Artificial Intelligence 174 (2010) 951–983

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Reasoning about cardinal directions between extended objects ☆

Weiming Liu [a], Xiaotong Zhang [b], Sanjiang Li [a,b,*], Mingsheng Ying [a,b]

[a] *Centre for Quantum Computation and Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia*
[b] *State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China*

A R T I C L E I N F O

Article history:
Received 6 July 2009
Received in revised form 18 May 2010
Accepted 18 May 2010

Keywords:
Qualitative spatial reasoning
Cardinal direction calculus
Connected regions
Consistency checking
Maximal canonical solution

A B S T R A C T

Direction relations between extended spatial objects are important commonsense knowledge. Recently, Goyal and Egenhofer proposed a relation model, known as the cardinal direction calculus (CDC), for representing direction relations between *connected* plane regions. The CDC is perhaps the most expressive qualitative calculus for directional information, and has attracted increasing interest from areas such as artificial intelligence, geographical information science, and image retrieval. Given a network of CDC constraints, the consistency problem is deciding if the network is realizable by connected regions in the real plane. This paper provides a cubic algorithm for checking the consistency of complete networks of basic CDC constraints, and proves that reasoning with the CDC is in general an NP-complete problem. For a consistent complete network of basic CDC constraints, our algorithm returns a '*canonical*' solution in cubic time. This cubic algorithm is also adapted to check the consistency of complete networks of basic cardinal constraints between possibly disconnected regions.

1. Introduction

Representing and reasoning with spatial information is of particular importance in areas such as artificial intelligence (AI), geographical information systems (GISs), robotics, computer vision, image retrieval, natural language processing, *etc.* While the numerical quantitative approach prevails in robotics and computer vision, it is widely acknowledged in AI and GIS that the qualitative approach is more attractive (see *e.g.* [6]).

A predominant part of spatial information is represented by relations between spatial objects. In general, spatial relations are classified into three categories: topological, directional, and metric (*e.g.* size, distance, shape, *etc.*). The RCC8 constraint language [34] is the principal topological formalism in AI, and has been extensively investigated by many researchers (see *e.g.* [37,35,43,7,47,46,24,25,23]). When restricted to simple plane regions, RCC8 is equivalent to the 9-Intersection Model (9IM) [9], which is a very influential relation model in GIS.

Unlike for topological relations, there are several competitive models for direction relations [10,11,2]. Most of these models approximate a spatial object by a point (*e.g.* its centroid) or a box. This is too crude in real-world applications such as describing directional information between two countries, say, Portugal and Spain. Recently, Goyal and Egenhofer [16,15] proposed a relation model, known as the cardinal direction calculus (CDC), for representing direction relations between connected plane regions. In the CDC the reference object is approximated by a box, while the primary object is

☆ This paper is an extended version of Xiaotong Zhang, Weiming Liu, Sanjiang Li, Mingsheng Ying, Reasoning with cardinal directions: An efficient algorithm, in: AAAI 2008, pp. 387–392.

* Corresponding author at: Centre for Quantum Computation and Intelligent Systems, Faculty of Engineering and Information Technology, University of Technology, Sydney, P.O. Box 123, Broadway, NSW 2007, Australia.

E-mail address: sanjiang.li@uts.edu.au (S. Li).

0004-3702/$ – see front matter
doi:10.1016/j.artint.2010.05.006

not approximated. This means that the exact geometry of the primary object could be used in the representation of the direction. This calculus has 218 basic relations, which is quite large when compared with the RCC8 and Allen's Interval Algebra [1]. Due to its expressiveness, the CDC has attracted increasing interest from areas such as AI [40,41,31], GIS [17], database [39], and image retrieval [19].

One basic criterion for evaluating a spatial relation model is the proper balance between its representation expressivity and reasoning complexity. While the reasoning complexity of the point-based and the box-based model of direction relations has been investigated in depth (see [26] and [2]), there are few works discussing the complexity of reasoning with the CDC.

One central reasoning problem with the CDC (and any other qualitative calculus) is the *consistency* (or *satisfaction*) problem. Other reasoning problems such as deriving new knowledge from the given information, updating the given knowledge, or finding a minimal representation can be easily transformed into the consistency problem [6]. In particular, given a complete network of CDC constraints

$$\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^{n} \quad \text{(each } \delta_{ij} \text{ is a CDC relation)} \tag{1}$$

over n spatial variables $v_1, \ldots, v_n$, the consistency problem is deciding if $\mathcal{N}$ is realizable by a set of n *connected* regions in the real plane. The consistency problem over the CDC is an open problem. Before this work, we did not know if there are efficient algorithms deciding if a set of CDC constraints are realizable. Even worse, we did not know if this is a decidable problem. Furthermore, we did not know how to construct a realization for a satisfiable set of CDC constraints.

This paper is devoted to solving these problems. We first show each consistent CDC network has a '*canonical*' solution (Theorem 3) and then devise a cubic algorithm for checking if a complete network of basic CDC constraints is consistent. When the network is consistent, this algorithm also generates a canonical solution. We further show that deciding the consistency of an arbitrary network of CDC constraints is an NP-Complete problem. This implies in particular that reasoning with the CDC is decidable.

Some restricted versions of the consistency problem have been discussed in the literature. Cicerone and di Felice [3] discussed the pairwise consistency problem, which decides when a pair of basic CDC relations (δ, δ') is consistent, *i.e.* when $\{v_1 \delta v_2, v_2 \delta' v_1\}$ is consistent. Skiadopoulos and Koubarakis [40] investigated the weak composition problem [7,24] of the CDC, which is closely related to the consistency problem of basic CDC networks involving only three variables.

The CDC algebra is defined over connected regions. A variant of the CDC was proposed in [41], where cardinal directions between possibly disconnected regions are defined in the same way. This calculus, termed the CDC_d in this paper, contains 511 basic relations. An $O(n^5)$ algorithm[1] was proposed in [41] for checking the consistency of basic constraints in the CDC_d, but the consistency problem over the CDC is still open. Recently, Navarrete et al. [31] tried to adapt the approach used in [41] to cope with connected regions, but their approach turns out to be incorrect (see Remark 3 in Section 6.1 of this paper).

The remainder of this paper proceeds as follows. Section 2 recalls basic notions in qualitative spatial/temporal reasoning and introduces the well-known Interval Algebra (IA) [1]. We introduce the CDC algebra in Section 3, where the connection between CDC and IA relations is established in a natural way. Section 4 introduces the notion of canonical solution of a consistent basic CDC network. Section 5 first proposes an intuitive $O(n^4)$ algorithm for consistency checking of complete basic networks and then improves it to $O(n^3)$. In Section 6, we first show local consistency is insufficient to decide the consistency of even basic CDC networks, and then apply our main algorithm to the pairwise consistency problem and the weak composition problem. In Section 7 we adapt the main algorithm for connected regions to solve consistency checking in two variants of the CDC. Section 8 discusses related work on the computational properties of other qualitative direction calculi. Conclusions are given in the last section.

Codes of the main algorithm are available via http://sites.google.com/site/lisanjiang/cdc, where we also provide illustrations for all 757 different consistent pairs of CDC basic relations and the illustration of the weak composition of $SW : W$ and $NE : E$. Interested readers may consult that webpage for detailed proofs of some minor results that are omitted in the present paper.

Table 1 summaries notations used in this paper.

2. Qualitative calculi: Basic notions and examples

Since Allen's Interval Algebra, the study of qualitative calculi or relation models has been a central topic in qualitative spatial and temporal reasoning. This section introduces basic notions and important examples of qualitative calculi.

2.1. Basic notions

Let D be a universe of temporal or spatial or spatial-temporal entities. We use small Greek symbols for representing relations on D. For a relation α on D and two elements x, y in D, we write $(x, y) \in \alpha$ or $x \alpha y$ to indicate that (x, y) is an instance of α. For two relations α, β on D, we define the complement of α, the intersection, and the union of α and β as follows.

[1] We note that this algorithm applies to any (possibly incomplete) set of basic constraints.

Table 1
Notations.

Notations	Meanings
$\alpha, \beta, \gamma, \delta, \theta$	relations (p. 952)
v_i, v_j	spatial variable or interval variable (p. 953)
$\mathcal{N}$	network of constraints (p. 953)
a, b, c	regions (p. 954)
$I_x(a), I_y(a)$	the x- and y-projective intervals of region a (Eq. (4))
$\mathcal{M}(a)$	the minimal bounding rectangle (mbr) of region a (Eq. (5))
χ	tile name variable (p. 955)
$\chi(a)$	tile χ of a (p. 955)
$\mathfrak{a} = \{a_i\}_{i=1}^n$	a set of regions a_i (p. 959)
$\iota^x(\delta), \iota^y(\delta)$	the x- and y-projective interval relations of δ (Eq. (10))
$\iota^x(\delta, \gamma)$	defined as $\iota^x(\delta) \cap \iota^x(\gamma)^\sim$ (Eq. (12))
ρ_{ij}^x	defined as $\iota^x(\delta_{ij}, \delta_{ji}) \setminus (\mathsf{m} \cup \mathsf{mi})$ (Eq. (14))
ρ_{ij}^y	defined as $\iota^y(\delta_{ij}, \delta_{ji}) \setminus (\mathsf{m} \cup \mathsf{mi})$ (Eq. (15))
$\mathcal{N}_x, \mathcal{N}_y$	the x- and y-projective IA networks of $\mathcal{N}$ (p. 958)
$\{I_i\}_{i=1}^n, \{J_i\}_{i=1}^n$	sets of intervals (p. 962)
m_i	rectangle (p. 962)
$S(\mathfrak{a})$	the frame of $\mathfrak{a}$ (Eq. (18))
$C(\mathfrak{a})$	the cell set of $\mathfrak{a}$ (Eq. (20))
c_{ij}	a cell in $C(\mathfrak{a})$ (Eq. (19))
$p_{ij}, p, p^{(k)}$	pixels, where $p_{ij} = [i, i+1] \times [j, j+1]$ (Definition 8)
a_i^r	the regularization of a_i (Eq. (21))
b_i	a special digital region contained in m_i (Eq. (25))
c_i	the connected component of b_i which has mbr m_i (Lemma 5)
$B(a)$	the Boolean matrix of digital region a (Eq. (30))

$$-\alpha = \{(x, y) \in D \times D\colon\ (x, y) \notin \alpha\},$$
$$\alpha \cap \beta = \{(x, y) \in D \times D\colon\ (x, y) \in \alpha \text{ and } (x, y) \in \beta\},$$
$$\alpha \cup \beta = \{(x, y) \in D \times D\colon\ (x, y) \in \alpha \text{ or } (x, y) \in \beta\}.$$

We write $\mathbf{Rel}(D)$ for the set of binary relations on D. Clearly, the 6-tuple $(\mathbf{Rel}(D);\ -, \cap, \cup, \varnothing, D \times D)$ is a Boolean algebra, where $\varnothing$ and $D \times D$ are the empty relation and the universal relation on D, respectively.

A finite set $\mathcal{B}$ of nonempty relations on D is *jointly exhaustive and pairwise disjoint* (JEPD) if any two entities in D are related by one and only one relation in $\mathcal{B}$. We write $\langle\mathcal{B}\rangle$ for the subalgebra of $\mathbf{Rel}(D)$ generated by $\mathcal{B}$, *i.e.* the smallest subalgebra of the Boolean algebra $\mathbf{Rel}(D)$ which contains $\mathcal{B}$. Clearly, relations in $\mathcal{B}$ are atoms in the Boolean algebra $\langle\mathcal{B}\rangle$. We call $\langle\mathcal{B}\rangle$ a *qualitative calculus* on D, and call relations in $\mathcal{B}$ *basic* relations of the calculus. A similar definition was given by Ligozat and Renz [27], where $\mathcal{B}$ was required to be closed under converse and contain id_D — the identity relation on D.

For two relations α, β on D, the converse of α and the composition of α and β are defined as usual.

$$\alpha^\sim = \{(y, x) \in D \times D\colon\ (x, y) \in \alpha\},$$
$$\alpha \circ \beta = \{(x, y) \in D \times D\colon\ (\exists z \in D)[(x, z) \in \alpha \text{ and } (z, y) \in \beta]\}.$$

For two relations α, β in $\langle\mathcal{B}\rangle$, it is possible that $\alpha^\sim$ or $\alpha \circ \beta$ is not in $\langle\mathcal{B}\rangle$, *i.e.*, they cannot be represented as the union of some relations in $\mathcal{B}$. We say a qualitative calculus $\langle\mathcal{B}\rangle$ is *closed under composition* (*closed under converse*, resp.) if the composition of any two relations (the converse of any relation, resp.) in $\langle\mathcal{B}\rangle$ is still a relation in $\langle\mathcal{B}\rangle$.

An important reasoning problem in a qualitative calculus $\langle\mathcal{B}\rangle$ is the consistency (or satisfaction) problem. Let $\mathcal{A}$ be a subset of $\langle\mathcal{B}\rangle$. A constraint over $\mathcal{A}$ has the form $(x\gamma y)$ with $\gamma \in \mathcal{A}$. For a set of variables $V = \{v_i\}_{i=1}^n$, and a set of constraints $\mathcal{N}$ involving variables in V, we say $\mathcal{N}$ is a *complete constraint network* (or *network* for short) if there exists a unique constraint $(v_i \gamma v_j)$ in $\mathcal{N}$ for each pair (i, j). A network $\mathcal{N}$ is said to be over $\mathcal{A}$ if each constraint in $\mathcal{N}$ is over $\mathcal{A}$. In particular, we say a network is a basic network if it is over $\mathcal{B}$. We stress that, in this paper, a *basic network of constraints* or a *network of basic constraints* is always *complete*, *i.e.* a basic constraint is specified for each pair of variables.

A constraint network $\mathcal{N} = \{v_i \gamma_{ij} v_j\}_{i,j=1}^n$ is *consistent* (or *satisfiable*) if there is an instantiation $\{a_i\}_{i=1}^n$ in D such that $(a_i, a_j) \in \gamma_{ij}$ holds for all $1 \leqslant i, j \leqslant n$. In this case, we call $\{a_i\}_{i=1}^n$ a solution of $\mathcal{N}$. The consistency problem over $\mathcal{A}$ is the decision problem of the consistency of constraint networks over $\mathcal{A}$.

2.2. The Interval Algebra

The Interval Algebra (IA) [1] is a qualitative calculus defined on the set of (closed and bounded) intervals in the real line. The IA is generated by a set $\mathcal{B}_{int}$ of 13 JEPD relations between intervals (see Table 2).

The IA is closed under converse and composition. It is a relation algebra in the sense of Tarski [44]. The computational complexity of reasoning with the IA has been extensively investigated by researchers in artificial intelligence (see [33,21]

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

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	$x^- = y^- < x^+ = y^+$

and references therein). In particular, Allen [1] introduced the important notion of path-consistency for networks of IA constraints and Valdés-Pérez [45] proved that path-consistency suffices to decide the consistency of basic IA networks.

The definitions of basic IA relations as given in Table 2 concern only the ordering of the endpoints of intervals. This suggests that different solutions of the same basic IA network respect the same ordering. In particular, we could choose intervals that have integer endpoints.

Definition 1 *(canonical set of intervals [28]).* Suppose $\mathfrak{l} = \{[l_i^-, l_i^+]\}_{i=1}^n$ is a set of intervals. Let $E(\mathfrak{l})$ be the set of endpoints of intervals in $\mathfrak{l}$. We say $\mathfrak{l}$ is a *canonical set of intervals* iff $E(\mathfrak{l}) = [0, M] \cap \mathbb{Z}$, where M is the largest number in $E(\mathfrak{l})$. A solution of a basic IA network is called a *canonical interval solution* if it is a canonical set of intervals.

Clearly, if $\mathfrak{l} = \{[l_i^-, l_i^+]\}_{i=1}^n$ is a canonical set of intervals, then each l_i^- (l_i^+) is an integer between 0 and $2n - 1$. Moreover, let M be the largest number in $E(\mathfrak{l})$. Then $M < 2n$ and for any $0 \leqslant l \leqslant M$ there exists i such that $l_i^- = l$ or $l_i^+ = l$. The following theorem shows that each consistent basic IA network has a unique canonical solution.

Theorem 1. *Suppose $\mathcal{N} = \{v_i \lambda_{ij} v_j\}_{i,j=1}^n$ is a basic IA network. If $\mathcal{N}$ is consistent, then it has a unique canonical interval solution.*

Proof. Suppose $\mathfrak{l} = \{[l_i^-, l_i^+]\}_{i=1}^n$ is a solution of $\mathcal{N}$. Write $\alpha_0 < \alpha_1 < \cdots < \alpha_{n^*}$ for the ordering of $E(\mathfrak{l}) = \{l_i^-, l_i^+\}_{i=1}^n$. Define $h : E(\mathfrak{l}) \to \{0, 1, \ldots, n^*\}$ as $h(x) = k$ if $x = \alpha_k$. Because only the ordering of endpoints of intervals matters in a solution, $\mathfrak{h} = \{[h(l_i^-), h(l_i^+)]\}_{i=1}^n$ is also a solution of $\mathcal{N}$. Moreover, it is clear that $E(\mathfrak{h}) = [0, n^*] \cap \mathbb{Z}$. Therefore, $\mathfrak{h}$ is a canonical interval solution of $\mathcal{N}$. Such a solution is clearly unique. □

3. Cardinal direction calculus

In this section we first introduce the cardinal direction calculus (CDC) [16,40] and then establish its connection with the Interval Algebra. In particular, we will associate two basic IA networks $\mathcal{N}_x$ and $\mathcal{N}_y$ with each basic CDC network $\mathcal{N}$, such that $\mathcal{N}$ is consistent only if $\mathcal{N}_x$ and $\mathcal{N}_y$ are consistent.

3.1. Direction relation matrix

The CDC is a qualitative calculus defined for extended objects in the plane.

Definition 2 *(plane region).* A subset a of the plane is called a *region* if a is a nonempty regular closed subset, *i.e.* if $a = \overline{a^\circ}$, where x° and $\bar{x}$ are the (topological) interior and the (topological) closure of a subset x of the plane, respectively.

Connectedness is an important topological property of plane regions.

Definition 3 *(connected region and simple region).* A region a is said to be *connected* if it has a connected interior a°, *i.e.*, for any nonempty open sets u, v in the plane, if $u \cup v = a^\circ$, then $u \cap v$ is nonempty. A connected region is called *simple* if is topologically equivalent to a closed disk.

By the above definition, a connected region has a connected interior. Regions with this property are often called *strongly connected* elsewhere. Note that the boundary of a connected region may still be disconnected. In contrast, simple regions have connected interior as well as connected boundary. Examples of plane regions are illustrated in Fig. 1.

For a bounded set b in the real plane, let

$$x^-(b) = \inf\{x\colon (x, y) \in b\}, \qquad x^+(b) = \sup\{x\colon (x, y) \in b\}, \tag{2}$$

$$y^-(b) = \inf\{y\colon (x, y) \in b\}, \qquad y^+(b) = \sup\{y\colon (x, y) \in b\}. \tag{3}$$

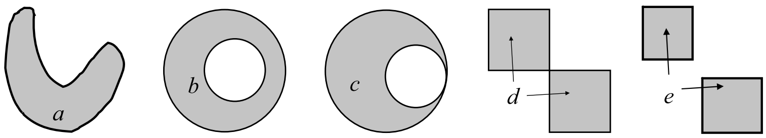


Fig. 1. Examples of plane regions, where a is a simple region, b, c are connected but non-simple regions, and d, e are not connected regions.

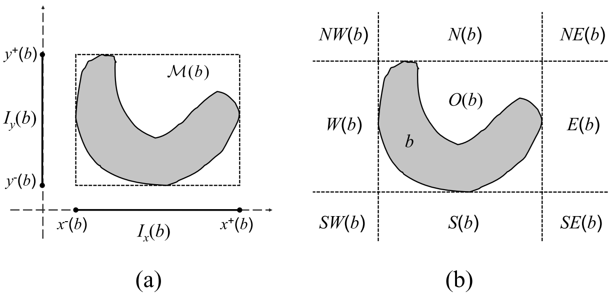
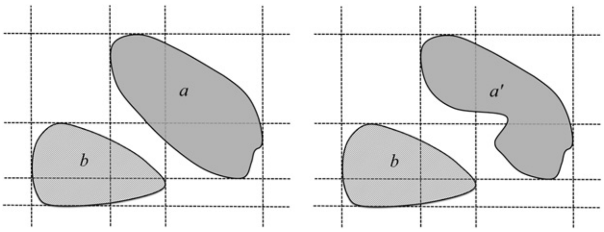


Fig. 2. A bounded connected region b (a) and its 9-tiles (b).



$$\mathrm{dir}(b,a)=\mathrm{dir}(b,a')=\begin{pmatrix}0&0&0\\1&1&0\\1&1&0\end{pmatrix},\ \mathrm{dir}(a,b)=\begin{pmatrix}0&1&1\\0&1&1\\0&0&0\end{pmatrix},\ \mathrm{dir}(a',b)=\begin{pmatrix}0&1&1\\0&0&1\\0&0&0\end{pmatrix}$$

Fig. 3. Illustrations of basic CDC relations.

We write

$$I_x(b)=\left[x^-(b),x^+(b)\right],\qquad I_y(b)=\left[y^-(b),y^+(b)\right]. \tag{4}$$

Let

$$\mathcal{M}(b)=I_x(b)\times I_y(b). \tag{5}$$

We call $\mathcal{M}(b)$ the *minimum bounding rectangle* (mbr) of b, and call $I_x(b)$ and $I_y(b)$ the x- and y-projection of b, respectively. Clearly, $\mathcal{M}(b)$ is the smallest rectangle which contains b and has sides parallel to the axes.

