



HAL
open science

Towards a qualitative representation for specifying natural language

Richard Dapoigny, Patrick Barlatier

► **To cite this version:**

Richard Dapoigny, Patrick Barlatier. Towards a qualitative representation for specifying natural language. 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Jun 2017, Annecy, France. 10.1109/CIVEMSA.2017.7995299 . hal-01850086

HAL Id: hal-01850086

<https://hal.science/hal-01850086>

Submitted on 26 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Towards a Qualitative Representation for Specifying Natural Language

Richard Dapoigny
LISTIC/Polytech' Annecy-Chambéry,
Domaine Universitaire
Po. Box 80439,
74944 Annecy le vieux cedex, France
Email: richard.dapoigny@univ-smb.fr

Patrick Barlatier
LISTIC/Polytech' Annecy-Chambéry,
Domaine Universitaire
Po. Box 80439,
74944 Annecy le vieux cedex, France
Email: patrick.barlatier@univ-smb.fr

Abstract—The imprecision of natural language such as X is before Y does not provide any information about the distance between X and Y . The human's way of reasoning often uses incomplete knowledge, and any formalization of natural language should be able to produce qualitative reasoning. For that purpose, we introduce a modeling for specifying cognitive geometry. More precisely, in this paper we tackle the problem of geometric reasoning using a formal representation that is built from an appropriate spatial theory which departs from classical geometry in many points. Instead of basic spatial entities of ordinary geometry such as points, lines and surfaces, we rather consider physical three-dimensional objects used to perceive space based on pointless geometry. Based on Tarski's mereogeometry of solids, we explain how the interpretation of mereogeometry within Lesniewski's framework of mereology leads to a coherent and simplified framework in which qualitative representations (e.g., solids, balls, concentricity, equidistance, etc.) can be expressed.

I. INTRODUCTION

The qualitative representation of information assumes that knowledge about spatial information is expressed with a finite set of qualitative relations. It has a wide number of applications in the fields of computer vision [16], geographic information system [25], autonomous robot navigation [23], natural language understanding [3] and the like. When quantitative descriptions are not available or computationally intractable, qualitative representations appear as fruitful substitutes. For example, a few set of primitives forms the basis of a rich spatial calculus (see e.g., [17]). Assessing situations require to specify relationships between objects of the domain. Then, it seems natural to represent qualitative information with relations.

Many approaches in qualitative representation are based on mereological theories. Mereology which is the theory of parts and wholes assumes that the part-of relation is the fundamental primitive [19]. However, mereology by itself is not a sufficient theory to take in account topological relations. For that purpose, mereotopology has been introduced as a common theory for integrating both topology and mereology [8], [24]. As underlined in [10] there are many situations in which mereotopology itself is lacking e.g., direction and orientation relations, distances and sizes or even, shapes. To cope with such deficiencies, mereogeometry which builds on the earlier work of Tarski [21], extends mereology with geometric

notions. Using spheres (or balls) as a primitive notion, and a small set of definitions, it captures full Euclidian geometry without quantitative relations. The spatial binary relations can be considered as constraints which restrict the spatial properties of objects, leading to efficient reasoning algorithms. In this paper, we explain how a coherent framework based on mereogeometry could be an expressive tool for representing qualitative spatial assertions in natural language.

II. BASIS OF MEREOGEOMETRY

We first recall the set of definitions suggested by Tarski to support the mereogeometry of solids itself referring to the theory Leśniewski known as (extensional) mereology. In the second subsection, we discuss some issues about Tarski's work while the last subsection presents the type-theoretical account of Tarski mereogeometry.

A. Leśniewski's Mereology

In the early 20th century, S. Leśniewski first proposed a higher-order logical theory [14] including (i) a logical part based on the basic category of propositions (S) and a single axiom about equivalence (\equiv) called protothetic and (ii) a calculus of name (a.k.a. Leśniewski's Ontology) which is centered on the category of name N together with a single axiom denoted ε , which belongs to the category $N \rightarrow N \rightarrow S$. Every name denotes either one object of the world (singular name), many objects (plural name) or no object at all (empty name). A single axiom is introduced for governing its behavior. Notice that there are no free variables in Leśniewski's systems and that there is only a general quantifier (for all) denoted \forall . The particular quantifier (for some) is expressed from the general quantifier as $\neg\forall x\neg(\phi(x))$, for any formula ϕ which depends on a single variable x . Following the work of Henkin [13], we assume that mereology can be fruitfully expressed in type theory. we describe all relations in the usual logic.

Axiom isEpsilon : $\forall A a, A \varepsilon a \equiv ((\neg\forall B, \neg(B \varepsilon A)) \wedge (\forall C D, ((C \varepsilon A) \wedge (D \varepsilon A)) \rightarrow (C \varepsilon D)) \wedge (\forall C, (C \varepsilon A) \rightarrow (C \varepsilon a)))$.

