On Redundant Topological Constraints* (Extended Abstract)

Matt Duckham
Department of Infrastructure Engineering
University of Melbourne

Sanjiang Li, Weiming Liu, Zhiguo Long
Centre for Quantum Computation & Intelligent Systems
Faculty of Engineering and Information Technology
University of Technology Sydney

Abstract

The Region Connection Calculus (RCC) is a well-known calculus for representing part-whole and topological relations. It plays an important role in qualitative spatial reasoning, geographical information science, and ontology. The computational complexity of reasoning with RCC has been investigated in depth in the literature. Most of these works focus on the consistency of RCC constraint networks. In this paper, we consider the important problem of redundant RCC constraints. For a set Γ of RCC constraints, we say a constraint (xRy) in Γ is *redundant* if it can be entailed by the rest of Γ. A *prime subnetwork* of Γ is a subset of Γ which contains no redundant constraints but has the same solution set as Γ. It is natural to ask how to compute a prime subnetwork, and when it is unique. In this paper, we show that this problem is in general intractable, but becomes tractable if Γ is over a tractable subclass of RCC. If $\mathcal{S}$ is a tractable subclass in which weak composition distributes over non-empty intersections, then we can show that Γ has a unique prime network, which is obtained by removing all redundant constraints from Γ. As a byproduct, we identify a sufficient condition for a path-consistent network being minimal.

1 Introduction

Qualitative spatial reasoning (QSR) is a common subfield of artificial intelligence and geographical information science, and has applications in GIS, cognitive robotics, high-level understanding of video data etc. The Region Connection Calculus (RCC) (Randell, Cui, and Cohn 1992) is perhaps the most well-known calculus for representing qualitative spatial information. Based on a binary connectedness relation, it defines a class of binary topological relations between regions in a connected topological space (e.g., the real plane). The RCC is an expressive formalism for representing topological information, and the computational complexity of reasoning with RCC has been investigated in depth in the literature. Most of these works focus on the consistency or satisfiablility of RCC constraint networks.

In this paper, we consider the important problem of redundant RCC constraints. Given a set Γ of RCC constraints, we say a constraint (xRy) in Γ is *redundant* if it can be entailed by the rest of Γ, i.e., removing (xRy) from Γ will not change the solution set of Γ. It is natural to ask when a network is redundant and how to get an irredundant subset without changing the solution set. We call a subset of Γ a *prime subnetwork* of Γ if it contains no redundant constraints but has the same solution set as Γ.

We show that it is in general co-NP hard to determine if a constraint is redundant in a network of RCC constraints, but Γ is over a tractable subclass, then a prime subnetwork can be found in $O(n^5)$ time. If in addition weak composition distributes over non-empty intersections of relations in $\mathcal{S}$, then Γ has a unique prime subnetwork, which is obtained by removing all redundant constraints from Γ.

As in the case of propositional logic formulas (Liberatore 2005), redundancy of RCC constraints "often leads to unnecessary computation, wasted storage, and may obscure the structure of the problem" (Belov et al. 2012). Finding a prime subnetwork can be useful in at least the following aspects: a) computing and storing the relationships between spatial objects and hence saving space for storage and communication; b) facilitating comparison between different constraint networks; c) handling the inconsistency by modifying critical constraints; d) unveiling the essential graphical structure of a network; and e) adjusting geometrical objects to meet topological constraints (Wallgrün 2012).

Motivational Example: Placename Footprints

To motivate our discussion, we focus briefly on one specific application to illustrate just one of our five uses of prime subnetworks: saving space for storage. Figure 1 gives a small example of a set of spatial regions formed by the geographic "footprints" associated with placenames in the Southampton area of the UK. The footprints are derived from crowd-sourced data, formed from the convex hull of the sets of coordinate locations at which individuals used the placenames on social media. Using such data sets in natural language placename searches frequently requires queries over the topological relationships between footprints (e.g., "is Clarence Pier in Southampton?"). Computing such relationships on-the-fly requires computationally intensive and slow geometric operations; by contrast Web-search queries demand rapid responses.

