

Towards Cross-checking WordNet and SUMO Using Meronymy

Javier Álvarez **German Rigau**

LoRea & IXA Groups
Computer Languages and Systems Department
University of the Basque Country (UPV/EHU)



- 1 Introduction
- 2 Cross-checking WordNet and Adimen-SUMO
- 3 Some Experimental Results
- 4 Conclusions and Future Work

Cross-checking knowledge sources

- This work is an initial study about:
 - ▶ Knowledge representation
 - ▶ Common Sense (world knowledge)
 - ▶ Reasoning
- In particular, we focus on:
 - ▶ WordNet (Fellbaum, 1998)
 - ▶ SUMO (Niles and Pease, 2001)
 - ▶ WN-SUMO Mapping (Niles and Pease, 2003)
- We expect all these knowledge sources to encode **correct world knowledge** (true knowledge).
- Despite being created manually, they are **not free of errors** and discrepancies.
- We apply a new Black-box strategy (Álvarez et al., 2017b) to the meronymy information encoded in these resources.

SUMO (Niles and Pease, 2001)

- IEEE Standard Upper Ontology Working Group
- SUMO syntax goes beyond first-order logic (FOL)
- SUMO cannot be directly used by FOL Automated Theorem Provers (ATPs) without a suitable transformation
- Different transformations of SUMO into FOL:
 - ▶ TPTP-SUMO (Pease and Sutcliffe, 2007)
 - ▶ Adimen-SUMO (Álvarez et al., 2012)

Adimen-SUMO I

- Following the line of (Horrocks and Voronkov, 2006)
- Obtained by applying a reengineering process to SUMO
 - ▶ With the help of ATPs (*Automated Theorem Provers*)
 - ▶ Around an 88% of the *core* of SUMO (top and middle levels) is translated
 - ▶ Domain ontologies are not used (by now)
 - ▶ The resulting ontology can be used in tasks that involve reasoning with commonsense knowledge
- The process of manually debugging the ontology is very costly
 - ▶ Only 64 manually created tests
 - ▶ Example:

```
( =>  
  (and  
    ($instance ?BRAIN Brain)  
    ($instance ?PLANT Plant))  
  (not  
    (properPart ?BRAIN ?PLANT)))
```

Adimen-SUMO II

- We have proposed different methodologies for the automatic debugging ontologies like Adimen-SUMO
 - ▶ Black-box testing strategies (Álvarez et al., 2015, 2017b)
 - ▶ White-box testing strategies (Álvarez et al., 2017a)
- More than 100 axioms from Adimen-SUMO has been improved

Black-box Testing I

- Introduced in (Álvarez et al., 2015) and fully described in (Álvarez et al., 2017b)
- Adaptation of the methodology for the design and evaluation of ontologies introduced in (Grüninger and Fox, 1995)
- Based on the use of **Competency Questions (CQs)**:
 - ▶ Problems that an ontology is expected to answer
- Its application is automatic by means of the use of ATPs
- Classification of (dual) problems (truth and falsity tests):
 - ▶ *Passing*: the ATPs are able to demonstrate a truth test
 - ▶ *Non-passing*: the ATPs are able to demonstrate a falsity test
 - ▶ *Unknown*: the ATPs produce no answer within a time limit

Black-box Testing II

- CQs are automatically created on the basis of few **Question Patterns** (QPs) by exploiting WordNet and its mapping into SUMO
- In (Álvarez et al., 2017b):
 - ▶ *antonym* and *event* (*agent*, *instrument* and *result*) relations
 - ▶ 11 QPs are proposed
 - ▶ More than 7,500 CQs are created
 - ▶ More than 43% of CQs are validated
 - ▶ Example:

```
(forall (?Y)
  (=>
    ($instance ?Y MusicalComposition)
    (exists (?X)
      ($instance ?X ComposingMusic)
      (result ?X ?Y))))
```

