A Framework for Qualitative Spatial Reasoning Based on the Connection Patterns of Regions

Kazuko TAKAHASHI and Takao SUMITOMO

School of Science&Technology, Kwansei Gakuin University, 2-1, Gakuen, Sanda, 669-1337, JAPAN (E-mail) ktaka@kwansei.ac.jp scbc0071@ksc.kwansei.ac.jp

Abstract

We propose a new framework called PLCA for qualitative spatial reasoning. PLCA provides a symbolic representation for the figure in twodimensional space, that focuses on the connections between regions. It is based on the simple objects: points, lines, circuits and areas. The entire figure is represented as a combination of these objects. Pairs of areas, circuits or lines never cross. PLCA provides not only mereological reasoning between regions, but also topological reasoning. Moreover, spatial semantic reasoning is possible by adding attributes, such as the properties that hold in the regions. We compare PLCA with existing qualitative spatial reasoning methods, such as RCC and the 9intersection model, and show that PLCA is upper compatible with them.

Keywords: qualitative spatial reasoning, regions, objects, RCC

1 Introduction

Recent advances of computer performance and network infrastructure have increased the opportunities for various users to use spatial data, such as figures and images. Since spatial data are generally stored and processed as numerical data, their processing requires more memory and time, as compared to processing text data. Therefore, an efficient mechanism for spatial data processing is desired. The more refined the data used are, the clearer and more precise the figure is. However, refined data are not always necessary unless a clear, precise figure is required. It can be sufficient to know the number of objects or the positional relationships of the objects in a figure depending on a user's purpose.

Qualitative spatial reasoning is a method that treats images or figures qualitatively, not quantitatively like numerical data, by extracting the information necessary for a user's purpose [Cohn *et al.*, 2001; Renz, 2002; Stock, 1997]. It has many applications, including Geographical Information Systems(GIS) and image processing. Region Connection Calculus(RCC) [Randell *et al.*, 1992] is one of the representatives of the theories for qualitative spatial reasoning. It considers



Figure 1: Simulation of fire

a space as a set of regions paying attention only to their relative positions. It is a simple, elegant theory that is suitable for mereological reasoning, but the abstraction is rather coarse when treating realistic problems. For example, consider the qualitative simulation of a fire burning from a certain spot. The area burning increases as time passes. We want to construct a barrier to stop it from spreading to the right. Figure 1 shows a qualitative model of this example. The circle shows the predicted region of burning at some instant, and the rectangle shows the barrier. The spread of the fire can be stopped in case (a), but not in case (b). In both cases, however, RCC regards the circle and the rectangle as *overlapping*. In GIS or image processing applications, such discrimination of the connection patterns of regions is essential in many cases.

In this paper, we propose a new framework called PLCA which provides a symbolic representation for spatial data using the simple objects: points(P), lines(L), circuits(C)and areas(A). In PLCA, no pair of areas has a part in common, and every pair is either point-connected, line-connected or disconnected. Symbolic representation enables compact information at the level that is suitable for the user's purpose and allows rapid processing. Its simple, clear data structure makes the system easy to implement and feasible. Moreover, if we add attributes such as size, shape or direction, or semantical properties (e.g. what an object stands for) to each object in PLCA, we can perform more fruitful reasoning. PLCA is more expressive than RCC since it distinguishes the connection patterns of regions, while RCC does not. It can provide detailed information, such as the connected segments of the borders and the number of pieces that constitute a region. PLCA has the same expressive power with another method of qualitative spatial reasoning, called the 9-intersection model. We discuss the correspondence of PLCA with these two representative methods, and show that PLCA is upper compatible with the existing qualitative spatial reasoning methods.

This paper is organized as follows. In section 2, we present the formal definition for PLCA. In section 3, we introduce an attributed PLCA and discuss its reasoning. In section 4, we compare our approach with other works. And finally, in section 5, we show the conclusion.

2 PLCA expression

2.1 Definition of Classes

PLCA has four basic components: *points*, *lines*, *circuits* and *areas*.

Point is defined as a primitive class.

