QUALITATIVE SPATIAL REASONING VIA 3-VALUED HETEROGENEOUS LOGIC

Konstantine Arkoudas, Selmer Bringsjord

Departments of Cognitive and Computer Science, RPI, Troy, NY, U.S.A.

Sangeet Khemlani

Department of Cognitive Science, Princeton University, NJ, U.S.A.

Abstract: Systems for qualitative spatial reasoning (QSR) are usually formulated as relation algebras, and reasoning in such systems is performed by constraint-satisfaction techniques. While this is often adequate, it is a rather inexpressive framework that cannot model and solve many spatial reasoning problems; it can also complicate the combination of different spatial formalisms, e.g., the combination of topological with metric primitives, or absolute orientation with relative orientation. Here we suggest an alternative approach, whereby spatial information is expressed in a rich quantified 3-valued logic, equipped with a novel semantics for dealing with incomplete information. Decidability is ensured by a systematic compilation into propositional logic and the use of SAT solvers. To illustrate, we define and implement a new system for two-dimensional positional reasoning that combines Frank's cardinal-direction calculus, the ip- op calculus for reasoning about relative orientation, and various new positional primitives. Unlike previous work, the system uses diagrams as well as symbolic formulas. In particular, the logic we introduce is *heterogeneous*, meaning that it combines symbolic and diagrammatic representation and inference.

1 A HYBRID SYSTEM FOR REASONING ABOUT ORIENTATION

Representing and reasoning about position and orientation is an active area of QSR, with applications ranging from robot navigation and geographic information systems to computational linguistics. Most of the existing systems are based either on absolute reference systems or on relative reference systems. For the latter, a reference axis is introduced by fixing a given *origin* and a *relatum*, and then the position of a given *referent* is described with respect to that axis. It is increasingly recognized that realistic scenarios demand the ability to handle both absolute and relative orientation.

The system we are about to introduce, *CDC* (for *C*ombined *D*irection *C*alculus), integrates:

1. an absolute-reference orientation system implementing Frank's cardinal-directions calculus (Frank, 1991); and

2. a relative-reference orientation system in which the reference axis is specified by an arbitrary origin and relatum, implementing Ligozat's flip-flop calculus (Ligozat, 1993).

We introduce several additional primitives that are not part of either of these two systems.

In what follows we demonstrate the system on a number of examples, starting with a problem from (Isli et al., 2001) that illustrates the need for combining absolute- and relative-orientation reasoning:

1. Viewed from Hamburg, Berlin is to the left of Paris, Paris is to the left of London, and Berlin is to the left of London.

2. Viewed from London, Berlin is to the left of Paris.

3. Hamburg is to the north of Paris, and north-west of Berlin.

4. Paris is to the south of London.

The first two premises are consistent. Indeed, if we `assert` the first two premises (i.e., insert them into the knowledge base) and then issue the command (`find-model`), *CDC* will automatically find and display the following spatial model:

Hamburg	Berlin	
London	Paris	

Likewise, the last two premises are mutually consistent. *CDC* automatically produces the following model for them:
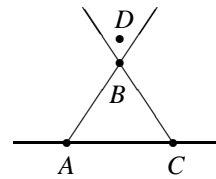
Hamburg		
London		Berlin
Paris		

Nevertheless, the conjunction of all four premises is inconsistent,[1] and *CDC* readily discovers this. There are two ways to demonstrate the inconsistency. One is to ask the system to find a model for the current knowledge base (which contains all four premises). If the knowledge base is inconsistent, as in this example, the system will report that no such model exists. The other is to ask whether the sentence `false` follows from the knowledge base. In this case *CDC* confirms that `false` indeed follows.

Consider next the *Indian-tent problem*, a rather simple problem that nevertheless presents challenges to several QSR systems (Röhrig, 1997, p. 229) and used as a benchmark by the SparQ-Toolbox (Wallgrün et al., 2006): There are four objects (points, regions, or whatever), *A*, *B*, *C*, and *D*, whose spatial arrangement is as follows:

1. Viewed from *A*, *C* is to the right of *B* (equivalently, *C* is to the right of the line from *A* to *B*).

2. Viewed from *C*, *D* is to the right of *B*.

3. Viewed from *A*, *D* is to the left of *B*.

The goal is to deduce that viewed from *C*, *D* is to the right of *A*. Geometrically, the configuration must be isomorphic to the following:

---

[1]The only reason the four premises are jointly inconsistent is because, in calculi of this sort, directions such as *north* and *northwest* are required to be mutually exclusive. That clearly represents a departure from ordinary usage, where the two are not only compatible (e.g., we say that Chicago is both northwest and north of Baltimore), but in fact one implies the other. The requirement is nevertheless customarily imposed because in the constraint-based paradigm that has dominated the field, the base relations of a QSR calculus must be mutually exclusive.