By extending the four edges of $\mathcal{M}(b)$, we partition the plane into nine tiles, denoted as $NW(b)$, $N(b)$, $NE(b)$, $W(b)$, $O(b)$, $E(b)$, $SW(b)$, $S(b)$, $SE(b)$ (see Fig. 2(b)). Note that each tile is a (bounded or unbounded) connected region, and the intersection of two tiles is of dimension lower than two.

Write

$$\textsc{TileName}=\{NW,N,NE,W,O,E,SW,S,SE\} \tag{6}$$

for the set of tile names. In the remainder of this paper, we often use the variable symbol χ to address a tile name in TileName. For a bounded region b, we use $\chi(b)$ to address a tile of b.

Since the partition only concerns the mbr of b, we have

Proposition 1. *For two bounded connected regions b, c, if $\mathcal{M}(b)=\mathcal{M}(c)$, then $\chi(b)=\chi(c)$ for each $\chi\in$ TileName. In particular, $\chi(b)=\chi(\mathcal{M}(b))$ for each $\chi\in$ TileName.*

The notion of direction relation matrix was first proposed by Goyal and Egenhofer [16] for representing the cardinal direction between extended spatial objects.

Table 3
Examples of valid direction relation matrices (M_1 and M_2) and invalid direction relation matrix (M_3) for bounded connected regions.

$M_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$M_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$	$M_3 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

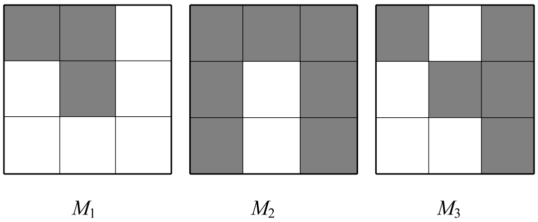


Fig. 4. Iconic representations of valid and invalid matrices in Table 3.

Definition 4 *(direction relation matrix).* Suppose a, b are two bounded connected regions. Take b as the reference object, and a as the primary object. The direction of a to b is encoded in a 3×3 Boolean matrix

$$\mathsf{dir}(a,b) = \begin{bmatrix} d^{NW} & d^{N} & d^{NE} \\ d^{W} & d^{O} & d^{E} \\ d^{SW} & d^{S} & d^{SE} \end{bmatrix}, \tag{7}$$

where

$$d^{\chi} = 1 \quad \Leftrightarrow \quad a^{\circ} \cap \chi(b) \neq \varnothing \quad \big(\chi \in \{NW, N, NE, W, O, E, SW, S, SE\}\big) \tag{8}$$

where a° is the interior of a (see Fig. 3). We call a 3×3 Boolean matrix M a *direction relation matrix*, or a *valid matrix*, if there exist bounded connected regions a, b such that $M = \mathsf{dir}(a,b)$.

Remark 1. The original description of the direction relation matrix in [16,15] lacks formality and does not consider limit cases. Formal definitions of this model appear in [40] and [3]. These two models differentiate only in the limit cases. Our definition of direction relation matrix is in accord with that of [40]. In Section 6.2 we will argue that this model is more appropriate than the one given in [3].

Remark 2. In their work [16,15], Goyal and Egenhofer considered cardinal directions between simple regions, which are, roughly speaking, connected regions without holes. This convention was adopted by other researchers, *e.g.* [40,3,31]. In this paper, we interpret spatial objects as bounded connected regions, which may have holes. This assumption is arguably more pragmatic. Moreover, as we will see in Section 7.2, the generalization from simple regions to connected regions does not affect the consistency of cardinal direction constraints. In particular, for each valid matrix M, there exist *simple* regions a, b such that $M = \mathsf{dir}(a,b)$.

Goyal and Egenhofer identified altogether 218 valid matrices.

Proposition 2. *(See [15].) A* 3×3 *Boolean matrix is a direction relation matrix if and only if it is nonzero and* 4*-connected,* i.e. *nonzero entries can be joined with vertical or horizontal lines without crossing zero entries.*

Table 3 gives examples of valid and invalid matrices.

Each valid matrix M uniquely determines a direction relation as follows:

$$\delta(M) = \big\{(a,b)\colon\ a, b \text{ are bounded connected regions and } \mathsf{dir}(a,b) = M\big\}. \tag{9}$$

In the remainder of this paper we make no distinction between a valid matrix M and the direction relation $\delta(M)$ it represents. Write $\mathcal{B}_{dir}$ for the set of cardinal directions represented by these valid matrices. Because each ordered pair of bounded connected regions determines a unique direction relation matrix, $\mathcal{B}_{dir}$ is a JEPD set of relations. The cardinal direction calculus (CDC) is defined to be the qualitative calculus generated by $\mathcal{B}_{dir}$ over the set of bounded connected regions.

As usual, we call a relation in $\mathcal{B}_{dir}$ a basic CDC relation. For a basic CDC relation, its matrix provides an intuitive representation. In particular, we can easily transform a valid matrix into an iconic representation of the basic relation [15]. Fig. 4 gives iconic representations for the matrices in Table 3. Note that the grey region in the icon of M_3 is not a connected region (cf. Definition 3). This explains why M_3 is not a valid matrix.

A different system of notations is introduced in [40] for CDC relations.

Definition 5 *(single-tile, multi-tile, component).* A basic CDC relation is called *single-tile* if its matrix has only one nonzero entry, and called *multi-tile* otherwise. We say a single-tile relation $[s^\chi]_{\chi\in\textsc{TileName}}$ is a *component* of a multi-tile relation $[d^\chi]_{\chi\in\textsc{TileName}}$, if $s^\chi \leqslant d^\chi$ for each tile name $\chi \in \textsc{TileName}$.

The nine single tile relations are written as NW (*northwest*), N (*north*), NE (*northeast*), W (*west*), O (*same*), E (*east*), SW (*southwest*), S (*south*), SE (*southeast*), respectively. A multi-tile relation δ is written as $\delta_1 : \delta_2 : \cdots : \delta_k$ ($k \leqslant 9$), where δ_i ($1 \leqslant i \leqslant k$) are all components of δ. Take the two valid relations in Fig. 4 as examples. We have $\delta(M_1) = NW : N : O$ and $\delta(M_2) = SW : W : NW : N : NE : E : SE$.

In order to describe the relative position of two connected regions a, b with the CDC, knowing the direction of b to a is not enough [15]. Fig. 3 shows an example, where the direction of b to a is the same as that of b to a' but the direction of a to b is different from that of a' to b, *i.e.* $\mathsf{dir}(b,a) = \mathsf{dir}(b,a')$ but $\mathsf{dir}(a,b) \neq \mathsf{dir}(a',b)$. This is drastically different from the IA and many other well-known qualitative calculi, where the basic relation of a to b is uniquely determined by that of b to a. We will investigate this pairwise consistency problem in detail in Section 6.2.

The next subsection establishes the connection between the CDC and the IA.

3.2. Projective IA networks

The connection between the CDC and the IA is established via the notion of projective interval relations.

Definition 6 *(projective interval relation).* For a basic CDC relation δ, the *x-projective interval relation* of δ is defined as

$$\iota^x(\delta) = \big\{\big(I_x(a), I_x(b)\big)\colon (a,b) \in \delta\big\}, \tag{10}$$

where $I_x(a)$ and $I_x(b)$ are the x-projective intervals of a and b, respectively.

Although $\iota^x(\delta)$ is called a projective interval relation, it is not immediately clear that $\iota^x(\delta)$ is indeed a relation in the IA. The following proposition, however, confirms this.

Proposition 3. *Suppose δ is a basic CDC relation. Then the x-projective interval relation $\iota^x(\delta)$ is one of the following IA relations*

$$\mathsf{p}\cup\mathsf{m},\quad \mathsf{s}\cup\mathsf{d}\cup\mathsf{f}\cup\mathsf{eq},\quad \mathsf{pi}\cup\mathsf{mi},\quad \mathsf{o}\cup\mathsf{fi},\quad \mathsf{oi}\cup\mathsf{si},\quad \mathsf{di}. \tag{11}$$

Proof. See Appendix A. □

Appendix A also gives a method for computing the projective interval relation of a basic CDC relation. Take the two valid relations in Table 3 as examples, we have $\iota^x(M_1) = \mathsf{o}\cup\mathsf{fi}$ and $\iota^x(M_2) = \mathsf{di}$.

For a pair of basic CDC relations (δ, γ), write

$$\iota^x(\delta,\gamma) = \iota^x(\delta) \cap \iota^x(\gamma)^\sim, \tag{12}$$

where $\iota^x(\gamma)^\sim$ is the converse of $\iota^x(\gamma)$ in the IA. The next proposition examines what kind of IA relation $\iota^x(\delta,\gamma)$ can be.

Proposition 4. *For a pair of basic CDC constraints (δ, γ), $\iota^x(\delta,\gamma)$ is either empty or an IA relation in $\mathcal{B}^*_{int}$, where*

$$\mathcal{B}^*_{int} = \{\mathsf{o},\mathsf{s},\mathsf{d},\mathsf{f},\mathsf{eq},\mathsf{fi},\mathsf{di},\mathsf{si},\mathsf{oi}\} \cup \{\mathsf{p}\cup\mathsf{m}, \mathsf{pi}\cup\mathsf{mi}\}. \tag{13}$$

Proof. See Appendix A. □

For example, if δ is the single-tile relation W, and γ is the multi-tile relation $NE : E : SE$, then $\iota^x(\delta,\gamma)$ is the non-basic IA relation $\mathsf{p}\cup\mathsf{m}$.

The next proposition follows immediately.

Proposition 5. *For a pair of basic CDC relations (δ, γ), if (a,b) is a solution of $\{v_1\delta v_2, v_2\gamma v_1\}$, then $(I_x(a), I_x(b)) \in \iota^x(\delta,\gamma)$.*

Proof. From $(a,b) \in \delta$ and $(b,a) \in \gamma$ we have $(I_x(a), I_x(b)) \in \iota^x(\delta)$ and $(I_x(b), I_x(a)) \in \iota^x(\gamma)$. Therefore, $(I_x(a), I_x(b)) \in \iota^x(\delta) \cap \iota^x(\gamma)^\sim = \iota^x(\delta,\gamma)$. □

This result can easily be extended to CDC constraint networks.

Proposition 6. *A basic CDC constraint network $\mathcal{N} = \{v_i\delta_{ij}v_j\}_{i,j=1}^n$ is consistent only if the IA constraint network $\{v_i\iota^x_{ij}v_j\}_{i,j=1}^n$ is consistent, where $\iota^x_{ij} = \iota^x(\delta_{ij},\delta_{ji})$.*

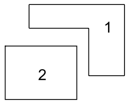
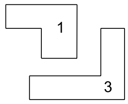
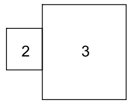
(i,j)	δ_{ij}	δ_{ji}	illus.	ρ^x_{ij}	ρ^y_{ij}
(1,2)	$\begin{pmatrix}0&1&1\\0&0&1\\0&0&0\end{pmatrix}$	$\begin{pmatrix}0&0&0\\1&1&0\\1&1&0\end{pmatrix}$		oi	oi
(1,3)	$\begin{pmatrix}1&1&0\\0&1&0\\0&0&0\end{pmatrix}$	$\begin{pmatrix}0&0&0\\0&0&1\\0&1&1\end{pmatrix}$		o	oi
(2,3)	$\begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix}$	$\begin{pmatrix}0&0&1\\0&0&1\\0&0&1\end{pmatrix}$		p	d

Fig. 5. A basic CDC network and its projective IA basic networks.

Proof. Suppose $\{a_i\}_{i=1}^n$ is a solution to $\mathcal{N}$. Then $\{I_x(a_i)\}_{i=1}^n$ is a solution to $\{v_i \iota^x_{ij} v_j\}_{i,j=1}^n$. This is because, by Proposition 5, we know $(I_x(a_i), I_x(a_j))$ is an instance of ι^x_{ij} for each pair of i, j. $\square$

Similar notions and results also apply to the y-direction.

The consistency of the IA network appearing in Proposition 6 can be further reduced to the consistency of a *basic* IA network.

Definition 7 *(projective IA networks).* Let $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ be a basic CDC network. For each i, j, define

$$\rho^x_{ij} = \iota^x_{ij} \setminus (\mathsf{m} \cup \mathsf{mi}), \tag{14}$$

$$\rho^y_{ij} = \iota^y_{ij} \setminus (\mathsf{m} \cup \mathsf{mi}), \tag{15}$$

where $\iota^x_{ij} = \iota^x(\delta_{ij}, \delta_{ji})$ and $\iota^y_{ij} = \iota^y(\delta_{ij}, \delta_{ji})$. We call $\mathcal{N}_x = \{v_i \rho^x_{ij} v_j\}_{i,j=1}^n$ and $\mathcal{N}_y = \{v_i \rho^y_{ij} v_j\}_{i,j=1}^n$ the x- and y-projective IA networks of $\mathcal{N}$, respectively.

The following result shows that the IA networks $\{v_i \iota^x_{ij} v_j\}_{i,j=1}^n$ and $\{v_i \iota^y_{ij} v_j\}_{i,j=1}^n$ are consistent iff $\mathcal{N}_x$ and $\mathcal{N}_y$ are consistent.

Proposition 7. *An IA network $\mathcal{N} = \{v_i \iota_{ij} v_j\}_{i,j=1}^n$ over $\mathcal{B}^*_{int}$ (see Eq. (13)) is consistent iff $\widehat{\mathcal{N}} = \{v_i \widehat{\iota}_{ij} v_j\}_{i,j=1}^n$ is satisfiable, where $\widehat{\iota}_{ij} = \iota_{ij} \setminus (\mathsf{m} \cup \mathsf{mi})$.*

Proof. We first observe that $\widehat{\mathcal{N}}$ is a refinement of $\mathcal{N}$. Therefore, if $\widehat{\mathcal{N}}$ is consistent, then $\mathcal{N}$ is also consistent. On the other hand, suppose $\{I_i = [u_i^-, u_i^+]\}_{i=1}^n$ is a solution to $\mathcal{N}$. We call a point $u \in M = \{u_i^-, u_i^+\}_{i=1}^n$ a meet point if there exist $i \neq j$ such that $u_i^+ = u_j^-$, *i.e.* I_i meets I_j. We can prove that $\widehat{\mathcal{N}}$ has a solution by using induction on the number K of meet points. Details are omitted. $\square$

As a corollary of Propositions 6 and 7, we know

Theorem 2. *A basic CDC network of constraints $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is consistent only if the projective IA networks $\mathcal{N}_x = \{v_i \rho^x_{ij} v_j\}_{i,j=1}^n$ and $\mathcal{N}_y = \{v_i \rho^y_{ij} v_j\}_{i,j=1}^n$ are consistent.*

The above theorem establishes a necessary condition for the consistency of basic CDC network.

Example 1. Fig. 5 specifies a basic CDC network $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^3$, and gives its projective IA networks $\mathcal{N}_x = \{v_i \rho^x_{ij} v_j\}_{i,j=1}^3$ and $\mathcal{N}_y = \{v_i \rho^y_{ij} v_j\}_{i,j=1}^3$. For each pair of $i \neq j$, a solution of $\{v_i \delta_{ij} v_j, v_j \delta_{ji} v_i\}$ is illustrated.

In the next section, we prove that each consistent CDC network $\mathcal{N}$ has a solution $\{a_i\}_{i=1}^n$ such that $\{I_x(a_i)\}_{i=1}^n$ and $\{I_y(a_i)\}_{i=1}^n$ are solutions to the projective IA networks $\mathcal{N}_x$ and $\mathcal{N}_y$, respectively.

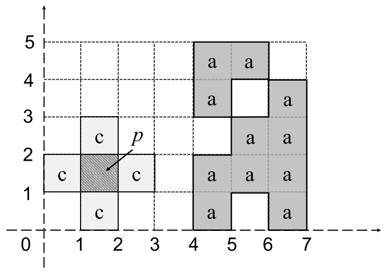


Fig. 6. A pixel p and its 4-neighbors (pixels with mark "c"), and a digital region a.

4. Canonical solution

Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a consistent basic CDC network. We show $\mathcal{N}$ has a canonical solution in a sense similar to that for IA networks.

Definition 8 *(pixel, digital region, digital solution).* A *pixel* is a rectangle $p_{ij} = [i, i+1] \times [j, j+1]$, where i, j are integers. A region a is *digital* if a is composed of pixels, *i.e.* $p_{ij} \cap a^\circ \neq \varnothing$ iff $p_{ij} \subseteq a$. A solution $\mathfrak{a} = \{a_i\}_{i=1}^n$ of a basic CDC network is *digital* if each a_i is a digital region.

Fig. 6 illustrates a pixel p and a digital region a.

Definition 9 *(canonical solution).* Let $\mathcal{N}$ be a consistent basic CDC network, and let $\mathcal{N}_x$ and $\mathcal{N}_y$ be the x- and y-projective IA networks of $\mathcal{N}$ (cf. Definition 7), respectively. A solution $\mathfrak{a} = \{a_i\}_{i=1}^n$ of $\mathcal{N}$ is said to be *canonical* if it is a digital solution and $\{I_x(a_i)\}_{i=1}^n$ and $\{I_y(a_i)\}_{i=1}^n$ are the canonical solutions of $\mathcal{N}_x$ and $\mathcal{N}_y$, respectively.

In this section, we show each consistent CDC network has a canonical solution. To this end, we first introduce some notions and properties of digital regions.

Definition 10 *(4-neighbors, 4-connected, connected component).* The 4-neighbors of a pixel p_{ij} are $p_{i-1,j}, p_{i+1,j}, p_{i,j-1}, p_{i,j+1}$. Two pixels p, q are said to be *4-connected* in a digital region a if there exists a series of $k+1$ pixels $p = p^{(0)}, p^{(1)}, \ldots, p^{(k-1)}, p^{(k)} = q$ in a such that $p^{(s)}$ is a 4-neighbor of $p^{(s-1)}$ for any $1 \leqslant s \leqslant k$. We say a digital region a is 4-connected if any two pixels contained in a are 4-connected in a. For digital regions a, c, we say c is a *connected component* of a if c is a 4-connected subset of a, and for any 4-connected subset c' of a, we have $c' \subseteq c$ or $c \cap c'$ contains no pixel.

For example, the 4-neighbors of p in Fig. 6 are the four pixels with mark "c". The digital region a composed of pixels with mark "a" is not 4-connected and has two 4-connected components.

The following proposition shows that 4-connectedness is equivalent to connectedness as far as digital regions are concerned.

Proposition 8. *A bounded digital region is a connected region iff it is* 4*-connected.*

Proof. Recall we require a connected region has a connected interior. This can be proved by using induction on the number of pixels contained in the digital region. Details are omitted. □

4.1. Regular solution

Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is a set of n bounded connected regions. We show $\mathfrak{a}$ can be *regularized* without changing the CDC relations between any two regions in $\mathfrak{a}$.

Write $[x_i^-, x_i^+] \times [y_i^-, y_i^+]$ for $\mathcal{M}(a_i)$, the mbr of a_i. Suppose

$$\alpha_0 < \alpha_1 < \cdots < \alpha_{n_x} \tag{16}$$

is the ordering of real numbers in $\{x_i^-, x_i^+ \colon 1 \leqslant i \leqslant n\}$, and

$$\beta_0 < \beta_1 < \cdots < \beta_{n_y} \tag{17}$$

is the ordering of real numbers in $\{y_i^-, y_i^+ \colon 1 \leqslant i \leqslant n\}$, where $n_x + 1$ and $n_y + 1$ are the cardinalities of $\{x_i^-, x_i^+ \colon 1 \leqslant i \leqslant n\}$ and $\{y_i^-, y_i^+ \colon 1 \leqslant i \leqslant n\}$, respectively. Extending edges of each rectangle $\mathcal{M}(a_i)$ until meeting the boundary of the rectangle $[\alpha_0, \alpha_{n_x}] \times [\beta_0, \beta_{n_y}]$, we partition $[\alpha_0, \alpha_{n_x}] \times [\beta_0, \beta_{n_y}]$ into cells. We next show these cells can be used to compose a solution of $\mathcal{N}$.

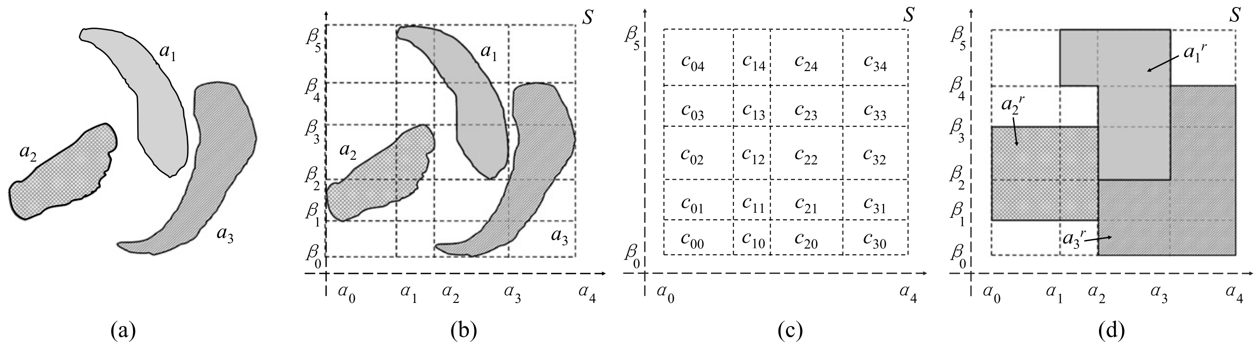


Fig. 7. Illustration of regularization.

