Qualitative Representation and Reasoning

Spatial Representation and Reasoning: RCC8 and Topology

Knowledge Representation and Reasoning

January 30, 2006

RCC8 and Topology – Outline

Motivation

We may want to state qualitative relationships between regions in space, for example:

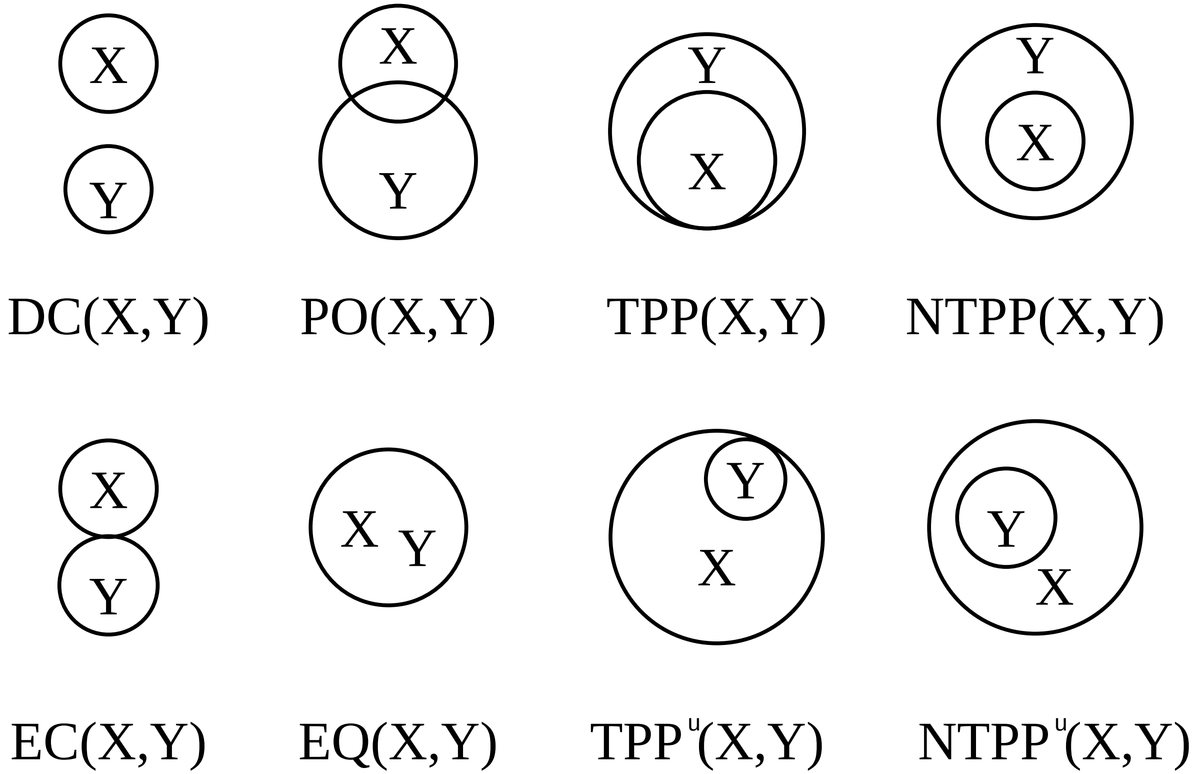
- “Region X touches region Y”
- “Germany and Switzerland have a common border”
- “Freiburg is located in Baden-Württemberg”

Possible Applications

- This can be useful when only *partial information* is available:
  - We may know that region X is not connected with region Y without knowing the shape and location of X and Y.
- We may want to *query* a database:
  - Show me all countries bordering the Mediterranean!
- We may want to state *integrity constraints*:
  - An island has to be located in the interior of a sea.

Qualitative Relations Between Regions: RCC8

Eight relations between regions:



Intuition

- Regions are some *“reasonable” non-empty subsets of space*.
- DC (disconnected) means that the two regions do not share any point at all.
- EC (externally connected) means that they only share *borders*.
- PO (partially overlapping) means that the two regions share *interior* points.
- TPP (tangential proper part) means that one region is a subset of the other sharing some points on the *borders*.
- NTPP (non-tangential proper part) same, but without sharing any bordering points.

Questions

- How can we formalize regions?
- How can we formalize these relations?
- Are they disjoint and exhaustive?
- Can we come up with a composition table?
- What is the computational complexity of reasoning with these relations?
- Can we identify a tractable fragment?

Point-Set Topology

Point-set topology is a mathematical theory that deals with properties of space independent of size and shape.

In topology, we can define notions such as

- interior and exterior points of regions,
- isolated points of regions,
- boundaries of regions,
- connected components of regions,
- connected regions,
- ...

