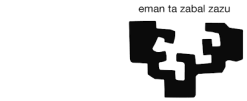
Towards Effective Correction Methods Using WordNet Meronymy Relations

Javier Álvez Itziar Gonzalez-Dios German Rigau

LoRea Group Ixa Group, HiTZ Center
University of the Basque Country (UPV/EHU)



12^{th} International Global WordNet Conference – GWC 2023

Donostia/San Sebastian, Basque Country – January 23-27, 2023

Commonsense Reasoning and Natural Language Processing

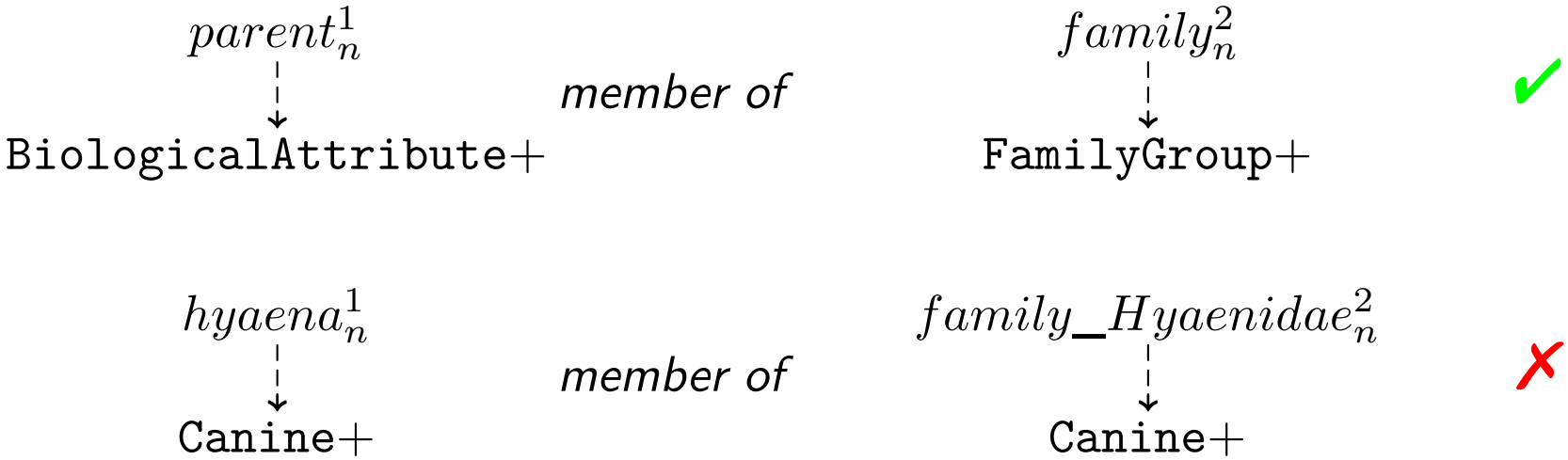
- Commonsense reasoning is applied in many tasks related to Natural Language Processing (NLP)
  - Recognizing Textual Entailment (RTE) (Bos and Markert, 2006; Dagan et al., 2013; Abzianidze, 2017)
  - Natural Language Inference (NLI) (Bowman et al., 2015)
  - Interpretable Semantic Textual Similarity (ISTS) (Lopez-Gazpio et al., 2017)
- WordNet (Fellbaum, 1998) is a semantic resource
  - One of the most frequently applied in NLP tasks
  - Interlinked with other resources
    - The EuroWordNet Top Ontology (TCO) (Rodríguez et al., 1998)
    - SUMO (Niles and Pease, 2001)
- Interlinking knowledge and lexical resources enables cross-checking them and validating the encoded knowledge

Automatically cross-checking knowledge and lexical resources

- A general framework for evaluating the competency of SUMO-based ontologies (Álvez et al., 2015, 2019)
  - Adimen-SUMO (Álvez et al., 2012)
  - Automatic creation of *competency questions* (CQs) (Grüninger and Fox, 1995)
    - On the basis of several predefined question patterns (QPs)
    - Instantiated according to WordNet and its mapping into SUMO (Niles and Pease, 2003)
  - Automatically evaluated by automated theorem provers (ATPs)

