The Computational Complexity of Topological Logics

Ian Pratt-Hartmann
University of Manchester

(Joint work with Roman Kontchakov and Michael Zakharyaschev)

CSE Departmental Colloquium
University at Buffalo, SUNY
May, 2010
• A **spatial logic** is a formal language with
  – variables ranging over ‘geometrical entities’
  – non-logical primitives denoting relations and operations defined over those geometrical entities.

• Any spatial logic is thus characterized by by three parameters:
  – a **logical syntax**:
    propositional logic, FOL, higher-order logic …
  – a signature of **non-logical (geometrical) primitives**:
    \( \text{conv}(x) \), \( c(x) \), \( C'(x, y) \), …
  – a **class of interpretations** (more on this below).

• A **topological logic** is a spatial logic whose non-logical primitives are all topological in character.
- Probably the best-known topological logic is the ‘RCC8’ language (Randall, Cui and Cohn, 1992), (Egenhofer 1991)

<table>
<thead>
<tr>
<th>DC($r_1, r_2$)</th>
<th>EC($r_1, r_2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PO($r_1, r_2$)</td>
<td>EQ($r_1, r_2$)</td>
</tr>
<tr>
<td>TPP($r_1, r_2$)</td>
<td>NTPP($r_1, r_2$)</td>
</tr>
</tbody>
</table>

- Example of a formula in this logic:

\[
(TPP(r_1, r_2) \land NTPP(r_1, r_3)) \rightarrow (PO(r_2, r_3) \lor TPP(r_2, r_3) \lor NTPP(r_2, r_3)).
\]
• If $X$ is a topological space, a frame on $X$ is a pair $(X, \mathcal{R})$, where $\mathcal{R}$ is a (non-empty) collection of subsets of $X$—called regions.

• For example, we can consider the frame $(X, \text{RC}(X))$ of regular closed sets in $X$. (A regular closed set is the closure of an open set).

• If $X$ is a topological space, $\text{RC}(X)$ is a Boolean algebra under natural operations:

\[
\begin{align*}
    r_1 + r_2 &= r_1 \cup r_2 \\
    r_1 \cdot r_2 &= \text{cl}(\text{int}(r_1 \cap r_2)) \\
    -r_1 &= \text{cl}(\text{cmp}(r_1))
\end{align*}
\]

So frames of the form $\text{RC}(X)$ are natural structures over which to interpret topological logics.
• Denote the class of frames \( \{(X, \text{RC}(X)) \mid X \text{ a topological space}\} \) by \( \text{REGC} \).

• And given an assignment of variables to regions of a frame in \( \text{REGC} \), the \( \text{RCC8} \)-primitives have natural formal interpretations:

\[
\begin{align*}
\text{DC}(r_1, r_2) & \quad \text{iff} \quad r_1 \cap r_2 = \emptyset \\
\text{TPP}(r_1, r_2) & \quad \text{iff} \quad r_1 \subseteq r_2 \text{ but } r_1 \nsubseteq \text{int}(r_2) \\
\text{NTPP}(r_1, r_2) & \quad \text{iff} \quad r_1 \subseteq \text{int}(r_2) \\
\ldots & \quad \ldots \quad \ldots
\end{align*}
\]

• This gives us notions of \textit{satisfiability} and \textit{validity} for formulas, with respect to either frames or, more generally, \textit{classes of frames}.

• We denote the satisfiability problem for \( \text{RCC8} \)-formulas over a frame-class \( \mathcal{K} \) by \( \text{Sat}(\text{RCC8}, \mathcal{K}) \).
• For example,

\[ \neg (\text{TPP}(r_1, r_2) \land \text{NTPP}(r_1, r_3)) \rightarrow \]
\[ (\text{PO}(r_2, r_3) \lor \text{TPP}(r_2, r_3) \lor \text{NTPP}(r_2, r_3)) \]

is not satisfiable over \text{REGC}.

\begin{center}
\begin{tikzpicture}
  \node (r1) at (0,0) {$r_1$};
  \node (r2) at (0,-3) {$r_2$};
  \node (r3) at (2,0) {$r_3$};
  \draw (r1) ellipse (1 and 0.5);
  \draw (r3) ellipse (2 and 1);
  \draw (r2) ellipse (1 and 0.5);
  \draw (r1) -- (r2);
  \draw (r2) -- (r3);
\end{tikzpicture}
\end{center}