Definition 11 *(frame, cell set, regularization).* Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is a set of bounded connected regions. Denote by

$$S(\mathfrak{a}) = [\alpha_0, \alpha_{n_x}] \times [\beta_0, \beta_{n_y}], \tag{18}$$

$$c_{ij} = [\alpha_i, \alpha_{i+1}] \times [\beta_j, \beta_{j+1}] \quad (0 \leqslant i < n_x, 0 \leqslant j < n_y), \tag{19}$$

$$C(\mathfrak{a}) = \{c_{ij}\colon\ 0 \leqslant i < n_x, 0 \leqslant j < n_y\}, \tag{20}$$

$$a_i^r = \bigcup \big\{c \in C(\mathfrak{a})\colon\ c \cap a_i^\circ \neq \varnothing\big\} \quad (i = 1, \ldots, n), \tag{21}$$

$$\mathfrak{a}^r = \big\{a_i^r\big\}_{i=1}^n. \tag{22}$$

We call $S(\mathfrak{a})$ the *frame* of $\mathfrak{a}$, call c_{ij} a *cell* of $\mathfrak{a}$, and call $C(\mathfrak{a})$ the *cell set* of $\mathfrak{a}$, and call $\mathfrak{a}^r$ the *regularization* of $\mathfrak{a}$ (see Fig. 7).

Proposition 9. *For a set of connected regions $\mathfrak{a} = \{a_i\}_{i=1}^n$, we have $a_i \subseteq a_i^r$ and $\mathcal{M}(a_i) = \mathcal{M}(a_i^r)$ for each i, and $S(\mathfrak{a}) = S(\mathfrak{a}^r)$ and $C(\mathfrak{a}) = C(\mathfrak{a}^r)$, where $\mathfrak{a}^r = \{a_i^r\}_{i=1}^n$ is the regularization of $\mathfrak{a}$.*

Proof. Straightforward. □

The regularization is also a solution.

Proposition 10. *Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is a solution to a basic CDC network $\mathcal{N}$. Then $\mathfrak{a}^r = \{a_i^r\}_{i=1}^n$ is also a solution to $\mathcal{N}$.*

Proof. It is clear that each a_i^r is connected. For each pair of i, j, to show $\mathsf{dir}(a_i, a_j) = \mathsf{dir}(a_i^r, a_j^r)$, it suffices to show for each tile name χ that $a_i^\circ \cap \chi(a_j) = \varnothing$ iff $(a_i^r)^\circ \cap \chi(a_j^r) = \varnothing$. Because $\mathcal{M}(a_j^r) = \mathcal{M}(a_j)$, we know by Proposition 1 that $\chi(a_j^r) = \chi(a_j)$ for any χ. Therefore, we need only show $a_i^\circ \cap \chi(a_j) = \varnothing$ iff $(a_i^r)^\circ \cap \chi(a_j) = \varnothing$. If $(a_i^r)^\circ \cap \chi(a_j) = \varnothing$, by $a_i \subseteq a_i^r$, we know $a_i^\circ \cap \chi(a_j) = \varnothing$. On the other hand, suppose $(a_i^r)^\circ \cap \chi(a_j) \neq \varnothing$. There exists a cell $c_{st} \subseteq \chi(a_j)$ such that $c_{st} \subseteq a_i^r$. This is possible iff $c_{st} \cap a_i^\circ$ is nonempty. Therefore, we know $a_i^\circ \cap \chi(a_j) \neq \varnothing$. □

Definition 12 *(regular solution).* A solution $\mathfrak{a} = \{a_i\}_{i=1}^n$ of a basic CDC network $\mathcal{N}$ is called *regular* if $\mathfrak{a}$ is the same as its regularization $\mathfrak{a}^r = \{a_i^r\}_{i=1}^n$.

The regularization of a solution is regular, *i.e.* $(a_i^r)^r = a_i^r$ for each $1 \leqslant r \leqslant n$. This is because $\mathfrak{a}^r$ and $\mathfrak{a}$ have the same frame and the same cell set.

Example 1 *(continued).* Fig. 7 illustrates how to transform a solution $\{a_1, a_2, a_3\}$ (Fig. 7(a)) of the constraint network specified in Fig. 5 into a regular solution $\{a_1^r, a_2^r, a_3^r\}$ (Fig. 7(d)). The frame and the cell set are illustrated in Fig. 7(b) and (c), respectively.

The following proposition asserts that canonical solutions (see Definition 9) are regular solutions.

Proposition 11. *Let $\mathcal{N}$ be a basic CDC network. Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is a canonical solution of $\mathcal{N}$. Then $\mathfrak{a}$ is a regular solution of $\mathcal{N}$.*

Proof. Construct the frame and the cell set of $\mathfrak{a}$ as in Definition 11. By Definition 9, we know $\{I_x(a_i)\}_{i=1}^n$ and $\{I_y(a_i)\}_{i=1}^n$ are the canonical solutions of $\mathcal{N}_x$ and $\mathcal{N}_y$. It is easy to see that $\alpha_i = i$ and $\beta_j = j$ for any $0 \leqslant i \leqslant n_x$ and $0 \leqslant j \leqslant n_y$. Moreover, each cell c_{ij} is exactly the pixel $p_{ij} = [i, i+1] \times [j, j+1]$. Because a_i is a digital region, it is clear $a_i^\circ \cap c_{st} = a_i^\circ \cap p_{st}$ is nonempty iff $p_{st} \subseteq a_i$. This implies $a_i = a_i^r$ for each $1 \leqslant i \leqslant n$. Therefore, $\mathfrak{a}$ is a regular solution of $\mathcal{N}$. □

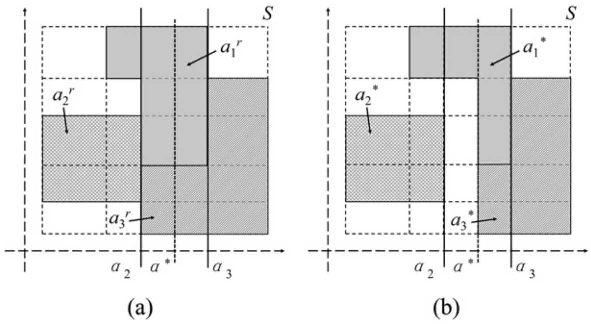


Fig. 8. Illustration of meet-freeing.

4.2. Meet-free solution

Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is a regular solution of a basic CDC network $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$. By Proposition 6 we know $\{I_x(a_i)\}_{i=1}^n$ is a solution of $\{v_i \iota_{ij}^x v_j\}_{i,j=1}^n$. Because $\rho_{ij}^x = \iota_{ij}^x \setminus (\mathrm{m} \cup \mathrm{mi})$, it is possible that $\{I_x(a_i)\}_{i=1}^n$ is not a solution of $\mathcal{N}_x = \{v_i \rho_{ij}^x v_j\}_{i,j=1}^n$. Take the basic CDC network $\mathcal{N}$ described in Fig. 5 as example. Fig. 7(d) gives a regular solution $\{a_i^r\}_{i=1}^3$ of $\mathcal{N}$. Since $I_x(a_2^r)$ and $I_x(a_3^r)$ meet at $x = \alpha_2$, $\{a_i^r\}_{i=1}^3$ is not a canonical solution of $\mathcal{N}$. To transform $\{a_i^r\}_{i=1}^3$ into a canonical solution of $\mathcal{N}$, we introduce the notion of meet-free solution.

Definition 13 *(meet-free solution).* A solution $\mathfrak{a} = \{a_i\}_{i=1}^n$ of $\mathcal{N}$ is *meet-free* if for any i, j, $I_x(a_i)$ does not meet $I_x(a_j)$, and $I_y(a_i)$ does not meet $I_y(a_j)$.

Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is a solution of $\mathcal{N}$ and $\mathfrak{a}^r = \{a_i^r\}_{i=1}^n$ is its regularization (cf. Fig. 7(a) and (d)). Suppose $\mathcal{M}(a_i) = [x_i^-, x_i^+] \times [y_i^-, y_i^+]$ and $\mathcal{M}(a_j) = [x_j^-, x_j^+] \times [y_j^-, y_j^+]$ meet at x direction, *i.e.* $x_i^+ = x_j^-$. Recall $\alpha_0 < \cdots < \alpha_{n_x}$ is the ordering of $\{x_i^-, x_i^+ \colon 1 \leqslant i \leqslant n\}$. We call $\alpha_k = x_i^+$ an *x-meet point* (cf. Fig. 8(a)). Clearly, $k > 0$ and

$$x_i^- \leqslant \alpha_{k-1} < x_i^+ = x_j^- = \alpha_k < \alpha_{k+1} \leqslant x_j^+.$$

We next show how to delete this meet point by transforming $\mathfrak{a}^r$ into another regular solution. Fig. 8 illustrates the process.

Write $\alpha^* = (\alpha_k + \alpha_{k+1})/2$. The line $x = \alpha^*$ divides each cell c_{kl} $(0 \leqslant l < n_y)$ into two equal parts, written in order c_{kl}^- and c_{kl}^+. For each $1 \leqslant s \leqslant n$ and each $0 \leqslant l < n_y$, if $c_{kl} \subseteq a_s^r$ but $c_{k-1,l} \not\subseteq a_s^r$ then delete c_{kl}^- from a_s^r (cf. Fig. 8(a)). The remaining part of a_s^r, written as a_s^*, is still connected, and it is straightforward to show that $\mathfrak{a}^* = \{a_s^*\}_{s=1}^n$ is also a regular solution of $\mathcal{N}$. Such a modification introduces no new meet points. Continuing this process for at most n times, we will have a solution that has no x-meet points. The same procedures can be applied to y-meet points. In this way we obtain a meet-free solution.

The meet-free solution has the same frame but different cell set as $\mathfrak{a}$.

Proposition 12. *Each consistent basic CDC network has a regular solution that is meet-free.*

Proof. See Appendix B. □

The next step is to transform regions appearing in a regular and meet-free solution into digital regions.

4.3. Canonical solution

Suppose $\mathfrak{a} = \{a_s\}_{s=1}^n$ is a regular and meet-free solution of $\mathcal{N}$. As in Section 4.1, we write $\mathcal{M}(a_i) = [x_i^-, x_i^+] \times [y_i^-, y_i^+]$, $S(\mathfrak{a}) = [\alpha_0, \alpha_{n_x}] \times [\beta_0, \beta_{n_y}]$, and $C(\mathfrak{a}) = \{c_{ij} \colon 0 \leqslant i < n_x, 0 \leqslant j < n_y\}$ for the frame and the cell set of $\mathfrak{a}$, respectively. Since $\mathfrak{a}$ is regular, each a_i is composed of a subset of cells in $C(\mathfrak{a})$.

We transform $\mathfrak{a} = \{a_s\}_{s=1}^n$ into a digital solution $\mathfrak{a}^+ = \{a_s^+\}_{s=1}^n$, where each a_s^+ is a digital region contained in $S(\mathfrak{a}^+) = [0, n_x] \times [0, n_y]$ such that

A pixel p_{ij} is contained in a_s^+ iff the cell c_{ij} in $C(\mathfrak{a})$ is contained in a_s.

That is,

$$a_s^+ = \bigcup \{p_{ij} \colon c_{ij} \subseteq a_s, 0 \leqslant i < n_x, 0 \leqslant j < n_y\}. \tag{23}$$

Clearly, a_s^+ is a connected digital region. Fig. 9 illustrates the process.

The following lemma shows that $\mathfrak{a}^+$ is also a solution of $\mathcal{N}$.

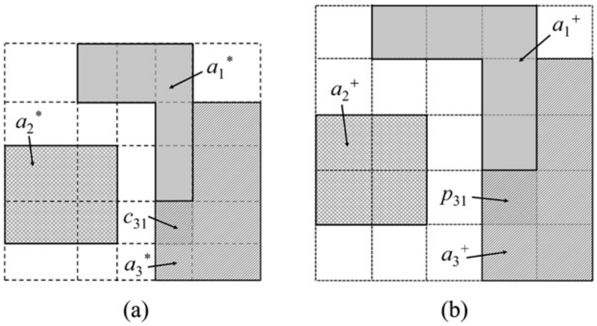


Fig. 9. Transform a regular solution (a) into a digital one (b).

Lemma 1. *Suppose $\mathfrak{a} = \{a_s\}_{s=1}^n$ is a regular meet-free solution of a basic CDC network $\mathcal{N}$. Then $\mathfrak{a}^+ = \{a_s^+\}_{s=1}^n$ constructed as above is also a regular and meet-free solution of $\mathcal{N}$.*

Proof. It is clear that $S(\mathfrak{a}^+) = [0, n_x] \times [0, n_y]$ and $C(\mathfrak{a}^+) = \{p_{st}\colon\ 0 \leqslant s < n_x, 0 \leqslant t < n_y\}$, and $p_{st} \cap (a_i^+)^\circ$ is nonempty iff p_{st} is contained in a_i^+. Hence, $(a_i^+)^r = a_i^+$ for each i, and $\mathfrak{a}^+$ is regular. It is easy to see $I_x(a_i^+) = [s_1, s_2]$ iff $I_x(a_i) = [\alpha_{s_1}, \alpha_{s_2}]$. Therefore, $I_x(a_i^+)$ meets $I_x(a_j^+)$ iff $I_x(a_i)$ meets $I_x(a_j)$. Similar conclusion holds for the y-direction. Since $\mathfrak{a}$ is meet-free, this implies that $\mathfrak{a}^+$ is also meet-free.

For each i, j and each tile name χ, by Eq. (23) we know a pixel p_{st} is contained in a tile $\chi(a_j^+)$ iff the corresponding cell c_{st} is contained in $\chi(a_j)$. Therefore, a pixel p_{st} is contained in $a_i^+ \cap \chi(a_j^+)$ iff c_{st} is contained in $a_i \cap \chi(a_j)$. By definition we have $\mathrm{dir}(a_i, a_j) = \mathrm{dir}(a_i^+, a_j^+)$ for any i, j. Therefore, the assignment $\mathfrak{a}^+ = \{a_s^+\}_{s=1}^n$ is also a solution of $\mathcal{N}$. □

The following lemma proves that $\mathfrak{a}^+$ is a canonical solution.

Lemma 2. *Suppose $\mathfrak{a} = \{a_s\}_{s=1}^n$ is a regular meet-free solution of a basic CDC network $\mathcal{N}$. Then $\{I_x(a_s^+)\}_{s=1}^n$ and $\{I_y(a_s^+)\}_{s=1}^n$ are the canonical solutions of the projective IA networks $\mathcal{N}_x$ and $\mathcal{N}_y$, respectively, where a_s^+ is the digital region defined in Eq. (23).*

Proof. Take the x-direction as example. Because $\mathfrak{a}^+ = \{a_i^+\}_{i=1}^n$ is a solution to $\mathcal{N}$, we know by Proposition 5 that $(I_x(a_i^+), I_x(a_j^+))$ is an instance of $\iota^x(\delta_{ij}) \cap \iota^x(\delta_{ji})^\sim$. Since $\mathfrak{a}^+$ is meet-free, $I_x(a_i^+)$ cannot meet $I_x(a_j^+)$ for any two i, j. This implies that $\mathfrak{I} = \{I_x(a_i^+)\}_{i=1}^n$ is a solution to the basic IA network $\mathcal{N}_x = \{v_i \rho_{ij}^x v_j\}_{i,j=1}^n$, where ρ_{ij}^x is defined by Eq. (14). By the choice of α_i (Eq. (16)), we can easily show $E(\mathfrak{I})$, the set of endpoints of intervals in $\mathfrak{I}$, equals to $[0, n_x] \cap \mathbb{Z}$. By Definition 1, $\mathfrak{I}$ is a canonical set of intervals. Therefore, it is the canonical solution of $\mathcal{N}_x$. □

As a corollary of Lemmas 1 and 2, we know $\mathfrak{a}^+$ is a canonical solution of $\mathcal{N}$.

Theorem 3. *Each consistent basic CDC network has a canonical solution.*

In the next subsection, we show how to construct a canonical solution directly.

4.4. Maximal canonical solution

Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a consistent basic CDC network. By Theorem 2, we know the projective IA networks $\mathcal{N}_x$ and $\mathcal{N}_y$ are consistent.

Lemma 3. *Let $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ be a consistent basic CDC network. Suppose $\{I_i\}_{i=1}^n$ and $\{J_i\}_{i=1}^n$ are the canonical interval solutions of $\mathcal{N}_x$ and $\mathcal{N}_y$, respectively. Let $m_i = I_i \times J_i$ for each i. If $\{a_i\}_{i=1}^n$ is a canonical solution of $\mathcal{N}$, then we have $a_i \subseteq \mathcal{M}(a_i) = m_i$ for $1 \leqslant i \leqslant n$.*

Proof. By Definition 9, we know $I_x(a_i) = I_i$ and $I_y(a_i) = J_i$ for each i. Therefore, $a_i \subseteq \mathcal{M}(a_i) = I_i \times J_i = m_i$. □

For each constraint δ_{ij} in $\mathcal{N}$, we assume

$$\delta_{ij} = \begin{bmatrix} d_{ij}^{NW} & d_{ij}^{N} & d_{ij}^{NE} \\ d_{ij}^{W} & d_{ij}^{O} & d_{ij}^{E} \\ d_{ij}^{SW} & d_{ij}^{S} & d_{ij}^{SE} \end{bmatrix}, \tag{24}$$

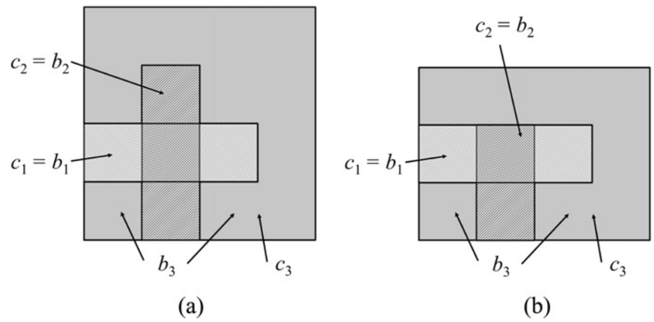


Fig. 10. Illustration of disconnected b_i.

where $d_{ij}^{\chi} \in \{0,1\}$ for each tile name $\chi \in$ TileName. Suppose $\{a_i\}_{i=1}^n$ is a canonical solution of $\mathcal{N}$. By Lemma 3, we know $\mathcal{M}(a_j) = m_j$ for each j. As a consequence of Proposition 1, we know $\chi(a_j) = \chi(\mathcal{M}(a_j)) = \chi(m_j)$ for each $\chi \in$ TileName. Moreover, if $d_{ij}^{\chi} = 0$, then $a_i^{\circ} \cap \chi(m_j) = \varnothing$ because $(a_i, a_j) \in \delta_{ij}$. This means that a_i does not contain pixels which are contained in $\chi(m_j)$. In other words, if $d_{ij}^{\chi} = 0$, then pixels contained in $\chi(m_j)$ are disallowed in a_i.

Lemma 4. *Let $\mathcal{N}$ and m_i be as in Lemma 3. For each i, define*

$$b_i = \bigcup \bigl\{ p_{st} \subseteq m_i\colon\ (\forall j)(\forall \chi)\bigl[d_{ij}^{\chi} = 0 \to p_{st} \not\subseteq \chi(m_j)\bigr]\bigr\}. \tag{25}$$

Suppose $\{a_i\}_{i=1}^n$ is a canonical solution of $\mathcal{N}$. Then $a_i \subseteq b_i$ for each i.

Note that if we write

$$D_i = \bigl\{ p_{st} \subseteq m_i\colon\ (\exists\, j, \chi) \text{ such that } p_{st} \subseteq \chi(m_j) \text{ and } d_{ij}^{\chi} = 0 \bigr\}, \tag{26}$$

then D_i is the set of all pixels in m_i that cannot appear in the instantiation of v_i in any canonical solution as a result of violating some constraint in $\mathcal{N}$. It is easy to see that

$$b_i = \bigcup \{ p_{st} \subseteq m_i\colon\ p_{st} \neq\in D_i \}. \tag{27}$$

The following example shows that b_i is not always a connected region.

Example 2. Let $\mathcal{N} = \{v_1(W:O:E)v_2,\ v_2(N:O:S)v_1,\ v_1 O v_3,\ v_3(N:NE:E:SE:S)v_1,\ v_2 O v_3,\ v_3(W:NW:N:NE:E)v_2\}$. By computing the canonical solutions of $\mathcal{N}_x$ and $\mathcal{N}_y$, we have

$$m_1 = [0,3] \times [1,2], \qquad m_2 = [1,2] \times [0,3], \qquad m_3 = [0,4] \times [0,4].$$

By excluding impossible pixels from m_i, we obtain the three digital regions b_1, b_2, b_3 as shown in Fig. 10(a). Note that $b_1 = m_1$ and $b_2 = m_2$ are both connected. However, b_3 has two connected components, one consists of the pixel p_{00}, the other is obtained by excluding p_{00} from b_3. Write c_3 for this component. It is clear that $\mathcal{M}(c_3) = m_3$. Let $c_1 = b_1$, $c_2 = b_2$. The assignment $\{c_1, c_2, c_3\}$ clearly satisfies all the constraints in $\mathcal{N}$.

In the following, we show that, if $\mathcal{N}$ is consistent, then the digital region b_i defined in Eq. (25) has a unique connected component (cf. Definition 10) c_i such that $\mathcal{M}(c_i) = \mathcal{M}(b_i) = m_i$. We first note that each connected sub-region of b_i is contained in a unique connected component of b_i. This is because different components of b_i have no common interior points.

