

Logical properties of foundational mereogeometrical relations in bio-ontologies

Thomas Bittner

*Departments of Philosophy and Geography, State University of New York at Buffalo,
New York State Center of Excellence in Bioinformatics and Life Sciences,
National Center for Geographic Information and Analysis
e-mail: bittner3@buffalo.edu*

Abstract. One aim of this paper is to improve the logical and ontological rigor of the OBO relation ontology by providing axiomatic specifications for logical properties of relations such as *part_of*, *located_in*, *connected_to*, *adjacent_to*, *attached_to*, etc. All of these relations are currently only loosely specified in OBO.

A second aim is to improve the expressive power of the relation ontology by including axiomatic characterizations of qualitative size relations such as (roughly-the-) *same-size-as*, *negligible-in-size-with-respect-to*, *same-scale*, etc. Those relations are important for comparing anatomical entities in a way that is compatible with the normal variations of their geometric properties. Moreover, qualitative size relations are important for distinguishing anatomical entities at different scales. Unfortunately, the formal treatment of those relations is difficult due to their context-dependent nature and their inherent vagueness. This paper presents a formalization that facilitates the separation of ontological aspects that are context-independent and non-vague from aspects that are context-dependent and subject to vagueness.

A third aim is to explicitly take into account the specific temporal properties of all of the relations and to provide a formalization that can be used as a basis for the formal representation of canonical anatomy as well as of instantiated anatomy.

All the relations and their properties are illustrated informally using a human synovial joint as a running example. At the formal level 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 accessible on <http://www.ifomis.org/bfo/fo1>.

Keywords: formal ontology, OBO, mereology, mereotopology, mereogeometry, qualitative representation and reasoning, vagueness, context

1. Introduction

There is widespread recognition that many existing biological and medical ontologies (or controlled vocabularies) can be improved by employing more rigorous logical methods (Rector and Horrocks, 1997; Schulze-Kremer, 1998; Smith *et al.*, 2003; Smith and Rosse, 2004; Rosse *et al.*, 2005). For this reason, the Open Biomedical Ontologies (OBO) consortium (OBO, 2006) has now added the criterion that the relations used to connect terms in OBO ontologies need to be applied in ways consistent with the OBO relation ontology (RO) (Smith *et al.*, 2005).

Unfortunately, the current version of the OBO relation ontology (the version published in (Smith *et al.*, 2005)) focuses on the definitions of class-level relations (i.e., relations between classes of entities such as the class of all right ventricles and the class of all hearts) and provides only a few axioms characterizing the relations between particular entities (individual-level relations) which are used to define the class-level relations.¹ Obviously, the logical properties of the class-level relations will not be fully specified until the individual-level relations used in their definitions are. Important individual-level relations whose logical properties are only loosely specified in (Smith *et al.*, 2005) are listed in Table 1. For example, Smith *et al.* (2005) only include axioms for reflexivity, antisymmetry, and transitivity for the parthood relation between particular entities (continuants, individuals, anatomical entities). These axioms are insufficient to

¹This is a general problem in current biomedical ontologies including the FMA (Rosse and Mejino, 2003), GALEN (Open-GALEN, 2003), and the Gene Ontology (The Gene Ontology Consortium, 2000), and in the literature on foundational relations in bio-medical ontologies including (Schulz *et al.*, 2000; Hahn *et al.*, 1998; Schulz and Hahn, 2001; Mejino *et al.*, 2003; Rogers and Rector, 2000).

specify the logical properties of `part_of` in a way that makes it possible to formally distinguish `part_of` from relations such as `contained_in`, `smaller_than`, `younger_than`, etc. (Bittner and Donnelly, 2007b).

name	RO relation			Symbol
	1. arg.	2. arg.	3. arg.	
<code>part_of</code>	continuant	continuant	time	P
<code>part_of</code>	region	region	–	P
<code>overlap</code>	continuant	continuant	time	O
<code>located_in</code>	continuant	region	time	L
<code>located_in</code>	continuant	continuant	time	$LocIn$
<code>contained_in</code>	continuant	continuant	time	$ContIn$
<code>adjacent_to</code>	continuant	continuant	time	Adj

Table 1

Relations of the relation ontology and the symbols used in the presented axiomatic theory to refer to those relations. (The argument types of the RO-relations are included since Smith *et al.* (2005) ‘overload’ `part_of` and `located_in`.)

In this paper an axiomatic theory is presented that more precisely specifies the logical properties of the relations in the OBO relation ontology. For example, unlike the RO, which introduces `adjacent_to` as a primitive relation without even giving axioms to specify the properties of this relations, in the presented ontology adjacency is defined by making use of some qualitative size relations which are not included in the OBO relation ontology. Thus in addition to offering a more precise analysis of the relations in the RO, the analysis shows how the RO might be expanded to incorporate more relations that are important for representing and reasoning about biological structures. The discussion in this paper shows how the qualitative size relation negligible-with-respect-to can be used formally to characterize OBO’s `adjacent_to` relation.

The formal ontology is developed within the framework of a temporal mereogeometry with location relations which includes a temporal mereology and a temporal mereotopology as sub-theories. The temporal mereogeometry is extended with qualitative size relations (Bittner and Donnelly, 2006, 2007a). The underlying temporal framework allows one to distinguish time-dependent (i.e., changeable) and permanent versions of all relations. This distinction is important since the OBO relation ontology is intended to be applicable in all of the following areas:

- (i) In canonical anatomy, where mereotopological relations as well as adjacency and attachment relations are typically permanent (bones do not break, ligaments are permanently attached to bones, etc.) while ordering and distance relations (the spatial arrangement of anatomical parts with respect to one another) often change over time (i.e., are not permanent – the heart beats, the jaw opens and closes, etc.) (Bittner and Goldberg, 2007).
- (ii) In clinical contexts where one deals with actually instantiated anatomical structures (Neal *et al.*, 1998) that do undergo non-normal changes of mereotopological, adjacency, and attachment relations (Ceusters and Smith, 2005). Actually instantiated anatomical structures encounter injuries and diseases but also undergo healing and surgery processes. All these processes may change mereotopological relations as well as adjacency and attachment relations. For example, bones get broken and are put back together, muscles and bones get detached and reattached again.
- (iii) To understand processes that occur in anatomical structures (canonical and instantiated) mereogeometrical relations between anatomical entities of different scale need to be taken into account. Relations between entities of different scale are often NOT permanent. It is normal for organisms to gain and to lose microscopic parts all the time.

In the presented ontology the formal foundations are laid for dealing with these different kinds of temporal behaviors.

2. A running example and requirements for the formal theory

It will be helpful to present the axiomatic theory in the context of a running example. This example is also meant to illustrate the complexities the presented formal theory has to deal with.

2.1. Anatomical structure and foundational relations

Figure 1 depicts the index finger of Joe Doe's left hand in a straight (Figure 1(a)) and in a bent position (Figure 1(b)). The major parts of the depicted finger include the proximal, middle, and distal phalanges and three hinged joints: distal interphalangeal (DIP), proximal interphalangeal (PIP), and metacarpophalangeal (MCP).

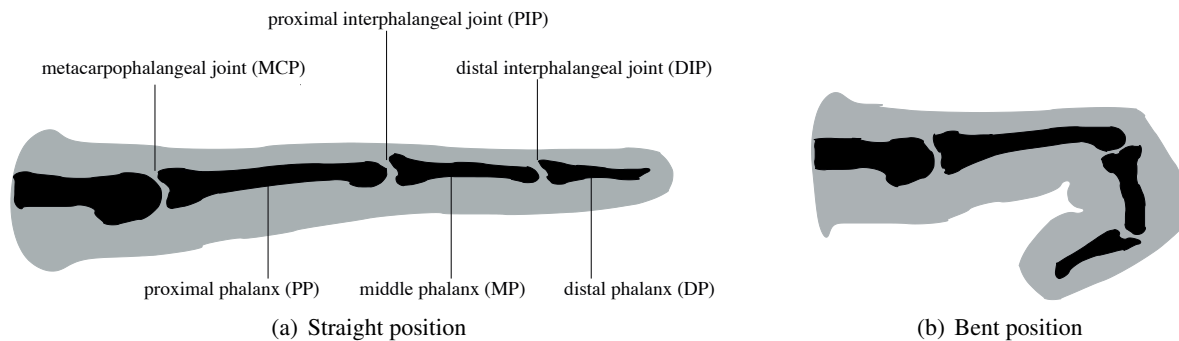


Fig. 1. Joe Doe's left index finger.

The examples used in what follows focus on the proximal interphalangeal joint (PIP) of Joe's left index finger (Joe's PIP for short). This is a synovial joint whose major parts are depicted in Figure 2(a): the bony part of the middle phalanx (MPB), the articular cartilage of the middle phalanx (AC-MP), the bony part of the proximal phalanx (PPB), the articular cartilage of the proximal phalanx (AC-PP), the ligament, the synovial membrane, the synovial cavity, and the synovial fluid. A synovial joint is a movable joint, which contains synovial fluid in a synovial cavity. The fluid acts as a lubricant to allow the surface of the cartilage-capped bones that meet in the joint cavity to slide freely along one another (Stevens and Lowe, 2005). The bony part of the middle phalanx (MPB) and articular cartilage of the middle phalanx (AC-MP) together with the the articular cartilage of the middle phalanx in the distal interphalangeal joint (DIP) form the middle phalanx (MP). Similarly, the bony part of the proximal phalanx (PPB) and the articular cartilage of the proximal phalanx (AC-PP) together with the the articular cartilage of the middle phalanx in the metacarpophalangeal joint (MCP) form the proximal phalanx (PP).

In the graph (Figure 2(b)) the salient material parts of the proximal interphalangeal joint of Joe's left index finger are indicated by the corresponding labels of the nodes. The solid edges indicate that the anatomical entities at the end of the edges are externally connected, i.e., have zero distance but do not overlap. The distance between extended spatial objects is here understood as the greatest lower bound of the distances between any point of the region occupied by the first object and any point of the region occupied by the second object. External connection holds, for example, between the synovial cavity and the synovial membrane, the articular cartilage of the proximal phalanx and the bony part of the proximal phalanx, and so on.

