Efficiently Characterizing Non-Redundant Constraints in Large Real World Qualitative Spatial Networks

Michael Sioutis  
Université d’Artois  
CRIL CNRS UMR 8188  
France  
sioutis@cril.fr

Sanjiang Li  
University of Technology, Sydney  
QCIS, FEIT  
Australia  
sanjiang.li@uts.edu.au

Jean-François Condotta  
Université d’Artois  
CRIL CNRS UMR 8188  
France  
condotta@cril.fr

Abstract

RCC8 is a constraint language that serves for qualitative spatial representation and reasoning by encoding the topological relations between spatial entities. We focus on efficiently characterizing non-redundant constraints in large real world RCC8 networks and obtaining their prime networks. For a RCC8 network $\mathcal{N}$ a constraint is redundant, if removing that constraint from $\mathcal{N}$ does not change the solution set of $\mathcal{N}$. A prime network of $\mathcal{N}$ is a network which contains no redundant constraints, but has the same solution set as $\mathcal{N}$. We make use of a particular partial consistency, namely, $\sqsubseteq$-consistency, and obtain new complexity results for various cases of RCC8 networks, while we also show that given a maximal distributive subclass for RCC8 and a network $\mathcal{N}$ defined on that subclass, the pruning capacity of $\sqsubseteq$-consistency and $\Diamond$-consistency is identical on the common edges of $G$ and the complete graph of $\mathcal{N}$, when $G$ is a triangulation of the constraint graph of $\mathcal{N}$. Finally, we devise an algorithm based on $\sqsubseteq$-consistency to compute the unique prime network of a RCC8 network, and show that it significantly progresses the state-of-the-art for practical reasoning with real RCC8 networks scaling up to millions of nodes.

1 Introduction

The Region Connection Calculus (RCC) is the dominant approach in Artificial Intelligence, and Knowledge Representation in particular, for representing and reasoning about topological relations [Randell et al., 1992]. RCC can be used to describe regions that are non-empty regular subsets of some topological space by stating their topological relations to each other. RCC8 is the constraint language formed by the following 8 binary topological base relations of RCC: disconnected ($\sqsubseteq$), externally connected (EC), equal (EQ), partially overlapping (PO), tangential proper part (TPP), tangential proper part inverse ($\sqsupseteq$TPP), non-tangential proper part ($\sqsupseteq$TPP), and non-tangential proper part inverse ($\sqsubseteq$TPP). These 8 relations are depicted in [Randell et al., 1992, Fig. 4].

The literature has mainly focused on the satisfiability problem [Renz and Nebel, 2001] and the minimal labeling problem (MLP) [Amaneddine et al., 2013; Liu and Li, 2012] of a RCC8 network. The satisfiability problem is deciding if there exists a solution, i.e., a spatial configuration satisfying the constraints of a given network, whilst the MLP is determining all the base relations participating in at least one solution for each of the constraints of that network. Recently, the important problem of deriving redundancy in a RCC8 network was considered and already well established in [Duckham et al., 2014; Li et al., 2015]. For a RCC8 network $\mathcal{N}$ a constraint is redundant, if removing that constraint from $\mathcal{N}$ does not change the solution set of $\mathcal{N}$. A prime network of $\mathcal{N}$ is a network which contains no redundant constraints, but has the same solution set as $\mathcal{N}$. Finding a prime network can be useful in many applications such as computing, storing, and compressing the relationships between spatial objects and hence saving space for storage and communication, facilitating comparison between different networks, merging networks [Condotta et al., 2009], aiding querying in spatially-enhanced databases [Nikolaou and Koubarakis, 2013; OGC, 2012], unveiling the essential network structure of a network (e.g., being a tree or of bounded treewidth [Bodirsky and Wölfel, 2011]), and adjusting geometrical objects to meet topological constraints [Wallgrün, 2012]. Due to space constraints, we refer the reader to [Li et al., 2015] for a well depicted real motivational example and further application possibilities.

In [Li et al., 2015], the complexity results and algorithms obtained rely on the use of $\Diamond$-consistency [Renz and Ligozat, 2005], which enforces consistency on all paths of length 2 in the complete graph of a given network. Hence, the approach offered becomes impractical for networks scaling up to a few tens of thousands of nodes as it is time consuming and most often hits the memory limit (preview Table 1). To solve this practical problem, we are concerned with efficiently characterizing non-redundant constraints in large real world RCC8 networks using partial reasoning. In particular,
we make the following contributions: (i) we provide new complexity results for various cases of RCC8 networks with respect to deriving their non-redundant constraints and obtaining their prime networks, using a particular partial consistency that enforces consistency on all paths of length 2 in a graph $G$, viz., $\xi$-consistency [Chmeiss and Condotta, 2011], while we also show that given a maximal distributive subclass for RCC8 and a network $\mathcal{N}$ defined on that subclass, the pruning capacity of $\xi$-consistency and $\circ$-consistency is identical on the common edges of $G$ and the complete graph of $\mathcal{N}$, when $G$ is a triangulation of the constraint graph of $\mathcal{N}$; (ii) as a byproduct, we show that $\xi$-consistency on a network defined on a maximal distributive subclass of relations equals to $\xi$-consistency [Amaneddine et al., 2013], which can have a positive impact on the algorithm for solving the MLP of a network that was devised in [Amaneddine et al., 2013]; (iii) given a maximal distributive subclass for RCC8 and a triangulation of its constraint graph, we implement an algorithm to compute the unique prime network of very large real world RCC8 networks scaling up to millions of nodes, and show that it goes well beyond the state-of-the-art.

