Towards automated knowledge-based mapping between individual conceptualisations to empower personalisation of Geospatial Semantic Web

Pragya Agarwal[1], Yongjian Huang[2], Vania Dimitrova[3]

School of Computing, University of Leeds, UK
Email: [1]pragya@comp.leeds.ac.uk, [2]scs4yh@comp.leeds.ac.uk, [3]vania@comp.leeds.ac.uk

Abstract. Geospatial domain is characterised by vagueness, especially in the semantic disambiguation of the concepts in the domain, which makes defining universally accepted geo- ontology an onerous task. This is compounded by the lack of appropriate methods and techniques where the individual semantic conceptualisations can be captured and compared to each other. With multiple user conceptualisations, efforts towards a reliable Geospatial Semantic Web, therefore, require personalisation where user diversity can be incorporated. The work presented in this paper is part of our ongoing research on applying commonsense reasoning to elicit and maintain models that represent users' conceptualisations. Such user models will enable taking into account the users' perspective of the real world and will empower personalisation algorithms for the Semantic Web. Intelligent information processing over the Semantic Web can be achieved if different conceptualisations can be integrated in a semantic environment and mismatches between different conceptualisations can be outlined. In this paper, a formal approach for detecting mismatches between a user's and an expert's conceptual model is outlined. The formalisation is used as the basis to develop algorithms to compare models defined in OWL. The algorithms are illustrated in a geographical domain using concepts from the SPACE ontology developed as part of the SWEET suite of ontologies for the Semantic Web by NASA, and are evaluated by comparing test cases of possible user misconceptions.

Keywords: *Semantic Web, Geospatial, Ontologies, User, Personalisation*

1. Introduction

'*The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation*.'- Tim Berners-Lee

The idea of a Semantic Web introduced by Berners- Lee *et al.* (2001) proposed 'a web of data that can be processed directly or indirectly by machines, bringing a higher degree of automation in exploiting data in a meaningful way'. The knowledge-based web where computers are empowered with machine-processable semantics and able to assist humans in their every day life (Berners-Lee *et al.* 2001) is becoming a reality due to the enormous research effort worldwide. Many services are now being automated so that they can understand the content on the web, are able to process a vast amount of knowledge, and provide more accurate and effective search of information resources (Ding et al. 2005). The recent developments in the Semantic Web have great potential for the geospatial community, in specific, because the focus on the incorporation of data semantics will lead to better retrieval and more reliable integration methods by tapping into the semantics during the search process on the web. Nevertheless, the basic semantic web and the technological developments are not targeted to the specific needs of the geospatial community. The idea of a more focussed 'Geospatial Semantic Web' has been recognised as a research priority by UCGIS (Fonseca and Sheth 2002). There is a distinct move away from structure and syntax in the geospatial community accompanied by an increased awareness that semantics is the backbone for a successful ontology to enable translation of data from different resources and users. Agarwal (2005) discuss in detail the problems associated with ontology development in the geo-spatial domain primarily due to semantic ambiguities. However, the use of geo-ontologies has not yet been tapped into to its full potential, as one absolute conceptualisation of the real world will not satisfy any reliable information processing of the geospatial data since the different individual semantic perspectives are not necessarily taken into account.

Egenhofer (2002) identified the need to support queries based on meanings and better definition of spatial terms across a number of disciplines, and the need to integrate multiple terminological ontologies as a backbone for an effective Geospatial Semantic Web (GSW). Support ontologies for a GSW will also need to be able to integrate multiple perspectives and multiple semantic conceptualisations for geographic terms and concepts. Success of a standardised geo-ontology for the semantic web will be determined by the level of acceptance by the users of the services- both experts and naïve, and the level to which the basic geo-ontology is semantically compatible with the users' conceptualisations. Users' preferences, expectations, goals and tasks differ while using the web for information resources. Moreover, people form different conceptual models of the world and these models dynamically change over time. The knowledge-enhanced web services are normally driven by some description of the world which is encoded in the system in the form of an ontology defined by knowledge engineers. The users' conceptualisation of the world may differ, sometimes significantly, from the conceptualisation encoded in the system. If not taken into account, the discrepancies between a user's and a system's conceptualisation may lead to the user's confusion and frustration when utilising Semantic Web services, which, in turn, can make these services less popular.

With the increasing reliance on the WWW (World Wide Web) and the technological developments in the search engines and the interface as well as the vast amount of spatial and earth sciences data and resources now located and available over the web, the number of people who use Semantic Web services is expanding, and hence, dealing with user diversity and providing personalisation functionality becomes paramount (Henze and Herrlich 2004, Dolog *et al.* 2003). Co-operation is the key word in the vision for a Semantic Web. There is a strong necessity to include the people as an axis in the design, development, and deployment of semantically enriched services. The development of a geo-spatial semantic web will also include management of the geo-ontologies and the need to update these and integrate these with the different conceptual models of the real world from the users. The UCGIS research initiative recognises the need for the GSW to be a collaborative process. The meaning of the geospatial information is shaped and changed by the interaction of people and systems. However, in most systems, there are no mechanisms to capture this interaction and context of the user. Therefore, there is a need for research in the direction where integration of ontologies can be carried out based on a co-ordinated effort to align these and identify the discrepancies between different semantic conceptualisations and multiple perspectives. Computational models are needed that can process the different terminological and semantic ontologies and process the semantic incompatibilities between users and the expert's geo-ontology. Personalisation of semantic web is required to exploit the user intentions and perspectives and for automated reasoning tools that can detect mismatches and discrepancies between the user ontologies and the expert ontology that forms the basis or the backbone for the web-based resources.

The one-size-fits-all-users approach to developing web applications is becoming outdated. Personalized information systems aim at giving the individual user optimal support in accessing, retrieving, and storing information. Personalisation of web-based services can be achieved from different perspectives. Many different research disciplines have contributed to explore personalization techniques and to evaluate their usefulness within various application areas: e.g. hypertext research has studied personalization in the area of so-called adaptive hypertext systems, artificial intelligence techniques have been widely used to cluster web data, usage data, and user data, reasoning and uncertainty management has been adopted to draw conclusions on appropriate system behaviour. Previous work, such as GLUE, has also attempted to apply machine learning approaches to ontology mapping on the semantic web using heuristics and multi-strategy learning approaches (Doan *et al.* 2002).

This paper outlines the development of an automated approach to user modelling and to aligning and co-ordinating different conceptualisations for achieving personalised data integration for a GSW. Algorithms are developed from the formal approach to compare different user models based in Ontology Web Language (OWL). Patterns of discrepancies between a user and a system's conceptualisation are analysed and a formal approach is proposed, based on Description Logic (DL), to define these patterns. OWL-based rules are then derived and implemented in a demonstration prototype that compares an expert ontology and a user's conceptualisation, both represented in OWL. The discrepancies are identified and registered as misconceptions between the different models and allows the identification of mismatches between individual conceptualisations in the domain. The innovativeness of this paper lies in proposing personalisation approach for a geospatial semantic web by formalising semantic mismatches, developing algorithms based on these formalisations, combining knowledge elicitation methods and user models with ontology mapping and integration approach, and developing test-bed for evaluation of core geo-ontologies against multiple conceptualisations to allow integration of different perspectives in information systems.

The rest of the paper is structured as follows. Section 2 proposes an argument for the semantic personalisation of web-based geo-spatial services. A brief review of previous methods and initiatives, relevant to the purposes of this paper, in user modelling and personalisation, and in ontology mapping is presented in section 3. The overall methodological framework for the approach outlined in this paper is presented in section 4, where the ontology domain is introduced, and the knowledge elicitation process is explained. The formalisation approach is outlined in section 4 and the basic assumptions and notations made clear. The demonstration of the algorithms based on the DL formalisations is presented using the domain ontology for semantic web and test cases of user conceptualisations in section 5, and the results summarised. The conclusions and future research directions are outlined in section 6.

2. Semantic Personalisation of the Geospatial Web

The personalization of web-based services has been a prime concern of the user-modelling community which deals with methods for gaining some understanding of users, i.e. a user model, and using that understanding to tailor the system's behaviour to the needs of individuals. Existing design paradigms in geo-spatial services need to be redefined to deal with the new challenges brought by the need to deal with a diverse user population having different preferences, goals, understanding of tasks, conceptual models, etc. Added to this is the vast number of sources provided by tracking the users' activities to discover patterns of using the web in different application areas. Furthermore, new diagnostic techniques and models are needed to capture the long-term development of users' capabilities, the dynamics of user's goals and conceptual understanding, the uncertainty and inconsistency of naive users' conceptualizations, and so on. The ambitious target is to offer manageable, extendible and standardized infrastructure for complementing and collaborating applications tailored to the needs of individual users. Without the benefit of deeper semantic or ontological knowledge about the underlying domain, personalization systems cannot handle heterogeneous and complex objects based on their properties and relationships. Nor can these systems possess the ability to automatically explain or reason about the user models or user recommendations. This realization points to an important research focus that combines the strengths of Web mining with semantic or ontological knowledge. Traditional personalization and adaptation architectures were suited to deal with closed-world assumption, where user modelling methods, such as overlay, bug library, constraint-based modelling and other marked discrepancies in a user and expert's semantics as erroneous, and often called them misconceptions. New approaches for open-world user modelling that facilitate elicitation of extended models of users, are able to deal with the dynamics of a user's conceptualization and that acknowledge semantic discrepancies and heterogeneity are required to effectively personalize the Semantic Web for the Geo-spatial community.

