ECAI 2012
Luc De Raedt et al. (Eds.)

doi:10.3233/978-1-61499-098-7-552

Here, There, but Not Everywhere: An Extended Framework for Qualitative Constraint Satisfaction

Weiming Liu and **Sanjiang Li**[1]

Abstract. Dealing with spatial and temporal knowledge is an indispensable part of almost all aspects of human activities. The qualitative approach to spatial and temporal reasoning (QSTR) provides a promising framework for spatial and temporal knowledge representation and reasoning. QSTR typically represents spatial/temporal knowledge in terms of qualitative relations (e.g., *to the east of*, *after*), and reasons with the knowledge by solving qualitative constraints. When formulating a qualitative constraint satisfaction problem (CSP), it is usually assumed that each variable could be "*here, there and everywhere*[2]." Practical applications e.g. urban planning, however, often require a variable taking values from a certain finite subset of the universe, i.e. require it to be '*here or there*'. This paper extends the classic framework of qualitative constraint satisfaction by allowing variables taking values from finite domains. The computational complexity of this extended consistency problem is examined for five most important qualitative calculi, viz. Point Algebra, Interval Algebra, Cardinal Relation Algebra, RCC-5, and RCC-8. We show that the extended consistency problem remains in NP, but when only basic constraints are considered, the extended consistency problem for each calculus except Point Algebra is already NP-hard.

1 INTRODUCTION

Spatial and temporal information is pervasive and increasingly involved in our everyday life. Many tasks in real or virtual world require sophisticated spatial and temporal reasoning abilities. Rapid progress in science and technology in this century presents new challenges for spatial and temporal reasoning. Taking spatial information as an example, on one hand, people now can easily acquire location information with the help of GPS-enabled mobile equipment and web GISs such as Google Maps. This has greatly increased the public's demands for location-based services. On the other hand, the development of technologies (such as remote sensing, medical imaging, and sensor networks) has brought us huge volumes of spatial data, which makes the phenomenon of '*rich data but poor knowledge*' particularly serious in the area of spatial knowledge management.

The qualitative approach to spatial and temporal reasoning (QSTR) has the potential to resolve the conflict between data and knowledge. This is because the main aims of QSTR research are to design (i) human comprehensible and cognitively plausible spatial logics (or query languages); and (ii) efficient algorithms for consistency checking (or query preprocessing). For intelligent systems, the ability to understand and process qualitative, vague or even inconsistent (textual, graphical or speech) information collected from human beings or the Web is very important. This is because 'the input and the output of spatial processes is often qualitative rather than quantitative' [11]. Typically, QSTR represents spatial/temporal information in terms of human comprehensible qualitative predicates (e.g. *partially overlaps*, *west of*, *after*) and reduces spatial reasoning to constraint satisfaction problems (CSPs). In the past three decades, QSTR has made significant progress, and prominent relation models such as the Interval Algebra [1] and the Region Connection Calculus RCC-8 [12] have been applied in areas such as natural language processing, geographical information systems, robotics, content-based image retrieval (see e.g. [2]).

But there is a growing consensus that breakthroughs are necessary in order to bring spatial/temporal reasoning theory closer to practical applications. One reason might be that the current qualitative reasoning scheme uses a rather restricted constraint language: constraints in a qualitative CSP are always taken from the *same* calculus and only relate variables from the *same* domain. This is quite undesirable, as constraints involving restricted variables and/or multiple aspects of information frequently appear in practical tasks such as urban planning and spatial query processing.

Consider the following example. Suppose you are recommended a restaurant in Sydney by a friend who had dinner there before. The spatial information about the restaurant may be like "it is *in* downtown and *close to* a MacDonald, and it is to the *west of* or *the southwest of* Central Station." In this example, topological, directional, and distance information appears together. Furthermore, while the position of the restaurant is kind of totally unknown, the position of Central Station is fixed as a landmark, and the position of downtown is also fixed somehow, but the position of "MacDonald" is only *finitely fixed* because there are several MacDonalds in Sydney downtown.

In this paper, we say a variable is *finitely restricted* if it can only take values from a finite subset of the universe in a qualitative calculus. While some recent works have considered how to reason with qualitative constraints from different calculi [4, 5, 7, 16], the importance of solving constraints that involve restricted variables has been totally neglected. The recent work "*solving qualitative constraints involving landmarks*" [8] is the first step toward this direction, where a *landmark* is interpreted as an outstanding element in the universe. In other words, a landmark is a restricted variable that can only take one value from the universe.

This paper aims to extend the qualitative CSP framework by allowing variables to be finitely restricted. In such a qualitative CSP, the constraints are taken from a given qualitative calculus but could be non-basic relations, and the domain of each variable is either the universe of the calculus or a (nonempty) finite subset of the universe. An important question is, *how does this extension affect the computational complexity of deciding the consistency of qualitative CSPs*?

[1] Centre for Quantum Computation & Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology Sydney, Autralia, emails: weiming.liu-1@student.uts.edu.au, sanjiang.li@uts.edu.au

[2] A song by The Beatles from the album Revolver.

This paper examines the effect for several most important qualitative calculi, viz. Point Algebra (PA) [15], Interval Algebra (IA) [1], Cardinal Relation Algebra (CRA) [6], RCC-5, and RCC-8 [12]. We show that the extended consistency problem remains in NP for each calculus, but even when only basic constraints are considered, the problem for each calculus except PA is already NP-hard.

The remainder of this paper proceeds as follows. Section 2 introduces basic notions in qualitative constraint solving as well as the five qualitative calculi discussed in this paper. The extended qualitative CSP framework is also presented here. Section 3 discusses the computational complexity of reasoning with Point Algebra, showing that the problem is NP-complete in general but is tractable if only basic networks are considered. Sections 4 and 5 prove that the extended consistency problems for basic networks in Cardinal Relation Algebra, Interval Algebra, and RCC-5 are all NP-complete. The last section concludes the paper.

