Paper VII

Qualitative Reasoning About Sets Applied to Spatial Reasoning

Peter Jonsson and Thomas Drakengren

ABSTRACT

We propose a polynomial-time algorithm for a restricted form of qualitative reasoning about sets. The algorithm decides satisfiability of set expressions of the type XrY where r is one of the three relations of subset, disjointness and non-equality. The algorithm also handles certain disjunctions of such expressions. With the aid of this algorithm, we construct a straightforward, tractable algorithm for the Horn fragment of the spatial theory RCC-5. The algorithm also enables us to show certain topological properties of RCC-5. For instance, we show that a result by Grigni *et al* about RCC-5 restricted to two-dimensional space does not hold for RCC-5 restricted to three-dimensional space.

1 Introduction

Reasoning about sets has proved itself useful in a number of areas in computer science. The typical AI examples are the so-called "terminological languages" [Brachman and Schmolze, 1985]. Outside the AI community, reasoning about sets has been employed in, for example, analyses of programs, logic programming and the study of finite automata. The main result of this paper is a tractable algorithm for deciding satisfiability of a class of set expressions. More specifically, we relate set variables with subset, disjointness and non-equality and we allow certain disjunctions of such set relations.

Even though the problem that we study appears to be severely restricted, it seems to have practical relevance. To demonstrate this, we show how it can be used in qualitative spatial reasoning. This kind of spatial reasoning has received much attention in the literature recently. The main reason for this is, probably, that spatial reasoning has proved to be applicable to real-world problems in, for example, geographical databases [Egenhofer, 1991; Grigni *et al.*, 1995] and molecular biology [Cui, 1994]. In both these applications, the size of the problem instances can be huge, so it is highly relevant to study the complexity of problems and algorithms.

A widespread framework for qualitative spatial reasoning is the RCC approach [Randell and Cohn, 1989; Randell *et al.*, 1992]. This approach is based on modelling qualitative spatial relations between regions using first-order logic. Of special interest is a subtheory of RCC, RCC-5. Bennett [1994] has shown that this subtheory has a much lower complexity figure than the full RCC framework but, unfortunately, the satisfiability problem for RCC-5 is still NP-complete [Renz, 1996].

This motivates the search for tractable subclasses of RCC-5. Renz [1996] has shown that there exists a large, maximal tractable subclass of RCC-5, denoted $\widehat{\mathcal{H}}_5$. $\widehat{\mathcal{H}}_5$ contains 28 elements out of 32 (87%). Renz's proof of the tractability of $\widehat{\mathcal{H}}_5$ relies on a complicated reduction in two steps: the problem is first reduced to a fragment of the modal logic S5 and, second, this modified problem is reduced to propositional Horn logic. We provide an alternative method for $\widehat{\mathcal{H}}_5$. We employ a simple reduction from the satisfiability problem for $\widehat{\mathcal{H}}_5$ to our tractable class of set expressions. This provides two advantages. Firstly, the resulting algorithm is very straightforward and does not rely on a translation of the problem to logic. Secondly, we can use the algorithm for showing topological properties of RCC-5. For example, we complement the findings in Grigni *et al.* [1995] on spatial reasoning with strong topological assumptions. They have shown that the satisfiability problem of a certain subset of RCC-5 is NP-complete if we require the sets to be connected sets in two-dimensional Euclidean space. However, we show that an extended version of this problem is tractable if the allowed sets are connected sets in *three*-dimensional Euclidean space.

In Sections 2 and 3 we introduce the algorithm for reasoning about sets and show that it is tractable. RCC-5 is introduced in Section 4 together with the new tractability proof of $\widehat{\mathcal{H}}_5$. The topological results are collected in Section 5. Finally, the last two sections discuss the results and conclude the paper.

2 Disjunctive Set Relations

Our approach to reasoning about sets is based on *set variables* and *relations* on these. We assume that we have a fixed universe of variable names for sets. Then, an S-*interpretation* is a function that maps set variables to non-empty sets. The reason for restricting ourselves to only consider non-empty sets is that we intend to use the algorithms for spatial reasoning. Hence, we want to disallow the pathological case of having empty objects. It may, however, be necessary to allow empty sets in other applications but this is left for future research.

We will relate set variables by the following three relations: subset ($\subseteq$), non-equality ($\neq$) and disjointness (disj). The relations $\subseteq$ and $\neq$ have their usual definitions and $X \mathsf{disj} Y$ holds iff $X \cap Y = \emptyset$.

