RCC-5 and its Tractable Subclasses

Peter Jonsson and Thomas Drakengren
Email: {petej,thodr}@ida.liu.se

Abstract

We investigate the computational properties of the spatial algebra RCC-5 which is a restricted version of the RCC framework. The satisfiability problem for RCC-5 is known to be NP-complete but not much is known about its approximately four billion subclasses. We provide a complete classification of satisfiability for all these subclasses into polynomial and NP-complete respectively. In the process, we identify all maximal tractable subalgebras which are four in total.

IDA Technical Report 1996
LiTH-IDA-R-96-25
ISSN-0281-4250

Department of Computer
and Information Science,
Linköping University,
S-581 83 Linköping, Sweden

1 Introduction

Qualitative spatial reasoning has received a constantly increasing amount of interest in the literature. A well-known framework for qualitative spatial reasoning is the so-called 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 are two subtheories of RCC, RCC-5 and RCC-8, which were first presented by Bennett [1994]. Bennett shows that it is sufficient to use propositional logics for reasoning about these subclasses. Hence, the reasoning becomes more efficient when compared to reasoning in a full first-order logic. Bennett's approach uses classical propositional logic for RCC-5 and intuitionistic propositional logic for RCC-8. Unfortunately, these logics are known to be computationally hard. The satisfiability problem for classical propositional logic and intuitionistic propoisitional logic is NP-complete [Garey and Johnson, 1979] and PSPACE-complete [Statman, 1979] respectively. However, the complexity of the underlying logic does not carry over in both cases; Renz [1996] has shown that the satisfiability problem for both RCC-5 and RCC-8 is NP-complete.

These findings motivate the search for tractable subclasses of RCC-5 and RCC-8. Nebel [1995] showed that reasoning with the basic relations in RCC-8 is a polynomial-time problem. Renz [1996] improved this result substantially by showing the following results:

- There exists a large, maximal subclass of RCC-8, denoted $\widehat{\mathcal{H}}_8$, which contains all basic relations and is polynomial. $\widehat{\mathcal{H}}_8$ contains 148 elements out of 256 (58%).

- There exists a large, maximal subclass of RCC-5, denoted $\widehat{\mathcal{H}}_5$, which contains all basic relations and is polynomial. $\widehat{\mathcal{H}}_5$ contains 28 elements out of 32 (87%). Furthermore, this is the unique, maximal subclass of RCC-5 containing all basic relations.

We will concentrate on RCC-5 in this article. The main result is a complete classification of all subclasses of RCC-5 with respect to tractability. The classification makes it possible to determine whether a given subclass is tractable or not by a simple test that can be carried out by hand or automatically. We have thus gained a clear picture of the tractability borderline in RCC-5. As is more or less necessary when showing results of this kind, the main proof relies on a case analysis performed by a computer. The number of cases considered was approximately 4×10^4. This figure is quite small when compared to the 4×10^6 cases that were necessary to prove the results in [Jonsson *et al.*, 1996]. However, it is still a rather large analysis if we compare it with the approximately 8000 cases needed to show the uniquness of the ORD-Horn class [Nebel and Bürckert, 1995]. The analysis cannot, of course, be reproduced in a research paper or be verified manually. Hence, we include a description of the program used. The program can be obtained from the authors.

The structure of the article is as follows: Section 2 defines RCC-5 and some auxiliary concepts. Section 3 contains the tractability proofs for four subclasses of RCC-5 and in Section 4 we show that these are the only maximal tractable subclasses of RCC-5. The article concludes with a brief discussion of the results.

2 The RCC-5 Theory

We follow Bennett [1994] in our definition of RCC-5. RCC-5 is based on the notions of *regions* and *binary relations* on them. A region p is a variable interpreted over the non-empty subsets of some fixed set. 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 $\mathcal{X}$ and use the discrete topology $\mathcal{T} = \langle \mathcal{X}, 2^{\mathcal{X}} \rangle$, where every subset of $\mathcal{X}$ 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*. The definition of these relations can be found in Table 1. Figure 1 gives 2-dimensional examples of the relations in RCC-5. 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 specified in Table 1.

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. Relations of special interest are the *null* relation $\varnothing$ (also denoted by $\perp$) and the *universal* relation $\mathbf{B}$ (also denoted $\top$).