2 PRELIMINARIES

In this section, we first recall several well-know qualitative calculi and basic notions in qualitative constraint solving, and then introduce the extended qualitative CSP framework.

2.1 Qualitative calculi

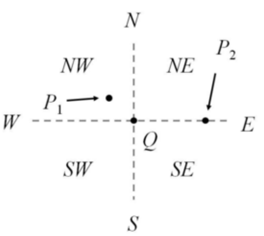
The qualitative approach to spatial and temporal knowledge representation and reasoning is mainly based on qualitative calculi. Suppose U is the universe of spatial or temporal entities. Write $\mathbf{Rel}(U)$ for the Boolean algebra of binary relations on U. This paper is concerned with binary *qualitative calculi* which are just finite Boolean subalgebras of $\mathbf{Rel}(U)$. For a relation α in a qualitative calculus $\mathcal{M}$, we call α a *basic* relation in $\mathcal{M}$ if it is an atom in $\mathcal{M}$.

We next recall the well-known Point Algebra (PA) [15, 14], Cardinal Relation Algebra (CRA) [3, 6], Interval Algebra (IA) [1], and RCC-5 and RCC-8 [12].

Definition 1 (Point Algebra [15])**.** Let U be the set of real numbers. The Point Algebra is the Boolean subalgebra generated by the jointly exhaustive and pairwise disjoint (JEPD) set of relations $\{<,>,=\}$, where $<,>,=$ are defined as usual.

PA contains eight relations, viz. the three basic relations $<,>,=$, the empty relation, and the four non-basic nonempty relations $\leq,\geq,\neq,?$, where $?$ stands for the universal relation.

Definition 2 (Cardinal Relation Algebra [3, 6])**.** Let U be the real plane. Define binary relations $NW, N, NE, W, EQ, E, SW, S, SE$ as in Table 1. The Cardinal Relation Algebra (CRA) is generated by these nine JEPD relations.



Relation	Definition
NW	$x < x', y > y'$
N	$x = x', y > y'$
NW	$x > x', y > y'$
W	$x < x', y = y'$
EQ	$x = x', y = y'$
E	$x > x', y = y'$
SW	$x < x', y < y'$
S	$x = x', y < y'$
SW	$x > x', y < y'$

Table 1. Definitions and illustrations of basic relations of Cardinal Relation Algebra, where in the left figure we have P_1 NW Q and P_2 E Q

The Cardinal Relation Algebra can be viewed as the Cartesian product of two Point Algebras.

Definition 3 (Interval Algebra [1])**.** Let U be the set of closed intervals on the real line. Thirteen binary relations between two intervals $x = [x^-, x^+]$ and $y = [y^-, y^+]$ are defined by the order of the four endpoints of x and y, see Table 2. The Interval Algebra is generated by these JEPD relations.

Relation	Symbol	Converse	Meaning
before	p	pi	$x^- < x^+ < y^- < y^+$
meets	m	mi	$x^- < x^+ = y^- < y^+$
overlaps	o	oi	$x^- < y^- < x^+ < y^+$
starts	s	si	$x^- = y^- < x^+ < y^+$
during	d	di	$y^- < x^- < x^+ < y^+$
finishes	f	fi	$y^- < x^- < x^+ = y^+$
equals	eq	eq	$x^- = y^- < x^+ = y^+$

Table 2. Basic IA relations and their converses, where $x = [x^-, x^+], y = [y^-, y^+]$ are two intervals.

Definition 4 (RCC-5 and RCC-8 Algebras[3])**.** Let U be the set of nonempty regular closed sets, or *regions*, in the real plane. The RCC-8 algebra is generated by the eight topological relations

$$\mathbf{DC}, \mathbf{EC}, \mathbf{PO}, \mathbf{EQ}, \mathbf{TPP}, \mathbf{NTPP}, \mathbf{TPP}^\sim, \mathbf{NTPP}^\sim,$$

where $\mathbf{DC}$, $\mathbf{EC}$, $\mathbf{PO}$, $\mathbf{TPP}$ and $\mathbf{NTPP}$ are defined in Table 3, $\mathbf{EQ}$ is the identity relation, and $\mathbf{TPP}^\sim$ and $\mathbf{NTPP}^\sim$ are the converses of $\mathbf{TPP}$ and $\mathbf{NTPP}$ respectively. See Figure 1 for illustration. The RCC-5 algebra is the sub-algebra of RCC-8 generated by the five part-whole relations

$$\mathbf{DR}, \mathbf{PO}, \mathbf{EQ}, \mathbf{PP}, \mathbf{PP}^\sim,$$

where $\mathbf{DR} = \mathbf{DC} \cup \mathbf{EC}$, $\mathbf{PP} = \mathbf{TPP} \cup \mathbf{NTPP}$, and $\mathbf{PP}^\sim = \mathbf{TPP}^\sim \cup \mathbf{NTPP}^\sim$.

Relation	Meaning	Relation	Meaning
DC	$a \cap b = \varnothing$	**TPP**	$a \subset b, a \not\subseteq b^\circ$
EC	$a \cap b \neq \varnothing, a^\circ \cap b^\circ = \varnothing$	**NTPP**	$a \subset b^\circ$
PO	$a \not\subseteq b, b \not\subseteq a, a^\circ \cap b^\circ \neq \varnothing$	**EQ**	$a = b$

Table 3. Topological interpretation of basic RCC-8 relations in the plane, where a, b are plane regions, and a°, b° are the interiors of a, b, respectively.

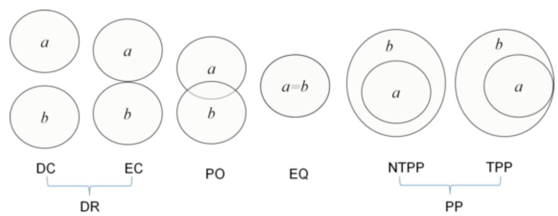


Figure 1. Illustration for basic relations in RCC-5 / RCC-8

[3] We note that the RCC algebras have interpretations in arbitrary topological spaces. In this paper, we only consider the interpretation in the real plane.