$\rightsquigarrow$ Topology seems to be the right formal framework.

Topology

Definition

A *topological space* is a pair $\mathcal{T} = (\mathcal{U}, \mathcal{O})$, where

- $\mathcal{U}$ is a nonempty set (the universe), and
- $\mathcal{O}$ is a set of subsets of $\mathcal{U}$ (the *open sets*)

such that the following conditions hold:

- $\emptyset \in \mathcal{O}$ and $\mathcal{U} \in \mathcal{O}$.
- If $O_1 \in \mathcal{O}$ and $O_2 \in \mathcal{O}$, then $O_1 \cap O_2 \in \mathcal{O}$.
- If $(O_i)_{i \in I}$ is a (possibly infinite) family of elements from $\mathcal{O}$, then

$$\bigcup_{i \in I} O_i \in \mathcal{O}.$$

Example: In Euclidian space, a set O is open if for each point $x \in O$ there is a ball surrounding x that is contained in O.

Terminology & Notation

Definition

- A set $N \subseteq \mathcal{U}$ is a *neighborhood* of a point x if there is an open set $O \in \mathcal{O}$ such that $x \in O \subseteq N$. Let $X \subseteq \mathcal{U}$ and $x \in \mathcal{U}$.
- $x \in \mathcal{U}$ is an *interior point* of X if there is a neighborhood N of x such that $N \subseteq X$.
- $x \in \mathcal{U}$ is a *touching point* of X if every neighborhood of x has a nonempty intersection with X.
- $x \in \mathcal{U}$ is a boundary point of X if x is a touching point of X and of its complement $\overline{X}$.

Notation:

- $i(X)$ is the set of interior points of X (the *interior* of X).
- $cl(X)$ is the set of touching points of X (the *closure* of X).
- $bd(X)$ is the set of boundary points of X.
- A set is *closed* if $X = cl(X)$.

Interior, Boundary, and Closure Operators

The function $i(\cdot)$ is an interior operator:

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

Note:

- X is *open* iff $X = i(X)$
- $cl(X) = \overline{i(\overline{X})}$
- $bd(X) = cl(X) \cap cl(\overline{X})$

From Interior Operators to Topologies and Back

Theorem

Let $\mathcal{U}$ be a set and $i\colon 2^{\mathcal{U}} \to 2^{\mathcal{U}}$ be an "interior operator". Define

$$\mathcal{O} := \{O \subseteq \mathcal{U} \mid O = i(O)\}.$$

Then $\mathcal{T} = (\mathcal{U}, \mathcal{O})$ is a topological space.

Proof.

Since $i(\mathcal{U}) = \mathcal{U}$ by (1), we have $\mathcal{U} \in \mathcal{O}$. Since $i(\emptyset) \subseteq \emptyset$ by (3), we have $i(\emptyset) = \emptyset$, and therefore $\emptyset \in \mathcal{O}$.

By (2), $\mathcal{O}$ is closed under pairwise intersection. From (2), it follows that $X \subseteq Y$ implies $i(X) \subseteq i(Y)$ (which we need below).

Let $O := \bigcup_{i \in I} O_i$, $O_i = i(O_i)$ for all i. Of course, $i(O) \subseteq O$. Clearly, $O_i \subseteq O$ for all i. Then $O_i = i(O_i) \subseteq i(O)$. Therefore $O = \bigcup_{i \in I} O_i \subseteq i(O)$. Hence, $O = i(O)$, i. e., $O \in \mathcal{O}$. Thus, $\mathcal{O}$ is closed under arbitrary unions. □

Topological Set Expressions and Their Interpretations

Topological set expressions describe subsets of a topological space:

$$s \longrightarrow X \mid \top \mid \bot \mid s' \sqcap s'' \mid s' \sqcup s'' \mid \bar{s} \mid \mathtt{I}s',$$

with set variables X, Y, Z.

Definition

A *topological interpretation* is a tuple $I = (\mathcal{T}, d)$, where $\mathcal{T} = (\mathcal{U}, \mathcal{O})$ is a topological space with an associated interior operator i and d is a function from set variables to subsets of $\mathcal{U}$.

d is extended to topological set expressions as follows:

$$\begin{aligned} d(\bot) &= \emptyset & d(\top) &= \mathcal{U} \\ d(s \sqcap s') &= d(s) \cap d(s') & d(s \sqcup s') &= d(s) \cup d(s') \\ d(\bar{s}) &= \mathcal{U} - d(s) & d(\mathtt{I}s) &= i(d(s)) \end{aligned}$$

Topological Set Constraints

Elementary set constraints:

$$s \doteq t \quad \text{or} \quad s \not\doteq t$$

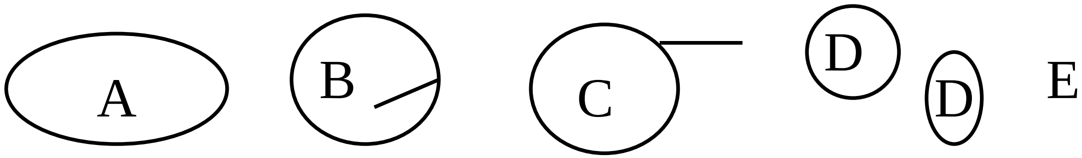
Complex set constraints: combinations using $\wedge$, $\vee$, and $\neg$.

