

A Formal Approach to Qualitative Reasoning on Topological Properties of Networks

Andrea Rodríguez^{1,3} and Claudio Gutierrez^{2,3}

¹ Department of Computer Science, Universidad de Concepción
`andrea@udec.cl`

² Department of Computer Science, Universidad de Chile
`cgutierrez@dcc.uchile.cl`

³ Center for Web Research, Universidad de Chile

Abstract. Qualitative reasoning uses a limited set of relevant distinctions of the domain to allow a flexible way of representing and reasoning about it. This work presents a conceptual framework for qualitative reasoning about information networks from a spatial-topological point of view. We consider the properties of connectivity and some topological invariants to describe the structural characteristics of and the topological relationships between networks. The paper presents a data model for networks which generalizes the notion of graph, founded in algebraic and topological considerations. Such conceptual tool can be useful in different domains, from social to technological networks.

1 Introduction

Topological properties are related to the concept of *connectivity*, upon which different relations may be defined; for example, *overlapping*, *inside*, *disjoint* and *meet*. An important extension beyond the power of traditional query languages for graph and networks is the incorporation of topological relationships into the primitives of query languages. These facts have already been recognized in the spatial domain, where topological relations have played an important role for spatial reasoning [9,1] and query languages [13].

A formal approach to this subject is beneficial for several reasons. The formalism serves as a tool to identify and derive systematically relationships while avoiding redundant and contradicting relations and notions, and helps proving the completeness of the set of relationships. The formal method can be applied to determine the relation between any two networks and to reason formally about them. Algorithms to determine relationships can be specified exactly, and mathematically sound models will help to define formally the relationships. The formalism can extend definitions on networks to more general concepts on networks.

Topological and spatial reasoning is a well established subject dealing with the development of formal models for defining and reasoning about topological characteristics of spatial objects and topological relations [10]. When trying to extend these techniques to sets of overlapping networks the following important issues arise:

- The standard notion of graph does not suffice to represent faithfully the information. Note that an edge could be part of a network which does not have both of its end-nodes (e.g. the sub-network of Latin American routes consisting of Chilean cities and LanChile’s routes: a flight going out of the country is an edge which has only one node in the sub-network).
- Graphs do not behave like sets of point nor as “point-less objects.” Hence from a topological point of view, the approach should be a mixture of point-set topology and point-less topology
- The classic set of operations on graphs must be enriched to be able to express in a flexible manner spatio-topological relations among networks.
- Standard spatial notions cannot be taken as on-the-shelf technology for networks.

This work presents a formal framework for topological reasoning on relationships among network properties, particularly focusing on connectivity aspects. It discusses the level of abstraction needed, that is, what are the “good” objects in this domain, and what are the “good” operations to act over these objects. It presents several operations on them and studies their properties. In addition, it introduces the main approaches used in spatial reasoning for defining topological relations between networks. Such definitions can be applied to broad application domains, such as social networks, technological networks, and conceptual or metadata networks. In particular we:

Related Work. To the best of our knowledge the subject of this paper has not been addressed formally. There is related work on spatial reasoning which is useful in our context, and that we describe below.

In the spatial domain, qualitative reasoning of topological relations has obtained particular attention from the research community, since it allows automatic reasoning based on a cognitively plausible representation of spatial concepts [10]. Most of the work on topological qualitative reasoning define ontologies of spatial entities, where some fundamental concepts are contact, parthood and boundary [4,1,9]. Stell and Worboys [14] present a theory of parthood and boundary that can be connected to different formalisms for topological relations. This formalism represents set of regions as a bi-Heyting algebra [7] and expresses certain important constructors on the regions purely in terms of the operations presented in the algebra. One of the examples they give is the algebra of graphs. We follow some of these ideas when looking for the right data structure for networks and operations over it, but consider a more general notion of graph.

Two well known ontologies for topological spatial relations are the Region Connected Calculus (RCC) [1] and the point-set topological model [9,8]. RCC is a logic-based formalization of topological relations that uses a basic connectivity relation between closed regions. The point-set topological model defines topological relations based on the set intersection of the interior, boundary and exterior of spatial objects. Such formalism uses relation algebra [6] to create an inference mechanism given by the composition of topological relations [11]. Although both models result in the same set of topological relations between spatial regions, they differ in their reasoning capabilities. While reasoning with relation algebras

has computational advantages, axiomatic theories are richer in their expressive power.

