Artificial Intelligence 225 (2015) 51–76

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

# On redundant topological constraints

Sanjiang Li [a,b,*], Zhiguo Long [a], Weiming Liu [c], Matt Duckham [d], Alan Both [d]

[a] *Centre for Quantum Computation & Intelligent Systems, University of Technology Sydney, Sydney, Australia*
[b] *Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China*
[c] *Baidu (China) Co., Ltd., Shanghai, China*
[d] *Department of Infrastructure Engineering, University of Melbourne, Melbourne, Australia*

ARTICLE INFO

*Article history:*
Received 13 May 2014
Received in revised form 26 March 2015
Accepted 29 March 2015
Available online 2 April 2015

*Keywords:*
Qualitative spatial reasoning
Region connection calculus
Redundancy
Prime subnetwork
Distributive subalgebra

ABSTRACT

Redundancy checking is an important task in the research of knowledge representation and reasoning. In this paper, we consider redundant qualitative constraints. For a set $\Gamma$ of qualitative constraints, we say a constraint $(xRy)$ in $\Gamma$ is *redundant* if it is entailed by the rest of $\Gamma$. A *prime subnetwork* of $\Gamma$ is a subset of $\Gamma$ which contains no redundant constraints and has the same solution set as $\Gamma$. It is natural to ask how to compute such a prime subnetwork, and when it is unique. We show that this problem is in general intractable, but becomes tractable if $\Gamma$ is over a tractable subalgebra $\mathcal{S}$ of a qualitative calculus. Furthermore, if $\mathcal{S}$ is a subalgebra of the Region Connection Calculus RCC8 in which weak composition distributes over nonempty intersections, then $\Gamma$ has a *unique* prime subnetwork, which can be obtained in cubic time by removing all redundant constraints simultaneously from $\Gamma$. As a by-product, we show that any path-consistent network over such a distributive subalgebra is minimal and globally consistent in a qualitative sense. A thorough empirical analysis of the prime subnetwork upon real geographical data sets demonstrates the approach is able to identify significantly more redundant constraints than previously proposed algorithms, especially in constraint networks with larger proportions of partial overlap relations.



## 1. Introduction

Qualitative spatial reasoning is a common subfield of artificial intelligence and geographical information science, and has applications ranging from natural language understanding [13], robot navigation [46], geographic information systems (GISs) [18], sea navigation [54], to high level interpretation of video data [48].

Typically, the qualitative approach represents spatial information by introducing a relation model on a domain of spatial entities, which could be points, line segments, rectangles, or arbitrary regions. In the literature, such a relation model is often called a *qualitative calculus* [34]. In the past three decades, dozens of spatial (as well as temporal) qualitative calculi have been proposed in the literature (cf. [11]). Among these, Interval Algebra (IA) [1] and the RCC8 algebra [41] are widely known as the most influential qualitative calculi for representing qualitative temporal and, respectively, spatial information. Other well-known qualitative calculi include Point Algebra (PA) [51], Cardinal Relation Algebra (CRA) [33], Rectangle Algebra (RA) [24], the RCC5 algebra [41], etc.

* Corresponding author at: Centre for Quantum Computation & Intelligent Systems, University of Technology Sydney, Sydney, Australia.
*E-mail addresses:* sanjiang.li@uts.edu.au (S. Li), zhiguo.long@student.uts.edu.au (Z. Long), liuweiming@baidu.com (W. Liu), matt@duckham.org (M. Duckham), aboth@student.unimelb.edu.au (A. Both).

http://dx.doi.org/10.1016/j.artint.2015.03.010

Using a qualitative calculus $\mathcal{M}$, we represent spatial or temporal information in terms of relations in $\mathcal{M}$, and formulate a spatial or temporal problem as a set of qualitative constraints (called a *qualitative constraint network*). A qualitative constraint has the form $(xRy)$, which specifies that the two variables $x, y$ are related by the relation $R$. The *consistency problem* is to decide whether a set of qualitative constraints can be satisfied simultaneously. The consistency problem has been investigated in depth for many qualitative calculi in the literature, see e.g., [51,50,33,40,39,43,42,14,55,37,28,35,45,30].

In this paper, we consider the important problem of redundant qualitative constraints. Given a set $\Gamma$ of qualitative constraints, we say a constraint $(xRy)$ in $\Gamma$ is *redundant* if it is entailed by the rest of $\Gamma$, i.e., removing $(xRy)$ from $\Gamma$ will not change the solution set of $\Gamma$. It is natural to ask when a network contains redundant constraints and how to get a non-redundant subset without changing the solution set. We call a subset of $\Gamma$ a *prime subnetwork* of $\Gamma$ if it contains no redundant constraints and has the same solution set as $\Gamma$.

