A Relation-Based Merging Operator for Qualitative Spatial Data Integration and Conflict Resolution

Jan Oliver Wallgrün, Frank Dylla



Contact Address:

Dr. Thomas Barkowsky SFB/TR 8 Universität Bremen P.O.Box 330 440 28334 Bremen, Germany

Tel +49-421-218-64233 Fax +49-421-218-64239 barkowsky@sfbtr8.uni-bremen.de www.sfbtr8.uni-bremen.de

A Relation-based Merging Operator for Qualitative Spatial Data Integration and Conflict Resolution

Jan Oliver Wallgrün¹ and Frank Dylla²

¹ IRTG GRK 1498 Semantic Integration of Geospatial Information ² SFB/TR 8 Spatial Cognition University of Bremen, Germany {wallgruen,dylla}@sfbtr8.uni-bremen.de

Abstract. We describe a distance-based approach to integrate spatial information from different sources that is given in the form of constraint networks over relations from a qualitative spatial calculus. The distance functions are based on the notion of conceptual neighborhood between the spatial relations. We opt for a merging approach that is relation-based instead of model-based and as a result is also able to relax inconsistent networks. We investigate the properties of the proposed merging operators and describe an algorithm for their computation.

1 Introduction

We address the problem of combining information from different sources as it arises in the context of qualitative spatial representations. Formalisms for representing spatial information and reasoning about space using sets of qualitative relations are investigated within the research area of qualitative spatial reasoning (QSR), an active subfield of AI research (see [1, 2] for an overview). This research has led to a multitude of so-called qualitative spatial (and temporal) calculi dealing with different aspects of space such as topology or direction, suitable for tasks in which precise quantitative information is not available or not desirable.

A spatial arrangement of objects (for instance objects from a spatial database) can be described using the formalism of a qualitative constraint network (QCN) over a given qualitative calculus. A QCN is a graph in which the objects are represented by the vertices and edges are labeled with relations from the calculus that describe the qualitative relations holding between the objects (cmp. Fig. 1). The relations can be interpreted as constraints restricting possible geometries which can be assigned to the objects.

Our goal is to develop a solution to the problem of combining several constraint networks over the same calculus in a way that the result is always consistent. This problem occurs, for instance, when merging information from several databases containing qualitative information or when combining the believes of

multiple agents (e.g., humans or robots) expressed in a qualitative way. One important particularity of the QCN formalism which has to be taken into account when defining suitable merging operators is that in contrast to similar merging problems in propositional logic it is often not possible to express all disjunctions of possible models without admitting additional models. This aspect will play an important role in the investigations presented in this paper.

In earlier work [3], we developed an approach to relax a single inconsistent constraint network until it becomes consistent. The approach is based on the idea of using conceptual neighborhood between the relations of a qualitative calculus [4] to define a distance over constraint networks [5,6]. The merging operators we describe in this paper are a direct extension of this idea to the more general problem of combining several QCNs. We also put the approach onto a solid theoretical basis by relating it to work on logic-based merging [7,8] as recently suggested in [9, 10]. In contrast to the merging operators defined in [9] which are *model-based* in the sense that the result only depends on the models of the input networks, we instead aim for a relation-based approach in which every relation contained in the input QCN is able to affect the merging result. We argue that this approach is advantageous in many application scenarios, in particular when considering spatial database integration where the spatial relations stored in the database are based on independent observations (for a concrete example, cf. Sec. 3). Furthermore, our approach allows us to also deal with inconsistent input networks, which is not directly possible in a modelbased framework. In addition to defining the merging operators, we show how the result can be computed by incrementally relaxing the input networks until their intersections become consistent.

In the remainder of the paper, we first lay out the background of qualitative calculi, constraint networks, and conceptual neighborhood. Then we describe our merging scenario, develop rationality criteria for our operators, and define the operators themselves. Finally, we describe an algorithm for calculating the merging results.

2 Qualitative Spatial Representation

In the following overview on representation and reasoning with qualitative spatial calculi, we restrict ourselves to calculi over binary relations. However, our approach can be adapted to relations of higher arity as well.

2.1 Qualitative Spatial Calculi

A qualitative spatial calculus C defines a set \mathcal{B}_{C} of spatial relations over a domain of spatial objects \mathcal{D}_{C} (e.g., points, lines, regions). For every pair of objects from the domain exactly one relation from this set of so-called *base relations* holds (i.e., \mathcal{B}_{C} is jointly exhaustive and pairwise disjoint). For example, C could define a set of cardinal directions north-of (N), northwest-of (NW), west-of (W), southwest-of (SW), etc. plus the *identity relation* equal (EQ) for points in the plane.

To be able to express incomplete or imprecise spatial knowledge, the qualitative spatial calculus actually is concerned with the so-called set of general relations $\mathcal{R}_{\mathcal{C}}$ containing all possible unions of base relations. For instance, given $r = \text{NE} \cup \text{N} \cup \text{NW}$, ArB would express that A is either to the northeast, north, or northwest of B. Complete ignorance is expressed by the universal relation $U = \bigcup_{b \in \mathcal{B}_{\mathcal{C}}} b$. Here, we adopt the often used way of notating general relations as sets of base relations instead of unions, meaning that $\mathcal{R}_{\mathcal{C}} = 2^{\mathcal{B}_{\mathcal{C}}}$ and that the relation above will be denoted as $A\{\text{NE}, \text{N}, \text{NW}\}B$. Another special relation is the empty relation \varnothing which cannot be realized by any pair of objects.

In addition to defining relations, a qualitative calculus also defines a set $\mathcal{O}_{\mathcal{C}} = \{ \cap, \cup, \bar{}, \vee, \circ \}$ of operations over $\mathcal{R}_{\mathcal{C}}$. \cap, \cup , and $\bar{}$ are the operations of intersection, union, and complement which keep their set-theoretic meaning. The unary operation $\bar{}$ is the converse operation which tells us the relation holding between B and A from the relation holding between A and B, e.g. $\{N\}^{\vee} = \{S\}$. The binary composition operation \circ yields the relation that has to hold between A and B as well as between B and C, e.g., $\{N\} \circ \{SW\} = \{NW, W, SW\}$.

