Mapping WordNet to Basic Formal Ontology using the KYOTO ontology

Selja Seppälä

Department of Philosophy, University at Buffalo

Summary

Background: Ontologies are used with natural language processing (NLP) tools to carry out ontology-related text manipulation tasks.

These tasks require coupling lexico-semantic resources with ontologies, such as upper-level ontologies that are extended by domain-specific ontologies.

Issue: No lexico-semantic resource available for the Basic Formal Ontology (BFO), widely used in the biomedical domain.

Hypothesis: A large portion of WordNet synsets, especially nouns and verbs, can be semi-automatically mapped to BFO using simple mapping rules.

Goal: Semi-automatically mapping WordNet 3.0 to BFO 2.0.

Outcome: A lexico-semantic resource that could be used in NLP tools.

Discussion 1

Indirect mapping

Issue: WordNet is too large to be manually mapped to BFO.

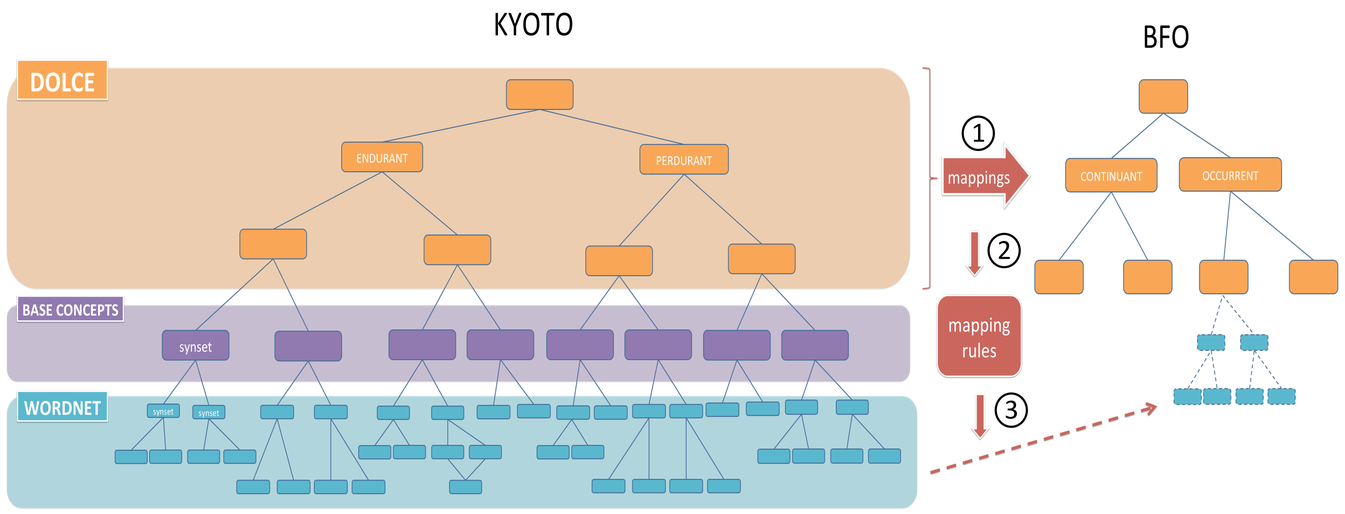
Possible solution: Mapping top levels of WordNet to BFO and propagating mappings downwards.

Subsequent issues:

- WordNet fails to comply with basic ontological principles.
- Would only cover nouns and verbs.

Solution: Use the KYOTO ontology, which also includes adjectives.

