ECAI 2014
T. Schaub et al. (Eds.)

doi:10.3233/978-1-61499-419-0-1127

A Concise Horn Theory for RCC8

Matthias Westphal and **Julien Hué** [1]

Abstract. RCC8 is a well-known constraint language for expressing and reasoning about spatial knowledge. We state a simple and concise Horn theory for RCC8 analogous to the ORD-Horn theory for temporal reasoning. This theory allows for expressing RCC8 and retains tractability of the well-known Horn reduct of RCC8. Further, it is much more adequate for practical purposes in the area of logic programming and surpasses previous attempts.

1 BACKGROUND

We assume the reader is familiar with standard FOL concepts, notation in the CSP literature, and syntactic interpretations [2, 6]. A constraint language Γ is simply a relational FO structure. The *Constraint Satisfaction Problem for* Γ, CSP(Γ), is the decision problem if some given primitive positive FO sentence formulated over the signature of Γ is satisfiable in Γ. An algorithm *solves* CSP(Γ) if it solves the decision problem.

The spatial reasoning formalism RCC8 [3] describes binary relations between abstract spatial regions. There are 8 relation primitives: equal *EQ*, disconnected *DC*, externally connected *EC*, partially overlaps *PO*, non-tangential proper part *NTPP*, tangential proper part *TPP*, and the converses *NTPPI*, *TPPI*. For example, the formalism allows to formally state that "Central Park is a (non-tangential proper) part of Manhattan". These eight relations are denoted in the signature σ_{RCC8A} of atomic RCC8 relation symbols. Further, RCC8 contains all disjunctions of these relation symbols forming the signature σ_{RCC8} of 256 relations. We denote by Γ the constraint language RCC8 on signature σ_{RCC8} constructed as an ω-categorical structure as in [2].

There are two important reducts of Γ: the σ_{RCC8A}-reduct Γ_{RCC8A}, and the $\sigma_{\widehat{\mathcal{H}}_8}$-reduct $\Gamma_{\widehat{\mathcal{H}}_8}$. Symbols of $\sigma_{\widehat{\mathcal{H}}_8}$ denote the RCC8 Horn relations. Renz and Nebel [4] gave a transformation of any instance of CSP($\Gamma_{\widehat{\mathcal{H}}_8}$) with n FO variables into an equivalent propositional Horn formula with $O(n^4)$ literals. Well-known results (e.g. [4]) are: $\sigma_{\mathrm{RCC8A}} \subsetneq \sigma_{\widehat{\mathcal{H}}_8} \subsetneq \sigma_{\mathrm{RCC8}}$, $|\sigma_{\widehat{\mathcal{H}}_8}| = 148$, and strong 3-consistency solves CSP($\Gamma_{\widehat{\mathcal{H}}_8}$). The reduct $\Gamma_{\widehat{\mathcal{H}}_8}$ is the largest maximal tractable subclass of RCC8 that includes all atomic relations. Further, CSP($\Gamma_{\widehat{\mathcal{H}}_8}$) $\in$ P while CSP(Γ) is NP-complete.

We present a syntactic interpretation of RCC8 and a concise Horn theory for solving CSP($\Gamma_{\widehat{\mathcal{H}}_8}$) in polynomial time. The theory is weaker than strong 3-consistency. The Herbrand expansion yields propositional Horn formulas with $O(n^2)$ literals and $O(n^3)$ clauses.

2 A CONCISE HORN THEORY

Bennett discussed a number of different representations for logical reasoning about spatial relations [1]. Among them is the RCC7 representation which consists of relations representable as equations in interior algebra. These 7 relations are: equality *EQ*, disconnectedness *DC*, discreteness *DR*, parthood P, non-tangential parthood *NTP*, and the converses P^{-1}, NTP^{-1}. Bennett found 115 distinct relations to be definable using purely conjunctive formulas if the complements of the relations are included.