Line is defined as a class that satisfies the following condition: for an arbitrary instance l of Line, l.points is an array $[p_1, p_2]$ where $p_1, p_2 \in Point$. A line has an inherent orientation. l^- means the opposite orientation $[p_2, p_1]$ if $l = [p_1, p_2]$. l^* denotes either l or l^- . Intuitively, a line is the edge connecting two (not always different) points. No two lines are allowed to cross.

Circuit is defined as a class that satisfies the following condition: for an arbitrary instance c of *Circuit*, *c.lines* is an array $[l_1^*, \ldots, l_n^*]$ where $l_1^*, \ldots, l_n^* \in Line(n \ge 1)$, $l_i^*.points = [p_i, p_{i+1}](1 \le i \le n)$ and $p_{n+1} = p_1$. Intuitively, a circuit is the closed circuit that is constructed by connecting n lines.

For $c_1, c_2 \in Circuit$, we introduce two new predicates lcand pc to denote that two circuits share line(s) and point(s), respectively. $lc(c_1, c_2)$ is true iff there exists $l \in Line$ such that $(l \in c_1.lines) \land (l \in c_2.lines)$. $pc(c_1, c_2)$ is true iff there exists $p \in Point$ such that $(p \in l_1.points) \land (p \in l_2.points) \land (l_1 \in c_1.lines) \land (l_2 \in c_2.lines)$. A circuit is a border between an area and its adjacent areas viewed from the side of that area. An array of lines itself can represent the concept of circuit, but we use the class Circuit for convenience.

Area is defined as a class that satisfies the following condition: for an arbitrary instance a of Area, a.circuits is a set $\{c_1, \ldots, c_n\}$ where $c_1, \ldots, c_n \in Circuit(n \ge 1)$, and $\forall c_i, c_j \in a.circuits; (i \ne j) \rightarrow (\neg pc(c_i, c_j) \land \neg lc(c_i, c_j))$. Intuitively, an area is a connected region which consists of exactly one piece¹. No two areas are allowed to cross. The final condition means that any pair of circuits which belong to the same area cannot share a point nor a line.

The *PLCA expression* \mathcal{E} is defined as a class that satisfies the following condition: for an arbitrary instance e of \mathcal{E} ,

- 1. *e.points* is a subset of *Point*
- 2. *e.lines* is a subset of *Line*
- 3. *e.circuits* is a subset of *Circuit*
- 4. e.areas is a subset of Area
- 5. *e.outermost* is the element of of *e.circuits*.

We assume that there exists a circuit in the outermost side of the figures that is called *outermost*. Intuitively, the entire space is encircled by the *outermost* circuit and divided into the number of areas.

Example 1.

We present the PLCA expression e for the figure shown in Figure 1(a). We assume that c_{outer} is the outermost circuit in the figure. There are six areas, eight circuits, nine lines and five points. Figure 2 shows the names of the objects. The total PLCA expression e is given as follows.

 $e.points = \{p_0, p_1, p_2, p_3, p_4\}$ $e.lines = \{l_0, l_1, l_2, l_3, l_4, l_5, l_6, l_7, l_8\}$ $e.circuits = \{c_{outer}, c_0, c_1, c_2, c_3, c_4, c_5, c_6\}$ $e.areas = \{a_0, a_1, a_2, a_3, a_4, a_5\}$ $e.outermost = c_{outer}$ $l_0.points = [p_0, p_0]$ $l_1.points = [p_4, p_1]$ $l_2.points = [p_1, p_2]$ $l_3.points = [p_2, p_3]$ $l_4.points = [p_3, p_4]$ $l_5.points = [p_1, p_4]$ $l_6.points = [p_2, p_1]$ $l_7.points = [p_3, p_2]$ $l_8.points = [p_4, p_3]$ $c_{outer}.lines = [l_0]$ $c_0.lines = [l_0^-]$ $c_1.lines = [l_1^-, l_5^-]$ $\begin{array}{l} c_2.lines = [l_2^-, l_6^-] \\ c_3.lines = [l_1, l_2, l_3, l_4] \end{array}$ $c_4.lines = [l_4^-, l_8^-]$ $c_5.lines = [l_3^-, l_7^-]$ $c_6.lines = [l_5, l_8, l_7, l_6]$ $a_0.circuits = \{c_6, c_0\}$ $a_1.circuits = \{c_1\}$ $a_2.circuits = \{c_2\}$ $a_3.circuits = \{c_3\}$ $a_4.circuits = \{c_4\}$ $a_5.circuits = \{c_5\}$

2.2 Consistency of PLCA

If the PLCA expression e satisfies all the following constraints, e is said to be *consistent*.

