Representation of Atypical Entities in Ontologies

Christophe Jouis (1, 2), Julien Bourdaillet (2)
(1) Université Paris III - Sorbonne Nouvelle
17, rue de la Sorbonne 75230 - Paris Cedex 05, France

&

(2) LIP6 (Laboratoire d'Informatique de Paris VI - Université Pierre et Marie Curie)
104, avenue du Président Kennedy- 75016 Paris, France
E-mail: cjouis@univ-paris3.fr, julien.bourdaillet@lip6.fr

Abstract

This paper is a contribution to formal ontology study. Some entities belong more or less to a class. In particular, some individual entities are attached to classes whereas they do not check all the properties of the class. To specify whether an individual entity belonging to a class is typical or not, we borrow the topological concepts of interior, border, closure, and exterior. We define a system of relations by adapting these topological operators. A scale of typicality, based on topology, is introduced. It enables to define levels of typicality where individual entities are more or less typical elements of a concept.

1. Introduction

Some entities belong more or less to a class. In particular, some individual entities are attached to classes whereas they do not check all the properties of the class.

To illustrate this phenomenon, let us consider the ontological network above (see Figure 1).
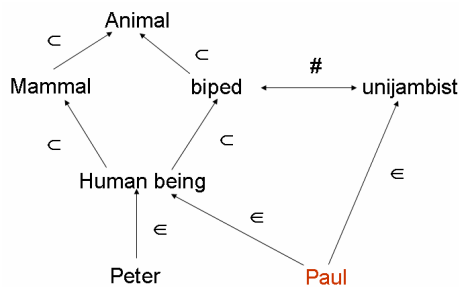

Figure 1: The element Paul does not satisfy all the properties of Human being

This network corresponds to the eight following declarative statements:

(1) Human beings are bipeds;
(2) Bipeds are animals;
(3) Mammals are animals;
(4) Human beings are mammals;
(5) Peter is a human being;
(6) Paul is a unijambist;
(7) Paul is a human being;
(8) One cannot be at the same time biped and unijambist.

Because Paul is a human being, he inherits all the typical properties of a human being, in particular to be a biped. A paradox is introduced by the statement (8) because "Human beings are bipeds" is a general fact but not a universal fact. The statement (1) means "Human beings in general are bipeds but there are exceptions to this law".

The same phenomenon is observed with distributive classes. Some subclasses are attached more or less to a general class because some of theirs elements may not check all the properties of this general class. To illustrate this phenomenon, let us consider the ontological network above (see Figure 2).
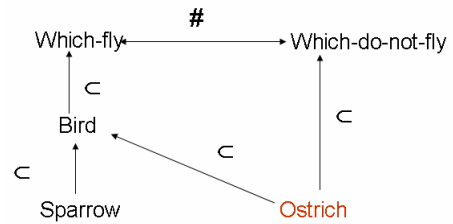

Figure 2: The subclass "Ostrich" does not satisfy all the properties of the class "Bird"

This network corresponds to the five following declarative statements:

(9) Birds fly;
(10) Sparrows are birds;
(11) Ostriches are birds;
(12) Ostriches do not fly;
(13) One cannot be at the same time a flying being and a not flying being.

For identical reasons to the first example, the statement (13) introduces a paradox.

2. Related Works

In Artificial Intelligence, the solution for this kind of problem is default reasoning: an individual A belonging to a concept F inherits concepts subsuming F except contrary indications. This technique of default reasoning led for example Reiter (Reiter 1980) to propose non-monotonic logics.

In terminology, traditionally, the problem of the atypical entities is solved with the idea of Frege (1893): a concept is seen like a function from a field to the set of the values of truth. The concept is used to decide which objects "fall

under the concept" and which do not; i.e. those where the concept applies and those where it does not. We define the extension of a concept F the set of all the individual entities which fall under the concept F, i.e.:

$$\text{Ext (F)} =_{def} \{A \,/F(A) = \text{True}\}$$

For instance, Ext("Human being") = {"Peter", "Paul", …}.

Two models close to our proposition in Section 3 are presented below.

B. Smith introduced mereotopology (Smith 1996). Mereology is the theory of parts and wholes. This theory is reformulated and detailed by Smith in order to serve as a foundation for topology.