A topological interpretation $I = (\mathcal{T}, d)$ satisfies a constraint:

$$I \models s \doteq t \quad \text{iff} \quad d(s) = d(t)$$
$$I \models s \not\doteq t \quad \text{iff} \quad d(s) \neq d(t)$$

As usual: model, satisfiability, equivalence, entailment, ...

What Kind of Regions Do We Want to Consider?



A and D are reasonable regions, B, C, and E are not

In other words, X is a region iff it is non-empty

$$X \dot{\neq} \bot$$

and regular, i. e., the closure of an open set:

$$X \doteq \overline{\mathtt{I}(\overline{\mathtt{I}X})}.$$

It is not necessary that a region is internally connected.

Applying the Topological Set Constraints to RCC8

The RCC8 relations are shorthands for topological set constraints:

$$\begin{aligned}
\mathrm{DC}(X,Y) &:= X \sqcap Y \doteq \bot \\
\mathrm{EC}(X,Y) &:= X \sqcap Y \not\doteq \bot \wedge \mathtt{I}X \sqcap \mathtt{I}Y \doteq \bot \\
\mathrm{PO}(X,Y) &:= \mathtt{I}X \sqcap \mathtt{I}Y \not\doteq \bot \wedge X \sqcap \overline{Y} \not\doteq \bot \wedge \overline{X} \sqcap Y \not\doteq \bot \\
\mathrm{EQ}(X,Y) &:= X \doteq Y \\
\mathrm{TPP}(X,Y) &:= X \sqcap \overline{Y} \doteq \bot \wedge X \sqcap \overline{\mathtt{I}Y} \not\doteq \bot \\
\mathrm{NTPP}(X,Y) &:= X \sqcap \overline{\mathtt{I}Y} \doteq \bot
\end{aligned}$$

In addition, each named region must satisfy non-emptiness and regularity.

$\rightsquigarrow$ It follows that the relations are disjoint and exhaustive.

Normal Form Constraints

- A topological set constraint is in normal form if it is $s \doteq \top$ or $s \not\doteq \top$.
- Every set constraint can be translated into normal form.
- $s \doteq t$ is equivalent to $(\bar{s} \sqcup t) \sqcap (\bar{t} \sqcup s) \doteq \top)$
  - $\mathsf{DC}(X,Y) = \overline{X} \sqcup \overline{Y} \doteq \top$
  - $\mathsf{EC}(X,Y) = \overline{X} \sqcup \overline{Y} \not\doteq \top \wedge \overline{\mathtt{I}X} \sqcup \overline{\mathtt{I}Y} \doteq \top$
  - ...

A Deduction Theorem

Notation: $s \sqsubseteq t$ stands for $\overline{s} \sqcup t \doteq \top$.

Theorem (Deduction Theorem, Nutt 99)

Let s, t be set expressions. Then

$$s \doteq \top \models t \doteq \top \text{ iff } \models \mathtt{I}s \sqsubseteq \mathtt{I}t.$$

Theorem (Convexity)

The conjunctive set constraint

$$s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top \wedge t_1 \not\doteq \top \wedge \ldots \wedge t_n \not\doteq \top$$

is satisfiable if and only if the following constraints are satisfiable for each $j \in \{1, \ldots, n\}$

$$s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top \wedge t_j \not\doteq \top.$$

Proof idea.

($\Leftarrow$) Construct models for each j and create a common model by taking

Topology and Modal Logic (1)

The modal logic S4 can be characterized by the following axiom schemata (with $\mathbf{I}$ instead of $\Box$ as the modal box operator)

- $\mathbf{I}\top \leftrightarrow \top$ (valid in all frames)
- $\mathbf{I}\varphi \rightarrow \varphi$ (valid in **T**-frames, reflexivity)
- $\mathbf{I}\varphi \wedge \mathbf{I}\psi \leftrightarrow \mathbf{I}(\varphi \wedge \psi)$ (valid in all frames)
- $\mathbf{I}\mathbf{I}\varphi \leftrightarrow \mathbf{I}\varphi$ (valid in **T4**-frames, transitivity & reflexivity)