- **1. constraint on P-L** For any $p \in Point$ there exists at least one line l such that $p \in l.points$.
- **2. constraint on L-C** For any $l \in Line$, there exist exactly two distinct circuits c_1, c_2 such that $l \in c_1.lines, l^- \in c_2.lines$.
- **3. constraint on C-A** For any $c \in Circuit$, there exists exactly one area *a* such that $c \in a.circuits$.

These constraints show that a PLCA expression does not allow the isolated point nor the isolated line.

2.3 **Two-Dimensional Realizability**

For a consistent PLCA expression, there exists a figure in n-dimensional space. In general, it may not be two-dimensional.

Example 2.

¹We use the term *area* instead of *region*, since *area* as used in this paper is a different entity from the *region* generally used in qualitative spatial reasoning.



Figure 2: Names of the objects



Figure 3: PLCA-expression unrealizable in two dimensional space

We give a part of a consistent PLCA expression e as follows.

$$e.lines = \{l_0, l_1, l_2\}$$

$$e.circuits = \{c_{outer}, c_0, c_1, c_2, c_3, c_4\}$$

$$e.areas = \{a_0, a_1\}$$

$$c_{outer}.lines = [l_0]$$

$$c_0.lines = [l_0]$$

$$c_1.lines = [l_1]$$

$$c_2.lines = [l_2]$$

$$c_3.lines = [l_1^-]$$

$$c_4.lines = [l_2^-]$$

$$a_0.circuits = \{c_0, c_3, c_4\}$$

$$a_1.circuits = \{c_1, c_2\}$$

There are only two areas a_0 and a_1 , but a_1 with two circuits should consist of one piece. The corresponding figure is shown in Figure 3. We cannot realize the expression in two-dimensional space.

We introduce concepts of *inner* and *outer circuits* to determine whether a corresponding figure exists in two-dimensional space for a given consistent PLCA expression.

- 1. The outermost is an outer circuit.
- If a line l(l⁻) belongs to an outer circuit, then the circuit to which l⁻(l) belongs to is an inner circuit, respectively.
- 3. There is a unique *inner circuit* that belongs to one area, and the other ones are *outer circuits*.
- In Example 2, circuits c_1 and c_2 , which are both included

in area a_1 , are inner circuits. Therefore, this expression does not satisfy the third condition.

The following proposition holds.

Proposition 2.1 For a consistent PLCA expression, there exists a figure in two-dimensional space iff each circuit can be identified as an inner or outer circuit.

2.4 The Equivalence of PLCA Expressions

For consistent PLCA expressions e_1, e_2 , if there exists a bijective function f such that satisfies the following conditions, then e_1 and e_2 are said to be *PLCA-equivalent*.

For $\forall p \in e_1.points$, $f(p) \in e_2.points$ For $\forall l \in e_1.lines$, $f(l) \in e_2.lines$ For $\forall c \in e_1.circuits$, $f(c) \in e_2.circuits$ For $\forall a \in e_1.areas$, $f(a) \in e_2.areas$ For $\forall l \in e_1.lines$, f(l.points) = f(l).pointsFor $\forall c \in e_1.circuits$, f(c.lines) = f(c).linesFor $\forall a \in e_1.areas$, f(a.circuits) = f(a).circuits

The following proposition holds.

Proposition 2.2 Let F_1 and F_2 be figures in two-dimensional space, e_1 and e_2 are the corresponding PLCA expressions, respectively. If F_1 and F_2 are homeomorphic, then e_1 and e_2 are PLCA-equivalent, and vice versa.

3 Reasoning on Attributed PLCA

3.1 Attributed PLCA

In most applications such as GIS, figure or image data include not only the spatial information, such as shape, position and size, but also semantic information, such as what the object stands for or which property holds on the region.

Since PLCA itself provides only spatial information, we add regional attributes to each object so that semantic information can also be represented. We define a *region* as a non-empty set of areas that have the same attribute.