2 Preliminaries

A (binary) qualitative temporal or spatial constraint language [Renz and Ligozat, 2005] is based on a finite set $B$ of jointly exhaustive and pairwise disjoint (JEPD) relations defined on a domain $D$, called the set of base relations. The base relations of set $B$ of a particular qualitative constraint language can be used to represent definite knowledge between any two entities with respect to the given level of granularity. $B$ contains the identity relation $\text{Id}$, and is closed under the converse operation $^{-1}$. Indefinite knowledge can be specified by unions of possible base relations, and is represented by the set containing them. Hence, $2^B$ represents the total set of relations. $2^B$ is equipped with the usual set-theoretic operations union and intersection, the converse operation, and the weak composition operation denoted by symbol $\circ$ [Renz and Ligozat, 2005]. In the case of RCC8 [Randell et al., 1992], as noted in Section 1, the set of base relations $B$ is the set \{DC,EC,PO,TPP,NTPP,TPP_i,NTPP_i,\}, with EQ being relation Id. RCC8 networks can be viewed as qualitative constraint networks (QCNs), defined as follows.

Definition 1 A QCN is a pair $\mathcal{N} = (V, C)$ where: $V$ is a non-empty finite set of variables; $C$ is a mapping that associates a relation $C(v, v') \in 2^B$ to each pair $(v, v')$ of variables $V \times V$. $C$ is such that $C(v, v) = \{\text{Id}\}$ and $C(v, v') = (C(v', v))^{-1}$.

In what follows, given a QCN $\mathcal{N} = (V, C)$ and $v, v' \in V$, $\mathcal{N}[v, v']$ will denote the relation $C(v, v')$. $\mathcal{N}[v, v']|_r$ with $r \in 2^B$ is the QCN $\mathcal{N}'$ defined by $\mathcal{N}'[v, v'] = r$, $\mathcal{N}'[v', v] = r^{-1}$, and $\mathcal{N}'[v, v'] = \mathcal{N}[v, v'] \cap (V \times V) \setminus \{(v', v), (v, v')\}$. Given a set of variables $V$, a graph $G = (V, E)$, and a graph $G' = (V', E')$, $G$ is a subgraph of $G'$, denoted by $G \subseteq G'$, if $V \subseteq V'$ and $E \subseteq E'$. A QCN $\mathcal{N} = (V, C)$ and a graph $G = (V, E)$, $\mathcal{N}$ is $\xi$-consistent [Chmeiss and Condotta, 2011] iff $\forall (v, v'), (v', v) \in E$, $G$ is defined by the empty relation $\emptyset$. A solution of $\mathcal{N}$ is a mapping $\sigma$ defined from $V$ to the domain $D$, yielding a spatial configuration, such that for every pair $(v, v')$ of variables in $V$, $(\sigma(v), \sigma(v'))$ can be described by $\mathcal{N}[v, v']$, i.e., there exists a base relation $b \in \mathcal{N}[v, v']$ such that the relation defined by $(\sigma(v), \sigma(v'))$ is $b$. Two QCNs are equivalent iff they admit the same set of solutions.

Definition 2 A QCN $\mathcal{N}$ is satisfiable if it admits a solution.

A sub-QCN $\mathcal{N}'$ of $\mathcal{N}$ is a QCN $(V, C')$ such that $\mathcal{N}'[v, v'] \subseteq \mathcal{N}[v, v'] \forall v, v' \in V$ where $\mathcal{N}[v, v'] \neq \emptyset$. If $b$ is a base relation, then $\{b\}$ is a singleton relation. An atomic QCN is a QCN where each constraint is a singleton relation. A scenario $S$ of $\mathcal{N}$ is an atomic satisfiable sub-QCN of $\mathcal{N}$. A partial scenario of $\mathcal{N}$ on $V' \subseteq V$ is a scenario restricted to constraints involving only variables of $V'$.

Definition 3 (Li et al., 2015) A QCN $\mathcal{N} = (V, C)$ is weak globally consistent iff, for any $V' \subseteq V$, every partial scenario of $\mathcal{N}$ on $V'$ can be extended to a partial scenario of $\mathcal{N}$ on $V \cup \{v\} \subseteq V$, for any $v \in V \setminus V'$.

A base relation $b \in \mathcal{N}[v, v']$ with $v, v' \in V$ is feasible (resp. un feasible) iff there exists (resp. there does not exist) a scenario $S$ of $\mathcal{N}$ such that $S[v, v'] = \{b\}$.

Definition 4 (Li et al., 2015) A QCN $\mathcal{N} = (V, C)$ is minimal iff $\forall v, v' \in V$ and $\forall b \in \mathcal{N}[v, v']$, $b$ is a feasible base relation of $\mathcal{N}$. The unique equivalent minimal sub-QCN of $\mathcal{N}$, called the minimal QCN of $\mathcal{N}$, is denoted by $\mathcal{N}_{\min}$.

A subclass of relations is a set $\mathcal{A} \subseteq 2^B$ closed under converse, intersection, and weak composition. In what follows, all the considered subclasses will contain the singleton relations of $2^B$.

Definition 5 A subclass $\mathcal{A} \subseteq 2^B$ is a tractable subclass if a QCN $\mathcal{N}$ comprising only relations from $\mathcal{A}$ is tractable. A subclass $\mathcal{A} \subseteq 2^B$ is a maximal tractable subclass if there is no other tractable subclass that properly contains $\mathcal{A}$.

Given three relations $r, r'$, and $r''$, we say that weak composition distributes over intersection if we have that $r \circ (r' \cap r'') = (r \cap r') \circ (r' \cap r'')$ and $(r' \cap r') \cap r = (r' \cap r') \circ (r' \cap r)$.

Definition 6 (Li et al., 2015) A subclass $\mathcal{A} \subseteq 2^B$ is a distributive subclass if weak composition distributes over non-empty intersections for all relations $r, r', r'' \in \mathcal{A}$. A subclass $\mathcal{A} \subseteq 2^B$ is a maximal distributive subclass if there is no other distributive subclass that properly contains $\mathcal{A}$.