Reminder: Interior operator

- $i(\mathcal{U}) = \mathcal{U}$
- $i(X) \subseteq X$
- $i(X) \cap i(Y) = i(X \cap Y)$
- $i(i(X)) = i(X)$

Topology and Modal Logic (2)

Define a translation function π from set expressions to S4 formulae as follows:

- $\pi(X) = X$
- $\pi(\overline{s}) = \neg\pi(s)$
- $\pi(s \sqcap t) = \pi(s) \wedge \pi(t)$
- $\pi(s \sqcup t) = \pi(s) \vee \pi(t)$
- $\pi(\mathtt{I}s) = \mathbf{I}\pi(s)$

A set expression s is called a *topological tautology* if $d(s) = \mathcal{U}$ for all topological interpretations $I = ((\mathcal{U}, \mathcal{O}), d)$.

Theorem (McKinsey & Tarski 48)

s is a topological tautology iff $\pi(s)$ is S4-valid.

Corollary

s is topologically satisfiable iff $\pi(s)$ is S4-satisfiable.

Topological Set Constraints and Modal Logic (1)

How can we use this result for conjunctive topological set constraints?

1. Using the convexity theorem, we only have to test the satisfiability of constraints of the form

$$C_j = (s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top \wedge t_j \not\doteq \top).$$

2. C_j is satisfiable iff $s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top \not\models t_j \doteq \top$. Equivalently, we can test $s \doteq \top \not\models t_j \doteq \top$, with $s = s_1 \sqcap \ldots \sqcap s_m$.

3. Using the deduction theorem, it suffices to check $\not\models \mathtt{I}s \sqsubseteq \mathtt{I}t_j$, i. e., whether $\overline{\mathtt{I}s} \sqcup \mathtt{I}t_j$ is not a tautology, i. e., whether $\mathtt{I}s \sqcap \overline{\mathtt{I}t_j}$ is satisfiable. Using the McKinsey-Tarski theorem, this amounts to test for S4-satisfiability of $\pi(\mathtt{I}s \sqcap \overline{\mathtt{I}t_j})$.

Topological Set Constraints and Modal Logic (2)

Theorem (Translation)

The formula

$$s_1 \doteq \top \wedge \ldots \wedge s_m \doteq \top \wedge t_1 \not\doteq \top \wedge \ldots \wedge t_n \not\doteq \top$$

is satisfiable if for each $j \in \{1, \ldots, n\}$*, the following formula is S4-satisfiable:*

$$\mathbf{I}\pi(s_1) \wedge \ldots \wedge \mathbf{I}\pi(s_m) \wedge \neg \mathbf{I}\pi(t_j).$$

Topological Set Constraints and Modal Logic (3)

Let $\Box$ and $\Diamond$ be K-modalities.

Proposition

Let $\varphi_1, \ldots \varphi_m, \psi_1, \ldots, \psi_n$ be multi-modal formulae not containing the K-operators $\Box$ and $\Diamond$. Then

$$\Box\varphi_1 \wedge \ldots \wedge \Box\varphi_m \wedge \Diamond\psi_1 \wedge \ldots \wedge \Diamond\psi_n$$

is satisfiable iff for all $j \in \{1, \ldots, n\}$ the formulae

$$\varphi_1 \wedge \ldots \wedge \varphi_m \wedge \psi_j$$

are satisfiable.

Proof idea.

Create from models satisfying the later formula a modal interpretation for the former formula. $\Box$

New Translation

Use a multi-modal logic for the translation. Extend π as follows:

- $\pi(s \doteq \top) = \Box \mathbf{I} \pi(s)$
- $\pi(s \dot{\neq} \top) = \Diamond \neg \mathbf{I} \pi(s)$
- $\pi(C_1 \wedge C_2) = \pi(C_1) \wedge \pi(C_2)$
- ...

This leads to the following translation of RCC8 constraints:

- $\pi(\mathsf{DC}(X,Y)) = \Box \mathbf{I} \neg (X \wedge Y)$
- $\pi(\mathsf{EC}(X,Y)) = \Box \mathbf{I} \neg (\mathbf{I}X \wedge \mathbf{I}Y) \wedge \Diamond \neg \mathbf{I} \neg (X \wedge Y)$
- ...

Theorem (Translation)

Let C be an arbitrary topological set constraint. Then C is satisfiable iff $\pi(C)$ is satisfiable.

Outlook

- We wanted to state qualitative relationships between spatial regions
- Semantics: Topology
- Language for describing relations: Topological set constraints
- . . . can be translated to modal logic (McKinsey & Tarski)
- Combination can be handled with another modality

$\rightsquigarrow$ Reasoning in RCC8?

$\rightsquigarrow$ Complexity?