The first conjunction of the right side of the equivalence prevents A from being an empty name, the second conjunction states the uniqueness of A while the last conjunction refers to a kind of convergence (anything which is A is also an a). Only two constant name definitions hold in Leśniewski's Ontology,

the universal name (universal) and the empty name (empty), each of them having the type N .

This logical system was the support of a theory called mereology, whose purpose was the description of the world with collective classes and mereological relations such as the so-called part-of relation. While set theory relies on the opposition between element and set, mereology is rather based on the opposition between part and whole. Mereology primarily assumes the distinction between the distributive and collective interpretations of a class. Whereas the ontology introduces names, where name is the distributive notion, the mereology rather relies on classes, where class is the collective notion. Notice that the introduction of plural names gives mereology an expressive power that goes beyond the capabilities of first-order logic (e.g., predicates may be introduced which have plural subjects). The most usual formalization introduces a mereological element called part (pt), as a primitive. Mereology is developed on a minimal collection of axioms and new primitive functors are defined whose the most important are that of (collective) class, denoted Kl (a.k.a sum) and el (for element of). In this axiomatization pt is taken as primitive, while el and Kl are introduced by definitions. Two axioms govern the behavior of the pt function:

Axiom *asymmetric_Part* : $\forall A B, A \varepsilon pt B \rightarrow B \varepsilon (distinct (pt A))$.
Axiom *transitive_Part* : $\forall A B C, A \varepsilon pt B \wedge B \varepsilon pt C \rightarrow A \varepsilon pt C$.

where *distinct* refers to a name corresponding to all objects that are distinct from ptA . The name forming function el i.e., being an element of, is specified as a definition with respects to:

Variable *MD1* : $\forall A B, A \varepsilon el B \equiv (A \varepsilon A \wedge (singular_equality A B \vee A \varepsilon pt B))$.

A similar definition formally introduces the Kl function.

Variable *MD2* : $\forall A a, A \varepsilon Kl a \equiv (A \varepsilon A \wedge (\neg \forall B, \neg(B \varepsilon a)) \wedge (\forall B, (B \varepsilon a \rightarrow B \varepsilon el A)) \wedge (\forall B, B \varepsilon el A \rightarrow \neg \forall C D, \neg((C \varepsilon a) \wedge (D \varepsilon el C) \wedge (D \varepsilon el B))))$.

In other words, A is the class of objects a if and only if the following conditions are met: (i) A is an object, (ii) every a is an element of the object A and (iii) for all B , if B is an element of the object A , then some element of the object B is an element of some a . Furthermore, an axiom states the uniqueness of the collective class. The semantics

$A \varepsilon A$	A is an object
$A \varepsilon B$	A is B
$A \varepsilon (pt B)$	A is part of the object B
$A \varepsilon \neg(pt B)$	A is not part of the object B
$A \varepsilon (el B)$	A is an element of the object B
<i>singular_equality</i> ($A B$)	A is the same object as B
$A \varepsilon (Kl a)$	A is the class of object a

TABLE I
SEMANTICS FOR BASIC EXPRESSIONS IN MEREOLGY.

underlying some basic expressions is given in table I. Many other useful constructs have been added to mereology such as collection (*coll*), sub-collection (*subcoll*), external (*ext*) or overlap (*ov*).¹

Variable *MD3* : $\forall P a, P \varepsilon coll a \equiv (P \varepsilon P \wedge \forall Q, Q \varepsilon el P \rightarrow \neg \forall C D, \neg(C \varepsilon a \wedge C \varepsilon el P \wedge D \varepsilon el C \wedge D \varepsilon el Q))$.
Variable *MD4* : $\forall P Q, P \varepsilon subcoll Q \equiv (P \varepsilon P \wedge \forall C,$

$C \varepsilon el P \rightarrow C \varepsilon el Q)$.

Variable *MD5* : $\forall P Q, P \varepsilon ext Q \equiv (P \varepsilon P \wedge \forall C, \neg(C \varepsilon el P \wedge C \varepsilon el Q))$.

Variable *MD7* : $\forall P Q, P \varepsilon ov Q \equiv (P \varepsilon P \wedge \neg \forall C, \neg(C \varepsilon el P \wedge C \varepsilon el Q))$.

B. Tarski's Definitions

Tarski starts out from some mereological notions such as being a part of and sum and then, provides a minimal set of axioms constituting the mereological part on which the rest of the theory relies [21]. This part will be substituted in the present work by the mereology developed in subsection II-A. Apart from this mereological part, Tarski takes *sphere* (which will be called *ball* here) as the only primitive notion that is specific to the geometry of solids. These notions are sufficient to formulate a set of nine geometrical definitions which constitutes the basis of mereogeometry.