When we `assert` these three premises and ask *CDC* to find a model, the system responds with the following diagram:

A		
	B	D
C		

Further, when we `query` whether it follows logically that *D* must be to the right of *A* from *C*'s perspective, *CDC* quickly responds affirmatively.

For our third and final example, suppose that we must arrange five objects (e.g., furniture pieces) *A*, *B*, *C*, *D*, and *E*, according to the following constraints:

1. *A* must not be adjacent to *C*.

2. Nothing is to the right of *E*.

3. If *D* and *A* are not adjacent, then *B* should be in the middle.

4. *D* is above all others.

5. *E* and *D* are adjacent.

When we ask *CDC* to find a model for these requirements, it promptly[2] returns the following diagram:

		D
A	B	E
	C	

The remainder of the paper is structured as follows. The next section contains a discussion of our overall approach to QSR in general terms. In section 3 we apply this methodology to define *CDC* rigorously. In section 4 we show how to carry out the SAT reduction for *CDC* and in general. Finally, section 5 concludes.

2 GENERAL METHODOLOGY

A spatial problem deals with a finite system of objects $s_1, \ldots, s_N$. Each object has a number of *attributes*, which typically represent spatial properties. While

---

[2]The current implementation of *CDC* solves all of the sample problems that appear in the paper in a fraction of a second (on an 2GHz IBM T2500 with 1GB of RAM). The complete source code along with a machine-readable test suite of numerous problems, including these examples, can be obtained by contacting the authors.

there may be several attributes in general, in practice there is often only one attribute of interest. In this case, as in many others, this attribute is *location*, which here consists of a pair of numeric coordinates that locate each object on a two-dimensional grid.

A *system state* is a function σ that maps each object s_i to a finite and non-empty set of attribute values. As a simple example, suppose we have three objects s_1, s_2, and s_3, to be located on a 2×2 grid. Then a system state σ might map s_1 to $(1,1)$, s_2 to $(2,1)$, and s_3 to $(1,2)$:

$$\begin{aligned} \sigma(s_1) &= \{(1,1)\}; \\ \sigma(s_2) &= \{(2,1)\}; \\ \sigma(s_3) &= \{(1,2)\}. \end{aligned} \quad (1)$$

We can depict σ diagrammatically as follows:

s_1	s_3
s_2	

Such a state is called a *world*, because it maps each object to a *unique* attribute value, in this case to a unique location.[3] Thus a world provides a maximal amount of information: it gives the precise attribute values (e.g., the precise locations) of all objects. Oftentimes, however, we do not know exact attribute values. For instance, we might know the precise location of s_1 (say, $(2,2)$), but for s_2 and s_3 we might only know that they are both on the top row, but without knowing their exact positions. That would be captured by the following state:

$$\begin{aligned} \sigma(s_1) &= \{(2,2)\}; \\ \sigma(s_2) = \sigma(s_3) &= \{(1,1),(1,2)\}. \end{aligned} \quad (2)$$

In the extreme case, we might have no information whatsoever about the locations of any of the objects:

$$\sigma(s_1) = \sigma(s_2) = \sigma(s_3) = \{(1,1),(1,2),(2,1),(2,2)\}.$$

So that is why states map objects to finite *sets* of attribute values, rather than single attribute values. Since set membership is disjunctive, this provides us with a technically convenient device for dealing with incomplete information. Moreover, the finiteness requirement ensures that we can encode the content of a state with a finite disjunction. For instance, state (2) can be represented by the CNF formula

$$loc(s_1,(2,2)) \wedge [loc(s_2,(1,1)) \vee loc(s_2,(1,2))] \\ \wedge [loc(s_3,(1,1)) \vee loc(s_3,(1,2))],$$

where the literal $loc(s_i,l)$ has the obvious meaning. Any state σ can be straightforwardly encoded by a CNF formula F_σ.

[3]Technically, each object is mapped to a singleton, but it is convenient to treat worlds as if they map objects directly to values.