Personalization functionality on the Semantic Web has to be implemented and applied to deal with user diversity. The open-world assumption of the Semantic web refers to the need to take into account user viewpoints ranging from domain experts to complete novices. Rather than a closed view of the world, the personalisation efforts for geo-spatial services design will ensure that the different perspectives and semantic conceptualisations of the real world are maintained as 'open'. The idea is to have a basic core ontology that encapsulates the primary concepts, terms, relations and properties and the services allow the users to access the knowledge base according to their individual conceptual models of the world. The approach defined in this paper is an effort to allow the system to reconcile the user conceptual model with the core ontology and therefore identify the discrepancies and similarities, and thereby allowing the system to identify the differences in the user conceptualisations with the so-called expert ontology. This will allow, first, for the development of systems that allow personalisation by incorporating user models and diversity and second, as a means to test any core ontologies that are developed as the basis for a geo-spatial services against user conceptualisations for discrepancies and thereby evaluate its reliability as a standard, re-usable ontology. Moreover, the personalisation approach allows flexibility and the possibility of using the user models to enrich the available information resources with shared semantics instead of relying on fixed ontologies available to the developers at the design stage.

In this paper, we focus on user modelling and alignment of different semantic models for personalisation. Using this approach, the research towards specification of well-defined standards and ontologies for inter-operability in the geo-spatial domain can be enhanced and personalised to provide extendibility, flexibility, interoperability, and re-usability. The underlying principle for the methodology adopted in this paper is that an ontology, whatever the scale or granularity, maps the tacit knowledge from the real world (Smith 2003), and makes this knowledge explicit by specification of relations and rules. Ontologies are also proposed as methods to resolve semantic heterogeneity in the geographic context (Hakimpour and Timpf 2001). Maedche and Staab (2002) say that ontologies play an important role for many knowledge-intensive applications by providing a source of precisely defined terms. Ontology, therefore, conceptualises and codifies the knowledge in a domain, and can be mapped as a knowledge domain. The proposed approach for reconciling user diversity and of creating an open system for geo-processing services consists of a formal mechanism by which the semantic mismatches between different models can be identified. The approach developed in this paper is based on a critical review of previous work carried out in ontology alignment and integration (Agarwal 2004b), a brief overview of which is presented in the following section.

3. Overview of existing approaches for ontology alignment and mapping

A number of methods are proposed for ontology integration and there is not always an agreement in this community on the meanings of terms that are used (Klein 2001). A review of different terms and their commonly employed definitions (Uschold *et al.* 1998, Klein 2001, Wache *et al.* 2001, Noy and Musen 2002) reveals that there is some confusion in meanings across various terminological frameworks. In general, integration results in a new ontology, maintaining relations and concepts that are consistent in the two ontologies, while merging results in a new version of the ontology by including all overlapping information from the ontologies. Although merging is less reductive than integration, Silva and Rocha (2003) say that merging is applicable for constructing single data repositories from various data sources and is time-consuming and complex. Alignment, on the other hand, is focussed on a concept-level approach and on finding corresponding semantic properties in the two ontologies (Wache *et al.* 2001). The mapping process during alignment is goal-oriented and since alignment is only carried out for parts of the ontology with corresponding concepts and semantics, the demands on resources, in time and system, are not as high as for merging. In the context of the approach adopted in this paper, '***alignment***' and '***mapping***' are most relevant, as these are the least reductive of all methods maintaining the semantic consistency and coherence of the original ontology while comparing and mapping across the different ontologies.

Resolution of semantic differences is more crucial than syntactic resolution for aligning ontologies, and conflicts in terminological mismatch are of greater concern while developing re-usable and shareable models from a comparison and similarity assessment of existing ontologies. For the personalisation of geospatial services, the aim is to find points of mismatches between different conceptual models to define ways of either reconciling these differences or using the variability in semantics to find the most suitable information source from the available resources. Therefore, mapping is goal-oriented, with a definitive articulation, and the issue of finding terms on which to align the ontology is not relevant. Alignment is identified here as an appropriate mechanism for knowledge mapping and for finding conceptual associations across the diverse range of ontologies because of the non-reductive nature of this method and because it allows for semantic explication. Merging and integration, as mapping processes, are not relevant because the aim is not to develop an integrated resource, but to compare and identify similarities and differences in ontologies for meanings and conceptualisations of different terms and concepts from different user models. In this case, the different ontologies or user models can be considered as version of the same ontology and hence the term '***ontology versioning***' can also be used for the techniques applied here. This will become clearer from the discussions that follow.

3.1 Existing methods and tools for mapping diverse knowledge resources-problems and limitations

Standardisation is required to achieve semantic inter-operability, and the major problems are associated with syntactic and domain mismatches between the ontologies that are mapped. Generally, ontology mapping methods employ logic subsumption and inheritance rules for similarity judgement and for semantic association between ontologies (Maedche and Staab 2001a). Alternative approaches have also been proposed (Bowers and Delcambre 2000, Melnik and Decker 2000, Bouquet *et al.* 2003) that consider the complexity in concept semantics and attempt to resolve mismatches and conflicts that can occur due to conceptualisation of inherently vague concepts and variability in linguistic and cognitive interpretations from the real world. A major problem faced in conceptual alignment is caused by variation in knowledge granularity and discourse domains for the concept across the different ontologies. Visser *et al.* (1998) classify this as a problem in '***conceptualisation***' where domain is interpreted in different ways. Such problems are solved by human interpretation and different conceptualisations can be used together by aligning the overlapping parts of the ontology.

Klein (2001) summarises the main issues involved in combining ontologies, among which is the occurrence of 'ontology level mismatches' which correspond to categories described above, such as domain coverage, concept scope, synonyms, homonyms, concept description, paradigm, and encoding. In addition, naming conflicts, as proposed by Bishr (1997) and Goh (1997), arising from semantic ambiguities in the use of homonyms and synonyms for concept description that cause problems in the specification of ontology, can cause problems in alignment and ontology comparison (Visser *et al.* 1998). Use of natural language specifications and the difference in detail in the ontology can cause problems in the extent to which instances, properties and relations are explicated for a concept, causing 'conceptualisation mismatch' (Klein 2001). Noy and Musen (2000) also state that finding terms that need to be (and can be) aligned is difficult.

In conceptual alignment, the syntactic differences are explicit, but difficult to reconcile in ontology comparison (Kitakami *et al.* 1996), especially with the tools and methods that are currently available. Semantic differences, on the other hand, are implicit, subtle, and more difficult to reconcile in the tools and methods without the use of human judgement and interpretation. Klein (2001) and Maedche and Staab (2002), however, point out that problems in conceptual integration and alignment are still under-developed in most of the proposed techniques. In particular, semantic mapping at the modal level (at the level of domain) still relies on human intervention and judgment to identify appropriate semantic measures and associations, and therefore, there is a need for more comprehensive methods for aligning concepts from ontologies.

Various approaches have been proposed in ontology engineering literature for mapping, merging and integrating ontologies, to achieve inter-operability and semantic homogeneity. McGuinness *et al.* (2000) propose the following generic steps be followed in the integration process:

1. find the instances where ontologies overlap;
2. item relate concepts that are semantically close via equivalence and subsumption relations; and
3. item check the consistency, coherency and non-redundancy of the results.