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 $\Im$ iff XB_iY is satisfied by $\Im$ 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 $\Im$ that satisfies every formula of Θ. Such a satisfying R-interpretation is called an *R-model* of Θ. The reasoning problem we will study is the following:

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

We denote this problem by RSAT. In the following, we often consider restricted versions of RSAT where the relations used in formulae in Θ are only from a subset S of R_5. In this case we say that Θ is a set of formulae over r and we use a parameter in the problem description to denote the subclass under consideration, *e.g.* RSAT(S).

Definition 2.1 Let $\mathbf{B} = \{\mathtt{DR}, \mathtt{PO}, \mathtt{PP}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$. RCC-5 consists of the set

$X\{\mathtt{DR}\}Y$	iff	$X \cap Y = \varnothing$
$X\{\mathtt{PO}\}Y$	iff	$\exists a, b, c : a \in X, a \notin Y, b \in X, b \in Y, c \notin X, c \in Y$
$X\{\mathtt{PP}\}Y$	iff	$X \subset Y$
$X\{\mathtt{PP}^{-1}\}Y$	iff	$X \supset Y$
$X\{\mathtt{EQ}\}Y$	iff	$X = Y$

Table 1: The five basic relations of RCC-5.

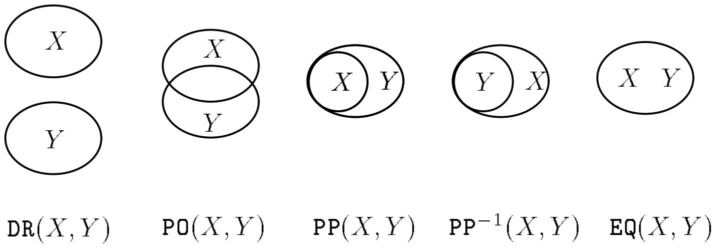

Figure 1: Pictorial example of the relations in RCC-5.

$R_5 = 2^{\mathbf{B}}$ and the operations unary *converse* (denoted by $^\smile$), binary *intersection* (denoted by $\cap$) and binary *composition* (denoted by $\circ$). They are defined as follows:

$$\begin{array}{lrll} \forall X, Y\colon & X R^\smile Y & \text{iff} & Y R X \\ \forall X, Y\colon & X(R \cap S)Y & \text{iff} & X R Y \wedge X S Y \\ \forall X, Y\colon & X(R \circ S)Y & \text{iff} & \exists Z : (X R Z \wedge Z S Y) \end{array}$$

It can easily be verified that $R \circ S = \bigcup\{B \circ B' | B \in R, B'' \in S\}$, *i.e.* composition is the union of the component-wise composition of basic relations.

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 with respect to satisfiability. The closure operation is similar to the closure operation for RCC-5 introduced by Renz [1996] but it does not pose the same restrictions on the given subclass.[1]

Definition 2.2 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. We proceed by showing that closures preserve polynomiality and NP-completeness.

Lemma 2.3 Let $\mathcal{S} \subseteq R_5$. Then RSAT($\overline{\mathcal{S}}$) can be polynomially transformed to RSAT($\mathcal{S}$).

[1] Renz's operation requires $\{\mathtt{EQ}\}$ to be a member of the subclass to be closed. The reason for this is that he studies binary constraint networks instead of sets of formulae, as we do.

Proof: Let $\mathcal{T} = \overline{\mathcal{S}} - \mathcal{S}$. Every element of $R \in \mathcal{T}$ is equivalent to some expression ε_R over $\mathcal{S}$ involving converse, intersection and composition. Let m be the maximum number of operators appearing in these expressions.

We show by induction that for any set of formulae Θ over $\overline{\mathcal{S}}$, we can construct a set Θ' over $\mathcal{S}$ such that $|\Theta'| \leq (2^m \times |\Theta|)$ and Θ is R-satisfiable iff Θ' is. Since m is fixed for a given $\mathcal{S}$, this is a polynomial-time transformation.

Basis step: $m = 1$. For any formula $(XRY) \in \Theta$ such that $R \in \mathcal{T}$ one of the following cases applies:

1. $R = S^{\smile}$ and $S \in \mathcal{S}$. In this case, the formula XRY in Θ is replaced by YSX.
2. $R = S \cap T$ and $S, T \in \mathcal{S}$. In this case, the formula XRY in Θ is replaced by the two formulae XSY, XTY.
3. $R = S \circ T$ and $S, T \in \mathcal{S}$. In this case, the formula XRY in Θ is replaced by the two formulae XSZ, ZTY where Z is a fresh variable.

