

A Spatio-Temporal Ontology for Geographic Information Integration

Thomas Bittner^{*,1,2,3,4}, Maureen Donnelly^{1,3}, and Barry Smith^{1,4}

¹Department of Philosophy, ²Department of Geography,

³New York State Center of Excellence in Bioinformatics and Life Sciences

⁴National Center for Geographic Information and Analysis (NCGIA)
State University of New York at Buffalo

(Received 00 Month 200x; in final form 00 Month 200x)

This paper presents an axiomatic formalisation of a theory of top-level relations between three categories of entities: individuals, universals, and collections. We deal with a variety of relations between entities in these categories, including the sub-universal relation among universals and the parthood relation among individuals, as well as cross-categorical relations such as instantiation and membership. We show that an adequate understanding of the formal properties of such relations – in particular their behavior with respect to time – is critical for geographic information processing.

The axiomatic theory is developed using *Isabelle*, a computational system for implementing logical formalisms. All proofs are computer-verified and the computational representation of the theory is available online.

Keywords: Geospatial ontology, spatio-temporal ontology, qualitative spatio-temporal reasoning, interoperability, axiomatic theories

1 Introduction

Geographic information-processing today faces two major problems: barriers to information integration caused by incompatible terminology systems, and a lack of interoperability among the available software systems (Bishr 1998, Visser *et al.* 2001, Fonseca *et al.* 2002a, Winter 2001, Fonseca *et al.* 2002b, Kuhn 2003, Duckham and Worboys 2005, Mancarella *et al.* 2004). The needed standardization on both levels is still far from being achieved (Winter and Nittel 2003, Agarwal 2005, Abdelmoty *et al.* 2005).

Different and often incompatible terminologies are used by different disciplinary and professional groups for expressing spatial information and for structuring spatial data. It is often the case that the same term has different meanings in different terminologies or that distinct terms in different terminologies have the same or overlapping meanings (Bishr 1998, Visser *et al.* 2001, Mancarella *et al.* 2004, Duckham and Worboys 2005). *Ontologies* are controlled, structured terminologies for which a semantics is provided in a well defined and unambiguous manner (Gruber 1993, Guarino 1998). Ontologies can thus be used, to overcome at least some of the problems caused by these various forms of semantic heterogeneity.

In this paper we focus on how to employ *logic-based top-level* ontologies for this purpose. Top-level ontologies specify the meanings of terms denoting the basic kinds of entity (such as ‘endurant’, ‘individual’, ‘universal’) as well as the basic kinds of relation such as ‘identical-to’, ‘part-of’, ‘connected-to’, ‘extension-of’, ‘member-of’, ‘instance-of’, ‘sub-universal-of’ etc. These are very general terms, which are needed in almost every domain. They are used to structure information and to define domain-specific terminology in domains as disparate as medicine, biology, and politics as well as in geo-spatial disciplines such as hydrology and environmental science (Smith 2003).

A logic-based ontology is a formalized theory (Copi 1979) consisting of axioms, definitions, and theorems. The terms of the terminology whose semantics is to be specified appear as names, predicate and relation symbols in the formal language of the logical theory. Logical axioms and definitions are then added to

*Corresponding author. Email: bittner3@buffalo.edu

express relationships between the entities and relations denoted by those symbols. Theorems are logical consequences of the axioms and definitions and make explicit additional logical properties of relations and interrelations between relations. Explicit knowledge about logical properties and interrelations between relations is fundamental to the implementation of automated reasoning. For example, it is well-known that most automated reasoning about binary topological relations (relations denoted by top-level terms such as ‘connected to’, ‘overlap’, ‘externally connected to’, ‘disconnected’, etc.) (Randell *et al.* 1992, Egenhofer and Franzosa 1991) is based on composition tables (Egenhofer 1991, Cohn and Hazarika 2001). Composition tables, however are just compact representations of the axioms, definitions, and theorems of some underlying logical theory.

This paper is structured as follows. We start by giving an informal but systematic account of basic categories of enduring entities such as mountains, planets, organisms, ecoregions, socio-economic units, and the basic kinds of top-level relations between them (parthood relations, sub-universal relations, etc.). We discuss the relevance of these categories of entities and relations to geographic information science and geographic information processing and we show how a more exact understanding helps to overcome problems in information processing which arise where different groups need to exchange and integrate information that is expressed in semantically heterogeneous ways. We also emphasize the specific temporal properties of the entities in the mentioned categories and the time-dependent character of certain types of the top-level relations.

We then develop a formal ontology that specifies the meanings of the top-level terms using an axiomatic theory which is a modified and extended version of the theory presented in (Bittner *et al.* 2004b). We then discuss the computational representation of the presented theory. We conclude with a discussion of the related literature and some ongoing research.

2 Top-level categories of entities and relations

In this paper we focus on independent endurants, entities which survive self-identically through time while undergoing changes of various sorts.¹ Following Bittner *et al.* (2004b) we distinguish three basic categories of entities in our treatment of independent endurants: (i) individual endurants (Napoleon, Napoleon’s heart, New York City, New York State, the Planet Earth); (ii) endurant universals (kinds, types or classes) (*human being*, *heart*, *human settlement*, *socio-economic unit*); and (iii) collections of individual endurants (the collection of counties in New York State, the collection of cells currently in your body, the collection of all human beings existing at a given time).

Note that individuals, universals, and collections have, according to the theory here advanced, different temporal properties. Individuals can gain and lose parts. For example, organisms gain and lose cells, a city gains and loses parts every time a building is built or demolished, the continent of North America gains and loses parts due to erosion processes, volcanic eruptions, etc. Universals gain and lose instances. For example, the universal *human being* gains or loses instances every time a person is born or dies. The universal *socio-economic unit* may gain or lose instances due to administrative reforms. Collections, on the other hand, are like sets in the mathematical sense in that they are identified through their members. Thus collections in our sense of this term cannot have different members at different times. For example, the collection of cells constituting your body at 1 pm on April 6 in 1999 is fixed. It cannot gain or lose members. As soon as a cell in this collection ceases to exist and/or a new cell is created in your body, your body is constituted by a *different* collection of cells.

In general many of the properties and relations treated in the remainder of this paper are *time-dependent*. Consider, for example, parthood and location relations. The Czech Republic was not part of the European Union in 2001 but it is part of the European Union in 2004. Similarly, the Auenwald in Leipzig was located in a singly connected spatial region 100 years ago. Today it consists of multiple disconnected patches of forest. This means that in an ontology of enduring entities it is insufficient to say of endurants x and y that a certain top-level relation holds between them, because the relations which hold will be different at

¹For discussions of perdurants (processes) and dependent endurants (qualities, roles, etc.) see Simons (1987), Sider (2001), Grenon and Smith (2004), Bittner *et al.* (2004a), Galton and Worboys (2005), Grenon and Smith (2004), Smith and Grenon (2004).

different points in time (Simons 1987, Lowe 2002). To take this time-dependent character of relations into account we include time instants as a fourth category in our ontology and add a temporal parameter to time-dependent relations (Thomson 1983, Simons 1987). For example, to take the time-dependent character of the parthood and location relations into account, we say that x is a part of y at time t_1 but x is no longer part of y at time t_2 , or that x is located in y at t_1 but is no longer located in y at t_2 .

Endurant individuals are tied to universals through the instantiation relation. For example, New York State is an instance of the universal *socio-economic unit*; the City of Buffalo is an instance of the universal *city*; I am an instance of the universal *human being*. The instantiation relation for enduring entities is time-dependent. For example, it is the case that I am an instance of the universal *adult* now, but it was not the case that I was an instance of this universal in 1966. Similarly, the City of Buffalo is an instance of *city* today but was not an instance of *city* in 1802, when it instantiated the universal *village*.

Certain collections are tied to universals through the relation extension-of. This relation too is time-dependent. The extension of a universal at a given time is the collection of individuals which instantiate the universal at that time. Not every universal has an extension at every time, as is shown by the case of extinct species. Moreover, not every collection is the extension of a universal. For example, the collection of the cells in your body now is not the extension of any universal now. Rather it is a proper sub-collection of the extension of the universal *cell*. Similarly, we hold that the collection of human beings in the City of Buffalo now is a proper sub-collection of the extension of the universal *human being* now. But there is no universal human-being-in-Buffalo. We will discuss this in more detail in Section 5.

Some collections consist of disjoint parts of some other individual, which jointly sum up to this individual at a given time. We call such collections *partitions* of the individual in question. For example the collection which includes as its members the US states and the District of Columbia in 2006 partitions the USA in 2006. The collection of counties of New York State in 2006 partitions New York State in 2006. The mentioned partitions consist of fiat parts (Smith and Varzi 2000), as also does the collection of US postal districts, or the collection formed by the northern and southern hemispheres of the Earth.

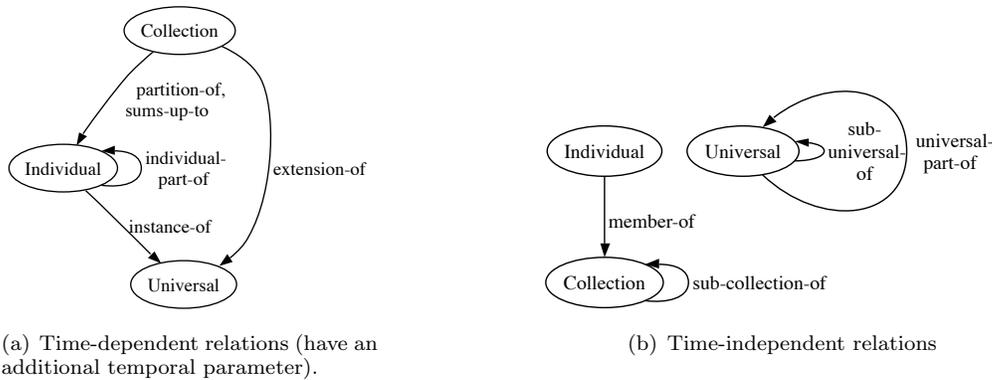


Figure 1. Top-level relations between the three basic categories of individuals, universals, and collections. Relations are represented by labeled arrows.

Given the three top-level categories of individuals, universals, and collections, we can distinguish the top-level relations according to the kinds of entities they relate as depicted in Figure 1. In Figure 1(a) *time-dependent* and in Figure 1(b) *time-independent* top-level relations are depicted. We will develop the formal theory in a modular fashion along the lines depicted in Figure 2. An arrow in the figure indicates that every axiom of the sub-theory developed in the section designated by the starting node is also an axiom of the sub-theory developed in the section designated by the end node.

In the remainder of this section we discuss two examples which demonstrate that the consistent usage of the terms depicted in Figure 1 helps to structure and exchange information and to improve the precision of definitions used in scientific discourse. The examples will provide further motivation for the formal ontology to be presented in the second part of this paper.

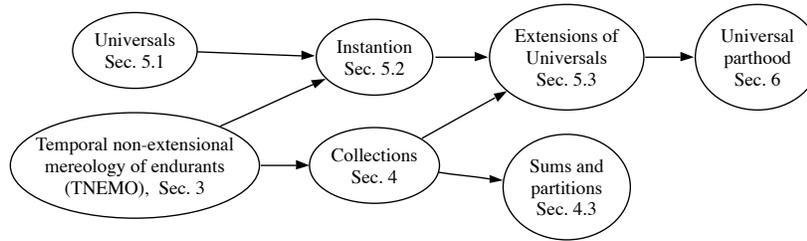


Figure 2. The hierarchical structure of the top-level ontology. An arrow from T_1 to T_2 means that every axiom of sub-theory T_1 is also an axiom of sub-theory T_2 .

2.1 Land use classification (Example 1)

As our first case study, we consider the following problem, which is an example of the sort of problem that arises within the planning process in Germany, where spatial data provided by the German government is classified according to the ATKIS-OK-250 terminology system. The problem is that this data needs to be integrated with other data, provided by the European Community, that is classified according to the CORINE land cover terminology system (Visser *et al.* 2001). To integrate these different data sets, we need to establish semantic relations between the terms in the ATKIS and CORINE systems.

In this context we need to distinguish (i) *domain specific* technical terms informally defined according to the respective standards and (ii) terms used in the definitions of the standards and in the data sets whose meaning needs to be defined in a domain-independent top-level ontology. In our examples we use the `typewriter` font for domain-specific technical terms from the ATKIS terminology and the `Sans Serif` font for domain-specific technical terms from the CORINE terminology. We use `SMALL CAPITAL LETTERS` to signify the terms whose semantics is defined in our top-level ontology. We underline terms, including relation terms, whose semantics is specified through a mapping to a term in our top-level ontology.

Assume a data set from the year 2000 with an entry expressed according to the ATKIS standard as the statement:

(A) Stadtwald-1 is a forest.

Assume that we have a second data set from 2000 derived from a satellite image classified according to the CORINE classification system. This dataset has an entry that, according to the CORINE standard, can be expressed as the statement:

(B) Stadtwald-2 is classified as a mixed forest.

Assume further that, according to the land cover classification of the CORINE standard, the relation between mixed forest and forest can be expressed as the statement:

(C) mixed forest is subsumed by forest.