From the mereologic primitive x is part of y (noted xPy), three relations are defined: x overlaps y (xOy ⇔ ∃z (zPx ∧ zPy)), x is a discrete from y (xDy ⇔ ¬xOy), and x is a point (Pt(x) ⇔ ∀y (yPx ⇒ y=x)).

A condition φ in a free variable x is satisfied iff the sentence φx is true for at least one value of x. σx(φx) represents the sum of all entities x where φx. This permits to define the set operations x∪y ⇔ σz (zPx ∨ zPy) and x∩y ⇔ σz (zPx ∧ zPy).

The topological primitive x is an interior part of y (noted xIPy) and the relation x crosses y (xXy ⇔ ¬xPy ∧ ¬xDy) enable to define x straddles Y (xSty ⇔ ∀z (xIPZ ⇒ zXy)) and the border as xBy ⇔ ∀z (zPx ⇒ zSty).

The closure is then defined as cl(x) ⇔ x∪σy(yBx). This definition of closure satisfies the three topology axioms (Kuratowski 1958): xPcl(x), cl(cl(x)) = cl(x), and cl(x∪y) = cl(x) ∪ cl(y).

We note also that an interior border is defined as xIBy ⇔ xIPy ∧ xBx and a neighbourhood of a point x is any entity y of which x is an interior part (see Section 5).

J.-P. Desclés (Desclés 1990, Desclés and alii 1991, Freund 2004, Desclés and alii 2005) introduced the concept of LDO (Logic of Determination of Objects) (Pascu 2001). With every concept F the following are canonically associated:

- An object called "typical object", τF which represents the concept F as an object. This object is completely undetermined, for instance τ"Human being" = "a man";
- A function δF defined on objects : the image-object is more determined than the argument-object for this function, for instance δ"Human being" τ "animal" = "a typical human animal";
- The intension of the concept, Int(F) conceived as the class of all concepts that the concept F "includes", that is a semantic network of concepts structured by the relation "IS-A", for instance Int("Human being") = {"Mammal", "Biped", "Animal", …};
- The expanse of the concept, Exp(F) which contains all "more or less determined objects" such that the concept F applies to;
- A part of the expanse is the extension of the concept, Ext(F) which contains all completely determined objects such that the concept F applies to.

LDO captures two kinds of objects: typical objects and atypical objects. Typical objects in Exp(F) inherit all concepts of Int(F); atypical objects in Exp(F) inherit only some concepts of Int(F).

3. Modeling using topology

3.1 Topology Basics

Let E be any set and let T be a family of subsets of E. Then T is a topology on E if

1. Both the empty set and E are elements of T.
2. Any union of elements of T is an element of T.
3. Any intersection of finitely many elements of T is an element of T.

If T is a topology on E, then E together with T is called a *topological space*. All sets in T are called *open*; note that not all subsets of E are in T. A subset of E is said to be *closed* if its complement is in T (i.e., it is open). A subset of E may be open, closed, both, or neither.

- A set U is called open if, intuitively speaking, starting from any point x in U one can move by a small amount in any direction and still be in the set U. In other words, the distance between any point x in U and the edge of U is always greater than zero. The empty set is open. The union of any number of open sets is open. The intersection of a finite set of open sets is open.
- A *closed* set is a set whose complement is open.
- The *interior* of a set S –in(S)–consists of all points which are intuitively "not on the edge of S". A point which is in the interior of S is an interior point of S. The notion of interior is in many ways dual to the notion of closure.
  The *closure* of a set S –cl(S)– consists of all points which are intuitively "close to S". A point which is in the closure of S is a point of closure of S. The notion of closure is in many ways dual to the notion of interior: cl(S) = co(in(co(S))) = co(ex(S)), where co (S) represents complementary to S.
- The *exterior* of a set –ex(S)– is the interior of its complement co(S): ex(S) = in(co(S)).
- The *boundary* (or frontier or border) of a set –bo(S)–is the set's closure minus its interior: bo(S) = cl(S) - in(S). Equivalently, the boundary of a set is the intersection of its closure with the closure of its complement: bo(S) = cl(S) ∩ cl(co(S)).
- A *neighborhood* of a point x is a set containing an open set which in turn contains the point x. More generally, a neighborhood of a set S is a set containing an open set which in turn contains the set S. A neighborhood of a point x is thus a neighborhood of the singleton set {x}. (Note that under this definition, the neighborhood itself need not be open. Many authors require that neighborhoods be open).

