GeospatialRules: A Datalog^{+} RuleML Rulebase for Geospatial Reasoning

RuleML 2014, Rule Challenge
Prague, Czech Republic, August 18–22, 2014

Gen Zou

Faculty of Computer Science,
University of New Brunswick, Fredericton, Canada

Outline

Outline

1. Background
2. Development of GeospatialRules
3. Conclusion and Future Work

Datalog$^+$

- Datalog$^+$ extends Datalog with rule heads allowing
    - Existential quantifiers
    - Equality
    - Falsity (for integrity rules)
- A family of decidable sublanguages of Datalog$^+$, Datalog$^\pm$, restricts rule bodies for ontological querying
- Deliberation RuleML 1.01 introduces Relax NG and XSD schemas for Datalog$^+$

Region Connection Calculus (RCC)

- RCC axiomatizes spatial relations between 'regions' in first-order logic
- Primitive binary relation $C(x, y)$: x connects with y
  - Reflexivity and symmetry of $C(x, y)$

$$
\begin{array}{ll}
(1) & \forall x : C(x, x) \\
(2) & \forall x, y : C(y, x) \rightarrow C(x, y)
\end{array}
$$

Defining Axioms of RCC Relations

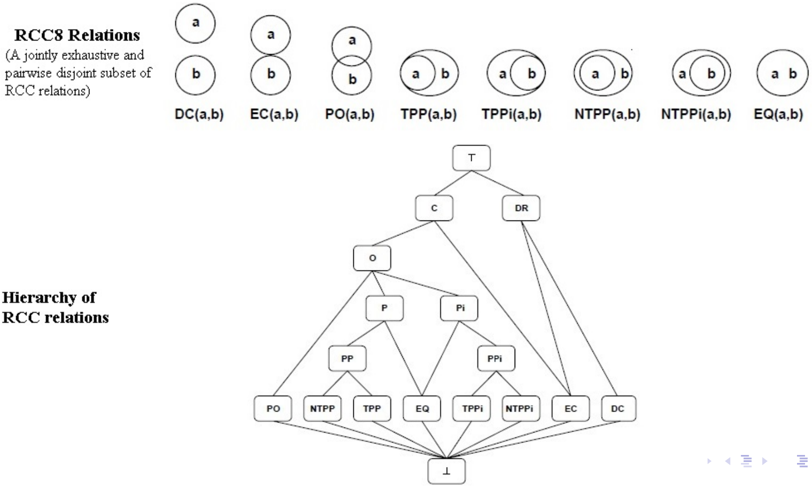
Relation Definition	Meaning
$C(x, y)$ (primitive)	x connects with y
(3) $DC(x, y) \equiv_{def} \neg C(x, y)$	x is disconnected from y
(4) $P(x, y) \equiv_{def} \forall z[C(z, x) \rightarrow C(z, y)]$	x is part of y
(5) $O(x, y) \equiv_{def} \exists z[P(z, x) \wedge P(z, y)]$	x overlaps with y
(6) $DR(x, y) \equiv_{def} \neg O(x, y)$	x is discrete from y
(7) $EC(x, y) \equiv_{def} C(x, y) \wedge DR(x, y)$	x is externally connected to y
(8) $PO(x, y) \equiv_{def} O(x, y) \wedge \neg P(x, y) \wedge \neg P(y, x)$	x partially overlaps with y
(9) $EQ(x, y) \equiv_{def} P(x, y) \wedge P(y, x)$	x is equal to y
(10) $PP(x, y) \equiv_{def} P(x, y) \wedge \neg P(y, x)$	x is a proper part of y
(11) $TPP(x, y) \equiv_{def} PP(x, y) \wedge \exists z[EC(z, x) \wedge EC(z, y)]$	x is a tangential proper part of y
(12) $NTPP(x, y) \equiv_{def} PP(x, y) \wedge \neg\exists z[EC(z, x) \wedge EC(z, y)]$	x is a nontangential proper part of y

Inverses of non-symmetric RCC relations:

$$\Phi i(x, y) \equiv_{def} \Phi(y, x), \Phi \in \{P, PP, TPP, NTPP\}$$

Graphical View of RCC Relations

Figures borrowed from Cohn, A., Bennett, B., Gooday, J., Gotts, N.: Qualitative spatial representation and reasoning with the Region Connection Calculus. GeoInformatica 1(3) (October 1997) 275–316.