Now suppose that we need to verify that the entity named ‘Stadtwald-1’ is classified correctly according to ATKIS by comparing the classification of what is putatively the same entity in the CORINE dataset. (A), (B), and (C) will imply that Stadtwald-1 is classified in the same way (and thus probably correctly) according to both ATKIS and CORINE under the following assumptions:

- (a0) ‘Stadtwald-1’ and ‘Stadtwald-2’ name two (possibly one) entities that overlap, i.e., share a common INDIVIDUAL PART.
- (a1) The terms **forest** and **forest** refer to the same UNIVERSAL: *forest*.
- (a2) The term **mixed forest** refers to the UNIVERSAL *mixed forest*.
- (a3) The phrase is a in (A) and the phrase is classified as in (B) both mean INSTANCE-OF – a relation that holds between an INDIVIDUAL and a UNIVERSAL (as in: I am an INSTANCE-OF the universal *human being*).
- (a4) The phrase is subsumed by in (C) refers to the relation SUBUNIVERSAL-OF.
- (a5) The relations INSTANCE-OF and SUBUNIVERSAL-OF have the following logical interrelationship that can be exploited for reasoning purposes: If *c* is a SUBUNIVERSAL-OF *d* and *x* is an INSTANCE-OF *c* then *x* is an INSTANCE-OF *d*.
- (a6) If two INSTANCES OF the UNIVERSAL *forest* overlap (i.e., share a common PART) then they are IDENTICAL.

The reasoning goes as follows:

Stadtwald-1 is an INSTANCE-OF *forest* (A, a1, a3). Stadtwald-2 is an INSTANCE-OF *mixed forest* (B, a1, a3). Since *mixed forest* is a SUBUNIVERSAL-OF *forest* (C, a4) it follows that every INSTANCE-OF *mixed forest* is an INSTANCE-OF *forest* (a5). Thus, in particular, Stadtwald-2 is an INSTANCE-OF *forest*. Since Stadtwald-1 and Stadtwald-2 OVERLAP (a0) and are both INSTANCES OF *forest*, it follows that Stadtwald-1 and Stadtwald-2 are identical (a6) and are classified in the same way, as *forest*, according to both ATKIS and CORINE. (For a formal proof see Figure 6 on page 24 of this paper.) (1)

Notice that we can justify that (a0) is a reasonable assumption by pointing to the overlap of the geo-referenced locations of the individuals Stadtwald-1 and Stadtwald-2 on standard maps. This is a common GIS practice. (a1–a2) are assumptions that need to be justified at the domain level based on formal definitions of the universals *forest* and *mixed forest*. (a3–a6) are assumptions that need to be justified by the axioms, definitions, and theorems of the underlying top-level ontology.

This example and in particular the assumptions (a0–a6) which we needed in order to derive the desired conclusions, show that it is important to have formal means:

- (i) to decide whether two individuals or two universals are identical independently of their name;
- (ii) to exactly specify the semantics of top-level terms used to refer to the relations that hold between and among individuals and universals. This includes the specification of logical properties and interrelationships that can be exploited for reasoning purposes.

In this paper we focus on (ii) by providing a collection of top-level terms with a well defined semantics. Those terms can then be used to specify the meaning of terms like those used in (A–C) and link them as demonstrated in (a3–a5). We will partly address (i) by giving general identity conditions for individuals, universals, and collections in our axiomatic theory. Usually, (i) will be addressed in domain ontologies which provide necessary and sufficient conditions for identifying and distinguishing domain specific universals of a given domain and the individuals that instantiate those universals.

Notice that our technique is quite different from other approaches which ignore the distinctions between top-level and domain-specific terms and which are based on direct mappings between domain-specific terminologies using semantic similarity measures, e.g., (Fonseca *et al.* 2000, 2002b,a, Kuhn 2003, Rodríguez and Egenhofer 2003, 2004, Yetongnon *et al.* 2006). We do not reject such approaches; we are however confident that the quality of their achieved results will be enhanced through greater rigor of the sort presented here.

2.2 Universal vs. instance level relations (Example 2)

In this example we assume that domain-specific terms such as ‘(land-)property’, ‘east’, ‘river’, ‘road’, ‘ford’, ‘Althener ford’, ‘Parthe River’ etc. have a well defined semantics specified in some domain ontology such

as the Spatial Data Standard (SDTS) (SDTS 1997). Thus here we focus on those underlined terms in the sentences (D-H) which will acquire their meanings through mappings to terms in our top-level ontology.

In the Spatial Data Standard (SDTS 1997) we find sentences such as

- (G) A ford is a shallow part of a river which can be easily crossed.
- (H) [Headwater is] the upper part of a river system ...

A data set is a representation of facts about geographic phenomena that also can be expressed by means of sentences like:

- (D) The eastern part of the Stadtwald is separated from its western part by a road.
- (E) The northern part of the property is cultivated while the southern part is kept uncultivated.
- (F) The Althener ford is part of the Parthe River.

(G) informally describes the meaning of the term ‘ford’ as it is used in (F). (H) informally describes the meaning of the term ‘headwater’. While (G) expresses a general relationship between fords and rivers, (F) expresses a fact about the relation between a particular ford and a particular river.

In (D-F) ‘part of’ refers to the parthood relation that holds between individual entities (parts of the Stadtwald, parts of this piece of land, parts of that river). We will refer to this relation as INDIVIDUAL-PART-OF (Smith and Rosse 2004, Donnelly and Bittner 2005, Smith *et al.* 2005, Donnelly *et al.* 2006). The intended interpretation of ‘part of’ in (G) and (H) is a relation that holds between universals. The STDS does not talk about parthood relations between a specific river and a specific ford. Similarly, statement (H) is not about the parthood relation between a particular headwater and a particular river system.

The intended interpretation of ‘part of’ in (H) is a relation R_H that holds between universals such as *headwater* and *river system* if and only if at every time t (i) every instance of the first universal (i.e., *headwater*) is an individual part of some instance of the second universal (i.e., *river system*) and (ii) every instance of the second universal (i.e., *river system*) has an instance of the first universal (i.e., *headwater*) as an individual part.

By contrast, the intended interpretation of ‘part of’ in (G) is a relation R_G that holds between the universals *ford* and *river* if and only if every instance of the universal *ford* is an individual part of some instance of the universal *river*. It is important to note that the intended meaning of ‘part of’ in (G) allows for the fact that there are rivers that do not have fords. Consequently it does not hold that every instance of the second argument (i.e., *river*) has an instance of the first argument (i.e., *ford*) as its individual part. Hence the relations R_H and R_G are truly distinct: condition (ii) of the definition of R_H is not satisfied for the relation used in (G). The relations R_H and R_G are structurally similar but distinct, i.e. the former is a sub-relation of the latter.

Another example of a sentence in which the intended interpretation of ‘part of’ is R_G is:

- (I) A waterfall is part of a watercourse.

Every waterfall is an individual part of some watercourse, but not every watercourse has a waterfall as an individual part.

Now consider the sentence:

- (J) A wall is part of a building.

The intended interpretation of ‘part of’ in (J) is a relation R_J that holds between two universals (e.g., *wall* and *building*) if and only if at every time t every instance of the second universal (e.g., *building*) has an instance of the first universal (e.g., *wall*) as an individual part. Neither R_H nor R_G are meant by ‘part of’ as used in (J), since not every instance of the first universal (*wall*) is an individual part of some instance of the second universal (*building*). For example, the Great Wall of China is a wall which is not an individual part of any building. Consequently we have $R_J(\textit{wall}, \textit{building})$ but not $R_H(\textit{wall}, \textit{building})$ and not $R_G(\textit{wall}, \textit{building})$.

Another example of a sentence in which the intended interpretation of ‘part of’ is the relation R_J is sentence (K):

(K) A tree is part of a forest.

Every instance of *forest* has some instance of *tree* as individual part but not every instance of *tree* is individual part of some instance of *forest*.

The relations R_H , R_G , and R_J belong to the group of *universal-level* parthood relations referred to by UNIVERSAL-PART-OF (Figure 1(b)).

3 Temporal mereology of individuals

In the second part of this paper we give an axiomatic characterisation of the relations depicted in Figure 1. We present the theory in a sorted first-order predicate logic with identity. We use the letters t, t_1, t_2, \dots as variables ranging over instants of time; w, x, x_1, y, z, \dots as variables ranging over independent enduring individuals; c, d, e, g as variables ranging over universals; and p, q, r, p_1, \dots as variables ranging over collections. The logical connectors $\neg, =, \wedge, \vee, \rightarrow, \leftrightarrow$ have their usual meanings (not, identical-to, and, or, if ... then, and if and only if (iff), respectively). We use the symbol \equiv for definitions. We write (x) to symbolise universal quantification (for all $x \dots$) and $(\exists x)$ to symbolise existential quantification (there is at least one $x \dots$). (See (Copi 1979) for an introduction.) All quantification is restricted to a single sort. Restrictions on quantification will be understood by conventions on variable usage. Leading universal quantifiers are omitted. Labels for axioms begin with ‘A’, labels for theorems begin with ‘T’, and labels for definitions begin with ‘D’.

We here develop a temporal version of mereology (Simons 1987) based on the ternary primitive \leq , an abbreviation for INDIVIDUAL-PART-OF. On the intended interpretation $x \leq_t y$ means: the individual enduring x is part of the individual enduring y at time-instant t . For example the Czech Republic is part of the European Union in 2007, but was not a part of the European Union in 2001. Similarly we can express the statement that this blood cell was part of my body yesterday, but is not a part of my body now. At all times at which I exist, I am a part (but not a proper part) of myself.

We now add two relation symbols, $<$ and O , whose intended interpretations are (instance-level) proper parthood and overlap among enduring individuals respectively. We then specify the meaning of the symbols $<$ and O relative to \leq by means of the definitions $D_{<}$ and D_O .

$$D_{<} \quad x <_t y \equiv x \leq_t y \wedge \neg y \leq_t x \qquad D_O \quad O \ xyt \equiv (\exists z)(z \leq_t x \wedge z \leq_t y)$$

The individual x is a proper part of the individual y at time instant t if and only if x is a part of y at t and y is not a part of x at t ($D_{<}$). The individual x overlaps the individual y at time t if and only if there exists an individual z such that z is a part of x at t , and z is a part of y at t (D_O). At this time Montana is a proper part of (and overlaps) the United States. Yellowstone National Park (YNP) overlaps Wyoming, Montana, and Idaho.

We introduce the symbol ‘ E ’ where the intended interpretation of ‘ $E \ x t$ ’ is: x exists at time t . Formally we define E in terms of \leq : $E \ x t$ holds if and only if x is a part of itself at t (D_E). Clearly, only an entity that exists at a given time can be a part of itself at that time. We also introduce the symbol ‘ \sim ’ where ‘ $x \sim_t y$ ’ is interpreted as meaning: x and y are *mereologically equivalent* at t . Intuitively, two entities are mereologically equivalent at time t if and only if they have exactly the same parts at t . (In a formal ontology that also includes the notion of location (Casati and Varzi 1995) we could say that mereologically equivalent entities *coincide*, i.e., occupy exactly the same space.) For example, at this time the City of Vienna and the Austrian Federal State Vienna are mereologically equivalent: they have exactly the same parts and so occupy the same region of space. However, they are clearly distinct in their non-mereological properties. For example, the City of Vienna, but not the Federal State of Vienna, is governed by a mayor.

Formally we define that $x \sim_t y$ holds if and only if x is an individual part of y at t and y is a part of x at t (D_{\sim}).

$$D_E \quad E xt \equiv x \leq_t x \qquad D_{\sim} \quad x \sim_t y \equiv x \leq_t y \wedge y \leq_t x$$

The water mass that constitutes Lake Erie at time t is mereologically equivalent to Lake Erie at time t . They coincide at t and have exactly the same parts at t . Notice, however, that the water mass that constitutes Lake Erie at time t and Lake Erie are distinct objects. This can be seen by considering their mereological structure at different times: In 10 years time Lake Erie will be constituted by a different water mass. Only few (if any) of the H_2O molecules that were parts of the water mass 10 years ago will still be part of the Lake.

The definitions $D_{<}$, D_O , and D_E specify the meanings of the symbols ' $<$ ', ' O ', and ' E ' in terms of ' \leq '. The meanings of the former depend on the meaning of the latter. The symbol ' \leq ', however, is *primitive*, which means that it is not defined in terms of other symbols. Its meaning is specified, rather, by means of axioms, which explicate properties of the parthood relation.

We add the following axioms including: every individual exists at some time ($AM1$); if x is a part of y at t then x and y exist at t ($AM2$); for a fixed t \leq_t is transitive ($AM3$)¹; if x exists at t and x is not part of y at t then there exists a z such that z is part of x at t and z and y do not overlap at t ($AM4$).

$$\begin{array}{ll} AM1 & (\exists t) E xt \\ AM2 & x \leq_t y \rightarrow E xt \wedge E yt \\ AM3 & x \leq_t y \wedge y \leq_t z \rightarrow x \leq_t z \\ AM4 & E xt \wedge \neg x \leq_t y \rightarrow (\exists z)(z \leq_t x \wedge \neg O zyt) \end{array}$$

$AM1$ amounts to requiring that every object exists at some time (thus we exclude objects that never exist). It does not imply, however, that every object needs to exist at all times. $AM2$ asserts that parthood can hold at time t only between objects that exist at time t . $AM3$ tells us that, for example, if at time t New York City is part of New York State and New York State is part of the United States, then it follows that New York City is part of the United States at this same time. $AM4$ tells us for example that if Montana is not part of New York State, then there exists at least one part of Montana that does not overlap New York State.

Notice that the fact that these examples seem to be so trivial to human beings tells us that axioms $AM1$ – $AM4$ capture some important aspect of the spatial domain that we humans take for granted. They are trivially true of the spatial domain, and it is therefore important that they are captured in a formal ontology that is designed to support automated reasoning in this domain. For computers, all truths however trivial, need to be made explicit.

