ECAI 2014
T. Schaub et al. (Eds.)

doi:10.3233/978-1-61499-419-0-997

A Logic of Part and Whole for Buffered Geometries

Heshan Du and **Natasha Alechina**[1]

Abstract. We propose a new qualitative spatial logic for reasoning about part-whole relations between geometries (sets of points) represented in different geospatial datasets, in particular crowd-sourced datasets. Since geometries in crowd-sourced data can be less inaccurate or precise, we buffer geometries by a margin of error or level of tolerance σ, and define part-whole relation for buffered geometries. The relations between geometries considered in the logic are: buffered part of (BPT), Near and Far. We provide a sound and complete axiomatisation of the logic with respect to metric models, and show that its satisfiability problem is NP-complete.

1 MOTIVATION

This work is motivated by our previous work [2] on integrating authoritative geospatial information and crowd-sourced or volunteered geospatial information. Geometry representations of the same location or place in different datasets are usually not exactly the same. Objects are also sometimes represented at different levels of granularity. For example, consider geometries of objects in Nottingham city centre given by the Ordnance Survey of Great Britain (OSGB) [6] and by the OpenStreetMap (OSM) [5] in Figure 1. The position and shape of the Prezzo Ristorante are represented differently in OSGB (light) and OSM (dark). The Victoria Shopping Centre is represented as a whole in OSM, and as several shops in OSGB.

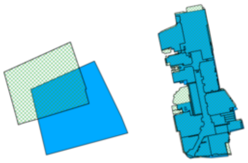

Figure 1. Prezzo Ristorante (left) and Victoria Shopping Centre (right)

In order to integrate the datasets, we need to determine which objects are the same and sometimes (as in the example of Victoria Shopping Centre) which objects in one dataset are parts of objects in another. One way to produce such matches is to use locations and geometries of objects, although of course we also use any lexical labels associated with the objects, such as names of restaurants etc. The generated matches are seen as assumptions, and are retractable if found incorrect. We check correctness of matches by checking their logical consistency. Some of the checks use ontology reasoning (if an object is classified as a restaurant in one dataset and as a bank in another, together with an axiom stating that the concepts of Restaurant and Bank are disjoint, a contradiction can be derived). Other checks are performed using spatial reasoning. In [3], we proposed a spatial logic LNF that contains relations of being buffered equal (BEQ), Near and Far to validate 'sameAs' matches of objects: if it is conjectured that a_1 is 'sameAs' b_1 and that a_2 is 'sameAs' b_2, then a contradiction can be derived if $NEAR(a_1, a_2)$ and $FAR(b_1, b_2)$. However, LNF is not appropriate for verifying 'partOf' matches. In this paper, we are proposing a logic where we can formalise for example the following simple argument: a cannot be part of b_1 and of b_2 where $FAR(b_1, b_2)$ holds. The main concepts of this logic, which we call a Logic of ParT and whole for Buffered geometries (LBPT), are explained in the next section. We also compare it to existing spatial logics.

[1] University of Nottingham, UK, email: hxd@cs.nott.ac.uk, nza@cs.nott.ac.uk

2 BPT, NEAR, FAR

For the application described in the previous section, we found it difficult to use formalisms such as RCC and other topology or mereology theories [1], since they presuppose accurate geometries or regions with sharp boundaries. Unlike existing models for spatial relations between indeterminate regions or objects with broad boundaries based on rough set theory [7], we could not define a certain inner region, because the same location can be represented using two disconnected polygons from authoritative and crowd-sourced geospatial datasets respectively, which requires that the whole region within the buffer [4] of a geometry is uncertain. The first logic we designed for debugging geometry matches, LNF [3], has the 'buffered equal' relation as a basic relation, which turns out to be less useful when the data is represented at different levels of abstraction (such as a shopping centre in one set and a collection of shops in another). In [3], we gave a sound and complete axiomatisation for LNF, but only with respect to geometries consisting of a single point. In this paper, we start with the 'buffered part of' relation as the basic relation, and interpret geometries as sets of points.