Besides this general framework for mapping information from ontologies, specific methods include a range of top-level to bottom-up approaches for aligning concepts and finding semantic associations between ontologies. Silva and Rocha (2003) propose '***semantic bridges'*** where certain mediator agents are used to define the mapping between ontologies. This method is employed in tools such as KRAFT (Preece *et al.* 1999) and MAFRA (Silva and Rocha 2003). However, Wache *et al.* (2001) argue that such mappings fail to maintain the semantics of the concepts as the user is allowed to propose mappings even if these conflict with the internal semantic arrangement of the ontologies. Heuristics-based methods rely on lexical relationships such as synonym, homonym and hypernym for mapping the similarity between concepts (Klein 2001). Such methods have been employed in tools such as OBSERVER (Mena *et al.* 2000). Other methods such as '***formal ontology methods'*** (Guarino and Welty 2000) rely on inheritance where the different ontologies are linked to a top-level ontology and mappings between the different ontologies are formed by inheriting a common super-class for all lower-level concepts. Such methods adhere closely to formalised frameworks but problems occur when concepts in different ontologies do not overlap or are not terminologically coincidental. Rodriguez and Egenhofer (1999, 2003) present a semantic similarity model for geographic data types where linguistic analysis and contextual variability in semantic heterogeneity is incorporated to assess semantic similarity between entity types.

Several semi-automated tools now exist that use one or more of the mapping methods discussed previously. There is no consensus on how ontologies should be mapped (Noy and Musen 2002) and this creates problems in determining the relative appropriateness of the different methods to an ontology mapping task. Ontology integration and alignment requires a certain amount of human interpretation and judgement in determining appropriate categories and concepts for mapping. Although semantic relations are important in mapping ontologies, most automated processes focus on making the instances correspond and human interpretation is required to ensure that the semantics and concepts correspond with each other (Kokla and Kavouras 2001). The basic, common aim of all the tools is to make the concepts, instances and relations explicit from an ontology. However, since they all differ in the nature of input data and the methods employed for mapping, it is difficult to compare them in a systematic manner. Duineveld *et al.* (2000) can be referred to for a review of commonly employed ontological engineering tools. Noy and Musen (2002) present a set of criteria for a comparison of these tools based on usability, knowledge expressiveness and inter-operability across different representation languages. It is proposed that as part of future work from this paper, an evaluation procedure with our proposed methodology with existing tools will be carried out.

To sum up, this exhaustive review has demonstrated that while previous research has attempted to describe mismatch patterns that may occur between two conceptualisations, the terminological framework is unclear, the descriptions provided are vague at times and there is a lack of formal descriptions of these patterns. In the following sections, we present a brief overview of the formalisation approach that forms the backbone of the work presented in this paper to capture possible mismatches between a user's and an expert's conceptual models.

4. Methodological Framework

Previous related work (Agarwal 2004b, Huang *et al.* 2005) has presented a review of mismatches and shown that discrepancies between conceptualisations are inevitable in domains that are largely unstructured, such as the geographic domain. A standard terminology is often difficult within such domains, as opposed to more structured domains, such as Law and Chemistry, and is largely dependent on the context of the use and user. Indeterminacy and ambiguity in meanings are key issues in the development of ontologies in such domains. Smith (1989) refers to this as the 'Tower of Babel' problem where the heterogeneity in terminology and meanings leads to conceptual and terminological incompatibilities. Commonsense notions and cognitive conceptualisations structure knowledge by determining much of the conceptual and semantic content within the main categories within an ontology representative of the domain. The inherent vagueness encapsulated in the terms and concepts leads to variability in the conceptualisations and causes conflicts and mismatches. Empirical results show that individual conceptualisations are characterised by semantic heterogeneity (Hameed et al. 2001, Agarwal 2004a). It is clear from the review of ontology mapping methods that no single approach is fully able to align ontologies using concepts, and syntactic and semantic differences are difficult to reconcile in semi-automated procedures requiring human intervention and judgement to resolve heterogeneity. It is also noted that a consideration of ontologies as multi-layered structures and resolution of semantic heterogeneity at different layers is more effective for conceptual alignment between ontologies, and a combination of different mapping methods provide better indication for meanings of concepts. The methodology employed in this paper is a hybrid approach that is based primarily on principles of *'semantic coordination'* (Bouquet *et al.* 2003), where instead of assuming generic abstract structures for aligning the different ontologies, an agreement on the meanings of concepts is realised by comparing how different knowledge models map onto each other. Ontology mismatch can be a measure of the level of heterogeneity inherent in the concept (Visser *et al.* 1998). In the methodology adopted here, different levels of semantic knowledge are considered, and semantic relations are made explicit for a comparison of the meanings of concepts in different hierarchical structures. Although 'semantic coordination' is distinguished from 'meaning negotiation' in some AI literature (Magnini *et al.* 2002), these are considered as equivalent terms in the approach adopted here. Section 4.1 details out the formalisation approach for defining the mismatches, along with the basic assumptions and notations. The prototype developed to reason with OWL-based ontologies is introduced in section 4.2. Section 4.3 outlines the domain, the specific ontology considered for the test scenario and the process adopted for capturing individual user conceptualisations and models. The demonstration of the tool developed in JAVA, on the test ontologies captured from the user study, and based on the formalisation approach presented in section 4.1, is summarised in section 4.4.

The methodological approach followed for the work presented in this paper can be outlined as follows:

1. define a formalisation that captures certain phenomena (misconceptions in our case)
2. implement a demonstration program that follows the formalisation; normally the demonstration are in a specific domain and for a specific problem
3. empirically test the demonstrator to verify the algorithms and the formalisation
4. fine-tune both the demonstrator (i.e. some problems might be due to implementation decisions rather than flaws in the formalisation) and the formalisation (i.e. there may be aspects of the phenomena that may have been missed or defined inappropriately).

4.1 Formalisation Approach

Formal approaches allow the design of algorithms at levels higher than the specific applications, and therefore, bring considerable insights into the design of intelligent system. We use Description Logic (DL) to formally define discrepancies between a user's and a system's conceptualisations. The formal descriptions can be followed in algorithms for user modelling in a variety of domains. In our formalisation approach, we define concept as ‘having meaning’ is distinguished from a ‘term’ that is a referent for the concept to the real world and therefore does not necessarily has semantic content. We hold that a concept is associated with four parts: term, definitions, instance and property (role). If two concepts match all of the four parts, then we consider that there is no misconception between them. At this stage of development, we have assumed that all the intentional meaning of a concept is reflected by and only reflected by its term, definitions and the properties. It is accepted that the semantic of the concept and the intentional meanings will also lie in the relations to other neighbouring concepts and this is expected to be taken into account in further development of the reasoning algorithms. One of possible solutions to that exceptional mismatch is to use ‘owl:sameAs’ to explicitly indicate that the two concepts are equal. The definition (or definiens) is expressed by the language of description logic, that is, we treat a concept as a set of individuals. Our approach for misconception identification is to first determine the relationship between two concepts by reasoning with their definitions and then check the term and properties for misconception.

We define five types of relationship between two concepts, namely, **equality**, **equivalence**, **subsumption**, **partial overlap** and **disjointness**. Equality, which indicates two concepts have exactly the same intentional meaning, is a special case of equivalence, which merely indicates two concepts have the same set of individuals. Equivalence is again a special case of subsumption, which shows one concept is a sub-class of another. Partial overlap refers to two concepts sharing part of instances yet not equivalent. Disjointness defines the relationship between two concepts without any common instance. The complete set of misconceptions and their formalisations have been presented in a related work (Huang *et al.* 2005) and can be referred to for further details. Here, we outline few example misconceptions and their definitions that were tested and identified in the user study to follow.

1. Mismatches based on equivalence

Two concept are considered equivalent if they have the same set of individuals, i.e. $C_u \equiv_{def} D$, $C_e \equiv_{def} D \vdash C_u \equiv C_e$, where D can be either atomic concept or combination of other concepts.

Term Mismatch $Term(C_u) \neq Term(C_e) \rightarrow$ Term Mismatch

Two concepts have the same sets of individuals; however, the concepts may have different intentional meaning. There are many examples in the space ontology. For instance, *edge* is equivalent to *boundary*, yet these two concepts have different intentional meanings in their own rights.

Attribute Mismatch $C_u \equiv_{def} D \cap \forall R \bullet E$, $C_e \equiv_{def} F \cap G \vdash R_u \equiv R_e$, where D, F, G can be any concepts.

$F \subseteq \forall R \bullet E$, $Term(C_u) = Term(C_e) \rightarrow$ Attribute Mismatch

This is so-called attribute assignment mismatch (Visser *et al.* 1997), which is a property misconception occurring when two properties are the same except the domains, with one being a subset of another. For instance, a user assigns to *AdministrativeRegion* the attribute of *haslocation*, which could be assigned to *Region* in the expert ontology.

Abstraction Mismatch $C_u \equiv_{def} (D_1 \cup D_2 \cup \ldots \cup D_m)$, $C_e \equiv_{def} (D_1 \cup D_2 \cup \ldots \cup D_m) \vdash C_u \equiv C_e$, where D_i can be either atomic concept or combination of other concepts

C_u does not exists $\rightarrow$ Abstraction Mismatch

This mismatch occurs when user has a concept whose abstraction does not exist in expert ontology. For example, *Coordinate* usually include *HorizontalCoordinate* and

VerticalCoordinate, but *Coordinate* could be missing in the user's conceptualisation, that is, the user is not aware that both of coordinates form the whole coordinate for a location.