The relation *adjacent_to* holds among material anatomical entities that are a very small but non-zero positive distance apart. (See also (Smith and Varzi, 2000; Bittner and Goldberg, 2007).) More precisely, the distance between two adjacent entities is non-zero, but negligible with respect to their size. Thus two adjacent anatomical entities have parts that are very close but are not connected. For example, the bony part of the proximal phalanx of Joe's left index finger (PPB) is adjacent to the synovial membrane of the proximal interphalangeal joint. This is because even though some of their parts are very close (so close

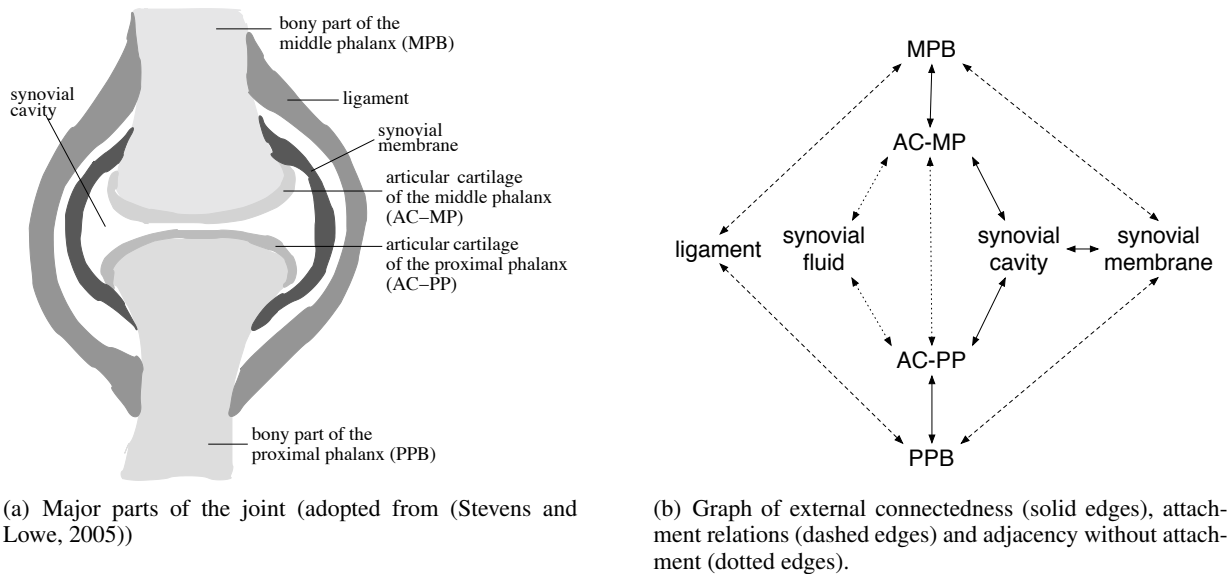


Fig. 2. The proximal interphalangeal joint of Joe Doe's left index finger.

that the synovial fluid remains inside the space enclosed by the membrane and the bone), there is always a small but positive distance between both objects. Similarly, the bony part of the proximal phalanx (PPB) is adjacent to the ligament of the proximal interphalangeal joint. It is also the case that the articular cartilage of the proximal phalanx (AC-PP) is adjacent to the articular cartilage of the medial phalanx (AC-MP). In the graph in Figure 2(b) the relation of adjacency is represented by dashed and dotted edges.

These examples show that, in order to be capable of representing adjacency relations between material parts of anatomical structures such as synovial joints, a bio-medical ontology needs to incorporate qualitative distance relations such as close-to (which will be analyzed in terms of the relation negligible-in-size-with-respect-to in Section 8).

2.2. Temporal change

The discussion in this subsections will illustrate that bio-medical ontologies need the formal resources to deal with a variety of time-dependent and time-independent (permanent) versions of foundational relations.

Time-dependent vs. permanent relations. As already mentioned, most mereotopological relations and many adjacency relations in canonical anatomy are permanent. That is, they hold between their relata at all times at which the relata exist. By ' x exists at time t ' is meant in the context of this paper that x is an anatomical entity that is part of a living organism. Barring a disruption in the normal functioning of Joe's PIP, all relations depicted in Figure 2(b) are permanent in the sense that these relations hold at all times at which the PIP as a whole exists as part of Joe's living body.

Since bio-medical ontologies are not only about canonical anatomical structures (e.g. synovial joints) but also about actually instantiated structures (e.g. Joe's PIP), one needs to take into account that, other than canonical anatomical structures, actually instantiated anatomical structures do undergo non-normal changes, i.e., encounter injuries and diseases. For example at time t the articular cartilages of the proximal and the medial phalanges (AC-PP and AC-MP) are adjacent. At time t_2 Joe gets into a fist fight and his PIP gets dislocated such that at t_2 the articular cartilages of the proximal and the medial phalanges are NOT adjacent. At time t_3 after surgery and a successful healing process the articular cartilages of the proximal and the medial phalanges are adjacent again.

Kinds of permanent adjacency relations. In addition to the general distinction of time-dependent and permanent relations it is important to distinguish permanent adjacency relations of different strength: The proximal phalanx of Joe's left index finger is permanently adjacent to the ligament of the proximal interphalangeal joint in the sense that the location of both is *fixed* with respect to one another. By contrast, the articular cartilage of the proximal phalanx (AC-PP) and the articular cartilage of the medial phalanx (AC-MP) are (permanently) adjacent, but both have the capability of *sliding* along one another. This can be seen easily by comparing the relative location of Joe's proximal and phalanges in Figures 1(a) and 1(b). In the graph in Figure 2(b), the kind of adjacency relation that permits sliding is represented by dotted edges. (Similar points are made by Bittner and Goldberg (2007) using the example of the human temporomandibular joint (TMJ).)

Granular parts. Relations between anatomical entities at different scale (at different levels of granularity) often are not permanent. Consider, for example, the synovial membrane at the inside of the synovial ligament of Joe's PIP and the particular formation of cells that constitutes this membrane at a given time. It is normal that the synovial membrane persists through time but gains and loses cells over time. Cells are *granular* parts, i.e., parts that are negligible in size with respect to their wholes. Cells are often also non-permanent parts, i.e., are outlived by the wholes they form. This example shows that, even in canonical anatomy, relations between anatomical entities at different levels of granularity are time-dependent and not permanent.

2.3. Formal representation

Conventions. For the reasons discussed above for each relation that holds between objects (continuants, anatomical entities) time-dependent and time-independent (permanent) versions are distinguished. Formally, the ontology is presented in a sorted first-order predicate logic with identity. Variables range over individual objects (anatomical entities that are parts of living organisms – a specific kind of continuant in the sense of the RO), regions of space, and instants of time. Variables x, y, z, w are used for objects, variables a, b, c, d for regions and variables t, t_1, t_2 for instants of time. It is assumed that the axiomatic apparatus that governs the sub-domain of time instants is an independent component of the relation ontology and is not addressed in this paper. Ontologically, time instances represent time slices of four-dimensional space time (Pianesi and Varzi, 1996; Bittner and Donnelly, 2004). For an overview of existing temporal formalisms see (Galton, 1987, 1999).

In order to be able to state definitions, axioms, and theorems in a readable way, the symbols listed in Table 1 are used in the theory for the respective RO-relations. The logical connectors $\neg, =, \wedge, \vee, \rightarrow, \leftrightarrow$ have their usual meanings: not, identical-to, and, or, if ... then, if and only if (iff). The symbol \equiv is used for definitions. (x) symbolizes universal quantification and $(\exists x)$ symbolizes existential quantification. Leading universal quantifiers are generally omitted. Names of axioms begin with the capital letter 'A', names of definitions begin with the capital letter 'D', and names of theorems begin with the capital letter 'T'.

Modality. In addition to the (temporal) distinction between time-dependent and permanent relations there is a (modal) distinction between relations that hold necessarily and relations that hold contingently. The temporal distinctions are orthogonal to the modal distinctions. There are necessary and contingent permanent relations and there are necessary and contingent time-dependent relations between anatomical objects. They are found in both, canonical as well as in instantiated anatomy. The distinction of modal aspects is important to specify further differences between canonical and instantiated anatomy. Unfortunately the formal representation of modal aspects of relations requires a more powerful representational framework than a first-order predicate logic with the sorts specified above.

To formally address the distinction between relations that hold necessarily and relations that hold contingently one has to (i) introduce additional sorts of variables for possible worlds and a corresponding axiomatic apparatus (Lewis, 1986) or (ii) to use a modal predicate logic (Hughes and Cresswell, 2004). Both options go beyond the scope of this paper.

3. Mereology of objects

In this section the part of the formal ontology which characterizes the logical properties of time-dependent and time-independent versions of the OBO relations *part_of* and *overlaps* is presented. An extended discussion of various options of developing temporal mereologies can be found in (Simons, 1987).

3.1. Time-dependent parthood

Following Bittner *et al.* (2004, to appear), in this sub-section a temporal version of mereology based on the ternary primitive P is presented, where $P\ xyt$ is interpreted as: object x is part of object y at time-instant t . For example this blood cell was part of my body yesterday, but it is not a part of my body now.

The relations of proper parthood and overlap among objects are defined in the usual way: x is proper part of y at t if and only if x is part of y at t and y is not part of x at t (D_{PP}); x overlaps y at t if and only if there is an object z such that z is part of x at t and z is part of y at t (D_O). P is used to distinguish the time-instants at which objects exist: x exists at time t (symbolically $E\ xt$) if and only if x is a part of itself at t (D_E).

$$\begin{array}{ll} D_{PP} & PP\ xyt \equiv P\ xyt \wedge \neg P\ yxt \\ D_O & O\ xyt \equiv (\exists z)(P\ zxt \wedge P\ zyt) \end{array} \qquad D_E \quad E\ xt \equiv P\ xxt$$

The following axioms are added. For every object there is some time at which it exists ($AM1$). At each time instant, the individual parthood relation is transitive ($AM2$). If x is a part of y at t then both x and y exist at t ($AM3$). If x exists at t and everything that overlaps x at t also overlaps y at t then x is a part of y at t ($AM4$).

$$\begin{array}{ll} AM1 & (\exists t)E\ xt \\ AM2 & P\ xyt \wedge P\ yzt \rightarrow P\ xzt \end{array} \qquad \begin{array}{ll} AM3 & P\ xyt \rightarrow E\ xt \wedge E\ yt \\ AM4 & E\ xt \wedge (z)(O\ zxt \rightarrow O\ zyt) \rightarrow P\ xyt \end{array}$$

One can prove that the following statements hold at all times: overlap is symmetric; proper parthood is asymmetric and transitive.²

Notice that there are significant differences between axioms $AM1$ - $AM4$ and the axioms for parthood in the relation ontology: (i) the RO does not have an existence predicate; (ii) the RO does not have axiom $AM4$; (iii) unlike the RO this ontology does not include an antisymmetry axiom. An existence predicate may not be so important in canonical anatomy, however it is important for representation of facts about instantiated anatomical structures in clinical contexts (the tumor did not exist at t_1 but it did exist at t_2). Consider axiom $AM4$. Using $AM4$ one can prove that if x is a proper part of y at t then there is a z such that z is a proper part of y at t and z does not overlap x at t ($TM1$). From this immediately follows that no object can have a single proper part. This, clearly, is an important logical property of the *part_of* relation that is not covered by the axioms of the RO. (See also (Simons, 1987; Varzi, 2003; Bittner and Donnelly, 2007b).)

$$TM1 \quad PP\ xyt \rightarrow (\exists z)(PP\ zyt \wedge \neg O\ xyt)$$

No axiom of antisymmetry for parthood among objects is included in the theory in order to leave open the possibility that there may be distinct objects which have exactly the same parts at a given time. For example, the FMA distinguishes between an anatomical entity and the tissue which constitutes that entity at a given time (Rosse and Mejino, 2003). Often, however, it is the case that if tissue x constitutes

²Theorems about standard properties of relations such as reflexivity, transitivity, symmetry, asymmetry, etc. are not explicitly listed in this paper. The theorems and their proofs, however, can be found in the computational representation of the theory. (See also Section 10.)

anatomical entity y at time t then x and y have the same parts at t . Consider, for example the synovial membrane of Joe's PIP and the tissue – the particular formation of cells – that constitutes this membrane at time t . At this time the synovial membrane and the tissue have exactly the same parts: the same cells, the same molecules, etc. but still one might want to hold that membrane and tissue are distinct entities since the synovial membrane may be constituted by different tissues (i.e., different cells) at different times. For an extended discussion see also (Bittner and Donnelly, 2007c; Donnelly and Bittner, 2008).