Two set variables related by one of these three relations is called an *atomic relation*. A *disjunctive set relation* (DSR) is a set of atomic relations.

An atomic relation XRY is satisfied by an S-interpretation I iff the relation $I(X)RI(Y)$ holds. A finite set $\Gamma = \{\gamma_1, \ldots, \gamma_n\}$ of DSRs is said to be S-satisfiable iff there exists an S-interpretation that satisfies at least one member of each γ_i, $1 \leq i \leq n$. Such a satisfying S-interpretation is called an S-model of Γ. To define tractable subclasses of these problems, we need to name some special classes of DSRs.

Definition 2.1 Let γ be a DSR. Then we let $\mathcal{S}(\gamma)$ denote the set of subset relations in γ, $\mathcal{NE}(\gamma)$ the set of non-equality relations, $\mathcal{D}(\gamma)$ the set of disjointness relations and $\mathcal{NS}(\gamma) = \mathcal{NE}(\gamma) \cup \mathcal{D}(\gamma)$. We say that γ is *subset relation* iff $|\mathcal{NS}(\gamma)| = 0$, γ is a *non-subset relation* iff $|\mathcal{S}(\gamma)| = 0$. If γ is a subset or a non-subset relation we say that γ is *homogeneous* and otherwise *heterogeneous*. If $|\gamma| = 1$ then γ is *atomic* and if $|\mathcal{S}(\gamma)| \leq 1$ then γ is

Horn. We extend these definitions to sets of relations in the obvious way. For example, if Γ is a set of DSRs and all $\gamma \in \Gamma$ are Horn, then Γ is Horn.

Let $\gamma = \{X \subseteq Y, X \,\mathsf{disj}\, Z, X \neq W\}$. Then $\mathcal{S}(\gamma) = \{X \subseteq Y\}$, $\mathcal{NE}(\gamma) = \{X \neq W\}$, $\mathcal{D}(\gamma) = \{X \,\mathsf{disj}\, Z\}$ and $\mathcal{NS}(\gamma) = \{X \,\mathsf{disj}\, Z, X \neq W\}$. Furthermore, γ is Horn but not atomic.

The satisfiability problem for DSRs is denoted DSRSAT and has the following definition:

INSTANCE: A finite set Γ of atomic DSRs.
QUESTION: Does there exist a model of Γ?

Later on, we will see that DSRSAT is NP-complete. However, we begin with the simpler problem ATOMDSRSAT which is DSRSAT restricted to atomic DSRs.

An ATOMDSRSAT instance can be represented by a labelled directed graph, where the nodes are set variables and the arcs are labelled by the relations $\subseteq$, $\neq$ and disj. From now on, let Γ be an arbitrary instance of ATOMDSRSAT and $G = \langle V, E \rangle$ the labelled directed graph representing it. Let G' be the graph obtained from G by removing arcs which are not labelled $\subseteq$. To be able to understand the algorithm in Figure 1, we need some auxiliary definitions.

Definition 2.2 A node $Z \in V$ is *null-defined* iff there exist $X, Y \in V$ such that

- there exists an arc $e \in E$ from X to Y labelled disj in G; and
- there exists a path from Z to X in G'; and
- there exists a path from Z to Y in G'.

We allow paths to be empty, *i.e.*, there always exists a path from X to X for arbitrary X. Hence, the case when we have $X \,\mathsf{disj}\, X$ is covered by this definition. The following lemma is immediate.

Lemma 2.3 If G contains a null-defined node, then there exists no S-model of G.

Proof: If Z is null-defined, then it must be interpreted as the empty set. Consequently, there exists no S-model of Z. □

Definition 2.4 Two nodes $X, Y \in V$ are *inconsistent* iff

- there exists an arc $e \in E$ from X to Y labelled $\neq$; and
- there exists a path from X to Y in G'; and
- there exists a path from Y to X in G'.

Lemma 2.5 If G contains two inconsistent nodes, then there exists no S-model of G.

Proof: If X and Y are inconsistent, then every S-model I must assign values to X and Y such that $I(X) = I(Y)$ and $I(X) \neq I(Y)$. Clearly, this implies the non-existence of an S-model of G. $\Box$

We can now show the tractability of AtomDSRSat and the NP-completeness of DSRSat.

Theorem 2.6 AtomDSRSat is polynomial.

Proof: The algorithm Atom-SAT is obviously polynomial. We claim that the algorithm accepts iff the given instance has an S-model.

Assume the algorithm rejects in line 3 or line 4. Then there exists no S-model by Lemmata 2.3 and 2.5.

