To appear in the *16th International Joint Conference on Artificial Intelligence (IJCAI'99)*, Stockholm, Sweden, August 1999.

Maximal Tractable Fragments of the Region Connection Calculus: A Complete Analysis

Jochen Renz
Institut für Informatik, Albert-Ludwigs-Universität
Am Flughafen 17, 79110 Freiburg, Germany

Abstract

We present a general method for proving tractability of reasoning over disjunctions of jointly exhaustive and pairwise disjoint relations. Examples of these kinds of relations are Allen's temporal interval relations and their spatial counterpart, the RCC8 relations by Randell, Cui, and Cohn. Applying this method does not require detailed knowledge about the considered relations; instead, it is rather sufficient to have a subset of the considered set of relations for which path-consistency is known to decide consistency. Using this method, we give a complete classification of tractability of reasoning over RCC8 by identifying two large new maximal tractable subsets and show that these two subsets together with $\widehat{\mathcal{H}}_8$, the already known maximal tractable subset, are the only such sets for RCC8 that contain all base relations. We also apply our method to Allen's interval algebra and derive the known maximal tractable subset.

1 Introduction

In qualitative spatial and temporal reasoning, knowledge is often represented by specifying the relationships between spatial or temporal entities. Of particular interest are disjunctions over a set of jointly exhaustive and pairwise disjoint (JEPD) relations. JEPD relations are also called *base relations*. Since any two entities are related by exactly one of the base relations, they can be used to represent definite information with respect to the given level of granularity. Indefinite information can be specified by disjunctions of possible base relations.

An important reasoning problem is deciding consistency of a set of constraints using these relations, which is in many cases NP-hard. Sometimes deciding consistency is tractable if only a subset of all possible disjunctions is used. If this subset contains all base relations, then instances of the NP-hard consistency problem can be solved by backtracking over tractable subinstances [Ladkin and Reinefeld, 1992; Nebel, 1997]. Larger tractable subsets often lead to more efficient solutions (cf. [Renz and Nebel, 1998]). The goal is to identify the boundary between tractable and NP-hard subsets, i.e., all maximal tractable subsets containing all base relations.

Two examples of these types of relations are Allen's interval algebra [Allen, 1983] mainly used for temporal reasoning and their spatial counterpart, Randell, Cui, and Cohn's [1992] Region Connection Calculus RCC-8. In the former case, the only maximal tractable subset containing all base relations has been identified [Nebel and Bürckert, 1995], in the latter case, only one maximal tractable subset has been identified so far [Renz and Nebel, 1999]. It was previously unknown whether there are others containing all base relations. For both subsets path-consistency is sufficient for deciding consistency. Tractability and sufficiency of path-consistency have been proven by reducing the consistency problem to a tractable propositional satisfiability problem which requires very detailed knowledge about the considered set of relations and complicated proofs.

We present a new method for proving tractability and showing sufficiency of path-consistency for this kind of problem which does not require detailed knowledge about the considered set of relations. Applying this method, we identify two large new maximal tractable subsets of RCC-8 and show that these subsets together with the already known maximal tractable subset are the only such sets for RCC-8 that contain all base relations. We consider only sets containing all base relations, since only these sets enable efficient backtracking algorithms.

The paper is structured as follows. In Section 2 we introduce RCC-8. In Section 3 we present the new method for proving tractability. In Section 4 we identify two large subsets of RCC-8 which are candidates for maximal tractable subsets, and apply our new method in Section 5 to show that both sets are tractable. In Section 6 the new method is used for finding a consistent scenario. In Section 7 we apply the method to Allen's interval algebra after which we discuss and summarize our results. Some of our proofs are computer assisted. The programs can be obtained from the author.

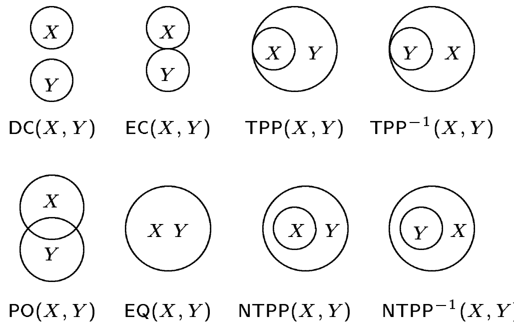

Figure 1: Examples for the eight base relations of RCC-8

2 The Region Connection Calculus RCC8

The Region Connection Calculus RCC-8 [Randell *et al.*, 1992] is a language for qualitative spatial representation and reasoning where *spatial regions* are regular subsets of a topological space. Regions themselves do not have to be internally connected, i.e., a region may consist of different disconnected pieces.

RCC-8 contains eight JEPD base relations and all 2^8 possible disjunctions thereof. The base relations are denoted by DC, EC, PO, EQ, TPP, NTPP, TPP^{-1}, and NTPP^{-1}, with the meaning of *DisConnected, Externally Connected, Partial Overlap, EQual, Tangential Proper Part, Non-Tangential Proper Part*, and their converses. Two-dimensional examples of the base relations are shown in Figure 1. In the following we write sets of base relations to denote disjunctions of base relations. The disjunction of all base relations is called the *universal relation*. The subset of RCC-8 consisting of the eight base relations is denoted by $\mathcal{B}$.

An important reasoning problem in this framework, denoted by RSAT, is deciding *consistency* of a set Θ of constraints of the form xRy, where R is an RCC-8 relation and x, y are spatial variables. The domain of spatial variables is the set of all spatial regions of the considered topological space. Θ is consistent if it has a *solution*, which is an assignment of spatial regions to the spatial variables of Θ in a way that all constraints are satisfied. When only relations of a specific subset $\mathcal{S} \subseteq$ RCC-8 are used in Θ, the corresponding reasoning problem is denoted by RSAT($\mathcal{S}$). RSAT is NP-complete in general. Consistency of Θ can be approximated by using an $O(n^3)$ time *path-consistency* algorithm, which makes Θ *path-consistent* by eliminating all the impossible labels (base relations) in every subset of constraints involving three variables [Mackworth, 1977]. If the empty relation occurs during this process, then Θ is inconsistent, otherwise the resulting set is path-consistent.