Outline

1. Background
2. Development of GeospatialRules
3. Conclusion and Future Work

Kinds of GeospatialRules

- Rules
  a) Datalog^{+} fragment of the first-order RCC defining axioms
  b) Additional rules which express part of the RCC knowledge that is not captured by a)
- Facts and Queries (special cases of Rules)

Mapping: RCC Relations $\longleftrightarrow$ GeospatialRules Relations

RCC Relation	GeospatialRules Relation
C	RCCConnected
DC	RCCDisconnected
P	RCCPartOf
O	RCCOverlapped
DR	RCCDiscrete
EC	RCCExternallyConnected
PO	RCCPartiallyOverlapped
EQ	RCCEqual
PP	RCCProperPartOf
TPP	RCCTangentialProperPartOf
NTPP	RCCNonTangentialProperPartOf
Pi	RCCInversePartOf
PPi	RCCInverseProperPartOf
TPPi	RCCInverseTangentialProperPartOf
NTPPi	RCCInverseNonTangentialProperPartOf

Identifying Datalog$^+$ Fragment of RCC Axioms

- Reflexivity and symmetry axioms of C(x, y) are kept unchanged

```
Forall ?X ( RCCConnected(?X ?X) )
Forall ?X ?Y (
  RCCConnected(?Y ?X) :- RCCConnected(?X ?Y)
)
```

- Datalog$^+$ fragment of other axioms is exposed using first-order transformations
  - Split "iff" definitions into two rules expressing "if" and "only if" parts separately
  - Transform each "only if" rule into one or more Datalog$^+$ rules
  - Take note of any "if" rules that cannot be expressed in Datalog$^+$, and transform the remaining ones into Datalog$^+$

Identifying Datalog$^+$ Fragment of RCC – Example

Split defining axiom of TangentialProperPartOf into three rules:

$$
\begin{aligned}
& \mathrm{TPP}(x, y) \equiv_{\mathrm{def}} \mathrm{PP}(x, y) \wedge \exists z[\mathrm{EC}(z, x) \wedge \mathrm{EC}(z, y)] \\
\equiv \quad & (\mathrm{TPP}(x, y) \rightarrow \mathrm{PP}(x, y) \wedge \exists z[\mathrm{EC}(z, x) \wedge \mathrm{EC}(z, y)]) \wedge \\
& (\mathrm{PP}(x, y) \wedge \exists z[\mathrm{EC}(z, x) \wedge \mathrm{EC}(z, y)] \rightarrow \mathrm{TPP}(x, y)) \\
\equiv \quad & (\mathrm{TPP}(x, y) \rightarrow \mathrm{PP}(x, y)) \wedge \\
& (\mathrm{TPP}(x, y) \rightarrow \exists z[\mathrm{EC}(z, x) \wedge \mathrm{EC}(z, y)]) \wedge \\
& (\forall z(\mathrm{PP}(x, y) \wedge \mathrm{EC}(z, x) \wedge \mathrm{EC}(z, y) \rightarrow \mathrm{TPP}(x, y)))
\end{aligned}
$$

Identifying Datalog$^+$ Fragment of RCC – Example

Identify Datalog$^+$ fragment of defining axiom of ProperPartOf:

$$
\begin{aligned}
& \mathrm{PP}(x, y) \equiv_{\text{def}} \mathrm{P}(x, y) \wedge \neg \mathrm{P}(y, x) \\
\equiv \quad & (\mathrm{PP}(x, y) \rightarrow \mathrm{P}(x, y) \wedge \neg \mathrm{P}(y, x)) \\
& \wedge (\mathrm{P}(x, y) \wedge \neg \mathrm{P}(y, x) \rightarrow \mathrm{PP}(x, y)) \\
\equiv \quad & (\mathrm{PP}(x, y) \rightarrow \mathrm{P}(x, y)) \wedge (\mathrm{PP}(x, y) \rightarrow \neg \mathrm{P}(y, x)) \\
& \wedge (\mathrm{P}(x, y) \wedge \neg \mathrm{P}(y, x) \rightarrow \mathrm{PP}(x, y)) \\
\equiv \quad & (\mathrm{PP}(x, y) \rightarrow \mathrm{P}(x, y)) \wedge (\mathrm{PP}(x, y) \wedge \mathrm{P}(y, x) \rightarrow \bot) \\
& \wedge (\mathrm{P}(x, y) \wedge \neg \mathrm{P}(y, x) \rightarrow \mathrm{PP}(x, y))
\end{aligned}
$$

The first two conjuncts, which are within Datalog$^+$ expressivity, are included in GeospatialRules as two rules. These rules express the "only if" part of the axiom. Note that the third conjunct is omitted since it is beyond Datalog$^+$ expressivity.