In the following we introduce a set of relations closely related to RCC7 and its use for defining further relations. Here, *EQ* is not included as a primitive because it can be defined using parthood. Further, we drop converse symbols and use the complements of *DC* and *DR*: connectedness *C* and overlap *O*. We refer to this set as RCC4. Its signature can be thought of as a shorthand notation of specific RCC8 relation symbols, i.e., we treat σ_{RCC4} as a proper subset of σ_{RCC8}.

Definition 1 (RCC4) *In the binary signature* $\sigma_{RCC4} := \{C, O, NTP, P\}$ *NTP denotes the relation symbol NTPP in* σ_{RCC8}*, C the relation symbol representing* $\neg DC(x,y)$ *over* Γ*, O the relation symbol representing* $\neg\big(DC(x,y) \vee EC(x,y)\big)$*, and P the symbol representing* $TPP(x,y) \vee NTPP(x,y) \vee EQ(x,y)$*. Further,* Γ_{RCC4} *denotes the* σ_{RCC4}*-reduct of* Γ.

The structure Γ_{RCC4} admits a FO interpretation of Γ. We first give the syntactic interpretation of the atomic RCC8 relation symbols (virtually the same map as provided in Table 5.4 of [1]).

Definition 2 (π_{RCC4}) *The syntactic interpretation* π_{RCC4} *of atomic RCC8 relation symbols in the RCC4 symbols is defined as follows:*

$$\begin{aligned}
\pi_{RCC4}(DC)(x,y) &:= \neg C(x,y)\\
\pi_{RCC4}(EC)(x,y) &:= C(x,y) \wedge \neg O(x,y)\\
\pi_{RCC4}(PO)(x,y) &:= O(x,y) \wedge \neg P(x,y) \wedge \neg P(y,x)\\
\pi_{RCC4}(EQ)(x,y) &:= P(x,y) \wedge P(y,x)\\
\pi_{RCC4}(TPP)(x,y) &:= P(x,y) \wedge \neg P(y,x) \wedge \neg NTP(x,y)\\
\pi_{RCC4}(NTPP)(x,y) &:= NTP(x,y)\\
\pi_{RCC4}(TPPI)(x,y) &:= \pi_{RCC4}(TPP)(y,x)\\
\pi_{RCC4}(NTPPI)(x,y) &:= \pi_{RCC4}(NTPP)(y,x)
\end{aligned}$$

It is easy to verify that this syntactic interpretation provides an interpretation of Γ_{RCC8A} in Γ_{RCC4}. The map can be easily extended to all RCC8 relations, thus Γ has an interpretation in Γ_{RCC4}. Note, although RCC4 is very close to RCC7, the formula $\neg EQ(x,y)$ is written as a disjunction over σ_{RCC4} – contrary to RCC7 where *EQ* is a primitive. From the RCC subsumption lattice [3] it follows:

Proposition 1 *Let* ψ *be a formula* $L_O(x,y) \wedge L_C(x,y) \wedge L_P(x,y) \wedge L_{NTP}(x,y) \wedge L_P(y,x) \wedge L_{NTP}(y,x)$ *in which* $L_R(x,y) = R(x,y)$ *or* $L_R(x,y) = \neg R(x,y)$*. Then* ψ *is either unsatisfiable in* Γ_{RCC4} *or equivalent to an atomic RCC8 relation.*

For convenience we consider the expansion of Γ_{RCC4} to $\Gamma_{\mathrm{RCC4+}}$ by adding a symbol $\overline{R}$ for the complement of each relation symbol R

[1] Albert-Ludwigs-Universität Freiburg, Department of Computer Science, Germany, email: westpham@informatik.uni-freiburg.de, hue@informatik.uni-freiburg.de

in σ_{RCC4} ($\Gamma_{\mathrm{RCC4+}}$ is still a reduct of Γ). Then, π_{RCC4} can be seen as primitive positive. In the following, we write $\pi_{\mathrm{RCC4+}}$ to denote the syntactic interpretation of each RCC8 relation symbol over $\Gamma_{\mathrm{RCC4+}}$ that uses CNF defining formulas consisting of prime implicates.

Computer analysis shows the following properties.