Renz and Nebel [1999] identified a tractable subset of RCC-8 (denoted by $\widehat{\mathcal{H}}_8$) containing all base relations which is maximal with respect to tractability, i.e., if any other RCC-8 relation is added to $\widehat{\mathcal{H}}_8$, the consistency problem becomes NP-complete. They further showed that enforcing path-consistency is sufficient for deciding RSAT($\widehat{\mathcal{H}}_8$). $\widehat{\mathcal{H}}_8$ contains 148 relations, i.e., about 60% of all RCC-8 relations.

3 A New Method for Proving Tractability of Sets of Relations

In this section we develop a new method for proving tractability of reasoning over disjunctions of a JEPD set of binary relations over a domain $\mathcal{D}$. Let $\mathcal{A}$ be a finite set of JEPD binary relations. We assume in the following that there is a relation $Id \in \mathcal{A}$ such that $Id(d, d)$ is satisfied for all $d \in D$. Id is called *identity relation*. The consistency problem CSPSAT($\mathcal{S}$) for sets $\mathcal{S} \subseteq 2^{\mathcal{A}}$ over a domain $\mathcal{D}$ is defined as follows [Renz and Nebel, 1999]:

Instance: A set V of variables over a domain D and a finite set Θ of binary constraints xRy, where $R \in \mathcal{S}$ and $x, y \in V$.

Question: Is there an instantiation of all variables in Θ such that all constraints are satisfied?

$\widehat{\mathcal{S}}$ denotes the closure of $\mathcal{S}$ under composition ($\circ$), intersection ($\cap$) and converse ($\smile$). CSPSAT($\widehat{\mathcal{S}}$) can be polynomially reduced to CSPSAT($\mathcal{S}$) if the universal relation is contained in $\mathcal{S}$ [Renz and Nebel, 1999]. Therefore, we consider in the following only sets $\mathcal{S} \subseteq 2^{\mathcal{A}}$ with $\mathcal{S} = \widehat{\mathcal{S}}$.

Before we present our new method for showing that CSPSAT($\mathcal{S}$) is tractable for a set $\mathcal{S}$, we define some terminology. A *refinement* of a constraint xRy is a constraint $xR'y$ such that $R' \subseteq R$. For instance, the constraint $x\{\mathsf{DC}, \mathsf{TPP}\}y$ is a refinement of the constraint $x\{\mathsf{DC}, \mathsf{EC}, \mathsf{PO}, \mathsf{TPP}\}y$. A refinement of a set of constraints Θ is a set of constraints Θ' that contains a refinement of every constraint of Θ. Every solution of Θ' is also a solution of Θ. A *consistent scenario* Θ_s of Θ is a refinement of Θ such that Θ_s is consistent and contains only constraints over relations of $\mathcal{A}$. We assume that a set of constraints Θ contains n ordered variables $x_1, \ldots, x_n$. The following definition will be central for our new method.

Definition 1 (reduction by refinement)
Let $\mathcal{S}, \mathcal{T} \subseteq 2^{\mathcal{A}}$. $\mathcal{S}$ can be reduced by refinement *to $\mathcal{T}$, if the following two conditions are satisfied:*

1. *for every relation $S \in \mathcal{S}$ there is a relation $T_S \in \mathcal{T}$ with $T_S \subseteq S$,*
2. *every path-consistent set Θ of constraints over $\mathcal{S}$ can be refined to a set Θ' by replacing $x_i S x_j \in \Theta$ with $x_i T_S x_j \in \Theta'$ for $i < j$,[1] such that enforcing path-consistency to Θ' does not result in an inconsistency.*

Lemma 2 *If path-consistency decides CSPSAT($\mathcal{T}$) for a set $\mathcal{T} \subseteq 2^{\mathcal{A}}$, and $\mathcal{S}$ can be reduced by refinement to $\mathcal{T}$, then path-consistency decides CSPSAT($\mathcal{S}$).*

Proof. Let Θ be a path-consistent set of constraints over $\mathcal{S}$. Since $\mathcal{S}$ can be reduced by refinement to $\mathcal{T}$, there is

[1]Note that we only refine constraints $x_i S x_j$ for $i < j$. This is no restriction, as by enforcing path-consistency the converse constraint $x_j S^{\smile} x_i$ will also be refined. Rather it offers the possibility of refining, e.g., converse relations to other than converse sub-relations, i.e., if, for instance, R is refined to r, we can refine $R^{\smile}$ to a relation other than $r^{\smile}$.

by definition a set Θ' of constraints over $\mathcal{T}$ which is a refinement of Θ such that enforcing path-consistency to Θ' does not result in an inconsistency. Path-consistency decides $\mathsf{CSPSAT}(\mathcal{T})$, so Θ' is consistent, and, hence, Θ is also consistent. ∎

Since path-consistency can be enforced in cubic time, it is sufficient for proving tractability of $\mathsf{CSPSAT}(\mathcal{S})$ to show that $\mathcal{S}$ can be reduced by refinement to a set $\mathcal{T}$ for which path-consistency decides $\mathsf{CSPSAT}(\mathcal{T})$. Note that for refining a constraint xSy ($S \in \mathcal{S}$) to a constraint xT_Sy ($T_S \in \mathcal{T}$), it is not required that T_S is also contained in $\mathcal{S}$. Thus, with respect to common relations, the two sets $\mathcal{S}$ and $\mathcal{T}$ are independent of each other.

We will now develop a method for showing that a set of relations $\mathcal{S} \subseteq 2^{\mathcal{A}}$ can be reduced by refinement to another set $\mathcal{T} \subseteq 2^{\mathcal{A}}$. In order to manage the different refinements, we introduce a *refinement matrix* that contains for every relation $S \in \mathcal{S}$ all specified refinements.

