ONSET: Automated Foundational Ontology Selection and Explanation

Zubeida Casmod Khan, C.Maria Keet

School of Mathematics, Statistics, and Computer Science, University of KwaZulu-Natal, South Africa and UKZN/CSIR-Meraka Centre for Artificial Intelligence Research, South Africa
zkhan@csir.co.za.keet@ukzn.ac.za

FAIR'12, Pretoria, South Africa, 15-16 November 2012

Overview

Introduction I

- Foundational ontologies are beneficial in ontology development [Borgo & Lesmo(2008), Keet(2011)].
- Benefits include:
  - Increases semantic interoperability between systems.
  - A starting point in ontology development.
  - Speeds up ontology development.
  - Improves quality of the system at hand.
- Although there are many foundational ontologies available:
  - DOLCE
  - BFO
  - GFO
  - SUMO
  - YAMATO and more

Introduction II

- Most developers do not know how to apply them.
- Reasons for not using them include:
  - Described as being either too abstract or expressive.
  - Too time consuming to learn.
  - Different foundational ontologies have conflicting philosophies and properties.

Problems to be addressed

- Most ontology development methodologies such as METHONTOLOGY, NeOn and On-To-Knowledge, do not include the usage of foundational ontologies.
- Ontology developers in general do not know which ontology to use.
  1. It is unknown which ontology is best for which scenario.
  2. Why should a particular foundational ontology be used for a scenario- parameters associated with it?

Solution of problem

We help with foundational ontology selection by:

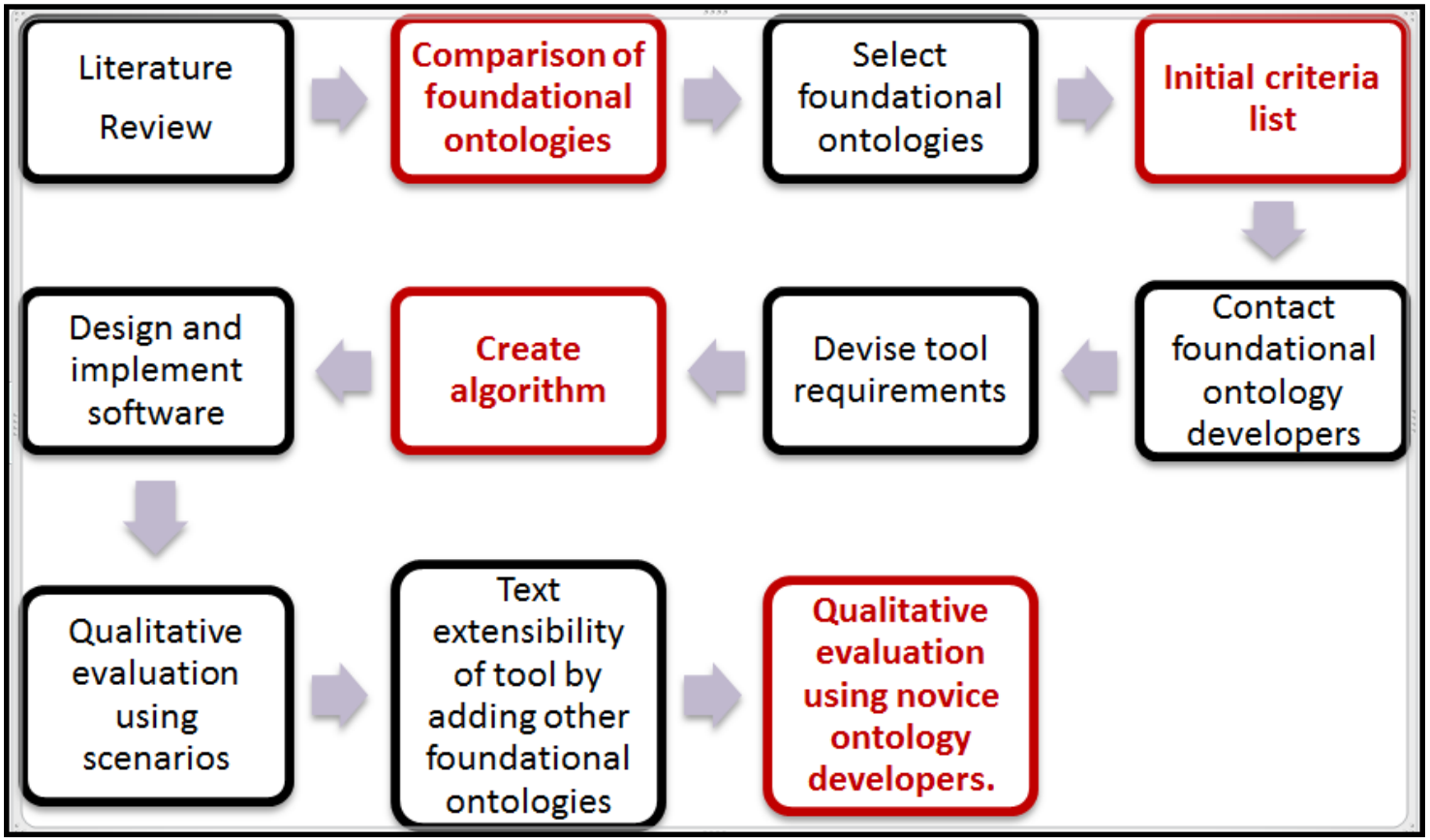
- Comparing existing foundational ontologies.
- Identifying important ontological criteria.
- Creating a method for automated foundational ontology selection using software.

This allows the user to:

1. Select a suitable foundational ontology based on the developer's input.
2. Generate an automated explanation for the selection.

We have realised the solution with ONSET.

Methodology



Literature Review

We conducted a literature review of about 60 published works. These included:

- Publications and documentation by the foundational ontology developers (e.g., the Wonderweb deliverable [Masolo et al.(2003)].)
- Comparative studies of foundational ontologies which were found to have focussed on only a limited amount of topics. [Ahmad & Lindgren(2010), Semy et al.(2004)]
- Case reports on usage in domain ontologies and applications [Semy et al.(2004), Smith et al.(2007)].

Foundational ontologies in ONSET

At present, we have the following foundational ontologies:

- DOLCE
- BFO
- GFO
- SUMO

GFO and SUMO were included in ONSET at a later stage, thereby proving ONSET's extensibility.

Categories of criteria and comparison I

The general list of criteria is divided into five categories:

- **Ontological Commitments**: Philosophical choices taken by foundational ontologies. e.g., an ontology of particulars or universals.
- **Representation Language**: Languages used to represent a domain ontology. e.g., OWL DL, KIF
- **Software engineering properties**: General properties associated with various foundational ontologies. e.g., dimensions, modularity
- **Subject Domain**: Existing domains expressed using foundational ontologies. e.g., engineering, geographical.