2.2 Qualitative constraint satisfaction problem

A qualitative calculus $\mathcal{M}$ provides a constraint language by using formulas of the form $v_i \alpha v_j$, where v_i, v_j are variables and α is a relation in $\mathcal{M}$. Formulas of the form $v_i \alpha v_j$ are called *constraints* (in $\mathcal{M}$). If α is a basic relation in $\mathcal{M}$, $v_i \alpha v_j$ is called a *basic constraint*. The consistency problem over $\mathcal{M}$ can then be formulated as below.

Definition 5. [2] Let $\mathcal{M}$ be a qualitative calculus on universe U. Suppose $\mathcal{S}$ is a subset of $\mathcal{M}$. The consistency problem CSPSAT($\mathcal{S}$) is defined as follows:

> **Instance**: A two tuple (V, Γ). Here V is finite set of variables $\{v_1, v_2, \cdots, v_n\}$ of variables, and Γ is a finite set of binary constraints $x\alpha y$, where $\alpha \in \mathcal{S}$ and $x, y \in V$.
> **Question**: Is there an instantiation $\nu : V \to U$ such that all constraints in Γ are satisfied?

If ν satisfies all constraints in Γ, then we say ν is a solution of Γ and say Γ is *consistent* or *satisfiable*.

Notation. We often write an instantiation $\nu : V \to U$ as an n-tuple $(\nu(v_1), \nu(v_2), \cdots, \nu(v_n))$.

The consistency problem as defined in Dfn. 5 has been investigated for many different calculi (see e.g. [1, 14, 10, 13, 9]). In particular, for Point Algebra, [14] shows that the consistency problem CSPSAT(PA) can be solved in $O(n^2)$, where n is the number of variables. For most other qualitative calculi, including IA, CRA, RCC-5, and RCC-8, the consistency problem CSPSAT($\mathcal{M}$) is NP-hard.

A set of constraints Γ is called a *basic constraint network* if Γ contains a basic constraint for each pair of variables. When only basic constraint networks are considered, however, the consistency problem over each of these four calculi becomes tractable.

2.3 Extended qualitative CSP

By Dfn. 5, in the classic consistency problem each variable can in principle take any value in the universe. In many practical applications, however, it is very common that we may have additional knowledge about some variables (cf. the restaurant and MacDonald example in introduction). It is therefore necessary to extend the qualitative CSP framework to allow restricted domains of variables.

Definition 6. Let $\mathcal{M}$ be a qualitative calculus on universe U. Suppose $\mathcal{S}$ is a subset of $\mathcal{M}$. The consistency problem CSPSAT$_f(\mathcal{S})$ is defined as follows, where the subscript 'f' stands for 'finite':

> **Instance**: A three tuple (V, Γ, D). Here V is a finite set of variables $\{v_1, v_2, \cdots, v_n\}$, D is an n-tuple $(D_1, D_2, \cdots, D_n)$ where each D_i is either U or a nonempty finite subset of U, and Γ is a finite set of binary constraints $x\alpha y$, where $\alpha \in \mathcal{S}$ and $x, y \in V$.
> **Question**: Is there an instantiation $\nu : V \to U$ such that $\nu(v_i) \in D_i$ for each i and all constraints in Γ are satisfied?

We say variable v_i appearing in such an instance is *finitely restricted* if its domain D_i is finite. If ν satisfies all constraints in Γ and $\nu(v_i) \in D_i$ for each i, then we say ν is a solution of (V, Γ, D) and say (V, Γ, D) is *consistent* or *satisfiable*.

As a special case, if D_i is required to be either the universe U or a singleton, we write the corresponding consistency problem CSPSAT$_s(\mathcal{S})$, where the subscript 's' denotes 'singleton'.

An instance of CSPSAT($\mathcal{S}$) is clearly an instance of CSPSAT$_f(\mathcal{S})$: we need only let each D_i to be the universe. We omit the proof of the following general result.

Proposition 1. *Suppose $\mathcal{B}$ is the set of basic relations in a qualitative calculus $\mathcal{M}$. Then we have*

i) CSPSAT($\mathcal{S}$) $\subset$ CSPSAT$_s(\mathcal{S})$ $\subset$ CSPSAT$_f(\mathcal{S})$*;*
ii) CSPSAT$_f(\mathcal{M})$ *is in NP if* CSPSAT$_f(\mathcal{B})$ *is in NP;*
iii) CSPSAT$_f(\mathcal{S})$ *is in NP if* CSPSAT$_s(\mathcal{S})$ *is in NP.*

Recall that the classic consistency problems for CRA, IA, RCC-5 and RCC-8 are all NP-complete. We have the following corollary.

Corollary 1. *The consistency problem* CSPSAT$_s(\mathcal{M})$ *and* CSPSAT$_f(\mathcal{M})$ *are all NP-hard for $\mathcal{M}$ being any one of IA, CRA, RCC-5, and RCC-8.*

The consistency problem CSPSAT$_s(\mathcal{S})$ was first introduced in [8], where a variable is called a *landmark* if its domain is a singleton, i.e., it is in fact *completely* determined. The main results of [8] can be summarised as follows, where we write $\mathcal{B}_\mathcal{M}$ (or simply $\mathcal{B}$) for the set of basic relations in qualitative calculus $\mathcal{M}$[4].