Lemma 5. *Let $\mathcal{N}$ and m_i be as in Lemma 3, and let b_i be as in Eq. (25). Suppose $\{a_i\}_{i=1}^n$ is a canonical solution of $\mathcal{N}$. Let c_i be the connected component of b_i which contains a_i. Then $\mathcal{M}(c_i) = m_i$.*

Proof. Because $a_i \subseteq c_i \subseteq b_i \subseteq m_i$ and $\mathcal{M}(a_i) = m_i$, we know $\mathcal{M}(c_i) = m_i$. □

The above lemma shows that, if $\mathcal{N}$ is consistent, then each b_i has a connected component c_i such that $\mathcal{M}(c_i) = \mathcal{M}(b_i) = m_i$. Is this component unique? The answer is yes! This is because, if two connected digital regions have the same mbr, then they share one pixel in common.

Lemma 6. *Let a and b be two connected digital regions. If $a^{\circ} \cap b^{\circ} = \varnothing$, then $\mathcal{M}(a) \neq \mathcal{M}(b)$.*

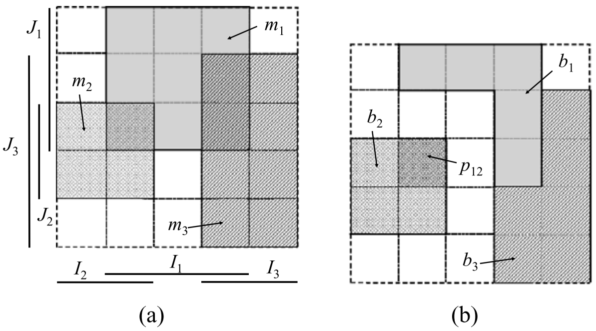


Fig. 11. Canonical interval solutions (a) and the maximal canonical solution (b).

Proof. Without loss of generality, suppose $\mathcal{M}(a) = [0, n_1] \times [0, n_2]$, where n_1, n_2 are positive integers. Suppose $a^\circ \cap b^\circ = \varnothing$ and $\mathcal{M}(a) = \mathcal{M}(b)$. Since $\mathcal{M}(a) = [0, n_1] \times [0, n_2]$, a contains a pixel p_{0s} and a pixel $p_{n_1-1,t}$ for some integers $0 \leqslant s, t < n_2$. Because a is a connected digital region, there is a path π (*i.e.* a sequence of 4-connected pixels in a) that connects p_{0s} to $p_{n_1-1,t}$ (cf. Proposition 8). Clearly, π separates $\mathcal{M}(a)$ into at least two disjoint components, each of which is contained in a rectangle smaller than $\mathcal{M}(a)$. Since $a^\circ \cap b^\circ = \varnothing$, b is contained in one such component. Therefore, $\mathcal{M}(b)$ is smaller than $\mathcal{M}(a)$. A contradiction. □

We note that the above property does not hold for general connected regions. Let $m = [0, 1] \times [0, 1]$. The diagonal of m separates m into two triangles a_1, a_2. Clearly, a_1 and a_2 have no common interior point, but they have the same mbr.

As a corollary of Lemma 6, we have

Lemma 7. *Let b be a digital region. There exists at most one connected component, written c, such that $\mathcal{M}(c) = \mathcal{M}(b)$.*

Proof. Suppose c_1 and c_2 are two different connected components of b. Clearly, $c_1^\circ \cap c_2^\circ = \varnothing$. By Lemma 6 we know $\mathcal{M}(c_1) \neq \mathcal{M}(c_2)$. This means no two different components have the same mbr. In particular, there exists at most one connected component which has the same mbr as b. □

If $\mathcal{N}$ is consistent, then by Lemma 5, we know such a unique component of b_i does exist. In fact, suppose $\{a_i\}_{i=1}^n$ is an arbitrary canonical solution of $\mathcal{N}$. Then the connected component c_i of b_i which contains a_i is the unique component of b_i such that $\mathcal{M}(c_i) = \mathcal{M}(b_i)$.

The following lemma shows that $\{c_i\}_{i=1}^n$ is also a solution of $\mathcal{N}$.

Lemma 8. *Let $\mathcal{N}$ and m_i be as in Lemma* 3*, and let b_i be as in Eq.* (25)*. Assume c_i is the unique connected component of b_i such that $\mathcal{M}(c_i) = m_i$. Then $\{c_i\}_{i=1}^n$ is a canonical solution of $\mathcal{N}$.*

Proof. To prove that $\{c_i\}_{i=1}^n$ is a solution of $\mathcal{N}$, we need only prove $\mathrm{dir}(c_i, c_j) = \delta_{ij}$ for each pair of i, j. By Proposition 1 and $\mathcal{M}(c_j) = m_j$, we know $\chi(c_j) = \chi(m_j)$ for each tile name $\chi \in \textsc{TileName}$. To show $\mathrm{dir}(c_i, c_j) = \delta_{ij}$, we need only prove that $c_i^\circ \cap \chi(m_j) = \varnothing$ iff $d_{ij}^\chi = 0$ for each $\chi \in \textsc{TileName}$.

Suppose $\{a_i\}_{i=1}^n$ is an arbitrary canonical solution of $\mathcal{N}$. By Lemma 5, we know $a_i \subseteq c_i$ for each i. If $c_i^\circ \cap \chi(m_j) = \varnothing$, then by $a_i \subseteq c_i$, we know $a_i^\circ \cap \chi(m_j) = \varnothing$. Because (a_i, a_j) is an instance of δ_{ij}, this implies $d_{ij}^\chi = 0$. On the other hand, if $c_i^\circ \cap \chi(m_j) \neq \varnothing$, then by $c_i \subseteq b_i$, we know b_i and $\chi(m_j)$ has a common pixel p_{st}. By the definition of b_i (cf. Eq. (25)) we know $d_{ij}^\chi \neq 0$.

This shows that $\{c_i\}_{i=1}^n$ is a solution of $\mathcal{N}$. Since each c_i is a connected digital region, $\{c_i\}_{i=1}^n$ is a digital solution of $\mathcal{N}$. By $\mathcal{M}(c_i) = m_i$, we know $I_x(c_i) = I_i$ and $I_y(c_i) = J_i$ for each i, where $\{I_i\}_{i=1}^n$ and $\{J_i\}_{i=1}^n$ are, as in Lemma 3, the canonical solutions of $\mathcal{N}_x$ and $\mathcal{N}_y$, respectively. This implies that $\{c_i\}_{i=1}^n$ is a canonical solution of $\mathcal{N}$ (see Definition 9). □

Actually, this is also the maximal canonical solution of $\mathcal{N}$.

Theorem 4. *Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a consistent basic CDC network. Let c_i be the unique connected component of b_i (cf. Eq.* (25)*) such that $\mathcal{M}(c_i) = m_i$. Then $\{c_i\}_{i=1}^n$ is the maximal canonical solution of $\mathcal{N}$.*

Proof. This is because, for any canonical solution $\{a_i\}_{i=1}^n$, by Lemma 5, we know $a_i \subseteq c_i$ for each i. □

An example is given in Fig. 11(a) to illustrate the procedures.

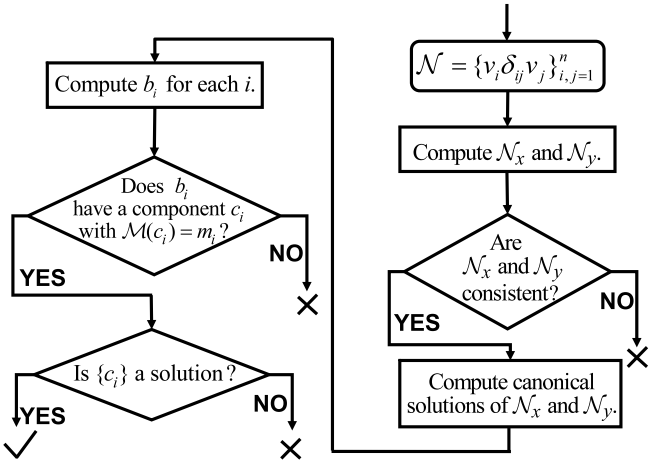


Fig. 12. Flowchart of the main algorithm.

Example 1 *(continued).* For the network specified in Fig. 5, we have

$$x_2^- < x_1^- < x_2^+ < x_3^- < x_1^+ < x_3^+, \tag{28}$$

$$y_3^- < y_2^- < y_1^- < y_2^+ < y_3^+ < y_1^+. \tag{29}$$

The canonical interval solutions of $\mathcal{N}_x$ and $\mathcal{N}_y$ are illustrated in Fig. 11(a). Note that d_{12}^O (cf. Eq. (32)) is 0. This excludes pixel $p_{12} \subseteq O(m_2) = m_2$ from m_1 (see Fig. 11(a)). Note that each b_i obtained in this example happens to be connected (see Fig. 11(b)). The maximal canonical solution of $\mathcal{N}$ is $\{b_1, b_2, b_3\}$.

So far we have shown that, if $\mathcal{N}$ is consistent, then its maximal canonical solution $\{c_i\}_{i=1}^n$ can be constructed as follows:

- Construct the canonical solutions $\{I_i\}_{i=1}^n$ and $\{J_i\}_{i=1}^n$ of the projective interval networks $\mathcal{N}_x$ and $\mathcal{N}_y$;
- Exclude impossible pixels from $m_i = I_i \times J_i$ and obtain a digital region $b_i \subseteq m_i$ (cf. Eq. (25));
- Find the unique connected component c_i of b_i such that $\mathcal{M}(c_i) = m_i$.

What if the consistency of $\mathcal{N}$ is unknown? The next example suggests we can use the above procedures to determine the consistency of $\mathcal{N}$.

Example 2 *(continued).* Suppose the constraint $v_2(N : O : S)v_1$ is replaced with $v_2(O : S)v_1$. See Fig. 10(b) for the illustration of $\{b_1, b_2, b_3\}$ in this case. Note that c_3 and c_2 fail to satisfy constraint $v_3(W : NW : N : NE : E)v_2$. By Theorem 4 we know the revised network cannot be consistent.

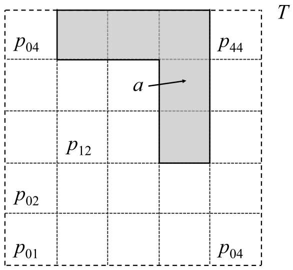
5. A consistency checking algorithm

In this section, we describe our algorithm for checking the consistency of basic CDC networks. Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a basic CDC network. Recall we assume $\mathcal{N}$ is *complete*, *i.e.* a basic constraint δ_{ij} is assigned for each pair of variables v_i, v_j. To examine the consistency of $\mathcal{N}$, we first compute the projective IA networks $\mathcal{N}_x$ and $\mathcal{N}_y$. If either is inconsistent, then $\mathcal{N}$ is inconsistent. Assume both $\mathcal{N}_x$ and $\mathcal{N}_y$ are consistent. We compute their canonical interval solutions $\{I_i\}_{i=1}^n$ and $\{J_i\}_{i=1}^n$, and construct a rectangle $m_i = I_i \times J_i$ for each i. We then continue to compute the digital region b_i according to Eq. (25) and determine if b_i has a component c_i such that $\mathcal{M}(c_i) = m_i$. If such a component does not exist for some i, then $\mathcal{N}$ is inconsistent. Otherwise, we check if $\{c_i\}_{i=1}^n$ is a solution of $\mathcal{N}$. If the answer is yes, then $\mathcal{N}$ is consistent and $\{c_i\}_{i=1}^n$ is its maximal canonical solution. If the answer is no, then $\mathcal{N}$ must be inconsistent.

Fig. 12 gives the flowchart of the algorithm.

5.1. An $O(n^4)$ consistency checking algorithm

In this section, we give a detailed description of our consistency checking algorithm, and analyze its computational complexity.



(a)

$$B(a) = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(b)

Fig. 13. A digital region a (a) and its Boolean matrix $B(a)$ (b).

Step 1. Projective IA networks

For a pair of basic CDC constraints $\{x_i \delta_{ij} x_j, x_j \delta_{ji} x_i\}$, the x- and y-projective IA relations ρ_{ij}^x and ρ_{ij}^y (Eqs. (14) and (15)) can be computed in constant time. A method is described in Appendix A. So the projective IA networks $\mathcal{N}_x$ and $\mathcal{N}_y$ can be constructed in $O(n^2)$ time (see Appendix A, Proposition 24). If $\mathcal{N}_x$ or $\mathcal{N}_y$ is inconsistent, which can be checked in cubic time by a path-consistency algorithm [1,45], then $\mathcal{N}$ is also inconsistent.

Example 1 *(continued).* For the basic CDC network specified in Fig. 5, we have

$$\mathcal{N}_x = \{v_1 \mathsf{oi} v_2, v_2 \mathsf{o} v_1, v_1 \mathsf{o} v_3, v_3 \mathsf{oi} v_1, v_2 \mathsf{p} v_3, v_3 \mathsf{pi} v_2\},$$

$$\mathcal{N}_y = \{v_1 \mathsf{oi} v_2, v_2 \mathsf{o} v_1, v_1 \mathsf{oi} v_3, v_3 \mathsf{o} v_1, v_2 \mathsf{d} v_3, v_3 \mathsf{di} v_2\}.$$

It is straightforward to prove that $\mathcal{N}_x$ and $\mathcal{N}_y$ are path-consistent.

Step 2. Canonical interval solutions

Suppose $\mathcal{N}_x$ and $\mathcal{N}_y$ are consistent. Their canonical solutions $\{I_i\}_{i=1}^n$ and $\{J_i\}_{i=1}^n$ can be constructed in cubic time by using Algorithm 1.

Algorithm 1 THE CANONICAL SOLUTION OF A BASIC IA NETWORK

Input: A basic consistent IA network $\mathcal{N} = \{v_i \rho_{ij} v_j\}_{i,j=1}^n$.
Output: The unique canonical solution of $\mathcal{N}$.
  $V \leftarrow \{x_i^-, x_i^+\}_{i=1}^n$;
  for each $p, q \in V$ **do**
    Calculate $Order[p, q]$ by constraints in $\mathcal{N}$ and Table 2;
  sort V by $Order$ ascending;
  $k \leftarrow 0$; $p \leftarrow$ the smallest element in V; $r[p] \leftarrow k$;
  while p is not the greatest element in V **do**
    $q \leftarrow$ the next element of p in V;
    if $Order[q, p] =$ '$>$' **then**
      $k \leftarrow k + 1$;
    $p \leftarrow q$, $r[p] \leftarrow k$;
  Output $\{[r[x_i^-], r[x_i^+]]\}_{i=1}^n$.

Write $I_i = [x_i^-, x_i^+]$ and $J_i = [y_i^-, y_i^+]$. By the definition of canonical interval solution, we know $x_i^-, x_i^+, y_i^-, y_i^+$ are integers between 0 and $2n - 1$. Write $n_x = \max\{x_i^+\}_{i=1}^n$ and $n_y = \max\{y_i^+\}_{i=1}^n$. Let $T = [0, n_x] \times [0, n_y]$ be the frame and define $m_i = [x_i^-, x_i^+] \times [y_i^-, y_i^+]$ for each i. Clearly, $m_i \subseteq T$.

Example 1 *(continued).* For the basic CDC network specified in Fig. 5, the canonical solutions of $\mathcal{N}_x$ and $\mathcal{N}_y$ are (see Fig. 11(a)) $\{I_1 = [1, 4], I_2 = [0, 2], I_3 = [3, 5]\}$ and $\{J_1 = [2, 5], J_2 = [1, 3], J_3 = [0, 4]\}$, respectively. Moreover, the frame $T = [0, 5] \times [0, 5]$, and $m_1 = [1, 4] \times [2, 5]$, $m_2 = [0, 2] \times [1, 3]$, $m_3 = [3, 5] \times [0, 4]$.

In the remainder of this paper, we often represent a digital region $a \subseteq T$ as an $n_x \times n_y$ Boolean matrix $B(a)$ as follows:

$$B(a)[k, l] = \begin{cases} 1, & \text{if } p_{kl} \subseteq a; \\ 0, & \text{otherwise.} \end{cases} \tag{30}$$

Example 3. Fig. 13(a) illustrates a digital region a contained in $T = [0, 5] \times [0, 5]$, and Fig. 13(b) shows the Boolean matrix $B(a)$ that represents a.

Table 4
Excluding impossible pixels from m_1.

$\begin{pmatrix}0&0&0&0&0\\0&0&0&0&0\\1&1&0&0&0\\1&1&0&0&0\\0&0&0&0&0\end{pmatrix}$	$\begin{pmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\1&1&0&0&0\end{pmatrix}$	$\begin{pmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&1&1&1\end{pmatrix}$
$B(T\cap O(m_2))$	$B(T\cap S(m_2))$	$B(T\cap SE(m_2))$
$\begin{pmatrix}0&0&0&0&0\\1&1&1&0&0\\1&1&1&0&0\\1&1&1&0&0\\1&1&1&0&0\end{pmatrix}$	$\begin{pmatrix}0&1&1&1&0\\0&1&1&1&0\\0&1&1&1&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$	$\begin{pmatrix}0&1&1&1&0\\0&0&0&1&0\\0&0&0&1&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$
$B(T\cap W(m_3))$	$B(m_1)$	$B(b_1)$

Note we address the elements of $B(a)$ from lower left corner to top right corner. In particular, the lower left corner of $B(a)$ in Fig. 13(a) is addressed as the $(0,0)$-element, *i.e.* $B(a)[0,0]$. This is in accord with the orderings of the endpoints in the canonical interval solutions.

Step 3. Excluding impossible pixels

As in Eq. (25), we write

$$b_i=\bigcup\bigl\{p_{st}\subseteq m_i\colon (\forall j)(\forall\chi)\bigl[d_{ij}^{\chi}=0\rightarrow p_{st}\not\subseteq\chi(m_j)\bigr]\bigr\}. \tag{31}$$

This means that a pixel p_{st} in m_i is contained in the digital region b_i iff p_{st} is not contained in any tile $\chi(m_j)$ with $d_{ij}^{\chi}=0$. To compute b_i, an intuitive method is checking for each pixel p in m_i whether p is contained in b_i.

Proposition 13. *For each pixel p contained in m_i, it needs at most $O(n)$ time to determine if p is in b_i.*

Proof. To determine if p is in b_i, by definition we need to check if p is not contained in $\chi(m_j)$ for all j,χ with $d_{ij}^{\chi}=0$. Note that it needs constant time to decide if a pixel is contained in a tile. Since there are at most $O(n)$ different tiles $\chi(m_j)$, it needs at most $O(n)$ time to determine if p is contained in b_i. □

Note that there are at most $O(n^2)$ pixels contained in m_i. As a consequence, it needs at most $O(n^3)$ time to compute b_i for each i. Therefore, we need at most $O(n^4)$ time to compute all b_i (see Algorithm 2).

Algorithm 2 EXCLUDING IMPOSSIBLE PIXELS

Input: A basic CDC network $\mathcal{N}=\{v_i\delta_{ij}v_j\}_{i,j=1}^{n}$, a canonical set of rectangles $\{m_i\}_{i=1}^{n}$ and its frame $T=[0,n_x]\times[0,n_y]$.
Output: A set of $n_x\times n_y$ Boolean matrices $\{B_i\}_{i=1}^{n}$.
  for each $1\leqslant i\leqslant n$ **do**
    for each pixel $p=[s,s+1]\times[t,t+1]$ in T **do**
      if p is in m_i **then**
        $B_i[s,t]\leftarrow 1$;
      else
        $B_i[s,t]\leftarrow 0$;
    for each $1\leqslant j\leqslant n$ and $j\neq i$ **do**
      for each $\chi\in$ TILENAME **do**
        if $\delta_{ij}^{\chi}=0$ **then**
          for each pixel $p=[s,s+1]\times[t,t+1]$ contained in $m_i\cap\chi(m_j)$ **do**
            $B_i[s,t]\leftarrow 0$;
  Output $\{B_i\}_{i=1}^{n}$.

Example 1 *(continued).* For the basic CDC network specified in Fig. 5, b_1 is the digital region obtained by excluding pixels $p_{12},p_{22},p_{13},p_{23}$ from $m_1=[1,4]\times[2,5]$, b_2 is exactly $m_2=[0,2]\times[1,3]$, and b_3 is the digital region obtained by excluding pixels p_{32},p_{33} from $m_3=[3,5]\times[0,4]$ (see Fig. 11(b)).

Take b_1 as example. Recall $T=[0,5]\times[0,5]$. There are four different $\chi(m_j)$ such that $T\cap\chi(m_j)$ is not a degenerate rectangle and $d_{1j}^{\chi}=0$, *viz.* $T\cap O(m_2)=[0,2]\times[1,3]$, $T\cap S(m_2)=[0,2]\times[0,1]$, $T\cap SE(m_2)=[2,5]\times[0,1]$, $T\cap W(m_3)=[0,3]\times[0,4]$ (see Fig. 11(a)). Table 4 illustrates the process of computing $B(b_1)$. The digital region b_1 can be computed by using Eq. (30), shown in Fig. 11. Note b_1 is connected.

In Section 5.2, however, we will show that this step can be improved to $O(n^3)$.

Step 4. Connected components

We further compute connected components of b_i for each i. Applying a general Breadth-First Search algorithm, we can find all connected components of b_i and determine if their mbrs are m_i in $O(n^2)$ time. Algorithm 3 is an optimised algorithm, which tries to find the desired connected component. For each pixel p at the bottom of m_i, from left to right, if p is contained in b_i, the algorithm finds the connected component of b_i which contains p. If this component has mbr m_i, then the algorithm returns this component as c_i and stops; otherwise, it goes to the next pixel at the bottom of m_i.