Ontological and lexical resources



WORDNET
Sets of synonymous English words (synsets) linked by semantic relations

BASE CONCEPTS
Subset of synsets subsuming all WN synsets and linked to ontology classes

DOLCE
Based on DOLCE-Lite-Plus, version 3.9.7, and OntoWordNet

Method

1. Mappings from DOLCE to BFO 2.0
2. Ruleset mapping KYOTO types to BFO based on 1
3. Get synsets, their base concepts and KYOTO types, and apply rules to map WordNet to BFO

Implementation example

Input

`immunity.n.02`

```
'Kyoto#condition__status-eng-3.0-13920835-n',
'Kyoto#state-eng-3.0-00024720-n',
```

```
'ExtendedDnS.owl#situation',
'ExtendedDnS.owl#non-agentive-social-object',
'ExtendedDnS.owl#social-object',
'DOLCE-Lite.owl#non-physical-object',
'DOLCE-Lite.owl#non-physical-endurant',
'DOLCE-Lite.owl#endurant',
'DOLCE-Lite.owl#spatio-temporal-particular',
'DOLCE-Lite.owl#particular'
```

Program tests if string in rules matches element in list

```
...
'#non-agentive-social-object > disposition'
'accomplishment > process'
'noun.act > process'
...
```

Output

```
immunity.n.02 > DISPOSITION
```

Evaluation and preliminary results

pos	synset	EXPECTED	MAPPED	Full mapping (correct BFO type)	Partial mapping (superordinate BFO type)	Incorrect WN-BFO mappings	No mapping
v	bring_around.v.02	process	occurrent		1		
v	immunize.v.02	process	process	1			
n	reserve.n.04	disposition	NO MATCH IN BFO				1
n	donor.n.02	role	role	1			
v	bandage.v.02	process	process	1			
a	neurotropic.a.01	disposition	quality			1	
n	venipuncture.n.01	process	process	1			
n	gauze.n.01	object	object	1			
v	vet.v.03	process	process	1			
n	coccidiosis.n.01	disposition	disposition	1			
n	contraindication.n.01	gdc	independent continuant		1		
n	plexor.n.01	object	object	1			

- Manual evaluation on 106 synsets marked 'medicine' in KYOTO.
- All nouns and verbs were correctly categorized.
- Incorrect mappings were mostly adjectives.

	Full mapping (correct BFO type)		Partial mapping (superordinate BFO type)		Total	
	occ.	%	occ.	%	occ.	%
Correct WN-BFO mappings	67	63.2	9	8.5	76	71.7
Incorrect WN-BFO mappings	-	-	-	-	29	27.4
No mapping	-	-	-	-	1	0.9
				Total	106	100

Discussion 2

Non-trivial mappings

Issues:

- DOLCE and BFO categories don't all align and their axioms may conflict.
- WordNet includes synsets that, in BFO terms, do not independently refer (e.g. `positive.a.04`).

Solutions:

- Extend the coverage of the rules.
- Ignore the axiomatizations.

Rationale:

- Work neither aimed at mapping DOLCE to BFO, nor at axiomatizing WordNet.
- Instead, answering the question: to what types of entities do WordNet synsets refer?
- Even a partial mapping should cover a large portion of WordNet, leaving a smaller subset of problematic cases.

Challenge: Providing BFO-compliant interpretations of unmatched WordNet synsets.

Conclusion and future work

Preliminary results:

Encouraging, but more work is needed to see if the method scales to the full WordNet.

Future work:

- Extending the evaluation set of medical synsets using hyponymy relations and other domain resources.
- Carrying out more thorough evaluations.
- Augmenting the mapping rules by exploiting other resources.

References

Arp, R., Smith, B., and Spear, A. (August 2015) *Building Ontologies with Basic Formal Ontology*, Cambridge, MA: MIT Press.

Fellbaum, C., editor (1998). *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

Hicks, A. and Herold, A. Cross-Lingual Evaluation of Ontologies with Rudify. In *Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Communications in Computer and Information Science, volume 128, 151–163, Springer, 2011.

Laparra, E., Rigau, G., and Vossen, P. (2012). Mapping WordNet to the Kyoto ontology. In *LREC*, 2584–2589.

Smith, B., Almeida, M., Bona, J. *et al.* (2012). *Basic Formal Ontology 2.0: DRAFT SPECIFICATION AND USER'S GUIDE.*

Vossen, P., Rigau, G., Agirre, *et al.* (2010). KYOTO: an open platform for mining facts. In *Proceedings of the 6th Workshop on Ontologies and Lexical Resources*, 1–10.

FNSNF Swiss National Science Foundation

University at Buffalo The State University of New York