A QCN $\mathcal{N}$ is $\circ$-consistent or closed under weak composition iff $\forall v, v', v'' \in V$ we have that $\mathcal{N}[v, v'] \subseteq \mathcal{N}[v, v''] \circ \mathcal{N}[v'', v']$. Given a QCN $\mathcal{N} = (V, C)$, $\circ$-consistency can be determined in $O(|V|^3)$ time. The $\circ$-consistent QCN of $\mathcal{N}$ is denoted by $\circ(\mathcal{N})$, and it is equivalent to $\mathcal{N}$. In what follows, all considered graphs are undirected. Given two graphs $G = (V, E)$ and $G' = (V', E')$, $G$ is a subgraph of $G'$, denoted by $G \subseteq G'$, iff $V \subseteq V'$ and $E \subseteq E'$. Given a QCN $\mathcal{N} = (V, C)$ and a graph $G = (V, E)$, $\mathcal{N}$ is $\xi$-consistent [Chmeiss and Condotta, 2011] iff for $\forall (v, v'), (v', v) \in E$, we have that $\mathcal{N}[v, v'] \subseteq \mathcal{N}[v, v'] \circ \mathcal{N}[v', v]$. Given a QCN $\mathcal{N} = (V, C)$ and a graph $G = (V, E)$, $\xi$-consistency can be determined in $O(|\Delta| \cdot |E|)$ time, where $\Delta$ is the maximum vertex degree of $G$. The $\xi$-consistent QCN of $\mathcal{N}$ is denoted by $\xi(\mathcal{N})$. A graph $G = (V, E)$ is a chordal graph (or triangulated graph)
We denote by \( \text{RCC8} \) contained in \( \text{RCC8} \) (\( \text{RCC8} \) in a \( \text{QCN} \)). Given a \( \text{RCC8} \) in a \( \text{QCN} \), then minimal distributive subclasses \( \text{RCC8} \) trivially inconsistent \( \text{QCN} \) and \( \text{RCC8} \) is weak globally consistent and minimal. Proposition 1 ([Sioutis and Koubarakis, 2012]) \( \text{RCC8} \) implication of chordal graphs in the satisfiability problem in \( \text{NP} \). \( \text{NP} \) -hard in general (cf. [Liu and Li, 2012]). However, for the two maximal distributive subclasses \( \text{D}^4_8 \) and \( \text{D}^6_8 \) for \( \text{RCC8} \), the minimal labeling problem becomes tractable [Li et al., 2015]. In particular we have the following result: Proposition 2 ([Li et al., 2015]) Let \( \mathcal{N} = (V, C) \) be a not trivially inconsistent \( \text{RCC8} \) defined on one of the maximal distributive subclasses \( \text{D}^4_8 \), or \( \text{D}^6_8 \). If \( \mathcal{N} \) is \( \omega \)-consistent, then \( \mathcal{N} \) is weak globally consistent and minimal.

Finally, let \( \mathcal{N} = (V, C) \) and \( \mathcal{N}' = (V', C') \) be two \( \text{QCNs} \) such that \( \mathcal{N}[v, v'] = \mathcal{N}'[v, v'] \) for all \( v, v' \in V \cap V' \). We denote by \( \mathcal{N}'' = \mathcal{N} \cup \mathcal{N}' \) the \( \text{QCN} \) \( \mathcal{N}'' = (V'', C'') \), where \( V'' = V \cup V' \), \( \mathcal{N}''[v, v'] = \mathcal{N}[v, v'] \) for all \( v, v' \in V \setminus V' \), \( \mathcal{N}''[v, v'] = \mathcal{N}'[v, v'] \) for all \( v, v' \in V' \setminus V \), and \( \mathcal{N}''[v, v'] = \mathcal{N}'[v, v'] \) for all \( v, v' \in V \setminus V' \).

3 Prime Network in \( \text{RCC8} \)

In this section we give the definitions of a redundant relation in a \( \text{QCN} \) \( \mathcal{N} \) of \( \text{RCC8} \) and a prime \( \text{QCN} \) of \( \mathcal{N} \), and provide new complexity results and algorithms for finding a prime \( \text{QCN} \) of \( \mathcal{N} \) when \( \mathcal{N} \) is defined on a non-tractable, a maximal tractable, or a maximal distributive subclass for \( \text{RCC8} \).

Given a \( \text{QCN} \mathcal{N} = (V, C) \), we say that \( \mathcal{N} \) entails a relation \( r(v, v') \in \mathcal{B} \), with \( v, v' \in V \), if for every solution \( \sigma \) of \( \mathcal{N} \), the relation defined by \( (\sigma(v), \sigma(v')) \) is a base relation \( b \) such that \( b \in r(v, v') \). A relation \( \mathcal{N}[v, v'] \) in \( \mathcal{N} \) is redundant if network \( \mathcal{N}[v, v'] \cup b \) entails \( \mathcal{N}[v, v'] \). Note that by definition every universal relation \( B \) in a \( \text{QCN} \) is redundant. Recalling the fact that the constraint graph of a \( \text{QCN} \) involves all the non-universal relations, we can obtain the following lemma:

Lemma 1 Given a \( \text{QCN} \mathcal{N} = (V, C) \) and its constraint graph \( G(N) = (V, E) \), a relation \( \mathcal{N}[v, v'] \) with \( v, v' \in V \) is redundant if \( (v, v') \notin E \).

We recall the definition of a reducible and a prime \( \text{QCN} \).

Definition 7 ([Li et al., 2015]) A \( \text{QCN} \mathcal{N} = (V, C) \) is reducible if it comprises a redundant relation other than relation \( B \), and irreducible otherwise. An equivalent irreducible sub-\( \text{QCN} \) of \( \mathcal{N} \), is called a prime \( \text{QCN} \) of \( \mathcal{N} \). If a prime \( \text{QCN} \) of \( \mathcal{N} \) is also unique, it is denoted by \( \mathcal{N}_{\text{prime}} \).

