Modal Logics for Qualitative Spatial Reasoning

BRANDON BENNETT, *Division of Artificial Intelligence, School of Computer Studies, University of Leeds, Leeds LS2 9JT, England. E-mail: brandon@scs.leeds.ac.uk*

Abstract

Spatial reasoning is essential for many AI applications. In most existing systems the representation is primarily numerical, so the information that can be handled is limited to precise quantitative data. However, for many purposes the ability to manipulate high-level qualitative spatial information in a flexible way would be extremely useful. Such capabilities can be proveded by logical calculi; and indeed 1st-order theories of certain spatial relations have been given [20]. But computing inferences in 1st-order logic is generally intractable unless special (domain dependent) methods are known.

0-order *modal* logics provide an alternative representation which is more expressive than classical 0-order logic and yet often more amenable to automated deduction than 1st-order formalisms. These calculi are usually interpreted as propositional logics: non-logical constants are taken as denoting propositions. However, they can also be given a nominal interpretation in which the constants stand for some kind of object. I show how 0-order logics can be given a spatial interpretation: constants denote regions and logical operators correspond to operations on regions which are important for characterising spatial situations.

Representing certain spatial concepts requires the introduction of modal operators, interpreted as functions generating regions related in specific ways to those denoted by their arguments. A significant example is the *convex-hull* operator whose value is the smallest convex region containing its argument. I investigate how this this operator can be captured in a multi-modal logic.

Keywords: spatial reasoning, ?????

1 Introduction

Spatial reasoning has a wide variety of potential applications in AI systems (e.g. spatial information plays a crucial role in robotics, geographical information systems, CAD/CAM, and systems used for medical analysis and diagnosis). In most existing computer systems representation and manipulation of spatial data is done *numerically.* Objects and regions are represented by sets of coordinates and information is extracted from this data by means of arithmetic and trigonometrical computations.

Numerical representation may be well suited for some purposes, in particular where the spatial information precisely describes some definite situation and where the output required from the system is itself primarily numerical. However, in many cases, useful spatial information does not describe a unique physical situation but qualitatively characterises a situation as being of a particular type. Extracting information from such data requires logical reasoning about the concepts involved in describing a situation; and hence requires a rigorous (formal) theory of qualitative spatial relationships.

It is not true that spatial reasoning has been neglected by mathematicians. Indeed the fields of geometry and topology are extremely well developed and are of direct relevance to automated reasoning about spatial situations. But the problem with nearly all mathematical theories is that they are too complex to reason with effectively. Topology is built upon a large amount of set theory so any naive reasoning algorithm based on standard formulations of topology will have as its search space virtually all of mathematics. Whilst rather more succinct axiomatisations of elementary geometry exist [23] these are still far too complex to be tackled by existing theorem proving techniques.

From a computational point of view, qualitative theories of spatial relations are relatively undeveloped. Nevertheless some significant work has been done. Randell and Cohn [19] and Randell et al. [20] specify a 1st-order theory of spatial regions based on a primitive relation of connectedness, $\mathsf{C}(x, y)$, together with a number of (quasi-Boolean) functions. Despite containing very few non-logical primitives this theory has been found to be quite expressive: indeed a large number of significant spatial relations can be defined exclusively in terms of the C relation [10]. Egenhofer [6] presents a much more limited framework in which a number of topological relations can be represented. He also shows how some simple inference rules can be used to generate the *composition* of any pair of these relations.[1]

The major problem in developing a useful formalism for reasoning about spatial information (indeed for any domain) is the trade off between expressive power and computational tractability. Whilst Egenhofer's representation does allow for certain inferences to be computed effectively, the scope of the theory is limited. On the other hand, although the formalism presented by Randell et al. [20] is very expressive, since it is presented in 1st-order logic, reasoning in the calculus is extremely difficult (however the use of pre-calculated composition tables for relations definable in the theory does enable certain kinds of inference to be computed efficiently).

The principal aim of this paper is to explore a framework for representing spatial information which is both expressive enough to be useful for solving real problems and is in some sense tractable. The formalisms which I suggest will provide such flexible and yet practical reasoning systems are *multi-modal 0-order logics*. Such calculi are normally regarded as *propositional* logics but as we shall see, a spatial interpretation of expressions in these formalisms can be given in which the non-logical constants refer to spatial *regions* rather than propositions.

It is common in computer science to equate tractability with polynomial-time computability. But to a logician this will probably seem an overly harsh restriction, since proof procedures in nearly all interesting logics are at least exponentially hard. Nevertheless, the formalisms presented here are *decidable* and hence far better to reason with than 1st-order theories (such as that given in [20]). Towards the end of this paper I shall suggest how modal representations could be utilised in effective reasoning systems by restricting the range of relations which can be expressed and customising proof procedures to limited classes of formulae.

The domain which we shall be concerned with is a limited but significant subdomain of spatial reasoning. Initially we shall look at formalisms which can represent topological relations such as those shown in figure 1 as well as Boolean-like relations between combinations of regions (e.g. x is the *sum* of y and z). In the latter half

[1] Given any two relations R_1 and R_2 taken from a set $\mathcal{B}$ of pairwise disjoint and mutually exhaustive relations. The composition, $R_1 \circ R_2$, of R_1 and R_2 is the disjunction/sum of all those relations $R_3 \in \mathcal{B}$ such that given $R_1(x, y)$ and $R_2(y, z)$ it is possible that $R_3(x, z)$.

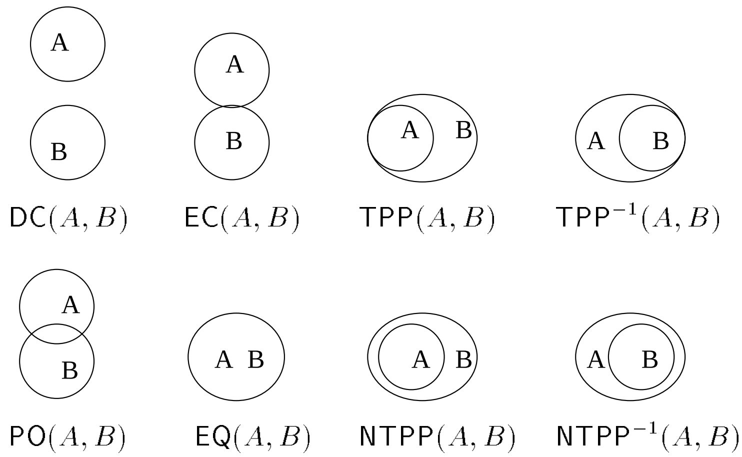

FIG. 1. An exhaustive set of disjoint relations definable in terms of C.

of the paper we shall examine the notion of convexity, which can be used to define various concepts of containment.

I shall explicitly refer to the relations shown in figure 1. These are: Dis-Connection, External Connection, Partial Overlap, Tangential Proper Part, Non-Tangential Proper Part and Equality. The notations used for these relations are given under the diagrams. The part relations, being asymmetric, have inverses denoted, R^{-1}. We shall also use some more general relations:

$$\begin{aligned}
\mathsf{DR}(x,y) &\equiv \mathsf{DC}(x,y) \vee \mathsf{EC}(x,y), \\
\mathsf{PP}(x,y) &\equiv \mathsf{TPP}(x,y) \vee \mathsf{NTPP}(x,y), \\
\mathsf{P}(x,y) &\equiv \mathsf{PP}(x,y) \vee \mathsf{EQ}(x,y) \quad \text{and} \\
\mathsf{TP}(x,y) &\equiv \mathsf{TPP}(x,y) \vee \mathsf{EQ}(x,y).
\end{aligned}$$

2 Models for modal logics

Currently the best known interpretations of modal logics are those in terms of *Kripke semantics* [12]. In Kripke semantics a model consists of a set of possible worlds together with an *accessibility relation* — a binary relation between worlds — associated with each modal operator. Propositions denote sets of possible worlds (the set of worlds in which they are true). A Kripke model, $\mathcal{M}$, is thus a structure $\langle W, R, P \rangle$, where W is a set of worlds, R is the accessibility relation, P is a mapping from natural numbers to subsets of W. P acts as an assignment to a denumerable set, $\{p_0, p_1, \ldots\}$ of propositional constants indexed by the natural numbers.

Such a model determines the truth of each modal formula at each possible world. Classical formulae are interpreted as follows:

- Atomic formulae, p_i are true in exactly the worlds in the set $P(i)$.
- Conjunctions, $\phi \wedge \psi$, are true in worlds where both ϕ and ψ are true.
- Disjunctions, $\phi \vee \psi$, are true in worlds where either ϕ or ψ (or both) is true.

- Negations, $\neg\phi$, are true in worlds where ϕ is not true.