Definition 2.1: The ball A is externally tangent to the ball B if (i) the ball A is disjoint from the ball B and (ii) given two balls X and Y containing as part the ball A and disjoint from the ball B , at least one of them is part of the other.

Definition 2.2: The ball A is internally tangent to the ball B if (i) the ball A is a proper part of the ball B and (ii) given two balls X and Y containing the ball A as a part and forming part of the ball B , at least one of them is a part of the other.

Definition 2.3: The balls A and B are externally diametrically tangent to the ball C if (i) each of the balls A and B is externally tangent to the ball C and (ii) given two balls X and Y disjoint from the ball C and such that A is part of X and B a part of Y , X is disjoint from Y .

Definition 2.4: The balls A and B are internally diametrically tangent to the ball C if (i) each of the balls A and B is internally tangent to the ball C and (ii) given two balls X and Y disjoint from the ball C and such that the ball A is externally tangent to X and B to Y , X is disjoint from Y .

Definition 2.5: The ball A is concentric with the ball B if one of the following conditions is satisfied: (i) the balls A and B are identical, (ii) the ball A is a proper part of B and besides, given two balls X and Y externally diametrically tangent to A and internally tangent to B , these balls are internally diametrically tangent to B and (iii) the ball B is a proper part of A and besides, given two balls X and Y externally diametrically tangent to B and internally tangent to A , these balls are internally diametrically tangent to A .

Definition 2.6: A point is the class of all balls which are concentric with a given ball.

Definition 2.7: The points A and B are equidistant from the point C if there exists a ball X which belongs as element to the point C and which satisfies the following condition: no ball Y belonging as element to the point A or to the point B is a part of X or is disjoint from X .

Definition 2.8: A solid is an arbitrary sum of balls.

Definition 2.9: The point P is an interior point of the solid B if there exists a ball A which is at the same time an element of the point P and a part of the solid B .

The 2D version of these definitions is illustrated in figure 1.

¹Due to a lack of space, we restrict the description to these three notions.

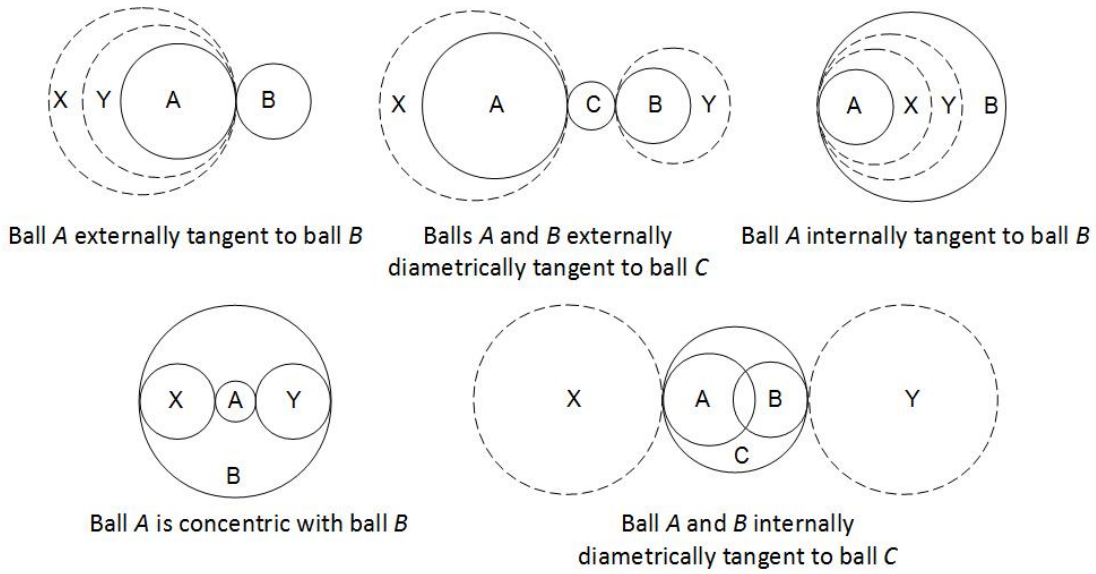


Fig. 1. Tarski's primitive definitions.

C. Some issues in Tarski's Work

While in the sketchy paper of Tarski [21] the logical background seems rather unclear [12], [5] we depart from all approaches stemming from a set-based framework and rather argue that an axiomatization of Tarski's work based solely on Leśniewski's Mereology is more coherent. Instead of introducing axioms for part-whole relations, we rather rely on the Leśniewski's mereology together with its set of axioms and theorems (in fact Tarski applies Leśniewski's mereology and does not extend it).