2.2 Qualitative Constraint Networks

A spatial arrangement of objects O_i can be described qualitatively based on a qualitative calculus \mathcal{C} by providing a set of relational facts using relations from $\mathcal{R}_{\mathcal{C}}$, e.g., $O_1\{N\}O_2, O_1\{E\}O_3, O_2\{S, SE\}O_3$, etc. The relations can be seen as constraints that restrict which values of $\mathcal{D}_{\mathcal{C}}$ can be assigned to the objects. We formally define such a qualitative spatial representation as a qualitative constraint network (QCN) in which the objects correspond to variables and the spatial relations correspond to constraints.

Definition 1 (Qualitative Constraint Network (QCN)) A qualitative constraint network over a qualitative calculus C is a pair (V, C) where:

- $-V = \{v_1, v_2, ... v_m\}$ is a set of variables
- $-C: V^2 \to \mathcal{R}_{\mathcal{C}}$ is a function mapping each pair of variables from V to a relation from $\mathcal{R}_{\mathcal{C}}$ where $C(v_i, v_j) = r \in \mathcal{R}_{\mathcal{C}}$ means that relation $v_i r v_j$ has to hold for the values assigned to v_i and v_j
- for all $1 \leq i, j \leq m$, $C(v_i, v_i) = id$ and $C(v_i, v_j) = C(v_j, v_i)^{\smile}$ holds (id is the identity relation of C).

In the remainder of this text we will also use the abbreviation C_{ij} for $C(v_i, v_j)$. One way to illustrate a qualitative constraint network is by a directed graph as shown in Fig. 1 containing a vertex for every variable v_i and one directed edge for every pair of variables v_i , v_j with i < j which is labeled by the corresponding relation. By convention, edges labeled with the universal relation U are omitted.

Naturally, the scene description provided by a QCN can be *consistent* or not. The QCN, hence, can be seen as a constraint satisfaction problem (CSP) in which the domain is typically infinite (e.g., points in \mathbb{R}^2). An assignment of values from $\mathcal{D}_{\mathcal{C}}$ to the variables v_i is a *solution* if it satisfies all constraints C_{ij} .

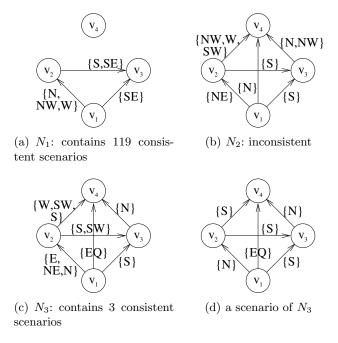


Fig. 1. Input QCNs N_1 to N_3 and a scenario of N_3 .

A QCN N is consistent if it has at least one solution. A QCN s is called atomic or a scenario if any C_{ij} consists of a single base relation. We say that a scenario s = (V, C') is a scenario of QCN N = (V, C) if all $C'_{ij} \subseteq C_{ij}$.

We will use a predicate consistent(N) to state that QCN N is consistent. We denote the set of all scenarios of N as $\langle N \rangle$ and the set of all consistent scenarios as $[\![N]\!]$. QCN will refer to the constraint network in which all constraints C_{ij} are U for a given set of variables V and a given calculus C) and, hence, $\langle QCN \rangle$ stands for the set of all possible scenarios given V and C. Figs. 1(a)–1(d) show several exemplary QCNs with cardinal direction constraints. The QCN in Fig. 1(d) is a scenario of the QCN in Fig. 1(c).

Deciding consistency of QCNs is NP-complete for many calculi but often tractable subalgebras are known. There exist two main methods for deciding consistency, both based on techniques developed for discrete CSPs. The so-called algebraic closure algorithm enforces a local consistency called path-consistency [12] and runs in $O(n^3)$ time for n variables. If algebraic closure is not sufficient to decide consistency for the relations occurring in the network, a backtracking search is performed [13] that recursively splits the constraints, until a level is reached which can be checked with algebraic closure.

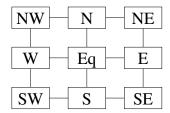


Fig. 2. A conceptual neighborhood graph for the cardinal direction calculus [11].

2.3 Conceptual Neighborhood

Our merging approach is based on the notion of similarity or distance between QCNs. Similarity is related to how the relations of the QCN can change, an aspect which is described by the notion of conceptual neighborhood introduced in [4]. Two base relations of a spatial calculus are conceptually neighbored, if they can be continuously transformed into each other without resulting in a third relation in between. For instance, N is conceptually neighbored to NW but not to W as one would have to pass through at least one other base relation (e.g., NW). The concrete conceptual neighborhood relation depends on the concrete set of continuous transformations one considers [4,3] which in turn need to be grounded in spatial change over time [14]. For this work, it is sufficient to assume that a suitable conceptual neighbor relation has been defined which is irreflexive and symmetric. It can be represented by the so-called conceptual neighborhood graph \mathcal{CNG} as illustrated in Fig. 2.

As proposed in [5], we will later use the shortest path distance between two base relations in the neighborhood graph to measure their similarity and extend this idea to complete scenarios.

3 Merging Qualitative Information

In the following, we describe operators for merging $n \ge 1$ QCNs $N_k = (V_k, C_k)$ over the same qualitative spatial calculus \mathcal{C} , representing information from different sources about a static arrangement of objects. Adopting a notation similar to that used in [9], the input is an n-tuple $\mathcal{N} = (N_1, N_2, ...N_n)$ referred to as the *input set*. We assume that the correct correspondences between the variables have already been established and—without loss of generality—that each N_k has the same set of variables $V = \{v_1, v_2, ...v_m\}$ which can be achieved in a preprocessing step. An exemplary merging problem could be to merge the QNCs N_1, N_2, N_3 from Fig. 1.

Two straightforward ways of combining QCNs are intergrating them conjunctively or disjunctively. Combining the QCNs conjunctively means to construct a new QCN by taking the intersection of the relations making up corresponding constraints.

Definition 2 (Intersection of QCNs) The intersection $N_1 \cap N_2$ of two QCNs $N_1 = (V, C_1)$ and $N_2 = (V, C_2)$ is a QCN N' = (V, C') with $C'(v_i, v_j) = C_1(v_i, v_j) \cap C_2(v_i, v_j)$ for all $1 \leq i, j \leq |V|$.