Definition 3 (refinement matrix) *A refinement matrix M of $\mathcal{S}$ has $|\mathcal{S}| \times 2^{|\mathcal{A}|}$ Boolean entries such that for $S \in \mathcal{S}$, $R \in 2^{\mathcal{A}}$, $M[S][R] = true$ only if $R \subseteq S$.*

For example, if the relation $\{\mathsf{DC}, \mathsf{EC}, \mathsf{PO}, \mathsf{TPP}\}$ is allowed to be refined only to the relations $\{\mathsf{DC}, \mathsf{TPP}\}$ and $\{\mathsf{DC}\}$, then $M[\{\mathsf{DC}, \mathsf{EC}, \mathsf{PO}, \mathsf{TPP}\}][R]$ is *true* only for $R = \{\mathsf{DC}, \mathsf{TPP}\}$ and for $R = \{\mathsf{DC}\}$ and $false$ for all other relations $R \in 2^{\mathcal{A}}$. M is called the *basic refinement matrix* if $M[S][R] = true$ if and only if $S = R$.

We propose the algorithm CHECK-REFINEMENTS (see Figure 2) which takes as input a set of relations $\mathcal{S}$ and a refinement matrix M of $\mathcal{S}$. This algorithm computes all possible path-consistent triples of relations R_{12}, R_{23}, R_{13} of $\mathcal{S}$ (step 4) and enforces path-consistency (using a standard procedure PATH-CONSISTENCY) to every refinement $R'_{12}, R'_{23}, R'_{13}$ for which $M[R_{ij}][R'_{ij}] = true$ for all $i, j \in \{1, 2, 3\}, i < j$ (steps 5,6). If one of these refinements results in the empty relation, the algorithm returns `fail` (step 7). Otherwise, the resulting relations $R''_{12}, R''_{23}, R''_{13}$ are added to M by setting $M[R_{ij}][R''_{ij}] = true$ for all $i, j \in \{1, 2, 3\}, i < j$ (step 8). This is repeated until M has reached a fixed point (step 9), i.e., enforcing path-consistency on any possible refinement does not result in new relations anymore. If no inconsistency is detected in this process, the algorithm returns `succeed`.

A similar algorithm, GET-REFINEMENTS, returns the revised refinement matrix if CHECK-REFINEMENTS returns `succeed` and the basic refinement matrix if CHECK-REFINEMENTS returns `fail`. Since $\mathcal{A}$ is a finite set of relations, M can be changed only a finite number of times, so both algorithms always terminate.

Lemma 4 *Let Θ be a path-consistent set of constraints over $\mathcal{S}$ and M a refinement matrix of $\mathcal{S}$. For every refinement Θ' of Θ with $x_iR'x_j \in \Theta'$ only if $x_iRx_j \in \Theta$, $i < j$, and $M[R][R'] = true$: if* CHECK-REFINEMENTS$(\mathcal{S}, M)$ *returns* `succeed`*, enforcing path-consistency to Θ' does not result in an inconsistency.*

Algorithm: CHECK-REFINEMENTS
Input: A set $\mathcal{S}$ and a refinement matrix M of $\mathcal{S}$.
Output: `fail` if the refinements specified in M can make a path-consistent triple of constraints over $\mathcal{S}$ inconsistent; `succeed` otherwise.

1. changes $\leftarrow$ *true*
2. *while* changes *do*
3. $\quad oldM \leftarrow M$
4. $\quad$ *for every* path-consistent triple $T = (R_{12}, R_{23}, R_{13})$ of relations over $\mathcal{S}$ *do*
5. $\quad\quad$ *for every* refinement $T' = (R'_{12}, R'_{23}, R'_{13})$ of T with $oldM[R_{12}][R'_{12}] = oldM[R_{23}][R'_{23}] = oldM[R_{13}][R'_{13}] = true$ *do*
6. $\quad\quad\quad T'' \leftarrow$ PATH-CONSISTENCY(T')
7. $\quad\quad\quad$ *if* $T'' = (R''_{12}, R''_{23}, R''_{13})$ contains the empty relation *then return* `fail`
8. $\quad\quad\quad$ *else do* $M[R_{12}][R''_{12}] \leftarrow true$, $M[R_{23}][R''_{23}] \leftarrow true$, $M[R_{13}][R''_{13}] \leftarrow true$
9. $\quad$ *if* $M = oldM$ *then* changes $\leftarrow false$
10. *return* `succeed`

Figure 2: Algorithm CHECK-REFINEMENTS

Proof. Suppose that CHECK-REFINEMENTS$(\mathcal{S}, M)$ returns `succeed` and GET-REFINEMENTS$(\mathcal{S}, M)$ returns the refinement matrix M'. Suppose further that enforcing path-consistency to Θ' results in an inconsistency which is detected first for the three variables x_a, x_b, x_c. Suppose that for every pair of variables x_i, x_j, $x_iR_{ij}x_j \in \Theta$ and $x_iR'_{ij}x_j \in \Theta'$. Enforcing path-consistency to Θ' can be done by successively enforcing path-consistency to every triple of variables of Θ'. Suppose that we start with the triple x_1, x_2, x_3. We have that $M'[R_{ij}][R'_{ij}] = true$ for every $i, j \in \{1, 2, 3\}, i < j$. After enforcing path-consistency to this triple we obtain the relations R''_{ij} for which again $M'[R_{ij}][R''_{ij}] = true$, otherwise CHECK-REFINEMENTS$(\mathcal{S}, M)$ would have returned `fail`. The same holds when we proceed with enforcing path-consistency to every triple of variables, we always end up with relations R^*_{ij} for which $M'[R_{ij}][R^*_{ij}] = true$. Therefore, for all possible relations $R^*_{ab}, R^*_{bc}, R^*_{ac}$ we can obtain by enforcing path-consistency to Θ' we have that $M'[R_{ij}][R^*_{ij}] = true$ for all $i, j \in \{a, b, c\}, i < j$. If this triple of relations were inconsistent, CHECK-REFINEMENTS$(\mathcal{S}, M)$ would have returned `fail`. ∎

If CHECK-REFINEMENTS returns `succeed` and GET-REFINEMENTS returns M', we have pre-computed all possible refinements of every path-consistent triple of variables as given in the refinement matrix M'. Thus, applying these refinements to a path-consistent set of constraints can never result in an inconsistency when enforcing path-consistency.

