

University of Groningen

Towards reference-aware FrameNet representations

Minnema, Gosse; Remijnse, Levi; Bos, Johan; Caselli, Tommaso; Fokkens, Antske; Nissim, Malvina; Postma, Marten; Vossen, Piek

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Minnema, G., Remijnse, L., Bos, J., Caselli, T., Fokkens, A., Nissim, M., Postma, M., & Vossen, P. (2020). *Towards reference-aware FrameNet representations: Bridging generic and specific event knowledge*. Poster session presented at GeCKo Symposium, Barcelona, Spain.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Towards reference-aware FrameNet representations: bridging generic and specific event knowledge

Gosse Minnema^a, Levi Remijnse^b, Johan Bos^a, Tommaso Caselli^a,
Antske Fokkens^b, Malvina Nissim^a, Marten Postma^b, Piek Vossen^b

^aUniversity of Groningen, The Netherlands

^bVU University Amsterdam, The Netherlands

{g.minnema, johan.bos, t.caselli, m.nissim}@rug.nl
{l.remijnse, antske.fokkens, m.postma, p.t.j.m.vossen}@vu.nl

1 Introduction

FrameNet (Baker et al., 2003) is a resource that encodes conceptual and linguistic knowledge in the form of *frames*: information packages defining word senses and semantic roles associated with a particular type of event, situation or concept. FrameNet is a rich resource for describing how events and situations can be conceptualized in language in different ways, but is limited by its focus on lexical semantics and lack of a notion of reference: a frame-semantic analysis of the event descriptions in (1) would tell us that both describe the same *event type* (i.e., a commercial transaction, conceptualized from two different perspectives), but not whether they in fact describe the same *event token* in the real world.

- (1) a. Yesterday, John sold Mary a book.
- b. A woman bought a novel in the shop.

To address this limitation, we are currently developing a new FrameNet-based resource, comprising a lexical database, annotated corpus and a semantic parser, that is ‘referentially enriched’ in two ways: frame annotations are linked, on one hand, to referential information from an ontology of real-world event tokens, and on the other hand to truth-conditional meaning representations.

2 Linking event token knowledge

‘Data-to-text’ We implement Vossen et al. (2018)’s ‘data-to-text’ method for data collection and event annotation: structured data about real-world event tokens of pre-specified types from knowledge bases such as Wikidata (Erxleben et al., 2014) is used as a starting point, and is then linked to texts known to describe the event tokens. In this way, we collect a large collection of texts that are linked to structured data describing the events referenced in the text. The

approach also addresses FrameNet’s data sparsity problem: whereas the original FrameNet corpus covers many different event types, but with a small number of annotations for every type, we limit the number of event types but make sure we get a sizeable number of annotations for every type.

Annotating explicit and implicit frames The availability of referential data allows for linking frames and semantic roles to event tokens and referential properties, and for annotating events that could not be annotated under standard FrameNet annotation because they lack an explicit lexical target, but are implied by the compositional semantics and/or pragmatics of the discourse. For example, a sentence like “he was shot and died” clearly describes a killing event, even though there is no single lexical item uniquely describing it. Instead, the event is ‘triggered’ compositionally by “shot” and “died”. Under standard FrameNet annotation it is not feasible to try to exhaustively annotate all events that are described in this way, but is doable in our framework, given that we already know which events we are looking for.

3 Parsing formal representations

The second goal of our project is to integrate frame annotations into Discourse Representation Structures (DRS) (Kamp and Reyle, 1993). Doing this, on one hand, allows for formally modeling event and role (co-)reference and pragmatic inference, and on the other hand adds rich conceptual information to the formal representations. Bos and Nissim (2008) laid the theoretical basis for combining FrameNet and DRT; we are currently working on implementing their ideas. In particular, we are working on post-processing the outputs of existing frame (e.g., Swayamdipta et al., 2017) and DRS parsers (Van Noord et al., 2018) for automatically creating combined representations.

References

- Collin F. Baker, Charles J. Fillmore, and Beau Cronin. 2003. The structure of the FrameNet database. *International Journal of Lexicography*, 16(3):281–296.
- Johan Bos and Malvina Nissim. 2008. Combining Discourse Representation Theory with FrameNet. In R. Rossini Favretti, editor, *Frames, corpora and knowledge representation*. Bononia University Press, Bologna.
- Fredo Erxleben, Michael Günther, Markus Krötzsch, Julian Mendez, and Denny Vrandečić. 2014. Introducing Wikidata to the Linked Data web. In *The Semantic Web – ISWC 2014*, pages 50–65, Cham. Springer International Publishing.
- H. Kamp and U. Reyle. 1993. *From Discourse to Logic: An Introduction To Model-Theoretic Semantics of Natural Language, Formal Logic and DRT*. Kluwer, Dordrecht.
- Swabha Swayamdipta, Sam Thomson, Chris Dyer, and Noah A. Smith. 2017. Frame-semantic parsing with Softmax-Margin Segmental RNNs and a syntactic scaffold. *arXiv preprint arXiv:1706.09528*.
- R. Van Noord, L. Abzianidze, A. Toral, and J. Bos. 2018. Exploring neural methods for parsing discourse representation structures. *Transactions of the Association for Computational Linguistics*, 6:619–634.
- Piek Vossen, Filip Ilievski, Marten Postma, and Roxane Segers. 2018. Do not annotate, but validate: a data-to-text method for capturing event data. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*.