

A learning environment for building and evaluating ontologies: case study of 2013 ontology building competition

Adrian Groza[†], Varga Bernadette[†] and Mario Vacca^{*}

[†] Intelligent Systems Group
Department of Computer Science
Technical University of Cluj-Napoca, Romania
Adrian.Groza@cs.utcluj.ro

^{*} Faculty of Information Technology
Computer Science and Statistics
Sapienza University of Rome, Italy
Mario.Vacca@uniroma1.it

ELSE 2014, 24-25 April, Bucuresti, Romania



Outline

- 1 **Dimensions of Ontology Evaluation**
- 2 Formal Specification
- 3 Results
- 4 Conclusions



ELSE 2014, 24-25 April, Bucuresti, Romania

Contributions

- Ontology Building Competition (BOC) is a competition for developing ontologies, where the evaluation and the final ranking are computed automatically, based on five dimensions: i) structural, ii) semantic, iii) design patterns, iv) worst practices, and v) competency questions.
- Both ontology and automatic assessment are two important sectors for the 21st century school (The BPM4ED project: Designing 21st century schools)
- The learning environment gradually assess the work of the students. Hence, during the semester students are aware of the strong and weak points of their practical work, and also on the current hierarchy in the designed classes
- The learning tool was developed in LISP and exploits the RacerPro reasoning capabilities.

Explanation

Ontology Building Competition



BOC2013

Main motivation



Tracks at BOC2013

25 teams, students year 4, Knowledge Based Systems

- Human Resource Management at IT Companies (12)
- Argument and Explanation Between Human Agents (13)

| Week | Activity |
|------|---|
| 1 | Racer server, RacerPorter, Family ontology |
| 2 | Racer user guide, Racer reference manual |
| 3 | Defining competency questions |
| 4 | Identifying related ontologies |
| 5 | Defining main concepts |
| 6 | Defining main roles |
| 7 | Populating ontology with individuals |
| 8 | Ontology evaluation, <i>First pre-evaluation</i> |
| 9 | Refining ontology: rules |
| 10 | Refining ontology: ontology design patterns |
| 11 | Removing inconsistencies, <i>Second pre-evaluation</i> |
| 12 | Refining ontology: answering to competitors' competency questions |
| 13 | <i>BOC 2013 competition</i> |
| 14 | Documenting ontology in Latex and presenting it |

Structural evaluation

Ontologies are analysed as graph structures:

- size (how many concepts, roles, individuals)
- depth and breadth of hierarchy
- density (average branching)
- balance (are all area equally developed?)
- modularization (what modules are defined?)

Semantic evaluation

These metrics are referred to as knowledge base metrics:

- Consistency, Cycles
 - Expressiveness (DL used)
 - Comprehensiveness (extent of target domain covered, e.g answering to **competency questions** (CQ).)
- 1 defined by evaluators and make public before competition
 - 2 each candidate can submit such CQ, formalised in Racer syntax, which would be public available before competition.

In this way, incentive exists for candidates to formalise CQs for which their ontology is able to provide answers, but the opponents may have difficulties.

Competency questions

30 queries in NRQL language

The available positions in a specified location and in domain:

```
(retrieve (?job) (and (?x Company) (?job (hasLocation Cluj)) (?x  
?job postedJob)))
```

Positions that require some Java knowledge:

```
(retrieve (?job) (and (?x Company) (?job (requiresExperience  
Java)) (?x ?job postedJob)))
```

All the employees inside a company, with more then a year
experience in C language,

```
(retrieve (?employee) (and (?x Developer) (?x (and  
(hasExperience 1 year) (knows C)) (?employee ?x hasCareer  
)))
```

Ontology Design Patterns

With this dimension, we try to automatically identify ontology design patterns (ODPs)¹ used by the competitors.

PP models partition of concepts. In an ontology, a partition is a concept which is divided into several disjoint sub-concepts.

```
(define-concept P (or C0 C1))
```

```
(disjoint C1 C2)
```

¹<http://ontologydesignpatterns.org/>

Worst Practices

231 WP that is, solutions unsuitable for the designing ontologies.

| | |
|---------------------------------------|--|
| Relationship "is" | Confusion with subclass relationship, membership of a class or same as |
| Recursive definition | Using an ontology element in its own definition. |
| Undefined inverse relationships | Having inverse relationships in the ontology, but they are not explicitly defined as such. |
| Lazy elements | Leaf concepts or roles that never appear in the application and do not have any instances. |
| Missing disjointness | Lack of disjoint axioms between classes or between properties that should be defined as disjoint (even-odd, prime-composite) |
| Individuals are not Classes | For instance, Madrid is an instance not a class |
| Redundancy of Disjoint Relation (RDR) | For instance, the concept is explicitly defined as disjoint with parent concept and also with its child concept. |
| Missing subclasses | Number of classes which have only one subclass. The situation indicates that either: i) the hierarchy is under-specified or ii) the distinction between the subclass is not appropriate. |
| Extra subclasses | Number of classes with more than 25 subclasses. Such a class is candidate for additional distinctions |