Let σ_1, σ_2 be system states. We say that σ_2 is an *extension* of σ_1, written $\sigma_2 \sqsubseteq \sigma_1$, iff $\sigma_2(s_i) \subseteq \sigma_1(s_i)$ for every $i = 1, \ldots, N$. If $\sigma_2 \sqsubseteq \sigma_1$ and $\sigma_1 \not\sqsubseteq \sigma_2$, then σ_2 is a *proper* extension of σ_1, written $\sigma_2 \sqsubset \sigma_1$. Thus, if $\sigma_2 \sqsubseteq \sigma_1$ then F_{σ_2} subsumes F_{σ_1}.

Note that oftentimes system states can (and should) be depicted diagrammatically. This is possible even in the presence of partial information (i.e., when the state is not a world), if we only introduce appropriate abstraction tricks and corresponding diagram-parsing conventions. For instance, if we place a question mark in a location to indicate that we do not know which object appears there, while an unoccupied location is simply left blank, then state (2) can be depicted as follows:

?	?
	s_1

The pervasive use of such diagrams is a distinguishing aspect of our approach. Indeed, in our work "system state" and "diagram" are used synonymously.

Let us now describe the syntax of the underlying logic. First, every object s_i is given a name c_i, and indeed for many purposes the objects can be identified with their names. A *term* is either an object name c_i or else a variable v. (To keep these apart, variables and constants start with lower- and upper-case letters, respectively.) Atomic sentences are of the form $(R\ t_1 \cdots t_k)$, where R is a *relation symbol* of arity k and $t_1 \cdots t_k$ are terms. There are also negations ($\text{not}\ p$), conjunctions and disjunctions ($\text{and/or}\ p_1 \cdots p_k$), conditionals and biconditionals ($\text{if/iff}\ p_1\ p_2$), and universal and existential quantifications ($\text{forall/some}\ v_1 \cdots v_k\ p$).

A specific system is largely determined by the stock of available relation symbols and their meaning. More precisely, to define a QSR system by this methodology, one must choose

1. a set of object attributes (as we remarked, a single attribute *location* suffices in many cases); and

2. a finite set of relation symbols $\mathcal{R}$, and their interpretations.

The interpretation of a symbol $R \in \mathcal{R}$ is a computable relation $\overline{R}$ on some attributes (typically on *location*). Thus, for instance, supposing that left is a binary relation symbol, $\overline{\text{left}}$ would be a binary relation on locations, defined, e.g., as follows:

$$\overline{\text{left}}((r_1,c_1),(r_2,c_2)) \equiv c_1 < c_2.$$

Then an atom such as (left B C) will be true in a given state σ iff the $\overline{left}$ relation definitely holds between all possible locations that σ assigns to the objects named B and C. (Recall that a state might map

an object to multiple locations.) Thus, e.g., assuming that A, B, and C are the names of the objects s_1, s_2, and s_3, respectively, the atom (left B C) is true in world (1), as s_2 is definitely to the left of s_3 in that state. Likewise, $(R\ t_1 \cdots t_k)$ will be false in a state σ iff $\overline{R}$ fails for *all* possible attribute values that σ assigns to the objects named by $t_1, \ldots, t_k$. Thus, e.g., (left A B) is false in state (1). But if $\overline{R}$ holds for *some* of these values and fails for others, then the truth value of $(R\ t_1 \cdots t_k)$ is *unknown*—the third value of the 3-valued semantics. Thus, e.g., the truth value of (left B C) is unknown in state (2), because it is true if s_2 assumes the location $(1,1)$ and s_3 assumes the location $(1,2)$, but false if s_2 is assigned to $(1,2)$ and s_3 to $(1,1)$. Given such interpretations for the relation symbols, any sentence p can be compiled into an equivalent formula F_p in propositional logic (in the context of the inference problems described below).

Let us be more precise. Recall that a term t is either a constant name c or a variable v. Thus, to evaluate a term t, i.e., to find out which system object it denotes, we need two pieces of information: a mapping from variables to objects, and a mapping from constants to objects. The mapping from constants to objects is usually fixed once and for all in the beginning of the session with the system. For any constant name c, we write $\overline{c}$ for the system object denoted by it through this initial mapping, e.g., $\overline{A} = s_1$, $\overline{B} = s_2$, $\overline{C} = s_3$. A variable mapping is a total function χ from the set of variables to the set of objects. Given such a mapping χ, the denotation of a term t is written as t^χ, and is defined as follows: If t is one of the constants, c, then $t^\chi = \overline{c}$; and if t is a variable v, then $t^\chi = \chi(v)$. We write $\chi[v \mapsto s_i]$ for the mapping that assigns s_i to v and agrees with χ everywhere else.

