Related DL-based standards (OWL, OWL2) are established

Many DL reasoners available
- FaCT++, pellet, HermiT, RacerPro, TrOWL

The user community is growing fast
- Swoogle searches over 10,000 online ontologies
- Larger and larger ontologies
  - SNOMED has 379,691 concepts
- more and more complicated ontologies
  - FMA (Foundational Model of Anatomy, OWL DL) has 41,647 concepts and 123,564 axioms

Impact: From domain applications to software engineering
Designing and Classifying Ontologies

Annotating Data with Ontological Vocabulary

[Photo source: talis.com]
Query Answering

Such as in SPARQL:

SELECT ?X
FROM <http://example.org/person.owl>
WHERE {?X friend ?Y .}

SELECT ?X
FROM <http://example.org/person.owl>
WHERE {?X friend ?Y .
?Y rdf:type Popular .}

OWL 2: Web Ontology Language

OWL 2 Full
OWL 2 DL
OWL 1 DL
OWL 2 QL
OWL 2 RL
OWL 2 EL

Undecidable
2NExpTime-Complete
NExpTime-Complete
PTime-Complete
In AC0
Three Approaches to Reasoning in OWL 2

1. Sound and complete reasoning in OWL 2 DL
2. Sound and complete reasoning in tractable profiles (RL, QL and EL)
3. Approximate reasoning in OWL 2 DL (from DL to tractable profiles)

OWL 2 DL: High Computational Complexity

- General class axioms $C \sqsubseteq D$ are widely used
  - such as the domain constraint
    - ObjectProperty(hasF domain (Person))
    - $\exists hasF \sqsubseteq Person$
- Expansion rules for general class axioms $C \sqsubseteq D$
  - If a new node $y$ is added, then add a general class description $D \cup -C$ in $L(y)$ for each general class axiom $C \sqsubseteq D$
- For example: if an ontology consists of 100 general class axioms
  - then theoretically there are $2^{100}$ possible expansions
  - for each node
OWL 2 DL: Ontology Classification is Challenging

- Classification is a complex reasoning service
  - Compute the class hierarchy
- Classification can be reduced to subsumption checking
  - Between every pair of classes
  - Which can be further reduced to class unsatisfiability checking
- For example, if an ontology consists of 100,000 classes
  - Then a naïve implementation needs
    - 100,000 * 100,000 class unsatisfiability checking

OWL 2 DL: ABox Reasoning Challenging Even for Ontologies in Profiles

| Ontology | \(\mathcal{C}V\) | \(\mathcal{R}V\) | \(\mathcal{L}V\) | \(|A|\) | TrOWL | Pellet | HermiT | FaCT++ |
|----------|----------------|----------------|----------------|------|-------|-------|--------|-------|
| VICODI   | 184            | 10             | 29614          | 114164 | 2.014 | 9.971 | 13.138 | timeout |
| NG-1     |                |                |                |       |       |       |        |        |
| NG-2     |                |                |                |       |       |       |        |        |
| NG-5     | 2748           | 413            | 47756          | 118458 | 28.947 | timeout | timeout | timeout |
| NG-8     |                |                | 78899          | 278365 | 63.833 | timeout | timeout | timeout |
| NG-13    |                |                | 97995          | 665304 | 143.288 | timeout | timeout | timeout |
Case Study 1: Profiles not enough

- Adding disjointness to SNOMED led to surprising results
  - many classes become inconsistent, e.g., *percutaneous embolization of hepatic artery using fluoroscopy guidance*
- Cause of inconsistencies identified as class *groin*
  - *groin* asserted to be subclass of both *abdomen* and *leg*
  - *abdomen* and *leg* are disjoint
  - modelling of *groin* (and other similar “junction” regions) identified as incorrect
- *Kaiser Permanente* extending SNOMED to express, e.g.:
  - *non-viral pneumonia* (negation)
  - *infectious pneumonia* is caused by a *virus* or a *bacterium* (disjunction)
  - *double pneumonia* occurs in two *lungs* (cardinalities)

---

Case Study 2: Profiles not enough

```
TwoUse

FreeTradeZone

Customer

Purchase Order

getCharges()

Product

Name: String

OWL

UML

<<owlRestriction>>

CountryFromEU

<<owlValue>>(hasValue = eu)

memberOfTradeZone : FreeTradeZone

<<owlRestriction>>

OrderFromEUCustomer

<<owlValue>>(someValuesFrom = CustomerFromACountry)

hasCustomer

DutyFreeOrder

<<owlRestriction>>

OrderFromEUCustomer

<<owlRestriction>>

CountryFromEU

<<owlRestriction>>

CustomerFromEUCountry
```