As argued by the author, Tarski's work can be given a meaningful account to some otherwise logically unclear parts of it, only if one assume a framework that is coherent with Leśniewski's logic. Henkin [13] was the first to implement Leśniewski's ideas in a computational framework, i.e., Church's type theory [7]. Therefore, using a type-theoretical background for substituting protothetic in Leśniewski's mereology respects the original ideas. Being a "part of" is said to be a primitive notion in Lesniewski's Mereology, but sum is not. We advocate that the references of Tarski to Leśniewski's work are syntactically exact (e.g., the word "sum" in definition 8 refers to the definition of "collection" in mereology [20]). We advocate for a uniform modeling in which balls and solids are names, and more precisely plurals in the sense of Leśniewski. Such an assumption yields a quantification over variables that are always of type N (name) and avoids the issues concerning the domain of discourse. In Leśniewski's systems, collective and distributive classes coexist [9]. It follows that instead of expressing Tarski's mereogeometry in a set-based framework [12], a more coherent picture can be proposed using Leśniewski's axiomatization. Collective classes are addressed with the Kl construct, while distributive classes require new categories and definitions (in the form of equivalences). The

Russellian notion of class that appears in definition 6, is solved using a definition in the sense of Leśniewski by constraining the quantification through concentricity.

III. FORMALIZING TARSKI'S MEREOTHEORY

A. From Protothetic to Type Theory

The underlying idea holds in the specification of Leśniewski's theory with the purpose of handling requests exactly as proofs through the use of a theorem prover. In order to reach our objectives, we substitute protothetic with a more computable logic based on λ -calculus as developed by Church in its Simple Type Theory. The seminal works of [13] and Andrews [1], [2] have stated the basic lines of this logic under the name $\mathcal{Q}0$. All Leśniewski's definitions are replaced with variable introduction and the λ -abstractor [13].

B. Expressing Definitions of Mereogeometry in Higher-order Logic

The first commitment concerns balls² and solids which are considered as constant plurals, since (i) they denote a constant plurality and (ii) they refer to collections of objects. In other words, they are instances of names (N) which are constrained to appear as the right argument of the copula (ε). In LO, constant plurals (e.g., "empty name" or "universal name") are defined in a similar way: their category must be defined first (i.e., N) and then, a definition explains their property using an equivalence. Since "balls" is a primitive, no specific property is required. Alternatively, a solid does not refer to a primitive name. As a consequence, it belongs to the category N and is defined as a particular collection of balls (see definition 8 below).

Variable *balls* : N .
Variable *solids* : N .

²Called spheres in Tarski's paper.

In such a way the domain of discourse is that of truth values and names (basic categories) with a restriction to balls and solids. Notice that the restrictions are not given by specifying appropriate types but rather by theorems which directly constrain their use. Using the εpt construct, the four axioms of mereology and derived theorems, then it is easy to prove that the relation εel is a partial order.

We first introduce the short-hand symbols \leq and $<$ which stand respectively for εel and εpt . Then, several relations among balls are defined such as concentricity, relying on the intended interpretation of the primitives. Points are defined as (mereological) collections of concentric balls. Equidistance among points makes use of properties of concentric balls while Euclidean axioms are able to constrain equidistance. We introduce successively:

Variable et	: $N \rightarrow N$.
Variable it	: $N \rightarrow N$.
Variable edt	: $N \rightarrow N \rightarrow N$.
Variable idt	: $N \rightarrow N \rightarrow N$.
Variable con	: $N \rightarrow N$.
Variable $point$: $N \rightarrow N$.
Variable $equid$: $N \rightarrow N \rightarrow N$.
Variable $ipoint$: $N \rightarrow N$.

The relations of external tangency (εet), internal tangency (εit), external diametricity (εedt), internal diametricity (εidt), concentricity (εcon), point ($point$), equidistance ($\varepsilon equi$) and interior point ($\varepsilon ipoint$) are defined using the already defined name functors.

Definition 1: external tangency.

Variable ET	: $\forall A B, A \varepsilon et B \equiv (A \varepsilon balls \wedge B \varepsilon balls \wedge A \varepsilon ext B \wedge \forall X Y, (X \varepsilon balls \wedge Y \varepsilon balls \wedge A \leq X \wedge X \varepsilon ext B \wedge A \leq Y \wedge Y \varepsilon ext B) \rightarrow (X \leq Y \vee Y \leq X))$.
---------------	--

Definition 2: internal tangency.

Variable IT	: $\forall A B, A \varepsilon it B \equiv (A \varepsilon balls \wedge B \varepsilon balls \wedge A < B \wedge \forall X Y, (X \varepsilon balls \wedge Y \varepsilon balls \wedge A \leq X \wedge X \leq B \wedge A \leq Y \wedge Y \leq B) \rightarrow (X \leq Y \vee Y \leq X))$.
---------------	--

Definition 3: external diametrical tangency.