3.2. Time-independent mereological relations

In terms of the time-dependent basic mereological relations one can define corresponding time-independent relations. Object x is a *permanent* part of object y if and only if whenever x or y exist, x is a part of y (D_{pP}). Object x is a *permanent proper part* of object y if and only if whenever x or y exist, x is a proper part of y (D_{pPP}). Entity x *permanently overlaps* entity y if and only if at all time at which x or y exist, x overlaps y (D_{pO}).

$$\begin{aligned} D_{pP} \quad pP \ xy &\equiv (t)((E \ xt \vee E \ yt) \rightarrow P \ xyt) \\ D_{pPP} \quad pPP \ xy &\equiv (t)((E \ xt \vee E \ yt) \rightarrow PP \ xyt) \\ D_{pO} \quad pO \ xy &\equiv (t)((E \ xt \vee E \ yt) \rightarrow O \ xyt) \end{aligned}$$

One can prove that permanent parthood is reflexive and transitive, that permanent proper parthood is asymmetric and transitive, and that permanent overlap is reflexive and symmetric. One can also prove that if x is a permanent part of y then x exists at t if and only if y exists at t (*TM2*), i.e., x and y exist at the same times. Similar theorems can be derived for all permanent relations introduced in this paper.

$$TM2 \quad pP \ xy \rightarrow (t)(E \ xt \leftrightarrow E \ yt)$$

As pointed out above, most parthood relations between macroscopic anatomical entities in canonical anatomy are permanent relations. The permanent parts of Joe's PIP of macroscopic scale are: the bony part of the middle phalanx (MPB), the articular cartilage of the middle phalanx (AC-MP), the bony part of the proximal phalanx (PPB), the articular cartilage of the proximal phalanx (AC-PP), the ligament, the synovial membrane, the synovial cavity, and the synovial fluid.

3.3. The importance of distinguishing time-dependent and permanent mereological relations

It is important to note that the logical relations between pP and pPP are not exactly analogous to those between P and PP : As stated in (D_{pPP}) x is a proper part of y at t if and only if x is a part of y at t and y is not a part of x at t . One can prove in the presented theory that if x is a permanent proper part of y then x is a permanent part of y and y is not a permanent part of x (*TM3*).

$$TM3 \quad pPP \ xy \rightarrow pP \ xy \wedge \neg pPyx$$

But one cannot prove the converse of this theorem, i.e., it is not always the case that if x is a permanent part of y and y is not a permanent part of x then x is a permanent proper part of y .

To see that the converse of *TM3* cannot be a theorem, consider Figure 3 which depicts Joe's left index finger (*JLIF*) at different points in time. At time t_1 , *FS* is proper part of *JLIF* as depicted in Figure 3(a). Suppose that Joe has an accident between t_1 and t_2 in which the distal part of his left index finger gets destroyed in a way such that at time t_2 Joe's left index finger is as depicted in Figure 3(b). Thus, while at t_1 *FS* is a proper part of *JLIF*, at t_2 *FS* is identical to *JLIF*. If Joe does not have any further accidents then the following holds: (i) whenever *FS* or *JLIF* exist, *FS* is a part of *JLIF* (i.e., *FS* is a permanent part of *JLIF*); (ii) it is not the case that whenever *FS* or *JLIF* exist, *JLIF* is a part of

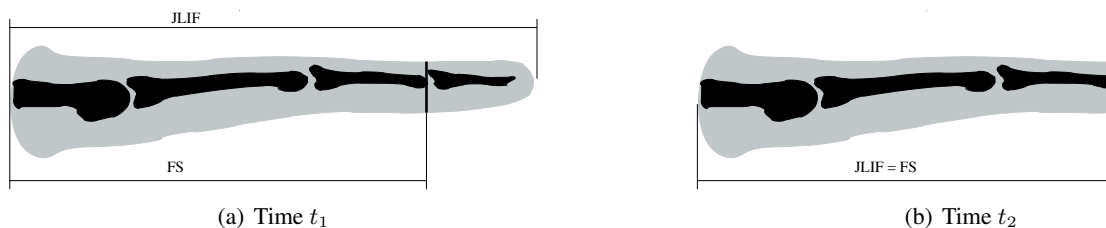


Fig. 3. Joe Doe's left index finger at different times.

FS (i.e., $JLIF$ is not a permanent part of FS). But clearly, FS is not a permanent proper part of $JLIF$. Hence the converse of $(TM3)$ cannot be a theorem.

Similarly, the logical relations between pO and pP are not exactly analogous to those between O and P (Donnelly, 2007). One can prove in the presented theory that if there is a z such that z is a permanent part of x and a permanent part of y then x and y permanently overlap ($TM4$). But one cannot prove the converse of this theorem. This is because permanently overlapping objects do not need to share permanent parts. Objects can overlap permanently by sharing different non-permanent parts at different times.

$$TM4 \quad (\exists z)(pP \ zx \wedge pP \ zy) \rightarrow pO \ xy$$

This discussion shows that it is important to explicitly distinguish permanent (parthood, overlap, etc.) relations from time-dependent (parthood, overlap, etc.) relations. This is particularly important since in ontologies formalized using description logics (Baader *et al.*, 2002) it is difficult to represent ternary relations (Grenon, 2006). Therefore often exclusively binary relation symbols are used. The temporal parameters of time-dependent relations are generally omitted and the symbols referring to time-dependent relations are assumed to be implicitly time-indexed. This means that at the level of the formal language time-dependent and permanent relations cannot be distinguished by the presence of a temporal parameter or the lack thereof. In formalisms that exclusively work with binary relations it is therefore important to make explicit whether a given binary relation symbol is implicitly time-indexed and refers to a time-dependent (parthood, overlap, etc.) relation, or whether a given binary relation symbol refers to a permanent (parthood, overlap, etc.) relation which, indeed, is binary.

4. Scale-sensitive mereogeometry of regions

The size of objects, their relative location with respect to one another, as well as connection, adjacency, and attachment relations – all holding at given times – are characterized in the presented ontology in terms of relations between the spatial regions at which these objects are located at those times (Casati and Varzi, 1999). As in the RO, spatial regions are assumed to be the parts of an independent non-changing background space in which all objects are located. Thus objects may change their location, their size, their shape by being located at different regions at different times, while the regions themselves do not change. For this reason the proposed ontology includes an *atemporal mereogeometry* of regions.

The ontology includes a mereogeometry of regions and not just a mereotopology, since adjacency is a kind of distance relation which cannot be specified in mere mereotopological terms. Using the mereogeometry, scale-dependent notions are introduced which are critical for the definition of adjacency as a relation implying close distance.

Technically, the mereological basis is a standard extensional mereology (Simons, 1987; Varzi, 1996). Following Bittner and Donnelly (2006, 2007a) this basis will be extended by a same-volume-size relation and a sphere primitive and scale will be introduced using the relations roughly-the-same-volume-size and negligible-in-volume-size. The resulting mereogeometry is similar in spirit to those developed by Tarski (1956), Bennett *et al.* (2000), and Schmidtke (2005). It differs from other approaches in the ways in which

it integrates relations such as roughly-the-same-size and negligible-in-size. An exact formal analysis of commonalities and differences of mereogeometries is not easy (Borgo and Masolo, 2007) and goes beyond the scope of this paper. An overview of other mereogeometries can be found in (Borgo and Masolo, 2007).

4.1. Mereology

Letters from the beginning of the alphabet are used as variables for regions and the Sans serif font for predicates whose parameters range over regions. A binary primitive \mathbf{P} is introduced, where $\mathbf{P} ab$ is interpreted as: region a is part of region b . In the standard ways the following definitions can be stated in terms of \mathbf{P} :

Region a is a proper part of region b if and only if a is part of b and b is not part of a ($D_{\mathbf{PP}}$); regions a and b overlap if and only if they share a common region as a part ($D_{\mathbf{O}}$); region c is the sum of a and b if and only if for all d , d overlaps c if and only if d overlaps a or d overlaps b (D_{+}); region c is the difference of b in a if and only if any object d overlaps c if and only if d overlaps some part of a and that does not overlap b (D_{-}). (If regions are modeled as point sets then ‘+’ is like set-union and ‘-’ is like set-difference.)

$$\begin{aligned} D_{\mathbf{PP}} \quad \mathbf{PP} ab &\equiv \mathbf{P} ab \wedge \neg \mathbf{P} ba \\ D_{\mathbf{O}} \quad \mathbf{O} ab &\equiv (\exists c)(\mathbf{P} ca \wedge \mathbf{P} cb) \\ D_{+} \quad +abc &\equiv (d)(\mathbf{O} dc \leftrightarrow (\mathbf{O} da \vee \mathbf{O} db)) \\ D_{-} \quad -abc &\equiv (d)(\mathbf{O} dc \leftrightarrow (\exists d_1)(\mathbf{P} d_1 a \wedge \neg \mathbf{O} d_1 b \wedge \mathbf{O} d_1 d)) \end{aligned}$$

The following axioms are included in the theory: \mathbf{P} is reflexive ($AR1$), \mathbf{P} antisymmetric ($AR2$); \mathbf{P} is transitive ($AR3$); if everything that overlaps u also overlaps v then u is a part of v ($AR4$); \mathbf{P} is reflexive ($AR1$); and if a is not a part of b then there is a region c which is the difference of b in a ($AR4$); for any regions a and b there is a region c that is the sum of a and b ($AR5$).

$$\begin{aligned} AR1 \quad \mathbf{P} aa & & AR4 \quad \neg \mathbf{P} ab \rightarrow (\exists c)(-abc) \\ AR2 \quad \mathbf{P} ab \wedge \mathbf{P} ba \rightarrow a = b & & AR5 \quad (\exists c)(+abc) \\ AR3 \quad \mathbf{P} ab \wedge \mathbf{P} bc \rightarrow \mathbf{P} ac & & \end{aligned}$$

The following theorems can be derived: if everything that overlaps a also overlaps b , then a is a part of b ($TR1$); a and b are identical if and only every c overlaps a if and only if c overlaps b ($TR2$). One can also prove that sums and differences are unique whenever they exist ($TR3$ - $TR4$). Thus $AR5$ in conjunction with $TR3$ ensure that summation is a functional operator (mapping any pair of regions to their unique sum).

$$\begin{aligned} TR1 \quad (c)(\mathbf{O} ca \rightarrow \mathbf{O} cb) \rightarrow \mathbf{P} ab & & TR3 \quad +abc_1 \wedge +abc_2 \rightarrow c_1 = c_2 \\ TR2 \quad a = b \leftrightarrow (c)(\mathbf{O} ca \leftrightarrow \mathbf{O} cb) & & TR4 \quad -abc_1 \wedge -abc_2 \rightarrow c_1 = c_2 \end{aligned}$$

Notice that the mereology of regions is quite different from the mereology of objects: \mathbf{P} is time independent, \mathbf{P} is antisymmetric and thus two regions are identical if and only if they have the same parts and overlap the same regions.

4.2. Size ordering

A binary primitive \sim is included in the theory, where, on the intended interpretation, $a \sim b$ holds if and only if regions a and b have the same volume size. In terms of \sim one can define that the size of a is less than or equal to the size of b if and only if there is a region c that is a part of b and has the same size as a (D_{\leq}).

$$D_{\leq} \quad a \leq b \equiv (\exists c)(c \sim a \wedge \mathbf{P} cb)$$

On the intended interpretation, $a \leq b$ holds if and only if the volume size of a is less than or equal to the volume size of b .