Mapping between WordNet and SUMO

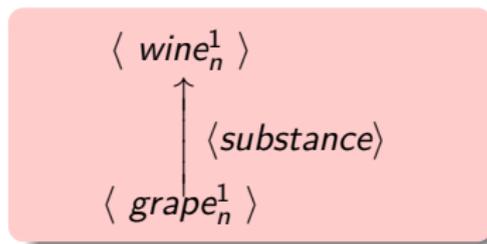
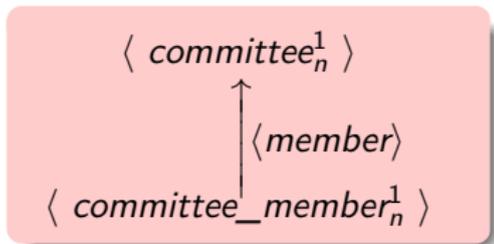
- Described in (Niles and Pease, 2003)
- It connects synsets of WordNet to terms of SUMO using 3 relations:
 - ▶ *equivalence* (=)
 - ▶ *subsumption* (+)
 - ▶ *instance* (@)
- Some examples:

$\langle \text{calcium}_n^1 \rangle$:	$[\text{Calcium}_c=]$
$\langle \text{calcium_oxide}_n^1 \rangle$:	$[\text{CompoundSubstance}_c+]$
$\langle \text{police_officer}_n^1 \rangle$:	$[\text{PoliceOfficer}_a=]$
$\langle \text{police_force}_n^1 \rangle$:	$[\text{PoliceOrganization}_c+]$

- 1 Introduction
- 2 Cross-checking WordNet and Adimen-SUMO
- 3 Some Experimental Results
- 4 Conclusions and Future Work

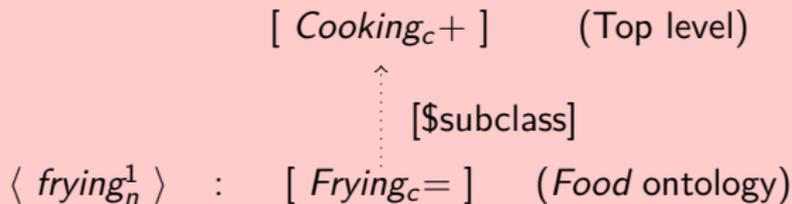
Meronymy Information in WordNet

- WordNet v3.0 provides 3 part-whole relations (22,187):
 - ▶ *part*: the general meronymy relation (9,097)
 - ▶ *member*: it relates particulars and groups (12,293)
 - ▶ *substance*: it relates physical matters and things (797)
- For example:



Exploiting the Mapping between WordNet and SUMO

- First, creating a mapping between WordNet and Adimen-SUMO:



- Propose a formal characterization of the mapping information:

< *male_horse*_n¹ > : [*Male*_a+] [*Horse*_c+]

▶ *Literal* interpretation:

(and
 (\$instance ?X Male)
 (\$instance ?X Horse))

▶ *Precise* interpretation:

(and
 (*attribute* ?X Male)
 (\$instance ?X Horse))

Question patterns for the Creation of CQs (I)

- Four different QPs depending on the used mapping relations (*precise* interpretation):
 - ▶ *equivalence*
 - ▶ *subsumption* or *instance*
- QPs are instantiated according to the mapping information of the synsets in the WordNet meronymy pairs.

Question patterns for the Creation of CQs (II)

- Applying the first QP (precise interpretation):

```
(exists (?X ?Y)
  (and
    <s_part ?X>
    <s_whole ?Y>
    (<SUMO_predicate> ?X ?Y)))
```

- to the following WN-SUMO meronymy relation:

```

< genus_malacosoma1n >      : [ Larvala+ ]
      ↑
      | <member>
      |
< malacosoma_americana1n > : [ Insectc+ ]
      ↑
      |
      | [memberr]
```

Question patterns for the Creation of CQs (III)

- Creates the following CQ:

```
(exists (?X ?Y)
  (and
    ($instance ?X Insect)
    (attribute ?Y Larval)
    (member ?X ?Y)))
```

Question patterns for the Creation of CQs (IV)

- Mapping of WordNet relations to Adimen-SUMO predicates, which have domain restrictions:

$\langle \text{part} \rangle$: [$part_r(Object_c \times Object_c)$]
 $\langle \text{member} \rangle$: [$member_r(SelfConnectedObject_c \times Collection_c)$]
 $\langle \text{substance} \rangle$: [$material_r(Substance_c \times CorpuscularObject_c)$]

- Many discrepancies arise during the instantiation of question patterns.
- 14,513 part relations from 22,187 (65%) do not hold domain conditions.
 - ▶ Example:

$\langle wine_n^1 \rangle$: [$Wine_c =$]
 \uparrow $\langle \text{substance} \rangle$ \uparrow
 $\langle \text{grape}_n^1 \rangle$: [$FruitOrVegetable_{c+}$]