We write $\models^{\mathcal{M}}_{\alpha} \phi$ to mean that formula ϕ is true at world α in model $\mathcal{M}$. A modal operator, $\Box$, is then interpreted as follows: in a model $\mathcal{M} = \langle W, R, P\rangle$

$$\models^{\mathcal{M}}_{\alpha} \Box\phi \quad \text{iff} \quad \models^{\mathcal{M}}_{\beta} \phi \quad \text{for all } \beta \in W \text{ s.t. } R(\alpha, \beta)$$

A vast spectrum of different modal operators can be specified by placing more or less general restrictions on the corresponding accessibility relation (often such restrictions are thought of as defining a logic rather than an operator but this is misleading since the possible worlds semantics allows any number of different operators to be encompassed in a single logical language). Furthermore, Kripke semantics allows one to specify operators whose logic seems to correspond well with intuitive properties of modal concepts employed in natural language. Indeed, a number of logics proposed for natural language modalities, which were originally specified proof theoretically (by axiom schemas intended to capture intuitive properties of modal concepts) can be captured very easily within the Kripke paradigm by quite simple restrictions on the accessibility relation.

Nevertheless, the apparent power and flexibility of Kripke style possible worlds semantics can be misleading. Many researchers in both AI and philosophical logic tend to think of possible worlds semantics as essentially based upon accessibility relations. However, whilst Kripke models may be appropriate for certain types of modal operator, in other cases a quite different structuring of possible worlds may be more natural.

2.1 Minimal and algebraic models

In this section I shall define an *algebraic model* structure, which will be the basis of the spatial logics developed in the rest of the paper. But first I present *minimal model* semantics as described by Chellas [4]. These are actually equivalent to algebraic models although they are more similar in structure to Kripke models; thus are intermediate between relational and algebraic models.

A *minimal model* is a structure $\mathcal{M} = \langle W, N, P\rangle$, where:

- W is a set of possible worlds.
- N maps each world to a set of sets of worlds.
- P maps each natural number (indexing a propositional constant) to a set of possible worlds.

The semantics of the classical connectives are as in the Kripke models but the interpretation of the modal operator $\Box$ is as follows: If α is a world in a minimal model $\mathcal{M} = \langle W, N, P\rangle$ then

$$\models^{\mathcal{M}}_{\alpha} \Box\phi \quad \text{iff} \quad \{\beta \mid \models^{\mathcal{M}}_{\beta} \phi\} \in N(\alpha)$$

The set $\{\beta \mid \models^{\mathcal{M}}_{\beta} \phi\}$ is the set of all worlds at which ϕ is true. This is called the truth set of ϕ. We may also regard this set of worlds as the *denotation* of the formula ϕ. I shall adopt this terminology and write $d(\phi)$ to refer to the denotation of ϕ (for atomic propositions $d(p_i) = P(i)$). Furthermore if formulae are also supposed to denote propositions then propositions become identified with the set of worlds in which they are true. Thus, if our modal operator is, for example, intended to be a *necessity* operator, the function N associates with each world the set of propositions (each of which is identified with a set of worlds) which are necessary at that world.

An *algebraic* model is similar to a minimal model but instead of having a function N mapping worlds to sets of sets of worlds, it has a function O mapping sets of worlds to sets of worlds. An algebraic model $\mathcal{A} = \langle W, O, P\rangle$ defines a modal operator according to the stipulation that:

$$\models^{\mathcal{A}}_{\alpha} \Box\phi \quad \text{iff} \quad \alpha \in O(d(\phi))$$

where $d(\phi)$ is again the denotation/truth-set of ϕ. This can be specified more succinctly simply as

$$d(\Box\phi) = O(d(\phi))$$

Each such model can be associated with a *modal algebra* [14]. This is a structure $\langle A, \cap, -, \mathbf{n}\rangle$, where $\langle A, \cap, -\rangle$ is a Boolean algebra and $\mathbf{n}$ is an extra (modal) operator.[2] An algebraic model $\langle W, O, P\rangle$ corresponds to an assignment of elements of the modal algebra $\langle 2^W, \cap, -, O\rangle$ to the propositions in P — the elements of the algebra being sets of possible worlds.

Minimal and algebraic models are equivalent in that given a minimal model we can straightforwardly construct an algebraic model in which the same formulae are true at each world and vice versa. From the two definitions of $\Box\phi$ we see that if a minimal and algebraic model agree on the denotations of all propositions at all worlds then $d(\phi) \in N(\alpha)$ iff $\alpha \in O(d(\phi))$. Therefore, $N(\alpha) = \{S \mid S \subseteq W \wedge \alpha \in O(S)\}$ and $O(S) = \{\alpha \mid \alpha \in W \wedge S \in N(\alpha)\}$. However, algebraic models are more uniform than minimal models because the semantics of modal operators are specified in essentially the same way as the classical connectives. They also seem to be more natural from the point of view of spatial interpretations.

Algebraic semantics is actually the oldest formal interpretation for modal logics: an algebraic interpretation of $S4$ was given by Tarski [24]; but, since Kripke's results [12], relational semantics has been given far more attention. The relationship between algebraic semantics and Kripke models was first studied by Lemmon [14, 15], who introduced the term 'modal algebra' (however, a theory of 'Boolean algebras with additional operators' had already been given by Jónsson and Tarski [11] and modal algebras are essentially a special case of these). More recently, Goldblatt [9] gives a detailed examination of different kinds of semantics for modal logics and the relationships between them.

2.2 Denotation based algebraic models

Since in an algebraic semantics the denotation function is of central importance it will be convenient to use in the sequel an alternative formulation of *model* in which this function occurs as a primary element. Thus, a (denotation based) algebraic model for a logic $\mathcal{L}$ is a structure $\langle \mathcal{U}, \mathcal{C}, d\rangle$ where:

---

[2] A modal algebra $\langle A, \cap, -, \mathbf{n}\rangle$ is normal iff $\mathbf{n}(x \cap y) = \mathbf{n}(x) \cap \mathbf{n}(y)$ and $\mathbf{n}(1) = 1$. Lemmon [14] shows how normal modal algebras correspond to Kripke models. Further results linking these structures are given by Goldblatt [9].

- $\mathcal{U}$ is a (non-empty) set
- $\mathcal{C}$ is a denumerable set of constants — the non-logical constants of $\mathcal{L}$
- d is a denotation function mapping (well-formed) formulae of $\mathcal{L}$ to subsets of $\mathcal{U}$.

d maps atomic formulae (i.e. elements of $\mathcal{C}$) to subsets of $\mathcal{U}$; and the denotations of complex formulae are determined by the denotations of the atomic formulae by means of recursive definitions of the logical connectives in $\mathcal{L}$. Formally

$$d(\Omega(\phi_1, \ldots, \phi_n)) = f_\Omega(d(\phi_1), \ldots, d(\phi_n))$$

where Ω is an n-ary connective in the language $\mathcal{L}$ and f_Ω maps n-tuples of subsets of $\mathcal{U}$ to subsets of $\mathcal{U}$. Characterising the semantics of a logical language $\mathcal{L}$ consists in specifying (meta-mathematically) permissible values of d for atomic propositions and the functions f_Ω corresponding to each connective Ω in $\mathcal{L}$.

3 Spatial interpretation of 0-order calculi

By far the best known interpretations of 0-order calculi are as propositional logics: the non-logical constants are regarded as denoting propositions and the connectives as operating on their (propositional) arguments to form more complex propositions. Within such a conception, the classical connectives are interpreted as expressing truth-functional combinations of their arguments, whilst modal operators are taken as asserting more subtle (non-truth functional) properties of their arguments. Many kinds of propositional modality have been studied: alethic modalities (necessity, possibility, contingency); propositional attitudes (knowledge, belief, certainty, etc.); deontic modalities (obligation, permission).

However, taking non-logical constants as denoting propositions is not the only way that 0-order calculi can be interpreted. Bennett [2] points out and makes use of non-propositional interpretations of both the standard classical 0-order calculus and also the 0-order intuitionistic calculus. Under these interpretations, the non-logical constants denote regions and the connectives correspond to operations forming new regions from their arguments. In fact this interpretation is compatible with many well known model theoretic accounts of 0-Order calculi, in which propositions are taken as denoting sets. These sets are often thought of as sets of possible worlds in which a proposition is true but they can also be regarded as sets of points (or atoms) making up a spatial region.

3.1 The classical calculus

In the case of the 0-order classical calculus (henceforth $\mathcal{C}_0$) such a semantics can be formally characterised as follows: a model for the logic $\mathcal{C}_0$ is a structure, $\langle \mathcal{U}, \mathcal{P}, d \rangle$, where $\mathcal{U}$ is a non-empty set, $\mathcal{P}$ is a denumerably infinite set of constants, and d is a denotation function which assigns to each constant in $\mathcal{P}$ a subset of $\mathcal{U}$. The domain of d is extended to all $\mathcal{C}_0$ formulae formed from the constants by stipulating that:

1. $d(\neg P) = \overline{d(P)}$
2. $d(P \wedge Q) = d(P) \cap d(Q)$
3. $d(P \vee Q) = d(P) \cup d(Q)$

where for any set S, $\overline{S}$ is the set of all elements of $\mathcal{U}$ which are not elements of S. Under this interpretation it can be shown that all tautologies denote the universe, $\mathcal{U}$, whatever the assignment of sets to the non-logical constants.