From the definitions and axioms the following can be proved as theorems $TM1$ – $TM6$ ²: If x is a proper part of y at t , then y is not a proper part of x at t ($TM1$); at no time is x a proper part of itself ($TM2$); if x is a part of y at t then x overlaps y at t ($TM3$); overlap is symmetric, i.e. if x overlaps y at t then y overlaps x at t ($TM4$); if x is part of y at t and x overlaps z at t then y overlaps z at t ($TM5$); and if x is a part of y at t and y and z do not overlap at t then x does not overlap z at t ($TM6$).

$$\begin{array}{ll} TM1 & x <_t y \rightarrow \neg y <_t x \\ TM2 & \neg x <_t x \\ TM3 & x \leq_t y \rightarrow O xyt \\ TM4 & O xyt \rightarrow O yxt \\ TM5 & x \leq_t y \wedge O xzt \rightarrow O yzt \\ TM6 & x \leq_t y \wedge \neg O yzt \rightarrow \neg O xzt \end{array}$$

We list these theorems as examples of how to make explicit the consequences of definitions and other assumptions using the deductive power of formal logic. Again, to a human being these theorems will seem to trivial. Consider the following examples (where time is assumed fixed): since Yellowstone National Park (YNP) overlaps Wyoming, Wyoming also overlaps YNP (by $TM4$). Wyoming is a part of the USA and

¹Notice that parthood in this most general sense is transitive (Simons 1987, Varzi 1996). There are however more specific parthood relations, for example, part-of-the-same-scale (Bittner and Donnelly 2006) or functional-part-of, that are not transitive (Varzi 2006).

²All theorems are computer verified. For details see Section 7.

YNP overlaps Wyoming. Therefore Wyoming overlaps the USA (by *TM5*). Wyoming is part of the USA and the USA does not overlap Germany. Therefore Wyoming does not overlap Germany (by *TM6*).

The properties of the relations that are made explicit by means of our axioms, definitions, and theorems can also be used to support automated reasoning about information stored, for example, in Geographic Information Systems. The prototypical cases include transitivity inference rules of the form

$$\begin{aligned} TIR & \text{ from } R(a, b) \text{ and } R(b, c), \text{ derive } R(a, c); \\ TIR_t & \text{ from } S(a, b, t) \text{ and } S(b, c, t), \text{ derive } S(a, c, t). \end{aligned}$$

Here ‘ $R(a, b)$ ’, ‘ $R(b, c)$ ’, ‘ $S(a, b, t)$ ’, ‘ $S(b, c, t)$ ’, etc. represent data entries in which the symbol ‘ R ’ is a meta-variable which can be replaced by the name of any time-independent transitive relation and ‘ S ’ is a meta-variable which can be replaced by the name of any time-dependent transitive relation (including ‘ \leq ’, ‘ $<$ ’, etc).

It is an important advantage of a logic-based ontology that it makes explicit which relations have the property of being transitive and thus can be used to support automated reasoning by means of such transitivity inference rules. Thus *AM3* tells us that a computer program can validly use the transitivity inference rule to derive ‘New York City is part of the USA’ from ‘New York City is part of New York State’ and ‘New York State is part of the USA’. Consequently, no GIS needs explicitly to store the information ‘New York City is part of the USA’ in addition to ‘New York City is part of New York State’ and ‘New York State is part of the USA’ since it can be derived automatically.

Notice that many relations are not transitive. Thus while Montana overlaps YNP and YNP overlaps Wyoming, Montana does not overlap Wyoming. Thus it is important to specify also which relations do not have the property of transitivity.

Some logical consequences of our axioms and definitions are as follows. From (D_{\sim}) it immediately follows that if x is mereologically equivalent to y then y is mereologically equivalent to x , i.e. \sim is symmetric (*TM7*). We can also prove that x is mereologically equivalent to itself at all times at which x exists (*TM8*) and that \sim is transitive for fixed times (*TM9*). Thus, at a given time t , \sim is an equivalence relation on the sub-domain of enduring individuals that exist at t .

$$\begin{aligned} TM7 & \quad x \sim_t y \rightarrow y \sim_t x & TM8 & \quad E xt \rightarrow x \sim_t x \\ TM9 & & & \quad x \sim_t y \wedge y \sim_t z \rightarrow x \sim_t z \end{aligned}$$

We can also prove: if x exists at t and everything that overlaps x at t overlaps y at t , then x is a part of y at t (*TM10*); if x and y exist at t , then x and y are mereologically equivalent at t if and only if every z overlaps x at t if and only if z overlaps y at t (*TM11*); if x and y exist at t , then x and y are mereologically equivalent at t if and only if every z is part of x at t if and only if z is part of y at t (*TM12*), i.e. two individuals are mereologically equivalent at t if and only if they have the same parts at t .

$$\begin{aligned} TM10 & \quad E xt \wedge (z)(O zxt \rightarrow O zyt) \rightarrow x \leq_t y \\ TM11 & \quad E xt \wedge E yt \rightarrow ((z)(O zxt \leftrightarrow O zyt) \leftrightarrow x \sim_t y) \\ TM12 & \quad E xt \wedge E yt \rightarrow ((z)(P zxt \leftrightarrow P zyt) \leftrightarrow x \sim_t y) \end{aligned}$$

Notice, that it does not follow from our axioms that if two individuals are parts of each other at a given time then they are identical (\leq is not antisymmetric). Nor does it follow that if two individuals overlap exactly the same things at a given time then they are identical (thus O is not extensional (Simons 1987)). Thus we allow for the possibility that two distinct individuals can have exactly the same parts at a given time. Again, the City of Vienna and the Austrian Federal State of Vienna have exactly the same parts, even though they are distinct.

We call the theory formed by *AM1–4 non-extensional temporal mereology* or ‘*TNEMO*’ for short. *TNEMO* specifies ‘ \leq ’ as meaning individual-part-of and it specifies the defined terms ‘ $<$ ’, ‘ O ’, ‘ E ’, and ‘ \sim ’ as meaning: individual-proper-part-of, overlaps, exists, and mereologically-equivalent respectively. Notice that further axioms may be needed to better approximate the meaning of ‘ \leq ’. See for example Varzi (1996)

for additional (atemporal) axioms.

TNEMO is distinct from the extensional temporal mereology presented in (Bittner *et al.* 2004b). Both theories share the definitions D_E and D_O and the standard axioms of transitivity (AM3) and strong supplementation (AM4). But because TNEMO has no axiom of antisymmetry we are able to introduce the predicate of mereological equivalence, which enables us to deal with coinciding but distinct individual endurants (which exist, intuitively, on different geospatial layers). We discuss these issues in more detail in (Bittner and Donnelly 2007d).

4 Collections

Collections of individual endurants are the second major category of entities in our ontology. Examples of collections include: the collection of Hispanic people in Buffalo's West Side as specified in the 2000 census records; the collection of people with an annual income higher than \$30,000 in a given postal district at a given time, and so on. In particular, we will consider two special sorts of collections: those that form partitions of individuals at given times and those that are the extensions of universals at given times.

We use the letters p, q, r as variables ranging over collections and we use ' \in ' to stand for the membership relation between individuals and collections. We use the notation ' $\{x_1, \dots, x_n\}$ ' to refer to a finite collection having exactly x_1, \dots, x_n as members. Since the disjoint categories of collections and individuals are represented by disjoint sorts in our theory, it follows that the relation \in is irreflexive and asymmetric, and that there are no collections of collections.

Collections comprehend in every case two or more individuals (AC1) (Burge 1977). Consequently there are no empty collections and no singleton collections. (Compare (Bunt 1985).) We require that two collections are identical if and only if they have the same members (AC2). This makes explicit the extensional character of collections.¹

$$AC1 \quad (\exists x)(\exists y)(x \in p \wedge y \in p \wedge \neg x = y) \qquad AC2 \quad p = q \leftrightarrow (x)(x \in p \leftrightarrow x \in q)$$

The collection p is a *sub-collection* of the collection q ($p \subseteq q$) if and only if every member of p is also a member of q (D_{\subseteq}).

$$D_{\subseteq} \quad p \subseteq q \equiv (x)(x \in p \rightarrow x \in q) \qquad \begin{array}{ll} TC1 & p \subseteq p \\ TC2 & p \subseteq q \wedge q \subseteq p \rightarrow p = q \\ TC3 & p \subseteq q \wedge q \subseteq r \rightarrow p \subseteq r \end{array}$$

We can prove that \subseteq is reflexive (TC1), antisymmetric (TC2), and transitive (TC3). Thus, \subseteq is a partial ordering. TC3 tells us that an automated reasoning system may validly apply the transitivity rule (TIR) to reason about data containing information about sub-collection relations among collections. TM4 tells us that an automated reasoner is permitted to infer that if two collections are sub-collections of each other then they are identical.

4.1 Fully, partly present, and non-present collections

Our treatment of collections makes them in some respects analogous to sets in the mathematical sense (Copi 1979). Thus collections are atemporal entities, i.e., they do not come into or go out of existence. The endurant individuals which are the members of collections, on the other hand, do exist at some times and fail to exist at others. Consider the collection p of cells in my body at some time instant t . We will say that

¹Thus in contrast to Bittner *et al.* (2004b) we require here that collections have at least two members. For a more comprehensive version of this theory of collections see (Bittner and Donnelly 2006).

this collection is *present* now to signify that all of its members exist at this moment. But at some later time t_2 , many of the cells that form p will no longer exist and p will then be only *partly present*. In 500 years no member of p will exist and p will be *non-present*. Similarly, consider the collection q of states that are spatially in Europe in 2005. This collection includes France and Lithuania but not Russia and Turkey since the latter are only partly in Europe. This collection is fully present in 2005, i.e., all its members exist in 2005. But it was only partly present in 1949 since, for example, Lithuania did not exist as a separate state at that time (because it was a part of the Soviet Union). 2000 years ago q was non-present, since at that time none of its members existed.

Following Bittner *et al.* (2004b) we introduce the symbols ‘ FP ’, ‘ PtP ’, and ‘ NP ’ with the interpretations ‘fully present at t ’, ‘partly present at t ’, and ‘non-present at t ’ for a given collection q . Formally we define that a collection p is fully present at t if and only if all its members exist at t (D_{FP}); p is partly present at t iff some of its members exist at t (D_{PtP}); and p is non-present at t if and only if none of its members exist at t (D_{NP}).

$$\begin{array}{ll} D_{FP} & FP\ pt \equiv (x)(x \in p \rightarrow E\ xt) \\ D_{PtP} & PtP\ pt \equiv (\exists x)(x \in p \wedge E\ xt) \end{array} \qquad D_{NP} \quad NP\ pt \equiv \neg PtP\ pt$$

Notice that, since every collection has at least two members ($AC2$), full presence is a special case of partial presence, i.e., we can prove that if p is fully present at t , then p is also partly present at t ($TC4$). We can also prove that if p is a sub-collection of q and q is fully present at t then p is fully present at t ($TC5$); if p is a sub-collection of q and q is non-present at t then p is non-present at t ($TC6$); and if p is a sub-collection of q and p is partly present at t then q is partly present at t ($TC7$).

$$\begin{array}{ll} TC4 & FP\ pt \rightarrow PtP\ pt \\ TC5 & p \subseteq q \wedge FP\ qt \rightarrow FP\ pt \end{array} \qquad \begin{array}{ll} TC6 & p \subseteq q \wedge NP\ qt \rightarrow NP\ pt \\ TC7 & p \subseteq q \wedge PtP\ pt \rightarrow PtP\ qt \end{array}$$

Notice that it would also be possible to develop a theory of collections in which collections only exist at times at which they are fully present. According to such a theory collections would behave temporally like individuals rather than like sets.

4.2 Discrete collections

The individuals in a given collection may or may not overlap at a given time. For example let p be the collection which includes my body and my heart as its only members. At this moment t , p is such that my body and my heart overlap (since at this moment my heart is part of my body). Similarly, the collection q which has as its members the territory of Canada and the territory of Quebec is such that its members overlap at this moment in time. Some collections, however, are formed by individuals all of which are at a given time pairwise disjoint, which means that they have no parts in common at that time. For example, all individuals in the collection of cells currently in my body are currently pairwise disjoint – thus the collection as a whole is currently discrete. The collection of planets currently in our solar system is currently discrete. The collection of the current states of the United States is currently discrete.

We introduce the symbol ‘ D ’ in our ontology and assert that collection p is *discrete* at time t if and only if the members p do not overlap at t (D_D):

$$D_D \quad D\ pt \equiv (x)(y)(x \in p \wedge y \in p \wedge O\ xyt \rightarrow x = y)$$

We can prove: non-present collections are discrete ($TC8$); if p is discrete at t then every sub-collection of p is discrete at t ($TC9$).

$$TC8 \quad NP \, pt \rightarrow D \, pt$$

$$TC9 \quad D \, pt \wedge q \subseteq p \rightarrow D \, qt$$

TC9 enables an automated reasoner to infer the discreteness of a sub-collection from the discreteness of the super-collection. Thus it is sufficient to explicitly represent that the collection of current US federal states is disjoint; that the collection of New England States is currently discrete can then be derived using *TC9*.

Notice that the same collection can be fully present and non-discrete today but fully present and discrete tomorrow. Today the collection q which has as its members the territory of Canada and the territory of Quebec is fully present and not-discrete. In the future, Quebec might leave Canada and become a separate sovereign state. At that time the territories of the Quebec and of Canada will still be members of q but discrete, i.e., non-overlapping, entities. Similarly, consider the collection p consisting of my body and my heart. After my death my body and my heart will remain members of p . They may also continue to exist as discrete entities: my heart in the body of another person and my body in the anatomy department of some medical school. A collection can also be non-discrete today and discrete tomorrow if members that overlap today shrink or cease to exist tomorrow.

4.3 Sums and partitions

The members of the collection of the territories of the current federal states of the USA and the District of Columbia currently *sum up to* the territory of the USA as a whole in the sense that everything that currently overlaps the territory of the USA currently overlaps at least one member of this collection and vice versa. Moreover the collection territories of the current federal states and the District of Columbia currently *partitions* the territory of the USA in the sense that its members currently sum up to the territory of the USA and, in addition, are currently pairwise disjoint.