3.2. Our Proposal

To specify whether an individual entity belonging to a class is typical or not, we borrow the topological concepts of interior, border, closure, and exterior. In topology these concepts are derived in the following operators: in, bo, cl and ex, respectively. There exist interesting properties of combination of the operators which enable the definition of an algebra (Kuratowski 1958, p. 24). In particular:

$\forall$ A a set,
$\quad$ in(A) $\subset$ A $\subset$ cl(A)
$\quad$ bo(A) $\subset$ cl (A) ; because cl(A) $=_{def}$ A $\cup$ bo(A)
$\quad$ A $\cap$ ex(A) = $\varnothing$
$\quad$ bo(A) = co(in(A)) $\cap$ co (ex(A)),
$\quad$ in (in(A)) = in (A)
$\quad$ cl (cl (A)) = cl (A), etc.

We define a system of relations by adapting these topological operators into the following algebraic relations: $\in$-in, $\in$-bo, $\in$-ex, $\subset$-in, $\subset$-bo and $\subset$-ex. We apply these relations to the extension of a class. For instance :
$\quad$ (X $\in$-in F) $\Leftrightarrow$ (X $\in$ in (Ext(F)).
Individual entities belonging to the extension of a class are more or less typical depending whether they are inside or at the border of the extension. We define the relations of inclusion and membership by distinguishing the interior, the border and the exterior of a class. The topological properties enable us to define rules of combination of the six relations $\in$-in, $\in$-bo, $\in$-ex, $\subset$-in, $\subset$-bo and $\subset$-ex in the next section[1]

3.3. Membership at the interior of a class

We define (X $\in$-in F) if and only if X inherits all the properties of F. Like the membership, the relation $\in$-in is irreflexive, asymmetric and intransitive. The relation $\in$-in is a specification of the relation $\in$, therefore:
$\quad$ Rule R1: (X $\in$-in F) $\Rightarrow$ (X $\in$ F)[2]
The heritage of property will be valid for the prototypic occurrences of the class, i.e.:
$\quad$ Rule R2: (X $\in$-in F) $\wedge$ (F $\subset$-in G) $\Rightarrow$ (X $\in$-in G)
$\quad$ R2bis: (X $\in$-in F) $\wedge$ (F $\subset$-bo G) $\Rightarrow$ (X $\in$-bo G)

3.4. Membership at the border of a class

We define (X $\in$-bo F) if and only if X is an atypical individual entity of F. Like the membership, the relation $\in$-bo is irreflexive, asymmetric and intransitive. The relation $\in$-bo is a specification of the relation $\in$, therefore:
$\quad$ Rule R3: (X $\in$-bo F) $\Rightarrow$ (X $\in$ F)[3]
The heritage of property is valid for atypical occurrences of a class, only if these ocurrencies belong to the border:
$\quad$ Rule R4: (X $\in$-bo F) $\wedge$ (F $\subset$-in G) $\Rightarrow$ (X $\in$-bo G)
$\quad$ Rule R5: (X $\in$-bo F) $\wedge$ (F $\subset$-bo G) $\Rightarrow$ (X $\in$-bo G)

3.5. Inclusion at the interior of a class

We define (F $\subset$-in G) if and only if F is a typical subclass of G. Like inclusion, $\subset$-in is irreflexive, asymmetric and transitive. The relation $\subset$-in is a specification of the relation $\subset$, therefore:
$\quad$ Rule R6: (X $\subset$-in F) $\Rightarrow$ (X $\subset$ F)[2]
The transitivity of inclusion results in rules of composition of the relations $\subset$-in and $\subset$-bo :
$\quad$ Rule R7: (F $\subset$-in G) $\wedge$ (G $\subset$-bo H) $\Rightarrow$ (F $\subset$-bo H)
$\quad$ Rule R8: (F $\subset$-bo G) $\wedge$ (G $\subset$-in H) $\Rightarrow$ (F $\subset$-bo H)