We first define the truth value of any given sentence w.r.t. a given world w and a given variable mapping χ, denoted $\mathcal{V}_{w/\chi}[p]$, as follows. Suppose first that p is an atomic sentence. If p is an identity $(=\ t_1\ t_2)$, then p is *true* iff $t_1^\chi = t_2^\chi$. For non-identities, $\mathcal{V}_{w/\chi}[(R\ t_1 \cdots t_k)]$ is defined as follows:

$$\begin{cases} \textit{true if } \overline{R}(w(t_1^\chi), \ldots, w(t_k^\chi)); \\ \textit{false otherwise.} \end{cases}$$

For non-atomic p, $\mathcal{V}_{w/\chi}[p]$ is defined in accordance with the strong 3-valued Kleene scheme, e.g., $\mathcal{V}_{w/\chi}[(\text{and } p_1\ p_2)]$ is *true* iff both $\mathcal{V}_{w/\chi}[p_1]$ and $\mathcal{V}_{w/\chi}[p_2]$ are *true*; *false* if one of them is *false*; and *unknown* otherwise. Universal and existential quantifications are desugared into conjunctions and disjunctions, respectively.

A *knowledge base* is a finite set of sentences β. A *context* is a pair $\gamma = (\beta, \sigma)$ consisting of a knowledge base β and a system state (diagram) σ. The fol-

lowing specifies the key notion of logical entailment in this framework: A world w *satisfies a sentence p* w.r.t. a variable mapping χ iff $\mathcal{V}_{w/\chi}[p] = \textit{true}$. This is denoted by writing $w \models_\chi p$. Likewise, w *satisfies a system state* σ, written $w \models \sigma$, iff $w \sqsubseteq \sigma$. We say that w satisfies a context $\gamma = (\beta, \sigma)$ w.r.t. a given χ, written $w \models_\chi (\beta, \sigma)$, iff $w \models_\chi p$ for all $p \in \beta$ and $w \models \sigma$. A context γ *entails a sentence p*, written $\gamma \models p$, iff $w \models_\chi \gamma$ implies $w \models_\chi p$ for every world w and variable mapping χ. Finally, γ *entails a system state* σ, written $\gamma \models \sigma$, iff $w \models_\chi \gamma$ implies $w \models \sigma$ for all w and χ.

With this background, we can describe the two types of inference supported in our framework as follows:

1. *Theorem proving*: Given a context γ, determine whether or not

   - a sentence p follows from γ; or
   - a state σ' follows from γ.

2. *Model finding*: Given a context γ, find a model for it, if one exists, or else report inconsistency. The system should be able to find as many distinct models for γ as possible.

For theorem proving, we encode the given context γ as a CNF formula F_γ, and check the satisfiability of $F_\gamma \wedge \mathcal{A} \wedge \neg F_p$ or that of $F_\gamma \wedge \mathcal{A} \wedge \neg F_{\sigma'}$, where $\mathcal{A}$ is a canonicity axiom that will be discussed later. For model-finding, we simply look for satisfying interpretations for $F_\gamma \wedge \mathcal{A}$.

We stress that grid-based numeric locations are *not* a necessary feature of this methodology. Locations could be data values of an arbitrary type, e.g., the thirteen relative regions of the Double-Cross Calculus (Freksa, 1992) determined by an arbitrary origin and relatum. Then a system state might map an object s_i to a set of "locations" such as $\{\textit{left-front}, \textit{right-back}\}$.

3 DEFINITION OF *CDC*

To define *CDC* in accordance with the preceding schema, we need to (a) specify the object attribute(s), and (b) specify the relation symbols and their interpretations. There is only one attribute, location, so for (a) we only need to specify the type of locations used in *CDC*. These will be cells on a two-dimensional grid. In particular, letting R and C denote the number of rows and columns of the grid, respectively,[4] we identify a location with an ordered pair (i, j) with

---

[4]Both dimensions of the grid (R and C) are adjustable parameters in our implementation; they can take any positive values.

$1 \leq i \leq R$ and $1 \leq j \leq C$. We write $\mathcal{L}$ for the set of all locations, namely, $\{(1,1),\ldots,(R,C)\}$. The top row and leftmost column are row 1 and column 1, while the bottom row and rightmost column are row R and column C, respectively. Thus, a system state here is a function

$$\sigma : \{s_1,\ldots,s_N\} \to [\mathcal{P}(\mathcal{L}) \setminus \{\emptyset\}]$$

that assigns a non-empty set of locations to every system object.