If an area belongs to some region, then all the circuits, lines and points that constitute the area also belong to the region. Note that a region may consist of multiple pieces.

Let Region be a class of region and RA(r), RC(r), RL(r), RP(r) denote the set of areas, circuits, lines and points which belong to region $r \in Region$,

respectively. In general, for a consistent two-dimensional realizable PLCA expression e, the attributed PLCA expression can be constructed as follows. For each r, first, we assign a subset of *e.areas* to r, and then, we define RA(r), RC(r), RL(r), RP(r).

$$\begin{aligned} RA(r) &= \{a \mid a \in r.areas\} \\ RC(r) &= \{c \mid c \in a.circuits \land a \in r.areas\} \\ RL(r) &= \left\{ l \mid \begin{array}{c} l \in c.lines \land c \in a.circuits \\ \land a \in r.areas \end{array} \right\} \\ RP(r) &= \left\{ p \mid \begin{array}{c} p \in l.lines \land l \in c.lines \\ \land c \in a.circuits \\ \land a \in r.areas \end{array} \right\} \end{aligned}$$

The PLCA expression in which each object is attached to an attribute in this way is called an *attributed PLCA expression*.

3.2 Topological Reasoning

If an attributed PLCA expression is a data form in the spatial database, the database can give the topological information about the regions. We show the definitions of several queries to the database.

When two regions are connected, we can decide how these are connected:

$$isPointConnected(r_1, r_2)$$

$$:= RP(r_1) \cap RP(r_2) \neq \phi$$

$$\land RL(r_1) \cap RL(r_2) = \phi$$

$$\land RA(r_1) \cap RA(r_2) = \phi$$

$$isLineConnected(r_1, r_2)$$

$$:= RL(r_1) \cap RL(r_2) \neq \phi$$

$$\land RA(r_1) \cap RA(r_2) = \phi$$

We can obtain the segments that are point-connected and lineconnected in the borders of two regions:

$$getPointConnects(r_1, r_2)$$

$$:= \begin{cases} d & d \in RP(r_1) \cap RP(r_2) \land \\ \neg \exists l(d \in l.points \land \\ l \in RL(r_1) \cap RL(r_2)) \end{cases}$$

$$getLineConnects(r_1, r_2)$$

$$:= \{l \mid l \in RL(r_1) \cap RL(r_2) \}$$

Furthermore, we can determine whether a region consists of one piece, and obtain the number of the connected parts if it consists of several disconnected pieces:

$$\begin{split} & isOnePiece(r) \\ & := \forall a_i \exists a_j \in RA(r); lc(a_i, a_j) \lor pc(a_i, a_j) \\ & getNumberOfParts(r) \\ & := k \quad where \; [\; r = \cup_i^k r_i \;] \land [\; \forall r_i; \; isOnePiece(r_i) \;] \land \\ & \; [\; \forall r_i \forall r_j; \neg isPointConnected(r_i, r_j) \land \\ & \; \neg isLineConnected(r_i, r_j) \;] \end{split}$$

Example 3.

If we make the following assignment on Example 1 and make regions r_1, r_2 and r_3 (Figure 4), then we can examine the topological relations between the regions.

$$r_1.areas = \{a_1, a_5\}$$

 $r_2.areas = \{a_2, a_4\}$
 $r_3.areas = \{a_3\}$



Figure 4: Topological reasoning



Figure 5: Attributed PLCA for fire simulation

For r_1 and r_2 , $isPointConnected(r_1, r_2)$ holds and $getPointConnects(r_1, r_2) = \{p_1, p_2, p_3, p_4\}$ is obtained. It means that r_1 and r_2 are connected with four points. Similarly, we can get the information such that r_1 and r_3 are connected with two lines, and that r_2 and r_3 are connected with two lines.

3.3 Semantic Reasoning

Example 4.

