Transactions in GIS, 2003, 7(2): 243–265

Research Paper

Revisiting the Concept of Geospatial Data Interoperability within the Scope of Human Communication Processes

Jean Brodeur
Natural Resources Canada
Centre for Topographic Information
and Centre for Research in Geomatics

Yvan Bédard
Geomatics Sciences Department
Université Laval, Québec and
Centre for Research in Geomatics

Geoffrey Edwards
Geomatics Sciences Department
Université Laval, Québec and
Centre for Research in Geomatics

Bernard Moulin
Computer Science Department
Université Laval, Québec and
Centre for Research in Geomatics

Abstract
Geospatial data interoperability has been the target of major efforts by standardization bodies (e.g. OGC, ISO/TC 211) and the research community since the beginning of the 1990s. It is seen as a solution for sharing and integrating geospatial data, more specifically to solve the syntactic, schematic, and semantic as well as the spatial and temporal heterogeneities between various representations of real-world phenomena. A few models have been proposed to automatically overcome heterogeneity of geospatial data and, as a result, increase the interoperability of geospatial data. However, the addition of a conceptual framework of geospatial data interoperability would contribute to understanding geospatial data interoperability, the appreciation of where existing contributions specifically apply, and would foster new contributions. In this paper, we revisit the concept of geospatial data interoperability within the broader scope of human communication and cognition. Human communication appears to be a rich framework since humans interoperate more easily than computers do. Accordingly, we present a conceptual framework of geospatial data interoperability that is broader in scope than existing frameworks and supported by several practical examples. An ontology of geospatial data interoperability is also introduced in order to refine the description of the conceptual framework. In such a communication-based framework, the notions of concept,

Address for correspondence: Jean Brodeur, Natural Resources Canada, Centre for Topographic Information, 2144 King Street West, Sherbrooke, Quebec J1J 2E8, Canada.
E-mail: Brodeur@RNCan.gc.ca

context, proximity, and ontology appear to be fundamental elements. These elements constitute a new approach to *geosemantic proximity*.

1 Introduction

For almost a decade, interoperability of geospatial data has been a prime concern in the geospatial information community. Software developers, data producers, and users aim at enabling the sharing and integration of geospatial data and geoprocessing resources (Kottman 1999). Organizations such as the OpenGIS Consortium Inc. (OGC) and ISO/TC 211, as well as the research community, have pioneered in laying out the current foundation for geospatial data interoperability. This community views interoperability as a solution to problems arising from syntactic, structural, and semantic heterogeneities, especially spatial and temporal heterogeneities, between data sources (Charron 1995, Bishr 1997, Laurini 1998, Ouksel and Sheth 1999, Sheth 1999).

Within the context of OGC, interoperability corresponds to "software components operating reciprocally to overcome tedious batch conversion tasks, import/export obstacles, and distributed resource access barriers imposed by heterogeneous processing environments and heterogeneous data" (McKee and Buehler 1998). Sondheim et al. (1999) describe interoperability as a non-imposed, bottom-up approach where heterogeneous systems, data models, and data sources deployed independently of one another, can exchange data and handle queries (and other processing requests) as well as make use of a common understanding of the data and requests.

Progress in geospatial interoperability is observed for syntactic heterogeneity (i.e. GIS format translation, spatial data structure transformation, such as raster to vector transformation) and geometric data-type definition. Progress is also observed for structural heterogeneity, that is, differences in the internal organization of GIS application data, the geodetic datum, map projections, and coordinate systems. Research in geospatial interoperability, however, must go beyond geometry-related and database-structure concerns to take into account semantics (Egenhofer 1999, Rodriguez 2000). Reconciling both semantic and geometric heterogeneities between different geospatial datasets describing the same phenomenon is deemed a major challenge. For example, it is still a problem to reconcile representations such as *Wetlands*, *Marshes/swamps*, *Marshes*, *Swamps*, *Marshes/fens*, *Milieux humides*, and *Marais* that are used to describe the same phenomena (see Table 1 for additional details).

The nature of the problem that initiated the research presented in this paper relates to users of geospatial data increasingly having to deal with numerous data sources to meet their specific needs. Examples of Canadian sources include the National Topographic Data Base produced by the Department of Natural Resources Canada (Natural Resources Canada 1996); the Street Network Files, the Digital Boundary Files, and the Digital Cartographic Files produced by Statistics Canada (Statistics Canada 1997); the VMap libraries produced for military purposes (VMap 1995); the National Atlas of Canada produced by the Department of Natural Resources Canada (Natural Resources Canada 1996); and several provincial topographic data sources that are usually carried out at larger scales (e.g. BC Ministry of Environment Lands and Parks (Geographic Data BC) 1992, New Brunswick 2000, OBM 1996, Québec 2000). Typically, each data source describes closely related topographical phenomena differently. See for instance *Waterbody*, *Lake*, *Lake/pond*, *Coastline*, *River/stream*, and *Canal* in Table 1 which lists

Table 1 Examples of phenomena abstracted differently in independent topographical databases

NTDB[1]	VMap[2]	BC Digital Baseline Mapping (BCDBM)[3]	ON Digital Topographic Database (ONDTD)[4]	BDTQ[5]	Information sur les terres et les eaux pour la province du Nouveau-Brunswick[6]
– Waterbody [2D] – Watercourse [1D] – Irrigation Canal [2D] – Navigable Canal [1D][2D] – Flooded area [2D] – Reservoir [2D] – Liquid depot/dump [0D][2D]	– Lake/Pond [2D] – Lake subject to inundation [2D] – River/stream [2D] – Coastline/shoreline [1D]	– Coastline [1D] – Ditch [1D] – Flooded land [1D] – Lake [1D] – River/stream [1D]	– Flooded land [0D] [1D] – Lake [0D] [1D] – River/stream [0D] [1D]	– Canal [1D][2D] – Cours d'eau [1D][2D] – Lac [2D] – Mare [2D]	– Canal [0D][1D] – Rivière – trait double [1D] – Lac (?) – Littoral (?) – Lac de rivière (?)
– Wetland [2D]	– Marsh/swamp [2D]	– Marsh [0D][1D] – Swamp [0D][1D]	– Marsh/Fen [0D] [1D]	– Milieu humide (végétation) [2D]	– Marais de canneberge (?) – Marais (?)
– Road [1D] – Limited access road [1D]	– Road [1D] – Car track [1D]	– Road [1D]	– Accesway [1D] – *Road* [1D]	– Voie de *communication* [1D] – Autoroute [1D] – Rue [1D] – Chemin [1D] – Route [1D]	– Artère (?) – Route collectrice (?) – Chemin local (?) – Chemin municipal (?) – Chemin d'accès aux ressources naturelles (?) – Route en construction (?) – Rue (?)
– Vegetation [2D]	– Trees [2D] – *Orchard/plantation* [2D] – Vineyard [2D]	– Wooded area [1D] – *Vineyard* [1D] – Orchard [1D] – Nursery [1D]	– Wooded area [0D] [1D]	– Milieu boisé [2D] – Verger (aires désignées) [2D]	– Clairière (?) – Bande défrichée (>100 m) (?) – Pépinière (?) – Verger (?) – Rangée d'arbres (>100 m) (?) – Zone boisée (>2 m haut) (?)
– Railroad [1D]	– Railroad [0D][1D] – Railroad siding/ railroad spur [1D]	– RailLine [1D]	– Rail line [1D]	– Voie ferrée [1D]	– Chemin de fer (?) – Triage de chemin de fer (?)
– Bridge [1D] – Obstacle to air *Navigation* [1D]	– Bridge/overpass/ viaduc (?)	– Bridge [1D][2D] – Trestle [1D][2D]	– Bridge (roadway) [0D][1D] – Bridge (railway) [0D][1D] – Culvert (roadway) [0D][1D] – Culvert (railway) [0D][1D]	– Pont [0D][1D] – Pont d'étagement [1D]	– Pont (?) – Ponceau (petit) (?)

Spatial pictogram descriptions: [0D]:0D; [1D]:1D; [2D]:2D; ?:unknown geometry; [0D] [1D]:multiple geometry; [0D][1D]:alternate geometry (see Bédard (1999b) and Brodeur et al. (2000) for more details). [1](Natural Resources Canada 1996); [2](VMap 1995); [3](BC Ministry of Environment Lands and Parks (Geographic Data BC) 1992); [4](OBM 1996); [5](Québec 2000); [6](New Brunswick 2000).

some examples of phenomena represented differently by different sources, both geometrically and semantically. Consequently, the retrieval of geospatial data complying with the user's needs and, subsequently, the data's integration into a coherent whole remains a crucial challenge. This kind of interoperability problem could have been addressed from different points of view (e.g. Harvey 1997). Our perspective is strongly influenced by the artificial intelligence and database modelling approaches to human communication and ontologies, which are connected but different from the philosophical perspective of these topics.