2 The (abstract) Model

In this section we introduce a general framework to model networks and present the algebraic properties of different categories of objects and their operations. We will use the basic graph terminology as in Diestel’s Book [12].

Definition 1 (Semigraph). Let $U = (V_U, E_U)$ be a graph.

1. A semigraph over U is a pair (V, E) , where $V \subseteq V_U$ and $E \subseteq E_U$.
2. A net is a semigraph (V, E) such that for each $uv \in E$ it holds that either $u \in V$ or $v \in V$.

Note that a *graph* (in the classical sense) is a semigraph such that $E \subseteq V \times V$. In what follows, there will be always a *universe* graph $U = (V_U, E_U)$ which will be the “space” on which the objects we deal with live in.

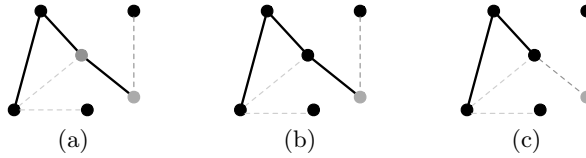


Fig. 1. Basic notions: (a) semigraph, (b) net, and (c) graph. Dark nodes and edges belong to the semigraph, net or graph, respectively.

Notations. Let $V \subseteq V_U$ be a set of nodes, let $E \subseteq E_U$ be a set of edges, and let G be an arbitrary semigraph. We will denote by V_G its set of nodes and by E_G its set of edges.

Use uv to denote the undirected edge $\{u, v\}$. A node v and an edge e are *incident* if $e = vw$ for some w . $\text{inc}(V)$ is the set of edges $\{uv \in E_U : u \in V \vee v \in V\}$. Similarly, $\text{inc}(E)$, is the set of nodes $\{v \in V_U : uv \in E\}$ $\text{sg}(V)$ will denote the semigraph $(V, \text{inc}(V))$. Similarly, $\text{sg}(E)$ will denote the semigraph $(\text{inc}(E), E)$, and $\text{sg}(G)$ will denote the semigraph $(V_G \cup \text{inc}(E), E_G \cup \text{inc}(V))$.

Definition 2 (Basic operations on semigraphs). Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be semigraphs.

1. The union of G_1 and G_2 (denoted $G_1 \cup G_2$) is the semigraph $(V_1 \cup V_2, E_1 \cup E_2)$.
2. The intersection of G_1 and G_2 (denoted $G_1 \cap G_2$) is the semigraph $(V_1 \cap V_2, E_1 \cap E_2)$.
3. The difference of G_1 and G_2 (denoted $G_1 - G_2$) is the semigraph $(V_1 - V_2, E_1 - E_2)$. In particular, the complement of G_2 , denoted G_2^c , is the semigraph $U - G_2$.

Using the fact that the product of two Boolean algebras with the operations defined pairwise is again a Boolean algebra we get:

Proposition 1. *The set of semigraphs with the operations of union, intersection and complement, together with 0 defined as (\emptyset, \emptyset) and $1 = (U, U \times U)$ is a Boolean algebra.*

The Algebraic Structure of Networks. One can enrich the Boolean Algebra structure of semigraphs described above by defining closure operators over semigraphs, and hence, a structure of Topological space (Kuratowski space).

A closure operator (cl) must satisfy some properties for each element of the domain. Two basic properties are $cl(\emptyset) = \emptyset$ and $G \subseteq cl(G)$. In addition, by property $cl(G \cup H) = cl(G) \cup cl(H)$, one needs to specify only cl over single nodes and single edges. From the idempotence property ($cl(cl(G)) = cl(G)$) it follows that $cl(v)$ should add no nodes (or do trivial things like adding nodes independent of v , e.g. the whole universe, all isolated nodes, etc.) Similarly for $cl(e)$ for an edge e . In fact, the only two natural choices for closure are: (1) $cl_E(G) = sg(E(G))$, and (2) $cl_V(G) = sg(V(G))$. But the topologies they generate are not essentially different:

Lemma 1. *Let T_E and T_V the topologies induced by the closure operators cl_E and cl_V respectively. Then G is open in T_E if and only if G is closed in T_V .*

Heyting Algebras Via Closure Operators. Every topology provides a complete Heyting algebra in the form of its open set lattice. The Heyting algebra is defined as follows: objects are open sets; operations are set-theoretical union and intersection; and the element $A \Rightarrow B$ is the interior of the union of $A^c \cup B$, where A^c denotes the complement of the open set A .