Jeff Z. Pan
Case Study 3: Profiles not enough

- Integrated PDDSL (IPDDSL)
  - Structural modelling with expressiveness of **OWL 2 DL**

```xml
DeviceType "Cisco_7603"

SubClassOf: pd hasConfiguration some ( pd hasSlot some ( pd hasCard some Cisco_7600 SIP ))

longName : "CISCO 7603 CHASSIS"
description : "The Cisco® 7603 Router is a high-performance...

allowed : { PossibleConfiguration "Cisco_7603_Configuration" (
  Slot "1" allowed: "Supervisor_Engine_2", "Supervisor_Engine_720" required: true
  Slot "2" allowed: "Supervisor_Engine_2", "Supervisor_Engine_720" "Catalyst_6500_Module"
  required : false
  Slot "3" allowed: "Catalyst_6500_Module" required : false

Device serialNumber : "cisco_7603" hasType : "Cisco_7603"
configuration :

  Slot id: "1": Card serialNumber: "supervisor_2_5" hasType: "WS-X6K-S2U-MSPC2"
  Slot id: "2":
  Slot id: "3":
```

Approximate Reasoning

- Input approximation
  - Semantic approximation
  - Syntactic approximation

- Approximate Deduction

- Outputs (Quality Control)
  - soundness preserving
  - completeness preserving
Semantic Approximation [AAAI2007]

- Strongest weaker approximation for QL ES(O, OWL 2 QL) of an OWL2 DL O is finite and unique.
- **Theorem 1**: Given an ontology O, a conjunctive query q(X) and an evaluation [X→S], if ES(O, OWL 2 QL) |= q_{[X→S]} then O |= q_{[X→S]}.
- **Theorem 2**: Given an ontology O_S, a database-style conjunctive query q(X) without non-distinguished variables and an evaluation [X→S], ES(O_S, OWL 2 QL) |= q_{[X→S]} iff O_S |= q_{[X→S]}.

Evaluation: Lehigh University Benchmark (Quill – QL reasoner in TrOWL)

- OWL 2 DL/ SHIQ reasoners [Motik and Sattler, 2006]
  - can answer queries for up to 4 universities
- Syntactic approximation systems [Guo et al., 2004]
  - more scalable
  - can be incorrect/incomplete
- Quill/ONTOSEARCH2 can handle at least 50 universities
  - Sound and complete results for all queries
- DLDB vs. Quill/ONTOSEARCH2
Faithful Syntactic Approximate Reasoning
[AAAI2010]

- Syntactic approximation from OWL2 DL to OWL2 EL
  - Minor syntactic gap results in major complexity difference
  - Using approximation to bridge the gap

<table>
<thead>
<tr>
<th>DL ROQ (large subset of OWL2 DL)</th>
<th>DL EL++ (large subset of OWL2 EL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T ⊨ A</td>
<td>C ⊆ D</td>
</tr>
<tr>
<td>C ⊆ D</td>
<td>r ⊆ s, r₁ ∘ ... ∘ rₙ ⊆ s</td>
</tr>
<tr>
<td>a : C</td>
<td>(a, b) : r</td>
</tr>
</tbody>
</table>

N2EXPTIME-complete  PTIME-complete

Example: How Does it Work

- Represent non-OWL2-EL concepts with fresh named concepts
  - E.g., ∀r.C subClassOf D → Aᵥr.C subClassOf D
- Maintain semantic relations for these named concepts
  - complementary relations
  - cardinality relations
- Additional tractable completion Rules (on top of the EL ones), e.g.
  - Handling complement
    - E.g. B subClassOf C ⇒ ¬C subClassOf ¬B
TrOWL is a tractable OWL2 reasoning infrastructure developed at Aberdeen [ESWC2010]
- It supports many pluggable reasoners
- It works with OWL API, Protégé 4, Jena and Semantic Mediawiki

Quality guaranteed transformations
- Faithful approximate reasoning
  - Quill: OWL 2 DL -> OWL 2 QL (semantic approximation) [AAAI2007]
  - REL: OWL 2 DL -> OWL 2 EL (syntactic approximation) [AAAI2010]
- Divide and conquer
  - Modularisation [ISWC2009]
  - Forgetting [ESWC2008,ISWC2009b]
  - Distributed summarisation [AAA12012]

Parallel ABox reasoning in EL [JIST2011]

Stream reasoning in EL [CIKM2011]

Local Closed world reasoning with NBox [JTST]

Challenges:
- Computing one justification for OWL 2 DL is costly
- Computing all justifications is NP-complete even for OWL 2 tractable profiles

Solutions
- Approximate justifications:
  - relatively small subsets rather than minimal subsets
- Tractable new OWL2 profiles for computing a diagnosis [TST2010]
Architecture of Guidance Services

- Dom/SE-Oneto Bridge
  - Transformation
  - Integrated modelling
    - OWL Modelling
    - Dom/SE Modelling
- Ontology
- User prepares integrated model
- Guidance Services

TrOWL: Some New Results
Parallel ABox Reasoning for EL

- Context based, lock free parallel algorithm

<table>
<thead>
<tr>
<th>Worker 1</th>
<th>Worker 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( { a } \subseteq \exists r.A )</td>
<td>( { a } \subseteq \exists r.A )</td>
</tr>
<tr>
<td>( A \subseteq C )</td>
<td>( \exists r.A \rightarrow \exists r.C )</td>
</tr>
</tbody>
</table>

scheduled

processed

Evaluation

- Benchmark
  - VICODI ontology
  - NotGalen TBox + synthetic ABox generated by SyGENiA