Accordingly, the next section of this paper reviews some fundamental notions of communication, cognitive science, ontology, and database modelling that support the proposed framework of interoperability presented in Section 3. In Section 4, we depict the five phases and three levels of the ontology of geospatial data interoperability supported by the framework. In Section 5, we develop the mapping between a concept and a conceptual representation. Section 6 introduces the approach we are working on to assess semantic proximity between a geospatial concept and a geospatial conceptual representation. This is called *geosemantic proximity* as it considers in a holistic way the semantic, spatial, and temporal descriptions of geospatial concepts and geospatial conceptual representations. Section 7 concludes this paper and indicates future work.

2 Interoperability and the Human Communication Process

The framework of geospatial data interoperability that will be discussed in the following sections is supported by some theoretical notions used in the fields of human communication and perception, ontology, data modelling, and, more specifically, by the notions of context and semantic proximity found in artificial intelligence. This section describes these concepts.

2.1 Communication Process

The study of the different aspects of communication between all kinds of systems originated in 1948 with Norbert Weiner's ideas concerning cybernetics and the numerous adaptations that followed, yielding new insights into human communication (Weiner 1950, Blake and Haroldsen 1975, Campbell 1982, Sowa 1984). We believe that the communication process between humans represents an ideal model of what interoperability should be. It begins when an individual has something in mind representing real-world phenomena and wants to communicate it to someone else.

The communication process is described as being composed of a source, a signal, a communication channel, a destination, a possible source of noise, and feedback. In the first stage of the communication process, the source has a representation of real-world phenomena that corresponds, for humans, to their cognitive model (Denes and Pinson 1971, Schramm 1971b, Bédard 1986, Logie and Denis 1991) and, for machines, to a part of their physical memory. This model is developed through the direct observation (detection and recognition of raw signals) of phenomena and from the observation of preprocessed, intentional semantic signals from others (Bédard 1986). The source selects information to be communicated, transforms it into signals such as words (spoken or written) or data, and organizes them into a message that is placed in the communication channel (Denes and Pinson 1971, Campbell 1982). This is known as the encoding process. It follows rules that are more or less formal depending on the context and the nature of the encoder (human or machine). The resulting signals or data are physical descriptions free of any intrinsic signification (Schramm 1971a, Cherry 1978, Campbell 1982, Bédard 1986). Signification is what signals evoke to the source and to the destination, respectively and, consequently, cannot be transmitted (Cherry 1978). Signals are the mediation component between the source and the destination, but their intended meaning is not embedded within the signal. Once the destination has received the message, the decoding process starts as it tries to *understand* the incoming signal, that is, to find its intended meaning, and the communication process ends with the creation of the destination's evoked concepts (Schramm 1971a, Bédard 1986). The communication process works properly when the destination's evoked concepts are sufficiently isomorphic to the source's concepts, that is, when both represent the same real-world phenomena. Feedback, or retroactive communication may be used to improve isomorphism. The notion of commonness (Schramm 1971a) is basic in the communication process: in other words, the source and destination shall have a common set of knowledge and signals to make the process work properly. The destination relies on signals and referents (i.e. knowledge and beliefs of the world) to recognize the message (Denes and Pinson 1971, Krech and Crutchfield 1971).

2.2 *Perception and Cognition*

In the communication process, as well as in the case of interoperability, perception and cognition play a leading role in building, structuring, and disseminating human information. As mentioned previously, human communication begins when someone wants to transmit information in their mind to someone else. As such, cognitive models are basic elements of human-to-human communication. They are built up from physical signals captured through our sensory systems, which generate *perceptual states* (Barsalou 1999). Then, the human selective attention extracts only a subset of interest among these perceptual states (Krech and Crutchfield 1971, Sears and Freedman 1971, Barsalou 1999) and stores this permanently in memory as *perceptual symbols* (Barsalou 1999).

The literature essentially recognizes two modes of perceptual symbol representations. The first corresponds to modal or analogical representations (Kosslyn 1980, Barsalou 1999, Kettani and Moulin 1999) and is isomorphic to the perceptual state such as an image captured by the visual sensory system. It consists of the reproduction or conversion of raw signals into memory. The second mode corresponds to amodal or propositional representations (Barsalou 1999, Kettani and Moulin 1999) and refers to *tacit knowledge* (Pylyshyn 1981). Inspired by logics, statistics, mathematics, and computer science, it corresponds to structures such as feature lists, frames, schemata, and semantic nets (Lehmann 1992, Barsalou 1999). Moreover, Barsalou (1999) brought a new definition of a perceptual symbol as a record of neural activation resulting from the perception process where the neural system, which is common to imagery and perception, underlies the conceptual knowledge. Perceptual symbols are more likely qualitative and functional, and are not stored independently from others in memory (Krech and Crutchfield 1971, Barsalou 1999).

In Barsalou's theory, a perceptual symbol corresponds to a *concept* and behaves like a *simulator* that generates conceptual representations (i.e. simulation of the concept). This notion of *simulator* is similar to a kind of dynamic translator generating translations of the concept on the fly for a specific use. Concepts are made of cognitive elements, which are not directly accessible, and a translator function that encapsulates these elements. The translator function reproduces these cognitive elements in the context of data processes. A concept can only be communicated via selected data elements translated into physical signals, which are conceptual representations. A huge literature in the field of semiology exists which defines rules about the best ways to create conceptual representations. For instance, a concept corresponding to "water area" can be translated in a number of conceptual representations such as "waterbody," "coastline," "lake," and "river/stream" represented by either surfaces or lines in different colors, shading and line styles (all constituting possible simulations of "water area"). The concept's translator also recognizes conceptual representations associated with the concept (Barsalou 1999). This is carried out by matching a simulation of the concept (i.e. a generated conceptual representation) with the incoming conceptual representation; when the matching fails, a new concept is instantiated.

2.3 Ontology and Conceptual Modelling for Database Development

The communication process, as a proper model to depict interoperability, involves real-world phenomena along with their descriptions: the different human cognitive models and physical models such as signals. Real-world phenomena, their identification, and description have been studied within the realms of ontology and conceptual modelling for database development.

In its philosophical meaning, ontology stands for the description of the world in itself (Peuquet et al. 1998); a model and an abstract theory of the world (Smith and Mark 1999); and, the science of being, of the type of entities, of properties, of categories, and of relationships that compose reality (Lehmann 1992, Peuquet et al. 1998, Smith and Mark 1999, Bittner and Edwards 2001). It is described from two perspectives. The first one, called formal ontology, refers to shared structures between scientific domains such as identity, plurality, and unity. The second one, called material ontology, relates to the conditions that are necessary to belong to an entity type within a given domain (Peuquet et al. 1998). In artificial intelligence (AI), Gruber (1993a, b) defines an ontology as "an explicit specification of a conceptualisation" and Guarino (1998) as "a logical theory accounting for the intended meaning of a formal vocabulary." AI definitions of ontology and the material ontology in philosophy tend to follow a similar objective. As shown in Table 1, there may be multiple ways of describing a single conceptualization. This is particularly reflected in Gruber's definition of ontology, which admits that each explicit description (i.e. specification, vocabulary) consists of one specific ontology. Guarino's definition goes further by considering the ontological relationship (i.e. the "intended meaning") that exists between a description (i.e. the vocabulary) and the concept it evokes. Consequently, we can admit a relationship between the philosophical and the AI notions of ontology only if we consider that "conceptualization" in the AI context corresponds to the philosophical definition of ontology (Rodriguez 2000). We can also accept that descriptions from multiple ontologies (as in AI) can ontologically refer to the same concept (or phenomenon). While there is an obvious connection between the philosophical and the AI definition of ontology, the latter is definitely considered as the main orientation of this paper. As such, we refer to ontology as a formal representation of phenomena with an underlying vocabulary including definitions and axioms that make the intended meaning explicit and describe phenomena and their interrelationships.