Clearly, if Θ is R-satisfiable then Θ' is and vice versa. Further, $|\Theta'| \leq 2^1 \times |\Theta|$.

Inductive step: We assume that the hypothesis holds for $m = k$ and assume that the maximum number of operators appearing in expressions ε_R for $R \in \mathcal{T}$ is $k+1$. Let $\mathcal{T}' \subseteq \mathcal{T}$ be the relations R such that the expressions ε_R involve $k+1$ operators. For all these relations we can find expressions ε'_R over $\overline{\mathcal{S}} - \mathcal{T}'$ that contain only one operator.

Applying now the above transformation for all $R \in \mathcal{T}'$ using ε'_R yields a set Θ' over $\overline{\mathcal{S}} - \mathcal{T}'$ of size $2 \times |\Theta|$ that is equivalent to Θ wrt. R-satisfiability. Applying the induction hypothesis gives that it is possible to construct a set Θ' of size $2^{k+1} \times |\Theta|$ that is equivalent to Θ wrt. R-satisfiability, which concludes the induction. $\square$

Corollary 2.4 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.

3 Tractable Subclasses of RCC-5

We begin this section by defining four tractable subalgebras of RCC-5, which can be found in Table 2. Later on, we show that these algebras are the only maximal tractable subalgebras of RCC-5. The tractability of the first algebra, R_5^{28}, has been established by Renz [1996]. The name R_5^{28} reflects the fact that the algebra contains 28 elements.

Theorem 3.1 RSAT(R_5^{28}) is polynomial.

The tractability of our second algebra, RSAT(R_5^{20}), can be settled quite easily. The algorithm can be found in Figure 3.

Lemma 3.2 The algorithm RSAT(R_5^{20})(Θ) accepts iff Θ has an R-model.

Proof: *if:* We show the contrapositive, *i.e.* if RSAT(R_5^{20})(Θ) rejects then Θ has no R-model. Clearly, the satisfiability of Θ is preserved under the transformations made in lines 7-10. Hence, if the algorithm rejects in line 5 or 6 then Θ is not satisfiable.

only-if: Consider the set Θ after the completion of line 11. We denote this set by Θ'. Obviously, Θ' is satisfiable iff the initial Θ was satisfiable. Also observe that line 7 ensures that Θ' does not relate any variables with `EQ`. Furthermore, line 8 guarantees that there is at most one relation that relates two variables.

Now, we construct an R-model M for Θ' as follows: Let V be the set of variables in Θ'. Assign to the members of V non-empty sets that are pairwise disjoint. For each relation of the type $X\{\mathtt{PO}\}Y$ or $X\{\mathtt{PO},\mathtt{EQ}\}Y$, introduce a new value $\alpha_{X,Y}$ and let $X = X \cup \{\alpha_{X,Y}\}$ and $Y = Y \cup \{\alpha_{X,Y}\}$. Clearly, two sets X, Y are disjoint (and are thus related by `DR`) under M unless $X\{\mathtt{PO}\}Y$ or $X\{\mathtt{PO},\mathtt{EQ}\}Y$ is in Θ. But in these cases, X and Y must not be disjoint. In fact, by introducing $\alpha_{X,Y}$, we have forced $X\{\mathtt{PO}\}Y$ to hold under M which satisfies formulae of the type $X\{\mathtt{PO}\}Y$ as well as formulae of the type $X\{\mathtt{PO},\mathtt{EQ}\}Y$. Hence, M is an R-model of Θ' which implies the R-satisfiability of Θ. $\Box$

Theorem 3.3 RSAT(R_5^{20}) is polynomial.

Proof: Algorithm RSAT(R_5^{20}) correctly solves the RSAT(R_5^{20}) problem by the previous lemma. Furthermore, the number of iterations are bounded from above by the number of variables and the number of formulae in the given instance and the tests can easily be performed in polynomial time. $\Box$

Next we show the tractability of RSAT(R_5^{17}).

Theorem 3.4 RSAT(R_5^{17}) is polynomial.