3.6. Inclusion at the border of a class

We define (F $\subset$-bo G) if and only if F is an atypical subclass of the class G. Like inclusion, $\subset$-bo is asymmetric and transitive. However, $\subset$-bo is irreflexive. The relation $\subset$-bo is a specification of the relation $\subset$, therefore:
$\quad$ Rule R9 : (X $\subset$-bo F) $\Rightarrow$ (X $\subset$ F)[3]
The transitivity of inclusion results in rules clarified in item 2.3., i.e.:
$\quad$ Rule R7: (F $\subset$-in G) $\wedge$ (G $\subset$-bo H) $\Rightarrow$ (F $\subset$-bo H)
$\quad$ Rule R8: (F $\subset$-bo G) $\wedge$ (G $\subset$-in H) $\Rightarrow$ (F $\subset$-bo H)

3.7. External membership and inclusion

We define (X $\in$-ex F) if and only if X cannot belong neither the interior nor the border of F (and in the same way recursively for the subclasses of F). Thanks to inheritance, the relation $\in$-ex is applied to more general classes:
$\quad$ Rule 10: (X $\in$-ex F) $\wedge$ (F $\subset$ G) $\Rightarrow$ (X $\in$-ex G)
The relation $\in$-ex is irreflexive, asymmetric and intransitive.
The relation $\subset$-ex corresponds to the relation of disjunction between classes, i.e. $\subset$-ex is irreflexive, symmetric and intransitive. The relation $\subset$-ex is propagated in the more specific classes:
$\quad$ Rule R11: (F $\subset$-ex G) $\wedge$ (H $\subset$ G) $\Rightarrow$ (H $\subset$-ex F).
$\quad$ Rule R12: (F $\subset$ G) $\wedge$ (H $\subset$-ex G) $\Rightarrow$ (H $\subset$-ex F).
$\quad$ Rule R13: (X $\in$-in F) $\wedge$ (H $\subset$-ex G) $\Rightarrow$ (X $\in$-ex G).

3.8. Topological interpretation of the two examples

Figure 3 represents an interpretation of Figure 1 using our topological relations. In particular, we notice that Paul is an atypical element of the class "Human being".
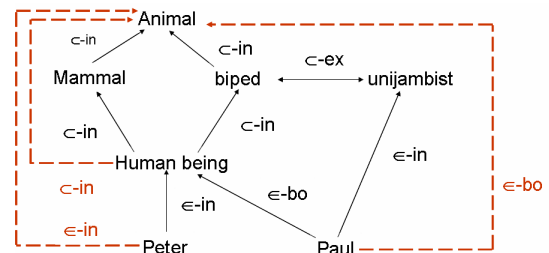

Figure 3 : Interpretation of Figure 1

---

[1] We consider that (X$\in$-cl F) $\Rightarrow$ (X $\in$ F) and (F $\subset$-cl G) $\Rightarrow$ (F $\subset$ G), because F $\subset$ cl(F) and (X$\in$-cl F) $\Leftrightarrow$ (X $\in$ cl(Ext(F)) and (F $\subset$ G) $\Leftrightarrow$ (F $\subset$ Ext(G)).

[2] Because $\forall$ A a set, in(A) $\subset$ A.

[3] Because $\forall$ A a set, bo(A) $\subset$ cl (A).

Dotted arrows represent some possible deductions thanks to the rules of combination we defined in the previous section.

For example:

>(14) *"Human being" is a typical class of "Animal" (statements 3 and 4 and ⊂-in is transitive);*
>
>(15) *"Peter" is a typical element of the "Animal" class (statement 14 and rule R2);*
>
>(16) *"Paul" is an atypical element of the "Animal" class (statement 14 and rule R4).*

Figure 4 represents an interpretation of Figure 2 using our topological relations. In particular, we notice that the class "Ostrich" is an atypical subclass of the class "Bird". Dotted arrows represent some possible deductions thanks to the rules of combination we defined in the previous section.

For example:

>(17) *The class "Sparrow" is a typical subclass of the class "Which fly" (statements 9 and 10 and ⊂-in is transitive);*
>
>(18) *The class "Ostrich" is an atypical subclass of the class "Which fly" (statements 9 and 11 and rule R8).*
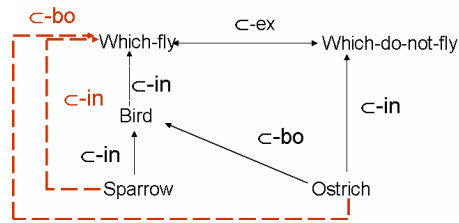

Figure 4: Interpretation of Figure 2

4. Combination table of the six relations

With respects to the rules defined in section 3, we define the combination table of all the possibilities. An element Rz of the table represents the combination of an element Rx in lines and an element Ry in columns.