Categories of criteria and comparison II

- **Applications**: Application scenarios of domain ontologies e.g., the Semantic Web, ontology based data access.

We have received feedback from the creators of DOLCE, BFO and GFO for their lists.

Foundational Ontology Comparative Tables

Term	DOLCE	BFO	GFO	SUMO
Universals vs. Particulars	Particulars	Universals	Universals, concepts, and symbols	Universals and particulars
Descriptive vs. Realist	Descriptive	Realist	Descriptive and Realist	Descriptive
Multiplicative vs. Reductionist	Multiplicative	Reductionist	Unclear	Multiplicative
Endurantism vs. Perdurantism	Endurantism and perduran-tism	Endurantism and perduran-tism	Endurantism and perduran-tism	Endurantism and perduran-tism

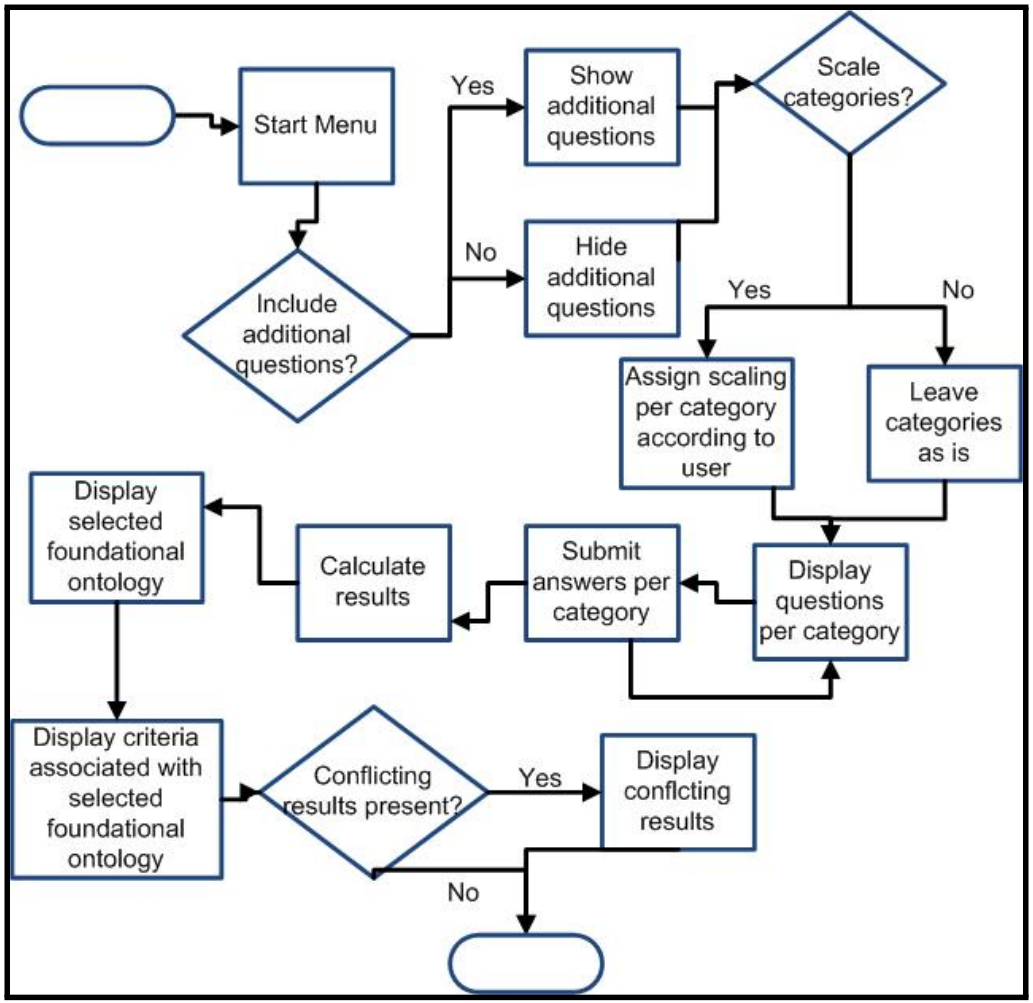
Table: A section of the comparative table of ontological commitments.
For full comparison see `http://www.meteck.org/files/onset/CriteriaLists.pdf`

Requirements

Outline of functional requirements:

1. Store answers
2. Selection
3. Explain selection
4. Conflicting answers
5. Provide ontology references
6. Additional questions
7. Optional scaling
8. Extensibility

Flow Diagram



Algorithm outline I

- Decides whether additional questions are to be displayed.
- Stores a scaling value for each category.

```
   output: Include additional questions?
 1 if input is yes then
 2 |   Show additional questions
 3 else
 4 |   Hide additional questions
 5 end
   output: Assign scaling per category?
 6 if input is yes then
 7 |   for i ← 0 to numOfCategories do
 8 |   |   Read scaling value;
 9 |   |   Store scaling value in ScalingValues[i];
10 |   end
11 else
12 |   ScalingValues[i] = 1;
13 end
```

Algorithm outline II

- Displays questions per category.
- Displays options per category.
- Applies scaling per category.
- Increments counter variable for respective foundational ontology.
- Stores the user's answers for each question.

```
14 for i ← 0 to numOfCategories do
15 |   for j ← 0 to numOfQuestionsPerCategory do
16 |   |   Display question;
17 |   |   Display options;
18 |   |   if option corresponds to DOLCE then
19 |   |   |   DOLCECount = DOLCECount+(1* ScalingValues[i]);
20 |   |   |   DOLCEANSWERS[j]= option text;
```

Algorithm outline III

- Compares counter variables to find largest counter variable.
- Uses the largest counter variable to select the appropriate foundational ontology.
- Uses arrays of ontological choices to display the motivation for the selected foundational ontology.