The following axioms are included: \sim is reflexive (AR6); \sim is symmetric (AR7); \sim is transitive (AR8); if a is part of b and a and b have the same size then b is part of a (AR9); for any a and b , the size of a is less than or equal to the size of b or the size of b is less than or equal to the size of a (AR10); if the size of a is less than or equal to the size of b and the size of b is less than or equal to the size of a , then a and b have the same size (AR11).

$$\begin{array}{ll}
 \text{AR6} & a \sim a \\
 \text{AR7} & a \sim b \rightarrow b \sim a \\
 \text{AR8} & a \sim b \wedge b \sim c \rightarrow a \sim c \\
 \text{AR9} & \mathbf{P} ab \wedge a \sim b \rightarrow \mathbf{P} ba \\
 \text{AR10} & a \leq b \vee b \leq a \\
 \text{AR11} & a \leq b \wedge b \leq a \rightarrow a \sim b
 \end{array}$$

In the theory one can prove: if a is identical to b , then a and b are of the same size (TR5); if a is part of b and b is part of a , then a and b have the same size (TR6); if a is part of b and a and b have the same size then a and b are identical (TR7); if a is a part of b , then the size of a is less than or equal to the size of b (TR8); \leq is reflexive (TR9); \leq is transitive (TR10); if the size of a is less than or equal to the size of b and b and c have the same size, then the size of a is less than or equal to the size of c (TR11); if c and a have the same size and the size of c is less than or equal to the size of b then the size of c is less than or equal to the size of a (TR12).

$$\begin{array}{ll}
 \text{TR5} & a = b \rightarrow a \sim b \\
 \text{TR6} & \mathbf{P} ab \wedge \mathbf{P} ba \rightarrow a \sim b \\
 \text{TR7} & \mathbf{P} ab \wedge a \sim b \rightarrow a = b \\
 \text{TR8} & \mathbf{P} ab \rightarrow a \leq b \\
 \text{TR9} & a \leq a \\
 \text{TR10} & a \leq b \wedge b \leq c \rightarrow a \leq c \\
 \text{TR11} & a \leq b \wedge b \sim c \rightarrow a \leq c \\
 \text{TR12} & c \sim a \wedge a \leq b \rightarrow c \leq b
 \end{array}$$

Thus, \sim is an equivalence relation, \leq is reflexive and transitive, and \sim , \leq , \mathbf{P} , and $=$ are logically interrelated in the expected ways. For more detailed discussions see (Bittner and Donnelly, 2006, 2007a).

4.3. Spheres

The primitive predicate \mathbf{S} is included in the theory, where ‘ $\mathbf{S} a$ ’ is interpreted as a is a sphere. In terms of \mathbf{S} one can define: Region a is *maximal with respect to b in c* if and only if (i) a , b , and c are spheres, (ii) a and b are non-overlapping parts of c , and (iii) every sphere u that has a as a part either is identical to a , overlaps b , or is not a part of c ($D_{\mathbf{Mx}}$) (Fig. 4(i)). Region a is a *concentric proper part of b* if and only if (i) a and b are spheres, (ii) a is a proper part of b and (iii) all spheres that are maximal with respect to a in b have the same size ($D_{\mathbf{CoPP}}$) (Fig. 4(ii)).

$$\begin{array}{ll}
 D_{\mathbf{Mx}} & \mathbf{Mx} abc \equiv \mathbf{S} a \wedge \mathbf{S} b \wedge \mathbf{S} c \wedge \mathbf{P} ac \wedge \mathbf{P} bc \wedge \neg \mathbf{O} ab \wedge \\
 & (d)(\mathbf{S} d \wedge \mathbf{P} ad \rightarrow (a = d \vee \mathbf{O} db \vee \neg \mathbf{P} dc)) \\
 D_{\mathbf{CoPP}} & \mathbf{CoPP} ab \equiv \mathbf{S} a \wedge \mathbf{S} b \wedge \mathbf{PP} ab \wedge (d)(e)(\mathbf{Mx} dab \wedge \mathbf{Mx} eab \rightarrow d \sim e)
 \end{array}$$

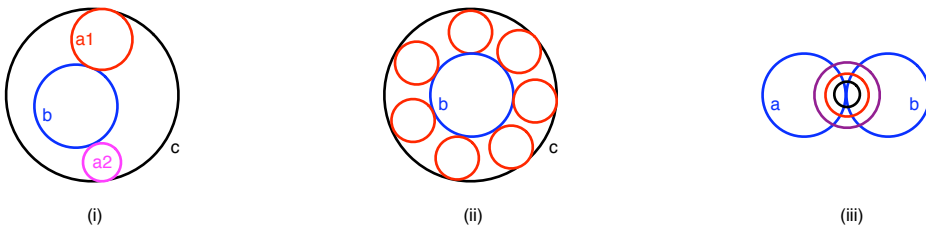


Fig. 4. Illustrations for (i) $D_{\mathbf{Mx}}$: $\mathbf{Mx} a_1bc \wedge \mathbf{Mx} a_2bc$, (ii) $D_{\mathbf{CoPP}}$: $\mathbf{CoPP} bc$, and (iii) $D_{\mathbf{C}}$: $\mathbf{C} ab$.

The following spheres are required to exist: Every region has a sphere as a part (AR12). Every sphere has a concentric proper part (AR13). If sphere a is a proper part of sphere b then there is a sphere c that is maximal with respect to a in b (AR14).

$$\begin{aligned} AR12 & (\exists c)(\mathbf{S} c \wedge \mathbf{P} ca) \\ AR13 & \mathbf{S} a \rightarrow (\exists b)(\mathbf{S} b \wedge \mathbf{CoPP} ba) \\ AR14 & \mathbf{S} a \wedge \mathbf{S} b \wedge \mathbf{PP} ab \rightarrow (\exists c)(\mathbf{Mx} cab) \end{aligned}$$

4.4. Connectedness relations between regions

Similar to Bennett *et al.* (2000) the connectedness relation is defined as follows: two regions a and b are *connected* if and only if there is a sphere c that overlap a and b and all spheres that are concentric proper parts of c also overlap a and b (D_C) (Fig. 4(iii)).

$$D_C \quad \mathbf{C} ab \equiv (\exists c)(\mathbf{S} c \wedge \mathbf{O} ca \wedge \mathbf{O} cb \wedge (d)(\mathbf{CoPP} dc \rightarrow (\mathbf{O} da \wedge \mathbf{O} db)))$$

On the intended interpretation, the connection relation \mathbf{C} holds between regions a and b if and only if the distance between them is zero (where the distance between regions is here understood as the greatest lower bound of the distance between any point of the first region and any point of the second region).

One can prove that \mathbf{C} is reflexive (TR13), symmetric (TR14), and that if a is part of b , then everything connected to a is connected to b (TR15). These are the usual axioms for \mathbf{C} in a mereotopological framework (Varzi, 1996).

$$\begin{aligned} TR13 & \mathbf{C} aa & TR15 & \mathbf{P} ab \rightarrow (c)(\mathbf{C}ca \rightarrow \mathbf{C}cb) \\ TR14 & \mathbf{C} ab \rightarrow \mathbf{C} ba & TR16 & \mathbf{O} ab \rightarrow \mathbf{C} ab \\ & & TR17 & \mathbf{P} ab \wedge \mathbf{C} ac \rightarrow \mathbf{C} bc \end{aligned}$$

In addition, the following theorems can be derived: if a and b overlap, then they are connected (TR16); if a is part of b and a is connected to c then b is connected to c (TR17).

The following relations among regions are defined using the connection relation: a and b are *externally connected* if and only if a and b are connected and a and b do not overlap (D_{EC}); a and b are *disconnected* if and only if a and b are not connected (D_{DC}); region c is self-connected if and only if any two regions that sum up to c are connected (D_{SC}).

$$\begin{aligned} D_{EC} \quad \mathbf{EC} ab & \equiv \mathbf{C} ab \wedge \neg \mathbf{O} ab & D_{SC} \quad \mathbf{SC} c & \equiv (a)(b)(+abc \rightarrow \mathbf{C} ab) \\ D_{DC} \quad \mathbf{DC} ab & \equiv \neg \mathbf{C} ab & & \end{aligned}$$

For example, the region which is occupied by my feet at this moment of time is not self-connected, whereas the region which is occupied by just my left foot is self-connected. One can prove that \mathbf{EC} and \mathbf{DC} are irreflexive and symmetric.

4.5. Characterizing scale differences

To formalize adjacency relations the capability to formally characterize differences in scale is required. For this purpose the theory of qualitative size relations of (Bittner and Donnelly, 2006, 2007a) is included into the ontology.³ This is done by introducing the primitive *roughly the same volume-size* (\approx) and associated definitions and axioms.

³Bittner and Donnelly (2006, 2007a) use techniques from Order of Magnitude Reasoning from Artificial Intelligence (Raiman, 1991; Dague, 1993b,a; Mavrovouniotis and Stephanopoulos, 1988). For alternative approaches to introducing qualitative size relations see also (Gerevini and Renz, 2002) and (Bennett, 2002).

The informal meaning of $r_1 \approx r_2$ for two specific regions r_1 and r_2 as, r_1 and r_2 have roughly the same volume, is at least intuitively relatively clear. What is meant by $a \approx b$ in general for arbitrary regions a and b is *context-dependent* and to a certain degree *vague*. That is, despite a given difference in volume size two large regions (context 1) may be of *roughly the same size*, while the same difference in volume size will make two much smaller regions (context 2) be of clearly distinct size. Moreover, in a single context there will be regions that clearly are of roughly the same size and other regions that are clearly of different size. In addition there will be *boundary cases*, i.e., regions that are neither clearly of roughly the same size nor clearly of different size (Keefe and Smith, 1996). The vagueness and context dependency of *roughly the same volume size* (\approx) generalizes to all notions defined in terms of \approx .

The focus of the formal ontology is on *logical* properties of \approx and relations defined in terms of \approx . These properties hold in all contexts and are not affected by the underlying vagueness and thus can be used for context-independent deductive reasoning. Context dependence and vagueness will be discussed in Section 9.

In terms of \approx the relations *negligible in size* (\ll) and *same scale* (\cong) between regions can be defined: Region a is *negligible in size with respect to* region b if and only if there are regions c_1 and c_2 such that (i) a and c_1 have the same size, (ii) c_1 is a part of b , (iii) c_2 is the difference of c_1 in b and (iii) c_2 has roughly the same volume-size as b (D_{\ll}). Regions a and b are of the *same scale* if and only if neither is negligible in size with respect to the other (D_{\cong}).

$$\begin{aligned} D_{\ll} \quad a \ll b &\equiv (\exists c_1)(\exists c_2)(c_1 \sim a \wedge \mathbf{P} c_1 b \wedge \neg b c_1 c_2 \wedge c_2 \approx b) \\ D_{\cong} \quad a \cong b &\equiv \neg(a \ll b) \wedge \neg(b \ll a) \end{aligned}$$

The following axioms are included: \approx is reflexive (AR15); \approx is symmetric (AR16); if a and b have roughly the same size and b and c have the same size, then a and c have roughly the same size (AR17); if a and b have roughly the same size and the size of a is less than or equal to the size of c and the size of c is less than or equal to the size of b , then c and a , as well as c and b , have roughly the same size (AR18). If a is negligible with respect to b and the size of b is less than or equal to the size of c , then a is negligible with respect to c (AR19).

$$\begin{aligned} AR15 \quad a &\approx a \\ AR16 \quad a &\approx b \rightarrow b \approx a \\ AR17 \quad a &\approx b \wedge b \sim c \rightarrow a \approx c \\ AR18 \quad a &\approx b \wedge a \leq c \wedge c \leq b \rightarrow (c \approx a \wedge c \approx b) \\ AR19 \quad a &\ll b \wedge b \leq c \rightarrow a \ll c \end{aligned}$$

Note that no transitivity axiom for \approx is included in the theory. In many of the intended models of the theory, it is possible to find regions c_1, \dots, c_n such that $a \approx c_1, c_1 \approx c_2, \dots$ and $c_n \approx b$ and but NOT $a \approx b$. Hence, adding a transitivity axiom for \approx would give rise to a version of the Sorites paradox (Hyde, 1996; van Deemter, 1995).