3.2 The intuitionistic calculus

Tarski [21] gives similar but slightly more complex semantics for the 0-order intuitionistic logic (henceforth $\mathcal{I}_0$). In addition to the usual set operators of union, intersection and complement, this requires an additional *interior* function, i. This is constrained to obey the axioms (see e.g. [13]) of an interior function, as employed in point-set topology; so the semantics can be seen as a *topological* interpretation of intuitionistic logic. A model for $\mathcal{I}_0$ is then a structure $\langle \mathcal{U}, i, \mathcal{P}, d \rangle$ where d now assigns to each constant an *open* subset of U (a set X such that $i(X) = X$) and its domain is extended to all $\mathcal{I}_0$ formulae as follows:

1. $d(\sim P) = i(\ \overline{d(P)}\ )$
2. $d(P \wedge Q) = d(P) \cap d(Q)$
3. $d(P \vee Q) = d(P) \cup d(Q)$
4. $d(P \Rightarrow Q) = i(\ \overline{d(P)} \cup d(Q)\ )$

This denotation function is such that all intuitionistic theorems denote $\mathcal{U}$ under any assignment of *open* sets to non-logical constants. Note that I use distinct symbols for negation and implication in the classical and intuitionistic languages (for classical implication I shall write ' $\rightarrow$ ') but for conjunction and disjunction I use the same symbols, since their interpretations are the same in both systems.

3.3 Semantic correspondence to proof theory

For a semantic interpretation to be faithful to a logic, the property of *theoremhood* and the relation of *entailment* must be definable in the semantics and these concepts must be shown to coincide with the corresponding proof-theoretic concepts. In the context of the algebraic set semantics considered here this will amount to establishing the following correspondences:

CT	$\vdash \phi$	iff	$\models d(\phi) = \mathcal{U}$
CE[3]	$\psi_1, \ldots, \psi_n \vdash \phi$	iff	$d(\psi_1) = \mathcal{U}, \ldots, d(\psi_n) = \mathcal{U} \ \models\ d(\phi) = \mathcal{U}$

where ' $\vdash$ ' is the derivability relation of the logic and '$\models$' is the entailment relation between properties of the algebraic model structures. **CT** and **CE** hold for both $\mathcal{C}_0$ and $\mathcal{I}_0$ under the interpretations given above (see [2] for further details).

These correspondences mean that the logic can straightforwardly be used to determine entailments between constraints on possible models which are specified in terms of equations of the form $\tau = \mathcal{U}$ where τ is a set-term corresponding to some formula, ϕ, in the logic (i.e. $\tau = d(\phi)$). For example,

$$\overline{d(A) \cap d(B)} = \mathcal{U},\ \overline{d(C)} \cup d(A) = \mathcal{U} \quad \models\ \overline{d(C) \cap d(B)} = \mathcal{U}$$

because $\neg(A \wedge B), C \rightarrow A \ \vdash\ \neg(C \wedge B)$ in $\mathcal{C}_0$.

3.4 Representing spatial relationships

In the languages $\mathcal{C}_0^+$ and $\mathcal{I}_0^+$ [2] constraints on the relative extensions of spatial regions are expressed by 0-order formulae. Each formula (of $\mathcal{C}_0$ or $\mathcal{I}_0$) corresponds (in accordance with the semantics given above) to a set-term in which constants denoting regions (sets of points) are combined by various set operators (the Boolean operators and, in $\mathcal{I}_0$, also the interior operator). Formulae are interpreted as describing spatial situations by placing constraints on possible values of the constants. Each formula is given either a positive or a negative interpretation depending upon whether it is a *model* or an *entailment* constraint (the reason for this terminology will be made clear below). In the former case the interpretation is that, in the situation being described, the denotation of the set-term equals the entire universe.

Table 1 shows how four spatial relations can be characterised with model constraints stated in terms of the classical propositional calculus.

TABLE 1. Definitions of Four Topological Relations in $\mathcal{C}_0$

Relation	*Description*	*Set Equation*	*Model Constraint*
$\mathsf{DR}(X,Y)$	X and Y are discrete	$\overline{X \cap Y} = \mathcal{U}$	$\neg(X \wedge Y)$
$\mathsf{P}(X,Y)$	X is part of Y	$\overline{X} \cup Y = \mathcal{U}$	$X \rightarrow Y$
$\mathsf{P}^{-1}(X,Y)$	Y is part of X	$X \cup \overline{Y} = \mathcal{U}$	$Y \rightarrow X$
$\mathsf{EQ}(X,Y)$	X and Y are equal	$(\overline{X} \cup Y) \cap (X \cup \overline{Y}) = \mathcal{U}$	$X \leftrightarrow Y$

In terms of this representation, we can see that the example given at the end of the last section corresponds to the inference:

$$\mathsf{DR}(A,B) \wedge \mathsf{P}(C,A) \ \vdash \ \mathsf{DR}(C,B)$$

3.5 Entailment constraints

A formula can also be given a negative interpretation by specifying it as an entailment constraint rather than a model constraint. The meaning of an entailment constraint is that the corresponding set-term does *not* denote the universe. Such negative constraints are essential for describing certain spatial situations in terms of the logics $\mathcal{C}_0^+$ and $\mathcal{I}_0^+$. For example if we want to say that one region, A is a *proper part* of another region, B, we specify that $\overline{A} \cup B = \mathcal{U}$ but $\overline{B} \cup A \neq \mathcal{U}$. Thus (again using $\mathcal{C}_0$) we would have $\neg A \vee B$ as a model constraint and $\neg B \vee A$ as an entailment constraint. (Note that the latter inequality could *not* be expressed by the model constraint $\neg(\neg B \vee A)$ because this corresponds to the constraint $B \cap \overline{A} = \mathcal{U}$ which is much stronger than what is required.)

An important use of entailment constraints is to ensure that regions involved in a situation description are *non-null*. If null-regions are allowed they have properties which may seem counter-intuitive (for example the null region is both part of and disconnected from any other region) and many useful and apparently sound inferences may not hold if it is allowed that one of the regions involved may be null. The requirement that a region is non-null is expressed by the inequality $\overline{X} \neq \mathcal{U}$, which corresponds to the entailment constraint $\neg X$ (in $\mathcal{C}_0$).

Since, specification of a spatial situation will generally require both positive and negative constraints, a situation description is represented by a pair of sets of formulae, $\langle \mathcal{M}, \mathcal{E} \rangle$, where $\mathcal{M}$ is a set of model constraints and $\mathcal{E}$ is a set of entailment constraints. Thus 'A is a proper part of B' (where A and B must be non-null) is represented by the expression

$$\langle \{\neg A \vee B\},\ \{\neg B \vee A,\ \neg A,\ \neg B\} \rangle$$

If $\mathcal{L}$ is a 0-order logic then the language obtained by extending the representation in this way will be called $\mathcal{L}^+$.

3.6 Spatial interpretation of the intuitionistic calculus

The intuitionistic calculus $\mathcal{I}_0$ is particularly expressive for describing spatial relations. This is because the interior function in the semantics enables one to use $\mathcal{I}_0$ formulae to specify constraints which distinguish between two regions being connected from their overlapping. This distinction is made possible by the following interpretation of the notions of *region*, *overlap* and *connection*:

- A *region* is identified with an open set of points. (So regions are denoted by propositional letters in the $\mathcal{I}_0$ representation.)
- Regions *overlap* if they share at least one point.
- Regions are *connected* if their *closures* share at least one point.

This interpretation is in accord with that suggested for the RCC theory in [20]. There is also a dual interpretation (which will be used in section 5.1) under which regions are identified with *closed* sets, which are connected if they share a point and overlap if they share an *interior* point. Table 2 shows how the set of eight relations shown in figure 1 can be represented by sets of model and entailment constraints specified by means of $\mathcal{I}_0$ formulae interpreted in accordance with the semantics given in section 3.2.