Experimental results on meronymy

- Low validation ratio (Álvez et al., 2017) due to:
  - Incorrect mappings
  - Discrepancies of knowledge
  - Limitations of ATPs
- Examples:



Objectives

- To shed light on the sources of difficulty when improving knowledge resources
- Discover strategies that lead to improving knowledge resources while reducing the human effort

Our Proposal

- Three approaches:
  - Mapping between SUMO and WordNet
    - Structural corrections, on the basis of the WordNet hierarchy
    - Opportunistic corrections, by performing an error analysis
  - Matching knowledge discrepancies
    - Improving Adimen-SUMO
  - Combining the above two approaches
- Evaluate the impact on unseen data
  - The WebChild project (Tandon et al., 2014, 2017)

Outline

- Description of the framework and initial state
- Mapping between SUMO and WordNet
- Matching knowledge discrepancies
- Combined approach
- Evaluation and discussion
- Conclusions and future work

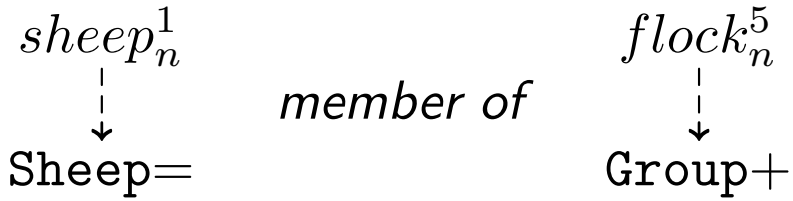
Description of the framework I

- WordNet:
  - *member* meronymy relation (12,293 pairs)
    - Also in the WebChild Project (39,127 pairs)
  - Base level concepts (BLCs)
  - Semantic file and domain
  - Mapping to TCO
- Adimen-SUMO:
  - *member* relation
- Mapping between WordNet and SUMO:
  - Two mapping relations: *subsumption* ($+$) and *equivalence* ($=$)
- Four QPs for the creation of CQs:
  - Two goals in first-order logic
- Evaluation of CQs using automated theorem provers:
  - Validated
  - Unvalidated
  - Unclassified

Description of the framework II

Example:

- WordNet pair and mapping to SUMO:



- Resulting CQ:

```
(forall (?X)
    (=>
        ($instance ?X Sheep)
        (exists (?Y)
            (and
                ($instance ?Y Group)
                (member ?X ?Y) ) ) ) )
```

Initial state

- From the 12,293 *member* pairs provided by WordNet:
  - Only 19 are validated
  - 11,963 pairs are unvalidated
    - Only 24 yield a correct CQ according to the SUMO domain restrictions
  - 311 remains unknown
- Conclusions:
  - A mere 0.15% of the WordNet knowledge about *member* is validated
  - Most of the unvalidated pairs yield semantically incorrect SUMO conjectures

Mapping between SUMO and WordNet

Structural corrections

- From 800 BLCs, we have inspected 200 BLCs
  - 50 manual corrections
- Corrections have been propagated
  - 3,883 automatic corrections
- Effort: 10 human-hours
- Example:
  - $fish\_genus^1_n$ ("*any of various genus of fish*") was mapped to `Fish`+ ("*A cold-blooded aquatic Vertebrate characterized by fins and breathing by gills. ...*")
  - We have corrected the mapping to `GroupOfAnimals`+ ("*Any Group which contains exclusively non-human members*")

Mapping between SUMO and WordNet

Opportunistic corrections

- We have analyzed the unclassified meronymy pairs
  - Grouped according to their mapping
  - Ordered by frequency
- We have proposed heuristics to automatically expand *ad hoc* corrections
  - 4 heuristics
  - 1,961 automatic corrections
  - Effort: 2 human-hours

Mapping between SUMO and WordNet

Opportunistic corrections: example

- Heuristic: given a synset s
  - Condition:
    - s is hyponym of $group_n^1$ in WordNet
    - s is connected to both `Plant`+ and `Group`+ in TCO
    - s is connected to a subclass of `Plant` in SUMO
    - Some of the words *family*, *genus*, *order*, *suborder*, *class*, *phylum*, *subphylum*, *kingdom*, *subkingdom*, *division*, *subdivision*, *algae*, *superfamily*, *subfamily*, *superorder* or *group* occurs in the gloss of s
  - Action:
    - Correct the mapping of s to `Group`+