Reasoning with RCC8

Reminder

- RCC8 is a relation calculus for expressing spatial/topological information.
- Topology is the right mathematical theory to give meaning to the RCC8 relations.
- Topological set constraints can be used to characterize the relations.
- Validity of topological set expressions is equivalent to S4-validity of translations of these set expressions.
- Using an additional K-modality, satisfiability of the topological set constraints can be tested

$\rightsquigarrow$ Reduction of spatial reasoning problems to modal logic reasoning problems.

What is the Role of the Dimension?

- Already mentioned: McKinsey & Tarski do not mention dimension at all.

⇝ It has been shown:

  - If an RCC8 CSP is topologically satisfiable, then it is satisfiable in 2 dimensions, provided we do not require regions to be internally connected.
  - If we require regions to be internally connected in 2 dimensions, then the problem is open.
  - If we allow 3 dimensions, then the issue whether regions are internally connected is not crucial anymore.

Upper Bounds – Using Results From Modal Logic

- Satisfiability in most modal logics (incl. K and S4 and multi-modal logics using these modalities) is PSPACE-complete.
- Upper bound from tableaux proofs: It suffices to explore one branch at a time in a depth-first manner and the depth is bounded polynomially by the size of the formula.
- Lower bound from reduction from QBF (quantified boolean formula).
- Deciding the satisfiability of topological set constraints is in PSPACE.

A Better Upper Bound for RCC8

- Since the topological set constraints (and hence the modal formulae) resulting from RCC8 constraints are very restricted, there might be hope that we can do better than PSPACE.
- Consider the nesting depth of **I** in modal formulae resulting from RCC8-formulae.
- Satisfiability of S4-formulae with a fixed nesting depth is NP-complete.

$\rightsquigarrow$ Guess a base relation (for each non-base relation) and then guess a satisfying interpretation.

Proposition

RCC8 satisfiability is in NP.

Constraint Propagation

- As in Allen's interval algebra, we may want to use constraint propagation instead of translating everything to modal logic.
- We need a composition table . . .
- . . . which could be computed using the modal logic encoding (and in fact, this has been done).
- Based on this table, we can then apply the path-consistency algorithm
- . . . and ask ourselves for which fragment of RCC8 it is complete.

Composition Table

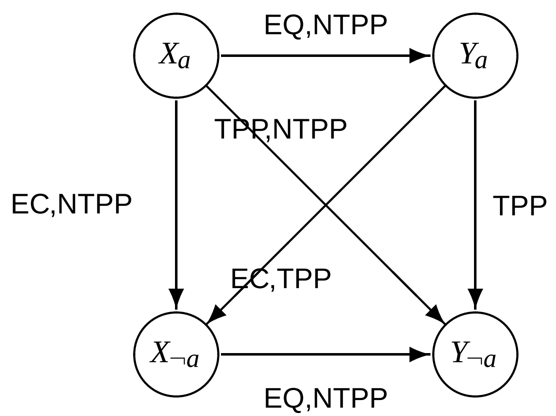
∘	DC	EC	PO	TPP	NTPP	TPP^{-1}	$NTPP^{-1}$	EQ
DC	*	DC,EC PO,TPP NTPP	DC,EC PO,TPP NTPP	DC,EC PO,TPP NTPP	DC,EC PO,TPP NTPP	DC	DC	DC
EC	DC,EC PO,TPP^{-1} $NTPP^{-1}$	DC,EC PO,TPP TPP^{-1},EQ	DC,EC PO,TPP NTPP	EC,PO TPP NTPP	PO TPP NTPP	DC,EC	DC	EC
PO	DC,EC PO,TPP^{-1} $NTPP^{-1}$	DC,EC PO,TPP^{-1} $NTPP^{-1}$	*	PO TPP NTPP	PO TPP NTPP	DC,EC PO, TPP^{-1} $NTPP^{-1}$	DC,EC PO,TPP^{-1} $NTPP^{-1}$	PO
TPP	DC	DC,EC	DC,EC PO,TPP NTPP	TPP NTPP	NTPP	DC,EC PO,TPP TPP^{-1},EQ	DC,EC PO,TPP^{-1} $NTPP^{-1}$	TPP
NTPP	DC	DC	DC,EC PO,TPP NTPP	NTPP	NTPP	DC,EC PO,TPP NTPP	*	NTPP
TPP^{-1}	DC,EC PO,TPP^{-1} $NTPP^{-1}$	EC,PO TPP^{-1} $NTPP^{-1}$	PO TPP^{-1} $NTPP^{-1}$	PO,EQ TPP TPP^{-1}	PO TPP NTPP	TPP^{-1} $NTPP^{-1}$	$NTPP^{-1}$	TPP^{-1}
$NTPP^{-1}$	DC,EC PO,TPP^{-1} $NTPP^{-1}$	PO TPP^{-1} $NTPP^{-1}$	PO TPP^{-1} $NTPP^{-1}$	PO TPP^{-1} $NTPP^{-1}$	PO,TPP^{-1} TPP,NTPP $NTPP^{-1}$,EQ	$NTPP^{-1}$	$NTPP^{-1}$	$NTPP^{-1}$
EQ	DC	EC	PO	TPP	NTPP	TPP^{-1}	$NTPP^{-1}$	EQ