Variable EDT	: $\forall A B C, A \varepsilon edt B C \equiv (A \varepsilon balls \wedge B \varepsilon balls \wedge C \varepsilon balls \wedge B \varepsilon et A \wedge C \varepsilon et A \wedge \forall X Y, (X \varepsilon balls \wedge Y \varepsilon balls \wedge B \leq X \wedge X \varepsilon ext A \wedge C \leq Y \wedge Y \varepsilon ext A) \rightarrow X \varepsilon ext Y)$.
----------------	--

Definition 4: internal diametrical tangency.

Variable IDT	: $\forall A B C, A \varepsilon idt B C \equiv (A \varepsilon balls \wedge B \varepsilon balls \wedge C \varepsilon balls \wedge B \varepsilon it A \wedge C \varepsilon it A \wedge \forall X Y, (X \varepsilon balls \wedge Y \varepsilon balls \wedge X \varepsilon ext A \wedge Y \varepsilon ext A \wedge B \varepsilon ext X \wedge C \varepsilon ext Y) \rightarrow X \varepsilon ext Y)$.
----------------	--

Definition 5: concentric balls.

Variable CON	: $\forall A B, A \varepsilon con B \equiv (A \varepsilon balls \wedge B \varepsilon balls \wedge singular_equality A B \vee (A < B \wedge \forall X Y, (X \varepsilon balls \wedge Y \varepsilon balls \wedge A \varepsilon edt X Y \wedge X \varepsilon it B \wedge Y \varepsilon it B) \rightarrow B \varepsilon idt X Y) \vee (B < A \wedge \forall X Y, (X \varepsilon balls \wedge Y \varepsilon balls \wedge B \varepsilon edt X Y \wedge X \varepsilon it A \wedge Y \varepsilon it A) \rightarrow A \varepsilon idt X Y))$.
----------------	--

Definition 6: point.

Variable $POINT$: $\forall P B, P \varepsilon (point B) \equiv (P \varepsilon P \wedge B \varepsilon balls \wedge \forall B', B' \varepsilon balls \wedge B' con B)$.
------------------	--

In other words, a point stands for the set of all those balls that are concentric with a given ball. Definition 7: equidistance.

Variable $EQUID$: $\forall A B C, A \varepsilon equid B C \equiv (A \varepsilon balls \wedge B \varepsilon balls \wedge C \varepsilon balls \wedge \exists X, (X \varepsilon balls \wedge X \varepsilon con A \wedge \neg \exists Y, (Y \varepsilon balls \wedge Y \varepsilon (union B C) \wedge Y \leq X \vee Y \varepsilon ext X))$.
------------------	--

Definition 8: solids.

Variable $TarskiD8$: $\forall A, A \varepsilon solids \equiv \exists B, (B \varepsilon B \wedge B \varepsilon coll balls \wedge A \varepsilon subcoll B)$.
---------------------	--

Definition 9: interior point.

Variable $IPOINT$: $\forall P X C, P \varepsilon ipoint X \equiv (X \varepsilon solids \wedge P \varepsilon point C \wedge \exists A', (A' \varepsilon balls \wedge A' \varepsilon P \wedge A' \leq X))$.
-------------------	---

C. The Axiom System

The axiom system of Tarski for the geometry of solid can be broadly divided in three parts, (i) axioms stating the existence of a correspondence between notions of the geometry of solids and notions of ordinary point geometry, (ii) two axioms establishing a correspondence between notions of the geometry of solids and topology and (iii) internal axioms that are derivable from Leśniewski's mereology. Axioms of the former part are:

Axiom 1: The notions of point and equidistance of two points to a third satisfy all axioms of ordinary Euclidean geometry of three dimensions.

More specifically, this axiom states that (i) points as they are introduced in definition 6 correspond to points of an ordinary point-based geometry and (ii) the relation $EQUID$ corresponds to an ordinary equidistance relation. With Π standing for mereogeometrical points, the structure $\langle \Pi, EQUID \rangle$ is a Pieri's structure [21]. Then, it can be proved that $\langle \Pi, EQUID \rangle$ is isomorphic to ordinary Euclidian geometry $\langle \mathbb{R}^3, EQUID^{\mathbb{R}^3} \rangle$ (see [12] p 500).

Axiom 2: If A is a solid, the class α of all interior points of A is a non-empty regular open set.

Axiom 3: If the class α of points is a non-empty regular open set, there exists a solid A such that α is the class of all its interior points.

The second axiomatic part relies on the structure $\langle \Pi, EQUID \rangle$. It follows that we are able to define the family of open balls $\mathcal{O}b_{\Pi}$ in it and then introduce in Π the family \mathcal{O}_{Π} of open sets together with appropriate topological operations of closure and openness such that $\langle \Pi, \mathcal{O}_{\Pi} \rangle$ is a topological space. If in $\langle \Pi, \mathcal{O}_{\Pi} \rangle$, we introduce the family $\mathcal{O}r_{\Pi}^0$ of all regular open sets excluding the empty set, then it can be proved that $\langle \mathcal{O}r_{\Pi}^0, \mathcal{O}b_{\Pi}, \subseteq \rangle$ is isomorphic to $\langle \mathcal{O}r_{\mathbb{R}^3}^0, \mathcal{O}b_{\mathbb{R}^3}, \subseteq \rangle$ (see [12] for more details). It follows that the resulting axiom system provides a minimal system that can serve as a basis for constructing spatial theories.