Proposition 2. *[8] The consistency problem* CSPSAT$_s(\mathcal{B}_\mathcal{M})$ *is in P for qualitative calculi $\mathcal{M}$ being one of PA, IA, CRA, and RCC-5, and is NP-complete for $\mathcal{M}$ being RCC-8.*

As a corollary, we have

Corollary 2. *The consistency problems* CSPSAT$_f(\mathcal{M})$ *and* CSPSAT$_s(\mathcal{M})$ *are all NP-complete for qualitative calculi $\mathcal{M}$ being one of IA, CRA, RCC-5, and RCC-8.*

It remains to determine:

- The complexities of CSPSAT$_f(PA)$ and CSPSAT$_f(\mathcal{B}_{\text{PA}})$.
- The complexities of CSPSAT$_f(\mathcal{B})$ for IA, CRA, RCC-5.

This paper is devoted to these problems. Our results are summarised in Table 4, where completely new results are underlined.

To prove the complexity results, we will need the following notion of finitely restricted sub-instance.

Definition 7. Suppose (V, Γ, D) is a CSPSAT$_f(\mathcal{S})$ instance in qualitative calculus $\mathcal{M}$, where $V = \{v_1, \cdots, v_n\}$, $D = (D_1, \cdots, D_n)$ and $\Gamma = \{v_i \alpha_{ij} v_j\}_{1 \le i,j \le n}$. Let $V' = \{v_i : D_i \neq U\}$ be the set of finitely restricted variables in V. Suppose $V' = \{v_{i_1}, v_{i_2}, \cdots, v_{i_k}\}$. Let $\Gamma' = \{v_{i_r} \alpha_{i_r i_s} v_{i_s}\}_{1 \le r,s \le k}$ and $D' = (D_{i_1}, D_{i_2}, \cdots, D_{i_k})$. Then (V', Γ', D') is an instance of CSPSAT$_f(\mathcal{S})$, and we call it the *finitely restricted sub-instance* of (V, Γ, D).

The following proposition follows directly from the fact that a consistent basic PA (IA or CRA) network is globally consistent. Note this result does not hold for RCC-5 or RCC-8.

Proposition 3. *An instance (V, Γ, D) of* CSPSAT$_f(\mathcal{B})$ *in PA (IA or CRA) is consistent iff (V, Γ) as a* CSPSAT($\mathcal{B}$) *instance is consistent and its finitely restricted sub-instance is also consistent.*

The NP-hardness results provided in the following sections are proved by polynomial reductions from 3-SAT. We fix some notations here. Suppose $\phi = \bigwedge_{j=1}^m c_j$ is a 3-SAT instance over propositional variables $p_1, p_2, \cdots, p_n$, where c_j is a clause with three literals. Write Var(c_j) for the set of propositional variables appearing in c_j, and write Var$^+(c_j)$ (Var$^-(c_j)$ resp.) for the set of propositional variables that occur as positive (negative resp.) literals in c_j. Formally,

$$\text{Var}^+(c_j) = \{p_i : p_i \text{ is a literal in } c_j\},$$
$$\text{Var}^-(c_j) = \{p_i : \neg p_i \text{ is a literal in } c_j\},$$
$$\text{Var}(c_j) = \text{Var}^+(c_j) \cup \text{Var}^-(c_j).$$

[4] Although PA and CRA were not considered in [8], we can prove for PA and CRA in exactly the same way as for IA that CSPSAT$_s(\mathcal{B})$ is in P.

	PA		CRA		IA		RCC-5		RCC-8	
$\mathcal{S}$	$\mathcal{B}_{\text{PA}}$	PA	$\mathcal{B}_{\text{CRA}}$	CRA	$\mathcal{B}_{\text{IA}}$	IA	$\mathcal{B}_{\text{RCC-5}}$	RCC-5	$\mathcal{B}_{\text{RCC-8}}$	RCC-8
CSPSAT$(\mathcal{S})$	P	P	P	NP-C	P	NP-C	P	NP-C	P	NP-C
CSPSAT$_s(\mathcal{S})$	P	P	P	NP-C	P	NP-C	P	NP-C	NP-C	NP-C
CSPSAT$_f(\mathcal{S})$	P	NP-C	NP-C	NP-C	NP-C	NP-C	NP-C	NP-C	NP-C	NP-C

Table 4. Computational complexities of the various consistency problems in PA, CRA, IA, RCC-5, and RCC-8, where $\mathcal{B}_{\mathcal{M}}$ is the set of basic relations of $\mathcal{M}$, and completely new results obtained in this paper are underlined

3 POINT ALGEBRA

In this section we prove that CSPSAT$_f(\mathcal{B}_{\text{PA}})$ is in P but CSPSAT$_f(PA)$ is NP-complete.

The following proposition shows that any consistent instance of CSPSAT$_f(\mathcal{B}_{\text{PA}})$ has a *minimal* solution in a sense. Note that we need keep only one variable from multiple equal variables (with the $=$ constraints), with the intersection of the domains of all the equal variables as its new domain. This procedure takes $O(n^2 + L)$ time and therefore we may safely assume $\Gamma = \{v_i < v_j\}_{1 \le i < j \le n}$.

Proposition 4. *Suppose (V, Γ, D) is an instance of* CSPSAT$_f(\mathcal{B}_{PA})$ *such that each D_i is a finite subset of U. If (V, Γ, D) is consistent, then there is a unique solution $(a_1, a_2, \cdots, a_n)$ such that $a_i \le a'_i$ ($1 \le i \le n$) holds for any other solution $(a'_1, a'_2, \cdots, a'_n)$. Furthermore, if $\Gamma = \{v_i < v_j\}_{1 \le i < j \le n}$, then*

- $a_1 = \min D_1$;
- $a_k = \min\{x \in D_k : x > a_{k-1}\}$ *for* $k = 2, 3, \cdots, n$.