(A Rx B) ∧ (B Ry C) ⟹ (A Rz C)	∈-bo	∈-in	∈-ex	⊂-in	⊂-bo	⊂-ex
∈-bo	NIL (1)	NIL (1)	NIL (1)	∈-bo (2)	∈-bo (7)	NIL (14)
∈-in	NIL (1)	NIL (1)	NIL (1)	∈-in (3)	∈-bo (8)	∈-ex (15)
∈-ex	NIL (1)	NIL (1)	NIL (1)	∈-ex (4)	∈-ex (4)	NIL (14)
⊂-in	NIL (5)	NIL (5)	NIL (5)	⊂–in (6)	⊂-bo (9)	⊂-ex (13)
⊂-bo	NIL (5)	NIL (5)	NIL (5)	⊂-bo (10)	⊂-bo (11)	⊂-ex (13)
⊂-ex	NIL (5)	NIL (5)	NIL (5)	⊂-ex (12)	⊂-ex (12)	NIL (14)

Comments:
(1) ∈ is intransitive ; (2) rule R4; (3) rule R2; (4); rule R10; (5) inheritance of properties is not in this direction; (6) ⊂-in is transitive; (7) rule R5; (8) rule R2bis; (9) rule R7; (10) rule R8 ; (11) ⊂-bo is transitive ; (12) rule R11 ; (13) rule R12; (14) we can not conclude; (15) rule R13.

5. Scale of typicality

In this section, a scale of typicality, based on topology, is introduced. It enables to define degrees of typicality where individual elements belonging to a class are more or less typical. The most typical elements are the elements where there exists no doubt on their class membership. Atypical elements miss some properties of class' typical elements which involve a reduction of their typicality degree.

For example, in the "Bird" class a "Sparrow" is more typical than a "Crow" because for the common sense (at least in France) birds are small. A "Hen" which flies hardly is less typical than a "Crow", but more than an "Ostrich" which does not. These differences involve a scale of typicality. To model this scale of typicality, the thickness of the border of a class is introduced (see Figure 5).
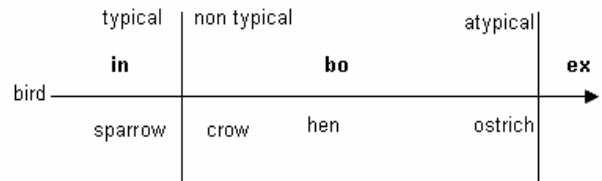

Figure 5: Thickness of the border of the "Birds" class permits the introduction of typicality levels of "Crow", "Hen" and "Ostrich".

The more elements with a low typicality degree are allowed to belong to a class, the more its border is thick. In our example, typicality degrees decrease with the loss of the flying property or in function of the common sense. Being given e, an integer constant arbitrarily set, which model the thickness of the border, it is then possible to define the interior in(F, e), the exterior ex(F, e), and the border bo(F, e) of a class F in function of e in the following way.

5.1. Interior of F in function of e, in(F, e)

if e = 0, in(F, e) = in(F) ;
if e > 0, in(F, e) ⊂ in(F) ;
x ∈ in(F, e) ⟺ ∃ n ∈ N(x) : n ⊂ in(F, e-1), where N(X) represents the set of the neighbourhoods of X.

5.2. Exterior of F in function of e, ex(F, e)

if e = 0, ex(F, e) = ex(F) ;
if e > 0, ex(F) ⊂ ex(F, e);
x ∈ ex(F, e) ⟺ ∃ n ∈ N(x) : n ∩ F = ∅ ∧ n ⊂ ex(F, e-1)

5.3. Border of F in function of e, bo(F, e)

if e = 0, bo(F,e) = bo(F) ;
if e > 0, bo(F) ⊂ bo(F,e) ;
x ∈ bo(F, e) ⟺ ∃ n ∈ N(x), n ∩ in(F, e) = ∅ ∧ n ∩ ex(F, e) = ∅.