Proof: Consider the algorithm RSAT(R_5^{17}) in Figure 3. If there exists X, Y such that $X \perp Y \in \Theta$ then Θ is not satisfiable. Otherwise, we can let all variables get the same value. Since `EQ` is a member of every relation that occurs in Θ, this interpretation is an R-model of Θ. $\Box$

To show the tractability of RSAT(R_5^{14}), we need a number of auxiliary results. First, note that an RSAT problem instance can be represented by a labelled directed graph, where the nodes are region variables and the arcs are labelled by relations between variables. From now on, let Θ be an arbitrary instance of RSAT(R_5^{14}) and $G = \langle V, E\rangle$ the labelled directed graph representing it. The forthcoming proofs are similar in spirit with some of the proofs appearing in [Drakengren and Jonsson, 1996].

Definition 3.5 A RCC-5 relation R is said to be an *acyclic* relation iff any cycle in any G with R on every arc is never satisfiable.

algorithm RSAT(R_5^{20}).
Input: An instance Θ of RSAT(R_5^{20}).
repeat
  $\Theta' \leftarrow \Theta$
  if $\exists X, R : XRX \in \Theta$ and $\mathtt{EQ} \notin R$ **then reject**
  if $\exists X, Y : X \bot Y \in \Theta$ **then reject**
  if $\exists X, Y : X \neq Y$ and $X\{\mathtt{EQ}\}Y \in \Theta$ **then** substitute Y for X in Θ
  if $\exists X, Y, R, S : XRY \in \Theta$ and $XSY \in \Theta$ **then**
    $\Theta \leftarrow (\Theta - \{XRY, XSY\}) \cup \{X(R \cap S)Y\}$
  if $\exists X, R : XRX \in \Theta$ and $\mathtt{EQ} \in R$ **then** $\Theta \leftarrow \Theta - \{XRX\}$
until $\Theta = \Theta'$
accept

algorithm RSAT(R_5^{17}).
Input: An instance Θ of RSAT(R_5^{17}).
if $\exists X, Y$ such that $X \bot Y \in \Theta$ **then reject**
else accept

algorithm RSAT(R_5^{9}).
Input: An instance G of RSAT(R_5^{9}).
Let G' be the graph obtained from G by removing arcs which are not labelled $\{\mathtt{PP}, \mathtt{EQ}\}$.
Find all strong components C in G'
for every arc e in G whose relation does not contain EQ **do**
  if e connects two nodes in some C **then reject**
accept

Figure 2: Algorithms for RSAT(R_5^{20}), RSAT(R_5^{17}) and RSAT(R_5^{9}).

The relation PP is an example of an acyclic relation while $\{\mathtt{PP}, \mathtt{EQ}\}$ is not acyclic. We continue by showing a few properties of acyclic relations.

Corollary 3.6 Let R be an acyclic relation. Then every relation $R' \subseteq R$ is acyclic.

Proof: Since taking subsets of R constrains satisfiability further, the result follows. $\square$

Corollary 3.7 Let R be an acyclic relation, and choose A such that $A \subseteq \{R' \mid R' \subseteq R\}$. Then, any cycle in G where every arc is labelled by some relation in A is unsatisfiable.

Proof: Same argument as in the previous corollary. $\square$

The following definition is needed in the forthcoming proofs.

Definition 3.8 Let $\mathcal{I}$ be an instance of the R-satisfiability problem, M a model for $\mathcal{I}$ and $r \in R_5$ a relation between two region variables X and Y in $\mathcal{I}$. Then r is said to be *satisfied as* r' for any relation $r' \in R_5$, such that $X r' Y$ is satisfied in M.

The definition may seem a bit cumbersome but the essence should be clear. We provide a small example for further clarification.

Example 3.1 Let X and Y be region variables related by $X\{\mathtt{PO},\mathtt{PP}\}Y$, and M a model where X is interpreted as $\{1,2\}$ and Y as $\{1,2,3\}$. Then in M, $\{\mathtt{PO},\mathtt{PP}\}$ is satisfied as $\{\mathtt{PP}\}$, but also as $\{\mathtt{PO},\mathtt{PP}\}$.

Lemma 3.9 Let R be an acyclic relation, and A, A' sets such that $A \subseteq \{R' \mid R' \subseteq R\}$ and $A' \subseteq \{a \cup \{\mathtt{EQ}\} \mid a \in A\}$. Then, every cycle C labelled by relations in $A \cup A'$ is satisfiable iff it contains only relations from A'. Furthermore, all relations in the cycle have to be satisfied as $\mathtt{EQ}$.