Theorem 5 *Let $\mathcal{S}, \mathcal{T} \subseteq 2^{\mathcal{A}}$, and let M be a refinement matrix of $\mathcal{S}$.* GET-REFINEMENTS$(\mathcal{S}, M)$ *returns the refinement matrix M'. If for every $S \in \mathcal{S}$ there is a $T_S \in \mathcal{T}$ with $M'[S][T_S] = true$, then $\mathcal{S}$ can be reduced by refinement to $\mathcal{T}$.*

Proof. By the given conditions, we can refine every path-consistent set Θ of constraints over $\mathcal{S}$ to a set Θ' of constraints over $\mathcal{T}$ such that $M'[R_{ij}][R'_{ij}] = true$ for every $x_i R_{ij} x_j \in \Theta$ and $x_i R'_{ij} x_j \in \Theta'$, $i < j$, $R'_{ij} \in \mathcal{T}$. M' is a fixed point with respect to GET-REFINEMENTS, i.e., GET-REFINEMENTS$(\mathcal{S}, M')$ returns M', thus it follows from Lemma 4 that enforcing path-consistency to Θ' does not result in an inconsistency. ∎

By Lemma 2 and Theorem 5 we have that the procedures CHECK-REFINEMENTS and GET-REFINEMENTS can be used to prove tractability for sets of relations.

Corollary 6 *Let $\mathcal{S}, \mathcal{T} \subseteq 2^{\mathcal{A}}$ be two sets such that path-consistency decides* $\mathsf{CSPSAT}(\mathcal{T})$*, and let M be a refinement matrix of $\mathcal{S}$.* GET-REFINEMENTS$(\mathcal{S}, M)$ *returns M'. If for every $S \in \mathcal{S}$ there is a $T_S \in \mathcal{T}$ with $M'[S][T_S] = true$, then path-consistency decides* $\mathsf{CSPSAT}(\mathcal{S})$*.*

In the following sections we apply this method to the Region Connection Calculus RCC-8 and to Allen's interval algebra for proving certain subsets to be tractable. For RCC-8 it will help us make a complete analysis of tractability by identifying two large new maximal tractable subsets. Later, it will turn out that the new method can be used for identifying the fastest possible algorithms for finding a consistent scenario.

4 Candidates for Maximal Tractable Subsets of RCC8

In order to identify maximal tractable subsets of a set of relations with an NP-hard consistency problem, two different tasks have to be done. On the one hand, tractable subsets have to be identified. On the other hand, NP-hard subsets have to be identified. Then, a tractable subset is maximal tractable if any superset is NP-hard. If tractability has not yet been shown for certain subsets, the NP-hardness results restrict the number of different subsets that still have to be analyzed for tractability.

In this section we present new NP-hardness results for RCC-8 and identify those subsets that are candidates for maximal tractable subsets, i.e., if any other relation is added to one of those candidates they become NP-hard. Before this, we summarize Renz and Nebel's [1999] NP-hardness proofs that were necessary for showing that $\widehat{\mathcal{H}}_8$ is a maximal tractable subset of RCC-8.

Lemma 7 (Renz and Nebel, 1999) $\mathsf{RSAT}(\mathcal{B} \cup \{R\})$ *is* NP*-hard if $R \in \mathcal{N} \subseteq$ RCC-8, where:*

$$\mathcal{N} = \{R \mid \{\mathsf{PO}\} \not\subseteq R \text{ and } \{\mathsf{(N)TPP}, \mathsf{(N)TPP}^{-1}\} \subseteq R\}.$$

(N)TPP *indicates either* TPP *or* NTPP.

Renz and Nebel [1999] proved maximality of $\widehat{\mathcal{H}}_8$ by showing that $\mathsf{RSAT}(\widehat{\mathcal{H}}_8 \cup \{\mathsf{EQ}, \mathsf{NTPP}\})$ is NP-hard and that the closure of $\widehat{\mathcal{H}}_8$ plus any relation of RCC-8 $\setminus (\widehat{\mathcal{H}}_8 \cup \mathcal{N})$ contains the relation $\{\mathsf{EQ}, \mathsf{NTPP}\}$. This NP-hardness proof, however, does not make use of all relations of $\widehat{\mathcal{H}}_8$, but only of the relations $\{\mathsf{EC}, \mathsf{TPP}\}$ and $\{\mathsf{EC}, \mathsf{NTPP}\}$. Thus, any set of RCC-8 relations that contains all base relations plus the relations $\{\mathsf{EC}, \mathsf{TPP}\}$, $\{\mathsf{EC}, \mathsf{NTPP}\}$, and $\{\mathsf{EQ}, \mathsf{NTPP}\}$ is NP-hard.

This result rules out a lot of subsets of RCC-8 to be tractable, but still leaves the problem of tractability open for a large number of subsets. We found, however, that it is not necessary that both relations, $\{\mathsf{EC}, \mathsf{TPP}\}$ and $\{\mathsf{EC}, \mathsf{NTPP}\}$, must be added to $\{\mathsf{EQ}, \mathsf{NTPP}\}$ in order to enforce NP-hardness. It is rather sufficient for NP-hardness that only one of them is added to $\{\mathsf{EQ}, \mathsf{NTPP}\}$.

This is shown in the following two lemmata. They are proven by reducing 3SAT, the NP-hard satisfiability problem of propositional formulas that contain exactly three literals per clause [Garey and Johnson, 1979], to the respective consistency problem. Both reductions are similar to the reductions given in [Renz and Nebel, 1999]: every literal as well as every literal occurrence L is reduced to a constraint $X_L R Y_L$ where $R = R_t \cup R_f$ and $R_t \cap R_f = \emptyset$. L is true iff $X_L R_t Y_L$ holds and false iff $X_L R_f Y_L$ holds. "Polarity constraints" enforce that $X_{\neg L} R_t Y_{\neg L}$ holds iff $X_L R_f Y_L$ holds, and *vice versa.* "Clause constraints" enforce that at least one literal of every clause is true.