• We can also interpret $\mathcal{RCC}_8$-formulas over smaller frame-classes:
  e.g.

\[ \text{RC}(\mathbb{R}), \quad \text{RC}(\mathbb{R}^2), \quad \text{RC}(\mathbb{R}), \quad \{\text{RC}(\mathbb{R}^n) \mid n \geq 1\}, \ldots \]

However, this makes no difference to the satisfiability/validity problem: $\text{Sat}(\mathcal{RCC}_8, \text{REGC}) = \text{Sat}(\mathcal{RCC}_8, \text{RC}(\mathbb{R}^n))$ for all $n \geq 1$. 
• Some simple facts:

**Theorem 1** \((\approx \text{Renz 1998})\). *The problem \text{Sat}(\text{RCC8}, \text{REGC}) is NP-complete.* Indeed, for any \(n \geq 0\),

\[
\text{Sat}(\text{RCC8}, \text{RC}(\mathbb{R}^n)) = \text{Sat}(\text{RCC8}, \text{REGC}).
\]

• Actually, by restricting the language somewhat, we get better complexities:
  
  – if we consider only conjunctions of \(\text{RCC8}\)-primitives, complexity of satisfiability goes down to \(\text{NLOGSPACE}\)
  
  – Various (larger) tractable fragments have been found (Nebel and Bürckert 1995), (Renz 1999), \ldots,

• Warning:

  [Regions need not be connected.]
• Now suppose we add $+, \cdot, -, 0$ and $1$ to $\mathcal{RCC}8$, yielding the language $\mathcal{BRCC}8$ (Wolter and Zakharyaschev, 2000), thus:

$$EC(r_1 + r_2, r_3) \rightarrow (EC(r_1, r_3) \lor EC(r_2, r_3)).$$

• But now, we can replace the $\mathcal{RCC}8$-predicates with the binary relations of equality ($=$) and contact:

$$C(r_1, r_2) \text{ iff } r_1 \cap r_2 = \emptyset.$$  

thus:

$$DC(r_1, r_2) \equiv \neg C(r_1, r_2)$$

$$TPP(r_1, r_2) \equiv r_1 \cdot -r_2 = 0 \land C(r_1, -r_2)$$

$$NTPP(r_1, r_2) \equiv r_1 \neq 0 \land \neg C(r_1, -r_2)$$

... ... ...

• For this reason, the language is now called, simply, $\mathcal{C}$. 
Some more simple facts:

**Theorem 2** (Wolter and Zakharyaschev, 2000). *The problem* \( \text{Sat}(C, \text{REGC}) \) *is NP-complete. For any* \( n \geq 1 \), *the problem* \( \text{Sat}(C, \text{RC}(\mathbb{R}^n)) \) *is PSPACE-complete.*

- The critical difference here is that the spaces \( \mathbb{R}^n \) are **connected**. (The PSPACE-hardness result generally applies when \( C \) is interpreted over the class of regular closed algebras of connected topological spaces.)

- Logics which cannot express the property of connectedness are of limited interest. So let’s add it!
• We employ a unary predicate $c$ with the semantics

\[ c(r) \text{ iff } r \text{ is connected} \]

• We consider the languages

  – $\mathcal{RCC8c}$: RCC8 plus the unary predicate $c$;
  – $\mathcal{Cc}$: W+Z’s language (i.e. $C$, $+$, $\cdot$, $-$, 0, 1) plus the unary predicate $c$;
  – $\mathcal{Bc}$: like $\mathcal{C}$, but without $C$.

• Example of an $\mathcal{RCC8c}$-formula in the 3 variables $r_1, r_2, r_3$:

\[ \bigwedge_{1 \leq i \leq 3} c(r_i) \land \bigwedge_{1 \leq i < j \leq 3} \text{EC}(r_i, r_j). \]
• Adding the predicate $c$ makes the logic much more sensitive to the underlying space.

• Example:

\[
\bigwedge_{1 \leq i \leq 3} c(r_i) \wedge \bigwedge_{1 \leq i < j \leq 3} \text{EC}(r_i, r_j)
\]

is not satisfiable in $\text{RC}(\mathbb{R})$ (because any realizing assignment would make $r_1$, $r_2$ and $r_3$ intervals); but it is satisfiable in $\text{RC}(\mathbb{R}^n)$ for $n \geq 2$.