In Figure 1 a \( \text{QCN} \) \( \mathcal{N} \) of \( \text{RCC8} \) and its prime \( \text{QCN} \) are depicted. Relation \( \{DC\} \) is irreducible as it can be entailed by \( \mathcal{N} \) \{v1, v2\}/B and, thus, can be replaced with relation \( B \) (that denotes the lack of the contradiction between two entities in a \( \text{QCN} \)).

3.1 \( \text{RCC8} \) networks in the general case

Given an arbitrary \( \text{QCN} \mathcal{N} \) of \( \text{RCC8} \), finding a prime \( \text{QCN} \) of \( \mathcal{N} \) is clearly at least as hard as determining if \( \mathcal{N} \) is reducible. We then have the following result, which relies on the fact that retrieving a solution for a \( \text{QCN} \) of \( \text{RCC8} \) is \( \text{NP} \)-hard in general [Renz and Nebel, 1999]:

Proposition 3 ([Li et al., 2015]) Let \( \mathcal{N} = (V, C) \) be a \( \text{QCN} \) of \( \text{RCC8} \). It is co-\( \text{NP} \)-complete to decide if \( \mathcal{N}[v, v'] \), with \( v, v' \in V \), is redundant in \( \mathcal{N} \).

However, we can have a stronger theoretical result than that for \( \text{QCNs} \) of \( \text{RCC8} \) of bounded treewidth. The width of a tree decomposition \( (T, \{X_1, \ldots, X_n\}) \) [Diestel, 2012] is \( \max_{1 \leq i \leq n} |X_i| - 1 \). The treewidth of a graph \( G \) is the minimum width possible for arbitrary tree decompositions of \( G \).

Theorem 1 ([Bodirsky and Wölf, 2011]) For any \( k \), the satisfiability problem for a \( \text{QCN} \) of \( \text{RCC8} \) of treewidth at most \( k \) can be solved in polynomial time.

A detailed algorithm for Theorem 1 that builds on the proof sketch of [Bodirsky and Wölf, 2011] is provided in [Huang et al., 2013]. In particular, the satisfiability check can be made in \( O(w^5|V|^{|V|\log|V|}) \) time for a \( \text{QCN} \mathcal{N} = (V, C) \) of \( \text{RCC8} \), where \( w \) is the treewidth of \( G(\mathcal{N}) \) and \( c \) a constant such that \( |B||V|^{|V|\cdot 1/2} \leq c|V|^2 \), with the algorithm provided in [Huang et al., 2013]. We then have the following result:

Proposition 4 Let \( \mathcal{N} = (V, C) \) be a \( \text{QCN} \) of \( \text{RCC8} \) of bounded treewidth. It takes polynomial time to decide if \( \mathcal{N}[v, v'] \), with \( v, v' \in V \), is non-redundant in \( \mathcal{N} \).

Proof. Let \( r = B \setminus \mathcal{N}[v, v'] \), viz., the set of base relations not included in \( \mathcal{N}[v, v'] \). To check if \( \mathcal{N}[v, v'] \) is non-redundant in \( \mathcal{N} \), we must check if \( \mathcal{N}[v, v'] \cup r \) is satisfiable. This satisfiability check can be made in polynomial time due to Theorem 1.

The obtained bound for the algorithm for Theorem 1 does not allow for practical applicability of Proposition 4, for which we are interested in this paper, but, nevertheless, is of theoretical merit. To check if a \( \text{QCN} \mathcal{N} \) of \( \text{RCC8} \) is reducible, we need to check \( O(E(G(\mathcal{N}))) \) relations as suggested by Lemma 1.

Thus, due to Proposition 4 we can obtain a prime network in polynomial time with a simple procedure similar to the one provided in [Li et al., 2015, p. 60].
3.2 RCC8 networks defined on a maximal tractable subclass

Given a QCN $\mathcal{N}$ of RCC8 that is defined on a maximal tractable subclass for RCC8, finding a prime QCN of $\mathcal{N}$ takes polynomial time. In particular, we have the following result:

**Proposition 5** Let $\mathcal{N} = (V, C)$ be a QCN of RCC8 defined on one of the maximal tractable subclasses $\mathcal{H}_8$, $\mathcal{C}_8$, or $\mathcal{Q}_8$, $G(\mathcal{N}) = (V, E')$ its constraint graph, and $G = (V, E)$ a chordal graph such that $G(\mathcal{N}) \subseteq G$. We can determine if a relation is non-redundant in $\mathcal{N}$ in $O(\delta \cdot |E|)$ time, where $\delta$ denotes the maximum vertex degree of $G$, and in $O(|E|)$ time find all non-redundant relations in $\mathcal{N}$. In addition, a prime QCN of $\mathcal{N}$ can be found in $O(\delta \cdot |E|)$ time.

**Proof.** Let $r = B \setminus N[v, v']$, where $N[v, v']$ with $(v, v') \in E'$ is a non-universal relation in $\mathcal{N}$. To check if relation $N[v, v']$ is non-redundant in $\mathcal{N}$, we must check if there exists a $b \in r$ such that $N[v, v'] \setminus \{b\}$ is satisfiable. If such $b$ exists, $N[v, v']$ is non-redundant in $\mathcal{N}$. Due to Proposition 1, satisfiability checking can be done in $O(\delta \cdot |E|)$ time with the $\mathcal{G}$-consistency operation, and as we can have a check for at most a constant number of $|B|$ base relations, the complexity remains $O(\delta \cdot |E|)$. By Lemma 1, we need only perform the non-redundancy check for a total of $O(|E'|)$ relations. Thus, finding all non-redundant relations takes $O(\delta \cdot |E| \cdot |E'|)$ time, and, as a consequence, we can obtain a prime network in $O(\delta \cdot |E| \cdot |E'|)$ time using a simple procedure like the one provided in [Li et al., 2015, p. 60].