TABLE 2. Some RCC Relations Defined in $\mathcal{I}_0^+$ (including the 8 relation basis)

Relation	*Model Constraint*	*Entailment Constraints*
$\mathsf{DC}(X,Y)$	$\sim X \vee \sim Y$	$\sim X,\ \sim Y$
$\mathsf{EC}(X,Y)$	$\sim(X \wedge Y)$	$\sim X \vee \sim Y, \sim X,\ \sim Y$
$\mathsf{PO}(X,Y)$	—	$\sim(X \wedge Y),\ X \Rightarrow Y,\ Y \Rightarrow X, \sim X, \sim Y$
$\mathsf{TPP}(X,Y)$	$X \Rightarrow Y$	$\sim X \vee Y,\ Y \Rightarrow X, \sim X, \sim Y$
$\mathsf{TPP}^{-1}(X,Y)$	$Y \Rightarrow X$	$\sim Y \vee X,\ X \Rightarrow Y, \sim X, \sim Y$
$\mathsf{NTPP}(X,Y)$	$\sim X \vee Y$	$Y \Rightarrow X, \sim X, \sim Y$
$\mathsf{NTPP}^{-1}(X,Y)$	$\sim Y \vee X$	$X \Rightarrow Y, \sim X, \sim Y$
$\mathsf{EQ}(X,Y)$	$X \Leftrightarrow Y$	$\sim X, \sim Y$
$\mathsf{C}(X,Y)$	—	$\sim X \vee \sim Y, \sim X, \sim Y$
$\mathsf{EQ}(X, \mathsf{sum}(Y,Z))$	$X \Leftrightarrow (Y \vee Z)$	$\sim X, \sim Y, \sim Z$

Let us consider, for example, the representations of the relations $\mathsf{DC}(X,Y)$ and $\mathsf{EC}(X,Y)$. If two regions share no points they cannot overlap (although they may be connected). In such a case the equation $i(\overline{X \cap Y}) = \mathcal{U}$ must hold; this can be represented by the $\mathcal{I}_0$ formula $\sim(X \wedge Y)$. In $\mathcal{I}_0$ (unlike $\mathcal{C}_0$) this formula is not equivalent to $\sim X \vee \sim Y$. The latter corresponds to the set-equation $i(\overline{X}) \cup i(\overline{Y}) = \mathcal{U}$, which says that the union of the exteriors of two regions exhaust the space. If the regions touch at one or more points, then these points of contact will not be in the exterior of either region so this equation will not hold. Hence the second (stronger) formula can be employed as a model constraint to describe the relation $\mathsf{DC}(X, Y)$. If the relation $\mathsf{EC}(X, Y)$ holds then the weaker constraint $\sim(X \wedge Y)$ holds but $\sim X \vee \sim Y$ must not hold, so this stronger formula is an entailment constraint.

The table also shows how the fundamental relation, C, of the RCC theory can be represented as well as the quasi-Boolean function sum (see [20]).

3.7 Extended 0-order calculi (EZOCs)

The interpretation of formulae as constraints can be applied to many logical calculi. And if the denotations of the constants are taken to be spatial regions (sets of atoms or points) then these constraints can be regarded as specifying spatial relationships between the regions. Whether these relationships are useful in describing spatial situations will depend upon whether the semantics of the logic reflects some structure which is relevant to significant features of the situations we wish to describe. If we have such a logic it is very likely that we will sometimes want to specify situations not only in terms of the constraints they satisfy but also in terms of their failure to satisfy certain constraints.

In this section I show how a logic can be extended to a language whose expressions consist of pairs of sets of formulae from the language of the original logic; one of these sets being interpreted as positive constraints and the other as negative constraints. I give definitions of consistency and entailment for the extended language in terms of entailments in the original logic and specify the conditions under which these definitions are correct.

Given a 0-order calculus $\mathcal{L}_0$ whose formulae can be interpreted as algebraic functions of sets, we can define an extension $\mathcal{L}_0^+$, whose expressions are pairs $\langle \mathcal{M}, \mathcal{E} \rangle$, where $\mathcal{M}$ and $\mathcal{E}$ are interpreted respectively as sets of constraints of the forms $\tau = \mathcal{U}$ and $\tau \neq \mathcal{U}$. Thus the expression is consistent unless the following entailment holds:

$$m_1 = \mathcal{U}, \ldots, m_j = \mathcal{U} \models_S e_1 = \mathcal{U} \vee \ldots \vee e_k = \mathcal{U} \qquad \textbf{(E)}$$

For most 0-order calculi it turns out that such an entailment can hold only if one of the disjuncts on the r.h.s. is itself entailed by the equations on the l.h.s.. The property of a theory/logic where disjunctions are only entailed if at least one disjunct is entailed is known as *convexity* [18]. If this convexity is established then one can check the entailment by checking whether any of the entailment constraints is individually entailed by the model constraints; and if the semantics is faithful to the logic $\mathcal{L}_0$ then the correspondence **CE** enables this entailment to be determined by checking the associated entailment in $\mathcal{L}_0$. I shall now give a general condition under which a logic has the required convexity property. This guarantees that the method of using model and entailment constraints presented in [2] can be applied to many other 0-order calculi.

3.8 Models and convexity

The proof of convexity will generally involve showing that, for any entailment of the form **E**, if for each disjunct on the r.h.s. there is counter-model that shows that that equation is not entailed by the l.h.s., there exists a counter-model to the entailment as a whole. I shall call a counter-model for a disjunct a DCM and a counter-model for the whole entailment an ECM. The obvious way to carry out such an existence proof is to show how given DCMs for each of the disjuncts we can somehow construct an ECM. It could be that there are logics for which such a construction is very complex but we shall see that for a large class of well-known logics a very simple construction is possible.

A model for an arbitrary 0-order language $\mathcal{L}_0$ will be a structure

$$\langle \mathcal{U}, \mathcal{F}, \mathcal{P}, d\rangle \qquad \text{where} \quad \mathcal{F} = \langle f_{\Omega_1}, \ldots, f_{\Omega_n}\rangle$$

Each operator Ω_i in the language will be interpreted as the function f_{Ω_i}.[4] These functions are mappings from subsets of $\mathcal{U}$ to subsets of $\mathcal{U}$. (For example, the classical negation is associated with the function from each set X ($\subseteq \mathcal{U}$) to the set $\{y \mid y \in \mathcal{U} \wedge y \notin X\}$.) In order to reflect the intended meanings of the operators, the structure $\langle f_{\Omega_1}, \ldots, f_{\Omega_n}\rangle$ must satisfy certain axioms Φ. I shall write $\Phi(\mathcal{F})$ to mean that $\mathcal{F}$ (a tuple of functions) satisfies these axioms.

To form an ECM from a set of DCMs we can often just gather the DCMs together to form a model which includes the DCMs as 'non-interacting' sub-models. The domain of the ECM will simply be the union of the domains of the DCMs. We will also have to construct new functions for the operators and for the assignment of sets to the atomic constants. The most straightforward way to interpret a function whose values are sets over a union of two domains is to simply evaluate it as the union of its denotations in each of the component domains. Given $f_1 : 2^{\mathcal{U}_1} \rightarrow 2^{\mathcal{U}_1}$ and $f_2 : 2^{\mathcal{U}_2} \rightarrow 2^{\mathcal{U}_2}$. I define $(f_1 \circledast f_2) : 2^{(\mathcal{U}_1 \cup \mathcal{U}_2)} \rightarrow 2^{(\mathcal{U}_1 \cup \mathcal{U}_2)}$ to be the function such that $(f_1 \circledast f_2)(X) = f_1(X_1) \cup f_2(X_2)$, where $X_1 = \{x \mid x \in X \wedge x \in \mathcal{U}_1\}$ and $X_2 = \{x \mid x \in X \wedge x \in \mathcal{U}_2\}$. I write $\langle \mathcal{F}_1 \circledast F_2\rangle$ to denote the n-tuple resulting from applying '$\circledast$' to corresponding pairs of functions in the n-tuples $\mathcal{F}_1$ and $\mathcal{F}_2$.

Using this notation we can define a *disjoint combination* of two models as follows:

$$(\langle \mathcal{U}_1, \mathcal{F}_1, \mathcal{P}, d_1\rangle \circledast \langle \mathcal{U}_2, \mathcal{F}_2, \mathcal{P}, d_1\rangle) = \langle \mathcal{U}_1 \cup \mathcal{U}_2, \mathcal{F}_1 \circledast \mathcal{F}_2, \mathcal{P}, d \circledast d\rangle$$

By successive application of this operation we can combine any finite number of models. The key feature of these constructed models is that the denotation of any term with respect to the combined model is simply the union of its denotation in the component models. This means in particular that if $\tau = \mathcal{U}$ in all component models then $\tau = \mathcal{U}$ in the combined model and, if $\tau \neq \mathcal{U}$ in any component, $\tau \neq \mathcal{U}$ in the combination. Consequently it is clear that combining DCMs for an entailment of the form **E** will produce an ECM for this entailment.

However, for this construction to be permissible, we must ensure that in joining models together we get an admissible model for the logic — i.e. a model that satisfies

[4] This characterisation of the semantics of operators is quite general because even if the operators are not normally interpreted directly with respect to an axiomatised set function the semantics can easily be recast in this form. For example in section 3.2 the only function in the model structures is i and the interpretations of '$\sim$' and '$\Rightarrow$' are defined indirectly in terms of further functions of i; but we can easily axiomatise $f_\sim$ by $f_\sim(x) = i(\overline{x})$ together with the usual axioms for i and similarly for the other connectives. An $\mathcal{I}_0$ model will then be a structure $\langle \mathcal{U}, \langle f_\wedge, f_\vee, f_\sim, f_\Rightarrow\rangle, \mathcal{P}, d\rangle$.

the axioms Φ. Thus if we are to use this method of model combination to show the convexity of the model theory of a logic with respect to entailments of the form E we need to show that:

$$\text{If} \quad \Phi(\mathcal{F}_1) \quad \text{and} \quad \Phi(\mathcal{F}_2) \quad \text{then} \quad \Phi(\langle \mathcal{F}_1 \circledast F_2 \rangle) \qquad (\mathbf{P})$$

In practice it seems that this property **P** holds for most well-known logics. The case of classical logic is trivial because all its connectives are interpreted by pure set functions, which always distribute over sub-domains in the desired way. For logics interpreted with respect to an interior function, **P** can be seen to hold as long as the space is allowed to contain disconnected subspaces. **P** also holds for modal logics with a Kripke semantics in which the accessibility relation can be disconnected. The common feature of the all these logics is that their models may contain essentially independent sub-models, such that denotations of operators are unions of their denotation in each sub model.[5]

3.9 An extended 0-order reasoning algorithm (EZORA)

If the operators of $\mathcal{L}_0$ satisfy the property **P** then consistency of the expressions of $\mathcal{L}_0^+$ can be defined in terms of the entailment relation, $\models_{\mathcal{L}_0}$, of $\mathcal{L}_0$ as follows:

$\langle \mathcal{M}, \mathcal{E} \rangle$ is consistent iff there is no formula $p \in \mathcal{E}$ such that $\mathcal{M} \models_{\mathcal{L}_0} p$
Otherwise it is inconsistent, in which case we write $\langle \mathcal{M}, \mathcal{E} \rangle \models_{\mathcal{L}_0^+} \bot$

This should make clear why the formulae in $\mathcal{E}$ are called 'entailment constraints'.