Proposition 2 *There are 42 RCC8 relations which have a primitive positive definition in Γ_{RCC4+}. All 148 RCC8 Horn relations have a Horn definition in Γ_{RCC4}.*

In the following we give a Horn theory over Γ_{RCC4} written as the Datalog program $\Pi^{\mathrm{RCC4^+}}$ with symbols in $\sigma_{\mathrm{RCC4+}}$ and a special symbol false denoting $\bot$. The program contains the necessary subsumption rules and rules on 3 variables that derive from the axiomatization of the atomic composition table of RCC8 given by Bodirsky and Wölfl [2] using $\pi_{\mathrm{RCC4+}}$ and simplifications. It solves $\mathrm{CSP}(\Gamma_{\mathrm{RCC4+}})$ by Proposition 1 and the results of [2].

$$
\begin{array}{rlrl}
 & P(x,x) & & \overline{NTP}(x,x) \\
C(x,y) & :\!-\; O(x,y) & O(x,y) & :\!-\; P(x,y) \\
P(x,y) & :\!-\; NTP(x,y) & \mathsf{false} & :\!-\; P(y,x), NTP(x,y) \\
O(x,y) & :\!-\; O(y,x) & C(x,y) & :\!-\; C(y,x) \\
\mathsf{false} & :\!-\; C(x,y), \overline{C}(x,y) & \mathsf{false} & :\!-\; NTP(x,y), \overline{NTP}(x,y) \\
\mathsf{false} & :\!-\; O(x,y), \overline{O}(x,y) & \mathsf{false} & :\!-\; P(x,y), \overline{P}(x,y) \\
P(x,y) & :\!-\; P(x,z), P(z,y) & C(x,y) & :\!-\; C(x,z), P(z,y) \\
O(x,y) & :\!-\; C(x,z), NTP(z,y) & O(x,y) & :\!-\; O(x,z), P(z,y) \\
NTP(x,y) & :\!-\; P(x,z), NTP(z,y) & NTP(x,y) & :\!-\; NTP(x,z), P(z,y)
\end{array}
$$

3 PROPERTIES AND EVALUATION

Analogous to the study of temporal formalisms in [6], we can analyze the relation of this new program to commonly used ones: strong 3-consistency ($\Pi^{\mathrm{RCC8}}_{\leq 3}$), the atomic composition table ($\Pi^{\mathrm{RCC8A}}_{\leq 3}$), and the augmented $\Pi^{\mathrm{RCC4^+}}$ that emulates positive unit resolution on Horn clauses in $\pi_{\mathrm{RCC4+}}$ ($\Pi^{\mathrm{RCC4^+}}_{RES}$). Additionally, each of them can be weakened by restricted inference to arcs of an instance that are part of some fixed chordal graph of an instance (denote by $|_{\widehat{G}}$). Finally, we can define a strict partial order on the programs, where $A \prec B$ if program A rules out more partial solutions than program B. For details, references, and discussion of these concepts see [6].

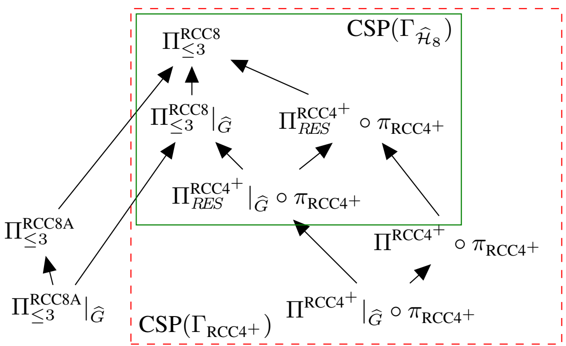

Figure 1. Lattice of propagation strength of Datalog programs given interpretations for RCC8. Frames indicate CSPs solved by the programs.

The lattice for $\prec$ is given in Figure 1 (where $A \leftarrow B$ if $A \prec B$) which also states which CSPs of reducts are solved by the programs.