Proof. Assume $\Gamma = \{v_i < v_j\}_{1 \le i < j \le n}$. This does not lose generality because we can combine variables related by $=$ constraints and then sort the variables by the $<, >$ constraints. Every D_i is a finite set, so (V, Γ, D) has only finitely many, say k, solutions. Suppose $(a_1^i, a_2^i, \cdots, a_n^i)(i = 1, 2, \cdots, k)$ enumerate all the solutions. Let $a_j = \min\{a_j^i\}_{1 \le i \le k}$. It is routine to prove that $(a_1, a_2, \cdots, a_n)$ is the minimal solution and satisfies the property. □

We propose a polynomial algorithm that solves CSPSAT$_f(\mathcal{B}_{\text{PA}})$. For an instance (V, Γ, D), we first transform it into its finitely restricted sub-instance (V', Γ', D'). We then decide the consistency of the sub-instance by attempting to compute $(a_1, a_2, \cdots, a_n)$ by equations in Proposition 4. If we fail in the k-th step due to the emptiness of $\{x \in D_k : x > a_{k-1}\}$, we may conclude that the sub-instance, and thus the original instance, is inconsistent. If we succeed computing $(a_1, a_2, \cdots, a_n)$, then it is a solution of the sub-instance and can be extended to a solution of the original instance. The soundness of the algorithm is clear by above argument.

Algorithm 1 SOLVING CSPSAT$_f(\mathcal{B}_{\text{PA}})$

Require: CSPSAT$_f(\mathcal{B}_{\text{PA}})$ instance (V, Γ, D)
Ensure: The consistency of (V, Γ, D)
  if Γ is not consistent **then Return** 'Inconsistent';
  for $v \in V$ **do**
    if D_v is the universe **then** remove v
  for $v, v' \in V$ **do**
    if $v = v'$ **then** $D_v \leftarrow D_v \cap D_{v'}$, remove v';
  Sort V to $v_1 < \cdots < v_{n'}$ by Γ, modify D correspondingly;
  $a_1 \leftarrow \min D_1$;
  for $k = 2, 3, \cdots, n'$ **do**
    if $a_{k-1} \ge \max D_k$ **then return** 'Inconsistent';
    $a_k \leftarrow \min\{x \in D_k : x > a_{k-1}\}$;
  return 'Consistent'.

Theorem 1. *Algorithm 1 solves the* CSPSAT$_f(\mathcal{B}_{PA})$.

We next analyse the complexity of the algorithm. Suppose there are n variables in V, and the sum of the cardinalities of all finite D_i is L. Then the input size is $O(n^2 + L)$ (n^2 constraints and L points). The following proposition shows the optimality of the algorithm.

Proposition 5. *The complexity of Algorithm 1 is $O(n^2 + L)$.*

Proof. The consistency of Γ can be computed in $O(n^2)$ time by Algorithm CSPAN proposed in [14]. Calculating the finitely restricted sub-instance takes $O(n + L)$ time, and merging equal variables takes $O(n^2 + L)$ time. Sorting left variables takes $O(n \log n)$ time. Let l_i be the cardinality of D_i. Then step '$a_1 \leftarrow D_1$' takes $O(l_1)$ time, and the i-th loop body takes $O(l_{i+1})$ time ($i = 1, 2, \cdots, n' - 1$). Therefore, the complexity of the algorithm is $O(n^2 + n \log n + l_1 + l_2 + \cdots + l_{n'}) = O(n^2 + L)$. □

Despite that both CSPSAT(PA) and CSPSAT$_f(\mathcal{B}_{\text{PA}})$ are in P, the next theorem shows that CSPSAT$_f(PA)$ is NP-hard.

Theorem 2. *The consistency problem* CSPSAT$_f(PA)$ *is an NP-complete problem.*

Proof. We provide a reduction from the NP-complete graph 3-coloring problem, which decides whether we can assign one of three colors to each node in a graph such that any edge in the graph connects two nodes with different colors.

Suppose (V, E) is a graph where $V = (v_1, \cdots, v_n)$ and $E \subseteq V \times V$ are respectively the node set and the edge set. Define a CSPSAT$_f(PA)$ instance (U, Γ, D) as follows:

$$U = \{u_1, \cdots, u_n\},$$
$$D = \{D_{u_1}, \cdots, D_{u_n}\}, \text{ where } D_{u_n} = \{1, 2, 3\},$$
$$\Gamma = \{u_i \ne u_j : (v_i, v_j) \in E\}.$$

Where a variable u_i is assigned a number k in $\{1, 2, 3\}$ corresponds to that node v_i is colored with the k-color, while a constraint in Γ, asserting two variables should be assigned different numbers, corresponds to that two nodes connected by an edge should be colored with different colors. It is clear that (V, E) can be 3-colored iff (U, Γ, D) is satisfiable. Therefore the consistency problem CSPSAT$_f(PA)$ is NP-hard, and hence NP-complete. □

Remark 1. The NP-hardness is due to the uncertainty of the $\ne$ constraints and the finiteness of the domains. In fact it can be proved that CSPSAT$_f(\mathcal{S})$ is in P for $\mathcal{S} = \{<, =, >, \le, \ge, ?\}$ (i.e., only the $\ne$ constraints are prohibited). A polynomial algorithm with the same complexity can be devised, except that the partial-ordered variables are embedded into a total order by topological sort.

4 CARDINAL RELATION ALGEBRA AND INTERVAL ALGEBRA

This section shows that CSPSAT$_f(\mathcal{B}_{\text{CRA}})$ is NP-hard by providing a polynomial reduction from 3-SAT. Note the reduction can also apply to CSPSAT$_f(\mathcal{B}_{\text{IA}})$ with slight modification.

Suppose $\phi = \bigwedge_{j=1}^{m} c_j$ is a 3-SAT instance over propositional variables $p_1, p_2, \cdots p_n$, where clause $c_j = l_{j,1} \vee l_{j,2} \vee l_{j,3}$. We construct in polynomial time an instance $(V_\phi, \Gamma_\phi, D_\phi)$ of $\text{CSPSAT}_f(\mathcal{B}_{\text{CRA}})$ such that $(V_\phi, \Gamma_\phi, D_\phi)$ has a solution ν if and only if ϕ is satisfied by a propositional assignment $\pi : \{p_1, p_2, \cdots p_n\} \to \{\textbf{true}, \textbf{false}\}$.