As applied to databases, a conceptual model is a simplified, abstract representation of a portion of reality resulting from a data-centered analysis of users' interests (Bédard 1999a, Simsion 2001). It results from reflective thinking to better understand that part of reality and to communicate information about it (Collongues et al. 1987, Simsion 2001). Bédard (1999b) mentions that conceptual models serve as tools for thinking, communication, development, and documentation. Conceptual models retain, organize, and store only features of interest in terms of general categories, object classes, properties, relationships, generalizations, aggregations, roles, constraints, behaviour, geometry, temporality, and so on, either in a lexical (e.g. EXPRESS, Bakus-Naur formalism) or graphical formalism (e.g. entity relationship formalism, UML) (Brodeur et al. 2000). Ideally a data dictionary defining the semantics of each schema component is included in a conceptual model, which makes the intended meaning of the modelled feature explicit. In a conceptual model, objects must be unique in the context of the database and, as such, characterized by only one combination of properties and relationships (Collongues et al. 1987, Simsion 2001). Because of the specific perspective for which a data model is elaborated or because of the experience of the data modeler, there is usually more than one data model to express the same part of reality (Collongues et al. 1987, Simsion 2001). The problem emerging from the existence of different conceptual models is to establish the relationship between models describing the same set of phenomena. Based on the definitions of data models and ontology, we see ontology as a theoretic layer underlying conceptual modelling providing the means to link classes and instances depicting the same part of reality differently but in similar ways. This linkage of classes and instances is made possible through the analysis of intrinsic and extrinsic properties (see Section 6) that give them their "identity."

2.4 Context

The situation in which a real-world phenomenon is perceived, abstracted, and used governs its description. This *context* affects the definition of concepts and conceptual representations. *Context* is widely recognized as a fundamental notion in semantic interoperability. It provides concepts and conceptual representations with real-world semantics (Kashyap and Sheth 1996, Ouksel and Sheth 1999, Wisse 2000). As reported in Kashyap and Sheth (1996), however, *context* has been associated with various other ideas such as knowledge for reasoning about other systems (Ouksel and Naiman 1993); signification, content, organization, and properties about data (Sciore et al. 1992); a characteristic associated with a partition of an ontology (Guha 1990); and membership in a database, relationship, export schema, or internal schema (Sheth and Kashyap 1992). Moreover, Frank's (2001) *five tiers of ontology* introduce the context as a main component of the social ontology. Context has influence in the conceptual level as well as in the implementation level. On the one hand, context drives how phenomena are perceived and abstracted that result in different object classes, properties, geometries, temporalities, relationships, and so on. On the other hand, it also acts at the implementation level for instance in specific data capture specifications such as "rapids depicted on a map by three points less than 100 metres apart and stretching over a distance of more than 100 metres are consolidated into a line." Following a conceptual orientation, *context* is in our approach associated with the manner an individual abstracts real-world phenomena; the description of these phenomena is organized into intrinsic and extrinsic properties of the corresponding concepts and conceptual representations.

Intrinsic properties refer to the literal meaning while extrinsic properties refer to the dependencies with other concepts or conceptual representations. For example, intrinsic properties of a road concept can be described by its classification, surface type, status, and geometry while extrinsic properties can be the relationships the road has with concepts such as built-up areas, bridges, dams, fords, tunnels, and the behaviour of the road in different situations. An important challenge in semantic interoperability of geospatial data resides in increasing *context*-based reasoning capabilities.

2.5 Semantic Proximity

In interoperability, as in the communication process, an incoming conceptual representation must be recognized and given a specific meaning. This implies matching the conceptual representation with a concept in the destination's cognitive model. The matching operation analyzes the context of both the conceptual representation and the concept to retrieve commonalities between them. This issue has been studied in AI thanks to the notion of semantic proximity which expresses the similarity between conceptual representations such as in semantic networks (Cohen 1982, Lehmann 1992), knowledge networks (Frankhauser et al. 1991, Frankhauser and Neuhold 1992), and context-based approaches (Kashyap and Sheth 1996, 1998). Generally speaking, a semantic network is an interconnected node–arc-like structure such as a frame system, a relational graph, a hierarchy (e.g. lattice, tree, or acyclic graph) in which nodes represent conceptual elements and arcs the relationships that exist between such elements (Cohen 1982, Lehmann 1992). Semantic proximity deals with the semantic relatedness between two different conceptual representations. The closest conceptual representation to a given one is typically the one having the smallest conceptual distance, i.e. the shortest path to a given conceptual representation in a representational space (Rodriguez 2000). Frankhauser et al. (1991) and Frankhauser and Neuhold (1992) implement knowledge networks in which conceptual representations (nodes) are linked to others by associations of the type generalization/specialization, negative association, or positive association. A coefficient expressing the strength between the two conceptual representations is assigned to each association. For two non-associated conceptual representations, a relationship and strength are inferred by traversing the network from one conceptual representation to the other, analyzing the nature and the strength of each relationship. In Kashyap and Sheth (1996, 1998), the semantic proximity between two conceptual representations corresponds to their comparison subjected to: (1) the context of comparison; (2) the abstraction or mapping used to associate their respective domain; (3) their domain (i.e. the possible values); and (4) their state (i.e. the values they hold at a given time). Predicates such as semantic resemblance, semantic relevance, semantic relation, semantic equivalence, and semantic incompatibility are used to express qualitatively the semantic proximity that exists between two conceptual representations.

3 A Conceptual Framework of Geospatial Data Interoperability

A few attempts have been made to automatically overcome semantic heterogeneity and increase the interoperability of geospatial data, notably the *Semantic Formal Data Structure* (Bishr 1997), the *Matching Distance* model (Rodriguez 2000), and the *Isis* approach (Benslimane 2001). In Bishr (1997), interoperability is defined as "the ability of a system or components of a system to provide information sharing and inter-application co-operative process control." Accordingly, the Semantic Formal Data Structure was elaborated to reconcile heterogeneous representations of a unique concept combining loosely coupled federated schemata and a Proxy Context mediator. In this approach, users formulate queries in their own vocabulary and submit them to the Context Mediator. Subsequently, the Context Mediator translates these queries according to a shared context definition (i.e. a federated schema) and passes the queries to the data source export schema to retrieve the data complying with the user's queries.

Rodriguez (2000) defines the semantic interoperability problem as the "identification of semantically similar objects belonging to different databases and the resolution of their schematic differences" (i.e. differences between the database schemata). She proposes the *Matching Distance* model to evaluate the semantic similarity between object classes of geospatial features. The variability and resemblance of class functions, parts, and attributes are analyzed to compute a quantitative measure of similarity between object classes. This model makes use of an acyclic graph with *is-a* and *part/whole* relationships, and ontologies in which class definitions are taken.

In the Isis approach (Benslimane 2001), interoperability is defined as an operation in which: (1) clients and data sources adopt a common representation model, (2) clients and data sources share a mutual understanding of common features, and (3) a process can dynamically transform one object representation into another, adapting the semantics and structure to the user's needs. This definition of interoperability has been tested in a mediation context-based approach in which context cooperation schemata, resulting from the interpretation of export schemata of independent databases with a unique reference context, are used to translate data from one database to another.

In the light of these somewhat different approaches, we believe that a more global framework would enhance understanding of geospatial data interoperability (Egenhofer 1999) and provide a theoretical foundation to better appreciate each contribution and foster new ones. In this section, we revisit the notion of interoperability within the broader scope of human communication and cognition. People interact using different representations of observable phenomena but regularly end up understanding each other. We believe that interoperability consists in a process similar to human communication (Darnell 1971; Denes and Pinson 1971; Lippmann 1971; Schramm 1971a, b; Bédard 1986) in which independent systems automatically manipulate, exchange, and integrate data coming from each other. This assumption motivated the creation of the conceptual framework of geospatial data interoperability, which is proposed hereafter, and of the concept of *geosemantic proximity* introduced in section 6.

First, let us assume the following situation. An individual, hereafter called a user agent (A_u), is looking for information about the hydrographic network in the area of the City of Sherbrooke. He or she launches a query on a search engine with the keywords "Lake," "River," and "Sherbrooke" targeting a geospatial database, hereafter called the data provider agent (A_{dp}). A_{dp} receives and interprets the request, searches for related information and, referring to the content it is aware of, sends a response to A_u. In other words, A_{dp} provides the main "Watercourses ◩" and "Waterbodies ⊡" in the vicinity of "Sherbrooke ⊡" (for instance "Lac des Nations ⊡," "Magog River ◩⊡," and "Saint-François River ⊡"). These elements correspond exactly to A_u's query.