• Example:

\[
\bigwedge_{1 \leq i < j \leq 5} c(r_{i,j}) \wedge \bigwedge_{\{i,j\} \cap \{k,\ell\}=\emptyset} \text{DC}(r_{i,j}, r_{k,\ell}) \wedge \bigwedge_{i \in \{j,k\}} \text{TPP}(r_i, r_{j,k})
\]

is not satisfiable in $\text{RC}(\mathbb{R}^2)$ (because any realizing assignment would induce a plane embedding of $K_5$); but it is satisfiable in $\text{RC}(\mathbb{R}^n)$ for $n \geq 3$. 

Various complexity results are known here

**Theorem 3** (Kontchakov, P-H, W+Z, forthcoming).

\[ \text{Sat}(\text{RCC8c, REGC}) \text{ is NP-complete (trivial);} \]
\[ \text{Sat}(\text{Cc, REGC}) \text{ is EXP\text{T}IME-complete;} \]
\[ \text{Sat}(\text{Bc, REGC}) \text{ is EXP\text{T}IME-complete.} \]

**Theorem 4.**

\[ \text{Sat}(\text{RCC8c, RC(}\mathbb{R}^n)) \text{ is NP-complete (n \geq 1)*;} \]
\[ \text{Sat}(\text{Bc, RC(}\mathbb{R})) \text{ is NP-complete ;} \]
\[ \text{Sat}(\text{Cc, RC(}\mathbb{R})) \text{ is PS\text{PACE}-complete;} \]
\[ \text{Sat}(\text{Bc, RC(}\mathbb{R}^n)) \text{ is EXP\text{T}IME-hard (n \geq 2);} \]
\[ \text{Sat}(\text{Cc, RC(}\mathbb{R}^n)) \text{ is EXP\text{T}IME-hard (n \geq 2).} \]

* Membership of \( \text{Sat}(\text{RCC8c, RC(}\mathbb{R}^2)) \) in NP is highly non-trivial (Shaefer, Sedgwick and Štefankovič).
• We may wish to distinguish between *connectedness* and *interior connectedness*:

![Diagram showing the distinction between connectedness and interior connectedness]

• We employ a unary predicate $c^\circ$ with the semantics

\[ c^\circ(r) \text{ iff } \text{int}(r) \text{ is connected} \]

• This gives us the further languages $\mathcal{RCC8c^\circ}$, $\mathcal{Bc^\circ}$, $\mathcal{Cc^\circ}$.

• Example of an $Cc^\circ$-formula

\[ c^\circ(\neg r_1) \land c^\circ(\neg r_2) \land \text{DC}(r_1, r_2) \land \neg c^\circ(\neg(r_1 + r_2)) \]
• The $Cc^\circ$-formula

$$c^\circ(-r_1) \land c^\circ(-r_2) \land DC(r_1, r_2) \land \neg c^\circ(-(r_1 + r_2))$$

is satisfiable over $\text{REGC}$, thus:

But it is not satisfiable over $\text{RC}(\mathbb{R}^n)$ for any $n$!

**Theorem 5.**

$Sat(\text{RCC8}c^\circ, \text{RC}(\mathbb{R}^n))$ is NP-complete ($n \geq 2$);

$Sat(Cc^\circ, \text{RC}(\mathbb{R}^n))$ is EXPTIME-hard ($n \geq 2$);*

$Sat(\text{Bc}^\circ, \text{RC}(\mathbb{R}^n))$ is NP-complete ($n \geq 3$).
• Actually, matters are even more delicate than this: $RC(\mathbb{R}^n)$ contains some very pathological sets:

• This prompts us to consider interpretations of spatial logics over collections of *tame* regions.

• Natural candidates for tame subalgebras of $RC(\mathbb{R}^n)$:
  – The regular closed *polyhedra* in $\mathbb{R}^n$, $RCP(\mathbb{R}^n)$:

  – The regular closed *semi-algebraic* subsets of $\mathbb{R}^n$, $RCS(\mathbb{R}^n)$.
• We consider first logics interpreted over 1-dimensional space.