One potential solution is to cache the topological rela-

*Future improvements and extended versions of this paper will be published in CORR at http://arxiv.org/abs/1403.0613

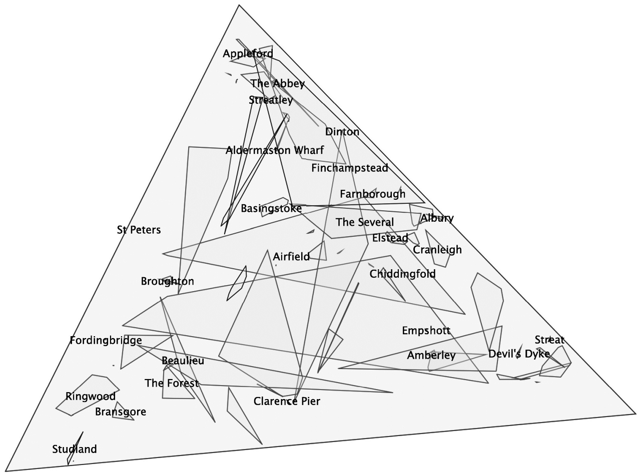

Figure 1: Examples of crowd-sourced geographic place-name "footprints" around Southampton, UK

tions between all footprints of interest. However, even the small example in Figure 1, the 84 footprints then require $84*83/2 = 3486$ stored relations. The moderate-sized footprint data set from which Figure 1 contains a total of 3474 footprints leads to a constraint network with $6,032,601$ relations. It is easy to see that as crowd-sourced data sources continue to grow, the volumes of such data is set to explode. In the case of footprints, many of the relationships can be inferred, and computing the prime subnetwork can reduce the number of stored relationships to be approximately linear in the number of footprints. In the case of the Southampton constraint network, 1324 redundant relations lead to a prime subnetwork with only 2150 relations needing to be stored. For the full data set, $5,713,563$ redundant relations lead to a prime subnetwork of just $319,038$ relations (in contrast to the full constraint network of more than 6 million relations).

In Section 2 we recall the RCC constraint language and then define the key notions of redundant constraint and prime subnetwork in Section 3. We present our major results in Section 4 and conclude the paper in Section 5.

2 RCC Constraint Language

The RCC was introduced in (Randell, Cui, and Cohn 1992). Let U be the set of nonempty regular closed sets of $\mathbb{R}^2$. We call each element in U a region. For two regions a, b, we say a is a *part* of b, written $a\mathbf{P}b$, if $a \subseteq b$; say a is *connected* to b, written $a\,\mathbf{C}\,b$, if $a \cap b \neq \varnothing$. Using $\mathbf{C}$ and $\mathbf{P}$, we define

$$\begin{aligned}
x\,\mathbf{PP}\,y &\equiv x\,\mathbf{P}\,y \wedge \neg(y\,\mathbf{P}\,x)\\
x\mathbf{O}y &\equiv (\exists z)(z\,\mathbf{P}\,x \wedge z\,\mathbf{P}\,y)\\
x\,\mathbf{DR}\,y &\equiv \neg(x\mathbf{O}y)\\
x\,\mathbf{PO}\,y &\equiv x\mathbf{O}y \wedge \neg(x\,\mathbf{P}\,y) \wedge \neg(y\,\mathbf{P}\,x)\\
x\,\mathbf{EQ}\,y &\equiv x\,\mathbf{P}\,y \wedge y\,\mathbf{P}\,x\\
x\,\mathbf{DC}\,y &\equiv \neg(x\mathbf{C}y)\\
x\,\mathbf{EC}\,y &\equiv x\,\mathbf{C}\,y \wedge \neg(x\mathbf{O}y)\\
x\,\mathbf{TPP}\,y &\equiv x\,\mathbf{PP}\,y \wedge (\exists z)(z\,\mathbf{EC}\,x \wedge z\,\mathbf{EC}\,y)\\
x\,\mathbf{NTPP}\,y &\equiv x\,\mathbf{PP}\,y \wedge \neg(x\,\mathbf{TPP}\,y)
\end{aligned}$$

Write $\mathbf{PP}^{-1}$, $\mathbf{TPP}^{-1}$ and $\mathbf{NTPP}^{-1}$ for the converses of $\mathbf{PP}$, $\mathbf{TPP}$ and $\mathbf{NTPP}$, respectively. Then

(1) $\quad \mathcal{B}_5 = \{\mathbf{DR}, \mathbf{PO}, \mathbf{EQ}, \mathbf{PP}, \mathbf{PP}^{-1}\}$