The redundancy problem (i.e., the problem of determining if a constraint is redundant in a network) is related to the minimal label problem (cf. [38,8,20,36,3]). A qualitative constraint network $\Gamma$ is called *minimal* if for each constraint $(xRy)$ in $\Gamma$, $R$ is the minimal (i.e., the *strongest*) relation between $x$, $y$ that is entailed by $\Gamma$. Roughly speaking, the minimal network removes 'redundant' or 'unnecessary' *basic relations* from each constraint, while the redundancy problem removes 'redundant' or 'unnecessary' *constraints* from the constraint network.

We show in this paper that it is in general co-NP hard to determine if a constraint is redundant in a *qualitative constraint network*. But if all constraints in $\Gamma$ are taken from a tractable subclass[1] $\mathcal{S}$ then a prime subnetwork can be found in polynomial time. For example, if $\mathcal{S}$ is a tractable subclass of RCC5 or RCC8 that contains all basic relations, then we can find a prime subnetwork in $O(n^5)$ time. Furthermore, if $\mathcal{S}$ is a subalgebra of RCC5 or RCC8 in which weak composition distributes over nonempty intersections, then $\Gamma$ has a unique prime subnetwork, which is obtained by removing all redundant constraints from $\Gamma$. We also devise a cubic time algorithm for computing this unique prime subnetwork, which has the same time complexity as the two approximate algorithms of Wallgrün [52].

As a by-product, we identify an important class of subalgebras of qualitative calculi, called *distributive subalgebras*. A subalgebra $\mathcal{D}$ of a qualitative calculus $\mathcal{M}$ is called *distributive* if weak composition distributes over nonempty intersections in $\mathcal{D}$. We show that any path-consistent network over a distributive subalgebra is weakly globally consistent and minimal, where *weakly global consistency* is a notion similar to but weaker than the well-known notion of global consistency (cf. Definition 5). For RCC8, we identify two maximal distributive subalgebras which are not contained in any other distributive subalgebras, one contains 41 relations and the other contains 64. The 41 relations contained in the first subalgebra are exactly the convex RCC8 relations identified in [8].

In this paper, we are mainly interested in topological constraints, as these are the most important kind of qualitative spatial information. A large part of our results can easily be transplanted to other qualitative calculi like PA, IA, CRA and RA. In particular, let $\mathcal{M}$ be one of PA, IA, CRA and RA and $\mathcal{S}$ a distributive subalgebra of $\mathcal{M}$ over which path-consistency implies consistency. Then we can show that any path-consistent network over $\mathcal{S}$ is globally consistent and minimal.[2] For ease of presentation, we state and prove these results only for RCC5 and RCC8, but indicate in Table 5 which result is applicable to which calculus.

### *1.1. Motivation*

As in the case of propositional logic formulas [32], redundancy of qualitative constraints "often leads to unnecessary computation, wasted storage, and may obscure the structure of the problem" [5].[3] Finding a prime subnetwork can be useful in at least the following aspects: a) computing and storing the relationships between spatial objects and hence saving space for storage and communication; b) facilitating comparison (or measure the distance) between different constraint networks; c) unveiling the essential network structure of a network (e.g., being a tree or a graph with a bounded tree-width); and d) adjusting geometrical objects to meet topological constraints [52].

To further motivate our discussion, we focus on one specific application to illustrate the application area a. and briefly explain how redundancy checking or finding a prime subnetwork helps to solve the application areas b–d.

Fig. 1 gives a small example of a set of spatial regions formed by the geographic "footprints" associated with placenames in the Southampton area of the UK. The footprints are derived from crowd-sourced data, formed from the convex hull of the sets of coordinate locations at which individuals used the placenames on social media (cf. [25]). Communicating and reasoning with the qualitative aspects of such data may require the storage and manipulation of large numbers of complex geometries with millions of vertices or large constraint networks with millions of relations.

Even for the small example in Fig. 1, the 84 footprints then require $84 * 83/2 = 3486$ stored relations. The moderate-sized footprint data set from which Fig. 1 is adapted contains a total of 3443 footprints which leads to a constraint network with 5,925,403 relations. Similarly, a moderate-sized geographic data set of only 1559 statistical areas in Tasmania, explored further in later sections, contains in total 3,093,551 vertices. In the case of both footprints and statistical areas, many of the relationships can be inferred, and computing the prime subnetwork can potentially reduce the number of stored

[1] Here a subclass $\mathcal{S}$ is *tractable* if the consistency of any constraint network defined over $\mathcal{S}$ can be determined in polynomial time.

[2] For PA, IA, CRA and RA, weakly global consistency is equivalent to global consistency.

[3] It is worth noting that redundancy can also enhance propagation during computation (cf. [10]).