• Consider the $\mathcal{RCC}8c$-formula

$$c(r_1) \land \bigwedge_{1 \leq i < j \leq 4} \text{EC}(r_i, r_j).$$

• This formula is satisfiable over $\text{RC}(\mathbb{R})$:

• But the only satisfying tuples are those in which some of the members have infinitely many components.

• That is, the formula is not satisfiable over $\text{RCP}(\mathbb{R})$. 
• Thus, we have shown:

\[ \text{Sat}(\text{RCC8c}, \text{RC}(\mathbb{R})) \neq \text{Sat}(\text{RCC8c}, \text{RCP}(\mathbb{R})) \]
\[ \text{Sat}(Cc, \text{RC}(\mathbb{R})) \neq \text{Sat}(Cc, \text{RCP}(\mathbb{R})). \]

• These problems do, however, have the same complexity:

**Theorem 6.**

\[ \text{Sat}(\text{RCC8c}, \text{RCP}(\mathbb{R})) \text{ is } \text{NP-complete}; \]
\[ \text{Sat}(Cc, \text{RCP}(\mathbb{R})) \text{ is } \text{PSPACE-complete}. \]

• On the other hand:

**Theorem 7.** \( \text{Sat}(Bc, \text{RCP}(\mathbb{R})) = \text{Sat}(Bc, \text{RC}(\mathbb{R})), \text{ and hence is } \text{NP-complete}. \)
• In two dimensions, we get a different pattern of sensitivity to tameness:

• For example, the $\mathcal{B}c^\circ$-formula

$$\bigwedge_{1 \leq i \leq 3} c^\circ(r_i) \land c^\circ\left(\sum_{1 \leq i \leq 3} r_i\right) \land \neg(c^\circ(r_1 + r_2) \lor c^\circ(r_1 + r_3))$$

is satisfiable over $\text{RC}(\mathbb{R}^2)$, thus,

but is unsatisfiable over $\text{RCP}(\mathbb{R}^2)$. 
• Thus, we have shown:

\[ \text{Sat}(Bc^\circ, \text{RC}(\mathbb{R}^2)) \neq \text{Sat}(Bc^\circ, \text{RCP}(\mathbb{R}^2)) \]

\[ \text{Sat}(Cc^\circ, \text{RC}(\mathbb{R}^2)) \neq \text{Sat}(Cc^\circ, \text{RCP}(\mathbb{R}^2)). \]

• Similarly (via a more elaborate construction):

\[ \text{Sat}(Bc, \text{RC}(\mathbb{R}^2)) \neq \text{Sat}(Bc, \text{RCP}(\mathbb{R}^2)) \]

\[ \text{Sat}(Cc, \text{RC}(\mathbb{R}^2)) \neq \text{Sat}(Cc, \text{RCP}(\mathbb{R}^2)). \]

**Theorem 8.**

\[ \text{Sat}(\text{RCC8c}\{\circ\}, \text{RCP}(\mathbb{R}^2)) = \text{Sat}(\text{RCC8c}\{\circ\}, \text{RC}(\mathbb{R}^2)). \]

**Theorem 9.**

\[ \text{Sat}(Bc, \text{RCP}(\mathbb{R}^n)) \text{ is EXPTime-hard } (n \geq 2); \]

\[ \text{Sat}(Cc, \text{RCP}(\mathbb{R}^n)) \text{ is EXPTime-hard } (n \geq 2); \]

\[ \text{Sat}(Cc^\circ, \text{RCP}(\mathbb{R}^n)) \text{ is EXPTime-hard } (n \geq 2); \]

\[ \text{Sat}(Bc^\circ, \text{RCP}(\mathbb{R}^2)) \text{ is EXPTime-hard; } \]

\[ \text{Sat}(Bc^\circ, \text{RCP}(\mathbb{R}^n)) \text{ is EXPTime-complete } (n \geq 3). \]
Conclusions

• We have explained what a topological logic (more generally, a spatial logic) is.

• We have reviewed some well-known results on RCC8, and considered the effect of adding connectedness predicates.

• We showed that even the simplest logics with connectedness:
  – are sensitive to the underlying space;
  – exhibit complex patterns of sensitivity to tameness in Euclidean spaces of different dimension;
  – reveal a complicated (and still, to some extent uncharted) complexity-theoretic landscape.

• The authoritative reference for more results on spatial logics