Assume to the contrary that the algorithm accepts. We sketch how to construct a model I of G over the set $\mathbb{N}$ of natural numbers.

Assume $V = \{v_1, \ldots, v_n\}$ and define the function $I : V \to 2^{\mathbb{N}}$ such as

$$I(v_i) = \{i\} \cup \{j \mid \text{ there is a path from } v_j \text{ to } v_i \text{ in } G'\}$$

for $1 \leq i \leq n$. We show that I is a model of G by considering the three ways two variables can be related.

1. $v_i \subseteq v_j$. Then there exists a path from v_i to v_j in G' which implies that $I(v_i) \subseteq I(v_j)$ by the construction of I.

2. $v_i \neq v_j$. First, note that v_i and v_j are not inconsistent since otherwise the algorithm would have rejected in line 3. If i is a member of $I(v_j)$, then $j \notin I(v_i)$ since, otherwise, v_i and v_j would be inconsistent. Hence, $I(v_i) \neq I(v_j)$ in this case. Assume to the contrary that i is not a member of $I(v_j)$. It follows immediately that $I(v_i) \neq I(v_j)$ since $i \in I(v_i)$.

1 **algorithm** Atom-SAT.
2 Input: An instance G of Atom-SAT.
3 **if** there exists a null-defined node in G **then reject**
4 **if** there exist inconsistent nodes in G **then reject**
5 **accept**

1 **algorithm** Horn-SAT.
2 Input: An instance Γ of HornDSRSat.
3 $A \leftarrow \bigcup\{\gamma \in \Gamma \mid |\mathcal{NS}(\gamma)| = 0\}$
4 **if** A is not satisfiable **then reject**
5 **if** $\exists\gamma \in \Gamma$ that blocks A and is a non-subset DSR **then reject**
6 **if** $\exists\gamma \in \Gamma$ that blocks A and is heterogeneous **then**
Compute Horn-SAT$((\Gamma - \{\gamma\}) \cup \mathcal{S}(\gamma)$
7 **accept**

Figure 1: The algorithms for AtomDSRSat *and* HornDSRSat.

3. $v_i \,\mathsf{disj}\, v_j$. Assume there exists a $k \in \mathbb{N}$ such that $k \in I(v_i)$ and $k \in I(v_j)$. By the construction of I, there must exist a path from v_k to v_i in G' and a path from v_k to v_j in G'. Obviously, this implies that v_k is null-defined. This leads to a contradiction since the algorithm did not reject in line 4. Consequently, $I(v_i) \,\mathsf{disj}\, I(v_j)$. □

Theorem 2.7 DSRSat is NP-complete.

Proof: Membership in NP follows from Theorem 2.6: Given a set Γ of DSRs, pick one set relation from each DSR in Γ and verify (in polynomial time) the satisfiability of these relations.

We prove NP-hardness by a polynomial-time reduction from the NP-complete problem Graph 3-Colourability. Let $G = \langle V, E\rangle$ be an arbitrary undirected graph. We construct a set of DSRs that are S-satisfiable iff G is 3-colourable. First we introduce three sets $C1$, $C2$ and $C3$ which are related as follows: $\{C1 \,\mathsf{disj}\, C2\}$, $\{C1 \,\mathsf{disj}\, C3\}$, $\{C2 \,\mathsf{disj}\, C3\}$. These sets will correspond to the three "colours" to paint the nodes of G with. For each $w \in V$, we introduce a set variable W and the following two DSRs: $\{W \subseteq C1,\ W \subseteq C2,\ W \subseteq C3\}$ and $\{C1 \subseteq W,\ C2 \subseteq W,\ C3 \subseteq W\}$.

This construction forces W to equal exactly one of the sets $C1$, $C2$ or $C3$. Finally, for each arc $\{x, y\} \in E$ we introduce the DSR $\{X \neq Y\}$. It is obvious that the resulting set of DSRs are S-satisfiable iff G is 3-colourable. Hence, DSRSAT is NP-complete. $\square$

3 The Horn Case

We study the S-satisfiability problem for Horn DSRs, denoted HORNDSRSAT, in this section. To show HORNDSRSAT tractable, the next definition is central.

Definition 3.1 Let A be a satisfiable set of atomic DSRs and let γ be a DSR. We say that γ *blocks* A iff for every $d \in \mathcal{NS}(\gamma)$, $A \cup \{d\}$ is not satisfiable.

Clearly, we can use the Atom-SAT algorithm to decide (in polynomial time) whether a DSR γ blocks a set A of atomic DSRs.