As noted in [Li et al., 2015], calculating $N_{\text{prime}}$ for a QCN $\mathcal{N}$ of RCC8 defined on a maximal tractable subclass is not always possible, as $\mathcal{N}$ may have more than one prime network. To obtain $N_{\text{prime}}$ we need $\mathcal{N}$ to be defined on a maximal distributive subclass, which we will study in the sequence.

3.3 RCC8 networks defined on a maximal distributive subclass

Given a QCN $\mathcal{N}$ of RCC8 that is defined on a maximal distributive subclass for RCC8, we will show that we can find $N_{\text{prime}}$ with an algorithm that goes well beyond the state-of-the-art algorithm presented in [Li et al., 2015]. To be able to do so, we will first present a result that is similar to a main result of [Blik and Sam-Haroud, 1999], namely, we will show that for a maximal distributive subclass for RCC8 and a QCN $\mathcal{N}$ defined on that subclass, the pruning capacity of $\mathcal{G}$-consistency and $\circ$-consistency is identical on the common edges of a chordal graph $G$ with $G(\mathcal{N}) \subseteq G$ and the complete graph of $\mathcal{N}$. As a consequence, we will also briefly explain the implication of this result in the minimal labelling problem (MLP) of a QCN of RCC8 [Amaneddine et al., 2013].

First, we recall the following lemma that will allow us to make a chordal graph complete by adding a single edge at a time and keeping all intermediate graphs chordal:

**Lemma 2 ([Blik and Sam-Haroud, 1999])** If $G = (V, E)$ is a non-complete chordal graph, then one can add a missing edge $(u, v)$ with $u, v \in V$ such that the graph $G' = (V, E \cup \{(u, v)\})$ is chordal and the graph induced by $X = \{x \mid (u, x), (x, v) \in E\}$ is complete.

We now proceed with proving the following result:

**Proposition 6** Let $\mathcal{N} = (V, C)$ be a not trivially inconsistent QCN of RCC8 defined on one of the maximal distributive subclasses $\mathcal{D}_8^1$, or $\mathcal{D}_8^4$, and $G = (V, E)$ a chordal graph such that $G(\mathcal{N}) \subseteq G$. Then, $\forall (v, v') \in E$ we have that $\circ(\mathcal{N})[v, v'] = \mathcal{G}(\mathcal{N})[v, v']$.

**Proof.** Suppose that $\mathcal{N}$ is $\mathcal{G}$-consistent. We consider only the case where $\mathcal{G}$-consistency did not result in the assignment of the empty relation $\emptyset$ for some edge in $G$, i.e., the case where $\mathcal{N}$ is satisfiable, as otherwise we would have that $\circ(\mathcal{N}) = \perp_{V, G}$, $\mathcal{G}(\mathcal{N}) = \perp_{V, \mathcal{G}}$, and as a consequence $\forall (v, v') \in E$ trivially have that $\circ(\mathcal{N})[u, v'] = \mathcal{G}(\mathcal{N})[v, v'] = \emptyset$, as desired. We will add to graph $G$ the missing edges one by one until it becomes complete. We will show that the relations that correspond to the new edges can be computed from existing relations, so that for each intermediate chordal graph $G'$ network $\mathcal{N}$ is $\mathcal{G}$-consistent, and for the final complete graph $K_V$ network $\mathcal{N}$ is therefore $\circ$-consistent, that is, $\circ$-consistent. Every edge is added according to Lemma 2 to retain a chordal graph at all times. Consider graph $G$ as shown below.

![Graph Diagram]

After adding edge $e = (v_{n-1}, v_j)$ to $G$ we obtain graph $G' = (V, E \cup \{e\})$. (As $e \notin E$, initially $N[v_{n-1}, v_j] = B$.) Let $X = \{x \mid (v_{n-1}, x), (x, v_j) \in E\}$ be the set of variables that induce a complete graph and are denoted in grey colour in our figure. We will show that the relation corresponding to edge $e$ can be computed as follows:

$$N[v_{n-1}, v_j] = \bigcap_{x \in X} N[v_{n-1}, x] \circ N[x, v_j] \quad (1)$$

Note that by construction graphs $A$ induced by $\{v_{n-1}\} \cup X$ and $B$ induced by $X \cup \{v_j\}$ are complete. Let us denote by $\mathcal{N}_A$ and $\mathcal{N}_B$ the QCNs of RCC8 that correspond to graphs $A$ and $B$ respectively. We need to show that network $\mathcal{N}$ is $\circ$-consistent. In particular, we need to show that every path $\pi$ of length 2 that goes through $v_{n-1}$ and $v_j$ is consistent, as every other path of length 2 is by assumption consistent. We need to consider two cases (reasoning is the same for symmetrical cases). With $x' \in X$, either $\pi = \langle v_{n-1}, v_j, x' \rangle$, or $\pi = \langle v_{n-1}, v_j, x' \rangle$. If $\pi = \langle v_{n-1}, v_j, x' \rangle$, we prove that for every base relation of $\mathcal{N}[v_{n-1}, x']$, there exist base relations for both $\mathcal{N}[v_{n-1}, v_j]$ as defined in (1) and $\mathcal{N}[x', v_j]$, so that the path is consistent (i.e., $\mathcal{N}[v_{n-1}, x'] \subseteq \mathcal{N}[v_{n-1}, v_j] \circ \mathcal{N}[x', v_j]$). As graph $A$ is complete, $\mathcal{N}_A$ is $\circ$-consistent, and therefore due to Proposition 2, $\mathcal{N}_A$ is weak globally consistent and minimal. Thus, for every base relation of $\mathcal{N}[v_{n-1}, x']$ there exist compatible base relations for all other constraints in $\mathcal{N}_A$, i.e., there exists a scenario for $\mathcal{N}_A$, denoted by $\mathcal{N}_A^S$. As $V(A) \cap V(B) = X$ induces a complete graph, the assignment of base relations for every constraint in $\mathcal{N}_A$ will define a partial scenario for $\mathcal{N}_B$ on $X$. Again, as graph $B$ is complete,
$N_D$ is $\varnothing$-consistent, and therefore due to Proposition 2, $N_D$ is weak globally consistent and minimal, thus, this partial scenario can be extended to a scenario for $N_D$, denoted by $N_D^3$. As such, $N[v_1,x']$ will be also characterized by a base relation. Finally, relations $N[v_1,v_2,x']$ and $N[x,v_1]$ for all $x \in X$ participating in (1) will be characterized by base relations, thus, there exists a base relation for $N[v_1,v_2]$ too as clearly $N_A^3 \cup N_B^3$ is $\varnothing$-consistent and, thus, satisfiable due to Proposition 1. If $\pi = (v_1,v_2,x')$, by equation (1) we know that for every base relation of $N[v_1,v_2]$ we can find base relations for $N[v_1,x']$ and $N[x,v_1]$, so that $N[v_1,v_2] \subseteq N[v_1,x'] \cup N[x,v_1]$ can hold.

Given a not trivially inconsistent QCN $N = (V,C)$ of RCC8 defined on one of the maximal distributive subclasses $D_8^{11}$, or $D_8^{34}$, and $G = (V,E)$ a chordal graph such that $G(N) \subseteq G$, due to Proposition 2 and Proposition 6 we have that $\forall (v,v') \in E$ every base relation $b \in \mathcal{P}(N)[v,v']$ is feasible. As a consequence we have the following result:

**Proposition 7** Let $N = (V,C)$ be a not trivially inconsistent QCN of RCC8 defined on one of the maximal distributive subclasses $D_8^{11}$, or $D_8^{34}$, and $G = (V,E)$ a chordal graph such that $G(N) \subseteq G$. Then, $N$ is $\varnothing$-consistent [Amaneddine et al., 2013] if it is $\varnothing$-consistent.

In [Amaneddine et al., 2013], given a QCN $N$ of RCC8 and a chordal graph $G = (V,E)$ such that $G(N) \subseteq G$, $\varnothing$-consistency is a partial consistency that is applied to characterize the feasible base relations $\forall N[u,v]$ with $(u,v) \in E$. It is shown that if $N$ is defined on a maximal tractable subclass $H_8$, $C_8$, or $Q_8$, computing $\varnothing$-consistency takes $O(\delta \cdot |E|^2)$ time, where $\delta$ is the maximum vertex degree of $G$. Proposition 7 suggests that if we restrict $N$ to a maximal distributive subclass $D_8^{11}$, or $D_8^{34}$, $\varnothing$-consistency can be computed in $O(\delta \cdot |E|)$ time. As $\varnothing$-consistency is used in an algorithm for solving the MLP in [Amaneddine et al., 2013], this result can have a positive impact on its performance, which is out of the scope of this paper to explore.

Let us now introduce a property that will hold for all the considered QCNs in what follows.

**Property 1** ([Li et al., 2015]) Let $N = (V,C)$ be a satisfiable QCN of RCC8. Then, $N$ will be said to satisfy the uniqueness property iff $\forall u,v \in V$, with $u \neq v$, we have that $N$ does not entail a relation $r \subseteq N[u,v]$ where $r = \{ EQ \}$.

The uniqueness property specifies that every relation in a QCN of RCC8 should be unique and not identical to any other relation. This property establishes a very weak condition in the sense that we can eliminate all entailed singleton \{EQ\} relations from a given QCN of RCC8 and retain its knowledge by reducing the spatial variables that are identical to each other into a single variable, maintaining all involved relations. This is a necessary condition to be able to obtain the unique prime network of a QCN [Li et al., 2015]. Before obtaining another result in this section, we recall the following lemma:

**Lemma 3** ([Li et al., 2015]) Let $N = (V,C)$ be a not trivially inconsistent QCN of RCC8 defined on one of the maximal distributive subclasses $D_8^{11}$, or $D_8^{34}$, and having the uniqueness property. If $N$ is $\varnothing$-consistent, then a relation $N[v,v']$, with $v, v' \in V$, is non-redundant in $N$ iff we have that $N[v,v'] \neq \bigcap \{ N[v,v'] \cup N[\nu,v'] \mid \nu \in V \setminus \{ v, v' \} \}$.

We will now prove a result that allows us to obtain the same set of non-redundant relations as the one provided by Lemma 3, more efficiently, through the use of chordal graphs.

**Proposition 8** Let $N = (V,C)$ be a not trivially inconsistent QCN of RCC8 defined on one of the maximal distributive subclasses $D_8^{11}$, or $D_8^{34}$, and having the uniqueness property, and $G = (V,E)$ a graph such that $G(N) \subseteq G$. If $G$ is chordal and $N$ is $\varnothing$-consistent, then a relation $N[v,v']$ is non-redundant in $G(N)$ (i.e., the $\varnothing$-consistent network $N$ as in Lemma 3), iff we have that $(v,v') \in E(G(N))$ and $N[v,v'] \neq \bigcap \{ N[v,v'] \cup N[\nu,v'] \mid \nu \in V \} \in E$.