For the operator cl_V the open sets are standard full subgraphs of U (complements of the closed sets in the topology T_V). Here the border (given by the topology) of G is the set of edges in the complement of G which are incident to G . The operator cl_E is the dual of the previous one.

The Heyting Algebra of Nets. Note that the objects defined in the Heyting algebras induced by the open sets of the topological spaces defined above were essentially graphs (or complements of graphs).

It is possible to extend the set of objects to be considered to semigraphs without loose edges (what we have called *nets*) and still having a structure of Heyting algebra by slightly modifying the operations of join and meet. Nets are operationally generated as follows: (1) choose a set of nodes V and (2) choose a set of edges incident to V . The induced operations (in order to be closed in this new universe) are the standard union and the meet $G_1 \wedge G_2$ defined as $(G_1 \cap G_2) \cap sg(V(G_1 \cap G_2))$ (observe that the standard intersection of semigraphs could leave isolated edges).

Proposition 2. *Let $R(U)$ be the set of nets over U . If we define $0 = \emptyset$ (the empty semigraph), $G_1 \vee G_2$ as the union, $G_1 \wedge G_2$ as defined above, and $G_1 \Rightarrow G_2$ as $sg(V_1^c) \cup G_2$, then $(R(U), \vee, \wedge, \Rightarrow, 0)$ is a Heyting algebra.*

Note that this is not a bi-Heyting algebra because the existence of nets R such that there are edges $uv \notin R$ with $u, v \in R$. Example: let $R = U - \{uv\}$. Then there is no unique minimal solution for $U \leq R \cup X$, because, for example, $X_1 = (\{u\}, uv)$ and $X_2 = (\{v\}, uv)$ are minimal solutions.

Interestingly, the problem described in the previous paragraph is the only barrier to have a bi-Heyting algebra.

Proposition 3. *Let $R^*(U)$ the set of full nets over U (i.e. nets such that the complement has no isolated edges, that is, it is again a net). If we define the same operations as in Proposition 2 and $1 = U$ and $G_1 \setminus G_2$ as $(sg(E_2^c) \cup G_2^c) \wedge G_1$, then $((U), \vee, \wedge, \Rightarrow, \setminus, 1, 0)$ is a bi-Heyting algebra.*

3 Possible Approaches to Define Connectivity in Networks

We examined the structure and operations over networks. In this section we will study the notion of when two objects (semigraphs) in this universe are “connected” or have “relationships”.

3.1 A Pure Topological Approach

Although pure topological notions are oriented to capture the concept of *continuity*, several notions from topology can be borrowed to speak of spatial notions [8,9]. We start from these notions, but will be interested in the notions of *connectivity* or *relationship*, and thus some concepts will naturally not be applicable in our context.

Following the approach of Egenhofer [8] we will build the framework on the notions of *boundary* and *interior*. Due to the particularities of our domain, we will add a third notion, that of *frontier*.

Definition 3 (Interior, Frontier, Boundary, Closure). *Let U be the universal graph, and H a semigraph in U .*

1. *The boundary of a semigraph H (in U), denoted $\partial(H)$, is the set of edges which are incident to H and its complement, i.e., the set of edges uv of U such that $u \in H$ and $v \notin H$. (Note that edges $uv \notin H$ with $u \in H$ and $v \in H$ are not in the boundary).
In particular, we define $\delta(H) = \partial(H) \cap H$ as the real boundary.*
2. *The frontier of a semigraph H (in U), denoted $\text{fr}(H)$, is the set of nodes of H adjacent to nodes not in H . (Or equivalently: the set of nodes of H incident to $\partial(H)$).
In particular, we define $\text{fr}'(H)$, the real frontier, as the subset of the nodes of $\text{fr}(H)$ incident to edges not in H .*
3. *The interior of a semigraph H (in U), denoted $\text{int}(H)$, is the semigraph consisting of all nodes and edges of H not incident with elements not in H .*
4. *The closure of a semigraph H (in U), denoted $\text{cl}(H)$, is the semigraph $H \cup \partial(H)$.*