Matching knowledge discrepancies

- From unclassified and unvalidated pairs, we have analyzed those with correct mapping
- Detect and solve the problems that prevent their validation
  - Most of them are related to organisms
- Interventions in Adimen-SUMO:
  - 3 corrected axioms:
    - Domain restriction of *member*
    - Characterization of `GroupOfAnimals` and `GroupOfPeople`
  - 1 new class:
    - `GroupOfPlants`
  - 12 new axioms:
    - Characterization of `GroupOfPlants`, `GroupOfAnimals`, `GroupOfPeople`, `AgeGroup`, `FamilyGroup`, `SocialUnit`, `EthnicGroup`, `BeliefGroup` and `Brood`
  - Effort: 2 human-hours

Combined approach

- The integration has required some updates
  - The mapping of 9 BLCs has been updated
    - Automatically propagated to 1,961 synsets
  - One heuristic for *member* has been updated
    - 2,411 automatic corrections
- Effort: almost 0 human-hours

Evaluation and discussion

Results

Data	Phase	Validated	Unvalidated T	Unvalidated C	Unknown
WordNet	Initial	19	11,963	24	311
	Mapping	29	6,561	5,811	5,703
	Knowledge	132	11,603	30	558
	Joint	10,071	808	58	1,414
WebChild	Initial	82	35,377	102	3,368
	Joint	18,569	3,526	136	17,032

Evaluation and discussion

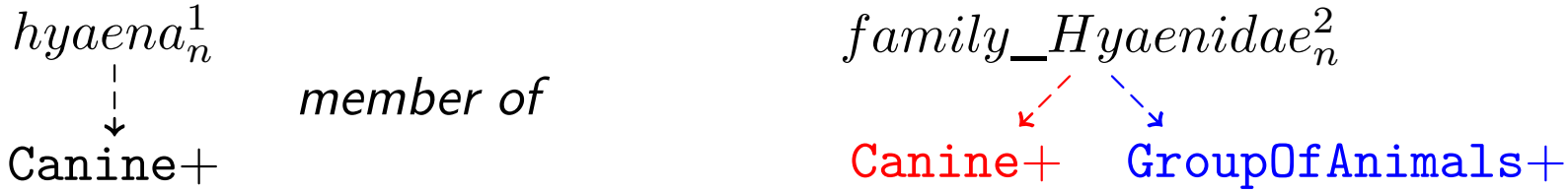
Performance metrics

Data	Phase	Recall	Precision	*F*1
WordNet	Initial	0.002	0.002	0.002
	Mapping	0.002	0.004	0.003
	Knowledge	0.011	0.011	0.011
	Joint	0.819	0.926	0.869
WebChild	Initial	0.002	0.002	0.002
	Joint	0.475	0.840	0.607

Evaluation and discussion

Example

- WordNet pair and mapping to SUMO:



- Resulting CQ:

```
(exists (?X ?Y)
    (and
        ($instance ?X Canine)
        ($instance ?Y GroupOfAnimals)
        (member ?X ?Y) ) )
```

Evaluation and discussion

Detailed Analysis

- Random sample of 5 pairs for each output (15 pairs in total)
- Validated pairs:
  - Good reasons: 4/5
  - Wrong pairs: 1/5
  - Example:
    - *genus*$^{2}_{n}$ ("*(biology) taxonomic group containing one or more species*") is incorrectly asserted to be member of *Carapidae*$^{1}_{n}$ ("*pearlfishes: related to the Brotulidae*")
- Unvalidated pairs:
  - Wrong pairs: 2/5
  - Mapping errors: 3/5
  - Example:
    - *Tremellales*$^{1}_{n}$ ("*fungi varying from gelatinous to waxy or even horny in texture*") is still mapped to `Fungus`
- Unclassified pairs:
  - Wrong pairs: 4/5
  - Missing knowledge: 1/5