Proof: *only-if:* Suppose that a cycle C is satisfiable and that it contains some relation from A. Induction on the number n of arcs in the cycle. For $n = 1$, we get a contradiction by the assumption. So, suppose for the induction that C contains $n+1$ arcs. Let M be an R-model for the relations in C. We cannot have that every relation in C is satisfied in M as some relation in A, by Corollary 3.7. Thus, some relation R' in C has to be satisfied as $\mathtt{EQ}$. But then we can collapse the two variables connected by R' to one variable, and we have a cycle with n nodes containing a relation from A. This contradicts the induction hypothesis.

if: Suppose that a cycle C contains only relations in A'. Then C can be satisfied by choosing $\mathtt{EQ}$ on every arc, thus forcing the variables to be identical. $\square$

After having studied acyclic relations, we will now turn our attention to *DAG-satisfying* relations. The formal definition is as follows.

Definition 3.10 A basic relation B is said to be *DAG-satisfying* iff any DAG (directed acyclic graph) labelled only by relations containing B is satisfiable.

A typical example of a DAG-satisfying relation is $\mathtt{EQ}$. Given a DAG labelled only by relations containing $\mathtt{EQ}$, we can always satisfy these relations by assigning some non-empty set S to all variables.

We can now show that $\mathtt{PP}$ is a DAG-satisfying relation.

Lemma 3.11 The basic relation $\mathtt{PP}$ is DAG-satisfying.

Proof: Before the actual proof, we need to define what a maximal node in a DAG is. Let G be an arbitrary DAG. Then a node v in G is *maximal* iff there are no arcs which start in v.

Now, let G be a DAG labelled only by relations containing PP. We show that G is satisfied by some R-model M. Induction on n which is the number of nodes in G. The case when $n = 0$ is trivial. Suppose that G has $n+1$ nodes and remove a maximal node g. By induction, the remaining graph $G' = \langle V', E' \rangle$ is satisfiable by a model M'. We shall now construct a model M of G, which agrees with M' on every variable in G'. Let $S = \bigcup\{M'(v) | v \in V'\} \cup \{\alpha\}$ where α is a fresh element. Let $M(g) = S$. Obviously, M is a model of G. □

We now state a simple result which holds for directed graphs in general.

Lemma 3.12 Let G be irreflexive[2] and having an acyclic subgraph D. Then those arcs of G which are not in D can be reoriented so that the resulting graph is acyclic.

Proof: Trivial. (A proof appears in [Drakengren and Jonsson, 1996].) □

By specialising this result, we get the next lemma.

Lemma 3.13 Let G be irreflexive with an acyclic subgraph D and let the arcs of D be labelled by relations containing PP, and the arcs not in D being labelled by relations containing PP and $\mathtt{PP}^{-1}$. Then G is R-satisfiable.

Proof: Reorient the arcs of G like in the previous lemma, yielding a DAG G'. In this construction, whenever an arc is reoriented, also invert the relation on that arc, so that G' is satisfiable iff G is. By the construction, only arcs containing both PP and $\mathtt{PP}^{-1}$ have been reoriented, so every arc in the DAG G' contains PP and, thus, since PP is DAG-satisfying, G' is satisfiable. Consequently, also G is satisfiable. □

We can now prove the correctness of the algorithm RSAT(R_5^9) Figure 3. However, this algorithm does not solve the RSAT(R_5^{14}) problem but a more restricted problem, namely the RSAT(R_5^9) problem. The subclass R_5^9 contains the elements $\{\{\mathtt{PP}, \mathtt{EQ}\}\} \cup \{R \cup \{\mathtt{PP}, \mathtt{PP}^{-1}\} \mid R \in R_5\}$. It can be verified that $R_5^{14} = \overline{R_5^9}$ so this is no limitation by Corollary 2.4.

Lemma 3.14 Algorithm RSAT(R_5^9) correctly solves RSAT(R_5^9).

Proof: Suppose the algorithm finds a strong component of G' (which then contain only the relation $\{\mathtt{PP}, \mathtt{EQ}\}$), in which two of the nodes are connected by an arc e, labelled by an relation R' which does not contain EQ. Then there exists a cycle C, in which the relation of every arc contains EQ, such that e connects two nodes in that C but e is not part of that cycle. By Lemma 3.9, C can be satisfied only by choosing the relation EQ on every arc in C, and since R' does not admit EQ, C is unsatisfiable.