This situation illustrates what interoperability should be between two agents. In this case, interoperability is associated with an *interpersonal communication* (Blake and

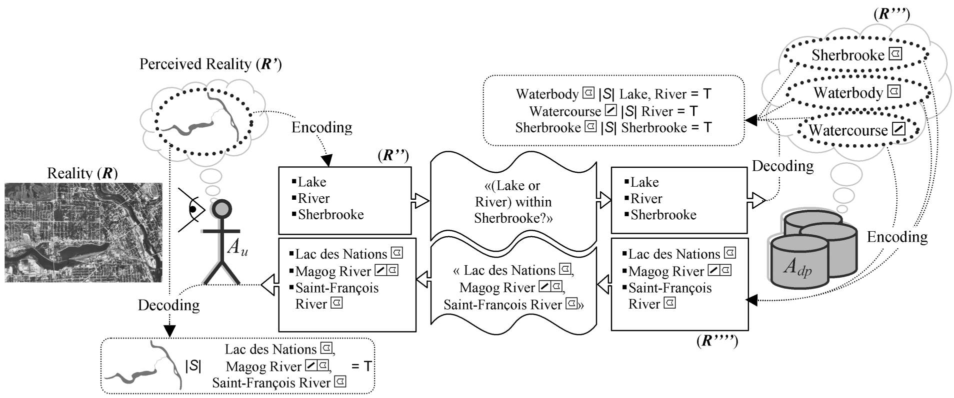

Figure 1 A conceptual framework of geospatial data interoperability (adapted from the communication process of geographic information systems in Bédard (1986))

Haroldsen 1975) (i.e. a dialogue-like communication) between two agents, each of them using its own vocabulary to express abstractions of real-world phenomena. As long as the two agents have a common background and a common set of symbols, they regularly end up understanding each other (Schramm 1971a, Bédard 1986).

Let us review this situation described as a communication process between the two agents. Figure 1 depicts in greater detail the interaction between them. First, there is the topographic reality as it exists at a given time and about which A_u is looking for hydrographic information (represented by ***R*** in the model).

Second, A_u's cognitive model of ***R*** is built from observed signals and his or her frame of reference–the set of rules and knowledge he or she used to abstract phenomena. A_u's cognitive model consists of properties that are judged significant. These properties are joined together and structured within concepts. A *concept* is a simplified version of a real-world phenomenon or part of it that does not exist in reality; it is entirely fictional (Sowa 1984). It is an abstract notion that denotes the "picture" an agent has in mind (Lippmann 1971, Schramm 1971a, Bédard 1986, Kettani and Moulin 1999). All concepts that A_u has in mind constitute his or her representation of reality, that is, his or her cognitive model, which is identified by ***R'*** in the framework.

As they are only abstractions, concepts cannot be communicated directly between agents and, as such, they must be transformed into physical representations. This is known in the communication process as the encoding operation. In this operation, only properties that adequately translate a concept in a given situation are selected. Then, these are transferred into signals of different types (words, abbreviations, punctuation, symbols, pictograms, etc.), which are aligned in a specific order according to a set of rules (i.e. a grammar) to build a conceptual representation. The encoded conceptual representation becomes the physical component depicting partly or wholly the concept to which it refers. This conceptual representation forms the third expression of the reality. It is illustrated by "Lake," "River," and "Sherbrooke" in Figure 1 and is identified by ***R''***. It designates the data transmitted and used for interoperability. Conceptual representations are released on the communication channel exempt of any of A_u's intended meaning and the message representing A_u's request – "(Lake or River) in Sherbrooke?" (see Figure 1) – travels to its destination.

At the destination, A_{dp} initiates the decoding operation, which aims at recognizing the received conceptual representations and at assigning them appropriate meaning. In our framework, this task is assigned to concepts (this will be discussed in more detail in Section 5). Under perfect conditions, conceptual representations will induce concepts in A_{dp} that are isomorphic to A_u's concepts. However, in most situations, conceptual representations will induce in A_{dp} concepts of similar meaning to A_u's concepts. The set of A_{dp}'s concepts constitutes the fourth expression of reality. It is denoted by ***R'''*** and illustrated by "Waterbody," "Watercourse," and "Sherbrooke" in the theoretical model (Figure 1).

When conceptual representations have been recognized, A_{dp} initiates the retrieval of information complying with A_u's interests. However, as is the case in ***R'***, concepts and even tokens (i.e. instances of concepts) matching A_u's request cannot be transmitted directly. Similarly to A_u's cognitive model, concepts consist of internal representations that are hidden to external agents. Consequently, concepts and tokens must be encoded into conceptual representations and placed in a reply message on the communication channel to reach A_u. These encoded conceptual representations constitute the fifth representation of the reality. It is denoted by ***R''''*** and illustrated by "Lac des Nations," "Magog River," and "Saint-François River" in Figure 1.

Once the reply reaches its destination, A_u starts the decoding operation in order to recognize the incoming message; he or she analyzes the conceptual representations included in the message to assess if they infer the previous concepts in ***R'***. If so, we say that interoperability occurred during the interaction between the two agents. This means that interoperability is not simply a one-directional communication process, but it consists of a bi-directional process as clearly demonstrated in the conceptual framework. It includes feedback in both directions which ensures that messages issued by A_u and A_{dp} have reached their destination and have been understood properly. We believe that this issue of bi-directional communication process takes on a fundamental character in the search for a solution to semantic interoperability of geospatial data.

4 Ontology of Geospatial Data Interoperability

As introduced in Section 2, ontology refers to a formal and accepted representation of phenomena with an underlying vocabulary, including definitions that make the intended meaning explicit. In the framework presented above, we have demonstrated that reality takes various configurations beginning with the reality itself, human beings' cognitive representations (or their physical counterpart in machines), and physical representations. The proposed conceptual framework introduces five different representations of the same reality (***R***, ***R'***, ***R''***, ***R'''***, and ***R''''***). Each of these representations is a distinct ontology of the reality that occurs in the proposed framework of geospatial data interoperability. Frank (2001) has already introduced a subdivision of ontology called the *five tiers of ontology*, in which a distinction is made between the different abstraction levels that an agent has to deal with when building its cognitive model, namely: physical reality, observation of physical world, objects with properties, social reality, and subjective knowledge. Here, we introduce a different and complementary subdivision of ontology for the purpose of geospatial data interoperability composed of the *five ontological phases of geospatial data interoperability* and the *three levels of ontology* presented hereafter. These new subdivisions will allow us to view interoperability from a different angle.

4.1 The Five Ontological Phases of Geospatial Data Interoperability

The *five ontological phases of geospatial data interoperability* consist of the five different facets of reality that appear in the framework of geospatial data interoperability. The first ontological phase consists of the reality itself (***R***), which is beyond description. Each phenomenon has its own identity that makes it distinguishable from the others.

The second ontological phase is A_u's cognitive model of the reality, ***R'***. It gathers all the concepts that take place in A_u's memory. This ontology results from direct observations (i.e. from our sensory motors) and indirect observations (i.e. mechanical sensors, information captured by other agents) of reality. It is a partial description of reality and can be viewed as a subset of ***R*** corresponding to A_u's *affordances* (Gibson 1979). This ontology is A_u's internal representation of reality.

The third ontological phase is the set of conceptual representations ***R''*** (objects and object classes) that are used to signify concepts of A_u's ontology. This ontology uses a vocabulary, which accurately specifies the intended meaning attributed to the different concepts. Each conceptual representation describes a concept within a specific context. Therefore, more than one conceptual representation can refer to a given concept, for example "Vegetation," "Tree," and "Wooded area" (see Table 1 for other examples). It is also possible that one conceptual representation refers to different concepts, depending on the context in which they are used. This refers to the notion of polysemy. For example, "Bridge" may refer to a road infrastructure, a hazard to air navigation, or even a hazard to marine navigation.

The fourth ontological phase consists of the set of A_{dp}'s concepts and refers to the database's internal representation of reality ***R'''***. In the theoretical framework, database agents such as A_{dp} behave as cognitive agents in which descriptions of concepts are internal representations of real-world phenomena that serve as interlingua in the interaction with other agents. Database concepts also include functions that can afford reasoning capabilities such as production and recognition of conceptual representations (this will be discussed in Section 5).

The fifth and last ontological phase consists of the conceptual representations ***R''''*** that A_{dp}'s concepts can produce. Like ***R''***, this ontological phase consists of physical representations which use a vocabulary that aims to deliver the intended meaning of A_{dp}'s concepts (***R'''***).