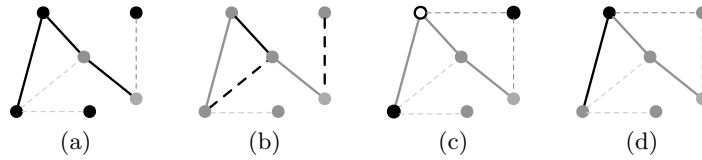


Fig. 2. Basic topological notions. Dark nodes and edges constitute the semigraph. (a) reference semigraph; (b) boundary (continuous dark line is the real boundary); (c) frontier (filled dark nodes form the real frontier); and (d) interior.

Proposition 4 (The boundary ∂)

1. $\partial(H)$ has no interior.
2. $\partial(H) = \text{cl}(H) \cap \text{cl}(G - H)$, i.e. the boundary of H is the intersection of the closure of H and its complement.
3. $\partial(H) = \partial(G - H)$, i.e. the boundary of H equals the boundary of its complement.
4. $\partial(H) = \emptyset$ iff $H = \emptyset$ or $H = G$.

Based on these definitions, one can derive topological relations between networks by considering the intersections between their topological invariants (i.e., interior, boundary and frontier), that is, the intersections between the components of networks that do not change under topological transformations (e.g. translation, scaling, and rotation).

Given any two connected nets H_1 and H_2 in U , let $\text{int}(H_1)$ and $\text{int}(H_2)$ be their interiors, and $D(H_1)$ and $D(H_2)$ be the union of their respective real frontiers with real boundary. Table 1 presents the eight matrices that derive the 10 possible 4-intersection matrices between nets.

3.2 The Region Connected Calculus RCC

RCC is a formalism for spatial reasoning that takes regions of space instead of points of classical geometry as primitives. For this, a primitive notion of connectivity is introduced by means of a binary predicate $C(x, y)$, whose semantics is that of “ x is connected to y .”

In the context of networks, the basic primitive is naturally defined as follows:

Definition 4. $C(x, y)$ is true iff there is a path from x to y in $x \cup y$ (where path is the standard notion in graph theory [12])

Note that if $x \cap y \neq \emptyset$ then $C(x, y)$, but the notion defined allow $C(x, y)$ to be true even though $x \cap y = \emptyset$.

The RCC definitions (we will use the RCC-8 framework) are axiomatized in standard first-order logic using quantifiers over variables ranging over the objects of the domain (regions in the spatial case) (see Table 2). The axioms for semigraphs coincides roughly with the naive intuition in the spatial domain.

Table 1. The basic 2x2 matrices upon which all possible 4-intersections matrices are derivable. Dark nodes and edges belong to one or both nets.

Matrices		Examples	Matrices		Examples
$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 0 & 0 \\ D(H_1) & 0 & 0 \end{matrix}$			$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 0 & 0 \\ D(H_2) & 0 & 1 \end{matrix}$		
$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 1 & 0 \\ D(H_1) & 0 & 0 \end{matrix}$			$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 1 & 0 \\ D(H_1) & 0 & 1 \end{matrix}$		
$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 1 & 0 \\ D(H_1) & 1 & 0 \end{matrix}$			$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 1 & 0 \\ D(H_1) & 1 & 1 \end{matrix}$		
$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 1 & 1 \\ D(H_1) & 1 & 0 \end{matrix}$			$\begin{matrix} \text{int}(H_2) & D(H_2) \\ \text{int}(H_1) & 1 & 1 \\ D(H_1) & 1 & 1 \end{matrix}$		

Table 2. RCC-8 definitions depending on the range of the quantification of the variables involved: over semigraphs and over nets. $\text{inc}(x, y)$ means x is incident to y .