**Proof.** By Lemma 1 and the fact that for a QCN $M$ of RCC8, $\varnothing(M)$ and $M$ are equivalent, it trivially follows that if $(v,v') \notin E(G(N))$ then $N[v,v']$ is redundant in $G(N)$. We will show that we can have the same set of non-redundant relations as the one suggested in Lemma 3. Let $N[v,v']$ with $(v,v') \in E(G(N))$ be a relation in $G(N)$, then by Proposition 6 we have that $N[v,v'] = (G(N), v,v')$ be the chordal graph that results by adding edge $e = \langle u,v \rangle$ to $G$ according to Lemma 2. It suffices to show that the next equation holds:

$$\bigcap \{ N[\nu,v'] \cup N[v,v'] \mid (v,v') \in E \} = \bigcap \{ N[v,v'] \cup N[\nu,v'] \mid (v,v') \notin E \}$$

(2) Equation (2) states that $N[v,v']$ is strictly contained in the intersection of all paths of length 2 that start with $v$ and end with $v'$ in $G$ iff $N[v,v']$ is strictly contained in the intersection of all paths of length 2 that start with $v$ and end with $v'$ in $G$. If (2) holds, it follows that we can add all edges necessary to obtain the paths of length 2 that start with $v$ and end with $v'$ and go through every $v'' \in V \setminus \{ v, v' \}$, and have that $\bigcap \{ N[\nu,v'] \cup N[v,v'] \mid (v,v') \in E \} = \bigcap \{ N[v,v'] \cup N[\nu,v'] \mid v'' \in V \setminus \{ v, v' \} \}$, so that the necessary and sufficient conditions specified in Lemma 3 and Proposition 8 are equivalent with respect to deciding the redundancy of $\varnothing(N)[v,v']$. We will assume that (2) does not hold and result in a contradiction. In particular, we will assume that $b$ is a base relation that is in the intersection of all paths of length 2 that start with $v$ and end with $v'$ in $G$, but not in the intersection of all paths of length 2 that start with $v$ and end with $v'$ in $G$. Clearly, this is only possible when $b \not\subseteq N[u,v] \cup N[u,x] \cup N[x,v']$, where new relation $N[u,v']$ according to (1) in the proof of Proposition 6, i.e., $N[u,v']$ is the intersection of all paths of length 2 that start with $u$ and end with $v'$ in $G$. Then, there exists a $x \in \{ u,v \}$ such that $b \not\subseteq N[v,x] \cup N[x,v']$. However, both $(v,x)$ and $(x,v')$ are in $G$ and we already know that $N[v,v'] \not\subseteq N[v,v']$ for all $v'' \in V$ such that $(v,v'') \notin E$, thus, $b \in N[v,x] \cup N[x,v']$. This is a contradiction, thus, (2) holds.

The set of non-redundant relations as provided by Lemma 3 and Proposition 8 is related to the unique prime network of a QCN $N$ of RCC8, viz., $N_{\varnothing}$, as follows:

**Theorem 2** ([Li et al., 2015]) Let $N = (V,C)$ be a satisfiable QCN of RCC8 defined on one of the maximal distributive
Algorithm 1: Delphysγ(\(\mathcal{N}\))

\[
\begin{aligned}
&\text{in : A satisfiable QCN } \mathcal{N} \text{ defined on } D_{64}^4 \text{ or } D_{8}^4, \\
&\text{output : } \chi, \text{ the set of non-redundant relations in } \circ(\mathcal{N}). \\
&\begin{align*}
\text{begin} & \\
&\chi \leftarrow \emptyset; \\
&G \leftarrow \text{Triangulation}(G(\mathcal{N})); \\
&\mathcal{N}' \leftarrow \circ(\mathcal{N}); \\
&Q \leftarrow \{(v, v') \mid (v, v') \in E(G(\mathcal{N}))\}; \\
&\text{while } Q \neq \emptyset \text{ do} \\
&\quad (v, v') \leftarrow Q.pop(); \\
&\quad \tau \leftarrow \emptyset; \\
&\quad \text{foreach } v'' \text{ such that } (v, v''), (v'', v') \in E(G) \text{ do} \\
&\quad \quad t \leftarrow \mathcal{N}'[v, v''] \circ \mathcal{N}'[v'', v']; \\
&\quad \quad \text{foreach } b \in B \text{ do} \\
&\quad \quad \quad \text{if } b \not\in \tau \text{ then } \tau \leftarrow \tau \cup \{b\}; \\
&\quad \quad \text{if } \tau \cup \mathcal{N}'[v, v'] \neq B \text{ then } \chi \leftarrow \chi \cup \{\mathcal{N}'[v, v']\}; \\
&\quad ; \\
&\text{end} \\
&\text{return } \chi; \\
\end{align*}
\end{aligned}
\]

<table>
<thead>
<tr>
<th>network</th>
<th>Delphys</th>
<th>Delphysγ</th>
<th>speedup (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>nuts</td>
<td>0.12s</td>
<td>0.11s</td>
<td>9.9%</td>
</tr>
<tr>
<td>adm1</td>
<td>25.536.93s</td>
<td>222.33s</td>
<td>99.1%</td>
</tr>
<tr>
<td>gadm1</td>
<td>140.685.26s</td>
<td>329.43s</td>
<td>99.8%</td>
</tr>
<tr>
<td>gadm2</td>
<td>9.18s</td>
<td>2.34s</td>
<td>74.5%</td>
</tr>
<tr>
<td>adm2</td>
<td>∞</td>
<td>1 060.51s</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1: Performance comparison on CPU time

subclasses \(D_{64}^4\), or \(D_{8}^4\), and having the uniqueness property. Further, let \(\chi\) be the set of non-redundant relations in \(\circ(\mathcal{N})\). Then, \(\forall u, v \in V\) we have that \(\mathcal{N}'\text{prime}[u, v] = \mathcal{N}[u, v]\) if \(\circ(\mathcal{N})[u, v] \in \chi\) and \(\mathcal{N}'\text{prime}[u, v] = B\) otherwise.