Lemma 8 *If* $(\mathcal{B} \cup \{\{\mathsf{EQ}, \mathsf{NTPP}\}, \{\mathsf{EC}, \mathsf{NTPP}\}\}) \subseteq \mathcal{S}$*, then* $\mathsf{RSAT}(\mathcal{S})$ *is* NP*-complete.*

Proof Sketch. Transformation of 3SAT to $\mathsf{RSAT}(\mathcal{S})$. $R_t = \{\mathsf{NTPP}\}$ and $R_f = \{\mathsf{EQ}\}$. Polarity constraints:

$$X_L\{\mathsf{EC}\}X_{\neg L}, Y_L\overline{\{\mathsf{DC}, \mathsf{EC}\}}Y_{\neg L},$$
$$X_L\{\mathsf{EC}, \mathsf{NTPP}\}Y_{\neg L}, X_{\neg L}\{\mathsf{EC}, \mathsf{NTPP}\}Y_L.$$

Clause constraints for each clause $c = \{i, j, k\}$: $Y_i\{\mathsf{NTPP}^{-1}\}X_j, Y_j\{\mathsf{NTPP}^{-1}\}X_k, Y_k\{\mathsf{NTPP}^{-1}\}X_i$. $\overline{\{\mathsf{DC}, \mathsf{EC}\}}$ is contained in $\widehat{\mathcal{S}}$. ∎

Lemma 9 *If* $(\mathcal{B} \cup \{\{\mathsf{EQ}, \mathsf{NTPP}\}, \{\mathsf{EC}, \mathsf{TPP}\}\}) \subseteq \mathcal{S}$*, then* $\mathsf{RSAT}(\mathcal{S})$ *is* NP*-complete.*

Proof Sketch. Transformation of 3SAT to $\mathsf{RSAT}(\mathcal{S})$. $R_t = \{\mathsf{NTPP}\}$ and $R_f = \{\mathsf{EQ}\}$. Polarity constraints:

$$X_L\{\mathsf{DC}, \mathsf{EC}, \mathsf{TPP}\}Y_{\neg L}, X_L\{\mathsf{DC}, \mathsf{NTPP}^{-1}\}X_{\neg L},$$
$$Y_L\{\mathsf{EC}, \mathsf{TPP}\}Y_{\neg L}, X_{\neg L}\{\mathsf{DC}, \mathsf{EC}, \mathsf{PO}, \mathsf{TPP}, \mathsf{NTPP}\}Y_L$$

Clause constraints for each clause $c = \{i, j, k\}$: $Y_i\{\mathsf{NTPP}^{-1}\}X_j, Y_j\{\mathsf{NTPP}^{-1}\}X_k, Y_k\{\mathsf{NTPP}^{-1}\}X_i$. All the above relations are contained in $\widehat{\mathcal{S}}$. ∎

Using these results, four other relations can be ruled out to be contained in any tractable subset of RCC-8, since the closure of the base relations plus one of the four relations contains $\{\mathsf{EQ}, \mathsf{NTPP}\}$ as well as $\{\mathsf{EC}, \mathsf{TPP}\}$.

Definition 10 *The two sets $\mathcal{NP}_8$ and $\mathcal{P}_8$ are defined as follows:*

- $\mathcal{NP}_8 = \mathcal{N} \cup \{\{\mathsf{EC}, \mathsf{NTPP}, \mathsf{EQ}\}, \{\mathsf{EC}, \mathsf{NTPP}^{-1}, \mathsf{EQ}\}, \{\mathsf{DC}, \mathsf{EC}, \mathsf{NTPP}, \mathsf{EQ}\}, \{\mathsf{DC}, \mathsf{EC}, \mathsf{NTPP}^{-1}, \mathsf{EQ}\}\}$.
- $\mathcal{P}_8 = \text{RCC-8} \setminus \mathcal{NP}_8$.

$\mathcal{NP}_8$ contains 76 relations, $\mathcal{P}_8$ contains 180 relations.

Corollary 11 $\mathsf{RSAT}(\mathcal{B} \cup \{R\})$ *is* NP*-hard, if $R \in \mathcal{NP}_8$.*

Lemmata 8 and 9 give us a sufficient but not necessary condition of whether a subset of $\mathcal{P}_8$ is NP-hard, namely, if $\mathcal{B} \subseteq \widehat{\mathcal{S}} \subseteq \mathcal{P}_8$ contains $\{\mathsf{EQ}, \mathsf{NTPP}\}$ and one of $\{\mathsf{EC}, \mathsf{TPP}\}$ or $\{\mathsf{EC}, \mathsf{NTPP}\}$, $\mathsf{RSAT}(\mathcal{S})$ is NP-hard, otherwise the complexity of $\mathsf{RSAT}(\mathcal{S})$ remains open.

We computed all subsets of $\mathcal{P}_8$ that are candidates for maximal tractable subsets with respect to the above two NP-hardness proofs, i.e., we enumerated all subsets $\mathcal{S} \subseteq \mathcal{P}_8$ with $\mathcal{S} = \widehat{\mathcal{S}}$ that contain all base relations and the universal relation, and checked whether they fulfilled the following two properties:

1. $\mathcal{S}$ does not contain ($\{\mathsf{EQ}, \mathsf{NTPP}\}$ and $\{\mathsf{EC}, \mathsf{TPP}\}$) or ($\{\mathsf{EQ}, \mathsf{NTPP}\}$ and $\{\mathsf{EC}, \mathsf{NTPP}\}$), and
2. the closure of $\mathcal{S}$ plus any relation of $\mathcal{P}_8 \setminus \mathcal{S}$ contains ($\{\mathsf{EQ}, \mathsf{NTPP}\}$ and $\{\mathsf{EC}, \mathsf{TPP}\}$) or ($\{\mathsf{EQ}, \mathsf{NTPP}\}$ and $\{\mathsf{EC}, \mathsf{NTPP}\}$).