IV. APPLICATIONS OF MEREOMETRY

Qualitative representations for cognitive systems are independent of specific values and granularities of representation. In such a way they allow for fast reasoning processes. Furthermore, qualitative knowledge should be robust under transformations. Using Tarki's mereogeometry (which itself includes

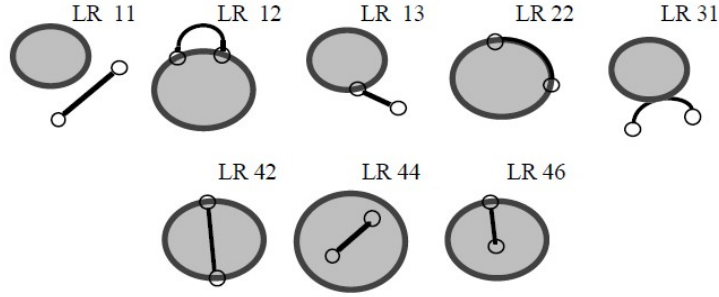


Fig. 2. An excerpt of the 19 line-region situations from the Egenhofer model.

Leśniewski's theories) it is possible to develop expressive and sound mereogeometries [6], [4], [12]. A point F is a fringe point relative to a solid A if no ball in F is either an element of or is exterior to A . With the notion of fringe point [12] and from definitions of Tarski's mereogeometry, we may simply introduce the *connection* relation as follows: solids A and B are connected iff they share at least a common fringe point.

Variable *FringePt* : $\forall B F A, F \in \text{Fringe } A \equiv (A \in \text{solids} \wedge B \in \text{balls} \wedge F \in (\text{point } B) \wedge \neg(B \leq A) \wedge \neg(B \in \text{ext } A))$.

Variable *C* : $\forall A B, \text{Connect } A B \equiv (A \in \text{solids} \wedge B \in \text{solids} \wedge \exists P, (P \in \text{Fringe } A \wedge P \in \text{Fringe } B))$.

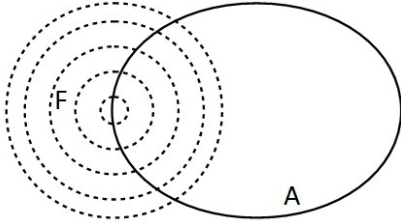


Fig. 3. A Fringe point.

where *Fringe* and *Connect* respectively belong to the categories $N \rightarrow N$ and $N \rightarrow N \rightarrow S$, \leq denotes the element-of relation, while *ext* refers to a mereological definition whose meaning is "external to". The connection function which requires the concept of point, is defined within mereogeometry and not in mereology as in usual mereotologies. The *between (bet)* definition provides the required property that holds when a given ball B stands between balls A and B [6].

Variable *BTW* : $\forall A B C, B \in \text{bet } A C \equiv (A \in \text{balls} \wedge B \in \text{balls} \wedge C \in \text{balls} \wedge \exists A' B' C', A' \in \text{balls} \wedge B' \in \text{balls} \wedge C' \in \text{balls} \wedge A' \in \text{con } A \wedge B' \in \text{con } B \wedge C' \in \text{con } C \wedge B' \in \text{edt } A' C')$.

We introduce then the notion of maximal ball in 2D solids. A ball B in a solid A is said to be maximal if there does not exist any other ball B' of which it is part.

Variable *MB* : $\forall B S, B \in \text{maxBall } S \equiv (S \in \text{solids} \wedge B \in \text{balls} \wedge B \leq S \wedge \forall B', (B' \in \text{balls} \wedge B' \leq S \wedge B \leq B') \rightarrow \text{singular_equality } B B')$

We can next introduce the notion of 1D solid as follows:

Variable *S1D* : $\forall A, A \in \text{solid_1D} \equiv \forall B, (B \in \text{Fringe } A \wedge B \in \text{ipoint } A)$.

As an application example, we show how to describe natural language based situations when a line "intersects" a given region in a geographic setting [18]. Different possibilities are summarized in figure 2. Each situation must involve a combination of the basic relations that have been described so far.

Situation LR11: the line is external to the region.

Variable *LR11* : $\forall A B, A \in \text{lr11 } B \equiv (A \in \text{solid_1D} \wedge B \in \text{solids} \wedge A \in \text{ext } B)$

Situation LR12: the external line touches the region in two extremal points.