Formally we define that individual y is a sum of the collection p at time t if and only if (i) p is fully present at t and (ii) every individual w overlaps y at t if and only if y overlaps some member of p at t . We will also say that x is a p -sum at t . A collection p is a *partition* of the individual y at time t if and only if (i) y is a p -sum at t and (iii) p is discrete at t . We introduce the symbols ‘*Sum*’ and ‘*PT*’ into the language of our ontology and specify their meanings as follows:

$$\begin{aligned} D_{Sum} \quad Sum \, pyt &\equiv FP \, pt \wedge (x)(O \, xyt \leftrightarrow (\exists z)(z \in p \wedge O \, xzt)) \\ D_{PT} \quad PT \, pyt &\equiv Sum \, pyt \wedge D \, pt \end{aligned}$$

An individual can be partitioned by multiple collections. For example, in addition to the collection of territories of current federal states and the District of Columbia, the territory of the USA is also partitioned by the current collection of territories of its counties.¹ Note that collections are distinct from mereological sums. Thus (ignoring, for the moment, the District of Columbia) the territory of the USA is identical with the mereological sum of the territories of the separate states but distinct from the collection of the territories of the separate states.

It immediately follows from definition D_{PT} that if x is partitioned by q at t then x is a q -sum at t . We can also prove that if two objects that are both sums of the members of the same collection at t then they are mereologically equivalent at t (*TC10*).

$$TC10 \quad Sum \, pxt \wedge Sum \, pyt \rightarrow x \sim_t y$$

Notice that we cannot prove that two objects that are sums of the members of the same collection at a given time are identical. Mereological summation is not a function that takes collections to unique individuals. Consider the City of Vienna, the Austrian Federal State of Vienna, and the collection, PDV, of postal

¹We here ignore the fact that in Louisiana counties are called ‘parishes’ and in Alaska ‘boroughs’.

districts of Vienna at some time t in 2005. Assume that at t the City of Vienna is a sum of PDV and that the Austrian State of Vienna is a sum of PDV. According to our theory this is consistent with the thesis that the City of Vienna and the Austrian State of Vienna are two distinct objects (which are also parts of each other at t).

We can also prove: if y is a sum of p at t , then every part of y at t overlaps at t some member of p , i.e., no part of y is discrete from all members of p (*TC12*); if x is a sum of p at t and y is a sum of q at t and p is a sub-collection of q , then x is part of y at t (*TC14*).

$$\begin{aligned} \text{TC12} \quad & \text{Sum } pyt \wedge x \leq_t y \rightarrow (\exists z)(z \in p \wedge O zxt) \\ \text{TC14} \quad & \text{Sum } pxt \wedge \text{Sum } qyt \wedge p \subseteq q \rightarrow x \leq_t y \end{aligned}$$

Theorem *TC12* tells us that if x is a sum of the collection p at t , then p exhausts x at t in the sense that there are no parts of x that are not overlapped by some member of p at t . Thus, if x is a partition of the collection p at t then the members of p are at t jointly exhaustive and pairwise disjoint. Hence the intuitions stated informally in the first paragraph of this subsection are ‘covered’ by our formal definitions.

Notice, that a given collection p may partition an individual x today and fail to partition the same individual x tomorrow (x may grow, members of p may shrink or cease to exist, etc). Many partitions, however are quite stable over time, since they are defined by fiat. Consider, for example, the partition of the planet Earth into northern and southern hemispheres. By definition, at no time at which the planet Earth exists can the collection whose members are its northern and southern hemispheres fail to partition the planet.

We call the theory which includes TNEMO and the axioms *AC1+2 non-extensional temporal mereology with sums* (TNEMO-S).

5 Universals, individuals, and collections

In this section we specify the logical properties of the relations that hold between individuals, universals, and collections. For this purpose we extend TNEMO-S by adding more primitives and corresponding axioms to our theory. We use variables c, d, e, g to range over universals such as *human being, federal state, mountain, forest, tree, plant*, and so forth.

5.1 The sub-universal relation

Universals are here assumed to form hierarchical tree structures ordered by sub-universal relations. Consider the ecoregion classification hierarchy depicted in Figure 3, where nodes represent universals and edges the subuniversal-of relation (Bailey 1983). The tree structure reflects the fact that the definition of each universal lower down in the hierarchy is formed by specifying its parent universal together with the relevant differentia that tell us what marks out instances of the defined universal (or ‘species’) within the wider parent universal (or ‘genus’), as in: *human* =_{df} *rational animal* where ‘rational’ is the differentia (Smith *et al.* 2004, Sorokine *et al.* 2006). Notice that differentia, on this Aristotelean approach to definitions, must always be pairwise disjoint. In addition, they may be such that the immediate sub-universals of a universal are jointly exhaustive (with respect to the immediate super-universal). Thus besides rational animals there are non-rational animals. In the ecoregion classification hierarchy we define *Humid Temperate Ecoregion* as a *Geographic Ecoregion* with humid temperate climate. Here the climate type is the differentia. In Table 1 we give some more example definitions for ecoregion universals with different types of differentia such as vegetation type and climax vegetation type. The ecoregion classification tree that can be constructed using these kinds of differentia is partly depicted in Figure 3.

According to the view defended here, this Aristotelian method of classification allows us to build classification systems which most closely resemble the hierarchical organization of the universals in reality. Of course there may be different ways of classifying, resulting in different classification trees and corresponding to different sub-universal relations. (Bittner (2007b) provides a detailed discussion of various classification

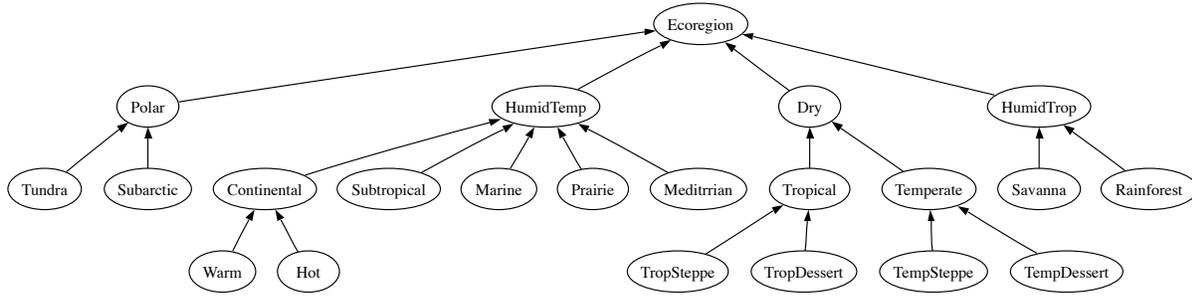


Figure 3. Classification of geographic ecoregions with respect to broad climatic similarity, definite vegetational affinities, etc. (Bailey 1983).

<i>Humid Temperate Ecoregion</i>	$=_{df}$	<i>Geographic Ecoregion</i> with humid temperate climate.
<i>Prairie Ecoregion</i>	$=_{df}$	<i>Humid Temperate Ecoregion</i> with prairie climate.
<i>Prairie Bushland Ecoregion</i>	$=_{df}$	<i>Prairie Ecoregion</i> with climax vegetation type Bushland.

Table 1. Definition of ecoregion universals using the Aristotelian method of classification.

schemes in the ecoregion realm.) Notice that the sub-universal relation will generate a lattice – in which a universal can have more than one immediate super-universal (multiple inheritance) – rather than a tree if either (i) different classification trees are mixed,¹ or (ii) universals are confused with collections and sub-universal relations are confused with sub-collection relations, or (iii) other mistakes are made as for example, those discussed in (Sorokine and Bittner 2005).²

In our formal theory we use the symbol ‘ \sqsubseteq ’ for the sub-universal relation. Like the sub-collection relation, \sqsubseteq is atemporal, i.e. it does not have a temporal parameter. We define the relations of proper sub-universal (\sqsubset) and taxonomic overlap (O_{\sqsubseteq}) in the obvious way in terms of \sqsubseteq . Universal c is a proper sub-universal of d if and only if c is a sub-universal of d and d is not a sub-universal of c (D_{\sqsubseteq}). Universals c and d taxonomically overlap if and only if either c is a sub-universal of d or d is a sub-universal of c ($D_{O_{\sqsubseteq}}$). We also introduce a predicate (*Root*), which picks out the root universal, which subsumes all universals (D_{root}).

$$\begin{array}{ll}
 D_{\sqsubseteq} & c \sqsubset d \equiv c \sqsubseteq d \wedge \neg d \sqsubseteq c \\
 D_{O_{\sqsubseteq}} & O_{\sqsubseteq} cd \equiv (c \sqsubseteq d \vee d \sqsubseteq c) \\
 D_{root} & Root c \equiv (g)(g \sqsubseteq c)
 \end{array}$$

We require that \sqsubseteq is reflexive, antisymmetric, and transitive (AU1-3). In addition we require that: if c is a proper sub-universal of d then there exists a universal e that is a proper sub-universal of d and which does not taxonomically overlap c (AU4); and if c and d share a common sub-universal then c and d taxonomically overlap (AU5). We finally add axiom AU6, which postulates the existence of a root universal. In the most general case, this will be the universal substance (or independent enduring).

$$\begin{array}{ll}
 AU1 & c \sqsubseteq c \\
 AU2 & (c \sqsubseteq d \wedge d \sqsubseteq c) \rightarrow c = d \\
 AU3 & (c \sqsubseteq d \wedge d \sqsubseteq e) \rightarrow c \sqsubseteq e \\
 AU4 & c \sqsubset d \rightarrow (\exists e)(e \sqsubseteq d \wedge \neg O_{\sqsubseteq} ec) \\
 AU5 & (\exists e)(e \sqsubseteq c \wedge e \sqsubseteq d) \rightarrow O_{\sqsubseteq} cd \\
 AU6 & (\exists c)Root c
 \end{array}$$

Axioms AU1–AU6 force the sub-universal relation to form a *tree* structure. AU5 ensures that a universal

¹As an example consider the universals *socio-economic unit* and *human settlement*. If we mix the classifications of socio-economic units and human settlements into a single classification structure, then the resulting structure will not be a tree, since neither *socio-economic unit* is a sub-universal of *human settlement* nor vice versa, though both have the universal *city* as a (proper) sub-universal.

²Those who insist that the hierarchical structure imposed by the sub-universal relation is indeed a lattice can fall back to the version of the theory presented in (Bittner *et al.* 2004b). In that theory lattice structures are permitted as long as what we call the no-partial-overlap principle (NPO) is not added to the theory.

does not have two immediate super-universals. Moreover, we can prove that there cannot be a universal with a single proper sub-universal or, more generally, that if c is a proper sub-universal of d then there exists a universal e that is a proper sub-universal of d and c and e have no sub-universal in common ($TU0$).¹

$$TU0 \quad c \sqsubset d \rightarrow (\exists e)(e \sqsubset d \wedge \neg(\exists f)(f \sqsubseteq e \wedge f \sqsubseteq c))$$

From axioms $AU1$ – $AU6$ we can also prove atemporal versions of $TM1$ – 6 which are omitted here. (In these theorems the relations part-of and sub-universal-of as well as overlap and taxonomic overlap correspond in the obvious ways.) We can also prove: there exists at most one root-universal ($TU1$); if everything that taxonomically overlaps c also taxonomically overlaps d then c is a sub-universal of d ($TU2$); and two universals c and d are identical if and only if every universal e overlaps c if and only if it overlaps d ($TU3$).

$$TU1 \quad \text{Root } c \wedge \text{Root } d \rightarrow c = d$$

$$TU2 \quad (e)(O_{\sqsubseteq} ec \rightarrow O_{\sqsubseteq} ed) \rightarrow c \sqsubseteq d$$

$$TU3 \quad (e)(O_{\sqsubseteq} ec \leftrightarrow O_{\sqsubseteq} ed) \leftrightarrow c = d$$

We call the theory formed by the axioms $AU1$ – 6 *Extensional Universal Mereology* (EUM).

5.2 Instantiation

Universals have various individuals at various times as their instances, i.e., the relation of instantiation is time-dependent. This relation holds between individuals, universals, and times (in that order). Notice that this does not conflict with the atemporal character of the sub-universal relation. The universal *Polar ecoregion* will remain a sub-universal of *Geographic ecoregion* even if a time is reached when no polar ecoregion exists on Earth.

We include the primitive relation *Inst* in our formal theory and write *Inst xct* to signify that the individual x instantiates the universal c at time-instant t . For example: New York City is an instance of the universal *city* now; the tundra ecoregion of Alaska is an instance of the universal *Tundra ecoregion* (Figure 3) now.

The relation *Inst* is irreflexive and asymmetric at every time. Since in our ontology universals and individuals are represented through variables of disjoint sorts we do not need to add explicit irreflexivity and asymmetry axioms for *Inst*. Axioms ($AI1$ – 2) mirror the relationship between instantiation and the sub-universal relation. $AI1$ tells us that if c is a sub-universal of d , then the instances of c at any given time are also instances of d at that time. For example, the universal *federal state* is a sub-universal of the universal *socio-economic unit*. Therefore every instance of *federal state* (e.g., New York State) is also an instance of *socio-economic unit*.