In order to encode instances to propositional CNF it has been common practice to consider the *support encoding of* $\Pi^{RCC8A}_{\leq 3}|_{\widehat{G}}$ (see [6] for references). However, this encoding is quite weak.

Proposition 3 *Unit propagation on the support encoding of* $\Pi^{RCC8A}_{\leq 3}$ *does* not *solve* $CSP(\Gamma_{RCC8A})$*, while unit propagation on the Herbrand expansion of* $\Pi^{RCC4^+}_{RES}|_{\widehat{G}} \circ \pi_{RCC4+}$ *solves* $CSP(\Gamma_{\widehat{\mathcal{H}}_8})$.

If **(a)** the support encoding of $\Pi^{\mathrm{RCC8A}}_{\leq 3}|_{\widehat{G}}$ is used SAT solvers have to rely on search to refute instances, but when **(b)** the Herbrand expansion of $\Pi^{\mathrm{RCC4^+}}_{RES}|_{\widehat{G}} \circ \pi_{\mathrm{RCC4+}}$ is used search is only necessary on non-Horn relations of instances in $\mathrm{CSP}(\Gamma)$. Further, both encodings result in propositional CNF with $O(n^2)$ literals and $O(n^3)$ clauses for instances with n FO variables, but interpretation and program of the new concise theory have less relation symbols and rules.

For a brief evaluation we consider 1 000 random instances of $\mathrm{CSP}(\Gamma)$ constructed from the $H(150, 15.5, 4.0)$ model with hard relations $\mathcal{NP}_8$ [5]. These instances have 150 FO variables and a constraint network degree of 15.5. The test setup is the same as in [6].[2]

	Runtime (seconds) of the Glucose-2.2 SAT solver				
SAT Encoding	Avg.	25-PCTL	50-PCTL	90-PCTL	99-PCTL
(a)	440.90	227.05	349.48	545.62	1 966.94
(b)	45.87	8.92	18.69	103.75	441.32

This is a factor 4 to 25 improvement. Further, the average size of an instance in DIMACS drops from **(a)** 360 MiB to **(b)** 160 MiB.

4 CONCLUSION

We presented a concise Horn theory for RCC8 that allows for expressing the entire RCC8 language and retains the tractability of the well-known Horn class of relations. The theory is of broad interest for incorporating RCC8 in logic programming, e.g. SAT and ASP. In particular, the theory allows for an improved $O(n^3)$ propositional representation that is superior to previous work.

For future work it would be interesting to consider propagators based on this theory in Constraint Programming. In particular, a potential graph based algorithm could form a fast propagator.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers and Stefan Wölfl for helpful comments. This work was supported by DFG (*SFB/TR 8 Spatial Cognition*, project R4-[LogoSpace]).

REFERENCES

[1] Brandon Bennett, 'Logical representations for automated reasoning about spatial relationships', *Ph.D. Thesis, University of Leeds*, (1997).
[2] Manuel Bodirsky and Stefan Wölfl, 'RCC8 is polynomial on networks of bounded treewidth', in *IJCAI*, ed., Toby Walsh, pp. 756–761. IJCAI/AAAI, (2011).
[3] David A. Randell, Zhan Cui, and Anthony G. Cohn, 'A spatial logic based on regions and connection', in *KR*, eds., Bernhard Nebel, Charles Rich, and William R. Swartout, pp. 165–176. Morgan Kaufmann, (1992).
[4] Jochen Renz and Bernhard Nebel, 'On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the region connection calculus', *Artificial Intelligence*, **108**(1-2), 69–123, (1999).
[5] Jochen Renz and Bernhard Nebel, 'Efficient methods for qualitative spatial reasoning', *JAIR*, **15**, 289–318, (2001).
[6] Matthias Westphal, Julien Hué, and Stefan Wölfl, 'On the propagation strength of SAT encodings for qualitative temporal reasoning', in *ICTAI*, pp. 46–54. IEEE, (2013).

[2] The tool to compute the encoding is available; see this publication's entry at `http://www.informatik.uni-freiburg.de/~ki/publications/`