We can also define entailment between $\mathcal{L}_0^+$ expressions. $\langle \mathcal{M}, \mathcal{E} \rangle \models_{\mathcal{L}_0^+} \langle \mathcal{M}', \mathcal{E}' \rangle$ holds if the following entailment between set equations holds:

$$\begin{array}{l} m_1 = \mathcal{U} \ \wedge \ \ldots \ \wedge m_h = \mathcal{U} \ \wedge e_1 \neq \mathcal{U} \ \wedge \ \ldots \ \wedge e_i \neq \mathcal{U} \\ \quad \models \\ m'_1 = \mathcal{U} \ \wedge \ \ldots \ \wedge m'_j = \mathcal{U} \ \wedge e'_1 \neq \mathcal{U} \ \wedge \ \ldots \ \wedge e'_k \neq \mathcal{U} \end{array}$$

This will be valid iff each of the conjuncts on the r.h.s. is entailed by the l.h.s. and such a conjunct will be entailed iff its negation is inconsistent with the l.h.s. (negating a constraint is achieved by changing its status from model constraint to entailment constraint or *vice versa*). This leads to the following definition of entailment in $\mathcal{L}_0^+$:

$\langle \mathcal{M}, \mathcal{E} \rangle \models_{\mathcal{L}_0^+} \langle \mathcal{M}', \mathcal{E}' \rangle$ if and only if:
for all $p \in \mathcal{M}'$ we have $\langle \mathcal{M}, \mathcal{E} \cup \{p\} \rangle \models_{\mathcal{L}_0^+} \bot$
and for all $q \in \mathcal{E}'$ we have $\langle \mathcal{M} \cup \{q\}, \mathcal{E} \rangle \models_{\mathcal{L}_0^+} \bot$

The procedure for determining entailments in an EZOC (including consistency checking as a special case) will be be called the Extended Zero-Order Reasoning Algorithm — EZORA for short.

[5] Determination of convexity properties of logics would probably be clearer and more decisive if conducted in terms of algebraic structures and their direct products. Goldblatt [9] has defined the disjoint union of modal frames and related it to products of modal algebras; but his analysis is limited to normal modalities.

4 Eliminating entailment constraints

The procedures for consistency checking and determination of entailments for the constraint calculi described above rely on the use of simple meta-level reasoning. In this section I explain how, by introducing into a 0-order calculus $\mathcal{L}_0$ a new operator, $\Box$, reasoning can be conducted at the object level of the enriched language. This language will be called $\mathcal{L}_0^{\Box}$

In reasoning with an extended 0-order language $\mathcal{L}_0^+$ the meanings of the two types of constraint are handled at the meta-level: Determination of entailments in these languages involves checking a number of different object-level entailments in the logic $\mathcal{L}_0$. A situation description is consistent if and only if none of its entailment constraints is entailed by the set of all model constraints. A natural question regarding these representations is whether it might be possible to extend the calculi involved so that the semantics of the two types of constraint was built directly into the object language. This would mean that computation of entailments could be carried out entirely at the object level.

Since in the original notation a model constraint X is interpreted as $X = \mathcal{U}$ an obvious solution is simply to introduce a new (modal) operator, $\Box$, such that $\Box X$ is interpreted as $X = \mathcal{U}$. An operator of this kind can be very easily characterised in terms of the *algebraic* semantics for modal logics described above. If $d(\alpha)$ is the denotation of a formula α, what we need is simply the operator such that:

- $d(\Box\,\alpha) = \mathcal{U}$ iff $d(\alpha) = \mathcal{U}$.
- $d(\Box\,\alpha) = \emptyset$ iff $d(\alpha) \neq \mathcal{U}$.

This operator is an $S5$ modal operator, since a formula $\Box\,\alpha$ is true in a model iff the formula is true at every point/world in the model. I shall call it a *strong-$S5$* operator because it does not allow the possibility, arising in the slightly weaker Kripke characterisation of $S5$, that there are worlds/points which are not relevant to evaluating the $\Box$ at a particular world (because the set of worlds is partitioned into clusters which are not accessible to each other).[6] With such an operator we can write $\Box\,\alpha$ to assert that α is a model constraint and $\neg\Box\,\alpha$ to assert that α is an entailment constraint.

Let us look at a simple example of spatial reasoning carried out in the classical 0-order calculus supplemented with a strong-$S5$ box operator. We shall consider the transitivity of the *proper part* relation, PP.

$$\mathsf{PP}(a,b) \wedge \mathsf{PP}(b,c) \models \mathsf{PP}(a,c)$$

$\mathsf{PP}(x,y)$ holds when $\overline{x} \cup y = \mathcal{U}$ but $\overline{y} \cup x \neq \mathcal{U}$. We also require that x and y are non-null (see [2]). Non-null constraints on regions can now be expressed as $\neg\Box\neg X$ for any region X (this could be written more succinctly as $\Diamond X$ but this notation would complicate the explanation of the following example). Thus the modal representation of $\mathsf{PP}(A,B)$ is:

$$\Box(A \rightarrow B) \wedge \neg\Box(B \rightarrow A) \wedge \neg\Box\neg A \wedge \neg\Box\neg B$$

[6] In most circumstances the strong and weak $S5$ operators cannot be distinguished at the object level. But the difference may sometimes be significant. For example a multi-modal logic may contain several distinct weak-$S5$ modalities but only one strong-$S5$ operator.

Hence the transitivity of PP corresponds to the entailment:

$$\Box(A \rightarrow B) \wedge \neg\Box(B \rightarrow A),\ \Box(B \rightarrow C) \wedge \neg\Box(C \rightarrow B),\ \neg\Box\neg A,\ \neg\Box\neg B,\ \neg\Box\neg C$$
$$\models\ \Box(A \rightarrow C) \wedge \neg\Box(C \rightarrow A) \wedge \neg\Box\neg A \wedge \neg\Box\neg C$$

In testing the validity of this entailment it is natural to proceed as follows. Since the r.h.s. is a conjunction, the sequent is valid iff each of the four sequents with the same l.h.s. but just one conjunct on the r.h.s. is valid. Of these four sequents, the two with $\neg\Box\neg A$ and $\neg\Box\neg C$ on the r.h.s. are trivially valid because these formulae also occur on the l.h.s.. To prove the validity of the other two, it is convenient to move all conjuncts on the l.h.s. which have an initial negation over to the right. We shall then have the following two sequents:

$$\Box(A \rightarrow B) \wedge \Box(B \rightarrow C) \models$$
$$\Box(A \rightarrow C) \vee \Box(B \rightarrow A) \vee \Box(C \rightarrow B) \vee \Box\neg A \vee \Box\neg B \vee \Box\neg C$$
$$\Box(A \rightarrow B) \wedge \Box(B \rightarrow C) \wedge \Box(C \rightarrow A) \models$$
$$\Box(B \rightarrow A) \vee \Box(C \rightarrow B) \vee \Box\neg A \vee \Box\neg B \vee \Box\neg C$$

We can verify this proof-theoretically by the application of just one modal rule (together with ordinary classical reasoning). This is the rule **RK** which holds in any *normal* modal logic:

$$\frac{(P_1 \wedge \ldots \wedge P_n) \rightarrow P}{(\Box P_1 \wedge \ldots \wedge \Box P_n) \rightarrow \Box P}\ [\mathbf{RK}]$$

This rule together with the deduction theorem means that

$$\text{if} \qquad P_1, \ldots, P_n \models P \qquad \text{then} \qquad \Box P_1, \ldots, \Box P_n \models \Box P$$

Application of this principle validates both of our sequents, since

$$A \rightarrow B,\ B \rightarrow C \models A \rightarrow C \qquad \text{and} \qquad A \rightarrow B,\ B \rightarrow C,\ C \rightarrow A \models B \rightarrow A.$$

Introduction of the new box operator to enable positive and negative constraints to be distinguished gives us a more uniform representation since, whereas previously the meaning of an expression was tied up essentially with the reasoning methods employed, in the new language, expressions have a clear algebraic interpretation. We need no longer concern ourselves with the distinction between model and entailment constraints but can now describe spatial situations simply by a set of modal formulae; and can reason about consistency and entailment directly in this object language.