Otherwise, every such strong component can be collapsed to a single variable, removing all arcs which start and end in the collapsed variable.

[2] A graph is irreflexive iff it has no arcs from a node v to the node v.

This transformation retains satisfiability using the same argument as above. After the collapsing, the subgraph obtained by considering only arcs labelled by $\{\mathtt{PP}, \mathtt{EQ}\}$ will be acyclic. Since the remaining arcs are labelled by relations containing both PP and $\mathtt{PP}^{-1}$, the graph is R-satisfiable by Lemma 3.13. (Note that the graph will be irreflexive since every node is contained in some strong component). □

Lemma 3.15 Algorithm RSAT(R_5^9) runs in linear time in the size of G.

Proof: Strong components can be found in linear time (see *e.g.* [Baase, 1988] and the remaining test can also be done in linear time. □

Theorem 3.16 RSAT(R_5^{14}) can be solved in polynomial time.

Proof: Follows immediately from the previous two lemmata. □

4 Classification of RCC-5

Before we can give the classification of RCC-5 we need two NP-completeness results.

Theorem 4.1 RSAT($\mathcal{S}$) is NP-complete if

1. $\mathcal{C}_1 = \{\{\mathtt{PO}\}, \{\mathtt{PP}, \mathtt{PP}^{-1}\}\} \subseteq \mathcal{S}$, or
2. $\mathcal{C}_2 = \{\{\mathtt{DR}, \mathtt{PO}\}, \{\mathtt{PP}, \mathtt{PP}^{-1}\}\} \subseteq \mathcal{S}$.

Proof: The proof for $\mathcal{C}_1$ appears in Renz [1996]. The proof for $\mathcal{C}_2$ is by polynomial-time reduction from RSAT($\mathcal{C}_1$). Let Θ be an arbitrary instance of RSAT($\mathcal{C}_1$). Construct the following set:

$$\Theta' = \{X\{\mathtt{PP}, \mathtt{PP}^{-1}\}Y \mid X\{\mathtt{PP}, \mathtt{PP}^{-1}\}Y \in \Theta\} \cup \{X\{\mathtt{DR}, \mathtt{PO}\}Y \mid X\{\mathtt{PO}\}Y \in \Theta\}.$$

Clearly, Θ' can be obtained from Θ in polynomial time and Θ' is an instance of RSAT($\mathcal{C}_2$). We show that Θ is satisfiable iff Θ' is satisfiable.

only-if: Assume that there exists an R-model I of Θ. It is not hard to see that I is also an R-model of I' since if $X\{\mathtt{PO}\}Y$ under I then $X\{\mathtt{DR}, \mathtt{PO}\}Y$ under I. Thus Θ' is R-satisfiable if Θ is R-satisfiable.

if: Assume the existence of an R-model I' that assigns the region variables of Θ' subsets of some set $\mathcal{U}$. Let α be an element such that $\alpha \notin \mathcal{U}$. We construct a new interpretation I as follows: $I(x) = I'(x) \cup \{\alpha\}$ for every variable x in Θ'. It can easily be seen that the following holds for I:

1. If $x\{\mathtt{DR}\}y$ under I' then $x\{\mathtt{PO}\}y$ under I.
2. If $x\{\mathtt{PO}\}y$ under I' then $x\{\mathtt{PO}\}y$ under I.