(2) $\quad \mathcal{B}_8 = \{\mathbf{DC}, \mathbf{EC}, \mathbf{PO}, \mathbf{EQ}, \mathbf{TPP}, \mathbf{NTPP}, \mathbf{TPP}^{-1}, \mathbf{NTPP}^{-1}\}$

are two jointly exhaustive and pairwise disjoint (JEPD) sets of relations, i.e., for any two regions $a, b \in U$, a, b is related by exactly one relation in $\mathcal{B}_l$ ($l = 5, 8$). We call the Boolean algebra generated by relations in $\mathcal{B}_l$ the RCCl algebra, which consists all relations that are unions of the basic relations in $\mathcal{B}_l$. For convenience, we denote a non-basic RCCl relation R as the subset of $\mathcal{B}_l$ it contains. For example, we write $\{\mathbf{DR}, \mathbf{PO}, \mathbf{PP}\}$ for the relation $\mathbf{DR} \cup \mathbf{PO} \cup \mathbf{PP}$, and write $\star_5$ and $\star_8$ for the universal relation.

The composition of two basic RCC5/8 relations is not necessarily a relation in RCC5/8. For two RCC5/8 relations R and S, we call the smallest relation in RCC5/8 that contains $R \circ S$ the *weak composition* of R and S, written $R \circ_w S$ (Düntsch, Wang, and McCloskey 2001; Li and Ying 2003).

RCC5/8 Constraint Network

An RCC5/8 constraint has the form (xRy), where x, y are variables taking values from U, the set of regions, R is an RCC5/8 relation (not necessarily basic). Given a set Γ of RCC5/8 constraints over variables $V = \{v_1, v_2, ..., v_n\}$, we say Γ is *consistent* or *satisfiable* if there is an assignment $\sigma : V \to U$ such that $(\sigma(v_i), \sigma(v_j))$ satisfies the constraint in Γ that relates v_i to v_j.

Without loss of generality, we assume Γ has the form $\{x_i R_{ij} x_j\}_{i,j=1}^n$, where, for any $1 \leq i, j \leq n$, there is a unique constraint R_{ij}, and $R_{ji} = R_{ij}^{\sim}$ and $R_{ii} = \mathbf{EQ}$. In this sense, we call Γ a *constraint network*. Let $\Gamma = \{x_i R_{ij} x_j\}_{i,j=1}^n$ and $\Gamma' = \{x_i R'_{ij} x_j\}_{i,j=1}^n$ be two RCC5/8 constraint networks. We say Γ and Γ' are *equivalent* if they have the same set of solutions; and say Γ *refines* Γ' if $R_{ij} \subseteq R'_{ij}$ for all (i, j). We say an RCC5/8 network Γ is a *basic* network if each constraint is either the universe relation or a basic relation; and say a basic network *complete* if there are no universal relations.

Suppose $\mathcal{S}$ is a subset of RCC5/8. We say an RCC5/8 network $\Gamma = \{v_i R_{ij} v_j\}$ is over $\mathcal{S}$ if $R_{ij} \in \mathcal{S}$ for every pair of variables v_i, v_j. The consistency problem over $\mathcal{S}$, written as $\mathsf{CSP}(\mathcal{S})$, is the decision problem of the consistency of an arbitrary constraint network over $\mathcal{S}$. It is well known that the consistency problem over RCC5/8, i.e., $\mathsf{CSP}(RCC5/8)$, is NP-complete and RCC8 has three maximal tractable subclasses that contain all basic relations (Renz 1999) and RCC5 has only one (Jonsson and Drakengren 1997).

We say a network $\Gamma = \{v_i R_{ij} v_j\}$ *path-consistent* if for every $1 \leq i, j, k \leq n$, we have

(3) $$R_{ij} \subseteq R_{ik} \circ_w R_{kj}.$$

A cubic algorithm, henceforth called the *path-consistency algorithm* or PCA, has been devised to enforce path-consistency. For any RCC5/8 network Γ, the PCA either detects inconsistency of Γ or returns a path-consistent network, written Γ_p, which is equivalent to Γ and known as the *algebraic closure* or *a-closure* of Γ (Ligozat and Renz 2004). It is easy to see that in this case Γ_p also refines Γ, i.e., we have $S_{ij} \subseteq R_{ij}$ for each constraint $(x_i S_{ij} x_j)$ in Γ_p.