One can prove: if a and b have the same size and b and c have roughly the same size, then a and c have roughly the same size (TR18); if a and b have the same size, then a and b have roughly the same size (TR19); if a is negligible with respect to b , then a is smaller than b (TR20); if the size of a is less than or equal to the size of b and b is negligible with respect to c , then a is negligible with respect to c (TR21); if a is a part of b and b is negligible with respect to c , then a is negligible with respect to c (TR22); if a is negligible with respect to b and b is part of c , then a is negligible with respect to c (TR23); \ll is transitive (TR24).

$$\begin{aligned} TR18 \quad a &\sim b \wedge b \approx c \rightarrow a \approx c & TR22 \quad \mathbf{P} ab \wedge b \ll c &\rightarrow a \ll c \\ TR19 \quad a &\sim b \rightarrow a \approx b & TR23 \quad a \ll b \wedge \mathbf{P} bc &\rightarrow a \ll c \\ TR20 \quad a &\ll b \rightarrow (a \leq b \wedge a \not\sim b) & TR24 \quad a \ll b \wedge b \ll c &\rightarrow a \ll c \\ TR21 \quad a &\leq b \wedge b \ll c \rightarrow a \ll c & & \end{aligned}$$

The same-scale relation \cong is reflexive and symmetric.

It follows that the relations \sim , \approx , \leq , \ll , \mathbf{P} , and $=$ are logically interrelated in the expected ways. For a more detailed discussion see (Bittner and Donnelly, 2006, 2007a).

4.6. Adjacency among regions

Using the mereogeometry with size and scale relations one can now define the adjacency relation between regions as close but non-zero distance: Region a is adjacent to region b if and only if: a and b are not connected and there is a region c such that (i) c is a sphere; (ii) c is negligible in size with respect to a and b , and (iii) c is connected to both a and b (D_{Adj}).

$$\begin{array}{l} D_{\text{Adj}} \quad \text{Adj } xyt \equiv \neg \mathbf{C} ab \wedge (\exists c)(\mathbf{S} c \wedge \mathbf{C} ca \wedge \mathbf{C} cb \wedge c \ll a \wedge c \ll b) \\ TR25 \quad \text{Adj } ab \wedge \mathbf{P} aa_1 \wedge \mathbf{P} bb_1 \wedge \mathbf{DC} a_1a_1 \rightarrow \text{Adj } a_1b_1 \end{array}$$

It follows that Adj is irreflexive and symmetric. One can also prove that if a is adjacent to b and a is part of a_1 and b is part of b_1 and a_1 and b_1 are disconnected then a_1 is adjacent to b_1 ($TR25$).

Note that what exactly counts as ‘close’ distance is context dependent and will in any case depend on the size of the related regions. This will be discussed in more detail in Section 8.

5. Location

To link the mereogeometry of regions to the sub-theory of objects, the ternary location relation L is introduced. On the intended interpretation $L xat$ means: object x is *exactly located* at region a at time t (Casati and Varzi, 1999). In other words, at time t , x takes up the whole region a but does not extend beyond a .

The following axioms for L are included in the theory: object x exists at t if and only if x is located at some region at t ($AL1$); if x is located at a at t , y is located at b at t , and x is part of y at t then a is part of b ($AL2$); it x is part of y at t and both, x and y are located at a at t then y is part of x at t ($AL3$).

$$\begin{array}{ll} AL1 & E xt \leftrightarrow (\exists a)L xat \\ AL2 & L xat \wedge L ybt \wedge P xyt \rightarrow P ab \\ AL3 & P xyt \wedge L xat \wedge L yat \rightarrow P yxt \\ TL1 & L xat \wedge L xbt \rightarrow a = b \end{array}$$

One can prove that if x is located at a at t and x is located at b at t then a and b are identical, i.e., at any time instant an object is located at at most one region ($TL1$). Note that the converse of $AL2$ is not a theorem of this theory, i.e., it is not provable that if x is located at a at t and y is located at b at t and a is a part of b then x is a part of y at t . To see that this cannot be a theorem consider the following example: The region at which the synovial fluid of Joe’s PIP is located at t is part of the region at which the synovial cavity of his PIP is located at t . However the synovial fluid is not part of the synovial cavity, since immaterial objects such as cavities cannot have material entities such as portions of synovial fluid as parts. Note also that, since the theory does not include an axiom of antisymmetry for P , the parthood relation between objects, it does not follow from $AL4$ that two objects that have the same parts at a time and that occupy the same region at that time are identical.

The location relations between objects are defined as follows (Donnelly, 2005): object x is located in object y at t if and only if the spatial region at which x is located at t is a part of the spatial region at which y is located at t (D_{LocIn}); object x is contained in object y at t if and only if x is located in y at t but x and y do not overlap at t (D_{ContIn}).

$$\begin{array}{l} D_{\text{LocIn}} \quad \text{LocIn } xyt \equiv (\exists a)(\exists b)(L xat \wedge L ybt \wedge P ab) \\ D_{\text{ContIn}} \quad \text{ContIn } xyt \equiv \text{LocIn } xyt \wedge \neg O xyt \end{array}$$

At time t , for example, the synovial fluid of Joe's PIP is located in the synovial cavity of his PIP. Since the fluid and the cavity do not overlap, it is also the case that the fluid is contained in the cavity. The proximal parts of the middle phalanx of Joe's left index finger (parts of the of the middle phalanx that are close to the proximal interphalangeal joint) are located in the middle phalanx. Since the proximal parts of the middle phalanx overlap the middle phalanx, the former is not contained in the latter. From Definition D_{LocIn} it immediately follows a weaker form of the converse of $AL2$: if x is located at a at t and x is located at b at t and a is a part of b then x is located in y at t .

At fixed times $LocIn$ is transitive. One can also prove: x exists at t if and only if x is located in itself at t ($TL2$); if x is part of y at t , then x is located in y at t ($TL3$); if x is located in y at t and y is part of z at t , then x is located in z at t ($TL4$); if x is part of y at t and y is located in z at t , then x is located in z at t ($TL5$); if x is part of y at t and y is located in x at t then y is part of x at t ($TL6$).

$$\begin{array}{ll}
TL2 & E\ xt \leftrightarrow LocIn\ xxt \\
TL3 & P\ xyt \rightarrow LocIn\ xyt \\
TL4 & LocIn\ xyt \wedge P\ yzt \rightarrow LocIn\ xzt \\
TL5 & P\ xyt \wedge LocIn\ yzt \rightarrow LocIn\ xzt \\
TL6 & P\ xyt \wedge LocIn\ yxt \rightarrow P\ yxt
\end{array}$$

It does not follow from the axioms that $ContIn$ is asymmetric or transitive. See also (Donnelly *et al.*, 2006; Schulz and Hahn, 2004) for more extended discussions of location and containment relations.

Object x is *permanently* located in object y if and only if whenever x or y exist, x is located in y (D_{pLocIn}). Object x *permanently* contained in object y if and only if whenever x or y exist, x is permanently contained in y ($D_{pContIn}$).

$$\begin{array}{ll}
D_{pLocIn} & pLocIn\ xy \equiv (t)((E\ xt \vee E\ yt) \rightarrow LocIn\ xyt) \\
D_{pContIn} & pContIn\ xy \equiv (t)((E\ xt \vee E\ yt) \rightarrow ContIn\ xyt)
\end{array}$$

One can prove that $pLocIn$ is reflexive and transitive. In addition, one can prove counterparts of theorems ($TL3 - TL6$) for $pLocIn$. One can also prove that x is permanently contained in y if and only if x is permanently located in y and x and y do not overlap at any time ($TL7$).

$$TL7 \quad pContIn\ xy \leftrightarrow (pLocIn\ xy \wedge (t)(\neg O\ xyt))$$

From $TL7$ it immediately follows that if x is permanently contained in y then x is permanently located in y and x and y do not overlap permanently. This theorem corresponds to the left-to-right direction of D_{ContIn} . However there is no theorem for permanent relations corresponding to right-to-left direction of D_{ContIn} in this theory. That is, the formula 'if x is permanently located in y and x and y do not overlap permanently then x is permanently contained in y ' is not a theorem. Again, this shows that there are subtle differences between time-dependent and permanent relations.

6. Connection relations between objects

In this section time-dependent and permanent topological relations between objects are defined in terms of the underlying mereogeometry of regions and the notion of location.

6.1. Time-dependent connection relations between objects.

Object x is connected to object y at time t if and only if x is located at a at t and y is located at b at t and a and b are connected (D_C). That is, object x is connected to object y at t if and only if the distance between the region at which x is located at t and the region at which y is located at t is zero. The definitions of external connectedness and disconnectedness between objects is similar to the definition of connectedness (D_{EC} , D_{DC}). Object x is self-connected at t if and only if x is located at a self-connected region at t (D_{SC}).

$$\begin{array}{ll}
D_C & C\ xyt \equiv (\exists a)(\exists b)(L\ xat \wedge L\ ybt \wedge \mathbf{C}\ ab) \\
D_{EC} & EC\ xyt \equiv (\exists a)(\exists b)(L\ xat \wedge L\ ybt \wedge \mathbf{EC}\ ab) \\
D_{DC} & DC\ xyt \equiv (\exists a)(\exists b)(L\ xat \wedge L\ ybt \wedge \mathbf{DC}\ ab)
\end{array}
\qquad
D_{SC} \quad SC\ xt \equiv (\exists a)(L\ xat \wedge \mathbf{SC}\ a)$$

At time t_1 the following connectedness relations hold among the anatomical parts of the finger depicted in Fig. 3(a): The proximal half of the middle phalanx of Joe's left index finger is connected to the middle phalanx and externally connected to the middle phalanx minus the proximal half⁴; The articular cartilage of the proximal phalanx is externally connected to the proximal phalanx; The articular cartilage of the proximal phalanx is externally connected to the synovial cavity; The synovial fluid and the synovial cavity are connected but not externally connected (see below). At time t the middle phalanx of Joe's left index finger is also self-connected.

C is symmetric at a time and EC and DC are irreflexive and symmetric at a time. One can prove counterparts of theorems (TR14-TR17) for parthood, overlap, and connection relations between objects. As an example, theorem (TC3), the counterpart of (TR15), is stated explicitly. In addition, the following theorems can be derived: x exists at t if and only if x is connected to itself at t (TC1); if x is connected to y at t and y is located in z at t then x is connected to z at t (TC2); if x is externally connected to y at t then x is connected to y at t and x and y do not overlap at t (TC4); at all times disconnected objects are not connected (TC5).

$$\begin{array}{ll}
TC1 & E\ xt \leftrightarrow C\ xxt \\
TC2 & C\ xyt \wedge LocIn\ yzt \rightarrow C\ xzt \\
TC3 & P\ xyt \rightarrow (z)(C\ zxt \rightarrow C\ zyt) \\
TC4 & EC\ xyt \rightarrow (C\ xyt \wedge \neg O\ xyt) \\
TC5 & DC\ xyt \rightarrow \neg C\ xyt
\end{array}$$

Notice that the logical relations between overlap and connection between objects are not exactly analogous to those between regions. In the sub-domain of regions it holds that a and b are externally connected if and only if a and b are connected but do not overlap. In the sub-domain of objects one can prove that if x is externally connected to y at t then x is connected to y at t and x and y do not overlap at t (TC4). But one cannot prove the converse direction, i.e., it is possible that two objects are connected and do not overlap at t but fail to be externally connected. Consider the synovial fluid (SF) and the synovial cavity (SC) of Joe's PIP at some fixed time t at which SF is contained in SC. Then SF is located in SC and SF and SC do not overlap (D_{ContIn}). Thus SF's region (the unique region at which SF is located at t) is a part of SC's region (the unique region at which SC is located at t) (D_{LocIn}). Hence SF's region overlaps SC's region. Consequently, SF's region is connected to SC's region (TR16). Form D_C it follows that SF is connected to SC at t . Thus SF and SC are an example of two objects that are connected and do not overlap. Assume that SF and SC are externally connected at t . It follows that SF's region and SC's region are externally connected (D_{EC}). Thus SF's region and SC's region are connected and do not overlap (D_{EC}). The latter contradicts the fact that SF's region and SC's region do overlap since SF is contained in SC. Hence SF and SC are not externally connected at t . This shows that it is possible that two objects are connected and do not overlap at t but fail to be externally connected at t and the converse of (TC4) cannot be a theorem of the presented axiomatic theory. For more theoretical background on the logical relations between overlap, connectedness, and location see (Donnelly, 2004a). For more bio-medical examples see (Donnelly, 2004b, 2005).