	R_5^{28}	R_5^{20}	R_5^{17}	R_5^{14}
$\bot$	$\bullet$	$\bullet$	$\bullet$	$\bullet$
$\{\mathtt{DR}\}$	$\bullet$	$\bullet$		
$\{\mathtt{PO}\}$	$\bullet$	$\bullet$		
$\{\mathtt{DR}, \mathtt{PO}\}$	$\bullet$	$\bullet$		
$\{\mathtt{PP}\}$	$\bullet$			$\bullet$
$\{\mathtt{DR}, \mathtt{PP}\}$	$\bullet$	$\bullet$		
$\{\mathtt{PO}, \mathtt{PP}\}$	$\bullet$			
$\{\mathtt{DR}, \mathtt{PO}, \mathtt{PP}\}$	$\bullet$	$\bullet$		
$\{\mathtt{PP}^{-1}\}$	$\bullet$			$\bullet$
$\{\mathtt{DR}, \mathtt{PP}^{-1}\}$	$\bullet$	$\bullet$		
$\{\mathtt{PO}, \mathtt{PP}^{-1}\}$	$\bullet$			
$\{\mathtt{DR}, \mathtt{PO}, \mathtt{PP}^{-1}\}$	$\bullet$	$\bullet$		
$\{\mathtt{PP}, \mathtt{PP}^{-1}\}$				$\bullet$
$\{\mathtt{DR}, \mathtt{PP}, \mathtt{PP}^{-1}\}$		$\bullet$		$\bullet$
$\{\mathtt{PO}, \mathtt{PP}, \mathtt{PP}^{-1}\}$	$\bullet$			$\bullet$
$\{\mathtt{DR}, \mathtt{PO}, \mathtt{PP}, \mathtt{PP}^{-1}\}$	$\bullet$	$\bullet$		$\bullet$
$\{\mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	$\bullet$
$\{\mathtt{DR}, \mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	
$\{\mathtt{PO}, \mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	
$\{\mathtt{DR}, \mathtt{PO}, \mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	
$\{\mathtt{PP}, \mathtt{EQ}\}$	$\bullet$		$\bullet$	$\bullet$
$\{\mathtt{DR}, \mathtt{PP}, \mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	
$\{\mathtt{PO}, \mathtt{PP}, \mathtt{EQ}\}$	$\bullet$		$\bullet$	
$\{\mathtt{DR}, \mathtt{PO}, \mathtt{PP}, \mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	
$\{\mathtt{PP}^{-1}, \mathtt{EQ}\}$	$\bullet$		$\bullet$	$\bullet$
$\{\mathtt{DR}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	
$\{\mathtt{PO}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$	$\bullet$		$\bullet$	
$\{\mathtt{DR}, \mathtt{PO}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$	$\bullet$	$\bullet$	$\bullet$	
$\{\mathtt{PP}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$			$\bullet$	$\bullet$
$\{\mathtt{DR}, \mathtt{PP}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$		$\bullet$	$\bullet$	$\bullet$
$\{\mathtt{PO}, \mathtt{PP}, \mathtt{PP}^{-1}, \mathtt{EQ}\}$	$\bullet$		$\bullet$	$\bullet$
$\top$	$\bullet$	$\bullet$	$\bullet$	$\bullet$

Table 2: The maximal tractable subalgebras of RCC-5.

3. If $x\{\mathtt{PP}\}y$ under I' then $x\{\mathtt{PP}\}y$ under I.

4. If $x\{\mathtt{PP}^{-1}\}y$ under I' then $x\{\mathtt{PP}^{-1}\}y$ under I.

Consequently, if $x\{\mathtt{PP},\mathtt{PP}^{-1}\}y$ under I' then $x\{\mathtt{PP},\mathtt{PP}^{-1}\}y$ under I and if $x\{\mathtt{DR},\mathtt{PO}\}y$ under I' then $x\{\mathtt{PO}\}y$ under I. Consequently, I is a model of Θ so Θ is R-satisfiable if Θ' is R-satisfiable. $\square$

The main theorem can now be stated and proven.

Theorem 4.2 For $\mathcal{S} \subseteq R_5$, RSAT($\mathcal{S}$) is polynomial iff $\mathcal{S}$ is a subset of some member of $R_P = \{R_5^{28}, R_5^{20}, R_5^{17}, R_5^{14}\}$, and NP-complete otherwise.

Proof: *if:* For each $r \in R_P$, RSAT(r) is polynomial as was shown in the previous section.
only-if: Assume there exists a subclass $\mathcal{S} \subseteq R_5$ such that RSAT($\mathcal{S}$) is polynomial and $\mathcal{S}$ is not a subset of any algebra in R_P. Without loss of generality, let $\mathcal{S}$ be such a class with the least number of elements. For each subalgebra r in R_P, choose a relation x such that $x \in \mathcal{S}$ and $x \notin r$. This can always be done since $\mathcal{S} \not\subseteq r$. Let X be the set of these relations. The following holds for X:

1. RSAT(X) is polynomial since $X \subseteq \mathcal{S}$.

2. X is not a subset of any algebra in R_P.

$\mathcal{S}$ is a minimal set satisfying (1) and (2) above. Hence, $|\mathcal{S}| \leq |X|$. Furthermore, $X \subseteq \mathcal{S}$ so $|\mathcal{S}| \leq |X|$ and $|\mathcal{S}| = |X|$. The set R_P contains four algebras so by the construction of X, $|X| \leq 4$. As a consequence, $|\mathcal{S}| \leq 4$.