To our surprise, this resulted in only three different subsets of $\mathcal{P}_8$, of which $\widehat{\mathcal{H}}_8$ is of course one of them. The other two subsets are denoted by $\mathcal{C}_8$ and $\mathcal{Q}_8$.[2]

$$\begin{aligned}
\widehat{\mathcal{H}}_8 &= \mathcal{P}_8 \setminus \{R \mid (\{\mathsf{EQ}, \mathsf{NTPP}\} \subseteq R \text{ and } \{\mathsf{TPP}\} \not\subseteq R) \\
&\qquad \text{or } (\{\mathsf{EQ}, \mathsf{NTPP}^{-1}\} \subseteq R \text{ and } \{\mathsf{TPP}^{-1}\} \not\subseteq R)\} \\
\mathcal{C}_8 &= \mathcal{P}_8 \setminus \{R \mid \{\mathsf{EC}\} \subset R \text{ and } \{\mathsf{PO}\} \not\subseteq R \text{ and} \\
&\qquad R \cap \{\mathsf{TPP}, \mathsf{NTPP}, \mathsf{TPP}^{-1}, \mathsf{NTPP}^{-1}, \mathsf{EQ}\} \neq \emptyset\} \\
\mathcal{Q}_8 &= \mathcal{P}_8 \setminus \{R \mid \{\mathsf{EQ}\} \subset R \text{ and } \{\mathsf{PO}\} \not\subseteq R \text{ and} \\
&\qquad R \cap \{\mathsf{TPP}, \mathsf{NTPP}, \mathsf{TPP}^{-1}, \mathsf{NTPP}^{-1}\} \neq \emptyset\}
\end{aligned}$$

$\mathcal{C}_8$ contains 158 different RCC-8 relations, $\mathcal{Q}_8$ contains 160 different relations. We have that $\widehat{\mathcal{H}}_8 \cup \mathcal{C}_8 = \mathcal{P}_8$.

Lemma 12 *For every subset $\mathcal{S}$ of* RCC-8*: If* $\mathsf{RSAT}(\mathcal{S})$ *is tractable, then $\mathcal{S} \subseteq \widehat{\mathcal{H}}_8$, $\mathcal{S} \subseteq \mathcal{C}_8$, or $\mathcal{S} \subseteq \mathcal{Q}_8$.*

Proof. By computing the closure of every set containing the base relations, the universal relation, and two arbitrary RCC-8 relations, the resulting set is either contained in one of the sets $\widehat{\mathcal{H}}_8$, $\mathcal{C}_8$, or $\mathcal{Q}_8$, or is NP-hard according to Lemma 8, Lemma 9, or Corollary 11. ■

So far we did not say anything about tractability of $\mathcal{C}_8$ or $\mathcal{Q}_8$ or subsets thereof. It might be that both sets are NP-hard or that there is a large number of different maximal tractable subsets that are contained in the two sets. What we obtained so far is, thus, only a restriction of the number of different subsets we have to check for tractability. However, we will see in the next section that both $\mathcal{C}_8$ and $\mathcal{Q}_8$ are in fact tractable.

5 A Complete Analysis of Tractability

In [Renz and Nebel, 1999] tractability of $\widehat{\mathcal{H}}_8$ was shown by reducing $\mathsf{RSAT}(\widehat{\mathcal{H}}_8)$ to HORNSAT. This was possible because $\widehat{\mathcal{H}}_8$ contains exactly those relations that are transformed to propositional Horn formulas when using the propositional encoding of RCC-8. This propositional Horn encoding of $\widehat{\mathcal{H}}_8$ was also used for proving that path-consistency decides $\mathsf{RSAT}(\widehat{\mathcal{H}}_8)$ [Renz and Nebel, 1999], by relating positive unit resolution, which is a complete decision method for propositional Horn formulas, to path-consistency. The propositional encoding of $\mathcal{C}_8$ and $\mathcal{Q}_8$ is neither a Horn formula nor a Krom formula (two literals per clause), so it is not immediately possible to reduce the consistency problem of these subsets to a tractable propositional satisfiability problem. Therefore we have to find other ways of proving tractability.

Let us have a closer look at the relations of the two sets $\mathcal{C}_8$ and $\mathcal{Q}_8$ and how they differ from $\widehat{\mathcal{H}}_8$.

Proposition 13 *For all relations $R \in \mathcal{P}_8 \setminus \widehat{\mathcal{H}}_8$ we have that $\{\mathsf{EQ}\} \subset R$. For all refinements R' of R with $R' \cup \{\mathsf{EQ}\} = R$ and $\{\mathsf{EQ}\} \not\subseteq R'$ we have that $R' \in \widehat{\mathcal{H}}_8$.*

$\mathcal{C}_8$ and $\mathcal{Q}_8$ are both subsets of $\mathcal{P}_8$, so if we can prove the following conjecture, then both sets are tractable:

Conjecture 14 *Let Θ be a path-consistent set of constraints over $\mathcal{C}_8$ or over $\mathcal{Q}_8$. If Θ' is obtained from Θ by eliminating the identity relation $\{\mathsf{EQ}\}$ from every constraint $xRy \in \Theta$ with $R \in \mathcal{P}_8 \setminus \widehat{\mathcal{H}}_8$, then enforcing path-consistency to Θ' does not result in an inconsistency.*

If we can prove this conjecture, then, by Proposition 13, Θ' contains only constraints over $\widehat{\mathcal{H}}_8$. Since path-consistency decides $\mathsf{RSAT}(\widehat{\mathcal{H}}_8)$, we then have that if Θ is path-consistent, Θ' is consistent and, therefore, Θ is also consistent, i.e., path-consistency decides consistency for sets of constraints over $\mathcal{C}_8$ and over $\mathcal{Q}_8$. In [Gerevini and Renz, 1998] it was shown that the relation $\{\mathsf{EQ}\}$ can always be eliminated from every constraint of a path-consistent set of constraints over $\widehat{\mathcal{H}}_8$ without changing consistency of the set. This was, again, shown by applying positive unit resolution to the propositional Horn encoding of $\widehat{\mathcal{H}}_8$, a method which is not applicable in our case. Instead, we can now apply the new method developed in Section 3, namely, we can check whether $\mathcal{C}_8$ and $\mathcal{Q}_8$ can be reduced by refinement to $\widehat{\mathcal{H}}_8$. Conjecture 14 gives the refinement matrix we have to check. We define this particular refinement matrix for arbitrary sets $\mathcal{S}$ of disjunctions over a set $\mathcal{A}$ of JEPD relations:

Definition 15 (identity-refinement matrix) *$M^{\neq}$ is the* identity-refinement matrix *of a set $\mathcal{S} \subseteq 2^{\mathcal{A}}$ if for every $S \in \mathcal{S}$, $M^{\neq}[S][S'] = true$ iff $S' = S \setminus Id$, where $Id \in \mathcal{A}$ is the identity relation.*

Proposition 16

- CHECK-REFINEMENTS$(\mathcal{C}_8, M^{\neq})$ *returns* `succeed`.
- CHECK-REFINEMENTS$(\mathcal{Q}_8, M^{\neq})$ *returns* `succeed`.