Algorithm 3 CALCULATING THE CONNECTED COMPONENT

Input: The $n_x \times n_y$ Boolean matrix B (of a digital region b), and a rectangle $m = [x^-, x^+] \times [y^-, y^+]$,
Output: If b has a connected component with mbr m, return its Boolean matrix C, otherwise return false.

```
for each x⁻ ⩽ r < x⁺ do
  if B[r, y⁻] = 1 then
    CurrentComponent ← ∅;
    Visiting ← {(r, y⁻)};
    MaxY ← y⁻; MinX ← r; MaxX ← r;
    while Visiting is not empty do
      let (s, t) be an element of Visiting;
      CurrentComponent ← CurrentComponent ∪ {(s, t)};
      Visiting ← Visiting \ {(s, t)};
      B[s, t] ← 0;
      MinX ← min(MinX, s);
      MaxX ← max(MaxX, s);
      MaxY ← max(MaxY, t);
      for each (s′, t′) in {(s − 1, t), (s + 1, t), (s, t − 1), (s, t + 1)} do
        if B[s′, t′] = 1 then
          Visiting ← Visiting ∪ {(s′, t′)};
  if MinX = x⁻ and MaxX = x⁺ − 1 and MaxY = y⁺ − 1 then
    for each (s, t) in CurrentComponent do
      C[s, t] ← 1;
    Output C;
Output false.
```

Note that b_i has at most one component whose mbr is m_i (Lemma 7). If no such component exists for some i, then $\mathcal{N}$ is inconsistent. Otherwise, let c_i be the unique connected component of b_i such that $\mathcal{M}(c_i) = m_i$.

Clearly, this step needs only $O(n^3)$ time.

Example 1 *(continued).* For the basic CDC network specified in Fig. 5, b_1, b_2, and b_3 happen to be connected (see Fig. 11(b)). This means, $c_i = b_i$ for $i = 1, 2, 3$.

Step 5. Checking a possible solution

The last step is checking if $\{c_i\}_{i=1}^n$ is a solution of $\mathcal{N}$. Note that if the answer is yes, then $\{c_i\}_{i=1}^n$ is the maximal canonical solution of $\mathcal{N}$ (see Theorem 4).

For each pair of c_i and c_j, we should check if $\mathsf{dir}(c_i, c_j) = \delta_{ij}$. Write

$$\delta_{ij} = \begin{bmatrix} d_{ij}^{NW} & d_{ij}^{N} & d_{ij}^{NE} \\ d_{ij}^{W} & d_{ij}^{O} & d_{ij}^{E} \\ d_{ij}^{SW} & d_{ij}^{S} & d_{ij}^{SE} \end{bmatrix}. \tag{32}$$

We need to check for each $\chi \in$ TILENAME whether the following equation holds

$$d_{ij}^{\chi} = 0 \quad \Leftrightarrow \quad c_i^{\circ} \cap \chi(m_j) = \varnothing. \tag{33}$$

Because c_i and m_j are all digital regions, $c_i^{\circ} \cap \chi(m_j)$ is nonempty iff there exists a pixel p which is contained in both c_i and $\chi(m_j)$. Therefore, we need only to check for each $\chi \in$ TILENAME whether the following equation holds

$$d_{ij}^{\chi} = 1 \quad \Leftrightarrow \quad c_i \cap \chi(m_j) \text{ contains a pixel.} \tag{34}$$

Recall c_i is a connected component of b_i and $\mathcal{M}(c_i) = \mathcal{M}(b_i) = m_i$. When $d_{ij}^{\chi} = 0$, by the definition of b_i, we know no pixel is contained in both b_i and $\chi(m_j)$. This means $c_i \cap \chi(m_j)$ contains no pixel. The condition in Eq. (34) always holds. On the other hand, suppose $d_{ij}^{\chi} = 1$. Note that the rectangle m_i has been computed in Step 2 and each tile $\chi(m_j)$ can be computed from m_j in constant time. Whether $m_i \cap \chi(m_j)$ contains a pixel can also be checked in constant time. This is because $m_i \cap \chi(m_j)$ is a (possibly degenerate) rectangle. If $m_i \cap \chi(m_j)$ contains no pixel, *i.e.*, it is a degenerate rectangle, then $c_i \cap \chi(m_j)$ contains no pixel. Suppose $m_i \cap \chi(m_j)$ is a non-degenerate rectangle. The next proposition shows that we can determine whether $c_i \cap \chi(m_j)$ contains a pixel in $O(n)$ time.

Proposition 14. *Let m_i, c_i be as constructed in Step 2 and Step 4, respectively. Let χ be a tile name. Suppose $d_{ij}^{\chi} = 1$ and $m_i \cap \chi(m_j)$ is a non-degenerate rectangle. Then whether $c_i \cap \chi(m_j)$ contains a pixel can be checked in $O(n)$ time.*

Proof. Write $m_i \cap \chi(m_j) = [x^-, x^+] \times [y^-, y^+]$. We need not check for every pixel p in $m_i \cap \chi(m_j)$ whether $p \subseteq c_i$. Instead, we need only check this for all *boundary* pixels of $m_i \cap \chi(m_j)$, where a pixel in a rectangle is called a boundary pixel if it has a 4-neighbor that is not contained in the rectangle. Formally, the boundary pixels of $m_i \cap \chi(m_j)$ has the form $p_{k,l}$ with $(k,l) \in H_1 \cup H_2$, where

$$H_1 = \{(k,l)\colon k \in \{x^-, x^+ - 1\} \text{ and } y^- \leqslant l < y^+\},$$
$$H_2 = \{(k,l)\colon x^- \leqslant k < x^+ \text{ and } l \in \{y^-, y^+ - 1\}\}.$$

We justify the above statement as follows. If $m_i \cap \chi(m_j) = m_i$, then by $\mathcal{M}(c_i) = m_i$ we know there exists a boundary pixel which is contained in c_i. Otherwise, $m_i \cap \chi(m_j)$ is a rectangle strictly contained in m_i. Because $\mathcal{M}(c_i) = m_i$, this implies that c_i contains a pixel p out of $m_i \cap \chi(m_j)$. If c_i contains no boundary pixel of $m_i \cap \chi(m_j)$, then c_i, as a connected digital region, contains no pixel of $m_i \cap \chi(m_j)$ at all.

Since $H_1 \cup H_2$ contains $O(n)$ pixels and checking if a pixel is contained in c_i needs constant time, $\mathrm{dir}(c_i, c_j) = \delta_{ij}$ can be checked in $O(n)$ time. □

Therefore, we can determine in $O(n)$ time whether $\mathrm{dir}(c_i, c_j) = \delta_{ij}$ for each pair of i, j. As a consequence, whether $\{c_i\}_{i=1}^n$ is a solution of $\mathcal{N}$ can checked in cubic time. This means Step 5 can be carried out in cubic time.

Example 1 *(continued).* For the basic CDC network specified in Fig. 5, we know $b_1 = c_1$ and $b_2 = c_2$ (see Fig. 11(b)). To show $\mathrm{dir}(b_1, b_2) = \delta_{12}$, we need to check if b_1 and $\chi(m_2)$ have common pixels for all χ with $d_{12}^\chi = 1$, *i.e.* $\chi = N, NE, E$. Since pixels p_{14}, p_{24}, and p_{32} are contained in $N(m_2)$, $NE(m_2)$, and $E(m_2)$, respectively. We know (b_1, b_2) is an instance of δ_{12}.

The above example shows that, applying our main algorithm to a consistent (complete) basic CDC network, we can construct its maximal canonical solution. In Section 6.1, we will give another example (Example 4), which illustrates how the main algorithm deals with inconsistent CDC networks.

Recall that only Step 3 needs at most $O(n^4)$ time, the algorithm determines the consistency of a (complete) basic CDC network in $O(n^4)$ time.

5.2. A cubic improvement

In this subsection, we improve the main algorithm to cubic. This is achieved by an $O(n^2)$ improvement for computing each b_i (see Eq. (31)) in Step 3. As a consequence, all b_i can be computed in cubic time.

For each i, recall b_i is defined as below (cf. Eq. (31)),

$$b_i = \bigcup \{p_{st} \subseteq m_i\colon (\forall j)(\forall \chi)[d_{ij}^\chi = 0 \rightarrow p_{st} \not\subseteq \chi(m_j)]\}. \tag{35}$$

The Boolean matrix $B(b_i)$ (see Eq. (30)) of b_i can be computed as follows, where $B(T \cap \chi(m_j))$ is the zero matrix if $T \cap \chi(m_i)$ is a degenerate rectangle.

Proposition 15. *For $0 \leqslant k < n_x$ and $0 \leqslant l < n_y$, we have*

$$B(b_i)[k,l] = \begin{cases} 1, & \text{if } B(m_i)[k,l] = 1 \text{ and } Q_i[k,l] = 0; \\ 0, & \text{otherwise}, \end{cases} \tag{36}$$

where

$$Q_i = \sum \{B(T \cap \chi(m_j))\colon d_{ij}^\chi = 0 \text{ and } j \neq i\}. \tag{37}$$

Proof. We first note that a pixel $p_{kl} \subseteq T$ is contained in a digital region $a \subseteq T$ iff $B(a)[k,l] = 1$. Suppose $B(m_i)[k,l] = 1$ and $Q_i[k,l] = 0$. The first equation implies that p_{kl} is a pixel contained in m_i. The latter is equivalent to saying that $B(T \cap \chi(m_j))[k,l] = 0$ for any $j \neq i$ and any χ with $d_{ij}^\chi = 0$. This means, for any $j \neq i$ and any χ, if $d_{ij}^\chi = 0$ then p_{kl} is not contained in $\chi(m_j)$. By the definition of b_i, this implies that $p_{kl} \subseteq b_i$. Hence $B(b_i)[k,l] = 1$.

Otherwise, suppose $B(m_i)[k,l] = 0$ or $Q_i[k,l] > 0$. We show $B(b_i)[k,l] = 0$. Note that by $b_i \subseteq m_i$ we know $B(b_i)[k,l] = 0$ if $B(m_i)[k,l] = 0$. Suppose $B(m_i)[k,l] = 1$ and $Q_i[k,l] > 0$. By the definition of Q_i, we know $B(T \cap \chi(m_j))[k,l] = 1$ for some j, χ such that $j \neq i$ and $d_{ij}^\chi = 0$. This implies that p_{kl} is contained in $\chi(m_j)$. Note that p_{kl} is also contained in m_i for $B(m_i)[k,l] = 1$. By the definition of b_i, we know p_{kl} is not contained in b_i. Hence $B(b_i)[k,l] = 0$. □

By the above proposition, we know $B(b_i)$ can be computed from $B(m_i)$ and Q_i in $O(n^2)$ time. We next show how to compute each Q_i in $O(n^2)$ time. We note that Q_i is not a Boolean matrix. Some of its entries may be integers larger than 1.

To this end, we introduce the following operations on matrices.

Table 5
A matrix N (a) and its difference matrix $\mathsf{diff}(N)$ (b) and $\mathsf{acc}(\mathsf{diff}(N))$, the cumulative matrix of $\mathsf{diff}(N)$ (c).

(a)	(b)	(c)
$\begin{pmatrix} 0&0&0&0&0\\ 0&1&1&1&0\\ 0&1&1&1&0\\ 0&1&1&1&0\\ 0&0&0&0&0 \end{pmatrix}$	$\begin{pmatrix} 0&0&0&0&0\\ 0&1&0&0&-1\\ 0&1&0&0&-1\\ 0&1&0&0&-1\\ 0&0&0&0&0 \end{pmatrix}$	$\begin{pmatrix} 0&0&0&0&0\\ 0&1&1&1&0\\ 0&1&1&1&0\\ 0&1&1&1&0\\ 0&0&0&0&0 \end{pmatrix}$

Definition 14 *(cumulative matrix and difference matrix).* For a matrix N, we define its *cumulative matrix* as

$$\mathsf{acc}(N)[k,l] = \sum_{t=0}^{k} N[t,l] \tag{38}$$

and its *difference matrix* as

$$\mathsf{diff}(N)[k,l] = \begin{cases} N[k,l], & \text{if } k = 0; \\ N[k,l] - N[k-1,l], & \text{otherwise.} \end{cases} \tag{39}$$

The cumulative matrix $\mathsf{acc}(N)$ can be computed column by column. We add the first column to the second one, and then add the updated second column to the third, *etc.* (see Table 5 for an example). In this way, cumulative matrix $\mathsf{acc}(N)$ can be computed linearly in the number of elements of N.

It is easy to verify that

- $N = \mathsf{acc}(\mathsf{diff}(N)) = \mathsf{diff}(\mathsf{acc}(N))$;
- the difference operation is additive, *i.e.* $\mathsf{diff}(N_1 + N_2) = \mathsf{diff}(N_1) + \mathsf{diff}(N_2)$.

Therefore, we have

$$\begin{aligned} Q_i &= \mathsf{acc}\big(\mathsf{diff}(Q_i)\big) \\ &= \mathsf{acc}\Big(\mathsf{diff}\Big(\sum\big\{B\big(T \cap \chi(m_j)\big)\colon d_{ij}^{\chi} = 0 \text{ and } j \neq i\big\}\Big)\Big) \\ &= \mathsf{acc}\Big(\sum\big\{\mathsf{diff}\big(B\big(T \cap \chi(m_j)\big)\big)\colon d_{ij}^{\chi} = 0 \text{ and } j \neq i\big\}\Big). \end{aligned}$$

Note that each digital region $T \cap \chi(m_j)$ is actually a (possibly degenerate) rectangle. The following proposition uses this property to show that the number of nonzero elements in $\mathsf{diff}(B(T \cap \chi(m_j)))$ is of order $O(n)$.

Proposition 16. *The number of nonzero elements in* $\mathsf{diff}(B(T \cap \chi(m_j)))$ *is fewer than* $4n$.

Proof. If $B(T \cap \chi(m_j))$ is the zero matrix, then $\mathsf{diff}(B(T \cap \chi(m_j)))$ is also the zero matrix. Otherwise, the nonzero elements in $B(T \cap \chi(m_j))$ compose a rectangle (cf. Table 5), *i.e.*, there exist $0 \leqslant x_1 \leqslant x_2 < n_x$ and $0 \leqslant y_1 \leqslant y_2 < n_y$ s.t.

$$B\big(T \cap \chi(m_j)\big)[k,l] = \begin{cases} 1, & \text{if } x_1 \leqslant k \leqslant x_2 \text{ and } y_1 \leqslant l \leqslant y_2; \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to prove that

$$\mathsf{diff}\big(B\big(T \cap \chi(m_j)\big)\big)[k,l] = \begin{cases} 1, & \text{if } k = x_1 \text{ and } y_1 \leqslant l \leqslant y_2; \\ -1, & \text{if } k = x_2 + 1 \text{ and } y_1 \leqslant l \leqslant y_2; \\ 0, & \text{otherwise.} \end{cases}$$

So if $x_2 < n_x - 1$, there are $(y_2 - y_1 + 1)$ '1's and '-1's in $\mathsf{diff}(B(T \cap \chi(m_j)))$; otherwise $x_2 = n_x - 1$, there are $(y_2 - y_1 + 1)$ '1's and none '-1's, with other elements being zeros. So there are at most $2 \times (y_2 - y_1 + 1) \leqslant 2 \times n_y < 4n$ nonzero elements in $\mathsf{diff}(B(T \cap \chi(m_j)))$. □

As a consequence, we know the sum of all these difference matrices can be computed in $O(n^2)$ time.

Proposition 17. *The matrices* Q_i *and* $B(b_i)$ *can be computed in* $O(n^2)$ *time for each* i.

Proof. By Proposition 16, we know each difference matrix $B(T \cap \chi(m_j))$ has at most $4n$ nonzero entries. To compute Q_i, we need only add up these matrices. Note that there are at most $9n$ different $B(T \cap \chi(m_j))$. We add these matrices one by

Table 6
Excluding impossible pixels from m_1: An efficient method.

$\begin{pmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 1&1&0&0&0\\ 1&1&0&0&0\\ 0&0&0&0&0 \end{pmatrix}$ $B(T\cap O(m_2))$	$\begin{pmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 1&1&0&0&0 \end{pmatrix}$ $B(T\cap S(m_2))$	$\begin{pmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&1&1&1 \end{pmatrix}$ $B(T\cap SE(m_2))$	$\begin{pmatrix} 0&0&0&0&0\\ 1&1&1&0&0\\ 1&1&1&0&0\\ 1&1&1&0&0\\ 1&1&1&0&0 \end{pmatrix}$ $B(T\cap W(m_3))$
$\begin{pmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 1&0&-1&0&0\\ 1&0&-1&0&0\\ 0&0&0&0&0 \end{pmatrix}$ $\mathsf{diff}(B(T\cap O(m_2)))$	$\begin{pmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 1&0&-1&0&0 \end{pmatrix}$ $\mathsf{diff}(B(T\cap S(m_2)))$	$\begin{pmatrix} 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&1&0&0 \end{pmatrix}$ $\mathsf{diff}(B(T\cap SE(m_2)))$	$\begin{pmatrix} 0&0&0&0&0\\ 1&0&0&-1&0\\ 1&0&0&-1&0\\ 1&0&0&-1&0\\ 1&0&0&-1&0 \end{pmatrix}$ $\mathsf{diff}(B(T\cap W(m_3)))$
$\begin{pmatrix} 0&0&0&0&0\\ 1&0&0&-1&0\\ 2&0&-1&-1&0\\ 2&0&-1&-1&0\\ 2&0&0&-1&0 \end{pmatrix}$ B_Σ	$\begin{pmatrix} 0&0&0&0&0\\ 1&1&1&0&0\\ 2&2&1&0&0\\ 2&2&1&0&0\\ 2&2&2&1&1 \end{pmatrix}$ $Q_1=\mathsf{acc}(B_\Sigma)$	$\begin{pmatrix} 0&1&1&1&0\\ 0&1&1&1&0\\ 0&1&1&1&0\\ 0&0&0&0&0\\ 0&0&0&0&0 \end{pmatrix}$ $B(m_1)$	$\begin{pmatrix} 0&1&1&1&0\\ 0&0&0&1&0\\ 0&0&0&1&0\\ 0&0&0&0&0\\ 0&0&0&0&0 \end{pmatrix}$ $B(b_1)$

one. Each time we need only change the values of at most $4n$ entries. This means Q_i can be computed in $O(n^2)$ time. By Proposition 15, the matrix $B(b_i)$ can also be computed in $O(n^2)$ time, given that we know $B(m_i)$ and Q_i. □

By Eq. (30), we know each b_i can also be computed in $O(n^2)$ time (cf. Algorithm 4). In this way, we improve Step 3 from $O(n^4)$ to cubic time.

Algorithm 4 Excluding Impossible Pixels, improved

Input: A basic CDC network $\mathcal{N}=\{v_i\delta_{ij}v_j\}_{i,j=1}^n$, a set of rectangles $\{m_i\}_{i=1}^n$ with frame $T=[0,n_x]\times[0,n_y]$.
Output: A set of $n_x\times n_y$ Boolean matrices $\{B_i\}_{i=1}^n$.

- **for** each $1\leqslant i\leqslant n$ **do**
  - let B_Σ be an $n_x\times n_y$ matrix initialized with zero;
  - **for** each $1\leqslant j\leqslant n$ and $j\neq i$ **do**
    - **for** each $\chi\in$ TileName **do**
      - **if** $\delta_{ij}^\chi=0$ and $m_i\cap\chi(m_j)$ contains at least one pixel **then**
        - suppose $m_i\cap\chi(m_j)$ is rectangle $[s_1,s_2]\times[t_1,t_2]$;
        - **for** each $t_1\leqslant k<t_2$ **do**
          - $B_\Sigma[s_1,k]\leftarrow B_\Sigma[s_1,k]+1$;
          - $B_\Sigma[s_2,k]\leftarrow B_\Sigma[s_2,k]-1$;
  - $Q_i\leftarrow\mathsf{acc}(B_\Sigma)$;
  - **for** each pixel $p=[s,s+1]\times[t,t+1]$ contained in T **do**
    - **if** p is contained in m_i and $Q_i[s,t]=0$ **then**
      - $B_i[s,t]\leftarrow 1$;
    - **else**
      - $B_i[s,t]\leftarrow 0$;
- Output $\{B_i\}_{i=1}^n$.

Example 1 *(continued).* In our running example, for $m_1=[1,4]\times[2,5]$, there are four different $\chi(m_j)$ such that $B(T\cap\chi(m_j))$ is not the zero matrix and $d_{1j}^\chi=0$, *viz.* $T\cap O(m_2)=[0,2]\times[1,3]$, $T\cap S(m_2)=[0,2]\times[0,1]$, $T\cap SE(m_2)=[2,5]\times[0,1]$, $T\cap W(m_3)=[0,3]\times[0,4]$ (see Fig. 11(a)). Write B_Σ for the sum of all these $\mathsf{diff}(B(T\cap\chi(m_j)))$. Its cumulative matrix is $Q_1=\mathsf{acc}(B_\Sigma)$. By Eq. (36), we know $B(b_1)[k,l]=1$ iff $B(m_1)[k,l]=1$ and $Q_1[k,l]=0$. Table 6 illustrates the process of computing $B(b_1)$. This can be compared with Table 4.

5.3. Beyond complete basic networks of CDC constraints

In the above two subsections, we have shown that the consistency of a (complete) basic network of CDC constraints can be determined in cubic time. Using a backtracking method, we immediately know that the consistency satisfaction problem of the CDC is an NP problem.

Lemma 9. *The consistency satisfaction problem of the CDC is an NP problem.*

Proof. Let $\mathcal{C}=\{v_ic_{ij}v_j\}_{i,j=1}^n$ be a set of CDC constraints. To determine if $\mathcal{C}$ is consistent, we need only branch each non-basic constraint c_{ij}, and then call our cubic algorithm to solve the basic network of CDC constraints. □

This problem is also NP-hard.