We claim that the algorithm in Figure 1 correctly solves HORNDSRSAT in polynomial time. The key lemma in proving the claim is the following.

Lemma 3.2 Let Γ be a set of arbitrary Horn DSRs. Let $A \subseteq \Gamma$ be the set of subset DSRs in Γ and let $D = \{D_1, \ldots, D_k\} \subseteq \Gamma$ be the set of DSRs that are not subset DSRs. Under the condition that A is S-satisfiable, Γ is S-satisfiable if D_i does not block A for any $1 \leq i \leq k$.

Proof: Pick one non-subset relation d_i out of every D_i such that $A \cup \{d_i\}$ is S-satisfiable. This is possible since no D_i blocks A. We show that $\Gamma' = A \cup \{d_1, \ldots, d_k\}$ is S-satisfiable and, hence, Γ is S-satisfiable.

Let G be the equivalent graph representation of Γ'. By the algorithm Atom-SAT, we know that there are only two circumstances that make G not satisfiable: (1) there exists a null-defined node in G; and (2) there exist inconsistent nodes in G. Hence, if G is not S-satisfiable then one of the two cases above apply.

Assume there exists a null-defined node Z in G. By the definition of null-defined, Z is null-defined also in some instance $A \cup \{d_i\}$, $1 \leq i \leq k$. In other words, d_i blocks A which contradicts our initial assumptions. Assume there exist two inconsistent nodes X and Y. Then there exists an instance $\Sigma = A \cup \{d_i\}$, $1 \leq i \leq k$ such that X and Y are inconsistent in Σ. Hence, d_i blocks A which leads to a contradiction. Consequently, Γ' is satisfiable and the lemma follows. $\square$

The soundness and completeness of Horn-SAT are settled in the forthcoming three lemmata.

Lemma 3.3 Let Γ be a set of Horn DSRs and let $C \subseteq \Gamma$ be the set of subset DSRs in Γ. If there exists a heterogeneous DSR $\gamma \in \Gamma$ such that γ blocks C, then Γ is S-satisfiable iff $(\Gamma - \{\gamma\}) \cup \mathcal{S}(\gamma)$ is S-satisfiable.

Proof: *if:* Trivial.
only-if: If Γ is S-satisfiable, then γ has to be S-satisfiable. Since γ blocks C, $\mathcal{S}(\gamma)$ must be S-satisfied in any solution of Γ. □

Lemma 3.4 Let Γ be a set of Horn DSRs. If Horn-SAT(Γ) accepts then Γ is S-satisfiable.

Proof: Induction over n, the number of heterogeneous DSRs in Γ.
Basis step: If $n = 0$ and Horn-SAT(Γ) accepts then the formulae in A are S-satisfiable and there does not exist any $\gamma \in \Gamma$ that blocks A. Consequently, Γ is S-satisfiable by Lemma 3.2.
Induction hypothesis: Assume the claim holds for $n = k$, $k \geq 0$.
Induction step: Γ contains $k+1$ heterogeneous DSRs. If Horn-SAT accepts in line 6 then $(\Gamma - \{\gamma\}) \cup \mathcal{S}(\gamma)$, which contains k heterogeneous DSRs, is S-satisfiable by the induction hypothesis. By Lemma 3.3, this is equivalent to Γ being S-satisfiable. If Horn-SAT accepts in line 7, then there does not exist any disequational or heterogeneous $\gamma \in \Gamma$ which blocks A. By Lemma 3.2, this means that Γ is S-satisfiable. □

Lemma 3.5 Let Γ be a set of Horn DSRs. If Horn-SAT(Γ) rejects then Γ is not S-satisfiable.

Proof sketch: Induction over the number of heterogeneous DSRs in Γ. The proof is similar to the proof of Lemma 3.4. □

Finally, we can show that Horn-SAT is a polynomial-time algorithm and, thus, show that HornDSRSat is polynomial.

Theorem 3.6 HornDSRSat is polynomial.

Proof: By Lemmata 3.4 and 3.5, it is sufficient to show that Horn-SAT is polynomial. The number of recursive calls is bounded by the number of heterogeneous DSRs in the given input. By Theorem 2.6, we can in polynomial time decide whether a set of atomic DSRs is S-satisfiable. Since we need only check a polynomial number of such systems in each recursion, the theorem follows. □

4 The Horn Fragment of RCC-5

We follow Bennett [1994] in our definition of RCC-5. RCC-5 is based on the notions *regions* and *binary relations* on them. A region X is a variable interpreted over the non-empty sets. It should be noted that we do not require the sets to be open sets in some topological space. This is no limitation since it is impossible to distinguish interior points from boundary points in RCC-5. Thus we can take any set $\mathcal{U}$ and use the discrete topology $\mathcal{T} = \langle \mathcal{U}, 2^{\mathcal{U}} \rangle$. Then every subset of $\mathcal{U}$ is an open set in $\mathcal{T}$.