Theorem 17 *Path-consistency decides* $\mathsf{RSAT}(\mathcal{C}_8)$ *as well as* $\mathsf{RSAT}(\mathcal{Q}_8)$.

Proof. It follows from Proposition 13 and Proposition 16 that Theorem 5 can be applied with $\mathcal{T} = \widehat{\mathcal{H}}_8$. Since path-consistency decides $\mathsf{RSAT}(\widehat{\mathcal{H}}_8)$, it follows

[2]The names were chosen, since all $\mathcal{P}_8$-relations not contained in $\mathcal{C}_8$ contain E<u>C</u>, and all $\mathcal{P}_8$-relations not contained in $\mathcal{Q}_8$ contain E<u>Q</u>.

from Corollary 6 that path-consistency decides $\mathsf{RSAT}(\mathcal{C}_8)$ as well as $\mathsf{RSAT}(\mathcal{Q}_8)$ ■

Together with Lemma 12 it follows that $\widehat{\mathcal{H}}_8$, $\mathcal{C}_8$, and $\mathcal{Q}_8$ are the only maximal tractable subsets of RCC-8 that contain all base relations.

Theorem 18 *For every subset $\mathcal{S}$ of* RCC-8 *that contains all base relations and the universal relation:* $\mathsf{RSAT}(\mathcal{S})$ *is tractable iff* $\mathcal{S} \subseteq \widehat{\mathcal{H}}_8$, $\mathcal{S} \subseteq \mathcal{C}_8$, *or* $\mathcal{S} \subseteq \mathcal{Q}_8$.

6 Finding a Consistent Scenario

Gerevini and Renz [1998] gave an $O(n^3)$ time algorithm for computing a consistent scenario for a set of constraints over $\widehat{\mathcal{H}}_8$. Their algorithm is based on first eliminating $\{\mathsf{EQ}\}$ from every constraint and then successively refining constraints to constraints over base relations in a particular order and enforcing path-consistency in between. As shown in the previous section it is also possible to eliminate $\{\mathsf{EQ}\}$ from all constraints over $\mathcal{C}_8$ and over $\mathcal{Q}_8$, so it is possible to apply Gerevini and Renz's algorithm also for $\mathcal{C}_8$ and $\mathcal{Q}_8$. By applying the method developed in Section 3 we can, however, improve this algorithm for all three maximal tractable subsets of RCC-8.

Lemma 19 $\widehat{\mathcal{H}}_8$, $\mathcal{C}_8$, *and* $\mathcal{Q}_8$ *can be reduced by refinement to* $\mathcal{B}$, *the set of all* RCC-8 *base relations.*

This gives us the possibility of computing a consistent scenario of a path-consistent set of constraints over a tractable subset by a simple table lookup scheme, which is the fastest possible way.

Lemma 20 *Let $\mathcal{S}$ be one of* $\widehat{\mathcal{H}}_8$, $\mathcal{C}_8$, $\mathcal{Q}_8$. *For every relation* $R \in \mathcal{S}$, *base*(R) *is the following base relation:*

(1) If $R \in \mathcal{B}$, *then base*$(R) = R$;

(2) else if $\{\mathsf{DC}\} \subseteq R$, *then base*$(R) = \{\mathsf{DC}\}$;

(3) else if $\{\mathsf{EC}\} \subseteq R$ *and* $\mathcal{S} = \mathcal{Q}_8$ *or* $\mathcal{S} = \widehat{\mathcal{H}}_8$, *then base*$(R) = \{\mathsf{EC}\}$;

(4) else if $\{\mathsf{PO}\} \subseteq R$, *then base*$(R) = \{\mathsf{PO}\}$;

(5) else if $\{\mathsf{NTPP}\} \subseteq R$ *and* $\mathcal{S} = \mathcal{C}_8$, *then base*$(R) = \{\mathsf{NTPP}\}$;

(6) else if $\{\mathsf{TPP}\} \subseteq R$, *then base*$(R) = \{\mathsf{TPP}\}$;

(7) else base$(R) = $ *base*$(R^{\smile})$.

Theorem 21 *A consistent scenario* Θ_s *of a path-consistent set* Θ *of constraints over* $\widehat{\mathcal{H}}_8$, $\mathcal{C}_8$, *or over* $\mathcal{Q}_8$ *can be computed in* $O(n^2)$ *time, by replacing every constraint* $xRy \in \Theta$ *with* $xR'y \in \Theta_s$, *where* $R' = $ *base*(R).

7 Applying the New Method to Allen's Interval Algebra

The method developed in Section 3 can also be applied to Allen's interval algebra [Allen, 1983] which consists of the 13 JEPD relations *before* ($<$), *meets* (m), *overlaps* (o), *starts* (s), *during* (d), *finishes* (f), their converse relations $>$, mi, oi, si, di, fi, and the identity relation *equal* ($=$) that describe the possible relationships between two convex intervals. The full algebra consists of the 2^{13} possible disjunctions of the base relations and has an NP-complete consistency problem. Tractable subclasses of the interval algebra for which path-consistency decides consistency were identified by Vilain et al. [1989], the "Pointisable" subclass $\mathcal{P}$ (about 2% of the full algebra), and by Nebel and Bürckert [1995], the "ORD-Horn" subclass $\mathcal{H}$ (about 10% of the full algebra) which is the only maximal tractable subclass of the interval algebra that contains all base relations.

Since the only maximal tractable subclass of the interval algebra containing all base relations has already been identified, we cannot provide any new results on that topic. We can, however, validate the usefulness of our new method by showing that tractability of the ORD-Horn subclass and sufficiency of path-consistency for deciding consistency for this subclass can also be proven using our method. For this we apply the same strategy as for $\mathcal{C}_8$ and $\mathcal{Q}_8$, namely, we use the identity-refinement matrix $M^{\neq}$ (cf. [Ligozat, 1996]).