Lemma 10. *The consistency satisfaction problem of the CDC is NP-hard.*

Proof. We prove this by reducing a known NP-hard problem to the consistency satisfaction problem of the CDC. Let $\mathcal{A}$ be the JEPD set of IA relations

$$\{\mathsf{p} \cup \mathsf{m} \cup \mathsf{pi} \cup \mathsf{mi}, \mathsf{o} \cup \mathsf{s} \cup \mathsf{d} \cup \mathsf{f} \cup \mathsf{eq} \cup \mathsf{fi} \cup \mathsf{si} \cup \mathsf{di} \cup \mathsf{oi}\}.$$

It can be proved that reasoning with $\mathcal{A}$ is already NP-hard. Actually, the NP-hardness of $\mathcal{A}$ is guaranteed by the work of Krokhin, Jeavons, and Jonsson [21]. We need only to show $\mathcal{A}$ is not contained in any of the eighteen maximal tractable subclasses of the IA (see [21, Table III]). The verification is straightforward and omitted here.

Reasoning with IA relations in $\mathcal{A}$ can be easily reduced to reasoning with the CDC. For a set of IA constraints

$$\mathcal{N} = \{v_i \lambda_{ij} v_j\}_{i,j=1}^{n}, \quad (\lambda_{ij} \in \mathcal{A}),$$

it is easy to see that $\mathcal{N}$ is satisfiable iff the set of CDC constraints

$$\mathcal{N}^* = \{v_i \varphi_{ij} v_j\}_{i,j=1}^{n}$$

is consistent, where φ_{ij} is the disjunction of the basic CDC relations in $\{W, E\}$ if $\lambda_{ij} = \mathsf{p} \cup \mathsf{m} \cup \mathsf{pi} \cup \mathsf{mi}$, and is the disjunction of the basic CDC relations in $\{O, W:O, O:E, W:O:E\}$ otherwise.

This shows that reasoning with the CDC is at least as hard as reasoning with $\mathcal{A}$. Therefore, reasoning with the CDC is also NP-hard. □

As a corollary, we know

Theorem 5. *Reasoning with the CDC is an NP-complete problem.*

A similar conclusion has been obtained for CDC_d (concerning possibly disconnected regions) in [41], where the authors use a reduction from the 3-SAT problem. There seems no direct connection between the NP-hardness of reasoning with CDC_d and the NP-hardness of reasoning with CDC.

6. Local consistency

As we have mentioned in Section 2.2, path-consistency is sufficient to decide the consistency of a basic IA network. In qualitative spatial and temporal reasoning, it is well-known that the usual definition of path-consistency based on composition is too strong. Several authors suggest using the notion of weak composition to define path-consistency (cf. [36,23]). It has been proved that the revised path-consistency is sufficient to decide the consistency of basic RCC8 networks (cf. [32,37,23]). For the CDC, the situation is rather complicated. This is because, unlike the IA and the RCC8, CDC basic relations are not closed under converse (see Fig. 3). A further adaption is necessary.

Definition 15 *(k-consistency).* Let D be a nonempty domain and let $\langle\mathcal{B}\rangle$ be a qualitative calculus on D. A basic network $\mathcal{N}$ over $\langle\mathcal{B}\rangle$ is said to be *k-consistent* iff every subnetwork of $\mathcal{N}$ involving k-variables is consistent. A basic network is said to be *path-consistent* if it is 3-consistent.

This definition of k-consistency, though very different from the classical one introduced by Freuder [12,13], is useful in the research of qualitative spatial and temporal reasoning, which typically uses a finite JEPD set of relations but an infinite domain.

In this section, we first show that, for any fixed $k > 0$, there exists a k-consistent basic CDC network which is inconsistent. Then we show how our main algorithm can be used to solve two special local consistency problems.

6.1. Local consistency is insufficient

Skiadopoulos and Koubarakis [40, Example 9] showed that there is a path-consistent but inconsistent basic CDC network. The following example shows that there is a 4-consistent basic CDC network which is inconsistent.

Example 4. Let a_1, a_2, a_3, a_4 be the squares illustrated in Fig. 14. Consider the basic CDC network $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^{5}$, where

- $\delta_{ij} = \mathrm{dir}(a_i, a_j)$ for $1 \leqslant i, j \leqslant 4$,
- $\delta_{i5} = O$ for $1 \leqslant i \leqslant 4$, and
- $\delta_{51} = N:NE:E$, $\delta_{52} = \delta_{53} = W:NW:N:NE:E$, $\delta_{54} = W:NW:N$.

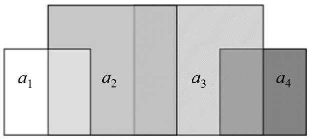


Fig. 14. A solution to the subnetwork over $\{v_1, v_2, v_3, v_4\}$ of $\mathcal{N}$ in Example 4.

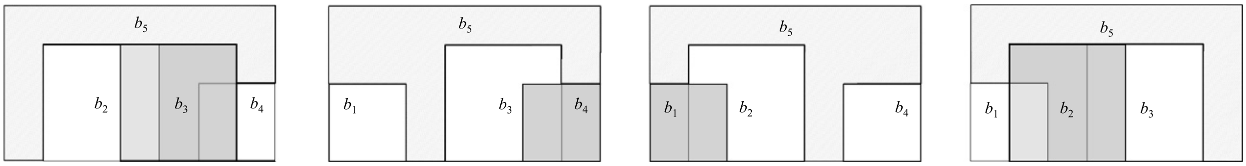


Fig. 15. Solutions to other subnetworks of $\mathcal{N}$ in Example 4 which involve four variables.

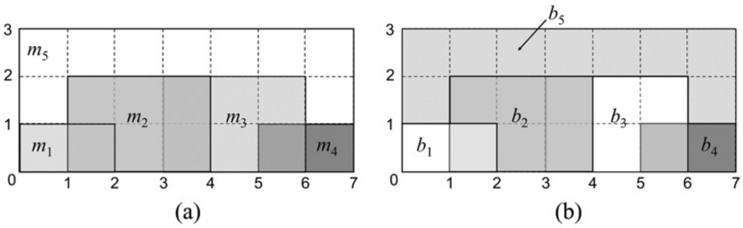


Fig. 16. Illustrations of m_i (a) and b_i (b) in Example 4.

Table 7
The projective IA networks in Example 4.

(i, j)	$(1,2)$	$(1,3)$	$(1,4)$	$(1,5)$	$(2,3)$	$(2,4)$	$(2,5)$	$(3,4)$	$(3,5)$	$(4,5)$
ρ^x_{ij}	o	p	p	s	o	p	d	o	d	f
ρ^y_{ij}	s	s	eq	s	eq	si	s	si	s	s

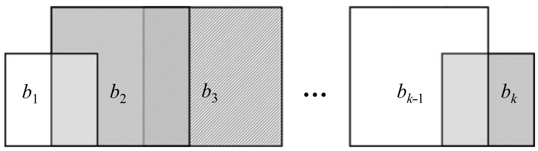


Fig. 17. A configuration of k squares.

Figs. 14 and 15 provide solutions to each subnetwork of $\mathcal{N}$ which involves four variables. Therefore, $\mathcal{N}$ is 4-consistent. We apply our main algorithm to determine the consistency of $\mathcal{N}$.

Step 1 computes the projective IA networks $\mathcal{N}_x = \{v_i \rho^x_{ij} v_j\}^n_{i,j=1}$ and $\mathcal{N}_y = \{v_i \rho^y_{ij} v_j\}^n_{i,j=1}$ (see Table 7).

Applying a path-consistency algorithm, we know that $\mathcal{N}_x$ and $\mathcal{N}_y$ are consistent. Using the algorithm in Step 2, we get the canonical solutions $\{I_i\}^5_{i=1}$, $\{J_i\}^5_{i=1}$ of $\mathcal{N}_x$ and $\mathcal{N}_y$, respectively, where

$$I_1 = [0,2], \qquad I_2 = [1,4], \qquad I_3 = [2,6], \qquad I_4 = [5,7], \qquad I_5 = [0,7],$$
$$J_1 = [0,1], \qquad J_2 = [0,2], \qquad J_3 = [0,2], \qquad J_4 = [0,1], \qquad J_5 = [0,3].$$

The rectangles $m_i = I_i \times J_i$ $(1 \leqslant i \leqslant 5)$ are illustrated in Fig. 16(a).

By the definition of b_i, we know no pixels should be excluded for $1 \leqslant i \leqslant 4$. This means $b_i = m_i$ for $1 \leqslant i \leqslant 4$. For b_5, however, pixels in $O(m_1)$, *viz.* p_{00} and p_{01}, are disallowed because $d^O_{51} = 0$. Similarly, pixels in $O(m_2)$, $O(m_3)$, $O(m_4)$ are also excluded. Fig. 16(b) illustrates the resulted digital regions. Because $\mathcal{M}(b_5) = [0,7] \times [1,3] \neq m_5$, the network is inconsistent according to our main algorithm.

In fact, for any positive integer k, there exists a k-consistent but inconsistent basic CDC network involving $k+1$ variables. Fig. 17 illustrates such a network, where k instead of four squares are used. This shows that k-consistency is insufficient for deciding the consistency of basic CDC networks.

Remark 3. Navarrete et al. [31] proposed an $O(n^4)$ algorithm REG-BCON for checking the consistency of a basic CDC network. It can be proved that any 4-consistent basic CDC network $\mathcal{N}$ would pass the algorithm REG-BCON. This is due to that the

NTB examination given there only involves four variables. Example 4 shows that REG-BCON is not always correct. Moreover, the 'Helly's Topological Theorem' used in [31] is unjustified and wrong.

Our main algorithm can be applied to solve two special local consistency problems, *i.e.* the pairwise consistency problem and the weak composition problem.

6.2. The pairwise consistency problem

Given that you know the relation of a to b, what about that of b to a? Mathematically speaking, this is the converse problem. The relation of b to a is the converse of that of a to b, and vice versa. Suppose $\mathcal{B}$ is a set of JEPD relations on D. The qualitative calculus $\langle\mathcal{B}\rangle$ is not necessarily closed under converse. This means, $\alpha^{\sim}$ could be a relation out of $\langle\mathcal{B}\rangle$ despite $\alpha \in \langle\mathcal{B}\rangle$.

As for the CDC, we have shown in Fig. 3 that a basic CDC relation may have more than one 'converses,' where for two basic relations α, β, we say α is a *converse* of β (in the CDC) if $\{v_1 \alpha v_2, v_2 \beta v_1\}$ is consistent.

The pairwise consistency problem in the CDC is the problem of deciding if $\{v_1 \delta_{12} v_2, v_2 \delta_{21} v_1\}$ is satisfiable for a pair of basic CDC relations δ_{12} and δ_{21}. This problem has been discussed by Cicerone and di Felice [3].

We next apply our main algorithm to solve the pairwise consistency problem. The first step computes ρ^x_{12} and ρ^y_{12} (see Eqs. (14) and (15)). If ρ^x_{12} or ρ^y_{12} is empty, then the program stops and returns 'inconsistent.' Otherwise, we go to Step 2. We construct the canonical solutions $\{I_1, I_2\}$ and $\{J_1, J_2\}$ to ρ^x_{12} and ρ^y_{12}, respectively. Let $m_i = I_i \times J_i$ for $i = 1, 2$. Write $I_i = [x_i^-, x_i^+]$ and $J_i = [y_i^-, y_i^+]$. Denote $n_x = \max\{x_1^+, x_2^+\}$, $n_y = \max\{y_1^+, y_2^+\}$. Clearly, $1 \leqslant n_x, n_y \leqslant 3$. Let $T = [0, n_x] \times [0, n_y]$. Then $T \subseteq [0, 3] \times [0, 3]$. Continuing as described in the main algorithm, we will determine if $\{v_1 \delta_{12} v_2, v_2 \delta_{21} v_1\}$ is consistent, and find the maximal canonical solution if it is consistent.

A specialized algorithm is implemented.[2] We obtain in total 757 consistent pairs of basic CDC relations. Among these consistent pairs (δ, δ'), one δ may correspond to multiple δ', *i.e.* δ may have multiple converses. In fact, among the 218 basic CDC relations, 119 have unique converse, 68 have two converses, 6 have four converses, 20 have eight converses, 4 have thirty converses, and one (*viz.* the single-tile relation O) has 198 converses.

Our result is unexpectedly different from that of [3], where Cicerone and di Felice obtained 2004 consistent pairs of basic CDC relations. A careful examination, however, shows that a similar but different model was used in [3].

When defining the direction relation matrix $\mathrm{dir}(a, b) = [d^\chi]_{\chi \in \mathrm{TileName}}$ of a to b (see Definition 4), we require $d^\chi = 1$ iff $a^\circ \cap \chi(b) \neq \varnothing$. While in [3], d^χ is 1 iff a has nonempty intersection with $\chi(b)$, *i.e.* $d^\chi = 1$ iff $a \cap \chi(b) \neq \varnothing$. We call these two definitions the *interior-based* and the *closure-based* direction relation matrix, respectively. In the following, we argue that the interior-based definition is more coherent than the closure-based one.

First, though it was not mentioned that $d^\chi = 1$ iff a has a common interior point with the tile $\chi(b)$ in the original definition of Direction Relation Matrix [16], Goyal and Egenhofer [17] defined the detailed direction relation matrix of a to b as a numerical matrix $\mathrm{dir}^*(a, b) = [(d^\chi)^*]_{\chi \in \mathrm{TileName}}$, where $(d^\chi)^*$ is interpreted as the ratio of the area of $a \cap \chi(b)$ and a. This is a natural extension of the interior-based direction relation matrix. From $\mathrm{dir}^*(a, b) = [(d^\chi)^*]_{\chi \in \mathrm{TileName}}$, we can obtain the coarse direction relation matrix $\mathrm{dir}(a, b) = [d^\chi]_{\chi \in \mathrm{TileName}}$ by setting $d^\chi = 1$ iff $(d^\chi)^* > 0$.

Second, the qualitative calculus introduced by the interior-based direction relation matrix is more desirable. For example, it is a natural requirement that the identity relation is contained in a unique basic CDC relation. For the interior-based definition, we have $\mathrm{dir}(a, a) = O$ for any connected region a. The following example, however, shows that this is not the case for the closure-based direction relation matrix.

Example 5. Take a as the square $[-1, 1] \times [-1, 1]$, and take a' as the unit disk centered at $(0, 0)$. Then $\mathcal{M}(a') = \mathcal{M}(a) = a$, but $\mathrm{dir}(a, a) \neq \mathrm{dir}(a', a')$ if we take the closure-based definition.

$$\mathrm{dir}(a, a) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \qquad \mathrm{dir}(a', a') = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

The interior-based definition is also consistent with that in [40,41].

6.3. The weak composition problem

The notion of weak composition plays a very important role in qualitative spatial and temporal reasoning [1,7,24,36]. For two basic CDC relations α, β, the weak composition $\alpha \circ_w \beta$ of α and β is defined to be the smallest relation in the CDC algebra which contains the composition $\alpha \circ \beta$. Since the CDC is a Boolean algebra, $\alpha \circ_w \beta$ is the union of all basic CDC relations it contains. For a basic CDC relation γ, it is easy to prove that

$$\gamma \subseteq \alpha \circ_w \beta \quad \Leftrightarrow \quad \gamma \cap (\alpha \circ \beta) \neq \varnothing. \tag{40}$$

[2] See https://sites.google.com/site/lisanjiang/cdc for illustrations of all consistent pairs.

Note that $\gamma \cap (\alpha \circ \beta)$ is nonempty iff the following set of basic CDC constraints

$$\mathcal{C} = \{v_1 \alpha v_2, v_2 \beta v_3, v_1 \gamma v_3\} \tag{41}$$

is consistent. We note that $\mathcal{C}$ is not a complete network. The constraint of, say, v_2 to v_1, is not specified. According to the previous subsection, α may have multiple converses. To apply our main algorithm, we need to extend $\mathcal{C}$ to a complete network:

$$\mathcal{C}^* = \{v_1 \alpha v_2, v_2 \alpha' v_1, v_2 \beta v_3, v_3 \beta' v_2, v_1 \gamma v_3, v_3 \gamma' v_1\}. \tag{42}$$

We then call our main algorithm to determine if the above completed network is consistent. If the answer is 'yes' for some α', β', γ', then γ is contained in the weak composition of α and β. Note that we need only to apply the main algorithm to those α', β', γ' such that (α, α'), (β, β'), (γ, γ') are consistent pairs. For example, the weak composition of $SW:W$ and $NE:E$ contains 206 out of the 218 basic CDC relations.[3]

The same problem has been considered in [15,40]. The main idea is to compute the weak composition *progressively*. Goyal [15] established the weak composition of two single-tile relations. Upon this, Theorem 1 of [40] establishes a rule for computing the weak composition of a single-tile relation and a basic relation. The correctness of this theorem is confirmed by our algorithm. Furthermore, Theorem 2 of [40] then gives a rule to compute the weak composition of two multi-tile relations. Two examples show that this rule is not always correct. Recently, this problem was fixed based on case by case analysis by Skiadopoulos and Koubarakis.[4] They also verified their results with ours.

7. Consistency checking for two variants of the CDC

The CDC algebra introduced in Definition 4 requires regions to be connected. This calculus has two variants in the literature. One, as introduced in [41], deals with cardinal direction relations between possibly disconnected regions; the other, as originally proposed in [16], deals with simple regions, *i.e.* connected regions that are topologically equivalent to closed disks. In this section, we show our consistency checking algorithm designed for connected regions (possibly with holes) can be adapted to cope with these two variants.

7.1. Cardinal directions between possibly disconnected regions

For two (possibly disconnected) regions a, b, similar to Definition 4, we write $\mathrm{dir}(a,b) = [d^{\chi}]_{\chi \in \textsc{TileName}}$ for the direction relation matrix of a to b, where d^{χ} is 1 if $a^{\circ} \cap \chi(b) \neq \varnothing$, and 0 otherwise. A 3×3 Boolean matrix M is valid if there exist two regions a, b such that $M = \mathrm{dir}(a,b)$. It is easy to see that all but the zero 3×3 Boolean matrices are valid. Each of these matrices represents a basic direction relation between possibly disconnected regions. We call the Boolean algebra generated by these JEPD relations the cardinal direction calculus for possibly disconnected regions, denoted as CDC_d.

Consistency checking in the CDC_d is similar to that in the CDC. Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a complete network of basic CDC_d constraints. Similar definitions of regular solutions, meet-free solutions, and canonical solutions can be defined in the CDC_d. Moreover, suppose $\mathfrak{a}$ is a solution to $\mathcal{N}$. We can transform $\mathfrak{a}$ into a canonical solution $\mathfrak{a}' = \{a'_i\}_{i=1}^n$ of $\mathcal{N}$. That is, $\mathfrak{a}'$ is a regular, meet-free, and digital solution, and $\{I_x(a'_i)\}_{i=1}^n$ and $\{I_y(a'_i)\}_{i=1}^n$ are canonical sets of intervals. We next show $\mathcal{N}$ has a maximal canonical solution. Actually, for each i, let b_i be the region obtained by deleting all disallowed pixels from $\mathcal{M}(a'_i)$ (see Eq. (25)). We assert that $\mathfrak{b} = \{b_i\}_{i=1}^n$ is the maximal canonical solution of $\mathcal{N}$.

Theorem 6. *Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a complete network of basic CDC_d constraints. If $\mathcal{N}$ is consistent, then $\{b_i\}_{i=1}^n$ is the maximal canonical solution of $\mathcal{N}$, where b_i is defined as in Eq.* (25)*.*

Proof. The proof is similar to that for Lemma 8. □

This shows, to construct the maximal canonical solution of a network of basic CDC_d constraints, we need not compute the connected components of b_i.

We next adapt our main algorithm to determine the consistency of a complete network $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ of basic CDC_d constraints. We first note that the projective interval relation of a basic CDC_d constraint can be computed in a similar way as that of a basic CDC constraint (see Remark 5 in Appendix A). As in the case of connected regions, Step 1 computes the projective IA networks $\mathcal{N}_x$ and $\mathcal{N}_y$. If either is inconsistent, then $\mathcal{N}$ is inconsistent. Otherwise, Step 2 constructs their canonical interval solutions, and Step 3 computes b_i for each i. If $\mathcal{M}(b_i) \neq m_i$ for some i then $\mathcal{N}$ is inconsistent. The above procedures, as in the connected case, need at most cubic time.

Since regions in a solution of CDC_d constraints are allowed to be disconnected, we need not compute the connected components of each b_i. Therefore, we go directly to Step 5, where we need to check if $\mathrm{dir}(b_i, b_j) = \delta_{ij}$ holds for each pair of

[3] See https://sites.google.com/site/lisanjiang/cdc for illustration.

[4] Personal communication, see also http://pelopas.uop.gr/~spiros/pubs.html.

i, j. Suppose $\delta_{ij} = [d_{ij}^{\chi}]_{\chi \in \text{TileName}}$. Similar to the connected case, we need only check whether $b_i^\circ \cap \chi(m_j)$ is nonempty for each χ with $d_{ij}^{\chi} = 1$.

To check if $b_i^\circ \cap \chi(m_j)$ is nonempty, it is sufficient to show that there is a pixel p contained in $b_i \cap \chi(m_j)$. Since b_i may be disconnected, checking all boundary pixels is, however, insufficient. Note that checking all pixels in $m_i \cap \chi(m_j)$ alone needs $O(n^2)$ time in the worst case for each pair i, j. This is undesirable since it will cost $O(n^4)$ time in total.

We next show this could also be simplified. For a digital region b contained in T, write $B(b)$ for the Boolean matrix that represents b (see Eq. (30)). For each $0 \leqslant k < n_x$ and each $0 \leqslant l < n_y$, write $M(b)[k,l]$ for the number of pixels contained in b and the rectangle $[0,k+1] \times [0,l+1]$. It is easy to see that $M(b)$ is an $n_x \times n_y$ integer matrix and

$$M(b)[k,l] = \sum \bigl\{B(b)[p,q]\colon\ 0 \leqslant p \leqslant k,\ 0 \leqslant q \leqslant l\bigr\} \tag{43}$$

for $0 \leqslant k < n_x$ and $0 \leqslant l < n_y$.