In the *five ontological phases of geospatial data interoperability*, we consider that each ontological phase includes a set of properties describing the *identity* of phenomena. This set of properties allows the binding of the different representations with the phenomena. Moreover, it is our interpretation that the *five tiers of ontology* (Frank 2001) deal more with the steps involved in cognition, which apply to ***R'*** and ***R'''*** specifically. In this regard, the *five ontological phases of geospatial data interoperability* deal more with reality and its different representations that occur in the interaction between two agents. Consequently, we feel that Frank's *five tiers of ontology* and our *five ontological phases of geospatial data interoperability* are complementary.

4.2 Levels of Ontology

Reality is usually abstracted and described with more or less detail depending on the accuracy needed in a given situation. Accordingly, the meaning of concepts and conceptual representations is described from more general to more specialized when used in a

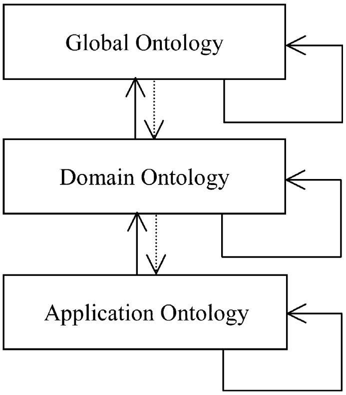

Figure 2 The *Three Levels of Ontology*

global context, within a scientific community or within a specific application, respectively. In the literature, authors refer typically to global ontology (Kashyap and Sheth 1996, 1998; Guarino 1998; Bergamaschi et al. 1999; Kahng and McLeod 1998; Sheth 1999; Smith 1999), domain ontology (Kashyap and Sheth 1996, 1998; Guarino 1998; Fowler et al. 1999; Sheth 1999; Smith 1999), and application ontology (Guarino 1998, Kahng and McLeod 1998, Kashyap and Sheth 1998, Smith 1999, Sycara et al. 1999). These *levels of ontology* (Figure 2) are characterized by a different granularity in the abstraction and description of phenomena. A comparison can be made with a dictionary describing a set of generic terms, a lexicon that is a brief dictionary specialized in a given science or technique, and a glossary that appears at the end of a book defining the specific meaning of terms used in the book. At the coarser level of granularity, global ontology compiles concepts or conceptual representations of a high and generic level of abstraction, independent of any specific domain. Examples of such ontologies are Wordnet (http://www.cogsci.princeton.edu/~wn/), CYC (http://www.cyc.com/), TERMIUM Plus (http://termiumplus.bureaudelatraduction.gc.ca), and *Le grand dictionnaire terminologique* (http://www.granddictionnaire.com). At the middle level of granularity, domain ontology makes an inventory of concepts or conceptual representations which are accepted and shared within an information community. An example of this level of ontology is the *National Standards for the Exchange of Digital Topographic Data: Topographic Codes and Dictionary of Topographic Features* (Canadian Council on Surveying and Mapping 1984), which compiles, defines, and structures a set of terms describing topographic phenomena. At the most detailed level of abstraction, an application ontology lists, defines, and organizes concepts or conceptual representations specific to an application. This kind of ontology is documented in many ways, for instance application schema (ISO/TC 211 2000), data dictionary, feature catalogue (ISO/TC 211 2001), repository (Brodeur et al. 2000), and data specification. Examples include the National Topographic Data Base – Standards and Specifications (Natural Resources Canada 1996) (http://scar.cits.rncan.gc.ca/bndt/), VMap Specifications (VMap 1995), British Columbia Specifications and Guidelines for Geomatics (BC Ministry of Environment Lands and Parks (Geographic Data BC) 1992), Ontario Digital Topographic Database-1:20,000, 1:10,000-A Guide for Users (OBM 1996), *Base de données topographiques du Québec (BDTQ) à l'échelle 1/20 000 – Normes de production* (Québec 2000), BD TOPO and

BD CARTO (http://www.ign.fr/fr/MP/BDGeo/), ATKIS-Digital Topographic Map 1:10,000 (http://www.atkis.de), and USGS-DLG (http://rockyweb.cr.usgs.gov/nmpstds/dlgstds.html).

As illustrated in Figure 2, the navigation between the different levels of ontology follows a bottom-up approach. An agent (A_u or A_{dp}) initiates its reasoning using its own knowledge and, if needed, that of other specific knowledge bases to which it has direct access. This corresponds to the application ontology level. When required, domain ontologies can be accessed to get shared conceptual representations within a specific community to facilitate communication between agents. Domain ontology can also be linked to other related domain ontologies to expand this level of knowledge. Finally, domain ontologies can access global ontologies to get conceptual representations of common usage. Again, such an ontology may be associated with others on the same level to expand the global representation and knowledge of reality.

Accordingly, we propose, as a complement of the framework, the *ontology of geospatial data interoperability* viewed as a two-dimensional subdivision. One dimension consists of the *five ontological phases of geospatial data interoperability* and the other consists of these *three levels of ontology*. This *ontology of geospatial data interoperability* consists of the various configurations of real-world phenomena descriptions that take place in interoperability. It shows the complexity and the components involved in geospatial data interoperability. As shown in our conceptual framework, geospatial data interoperability is not simply the task of being able to access geospatial data in a given format and schema and use it with a GIS system. Even if the geospatial data is transferred properly on your GIS system, they have to mean something otherwise they are useless. Geospatial data interoperability encompasses various abstractions and understandings of geospatial phenomena that are in interactions thanks to the communication process in which multiple ontologies of different granularities have to be considered in every phase of geospatial data interoperability. As such, the *ontology of geospatial data interoperability* helps to grasp and describe as a whole the scope of interoperability of geospatial data. In addition, along with the conceptual framework, it helps to understand all the relationships that exist between the real-world phenomena and their various descriptions. The ontology of geospatial data interoperability is then organized as shown in Figure 3, in which O_I^J identifies one component of the ontology signifying the level I of the ontological phase J. However, each component of this ontology may have different levels of relevance according to the different situation in which different geospatial databases take part.

5 Relationship between Concept and Conceptual Representation

With respect to the interoperability of geospatial data, encoding and decoding functions are crucial components since they are responsible for generating and recognizing geospatial conceptual representations, respectively. They are, to some extent, translation functions. Generally speaking, translators have been typically implemented as middleware components performing the conversion of a dataset from one data model and data format to another. Such an approach assumes that a correlation between data models and structures is already available or can be made in a timely and practical manner. This situation is acceptable when dealing with small databases. This is not the case when navigating on the Internet and dealing with a large number of datasets, since we have

Ontology	*R*	*R'*	*R''*	*R'''*	*R''''*
Global	O_G^R	$O_G^{R'}$	$O_G^{R''}$	$O_G^{R'''}$	$O_G^{R''''}$
Domain	O_D^R	$O_D^{R'}$	$O_D^{R''}$	$O_D^{R'''}$	$O_D^{R''''}$
Application	O_A^R	$O_A^{R'}$	$O_A^{R''}$	$O_A^{R'''}$	$O_A^{R''''}$

Figure 3 Ontology of Geospatial Data Interoperability

to deal with a practically infinite number of representations of the reality. Consequently, encoding and decoding functions are strongly tied to concepts in the framework of geospatial data interoperability outlined in Section 3. Therefore, a *concept* consists of the set of knowledge with the accompanying processes that an agent maintains about a phenomenon, which generate and recognize different representations of the concept. This position is supported by Barsalou's theory:

> ". . . a *concept* is equivalent to a simulator. It is the knowledge and accompanying processes that allow an individual to represent some kind of entity or event adequately. A given simulator can produce limitless simulations of a kind, with each simulation providing a different *conceptualization* of it. Whereas a concept represents a kind generally, a conceptualization provides one specific way of thinking about it. . . . Once a simulator becomes established in memory for a category, it helps identify members of the category on subsequent occasions . . ." (Barsalou 1999)

This approach to the assessment of geospatial data interoperability is in itself different from those already delineated in the geospatial information community, namely in Benslimane (2001) and Bishr (1997), and appears to be better aligned with human communication and cognition.

Accordingly, a concept appearing in ***R'*** and ***R'''*** behaves as a "simulator" which can generate different simulations of itself, that is, conceptual representations, as well as recognize a conceptual representation that is bound to it. Essentially, a concept of the brain or the machine is made of hidden data elements that are encapsulated by a simulation function (shown in Figure 1 by a dotted ellipse that encompasses the concept). The simulation function forms the main interface to access a concept.