We assume that we have a fixed universe of variable names for regions. Then, an *R-interpretation* is a function that maps region variables to the non-empty subsets of some set.

Given two interpreted regions, their relation can be described by exactly one of the elements of the set $\mathbf{B}$ of five *basic RCC-5 relations.* These relations are defined as follows: for all X and Y, $X\{\texttt{DR}\}Y$ iff $X \cap Y = \emptyset$, $X\{\texttt{PO}\}Y$ iff $\exists a, b, c : a \in X, a \notin Y, b \in X, b \in Y, c \notin X, c \in Y$, $X\{\texttt{PP}\}Y$ iff $X \subset Y$, $X\{\texttt{PP}^{-1}\}Y$ iff $X \supset Y$ and $X\{\texttt{EQ}\}Y$ iff $X = Y$.

A formula of the form XBY where X and Y are regions and $B \in \mathbf{B}$, is said to be satisfied by an R-interpretation iff the interpretation of the regions satisfies the relations as specified above. To express indefinite information, unions of the basic relations are used, written as sets of basic relations, leading to 2^5 binary RCC-5 relations. Naturally, a set of basic relations is to be interpreted as a disjunction of the basic relations. The set of all RCC-5 relations $2^{\mathbf{B}}$ is denoted by R_5.

A formula of the form $X\{B_1, \ldots, B_n\}Y$ is called a RCC-5 formula. Such a formula is satisfied by an R-interpretation I iff XB_iY is satisfied by I for some i, $1 \leq i \leq n$. A finite set Θ of RCC-5 formulae is said to be *R-satisfiable* iff there exists an R-interpretation I that satisfies every formula of Θ. Such a satisfying R-interpretation is called an *R-model* of Θ. The reasoning problem RSAT has the following definition:

INSTANCE: A finite set Θ of RCC-5 formulae.
QUESTION: Does there exist an R-model of Θ?

We will also consider restricted versions of RSAT where the relations used in formulae in Θ are only from a subset S of R_5. We denote such a problem RSAT(S).

Definition 4.1 Let $\mathbf{B} = \{\texttt{DR}, \texttt{PO}, \texttt{PP}, \texttt{PP}^{-1}, \texttt{EQ}\}$. RCC-5 consists of the set

$R_5 = 2^{\mathbf{B}}$ and the operations unary *converse* (denoted by $\smile$), binary *intersection* (denoted by $\cap$) and binary *composition* (denoted by $\circ$). They are defined such that for all X, Y,

$XR^{\smile}Y$ iff YRX;

$X(R \cap S)Y$ iff $XRY \wedge XSY$; and

$X(R \circ S)Y$ iff $\exists Z : (XRZ \wedge ZSY)$.

Next, we introduce a *closure* operation. The closure operation transforms a given subclass of R_5 to one that is polynomially equivalent to the original subclass wrt. satisfiability.

Definition 4.2 Let $\mathcal{S} \subseteq R_5$. Then we denote by $\overline{\mathcal{S}}$ the *closure* of $\mathcal{S}$, defined as the least subalgebra containing $\mathcal{S}$ closed under converse, intersection and composition.

A program for computing closures can be obtained from the authors. The easy proof of the next lemma is omitted.

Lemma 4.3 Let $\mathcal{S} \subseteq R_5$. RSAT($\mathcal{S}$) is polynomial iff RSAT($\overline{\mathcal{S}}$) is polynomial. RSAT($\mathcal{S}$) is NP-complete iff RSAT($\overline{\mathcal{S}}$) is NP-complete.

We will now focus upon the tractable subclass $\widehat{\mathcal{H}}_5$ which contain all relations in R_5 except $\{\mathtt{PP}, \mathtt{PP}^{-1}\}$, $\{\mathtt{DR}, \mathtt{PP}, \mathtt{PP}^{-1}\}$, $\{\mathtt{PP}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$ and $\{\mathtt{DR}, \mathtt{PP}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$. For our purposes, another definition is more convenient. Let

$$\mathcal{H}_B = \{\{\mathtt{PO}\}, \{\mathtt{DR}, \mathtt{EQ}\}, \{\mathtt{PP}, \mathtt{EQ}\}\}.$$

By a machine-assisted analysis, it can easily be shown that $\widehat{\mathcal{H}}_5 = \overline{\mathcal{H}_B}$.