Variable *LR12* : $\forall A B, A \in \text{lr12 } B \equiv (A \in \text{solid_1D} \wedge B \in \text{solids} \wedge \exists F F', ((F \in \text{Fringe } A \wedge F \in \text{Fringe } B \wedge F' \in \text{Fringe } A \wedge F' \in \text{Fringe } B) \wedge \forall P, (P \in \text{Fringe } A \wedge P \in \text{Fringe } B) \rightarrow P \in \text{con } F \vee P \in \text{con } F')) \wedge \exists P', P' \in \text{ipoint } A \wedge P' \in \text{ext } B)$

Situation LR13: the external line touches the region in a single extremal point.

Variable *LR13* : $\forall A B, A \in \text{lr13 } B \equiv (A \in \text{solid_1D} \wedge B \in \text{solids} \wedge \exists F, (F \in \text{Fringe } A \wedge F \in \text{Fringe } B) \wedge \forall P, (P \in \text{Fringe } A \wedge P \in \text{Fringe } B) \rightarrow P \in \text{con } F)$

Situation LR22: the line is a part of the border of the region.

Variable *LR22* : $\forall A B, A \in \text{lr22 } B \equiv (A \in \text{solid_1D} \wedge B \in \text{solids} \wedge \forall F, (F \in \text{Fringe } A \wedge F \in \text{Fringe } B))$

Situation LR31: the line touches the region on a part of its border.

Variable *LR31* : $\forall A B, A \in \text{lr31 } B \equiv (A \in \text{solid_1D} \wedge B \in \text{solids} \wedge \exists F, (F \in \text{Fringe } A \wedge F \in \text{Fringe } B))$

Situation LR42: line internal to the region with two contact points on the border.

Variable *LR42* : $\forall A B, A \in \text{lr42 } B \equiv (A \in \text{solid_1D} \wedge B \in \text{solids} \wedge \exists F F', (F \in \text{Fringe } A \wedge F \in \text{Fringe } B \wedge F' \in \text{Fringe } A \wedge F' \in \text{Fringe } B) \wedge \forall P, ((P \in \text{Fringe } A \wedge P \in \text{Fringe } B) \rightarrow P \in \text{con } F \vee P \in \text{con } F')) \wedge \forall P, (P \in \text{ipoint } A \wedge P < B) \vee P \in \text{Fringe } B)$

Situation LR44: line internal to the region.

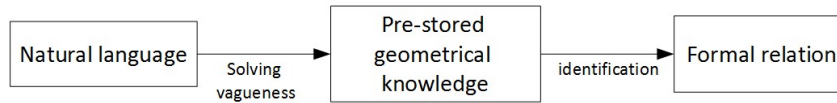


Fig. 4. Overview of the proposed architecture.

Variable $LR44 : \forall A B, A \in lr44 B \equiv (A \in solid_ID \wedge B \in solids \wedge A < B)$

Situation LR46: line internal to the region with one contact point on the border.

Variable $LR46 : \forall A B, A \in lr46 B \equiv (A \in solid_ID \wedge B \in solids \wedge \exists F, (F \in Fringe A \wedge F \in Fringe B \wedge \forall P, (P \in Fringe A \wedge P \in Fringe B) \rightarrow P \in con F) \wedge \forall P, (P \in ipoint A \wedge P < B) \vee P \in Fringe B)$

Based on previous works, we specify the semantics of spatial relations in natural language in a two-step process (see fig. 4). In the first part, all possible sentences which characterize formal spatial relations which might apply to the spatial setting are stored in a file. The first step involves an efficient algorithm for sentence similarity [15] whose purpose is to address the problem of vagueness inherent in natural language. The sentence-similarity algorithm has been implemented with very low computing time (\approx 1ms for two sentences of about 20 words). The best score resulting from the algorithm extracts one of the relations described above in a second step. Then this relation can be inserted in an automated theorem prover for further reasoning.

V. CONCLUSION

Generalizing solids to spatial regions, geometrical theories based on mereology present an appealing impact on spatial theories. The theory of Tarski, has been proved to be semantically complete with regards to the models expressed in terms of R^n and has been axiomatized by Bennett [4] using a set-based interpretation. We have shown in [11] that this assumption (i.e., interpreting $A \in pt B$ as $A \subseteq B$) considerably weakens the logical power of Leśniewski's framework. We have set a logical foundation for qualitative spatial representation with the following properties: (i) the proposed set of structures featuring geometrical entities and relations relies on Tarski's mereogeometry, (ii) it has a model in ordinary three-dimensional Euclidian geometry [22], (iii) it is coherent with Leśniewski's mereology and does not suffer the defects cited in [5] and (iv) it will serve as a basis for spatial reasoning with full compliance with Leśniewski's systems. Further works will focus on automated reasoning with the set of spatial relations.