$AI2$ tells us that if two universals share an instance x at some time t , then the universals taxonomically overlap. For example, in 2005 the tundra ecoregion of Alaska is an instance of both the universal *Tundra ecoregion* and of the universal *Geographic ecoregion*, both of which taxinomically overlap, since *Tundra ecoregion* is a sub-universal of *Geographic ecoregion*.

$$AI1 \quad c \sqsubseteq d \rightarrow (\text{Inst } xct \rightarrow \text{Inst } xdt)$$

$$AI2 \quad (\text{Inst } xct \wedge \text{Inst } xdt) \rightarrow O_{\sqsubseteq} cd$$

The axioms $AI3$ – 5 mirror the interaction between instantiation and existence of individuals in time. $AI3$ tells us that if x is an instance of a universal at t then x exists at t . For example, Napoleon Bonaparte is not an instance of the universal *human being* in 2005 since he does not exist in 2005. However he was an instance of *human being* in 1815. $AI4$ states that every universal is instantiated at some time. Thus we

¹ $TU0$ seems to be violated in classification systems in which universals with a single sub-universal are postulated. Sorokine and Bittner (2005) investigated this phenomenon in the context of ecoregion classifications and showed that in classification systems that violate $TU0$ either the sub-universal relation is confused with the instantiation relation, or universals that do not have instances in a given spatial location are neglected.

do not allow for universals such as ‘unicorn’ which do not have instances at any time. *AI5* states that at every time at which an individual exists it is an instance of some universal.¹

$$\begin{array}{ll} \text{AI3} & \text{Inst } xct \rightarrow E xt \\ \text{AI4} & (\exists t)(\exists x)(\text{Inst } xct) \\ \text{AI5} & E xt \rightarrow (\exists c)\text{Inst } xct \end{array}$$

5.3 Extensions

Every universal that has an instance at time t also has an extension at t . The extension of universal c at t is an individual object if and only if c has a single instance at t . The extension of c at t is a collection if and only if c has at least two instances at t . At the formal level we correspondingly introduce the symbols Ext_C and Ext_I . $Ext_C pct$ holds if and only if for all x , x is a member of p if and only if x instantiates c at t (D_{Ext_C}). $Ext_I xct$ holds if and only if x is an instance of c at t and all instances of c at t are identical to x (D_{Ext_I}).

$$\begin{array}{ll} D_{Ext_C} & Ext_C pct \equiv (x)(x \in p \leftrightarrow \text{Inst } xct) \\ D_{Ext_I} & Ext_I xct \equiv \text{Inst } xct \wedge (y)(\text{Inst } yct \rightarrow x = y) \end{array}$$

We then require that if universal c has an instance x at t , then either x is the extension of c at t or there is a collection p that is the extension of c at t (*AE1*).

$$\text{AE1} \quad \text{Inst } xct \rightarrow (Ext_I xct \vee (\exists p)(Ext_C pct))$$

We can prove that: if collection p is the extension of a universal at time t , then p is fully present at t (*TC15*); if individual x is an extension of universal c at t and individual y is an extension of c at t then x and y are identical (*TC16*), and that if collection p an extension of universal c at t and collection q is an extension of universal c at t , then p and q are identical (*TC17*).

$$\begin{array}{ll} \text{TC15} & Ext_C pct \rightarrow FP pt \\ \text{TC16} & Ext_I xct \wedge Ext_I yct \rightarrow x = y \\ \text{TC17} & Ext_C pct \wedge Ext_C qct \rightarrow p = q \end{array}$$

Thus at all times at which the universal c has an extension it has a unique extension. This extension may be either a single object or a collection. Hence, we are allowed to refer to *the* extension of a universal c at all times at which c has an instance.

For universals with two or more instances at a given time there is clearly a correspondence between the sub-universal structure of universals and the sub-collection structure of their extensions. We can indeed prove that if c is a sub-universal of d and p is the extension of c at t and q is the extension of d at t then p is a sub-collection of q (*TC19*).

$$\text{TC19} \quad c \sqsubseteq d \wedge Ext_C pct \wedge Ext_C qdt \rightarrow p \subseteq q$$

Notice however that there may be points in time where *distinct* universals have identical extensions. For example, if at some point in time all mammals except whales are extinct, then the extensions of the universals *mammal* and *whale* at that time are identical even though the corresponding universals are distinct.

¹Note that we here do not add an axiom requiring that two universals that have the same instances at all times are identical. Thus in contrast to (Bittner *et al.* 2004b) we leave open the possibility that two distinct universals may have exactly the same instances at all times. In a modal framework one usually demands that two universals are identical if and only if they have the same instances at all times and in all possible worlds (Oliver 1996).

We call the theory that extends EUM by the axioms *AI1*–*6* and *AE1* *Extensional Universal Mereology with Instantiation and Extensions* (EUMIE).

5.4 Discrete universals

Many universals are such that distinct instances do not overlap. We call such universals *discrete*. Examples of discrete universals include *federal state*, *human being*, *tree*, *forest*, *city*, *car*. No distinct instances of *federal state* overlap. No distinct instances of *forest* overlap. Universals which have distinct instances that do overlap include *socio-economic unit*, *ecoregion*, *human body part*. For example, the State of New York and Niagara County – both instances of *socio-economical unit* – overlap. Similarly, the my left hand and my left arm – both instances of *human body part* – overlap.

Formally we define that the universal c is discrete if and only if all of its extensions are discrete (D_{DU}).

$$D_{DU} \quad DU\ c \equiv (t)(p)(Ext_C\ pct \rightarrow D\ pt)$$

We can prove that at no time do distinct instances of a discrete universal overlap ($TC20$).

$$TC20 \quad DU\ c \wedge (\exists t)(Inst\ xct \wedge Inst\ yct \wedge O\ xyt) \rightarrow x = y$$

Consider the individuals referred to by the names ‘Stadtwald-1’ and ‘Stadtwald-2’ in Example 1 (Sec. 2.1) above. If both individuals overlap at some time (e.g., share at least one tree in 2000), then according to $TC20$, both individuals are identical, since both are instances of the discrete universal *forest*.

6 Universal parthood

We saw above in Example 2 (Sec. 2.2) that there, are besides the parthood relation between individuals, also parthood relations on the level of universals (*waterfall* part of *river*, *wall* part of *house*, and so on). In this section we combine the theories of non-extensional temporal mereology with sums, TNEMO-S, and the extensional universal mereology with instantiation and extensions, EUMIE, and introduce universal parthood relations. We call the resulting theory *Non-Extensional Temporal Mereology with Sums, and Universals*, TNEMO-S-U. We will then finally be able to formally characterize the relations R_G , R_J and R_H employed in Example 2.

6.1 Kinds of universal parthood relations

In the formal theory we introduce the predicates UP_1 , UP_2 and UP_{12} for the universal parthood relations. These predicates acquire their meaning through definitions that are based exclusively on the primitives \leq and $Inst$ (Donnelly and Bittner 2005, Donnelly *et al.* 2006, Smith *et al.* 2005):

$$D_{UP_1} \quad UP_1\ cd \equiv (t)(x)(Inst\ xct \rightarrow (\exists y)(Inst\ ydt \wedge x \leq_t y))$$

for every time t : every instance of c at t is an individual part at t of some instance of d at t

$$D_{UP_2} \quad UP_2\ cd \equiv (t)(y)(Inst\ yct \rightarrow (\exists x)(Inst\ xct \wedge x \leq_t y))$$

for every time t : every instance of d at t has some instance of c at t as an individual part at t

For example we have $UP_1(\textit{waterfall}, \textit{river})$, i.e., for all times t any instance of *waterfall* at t is an individual part at t of some instance of *river* at t . Similarly, $UP_2(\textit{wall}, \textit{house})$, and so on.

We now can use the universal parthood predicates UP_1 and UP_2 and the discrete-universal predicate DU as a basis to define a set of further predicates as conjunctions of the base predicates. From all the available possibilities we will discuss the following:

$$\begin{array}{ll}
D_{UP_{12}} & UP_{12} cd \equiv UP_1 cd \wedge UP_2 cd \\
D_{DUP_1} & DUP_1 cd \equiv UP_1 cd \wedge DU c \wedge DU d \\
D_{DUP_2} & DUP_2 cd \equiv UP_2 cd \wedge DU c \wedge DU d \\
D_{DUP_{12}} & DUP_{12} cd \equiv DUP_1 cd \wedge DUP_2 cd
\end{array}$$

The predicate UP_{12} is defined as the conjunction of UP_1 and UP_2 . For example we have $(UP_{12}(\text{capital}, \text{state}))$, i.e., for all times t every instance of state at t is an individual part at t of some instance of state at t and for all t every instance of state at t has some instance of capital at t as an individual part at t .

The relations picked out by DUP_1 , DUP_2 , and DUP_{12} are restrictions of the relations picked out by UP_1 , UP_2 , and UP_{12} to discrete universals. On the intended interpretation DUP_1 , DUP_2 , and DUP_{12} pick out respectively the relations R_G, R_J and R_H in Example 2 (Sec. 2.2). More examples and counter-examples of universal parthood relations between discrete universals are listed in Table 2.

relation	example	counter-example
DUP_1	$(\text{ford}, \text{river}), (\text{waterfall}, \text{watercourse})$	$(\text{wall}, \text{building})$
DUP_2	$(\text{wing}, \text{airplane}), (\text{wall}, \text{building})$	$(\text{ford}, \text{river})$
DUP_{12}	$(\text{human head}, \text{human body}), (\text{capital}, \text{state})$ $(\text{state}, \text{federal state}), (\text{federal state}, \text{federation})$	$(\text{wall}, \text{building})$

Table 2. Universal parthood relations.

6.2 Composition and transitivity reasoning

Transitivity reasoning.. We can prove a number of theorems that tell us how to validly combine partial information about universal parthood relations. We start with theorems involving the *same* universal parthood relation, proving that UP_1 , UP_2 , and UP_{12} , as well as DUP_1 , DUP_2 , and DUP_{12} are all transitive. That is, if c is a universal part of d (in the sense of UP_i) and d is a universal part of e (in the sense of UP_i), then c is a universal part of e (in the sense of UP_i) where the index i takes the values ‘1’, ‘2’, and ‘12’ respectively ($TPI1-3$). Similarly for DUP_1 , DUP_2 , and DUP_{12} ($TPI3-6$).

$$\begin{array}{ll}
TUP1-3 & UP_i cd \wedge UP_i de \rightarrow UP_i ce \quad \text{for } i \in \{1, 2, 12\} \\
TUP3-6 & DUP_i cd \wedge DUP_i de \rightarrow DUP_i ce \quad \text{for } i \in \{1, 2, 12\}
\end{array}$$

Consider the universals *county*, *federal state*, and *federation*. It is easy to verify that we have:

$$DUP_{12}(\text{county}, \text{federal state}) \text{ and } DUP_{12}(\text{federal state}, \text{federation})$$

and, by transitivity, $DUP_{12}(\text{county}, \text{federation})$. Thus, theorem $TUP6$ tells us that the facts $DUP_{12}(\text{county}, \text{federal state})$ and $DUP_{12}(\text{federal state}, \text{federation})$ can be combined to infer $DUP_{12}(\text{county}, \text{federation})$ by means of reasoning that exploits the validity of theorem $TPI6$ as discussed above together with the transitivity inference rule (TIR on page 9). Obviously, this kind of reasoning is important in integrating geographical databases at different scales, since it allows us to make explicit information that is implicit in the statements $DUP_{12}(\text{county}, \text{federal state})$ and $DUP_{12}(\text{federal state}, \text{federation})$.

Relation composition.. We now consider theorems about the combination of statements involving *distinct* universal parthood relations. Transitivity reasoning, i.e., reasoning that exploits the validity of theorems $TUP1-6$ and thus employs the transitivity property of a single relation, cannot be used to derive conclusions from premises which include facts about distinct relations. Thus we cannot derive valid conclusions from the premises

$$UP_1 cd \text{ and } UP_{12} de \quad \text{or} \quad UP_1 cd \text{ and } DUP_1 de$$

by means of transitivity reasoning, since UP_1 , DUP_1 , and UP_{12} are distinct relations. In such cases a more general form of reasoning based on the *composition* of relations is required. Relation composition has the form:

RC From $R(c, d)$ and $S(d, e)$, derive $T(c, e)$

‘ R ’, ‘ S ’, and ‘ T ’ are symbols referring to possibly distinct binary relations. One can see that the transitivity inference rule (TIR on page 9 of this paper) is a special kind of relation composition rule, where ‘ R ’, ‘ S ’, and ‘ T ’ refer to the same relation. Reasoning by relation composition is widely used in spatial reasoning about topological relations between individual spatial objects and regions (Egenhofer 1991, Egenhofer and Sharma 1993, Cohn *et al.* 1997). The importance of the composition of universal parthood relations for (bio and geo-) ontologies has been pointed out in Schulz *et al.* (2000), Spackman (2001), Schulz and Hahn (2004), Donnelly and Bittner (2005).

In the formal theory it immediately follows from the definitions of the various universal parthood predicates that the stronger UP_{12} predicate implies the weaker predicates UP_1 and UP_2 ($TUP7$). Similarly, a restricted universal parthood relation like DUP_i implies the non-restricted universal parthood relation UP_i ($TUP8-10$).

$$\begin{array}{l} TUP7 \quad UP_{12} cd \rightarrow UP_1 cd \wedge UP_2 cd \\ TUP8-10 \quad DUP_i cd \rightarrow UP_i cd \quad \text{for } i \in \{1, 2, 12\} \end{array}$$

Table 3 summarizes the various ways of validly combining partial information about universal parthood relations based on Theorems $TUP1-TUP10$. Assume that $UP_1 cd$ and $UP_{12} de$ hold, we can then derive $UP_1 de$ from $UP_{12} de$ using theorem $TUP14$. We then can apply theorem $TUP5$ to derive $UP_1 ce$ (the entry in the table) from $UP_1 cd$ and $UP_1 de$. Similarly, assume that $UP_1 cd$ and $DUP_1 de$ hold, we can derive $UP_1 de$ from $DUP_1 de$ using theorem $TUP15$. We then can apply theorem $TUP5$ to derive $UP_1 ce$ (the entry in the table) from $UP_1 cd$ and $UP_1 de$. Table 3 shows that we cannot validly combine partial information about universal parthood relations between arbitrary relations. For example we cannot prove any theorem that supports valid inferences from $UP_1 cd$ and $UP_2 de$.

