# Towards Cross-checking WordNet and SUMO Using Meronymy

Javier Álvez **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 (Álvez 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 (Álvez 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 (Álvez et al., 2015, 2017b)
  - White-box testing strategies (Álvez et al., 2017a)
- More than 100 axioms from Adimen-SUMO has been improved

# Black-box Testing I

- Introduced in (Álvez et al., 2015) and fully described in (Álvez 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 (Álvez 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:

$$\begin{array}{rcl}
\langle calcium_n^1 \rangle & : & [Calcium_c =] \\
\langle calcium_oxide_n^1 \rangle & : & [CompoundSubstance_c +] \\
\langle police_officer_n^1 \rangle & : & [PoliceOfficer_a =] \\
\langle police_force_n^1 \rangle & : & [PoliceOrganization_c +]
\end{array}$$

# 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:

$$\begin{tikzcd} \langle \mathit{committee}^1_n \rangle \\ \langle \mathit{committee\_member}^1_n \rangle \arrow[u, "\langle \mathit{member} \rangle"'] \end{tikzcd}$$

$$\begin{tikzcd} \langle \mathit{wine}^1_n \rangle \\ \langle \mathit{grape}^1_n \rangle \arrow[u, "\langle \mathit{substance} \rangle"'] \end{tikzcd}$$

# Exploiting the Mapping between WordNet and SUMO

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

$$\begin{array}{ccccc} & & [\ \mathit{Cooking}_c{+}\ ] & & \text{(Top level)} \\ & & \Big\uparrow\ [\$\text{subclass}] & & \\ \langle\ \mathit{frying}^1_n\ \rangle & : & [\ \mathit{Frying}_c{=}\ ] & & (\mathit{Food}\ \text{ontology}) \end{array}$$

- Propose a formal characterization of the mapping information:

$$\langle\ \mathit{male\_horse}^1_n\ \rangle\ :\ [\mathit{Male}_a{+}]\ [\mathit{Horse}_c{+}]$$

  - *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:

$$\begin{array}{ccc} \langle\ genus\_malacosoma^1_n\ \rangle & : & [\ Larval_a+\ ] \\ \uparrow \langle member \rangle & & \vdots\ [member_r] \\ \langle\ malacosoma\_americana^1_n\ \rangle & : & [\ Insect_c+\ ] \end{array}$$

# 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:

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

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

$$\begin{array}{ccc} \langle\ wine^1_n\ \rangle & : & [\ Wine_c{=}\ ] \\ \uparrow \langle substance \rangle & & \uparrow [\cancel{material_r}] \\ \langle\ {\color{red} grape^1_n}\ \rangle & : & [\ FruitOrVegetable_c{+}\ ] \end{array}$$

  - 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%)

# 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 insuficient execution time?

# Knowledge Validation

- Example:

$$\begin{tikzcd} \langle \mathit{police\_force}^1_n \rangle & : & [\ \mathit{PoliceOrganization}_c + \ ] \\ \langle \mathit{police\_officer}^1_n \rangle \arrow[u, "\langle \mathit{member} \rangle"'] & : & [\ \mathit{PoliceOfficer}_a = \ ] \arrow[u, dotted, "{[\mathit{member}_r]}"'] \end{tikzcd}$$

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

```
(forall (?Y)
    (=>
        (attribute ?Y PoliceOfficer)
        (exists (?X)
            (and
                ($instance ?X PoliceOrganization)
                (member ?X ?Y)))))
```

# Detection of Mapping Mismatches

- Example:

$$\begin{array}{ccc} \langle\ genus\_malacosoma^1_n\ \rangle & : & [\ Larval_a +\ ] \\ \Big\uparrow \langle member \rangle & & \Big\uparrow [member_r] \\ \langle\ malacosoma\_americana^1_n\ \rangle & : & [\ Insect_c +\ ] \end{array}$$

- Reason:
  - The attribute *Larval*$_a$ cannot be applied to groups in Adimen-SUMO

# Detection of WordNet Issues

- Example:

$$\begin{tikzcd}
\langle\ cell_n^2\ \rangle & : & [\ Cell_c{=}\ ] \\
\langle\ cell_nucleus_n^1\ \rangle \arrow[u, "\langle part \rangle"'] & : & [\ CellNucleus_c{=}\ ] \arrow[u, dotted, "{[part_r]}"']
\end{tikzcd}$$

- 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))))))
```

# Detection of Adimen-SUMO Issues

- Example:

$$\begin{tikzcd} \langle\ water\_ice^2_n\ \rangle & : & [\ Solid_a +\ ] \\ \langle\ water^1_n\ \rangle \arrow[u, "\langle substance \rangle"'] & : & [\ Water_c =\ ] \arrow[u, dotted, "{[material_r]}"'] \end{tikzcd}$$

- Problem:
  - The application of subattributes of *PhysicalState*$_A$ (as *Solid*$_a$) was restricted to be only! a property of *Substance*$_c$:

```
(forall (?OBJ)
   (<=>
      ($instance ?OBJ Substance)
      (exists (?ATTR)
         (and
            ($instance ?ATTR PhysicalState)
            (attribute ?OBJ ?ATTR)))))
```

# Summary

- Reported in (Álvez and Rigau, 2018)

| **SUMO relations** | **CQs** | | | | | |
|---|---|---|---|---|---|---|
| | QP #1 | QP #2 | QP #3 | QP #4 | **Total** | |
| *part$_r$* | +599<br>-6 | +56<br>-0 | +162<br>-1 | +8<br>-5 | **+825**<br>**-12** | **42.09%**<br>**0.61%** |
| *member$_r$* | +10<br>-9 | +1<br>-0 | +1<br>-0 | +0<br>-0 | **+12**<br>**-9** | **9.92%**<br>**7.44%** |
| *material$_r$* | +17<br>-0 | +1<br>-2 | +2<br>-0 | +0<br>-0 | **+17**<br>**-2** | **26.56%**<br>**3.13%** |
| **Total** | **+626**<br>**-15** | **+58**<br>**-2** | **+165**<br>**-1** | **+8**<br>**-5** | **+857**<br>**-23** | **39.95%**<br>**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*.

# 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/eprover/eprover`

# 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. Álvez and G. Rigau. Complete and consistent annotation of WordNet using the Top Concept Ontology. In *Proc. of the 11$^{th}$ Language Resources and Evaluation Conf. (LREC 2018)*. European Language Resources Association (ELRA), 2018.

J. Álvez, 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. Álvez, P. Lucio, and G. Rigau. Improving the competency of first-order ontologies. In K. Barker and J. M. Gómez-Pérez, editors, *Proc. of the 8$^{th}$ 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. Álvez, 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. Álvez, 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.