Additional Rules

- Compensate for a portion of information that is lost when the Datalog^{+} fragment was identified
- Subsumptions between RCC relations, e.g.

```
Forall ?X ?Y (
  RCCConnected(?X ?Y) :- RCCOverlapped(?X ?Y)
)
```

- Rules from the RCC8 composition table, e.g.

```
Forall ?X ?Y ?Z (
  RCCNonTangentialProperPartOf(?X ?Z) :-
    And(RCCPartOf(?X ?Y)
        RCCNonTangentialProperPartOf(?Y ?Z))
)
```

Facts

- Describe geospatial relationships among regions in North America, as well as the Pacific and the Atlantic oceans

```
RCCProperPartOf(USA NorthAmerica)
Ocean(Pacific)
Ocean(Atlantic)
State(California)
RCCProperPartOf(California USA)
RCCExternallyConnected(Pacific USA)
RCCExternallyConnected(Pacific California)
State(Virginia)
RCCProperPartOf(Virginia USA)
RCCExternallyConnected(Atlantic USA)
RCCExternallyConnected(Atlantic Virginia)
```

Sample Query

English Query: "*Which states [in the rulebase] are tangential proper parts of the USA*?"

GeospatialRules Query:

```
And(State(?St) RCCTangentialProperPartOf(?St USA))
```

Answer: `?St=California, ?St=Virginia`

Rule applied:

```
Forall ?X ?Y ?Z (
  RCCTangentialProperPartOf(?X ?Y) :-
    And(RCCProperPartOf(?X ?Y)
        RCCExternallyConnected(?Z ?X)
        RCCExternallyConnected(?Z ?Y))
)
```

RuleML/XML Version of GeospatialRules

- GeospatialRules rulebase (42 rules, 11 facts, 3 queries) available online
  `http://deliberation.ruleml.org/1.01/exa/RulebaseCompetition2014/GeospatialRulesRCC.ruleml`
- Can be validated against Deliberation RuleML 1.01 Relax NG schema (Relaxed Serialization for Datalog$^+$)
  `http://deliberation.ruleml.org/1.01/relaxng/datalogplus_min_relaxed.rnc`
- Can be normalized and validated against Deliberation RuleML 1.01 XSD schema for Datalog$^+$
  `http://deliberation.ruleml.org/1.01/xsd/datalogplus_min_normal.xsd`
- Mathematical English definitions of each RCC relation are included as comments

Outline

1. Background
2. Development of GeospatialRules
3. Conclusion and Future Work

Conclusion

- We developed a Datalog$^+$ rulebase, GeospatialRules, which can be used in geospatial applications for reasoning with RCC relations (e.g., in RCC8)
- GeospatialRules consists of rules, facts, and queries
- Subsumptions between RCC relations are preserved in GeospatialRules
- An XML version of GeospatialRules complying to the Deliberation RuleML 1.01 standard is provided
- Demo available online
  `http://wiki.ruleml.org/index.php/GeospatialRules_Demo`

Future Work

- Create new GeospatialRules versions for graph queries
  - Represent regions by globally unique object identifiers
  - Use graph versions of binary RCC relations as slot names
  - Express transformation from relational to graph version of RCC as rules in PSOA RuleML
  - Create new versions of GeospatialRules
    - Mediator-style: Extend rulebase with transformation rules from relational to graph schema
    - Warehouse-style: Convert rulebase into a pure graph version using transformation rules
- Compare GeospatialRules versions using PSOA RuleML implementation based on PSOA2Prolog
  - Mediator-Warehouse trade-off
  - Readability assessment (collaboration with geospatial experts)
  - Efficiency exploration (translation and execution times, memory requirements)