Let F and B denote the attributes of fire burning and the barrier, respectively. Assume that we add semantics to Example 1 so that $\{a_1, a_3, a_5\}$ shows the predicted region r_F of burning at some instant, and $\{a_2, a_3, a_4\}$ shows the region r_B of the barrier. (Figure 5(a)).

$$r_F.areas = \{a_1, a_3, a_5\}$$

 $r_B.areas = \{a_2, a_3, a_4\}$

 $r_{F \wedge B}$ shows the region with both attributes F and B, and $r_{\neg B}$ shows the region without attribute B. Then these regions are:

$$r_{F \land \neg B}.areas = \{a_1, a_5\}$$
$$r_{F \land B}.areas = \{a_3\}$$

By the queries to the database, we find that $isOnePiece(r_{F \wedge \neg B})$ does not hold and that $getNumberOfParts(r_{F \wedge \neg B})$ = 2.It means that $r_{F \wedge \neg B}$ consists of two parts. Moreover, since $lc(a_1, a_3)$ and $lc(a_3, a_5)$ hold, we can deduce that it is a_3 that breaks $r_{F \wedge \neg B}$ into two disconnected parts.

Example 5.



Figure 6: The basic relations of RCC-8

Assume that we make the following assignment for the example shown in Figure 5(b).

$$r_F.areas = \{a_1, a_3\}$$
$$r_B.areas = \{a_2, a_3\}$$

Then, $r_{F \wedge \neg B}$.areas = {a₁} and isOnePieace($r_{F \wedge \neg B}$) holds. It means that $r_{F \wedge \neg B}$ consists of one piece.

Both Examples 4 and 5 are the attributed PLCA expressions for the example of the simulation of a fire burning shown in Section 1. Comparing these examples, the predicted region of burning is divided by the barrier in case of Example 4, that is, if the barrier is constructed in the position shown in Figure 5(a), the fire can be stopped from spreading to the right, while in case of Example 5, it cannot. Conversely, we can determine the position of the barrier by checking whether the target region consists of one piece or not. This type of reasoning can be applied to the decision of the position for a new construction.

4 Comparison with Other Works

4.1 RCC

RCC is a theory that considers a space as a set of regions, paying attention only to their relative positions [Randell *et al.*, 1992]. Figure 6 shows the basic relations of RCC-8, a variant of RCC. These relations are pairwise disjoint and jointly exhaustive.

Region in an attributed PLCA is the same concept as the 'region' in RCC. Table 4.1 shows the relationship between the attributed PLCA and RCC. In this table, Y/N mean that there exists an element in the corresponding part. Although the number of all the possible combination is 2^4 , the case that does not appear in this table is an inconsistent PLCA expression. This table shows that there exists a mapping from PLCA to RCC, and that PLCA provides mereological reasoning which are performed in RCC.

Several studies on the extension of RCC have been done. Gotts extended pure RCC to express not only mereology but also topological shape [Gotts, 1994]. Borgo et al proposed a system which pays attention to the connection of regions [Borgo *et al.*, 1998] and Donnelly introduced the concept of *coincidence* into the relation *overlapping* [Donnelly, 2003]. Since these theories invoke new predicates for distinguishing figures at the lower level, the resulting system is very complicated and hard to implement. Simultaneously, various axioms are introduced, however, the completeness of the system is not proved.

PLCA is easy to implement and extend since it is not based on axiomatic systems.

4.2 9-intersection Model

Egenhofer et al. proposed the framework in which regions, lines and points are taken as basic objects for describing topological relations [Egenhofer and Herring, 1990]. They divided each object into three parts of inner, border and outer, and distinguished the topological relationships between two objects by representing them as a 3*3 matrix called the 9intersection, the elements of which show whether the intersection of each part is empty or not. Line-line relation and line-region relation are also expressed in this form. The 9intersection model has almost the same expressive power as PLCA but for the orientations of lines.

They also proposed a model that incorporates the concept of orientations of lines [Egenhofer and Herring, 1990; Egenhofer and Franzosa, 1995]. Other studies have extended these models to handle regions with holes [Egenhofer et at., 1994], splitting ratios [Nedas and Egenhofer, 2004] and so on. In these models, overlapping of primitive objects is allowed. Since there are many variations of overlapping patterns, different data structure is required in addition to 9-intersection to distinguish these patterns in the extended models. On the other hand, in PLCA, primitive objects without overlapping have the sufficient information on connection patterns, and new data structure can be defined in a bottom-up manner.