2. Misconceptions based on subsumption
Subsumption shows that one concept is a sub-class of another.

Structure Mismatch $C_u \equiv_{def} (D_1 \cap D_2 \cap \ldots \cap D_m \cap \forall R_1 \bullet F_1 \cap \forall R_2 \bullet$ © $\cap \ldots \cap \forall R_k \bullet F_k)$, $C_e \equiv_{def} (E_2 \cap \ldots \cap E_n \cap \forall S_1 \bullet G_1 \cap \forall S_2 \bullet G_1 \cap \ldots \cap \forall S_1 \bullet G_l) \vdash C_u \subseteq C_e$, where D_i, E_i can be either atomic concept or union of other concepts

$\forall$ i, $1 \leq i \leq m$, $\exists$ j, $1 \leq j \leq q$ n, $D_i \rightarrow E_j$, and
$\forall$ i, $1 \leq i \leq k$, $\exists$ j, $1 \leq j \leq l$, $R_j = S_j$, $F_j \rightarrow G_j$

$Term(C_u) = Term(C_e) \rightarrow$ Structure Mismatch

The description is similar to a subsumption problem. The only difference is the last condition, which indicates concept subsumption with structure mismatch or Definiens Mismatch. For instance, the user may define *Top* as *maximalheight*(*Top* $\equiv_{def} \forall$ *hasHeight* $\bullet$ *Maximum*) whereas expert ontology defines *Top* as *Maximum* with *updirection*(*Top* $\equiv_{def}$ *Maximum* $\cap \forall$ *hasDirection* $\bullet$ *Up*).

4.2 A prototype for discovering OWL-based mismatch patterns

Based on the formal descriptions of mismatches, we have implemented algorithms to capture a user's misconceptions defined as the discrepancies between the user's and the expert's perspective of the world. Because the misconception patterns were defined in Description Logic, they could easily be applied to conceptualisations defined in OWL. It must be noted that although OWL allows the semantics of geospatial data to be explicitly defined using ontologies, it is still limited in being able to provide direct support for representing the semantics of the procedures for processing geospatial data (Chen *et al.* 2004). Nevertheless, OWL is a common proposed standard for ontology creation and development for the semantic web and therefore we use it here as a standardised reasoning mechanism for developing our algorithms.

Ontology Web Language (OWL) is a further modification on the DAML+OIL language (Connolly *et al.* 2001) with an even richer representational framework than that of XML or RDF, and developed specifically for ontology specification. Syntactically, OWL is a vocabulary extension of RDF, but specifies semantic content on top of the RDF graph. Semantic restraints are imposed in OWL on terms and concepts. This means that the meanings are included, along with relationships and objects, and a richer set of specifications such as disjointness, cardinality, equality, symmetry and properties can be stated in an OWL specification (McGuinness and van Harmelen 2003). OWL-DL supports the OWL representation in description logic specification for use with reasoning systems, and OWL-lite is a sublanguage of OWL-DL but with relatively simple features targeted at novices and light application ontologies. In OWL, 'classes' define concepts from the domain, 'individuals' represent specific instances and 'properties' define the values that each individual can take (Martin *et al.* 2003). The five types of relationship are defied in OWL, as follows:

owl:same	**equality**
owl:eq	**equivalence**
owl:subsume	**subsumption**
owl:po	**partial overlap**
owl:dj	**disjointness**

For the implementation of the algorithms, we have used rule-based OWL inference engine Jena2 (Carroll *et al.* 2004). Jena can be downloaded from http://jena.sourceforge.net, and a comprehensive documentation is available there. Jena is a semantic Web toolkit. As presented in the Jena Application API, it has three major functionalities: RDF manipulation, RDQL implementation and ontology reasoning. Ontology reasoning was the latest feature of Jena and was assumed to be one of its characteristics, for the API, for reasoning can work on OWL-based files, which have been set as the standard ontology language by W3C.

Following the triple-based nature of Jena, we have defined a set of rules to capture the mismatches defined in the formalisation. For example, to check for *Structure Mismatch* for domain based on *subsumption* relation between two concepts, the following rule will be passed to Jena:

```
[rule3: (?C rdfs:domain ?D), (?D rdfs:subClassOf
?E),  (?E rdfs:subClassOf ?F)-> (?C rdfs:domain
?F) ]");
```

The rule states that if an object property C has a Class D as its domain and D is a subclass of E, which is a subclass of F, then C takes class F as another domain. For instance, in the space ontology, the object property *hasCapital* takes *City* as its domain and *City* is a subclass of *AdministrativeRegion*, which is a subclass of *Region*. Thus, we can deduce that *hasCaptial* also takes *Region* as its domain, which corresponds to the user’s perspective.

4.3 Case Ontology and Domain

There is a lack of comprehensive geo-ontologies for the Semantic Web. This is partly because of the vagueness and ambiguity in the geographic domain and because many of geographic concepts and terminology are anchored in human cognition. This means that there might be different individual understandings for the terms and concepts employed in an ontology. The SWEET (Semantic Web for Earth and Environmental Technology) suite of ontologies are constructed in OWL to provide a basic integrated framework and upper level ontology as a basis for a common semantic framework for the GeoSciences. The different ontologies within this suite are: EARTHREALM, Numerics, PHENOMENA (any transient features), Physical Properties, SPACE, Physical Substances, Time, Units and Dataset properties.

The web interface created for SWEET (http://sweet.jpl.nasa.gov) is aimed at supporting user intervention and is based on the assertion that a comprehensive ontology should include collaborative capabilities and community participation, thus allowing the users to update the terms and concepts and including their own conceptualisations in the existing knowledge base. These are hierarchical ontologies, for example, 'hydrosphere' is a parent concept for 'surface water' which is a parent for 'river' which is a parent for 'Mississippi River'. The Global Change Master Directory (GCMD) was used along with keywords from the Earth Science Modelling Framework (ESMF) to populate the ontologies. The SPACE ontology contains the maximum relevant concepts for spatial divisions and locations in the geographic context, and includes terminology specific to the spatial domain, focussing on spatial extents, such as country, equator, boundary, and relations such as has capital, has location, top of, north of etc. It works in coordination with the numerics ontology where the spatial extents and relations are modelled on a numerical scale. Table 1 shows an example of how a concept such as '*Region'* is conceptualised in the SPACE ontology. Figure 1 shows a slice of SPACE ontology.

Table 1. Conceptualisation of 'Region' in SPACE ontology

space#region	physical, material
< supertype space#NumericGeometricObject_2D; space#SpatialObject	
>subtype space#AdministrativeRegion	
attributes (direct and inherited) space#inside	
disjoint space#NumericGeometricObject_3D	

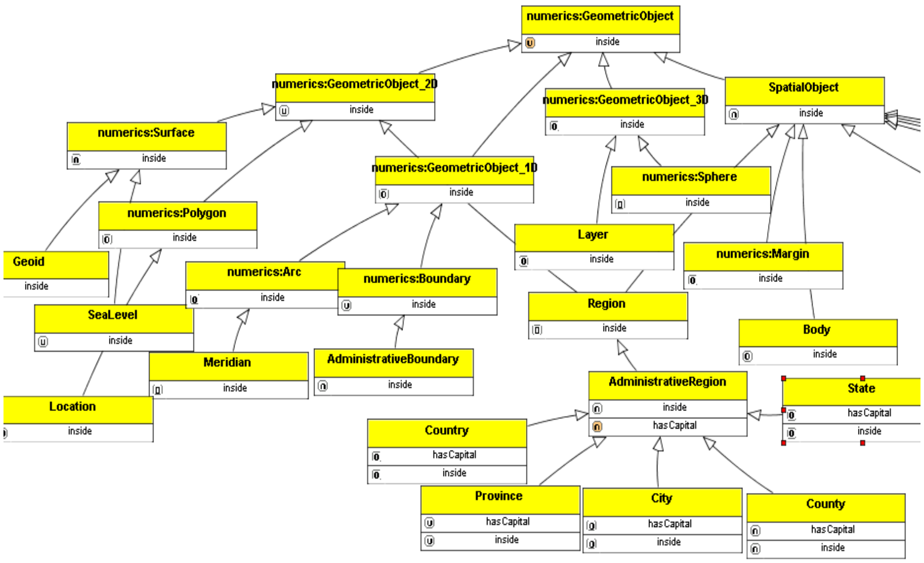

Fig. 1. Slice of SPACE ontology showing classes, subclasses and relations for REGION (created in Protégé using ezOWL)

4.4 Capturing User Semantics and Conceptualisations

The availability of geo-spatial knowledge resources on the web enables members of the public to take advantage of trusted knowledge built by domain experts, e.g. for planning travel routes and for accessing weather information. Users access the web services with different goals, often; these services require integration of the various different resources to provide a comprehensive result for the user search for their specific requirements. For example, in a 'what is in my backyard' service provided by the Environment Agency (EA), members of the general public can see what pollutants may be scattered across their neighbourhood. End-users will have their own contexts of use: property evaluation, ecology, etc. and for a member of the public, a general interest (based on a topographic view of different areas in the city). Each could potentially view the data provided by the others but form their own conceptual understanding of the location-based information. Automating the mapping of multiple conceptualisations and personalisation of web-based services will also facilitate pervasive computing in mobile services and enable more effective use of mobile GIS services.