Theorem 4.4 RSAT($\widehat{\mathcal{H}}_5$) is polynomially reducible to HORNDSRSAT.

Proof: RCC-5 formulae of the form $X\{\mathtt{PP}, \mathtt{EQ}\}Y$ can obviously be replaced by the Horn DSR $X \subseteq Y$. Likewise, an RCC-5 formula of the form $X\{\mathtt{DR}, \mathtt{EQ}\}Y$ can be replaced by the two Horn DSRs $X\{\,\mathsf{disj}\,, \subseteq\}Y$ and $Y\{\,\mathsf{disj}\,, \subseteq\}X$. Finally, RCC-5 formulae of the type $X\{\mathtt{PO}\}Y$ can be replaced by the following six DSRs: $A\{\subseteq\}X$, $A\{\,\mathsf{disj}\,\}Y$, $B\{\subseteq\}X$, $B\{\subseteq\}Y$, $C\{\,\mathsf{disj}\,\}X$ and $C\{\subseteq\}Y$, where A, B and C denote fresh set variables.

We have thus shown how RSAT($\mathcal{H}_B$) can be reduced to HORNDSRSAT in polynomial time. Since we know that $\widehat{\mathcal{H}}_5 = \overline{\mathcal{H}_B}$, the theorem follows from Lemma 4.3. □

Note that the Horn DSR framework allows for spatial formulae which are not formulae in RCC-5. For instance, the disjunction $X\{\texttt{PO}\}Y \lor X\{\texttt{PP}\}Z$ cannot be expressed in RCC-5, even though both PO and PP are in RCC-5. However, the expression is easily converted to a set of Horn DSRs.

5 Topological Properties of RCC-5

The goal of this section is to study some topological properties of RCC-5. First, we show that RCC-5 can always be interpreted over the open sets of the real line under the natural topology. Then, we study some more restrictive topological assumptions and their impact on the complexity of RCC-5.

The following theorem follows immediately from the model construction in Theorem 2.6.

Theorem 5.1 Let Γ be an S-satisfiable, finite set of DSRs. Then Γ has a model that maps region variables to finite subsets of $\mathbb{N}$.

We say that a model that assigns finite subsets of $\mathbb{N}$ to region variables is a $\mathbb{N}$-model. Recall that the natural topology of $\mathbb{R}$ has the set of all open intervals as its basis. This implies that every finite union of open intervals is an open set under this topology.

Theorem 5.2 Let Θ be a R-satisfiable set of RCC-5 formulae. Then Θ has a model of open sets (under the natural topology) in $\mathbb{R}$.

Proof: Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ be an arbitrary, satisfiable, finite set of RCC-5 formulae. Since Θ is R-satisfiable then for each i, $1 \leq i \leq n$, there exists an atomic RCC-5 relation $\alpha_i \in \gamma_i$ such that $A = \{\alpha_1, \ldots, \alpha_n\}$ is R-satisfiable. Observe that A is an instance of RSAT($\widehat{\mathcal{H}}_5$). Hence, we can polynomially reduce A to an equivalent DSRSAT instance B by Theorem 4.4. By Theorem 5.1, B has a $\mathbb{N}$-model I. Hence, there also exists a model I' from the variables in A to the finite subsets of $\mathbb{N}$. Consider the following interpretation:

$$I''(v) = \{y \in \mathcal{R} \mid x \in I'(v) \land x < y < x + 1/2\}.$$

Clearly, I'' is a model of A since I' is a model of A. Furthermore, $I'(v)$ is the union of open intervals in $\mathbb{R}$. Thus, $I'(v)$ is an open set in $\mathcal{R}$ under the natural topology, which concludes the proof. □

It should be clearly understood that the real line equipped with the natural topology is a very weak topological assumption for a spatial theory. Hence, it is interesting to investigate stronger topological assumptions. Such studies have been performed by, for instance, Grigni *et al.* [1995] and Lemon [1996].

Grigni *et al.* [1995] have studied spatial reasoning under such an assumption; they require the regions to be connected sets in two-dimensional Euclidean space. This makes the problem appreciably harder. They show that RCC-5[1] restricted to only atomic relations is NP-hard under these restrictions.

Even more surprising, if we only use the relations DR and PO, the problem remains NP-hard. In this section, we will show that this result does not hold if we allow the regions to be connected sets in three-dimensional Euclidean space (E^3). We say that such a model is a 3DC-model.