Lower Bound: Proving NP-Hardness

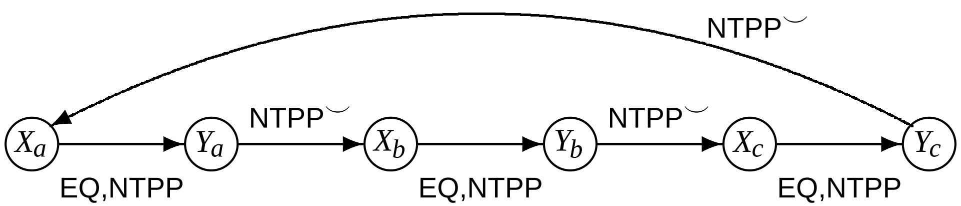
- Idea: Reduction from 3-SAT
- 3-SAT structure
  1. Literals a, b, c: can be true or false
  2. Complementary literals: a is true iff $\neg a$ is false
  3. Clauses $l_1 \vee l_2 \vee l_3$: at least one literal must be true
- RCC8-CSP
  1. Truth value constraints $X_a\{R_t, R_f\}Y_a$: Either $X_a\{R_t\}Y_a$ or $X_a\{R_f\}Y_a$ holds
  2. Polarity constraints: $X_a\{R_t\}Y_a$ holds iff $X_{\neg a}\{R_f\}Y_{\neg a}$ holds
  3. Clause constraints: At least one of $X_{l_1}\{R_t\}Y_{l_1}$, $X_{l_2}\{R_t\}Y_{l_2}$, or $X_{l_3}\{R_t\}Y_{l_3}$ holds

The Reduction

- Relations: $R_t = NTPP$, $R_f = EQ$
- Polarity constraints:



- Clause constraints:



- RCC8 sat.⇒3-SAT: follows from reduction
- 3-SAT⇒RCC8 sat.: Construction of model for Θ_ϕ for each positive 3-SAT instance ϕ

Tractable Fragments?

- As in the case of Allen's interval calculus, we may ask for maximal tractable subsets . . .
- Again, one can identify relations that can be encoded by Horn formulae.
- Idea: Consider relations that can be expressed in a way such that we have to consider only Horn formulae inside all worlds.
- Idea: Try to restrict the number of worlds to consider to a poly. number
- 148 Horn relations $\mathcal{H}_8$, which forms again a maximal subset.
- Path consistency is refutation complete for $\mathcal{H}_8$.
- There are 2 additional maximal subsets that allow for poly. satisfiability testing!