Obviously, when using the intersection to combine two or more QCNs, the resulting QCN can be inconsistent, even when the input networks themselves are all consistent. This is the case if the QCNs do not share a consistent scenario. Hence, intersection in general is too strict to serve as a suitable merging operator.

On the other hand, combining networks disjunctively means to take the union over corresponding relations.

Definition 3 (Union of QCNs) The union $N_1 \cup N_2$ of two QCNs $N_1 = (V, C_1)$ and $N_2 = (V, C_2)$ is a QCN N' = (V, C') with $C'(v_i, v_j) = C_1(v_i, v_j) \cup C_2(v_i, v_j)$ for all $1 \le i, j \le |V|$.

Using the union, consistent scenarios of the input networks are preserved but new ones may appear and the result will often be very unspecific reducing its usability. In addition, if all input networks are inconsistent, the resulting QCN may still be inconsistent.

In this work, we are interested in merging operators that are guaranteed to return a consistent result even when the input QCNs are not consistent (we only assume that all $C_{ij} \neq \emptyset$). Adopting the idea of distance-based merging [8, 7], we want our solution to be based on those models (consistent scenario in our case) that are as close as possible to all input networks simultaneously in a way that we will explain below. A main difference to existing work on merging QCNs [9, 10] is that we assume that all relations in the QCN can be considered independent and equally reliable information pieces that can have an effect on the result of the merging, while in the other approaches the result only depends on relations belonging to consistent scenarios. Consequentially, we will refer to our approach as relation-based in contrast to the model-based paradigm employed in the other approaches. To make this difference more clear, consider the example shown in Fig. 3. For convenience we introduce a compact notation for QCNs with three variables: A QCN N = (V, C) with variables $v_1, v_2,$ and v_3 is written as a triple of constraints $N = (C_{12}, C_{13}, C_{23})$. The input set in the example consists of the two QCNs $N_1 = (\{S\}, \{S\}, \{S, SE\})$ and $N_2 = (\{SW\}, \{S\}, \{SE\})$. N_1 has one consistent scenario, namely ($\{S\},\{S\},\{S\}$), while N_2 itself is a consistent scenario. As the model-based approach presented in [9] basically ignores relations that are not part of a consistent scenario such as SE in C_{23} of N_1 , it would consider the consistent scenarios ($\{SW\}, \{S\}, \{SE\}$) and ($\{S\}, \{S\}, \{S\}$) as equally plausible resolutions of the conflicts between the two QCNs. However, while the former scenario can be explained by a single small observation error in N_1 (C_{12} should have been SW instead of S), the latter would mean that there have been two small observation errors in N_2 . In contrast to such model-based operators, the merging operators we are going to define in Sec. 3.2 will treat the scenario $(\{SW\}, \{S\}, \{SE\})$ as a more plausible explanation.

Before we introduce the operators themselves, we start by formulating rationality criteria for relation-based QCN merging operators. As mentioned in

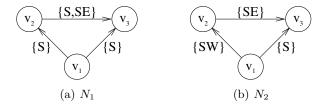


Fig. 3. Merging example: Input QCNs N_1 and N_2 .

the introduction, one particularity of the QCN merging scenario distinguishing it, for instance, from merging problems in propositional logic is that it is not always possible to combine models (or here consistent scenarios) into a single representation without obtaining additional models. Unfortunately, in many situations (e.g., merging databases) maintaining multiple hypotheses is undesirable or infeasible because of the additional complexity of tracking multiple hypotheses about the state of the world simultaneously. Hence, we define relation-based merging operators $\Delta(\mathcal{N})$ with $\mathcal{N} = \{N_1,...N_n\}$ that take an input set and return a single QCN and investigate how this requirement and the established rationality criteria fit together.

3.1 Rationality Criteria

To define the rationality criteria for our merging scenario, we follow criteria developed for information merging in a propositional setting (criteria (A1)–(A6) in [15] and (IC1)–(IC6) in [7]). Due to the special properties of QCNs and the fact that we are aiming at merging operators which are relation-based instead of model-based, we have to adapt the criteria leading to criteria (Q1)–(Q6) below. The resulting set of criteria turns out to be a specialization of the generic criteria for QCN merging (N1)–(N6) described in [10] but without assuming consistency of the input QCNs. We will point out where we make stronger demands tailored towards our particular merging approach.

The most basic requirement is that the merging result is a consistent QCN. Therefore we demand that $\Delta(\mathcal{N})$ always has to be consistent. In contrast, instantiating (N1) in [10] for our case would only demand that a QCN is returned but not necessarily a consistent one.

(Q1) $consistent(\Delta(\mathcal{N}))$

If the intersection of the input QCNs already is consistent, Δ should yield exactly this intersection. Again, the corresponding criterion (N2) in [10] would only make a weaker demand allowing the merging result to be inconsistent.

(Q2) if
$$consistent(\bigcap N_i)$$
 then $\Delta(\mathcal{N}) = \bigcap N_i$

The third criterion defined in [15] formalizes the 'irrelevance of syntax'. Concerned with defining criteria for model-based merging operators, they demand that the result of merging should only depend on the models of the input knowledge bases. In our relation-based case, it only makes sense to demand a significantly weakened version of the third criterion which basically claims that the order of input networks should not affect the result. For this, we define when two input sets are equivalent.

Definition 4 (Equivalence (\equiv) of input sets) Two input sets of QCNs $\mathcal{N} = (N_1, ..., N_n)$ and $\mathcal{N}' = (N'_1, ..., N'_n)$ are equivalent $(\mathcal{N} \equiv \mathcal{N}')$ iff there exists a bijection f between \mathcal{N} and \mathcal{N}' such that $\langle \langle N_k \rangle \rangle = \langle \langle f(N_k) \rangle \rangle$ for $1 \leq k \leq n$.

(Q3) if
$$\mathcal{N}_1 \equiv \mathcal{N}_2$$
 then $\Delta(\mathcal{N}_1) \equiv \Delta(\mathcal{N}_2)$

The fourth criterion is concerned with fairness of the merging operator stating that it must not give preference to one of the input knowledge bases. When merging two QCNs N_1 and N_2 and there is a scenario s part of the merging result which is also a scenario of N_1 , the same must hold for a scenario t of N_2 .

$$(\mathbf{Q4}) \ \exists s : s \in \langle \langle \Delta((N_1, N_2)) \rangle \land s \in \langle \langle N_1 \rangle \rangle \Leftrightarrow \exists t : t \in \langle \langle \Delta((N_1, N_2)) \rangle \land t \in \langle \langle N_2 \rangle \rangle$$

With the fifth property we demand that if we merge two input sets \mathcal{N}_1 and \mathcal{N}_2 individually and there is a scenario s part of both merging results, this scenario must also be part of the result of merging the input set resulting from combining the QCNs from \mathcal{N}_1 and \mathcal{N}_2 into a single set (written as $\mathcal{N}_1 \sqcup \mathcal{N}_2$).