- ▶ Reason: the first argument of $material_r$ is restricted to be $Substance_c$, which is incompatible with $FruitOrVegetable_c$
- So, we concentrate on the remaining 7,674 relations (35%)

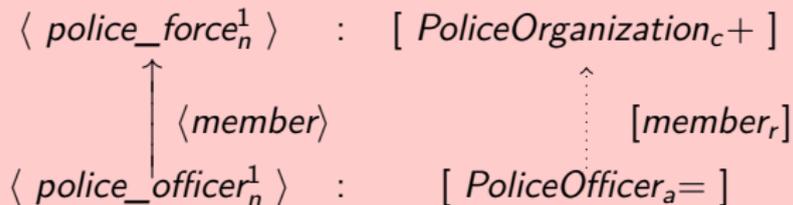
- 1 Introduction
- 2 Cross-checking WordNet and Adimen-SUMO
- 3 Some Experimental Results**
- 4 Conclusions and Future Work

Creating CQs and applying ATPs

- We apply the 4 QPs to the 7,674 relations allowing to create 2,145 CQs.
- When testing these CQs using ATPs such as Vampire (Kovács and Voronkov, 2013) or E-prover (Schulz, 2002):
 - ▶ *Passing*: knowledge validation
 - ▶ *Non-passing*: knowledge mismatches
 - WN-SUMO mapping issues
 - WordNet issues
 - SUMO issues
 - ▶ *Unknown*: Missing knowledge ... or insufficient execution time?

Knowledge Validation

- Example:



- Reason:

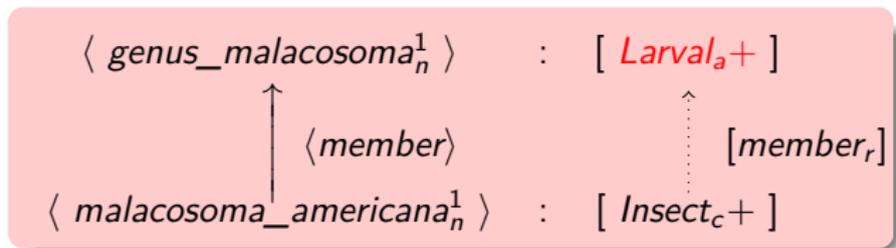
- ▶ The resulting CQ is entailed by Adimen-SUMO:

```

(forall (?Y)
  ( $\Rightarrow$ 
    (attribute ?Y PoliceOfficer)
    (exists (?X)
      (and
        ($instance ?X PoliceOrganization)
        (member ?X ?Y))))))
  
```

Detection of Mapping Mismatches

- Example:

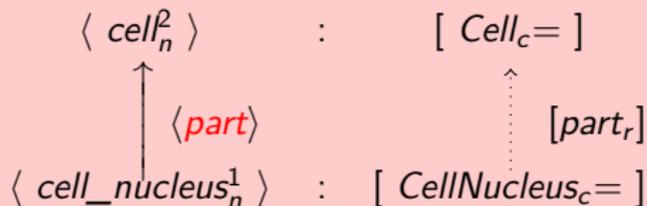


- Reason:

- ▶ The attribute *Larval*_a cannot be applied to groups in Adimen-SUMO

Detection of WordNet Issues

- Example:



- Reason:

- ▶ Some cells lack a nucleus, as stated by the following Adimen-SUMO axiom:

```

(forall (?C)
  (=>
    ($instance ?C RedBloodCell)
    (not (exists (?N)
      (and
        ($instance ?N CellNucleus)
        (part ?N ?C)))))))
  
```


Summary

- Reported in (Álvarez and Rigau, 2018)

SUMO relations	CQs				Total	
	QP #1	QP #2	QP #3	QP #4		
<i>part_r</i>	+599	+56	+162	+8	+825	42.09%
	-6	-0	-1	-5	-12	0.61%
<i>member_r</i>	+10	+1	+1	+0	+12	9.92%
	-9	-0	-0	-0	-9	7.44%
<i>material_r</i>	+17	+1	+2	+0	+17	26.56%
	-0	-2	-0	-0	-2	3.13%
Total	+626	+58	+165	+8	+857	39.95%
	-15	-2	-1	-5	-23	1.07%

- ▶ 857 *Passing* CQs (39.95% of total) enable to validate the knowledge of WordNet, SUMO and their mapping
- ▶ *part* is better aligned and covered (825 truth-tests, 42.09%) than *member* (only 12 truth-tests, 9.92%) and *substance* (17 truth-tests, 26.56%)
- ▶ Different issues are detected (23 falsity-tests, 1.07%)
- ▶ More than 60% of the total is *Unknown*.