As shown in Figure 2.a, by buffering the solid circle c by σ, where σ indicates the margin of error or level of tolerance, we obtain a larger circle, denoted as $buffer(c, \sigma)$, where every point is within σ distance from c. For a geometry c which is possibly represented inaccurately within the margin of error σ, the actual accurate representation is assumed to be somewhere within $buffer(c, \sigma)$. A geometry g is buffered part of a geometry h, if g is within $buffer(h, \sigma)$ (Figure 2.b).

We also have NEAR and FAR relations in the logic LBPT. They formalise concepts of being 'possibly connected' (given a possible displacement by σ) and 'definitely disconnected' (even if displaced by σ) respectively. Two geometries are possibly connected iff their σ

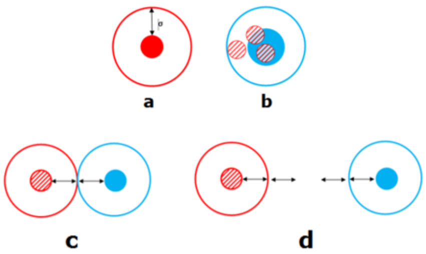

Figure 2. **a.** a buffer; **b.** three dashed circles are buffered part of (BPT) the solid circle; **c.** NEAR; **d.** FAR

buffers are connected. Figure 2.c and Figure 2.d show the boundary case of being NEAR, where $distance(g,h) = 2\sigma$ (their buffers are externally connected) and the case where two geometries are far apart and cannot possibly correspond to connected objects respectively.

3 SYNTAX, SEMANTICS AND AXIOMS

The language $L(LBPT)$ contains a set of individual names (terms), three binary predicates $NEAR$, FAR and BPT, and logical connectives, $\neg, \wedge, \vee, \rightarrow$.

Applying predicate letters to terms yields atomic formulas, e.g. $BPT(a,b)$. Every atomic formula is a well-formed formulas (wffs). If α and β are wffs, then $\neg\alpha$, $\alpha \wedge \beta$, $\alpha \vee \beta$, $\alpha \rightarrow \beta$ are wffs.

The language $L(LBPT)$ contains binary predicates $NEAR$, FAR and BPT. The formal definition of the alphabet is as follows:

We interpret the logic over models which are based on a metric space (similar to other spatial logics, such as [8] and [1]).

Definition 1 (Metric Space) *A* metric space *is a pair* (Δ, d)*, where* Δ *is a set and* d *is a metric on* Δ*, i.e., a function* $d : \Delta \times \Delta \longrightarrow \mathbb{R}_{\geq 0}$ *such that for any* $x, y, z \in \Delta$*, the following holds:*

1. $d(x,y) = 0$ *iff* $x = y$;
2. $d(x,y) = d(y,x)$;
3. $d(x,z) \leq d(x,y) + d(y,z)$.

Definition 2 (Metric Model) *A metric model* M *is a tuple* (Δ, d, I, σ)*, where* (Δ, d) *is a metric space,* I *is an interpretation function which maps each constant to a set of elements in* Δ*, and* $\sigma \in \mathbb{R}_{>0}$ *is the margin of error. The notion of* $M \models \phi$ *(*ϕ *is true in model* M*) is defined as follows:*

- $M \models BPT(a,b)$ *iff* $\forall p_a \in I(a) \exists p_b \in I(b) : d(p_a, p_b) \in [0, \sigma]$;
- $M \models NEAR(a,b)$ *iff* $\exists p_a \in I(a) \exists p_b \in I(b) : d(p_a, p_b) \in [0, 2\sigma]$;
- $M \models FAR(a,b)$ *iff* $\forall p_a \in I(a) \forall p_b \in I(b) : d(p_a, p_b) \in (4\sigma, +\infty)$;
- $M \models \neg\alpha$ *iff* $M \not\models \alpha$;
- $M \models \alpha \wedge \beta$ *iff* $M \models \alpha$ *and* $M \models \beta$;
- $M \models \alpha \vee \beta$ *iff* $M \models \alpha$ *or* $M \models \beta$;
- $M \models \alpha \rightarrow \beta$ *iff* $M \not\models \alpha$ *or* $M \models \beta$

where a, b *are individual names,* α, β *are wffs.*

A formula α is valid ($\models \alpha$) if for every metric model M, $M \models \alpha$. The logic LBPT is the set of all valid formulas of $L(LBPT)$.