On the other hand it is not clear that this enriched language is more desirable from the computational point of view. Introduction of the new operator makes the language far more expressive and consequently much harder to reason with. The effectiveness of the original $\mathcal{I}_0^+$ representation was in large point due to its lack of expressive power.

However, we have seen that as long as the new modal operator is only used to express what was previously expressed by means of the model/entailment constraint distinction, then all $\Box$ operators will only occur either up front or negated up front in the set of formulae describing a situation; and it seems likely that the optimal approach to reasoning with such formula sets is to mimic the EZORA algorithm (used for reasoning in the extended 0-order languages) described above. Specifically this means rewriting the sequents (according to simple classical principles) to obtain sets of sequents in which all formulae have a single $\Box$ at the front: the l.h.s. is a conjunction and the r.h.s. a disjunction of such formulae. Once the sequents are in this form, it is easy to see that the sequents which correspond to entailments verifiable by the extended 0-order reasoning algorithm can all be proved using only the modal rule **RK** together with classical reasoning.

As we know that EZORA is sound and complete we can conclude that only the rule **RK** is needed to prove all entailments in $\mathcal{L}_0^{\Box}$ involving formulae in which the $\Box$ occurs either up-front or negated up-front. Nevertheless the more intuitive interpretation of the modal operator in this context is as the strong-$S5$ operator. Later on (section 6.1) we shall see the strong-$S5$ operator used in contexts where an operator satisfying only the **RK** rule would not suffice.

5 The $S4$ modality as an interior operator

In describing many relationships which can hold between spatial regions it is necessary to have some means of distinguishing the surface of a region from its interior. For instance we can say that one region is *externally connected* to another if the two regions share a boundary point but do not share any interior points. If they do share interior points we can say that they *overlap*. And, even if we do not want to reduce regions to points in our representation language, we shall still need to be able to make such distinctions.

As we saw in section 3.2 the intuitionistic calculus, $\mathcal{I}_0$, can be interpreted in terms of Boolean operators plus an interior function. This means that the calculus can be used to represent a significant family of spatial relations. One drawback with this representation is that no logical operator corresponding to the interior function appears explicitly in the language: the function occurs in the interpretations of intuitionistic negation and implication and is only referred to indirectly in logical formulae used to represent spatial constraints. In this section I explain how the modal logic $S4$ can be used as a spatial representation in which the modal operator corresponds directly to the interior function. The topological interpretation of $S4$ is not new (it can be directly inferred from the results presented in [24] and [22]), however, as far as I know it has not actually been used as a basis for spatial reasoning.

It has long been known (see [7]) that formulae of the intuitionistic propositional calculus can be translated into modal formulae in such a way that an intuitionistic formula is a theorem if and only if the resulting modal formula is valid in the logic $S4$. The translation can be specified in terms of a recursive function, m, as follows:

$$\begin{array}{lcl} m(p) & = & \Box p \quad (\text{if } p \text{ is atomic}) \\ m(\sim p) & = & \Box\neg m(p) \\ m(p \vee q) & = & m(p) \vee m(q) \\ m(p \wedge q) & = & m(p) \wedge m(q) \\ m(p \Rightarrow q) & = & \Box(m(p) \rightarrow m(q)) \end{array}$$

If we compare this to the interpretation of $\mathcal{I}_0$ given in section 3.2 we see that it is the same except that the $\Box$ operator occurs in place of i and each atomic proposition is also preceded by an additional $\Box$ operator. The latter difference arises because in the interpretation of $\mathcal{I}_0$ it is required that all constants (and hence all formulae) denote *open* subsets of the universe, whereas in a modal logic atomic formulae can normally be assigned arbitrary sets of worlds/points. Thus intuitionistic formulae correspond only to (a subset of) *necessary* $S4$ formulae.

TABLE 3. $S4$ Encoding of some RCC Relations

Relation	*Model Constraint*	*Entailment Constraints*[7]
$\mathsf{DC}(X,Y)$	$\Box\neg\Box X \vee \Box\neg\Box Y$	—
$\mathsf{EC}(X,Y)$	$\Box\neg(\Box X \wedge \Box Y)$	$\Box\neg\Box X \vee \Box\neg\Box Y$
$\mathsf{PO}(X,Y)$	—	$\Box\neg(\Box X \wedge \Box Y)$, $\Box(\Box X \rightarrow \Box Y)$, $\Box(\Box Y \rightarrow \Box X)$
$\mathsf{TPP}(X,Y)$	$\Box(\Box X \rightarrow \Box Y)$	$\Box\neg\Box X \vee \Box Y$, $\Box(\Box Y \rightarrow \Box X)$
$\mathsf{TPP}^{-1}(X,Y)$	$\Box(\Box Y \rightarrow \Box X)$	$\Box\neg\Box Y \vee \Box X$, $\Box(\Box X \rightarrow \Box Y)$
$\mathsf{NTPP}(X,Y)$	$\Box\neg\Box X \vee \Box Y$	$\Box(\Box Y \rightarrow \Box X)$
$\mathsf{NTPP}^{-1}(X,Y)$	$\Box\neg\Box Y \vee \Box X$	$\Box(\Box X \rightarrow \Box Y)$
$\mathsf{EQ}(X,Y)$	$\Box(\Box X \leftrightarrow \Box Y)$	—
$\mathsf{C}(X,Y)$	—	$\Box\neg\Box X \vee \Box\neg\Box Y$
$\mathsf{EQ}(X,\mathsf{sum}(Y,Z))$	$\Box(\Box X \leftrightarrow (\Box Y \vee \Box Z))$	—

Using this interpretation the intuitionistic representation of spatial relations given in [2] can directly be translated into an $S4$ modal representation. The result of such a translation is shown in table 3. This encoding can either be used with the EZORA reasoning algorithm or the model-constraint/entailment-constraint distinction can be eliminated by adding an additional $\Box$ operator, according to the method described in section 4. In the latter case the additional '$\Box$' operator can be treated simply as another $S4$ operator, since $S4$ (like any normal modal logic) obeys the requisite rule **RK**.

5.1 From axioms for i to modal schemata

I shall now examine the relation between $S4$ and topology directly — without considering their correspondence to $\mathcal{I}_0$. This will enable us to see clearly the relation between: a function constrained to obey certain 1st-order axioms; and a modal operator, in a 0-order calculus, which satisfies certain modal schemata. I shall show the properties of the interior operator, i, as it is conceived of in point-set topology are mirrored by the $\Box$ operator of the modal logic $S4$.

Point-set topology can be formulated in a number of ways. One of the most straightforward is to take as primitive an *interior* operator, i, which maps subsets of a topological space to their interiors. A topological space can thus be described by a structure $\langle \mathcal{U}, i\rangle$.To be an interior operator i must satisfy the following axioms:

1. $i(X) \subseteq X$
2. $i(i(X)) = i(X)$
3. $i(\mathcal{U}) = \mathcal{U}$

[7] Non-null constraints have been omitted. For each region X involved in a relation we should also add the entailment constraint $\Box\neg\Box X$, corresponding the intuitionistic constraint $\sim X$.

4. $i(X \cap Y) = i(X) \cap i(Y)$

where X and Y are any subsets of $\mathcal{U}$.

Since modal operators (under the algebraic interpretation) can be conceived of as operators on sets, we might hope that these axioms could be translated into modal schemas.[8] Indeed this is the case. To perform the translation we simply use the interpretation of the classical connectives as set operations given in section 3.1. We can then specify, for a new modal operator, schemata whose interpretations are exactly the axioms given above. In the sequel I shall use the symbol '$\mathbf{I}$' to represent the modal interior operator; this will avoid confusion with the strong-$S5$ operator, for which the symbol '$\Box$' will be reserved. The properties of $\mathbf{I}$ can thus be stated by following schemata:

1. $\mathbf{I}\, X \rightarrow X$
2. $\mathbf{I}\,\mathbf{I}\, X \leftrightarrow \mathbf{I}\, X$
3. $\mathbf{I}\,\top \leftrightarrow \top$ (where $\top$ is any tautology)
4. $\mathbf{I}(X \wedge Y) \leftrightarrow \mathbf{I}\, X \wedge \mathbf{I}\, Y$

Each of these corresponds directly to a well-known modal schema:

T. $\Box P \rightarrow P$

4. $\Box P \rightarrow \Box\Box P$

N. $\Box \top$

R. $\Box(P \wedge Q) \leftrightarrow (\Box P \wedge \Box Q)$

$\mathbf{I}$ schema 2 is actually stronger than the modal schema **4** but the bi-conditional can be weakened to a conditional in the presence of $\mathbf{I}$ schema 1. Also, $\mathbf{I}$ schema 3 is clearly equivalent to simply $\mathbf{I}\,\top$. It is well-known (see [4]) that the weakest modal logic containing these schemas is $S4$, so this is the logic we need to capture the interior operator.

