

Frame- and Entity-Based Knowledge for Common-Sense Argumentative Reasoning

Teresa Botschen* Daniil Sorokin* Iryna Gurevych

Ubiquitous Knowledge Processing Lab (UKP) and Research Training Group AIPHES
Department of Computer Science, Technische Universität Darmstadt
www.ukp.tu-darmstadt.de

Abstract

Common-sense argumentative reasoning is a challenging task that requires holistic understanding of the argumentation where external knowledge about the world is hypothesized to play a key role. We explore the idea of using event knowledge about prototypical situations from FrameNet and fact knowledge about concrete entities from Wikidata to solve the task. We find that both resources can contribute to an improvement over the non-enriched approach and point out two persisting challenges: first, integration of many annotations of the same type, and second, fusion of complementary annotations. After our explorations, we question the key role of external world knowledge with respect to the argumentative reasoning task and rather point towards a logic-based analysis of the chain of reasoning.

1 Introduction

Recently, Habernal et al. (2018) introduced a challenging dataset for Argument Reasoning Comprehension (ARC) used in the SemEval-2018 shared task. After reviewing the participating systems, they hypothesize that *external world knowledge may be essential for ARC*.¹ We explore enriching models with event and fact knowledge on ARC to investigate into this hypothesis.

Semantic tasks profit from external knowledge: language understanding requires more complex knowledge than that contained in current systems and word embeddings. For the task of semantic plausibility, Wang et al. (2018) reveal the failure of models only relying on distributional data, whilst the injection of world knowledge helps. Glockner et al. (2018) point out the deficiency of state-of-the-art approaches for understanding entailment on

the large-scale SNLI corpus (Stanford Natural Language Inference) (Bowman et al., 2015). In their study, the model incorporating external lexical information from WordNet, KIM (Knowledge-based Inference Model) (Chen et al., 2018), does not yield the awaited improvements — where the crucial point might be WordNet (Miller, 1995) which does not contain explicit world knowledge in the form of event- and fact-based knowledge. Previous work argues that information in WordNet overlaps with word embeddings (Zhai et al., 2016), therefore we focus on other types of knowledge in our work.

Complementary sources of external knowledge: we experiment using the lexical-semantic resource FrameNet (FN) and the knowledge base Wikidata (WD). These resources provide information beyond the lexical relations encoded in WordNet and thus have a potential to enhance the underlying model with other kind of external world knowledge. On the one hand, FN provides qualitative event-knowledge about prototypical situations. Thus, identifying frames (FrameId) unveils the situation or action that is happening. On the other hand, WD provides fact-knowledge about the concrete entities. So, linking entities to a knowledge base (EntLink) disambiguates the participants. Furthermore, both resources provide meta-knowledge about how their frames or entries relate to each other.

Multiple levels of knowledge processing help: combining several kinds of annotations benefits question answering (Khashabi et al., 2018), external knowledge about synonyms enhances inference (Chen et al., 2018), and jointly modeling several tasks (e.g., frame-semantic parsing and dependency parsing) is fruitful (Peng et al., 2018). In particular, the idea of connecting event semantics and fact knowledge was confirmed by Guo et al. (2016): they jointly formalize semantic role labeling and relation classification and thereby improve upon PropBank semantic role labeling.

* First and second authors contributed equally to this work.

¹SemEval-2018 Task 12: <https://competitions.codalab.org/competitions/17327>

Outline In this paper, we investigate whether external information in terms of event-based frames (FN) and fact-based entities (WD) can contribute to holistic understanding of the argumentation in the ARC task. First, we examine the effect of both annotations separately and second, we explore whether a joint annotation benefits from the inherent complementarity of the schemata in FN and WD and eventually leads to a better annotation coverage. We enhance the baseline model provided with the ARC task in order to contrast three system configurations: ‘+FN’, ‘+WD’ and ‘+FN/WD’.

Contributions We (1) present a proof of concept for semantic enrichment for the ARC task, (2) identify the importance of advanced combinations of complementary semantic annotations and (3) question the key role of external world knowledge with respect to ARC.

Code The code for the experiments with the enriched model is available online: <https://github.com/UKPLab/emnlp2018-argmin-commonsense-knowledge>

2 Our Approach: Semantic Enrichment for Argument Reasoning Comprehension (ARC)

First, we explain the ARC task together with the baseline that we will build upon (cf. Sec. 2.1). Second, we review our two external knowledge sources, FN and WD, and comment on their complementarity (cf. Sec. 2.2, 2.3, 2.4). Finally, we present our approach with preprocessing and the actual model enrichment (cf. Sec. 2.5, 2.6).