- 1 Introduction
- 2 Cross-checking WordNet and Adimen-SUMO
- 3 Some Experimental Results
- 4 Conclusions and Future Work**

Conclusions

- Framework and benchmark for formal commonsense reasoning
- More than 10,000 CQs available (around 60% *Unknown*)!
- First steps cross-checking of WordNet, Adimen-SUMO and its mapping:
 - ▶ Validation of some pieces of knowledge
 - ▶ Detection of knowledge mismatches
 - ▶ Detection of missing knowledge
- Resources are ready for its application to practical NLP tasks
- <http://adimen.si.ehu.es/web/AdimenSUMO>
- <https://vprover.github.io/>
- <https://github.com/e prover/e prover>

Future Work

- Improving the WN-SUMO mapping
- Extending our proposal to additional WordNet relations
- Automatically derive new SUMO axioms from WordNet knowledge

Bibliography I

- J. Álvarez and G. Rigau. Complete and consistent annotation of WordNet using the Top Concept Ontology. In *Proc. of the 11th Language Resources and Evaluation Conf. (LREC 2018)*. European Language Resources Association (ELRA), 2018.
- J. Álvarez, P. Lucio, and G. Rigau. Adimen-SUMO: Reengineering an ontology for first-order reasoning. *Int. J. Semantic Web Inf. Syst.*, 8(4):80–116, 2012.
- J. Álvarez, P. Lucio, and G. Rigau. Improving the competency of first-order ontologies. In K. Barker and J. M. Gómez-Pérez, editors, *Proc. of the 8th Int. Conf. on Knowledge Capture (K-CAP 2015)*, pages 15:1–15:8. ACM, 2015. ISBN 978-1-4503-3849-3. doi: 10.1145/2815833.2815841.
- J. Álvarez, M. Hermo, P. Lucio, and G. Rigau. Automatic white-box testing of first-order logic ontologies. *CoRR*, abs/1705.10219, 2017a. URL <http://arxiv.org/abs/1705.10219>.
- J. Álvarez, P. Lucio, and G. Rigau. Black-box testing of first-order logic ontologies using WordNet. *CoRR*, abs/1705.10217, 2017b. URL <http://arxiv.org/abs/1705.10217>.
- C. Fellbaum, editor. *WordNet: An Electronic Lexical Database*. MIT Press, 1998.

Bibliography II

- M. Grüninger and M. S. Fox. Methodology for the design and evaluation of ontologies. In *Proc. of the Workshop on Basic Ontological Issues in Knowledge Sharing (IJCAI 1995)*, 1995.
- I. Horrocks and A. Voronkov. Reasoning support for expressive ontology languages using a theorem prover. In J. Dix et al., editor, *Foundations of Information and Knowledge Systems*, LNCS 3861, pages 201–218. Springer, 2006.
- L. Kovács and A. Voronkov. First-order theorem proving and Vampire. In N. Sharygina and H. Veith, editors, *Computer Aided Verification*, LNCS 8044, pages 1–35. Springer, 2013. ISBN 978-3-642-39798-1.
- I. Niles and A. Pease. Towards a standard upper ontology. In Guarino N. et al., editor, *Proc. of the 2nd Int. Conf. on Formal Ontology in Information Systems (FOIS 2001)*, pages 2–9. ACM, 2001. doi: 10.1145/505168.505170.
- I. Niles and A. Pease. Linking lexicons and ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In H. R. Arabnia, editor, *Proc. of the IEEE Int. Conf. on Inf. and Knowledge Engin. (IKE 2003)*, volume 2, pages 412–416. CSREA Press, 2003. ISBN 1-932415-08-4.

Bibliography III

- A. Pease and G. Sutcliffe. First-order reasoning on a large ontology. In Sutcliffe G. et al., editor, *Proc. of the Workshop on Empirically Successful Automated Reasoning in Large Theories (CADE-21)*, CEUR Workshop Proceedings 257. CEUR-WS.org, 2007.
- S. Schulz. E - A brainiac theorem prover. *AI Communications*, 15(2-3):111–126, 2002. ISSN 0921-7126.