	$UP_1 de$	$UP_2 de$	$UP_{12} de$	$DUP_1 de$	$DUP_2 de$	$DUP_{12} de$
$UP_1 cd$	$UP_1 ce$	–	$UP_1 ce$	$UP_1 ce$	–	$UP_1 ce$
$UP_2 cd$	–	$UP_2 ce$	$UP_2 ce$	–	$UP_2 ce$	$UP_2 ce$
$UP_{12} cd$	$UP_1 ce$	$UP_2 ce$	$UP_{12} ce$	$UP_1 ce$	$UP_2 ce$	$UP_{12} ce$
$DUP_1 cd$	$UP_1 ce$	–	$UP_1 ce$	$DUP_1 ce$	–	$DUP_1 ce$
$DUP_2 cd$	–	$UP_2 ce$	$UP_2 ce$	–	$DUP_2 ce$	$DUP_2 ce$
$DUP_{12} cd$	$UP_1 ce$	$UP_2 ce$	$UP_{12} ce$	$DUP_1 ce$	$DUP_2 ce$	$DUP_{12} ce$

Table 3. Composition of universal parthood relations.

6.3 Universal parthood and sub-universalhood

Finally we can prove a number of theorems about the interrelationships between universal parthood and the sub-universal relation. These theorems are useful since they allow us to combine partial information concerning facts about universal parthood and sub-universal relations. Given a particular classification hierarchy it is important to know whether information about universal parthood relations can be propagated up or down the classification hierarchy. For example, we might know that universal c is a universal part of universal d (in the sense of UP_1 , UP_2 , etc.) and that e is a sub-universal of c . The question may then arise – for example when applying definitions and axioms from a data-standard like ATKIS to a particular data set – as to whether we are permitted to derive that e is a universal part of d .

We first state the theorems formally and then provide english translations and simple examples. Theorems $TUP11-18$ are specific versions of the more general theorems discussed in Donnelly *et al.* (2006). Theorems $TUP11, 12, 15$, and 18 support valid reasoning using relation composition (RC).

<p><i>TUP11</i> $e \sqsubseteq c \wedge UP_1 cd \rightarrow UP_1 ed$</p> <p><i>TUP12</i> $d \sqsubseteq e \wedge UP_1 cd \rightarrow UP_1 ce$</p> <p><i>TUP13</i> $e \sqsubseteq d \wedge UP_2 cd \rightarrow UP_2 ce$</p> <p><i>TUP14</i> $c \sqsubseteq e \wedge UP_2 cd \rightarrow UP_2 ed$</p>	<p><i>TUP15</i> $e \sqsubseteq c \wedge UP_{12} cd \rightarrow UP_1 ed$</p> <p><i>TUP16</i> $c \sqsubseteq e \wedge UP_{12} cd \rightarrow UP_2 ed$</p> <p><i>TUP17</i> $e \sqsubseteq d \wedge UP_{12} cd \rightarrow UP_2 ce$</p> <p><i>TUP18</i> $d \sqsubseteq e \wedge UP_{12} cd \rightarrow UP_1 ce$</p>
---	---

As an illustration consider Theorems (*TUP11-13*), which can be rendered in English as follows:

(*TUP11*) If e is a sub-universal of c and c is a universal part of d in the sense of UP_1 (every instance of c is an individual part of some instance of d), then e is a universal part of d in the sense of UP_1 (*TIP11*).

Thus, from the facts that *index finger* is a sub-universal of *finger* and *finger* is universal part of *hand* (in the sense of UP_1), we can validly derive that *index finger* is universal part of *hand* (in the sense of UP_1).

(*TUP12*) If c is a universal part of d in the sense of UP_1 and d is a sub-universal of e then c is a universal part of e in the sense of UP_1 .

Thus from the facts that *watercourse* is a sub-universal of *waterbody* and *waterfall* is universal part of *watercourse* (in the sense of UP_1) we can validly derive that *waterfall* is universal part of *waterbody* (in the sense of UP_1).

(*TUP13*) If c is a universal part of d in the sense of UP_2 and e is a sub-universal of d , then c is a universal part of e in the sense of UP_2 .

Thus from the facts that *wing* (airplane wings, bird wings, etc.) is universal part of *airplane* (in the sense of UP_2) and that *passenger airplane* is a sub-universal of *airplane* we can validly derive that *wing* is universal part of *passenger airplane* (in the sense of UP_2).

Obviously, theorems similar in structure to theorems (*TUP11-18*) can be derived for universal parthood between discrete universals. We here discuss two such theorems:

$$TDUP19 \quad d \sqsubseteq e \wedge DUP_1 cd \rightarrow UP_1 ce \qquad TDUP20 \quad c \sqsubseteq e \wedge DUP_2 cd \rightarrow UP_2 ed$$

TDUP19 tells us that if d is a sub-universal of e , and if c and d are discrete, and c is universal part of d in the sense of DUP_1 then c is universal part of e in the sense of UP_1 . Notice that e may or may not be discrete. (We know from *TC9* that every sub-collection of a discrete collection is discrete, but a super-collection of a discrete collection may or may not be discrete.) For this reason we cannot prove $DUP_1 ce$. Analogously for *TPI20*.

7 Computational realization

The axiomatic theory TNEMO-S-U presented above is part of Basic Formal Ontology (BFO), a top-level ontology developed by the Ontology Research Group at Buffalo and by the Institute for Formal Ontology and Medical Information Science in Saarbruecken (<http://www.ifomis.uni-saarland.de/bfo>). Significant parts of BFO are implemented in *Isabelle* a computational system for implementing logical formalisms (Paulson 1994, Nipkow *et al.* 2002). *Isabelle* is public domain software and can be downloaded for a wide range of operating systems from the *Isabelle* website (Paulson and Nipkow 2005). A hierarchical representation of the resulting BFO sub-theory structure is shown in Figure 4.

The computational representation of TNEMO-S-U, including all definitions and axioms and the proofs of all theorems discussed above, can be accessed at (Bittner 2007a). In the remainder of this section we briefly discuss this computational representation.

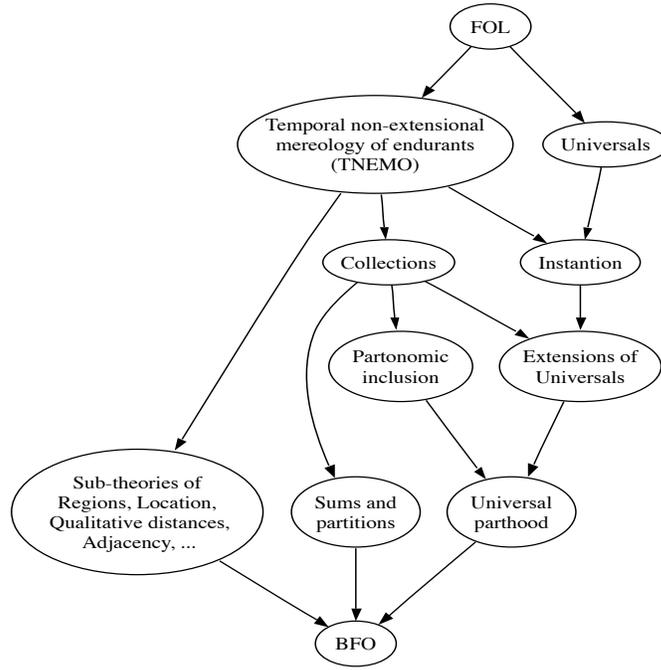


Figure 4. The hierarchical sub-theory structure of BFO. (An arrow from T_1 to T_2 means that every axiom of T_1 is also an axiom of T_2 .)

7.1 Formal specification of TNEMO-S-U

Figure 5 depicts a portion of the computational representation of the BFO sub-theory TNEMO that – as stated on line one – extends the Isabelle theory FOL (an Isabelle implementation of a sorted first-order predicate logic with identity). On the following lines the two different sorts (also called types) that are used in TNEMO are introduced: endurant individuals (\mathbf{Ob}) and time instants (\mathbf{Ti}). Both are derived from the Isabelle type “term”.

In the section *consts* of the code fragment in Figure 5 the predicate symbols and their signatures are introduced. The expression

```
O :: "Ob => Ob => Ti => o"
```

tells us that ‘O’ (for overlap) is a ternary (three-place) predicate symbol in which the first and second parameter are of sort \mathbf{Ob} (endurant individuals) and the third parameter is of sort \mathbf{Ti} (time-instants). The fourth parameter \mathbf{o} is the computational representation of the fact that $\mathbf{O}(x,y,t)$ is a ternary predicate that is either true or false.

The axioms *P_exists1*, *P_exists2*, *P_trans*, and *P_suppl* of the section *axioms* of Figure 5 are, respectively, the axioms AM1-AM4 of TNEMO as discussed in Section 3. Thus

```
P_exists1: "(ALL x. (EX t. E(x,t)))"
```

is interpreted as: “the axiom labeled *P_exists1* states that for all endurants x there is a time t such that x exists at t ”. Notice that the type information is inferred by the system automatically.

In the section *defs* of Figure 5 some definitions of TNEMO are represented. The definition *E_def* is the definition for the ‘exists’ predicate and corresponds to definition D_E of TNEMO in Section 3. Thus

```
E_def: "E(x,t) == P(x,x,t)"
```

is interpreted as: “the definition labeled *E_def* states that endurant x exists at time t if and only if x is part of itself at t ”.

An important feature of Isabelle as a generic system for implementing logical formalisms is that it allows

```

theory TNEMO = FOL:

typedecl Ob
typedecl Ti

arities Ob :: "term"
        Ti :: "term"

consts
O :: "Ob => Ob => Ti => o"
P :: "Ob => Ob => Ti => o"
PP :: "Ob => Ob => Ti => o"
D :: "Ob => Ob => Ti => o"
E :: "Ob => Ti => o"
Me :: "Ob => Ob => Ti => o"

axioms
P_exists1: "(ALL x. (EX t. E(x,t)))"
P_exists2: "(ALL x y t. (P(x,y,t) -->(E(x,t) & E(y,t))))"
P_trans: "(ALL x y z t. (P(x,y,t) & P(y,z,t) --> P(x,z,t)))"
P_ssuppl: "(ALL x y t. ((E(x,t) & ~P(x,y,t)) --> (EX z. (P(z,x,t) & ~O(z,y,t)))))"

defs
E_def: "E(x,t) == P(x,x,t)"
O_def: "O(x,y,t) == (EX z. (P(z,x,t) & P(z,y,t)))"
PP_def: "PP(x,y,t) == P(x,y,t) & ~P(y,x,t)"
D_def: "D(x,y,t) == ~O(x,y,t)"
Me_def: "Me(x,y,t) == (P(x,y,t) & P(y,x,t))"

theorem Me_refl: "E(x,t) ==> Me(x,x,t)"
apply(unfold Me_def)
apply(unfold E_def)
apply(auto)
done

```

Figure 5. Declarations, axioms, definitions, and a theorem of the non-extensional mereology TNEMO of (Bittner 2007a).

us to prove theorems semi-automatically. In fact many theorems of BFO can be proved with very little human assistance. Consider the theorem labeled `Me_refl` in Figure 5, which corresponds to Theorem *TM8* in Section 3. A proof in Isabelle is a sequence of applications of logical rules using the `apply` command. Every proof ends with the keyword `done`. (Nipkow *et al.* 2002) is a tutorial on how to prove theorems in Isabelle which requires only limited background in formal logic and can be downloaded from the Isabelle website. The recipe for the proof of theorem `Me_refl`, for example, is read as follows: (1) replace `Me(x,x,t)` by the right hand side of definition `Me_def`; (2) replace `E(x,t)` by the right hand side of definition `E_def`; and (3) search for a proof automatically. Theorems can also be proved explicitly by stepwise application of logical derivation rules, as demonstrated in the proof of theorem `O_imp_O_imp_P` (Theorem *TM10* in Section 3) in the module TNEMO of (Bittner 2007a). The important point is that if Isabelle ‘compiles’ a theory file, then this means that all the proofs are machine-verified.

7.2 How to use the implemented theory?

The Isabelle-based computational representation of BFO can be used in at least three ways, which we discuss in successive order below.

- (i) As a reference ontology to integrate domain-specific ontologies and terminologies;
- (ii) As the basis of a more detailed theory that includes additional theorems making further consequences of the current axioms explicit and also verifies additional reasoning rules;
- (iii) As a basis of an extended theory that has more primitives and hence more axioms. The resulting theory

may be a more comprehensive top-level ontology or it may be a more specific *domain ontology*.

7.2.1 TNEMO-S-U as a reference ontology. In Example 1 (Section 2.1) we demonstrated informally how a top-level ontology can serve as a reference ontology to integrate two data sets: one structured using the ATKIS terminology, the other using the CORINE terminology. We now can use the implemented theory to demonstrate how to link the two terminologies and the corresponding data sets formally and to verify the informal reasoning provided in Example 1 (labeled (1) on page 5).