The overall purpose of the experimental work is to conduct more appropriate and comprehensive evaluation of the algorithms and ontology mapping procedures (see Sampson 2005) based on real data of multiple user conceptualisations. Because the focus is on the geographic domain where many concepts and categories have a large cognitive anchoring, and the semantic content is, therefore formed by individual conceptualisations, it is feasible to conduct studies where users (naive domain users) are asked to identify the main concepts and their relations, on the basis of which user conceptual models can be extracted.

4.4.1 Experimental Framework

Previous related work (Agarwal 2004a) has shown ways in which user perspectives, especially for semantic content of the domain, can be extracted through mapping individual conceptualisations. For the work presented in this paper, the primary aim was to extract user models and semantic conceptualisations for the concepts present in the 'expert' ontology that is aimed as a knowledge resource forming the basis for a comprehensive geo-ontology for a web-based user interface. In this test case, SPACE ontology from the SWEET suite is used. The experiments at this stage were manually implemented and the questionnaire was designed after identifying a set of concepts and relations from the ontology forming its semantic content. For this, SWOOP (http://www.mindswap.org/2004/SWOOP/) was used as an exploration tool. The concepts were identified based on previous related work (Galton 2001, Agarwal 2004a) that have shown the inherent ambiguity in several of these concepts, such as region, boundary, and location due to the commonsense reasoning involved in forming meanings of these concepts and anchoring them to the real world objects. These concepts were selected because of their links to the real world as well as to the human commonsense reasoning and therefore these form valid focus within this test case. Also, these concepts were identified to be commonly used to extract web-based geographic information, for example, ‘show all the hospitals in my region’ or ‘show the boundary of the most economically developed area’, or the ‘boundary for the flood prone area in my region’, ‘show information about pollution level near the location of a specific industry’.

A part of the questionnaire that was sent to the end users is shown in Table 1. The questionnaire included a list of the relevant concepts without making any inherent hierarchy or relations apparent. The spatial concepts and the respective sub and super classes included in the questionnaire are ***Region, Zone, Spatial Object, Administrative Region, Geometrical Object, Boundary, Edge, Administrative Boundary, Country, State, City, Political Division, Position, Location, distance.*** It has to be pointed out that the design of the questionnaire was in no way aimed at capturing a complete conceptual model of the end-users for the geographic domain. It was instead aimed at (and enabled) capturing of partial conceptual models of the users, and was focussed solely on the concepts that were delineated for the purpose of the study.

Along with this, detailed definitions along with examples were provided for subclass, superclass, property and synonym to minimise any individual biases in interpretation of these terms. **Synonym** of a term ‘t’ was defined as ‘Similar in meaning to ‘t’, such as *table is synonym with desk*’. **Subclass** of a term ‘t’ is defined to hold ‘when a term is a child term of ‘t’, such as a *coffee table is a subclass of a table*’. **Superclass** of term ‘t’ exist ‘when a term is a parent of term ‘t’, such as *furniture is a superclass of table*’, and **Property** of term ‘t’ is stated to be ‘when a term is a characteristic of term ‘t’, such as *‘has legs’ is a property of table*’.

The questionnaire were sent to a wide range of end-users, from a cross section of disciplinary backgrounds, including Geography, Information management, Linguistics and Computer Science. Although the user responses were treated as anonymous, some personal information was also requested. This included previous experiences in using web-based services for geographic information, the websites that were used and examples of problems that were faced in using web services for geographic resources and information. Most of the users admitted to having used the web for weather-related and travel information for specific regions, as well as using the web for downloading demographic data, services and environmental information for their neighbourhoods and localities.

Clear instructions were provided on how to complete the questionnaire as well as detailed explanations given. The questionnaire sample is provided in Appendix A. There has been a vast amount of work previously done in the GIScience field to capture individual conceptualisation and mental models. As this was a test scenario, detailed control experimental settings were not practical. In addition, the primary focus in this study was on testing the methodology for alignment of diverse models, and therefore email communication was used. The respondents were asked to work independently and it was expected that the simple, self-explanatory design of the questionnaire enabled the respondents to express their internal semantic conceptualisation and understanding of the different concepts. The responses were compiled following their receipt by email.

5. Results from the user studies

The questionnaires resulted in a list of user conceptualisations of semantics, relations and properties for the concepts of region, boundary and location from the concepts extracted from the SPACE ontology, as well as individual semantic notions listing concepts and variations not included in the SPACE ontology. The user conceptualisations were aligned manually for the different conceptualisations of the concepts employed in the test questionnaires. The aim is to automate this procedure and future work will strive in this direction. Here, the conceptualisations of ‘Region’ are used for demonstration purposes. The subclasses, superclasses, properties and synonyms were listed for individual concepts. Example user ontology is constructed in OWL based on the aligned model resulting from the summary of the questionnaires. A sample of this ontology is shown below in Appendix B.

The user ontology constructed from the aligned model in OWL, expert ontology (in this case, SPACE ontology) and a user concept is taken as an input in the automated tool for ontology alignment. The program will first collect all the information related to the user’s central concept (*Region* in our experiment). These include the subclasses, super classes, synonyms and properties (with range). Secondly, it will search for the corresponding concepts and properties in the expert ontology, based on the rules of pattern matching and formalisations of misconceptions, examples of which are provided in section 4.1 and in further detail in a related publication (Huang *et al.* 2005). Finally, the user's perspectives on *Region* will be mapped to the corresponding concepts and properties in expert ontology. The perspectives that have no correspondents are reported as misconceptions or mismatch. Figures 2 and 3 show the interface of the automated tool, showing pull-down menus for concepts in user and expert ontology, an initialisation button, and a window that shows the misconceptions after aligning, mapping and comparing the two input ontologies.

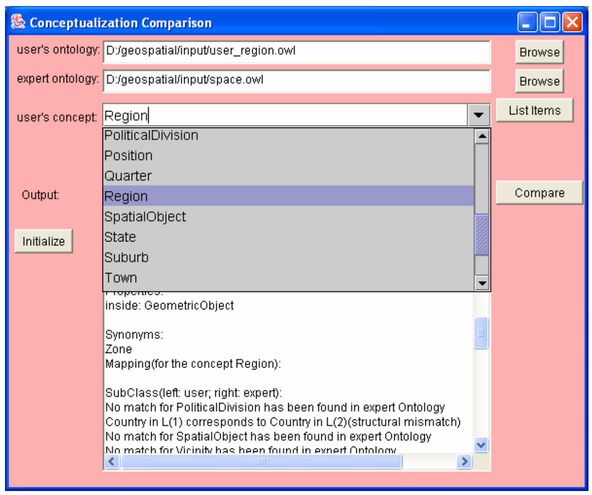

Fig. 2. The *Conceptualisation Comaprison* Interface showing the pull-down menus for selecting relevant concepts from the ontologies

The approach adopted in this test case to demonstrate the effectiveness of the automated prototype for comparing individual semantic conceptualisations is based on a comparison of the central concept from the user ontology with all the concepts in the expert ontology. This differs from the approach that we adopted in a related work (Huang *et al.* 2005) where artificially constructed ontologies were used to test the tool and therefore, concepts, both from the user as well as the expert ontology, were specified for comparison to limit the computational complexity. In this test case, the user ontology, as shown in Appendix B, was developed from user studies, focussed on semantic conceptualisation for a single geographic concept, and the comparison was based on a single central concept 'Region'. Unlike other real-world ontologies, which usually consist of a large number of inter-related concepts, this user ontology, as a demonstration, was limited to one concept and the relevant semantic relations and properties, as extracted from user questionnaires. Therefore, in this case, the program compared the complete user ontology with the expert ontology, and looked for all relevant concepts. However, as the scale of the user ontology increases with further concepts, relations and semantic properties, the computing complexity will increase significantly, because the program will compare every concept and property related to each concept in the user ontology against a the total number of concepts and properties in the expert ontology.