This simulation function performs the encoding to produce conceptual representations such as those appearing in ***R''*** and ***R''''***. It is in some way a translation process that goes from a hidden and more neutral representation – the concept – to a language-dependent representation – the conceptual representation. This function selects and puts together properties that adequately describe the concept in a specific situation. It uses a vocabulary, punctuation, and grammar in order to build conceptual representations.

In order to produce conceptual representations, the concept's simulation function searches to find the best way to describe the concept in a given situation. As such, the function has to take into consideration other concepts of similar meaning. These concepts are abstracted from a different context and are all organized in the same ontology.

Like perceptual symbols (see Section 2), concepts are not stored independently of others in ***R'*** and ***R'''***. On the contrary, they are to some extent a kind of attractor within a dynamic network (as an ontology structure). When a new concept is introduced into the network, existing similar concepts try to attract this new concept, to place it nearby, and to build the necessary links with it. Hence, similar concepts are clustered together, expressing some sort of proximity. On the one hand, a concept is defined as a discrete notion (Sowa 1984) that takes part in an ontology as a phenomenon in reality. On the other hand, ontology is seen as a more continuous but yet partial representation of reality linking multiple concepts together. The concept's simulation function takes advantage of this ontology structure to produce and recognize conceptual representations that are best suited to any given situation.

The simulation function also implements the decoding operation, that is, recognition of conceptual representations that are bound to the concept. As stated by Barsalou (1999), ". . . if the simulator for a category can produce a satisfactory simulation of a perceived entity, the entity belongs in the category." So, in order to recognize a conceptual representation as a member of a concept, a concept must be able to produce such a conceptual representation. As illustrated in Figure 1, "Waterbody " and "Watercourse " are A_{dp}'s concepts that can produce the "Lake " and "River " conceptual representations and, as such, "Lake " and "River " are bound to the "Waterbody " and "Watercourse ?" concepts.

A conceptual representation describes a concept within a specific context. Context is introduced here as a metaconcept that is omnipresent in the representation of real-world phenomena (Wisse 2000). A context is as imaginary and fictitious as is a concept. It consists of elements that influence the use of a concept and provide its real signification. Like Ouksel and Sheth (1999), we consider the context as the main vehicle that provides real-world semantics. The context description is usually embedded in the components defining and characterizing conceptual representations (Wisse 2000) such as object classes, properties, geometries, temporalities, domains, relationships, behaviours, and memberships to datasets or ontologies.

Two conceptual representations of the same concept express a contextual variation. Context adds a degree of freedom to the ontological representation of concepts. It is the context, which drives us to use different conceptual representations for the description of real-world phenomena, and, for that reason, it is also a notion related to ontology. Let us return to the "Bridge" example, which is a concept describing a road infrastructure, a hazard to air navigation, or a hazard to marine operations (in three different contexts). In this example, context consists of the elements that influence the use of the concept and that specify its meaning. As we can observe in the example, the description of the context is typically embedded in the properties of the conceptual representations. Each conceptual representation has its own specific properties, such as structure type, the elevation of the highest point, or the clearance between the watercourse and the bridge, respectively. Independently of their specific descriptions, they all refer to the same concept, which ontologically links all of these representations.

The framework and related ideas presented so far aim at situating the broad picture of geospatial data interoperability. In the next section, we define the notion of *geosemantic proximity* and identify where it applies within our framework.

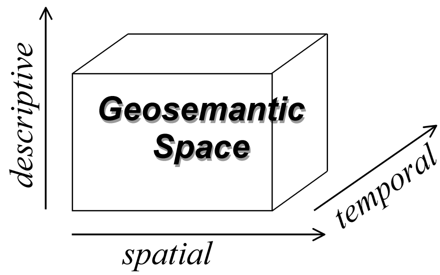

Figure 4 *Geosemantic Space*

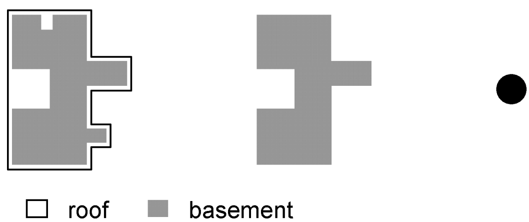

Figure 5 Various semantics of building's geometric representations

6 *Geosemantic Proximity*

The simulation function presented in the previous section generates and recognizes conceptual representations and, thus, evaluates the semantic proximity that stands between a concept and a conceptual representation, namely between ***R'*** and ***R''***, ***R''*** and ***R'''***, ***R'''*** and ***R''''***, and ***R''''*** and ***R'***. In the semantic interoperability of geospatial data, Bishr (1997) and Rodriguez (2000) mentioned that schematic heterogeneity can be solved only between semantically similar representations. Because it evaluates the semantic proximity, the simulation function becomes a key element of the conceptual framework and also a prerequisite to solving schematic heterogeneity. Consequently, we develop the notion of *geosemantic proximity* to concurrently assess the semantic, spatial, and temporal similarities (as components of a *geosemantic space*; see Figure 4) between a geospatial concept and a geospatial conceptual representation. Even if a geospatial concept and a geospatial conceptual representation have the same semantics, their spatial and temporal definitions may differ in several ways (Figure 5). For instance, the geometry of a building in a dataset may refer to the precise footprint of the basement while the geometry of a building in another dataset may refer to the precise footprint of the roof. The footprint may be geometrically delineated with more or less details because of different geometric depiction constraints. Also, in a third dataset, the same building may be represented as a point and be represented in a fourth dataset as a generalized surface, where all details smaller than 10% of the width or length of the building are ignored because of the difference in geometric granularity (e.g. different scales). All these datasets provide buildings with the same semantics from a pure "object-class/attribute" point of view (as it is usually considered in semantics proximity analysis) but they don't have different meanings from a geometric point of view and such difference is explicitly taken into account by what we called *geosemantic proximity* analysis.

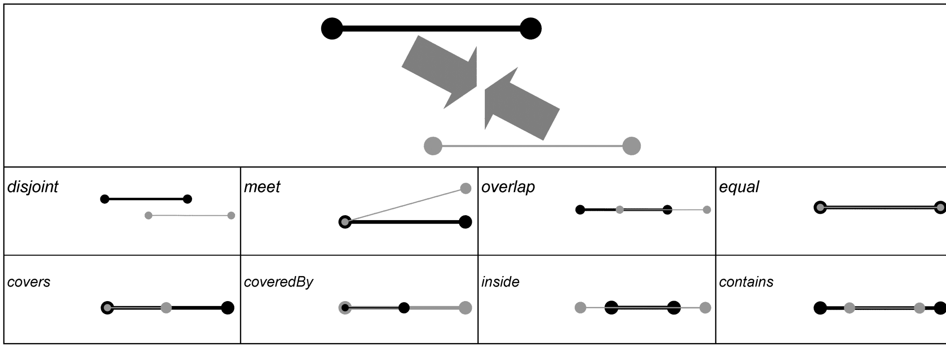

Figure 6 *Geosemantic Proximity* Predicates (where K is a concept and L is a conceptual representation)

According to this approach, work is currently underway to develop a methodology and a computational model of *geosemantic proximity* that will end up with *geosemantic proximity predicates* that are homomorphic with current spatial and temporal topological predicates (Allen 1983, Egenhofer 1993, Clementini and Di Felice 1994, Egenhofer et al. 1994). In this methodology, geospatial concepts and geospatial conceptual representations are compared to segments on a semantic axis made of an interior and a boundary, and *geosemantic proximity* consists essentially of the intersection of their respective contexts. On the one hand, the interior of a concept consists of its intrinsic properties that are components providing literal meaning (e.g. identification, attributes, attribute values, geometries, temporalities, and domain). On the other hand, the boundary of a concept consists of its extrinsic properties that are components providing meaning through relationships with other concepts (e.g. semantic, spatial, and temporal relationships as well as behaviours). Consequently, intersection between intrinsic and extrinsic properties leads to a set of *geosemantic proximity predicates*, as illustrated in Figure 6. In a way similar to human reasoning, *geosemantic proximity* is then assessed qualitatively taking into account the contexts of the respective representations (Kashyap and Sheth 1996).