- Environment
  - AWS EC2 cloud computing, 64-bit Linux, 7G RAM, each worker \( \approx 2.5\text{-}3.0 \text{ GHz} \)

<table>
<thead>
<tr>
<th>Off-the-shelf Reasoners</th>
<th>Pellet</th>
<th>HermiT</th>
<th>FaCT++ 1 worker</th>
<th>2 workers</th>
<th>4 workers</th>
<th>6 workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology</td>
<td>( \mathcal{L}_A )</td>
<td>( \mathcal{R}_A )</td>
<td>( \mathcal{L}_V )</td>
<td>( \mathcal{R}_V )</td>
<td>( \mathcal{L}_T )</td>
<td>( \mathcal{R}_T )</td>
</tr>
<tr>
<td>VICODI</td>
<td>184</td>
<td>10</td>
<td>29614</td>
<td>114164</td>
<td>2.014</td>
<td>9.971</td>
</tr>
<tr>
<td>NG-1</td>
<td>4236</td>
<td>8088</td>
<td>4.284</td>
<td>210.945</td>
<td>307.93</td>
<td>2.399</td>
</tr>
<tr>
<td>NG-2</td>
<td>12161</td>
<td>23976</td>
<td>9.342</td>
<td>757.379</td>
<td>timeout</td>
<td>3.025</td>
</tr>
<tr>
<td>NG-5</td>
<td>47756</td>
<td>118458</td>
<td>28.947</td>
<td>timeout</td>
<td>timeout</td>
<td>6.427</td>
</tr>
<tr>
<td>NG-8</td>
<td>78899</td>
<td>278365</td>
<td>63.833</td>
<td>timeout</td>
<td>timeout</td>
<td>12.604</td>
</tr>
<tr>
<td>NG-13</td>
<td>97995</td>
<td>665304</td>
<td>143.288</td>
<td>timeout</td>
<td>timeout</td>
<td>23.609</td>
</tr>
</tbody>
</table>
Local Closed World Reasoning

- **OWA (ontologies only cover key aspects of the world)**
  - Is Pepper Salad SpicyFood?
    - UNKNOWN

- **CWA (complete information about the world)**
  - Is Pepper Salad SpicyFood?
    - No, because
    - "SpicyFood={Curry Chicken, Spicy Grilled Shrimp}"

Explicit CWA vs. Implicit CWA

- **Is Spicy Grilled Shrimp the only SpicyFood?**
  - SpicyFood={Spicy Grilled Shrimp}?
  - No, because of background knowledge:
    - "MinorSpicyFood" is SpicyFood

- **CWA should support necessary reasoning**
Using SPARQL to support CWA

- Does **not** close ontological vocabulary.
- **Walk around**: Realised by testing for the absence of a pattern in a graph
- Get all food not known to be spicy

```sparql
SELECT ?dish
WHERE {
  ?dish rdf:type Food .
  FILTER NOT EXISTS { ?dish rdf:type SpicyFood}
}
```

- **Problem**: curry chicken is included in the answer set
  (if we use SPARQL without reasoning)

---

Negation as failure Box (NBox) [Ren et. al, 2010]

1. **To allow inference** w.r.t. the closed classes and properties
   - O=(T, A, N)
   - N = {SpicyFood, VegeFood} is the NBox in O
2. **To provide restricted forms of non-monotonic reasoning**
   - so that it does **not** increase the complexity of reasoning for OWL 2 DL
NBox Reasoning

- \((T, A, N) \models x: \neg B \iff (T, A) \not\models x:B\)
  - E.g., Salmon is neither VegeFood, nor SpicyFood
- \(\neg B\) is equivalent to \textbf{not} \(B\)
- \(B\) is equivalent to \(K_B\)

- Using classical reasoning to retrieve instance of predicates
  - E.g., Pepper is VegeFood
- Using \textit{nominals} to close predicates
  - E.g., VegeFood = \{Pepper, Salad\}
- Adding axioms back to ontology for incremental reasoning
  - Yuting orders Pepper!

Challenges for NBox Reasoning

- Challenge 1: Ontologies with nominals are harder to reason with
  - Using \textit{approximate reasoning} technologies [Ren et. al, 2010b] to reduce to a tractable DL
  - Identify safe consitions for tractable DLs, such as EL and DL-Lite [Lutz et al. 2012]

- Challenge 2: \textbf{Incremental reasoning} is usually difficult for expressive DLs
  - EL supports tractable incremental reasoning services! [Ren and Pan, 2011]

NBox reasoning is \textbf{already available} in the public release of TrOWL
Summary

- Promising approaches to reasoning with OWL 2 ontologies
  - Reasoning in tractable profiles
  - Approximate reasoning

- OWL-DBC
  - Connecting TrOWL with Oracle’s native inference engine
  - Exploit state of the art ontology reasoning and database optimisation technologies

Thank you!

Acknowledgement:
Yuan Ren
Nophadol Jekjantuk
Edward Thomas
Yuting Zhao