Given the Boolean matrix $B(b)$, $M(b)$ can be computed in $O(n^2)$ time by iteratively adding the k-th column to the $(k+1)$-th, and then iteratively adding the p-th row to the $(p+1)$-th.

$M(b)$ gives a way to compute the number of pixels contained in b and $[0,k+1] \times [0,l+1]$. The following proposition concerns pixels contained in b and an arbitrary rectangle.

Proposition 18. *Let $r = [x^-, x^+] \times [y^-, y]$ be a rectangle and let b be a digital region b, both contained in $T = [0, n_x] \times [0, n_y]$. Then $b \cap r$ contains a pixel iff*

$$\begin{aligned} &M(b)\bigl[x^- - 1, y^- - 1\bigr] + M(b)\bigl[x^+ - 1, y^+ - 1\bigr] \\ &\quad > M(b)\bigl[x^+ - 1, y^- - 1\bigr] + M(b)\bigl[x^- - 1, y^+ - 1\bigr], \end{aligned} \tag{44}$$

where $M(b)[-1,l] = M(b)[k,-1] = 0$.

Proof. Because $M(b)[k,l]$ denotes the number of pixels contained in both b and $[0,k+1] \times [0,l+1]$, the number of pixels in b which are contained in the rectangle $[x^-, x^+] \times [y^-, y^+]$ is $M(b)[x^+ - 1, y^+ - 1] - (M(b)[x^- - 1, y^+ - 1] + M(b)[x^+ - 1, y^- - 1]) + M(b)[x^- - 1, y^- - 1]$. The conclusion follows directly. □

The following proposition asserts that, given $M(b_i)$, it needs only constant time to decide whether $b_i^\circ \cap \chi(m_j)$ is nonempty for each χ with $d_{ij}^{\chi} = 1$.

Proposition 19. *Let m_i, b_i be as constructed in Step 2 and Step 3, respectively, of the main algorithm for possibly disconnected regions. Let χ be a tile name. Suppose $d_{ij}^{\chi} = 1$. Assume $M(b_i)$ has been computed. Then whether $b_i^\circ \cap \chi(m_j)$ is nonempty can be checked in constant time.*

Proof. Note that $b_i^\circ \cap \chi(m_j)$ is nonempty iff there is a pixel contained in both b_i and $\chi(m_j)$. Suppose $T \cap \chi(m_j) = [x^-, x^+] \times [y^-, y^+]$. Since $b_i \subseteq T$, we have

$$b_i \cap \chi(m_j) = b_i \cap \bigl(T \cap \chi(m_j)\bigr) = b_i \cap \bigl[x^-, x^+\bigr] \times \bigl[y^-, y^+\bigr].$$

By Proposition 18, we know $b_i \cap \chi(m_j)$ contains a pixel iff Eq. (44) holds, which can be checked in constant time. □

Since there are at most $9n$ different rectangles $T \cap \chi(m_j)$, given $M(b_i)$, whether $\mathsf{dir}(b_i, b_j) = \delta_{ij}$ can be checked in $O(n)$ time. Recall that each $M(b_i)$ can be computed in $O(n^2)$ time. This implies whether $\{b_i\}_{i=1}^n$ is a solution of $\mathcal{N}$ can be checked in cubic time. Therefore, Step 5 and the whole algorithm can be finished in $O(n^3)$ time.

Remark 4. Skiadopoulos and Koubarakis [41] proposed an algorithm for determining the consistency of basic constraints in the CDC_d. This algorithm applies to possibly incomplete sets of basic constraints. Let $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ be a network of CDC_d constraints such that each δ_{ij} is either basic or the universal relation. The algorithm in [41] can decide in $O(n^5)$ time whether $\mathcal{N}$ is consistent. In this sense, it is more general than our cubic algorithm (when applied to CDC_d).

7.2. Cardinal directions between simple regions

In the definition of direction relation matrix (Definition 4), we assume connected regions. It is possible that these connected regions *may have holes*. In the original work of Goyal and Egenhofer [16,17], objects are represented as *simple regions*, *i.e.* regions that are topologically equivalent to closed disks. Because each direction relation matrix between connected regions can be realized by a pair of simple regions, the set of cardinal direction relations between simple regions is the same as that between connected regions. We write CDC_s for the qualitative calculus generated by these cardinal direction relations on the set of simple regions.

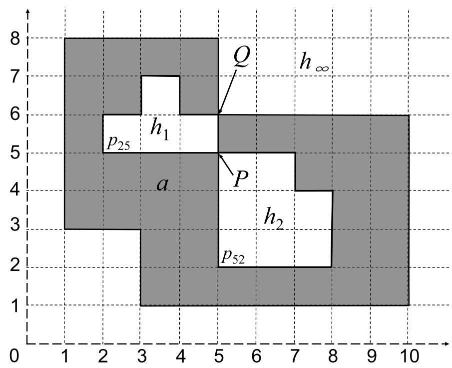


Fig. 18. A digital region a with two contact points $P = (5,5)$ and $Q = (5,6)$, where h_1 and h_2 are two holes of a, h_∞ is the closure of the unbounded component of the exterior of a.

Table 8
Four types of contact points, where h, a, x denote a pixel that is contained in a hole of a, in a, and in neither, respectively.

$\begin{array}{c\|c} h & a \\ \hline a & x \end{array}$	$\begin{array}{c\|c} a & h \\ \hline x & a \end{array}$	$\begin{array}{c\|c} x & a \\ \hline a & h \end{array}$	$\begin{array}{c\|c} a & x \\ \hline h & a \end{array}$

In this subsection, we show the difference between simple regions and connected regions does not affect the consistency of cardinal direction constraints. In particular, we prove that each consistent network of CDC constraints has a solution only using simple regions. The idea is to transform connected regions (possibly with holes) in the maximal canonical solution into simple regions without changing their relations.

We first note that a connected digital region a has at most finite holes, where a *hole* of a is the closure of a bounded connected component of the exterior of a. It is easy to see that all holes of a digital region are digital and simple regions. We call a point P a *contact point* of a digital region a if exactly two of the four pixels around P are contained in a and these two pixels are diagonally adjacent (cf. Fig. 18).

The following two lemmas study properties of contact points in the maximal canonical solution.

Lemma 11. *Let $\mathfrak{a} = \{a_i\}_{i=1}^n$ be the maximal canonical solution of a basic CDC network $\mathcal{N}$. Suppose $P = (k,l)$ is contact point of a_i. Let p be a pixel around P which is contained in a hole h of a_i. Then there exists $j \neq i$ such that $p \subseteq \mathcal{M}(a_j) \subseteq h$.*

Proof. Omitted. □

Because a maximal canonical solution is a meet-free solution, the mbrs of any two regions in a maximal canonical solution do not meet at a point (cf. Definition 13). The following lemma is a consequence of this observation and Lemma 11.

Lemma 12. *Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is the maximal canonical solution of a basic CDC network $\mathcal{N}$. If $P = (k,l)$ is a contact point of a_i, then among the four pixels around P, one is contained in a hole of a_i, two are contained in a_i, the other is contained in neither a_i nor its holes.*

Proof. Omitted. □

By the above lemma, there are four types of contact points. See Table 8. For convenience, we denote each type as a 4-tuple of symbols taken from $\{h, a, x\}$. For the four pixels, we start from the top left corner and go clockwise. These four types are written as ($haxa$), ($ahax$), ($xaha$), and ($axah$), respectively.

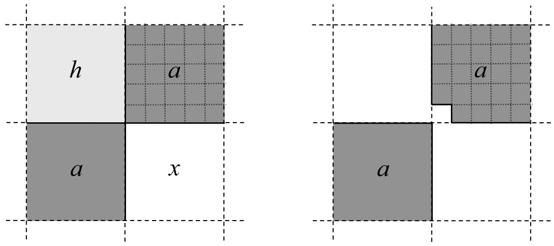
A contact point can be removed by deleting a sub-pixel from the a-pixel which follows the h-pixel according to the sequence of the type of the contact point (see Fig. 19). Once all contact points are removed from each a_i, the remaining holes are quite easy to cope with.

We need the following lemma to prove the main result of this subsection.

Lemma 13. *For two digital regions a, b, if a', b' satisfies the following equation*

$$a' \subseteq a \quad \text{and} \quad \mathcal{M}(b) = \mathcal{M}(b') \quad \text{and} \quad (\forall p)\big[p \subseteq a \to \big(a'\big)^\circ \cap p \neq \varnothing\big], \tag{45}$$

then $\mathsf{dir}(a,b) = \mathsf{dir}(a',b')$.



(a) (b)

Fig. 19. Remove a type (*haxa*)-contact point.

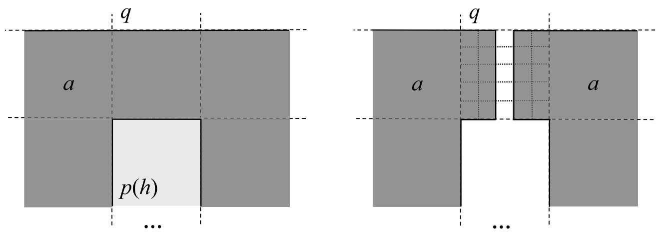


Fig. 20. Transform a connected digital region without contact points into a simple region.

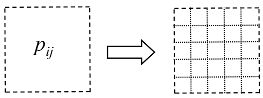


Fig. 21. Subdivision of a pixel into 25 sub-pixels.

Proof. Omitted. □

Theorem 7. *Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a consistent network of basic CDC constraints. Then it has a solution $\{a_i\}_{i=1}^n$ such that each a_i is a simple region.*

Proof. Suppose $\mathfrak{a} = \{a_i\}_{i=1}^n$ is the maximal canonical solution of $\mathcal{N}$. Then each a_i is a connected digital region, which may have holes. We assert that, for each i, there exists a simple region a_i' such that $a_i' \subseteq a_i$ and $\mathcal{M}(a_i') = \mathcal{M}(a_i)$, and $\mathsf{dir}(a_i, a_j) = \mathsf{dir}(a_i', a_j)$ and $\mathsf{dir}(a_j, a_i) = \mathsf{dir}(a_j, a_i')$ for any $j \neq i$.

We first subdivide each pixel into 25 equal sub-pixels (see Fig. 21).

For a contact point $P = (k,l)$ of a_i, without loss of generality, assume P has type (*haxa*). Removing the 1/25 sub-pixel which contains P from a_i, we obtain (after necessary regularization)[5] a new region (see Fig. 19). This procedure can be applied to all contact points of a_i at the same time. Write a_i^* for the resulted region. Since each pixel p of a_i has at most four contact points, the revised region a_i^* contains at least 21 sub-pixels of p. This implies that $a_i^* \subseteq a_i$ and $\mathcal{M}(a_i^*) = \mathcal{M}(a_i)$. It is routine to check that a_i^* is still connected but has no contact points.

Because a_i^* has no contact points, any two holes of a_i^* are disjoint, and any hole of a_i^* is disjoint from the closure of the unbounded component of the exterior of a_i^*. For each hole h of a_i^*, select the pixel $p(h)$ in h which has the highest y-index and whose left 4-neighbor is out of h. We cut a slot from a_i^* to connect the hole h with the exterior of a_i^* as follows. Let q be the lowest pixel which is above $p(h)$ but not contained in a_i^*. We delete the middle column of the sub-pixels from those pixels between $p(h)$ and q (see Fig. 20). After necessary regularization, we obtain another connected region which has fewer holes than a_i^*. Applying this operation to holes of a_i^* one by one, we obtain a connected region a_i', which has no holes. That is, a_i' is a simple region. It is easy to see that $a_i' \subseteq a_i$ and $\mathcal{M}(a_i') = \mathcal{M}(a_i)$. Moreover, for each pixel p contained in a_i, since p has sub-pixels that are contained in a_i', we know $(a_i')^\circ \cap p$ is nonempty. By Lemma 13, we know $\mathsf{dir}(a_i', a_j') = \mathsf{dir}(a_i, a_j)$ for any $j \neq i$. This means $\{a_i'\}_{i=1}^n$ is a solution of $\mathcal{N}$ which consists of simple regions. □

As a consequence, we know the satisfaction problem over the CDC_s can be determined in the same way as that over the CDC.

Theorem 8. *The consistency of a network of basic CDC_s constraints can be determined in cubic time.*

[5] Here regularization means the topological regularization of a set a, which is defined as $\overline{a^\circ}$.

8. Related work

More than a dozen qualitative direction/orientation calculi have been proposed in the literature. We refer the reader to [6,38] for more information. In this section, we focus on the computational properties of these calculi.

The CDC (for extended objects) is an extension of the projection-based cardinal algebra for point objects [10]. The computational properties of the latter has been studied by Ligozat [26]. In particular, he identified a maximal tractable subset of the cardinal algebra. Several researchers use boxes or rectangles to approximate extended objects and represent the direction relation between two rectangles by a pair of interval relation, see *e.g.* [18,2]. This calculus, known as the Rectangle Algebra (RA), is the two-dimensional extension of the IA. Balbiani et al. identified a tractable subclass of the RA which is larger than the product of the ORD-Horn class [33] of the IA.

Recently, Skiadopoulos et al. [42] introduced a family of directional models for extended objects. These models approximate the reference object by its mbr and partition the plane into five (instead of nine) tiles. The authors also gave methods for computing the converse and the composition of directional relations. The general consistency problem is left open. Unlike the above mentioned binary calculi, Clementini and Billen [4] introduced a ternary projective relation calculus for extended objects. The model describes the relative relation of one primary object to two reference objects by partitioning the plane into several zones with respect to the reference objects. Compared with the aforementioned direction calculi, this model does not have an extrinsic frame of reference. Rules of permutation and composition of relations in this ternary calculus were further established in [5].

As for point objects, a variety of ternary orientation calculi have been established, see *e.g.* [11,20,30,8]. Dylla and Wallgrün [8] examined most of these calculi and proved that many well-known orientation calculi can be expressed in the more general Oriented Point Relation Algebra ($\mathcal{OPRA}_m$ [30]). In addition, they demonstrated that the mapping can be used to determine composition tables of other calculi from that of an $\mathcal{OPRA}_m$. The composition-based reasoning is, however, incomplete for reasoning in most of these calculi [20,29]. Even worse, we do not know if the consistency problem is an NP problem for all but the calculus introduced in [20]. We regard this as a major challenge in the theoretical research and application of qualitative orientation calculi.

9. Conclusion

This paper provided a cubic algorithm for checking the consistency of complete basic CDC networks, which was earlier observed as impossible [31] for connected regions. If a basic CDC network is consistent, our algorithm also generates the maximal canonical solution. This general algorithm was then applied to solve the pairwise consistency problem and the weak composition problem.

Although devised to solve cardinal directional constraints between connected regions, our main algorithm can also be adapted to cope with cardinal directional constraints between possibly disconnected regions as well as those between simple regions. For a complete network of basic CDC_d constraints, our algorithm determines in cubic time if it is consistent. As for cardinal direction constraints over simple regions, we proved that each consistent basic CDC network has a solution using only simple regions. This suggests that the CDC does not distinguish between simply connectivity and connectedness.

It is worth mentioning that our cubic algorithm only works for complete basic CDC (or CDC_d) networks. This is unlike the $O(n^5)$ algorithm in [41], which can be applied to solve possibly incomplete sets of basic CDC_d constraints. It is still unknown if there exist efficient algorithms for checking the consistency of possibly incomplete sets of basic CDC constraints.

Most potential applications of qualitative spatial reasoning require multiple aspects of space. Combining spatial constraints of different calculi is a very important problem in the research of qualitative spatial reasoning. Some work has been done in this direction (see *e.g.* [14,22,28]). In particular, [28] points out that reasoning with basic RCC8 and basic RA [2] constraints is in P, but reasoning with basic RCC8 and basic CDC_d constraints is NP-Complete. Note that spatial variables in the RCC8, the RA, and the CDC_d are all interpreted over possibly disconnected regions. It is still open if similar results hold for connected regions. In particular, we do not know whether reasoning with basic RCC8 and basic CDC constraints is still decidable if spatial variables are interpreted over connected regions.

Acknowledgements

We thank the anonymous referees for their helpful suggestions. This work was partially supported by an ARC Future Fellowship (FT0990811) and a MSRA Young Professorship, and by the NSF of China (60673105, 60736011), and the National Basic Research Program of China (2007CB311003).

Appendix A. Direction relation vectors and projective interval relations

In this section, we introduce a method to compute the projective IA networks $\mathcal{N}_x = \{v_i \rho^x_{ij} v_j\}^n_{i,j=1}$ and $\mathcal{N}_y = \{v_i \rho^y_{ij} v_j\}^n_{i,j=1}$ of a basic CDC network $\mathcal{N} = \{v_i \delta_{ij} v_j\}^n_{i,j=1}$.

We begin with the one-dimensional counterpart of CDC relations.

Definition 16 *(direction relation vector).* Suppose $I = [x^-, x^+]$ and $J = [y^-, y^+]$ are two intervals. Interval J partitions the real line into three parts $L_1 = (-\infty, y^-]$, $L_2 = (y^-, y^+)$, and $L_3 = [y^+, +\infty)$. The direction of I to J is encoded in a Boolean vector $\mathsf{dir}(I, J) = (d_1, d_2, d_3)$, where $d_i = 0$ iff $(x^-, x^+) \cap L_i = \varnothing$. In this case, we call (d_1, d_2, d_3) a *direction relation vector*.

Clearly, a Boolean vector $t = (t_1, t_2, t_3)$ is a direction relation vector iff there exist two intervals I, J such that $t = \mathsf{dir}(I, J)$. The following proposition gives a characterization of direction relation vectors.

Proposition 20. *A Boolean vector $t = (t_1, t_2, t_3)$ is a direction relation vector if and only if $t \neq (0, 0, 0)$ and $t \neq (1, 0, 1)$.*

Interestingly, each direction relation vector actually represents an IA relation.

Proposition 21. *For two intervals I, J, $t = (t_1, t_2, t_3)$ is the direction relation vector of I to J iff (I, J) is an instance of α_t, where*

t	$(1, 0, 0)$	$(0, 1, 0)$	$(0, 0, 1)$	$(1, 1, 0)$	$(0, 1, 1)$	$(1, 1, 1)$
α_t	p ∪ m	s ∪ d ∪ f ∪ eq	pi ∪ mi	o ∪ fi	oi ∪ si	di

Proof. Take $t = (1, 0, 0)$ as example. For $I = [x^-, x^+]$ and $J = [y^-, y^+]$, $\mathsf{dir}(I, J) = (1, 0, 0)$ iff $x^- < y^-$, $(x^-, x^+) \cap (y^-, y^+) = \varnothing$, and $x^+ \leqslant y^+$ hold. This is equivalent to saying that $x^+ \leqslant y^-$, which is possible iff $(I, J) \in$ p ∪ m. □

In what follows, we call an IA relation a *vector IA relation* if it is the IA relation represented by a direction relation vector. We make no difference between a direction relation vector and the IA relation it represents. By the above proposition, we know there are six vector IA relations, *viz.*

p ∪ m, s ∪ d ∪ f ∪ eq, pi ∪ mi, o ∪ fi, oi ∪ si, di.

Note that a vector IA relation is in general non-basic, but a pair of vector IA relations are more precise. For example, from $\mathsf{dir}(I, J) = (0, 1, 0)$, we are not sure whether (I, J) is in s, or d, or f, or eq. Assuming $\mathsf{dir}(J, I)$ is also given, then it is easy to see that the IA relation between I, J is definite, *i.e.* a basic IA relation.

The following proposition summaries the correspondence between pairs of direction relation vectors and IA relations.

Proposition 22. *For a pair of direction relation vectors (s, t) and two intervals I, J, we have $s = \mathsf{dir}(I, J)$ and $t = \mathsf{dir}(J, I)$ iff (I, J) is an instance of the basic IA relation in the cell specified by (s, t) in Table* 9.

Proposition 22 shows that all basic IA relations except 'meets' and 'before' (and their converses) can be represented as pairs of direction relation vectors.

Proposition 23. *Suppose $\delta = [d^\chi]_{\chi \in \text{TILENAME}}$ is a basic CDC relation. Then the x-projective interval relation $\iota^x(\delta)$ is the IA relation associated to the vector (d_1, d_2, d_3), i.e. $(I, J) \in \iota^x(\delta)$ iff $\mathsf{dir}(I, J) = (d_1, d_2, d_3)$, where*

$$d_1 = \max\{d^{NW}, d^W, d^{SW}\}, \qquad d_2 = \max\{d^N, d^O, d^S\}, \qquad d_3 = \max\{d^{NE}, d^E, d^{SE}\}.$$

Proof. Omitted. □

Table 9
Pairs of vector IA relations.