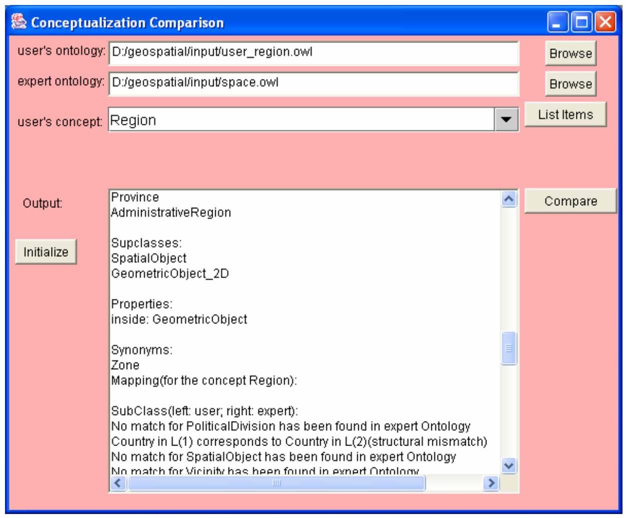

Fig. 3. The *Conceptualisation Comparison* Interface allowing the visualization of mismatches between the input ontologies based on DL algorithms and OWL-based reasoning

Below is the result from the automated mapping, employing OWL based reasoning, using the *Conceptualisation Comparison* tool described above, developed based on the DL algorithms, between the aligned user model and the expert ontology (SPACE ontology) for the concept 'Region'.

SubClass(left: user; right: expert):
No match for PoliticalDivision has been found in expert Ontology
Country in L(1) corresponds to Country in L(2)(structural mismatch)
No match for SpatialObject has been found in expert Ontology
No match for Vicinity has been found in expert Ontology
No match for Village has been found in expert Ontology
City in L(1) corresponds to City in L(2)(structural mismatch)
No match for Zone has been found in expert Ontology
No match for Area has been found in expert Ontology
No match for Quarter has been found in expert Ontology
No match for Ward has been found in expert Ontology

No match for Borough has been found in expert Ontology
No match for Town has been found in expert Ontology
State in L(1) corresponds to State in L(2)(structural mismatch)
No match for EDU has been found in expert Ontology
No match for Patch has been found in expert Ontology
No match for GeometricalObject has been found in expert Ontology
County in L(1) corresponds to County in L(2)(structural mismatch)
No match for Suburb has been found in expert Ontology

SuperClass(left: user; right: expert):
No match for Continent has been found in expert Ontology
No match for State has been found in expert Ontology
SpatialObject match SpatialObject
No match for GeometricalObject has been found in expert Ontology
No match for AdministrativeRegion has been found in expert Ontology
No match for City has been found in expert Ontology
No match for Country has been found in expert Ontology

Synonyms(left: user; right: expert):
No match for District has been found in expert Ontology
No match for Region has been found in expert Ontology
No match for Position has been found in expert Ontology
No match for Country has been found in expert Ontology
No match for State has been found in expert Ontology
Zone match Zone
No match for PoliticalDivision has been found in expert Ontology
No match for Area has been found in expert Ontology

Properties:
No match for hasPosition has been found in expert Ontology
No match for hasContext has been found in expert Ontology
No match for hasArea has been found in expert Ontology
isPartOf(City) corresponds to City in L(2) in expert Ontology
No match for hasEdge has been found in expert Ontology
No match for hasBoundary has been found in expert Ontology
No match for hasCommonCharacteristics has been found in expert Ontology
No match for hasLocation has been found in expert Ontology
isA(SpatialObject) corresponds to Spatial Object as Superclass in expert Ontology
No match for hasDistance has been found in expert Ontology
hasCapital corresponds to hasCaptial(City) in expert Ontology
No match for hasClimate has been found in expert Ontology

The primary misconceptions that have been identified by the program consist of **Abstraction Mismatch, Structural Mismatch** and **Attribute Mismatch** (see section 4.1 for formalizations). However, some limitations were also noted in the demonstration tool. The program fails to report all misconceptions which were apparent on manual inspection. For example, in the user's OWL file, 'AdministrativeRegion' is both subclass and super class of 'Region', which can be interpreted in two ways: 1. 'AdministrativeRegion' is equivalent to 'Region' and 2. there exist conflicting relationships on the concept. The program, which has discovered that 'AdministrativeRegion' is a subclass of 'Region', however, has no means to solve confusion and ambiguity from user's perspective. For the latter reason, the user's ontology has some properties such as *isPartOf(City)*, which can be arguably translated as a subclass of *City*, which is actually a subclass of *'Region'* in expert ontology. This is partly because of the inherent organisation of the user's OWL file and partly because the program at this stage lacks robust mechanisms to also handle the semantic meaning of properties along with its capacity to reason with semantic meanings of concepts. So, although the innovativeness of the formalization approach allows us to make the semantic misconceptions between concepts apparent, the tool needs further development to enable identification of semantic mismatches also at the property level. The systematic methodological approach has facilitated the evaluation of the demonstration tool and identify areas where it needs fine-tuning to make the formalizations more worthwhile in identification of semantic mismatched between individual user conceptualizations in the domain.

6. Conclusions and Future Work

The work presented in this paper is part of our ongoing research on applying commonsense reasoning to elicit and maintain models that represent users' conceptualisations of the real world. Such user models will enable taking into account the users' perspective of the world and will empower personalisation algorithms for the Semantic Web. A formal approach for detecting mismatches between a user's and an expert's conceptual model is outlined. The formalisation is used as the basis to develop algorithms to compare two conceptualisations defined in OWL. The algorithms are illustrated in a geographical domain using a geo-ontology in OWL developed as part of the SWEET initiative for the Semantic Web by NASA, and have been tested by using test cases of possible user misconceptions.

A number of possible benefits that the above approach can afford to the development of a geospatial semantic web is foreseen. The approach defined in this paper is an effort to allow the system to reconcile the user conceptual model with the core ontology and therefore identify the discrepancies and similarities, and thereby allowing the system to identify the differences in the user conceptualisations with the so-called expert ontology. This will allow, first, for the development of systems that allow personalisation by incorporating user models and diversity; second, this approach can be used to test core ontologies developed as the basis for a geo-spatial system/service against user conceptualisations for discrepancies. This will be useful in evaluating the reliability of ontologies for standardisation and re-usability. Moreover, the personalisation approach allows the possibility of using the user models to enrich the available information resources with shared semantics instead of relying on *'fixed'* ontologies available to the developers at the design stage.

The Semantic Web paradigm requires the deployment of appropriate user modelling approaches that capture and maintain different user perspectives. At this stage, the identification of suitable concepts from the core ontology and the capturing of the user conceptualisations, as well as development of user ontologies from the results is manual and requires human intervention. It is proposed that this process be deployed on a web-based service and be largely automated. Future work will develop on this work to develop automated web-based interfaces that can use the different semantics, detect the semantic mismatches and process the information available to integrate the knowledge resources based on individual conceptualisations of the domain. The long-term goal of our research is to apply commonsense reasoning approaches to capture and maintain users' conceptual models and to use these models for personalised, semantically-enhanced search on the web. For this, we consider that the domain expertise is encoded in some ontology (or several ontologies) pre-defined by domain experts and knowledge engineers. This expertise is used to guide the intelligent behaviour of the system and is combined with some model of the user that corresponds to the user's conceptualisation of the domain. Future work is also envisaged in carrying out more robust user modelling exercises to capture individual conceptualisations in the geographic domain. These methods will be used along with the algorithms and formal approaches described in this paper to test and fine-tune the algorithms.

Future work will also include incorporation of uncertainties in user models and semantic conceptualisations, and target more complex mappings and mismatches. We are also looking into the possibilities of using the mismatch detection algorithms in combination with additional reasoning to deal with vagueness and heterogeneity problems. The aim is also to explore the possibility of including other ontology language, standards and reasoning methods (for example, SWRL, RDF and XML) within these algorithms. For this purpose, transferability between different web languages will be carried out for geographic concepts. It is envisaged that work in this direction will also enable a comprehensive evaluation and comparison of ontology mapping tools that are currently available and assess their applicability for handling geospatial semantics.

References

Agarwal, P.,2004a, Contested nature of 'place': knowledge mapping for resolving ontological distinctions between geographical concepts, In Egenhofer, M., Freksa, C., Harvey, M. (Eds.) Third International Conference, GIScience 2004, Lecture Notes in Computer Science, Springer-Verlag, Berlin, LNCS 3234, pp. 1-21.

Agarwal, P., 2004b, RO-GEO-PLACE, PhD Thesis, University of Nottingham, 418 pp.

Agarwal, P., 2005, Ontological considerations in GIScience, International Journal of Geographical Information Science (IJGIS), Vol. 19, No. 5, pp. 501 - 535.