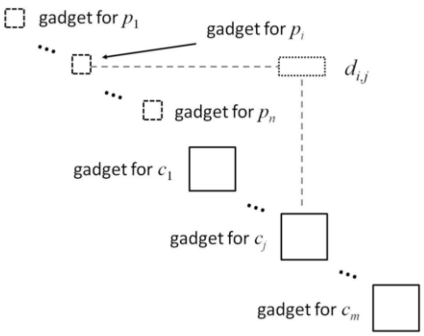


Figure 2. Overview of the configuration of all spatial variables in CRA, where we assume $p_i \in \text{Var}(c_j)$

We introduce for each propositional variable p_i and each propositional clause c_j a 'gadget', which is a small component in the constructed instance $(V_\phi, \Gamma_\phi, D_\phi)$ that simulates p_i or c_j. The gadget for p_i is a spatial variable v_i, and the gadget for c_j has five spatial variables $u_{j,1}, u_{j,2}, u_{j,3}, x_j$ and y_j. The domains of v_i and $u_{j,s}$ $(s = 1, 2, 3)$ all consist of two points, simulating the truth value assigned to p_i and $l_{j,s}$ (the s-th literal in clause c_j). Spatial variables x_j and y_j are used to exclude a certain configuration of $u_{j,1}, u_{j,2}$ and $u_{j,3}$ which corresponds to the truth valued assignments that falsify clause c_j. We also introduce spatial variable $d_{i,j}$ to propagate the value of v_i to that of $u_{j,s}$, assuming $l_{j,s}$ is p_i or $\neg p_i$. Figure 2 provides an overview for the relative locations of the domains of spatial variables v_i (the i-th top dashed box), $u_{j,s}, x_j, y_j$ (the j-th top solid box), and $d_{i,j}$ (the dots box at top right).

The domain of v_i is $D_{v_i} = \{V_i^+, V_i^-\}$, and the domain of $u_{j,s}$ is $D_{u_{j,s}} = \{U_{j,s}^+, U_{j,s}^-\}$ $(1 \leq j \leq m, s = 1, 2, 3)$ (see Figure 3 (a) for illustration). We intend to translate a propositional assignment π into a spatial assignment ν, such that $\pi(p_i) = \textbf{true}$ if $\nu(v_i) = V_i^+$ and $\pi(p_i) = \textbf{false}$ if $\nu(v_i) = V_i^-$. Assuming the literal $l_{j,s}$ is either p_i or $\neg p_i$, the point assigned to $u_{j,s}$ should be somehow decided by the point assigned to v_i. In details, we would like to require that

- If $l_{j,s} = p_i$, then $\nu(u_{j,s}) = U_{j,s}^+$ iff $\nu(v_i) = V_i^+$;
- If $l_{j,s} = \neg p_i$, then $\nu(u_{j,s}) = U_{j,s}^+$ iff $\nu(v_i) = V_i^-$.

The spatial variable $d_{i,j}$ is introduced to transfer the position of v_i to that of $u_{j,s}$. The domain of $d_{i,j}$ is specified as $D_{d_{i,j}} = \{D_{i,j}^+, D_{i,j}^-\}$, where the positions of $D_{i,j}^+$ and $D_{i,j}^-$ depend on $l_{j,s}$ being p_i or $\neg p_i$ (see Figure 3) (c)(d). With constraints

$$v_i \ \mathrm{W} \ d_{i,j}, \quad d_{i,j} \ \mathrm{N} \ u_{j,s},$$

it is clear that the requirements are fulfilled (cf. Figure 3 (c)(d)).

Clause $c_j = l_{j,1} \vee l_{j,2} \vee l_{j,3}$ is not satisfied by a propositional assignment π if and only if none of $l_{j,s}$ is satisfied by π, which corresponds to that $u_{j,s}$ all take positions $U_{j,s}^-$. Such configurations are excluded by spatial variables x_j and y_j in the gadget for c_j. Their domains are respectively $D_{x_j} = \{X_j^1, X_j^2, X_j^3\}$ and $D_{y_j} = \{Y_j^1, Y_j^2\}$, depicted in Figure 4. The following constraints about x_j and y_j are imposed:

$$c_{j,1} \ \mathrm{NW} \ x_j, \quad x_j \ \mathrm{NW} \ c_{j,2}, \quad x_j \ \mathrm{NW} \ y_j, \quad y_j \ \mathrm{NW} \ c_{j,3}.$$

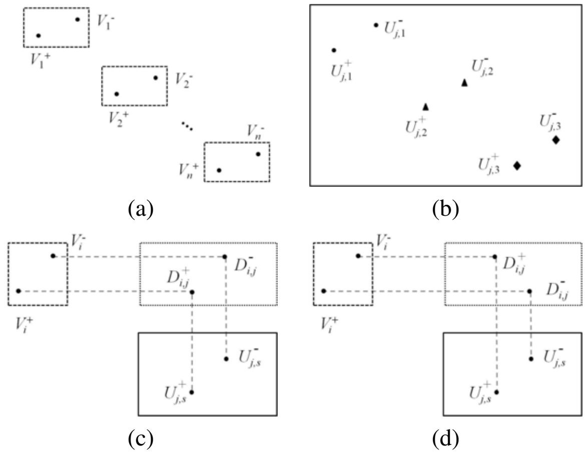


Figure 3. Illustrations of the domains of (a) v_i, (b) $u_{j,s}$, (c) (d) $d_{i,j}$, where $l_{j,s} = p_i$ in (c) and $l_{j,s} = \neg p_i$ in (d)

The first two constraints imply that $\nu(x_j) = X_j^1$ if $\nu(u_{j,1}) = U_{j,1}^-$ and $\nu(u_{j,2}) = U_{j,2}^-$. The latter two further imply that, if $\nu(x_j) = X_j^1$ and $\nu(u_{j,3}) = U_{j,3}^-$, then y_j is not realisable, because neither Y_j^1 nor Y_j^2 can satisfy the constraints. These four constraints together enforce that y_j is not realisable if $\nu(u_{j,s}) = U_{j,s}^-$ hold for $s = 1, 2, 3$. Furthermore, it can be verified that x_j and y_j are realisable in any other configuration of $u_{j,1}, u_{j,2}, u_{j,3}$.