Relation	Interpretation	Quantif. over semigraphs
$DC(x, y)$	x is disconnected from y x and y in $x \cup y$	No path between
$P(x, y)$	x is a part of y	$x \subseteq y$
$PP(x, y)$	x is a proper part of y	$x \subset y$
$EQ(x, y)$	x is equivalent with y	$x = y$
$O(x, y)$	x overlaps y	$x \cap y \neq \emptyset$
$DR(x, y)$	x is discrete from y	$x \cap y = \emptyset$
$PO(x, y)$	x partially overlaps y	$x \cap y \neq \emptyset \wedge x \not\subseteq y \wedge y \not\subseteq x$
$EC(x, y)$	x is externally connected to y	$x \cap y = \emptyset \wedge \text{inc}(x, y)$
$TPP(x, y)$	x is a tangential proper part of y	$x \subset y \wedge \text{inc}(x, y^c)$
$NTTP(x, y)$	x is a nontangential proper part of y	$x \subset y \wedge \neg \text{inc}(x, y^c)$

The problems are subtleties centered on non-existent edges between two nodes of the domain. In this framework it is more evident the insufficiency of graphs as basic data structure for qualitative reasoning. In such case, $C(x, y)$ must be defined as intersection of nodes and the definitions would collapse into standard set theoretical notions among nodes.

4 Conclusions

This work presents a formal framework for qualitative reasoning about topological properties of networks. It studies the structure of sets of overlapping networks

from a spatio-topological point of view, defines a data structure and operations associated with networks, and states and proves main properties of them.

For future work, we plan to select a standard set of operations that serves as basic for a query language design and that can relate the abstract model to the approaches for defining the connectivity in networks.

Acknowledgements. This work has been funded by Nucleus Millenium Center for Web Research, Grant P04-067-F, Mideplan, Chile.

References

1. D. A. Randell, Z. Cui, A. G. Cohn, *A Spatial logic based on regions and connection*, Proc. 3rd. Int. Conf. on Knowledge Representation and Reasoning, Morgan Kaufmann, San Mateo, pp. 165-176.
2. A. G. Cohn, B. Bennett, J. Gooday, N. M. Gotts, *Qualitative Spatial Representation and Reasoning with the Region Connection Calculus*, Geoinformatica, 1, 1-44 (1997)
3. J. de Kleer, J. Brown, *A Qualitative Physics Based on Confluences*. Artificial Intelligence 24, 7-83, (1984)
4. B. Smith, *Mereotopology - A Theory of parts and boundaries*, Data and Knowledge Engineering, 20, 287-303 (1996)
5. J. Sharma, D. Fleweling, M. Egenhofer, *A Qualitative Spatial Reasoner*, International Symposium on Spatial Data Handling, pp. 665-681, September 1994.
6. A. Tarski, *On The Calculus of Relations*, Journal of Symbolic Logic, 6(3), 73-89 (1941)
7. S. Vickers, *Topology via Logic*, Cambridge University Press, 1989.
8. M. J. Egenhofer, *A Formal Definition of Binary Topological Relationships*, Lecture Notes in Computer Science, Vol. 367, pp. 457-472, June 1989.
9. M. J. Egenhofer, R. Franzosa *Point-Set Topological Spatial Relations*, International Journal of Geographic Information Science, 5(2), 161-174 (1991)
10. Loiviero Stock *Spatial and Temporal Reasoning*, Kluwer Academic Publishers, 1997.
11. M. J. Egenhofer *Deriving the Composition of Binary Topological Relations*, Journal of Visual Languages and Computing, 5, 133-149 (1994)
12. R. Diestel, *Graph Theory*, Springer, New York, 1997.
13. E. Clementini, J. Sharma, M. Egenhofer, *Modeling Topological Spatial Relations: Strategies for Query Processing*, International Journal of Computer and Graphics, 18 (6), 815-822 (1994)
14. J. G. Stell, M. F. Worboys, *The Algebraic Structure of Sets of Regions*, Lecture Notes in Computer Science, Vol. 1329, pp. 163-174, October 1997.