Finally note that the converse of (TC5) cannot be a theorem of the presented theory since from $\neg C\ xyt$ it does not follow that x or y are located at some region at t .

6.2. Permanent connection relations between objects

Permanent connection relation between objects can be defined as follows: x is permanently connected to y if and only if whenever x or y exist, x is connected to y (D_{pC}); x is permanently externally connected

⁴To keep the example simple it is assumed that there is a unique crisp fiat boundary separating the left half and the right half of the middle phalanx.

to y if and only if whenever x or y exist, x is externally connected to y (D_{pEC}); x and y are permanently disconnected if and only if whenever x or y exist, x and y are disconnected (D_{pDC}); x is permanently self-connected if and only if whenever x exists, x is self-connected D_{pSC} .

$$\begin{array}{ll} D_{pC} & pC \, xy \equiv (t)((E \, xt \vee E \, yt) \rightarrow C \, xyt) \\ D_{pEC} & pEC \, xy \equiv (t)((E \, xt \vee E \, yt) \rightarrow EC \, xyt) \\ D_{pDC} & pDC \, xy \equiv (t)((E \, xt \vee E \, yt) \rightarrow DC \, xyt) \end{array} \qquad \begin{array}{ll} D_{pSC} & pSC \, x \equiv (t)(E \, xt \rightarrow SC \, xt) \\ TC6 & pDC \, xy \rightarrow (t)(\neg C \, xyt) \end{array}$$

One can prove that pC is reflexive and symmetric and that pEC is irreflexive and symmetric. One can also prove theorems corresponding to (TR14–TR17) as well as theorems corresponding to (TC2–TC5). One can also prove that if x and y permanently disconnected then they are not connected at all times (TC6).

As pointed out above, connection and external connectedness between anatomical entities in canonical anatomy are of permanent nature. In Figure 2(b) permanent external connectedness relations among parts of Joe’s PIP are represented as solid edges. In canonical anatomy most organs are permanently self-connected.

7. Qualitative size relations between objects

Qualitative size relations are important for comparing anatomical entities and for distinguishing entities at different scales. One can use the qualitative size and scale relations between regions to define qualitative size and scale relations between objects. Qualitative size and scale relations between objects share many properties with their counterparts in the realm of regions. But there are also differences between the different kinds of relations.

7.1. Crisp qualitative size relations

Objects x and y have the same size at time t , $x \sim_t y$, if and only if x is located at a at t and y is located at b at t and a and b have the same size (D_{\sim_t}). Similarly for *less than or equal* in size to (D_{\leq_t}).

$$\begin{array}{ll} D_{\sim_t} & x \sim_t y \equiv (\exists a)(\exists b)(L \, xat \wedge L \, ybt \wedge a \sim b) \\ D_{\leq_t} & x \leq_t y \equiv (\exists a)(\exists b)(L \, xat \wedge L \, ybt \wedge a \leq b) \end{array}$$

To distinguish the binary relation symbols that refer to relations between regions from the corresponding ternary relation symbols that refer to time-dependent relations between objects, the latter are written as time-indexed relation symbols.

Corresponding to axioms (AR6-AR11) and theorems (TR9 - TR12) one can prove that at a given time t the following holds: the same-size relation between objects, \sim_t , is reflexive (when restricted to objects that exist at time t), symmetric and transitive; the size ordering between objects, \leq_t , is reflexive (when restricted to objects that exist at time t) and transitive; the relations \sim_t and \leq_t can be composed in the expected ways.

One can also prove: if x is a part of y at t and x and y are of the same size at t then y is a part of x at t (TO1); if x is a part of y at t and y is a part of x at t then x and y are of the same size at t (TO2). Moreover, if x exists at t and y exists at t then either x is less than or equal in size to y or y is less than or equal in size to x , i.e., at a given time t , \leq_t is a total ordering on objects existing at t (TO3).

$$\begin{array}{ll} TO1 & P \, xyt \wedge x \sim_t y \rightarrow P \, yxt \\ TO2 & P \, xyt \wedge P \, yxt \rightarrow x \sim_t y \\ TO3 & E \, xt \wedge E \, yt \rightarrow x \leq_t y \vee y \leq_t x \end{array}$$

Note that there is no counterpart to theorem (TR7) in the realm of objects. That is, a formula stating that ‘if x is part of y at t and x and y have the same size at t then x and y are identical’ is not a theorem of this theory. To see this consider, again, the synovial membrane of Joe’s PIP and the tissue that constitutes this membrane at time t . As pointed out above, the membrane and the tissue are two distinct objects which, at time t , have exactly the same parts. Therefore one is a part of the other at t and, by theorem (TO2), both are of exactly the same size at t . Hence a counterpart to theorem (TR7) in the realm of objects cannot be a theorem of the theory, since this would force the membrane and the tissue to be identical.

Permanent size relations between objects can be defined as follows:

$$D_{\approx} \quad x \approx y \equiv (t)((E xt \vee E yt) \rightarrow x \sim_t y) \qquad D_{\leq} \quad x \leq y \equiv (t)((E xt \vee E yt) \rightarrow x \leq_t y)$$

Corresponding to axioms (AR6-AR9), axiom (AR11), and theorems (TR9 - TR12) one can prove: the permanent same-size relation between objects, \approx , is reflexive, symmetric and transitive; the permanent size ordering between objects, \leq , is reflexive and transitive; the relations \approx and \leq can be composed in the expected ways.

One can also prove that if x is a permanent part of y and x and y are permanently of the same size then y is a permanent part of x (TO4) and that if x is a permanent part of y and y is a permanent part of x then x and y are permanently of the same size (TO5).

$$TO4 \quad pP xy \wedge x \approx y \rightarrow pP yx \qquad TO5 \quad pP xy \wedge pP yx \rightarrow x \approx y$$

Note, again, there is no counterpart to theorem (TR7) in the realm of permanent relations between objects. In addition it is not the case that for all x and y , either x is permanently less than or equal in size to y or y is permanent less than or equal in size to x . That is, \leq is not a total ordering on the sub-domain of objects.

7.2. Vague qualitative size relations

Objects x and y have *roughly the same size* at time t , $x \approx_t y$, if and only if x is located at a at t and y is located at b at t and a and b have *roughly the same size* (D_{\sim}). The definitions for *negligible in size* and *same scale* are similar.

$$\begin{aligned} D_{\approx} \quad x \approx_t y &\equiv (\exists a)(\exists b)(L xat \wedge L ybt \wedge a \approx b) \\ D_{\ll} \quad x \ll_t y &\equiv (\exists a)(\exists b)(L xat \wedge L ybt \wedge a \ll b) \\ D_{\cong} \quad x \cong_t y &\equiv (\exists a)(\exists b)(L xat \wedge L ybt \wedge a \cong b) \end{aligned}$$

As above, relations between objects are written as time-indexed relations.

Corresponding to axioms (AR15-AR17, AR19) and theorems (TR18 and TR21-TR24) one can prove that at a given time t : the roughly-the-same-size relation \approx_t is reflexive – for objects that exist at time t – and symmetric; the negligible-in-size relation \ll_t is irreflexive, asymmetric, and transitive; the relations \sim_t and \approx_t can be composed in the expected ways; the relations \leq_t and \ll_t can be composed in the expected ways; the relations P and \ll_t can be composed in the expected ways. One can also prove that the same-scale relation \cong_t is reflexive and symmetric.

Permanent size relations between objects can be defined as follows:

$$\begin{aligned} D_{\approx} \quad x \approx y &\equiv (t)((E xt \vee E yt) \rightarrow x \approx_t y) \\ D_{\ll} \quad x \ll y &\equiv (t)((E xt \vee E yt) \rightarrow x \ll_t y) \\ D_{\cong} \quad x \cong y &\equiv (t)((E xt \vee E yt) \rightarrow x \cong_t y) \end{aligned}$$

Corresponding to axioms (AR15-AR17, AR19) and theorems (TR18 and TR21-TR24) one can prove that: the permanent roughly-the-same-size relation \approx is reflexive and symmetric; the permanent negligible-in-size relation \ll is irreflexive, asymmetric, and transitive; the relations \approx and \approx can be composed in the

expected ways; the relations $\overline{\leq}$ and $\overline{\ll}$ can be composed in the expected ways; the relations pP and $\overline{\ll}$ can be composed in the expected ways. One can also prove that the permanent same-scale relation $\overline{\cong}$ is reflexive and symmetric.

8. Adjacency and attachment

Objects x and y are adjacent_to one another when x and y are a *negligible* but non-zero distance apart. Here, what exactly counts as negligible distance may vary from context to context and will in any case depend on the size of x and y . For example, the articular cartilages of the proximal and the medial phalanges of Joe's PIP are adjacent and thus have some minimal positive distance that is negligible with respect to their size. However two Hydrogen molecules that are the same distance apart would not count as adjacent because this distance is not negligible with respect to the size of a Hydrogen molecule. Thus, the adjacency relation can be characterized meaningfully only meaningful if one takes the size of the related entities into account. To formally characterize the adjacency relation, the mereogeometry of regions and its capability to take into account qualitative size and scale relations is utilized.

8.1. Adjacency

Consider the proximal interphalangeal joint (PIP) of Joe Doe's left index finger as depicted in Figure 2. The articular cartilage of the middle phalanx (AC-MP) is adjacent to the articular cartilage of the proximal phalanx (AC-PP) in the sense that although the two objects are separated (not connected), the distance between them is negligible. Formally, one can define adjacency between objects in terms of adjacency of the regions at which they are located: object x is *adjacent_to* object y at time t if and only if there are regions a and b such that x is located at a at t and y is located at b at t and a and b are adjacent (D_{Adj}). Object x is permanently adjacent to object y if and only if at all times at which x or y exist, x is adjacent to y (D_{pAdj}).

$$\begin{aligned} D_{Adj} \quad Adj \, xyt &\equiv (\exists a)(\exists b)(L \, xat \wedge L \, ybt \wedge \mathbf{Adj} \, ab) \\ D_{pAdj} \quad pAdj \, xy &\equiv (t)((E \, xt \vee E \, yt) \rightarrow Adj \, xyt) \end{aligned}$$

It immediately follows that Adj and $pAdj$ are irreflexive and symmetric. Clearly, the middle phalanx of Joe's PIP is not adjacent to itself, but from the fact that the articular cartilage of the middle phalanx (AC-MP) is (permanently) adjacent to the articular cartilage of the proximal phalanx (AC-PP) it does follow that AC-PP is also (permanently) adjacent to AC-MP. Permanent adjacency relations are depicted as dashed and dotted edges in the graph of Figure 2(b).

One can prove that if at a given time two disconnected objects have adjacent parts then the objects themselves are adjacent at that time. That is, if x is adjacent to y at t and x is part of x_1 at t and y is part of y_1 at t and x_1 and y_1 are disconnected at t then x_1 is adjacent to y_1 at t ($TA1$). One can also prove that if two objects that are permanently disconnected have permanently adjacent parts then the objects themselves are permanently adjacent. That is, if x is permanently adjacent to y and x is a permanent part of x_1 and y is a permanent part of y_1 and x_1 and y_1 are permanently disconnected then x_1 is permanently adjacent to y_1 ($TA2$).

$$\begin{aligned} TA1 \quad Adj \, xyt \wedge P \, xx_1t \wedge P \, yy_1t \wedge DC \, x_1y_1t &\rightarrow Adj \, x_1y_1t \\ TA2 \quad pAdj \, xy \wedge pP \, xx_1 \wedge pP \, yy_1 \wedge pDC \, x_1y_1 &\rightarrow pAdj \, x_1y_1 \end{aligned}$$

8.2. Attachment

Consider, again, the proximal interphalangeal joint (PIP) of Joe Doe's left index finger as depicted in Figure 2. One can distinguish the mere permanent adjacency relation between the articular cartilage of the middle phalanx (AC-MP) and the articular cartilage of the proximal phalanx (AC-PP) from the stronger *fixed permanent adjacency* (or attachment) relation between the middle phalanx (MPB) and the synovial membrane of the PIP. AC-MP and AC-PP are permanently adjacent but they can slide along one another, i.e., AC-MP is adjacent to different permanent parts of AC-PP at different times and vice versa (Figures 1(a) and 1(b)).