This notion of *geosemantic proximity* is being elaborated based on ontology (Guarino 1999; Guarino and Welty 2000a, b), *fiat* boundaries (Smith 1994, Casati et al. 1998, Smith and Mark 1999, Smith and Varzi 2000), theories of temporal (Allen 1983) and spatial topology (Egenhofer 1993, Clementini and Di Felice 1994, Egenhofer et al. 1994), context (Kashyap and Sheth 1996, Bishr 1997, Wisse 2000), and semantic similarity (Kashyap and Sheth 1996, Ouksel and Sheth 1999, Rodriguez 2000). However, its detailed description is beyond the scope of this paper, hence it will be addressed in more detail elsewhere.

7 Conclusions

In this article, we have revisited the definitions of geospatial data interoperability and proposed a conceptual framework based on the cognitive and communication sciences. The interpersonal communication process between two agents, including the underlying internal representation of concepts along with encoding and decoding operations, appears to provide a rich framework to better understand the issues involved in geospatial data interoperability, especially when extended to human-to-computer communication and computer-to-computer communication. Central notions involved in this communication-based framework are concept, conceptual representations, ontology, context, and proximity.

The description of the conceptual framework is improved with an ontology of geospatial data interoperability, which is presented in two dimensions: the *five ontological phases of geospatial data interoperability* and *three levels of ontology*. The former describes the different configurations of reality involved in geospatial data interoperability. The latter consists of a subdivision of the different levels of granularity used in the description of real-world phenomena, typically identified by *global ontology*, *domain ontology*, and *application ontology*.

In the proposed framework, two elements characterize the idea of concept. First, the data component is hidden and not directly accessible by other agents. Second, a "simulation" function encapsulates the data component and essentially acts as the main interface for accessing the concept. This simulation function performs the encoding and decoding operation as found in the communication process in order to produce or recognize conceptual representations. It appears to be a fundamental element for the assessment of geospatial data interoperability. In addition, *geosemantic proximity* is a constituent component of the simulation function. Finally, conceptual representations denote physical representations, which serve as mediating components between agents. They are essentially context-dependent, conveying a concept in a specific situation.

In furthering this research, the theoretical model of *geosemantic proximity* will be developed in detail. This will necessitate a formalization of geospatial concepts and conceptual representations that must be aligned with the notion of context. Then a prototype will be designed and implemented as validation and as the experimental phase of this research. The expected results of this research should lead to significant progress concerning the assessment of geospatial data interoperability.

Acknowledgements

The authors wish to acknowledge the contribution of Natural Resources Canada – Centre for Topographic Information in supporting the first author for this research and of the GEOIDE Network of Centres of Excellence in geomatics, project DEC#2 (Designing the Technological Foundations of Spatial Decision-making with the World Wide Web), as well as the contribution of Mike Major for the English editorial review.

References

Allen J F 1983 Maintaining knowledge about temporal intervals. *Communication of the ACM* 26: 832–43

Barsalou L W 1999 Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–609

BC Ministry of Environment Lands and Parks (Geographic Data BC) 1992 *Digital Baseline Mapping at 1:20,000* (Release 2.0). Vancouver, Province of British Columbia, BC Ministry of Environment, Lands and Parks

Bédard Y 1986 A Study of the Nature of Data Using a Communication-based Conceptual Framework of Land Information. Unpublished PhD Dissertation, University of Maine

Bédard Y 1999a Principles of spatial database analysis and design. In Longley P A, Goodchild M F, Maguire D J, and Rhind D W (eds) *Geographical Information Systems: Principles, Techniques, Applications and Management*. New York, John Wiley and Sons: 413–24

Bédard Y 1999b Visual modelling of spatial database towards spatial PVL and UML. *Geomatica* 53: 169–86

Benslimane D 2001 Interopérabilité de SIG: la solution Isis. *Revue internationale de géomatique* 11: 7–42

Bergamaschi S, Castano S, and Vincini M 1999 Semantic integration of semi-structured and structured data sources. *Sigmod Record* 28(1): 54–60

Bishr Y 1997 Semantics Aspects of Interoperable GIS. Enschede, International Institute for Geo-Information Science and Earth Observation (ITC) PhD Dissertation

Bittner T and Edwards G 2001 Toward an ontology for geomatics. *Geomatica* 5: 475–90

Blake R H, and Haroldsen E O 1975 *A Taxonomy of Concepts in Communication*. New York, Hastings House Publishers

Brodeur J, Bédard Y, and Proulx M J 2000 Modelling geospatial application databases using UML-based repositories aligned with international standards in geomatics. In Li K J, Makki K, Pissinou N, and Ravada S (eds) *Proceedings of Eighth ACM Symposium on Advances in Geographic Information Systems (ACMGIS)*, Washington, DC, ACM Press: 39–46

Campbell J 1982 *Grammatical Man: Information, Entropy, Language, and Life*. New York, Simon and Schuster

Canadian Council on Surveying and Mapping 1984 *National Standards for the Exchange of Digital Topographic Data: Topographic Codes and Dictionary of Topographic Features*. Ottawa, Topographical Survey Division, Surveys and Mapping Branch, Energy, Mines and Resources Canada

Casati R, Smith B, and Varzi A C 1998 Ontological tools for geographic representation. In Guarino N (ed) *Formal Ontology in Information Systems*. Amsterdam, IOS Press: 77–85

Charron J 1995 Développement d'un processus de sélection des meilleures sources de données cartographiques pour leur intégration à une base de données à référence spatiale. Mémoire de maîtrise, Université Laval

Cherry C 1978 *On Human Communication: A Review, A Survey, and A Criticism* (Third Edition). Cambridge, MA, MIT Press

Clementini E and Di Felice P 1994 A model for representing topological relationship between complex geometric features in spatial databases. *Information Sciences*: 1–17

Cohen P 1982 Model of cognition: Overview. In Cohen P R and Feigenbaum E A (eds) *The Handbook of Artificial Intelligence*. Stanford, CA, HeirisTech Press: 1–10

Collongues A, Hugues J, and Laroche B 1987 *Merise: Méthode de conception* (2e). Paris, Bordas

Darnell D K 1971 Information theorie. In DeVito J A (ed) *Communication: Concepts and Processes*. Englewood Cliffs, NJ, Prentice-Hall: 37–45

Denes P B, and Pinson F N 1971 The speech chain. In DeVito J A (ed) *Communication: Concepts and Processes*, Englewood Cliffs, NJ, Prentice-Hall: 3–11

Egenhofer M 1993 A model for detailed binary topological relationships. *Geomatica*, 47: 261–273

Egenhofer M, Mark D M, and Herring J R (eds) 1994 The 9-Intersection: Formalism and Its Use for Natural-language Spatial Predicates. Santa Barbara, CA, University of California, National Center for Geographic Information and Analysis Technical Report No 94-1

Egenhofer M J 1999 Introduction: Theory and concepts. In Goodchild M F, Egenhofer M, Fegeas R, and Kottman C (eds) *Interoperating Geographic Information Systems*. Norwell, MA, Kluwer: 1–4

Fowler J, Perry B, Nodine M, and Bargmeyer B 1999 Agent-based semantic interoperability in InfoSleuth. *Sigmod Record* 28: 60–7

Frank A U 2001 Tiers of ontology and consistency constraints in geographic information systems. *International Journal of Geographical Information Science* 15: 667–78

Frankhauser P, Kracker M, and Neuhold E 1991 Semantic versus structural resemblance of classes. *Sigmod Record* 20: 59–63

Frankhauser P, and Neuhold E J 1992 Knowledge based integration of heterogeneous databases. In Hsiao D K, Neuhold E J, and Saks-Davis R (eds) *Proceedings of IFIP WG2.6 Database*

Semantics Conference on Interoperable Database Systems (DS-5)/IFIP Transaction (A-25). Lorne, Australia, Elsevier Science: 155–75

Gibson J J 1979 *The Ecological Approach to Visual Perception*. Boston, MA, Houghton Mifflin

Gruber T R 1993a Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Stanford, CA, Stanford University, Knowledge Systems Laboratory Technical Report No KSL 93-04

Gruber T R 1993b A Translation Approach to Portable Ontology Specification. Stanford, CA, Stanford University, Knowledge Systems Laboratory Technical Report No KSL 92-71

Guarino N 1998 Formal ontology and information systems. In Guarino N (ed) *Proceedings of Formal Ontology in Information Systems (FOIS '98)*. Amsterdam, IOS Press: 3–15

Guarino N 1999 The role of identity conditions in ontology design. In Freksa C and Mark D M (eds) *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*. Berlin, Springer-Verlag Lecture Notes in Computer Science No 1661: 221–34