For part (b): *CDC* has 24 relation symbols, 15 of which are binary, 8 are ternary, and one is unary. The unary relation is `middle`; (`middle` t) holds iff the object denoted by t is located at the center of the absolute reference system. The following are the binary relations: `north`, `south`, `east`, `west`, `north-west`, `north-east`, `south-west`, `south-east`, `above`, `below`, `left`, `right`, `diag`, `adjacent`, and `same-location`. In addition, there is the equality symbol: (`=` s t) iff s and t denote the same object. The ternary relations are those of the flip-flop calculus (`ff-right`, `ff-left`, `ff-front`, `ff-back`, `ff-inside`, `ff-start`, `ff-end`), and an extra ternary relation `between`.

We now come to the interpretations of these symbols. For each symbol R, $\overline{R}$ is a relation of the same arity on $\mathcal{L}$. Thus, for instance, $\overline{\text{above}}$ is a binary relation on $\mathcal{L}$. Specifically, $\overline{\text{above}}((r_1,c_1),(r_2,c_2))$ iff $r_1 < r_2$. We illustrate with the interpretations of a few more of the binary primitives:

$$\overline{\text{west}}((r_1,c_1),(r_2,c_2)): r_1 = r_2 \text{ and } c_1 < c_2$$
$$\overline{\text{adjacent}}((r_1,c_1),(r_2,c_2)): [r_1 = r_2 \text{ and }$$
$$|c_1 - c_2| = 1] \text{ or } [c_1 = c_2 \text{ and } |r_1 - r_2| = 1]$$

The interpretations of the rest should be obvious. The only somewhat tricky case is $\overline{\text{diag}}$, which holds for positions that are located diagonally.

For the base relations of the flip-flop calculus, we transform locations (r,c) into Cartesian coordinates (x,y), where $x = c$ and $y = R - r + 1$. Then, given an origin (r_1,c_1), a relatum (r_2,c_2), and a referent (r_3,c_3), with Cartesian coordinates (x_1,y_1), (x_2,y_2), and (x_3,y_3), respectively, we compute the slope and constant of the line from the origin to the relatum, and then determine the relative location of the referent by analytic geometry. For instance, writing $b? \to e_1; e_2$ for the conditional expression that denotes the value of e_2 (e_3) is b is true (false), the following interprets *ff-right*:

$$\overline{\textit{ff-right}}((r_1,c_1),(r_2,c_2),(r_3,c_3)) \equiv (c_1 = c_2)? \to$$
$$[r_2 \leq r_1? \to (c_3 > c_1); (c_3 < c_1)];$$
$$[(c_2 > c_1)? \to y_3 < y'; y_3 > y']$$

where $y' = (slope \cdot x) + constant$,

$$slope = y_2 - y_1/x_2 - x_1,$$

and $constant = y_1 - (slope \cdot x_1)$.

Finally, note that the semantics do not preclude worlds in which multiple objects are in the same location. In practice, our implementation rules out such worlds by adding the following sentence to the global knowledge base:

```
(forall x y
  (if (same-location x y) (= x y)))
```

4 TRANSLATION TO SAT

Recall that N, R, and C are the numbers of objects, rows, and columns, respectively. Our translation uses two basic types of Boolean variables, *location–i–r–c*, asserting that object i is in location (r,c), and *eq–i–i′*, asserting that objects i and $i′$ are identical. We define two additional variables in terms of *location*, *row–i–r* and *col–i–c*, asserting that object i is in row r and column c, respectively. There are, therefore, $N \cdot R \cdot C + N^2 + N \cdot (R+C)$ variables. For greater readability, we write variables of the form *location–i–r–c* and *eq–i–i′* as $location(i,(r,c))$ and $eq(i,i′)$, respectively, and likewise for *row* and *col*. The following axiom defines *row*:

$$\bigwedge_{i=1}^{N} \left[\bigwedge_{r=1}^{R} row(i,r) \Leftrightarrow \bigvee_{c=1}^{C} location(i,(r,c)) \right]$$

The definition of *col* is similar.