(Q5) if
$$s \in \langle\!\langle \Delta(\mathcal{N}_1) \rangle\!\rangle$$
 and $s \in \langle\!\langle \Delta(\mathcal{N}_2) \rangle\!\rangle$ then $s \in \langle\!\langle \Delta(\mathcal{N}_1 \sqcup \mathcal{N}_2) \rangle\!\rangle$

Finally, in (Q6) we demand that if $\Delta(\mathcal{N}_1)$ and $\Delta(\mathcal{N}_2)$ have a common scenario, the reverse direction of (Q5) is also true. Taken together (Q5) and (Q6) state that if for two input sets the merging agrees on certain scenarios, these scenarios should be exactly the scenarios of the resulting QCN of the combined input set.

(Q6) if a
$$t$$
 exists with $t \in \langle\!\langle \Delta(\mathcal{N}_1) \rangle\!\rangle$ and $t \in \langle\!\langle \Delta(\mathcal{N}_2) \rangle\!\rangle$
then $(s \in \langle\!\langle \Delta(\mathcal{N}_1 \sqcup \mathcal{N}_2) \rangle\!\rangle) \Rightarrow s \in \langle\!\langle \Delta(\mathcal{N}_1) \rangle\!\rangle \land s \in \langle\!\langle \Delta(\mathcal{N}_2) \rangle\!\rangle)$

We now proceed by defining our relation-based merging operators for QCNs and will later discuss to what extent they satisfy the rationality criteria defined here.

3.2 The Merging Operators

Above, we introduced the conceptual neighborhood graph as a way to measure the distance or similarity of the base relations of a calculus, assuming that variations are caused by imperfect observations. Seeing the conceptual neighborhood graph $\mathcal{CNG}_{\mathcal{C}}$ of a calculus \mathcal{C} as an undirected graph (cmp. Sec. 2.3), we now define the distance $d_{B\leftrightarrow B}$ between the two base relations $b_i, b_j \in \mathcal{B}_{\mathcal{C}}$ as the shortest path distance between the corresponding nodes in the graph:

$$d_{B \leftrightarrow B}(b_i, b_j) = \text{shortest path distance between}$$

 $b_i \text{ and } b_j \text{ in } \mathcal{CNG}_{\mathcal{C}}$ (1)

The next step is to define the distance between two atomic qualitative constraint networks s = (V, C) and s' = (V, C') over the same set of m variables and the same calculus. For this we need an aggregation operator that determines how the distances between constraints in s_i, s_j given by $d_{B \leftrightarrow B}(b_i, b_j)$ are combined. Candidates for this aggregation operator which we will denote as a are the sum or the max operator. The distance itself is defined as:

$$d_{S \leftrightarrow S}^{\textcircled{0}}(s, s') = \underset{1 \leqslant i < j \leqslant m}{\textcircled{0}} d_{B \leftrightarrow B}(C_{ij}, C'_{ij})$$
 (2)

The notion behind our merging operators $\Delta(\mathcal{N})$ is that the resulting QCN is built from the consistent scenarios that are closest to the input networks together with all inconsistent scenarios that are at most as distant as these consistent scenarios. Therefore, we further need to define the distance between a scenario and a general constraint network and based on that the distance between a scenario and the set of input networks (\mathcal{N}) .

For determining how close a scenario s is to a constraint network N we consider all scenarios of N and take the distance to the closest one. The resulting distance $d_{S \leftrightarrow N}^{\circledcirc}(s,N)$ is then given by

$$d_{S \leftrightarrow N}^{\circledcirc}(s, N) = \min_{s' \in \langle N \rangle} d_{S \leftrightarrow S}^{\circledcirc}(s, s') \tag{3}$$

To measure the distance between a scenario s and the set \mathcal{N} of all input networks N_i , we need to aggregate over the individual distances $d_{S \leftrightarrow N}^{\circledcirc}(s, N_i)$. To do this, we introduce another aggregation operator \circledast . Again, sum and max seem to be natural candidates for this aggregation operator. In the general case, the resulting distance is given by

$$d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(s, \mathcal{N}) = \underset{1 \le k \le n}{\circledast} d_{S \leftrightarrow N}^{\otimes}(s, N_k) \tag{4}$$

To construct the final merging result we take the set $S^{\textcircled{@},\textcircled{\circledast}}(\mathcal{N})$ of all scenarios that are closer or as close to \mathcal{N} as the closest consistent scenarios wrt. $d_{S\leftrightarrow\mathcal{N}}^{\textcircled{@},\textcircled{\circledast}}$.

$$S^{\tiny\textcircled{@},\textcircled{\circledast}}(\mathcal{N}) = \{ s \in \langle\!\langle QCN \rangle\!\rangle \mid \forall s' \in \llbracket QCN \rrbracket : d_{S \leftrightarrow \mathcal{N}}^{\textcircled{@},\textcircled{\circledast}}(s',\mathcal{N}) \geqslant d_{S \leftrightarrow \mathcal{N}}^{\textcircled{@},\textcircled{\circledast}}(s,\mathcal{N}) \} \quad (5)$$

As the final step, the resulting QCN is constructed by taking the union of all the scenarios in $S^{\otimes,\circledast}(\mathcal{N})$.

$$\Delta^{\hat{\otimes}, \circledast}(\mathcal{N}) = \bigcup_{s \in S^{\hat{\otimes}, \circledast}(\mathcal{N})} s \tag{6}$$

As discussed previously, the final union step, may lead to additional scenarios in $\Delta^{\otimes,\circledast}(\mathcal{N})$ that are not contained in $S^{\otimes,\circledast}(\mathcal{N})$ which is the price one has to pay to end up with a single QCN.