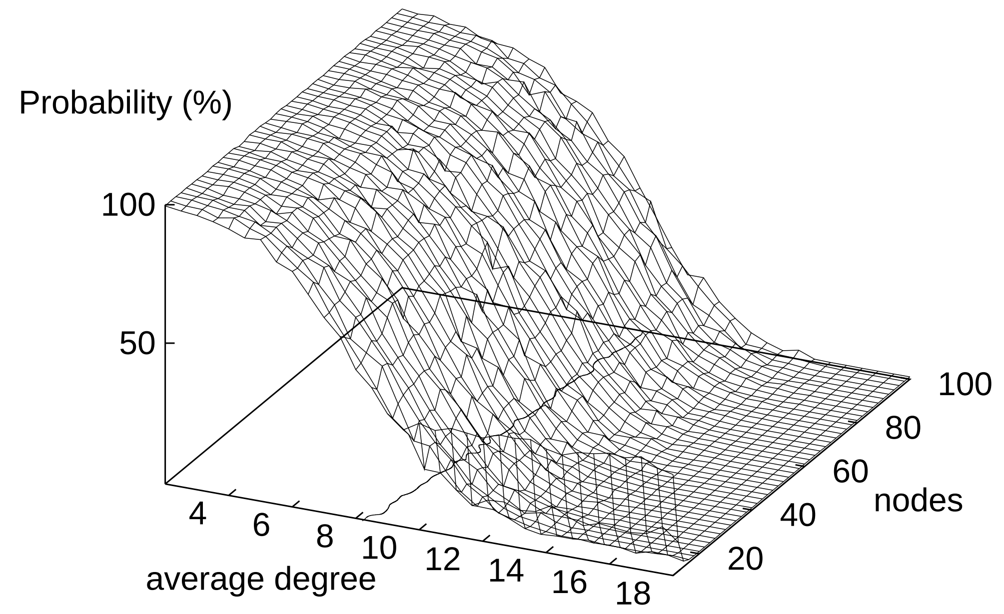
Some Experiments

- How difficult is the RCC8 satisfiability problem in practice?
- Are there particularly difficult instances?
  - ⇝ Where is the phase transition region?
  - ⇝ Cheeseman et al [IJCAI 91] conjectured that for all NP-complete problems there exists a parameter such that when changing this parameter there exists a very small range – the *phase transition region* – where the probability of satisfiability of randomly generated instances changes from 1 to 0. They also conjectured that in this area one finds many hard instances.
- How well does the path consistency method approximate satisfiability?
- Can $\mathcal{H}_8$ be used to speed up the satisfiability testing?

Generating Instances

- Randomly generating instances according to the following parameters:
  - Number of nodes n
  - Average number of constraints d: $(nd/2$ out of $n(n-1)/2$ possible constraints
  - Average number of base relations l per constraint
  - Allowed constraints
    - $A(n,d,l)$: all RCC8 relations
    - $H(n,d,l)$: only relations out of RCC8 $- \mathcal{H}_8$

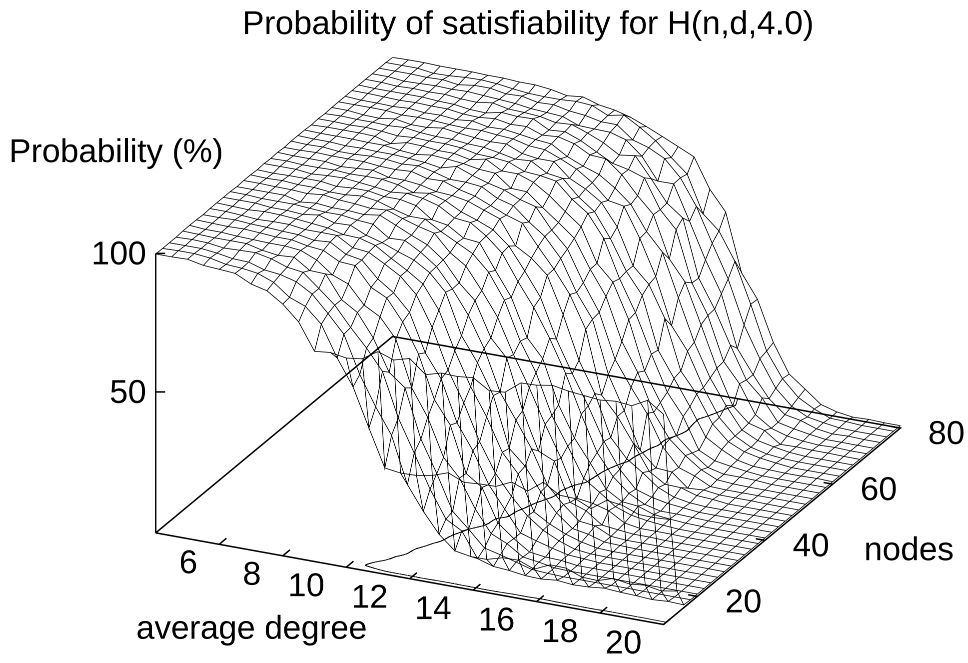
Phase Transition for $A(n, d, 4)$



500 instances per data point

- Phase transition for $A(n, d, 4)$ between $d = 8$ and $d = 10$ for $10 \leq n \leq 100$.

Phase Transition for $H(n, d, 4)$



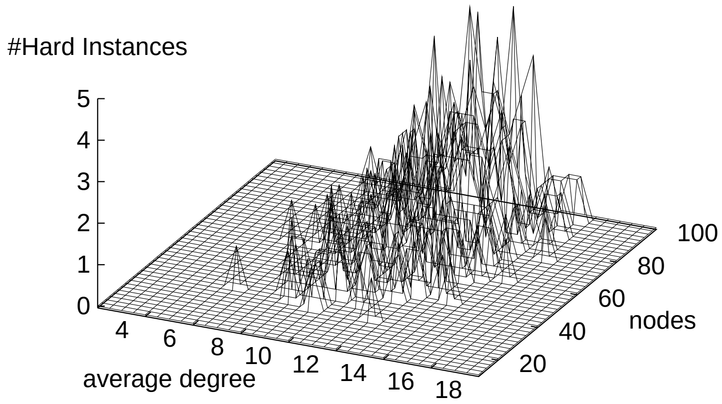
500 instances per data point

- Phase transition for $H(n, d, 4)$ between $d = 10$ and $d = 15$ for $10 \leq n \leq 80$.