$S4$ is obtained from classical propositional logic by adding the schemas:

K. $\Box(P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q)$;

T. $\Box P \rightarrow P$; and **4.** $\Box P \rightarrow \Box\Box P$;

and the rule of necessitation: $\dfrac{\vdash P}{\vdash \Box P}$ **[RN]** . Schemas **T** and **4** correspond to required properties of the interior operator. It can also be proved (see [4] theorem 4.3 case (4)) that any modal theory containing **N** and **R** is a *normal* modal theory, and hence obeys the schema **K**, as long as it is closed under the rule of equivalence: $\dfrac{\vdash P \leftrightarrow Q}{\vdash \Box P \leftrightarrow \Box Q}$ **[RE]** . Clearly any (extensional) function must conform to **RE** since a function applied to two terms which have the same value must have the same result.

That the 1st-order constraints on an interior function correspond to modal schemata defining S is not at all surprising given the relation between $\mathcal{I}_0$ and $S4$ described in

[8] This type of translation is related to the methods presented in [16] and [17] for finding model-theoretic constraints corresponding to modal schemata. However in our case we are going in the opposite direction. Another difference is that I am working with an algebraic semantics rather than the Kripke and minimal model semantics considered in the above cited works. It seems that the algebraic semantics makes the translation problem easier.

the previous section. However, the transformation of 1st-order axioms into modal schemata is a general technique for encoding spatial (and other) concepts into 0-order representations. This will be illustrated in the next section.

Since the $S4$ modality can be interpreted directly as an interior function over a topological space, we can use this interpretation to give a more direct encoding of the RCC relations into model and entailment constraints. For this purpose it is more straightforward to use the *dual* interpretation for connectedness (as mentioned in section 3.6), under which regions are closed sets which are connected if they share a point and overlap if they share an interior point. (In the $\mathcal{I}_0$ representation the other interpretation is more convenient because of the linkage between complementation and the interior function under the $\mathcal{I}_0$ coding.) The resulting $S4$ representation is shown in table 4. The new encoding is considerably simpler than that given in table 3 which was constructed indirectly via the $\mathcal{I}_0^+$ representation and the coding of $\mathcal{I}_0$ into $S4$.

TABLE 4. A simpler $S4$ Encoding of the RCC Relations

Relation	*Model Constraint*	*Entailment Constraints*
$\mathsf{DC}(X,Y)$	$\neg(X \wedge Y)$	$\neg X,\ \neg Y$
$\mathsf{EC}(X,Y)$	$\neg(\mathbf{I}X \wedge \mathbf{I}Y)$	$\neg(X \wedge Y),\ \neg X,\ \neg Y$
$\mathsf{PO}(X,Y)$	—	$\neg(\mathbf{I}X \wedge \mathbf{I}Y),\ X \rightarrow Y,\ Y \rightarrow X,\ \neg X,\ \neg Y$
$\mathsf{TPP}(X,Y)$	$X \rightarrow Y$	$X \rightarrow \mathbf{I}Y,\ Y \rightarrow X,\ \neg X,\ \neg Y$
$\mathsf{TPP}^{-1}(X,Y)$	$Y \rightarrow X$	$Y \rightarrow \mathbf{I}X,\ X \rightarrow Y,\ \neg X,\ \neg Y$
$\mathsf{NTPP}(X,Y)$	$X \rightarrow \mathbf{I}Y$	$Y \rightarrow X,\ \neg X,\ \neg Y$
$\mathsf{NTPP}^{-1}(X,Y)$	$Y \rightarrow \mathbf{I}X$	$X \rightarrow Y,\ \neg X,\ \neg Y$
$\mathsf{EQ}(X,Y)$	$X \leftrightarrow Y$	$\neg X,\ \neg Y$
$\mathsf{C}(X,Y)$	—	$\neg(X \wedge Y),\ \neg X,\ \neg Y$
$\mathsf{EQ}(X,\mathsf{sum}(Y,Z))$	$X \leftrightarrow (Y \vee Z)$	$\neg X,\ \neg Y$

6 Modal representation of convexity

We have seen how the topological interior function corresponds to the $S4$ modal box operator. Such a correspondence may suggest that other useful functions of spatial regions can be captured by modal operators in a 0-order calculus. In the remainder of this paper I shall investigate how the notion of the *convex-hull* (see [20]) of a region could be represented by means of a modal operator. By the convex-hull of a region I mean the smallest convex region of which it is a part. If one were to stretch an elastic membrane round a region then the convex-hull would be the whole of the region enclosed. Figure 2 shows convex-hulls of two regions in 2 dimensions (region B is a two piece region).

Randell et al. [20] have shown how various notions of containment can be defined in terms of convex-hull. For instance, we can say that $\mathsf{INSIDE}(X,Y)$ holds if X does not overlap Y but is part of the convex-hull of Y.

Before thinking about possible modal representations, we need to have a good idea of what properties we expect convex-hulls to exhibit. The following 1st-order axioms

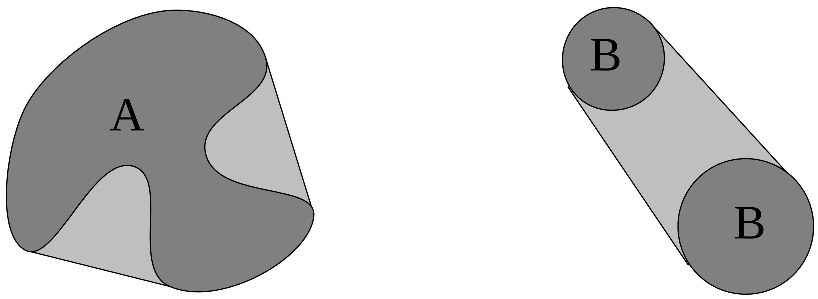

FIG. 2. Illustration of convex-hulls in 2 dimensions

specify properties of the convexity operator, 'conv':

1. $\forall x[\mathsf{conv}(\mathsf{conv}(x)) = \mathsf{conv}(x)]$
2. $\forall x \mathsf{TP}(x, \mathsf{conv}(x))$
3. $\forall x \forall y[\mathsf{P}(x, y) \;\rightarrow\; \mathsf{P}(\mathsf{conv}(x), \mathsf{conv}(y))]$
4. $\forall x \forall y[\mathsf{conv}(x) = \mathsf{conv}(y) \;\rightarrow\; \mathsf{C}(x, y)]$ [9]
5. $\forall x \forall y[\mathsf{prod}(\mathsf{conv}(x), \mathsf{conv}(y)) = \mathsf{conv}(\mathsf{prod}(\mathsf{conv}(x), \mathsf{conv}(y)))]$ [10]

Of these the first four were presented in [2], where further comments can be found. The fifth axiom states that the prod (i.e. intersection — see [20]) of two convex regions is itself convex. This list is not guaranteed to be a complete axiomatisation of the conv operator: it is very difficult to be sure that a set of axioms fully capture a concept unless we have a formal model (or set of models) within which the concept is defined and show that the axioms are sound and complete with respect to that model (those models). Investigating such models is the subject of ongoing work.

6.1 conv *as a modal operator*

We would like to represent conv as a modal operator in a 0-order calculus. This calculus will be a multi-modal language containing the usual classical connectives (interpreted algebraically according to section 3.1) plus three modal operators:

$\mathbf{I}$ an interior operator, constrained to behave exactly as the $S4$ modality,

$\Box$ the strong-S5 operator,

$\bigcirc$ the convexity operator, whose properties are to be specified.

To fix the meaning of the new operator, we need to find 0-order axiom schemata (or rule schemata) to enforce the desired properties of $\bigcirc$. These schemata will correspond to the 1st-order axioms given above. I do not know of a general method for performing this kind of transformation and it seems unlikely that such a method exists. However, in each case we can see that under the algebraic interpretations of the logical operators the schemata are equivalent to the axioms.

---

[9] Actually this is not necessarily true for infinite regions.

[10] It is possible that this may be derivable from the other axioms plus the definition of prod [20] but, if this is so, it is not obvious.

The schema corresponding to axiom 1 is very simple:

$$\bigcirc \bigcirc X \leftrightarrow \bigcirc X \qquad (Sch1)$$

Axiom 2 is a little harder to represent as a modal schema. $\mathsf{TP}(X,Y)$ means that X is a *tangential part* of X. This holds if either X is a tangential *proper* part of Y or X is equal to Y. Thus to represent this we use the encoding for $\mathsf{TP}(X,Y)$ given in table 4 but drop the second entailment constraint $Y \rightarrow X$ which would ensure that X and Y are non equal. Hence, using the strong-$S5$ $\Box$ rather than the model/entailment-constraint distinction, axiom 2 can be represented by the schema

$$\Box(X \rightarrow \bigcirc X) \ \wedge \ \neg\Box(X \rightarrow \mathbf{I} \bigcirc X), \qquad (Sch2)$$

which says that all regions are part of their convex-hull but not part of the interior of their convex-hull. We may note that the initial $\Box$ in the first conjunct is redundant, since it is implicit in modal axiom schemata that they are true in all possible worlds — or, in the context of algebraic semantics, that their denotation is $\mathcal{U}$.