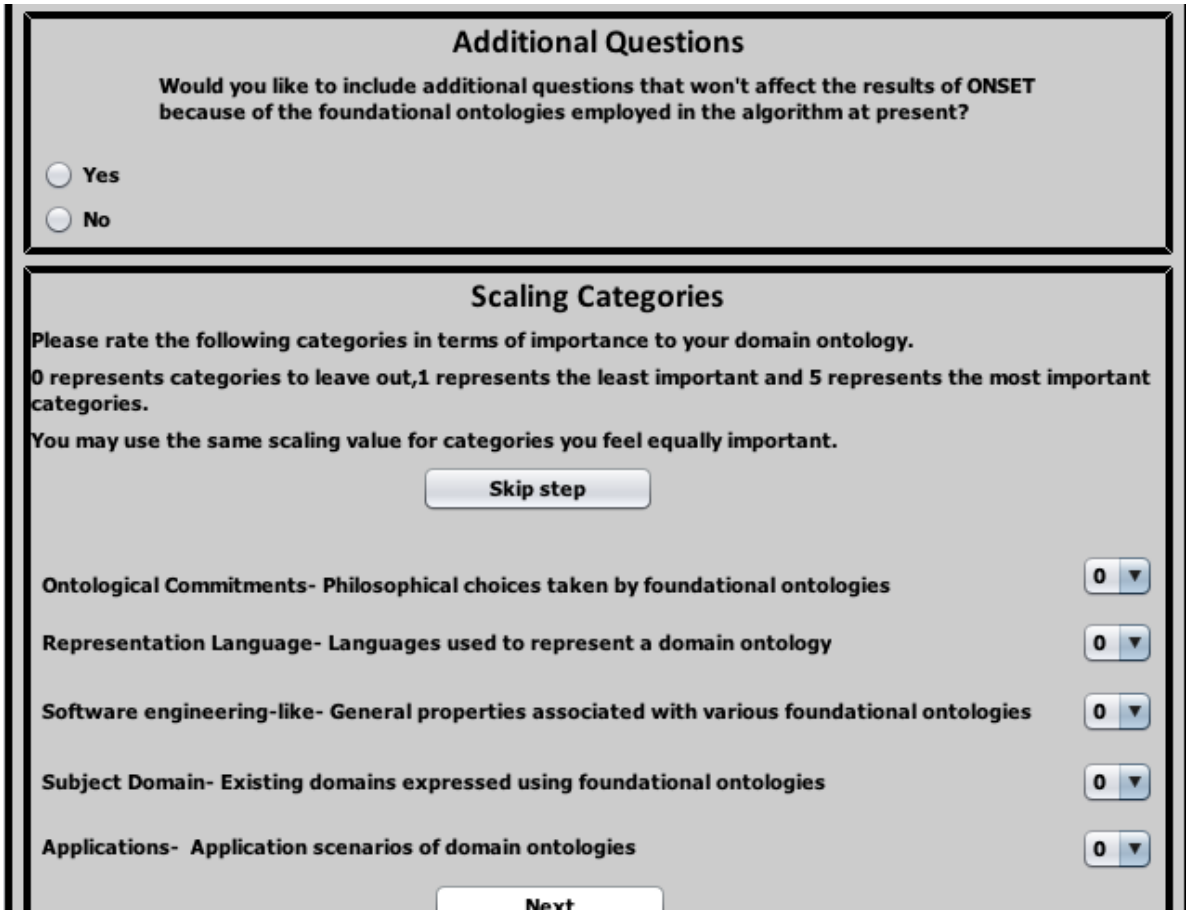
```
1 if DOLCECount > (BFOCount AND GFOCount AND SUMOCount) then
  |   output: Selected FO is FO
  |   output: Reasons why DOLCE was chosen:
2 |   for i ← 0 to DOLCEAnswers.length do
3 |   |   if DOLCEAnswers[i]! = null then
  |   |   |   output: DOLCEAnswers[i]
4 |   |   end
5 |   end
```

Algorithm outline IV

- Displays a list of existing subject domain ontologies for the selected foundational ontology.
- Uses arrays of ontological choices to display conflicting results.

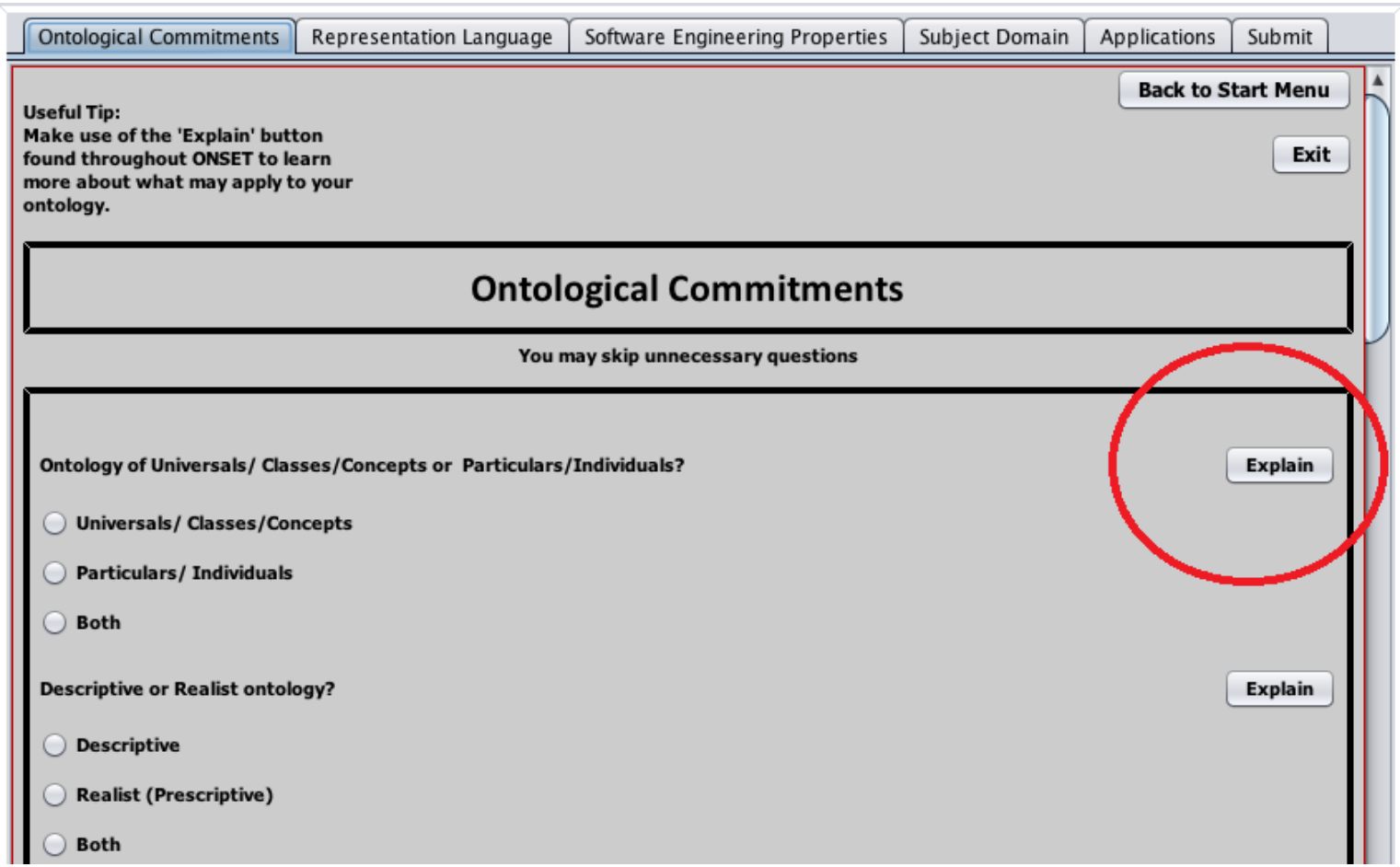
```
    output: Existing ontologies with the specified subject domain
6   for k ← 0 to DOLCEDomain.length do
        output: DOLCEDomain[k]
7   end
8   if BFOAnswers OR GFOAnswers OR SUMO not empty then
        output: Conflicting results: The tool has detected that some of your
                criteria matches with other foundational ontologies.
9       for i ← 0 to BFOAnswers.length do
10          if BFOAnswers[i]! = null then
                output: BFOAnswers[i]
11          end
```