Berners-Lee, T., Hendler, J., and Lassila, O., 2001, The Semantic Web, *The Scientific American* 284: 34-43

Bishr, Y. A., 1997, Semantic Aspects of Interoperable GIS, ITC Publication No. 56, Enschede, NL.

Bouquet, P., Serafini, L. and Zanobini, S., 2003, Semantic coordination: a new approach and an application, Proceedings of 2nd International Semantic Web Conference (ISWC2003), Sanibel Island, Florida, 20-23 October, 2003.

Bowers, S. and Delcambre, L., 2000, Representing and transforming model-based information, In First Workshop on the Semantic Web at the Fourth European Conference on Digital Libraries, Lisbon, Portugal, September 18–20, 2000.

Carroll, J. J., Ian Dickinson, Dollin, C., Reynolds, D., Seaborne, A., Wilkinson, K., (2004) Jena: implementing the semantic web recommendations. In Proceedings of the 13th international World Wide Web conference, pages 74–83. ACM Press.

Doan, A., Madhavan, J., Domingos, P., Halvey, A., 2002, *Learning to map between ontologies on the semantic web*. In Proceedings of WWW'02, pages 662—673.

Dolog, P., Henze, N., Nejdl, W., and Sintek, M., 2003, Towards the Adaptive Semantic Web. 1st Workshop on Principles and Practice of Semantic Web Reasoning (PPSWR'03), co-located with International Conference on Logic Programming (ICLP'03), Dec. 8, 2003, Mumbai, India.

Duineveld, A. J., Stoter, R., Weiden, M. R., Kenepa, B. and Benjamins, V. R., 2000, Wondertools? a comparative study of ontological engineering tools. International Journal of Human-Computer Studies, 52(6): 1111-1133.

Egenhofer, M.J. (2002) Toward the Semantic Geospatial Web. In *Proceedings of the Tenth ACM International Symposium on Advances in Geographic Information Systems,* McLean, Virginia.

Fonseca, F., and Sheth, A., 2003. *The Geospatial Semantic Web,* UCGIS Short-

Goh, C. H., 1997, Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Sources. Doctoral Dissertation, MIT, MA.

Galton, A., 2001, On the Ontological Status of Geographical Boundaries, Meeting on Funadamental Questions in GIScience, Manchester.

Guarino, N. and Welty, C., 2000, A Formal Ontology of Properties, Proceedings of the ECAI-00 Workshop on Applications of Ontologies and Problem Solving Methods, Berlin, Germany, pp. 121-128.

Hakimpour, F. and Timpf, S., 2001, Using Ontologies for Resolution of Semantic Heterogeneity in GIS, International Conference on AGILE, April 19-21, 2001, Brno.

Hameed, A., Sleeman, D. and Preece, A., 2001, Detecting Mismatches among Experts' Ontologies Acquired through Knowledge Elicitation, R&D in Intelligent Systems XVIII, Proceedings of ES2001:the 21st SGES International Conference on Knowledge BasedSystems and Applied Artificial Intelligence. Springer-Verlag, London, Cambridge, U.K., pp. 9–24.

Henze, N., Herrlich, M., 2004, The Personal Reader: A Framework for Enabling Personalization Services on the Semantic Web. Proceedings of the Twelfth GI-Workshop on Adaptation and User Modeling in interactive Systems (ABIS 04), October 2004, Berlin, Germany.

Huang, Y, Dimitrova, V. Agarwal, P., 2005, Detecting mismatches between a user's and a system's conceptualizations to empower user modelling for the Semantic Web, PersWeb 05 workshop, User Modelling Conference 28-29 July 2005, Edinburgh, in press.

Kitakami, H., Mori, Y. and Arikawa, M., 1996, An intelligent system for integrating autonomous nomenclature databases in semantic heterogeneity, In Database and Expert System Applications, DEXA'96. Lecture Notes in Computer Science, Zurich, Switzerland, pp. 187–196.

Klein, M., 2001, Combining and relating ontologies: an analysis of problems and solutions, IJCAI-2001 Workshop on Ontologies and Information Sharing, Seattle, WA, pp. 53–62.

Klein, M., 2001, Combining and relating ontologies: an analysis of problems and solutions, IJCAI-2001 Workshop on Ontologies and Information Sharing, Seattle, WA, pp. 53–62.

Kokla, M. and Kavouras, M., 2001, Fusion of top-level and geographic domain ontologies based on context formation and complementarity. International Journal of Geographic Information Science, 15(7): 679-687.

Li Ding, Tim Finin, Anupam Joshi (2005) *Analyzing Social Networks on the Semantic Web* , IEEE Intelligent Systems, vol. 1 (9) .

Maedche, A. and Staab, S., 2002, Measuring Similarities between Ontologies. In: A. Gómez-Pérez and V. R. Benjamins (Editors), Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web: 13th International Conference, EKAW 2002. Lecture Notes in Computer Science, Springer Verlag Heidelberg, Siguenza, Spain, October 1-4, 2002, pp. 251-263.

Magnini, B., Serafini, L. and Speranza, M., 2002, Using NLP Techniques for Meaning Negotiation, Proceedings of VIII Convegno AI*IA, Siena, Italy, 11-13 September 2002.

Martin, D., Burstein, M., Denker, G., Hobbs, J., Kagal, L., Ora Lassila, McDermott, D., McIlraith, S., Paolucci, M., Parsia, B., Payne, T., Sabou, M., Sirin, E., Solanki, M., Srinivasan, N. and Sycara, K., 2003, DAML-S (and OWL-S) 0.9 draft release. http://www.daml.org/services/daml-s/0.9/.

McGuinness, D. L. and vanHarmelen, F., 2003, OWL Web Ontology Language Overview: W3C Candidate Recommendation 18 August 2003, http://www.w3.org/TR/owl-features/./ , 2003.

McGuinness, D. L., Fikes, R., Rice, J. and Wilder, S., 2000, An environment for merging and testing large ontologies. In: A. G. Cohn, F. Giunchiglia and B. Selman (Editors), KR2000: Principles of Knowledge Representation and Reasoning. Morgan Kaufmann, San Francisco, pp. 483–493.

Melnik, S. and Decker, S., 2000, A layered approach to information modeling and interoperability on the web, In Electronic proceedings of the ECDL 2000 Workshop on the Semantic Web, Lisbon, Portugal, September 21 2000.

Mena, E., Illarramendi, A., Kashyap, V. and Sheth, A. P., 2000, OBSERVER: An Approach for Query Processing in Global Information Systems Based on Interoperation Across Pre-Existing Ontologies. Distributed and Parallel Databases, 8(2): 223-271.

Noy, N. F. and Musen, M. A., 2002, Evaluating Ontology-Mapping Tools: Requirements and Experience, Proceedings of the Workshop on Evaluation of Ontology Tools at EKAW'02 (EON2002). Available as SMI technical report SMI-2002-0936, Siguenza, Spain, October 2002.

Preece, A. D., Hui, K.-Y., Gray, W. A., Marti, P., Bench-Capon, T. J. M., Jones, D. M. and Cui, Z., 1999, The KRAFT Architecture for Knowledge Fusion and Transformation, 19th SGES International Conference on Knowledge-based Sytesm and Applied Artificial Intelligence (ES'99). Springer, Berlin.

Rodriguez, A. and Egenhofer, M. J., 1999, Assessing Semantic Similarities Among Geospatial Feature Class Definitions, International Conference on Interoperating Geographic Information Systems. Berlin, Springer. Lecture Notes in Computer Science 1580, Zurich, pp. 189-202.

Rodríguez, A. and Egenhofer, M., 2003, Determining Semantic Similarity Among Entity Classes from Different Ontologies. IEEE Transactions on Knowledge and Data Engineering, 12(2): 442-456.

Silva, N. and Rocha, J., 2003, MAFRA – An Ontology MApping FRAmework for the Semantic Web, Proceedings of the 6th International Conference on Business Information Systems, UCCS, Colorado Springs (CO), USA, May 2003.

Silva, N. and Rocha, J., 2003, MAFRA – An Ontology MApping FRAmework for the Semantic Web, Proceedings of the 6th International Conference on Business Information Systems, UCCS, Colorado Springs (CO), USA, May 2003.

Smith, B., 1989, Logic and Formal Ontology. In: J. N. Mohanty and W. McKenna (Editors), Husserl's Phenomenology: A Textbook. University Press of America, Lanham, pp. 29-67.