Axiom 3, which states that if X is part of Y then $\bigcirc X$ is part of $\bigcirc Y$ can be represented by

$$\Box(X \rightarrow Y) \rightarrow (\bigcirc X \rightarrow \bigcirc Y). \qquad (Sch3)$$

This requires some explanation. In general, where we have a 1st-order axiom of the form $p \rightarrow q$, this will be translated by $\Box\,\tau(p) \rightarrow \tau(q)$ (where $\tau(\alpha)$ is the representation of α), which ensures that if $\tau(p) = \mathcal{U}$ then $\tau(q) = \mathcal{U}$. Note that we do not need $\Box\,\tau(p) \rightarrow \Box\,\tau(q)$ because the antecedent must either denote $\emptyset$, in which case the schema is trivially satisfied, or it denotes $\mathcal{U}$, in which the consequent must also denote $\mathcal{U}$. If we were to write simply $\tau(p) \rightarrow \tau(q)$ this would represent the stronger requirement that $\tau(p)$ is always a subset of $\tau(q)$ whether or not $\tau(p) = \mathcal{U}$.

Using a similar transformation axiom 4 can be straightforwardly represented by:

$$\Box(\bigcirc X \leftrightarrow \bigcirc Y) \ \rightarrow \ \neg\Box\neg(X \wedge Y) \qquad (Sch4)$$

$\neg\Box\neg(X \wedge Y)$ corresponds to the entailment constraint representing $\mathsf{C}(X,Y)$ and asserts that X and Y share at least one point.

Finally axiom 5 can be straightforwardly captured by:

$$\bigcirc(\bigcirc X \wedge \bigcirc Y) \leftrightarrow (\bigcirc X \wedge \bigcirc Y) \qquad (Sch5)$$

It should be noted that the strong-$S5$ operator, $\Box$, is not needed if we specify the logic by means of *rule* schemata rather than only *axiom* schemata. For example, *Sch3* becomes:

$$\frac{\vdash X \rightarrow Y}{\vdash \bigcirc X \rightarrow \bigcirc Y}\,[\bigcirc\ \mathbf{Mon}]$$

which tells us that $\bigcirc$ is monotonic with respect to the part relation (i.e. $\rightarrow$).

The second conjunct of *Sch2* would correspond to the rule:

$$\frac{\vdash X \rightarrow \mathbf{I} \bigcirc X}{\vdash \bot}\,[\bigcirc\,\mathsf{TP}]$$

and *Sch4* to the rule:

$$\frac{\vdash (\bigcirc X \leftrightarrow \bigcirc Y) \wedge \neg(X \wedge Y)}{\vdash \bot} [\leftrightarrow \bigcirc \mathsf{C}].$$

7 Summary

I have presented a preliminary exploration of the potential applications of modal logics to qualitative spatial reasoning. The approach is a generalisation of the intuitionistic representation, $\mathcal{I}_0^+$, presented in [2] and provides a much richer language for expressing spatial information. I have shown how the meta-level reasoning algorithm [2] can be dispensed with by adding a new *strong-S5* operator to the object language and how a modal interior operator, $\mathbf{I}$, (equivalent to the $S4$ $\Box$) provides for more direct encoding of spatial relationships than $\mathcal{I}_0^+$. Finally, I have examined how a modal *convex-hull* operator could be defined by translating a 1st-order axiomatisation into modal schemata.

8 Further work

Whilst modal representations of spatial relations can be shown to have a theoretical advantage over 1st-order representations (namely that decision procedures are known for the modal languages), nevertheless doubts may remain as to whether the modal representations could ever be of practical use. After all a decision procedure does not necessarily provide us with an *effective* means of computation. Ideally we would like to have polynomial algorithms for spatial reasoning. Recently, a lot of research has been directed towards the need for more efficient modal reasoning systems [25], [1], [3], [5]. If the modal approach to qualitative reasoning is to be of practical use it will be necessary to demonstrate that the modal representations can be effectively manipulated. One way to do this would be to identify tractable sub-languages of modal calculi, which are capable of representing significant sets of spatial relations.

Another important direction for further work is to investigate how the expressive power of the representation can be extended. In section 6.1 I showed how properties of the convex-hull operator can be captured by means of modal schemata; and this technique could be applied to other spatial concepts. However, the method is somewhat *ad hoc* and does not provide us with a direct interpretation of the operator, in terms of model structures. To do this we shall need richer mathematical structures as models. I am currently looking at how modal calculi can be interpreted in Cartesian spaces over Euclidean fields. These have much more structure than topological spaces and allow many non-topological spatial relations to be defined. In particular convex-hull can be characterised in terms of the *betweenness* relation which can easily be defined in these models (see [23]).

Acknowledgements

This work was partially supported by the CEC under the Basic Research Action MEDLAR 2, Project 6471 and by the EPSRC under grants GR/H 78955 and GR/G64388.

References

[1] Y. Auffray, P. Enjalbert, and J-J. Herbrard. Strategies for modal resolution. *Journal of Automated Reasoning*, (6), 1990.

[2] B. Bennett. Spatial reasoning with propositional logics. In J Doyle, E Sandewall, and P Torasso, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference (KR94)*, San Francisco, CA., 1994. Morgan Kaufmann.

[3] L. Catach. Tableaux: A general theorem prover for modal logics. *Journal of Automated Reasoning*, (7), 1991.

[4] B F Chellas. *Modal Logic: An Introduction*. Cambridge University Press, 1980.

[5] S. Demri. Efficient strategies for automated reasoning in modal logics. In C. MacNish, D. Pearce, and L.M. Pereira, editors, *Logics in Artificial Intelligence*, volume 833 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, 1994. Proceedings of the workshop JELIA'94.

[6] M Egenhofer. Reasoning about binary topological relations. In O Gunther and H J Schek, editors, *Proceedings of the Second Symposium on Large Spatial Databases, SSD'91 (Zurich, Switzerland). Lecture Notes in Computer Science 525*, pages 143–160, 1991.

[7] M.C. Fitting. *Intuitionistic Logic Model Theory and Forcing*. North-Holland, 1969.

[8] M.C. Fitting. *Proof Methods for Modal and Intutionistic Logics*. Reidel, 1983.

[9] R. Goldblatt. Metamathematics of modal logic. In *Metamathematics of Modality*, chapter 1. CLSI Publications, 1993.

[10] N M Gotts. How far can we 'C'? defining a 'doughnut' using connection alone. In J Doyle, E Sandewall, and P Torasso, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference (KR94)*, San Francisco, CA., 1994. Morgan Kaufmann.

[11] B. Jónsson and A. Tarski. Boolean algebras with operators, part I. *American Journal of Matehematics*, 73:891–939, 1951.

[12] S. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24:1–14, 1959.

[13] K. Kuratowski. *Introduction to Set Theory and Toplolgy*. Pergamon Press, 2nd edition, 1972.

[14] E J Lemmon. Algebraic semantics for modal logics I. *Journal of Symbolic Logic*, 31:46–65, 1966.

[15] E J Lemmon. Algebraic semantics for modal logics II. *Journal of Symbolic Logic*, 31:191–218, 1966.

[16] R Nossum and D Gabbay. Semantical properties of some modal systems of logic. In *4th Scandinavian Conference on Artificial Intelligence*, 1993.

[17] Hans Jürgen Ohlbach. Synthesizing semantics for extensions of propositional logic. Technical Report MPI-I-94-225, Max-Planck-Institut für Informatik, Saarbrücken, Germany, 1994.

[18] D.C. Oppen. Complexity, convexity and combinations of theories. *Theoretical Computer Science*, (12):291–302, 1980.

[19] D A Randell and A G Cohn. Modelling topological and metrical properties of physical processes. In R Brachman, H Levesque, and R Reiter, editors, *Proceedings 1st International Conference on the Principles of Knowledge Representation and Reasoning*, pages 55–66, Los Altos, 1989. Morgan Kaufmann.

[20] D A Randell, Z Cui, and A G Cohn. A spatial logic based on regions and connection. In *Proceedings 3rd International Conference on Knowledge Representation and Reasoning*, pages 165–176, San Mateo, 1992. Morgan Kaufmann.

[21] A. Tarski. Foundations of the geometry of solids. In *Logic, Semantics, Metamathematics*, chapter 2. Oxford Clarendon Press, 1956. trans. J.H. Woodger.

[22] A. Tarski. Sentential calculus and topology. In *Logic, Semantics, Metamathematics*, chapter 17. Oxford Clarendon Press, 1956. trans. J.H. Woodger.

[23] A. Tarski. What is elementary geometry? In L.E.J. Brouwer, E.W. Beth, and A. Heyting, editors, *The Axiomatic Method (with special reference to geometry and physics)*, pages 16–29, Amsterdam, 1959. North-Holland.

[24] A. Tarski and J.C.C. McKinsey. Some theorems about the sentential calculi of Lewis and Heyting. *Journal of Symbolic Logic*, 13:1–15, 1948.

[25] L. A. Wallen. *Automated Deduction in Nonclassical Logics*. MIT Press, 1990.

Received 10 October 1994. Revised 21 February 1995