Let $R_3 = 2^{\{\mathtt{DR},\mathtt{PO},\mathtt{EQ}\}}$. We define the computational problem, which we will denote R3SAT, as follows:

INSTANCE: A finite set Θ of R_3 formulae.
QUESTION: Does there exist an 3DC-model of Θ?

Our initial observation is that if a set Θ of R_3 formulae has a 3DC-model, then it has an R-model (trivially). By showing the converse, *i.e.* if Θ has an R-model then it has a 3DC-model, we have also shown the polynomiality of R3SAT since $R_3 \subseteq \widehat{\mathcal{H}}_5$.

In the next lemma, we show how to modify an $\mathbb{N}$-model into a 3DC-model.

Lemma 5.3 Let Θ be a finite set of R_3 formulae with an $\mathbb{N}$-model. Then there exists a 3DC-model of Θ.

Proof: Let S be a finite set of natural numbers. We begin by constructing a three-dimensional connected set fork(S) with certain properties. The definition of fork is as follows:

$$\begin{aligned} \mathrm{fork}(S) &= \{(s,y,0) \mid s \in S \wedge 0 \leq y \leq 1\} \cup \\ &\quad \{(x,1,0) \mid \min(S) \leq x \leq \max(S)\}. \end{aligned}$$

[1] In their terminology, RCC-5 is named "medium resolution".

An example of a fork set can be found in Figure 2. Clearly, fork sets are connected.

Let I be an $\mathbb{N}$-model of Θ and assume Θ contains the region variables $X = \{X_1, \ldots, X_n\}$. Let $\alpha : X \to \{y \in \mathcal{R} \mid 0 \leq y < 2\pi\}$ be an arbitrary function satisfying the following requirement: $\alpha(X_i) = \alpha(X_j)$ iff $I(X_i) = I(X_j)$. We shall now construct a 3DC-model I' of Θ: define $I'(X_i)$ as $\text{fork}(I(X_i))$ rotated $\alpha(X_i)$ radians around the x-axis.

It remains to show that I' is a model of Θ. We have three cases:

1. $X_i\{\texttt{DR}\}X_j$ under I. Then $I(X_i) \neq I(X_j)$ so $\alpha(X_i) \neq \alpha(X_j)$. Consequently, $X_i\{\texttt{DR}\}X_j$ under I'.

2. $X_i\{\texttt{PO}\}X_j$ under I. Then there exist natural numbers a, b, c such that

$$a \in I(X_i), a \notin I(X_j)$$
$$b \in I(X_i), b \in I(X_j)$$
$$c \notin I(X_i), c \in I(X_j)$$

By the construction of fork, it follows that

$$(a, 0, 0) \in \text{fork}(I(X_i)), (a, 0, 0) \notin \text{fork}(I(X_j))$$
$$(b, 0, 0) \in \text{fork}(I(X_i)), (b, 0, 0) \in \text{fork}(I(X_j))$$
$$(c, 0, 0) \notin \text{fork}(I(X_i)), (c, 0, 0) \in \text{fork}(I(X_j)).$$

Now, observe that the points we refer to are on the x-axis which is invariant under the rotation imposed by α. Hence, $X_i\{\texttt{PO}\}X_j$ under I'.

3. $X_i\{\texttt{EQ}\}X_j$ under I. Then $I(X_i) = I(X_j)$ so $\alpha(X_i) = \alpha(X_j)$. Furthermore, $\text{fork}(I(X_i)) = \text{fork}(I(X_j))$ which implies that $X_i\{\texttt{DR}\}X_j$ under I'.

We have thus shown that I' is a 3DC-model of Θ and the lemma follows. □

Theorem 5.4 R3SAT is polynomial.

Proof: Let Θ be an instance of R3SAT. Since $R_3 \subseteq \widehat{\mathcal{H}}_5$, it is tractable to decide whether Θ has an R-model or not. If Θ does not have an R-model, then Θ cannot have a 3DC-model. Assume to the contrary that Θ has an R-model. Then, by Theorem 5.1, Θ has an $\mathbb{N}$-model. This implies, by Lemma 5.3, the existence of a 3DC-model of Θ. □

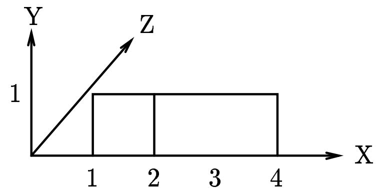

Figure 2: The set fork($\{1, 2, 4\}$).