Figure 6 shows the theory module `BFO_Example1` which formalizes the linking of certain ATKIS and CORINE terms and data items in the corresponding data sets. This theory extends BFO and thus can use all of BFO's axioms, definitions, and theorems. One can see that we use the theorems `IsA_trans_rule`, `Inst_IsA_rule`, and `DUn_and_0_impl_Id` (corresponding to theorems *TC3*, *TC20*, and axiom *AI1* of TNEMO-S-U). Sentences A, B, and C as well as the assumptions (a0–a4, a6) of Example 1 are represented as additional axioms. The theorems `align_ATKIS_to_BFO`, `align_CORINE_to_BFO` formally link ATKIS and CORINE terms to top-level relations in BFO. The theorem `Stadtwald_ATKIS = Stadtwald_CORINE` formally proves that both data items, `Stadtwald_ATKIS` and `Stadtwald_CORINE`, refer to the same entity, classified as *forest*, as described in Example 1.

7.2.2 Refining TNEMO-S-U. In principle the theory TNEMO-S-U (and thus BFO) has an infinite number of theorems. In this paper we discussed only certain representative theorems that are important for practical purposes such as transitivity reasoning and reasoning based on relation composition. Other theorems may be derived using the resources of TNEMO-S-U and its Isabelle representation. For example, the BFO module `PartonomicInclusion` contains theorems that formalize reasoning about hierarchical spatial subdivisions at different levels of granularity (e.g., postal districts, counties, federal states, federations, etc.).

7.2.3 Building a more comprehensive top-level ontology. This paper focussed on the mereology of independent endurants and thus it needs extending to yield an account of a top-level ontology: it lacks top-level relations such as *located-in*, *connected-to*, *close-to*, *adjacent-to* and top-level categories such as *perdurant* (process), *quality*, etc. (Smith and Grenon 2004, Grenon and Smith 2004). In (Bittner and Donnelly 2007a) we present our formal theory of qualitative size and distance relations. In (Bittner and Donnelly 2007b) we discuss how to incorporate the relations *located-in*, *connected-to*, and *adjacent-to* into BFO. The corresponding computational representations can be found at (Bittner 2007a) as indicated in Figure 4. It is also obvious that the general framework presented here can be extended by incorporating computational representations of the ontological theories of perdurants and qualities as presented for example in (Simons 1987, Sider 2001, Grenon and Smith 2004, Bittner *et al.* 2004a, Galton and Worboys 2005). An OWL-based implementation of some of those aspects can be found at <http://www.ifomis.uni-saarland.de/bfo>.

7.2.4 Building specific domain ontologies. Consider the code in Figure 6. We introduced the primitives `mixed-forest` and `forest` without any definitions or axioms. Those definitions and axioms could be added to the theory file using the vocabulary of BFO. But they would not be parts of the top-level ontology since they belong to the specific *domain* of geography.

The methodology of building domain ontologies by extending the underlying top-level ontology automatically ensures that the shared top-level terms and relations appearing in domain ontologies are used in a consistent, compatible, and provably correct way. The entire framework can in this way be used to integrate data from distinct though related domains, for example from geology, environmental planning, oil chemistry, and the soil sciences. How to use the presented framework to build domain ontologies is demonstrated in more detail in the context of ecoregion classifications in (Bittner 2007b). Further examples of such use in a variety of biomedical domains are outlined in (Smith *et al.* 2007).

```

theory BFO_Example1 = BFO:

consts
  T2000 :: "Ti"
  forest :: "Un"
  mixed_forest :: "Un"
  Stadtwald_ATKIS :: "Ob"
  forest_ATKIS :: "Un"
  isA_ATKIS :: "Ob => Un => Ti => o"
  Stadtwald_CORINE :: "Ob"
  is_classified_as_CORINE :: "Ob => Un => Ti => o"
  mixed_forest_CORINE :: "Un"
  is_subsumed_by_CORINE :: "Un => Un => o"

axioms
  SentenceA : "E(Stadtwald_ATKIS,T2000) & isA_ATKIS(Stadtwald_ATKIS,forest_ATKIS,T2000)"
  SentenceB : "E(Stadtwald_CORINE,T2000) &
    is_classified_as_CORINE(Stadtwald_CORINE,mixed_forest_CORINE,T2000)"
  SentenceC : "is_subsumed_by_CORINE(mixed_forest,forest)"

  overlap_Stadtwald_ATKIS_CORINE : "O(Stadtwald_ATKIS,Stadtwald_CORINE,T2000)" (* (a0) *)
  discrete_Universal_forest : "DUn(forest)" (* (a6) *)

  align_is_classified_as_CORINE_to_BFO :
    "is_classified_as_CORINE(x,c,t) ==> Inst(x,c,t)" (* (a3) *)
  align_isA_ATKIS_to_BFO : "isA_ATKIS(x,c,t) ==> Inst(x,c,t)" (* (a3) *)
  align_is_subsumed_by_CORINE_to_BFO : "is_subsumed_by_CORINE(c,d) ==> IsA(c,d)" (* (a4) *)
  align_forest_ATKIS : "IsA(forest_ATKIS,forest)" (* (a1) *)
  align_mixed_forest_CORINE : "IsA(mixed_forest_CORINE,mixed_forest)" (* (a2) *)

theorem align_ATKIS_to_BFO: "Inst(Stadtwald_ATKIS,forest,T2000)"
  apply(insert SentenceA)
  apply(insert align_forest_ATKIS)
  apply(erule conjE)
  apply(drule align_isA_ATKIS_to_BFO)
  apply(drule_tac c="forest_ATKIS" and d="forest" in Inst_IsA_rule)
  apply(auto)
done

theorem align_CORINE_to_BFO: "Inst(Stadtwald_CORINE,forest,T2000)"
  apply(insert SentenceB, insert SentenceC)
  apply(insert align_mixed_forest_CORINE)
  apply(drule align_is_subsumed_by_CORINE_to_BFO)
  apply(auto)
  apply(drule align_is_classified_as_CORINE_to_BFO)
  apply(drule IsA_trans_rule [of "mixed_forest_CORINE" "mixed_forest" "forest"])
  apply(assumption)
  apply(drule Inst_IsA_rule [of "mixed_forest_CORINE" "forest"]) (* (a5) *)
  apply(auto)
done

theorem "Stadtwald_ATKIS = Stadtwald_CORINE"
  apply(insert align_ATKIS_to_BFO)
  apply(insert align_CORINE_to_BFO)
  apply(insert overlap_Stadtwald_ATKIS_CORINE)
  apply(insert discrete_Universal_forest)
  apply(insert DUn_and_O_impl_Id [where x="Stadtwald_ATKIS" and
    y="Stadtwald_CORINE" and c="forest" and t="T2000"])
  apply(auto)
done
end

```

Figure 6. The reasoning example (1) on page 5 as a formalized Isabelle theory. The labels (a0) – (a6) correspond to the respective labels on page 5.

7.3 Automated reasoning and rapid prototyping

Isabelle is a development tool optimized towards expressively and not towards efficient reasoning. (Isabelle's expressive power goes well beyond the expressive power of First-Order Logic.) Once one has developed a highly expressive theory, less expressive logics with better computational properties can be used to implement certain portions of the full theory for specific purposes. See (Bittner and Donnelly 2007c) for an extended discussion of how to use logics of different expressive power to deal with the trade off between expressive power and computational complexity. An OWL-based implementation of BFO can be found at <http://www.ifomis.uni-saarland.de/bfo>.

Although we use first-order predicate logic as object language for our top-level ontology, we use Isabelle for its computational realization. Isabelle is based on the functional language ML and has many features of other functional languages including HASKELL-style axiomatic type classes (Thompson 1999). There is a ten year history in which software tools such as HASKELL have been used successfully for the representation of geospatial ontologies and as tools for rapid prototyping of GIS data structures using algebraic specifications (Frank and Kuhn 1999, Frank 2001, Winter and Nittel 2003). In our use of Isabelle we go one step further along this road. The choice of the Isabelle tool will allow us to combine the merits of full first-order predicate logic as the language to express our top-level ontology, with the merits of strongly typed functional languages as tools for algebraic specifications.

Appendix: Related work

In this paper we presented a logic-based top-level ontology, which can be used as a tool to specify the semantics of top-level terms used in geographic domain ontologies. We also described the computational representation of the presented ontology. Our work draws on but goes beyond partial solutions in works such as (Simons 1987, Casati and Varzi 1999, Varzi 2003, 1996, Neuhaus *et al.* 2004, Bittner *et al.* 2004b, Grenon and Smith 2004, Bittner *et al.* 2004a, Guarino and Welty 2000b). Our work is complementary to work on semantic similarity measures such as (Fonseca *et al.* 2000, 2002b,a, Rodríguez and Egenhofer 2003, 2004).

Our focus on independent endurants and the relations between them allowed us: (a) to consider in greater detail the temporal behavior of the different categories of independent endurants; (b) to focus on geographically important notions such as partitions, mereological equivalence, and discreteness; (c) to focus on the introduction of mereological relations between universals which, as pointed out in Example 2 (Sec. 2.2), are important for geographic data standards. Our work benefits from a long tradition in philosophy which deals with questions of identity and change over time and the semantics of relations such as part-of, subclass-of, and instance-of (Simons 1987, Casati and Varzi 1999, Thomson 1983), and from recent work in knowledge representation, for example, by Guarino *et al.* (Guarino and Welty 2000b,a). This paper contributes to this literature by developing and implementing an axiomatic theory of time-dependent spatial relations in a logically rigorous manner in the context of reasoning about geographic information.

Since we focus only on independent endurants our ontology is incomplete. However we have outlined in Section 7 how the presented theory can be extended using, for example, the work described in (Hornsby and Egenhofer 2000) on the explicit description of change with respect to states of individual objects, and in (Grenon and Smith 2004), which proposes a spatio-temporal ontology incorporating interrelations between enduring entities and the processes in which they participate.

Our logic-based top-level ontology is a formalized logical theory consisting of axioms, definitions, and theorems which are expressed in the first-order predicate logic (FOL). In this it goes beyond those terminology systems specified using *non-logical* ontologies. Examples are ontologies stated in natural language as in the STDS and in the various ISO standards. Bittner *et al.* (2005) shows how the latter fail to provide the required rigor of the specification of the semantics of top-level terms of the sort discussed in this paper.

Alternative FOL-based top-level ontologies include the DOLCE (Gangemi *et al.* 2003, Masolo *et al.* 2004) and the SUMO top-level ontologies (Niles and Pease 2001). DOLCE is similar in spirit to BFO and thus to TNEMO-S-U. DOLCE, too, rests on a non-extensional temporal mereology. SUMO, on the other

hand, includes an atemporal extensional mereology, including time only indirectly through a HOLDS-AT predicate. This means that it cannot deal with more complex notions such as mereological equivalence. Both DOLCE and SUMO are much broader in scope when compared to our theory, but also much less detailed in their analysis of those mereological notions. TNEMO-S-U can also be easily extended to comprehend a variety of spatial relations which are relevant for geographic ontologies, as demonstrated in (Bittner and Donnelly 2007a,b).

For all the mereological similarities between DOLCE and TNEMO-S-U, both theories are clearly distinct. Neither DOLCE nor SUMO explicitly includes mereological relations between universals such as universal parthood. Moreover, DOLCE lacks the existence predicate and the corresponding axioms. It is the existence predicate, however, that allows us to formalize the important distinction between temporary and permanent spatial relations. For example, most trees are temporary parts of the forests they constitute, since forests usually outlive their individual trees. On the other hand the boundary of a forest is a permanent part of that forest, since both forest and its boundary must exist at the same times. These issues are discussed in detail in (Bittner and Goldberg 2007) and (Bittner and Donnelly 2007b).

Acknowledgements

Smith's work on this paper was funded in part by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant 1 U 54 HG004028. Information on the National Centers for Biomedical Computing can be found at <http://nihroadmap.nih.gov/bioinformatics>.

References

- ABDELMOTY, A.I., SMART, P.D., JONES, C.B., FU, G. and FINCH, D., 2005, A critical evaluation of ontology languages for geographic information retrieval on the Internet. *Journal of Visual Languages & Computing*, **16**, 331–358.
- AGARWAL, P., 2005, Ontological considerations in GIScience. *International Journal of Geographical Information Science*, **19**, 501–536.
- BAILEY, R.G., 1983, Delineation of ecosystem regions. *Environmental Management*, **7**, 365–373.
- BISHR, Y., 1998, Overcoming the semantic and other barriers to GIS interoperability. *International Journal of Geographical Information Science*, **12**, 299–314.
- BITTNER, T., “A Computational Realisation of Basic Formal Ontology (BFO). <http://www.ifomis.org/bfo/fo1>”, 2007a.
- BITTNER, T., 2007b, From top-level to domain ontologies: Ecosystem classifications as a case study. In *Proceedings of the Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. International Conference (COSIT 2007)*, S. Winter, M. Duckham, L. Kulik and B. Kuipers (Eds), pp. 61–77.
- BITTNER, T. and DONNELLY, M., 2006, A theory of granular parthood based on qualitative cardinality and size measures. In *Proceedings of the Proc. 4th Int. Conference on Formal Ontology in Information Systems*, B. Bennett and C. Fellbaum (Eds), pp. 65–76.
- BITTNER, T. and DONNELLY, M., 2007a, A formal theory of qualitative size and distance relations. In *Proceedings of the Proceedings of the 21 International Workshop on Qualitative Reasoning, QR2007*, C. Price (Ed.).
- BITTNER, T. and DONNELLY, M., 2007b, Logical properties of foundational mereotopological and adjacency relations in bio-ontologies. *Department of Philosophy, SUNY Buffalo*.
- BITTNER, T. and DONNELLY, M., 2007c, Logical properties of foundational relations in bio-ontologies. *Artificial Intelligence in Medicine*, **39**, 197–216.
- BITTNER, T. and DONNELLY, M., 2007d, A temporal mereology for distinguishing objects and portions of stuff. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence (AAAI-07)*, R. Holte and A. Howe (Eds).