REFERENCES

- [1] P. Andrews, A reduction of the axioms for the theory of propositional types, *Fundamenta Mathematicae*, 52 (1963) 345–350.
- [2] P. Andrews, An Introduction to Mathematical Logic and Type Theory: to Truth through Proof, Applied Logic Series, 27, Kluwer Academic Publishers, 2nd edition, 2002.
- [3] M. Aurnague and L. Vieu, *A three-level approach to the semantics of space*, In C. Zelinsky-Wibbelt, editor, *The semantics of prepositions - from mental processing to natural language processing*, Berlin, 1993.
- [4] B. Bennett, *A categorical axiomatisation of region-based geometry*. *Fundamenta Informaticae* 46(1-2), 145–158, 2001.
- [5] A. Betti and I. Loeb, *On Tarski's foundations of the geometry of solids*. *Bulletin of Symbolic Logic*, 18(2), 230–260, 2012.
- [6] S. Borgo, N. Guarino and C. Masolo, *A pointless theory of space based on strong congruence and connection*. *Procs. of 5th Int. Conf. on Principle of Knowledge Representation and Reasoning (KR'96)*, Morgan Kaufmann, 220–229, 1996.
- [7] A. Church, *A formulation of the simple theory of types*. *Journal of Symbolic Logic*, 5, 56–68, 1940.
- [8] B.L. Clarke, *A Calculus of Individuals Based on 'Connection'*, *Notre Dame Journal of Formal Logic*, 22(3) (1981) 204–218.
- [9] R. E. Clay *Relation of Leśniewski's Mereology to Boolean Algebra*. *The Journal of Symbolic Logic* 39(4), 638–648, 1974.
- [10] A.G. Cohn, J. Renz, *Qualitative Spatial Representation and Reasoning*, *Handbook of Knowledge Representation*, 551–596, 2008.
- [11] R. Dapoigny and P. Barlatier, *A Coq-based axiomatization of Tarski's Mereogeometry*, *COSIT'2015*, 108–109, 2015.
- [12] R. Gruszczynski and R. Pietruszczak, *Full Development of Tarski's Geometry of Solids*. *The Bulletin of Symbolic Logic* 14(4), 481–540, 2008.
- [13] L. Henkin, *A theory of propositional types*. *Fundamenta Mathematicae*, 52 323–334. *Errata*, 53, 119, 1963.
- [14] S. Leśniewski, *Podstawy ogólnej teorii mnogości. I*, Moskwa: Prace Polskiego Kola Naukowego w Moskiewie, Sekcja matematyczno-przyrodnicza, 1916 (English translation by D. I. Barnett: *Foundations of the General Theory of Sets. I*, in S. Leśniewski, *Collected Works*, ed. S. J. Surma, J. Srzednicki, D. I. Barnett, and F. V. Rickey, Dordrecht: Kluwer, 1, 129–173, 1992).
- [15] Y. Li, D. McLean, Z.A. Bandar, J. D. O'Shea, and K. Crockett, *Sentence Similarity Based on Semantic Nets and Corpus Statistics* *IEEE Trans. on Knowledge and Data Engineering*, 18(8), 1138–1150, 2006.
- [16] H. A. Mallot and J. S. Allen, *Computational Vision: Information Processing in Perception and Visual Behavior*, MIT Press, Cambridge, MA, 2000.
- [17] D. A. Randell, Z. Cui, and A. G. Cohn, *A spatial logic based on regions and connection*, In *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*, 165–176, Morgan Kaufmann, 1992.
- [18] Shariff, A.R., Egenhofer, M.J., Mark, D.M.: *Natural language spatial relations between linear and areal objects: The topology and metric of English language terms*, *Int. Journal of Geographical Information Science* 12(3) 215–246 1998.
- [19] P. Simons. *Parts: A Study In Ontology*. Clarendon Press, Oxford, 1987.
- [20] V.F. Siniśi, *Leśniewski's foundations of mathematics*. *Topoi*, 2(1), 3–52, 1983.
- [21] A. Tarski, *Les fondements de la géométrie des corps (Foundations of the geometry of solids)*. in *Księga Pamiątkowa Pierwszego Polskiego Zjazdu Matematycznego*, 7, 29–33, 1929
- [22] A. Tarski, *Foundations of the geometry of solids*. in *Logics, Semantics, Metamathematics. Papers from 1923-1938 by Alfred Tarski*, Clarendon Press, 1956.
- [23] S. Thrun, *Robotic mapping: A survey* In G. Lakemeyer and B. Nebel, editors, *Exploring Artificial Intelligence in the New Millenium*, Morgan Kaufmann, 2002.
- [24] A. Varzi, *Parts, wholes and part-whole relations: the prospects of mereotopology*, *Data and Knowledge Engineering*, 20(3):259–286, 1996.
- [25] M. F. Worboys and M. Duckham, *Geographic Information Systems: A Computing Perspective (2nd Edition)*, CRC Press, Boca Raton, Florida, 2004.