Proposition 1. *Let $\mathcal{S}$ be a tractable subclass of RCC5/8 which contains all basic relations. An RCC5/8 network Γ over $\mathcal{S}$ is consistent if applying PCA to Γ does not result inconsistency.*

In particular, we have

Proposition 2. *A basic RCC5/8 network Γ is consistent if it is path-consistent.*

Distributive Subalgebra

Write $\widehat{\mathcal{B}}_5$ for the closure of $\mathcal{B}_5$ under converse, intersection, and weak composition in RCC5. Then $\widehat{\mathcal{B}}_5$ contains the basic relations as well as

$$\{\mathbf{PO},\mathbf{PP}\},\{\mathbf{PO},\mathbf{PP}^{-1}\},\{\mathbf{PO},\mathbf{PP},\mathbf{PP}^{-1},\mathbf{EQ}\},$$
$$\{\mathbf{DR},\mathbf{PO},\mathbf{PP}\},\{\mathbf{DR},\mathbf{PO},\mathbf{PP}^{-1}\},\{\mathbf{DR},\mathbf{PO}\},\star_5,$$

where $\star_5 = \{\mathbf{DR},\mathbf{PO},\mathbf{PP},\mathbf{PP}^{-1},\mathbf{EQ}\}$. It is interesting to note that in $\widehat{\mathcal{B}}_5$ the weak composition operation is *distributive* over nonempty intersections in the following sense.

Lemma 3. *Let R,S,T be three relations in $\widehat{\mathcal{B}}_5$. Suppose $S\cap T$ is nonempty. Then we have*

$$R \circ_w (S\cap T) = R\circ_w S \cap R\circ_w T \tag{4}$$
$$(S\cap T)\circ_w R = S\circ_w R \cap T\circ_w R. \tag{5}$$

In what follows, we call such a subclass a distributive subalgebra. Formally, we have

Definition 4. Let $\mathcal{S}$ be a subclass of RCC5/8. We say $\mathcal{S}$ is a *distributive subalgebra* if

- $\mathcal{S}$ contains all basic relations, and is closed under converse, weak composition, and intersection;
- weak composition distributes over nonempty intersections of relations in $\mathcal{S}$.

3 Redundant Constraint and Prime Subnetwork

We give a formal definition of redundant constraints.

Definition 5. Suppose Γ is an RCC5/8 network over variables $V = \{v_1,...,v_n\}$. We say Γ *entails* a constraint (v_iRv_j), written $\Gamma \models (v_iRv_j)$, if for every solution $\{a_1,...,a_n\}$ of Γ we have $(a_i,a_j)\in R$. A constraint (v_iRv_j) in Γ is *redundant* if $\Gamma\setminus\{(v_iRv_j)\}$ entails (v_iRv_j). We say Γ is *reducible* if it has a redundant constraint, and *irreducible* otherwise. We say a subset Γ' of Γ is a *prime* subnetwork of Γ if Γ' is irreducible and equivalent to Γ.

Each universal constraint $(v_i \star v_j)$ in Γ is, by definition, a redundant constraint in Γ.

Given an RCC5/8 network Γ, a very interesting question is, *how to find a prime subnetwork of* Γ? This problem is clearly at least as hard as determining if Γ is reducible. Similar to the case of propositional logic formulars (Liberatore 2005), we have the following result for RCC5/8.

Proposition 6. *Let Γ be an RCC5/8 network and suppose (xRy) is a constraint in Γ. It is co-NP-complete to decide if (xRy) is redundant in Γ.*

A naive method to obtain a prime subnetwork is to remove redundant constraints one by one from Γ until we get an irreducible network. Suppose we have an oracle which can tell if a constraint is redundant. Then in an additional $O(n^2)$ time we can find an irreducible network that is equivalent to Γ by removing several constraints from Γ.

Despite that it is in general intractable to determine if a constraint is redundant, we have a polynomial time procedure if the constraints are all taken from a tractable subclass.

Proposition 7. *Let $\mathcal{S}$ be a tractable subclass of RCC5/8 that contains all basic relations. Suppose Γ is a network over $\mathcal{S}$. Then in $O(n^3)$ time we can determine whether a constraint is redundant in Γ and in $O(n^5)$ time find all redundant constraints of Γ. In addition, a prime subnetwork for Γ can be found in $O(n^5)$ time.*

Definition 8. The *core* of an RCC5/8 network Γ, written Γ_c, is defined to be the set of non-redundant constraints in Γ.