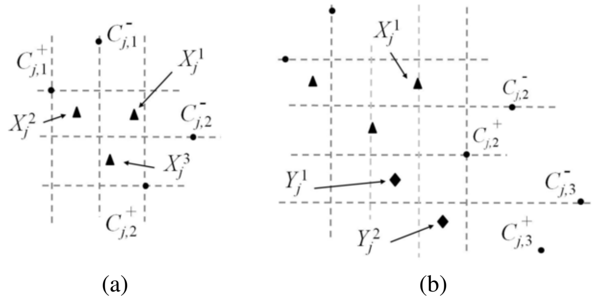


Figure 4. Illustrations for the domain of (a) x_j and (b) y_j

Now we have finished the construction of the $\text{CSPSAT}_f(\mathcal{B}_{\text{CRA}})$ instance $(V_\phi, \Gamma_\phi, D_\phi)$. In summary,

$$\begin{aligned} V_\phi =& \{v_1, \cdots, v_n\} \cup \{u_{j,1}, u_{j,2}, u_{j,3}, x_j, y_j\}_{1 \leq j \leq m} \\ & \cup \{d_{i,j} : p_i \in \text{Var}(c_j)\}, \\ D_{v_i} =& \{V_i^+, V_i^-\}, \quad D_{u_{j,s}} = \{U_{j,s}^+, U_{j,s}^-\} \ (s = 1, 2, 3), \\ D_{x_j} =& \{X_j^1, X_j^2, X_j^3\}, \ \ D_{y_j} = \{Y_j^1, Y_j^2\}, \ \ D_{d_{i,j}} = \{D_{i,j}^+, D_{i,j}^-\}. \end{aligned}$$

Note that only a part of the constraints between variables in V_ϕ have been explicitly expressed. The rest of the constraints can be trivially inferred from the configurations of the domains given in Figures 2 - 4, and the details are omitted due to space restriction.

Following the idea clarified in the above construction, the following proposition can be proved without difficulty by translating a satisfying truth assignment π of ϕ to a solution ν of $(V_\phi, \Gamma_\phi, D_\phi)$, and vice versa.

Proposition 6. *Given a 3-SAT instance ϕ, suppose $(V_\phi, \Gamma_\phi, D_\phi)$ is the $\text{CSPSAT}_f(\mathcal{B}_{CRA})$ instance as constructed above. Then ϕ is satisfiable if and only if $(V_\phi, \Gamma_\phi, D_\phi)$ is consistent.*

Note that for an 3-SAT instance ϕ with n propositional variables and m clauses, the $\textsc{cspsat}_f(\mathcal{B}_{\text{CRA}})$ instance $(V_\phi, \Gamma_\phi, D_\phi)$ contains in total $n + 8m$ spatial variables, $2n + 17m$ specified points in all domains, and $(n + 8m)^2$ basic CRA constraints. Therefore, the reduction is polynomial. Together with Proposition 2, we have

Corollary 3. *The consistency problem* $\textsc{cspsat}_f(\mathcal{B}_{\mathit{CRA}})$ *is an NP-complete problem.*

Now we turn to Interval Algebra. An interval $[a, b]$ naturally corresponds to the point (a, b) on the half plane. Based on this observation, we may translate the above reduction into another reduction from 3-SAT to the consistency problem $\textsc{cspsat}_f(\mathcal{B}_{\text{IA}})$ (the details are omitted here). Therefore we get the following corollary.

Corollary 4. *The consistency problem* $\textsc{cspsat}_f(\mathcal{B}_{\mathit{IA}})$ *is an NP-complete problem.*

5 RCC-5 AND RCC-8

It is proved in [8] that the consistency problem $\textsc{cspsat}_s(\mathcal{B}_{\text{RCC-8}})$ is NP-hard but $\textsc{cspsat}_s(\mathcal{B}_{\text{RCC-5}})$ is tractable [5]. By Proposition 1, we know $\textsc{cspsat}_f(\mathcal{B}_{\text{RCC-8}})$ is NP-hard and $\textsc{cspsat}_f(\mathcal{B}_{\text{RCC-5}})$ is in NP. This section further shows that $\textsc{cspsat}_f(\mathcal{B}_{\text{RCC-5}})$ is NP-hard.

Suppose $\phi = \bigwedge_{j=1}^{m} c_j$ is a 3-SAT instance that involves n propositional variables $p_1, p_2, \cdots p_n$. We first take $2n$ pairwise disjoint rectangles $r_1^+, r_1^-, \cdots, r_n^+, r_n^-$ on the plane, and then construct a $\textsc{cspsat}_f(\mathcal{B}_{\text{RCC-5}})$ instance (V, Γ, D) in which the domains are all unions of the selected rectangles, where we suppose $\text{Var}(c_j) = \{p_{j_1}, p_{j_2}, p_{j_3}\}$ and where $a^*_{j_k}$ is $r^-_{j_k}$ if $p_{j_k} \in \text{Var}^+(c_j)$, and is $r^+_{j_k}$ if $p_{j_k} \in \text{Var}^-(c_j)$, for $k = 1, 2, 3$.