Using ONSET I

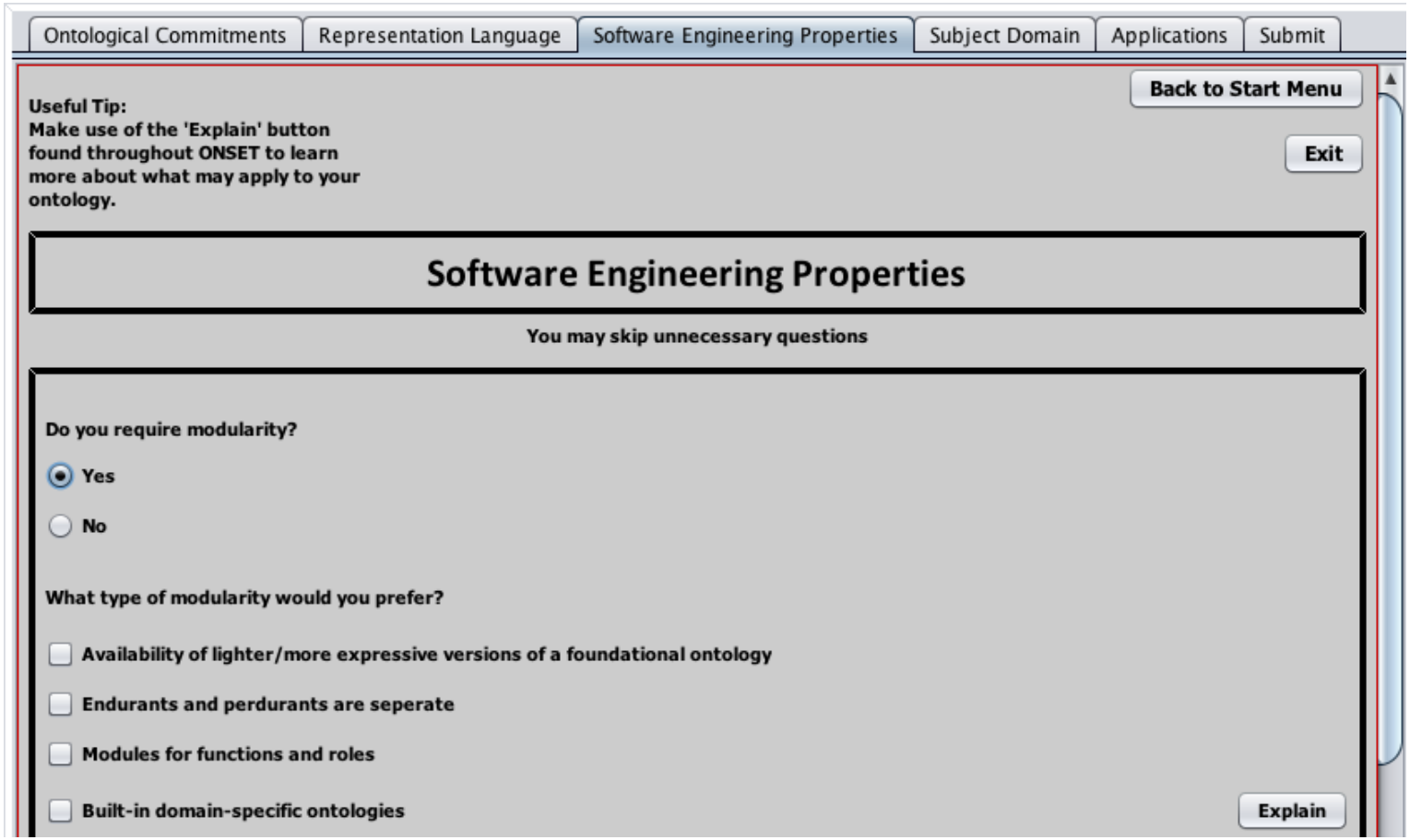


Using ONSET II

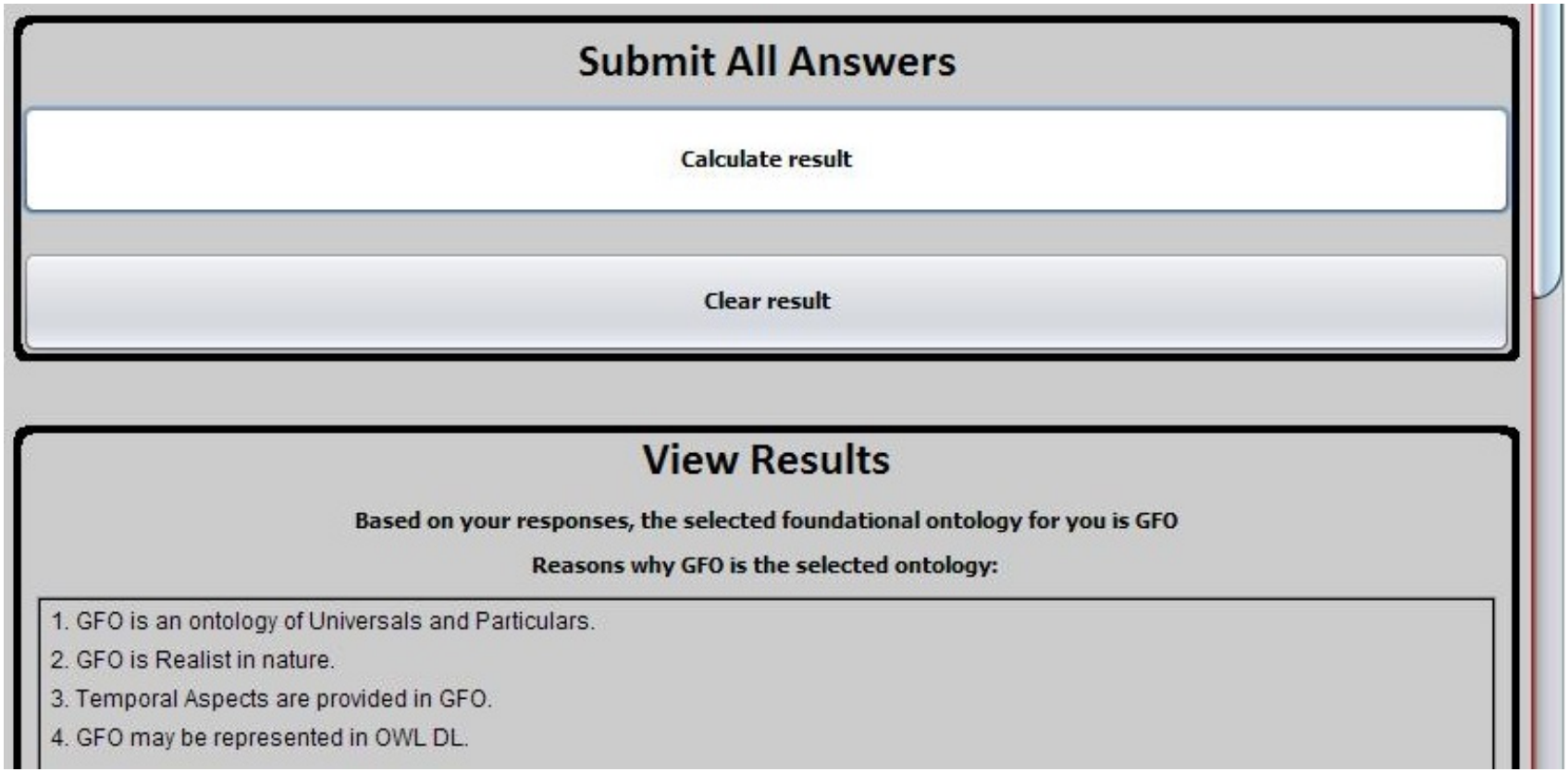
Using ONSET III



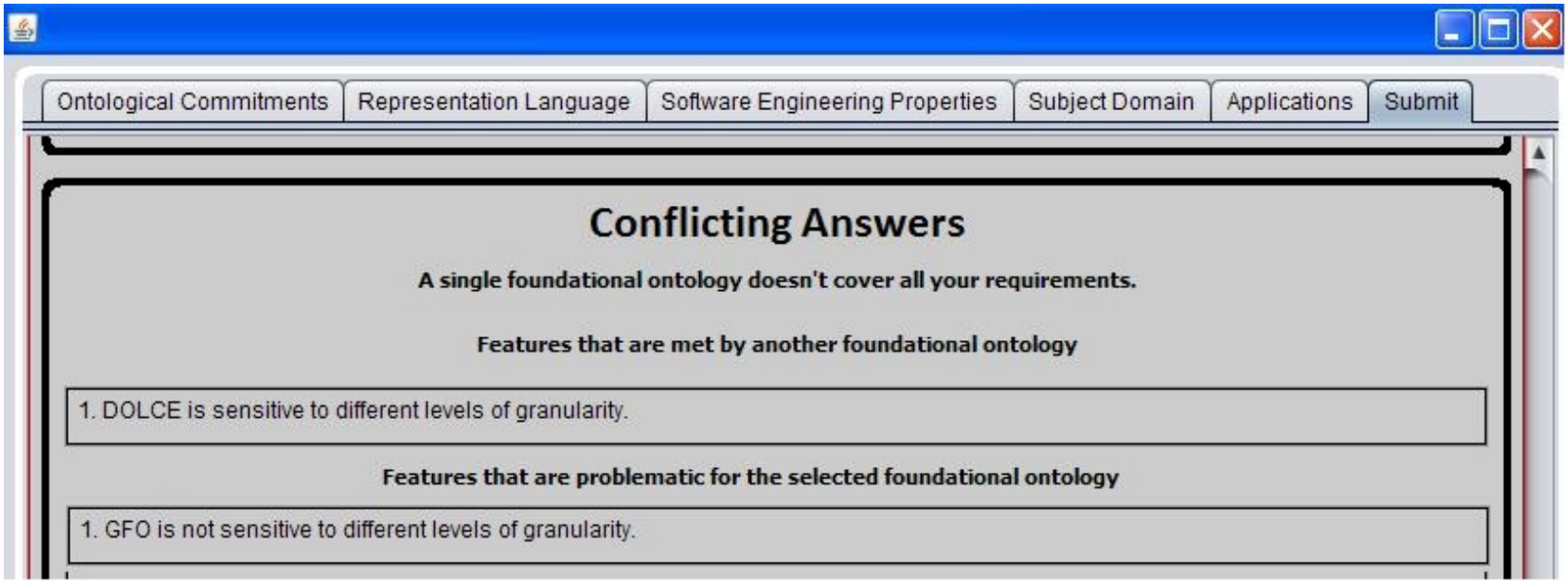
Using ONSET IV



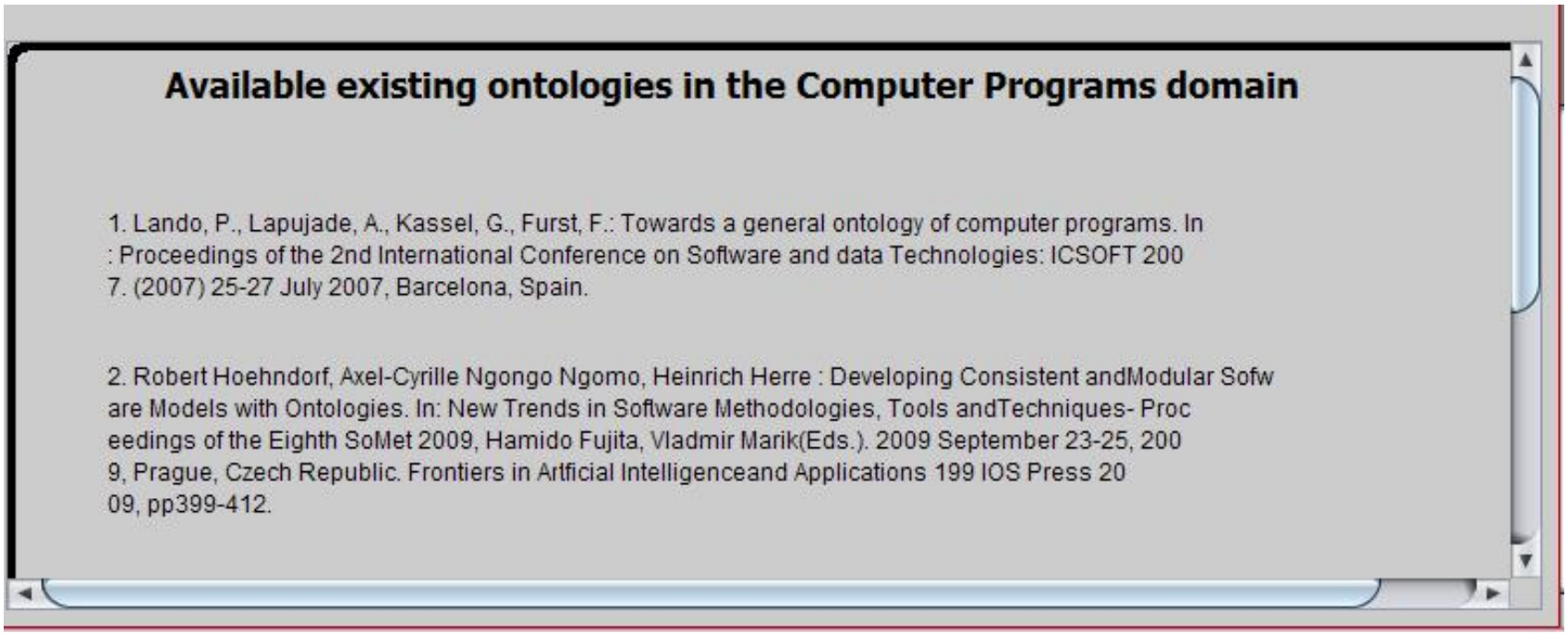
Using ONSET V



Using ONSET VI



Using ONSET VII



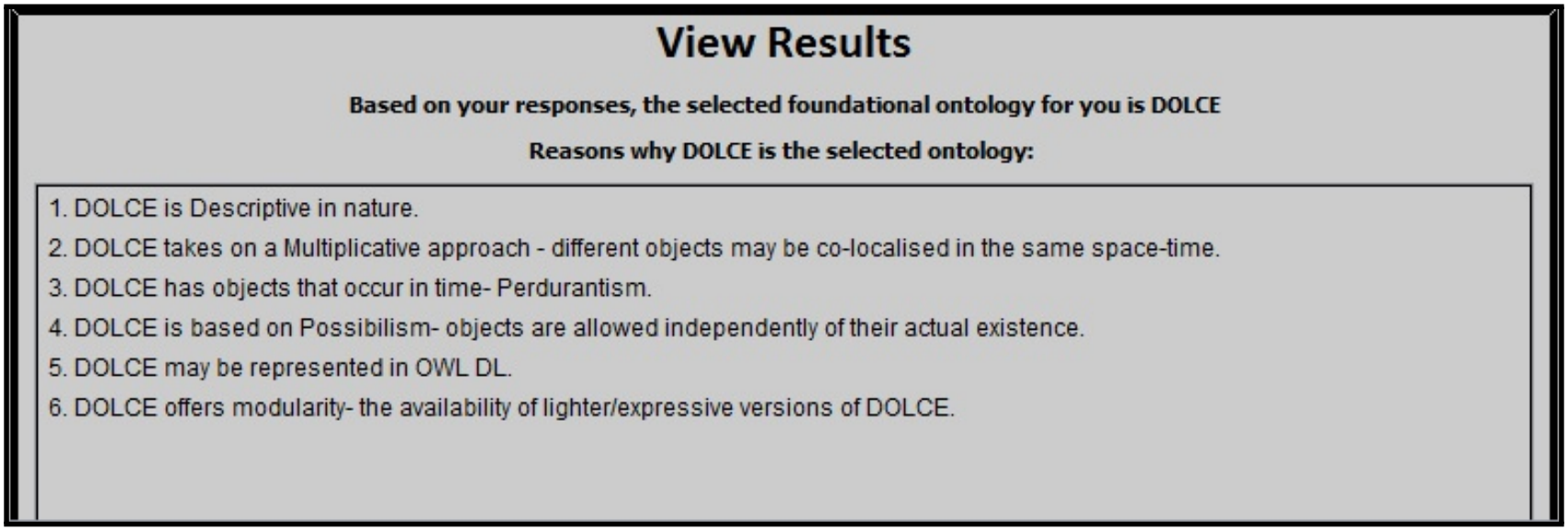
Existing scenario

We use an existing scenario [Oberle(2006)] for ontology development to evaluate ONSET's functionality.

Example (Semantic Management of Middleware)

For an application of the semantic web, the requirements are : descriptiveness, a multiplicative approach, possibilism, perdurantism, modularity (lightweight versions) and an executable language.

Output of ONSET: Existing scenario



View Results

Based on your responses, the selected foundational ontology for you is DOLCE

Reasons why DOLCE is the selected ontology:

1. DOLCE is Descriptive in nature.
2. DOLCE takes on a Multiplicative approach - different objects may be co-localised in the same space-time.
3. DOLCE has objects that occur in time- Perdurantism.
4. DOLCE is based on Possibilism- objects are allowed independently of their actual existence.
5. DOLCE may be represented in OWL DL.
6. DOLCE offers modularity- the availability of lighter/expressive versions of DOLCE.