3.3 Properties of the Operators

We now consider the rationality criteria defined in Sec. 3.1 and test whether they are satisfied by $\Delta^{\circledcirc,\circledast}(\mathcal{N})$. If this is not the case, we investigate to what degree this is due to the final union step, meaning that the criteria would at least hold for the scenarios contained in $S^{\circledcirc,\circledast}(\mathcal{N})$. Unless stated otherwise, the considerations in the following hold for all combinations of $\circledcirc,\circledast\in\{\Sigma,\max\}$.

Theorem 1 $\Delta^{\otimes, \circledast}$ satisfies (Q1).

Proof. $\langle\!\langle QCN \rangle\!\rangle$ contains at least one consistent scenario for every qualitative calculus (for example, with all constraints set to id). As a result, $S^{\otimes,\circledast}(\mathcal{N})$ contains at least one consistent scenario which is then also a consistent scenario of $\Delta^{\otimes,\circledast}(\mathcal{N})$.

Theorem 2 $\Delta^{\otimes, \circledast}$ satisfies (Q2).

Proof. (1) s is a scenario of $\bigcap N_i$ iff s is a scenario of all N_i . (2) $d_{S \leftrightarrow N}^{\circledcirc}(s, N_i) = 0$ iff $s \in \langle N_i \rangle$ (for both $\circledcirc = \Sigma$ and $\circledcirc = \max$). Based on (1) and (2), $s \in \langle \bigcap N_i \rangle$ iff $d_{S \leftrightarrow N}^{\circledcirc, \circledast}(s, \mathcal{N}) = 0$ (for both $\circledcirc = \Sigma$ and $\circledcirc = \max$). From this and the definition of $S^{\circledcirc, \circledast}$ it follows that if $\bigcap N_i$ is consistent then $s \in \langle \bigcap N_i \rangle \Leftrightarrow s \in S^{\circledcirc, \circledast}(\mathcal{N})$. Taking the union of all scenarios in $S^{\circledcirc, \circledast}(\mathcal{N})$ then only reconstructs $\bigcap N_i$ from its scenarios and hence $\Delta^{\circledcirc, \circledast}(\mathcal{N}) = \bigcap N_i$.

Theorem 3 $\Delta^{\otimes,\circledast}$ satisfies (Q3).

Proof. All considered operators for \circledast in Eq. 4 are commutative and associative so that the order of the N_i has no effect on the result.

Proving (Q4) is significantly more complex, mainly due to the final union step. Therefore, we start by showing that $S^{\circledcirc,\circledast}$ satisfies a corresponding version of (Q4).

Theorem 4
$$\exists s : s \in S^{@,\circledast}((N_1, N_2)) \land s \in \langle N_1 \rangle \Leftrightarrow \exists t : t \in S^{@,\circledast}((N_1, N_2)) \land t \in \langle N_2 \rangle holds.$$

Proof. All distance functions in Eqs. 1–4 are symmetric. Hence, for every scenario s in $S^{@,\circledast}((N_1,N_2))$ that is a scenario of N_1 there has to be a scenario t of N_2 with the same distance $d_{S \leftrightarrow \mathcal{N}}^{@,\circledast}$ which therefore is also contained in $S^{@,\circledast}((N_1,N_2))$.

We now first look at the case that $\circledast = \sum$ and define two auxiliary theorems. First we show that for two scenarios s,t, no scenario u exists with a smaller distance $d_{S \leftrightarrow \mathcal{N}}^{\otimes, \Sigma}$ to the input set (s,t) than the direct scenario distance between s and t. We here introduce the notation $d_{ij}^{s,t}$ for the distance $d_{B \leftrightarrow B}$ between the base relation forming the constraint C_{ij} in s and the corresponding base relation in t.

Theorem 5 For arbitrary scenarios $s, t, u, d_{S \leftrightarrow N}^{\textcircled{@}, \Sigma}(u, (s, t)) \ge d_{S \leftrightarrow S}^{\textcircled{@}}(s, t)$.

Proof. $rianlge = \sum (\text{case 1}): d_{S \leftrightarrow \mathcal{N}}^{\sum, \sum}(u, (s, t))$ can be rewritten as $\sum_{i,j} d_{ij}^{s,u} + d_{ij}^{u,t}$. As $d_{B \leftrightarrow B}$ satisfies the triangle inequation, this can never be smaller than $d_{S \leftrightarrow S}^{\sum}(s, t) = \sum_{i,j} d_{ij}^{s,t}$.

From this result, we get the interesting property that $S^{\otimes,\Sigma}((N_1,N_2))$ always contains at least one scenario of N_1 and of N_2 .

Theorem 6 If $s \in \langle N_1 \rangle$ and $t \in \langle N_2 \rangle$ and $d_{S \leftrightarrow S}^{\otimes}(s,t)$ is minimal over all scenarios of N_1 and N_2 then $s, t \in S^{\otimes, \Sigma}((N_1, N_2))$.

Proof. From Theorem 5 it follows that no scenario can have a smaller distance $d_{S \leftrightarrow N}^{\underline{\otimes}, \underline{\sum}}$ to (N_1, N_2) than the smallest distance $d_{S \leftrightarrow S}^{\underline{\otimes}}(s, t)$ with $s \in \langle N_1 \rangle$ and $t \in \langle N_2 \rangle$. The fact that

$$d_{S \leftrightarrow N}^{\textcircled{a}, \sum}(s, (N_1, N_2)) = d_{S \leftrightarrow N}^{\textcircled{a}, \sum}(t, (N_1, N_2)) = d_{S \leftrightarrow S}^{\textcircled{a}}(s, t)$$

proves the theorem.

From this theorem, it directly follows that:

Corollary 1 $\Delta^{\otimes, \sum}$ satisfies (Q4).

We now turn to the case \circledast = max, starting with two auxiliary theorems. In the following, we say that a scenario u lies between two other scenarios s and t if each base relation in u is one from the shortest path connecting the corresponding base relations in s and t. We show that for such a scenario there exists a mirrored one with the same overall distance to (s,t).

Theorem 7 For every scenario u lying between scenarios s and t, there exists a scenario v with $d_{ij}^{s,v} = d_{ij}^{t,u}$ and $d_{ij}^{t,v} = d_{ij}^{s,u}$ and $d_{S \leftrightarrow \mathcal{N}}^{\odot, \oplus}(u, (s, t)) = d_{S \leftrightarrow \mathcal{N}}^{\odot, \oplus}(v, (s, t))$.