It is easy to see that the core of Γ is contained in every prime subnetwork of Γ. *Are prime subnetworks unique*? In general this is not the case.

In the following we assume without loss of generality that Γ has the following property:

$$(\forall i,j)[(i\neq j)\rightarrow(\Gamma\not\models(v_i\,\mathbf{EQ}\,v_j))]. \tag{6}$$

In the next section we show that, if Γ is a constraint network over a distributive subalgebra of RCC5/8, then Γ_c is the unique prime network of Γ. This is quite surprising, as, in general, knowing that (xRy) and (uSv) are both redundant in Γ does not imply that (uSv) is redundant in $\Gamma\setminus\{(xRy)\}$.

4 Networks over a Distributive Subalgebra

In this section, we assume $\mathcal{S}$ is a distributive subalgebra of RCC5/8. Let Γ be a consistent network over $\mathcal{S}$ which satisfies (6). We show that Γ_c is equivalent to Γ and hence the unique prime network of Γ.

Definition 9 (cf. (Chandra and Pujari 2005; Liu and Li 2012)). Suppose $\Theta = \{v_iT_{ij}v_j\}_{1\leq i,j\leq n}$ is an RCC5/8 network. We say Θ is *minimal* if for every pair of variables v_i,v_j $(i\neq j)$ and every basic relation α in T_{ij}, there exists a solution $\{a_1,a_2,\cdots,a_n\}$ of Θ such that (a_i,a_j) is an instance of α.

Each consistent RCC5/8 constraint network has a unique minimal network, but it is in general NP-hard to compute it.

Notation: We write Γ_m for the minimal network of Γ, Γ_p for the a-closure of Γ, and Γ_c for the core of Γ.

To prove that Γ_c is equivalent to Γ, we need two important results. The first result, stated in Theorem 10, shows that Γ_m is exactly Γ_p. The second result, stated in Proposition 11, shows that a particular constraint (xRy) is redundant in Γ iff its corresponding constraint in Γ_p is redundant. Our main result, stated in Theorem 12, then follows immediately.

Theorem 10. *Let $\mathcal{S}$ be a distributive subalgebra of RCC5/8. Suppose Γ is a consistent network over $\mathcal{S}$ and Γ_p its a-closure. Then Γ_p is Γ_m, the minimal network of Γ.*

Proposition 11. *Let $\mathcal{S}$ be a distributive subalgebra of RCC5/8. Suppose Γ is a consistent network over $\mathcal{S}$ which satisfies* (6)*. Assume that (xRy) and (xSy) are the constraints from x to y in Γ and Γ_p respectively. Then (xRy) is redundant in Γ iff (xSy) is redundant in Γ_p.*

Recall that Theorem 10 asserts that Γ_p is minimal. As a consequence, we have our main result.

Theorem 12. *Let $\mathcal{S}$ be a distributive subalgebra of RCC5/8. Suppose Γ is a consistent network over $\mathcal{S}$ which satisfies* (6) *and Γ_c the core of Γ. Then Γ_c is equivalent to Γ and hence the unique prime network of Γ.*

Recall that Proposition 7 shows that the core of an RCC5/8 network over a tractable subclass can be found in $O(n^5)$ time. In the next subsection we show this can be improved if the network is over a distributive subalgebra.

A Cubic Algorithm for Computing the Core of Γ

We first consider the special case when Γ is path-consistent.

Lemma 13. *Let $\mathcal{S}$ be a distributive subalgebra of RCC5/8. Suppose Γ is a path-consistent network over $\mathcal{S}$. Then a constraint $(v_i R_{ij} v_j)$ is redundant in Γ iff $R_{ij} = \bigcap\{R_{ik} \circ_w R_{kj} : k \neq i, j\}$, i.e., R_{ij} is the intersection of the weak compositions of all paths from v_i to v_j which have length* 2*.*

Suppose Γ is a consistent network over a distributive subalgebra of RCC5/8 and satisfies (6). Proposition 11 and Lemma 13 suggest a simple way for computing Γ_c, the unique prime network of Γ. By Proposition 11, a constraint $(v_i R_{ij} v_j)$ in Γ is redundant iff the corresponding constraint $(v_i S_{ij} v_j)$ in Γ_p is redundant. Furthermore, Lemma 13 shows that $(v_i S_{ij} v_j)$ is redundant in Γ_p iff S_{ij} is the intersection of all $S_{ik} \circ_w S_{kj}$ ($k \neq i, j$). We hereby have the following cubic algorithm for finding all redundant constraints in Γ. For each constraint $(v_i S_{ij} v_j)$, to verify if $S_{ij} = \bigcap\{S_{ik} \circ_w S_{kj} : k \neq i, j\}$, we introduce a relation P_{ij} which consists of all basic relations α that are not in $S_{ik} \circ_w S_{kj}$ for some $k \neq i, j$ and then check if $P_{ij} \cup S_{ij}$ is the universal relation.