$$
\begin{aligned}
&V = \{v_1, v_2, \cdots, v_n\} \cup \{u_1, \cdots u_m\},\\
&D = \{V_1, V_2, \cdots, V_n, U_1, U_2, \cdots, U_m\}, \text{where}\\
&\quad V_i = \{r_i^+, r_i^-\},\\
&\quad U_j = \{a_{j_1} \cup a_{j_2} \cup a_{j_3} : a_{j_k} \in V_{j_k}\} - \{a^*_{j_1} \cup a^*_{j_2} \cup a^*_{j_3}\},\\
&\Gamma = \{v_i \mathbf{DR} v_j : i \neq j\}\\
&\quad \cup \{u_i \mathbf{DR} u_j : \text{Var}(c_i) \cap \text{Var}(c_j) = \varnothing\}\\
&\quad \cup \{u_i \mathbf{PO} u_j : \text{Var}(c_i) \cap \text{Var}(c_j) \neq \varnothing\}\\
&\quad \cup \{v_i \mathbf{PP} u_j : p_i \in \text{Var}(c_j)\} \cup \{v_i \mathbf{DR} u_j : p_i \notin \text{Var}(c_j)\}.
\end{aligned}
$$

The above construction relates the truth value assigned to propositional variable p_i (**true** or **false**) to the region assigned to spatial variable v_i (r_i^+ or r_i^-). Spatial variable u_j is required to contain spatial variables in set $\text{Var}(c_j) = \{p_{j_1}, p_{j_2}, p_{j_3}\}$ by the **PP** constraints. In fact, u_j is forced to be exactly the union of variables in $\text{Var}(c_j)$ according to the first term in U_j. Moreover, the second term in u_j forbids the one configuration among the eight which corresponds the truth assignment that falsifies clause c_j.

It is routine to verify that ϕ is satisfiable if and only if the $\textsc{cspsat}_f(\mathcal{B}_{\text{RCC-5}})$ instance (V, Γ, D) is consistent. This reduction is polynomial therefore $\textsc{cspsat}_f(\mathcal{B}_{\text{RCC-5}})$ is NP-hard.

Theorem 3. *The problem* $\textsc{cspsat}_f(\mathcal{B}_{\mathit{RCC\text{-}5}})$ *is NP-complete, if the regions mentioned in the instance are (general) polygons.*

[5] The regions in the finite-specified domains, or landmarks, in the instance are required to be (general) polygons.

6 CONCLUSION AND FUTURE WORK

One major difference between qualitative CSPs and classical CSPs is that the domain of a qualitative CSP is usually infinite, while that of a classical CSP is always finite. In this paper we proposed an extended qualitative CSP framework to support finite domains. In the extended framework, a spatial/temporal variable could take values from a finite domain or even a singleton. This reflects practical demands in applications such as urban planning and spatial query processing where we often have additional knowledge about certain variable, which restricts the possible candidates of the variable to a finite set. We believe this extension is necessary to bring QSTR closer to real-world applications.

The computational complexity of the extended consistency problem has been completely studied for five very popular qualitative calculi. The results are summarised in Table 4, where for each calculus, we determined whether each of the six variants of the consistency problem is in P or NP-complete.

Future work will consider the combination of multiple spatial/temporal calculi and approximated methods for solving the extended qualitative CSPs.

ACKNOWLEDGEMENTS

This work was partially supported by an ARC Discovery Project (DP120104159) and an ARC Future Fellowship (FT0990811).

REFERENCES

[1] James F. Allen, 'Maintaining knowledge about temporal intervals', *Commun. ACM*, **26**(11), 832–843, (1983).

[2] A G Cohn and J Renz, 'Qualitative spatial reasoning', in *Handbook of Knowledge Representation*, eds., Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, Elsevier, (2007).

[3] Andrew U. Frank, 'Qualitative spatial reasoning with cardinal directions', in *ÖGAI*, pp. 157–167, (1991).

[4] Alfonso Gerevini and Jochen Renz, 'Combining topological and size information for spatial reasoning', *Artif. Intell.*, **137**(1-2), 1–42, (2002).

[5] Sanjiang Li, 'Combining topological and directional information for spatial reasoning', in *IJCAI*, pp. 435–440, (2007).

[6] Gerard Ligozat, 'Reasoning about cardinal directions', *J. Vis. Lang. Comput.*, **9**(1), 23–44, (1998).

[7] Weiming Liu, Sanjiang Li, and Jochen Renz, 'Combining rcc-8 with qualitative direction calculi: Algorithms and complexity', in *IJCAI*, pp. 854–859, (2009).

[8] Weiming Liu, Sheng sheng Wang, Sanjiang Li, and Dayou Liu, 'Solving qualitative constraints involving landmarks', in *CP*, pp. 523–537, (2011).

[9] Weiming Liu, Xiaotong Zhang, Sanjiang Li, and Mingsheng Ying, 'Reasoning about cardinal directions between extended objects', *Artif. Intell.*, **174**(12-13), 951–983, (2010).

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

[11] D. Papadias and T. Sellis, 'Qualitative representation of spatial knowledge in two-dimensional space', *The VLDB Journal*, **3**(4), 479–516, (1994).

[12] David A. Randell, Zhan Cui, and Anthony G. Cohn, 'A spatial logic based on regions and connection', in *KR*, pp. 165–176, (1992).

[13] Jochen Renz and Bernhard Nebel, 'On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the region connection calculus', *Artif. Intell.*, **108**(1-2), 69–123, (1999).

[14] Peter van Beek, 'Reasoning about qualitative temporal information', *Artif. Intell.*, **58**(1-3), 297–326, (1992).

[15] Marc B. Vilain and Henry A. Kautz, 'Constraint propagation algorithms for temporal reasoning', in *AAAI*, pp. 377–382, (1986).

[16] Stefan Wölfl and Matthias Westphal, 'On combinations of binary qualitative constraint calculi', in *IJCAI*, pp. 967–973, (2009).