Another difference between the approaches based on the 9-intersection model and PLCA is the representation of the relations of the objects. In their approaches, the entire figure is represented in the form of a set of binary relations. If the figure data contain n objects, we have to assert ${}_{n}C_{2}$ relations, to avoid nondeterminacy. On the other hand, in PLCA, the entire figure is represented in a form in which all the objects are related.

5 Conclusion

We have proposed a new framework for qualitative spatial reasoning called PLCA, which is focused on the connection pattern of regions. Since it is based on simple objects, it is easy to implement the system.

In a PLCA expression, the entire figure is represented in a form in which all the objects are related, and any figure in two-dimensional space can be treated in a unified manner. Therefore, it is easy to extend the system.

If regional attributes are added to PLCA and several objects are combined, spatial semantic reasoning [Takahashi, 2003; 2004] can be performed.

We have also compared PLCA with the existing qualitative spatial reasoning methods, RCC and the 9-intersection model, and showed that PLCA is upper compatible with them.

We have implemented basic components of PLCA. In the future, we plan to construct higher level reasoning sys-

$RA(r_1) \cap RA(r_2)$	Ν	Ν	Y	Y	Y	Y	Y	Y
$RP(r_1) \cap RP(r_2)$	Ν	Y	N	Y	Ν	Y	Y	Y
$RA(r_1) \setminus RA(r_2)$	Y	Y	Y	Y	N	Ν	Y	Ν
$RA(r_2) \setminus RA(r_1)$	Y	Y	N	Ν	Y	Y	Y	Ν
RCC-8 relation	DC	EC	NTPP	TPP	NTPPi	TPPi	РО	EQ

Table 1: Attributed PLCA and RCC

tem on this framework, and to extend this framework to n-dimensional spaces.

References

- [Borgo et al., 1998] Stefano Borgo, Nicola Guarino and Claudio Masolo. A pointless theory of space based on strong connection and congruence. Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning (KR96), pages 220– 229. Morgan Kaufmann Publishers, 1996.
- [Cohn *et al.*, 2001] Anthony G. Cohn and Shyamanta M. Hazarika. Qualitative spatial representation and reasoning: an overview. *Fundamental Informaticae*, 46(1-1):1–29, 2001.
- [Donnelly, 2003] M. Donnelly. Layered mereotopology. In Eighteenth International Joint Conference on Artificial Intelligence (IJCAI 2003), pages 1269–1274, 2003.
- [Egenhofer and Herring, 1990] M. Egenhofer and J. Herring. Categorizing binary topological relations between regions, lines and points in geographic databases. *Department of Surveying Engineering, University of Maine*.
- [Egenhofer et at., 1994] M. Egenhofer and E. Clementini and P. Di Felis. Topological relations between regions with holes. *International Journal of Geographical Information Systems*, 8(2):129–144, 1994.
- [Egenhofer and Franzosa, 1995] M. Egenhofer and R. Franzosa. On the equivalence of topological relations. *International Journal of Geographical Information Systems*, 9(2):133–152, 1995.
- [Gotts, 1994] N. M. Gotts. How far can we 'C'? Defining a 'doughnut' using connection alone. In Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning (KR94), pages 246– 257, Morgan Kaufmann Publishers, 1994.
- [Nedas and Egenhofer, 2004] K.Nedas and M.Egenhofer Splitting ratios: Metric details of topological line-line relations. In Seventeenth International FLorida Conference on Artificial Intelligence (FLAIRS04), 2004.
- [Randell et al., 1992] D. Randell, Z. Cui, and A. Cohn. A spatial logic based on regions and connection. In Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR92), pages 165–176, Morgan Kaufmann Publishers, 1992.
- [Renz, 2002] Jochen Renz. Qualitative Spatial Reasoning with Topological Information. LNAI-2293, Springer Verlag, 2002.

- [Stock, 1997] Oliviero Stock, editor. *Spatial and Temporal Reasoning*. Kluwer Academic Press, 1997.
- [Takahashi, 2003] Kazuko Takahashi. Reasoning about propagation of properties over regions. In *Journal of Universal Computer Science*, 9(9):1030–1045, 2003.
- [Takahashi, 2004] Kazuko Takahashi. Reasoning on spatial constraints over regions, In *Florida International Conference on Artificial Intelligence (FLAIRS04)*, pages 801– 806, 2004.