6 Discussion

Results by Drakengren and Jonsson [1996] show that algorithms similar to Horn-SAT can be used for tractable inference in intuitionistic logic. This opens the way for some intriguing new research. Bennett [1994] has shown that a subtheory of RCC, RCC-8, can be reduced to propositional intuitionistic logic. RCC-8 is more expressive than RCC-5 since it makes it possible to distinguish between boundary and interior points. There are, for instance, two versions of the `PP` relation, one that allow boundary contact and one that disallows such contact.

By employing the connection between RCC-8 and intuitionistic logic established by Bennett, Nebel [1995] discovered the first published tractable subclass of RCC-8. These results were extended by Renz [1996] who isolated a large tractable subclass of RCC-8, denoted $\widehat{\mathcal{H}}_8$. This was accomplished by studying the connection between RCC-8 and yet another logic, a *multimodal* logic. Probably, there is a connection between spatial reasoning, qualitative reasoning about sets and (intuitionistic and/or modal) logic which certainly deserves further study.

7 Conclusions

We have proposed a tractable algorithm for qualitative reasoning about sets. The algorithm can decide the satisfiability of set variables related with subset, disjointness and non-equality. Furthermore, the algorithm allows disjunctions of such expressions under one restriction: there can be at most one subset relation in each disjunction.

To demonstrate the algorithm, we have applied it to problems in spatial reasoning. We have thus shown the existence of a straightforward, polynomial-time algorithm for the Horn fragment of RCC-5 [Renz, 1996]. Furthermore, we have found a subclass of RCC-5, R_3, for which it is tractable to decide whether there exist connected sets in three-dimensional Euclidean space satisfying the given restrictions or not. This finding complements the results in Grigni *et al.* [1995] showing a subclass of R_3 to be NP-complete if we require the sets to be connected sets in two-dimensional Euclidean space.

References

[Bennett, 1994] B. Bennett. Spatial reasoning with propositional logics. In J. Doyle, E. Sandewall, and P. Torasso, editors, *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR '94)*, pages 165–176, Bonn, Germany, May 1994. Morgan Kaufmann.

[Brachman and Schmolze, 1985] R. Brachman and J. Schmolze. An overview of the KL-ONE knowledge representation system. *Cognitive Science*, 9(2):171–216, 1985.

[Cui, 1994] Z. Cui. Using interval logic for order assembly. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pages 103–111. AAAI Press, 1994.

[Drakengren and Jonsson, 1996] Thomas Drakengren and Peter Jonsson. Reasoning about set constraints applied to tractable inference in intuitionistic logic. Submitted journal manuscript, 1996.

[Egenhofer, 1991] M. J. Egenhofer. Reasoning about binary topological relations. In O. Gunther and H. J. Schek, editors, *Advances in Spatial Databases*, pages 143–160. Springer-Verlag, 1991.

[Grigni *et al.*, 1995] M. Grigni, D. Papadias, and C. Papadimitriou. Topological inference. In Chris Mellish, editor, *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI '95)*, pages 901–906, Montréal, PQ, Canada, August 1995. Morgan Kaufmann.

[Lemon, 1996] O. Lemon. Semantical foundations of spatial logics. In J. Doyle and L. Aiello, editors, *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR '96)*, pages 212–219, Cambridge, MA, USA, October 1996. Morgan Kaufmann.

[Nebel, 1995] Bernhard Nebel. Computational properties of qualitative spatial reasoning: First results. In I. Wachsmuth, C-R. Rollinger, and W. Brauer, editors, *KI-95: Advances in Artificial Intelligence*, pages 233–244, Bielefeld, Germany, September 1995. Springer-Verlag.

[Randell and Cohn, 1989] D. A. Randell and A. G. Cohn. Modelling topological and metrical properties of physical processes. In Ronald J. Brachman, Hector J. Levesque, and Raymond Reiter, editors, *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning (KR '89)*, pages 55–66, Toronto, ON, Canada, May 1989. Morgan Kaufmann.

[Randell *et al.*, 1992] D. A. Randell, Z. Cui, and A. G. Cohn. A spatial logic based on regions and connection. In Bill Swartout and Bernhard Nebel, editors, *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR '92)*, pages 165–176, Cambridge, MA, USA, October 1992. Morgan Kaufmann.

[Renz, 1996] Jochen Renz. Qualitatives räumliches Schließen: Berechnungseigenschaften und effiziente Algorithmen. Master's thesis, Fakultät für Informatik, Universität Ulm, 1996. Available from `http://www.informatik.uni-freiburg.de/~sppraum`.