Proof. v can be built by replacing each base relation in u by the mirrored one along the connecting shortest path. Since both max and \sum are commutative the distance $d_{S \leftrightarrow \mathcal{N}}^{\circledcirc, \circledcirc}$ to (s,t) does not change.

Theorem 8 For every scenario $u \in S^{\otimes,\max}((s,t))$ with $d_{ij}^{s,u} = 0$ there exists a scenario $v \in S^{\otimes,\max}((s,t))$ with $d_{ij}^{t,v} = 0$.

Proof. Either u lies between s and t or there exists a scenario $u' \in S^{\otimes, \max}((s,t))$ with $d_{ij}^{s,u'} = 0$ lying between s and t because $d_{S \leftrightarrow \mathcal{N}}^{\otimes, \max}(u',(s,t)) \leqslant d_{S \leftrightarrow \mathcal{N}}^{\otimes, \max}(u',(s,t))$ for such a u'. According to Theorem 7, there also exists the mirrored scenario v in $S^{\otimes, \max}((s,t))$ which has $d_{ij}^{t,v} = 0$.

In Theorem 4, we showed that $S^{\otimes, \circledast}$ satisfies (Q4). From Theorem 8 it now follows that the final union step preserves this property.

Theorem 9 $\Delta^{\otimes,\max}$ satisfies (Q4).

Proof. The union step can only generate a scenario s of N_1 if for every i, j there exists at least one scenario x in $S^{\otimes,\max}((N_1,N_2))$ with $d_{ij}^{s,x}=0$. However, according to Theorem 8 for every such scenario there exists a scenario t in $S^{\otimes,\max}((N_1,N_2))$ with $d_{ij}^{t,y}=0$ to a scenario y of N_2 . Therefore, if the union step generates a scenario of N_1 , it will also generate one of N_2 .

Unfortunately, for (Q5), a simple counterexample shows that $\Delta^{\circledcirc,\circledast}$ does not satisfy the criterion.

Theorem 10 (Q5) does not hold for $\Delta^{\otimes, \circledast}$.

Proof. For $\circledast = \sum$ and $\textcircled{a} = \sum$, the following is a counter example: QCNs

$$N_1 = (\{S\}, \{SW\}, \{S\}),$$

$$N_2 = (\{SW\}, \{SW\}, \{SE\}), \text{ and }$$

$$N_3 = (\{SW\}, \{SW\}, \{SW\})$$

are given. With $\mathcal{N}_1 = (N_1)$ and $\mathcal{N}_2 = (N_2, N_3)$, the scenario ($\{SW\}, \{SW\}, \{SW\}$) is contained in $\Delta^{\Sigma,\Sigma}(\mathcal{N}_1)$ (result of the union of other scenarios in $S^{@,\circledast}(\mathcal{N}_1)$) and in $\Delta^{\Sigma,\Sigma}(\mathcal{N}_2)$ but not in $\Delta^{\Sigma,\Sigma}(\mathcal{N}_1 \sqcup \mathcal{N}_2)$.

As we see, the counterexample makes use of the union step to generate a scenario which breaks the criterion. Indeed, in Theorem 11 we show that $S^{@,\circledast}$ does satisfy a corresponding version of (Q5).

Theorem 11 $s \in S^{@,\circledast}(\mathcal{N}_1) \land s \in S^{@,\circledast}(\mathcal{N}_2) \Rightarrow s \in S^{@,\circledast}(\mathcal{N}_1 \sqcup \mathcal{N}_2)$ holds.

Proof. For $\circledast \in \{\sum, \max\}$, $d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(s, \mathcal{N}_1 \sqcup \mathcal{N}_2) = d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(s, \mathcal{N}_1) \circledast d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(s, \mathcal{N}_2)$ holds. Let us assume that s is a consistent scenario (case 1) with $s \in S^{\otimes, \circledast}(\mathcal{N}_1)$ and $s \in S^{\otimes, \circledast}(\mathcal{N}_2)$. Then for every other consistent scenario t,

$$d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(t, \mathcal{N}_1) \geqslant d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(s, \mathcal{N}_1)$$

and

$$d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(t, \mathcal{N}_2) \geqslant d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(s, \mathcal{N}_2)$$

has to hold and as a result also

$$d_{S \leftrightarrow \mathcal{N}}^{@,\circledast}(t, \mathcal{N}_1 \sqcup \mathcal{N}_2) \geqslant d_{S \leftrightarrow \mathcal{N}}^{@,\circledast}(s, \mathcal{N}_1 \sqcup \mathcal{N}_2).$$

Hence, s has to be contained in $S^{\circledcirc,\circledast}(\mathcal{N}_1 \sqcup \mathcal{N}_2)$. If s is an inconsistent scenario (case 2), we know from the definition of $S^{\circledcirc,\circledast}$ that for every consistent scenario t in $S^{\circledcirc,\circledast}(\mathcal{N}_1 \sqcup \mathcal{N}_2)$,

$$d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(s, \mathcal{N}_1) \leqslant d_{S \leftrightarrow \mathcal{N}}^{\otimes, \circledast}(t, \mathcal{N}_1)$$

and

$$d_{S \leftrightarrow \mathcal{N}}^{\textcircled{\otimes}, \textcircled{\circledast}}(s, \mathcal{N}_2) \leqslant d_{S \leftrightarrow \mathcal{N}}^{\textcircled{\otimes}, \textcircled{\circledast}}(t, \mathcal{N}_2)$$

and as a result also

$$d_{S \leftrightarrow \mathcal{N}}^{@, \circledast}(s, \mathcal{N}_1 \sqcup \mathcal{N}_2) \leqslant d_{S \leftrightarrow \mathcal{N}}^{@, \circledast}(t, \mathcal{N}_1 \sqcup \mathcal{N}_2)$$

holds. Hence, s in this case also has to be contained in $S^{\otimes, \circledast}(\mathcal{N}_1 \sqcup \mathcal{N}_2)$.

As a result, the loss of this property is a direct consequence of the demand that the merging result has to be a single QCN implemented by the final union step.