An algorithm based on Lemma 3 to obtain the set of non-redundant relations was provided in [Li et al., 2015] with a time complexity of \(O(|V|^3)\) for a given QCN \(\mathcal{N} = (V, C)\) of RCC8, which we here call Delphys. Due to Proposition 8, we can have a new algorithm with a time complexity of \(O(\delta \cdot |E|)\) for a chordal graph \(G = (V, E)\) such that \(G(\mathcal{N}) \subseteq G\), where \(\delta\) denotes its maximum vertex degree, as the triangulation procedure in line 3 is linear in the size of \(G\) [Parter, 1961; Diestel, 2012] and \(\delta\)-consistency in line 4 dominates the overall complexity. Our new algorithm is presented in Algorithm 1, under the name Delphysγ.

4 Experimental Evaluation

In this section, we compare the performance of Delphysγ with that of Delphys [Li et al., 2015] using a real dataset. The experiments were carried out on a computer with an Intel Core 2 Quad Q9400 processor with a CPU frequency of 2.66 GHz per core, 8 GB RAM, and the Precise Pangolin x86_64 OS. Both Delphysγ and Delphys were written in pure Python and run with Red Hat Python 2.4.0 (http://pypy.org). Only one of the CPU cores was used.

We consider the dataset of real network instances that was originally introduced in [Nikolaou and Koubarakis, 2014]:

- **nuts**: a nomenclature of territorial units using RCC8 relations that contains 2 235/3 176 nodes/edges.\(^1\)
- **adm1**: a network that describes the administrative geography of Great Britain using RCC8 relations [Goodwin et al., 2008] and contains 11 761/44 832 nodes/edges.
- **gadm1**: a network that describes the German administrative units using RCC8 relations and contains 42 749/159 600 nodes/edges.
- **gadm2**: a network that describes the world’s administrative areas using RCC8 relations and contains 276 727/589 573 nodes/edges (http://gadm.geocov.org/).
- **adm2**: a network that describes the Greek administrative geography using RCC8 relations and contains 1 732 999/5 236 270 nodes/edges.\(^2\)

The aforementioned network instances are satisfiable. They comprise relations that are properly contained in any of the two maximal distributive subclasses \(D_{64}^4\) and \(D_{8}^4\) for RCC8. (Also, some identical regions were properly amalgamated to satisfy the uniqueness property.)

The results on the performance of Delphysγ and Delphys are shown in Table 1.\(^3\) Note that symbol ∞ signifies that a reasoner hits the memory limit. The speedup for Delphysγ reaches as high as nearly 100% for the cases where Delphys was actually able to reason with the networks (e.g., gadm1).

Table 2 shows the decrease that we can achieve with respect to the total number of non-redundant relations that we can obtain from an initial network, which allows one to construct sparse constraint graphs that boost various graph related tasks such as storing, querying, representing, and reasoning. Note that the constraint graphs of the initial networks are sparse, thus, a lot of redundancy is already avoided. Still, for the biggest network of the dataset, namely, adm2, the decrease is around 66%, which is almost linear to the number of its vertices, confirming a similar observation in [Li et al., 2015].

5 Conclusion

We made use of \(\delta\)-consistency [Chmeiss and Condotta, 2011] to obtain new complexity results for various cases of RCC8 networks with respect to deriving their non-redundant constraints and obtaining their prime networks, while we also

---

\(^1\)Retrieved from: http://www.linkedopendata.gr/

\(^2\)As Delphys relies on \(\circ\)-consistency and consequently uses an adjacency matrix to store the complete graph of a given network that forbids it from being able to consider networks of more than a few tens of thousands of nodes, its operation was appropriately aided with the simple decomposition scheme of [Sioutis et al., 2015b], as described in [Sioutis et al., 2015a]. The same decomposition scheme was applied to Delphysγ too, to maintain a fair comparison. However, Delphysγ did not really benefit from that scheme, as it uses an adjacency list to store a chordal graph of a given network and it was already very time/memory efficient for the involved instances.
showed that given a maximal distributive subclass for RCC8 and a network $\mathcal{N}$ defined on that subclass, the pruning capacity of $G$-consistency and $\diamond$-consistency is identical on the common edges of $G$ and the complete graph of $\mathcal{N}$, when $G$ is a triangulation of $\mathcal{G}(\mathcal{N})$. As a byproduct, we showed that $G$-consistency on a network defined on a maximal distributive subclass of relations equals to $\diamond$-consistency [Amaneddine et al., 2013]. Given a maximal distributive subclass for RCC8, we devised an algorithm based on $G$-consistency to compute the unique prime network of a RCC8 network, and showed that it significantly progresses the state-of-the-art for practical reasoning with very large real RCC8 networks.

A Generalization to other Calculi

As RCC5 is a sublanguage of RCC8 where no significance is attached to boundaries of regions [Randell et al., 1992], all results in this paper immediately carry over to RCC5 with respect to its own particularities, such as its maximal tractable and maximal distributive subclasses.

All results, except that of Theorem 2 which uses some particular algebraic properties of RCC5/8, can be applied to several other qualitative spatial constraint calculi, such as the Cardinal Direction Calculus [Frank, 1991; Ligozat, 1998], the Block Algebra [Guesgen, 1989], and even the Interval Algebra [Allen, 1983] when viewed as a spatial calculus, again, with respect to their own particularities. In fact, it can be shown that the aforementioned results can be applied to any qualitative spatial constraint calculus that is a relation algebra and has patchwork for networks defined on one of its maximal tractable subclasses of relations [Huang, 2012]. This is due to the fact that the sound use of $G$-consistency, when $G$ is a triangulation of the constraint graph of a given network, implicitly considers patchwork for the involved qualitative spatial constraint calculus [Amaneddine et al., 2013, Property 1].

References