The relation between the middle phalanx (MPB) and the synovial membrane of the PIP is different from the mere permanent adjacency relation: they are *fixed adjacent* (or attached) in the sense that the middle phalanx (MPB) and the synovial membrane cannot slide with respect to each other. This means that at all times at which the joint as a whole exists, the *same* permanent parts of the middle phalanx and the synovial membrane are in the permanent adjacency relation. The permanent adjacency is due to the fact that the loose connective tissue fibers of the synovial membrane attach to the periphery of the articular cartilage (Mejino, 2007). Similarly, the middle phalanx (MBP) and the ligament of the PIP are in the attachment relation because some of the connective tissue fibers of the ligament intermingle with (or are embedded in) the outer layer of the bone (Mejino, 2007).

These examples show that the fixed nature of attachment relations is often due to what Galton (2000) calls *interlocking*, i.e., "their shapes and relative position are such that under a wide range of conditions they cannot become separated" (Galton, 2000, p. 150). However, attachment is strictly different from the relation of external connectedness in the sense of D_{EC} , since both objects have distinct non-overlapping boundaries, i.e., they have a very small but non-zero distance.

Formally one can define: x is *attached* (or fixed permanently adjacent) to y if and only if x and y are permanently disconnected and there exist permanent proper parts x_1 and y_1 of x and y that are permanently negligible in size with respect to x and y and x_1 and y_1 are permanently adjacent to one another (D_{Att}).⁵

$$D_{Att} \quad Att \, xy \equiv pDC \, xy \wedge (\exists x_1)(\exists y_1)(pPP \, x_1x \wedge pPP \, y_1y \wedge x_1 \lll x \wedge y_1 \lll y \wedge pAdj \, x_1y_1)$$

It follows that Att is irreflexive and symmetric. One can also prove that attachment implies permanent adjacency ($TA3$).

$$TA3 \quad Att \, xy \rightarrow pAdj \, xy$$

Attachment relations are depicted as dashed edges in the graph of Figure 2(b). The permanent adjacency relations are depicted as dotted edges in the graph of Figure 2(b).

9. Context- and precification- dependent interpretation of vague relations

The discussions in Sections 4.5, 7 and 8 have focused on the logical properties of the roughly-the-same-volume-size relation, \approx , and the relations that were defined in terms of \approx such as \lll , \cong , Adj , Att , etc. These logical properties, captured in the axioms, definitions, and theorems of the presented axiomatic theory, are context-independent. That is, they hold in all models of the theory. Moreover the logical properties are not affected by the underlying vagueness that characterizes the relations denoted by \approx , \lll , \cong , Adj , Att , etc.⁶ Thus the logical properties of the relations can be used for context- and vagueness-independent deductive reasoning.

For practical purposes, on the other hand, it is often important to take the context into account and to specify the intended interpretations of formal relations in terms of numerical constraints. These con-

⁵Revised from its original version based on (Donnelly, 2007).

⁶Vagueness is understood as a semantic phenomenon in the sense of (?).

straints, however, may be subject to vagueness and will be specific to certain classes of application domains and certain contexts within those domains. Based on (Bittner and Donnelly, 2006), one specific class of interpretations, which seems to be relevant in many practical application domains, will be discussed in the remainder of this section.

9.1. A class of interpretations

In the considered class of interpretations the relations between the intended meaning of the relations \approx , \ll , \cong , Adj , Att can be specified by means of a class of simple constraints. Let $\|a\|$ and $\|b\|$ be functions in the meta-language of the presented axiomatic theory that yield the exact volume-size of the regions a and b and let ω be a parameter which determines the canonical interpretations of the primitives \approx and the defined relations \ll , \cong , Adj , Att , etc. The constraints that characterize the considered class of interpretations are specified as follows:

- The context parameter ω ranges between 0 and 0.5, i.e., $0 < \omega < 0.5$;
- Regions a and b have roughly the same size if and only if $1/(1 + \omega) \leq \|a\|/\|b\| \leq 1 + \omega$;
- Region a is negligible in size with respect to b if and only if $\|a\|/\|b\|$ is smaller than $\omega/(1 + \omega)$;
- Regions a and b are of the same scale if and only if $\|a\|/\|b\| \geq \omega/(1 + \omega)$, and $\|b\|/\|a\| \geq \omega/(1 + \omega)$.

Since all constraints contain the parameter ω , it follows that what counts as ‘roughly the same size’, ‘negligible’, ‘adjacent’, etc. depends on the context- and precification- dependence of \approx represented by different possible choices for ω . As an example consider Table 2. Let JB be the region at which Joe’s body is located at time t and let JB have a volume of 70 liter. Similarly, let JH be the region at which Joe’s heart is located at time t and let JH have a volume of 0.3 liter. JH is negligible in size with respect to JB for choices of ω larger than 0.0042 and non-negligible otherwise. Regions at which cells are located (average size $400 * 10^{-15}$) are negligible in size with respect to JB for all choices of ω listed in the table.

ω	$JB \approx b$	$b \leq JB \wedge b \cong JB$	$b \ll JB$
0.2	$58.333 \leq \ b\ \leq 84$	$11.666 \leq \ b\ \leq 70$	$\ b\ < 11.666$
0.1	$63.636 \leq \ b\ \leq 77$	$6.363 \leq \ b\ \leq 70$	$\ b\ < 6.363$
0.01	$69.307 \leq \ b\ \leq 70.7$	$0.693 \leq \ b\ \leq 70$	$\ b\ < 0.693$
0.001	$69.93 \leq \ b\ \leq 70.07$	$0.0699 \leq \ b\ \leq 70$	$\ b\ < 0.0699$

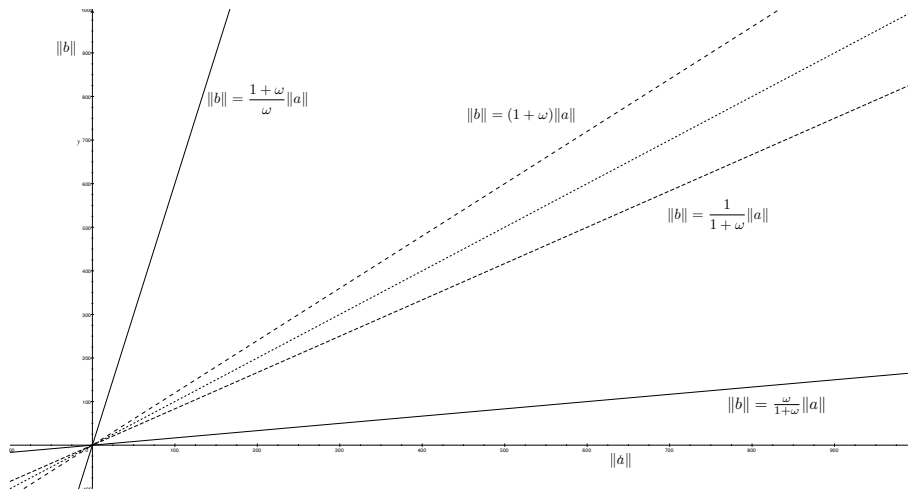
Table 2

The parameter ω determines which regions are roughly the same size and which sizes of regions are negligible with respect to others. Volume in liter: region of Joe’s body at t (JB) = 70 liter, region of Joe’s heart at t (JH) = 0.3 liter, region of an average cell at t (JC) = $400 * 10^{-15}$ liter. (Bittner and Donnelly, 2006)

More generally, consider Figure 5. If a and b have roughly the same size, then $(\|a\|, \|b\|)$ represents a point lying within the area delimited by the dashed lines. If a is negligible with respect to b , then $(\|a\|, \|b\|)$ represents a point lying between the positive vertical axis and the solid diagonal line near the vertical axis. If a and b are of the same scale, then $(\|a\|, \|b\|)$ represents a point lying within the area delimited by the solid lines.

Now consider a fixed entity a and imagine that different values of ω are appropriate for different contexts. The smaller the value of ω , the smaller the value of $|\|a\| - \|b\||$ must be for a to count as close in size to b and the larger $\|b\|$ must be for a to count as negligible in size with respect to b . To picture this situation graphically: the smaller the value of ω , the narrower the corridor between the dashed diagonal lines in Figure 5 and also the narrower the corridor between the solid diagonal lines and the positive vertical and horizontal axes.

Note that the constraints in the discussed class of interpretations are quite simple. In some application domains the relationships between $\|a\|$, $\|b\|$ and ω may be more complex than considered here.

Fig. 5. Graph for $\omega = 0.2$.

9.2. At the boundary between ontology and science

To determine context parameters and numerical constraints that capture the intended interpretations of relations such as \approx , \ll , \cong , *Adj*, *Att*, etc. for a given domain (such as gross-level anatomy) is NOT a matter of ontology or logic. To determine admissible ranges of precification parameters and numerical constraints is a scientific and empirical question. This may be a difficult task for at least two reasons:

1. Due to the qualitative character of canonical anatomy (Neal *et al.*, 1998; Bittner and Goldberg, 2007) and normal variations between individual instantiations of canonical anatomy it may be difficult to determine numerical admissible ranges parameters and constraints in biological/anatomical domains.
2. Relations relations such as \approx , \ll , \cong , *Adj*, *Att*, etc. are vague, thus assigning precise ranges of admissible parameters for fixing the interpretations may be impossible due to higher-order vagueness (Keefe and Smith, 1996).

For those reasons in many actual practical contexts there may not exist precise ranges numerical parameters that can be discovered. In such cases it may be necessary to fix the admissible ranges of numerical parameters by *fiat*.

In fact, fixing the admissible ranges of numerical parameters by fiat is a quite common practice in biology and in the medical sciences. For example the exact normal blood pressure slightly varies from individual to individual. Moreover ‘high blood pressure’ is a vague notion and thus no crisp boundary between normal and high exists. For practical reasons, however, crisp parameters have been introduced by fiat such that a blood pressure of 120/80 is normal and 140/90 or higher counts as high blood pressure. Similar fiat conventions need to be introduced to fix parameters and numerical constraints that capture the admissible interpretations of the relations \approx , \ll , \cong , *Adj*, *Att*, etc. As pointed out above this is not a matter of ontology or logic but a matter of science and of practicing medicine.

10. Computational representation

Current bio-medical ontologies are usually computationally represented in description logic (Baader *et al.*, 2002) languages such as OWL (Horrocks *et al.*, 2003) or the language of the OBO foundry (Smith *et al.*, 2007). Unfortunately, current description logic based languages lack the expressive power that is required to represent complex relationships between relations of the sort discussed in this paper (Bittner and Donnelly, 2005, 2007b). For this reason it is important to understand computational representations of the OBO ontologies as consisting of two complementary components:

1. A description logic (e.g. OWL, OBO format, etc.) based component that enables automatic reasoning and constrains meaning as much as possible;
2. A first- (or higher-) order logic based component that serves as meta-data and makes explicit logical properties of top-level categories and relations. (In particular the logical properties of relations that cannot be expressed in computationally efficient description logics.)

Existing publications on bio-medical ontologies mostly focus on the first component. For this reason the computational representation of the axiomatic theory presented above is used as an example of an ‘implementation’ of the second component.