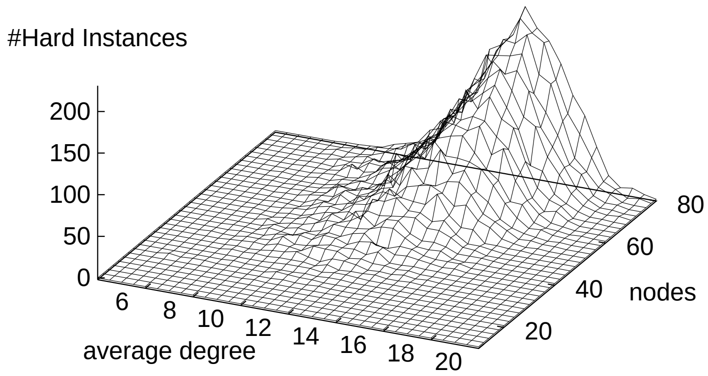
Hard Instances . . .

. . . using more than 10,000 search nodes

Number of hard instances for A(n,d,4.0)



Number of hard instances for H(n,d,4.0)

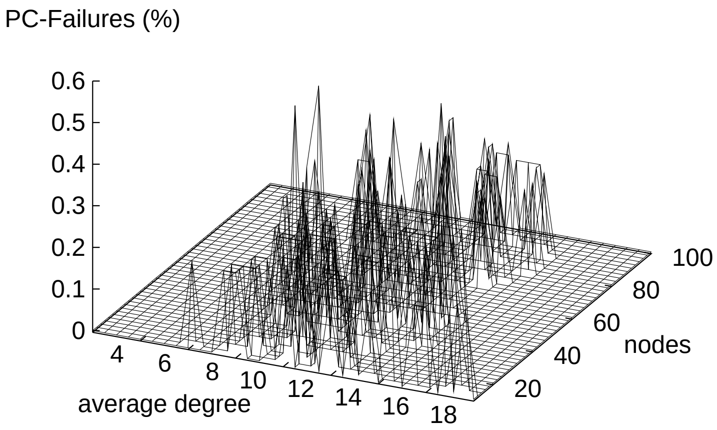


500 instances per data point

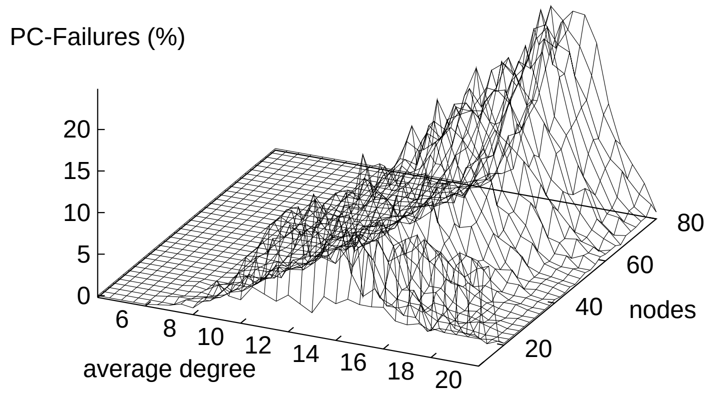
Quality of Path Consistency. . .

. . . measured as the percentage of path consistent but unsatisfiable CSPs

Percentage points of incorrect PCA answers for A(n,d,4.0)



Percentage points of incorrect PCA answers for H(n,d,4.0)



500 instances per data point

Outlook & Open Problems

- RCC8 is the spatial-topological counterpart of Allen's interval calculus.
- Formalization can be done using topology and – because of McKinsey & Tarski's result – modal logic.
- Computationally well behaved.
- In contrast to Allen's calculus no applications so far.
- Combinations of RCC8 with other constraint spatial calculi.
- Combining RCC8 and Allen's interval calculus to form a temporal-spatial calculus.
- Are there other interesting spatial calculi?

Literature I

B. Bennett.
Spatial Reasoning with propositional logic.
Principles of Knowledge Representation and Reasoning: Proceedings of the 4th International Conference (KR-94), 1994, 51-62.

W. Nutt.
On the translation of qualitative spatial reasoning problems into modal logics.
Advances in Artificial Intelligence, Proc. 23rd Annual German Conference on Artificial Intelligence, KI'99, Bonn (Germany), September 1999.

J. Renz & B. Nebel.
On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the Region Connection Calculus.
Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI'97), August 1997, 522-527.

J. Renz & B. Nebel,
Efficient Methods for Qualitative Spatial Reasoning,
Proceedings of the 13th European Conference on Artificial Intelligence (ECAI'98), August 1998, 562-566.