To weed out unintended models, we must ensure that for every $i \in 1,\ldots,N$ there is some $l \in \mathcal{L}$ such that *location(i,l)*; i.e., every object occupies *some* location:

$$\bigwedge_{i=1}^{N} \left[\bigvee_{l \in \mathcal{L}} location(i,l) \right]$$

Furthermore, *location* must be univalent, i.e., no object can occupy more than one location:

$$\bigwedge_{i=1}^{N} \bigwedge_{l \in \mathcal{L}} \left[location(i,l) \Rightarrow \bigwedge_{l' \in \mathcal{L} \setminus \{l\}} \neg location(i,l') \right]$$

We also postulate the following two axioms formalizing the semantics of the identity relation:

$$\left[\bigwedge_{i=1}^{N} eq(i,i) \right]$$

and

$$\bigwedge_{i=1}^{N} \left[\bigwedge_{j \in \{1,\ldots,N\} \setminus \{i\}} \neg eq(i,j) \right]$$

We write $\mathcal{A}$ for the conjunction of all of the above

$$\texttt{below}^{\mathcal{T}}(s_i, s_j) \;=\; \bigwedge_{x=1}^{R} \left[row(i,x) \Rightarrow \bigvee_{y=x+1}^{R} row(j,y) \right]$$

$$\texttt{ff-right}^{\mathcal{T}}(s_i, s_j, s_k) \;=\; \bigwedge_{(l_1, l_2) \in \mathcal{L}^2} \left[location(i, l_1) \wedge location(j, l_2) \Rightarrow \bigvee_{l_3 \in R(l_1, l_2)} location(k, l_3) \right]$$

$$\text{where } R(l_1, l_2) = \{ l_3 \in \mathcal{L} \mid \textit{ff-right}(l_1, l_2, l_3) \}$$

$$\texttt{north-west}^{\mathcal{T}}(s_i, s_j) \;=\; \texttt{above}^{\mathcal{T}}(s_i, s_j) \wedge \texttt{left}^{\mathcal{T}}(s_i, s_j)$$

Figure 1: SAT translation of sample base relations.

axioms, including the definitions of *row* and *col*. By an *interpretation I* we mean a function that assigns a truth value to every atom of the form *location–i–r–c* and *eq–i–i'*, $i, i' \in \{1, \ldots, N\}$, $r \in \{1, \ldots, R\}$, $c \in \{1, \ldots, C\}$. We write $I \models F$, where F is a propositional formula over this set of variables, to mean that I satisfies F, in the usual sense of propositional-logic semantics. We say that an interpretation I is *canonical* iff $I \models \mathcal{A}$. We are only interested in canonical interpretations.

A system state (diagram) σ can be encoded as a propositional formula F_σ as follows:

$$F_\sigma = \bigwedge_{i=1}^{N} \left[\bigvee_{l \in \sigma(s_i)} location(i, l) \right]$$

Alternatively—and more efficiently—we can encode σ as the conjunction of all unit clauses that state where an object cannot be. We can now define the main translation function $\mathcal{T}$ that takes a sentence p and a variable mapping χ and produces a formula in propositional logic (over the aforementioned set of variables) that captures the 3-valued-logic semantics of p in a sense made rigorous by the theorem below. For atoms, $\mathcal{T}((R\ t_1 \cdots t_k), \chi) = R^{\mathcal{T}}(t_1^\chi, \ldots, t_k^\chi)$, where $R^{\mathcal{T}}$ is defined for some sample R in figure 1. Boolean combinations are straightforward, e.g.,

$$\mathcal{T}((\texttt{and}\ p_1\ p_2), \chi) = \mathcal{T}(p_1, \chi) \wedge \mathcal{T}(p_2, \chi),$$

and universal (existential) quantifications are reduced to conjunctions (disjunctions), e.g.,

$$\mathcal{T}((\texttt{forall}\ v\ p), \chi)$$

is defined as the conjunction of all $\mathcal{T}(p, \chi[v \mapsto s_i])$ for $i = 1, \ldots, N$.

For a knowledge base β and mapping χ, $\mathcal{T}(\beta, \chi) = \{\mathcal{T}(p, \chi) \mid p \in \beta\}$. Note that the size of the clause set for base relations is $O(G^3)$, where G is the size of the grid (i.e., $R \cdot C$).

Writing **Sat**[S] and **UnSat**[S] to mean that S is satisfiable and unsatisfiable, respectively, we have:

Theorem 1. *Pick an arbitrary χ. Then:*

$(\beta, \sigma) \models p$ *iff* **UnSat**$[\{\mathcal{A}, F_\sigma, \neg \mathcal{T}(p, \chi)\} \cup \mathcal{T}(\beta, \chi)]$;

$(\beta, \sigma) \models \sigma'$ *iff* **UnSat**$[\{\mathcal{A}, F_\sigma, \neg F_{\sigma'}\} \cup \mathcal{T}(\beta, \chi)]$;

(β, σ) *has a model iff* **Sat**$[\{\mathcal{A}, F_\sigma\} \cup \mathcal{T}(\beta, \chi)]$.