Guarino N, and Welty C 2000a A formal ontology of properties (Preliminary version). In Dieng R and Corby O (eds) *Proceedings of the Twelfth International Conference on Knowledge Engineering and Knowledge Management: Methods, Models and Tools*. Berlin, Springer-Verlag: 97–122

Guarino N, and Welty C 2000b Identity, unity, and individuation: Towards a formal toolkit for ontological analysis. In Werner H (ed) *Proceedings of ECAI-2000: The European Conference on Artificial Intelligence*. Amsterdam, IOS Press: 219–23

Guha R V 1990 *Micro-theories and Contexts in Cyc. I. Basic Issues*. Austin, TX, Micro-electronics and Computer Technology Corporation

Harvey F 1997 Improving multi-purpose GIS design: Participative design. In Hirtle S C and Frank A U (eds) *Spatial Information Theory: A Theoretical Basis for GIS*. Berlin, Springer-Verlag Lecture Notes in Computer Science No 1329: 313–28

ISO/TC 211 2000 *CD.3 19109 Geographic Information: Rules for Application Schema (N1009)*. Geneva, International Organization for Standards

ISO/TC 211 2001 *ISO/DIS 19110 Geographic Information: Feature Cataloguing Methodology*. Geneva, International Organization for Standards

Kahng J and McLeod D 1998 Dynamic classificational ontologies: Mediation of information sharing in cooperative federated database systems, context and ontologies. In Papazoglou M P and Schlageter G (eds) *Cooperative Information Systems-Trends and Directions*. San Diego, CA, Academic Press: 179–203

Kashyap V and Sheth A 1996 Semantic and schematic similarities between database objects: A context-based approach. *The VLDB Journal* 5: 276–304

Kashyap V and Sheth A 1998 Semantic heterogeneity in global information systems: The role of metadata, context and ontologies. In Papazoglou M P and Schlageter G (eds) *Cooperative Information Systems-Trends and Directions*. San Diego, CA, Academic Press: 139–78

Kettani D, and Moulin B 1999 A spatial model based on the notions of spatial conceptual map and of object's influence areas. In Freska C and Mark D M (eds) *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Systems*. Berlin, Springer-Verlag Lecture Notes in Computer Science No 1661: 401–15

Kosslyn S M 1980 *Image in Mind*. Cambridge, MA, Harvard University Press

Kottman C 1999 The Open GIS Consortium and progress toward interoperability in GIS. In Goodchild M, Egenhofer M, Fegeas R, and Kottman C (eds) *Interoperating Geographic Information Systems*. Norwell, MA, Kluwer: 39–54

Krech D, and Crutchfield R S 1971 Perceiving the world. In Schramm W and Robert D F (eds) *The Process and Effects of Mass Communication*. Champaign-Urbana, IL, University of Illinois Press: 235–64

Laurini R 1998 Spatial multi-database topological continuity and indexing: A step towards seamless GIS data interoperability. *International Journal of Geographic Information Science* 12: 373–402

Lehmann F 1992 Semantic networks. *Computers and Mathematics with Applications* 23: 1–50

Lippmann W 1971 The world outside and the pictures in our heads. In Schramm W and Robert D F (eds) *The Process and Effects of Mass Communication*. Champaign-Urbana, IL, University of Illinois Press: 265–86

Logie R H and Denis M (eds) 1991 *Mental images in Human Cognition*. Amsterdam, North-Holland

McKee L and Buehler K (eds) 1998 *The OpenGIS Guide*. Wayland, MA, OpenGIS Consortium Inc

Natural Resources Canada 1996 *National Topographic Data Base: Standards and Specifications* (Edition 3.1). Sherbrooke, Quebec, Centre for Topographic Information-Sherbrooke

New Brunswick 2000 *Guide d'utilisation de la Base de données topographiques numériques (BDTN) du Nouveau-Brunswick* (1.1.2). Fredericton, NB, Services Nouveau-Brunswick

OBM 1996 *Ontario Digital Topographic Database – 1:10,000, 1:20,000 – A Guide for User*. Toronto, Ontario Ministry of Natural Resources

Ouksel A M and Naiman C 1993 Coordinating context building in heterogeneous information systems. *Journal of Intelligent Information Systems* 3: 151–83

Ouksel A M and Sheth A 1999 Semantic interoperability in global information systems: A brief introduction to the research area and the special section. *Sigmod Record* 28(1): 5–12

Peuquet D, Smith B, and Brogaard B 1998 The ontology of fields. In *Proceedings of the Summer Assembly of the University Consortium for Geographic Information Science*, Bar Harbor, Maine

Pylyshyn Z W 1981 The imagery debate: Analogue media versus tacit knowledge. *Psychological Review* 88: 16–45

Québec 2000 *Base de données topographiques du Québec (BDTQ) à l'échelle de 1/20 000 – Normes de production* (Version 1.0). Québec City, QC, Ministère des Ressources natural Direction générale de l'information géographique

Rodriguez M A 2000 Assessing Semantic Similarity Among Entity Classes. Unpublished PhD Dissertation, University of Maine

Schramm W 1971a How communication works. In DeVito J A (ed)*Communication: Concepts and Processes*. Englewood Cliffs, NJ, Prentice-Hall: 12–21

Schramm W 1971b The nature of communication between humans. In Schramm W and Robert D F (eds) *The Process and Effects of Mass Communication*. Champaign-Urbana, IL, University of Illinois Press: 3–53

Sciore E, Siegel M, and Rosenthal A 1992 Context inter-change using meta-attributes. In Finin T W, Nicholas C K, and Yesha Y (eds) *Proceedings of the First International Conference on Information and Knowledge Management (CIKM)*, Baltimore, Maryland

Sears D O and Freedman J L 1971 Selective exposure to information: A critical review. In Schramm W and Robert D F (eds) *The Process and Effects of Mass Communication*. Champaign-Urbana, IL, University of Illinois Press: 209–34

Sheth A 1999 Changing focus on interoperability in information systems: From systems, syntax, and structure to semantics. In Goodchild M, Egenhofer M, Fegeas R, and Kottman C (eds) *Interoperating Geographic Information Systems*. Norwell, MA, Kluwer: 5–29

Sheth A and Kashyap V 1992 So far (schematically) yet so near (semantically). In Hsiao D K, Neuhold E J, and Saks-Davis R (eds) *Proceedings of IFIP WG2.6 Database Semantics Conference on Interoperable Database Systems (DS-5)*. Lorne, Australia, Elsevier: 283–312

Simsion G C 2001 *Data Modeling Essentials: Analysis, Design, and Innovation* (2nd Edition) Scottsdale, AR, Coriolis

Smith B 1994 Fiat objects. In Guarino N, Vieu L, and Pribbenow S (eds) *Proceedings of Parts and Wholes: Conceptual Part-Whole Relations and Formal Mereology, Eleventh European Conference on Artificial Intelligence*. Amsterdam: 15–23

Smith B and Mark D 1999 Ontology with human subjects testing: An empirical investigation of geographic categories. *American Journal of Economics and Sociology* 58: 245–72

Smith B and Varzi A C 2000 Fiat and bona fide boundaries. *Philosophy and Phenomenological Research* 60: 401–20

Smith K and Obrst L 1999 Unpacking the semantics of source and usage to perform semantic reconciliation in large-scale information systems. *Sigmod Record* 28(1): 26–31

Sondheim M, Gardels K, and Buehler K 1999 GIS interoperability. In Longley P A, Goodchild M F, Maguire D J, and Rhind D W (eds) *Geographical Information Systems: Principles, Techniques, Applications and Management*, New York, John Wiley and Sons: 347–58

Sowa J F 1984 *Conceptual Structures: Information Processing in Mind and Machine*. Reading, MA, Addision-Wesley

Statistics Canada 1997 *Digital Boundary File and Digital Cartographic File 1996 Census (Reference Guide)*. Ottawa, Minister of Industry

Sycara K, Klusch M, Widoff S, and Lu J 1999 Dynamic service matchmaking among agents in open information environments. *Sigmod Record* 28(1): 47–53

VMap 1995 Vector Map (VMap): Level 1, Mil-V-89033. Bethesda, MD, U.S. National Imagery and Mapping Agency

Weiner N 1950 *The Human Use of Human Beings: Cybernetics and Society*. Boston, MA, Houghton and Mifflin

Wisse P 2000 *Metapattern: Context and Time in Information Models*. Reading, MA, Addison-Wesley