, Morro Bay, California, USA.

Bowers, S. and B. Ludäscher, 2004: An Ontology-Driven Framework for Data Transformation in Scientific Workflows. In: *Proc. of the 1st International Workshop on Data Integration in the Life Sciences (DILS'04)*, Leipzig, Germany, pp. 1-16.

Buyong, T., A. Frank and W. Kuhn, 1991. A Conceptual Model of Measurement-Based Multipurpose Cadastral Systems. *Journal of the Urban and Regional Information System Association,* 3(2), pp. 35-49.

Chrisman, N., 2002. *Exploring Geographic Information Systems*. 2nd edition, Wiley, London, pp. 19-23.

Cohen, P., 1985. *Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach*. Pitman, London.

Costa, P. and K. Laskey, 2006. PR-OWL: A Framework for Probabilistic Ontologies. In: *Proc. of Formal Ontology and Information Systems (FOIS 2006),* Baltimore, Maryland, IOS Press, Amsterdam, pp. 237-249.

Comber, A., 2002. *Automated Land Cover Change*. PhD Thesis, University of Aberdeen, Scotland.

Dempster, A., 1967. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2), pp. 325-399.

Dempster, A., 1968. A generalization of Bayesian inference. *Journal of the Royal Statistical Society*, B (30), pp. 205-247.

Fine, K., 1975. Vagueness, Truth and Logic. *Synthese*, 30, pp. 265-300.

Fisher, P., 1999. Models of Uncertainty in Spatial Data. In: *Geographical Information Systems – Principles and Technical Issues*. Wiley & Sons, New York, Vol. 1, pp. 191-205.

Fisher, P., A. Comber and R. Wadsworth 2006. Approaches to Uncertainty in Spatial Data. In: *Fundamentals of Spatial Data Quality*, ISTE, Ltd., London, pp. 43-59.

Frank, A., 2006. Distinctions Produce a Taxonomic Lattice: Are These the Units of Mentalese? In: *Proc. of Formal Ontology and Information Systems (FOIS 2006),* Baltimore, Maryland, IOS Press, Amsterdam, pp. 27-38.

Gärdenfors, P., 2000. *Conceptual Spaces*. MIT Press, London, pp. 1-30.

Goodchild, M., 2002. Measurement-based GIS. In: *Spatial Data Quality*, Taylor & Francis, New York, pp. 5-17.

Goodchild, M., 2004. Special Issue on Measurement-Based GIS. *Journal of Geographical Systems*, 6(4), pp. 323-428.

Gu T., H. Pang, D. Zhang, 2004. A Bayesian Approach for Dealing with Uncertain Contexts. In: *Proc. of Second International Conference on Pervasive Computing,* Vienna, Austria, Austrian Computer Society, pp. .

Guarino, N., 1998. *Formal Ontology in Information Systems. In: Proc. of Formal Ontology and Information Systems (FOIS'98)*, Trento, Italy. IOS Press, Amsterdam, pp. 3-15.

Guizzardi, G., H. Herre and G. Wagner, 2002. On the General Ontological Foundations of Conceptual Modelling. In: *Proc. of the 1st International Conference on Ontologies, Databases and Applications of Semantics* (ODBASE 2002).

Iliffe, J., 2000. *Datums and Map Projections*. Whittles Publishing, Caithness, p. 52.

Kamp, H., 1981. The Paradox of the Heap. In: *Aspects of Philosophical Logic*, Reidel, Dordrecht, pp. 225-277.

Keefe, R. and P. Smith, 1997. *Vagueness: A Reader*. MIT Press, London.

Klir, G. and B. Yuan, 1995. *Fuzzy Sets and Fuzzy Logic. Theory and Applications*. Englewood Cliffs, Prentice Hall, London.

Kuhn, W., 2003. Semantic Reference Systems. *International Journal of Geographic Information Science* (Guest Editorial), 17(5), pp. 405-409.

Kuhn, W., 2005. Geospatial Semantics: Why, of What, and How? *Journal on Data Semantics, Special Issue on Semantic-based Geographical Information Systems,* Springer Lecture Notes in Computer Science 3534, pp. 1-24.

Kuhn, W. and M. Raubal, 2003. Implementing Semantic Reference Systems. In: *Proc. of the 6th AGILE Conference on Geographic Information Science (AGILE 2003)*, Lyon, France, Presses Polytechniques et Universitaires Romandes, pp. 63-72.

Manning, C. and H. Schutze, 2001. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, Massachusetts.

Masolo, C., S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, 2003. “WonderWeb Deliverable D18, Ontology Library (final)”, Padova, Italy. http://wonderweb.semanticweb.org/deli erables/documents/D18.pdf (accessed 10 Jan. 2007).

Navratil, G. and A. Frank, 2006. What does Data Quality mean? An Ontological Framework. In: *Proc. of* Angewandte *Geoinformatik (AGIT* 2006), Salzburg, Austria, Wichmann, pp. 494-503.

OGC, 2003. *The OpenGIS Guide - Introduction to Interoperable Geoprocessing and the OpenGIS Specification (3rd Edition)*. OpenGIS Consortium, Wayland, Massachusetts.

Probst, F., 2006. An Ontological Analysis of Observations and Measurements. In: *Proc. of the 4th. International Conference on Geographic Information Science (GIScience 2004)*, Münster, Germany, pp. 304-320.

Probst, F. and M. Espeter, 2006: Spatial Dimensionality as Classification Criterion for Qualities. In: *Proc. of Formal Ontology and Information Systems (FOIS 2006),* Baltimore, Maryland, IOS Press, Amsterdam, pp. 77-88.

Russell, S. and P. Norvig, 2003. *Artificial Intelligence - A Modern Approach.* 2nd edition, Prentice Hall, New Jersey, pp. 462 - 583.

Shafer, G,. 1976. *A Mathematical Theory of Evidence.* Princeton University Press, Princeton.

Shannon, C., 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27, pp. 379–423 & 623–656, July & October.

Scheffler, I., 1979. *Beyond the Letter – A Philosophical Inquiry into Ambiguity, Vagueness and Metaphor in Language.* Routledge & Kegan Paul, London.

Sorensen, R., 1988. *Blindspots*. Clarendon Press, Oxford.

Sowa, J., 1999. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks Cole Publishing Co., Pacific Grove, California, pp. 506-510.

Tomai, E. and M. Kavouras, 2004. From “Onto-GeoNoesis” to “Onto-Genesis”: the Design of Geographic Ontologies. *Geoinformatica*, 8(3), pp. 285-302.

Vardi, M., 1986. On Epistemic Logic and Logical Omniscience. In: *Proc. of Theoretical Aspects of Reasoning about Knowledge*, Los Altos, California, Morgan Kaufmann, pp. 293-305.

Veregin, H., 1999. Data Quality Parameters. In: *Geographical Information Systems – Principles and Technical Issues*. Wiley & Sons, New York. Vol. 1, pp. 177-189.

Zadeh L., 1965. Fuzzy Sets. *Information and Control*, 8, pp. 338-353.