5 Conclusion

In this paper, we have systematically investigated the computational complexity of redundancy checking for RCC5/8 constraints. Although it is in general co-NP-complete, we have shown that a prime network can be found in $O(n^5)$ for any consistent network over a tractable subclass of RCC5/8. If the constraints are taken from a distributive subalgebra, we proved that the core of the constraint network is the unique prime network and can be found in cubic time.

Some of these results (e.g., all results before Theorem 10) can be applied to many other qualitative calculi e.g. Interval Algebra (Allen 1983) immediately, but Proposition 11 and Theorem 12 do use the particular algebraic properties of RCC5/8. Future work will consider how to extend our results to these qualitative calculi.

Acknowledgments

The authors are grateful to Ross Purves, University of Zürich, for helping to develop the footprint data set illustrated in Fig. 1. This work was partially supported by ARC (DP120103758, DP120104159) and NSFC (61228305).

Algorithm 1: Algorithm for finding redundant constraints in constraint networks over a distributive subalgebra $\mathcal{S}$ of RCC5/8, where $\star_l$ is the universal relation in RCCl.

Input: a consistent constraint network
$\Gamma = \{v_i R_{ij} v_j : 1 \leq i, j \leq n\}$ over $\mathcal{S}$.
Output: $Redun$: the set of redundant constraints of Γ.

$Redun \leftarrow \varnothing$;
$\Gamma_p =$ the a-closure of Γ;
for *each constraint* $(v_i S_{ij} v_j) \in \Gamma_p$ **do**
  $P_{ij} \leftarrow \varnothing$;
  for *each variable* $v_k \in V \setminus \{v_i, v_j\}$ **do**
    $temp_{ij} = S_{ik} \circ_w S_{kj}$;
    for *each basic RCC5/8 relation* α **do**
      if $\alpha \notin temp_{ij}$ **then**
        $P_{ij} \leftarrow P_{ij} \cup \{\alpha\}$;
      end
    end
  end
  if $P_{ij} \cup S_{ij} = \star_l$ **then**
    $Redun \leftarrow Redun \cup \{(v_i R_{ij} v_j)\}$;
  end
end

References

Allen, J. F. 1983. Maintaining knowledge about temporal intervals. *Commun. ACM* 26(11):832–843.

Belov, A.; Janota, M.; Lynce, I.; and Marques-Silva, J. 2012. On computing minimal equivalent subformulas. In *CP*, 158–174.

Chandra, P., and Pujari, A. K. 2005. Minimality and convexity properties in spatial CSPs. In *ICTAI*, 589–593.

Düntsch, I.; Wang, H.; and McCloskey, S. 2001. A relation-algebraic approach to the Region Connection Calculus. *Theor. Comput. Sci.* 255(1-2):63–83.

Jonsson, P., and Drakengren, T. 1997. A complete classification of tractability in RCC-5. *Journal of Artificial Intelligence Research* 6.

Li, S., and Ying, M. 2003. Region Connection Calculus: Its models and composition table. *Artif. Intell.* 145(1-2):121–146.

Liberatore, P. 2005. Redundancy in logic I: CNF propositional formulae. *Artif. Intell.* 163(2):203–232.

Ligozat, G., and Renz, J. 2004. What is a qualitative calculus? A general framework. In *PRICAI*, 53–64.

Liu, W., and Li, S. 2012. Solving minimal constraint networks in qualitative spatial and temporal reasoning. In *CP*, 464–479.

Randell, D. A.; Cui, Z.; and Cohn, A. G. 1992. A spatial logic based on regions and connection. In *KR*, 165–176.

Renz, J. 1999. Maximal tractable fragments of the Region Connection Calculus: A complete analysis. In *IJCAI*, 448–455.

Wallgrün, J. O. 2012. Exploiting qualitative spatial reasoning for topological adjustment of spatial data. In *SIGSPATIAL/GIS*, 229–238.