- BITTNER, T., DONNELLY, M. and SMITH, B., 2004a, Endurants and Perdurants in Directly Depicting Ontologies. *AI Communications*, **14**, 247–258.
- BITTNER, T., DONNELLY, M. and SMITH, B., 2004b, Individuals, Universals, Collections: On the Foundational Relations of Ontology. In *Proceedings of the Proceedings of the third International Conference on Formal Ontology in Information Systems, FOIS04*, A. Varzi and L. Vieu (Eds) (IOS Press), pp. 37–48.
- BITTNER, T., DONNELLY, M. and WINTER, S., 2005, Ontology and semantic interoperability. In *Large-scale 3D data integration: Problems and challenges*, D. Proserpi and S. Zlatanova (Eds) (CRCpress (Taylor & Francis)), pp. 139–160.
- BITTNER, T. and GOLDBERG, L.J., 2007, The qualitative and time-dependent character of spatial relations in biomedical ontologies. *Bioinformatics*, doi: **10.1093/bioinformatics/btm155**.
- BUNT, H.C., 1985, *Mass terms and model-theoretic semantics* (New York : Cambridge University Press).
- BURGE, T., 1977, A Theory of Aggregates. *Nous*, **11**, 97–117.
- CASATI, R. and VARZI, A.C., 1999, *Parts and Places* (Cambridge, MA: MIT Press.).
- CASATI, R. and VARZI, A., 1995, The Structure of Spatial Localization. *Philosophical Studies*, **82**, 205–239.
- COHN, A.G. and HAZARIKA, S.M., 2001, Qualitative Spatial Representation and Reasoning: An Overview. *Fundamenta Informaticae*, **46**, 1–29.
- COHN, A., BENNETT, B., GOODDAY, J. and GOTTS, N., 1997, Qualitative Spatial Representation and Reasoning with the Region Connection Calculus. *geoinformatica*, **1**, 1–44.
- COPI, I., 1979, *Symbolic Logic* (Upper Saddle River, NJ 07458: Prentice Hall).
- DONNELLY, M. and BITTNER, T., 2005, Spatial relations between classes of individuals. In *Proceedings of the Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. International Conference (COSIT 2005)*, D. Mark and T. Cohn (Eds), no. 3693 in Lecture Notes in Computer Science (Springer Verlag), pp. 182 – 199.
- DONNELLY, M., BITTNER, T. and ROSSE, C., 2006, A formal theory for spatial representation and reasoning in bio-medical ontologies. *Artificial Intelligence in Medicine*, **36**, 1–27.
- DUCKHAM, M. and WORBOYS, M., 2005, An algebraic approach to automated geospatial information fusion. *International Journal of Geographical Information Science*, **19**, 537–557.
- EGENHOFER, M., 1991, Reasoning about Binary Topological Relations. In *2nd Symposium on Large Spatial Databases, SSD'91*, 525 of *Lecture Notes in Computer Science* (Zurich, Switzerland: Springer-Verlag), pp. 143–160.
- EGENHOFER, M. and SHARMA, J., 1993, Topological Consistency. In *Proceedings of the Proceedings of the 5th International Symposium on Spatial Data Handling*, P. Bresnahan, E. Corwin and D. Cowen (Eds), 1 (Charleston: IGU Commission of GIS), pp. 335–343.
- EGENHOFER, M.J. and FRANZOSA, R.D., 1991, Point-set topological spatial relations. *International Journal of Geographical Information Systems*, **5**, 161–174.
- FONSECA, F., EGENHOFER, M., AGOURIS, P. and CÂMARA, G., 2002a, Using Ontologies for Integrated Geographic Information Systems. *Transactions in GIS*, **6**, 231–257.
- FONSECA, F., EGENHOFER, M., DAVIS, C., and BORGES, K., 2000, Ontologies and Knowledge Sharing in Urban GIS. *Computer, Environment and Urban Systems*, **24**, 251–272.
- FONSECA, F., EGENHOFER, M., DAVIS, C. and CÂMARA, G., 2002b, Semantic Granularity in Ontology-Driven Geographic Information Systems. *Annals of Mathematics and Artificial Intelligence*, **36**, 121–151.
- FRANK, A.U., 2001, Tiers of ontology and consistency constraints in geographical information systems. *International Journal of Geographical Information Science*, **15**, 667–678.
- FRANK, A.U. and KUHN, W., 1999., A Specification Language for Interoperable GIS. In *Interoperating Geographic Information Systems*, M.F. Goodchild, M. Egenhofer, R. Fegeas and C. Kottman (Eds) (Norwell, MA: Kluwer, Norwell, MA), pp. 123–132.
- GALTON, A. and WORBOYS, M., 2005, Processes and Events in Dynamic Geo-Networks. In *Proceedings of the GeoSpatial Semantics: Proceedings of First International Conference, GeoS 2005*, A. Rodríguez, I. Cruz, S. Levashkin and M. Egenhofer (Eds), 3799 of *Lecture Notes in Computer Science* (Springer), pp. 45–59.
- GANGEMI, A., GUARINO, N., MASOLO, C., OLTRAMARI, A. and SCHNEIDER, L., 2003, Sweetening

- Ontologies with DOLCE. *AI Magazine*, **23**, 13–24.
- GRENON, P. and SMITH, B., 2004, SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition and Computation*, **4**, 69–103.
- GRUBER, T., 1993, A translation approach to portable ontology specification. *Knowledge Acquisition*, **5**, 199–220.
- GUARINO, N., 1998, Formal Ontology and Information Systems. In *Proceedings of the Formal Ontology and Information Systems, (FOIS'98)*, N. Guarino (Ed.) (IOS Press), pp. 3–15.
- GUARINO, N. and WELTY, C., 2000a, A Formal Ontology of Properties. In *Proceedings of the Proceedings of EKAW-2000: The 12th International Conference on Knowledge Engineering and Knowledge Management*, R. Dieng and O. Corby (Eds), LNCS (Spring-Verlag).
- GUARINO, N. and WELTY, C., 2000b, Identity, Unity, and Individuation: Towards a Formal Toolkit for Ontological Analysis. In *Proceedings of the Proceedings of ECAI-2000: The European Conference on Artificial Intelligence* (IOS Press, Amsterdam).
- HORNSBY, K. and EGENHOFER, M., 2000, Identity-Based Change: A Foundation for Spatio-Temporal Knowledge Representation. *International Journal of Geographical Information Science*, **14**, 207–224.
- KUHN, W., 2003, Semantic reference systems. *International Journal of Geographical Information Science*, **17**, 405–409.
- LOWE, E.J., 2002, *A survey of Metaphysics* (Oxford University Press).
- MANCARELLA, P., RAFFAET, A., RENSO, C. and TURINI, F., 2004, Integrating knowledge representation and reasoning in Geographical Information Systems. *International Journal of Geographical Information Science*, **18**, 417–447.
- MASOLO, M., BORGO, S., GANGEMINI, A., GUARINO, N., OLTRAMARI, A. and OLTRAMARI, A., 2004, WonderWeb Deliverable D18 – Ontology Library (final). Technical report, ISTC-CNR.
- NEUHAUS, F., GRENON, P. and SMITH, B., 2004, A formal theory of substances, qualities, and universals. In *Proceedings of the Proceedings of the third International Conference on Formal Ontology in Information Systems, FOIS04*, A. Varzi and L. Vieu (Eds), 114 of *Frontiers in Artificial Intelligence and Applications* (IOS Press), pp. 49–59.
- NILES, I. and PEASE, A., 2001, Towards a standard upper ontology. In *Proceedings of the FOIS '01: Proceedings of the international conference on Formal Ontology in Information Systems*, Ogunquit, Maine, USA (New York, NY, USA: ACM Press), pp. 2–9.
- NIPKOW, T., PAULSON, L.C. and WENZEL, M., 2002, *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*, LNCS Vol. 2283 (Springer).
- OLIVER, A., 1996, The Metaphysics of Properties. *Mind*, **105**, 1–80.
- PAULSON, L. and NIPKOW, T., “Isabelle homepage: <http://isabelle.in.tum.de/>”, 2005.
- PAULSON, L.C., 1994, *Isabelle: A Generic Theorem Prover* (Springer Verlag).
- RANDELL, D., CUI, Z. and COHN, A., 1992, A Spatial Logic Based on Regions and Connection. In *Proceedings of the Principles of Knowledge Representation and Reasoning. Proceedings of the Third International Conference (KR92)*, B. Nebel, C. Rich and W. Swartout (Eds) (Morgan Kaufmann), pp. 165–176.
- RODRÍGUEZ, A. and EGENHOFER, M., 2003, Determining Semantic Similarity Among Entity Classes from Different Ontologies. *IEEE Transactions on Knowledge and Data Engineering*, **15**, 442–456.
- RODRÍGUEZ, A. and EGENHOFER, M., 2004, Comparing Geospatial Entity Classes: An Asymmetric and Context-Dependent Similarity Measure. *International Journal of Geographical Information Science*, **18**, 229–256.
- SCHULZ, S. and HAHN, U., 2004, Parthood as Spatial Inclusion – Evidence from Biomedical Conceptualizations. In *Proceedings of the Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning (KR2004)*, pp. 55–63.
- SCHULZ, S., HAHN, U. and ROMACKER, M., 2000, Modeling anatomical spatial relations with description logics.. In *Proceedings of the AMIA 2000 – Proceedings of the Annual Symposium of the American Medical Informatics Association*, pp. 779 – 783.
- SDTS, 1997, American National Standard for Information Systems - Spatial Data Transfer Standard (SDTS) - Part 2, Spatial Features. Technical report.

- SIDER, T., 2001, *Four-Dimensionalism* (Clarendon Press, Oxford).
- SIMONS, P., 1987, *Parts, A Study in Ontology* (Oxford: Clarendon Press).
- SMITH, B., 2003, Ontology: An Introduction. In *Proceedings of the Blackwell Guide to the Philosophy of Computing and Information Blackwell Guide to the Philosophy of Computing and Information, 2003, 155–166*, L. Floridi (Ed.) (Oxford: Blackwell), chap. Ontology, pp. 155–166.
- SMITH, B., “Basic Formal Ontology”, 2007.
- SMITH, B., ASHBURNER, M. and ROSSE, C., 2007, The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, **25**, doi:10.1038/nbt1346.
- SMITH, B., CEUSTERS, W., KLAGGES, B., KÖHLER, J., KUMAR, A., LOMAX, J., MUNGALL, C., NEUHAUS, F., RECTOR, A. and ROSSE, C., 2005, Relations in Biomedical Ontologies. *Genome Biology*, **6**, r46.
- SMITH, B. and GRENON, P., 2004, The Cornucopia of Formal-Ontological Relations. *Dialectica*, **58**, 279–296.
- SMITH, B., KOEHLER, J. and KUMAR, A., 2004, On the Application of Formal Principles to Life Science Data: a Case Study in the Gene Ontology. In *Proceedings of the Data Integration in the Life Sciences*, E. Rahm (Ed.), 2994 of *LNBI* (Springer Verlag), pp. 79–94.
- SMITH, B. and ROSSE, C., 2004, The Role of Foundational Relations in the Alignment of Biomedical Ontologies. In *Proceedings of the Proceedings of the 11th World Congress on Medical Informatics*, M. Fieschi, E. Coiera and Y.J. Li (Eds), pp. 444–448.
- SMITH, B. and VARZI, A., 2000, Fiat and Bona Fide Boundaries. *Philosophy and Phenomenological Research*, **60**, 401–420.
- SOROKINE, A. and BITTNER, T., 2005, Understanding taxonomies of ecosystems: a case study. In *Proceedings of the Developments in Spatial Data Handling*, P. Fisher (Ed.) (Springer Verlag, Berlin), pp. 559–572.
- SOROKINE, A., BITTNER, T. and RENSCHER, C., 2006, Ontological investigation of ecosystem hierarchies and formal theory for multiscale ecosystem classifications. *geoinformatica*, **10**, 313–335.
- SPACKMAN, K., 2001, Normal forms for description logic expressions of clinical concepts in SNOMED RT. *Journal of the American Medical Informatics Association*, pp. 627–631.
- THOMPSON, S., 1999, *Haskell: The Craft of Functional Programming*, 2 (Addison-Wesley).
- THOMSON, J.J., 1983, Parthood and Identity Across Time. *Journal of Philosophy*, **80**, 201–220.
- VARZI, A., 1996, Parts, Wholes, and Part-Whole Relations: The Prospects of Mereotopology. *Data and Knowledge Engineering*, **20**, 259–86.
- VARZI, A., 2003, Mereology. In *Stanford Encyclopedia of Philosophy*, E.N. Zalta (Ed.) (Stanford: CSLI (internet publication)).
- VARZI, A.C., 2006, A Note on the Transitivity of Parthood. *Applied Ontology*, **1**, 141–146.
- VISSER, U., STUCKENSCHMIDT, H., WACHE, H. and VOEGELE, 2001, Using Environmental Information Efficiently: Sharing Data and Knowledge from Heterogeneous Sources. In *Environmental Information Systems in Industry and Public Administration*, C. Rautenstrauch and S. Patig (Eds) (Hershey, PA, IDEA Group), pp. 41–74.
- WINTER, S., 2001, Ontology: Buzzword or paradigm shift in GI science?. *International Journal of Geographical Information Science*, **15**, 587–590.
- WINTER, S. and NITTEL, S., 2003, Formal information modelling for standardisation in the spatial domain. *International Journal of Geographical Information Science*, **17**, 721–741.
- YETONGNON, K., SUWANMANEE, S., BENSLIMANE, D. and CHAMPIN, P.A., 2006, A web-centric semantic mediation approach for spatial information systems. *Journal of Visual Languages & Computing*, **17**, 1–24.