To show that $\mathcal{S}$ does not exist, a machine-assisted case analysis of the following form was performed:

1. Generate all subsets of R_5 of size ≤ 4. There are $\sum_{i=0}^{4} \binom{32}{i} = 41449$ such subsets.

2. Let $\mathcal{T}$ be such a set. Test: $\mathcal{T}$ is a subset of some subalgebra in R_P or $\mathcal{C}_i \subseteq \overline{\mathcal{T}}$ for some $i \in \{1,2\}$.

The test succeeds for all $\mathcal{T}$. Hence, either RSAT($\mathcal{S}$) is NP-complete or $\mathcal{S}$ is a subset of some member of R_P. Both cases contradict our initial assumptions so $\mathcal{S}$ cannot exist. $\square$

5 Discussion

The complexity of problems in AI has, almost without exception, been studied in an *ad hoc* manner. Papers have reported complexity results for problems that the author believes is interesting for the AI community. However, these papers have not always given evidence that this is the case. The more systematic approach of building complete classifications has several advantages over this traditional method. First of all, the question of relevance drops entirely. If the problem class under consideration is regarded relevant, then its tractable subclasses are relevant. Furthermore, a complete classification is far more satisfactory from a scientific point of view; to gain a clear picture of the borderline between tractability and intractability has an intrinsic scientific value. The major drawback with the approach is its inherent difficulty. Clearly, it can be an extremely non-trivial task to provide complete classifications even if one uses computer-generated proofs.

The work reported in this article can be extended in many different ways. One obvious extension is to study other computational problems than the RSAT problem. Renz [1996] has studied two problems, RMIN and RENT, on certain subclasses of RCC-5 and RCC-8. The problem RMIN is to decide if a set of spatial formulae Θ is minimal, *i.e.* whether all basic relations in every formula of Θ can be satisfied or not. The problem RENT is to decide whether a formula XRY is entailed by a set of spatial formulae. One open problem is to classify the subclasses of RCC-5 with respect to these two problems.

Another interesting research direction is to completely classify other spatial algebras, such as RCC-8. RCC-8 contains $2^{256} \approx 10^{77}$ relations so the question whether this is feasible or not remains to be answered. A (probably) easier problem is to study the class $\mathcal{C}_0$ suggested by Bennett [1994] which only contains four relations. Certainly, there are also other spatial algebras worth studying.

6 Conclusions

We have studied computational properties of RCC-5. All of the 2^{32} possible subclasses are classified with respect to whether their corresponding satisfiability problem is tractable or not. The classification reveals that there are four maximally tractable subclasses of the algebra.

References

[Baase, 1988] Sara Baase. *Computer Algorithms: Introduction and Analysis.* Addison Wesley, Reading, MA, 2nd edition, 1988.

[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 on Knowledge Representation and Reasoning (KR-94)*, pages 165–176, Bonn, Germany, May 1994. Morgan Kaufmann.

[Drakengren and Jonsson, 1996] Thomas Drakengren and Peter Jonsson. Maximal tractable subclasses of Allen's interval algebra: Preliminary report. In *Proceedings of the 13th (US) National Conference on Artificial*

Intelligence (AAAI-96), pages 389–394, Portland, OR, USA, August 1996. American Association for Artificial Intelligence.

[Garey and Johnson, 1979] Michael Garey and David Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, New York, 1979.

[Jonsson *et al.*, 1996] Peter Jonsson, Thomas Drakengren, and Christer Bäckström. Tractable subclasses of the point-interval algebra: A complete classification. In J. Doyle and L. Aiello, editors, *Proceedings of the 5th International Conference on Principles on Knowledge Representation and Reasoning (KR-96)*, Cambridge, MA, USA, October 1996. Morgan Kaufmann.

[Nebel and Bürckert, 1995] Bernhard Nebel and Hans-Jürgen Bürckert. Reasoning about temporal relations: A maximal tractable subclass of Allen's interval algebra. *Journal of the ACM*, 42(1):43–66, 1995.

[Nebel, 1995] B Nebel. Computational properties of qualitative spatial 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 on 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 on 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 Algortithmen. Master thesis report, Fakultät für Informatik, Universität Ulm, 1996.

[Statman, 1979] R Statman. Intuitionistic logic is polynomial-space complete. *Theoretical Computer Science*, 9(1):67–72, 1979.