Conclusions and future work

- Conclusions:
  - Correcting knowledge resources is time-consuming and a manual task
  - It is worth correcting the ontology
- Future work:
  - From our detailed analysis:
    - Characterization of metonymy
    - Correction of the knowledge
    - Analysis of the inheritance of relations
  - Experimentation with other datasets:
    - BLESS (Baroni and Lenci, 2010)
    - Morphosemantic links (Fellbaum et al., 2009)
  - New mapping between WordNet and Adimen-SUMO on the basis of formulas

Bibliography I

L. Abzianidze. LangPro: Natural language theorem prover. In L. Specia, M. Post, and M. Paul, editors, *Proc. of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (EMNLP 2017)*, pages 115–120. Association for Computational Linguistics, 2017. URL `http://aclweb.org/anthology/D17-2020`.

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, P. Lucio, and G. Rigau. Black-box testing of first-order logic ontologies using WordNet. *CoRR*, abs/1705.10217, 2017. URL `http://arxiv.org/abs/1705.10217`.

Bibliography II

J. Álvez, P. Lucio, and G. Rigau. A framework for the evaluation of SUMO-based ontologies using WordNet. *IEEE Access*, 7:36075–36093, 2019. ISSN 2169-3536. doi: 10.1109/ACCESS.2019.2904835.

M. Baroni and A. Lenci. Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, 36(4):673–721, 2010. URL `http://dblp.uni-trier.de/db/journals/coling/coling36.html#BaroniL10`.

J. Bos and K. Markert. Recognising textual entailment with robust logical inference. In J. Quiñonero-Candela, I. Dagan, B. Magnini, and F. d'Alché Buc, editors, *Machine Learning Challenges. Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Tectual Entailment*, pages 404–426, Berlin, Heidelberg, 2006. Springer. ISBN 978-3-540-33428-6.

S. R. Bowman, G. Angeli, C. Potts, and C. D. Manning. A large annotated corpus for learning natural language inference. In *Proc. of the Conf. on Empirical Methods in Natural Language Processing (EMNLP 2015)*, pages 632–642. Association for Computational Linguistics, 2015.

Bibliography III

I. Dagan, D. Roth, M. Sammons, and F. M. Zanzotto. Recognizing textual entailment: Models and applications. *Synthesis Lectures on Human Language Technologies*, 6(4):1–220, 2013.

C. Fellbaum, editor. *WordNet: An Electronic Lexical Database*. MIT Press, 1998.

C. Fellbaum, A. Osherson, and P. Clark. Putting semantics into WordNet's "Morphosemantic" Links. In Z. Vetulani and H. Uszkoreit, editors, *Human Language Technology. Challenges of the Information Society*, LNCS 5603, pages 350–358. Springer, 2009. ISBN 978-3-642-04235-5.

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. Lopez-Gazpio, M. Maritxalar, A. Gonzalez-Agirre, G. Rigau, L. Uria, and E. Agirre. Interpretable semantic textual similarity: Finding and explaining differences between sentences. *Knowledge-Based Systems*, 119:186 – 199, 2017. ISSN 0950-7051.

Bibliography IV

I. Niles and A. Pease. Towards a standard upper ontology. In N. Guarino 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.

H. Rodríguez, S. Climent, P. Vossen, L. Bloksma, W. Peters, A. Alonge, F. Bertagna, and A. Roventini. The top-down strategy for building EuroWordNet: Vocabulary coverage, base concepts and top ontology. In *EuroWordNet: A multilingual database with lexical semantic networks*, pages 45–80. Springer, 1998.

N. Tandon, G. de Melo, F. Suchanek, and G. Weikum. WebChild: Harvesting and organizing commonsense knowledge from the Web. In *Proc. of the 7th ACM Int. Conference on Web Search and Data Mining (WSDM '14)*, pages 523–532. Association for Computing Machinery, 2014. doi: 10.1145/2556195.2556245.

Bibliography V

N. Tandon, G. de Melo, and G. Weikum. WebChild 2.0 : Fine-grained commonsense knowledge distillation. In M. Bansal and H. Ji, editors, *Proc. of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017)*, pages 115–120, 2017. ISBN 978-1-945626-71-5. doi: 10.18653/v1/P17-4020.