Outline

- 1 Dimensions of Ontology Evaluation
- 2 Formal Specification**
- 3 Results
- 4 Conclusions



ELSE 2014, 24-25 April, Bucuresti, Romania

Formal specification - TBox of BOC2013

1. $Team \sqsubseteq (< 3)hasMember.Student \sqcap \exists hasDeveloped.Ontology$
2. $Competition \sqsubseteq \exists hasTrack.Track \sqcap (> 1)hasTeam.Team$
3. $Undergraduate \sqcup InternshipSt \sqcup DiplomaSt \sqsubseteq Student$
4. $Competitor \equiv Student \sqcap \exists hasTask.CreateOntology$
5. $Organiser \equiv Student \sqcap \exists hasTask.DefineMetrics$
6. $StructuralMetric \sqsubseteq Metric$
7. $SemanticMetric \sqsubseteq Metric$
8. $DesignPatterns \sqsubseteq Metric$
9. $WorstPractices \sqsubseteq Metric$
10. $SizeMetric \sqsubseteq StructuralMetric$
11. $StructuralPattern \sqsubseteq DesignPattern$
12. $ArchitecturalPattern \sqsubseteq StructuralPattern$
13. $LogicalPattern \sqsubseteq StructuralPattern$
14. $CorrespondencePattern \sqsubseteq DesignPattern$
15. $ContentPattern \sqsubseteq DesignPattern$
16. $ReasoningPattern \sqsubseteq DesignPattern$
17. $PresentationPattern \sqsubseteq DesignPattern$
18. $LexicoSyntacticPattern \sqsubseteq DesignPattern$
19. $CompetencyQuestion \sqsubseteq SemanticMetric$

Formal specification - ABox of BOC2013

(instance c2013 Competition)
(instance hr Track)
(instance ae Track)
(related c2013 hr has-track)
(related c2013 ae has-track)
(attribute-filler c2013 2013 holds-in)
(instance no-of-concepts SizeMetric)
(instance no-of-instances SizeMetric)
(instance mircea-marius Competitor)
(instance o1 Ontology)
(instance mister Team)
(related mister mircea-marius has-member)
(related mister o1 has-team)
(related o1 hr in-track)
(attribute-filler mister "user-mister" has-username)
(attribute-filler mister "pass-mister" has-password)
(attribute-filler o1 20.0 has-no-of-concepts-score)
(attribute-filler o1 15.0 has-no-of-individuals-score)

Querying the competition

```
;;all teams (concept-instances Team)
;;all students (concept-instances Student)
;;all competitors (concept-instances Competitor)
;;all organisers (concept-instances Organiser)
;;all tracks (concept-instances Track)
;; all ontologies in hr track (individual-fillers hr (inv in-track))
;; all ontologies in ae track (individual-fillers ae (inv in-track))
;;all tasks (concept-instances Task)
(describe-individual o1)
(describe-individual mister)
```

Outline

- 1 Dimensions of Ontology Evaluation
- 2 Formal Specification
- 3 Results**
- 4 Conclusions



ELSE 2014, 24-25 April, Bucuresti, Romania

Structural evaluation

- no. of concepts: [18,65] for HR, [16,63] for AE, average 37.2
- no. of roles [14,79], average 33
- no. of instances [10,723]
- instances/concept [0.10,13.54]
- re-using - 1 ontology



Evaluation results

Semantic evaluation

- all the ontologies were consistent
- 60% have cycles, average 7.63 cycles/ontology

Design patterns

- 33% have the partition design pattern

Worst practices

- average 18.63 empty classes

Competency questions

- 30 CQ in NRQL

Outline

- 1 Dimensions of Ontology Evaluation
- 2 Formal Specification
- 3 Results
- 4 Conclusions**



ELSE 2014, 24-25 April, Bucuresti, Romania

BOC2014

- <http://cs-gw.utcluj.ro/~adrian/BOC2013/>
- For <http://cs-gw.utcluj.ro/~adrian/BOC2014/> we are considering:
 - 30 June 2014
 - tracks: **Romanian Tourism Ontology** (17) and **Vehicular Networks** (13)

Thank you!