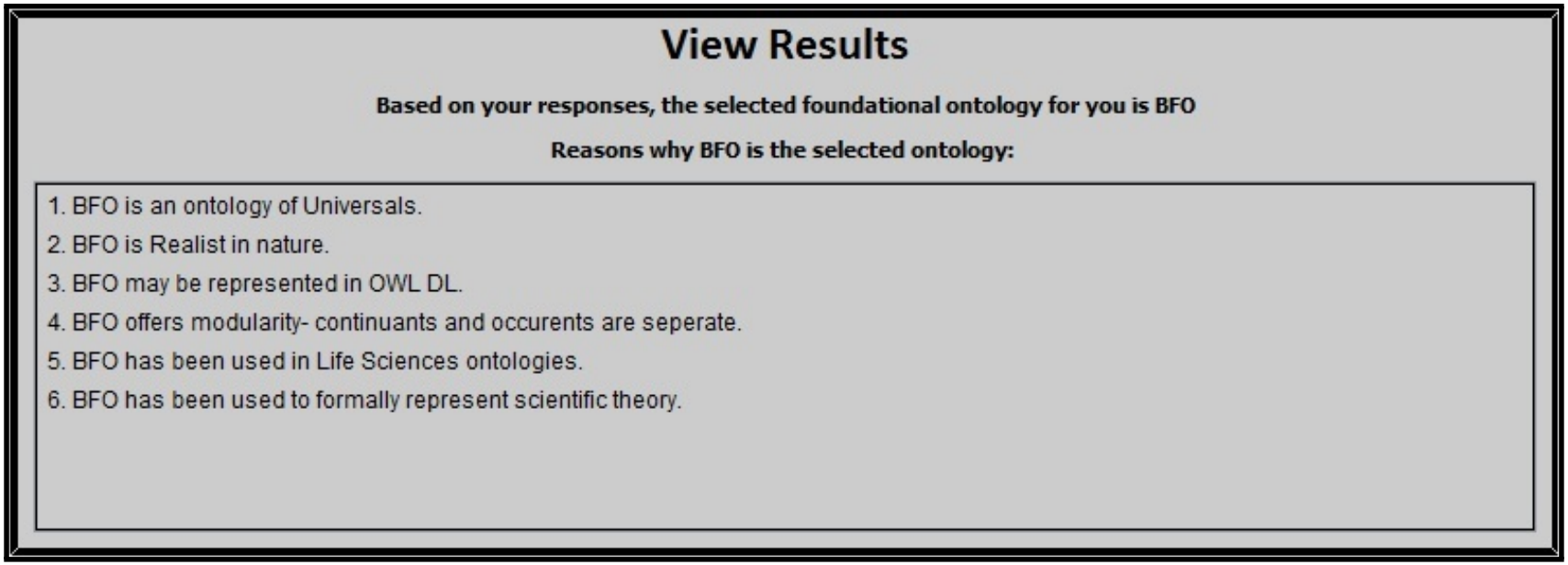
Simulated scenario

We use a simulated scenario to assess the scaling in ONSET.

Example (Scaling effects)

Assume that there is an ontology to be created with the following requirements: an ontology of universals, realist in nature, to be represented in OWL DL, modularity (endurants and perdurants separate, built-in domain specific ontologies), applying it to formally represent a scientific theory and a domain of life sciences.

Output of ONSET: Simulated scenario without scaling

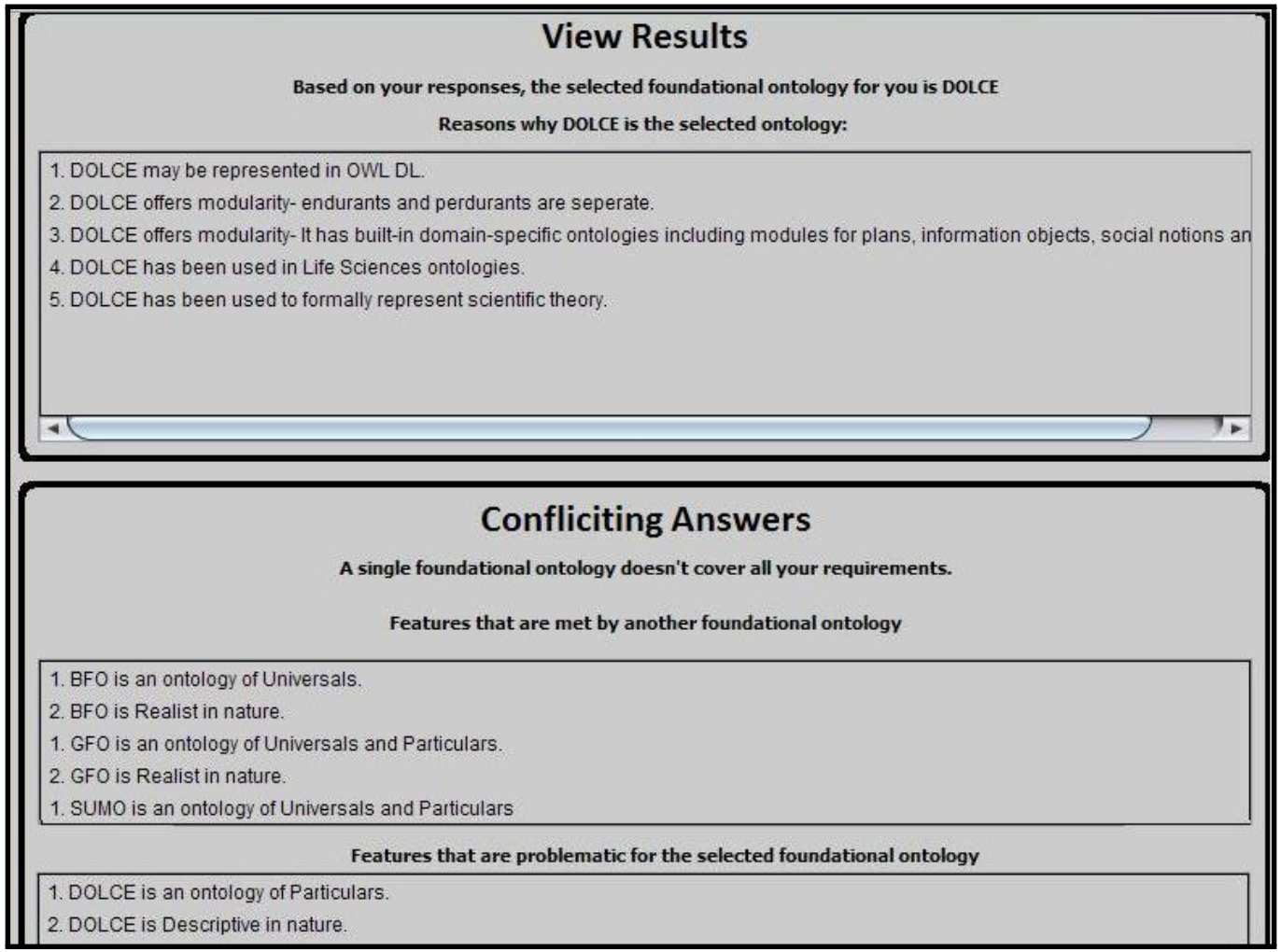


Scaling categories

We then use the same input but assign priorities for the categories in the scaling section in ONSET. We assign:

- ontological commitments = 1
- representation languages = 5
- software engineering properties = 3
- subject domain = 5
- applications = 4

Output of ONSET: Simulated scenario with scaling



View Results

Based on your responses, the selected foundational ontology for you is DOLCE

Reasons why DOLCE is the selected ontology:

1. DOLCE may be represented in OWL DL.
2. DOLCE offers modularity- endurants and perdurants are seperate.
3. DOLCE offers modularity- It has built-in domain-specific ontologies including modules for plans, information objects, social notions an
4. DOLCE has been used in Life Sciences ontologies.
5. DOLCE has been used to formally represent scientific theory.

Conficliting Answers

A single foundational ontology doesn't cover all your requirements.