(We assume without loss of generality that sentences in the above theorem are closed, i.e., have no free variables. If one does, we can consider its universal closure instead, since the semantics ensure that the two are equivalent.) This result completes the reduction of *CDC* to SAT, and enables us to determine whether an arbitrary sentence or diagram follows from the current context. For a given inference problem, our implementation carries out this propositional encoding and then proceeds in three stages. First it translates the produced formulas into CNF; it then translates the CNF into DIMACS format; and it finally invokes a SAT solver on the DIMACS input (currently RSat).

It should be noted that the reduction to SAT can be carried out automatically not just for *CDC* but for *any* QSR system adhering to the approach we have outlined in this paper, given the interpretations of its relation symbols. We sketch out the relevant technique below. Suppose for simplicity that there is only one attribute (this is not an essential restriction), which, in the context of a specific inference problem, can only take values from a finite set A. Then we introduce variables of the form *att-i-v* for $i = 1, \ldots, N$, $v \in A$. Now let $\overline{R}$ be an interpreted relation of arity $k + 1$, $k \geq 0$. Given values $v_1, \ldots, v_k \in A$, define the *projection* $\overline{R} \downarrow (v_1, \ldots, v_k)$ as follows:

$$\overline{R} \downarrow (v_1, \ldots, v_k) = \{ v \in A \mid \overline{R}(v_1, \ldots, v_k, v) \}.$$

Then we define $R^{\mathcal{T}}(s_{i_1}, \ldots, s_{i_k})$ as follows:

$$\bigwedge_{(v_1, \ldots, v_k) \in A^k} \{ att\text{-}i_1\text{-}v_1, \ldots, att\text{-}i_k\text{-}v_k \} \Rightarrow$$

$$\left[\bigvee_{v \in \overline{R} \downarrow (v_1, \ldots, v_k)} att\text{-}i_{k+1}\text{-}v \right]$$

This is essentially the same general scheme that was used in the translation of *CDC*. (Derived attributes such as *row* and *col* are convenient but inessential abbreviations.)

5 COMPARISON WITH PREVIOUS APPROACHES AND CONCLUSIONS

Historically, most of the work in QSR has stemmed from and was heavily influenced by Allen's calculus (Allen, 1983). Although some important early work was couched in first-order logic (Randell et al., 1992), by and large, following Allen it has been widely thought that an expressive reasoning framework for QSR powered by a general-purpose inference procedure would be infeasible. Accordingly, expressivity and reasoning completeness have been sacrificed in the interest of efficiency. With few exceptions, QSR systems are couched as relation algebras, and reasoning in such systems is performed by CSP techniques on networks of objects constrained by binary (or occasionally ternary) base relations.

In the wake of the remarkable progress that has been achieved in SAT-solving technology over the last decade, this approach has become questionable. The general-purpose reasoning provided by off-the-shelf SAT-solvers is now powering systems that solve extremely demanding problems, not only in hardware and software verification, but in AI as well (e.g., for planning and scheduling). That QSR could also stand to benefit from this progress is suggested by the following observation: The reasoning required in many practical QSR applications is model-based, dealing with a finite set of objects (regions, points, lines, time intervals, or arbitrary objects in a scene), each having a finite number of possible spatial-attribute values. Therefore, one can retain first-order logic and still achieve decidability through propositionalization, by restricting the universe of discourse to the set of objects in question and then deciding entailment through off-the-shelf SAT solvers.

By comparison to the CSP tradition, the approach we have suggested in this paper can offer the following advantages:

1. Increased expressivity: The full expressive power of first-order logic is available, allowing for much more natural modeling of spatial information. Anything that could be modeled with relational constraints can be expressed in first-order logic, but the converse is not true. Many problems that could not be solved—or even expressed—in

pure constraint-based calculi can be directly formulated and solved in the present setting. The furniture-arrangement problem from section 1, for instance, is beyond the reach of current QSR systems, but it is readily formulated and solved in *CDC*.

2. Higher level of abstraction: In the present approach there is no need to compute transitivity tables or to devise or modify path-consistency algorithms. These are laborious processes—often left unfinished for many systems–that are necessitated largely by the idiosyncrasies of the underlying reasoning mechanism. When defining a QSR system in our approach, one can focus on the purely logical aspects of the primitive relations and relegate the reasoning to the SAT solver. It is also not necessary to require the primitive relations of the system to be JEPD (jointly exhaustive and pairwise disjoint), a requirement that can have somewhat awkward modeling consequences (see footnote 1).