For (Q6), however, it can be shown that not only $\Delta^{\otimes, \circledast}$ does not satisfy the criterion but that this is also the case for $S^{\otimes, \circledast}$ and, hence, not a result of the final union step. It currently seems possible that it is rather the result of including inconsistent scenarios with a distance smaller than the closest consistent scenario. While it would be counterintuitive not to include these scenarios when inconsistent scenarios with a larger distance are included, the properties of such alternatives need to be investigated in more detail as part of future research.

Theorem 12 (Q6) does not hold for $\Delta^{\textcircled{@},\textcircled{\$}}$ (and also not for $S^{\textcircled{@},\textcircled{\$}}$).

Proof. For both $\circledast \in \{\sum, \max\}$ (with $\boxtimes = \sum$), the following is a counter example: N_1, N_2, N_3, N_4 are QCNs over three variables and

$$N_1 = (\{SW\}, \{SW\}, \{S\}),$$

$$N_2 = (\{SW\}, \{SW\}, \{SW\}),$$

$$N_3 = (\{S\}, \{SW\}, \{SW\}),$$

$$N_4 = (\{S\}, \{S\}, \{S\}).$$

With $\mathcal{N}_1=(N_1,N_2,N_3)$ and $\mathcal{N}_2=(N_4)$, the scenario $(\{S\},\{SW\},\{S\})$ is contained in $S^{\Sigma,\circledast}(\mathcal{N}_1\sqcup\mathcal{N}_2)$ (and $\Delta^{\Sigma,\circledast}(\mathcal{N}_1\sqcup\mathcal{N}_2)$) but not in $S^{\Sigma,\circledast}(\mathcal{N}_2)$ (nor in $\Delta^{\Sigma,\circledast}(\mathcal{N}_2)$).

Overall, we have one criterion (Q6) that our merging operators do not satisfy in general and an additional one (Q5) where this is a consequence of the demand of getting a single QCN as result. While for the first problem alternatives may exist, it seems questionable whether suitable operators can be found which do not suffer from similar problems caused by the final union step.

Merging operators are often classified into majority or arbitration operators. We briefly state here without proof that $\Delta^{\otimes,\max}$ satisfies the notion of an arbitration operator (see (A7) in [15]), while $\Delta^{\otimes,\sum}$ is preferable if a majority-based resolution is desired (see (M7) in [15]).

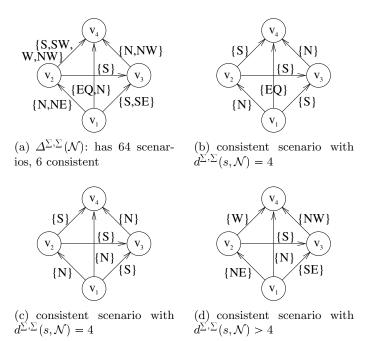


Fig. 4. The solution of $\Delta^{\Sigma,\Sigma}(\mathcal{N})$ according to Alg.1.

4 An Algorithm to Compute $oldsymbol{\Delta}^{\tiny{\textcircled{@}},\textcircled{\circledast}}(\mathcal{N})$

In this section, we describe an algorithm to compute $\Delta^{\circledcirc,\circledast}(\mathcal{N})$. While the time complexity is still exponential in the worst-case, the algorithm is based on the following two notions in order to significantly improve its performance in practice, in particular when the input QCNs are rather close to each other: (1) candidate scenarios are considered in order of increasing distance to \mathcal{N} as given by $d_{S \leftrightarrow \mathcal{N}}^{\circledcirc,\circledast}(s,\mathcal{N})$, and (2) the expensive consistency checking is delayed as long as possible and does not have to consider individual scenarios. The generation of scenarios in order of increasing distance is based on a set of relax functions. The function $relaxC(C, d_C)$ yields the relation consisting of all base relations which have minimal distance d_C for $d_C \geqslant 0$ to a base relation in constraint C_{ij} .

$$relaxC(C_{ij}, d_C) = \left\{ b \in \mathcal{B}_C \mid \min_{b' \in C_{ij}} d_{B \leftrightarrow B}(b, b') = d_C \right\}$$
 (7)

Based on it, the function $relaxN^{\circledcirc}(N,d_N)$ yields a set of networks in which constraints C_{ij} have been changed using relaxC with parameter e_{ij} so that aggregation with \circledcirc over e_{ij} yields d_N .

$$relax N^{\circledcirc}(N = (V, C), d_N) =$$

$$\left\{ N' = (V, C') \mid C'_{ij} = relax C(C_{ij}, e_{ij}) \land \bigotimes_{1 \leq i,j \leq m} e_{ij} = d_N \right\}$$
(8)

Algorithm 1 Merging algorithm

```
procedure \Delta^{\otimes, \circledast}(\mathcal{N})
 1: S \leftarrow (V, C) with all C_{ij} = \emptyset
 2: d_{\mathcal{N}} \leftarrow 0; consistent \leftarrow false
 3: repeat
          R \leftarrow relax \mathcal{N}^{\otimes, \circledast}(\mathcal{N}, d_{\mathcal{N}})
 4:
          for all (N'_1,...N'_n) \in R do
 5:
              I \leftarrow \bigcap_{i=1}^{n} N_i'
 6:
              if \forall i, j : C_{ij} \neq \emptyset holds for I then
 7:
 8:
                   if consistent(I) then consistent \leftarrow true end if
 9:
10:
               end if
           end for
11:
           d_{\mathcal{N}} \leftarrow d_{\mathcal{N}} + 1
12:
13: until consistent
14: return S
```

Hence, each scenario $s \in \langle\!\langle QCN \rangle\!\rangle$ is a scenario of a network in $relaxN^{\circledcirc}(N,d_N)$ if and only if $d_{S \leftrightarrow N}^{\circledcirc}(s,N) = d_N$. The function $relax\mathcal{N}^{\circledcirc,\circledast}(\mathcal{N},d_{\mathcal{N}})$ then yields a set of modified input sets

The function $relax \mathcal{N}^{\circledcirc,\circledast}(\mathcal{N}, d_{\mathcal{N}})$ then yields a set of modified input sets where each modified network N_k has been modified using $relax N^{\circledcirc}(N_k, e_k)$ such that the e_k are aggregated with \circledast to $d_{\mathcal{N}}$.