Features that are met by another foundational ontology

1. BFO is an ontology of Universals.
2. BFO is Realist in nature.
1. GFO is an ontology of Universals and Particulars.
2. GFO is Realist in nature.
1. SUMO is an ontology of Universals and Particulars

Features that are problematic for the selected foundational ontology

1. DOLCE is an ontology of Particulars.
2. DOLCE is Descriptive in nature.

ONSET's experimental evaluation I

Does ONSET makes a difference in ontology selection? (correctness and time)

Set-up

- Lecture on foundational ontologies (1.5 hours).
- Announcement of experiment to be held in a week.
- Divide class into groups A an B randomly.
- Give both groups a soft-copy of instructions and tasks:
  - Group A: complete the given tasks in the prescribed time by using their lecture notes and resources found on the internet.
  - Group B: complete the given tasks in the prescribed time using same resources as Group A and by and using ONSET.

ONSET's experimental evaluation II

- Participants upload answers to capture timing.
- Evaluate the tasks:
  - Assessing and comparing the quality of answers of each group.
  - Comparing the time taken to complete the tasks of each group
  - Collecting and analysing user opinions of all participants.
- We use an accuracy measure to assess the quality of answers.

Tasks in experiment

- Five scenarios for domain ontology development,where for each scenario, participants had to select a foundational ontology and motivate.
- Open-ended questions asking their opinion on the experiment.
- For Group B only, feedback on the tool.

ONSET's experimental evaluation III

Example (Use Case 1: Ontology of heart disease)

You are to develop an ontology of heart diseases. The ontology must capture the intrinsic nature of the real world only. As such, entities that are not extended in space and time must not be found in the ontology. Possible future conditions that are predicted and previous conditions of the heart must be modelled in the ontology. Since it is a biological ontology, you wish to register it with the OBO foundry to allow reuse and integration with other ontologies. This ontology must be modelled in OWL 2 EL.

For the above use case, select a foundational ontology and provide all possible reasons for that choice.

ONSET's experimental evaluation IV

Example (Use Case 1: Ontology of heart disease)

You are to develop an ontology of **heart diseases**. The ontology must capture the **intrinsic nature of the real world only**. As such, **entities that are not extended in space and time** must not be found in the ontology. **Possible future conditions** that are predicted and **previous conditions** of the heart must be modelled in the ontology. Since it is a biological ontology, you wish to **register it with the OBO foundry** to allow reuse and integration with other ontologies. This ontology must be **modelled in OWL 2 EL.**

ONSET's experimental evaluation V

Summary of use case

- Heart diseases = biomedical domain
- Intrinsic nature of the real world only = realist ontology
- Entities that are not extended in space and time = abstract entities
- Possible future conditions and previous conditions = eternalist stance
- Register it with the OBO foundry = register it with the OBO foundry
- Modelled in OWL 2 EL = modelled in OWL 2 EL

Results I

Scenario	**Group A Average**	**Group B (ONSET) Average**
1.Ontology of heart diseases	22%	52%
2.Ontology for the integration of databases of a manufacturing factory	16%	43%
3.Ontology of economic systems	20%	48%
4.Ontology of banks	16%	37%
5.Ontology for conceptual data models	8%	51%
All Scenarios	**16%**	**46%**

Table: Comparing the accuracy averages of answers by Group A and Group B.

Results II

- Group B (ONSET) was about 3 times as accurate as Group A in ontology selection.
- Group A had completed their tasks in, on average 105 minutes while Group B (ONSET) had completed theirs in, on average 97 minutes.
- ONSET does assist in ontology selection and the time taken for ontology selection is somewhat less when using ONSET.
- Group B (ONSET) was impressed with ONSET.
- All participants from Group B (ONSET) who provided feedback felt that the task would not have been as easily performed without the help of ONSET.

Results III

- ONSET was well received at MAIS'11 resulting in it being used in ontology engineering courses at UKZN and UNISA.
- We received positive feedback from the foundational ontology developers.

Discussion I

- ONSET lets the user select a foundational ontology and provides an explanation for it.
- The method designed can be used at the start of ontology development, during improvement of an existing ontology with a foundational ontology, and for ontology interoperability scenarios.
- ONSET raises some new questions and issues:
  - If the envisioned WFOL [Masolo et al.(2003)] existed, foundational ontology selection would have been less of an issue.
  - The notion of 'foundational' is based on the assumption that there is just one foundational ontology.

Discussion II

- However, we observe that conflicting answers may arise to the extent where sometimes more than one foundational ontologies is applicable.
- Furthermore, the foundational ontology choice may change when the scaling of categories is altered.

- ONSET can be used to investigate these new issues that have emerged.

Conclusion and future works

- ONSET solves the problem of selecting a foundational ontology, by using user requirements and foundational ontology criteria.
- It provides the ontology developer with the motivation for using a particular foundational ontology.
- The compiled, verified criteria list and software implementation of them is the first paper-based and software-based approach in foundational ontology selection.
- The functionality, quality and usefulness of ONSET has been evaluated, with positive results.

Conclusion and future works

- Add YAMATO and gist foundational ontologies to ONSET.
- Extend ontological commitments tab with CQs from NeOn methodology.
- We are currently working on a foundational ontology repository which ONSET will be added to.
- Including ONSET in an existing ontology development methodology.

References I

Ahmad, F. & Lindgren, H. (2010).
Selection of foundational ontology for collaborative knowledge modeling in healthcare domain.
In Proc. of AIMSA'10, vol. 6304 of *LNCS*, pp. 261–262. (Springer).

Borgo, S. & Lesmo, L. (2008).
The attractiveness of foundational ontologies in industry.
In Proc. of FOMI'08, pp. 1–9. (Amsterdam, The Netherlands: IOS Press).

Keet, C. M. (2011).
The use of foundational ontologies in ontology development: an empirical assessment.
In Proc. of ESWC'11, vol. 6643 of *LNCS*, pp. 321–335. (Springer).
Heraklion, Crete, Greece, 29 May-2 June, 2011.

Masolo, C. et al. (2003).
Ontology library.
WonderWeb Deliverable D18 (ver. 1.0, 31-12-2003).
http://wonderweb.semanticweb.org.

Oberle, D. (2006).
Semantic Management of Middleware, vol. 1 of *Semantic Web and Beyond*.
(New York: Springer).

References II

Semy, S. K. et al. (2004).

Toward the use of an upper ontology for us government and us military domains: An evaluation. Technical Report MTR 04B0000063, The MITRE Corporation.

Smith, B. et al. (2007).

The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. Nature Biotechnology, 25, 1251–1255.

- To download ONSET (.jar file), user manual, complete criteria lists and use cases go to `http://www.meteck.org/files/onset/`.