3. Built-in mechanisms for dealing with incomplete spatial knowledge: The semantics of the present framework are based on an intuitive new 3-valued logic that is particularly apt for modeling incomplete spatial information. We have shown how to compile these semantics into propositional logic.

4. Extensibility: New dimensions of spatial representation and reasoning can be incorporated with relatively little effort. The relative-orientation primitives of the flip-flop calculus, for instance, were added to the cardinal-direction primitives of Frank's calculus in less than two hours. By contrast, combining these two systems in a constraint-based algebraic setting was a major research challenge that by itself merited publication (Isli et al., 2001). Similar systems could be implemented for, e.g., topological inference.

5. Orthogonal efficiency improvements: Progress in SAT-solving technology is rapid, and should translate into corresponding efficiency gains for SAT-based QSR systems.

6. Prominent role for diagrams: Diagrams play a crucial role in spatial cognition, but so far they have been largely absent from QSR systems, which are usually entirely algebraic, even though QSR is recognized as "especially suited for applications that involve interaction with humans, as they provide an interface based on human spatial concepts" (Wallgrün et al., 2006, p. 39). The system we have presented can accept diagrammatic input, including incompletely specified diagrams, and can also present output diagrammat-

ically. Moreover, the underlying framework provides a general formal notion of diagrams.

7. Heterogeneous proofs: In addition to automating reasoning tasks such as model-finding and theorem-proving, the present framework allows for proofs that express spatial reasoning.[5] These are given in a heterogeneous framework that is specifically designed to combine visual and symbolic reasoning. None of the present systems allow for proofs, let alone heterogeneous proofs. Nevertheless, proofs are not only interesting in their own right, but they could also play an important role in human-machine interaction, since they can serve as *explanations* of spatial reasoning.

In the near future we plan to integrate additional spatial primitives (particularly topological ones); improve the efficiency of the SAT encoding; pursue additional optimizations (e.g., cache the canonicity clauses after the first translation, instead of reencoding them on every query); and evaluate the system's performance on a wider range of problems.

REFERENCES

Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843.

Arkoudas, K. and Bringsjord, S. (2009). Vivid: An AI Framework for Heterogeneous Problem Solving. *Artificial Intelligence*. In press, doi:10.1016/j.artint.2009.06.002.

Frank, A. U. (1991). Qualitative Spatial Reasoning with Cardinal Directions. In Kaindl, H., editor, *Proc. 7th Austrian Conference on Artificial Intelligence*, pages 157–167. Springer.

Freksa, C. (1992). Using Orientation Information for Qualitative Spatial Reasoning. In Frank, A. U., Campari, I., and Formentini, U., editors, *Theories and methods of spatio-temporal reasoning in geographic space*, pages 162–178. Springer.

Isli, A., Haarslev, V., and Möller, R. (2001). Combining cardinal direction relations and relative orientation relations in Qualitative Spatial Reasoning. Technical Report FBI-HH-M-304/01, Fachbereich Informatik, Universitöt Hamburg.

Ligozat, G. F. (1993). Qualitative triangulation for spatial reasoning. In Frank, A. U. and Campari, I., editors, *Spatial Information Theory. A Theoretical Basis for GIS. European Conference, COSIT'93*, volume 716 of *LNCS*, pages 54–68. Springer.

Randell, D. A., Cui, Z., and Cohn, A. G. (1992). A spatial logic based on regions and connection. In *Proceedings of the Third International Conference on Knowledge Representation and Reasoning*, pages 165–176.

Röhrig, R. (1997). Representation and Processing of Qualitative Orientation Knowledge. In *Proceedings of the 21st Annual German Conference on Artificial Intelligence*, volume 1303 of *LNCS*, pages 219–230.

Wallgrün, J. O., Frommberger, L., Wolter, D., Dylla, F., and Freksa, C. (2006). Qualitative Spatial Representation and Reasoning in the SparQ-Toolbox. In Barkowsky, T., Knauff, M., Ligozat, G., and Montello, D. R., editors, *Spatial Cognition*, volume 4387 of *Lecture Notes in Computer Science*, pages 39–58. Springer.

---

[5]We have not had space to cover that aspect of the system here, but we discuss it in detail elsewhere (Arkoudas and Bringsjord, 2009).