The following calculus (that we will also refer to as LBPT) has been proved sound and complete for LBPT:

A 0 All tautologies of classical propositional logic
A 1 $BPT(a,a)$;
A 2 $NEAR(a,b) \rightarrow NEAR(b,a)$;
A 3 $FAR(a,b) \rightarrow FAR(b,a)$;
A 4 $BPT(a,b) \wedge BPT(b,c) \rightarrow NEAR(c,a)$;
A 5 $BPT(b,a) \wedge BPT(b,c) \rightarrow NEAR(c,a)$;
A 6 $BPT(b,a) \wedge NEAR(b,c) \wedge BPT(c,d) \rightarrow \neg FAR(d,a)$;
A 7 $NEAR(a,b) \wedge BPT(b,c) \wedge BPT(c,d) \rightarrow \neg FAR(d,a)$;
MP Modus ponens: $\phi,\ \phi \rightarrow \psi \ \vdash\ \psi$.

The notion of derivability $\Gamma \vdash \phi$ in LBPT is standard. A formula ϕ is LBPT-derivable if $\vdash \phi$. A set Γ is (LBPT) inconsistent if for some formula ϕ it derives both ϕ and $\neg\phi$.

4 SOUNDNESS, COMPLETENESS, DECIDABILITY AND COMPLEXITY

We proved that LBPT is sound and complete for metric models.

Theorem 1 (Soundness) *Every LBPT derivable formula is valid:* $\vdash \phi \Rightarrow \models \phi$.

Theorem 2 (Completeness) *Every LBTP valid formula is derivable:* $\models \phi \Rightarrow \vdash \phi$.

For completeness, we actually proved that given a finite consistent set of formulas, we can build a satisfying metric model for it. This shows that $\not\vdash \phi \Rightarrow \not\models \phi$ and by contraposition we get completeness. From the bound on the size of the satisfying model, we also have the following theorem:

Theorem 3 (Decidability and Complexity) *The LBPT satisfiability problem is NP-complete.*

Detailed proofs can be found here: `www.cs.nott.ac.uk/~hxd/report/lbpt.pdf`.

5 CONCLUSION

We presented a logic LBPT which formalizes the concepts of being 'possibly part of' (BPT), 'possibly connected' (NEAR) and 'definitely disconnected' (FAR). We provided a sound and complete axiomatisation of it with respect to metric models and showed that its satisfiability problem is NP-complete. A LBPT reasoner is under development and testing, for validating 'partOf' matches between objects from authoritative and crowd-sourced geospatial datasets.

REFERENCES

[1] M. Aiello, I. Pratt-Hartmann, and J. Benthem, *Handbook of Spatial Logics*, Springer, 2007.
[2] H. Du, N. Alechina, M.J. Jackson, and G. Hart, 'Matching Formal and Informal Geospatial Ontologies', in *Geographic Information Science at the Heart of Europe*, Lecture Notes in Geoinformation and Cartography, 155–171, Springer, (2013).
[3] H. Du, N. Alechina, K. Stock, and M.J. Jackson, 'The Logic of NEAR and FAR', in *Conference On Spatial Information Theory, COSIT 2013*, volume 8116 of *LNCS*, pp. 475–494. Springer, (2013).
[4] ISO Technical Committe 211, 'ISO 19107:2003 Geographic information – Spatial schema', Technical report, International Organization for Standardization (TC 211), (2003).
[5] OpenStreetMap. The Free Wiki World Map. http://www.openstreetmap.org, 2012.
[6] Ordnance Survey. Ordnance Survey. http://www.ordnancesurvey.co.uk/oswebsite, 2012.
[7] Z. Pawlak, L. Polkowski, and A. Skowron, 'Rough Set Theory', in *Wiley Encyclopedia of Computer Science and Engineering*, (2008).
[8] F. Wolter and M. Zakharyaschev, 'Reasoning about Distances', in *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI-03)*, pp. 1275–1282, (2003).