10.1. Development and representation of the theory

The presented theory is part of the top-level ontology ‘Basic Formal Ontology’ (BFO). BFO is being developed using *Isabelle*, a computational system for implementing logical formalisms (Nipkow *et al.*, 2002). Isabelle is public domain software and can be downloaded for a wide range of operation systems from the Isabelle website <http://isabelle.in.tum.de/>. The computational representation of BFO consists of several hierarchically organized sub-theories. The sub-theories that were discussed in this paper are listed in Table 3.

sub-theory	discussed in	content	OBO-relations
TNEMO	Section 3	non-extensional temporal mereology of objects: temporary and permanent parthood and overlap	part_of, overlap
EMR	Section 4.1	extensional atemporal mereology of regions: parthood and overlap	part_of, overlap
QSizeR, RGB	Section 4.2– 4.5	scale-sensitive mereo-geometry of regions	–
TORL	Section 5	temporary and permanent location and containment	located_in, contained_in
TMTL	Section 6	temporary and permanent connection	–
QSizeO	Section 7	temporary and permanent size relations	–
Adjacency	Section 8	temporary and permanent adjacency; attachment	adjacent_to

Table 3

Sub-theories of the computational representation of BFO and where they are discussion in this paper.

For example, Figure 6 depicts a portion of the computational representation of the sub-theory TORL – the sub-theory which characterizes the location relations as presented in Section 5. The first line states that TORL extends the theory TNEMO – the non-extensional temporal mereology of objects as presented in Section 3, and the theory EMR – the extensional mereology of regions as presented in Section 4. Both sub-theories in turn extend the Isabelle theory FOL (an Isabelle-implementation of a sorted first order predicate logic with identity).

In the section `consts` the predicate symbols and their signatures are introduced. `L :: "Ob => Rg => Ti => o"` specifies that `L` (for location) is a ternary (three-place) predicate symbol whose first parameter is of type `Ob` (objects), the second parameter is of type `Rg` (regions), and the third parameter is of type `Ti` (time-instants). (The fourth parameter `o` is the computational representation of the fact that $L(x, a, t)$ is a predicate that is either true or false.) In the section `axioms` the axioms of the sub-theory are given. The axioms `L_exists` and `L_P_PR` are the axioms AL1 and AL2 of TORL as discussed in Section 4. In the section `defs` the definitions of the sub-theory are given. The definition `LocIn_def` is the definition for the *LocIn* predicate and corresponds to definition D_{LocIn} of TORL. (For details of the syntax see (Nipkow *et al.*, 2002).)

Isabelle, as a computational system for implementing logical formalisms, provides a range of tools that are useful for the development of formal ontologies. Firstly, it includes a number of ‘object logics’ that can be used as foundations for the development of a formal ontology. In the theory presented here the object logic FOL was used. Other object logics include first order and higher order versions of set theory, theories of lists, natural numbers, and others (Nipkow *et al.*, 2002).

```

theory TORL imports TNEMO EMR
begin
consts
  L :: "Ob => Rg => Ti => o"
  LocIn :: "Ob => Ob => Ti => o"
  ...
axioms
  L_exists: "(ALL x t. (E(x,t) <-> (EX a. L(x,a,t))))"
  L_P_PR: "(ALL x y a b t. (L(x,a,t) & L(y,b,t) & P(x,y,t) --> PR(a,b)))"
  ...
defs
  LocIn_def: "LocIn(x,y,t) == (EX a b. (L(x,a,t) & L(y,b,t) & PR(a,b)))"
  ...
theorem LocIn_and_P_imp_LocIn: "[|LocIn(x,y,t);P(y,z,t)|] ==> LocIn(x,z,t)"
  apply(insert P_imp_LocIn)
  apply(insert LocIn_trans)
  apply(auto)
done
...
end

```

Fig. 6. Computational representation of a portion of the sub-theory TORL.

Secondly, Isabelle is also an environment for automatic and interactive theorem proving. All theorems of BFO discussed above were proved in Isabelle and their proofs can be found in the respective theory files. In fact many theorems can be proved in Isabelle with little human assistance. Consider the theorem labeled `LocIn_and_P_imp_LocIn` in Figure 6 which corresponds to Theorem (TL4) in Section 5. A proof in Isabelle is a sequence of application of logical rules using the `apply` command. (For details see (Nipkow *et al.*, 2002).) The proof of theorem `LocIn_and_P_imp_LocIn` (Theorem TL4), for example, is read as follows: (1) use theorem `P_imp_LocIn` (Theorem TL3); (2) use theorem `LocIn_trans` (transitivity of *LocIn*); and (3) search for a proof automatically. Every successful proof ends with the keyword `done`. Theorems can also be proved more or less step by step as demonstrated in the proof of theorem `LocIn_trans` in module TORL. The important point is that if Isabelle ‘compiles’ a theory module then all the proofs are machine-verified, i.e., correct.

Thirdly, Isabelle also provides support for the documentation of axiomatic theories and for publishing theories on the internet. The Isabelle-generated HTML documentation of BFO can be accessed at <http://www.ifomis.org/bfo/fol>.

10.2. Automated reasoning

Isabelle is a system to *design* axiomatic theories and as such has an expressive power that goes well beyond the expressive power of First Order Logic. Once one has designed a theory and verified all theorems, less expressive logics can be used to implement certain portions of the full theory to facilitate automatic reasoning. Less expressive logics like description logics (Baader *et al.*, 2002), for example, have better computational properties and can be used for reasoning about large data sets. For example, all axioms and theorems of BFO that have the form⁷

$$\begin{aligned}
 (a) \quad & R(x,y) \wedge R(y,z) \rightarrow R(x,z) \quad (\text{all transitive relations}) \\
 (b) \quad & R(x,y) \wedge S(y,z) \rightarrow R(x,z) \quad (TL4, TC2, TR11, AR19, TR20, \dots) \\
 (c) \quad & R(x,y) \wedge S(y,z) \rightarrow S(x,z) \quad (TL5, TC3, TR12, TR18, TR19, \dots)
 \end{aligned} \tag{1}$$

can be used as axioms in description logics that include a rule composition operator (Baader *et al.*, 2006). Axioms and theorems of the form (1a) facilitate transitivity reasoning. Axioms and theorems of the forms

⁷*R* and *S* in (1) are meta-variables for relations. Possible temporal parameters are omitted.

(1b) and (1c) facilitate reasoning by *relation composition*. These kinds of reasoning have been identified as being particularly important for bio-ontologies (Spackman, 2000).

Notice, however, that in those less-expressive logics one ‘only’ use these axioms and theorems to support certain forms of automated reasoning. They do not verify the validity of those inferences nor do they specify the semantics of the terms used. Computationally efficient logics often lack the expressive power that is required for those tasks. Thus, as pointed out above, one needs both: (i) computationally efficient logics of restricted expressive power for certain forms of automated reasoning (Levesque and Brachman, 1985; Bittner and Donnelly, 2007b), and (ii) highly expressive logical tools like Isabelle to *design* formal ontologies as axiomatic theories expressed in first order logic.

Isabelle is not the only computational system that can be used for the development of formal ontologies. Other computational frameworks exist or are being developed. In the future a system such as HETS (Mossakowski *et al.*, 2007), that integrates formal systems of different expressive powers, may be needed. This is because such a system may not only facilitate the development of formal ontologies in languages with high expressive power but may also provide automatic means to integrate those ontologies with ontologies represented in computationally efficient languages.

11. Conclusions

To improve the logical and ontological expressiveness and rigor of the OBO relation ontology, a formal specification of the logical properties of the foundational relations listed in Table 1 was provided as an axiomatic theory within the framework of First Order Logic. In addition a computational representation of this theory was created which enabled the verification of the theory and its publication on the internet. This in turn enables others to refine and to extend this theory, or to extract useful axioms and theorems for automatic reasoning in less expressive but computationally tractable logics.

This paper also demonstrated how to expand the relation ontology (a) by incorporating qualitative size and distance relations and (b) by incorporating the distinctions between time-dependent and permanent versions of foundational relations. The importance of both expansions for representing and reasoning about biological structures was demonstrated using the running example of the proximal interphalangeal joint (PIP).

In the formal theory three classes of relations can be distinguished: relations among regions, time-dependent relations among objects, and permanent relations among objects. There are corresponding relations in the different classes. For example, there is a *part_of* relation among regions and there are time-dependent as well as permanent *part_of* relations among objects. Similarly for many other mereogeometrical relations. It was shown that corresponding relations share many properties but that there are also important differences in the sense that the logical relationships between some of the relations among regions are not exactly analogous to those between their counterparts among objects. Moreover, in the realm of objects the logical relationships between some of the permanent relations are not exactly analogous to those between their time-dependent counterparts. A summary of some important differences is given in Table 4. It is important to be aware of those differences, since many bio-medical ontologies that are expressed in less expressive but computationally tractable logics fail to explicitly distinguish between the different kinds of relations.

The attachment relation discussed in Section 8 is different from the other relations in the following senses: (i) Attachment is the only permanent relation and does not have a corresponding time-dependent relation; (ii) Similarly to containment (*ContIn*) and unlike all other permanent relations discussed above, attachment does not have a corresponding relation in the realm of regions; (iii) Attachment (and adjacency) relations are qualitative distance relations and require a mereogeometrical framework.

There are a number of open questions that need to be addressed in future work. One of the most important is the integration of context- and precification- dependent scientific variables (such as the parameter ω in Section 9) into formal ontologies. The presented approach of separating context- and precification-dependent scientific questions from context- and precification-independent ontological questions is ap-

relations between regions	relations between objects	
	time-dependent relations	permanent relations
$PP\ ab \leftrightarrow P\ ab \wedge \neg P\ ba$	$PP\ xyt \leftrightarrow P\ xyt \wedge \neg P\ yxt$	$pPP\ xy \rightarrow pP\ xy \wedge \neg pPyx$
$O\ ab \leftrightarrow (\exists c)(P\ ca \wedge P\ cb)$	$O\ xyt \leftrightarrow (\exists z)(P\ zxt \wedge P\ zyt)$	$(\exists z)(pP\ zx \wedge pP\ zy) \rightarrow pO\ xy$
–	$ContIn\ xyt \leftrightarrow LocIn\ xyt \wedge \neg O\ xyt$	$pContIn\ xy \rightarrow pLocIn\ xy \wedge \neg pO\ xy$
$EC\ ab \leftrightarrow (C\ ab \wedge \neg O\ ab)$	$EC\ xyt \rightarrow (C\ xyt \wedge \neg O\ xyt)$	$pEC\ xy \rightarrow (pC\ xy \wedge \neg pO\ xy)$
$DC\ ab \leftrightarrow \neg C\ ab$	$DC\ xyt \rightarrow \neg C\ xyt$	$pDC\ xy \rightarrow \neg pC\ xy$
$a \leq b \vee b \leq a$	$E\ xt \wedge E\ yt \rightarrow (x \leq_t y \vee y \leq_t x)$	–
$P\ ab \wedge P\ ba \rightarrow a = b$	–	–
$P\ ab \wedge a \sim b \rightarrow a = b$	–	–

Table 4

Comparing logical relationships among different kinds of relations: relations among regions, time-dependent relations among objects, and permanent relations among objects. Only the logically strongest theorem of a given form for a particular kind of relation is displayed. (If there is one.)

appropriate for the *development* and *verification* of formal ontologies. However for the *use* of those ontologies, methods need to be found to integrate context- and precification- dependent scientific variables into formal ontologies.

Acknowledgments

Maureen Donnelly's comments on earlier versions of this paper are gratefully acknowledged. Helpful comments by A. Galton, J. L. V. Mejino, B. Smith, and anonymous reviewers are also gratefully acknowledged. All remaining errors are solely the responsibility of the author.