$$relax \mathcal{N}^{\otimes, \circledast}(\mathcal{N} = (N_1, N_2, ... N_n), d_{\mathcal{N}}) =$$

$$\left\{ (N'_1, ..., N'_n) \mid N'_k = relax N^{\otimes}(N_k, e_k) \wedge \underset{1 \leq k \leq n}{\circledast} e_k = d_{\mathcal{N}} \right\}$$

$$(9)$$

Given this set of relax functions the basic version of our algorithm proceeds as follows (see Algorithm 1): The outer loop increases $d_{\mathcal{N}}$ and considers scenarios s with $d_{S \leftrightarrow \mathcal{N}}^{\otimes, \oplus}(s, \mathcal{N}) = d_{\mathcal{N}}$. The algorithm stops when there is at least one consistent scenario among these. All scenarios are collected in the QCN S which in the end will contain the merging result. To find the scenarios with $d_{S \leftrightarrow \mathcal{N}}^{\otimes, \oplus}(s, \mathcal{N}) = d_{\mathcal{N}}$, the following happens inside the outer loop: Relaxed input sets are generated with $relax\mathcal{N}^{\otimes, \oplus}(\mathcal{N}, d_{\mathcal{N}})$ for the given $d_{\mathcal{N}}$. For every element $N' = (N'_1, ... N'_n)$ of the resulting set, the algorithm takes the intersection over all N'_k . It can be shown, that $d_{S \leftrightarrow \mathcal{N}}^{\otimes, \oplus}(s, \mathcal{N}) = d_{\mathcal{N}}$ holds for a scenario s if and only if s is a scenario of one of the networks I generated in this inner loop. Hence, when the intersection does not contain empty constraints (which would mean it does not have scenarios at all), we add all scenarios of I to S through the union operation in line 8. In addition, it is checked whether I is consistent using standard QSR techniques. If consistent, we know that we have found at least one consistent scenario and the algorithm will stop after all remaining tuples in R have been processed (see line 5).

Fig. 4(a) shows an example of employing Algorithm 1 to compute $\Delta^{\Sigma,\Sigma}$ for merging the QCNs N_1 to N_3 from Fig. 1. The resulting network contains 64 scenarios, six of them consistent (for comparison the input QCNs have 2187, 6,

and 18 scenarios). The minimum distance is three (four inconsistent scenarios) and the maximum distance is four. Two of the consistent scenarios (Figs. 4(b) and 4(c)) have distance four, the other four consistent scenarios with distance > 4 result from the final union step (e.g. Fig. 4(d)).

5 Conclusions

We have introduced a family of distance-based operators for merging qualitative spatial information from different sources given in the form of qualitative constraint networks. The operators are relation-based in the sense that they treat every relation as an independent piece of information that may affect the result and can be applied to inconsistent input networks. We analyzed to which degree the operators satisfy the previously established rationality criteria. Deviations from the criteria are partially due to the fact that QCNs cannot express all disjunctions of scenarios without leading to additional scenarios. Nevertheless, alternatives need to be evaluated as part of future research. We also presented an algorithm for computing the merging result by incrementally relaxing the input networks and delaying expensive consistency checking as long as possible in order to increase the average-case efficiency compared to a naive implementation. The next step is to perform an experimental evaluation in which the algorithm is applied to real integration problems.

Acknowledgements

The authors would like to thank Diedrich Wolter for helpful comments and suggestions. Financial support by the Deutsche Forschungsgemeinschaft under grants IRTG GRK 1498 Semantic Integration of Geospatial Information and SFB/TR 8 Spatial Cognition is gratefully acknowledged.

References

- Cohn, A.G., Hazarika, S.M.: Qualitative spatial representation and reasoning: An overview. Fundamenta Informaticae 46(1-2) (2001) 1–29
- Renz, J., Nebel, B.: Qualitative spatial reasoning using constraint calculi. In Aiello, M., Pratt-Hartmann, I.E., van Benthem, J.F., eds.: Handbook of Spatial Logics. Springer (2007) 161–215
- Dylla, F., Wallgrün, J.O.: Qualitative spatial reasoning with conceptual neighborhoods for agent control. Journal of Intelligent and Robotic Systems 48(1) (2007) 55–78
- Freksa, C.: Conceptual neighborhood and its role in temporal and spatial reasoning. In Singh, M., Travé-Massuyès, L., eds.: Decision Support Systems and Qualitative Reasoning. (1991) 181 – 187
- 5. Bruns, H.T., Egenhofer, M.J.: Similarity of spatial scenes. In: Seventh International Symposium on Spatial Data Handling, Taylor & Francis (1996) 173–184

- Moratz, R., Freksa, C.: Spatial reasoning with uncertain data using stochastic relaxation. In Brauer, W., ed.: Fuzzy-Neuro Systems 98. Infix; Sankt Augustin (1998) 106–112
- Grégoire, É., Konieczny, S.: Logic-based approaches to information fusion. Information Fusion 7(1) (2006) 4–18
- 8. Lin, J., Mendelzon, A.O.: Merging databases under constraints. International Journal of Cooperative Information Systems 7 (1996) 55–76
- Condotta, J.F., Kaci, S., Schwind, N.: A framework for merging qualitative constraints networks. In Wilson, D., Lane, H.C., eds.: Proceedings of the Twenty-First International Florida Artificial Intelligence Research Society Conference, May 15-17, 2008, Coconut Grove, Florida, USA, AAAI Press (2008) 586–591
- Condotta, J.F., Kaci, S., Marquis, P., Schwind, N.: Merging qualitative constraints networks using propositional logic. In: 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty(ECSQARU'09). (2009)
- Goyal, R.K., Egenhofer, M.J.: Similarity of cardinal directions. In Goos, G., Hartmanis, J., van Leeuwen, J., eds.: Advances in Spatial and Temporal Databases. Volume 2121 of LNCS. Springer-Verlag Berlin Heidelberg (2001) 36–55
- Mackworth, A.K.: Consistency in networks of relations. Artificial Intelligence 8 (1977) 99–118
- 13. Nebel, B.: Solving hard qualitative temporal reasoning problems: Evaluating the efficiency of using the ord-horn class. CONSTRAINTS **3**(1) (1997) 175–190
- 14. Galton, A.: Qualitative Spatial Change. Oxford University Press (2000)
- 15. Konieczny, S., Pérez, R.P.: On the logic of merging. In: Sixth International Conference on Principles of Knowledge Representation and Reasoning (KR'98). (1998) 488–498