Proposition 22 GET-REFINEMENTS$(\mathcal{H}, M^{\neq})$ *returns* $M^{\mathcal{P}}$ *which has the following property: For every* $R \in \mathcal{H}$ *there is an* $R' \in \mathcal{P}$ *such that* $M^{\mathcal{P}}[R][R'] = $ *true.*

This was computed in less than 25 minutes on a Sun Sparc Ultra1. We can now apply Theorem 5:

Theorem 23 $\mathcal{H}$ *can be reduced by refinement to* $\mathcal{P}$.

Not all of the 188 pointisable relations are necessary for this refinement, but only 30 of them which are obtained according to the following refinement scheme.

Lemma 24 *For every relation* $R \in \mathcal{H}$, *point*(R) *is the following pointisable relation where* $R^{\neq} = R \setminus \{=\}$ *and* $R' = R \cap \{<, >, o, oi, d, di\}$:

(1) If $R = \{=\}$, *then point*$(R) = R$;

(2) else if $R' \neq \{\}$, *then point*$(R) = R'$;

(3) else point$(R) = R^{\neq}$.

Using this refinement scheme, every path-consistent set Θ of constraints over ORD-Horn can be refined to a path-consistent set Θ' by replacing every constraint $xRy \in \Theta$ with $xR'y \in \Theta'$, where $R' = $ *point*(R). This can be useful for some tasks such as finding a consistent scenario (cf. [Gerevini and Cristani, 1997]).

8 Discussion & Further Work

The method we developed for proving tractability of reasoning over sets of relations is very simple, does not require detailed knowledge about the considered relations, and seems to be very powerful. The only difficulty of this method is finding an adequate refinement matrix. However, the simple heuristic of eliminating all identity relations was successful for all maximal tractable fragments of RCC-8 and Allen's interval algebra which contain all base relations. This leads to an interesting question, namely, whether it is a general property of sets of relations containing all base relations for which path-consistency decides consistency that the identity relation can be eliminated from all constraints of a path-consistent set without making the set inconsistent.

The complete analysis of tractability for RCC-8 gives the possibility to develop more efficient algorithms for deciding consistency. As Renz and Nebel [1998] have shown in an empirical study, running different strategies in parallel is very effective; almost all apparently hard instances of the phase-transition region up to a certain size were solved in a few seconds. The two maximal tractable subsets of RCC-8 identified in this paper suggest several new strategies that should be empirically studied. Both subsets are larger than $\widehat{\mathcal{H}}_8$, but their average branching factor is higher ($\mathcal{C}_8$: 1.523, $\mathcal{Q}_8$: 1.516, $\widehat{\mathcal{H}}_8$: 1.438).

9 Summary

We developed a general method for proving sufficiency of path-consistency for deciding consistency of disjunctions over a set of JEPD relations. We applied this method to the Region Connection Calculus RCC-8 and identified two large new maximal tractable subsets. Together with $\widehat{\mathcal{H}}_8$, the already known maximal tractable subset, these are the only such sets for RCC-8 that contain all base relations and can, hence, be used to increase efficiency of backtracking algorithms for reasoning over NP-hard subsets. We also applied our method to Allen's interval algebra which resulted in a simple proof of tractability of reasoning in the ORD-Horn subclass.

Acknowledgments

This research was supported by DFG as part of the project FAST-QUAL-SPACE, which is part of the DFG special research effort on "Spatial Cognition". I want to thank Christoph Dornheim and Bernhard Nebel for their helpful comments.

References

[Allen, 1983] James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.

[Garey and Johnson, 1979] Michael R. Garey and David S. Johnson. *Computers and Intractability—A Guide to the Theory of NP-Completeness.* Freeman, San Francisco, CA, 1979.

[Gerevini and Cristani, 1997] Alfonso Gerevini and Matteo Cristani. On finding a solution in temporal constraint satisfaction problems. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence*, pages 1460–1465, Nagoya, Japan, 1997.

[Gerevini and Renz, 1998] Alfonso Gerevini and Jochen Renz. Combining topological and qualitative size constraints for spatial reasoning. In *Proceedings of the 4th International Conference on Principles and Practice of Constraint Programming*, Pisa, Italy, 1998. Springer.

[Ladkin and Reinefeld, 1992] Peter B. Ladkin and Alexander Reinefeld. Effective solution of qualitative interval constraint problems. *Artificial Intelligence*, 57(1):105–124, 1992.

[Ligozat, 1996] Gerard Ligozat. A new proof of tractability for ORD-Horn relations. In *Proceedings of the 13th National Conference of the American Association for Artificial Intelligence*, pages 715–720, Portland, OR, August 1996. MIT Press.

[Mackworth, 1977] Alan K. Mackworth. Consistency in networks of relations. *Artificial Intelligence*, 8:99–118, 1977.

[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 Association for Computing Machinery*, 42(1):43–66, 1995.

[Nebel, 1997] Bernhard Nebel. Solving hard qualitative temporal reasoning problems: Evaluating the efficiency of using the ORD-Horn class. *CONSTRAINTS*, 3(1):175–190, 1997.

[Randell *et al.*, 1992] David A. Randell, Zhan Cui, and Anthony G. Cohn. A spatial logic based on regions and connection. In *Principles of Knowledge Representation and Reasoning: Proceedings of the 3rd International Conference*, pages 165–176, Cambridge, MA, 1992. Morgan Kaufmann.

[Renz and Nebel, 1998] Jochen Renz and Bernhard Nebel. Efficient methods for qualitative spatial reasoning. In *Proceedings of the 13th European Conference on Artificial Intelligence*, pages 562–566, Brighton, UK, August 1998. Wiley.

[Renz and Nebel, 1999] Jochen Renz and Bernhard Nebel. On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the Region Connection Calculus. *Artificial Intelligence*, 108(1-2):69–123, 1999.

[Vilain *et al.*, 1989] Marc B. Vilain, Henry A. Kautz, and Peter G. van Beek. Constraint propagation algorithms for temporal reasoning: A revised report. In D. S. Weld and J. de Kleer, editors, *Readings in Qualitative Reasoning about Physical Systems*, pages 373–381. Morgan Kaufmann, San Mateo, CA, 1989.