These definitions remain compatible with the classical operators of topology by defining:

$$ex(F) = \cap_{i=0..e} ex(F, i)$$
$$in(F) = \cap_{i=0..e} in(F, i)$$
$$bo(F) = \cap_{i=0..e} bo(F, i)$$

Because the border has a thickness, it is then possible to define the interior border and the exterior border in the following way:

$x \in bo_{in}(F, e) \Leftrightarrow$
$\exists\, n_e \in N(x),\, \forall\, n \in N(F) : n_e \subset n \wedge n_e \cap in\,(F, e) \neq \varnothing.$

$x \in bo_{ex}(F, e) \Leftrightarrow$
$\exists\, n_e \in N(x),\, \forall\, n \in N(F) : n_e \subset n \wedge n_e \cap ex\,(F, e) \neq \varnothing.$

The interior border represents non-typical elements, i.e. elements with a lower typicality degree than typical elements of a class. The exterior border represents atypical elements which do not inherit all the properties of a class (see Figure 6).
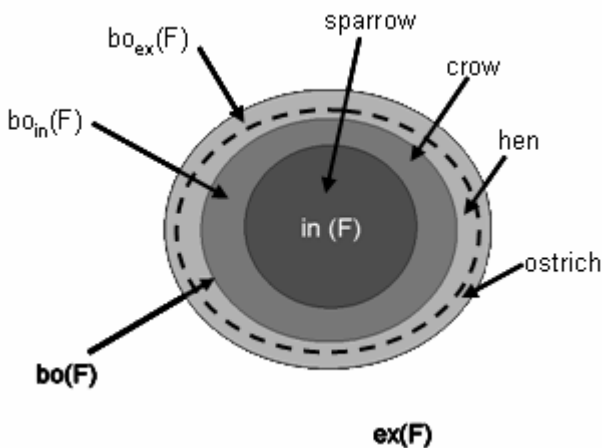

Figure 6: A concept in which the border has a thickness

6. Conclusion

In this paper, the topological concepts of interior, border, closure and exterior are used to specify whether an individual entity belonging to a class is typical or not. By adapting these operators, a system of relations is defined. A scale of typicality is introduced. It enables to define levels of typicality where individual entities are more or less typical element of a class.

This model can be used by ontology builders during the modelisation process or maintenance. When a certain size is achieved by an ontology and an atypical entity is discovered, the cost of ontology redesign may be too expensive. This model facilitates the ontology maintenance by avoiding redesign.

7. References

Frege, G. (1893). *Logical Investigations*, ed. Peter Geach, Blackwells, 1975.

Desclés J.-P., Pascu A. (2005). Logic of Determination of Objects: the meaning of variable in quantification, *Flairs 2005*, edited by AAAI Press, Florida, pp. 610-616.

Desclés, J.-P. and Kanellos, I. (1991). La notion de typicalité: une approche formelle. In *Sémantique et Cognition, sous la direction de D. Dubois*, pp. 225--244, Paris. Editions CNRS.

Desclés, J.-P. (1990). *Langages Applicatifs, Langues Naturelles et Cognition*. Paris: Hermès.

Freund M., Desclés J.-P., Pascu A., Cardot J. (2004), "Typicality, contextual inferences and object determination logic", FLAIRS 04, Miami, pp. 491--495.

Grosso, W. (1999). Formal Aspects Of Protégé, *Fourth International Protégé Workshop*, July, 1999, on line http://protege.stanford.edu/doc/presentations.html.

Jouis, C., (2002), "Logic of Relationships", Chapitre II.8. of the Volume "*The Semantics of Relationships*", Rebecca Green, Carol A. Bean & Sung Hyon Myaeng Eds, Kluwer Academic Publishers, Dordrecht.

Jouis, C. (2007). Un système logique pour les relations sémantiques entre concepts. In *Proceedings of Toth Terminologie & Ontologie : Théories et Applications*, Annecy. 1er juin 2007, pp.15--37.

Kuratowski, C. (1958). *Topologie*. 2 vol. Panstwowe Wydawnie two Naukowe, Varsovie.

Pascu, A. (2001). *Logique de Détermination d'Objets : concept de base et mathématisation en vue d'une modélisation objet*. PhD diss., Département de Logique, Informatique et Cognition (LALICC), Université de Paris-Sorbonne.

Pickover Clifford A. (2006). *The Möbius Strip: Dr. August Möbius's Marvelous Band in Mathematics, Games, Literature, Art, Technology, and Cosmology*. Thunder's Mouth Press.

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13, pp. 81--132.

Smith B. (1996). Mereotopology: A Theory of Parts and Boundaries. *Data and Knowledge Engineering* 20, pp. 287--303.