Smith, B., 2003, Ontology: Philosophical and Computational. In: L. Floridi (Editor), The Blackwell Guide to the Philosophy of Computing and Information. Blackwell, Oxford.
Term Research Priority White Paper, Leesburg, VA, USA, University Consortium for Geographic Information Science, www.ucgis.org , Research Priorities. pp.2.
Uschold, M., Healy, M., Williamson, K., Clark, P. and Woods, S., 1998, Ontology reuse and application. In: N. Guarino (Editor), Formal Ontology in Information Systems (FOIS'98). IOS Press, Amsterdam, Treno, Italy, June 6-8, 1998.
Visser, P. R. S., Jones, D. M., Bench-Capon, T. J. M. and Shave, M. J. R., 1998, Assessing Heterogeneity by Classifying Ontology Mismatches. In: N. Guarino (Editor), Proceedings of Formal Ontology in Information Systems, FOIS' 98. IOS Press, Amsterdam, The Netherlands, Trento, Italy, pp. 148-162.
Wache, H., Vögele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H. and Hübner, S., 2001, Ontology-based Integration of Information - A Survey of Existing Approaches, Proceedings of IJCAI-01 Workshop: Ontologies and Information Sharing, Seattle, WA, pp. 108-117.

Appendix A: Questionnaire

Terms	Define using one sentence what this term means to you	Name similar terms from the list above and write the corresponding relation (list number) with the term from the list given here: 1. Synonym 2. Subclass 3. Superclass 4. Property 5. Other (pls. specify)	Name any other similar terms that might not be present in the list above and write the corresponding relation (list number) with the term from the list given here: 1. Synonym 2. Subclass 3. Superclass 4. Property 5. Other (pls. specify)
Region			
Boundary			
Spatial Object			
Location			

APPENDIX B: Aligned User Ontology for 'Region'


```
<?xml version="1.0"?>
<rdf:RDF
    xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns="http://www.owl-ontologies.com/unnamed.owl#"
  xml:base="http://www.owl-ontologies.com/unnamed.owl">
  <owl:Ontology rdf:about="">
    <owl:imports rdf:resource="http://protege.stanford.edu/plugins/owl/protege"/>
  </owl:Ontology>
  <owl:Class rdf:ID="Suburb">
    <rdfs:subClassOf>
      <owl:Class rdf:ID="Region"/>
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="District">
    <owl:equivalentClass>
      <owl:Class rdf:about="#Region"/>
    </owl:equivalentClass>
  </owl:Class>
  <owl:Class rdf:ID="Position">
    <owl:equivalentClass>
      <owl:Class rdf:about="#Region"/>
    </owl:equivalentClass>
  </owl:Class>
  <owl:Class rdf:ID="County">
    <rdfs:subClassOf>
      <owl:Class rdf:about="#Region"/>
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="CountrySub">
    <rdfs:subClassOf>
      <owl:Class rdf:about="#Region"/>
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="Area">
    <rdfs:subClassOf>
      <owl:Class rdf:about="#Region"/>
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="Town">
    <rdfs:subClassOf>
      <owl:Class rdf:about="#Region"/>
    </rdfs:subClassOf>
```

```
    </owl:Class>
    <owl:Class rdf:ID="AdministrativeRegion"/>
    <owl:Class rdf:ID="Village">
      <rdfs:subClassOf>
        <owl:Class rdf:about="#Region"/>
      </rdfs:subClassOf>
    </owl:Class>
    <owl:Class rdf:ID="Borough">
      <rdfs:subClassOf>
        <owl:Class rdf:about="#Region"/>
      </rdfs:subClassOf>
    </owl:Class>
    <owl:Class rdf:ID="Continent"/>
    <owl:Class>
      <owl:unionOf          rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-
ns#nil"/>
    </owl:Class>
    <owl:Class rdf:ID="ZoneEq">
      <owl:equivalentClass>
        <owl:Class rdf:about="#Region"/>
      </owl:equivalentClass>
    </owl:Class>
    <owl:Class rdf:ID="Vicinity">
      <rdfs:subClassOf>
        <owl:Class rdf:about="#Region"/>
      </rdfs:subClassOf>
    </owl:Class>
    <owl:Class rdf:ID="CountryEq">
      <owl:equivalentClass>
        <owl:Class rdf:about="#Region"/>
      </owl:equivalentClass>
    </owl:Class>
    <owl:Class rdf:ID="PoliticalDivision">
      <rdfs:subClassOf>
        <owl:Class rdf:about="#Region"/>
      </rdfs:subClassOf>
    </owl:Class>
    <owl:Class rdf:ID="GeometricalObject"/>
    <owl:Class rdf:ID="Patch">
      <rdfs:subClassOf>
        <owl:Class rdf:about="#Region"/>
      </rdfs:subClassOf>
    </owl:Class>
    <owl:Class rdf:ID="Ward">
      <rdfs:subClassOf>
        <owl:Class rdf:about="#Region"/>
```

```
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Country"/>
<owl:Class rdf:ID="GeometricalObjectSub">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Region"/>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="AreaEq">
  <owl:equivalentClass>
    <owl:Class rdf:about="#Region"/>
  </owl:equivalentClass>
</owl:Class>
<owl:Class rdf:ID="SpatialObjectSub">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Region"/>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="CitySub">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Region"/>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="Quarter">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Region"/>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="RegionEq">
  <owl:equivalentClass>
    <owl:Class rdf:about="#Region"/>
  </owl:equivalentClass>
</owl:Class>
<owl:Class rdf:ID="State"/>
<owl:Class rdf:ID="EDU">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Region"/>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="PoliticalDivisionEq">
  <owl:equivalentClass>
    <owl:Class rdf:about="#Region"/>
  </owl:equivalentClass>
</owl:Class>
<owl:Class rdf:ID="StateEq">
  <owl:equivalentClass>
    <owl:Class rdf:about="#Region"/>
```

```
  </owl:equivalentClass>
</owl:Class>
<owl:Class rdf:ID="Zone">
  <rdfs:subClassOf>
    <owl:Class rdf:about="#Region"/>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="City"/>
<owl:Class rdf:about="#Region">
  <owl:equivalentClass rdf:resource="#ZoneEq"/>
  <owl:equivalentClass rdf:resource="#District"/>
  <owl:equivalentClass rdf:resource="#PoliticalDivisionEq"/>
  <rdfs:subClassOf rdf:resource="#Continent"/>
  <owl:equivalentClass rdf:resource="#StateEq"/>
  <rdfs:subClassOf rdf:resource="#State"/>
  <owl:equivalentClass rdf:resource="#Position"/>
  <owl:equivalentClass rdf:resource="#CountryEq"/>
  <rdfs:subClassOf>
    <owl:Class rdf:ID="SpatialObject"/>
  </rdfs:subClassOf>
  <owl:equivalentClass rdf:resource="#AreaEq"/>
  <rdfs:subClassOf rdf:resource="#GeometricalObject"/>
  <owl:equivalentClass rdf:resource="#RegionEq"/>
  <rdfs:subClassOf rdf:resource="#AdministrativeRegion"/>
  <rdfs:subClassOf rdf:resource="#City"/>
  <rdfs:subClassOf rdf:resource="#Country"/>
</owl:Class>
<owl:Class rdf:ID="StateSub">
  <rdfs:subClassOf rdf:resource="#Region"/>
</owl:Class>
<owl:ObjectProperty rdf:ID="hasPosition">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Region"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasContext">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Region"/>
      </owl:unionOf>
    </owl:Class>
```

```
 </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasArea">
 <rdfs:domain>
  <owl:Class>
   <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#Region"/>
   </owl:unionOf>
  </owl:Class>
 </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="isPartOf">
 <rdfs:domain>
  <owl:Class>
   <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#Region"/>
   </owl:unionOf>
  </owl:Class>
 </rdfs:domain>
 <protege:allowedParent rdf:resource="#City"/>
 <rdfs:range rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasEdge">
 <rdfs:domain>
  <owl:Class>
   <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#Region"/>
   </owl:unionOf>
  </owl:Class>
 </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasBoundary">
 <rdfs:domain>
  <owl:Class>
   <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#Region"/>
   </owl:unionOf>
  </owl:Class>
 </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasLocation">
 <rdfs:domain>
  <owl:Class>
   <owl:unionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#Region"/>
   </owl:unionOf>
  </owl:Class>
```

```xml
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasCommonCharacteristics">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Region"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="isA">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Region"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <protege:allowedParent rdf:resource="#SpatialObjectSub"/>
  <rdfs:range rdf:resource="http://www.w3.org/2002/07/owl#Class"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasDistance">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Region"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasCapital">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Region"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasClimate">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Region"/>
      </owl:unionOf>
```

```
      </owl:Class>
    </rdfs:domain>
  </owl:ObjectProperty>
</rdf:RDF>

<!-- Created with Protege (with OWL Plugin 1.2, Build 161)
http://protege.stanford.edu -->
```