$s\backslash t$	$(1, 0, 0)$	$(0, 1, 0)$	$(0, 0, 1)$	$(1, 1, 0)$	$(0, 1, 1)$	$(1, 1, 1)$
$(1, 0, 0)$	∅	∅	p, m	∅	∅	∅
$(0, 1, 0)$	∅	eq	∅	f	s	d
$(0, 0, 1)$	pi, mi	∅	∅	∅	∅	∅
$(1, 1, 0)$	∅	fi	∅	∅	o	∅
$(0, 1, 1)$	∅	si	∅	oi	∅	∅
$(1, 1, 1)$	∅	di	∅	∅	∅	∅

As a corollary, we have

Proposition 3. *Suppose δ is a basic CDC relation. Then the x-projective interval relation $\iota^x(\delta)$ is one of the following IA relations*

$$\mathsf{p} \cup \mathsf{m}, \quad \mathsf{s} \cup \mathsf{d} \cup \mathsf{f} \cup \mathsf{eq}, \quad \mathsf{pi} \cup \mathsf{mi}, \quad \mathsf{o} \cup \mathsf{fi}, \quad \mathsf{oi} \cup \mathsf{si}, \quad \mathsf{di}. \tag{46}$$

Proof. This is because, by Proposition 23, $\iota^x(\delta)$ is an vector IA relation. Checking Proposition 21, we know an vector IA relation must be an IA relation in Eq. (46). □

By Proposition 22 and Proposition 23, we have

Proposition 4. *For a pair of basic CDC constraints (δ, γ), $\iota^x(\delta, \gamma)$ is either empty or an IA relation in $\mathcal{B}^*_{int}$, where*

$$\mathcal{B}^*_{int} = \{\mathsf{o}, \mathsf{s}, \mathsf{d}, \mathsf{f}, \mathsf{eq}, \mathsf{fi}, \mathsf{di}, \mathsf{si}, \mathsf{oi}\} \cup \{\mathsf{p} \cup \mathsf{m}, \mathsf{pi} \cup \mathsf{mi}\}. \tag{47}$$

Proof. Recall $\iota^x(\delta, \gamma) = \iota^x(\delta) \cap \iota^x(\gamma)^\smile$. Because $\iota^x(\delta)$ and $\iota^x(\gamma)$ are two vector IA relations. According to Table 9, we know the intersection of two vector IA relations is either empty or a relation in Eq. (13). □

For a pair of basic CDC constraints (δ, γ), the projective IA relation $\iota^x(\delta, \gamma)$ can be computed in constant time:

1. Compute $\iota^x(\delta)$ and $\iota^x(\gamma)$ according to Proposition 23;
2. Check Table 9 and determine $\iota^x(\delta) \cap \iota^x(\gamma)^\smile$.

The projective IA networks of a basic CDC network can be computed in $O(n^2)$ time.

Proposition 24. *Suppose $\mathcal{N} = \{v_i \delta_{ij} v_j\}_{i,j=1}^n$ is a basic CDC network. Then its projective IA networks $\mathcal{N}_x = \{v_i \rho^x_{ij} v_j\}_{i,j=1}^n$ and $\mathcal{N}_y = \{v_i \rho^y_{ij} v_j\}_{i,j=1}^n$ can be computed in $O(n^2)$ time.*

Proof. This is because for each pair of (i, j), ι^x_{ij} and ι^y_{ij} can be computed in constant time. By $\rho^x_{ij} = \iota^x_{ij} \setminus \{\mathsf{m} \cup \mathsf{mi}\}$ and $\rho^y_{ij} = \iota^y_{ij} \setminus \{\mathsf{m} \cup \mathsf{mi}\}$, we know ρ^x_{ij} and ρ^y_{ij} can also be computed in constant time. Since there are $O(n^2)$ pairs, it is clear $\mathcal{N}_x$ and $\mathcal{N}_y$ can be computed in $O(n^2)$ time. □

Remark 5. For the CDC_d, *i.e.* the Cardinal Direction Calculus for possibly disconnected regions, we can compute the projective interval relations in a similar way. In particular, similar to Proposition 23, the x-projective interval relation $\iota^x(\delta)$ of a basic CDC_d relation $\delta = [d^\chi]_{\chi \in \text{TileName}}$ is the IA relation associated to the vector (d_1, d_2, d_3), where $d_1 = \max\{d^{NW}, d^W, d^{SW}\}$, $d_3 = \max\{d^{NE}, d^E, d^{SE}\}$, and d_2 is 1 if both d_1 and d_3 are 1, and is $\max\{d^N, d^O, d^S\}$ otherwise. For example, if $\delta = W : NE$, then $\iota^x(\delta)$ is the IA relation associated to the vector $(1, 1, 1)$, which is di according to Proposition 21. Except the above modification, all the other procedures for computing the projective IA networks of a basic CDC_d network are exactly the same as that in CDC. This means the projective IA networks can also be computed in $O(n^2)$ time for CDC_d.

Appendix B. Proof of Proposition 12

Proposition 12. *Each consistent basic CDC network has a regular solution that is meet-free.*

Proof. Let $\mathfrak{a} = \{a_i\}_{i=1}^n$ be a regular solution of $\mathcal{N}$. Suppose $C(\mathfrak{a}) = \{c_{st}\colon\ 0 \leqslant s < n_x, 0 \leqslant t < n_y\}$ is the cell set of $\mathfrak{a}$, where $c_{st} = [\alpha_s, \alpha_{s+1}] \times [\beta_t, \beta_{t+1}]$ (see Eqs. (16)–(22) for definitions of $\alpha_s, \beta_t, a_i^r, c_{st}$ etc.). Because $\mathfrak{a}$ is regular, we know $a_i = a_i^r = \bigcup\{c_{st} \in C(\mathfrak{a})\colon\ a^\circ \cap c_{st} \neq \varnothing\}$. We show the meet-freeing process described in Section 4.2 can remove all meet-points while changing no CDC relations.

Take the x-direction as example. We prove this by using induction on m_x, the number of x-meet points of $\mathfrak{a}$. Suppose the hypothesis holds for any regular solutions with at most $m - 1$ x-meet points. Assume $\mathfrak{a} = \{a_i\}_{i=1}^n$ has $m > 0$ x-meet points. We show $\mathfrak{a}$ can be transformed into another regular solution with fewer x-meet points.

Suppose α_k is the largest x-meet point, where $0 < k < n_x$. Write $\alpha^* = (\alpha_k + \alpha_{k+1})/2$. The line $x = \alpha^*$ divides each cell c_{kl} $(0 \leqslant l < n_y)$ into two equal parts, written in order c^-_{kl} and c^+_{kl}. For each $1 \leqslant i \leqslant n$ and each $0 \leqslant l < n_y$, if $c_{kl} \subseteq a_i$ but $c_{k-1,l} \not\subseteq a_i$ then delete c^-_{kl} from a_i (cf. Fig. 8(a)). Write a_i^* for the remaining part of a_i, *i.e.*

$$a_i^* = \bigcup\{c_{st} \subseteq a_i\colon\ s \neq k\} \cup \bigcup\{c^+_{kt}\colon\ c_{kt} \subseteq a_i\} \cup \bigcup\{c^-_{kt}\colon\ c_{kt}, c_{k-1,t} \subseteq a_i\}. \tag{48}$$

Clearly, each a_i^* is contained in a_i. We claim $\mathfrak{a}^* = \{a_i^*\}_{i=1}^n$ is a regular solution of $\mathcal{N}$ which has fewer x-meet points.

Write $C^0 = \{c_{st} \in C(\mathfrak{a}){:}\ s \neq k\}$, $C^+ = \{c_{kt}^+{:}\ c_{kt} \in C(\mathfrak{a})\}$, $C^- = \{c_{kt}^-{:}\ c_{kt} \in C(\mathfrak{a})\}$. It is clear that $C(\mathfrak{a}^*)$, the cell set of $\mathfrak{a}^*$, is the union of C^0, C^+, and C^-. Moreover, each a_i^* is composed of cells in $C(\mathfrak{a}^*)$. This shows that $\mathfrak{a}^*$ is also regular.

Suppose $I_x(a_i) = [x_i^-, x_i^+]$. We can easily prove that $I_x(a_i^*) = I_x(a_i)$ if $x_i^- \neq \alpha_k$ and $I_x(a_i^*) = [\alpha^*, x_i^+] \subset I_x(a_i)$ if $x_i^- = \alpha_k$. It is then clear that two intervals $I_x(a_i^*)$ and $I_x(a_j^*)$ meet iff $I_x(a_i)$ and $I_x(a_j)$ meet. Since α_k and α^* cannot not be x-meet points of $\mathfrak{a}^*$, we know $\mathfrak{a}^*$ has fewer x-meet points than $\mathfrak{a}$.

We next prove that each a_i^* is a connected region. Similar to digital regions, we can introduce the notions of 4-neighbors and 4-connectedness in a cell set (cf. Definition 10), and prove that a_i^* is a connected region iff it is 4-connected (cf. Proposition 8).[6]

If $I_x(a_i)$ is contained in $[0, \alpha_k]$ or $[\alpha_{k+1}, \alpha_{n_x}]$, then $a_i^* = a_i$. Moreover, if $I_x(a_i)$ is contained in $[\alpha_k, \alpha_{n_x}]$, then $a_i^* = [\alpha^*, \alpha_{n_x}] \times [\beta_0, \beta_{n_y}] \cap a_i$. In these cases, a_i^* is clearly a connected region. Suppose $[\alpha_{k-1}, \alpha_{k+1}] \subseteq I_x(a_i)$. We show a_i^* is 4-connected. We first observe that there exists $0 \leqslant t < n_y$ such that $c_{k-1,t}, c_{kt}$ are both contained in a_i. This is due to the 4-connectedness of a_i in the cell set $C(\mathfrak{a})$. Second, we show each $C(\mathfrak{a}^*)$-cell c in a_i^* is 4-connected to $c_{k-1,t}$, hence 4-connected to c_{kt}^- and c_{kt}^+, in a_i^*.

Without loss of generality, we suppose $c = c_{kt'}^+ \in C^+$. By the 4-connectedness of a_i in $C(\mathfrak{a})$, we know $c_{k-1,t}$ is 4-connected to $c_{kt'}$ in a_i. Suppose $c_{k-1,t} = c^{(0)}, c^{(1)}, \ldots, c^{(m)} = c_{kt'}$ is a series of pairwise different $C(\mathfrak{a})$-cells in a_i such that $c^{(i+1)}$ is a 4-neighbor of $c^{(i)}$ for $i = 0, \ldots, m-1$. We replace each $c^{(i)}$ with one or two $C(\mathfrak{a}^*)$-cells if $c^{(i)}$ is not in $C(\mathfrak{a}^*)$. We replace $c^{(m)}$ by c_{kt}^-, c_{kt}^+ in order if $c^{(m-1)}$ is the left 4-neighbor of $c_{kt'}$; and by c_{kt}^+ otherwise. For $1 \leqslant i < m$, if $c^{(i-1)}$ is the left 4-neighbor of $c^{(i)} = c_{kt_1}$, we replace $c^{(i)}$ with $c_{kt_1}^-, c_{kt_1}^+$ in order; if $c^{(i+1)}$ is the left 4-neighbor of $c^{(i)}$, we replace $c^{(i)}$ with $c_{kt_1}^+, c_{kt_1}^-$ in order; and otherwise, replace $c^{(i)}$ with c_{kt}^+. It is straightforward to prove the revised series of $C(\mathfrak{a}^*)$-cells are 4-connected in a_i^*. This shows that each a_i^* is a connected region.

We then prove that $\mathrm{dir}(a_i^*, a_j^*) = \mathrm{dir}(a_i, a_j)$ for each pair of $i \neq j$. This is equivalent to proving that

$$\bigl(\exists c \in C(\mathfrak{a})\bigr) c \subseteq a_i \cap \chi(a_j) \quad \text{iff} \quad \bigl(\exists c^* \in C\bigl(\mathfrak{a}^*\bigr)\bigr) c^* \subseteq a_i^* \cap \chi\bigl(a_j^*\bigr) \tag{49}$$

holds for all $1 \leqslant i, j \leqslant n$ and all $\chi \in \textsc{TileName}$. For each $c \in C(\mathfrak{a})$, define $c^* = c$ if $c \in C^0$ and define $c^* = c_{kt}^+$ if $c = c_{kt}$. It is straightforward to prove that $c \subseteq \chi(a_j)$ iff $c^* \subseteq \chi(a_j^*)$. It is also clear that $c \subseteq a_i$ iff $c^* \subseteq a_i^*$. This shows that if there exists $c \in C(\mathfrak{a})$ s.t. $c \subseteq a_i \cap \chi(a_j)$ then there exists c^* in $C(\mathfrak{a}^*)$ s.t. $c^* \subseteq a_i^* \cap \chi(a_j^*)$. On the other hand, suppose there exists $c' \in C(\mathfrak{a}^*)$ s.t. $c' \subseteq a_i^* \cap \chi(a_j^*)$. If $c' \in C_0$, then $c' \subseteq a_i \cap \chi(a_j)$; if $c' = c_{kt}^+$ for some t, then $c_{kt} \subseteq a_i \cap \chi(a_j)$. If $c' = c_{kt}^-$ for some t, then by $c' \subseteq a_i^*$ and the definition of a_i^*, both $c_{kt}, c_{k-1,t}$ are contained in a_i. By the construction of a_j^*, we know $\mathcal{M}(a_j^*) = \mathcal{M}(a_j) \cap [\alpha^*, x_j^+] \times [0, n_y]$ if $x_j^- = \alpha_k$ and $\mathcal{M}(a_j^*) = \mathcal{M}(a_j)$ otherwise. It is easy to see that either $c_{kt} \subseteq \chi(a_j)$ or $c_{k-1,t} \subseteq \chi(a_j)$. This implies that either $c_{kt} \subseteq a_i \cap \chi(a_j)$ or $c_{k-1,t} \subseteq a_i \cap \chi(a_j)$. Therefore, Eq. (49) holds for every i, j, χ. This shows that $\mathfrak{a}^*$ is a regular solution of $\mathcal{N}$ which has fewer x-meet points than $\mathfrak{a}$. By the induction hypothesis, we know $\mathfrak{a}$ can be transformed into a regular meet-free solution. □

References

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

[2] P. Balbiani, J.-F. Condotta, L. Fariñas del Cerro, A new tractable subclass of the rectangle algebra, in: Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI-99), 1999, pp. 442–447.

[3] S. Cicerone, P. di Felice, Cardinal directions between spatial objects: the pairwise-consistency problem, Information Sciences 164 (1–4) (2004) 165–188.

[4] E. Clementini, R. Billen, Modeling and computing ternary projective relations between regions, IEEE Transactions on Knowledge and Data Engineering 18 (6) (2006) 799–814.

[5] E. Clementini, S. Skiadopoulos, R. Billen, F. Tarquini, A reasoning system of ternary projective relations, IEEE Transactions on Knowledge and Data Engineering 22 (2) (2010) 161–178.

[6] A.G. Cohn, J. Renz, Qualitative spatial reasoning, in: F. van Harmelen, V. Lifschitz, B. Porter (Eds.), Handbook of Knowledge Representation, Elsevier, 2007.

[7] I. Düntsch, H. Wang, S. McCloskey, A relation-algebraic approach to the region connection calculus, Theoretical Computer Science 255 (1–2) (2001) 63–83.

[8] F. Dylla, J.O. Wallgrün, On generalizing orientation information in $\mathcal{OPRA}_m$, in: Proceedings of the 29th Annual German Conference on AI (KI-06), 2006, pp. 274–288.

[9] M.J. Egenhofer, Reasoning about binary topological relations, in: Proceedings of the 2nd International Symposium on Advances in Spatial Databases (SSD'91), Springer-Verlag, London, UK, 1991, pp. 143–160.

[10] A.U. Frank, Qualitative spatial reasoning with cardinal directions, in: Proceedings of the 7th Austrian Conference on Artificial Intelligence (ÖGAI-91), 1991, pp. 157–167.

[11] C. Freksa, Using orientation information for qualitative spatial reasoning, in: Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, in: Lecture Notes in Computer Science, vol. 639, Springer, 1992, pp. 162–178.

[12] E.C. Freuder, Synthesizing constraint expressions, Communications of the ACM 21 (11) (1978) 958–966.

[13] E.C. Freuder, A sufficient condition for backtrack-free search, Journal of the ACM 29 (1) (1982) 24–32.

[14] A. Gerevini, J. Renz, Combining topological and size information for spatial reasoning, Artificial Intelligence 137 (1) (2002) 1–42.

[15] R.K. Goyal, Similarity assessment for cardinal directions between extended spatial objects, PhD thesis, The University of Maine, 2000.

[6] Another possible way is to transform these a_i^* into digital region as in Eq. (23).

[16] R.K. Goyal, M.J. Egenhofer, The direction-relation matrix: A representation of direction relations for extended spatial objects, in: UCGIS Annual Assembly and Summer Retreat, Bar Harbor, ME, 1997.
[17] R.K. Goyal, M.J. Egenhofer, Similarity of cardinal directions, in: Proceedings of the 7th International Symposium on Advances in Spatial and Temporal Databases (SSTD 2001), 2001, pp. 36–58.
[18] H.W. Guesgen, Spatial reasoning based on Allen's temporal logic, Technical report, International Computer Science Institute, 1989.
[19] P.W. Huang, C.H. Lee, Image database design based on 9D-SPA representation for spatial relations, IEEE Transactions on Knowledge and Data Engineering 16 (12) (2004) 1486–1496.
[20] A. Isli, A.G. Cohn, A new approach to cyclic ordering of 2D orientations using ternary relation algebras, Artificial Intelligence 122 (1–2) (2000) 137–187.
[21] A. Krokhin, P. Jeavons, P. Jonsson, Reasoning about temporal relations: The tractable subalgebras of Allen's interval algebra, Journal of the ACM 50 (5) (2003) 591–640.
[22] S. Li, Combining topological and directional information for spatial reasoning, in: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-07), 2007, pp. 435–440.
[23] S. Li, H. Wang, RCC8 binary constraint network can be consistently extended, Artificial Intelligence 170 (1) (2006) 1–18.
[24] S. Li, M. Ying, Region connection calculus: its models and composition table, Artificial Intelligence 145 (1–2) (2003) 121–146.
[25] S. Li, M. Ying, Generalized region connection calculus, Artificial Intelligence 160 (1–2) (2004) 1–34.
[26] G. Ligozat, Reasoning about cardinal directions, Journal of Visual Languages and Computing 9 (1) (1998) 23–44.
[27] G. Ligozat, J. Renz, What is a qualitative calculus? A general framework, in: Proceedings of the 8th Pacific Rim Trends in Artificial Intelligence (PRICAI-04), 2004, pp. 53–64.
[28] W. Liu, S. Li, J. Renz, Combining, RCC-8 with qualitative direction calculi: Algorithms and complexity, in: Proceedings of the 21th International Joint Conference on Artificial Intelligence (IJCAI-09), AAAI, 2009.
[29] D. Lücke, T. Mossakowski, D. Wolter, Qualitative reasoning about convex relations, in: Proceedings of the international conference on Spatial Cognition VI: Learning, Reasoning, and Talking about Space, Springer, 2008, p. 440.
[30] R. Moratz, Representing relative direction as a binary relation of oriented points, in: Proceedings of the 17th European Conference on Artificial Intelligence (ECAI-06), 2006, p. 407.
[31] I. Navarrete, A. Morales, G. Sciavicco, Consistency checking of basic cardinal constraints over connected regions, in: Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-07), 2007, pp. 495–500.
[32] B. Nebel, Computational properties of qualitative spatial reasoning: First results, in: Proceedings of the 19th Annual German Conference on Artificial Intelligence (KI-95), 1995, pp. 233–244.
[33] B. Nebel, H.-J. Bürckert, Reasoning about temporal relations: A maximal tractable subclass of Allen's interval algebra, Journal of the ACM 42 (1) (1995) 43–66.
[34] D.A. Randell, Z. Cui, A.G. Cohn, A spatial logic based on regions and connection, in: Proceedings of the 3rd International Conference on Knowledge Representation and Reasoning (KR-92), 1992, pp. 165–176.
[35] J. Renz, Maximal tractable fragments of the Region Connection Calculus: A complete analysis, in: Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI-99), 1999, pp. 448–454.
[36] J. Renz, G. Ligozat, Weak composition for qualitative spatial and temporal reasoning, in: Proceedings of the 11th International Conference on Principles and Practice of Constraint Programming (CP 2005), Springer, 2005, pp. 534–548.
[37] J. Renz, B. Nebel, On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the Region Connection Calculus, Artificial Intelligence 108 (1999) 69–123.
[38] J. Renz, B. Nebel, Qualitative spatial reasoning using constraint calculi, in: M. Aiello, I. Pratt-Hartmann, J. van Benthem (Eds.), Handbook of Spatial Logics, Springer, 2007, pp. 161–215.
[39] S. Skiadopoulos, C. Giannoukos, N. Sarkas, P. Vassiliadis, T. Sellis, M. Koubarakis, Computing and managing cardinal direction relations, IEEE Transactions on Knowledge and Data Engineering 17 (12) (2005) 1610–1623.
[40] S. Skiadopoulos, M. Koubarakis, Composing cardinal direction relations, Artificial Intelligence 152 (2) (2004) 143–171.
[41] S. Skiadopoulos, M. Koubarakis, On the consistency of cardinal direction constraints, Artificial Intelligence 163 (1) (2005) 91–135.
[42] S. Skiadopoulos, N. Sarkas, T. Sellis, M. Koubarakis, A family of directional relation models for extended objects, IEEE Transactions on Knowledge and Data Engineering 19 (8) (2007) 1116.
[43] J.G. Stell, Boolean connection algebras: a new approach to the region-connection calculus, Artificial Intelligence 122 (1) (2000) 111–136.
[44] A. Tarski, On the calculus of relations, Journal of Symbolic Logic 6 (1941) 73–89.
[45] R.E. Valdés-Pérez, The satisfiability of temporal constraint networks, in: Proceedings of the 6th National Conference on Artificial Intelligence (AAAI-87), 1987, pp. 256–260.
[46] F. Wolter, M. Zakharyaschev, Spatial reasoning in RCC-8 with Boolean region terms, in: Proceedings of the 14th European Conference on Artificial Intelligence (ECAI 2000), 2000, pp. 244–250.
[47] F. Wolter, M. Zakharyaschev, Spatio-temporal representation and reasoning based on RCC-8, in: Proceedings of the 7th Conference on Principles of Knowledge Representation and Reasoning (KR 2000), 2000, pp. 3–14.