2.1 ARC Task

The ARC task (Habernal et al., 2018) is formulated as follows: given a debate title (a), claim (b) and reason (c), a system chooses the correct warrant (i) over the other (ii), see Figure 1. The warrants vary only slightly, e.g., by a single negation. The argumentation chain is sophisticated and uses logical reasoning and language understanding. In order to automatically draw the correct decision, a holistic

(a)	title: Can companies be trusted?
(b)	claim: Companies can't be trusted.
(c)	reason: Corporations have only goal: to make a profit.
(i)	warrant: they do not have to satisfy customers to make a profit
(ii)	warrant: they have to satisfy customers to make a profit

Figure 1: An instance of the ARC corpus, illustrating title (a), claim (b), reason (c) and the warrants (i) and (ii).

understanding of the context of both, claim and reason, is crucial – for which Habernal et al. (2018) recommend the inclusion of external knowledge.

Baseline The baseline provided by Habernal et al. (2018) is an intra-warrant attention model that reads in *Word2Vec* vectors (Mikolov et al., 2013) of all words in (a-c) and adapts attention weights for the decision between (i) and (ii).

Shared task winner The shared task winner, GIST (Choi and Lee, 2018), transfers inference knowledge (SNLI, Bowman et al., 2015) to the task of ARC and benefits from similar information in both datasets.

Our approach in contrast to GIST Our approach extends the baseline model with two external knowledge schemata, FN and WD, to explore their effects. The intuition can be explained with the instance in Figure 1: FN could be helpful by disambiguating ‘companies’ and ‘corporations’ to the same frame with meta-knowledge how it relates to other frames and WD could be of additional help by mapping them to entities with detailed information and examples for such institutions. We focus on utilizing the two knowledge schemata of FN and WD and thus, our interest is orthogonal to GIST. The advantage of our approach is independence of domain and task, which becomes especially relevant in scenarios lacking large-scale support data.

2.2 FrameNet’s Event Knowledge

The Berkeley FrameNet (Baker et al., 1998; Ruppenhofer et al., 2016) is an ongoing project for manually building a large lexical-semantic resource with expert annotations. It embodies the theory of frame semantics (Fillmore, 1976): frames capture units of meaning corresponding to prototypical situations. It consists of two parts, a lexicon that maps predicates to frames they can evoke, and fully annotated texts. For example, the verb *buy* can evoke either the frame *Commerce.buy* or *Fall.for*, depending on the context (buying goods versus buying a lie). Furthermore, the lexicon gives access to frame-specific role-labels (e.g., *Buyer*, *Goods* or *Deception*, *Victim*) as applied in semantic role labeling. Finally, FN specifies high-level relations (e.g., *inherit*, *precede*) between frames, forming a hierarchy with a collection of (*frame,relation,frame*)-triples. We use FN 1.5 which contains ~1K frames and ~12K distinct predicate-frame combinations.²

²framenet.icsi.berkeley.edu/fndrupal

2.3 Wikidata’s Fact Knowledge

Wikidata is a collaboratively constructed knowledge base that encodes world knowledge in the form of binary relations. It contains more than 40 million entities and 350 million relation instances.³ The binary relations express both semantic and ontological connections between the entities (e. g. CAPITAL(Hawaii, Honolulu); INSTANCE OF(Hawaii, location)). Wikidata includes an ontology of fine-grained classes and is interlinked with other semantic web resources. A broad community curation, similar to Wikipedia, ensures a higher data quality compared to other knowledge bases (Färber et al., 2015). Formally, Wikidata can be described as a graph $W = (E, R, I)$, where E is a set of entities, R is a set of binary relation types and I is a collection of relation instances encoded as $r(e_1, e_2)$, $r \in R$, $e_1, e_2 \in E$.

2.4 Complementarity of Annotations

Work on event semantics hints at two annotation types complementing each other: additional information about participants benefits event prediction (Ahrendt and Demberg, 2016; Botschen et al., 2018) and context information about events benefits the prediction of implicit arguments and entities (Cheng and Erk, 2018). The complementarity is further affirmed by efforts on aligning WD and the FN lexicon: the best alignment approach only maps 37% of the total WD properties to frames (Mousselly-Sergieh and Gurevych, 2016). The complementarity of FN and WD annotations is the reason for also testing a model with the joint annotation ‘+FN/WD’.

2.5 Preprocessing - Obtaining Annotations

We use two freely available systems to obtain semantic annotations for the claim (b), the reason (c) and the alternative warrants (i, ii): the frame identifier by Botschen et al. (2018) for frame annotations and the entity linker by Sorokin and Gurevych (2018). We employ pre-trained vector representations to encode information from FN and WD. We use the pre-trained frame embeddings (50-dim.) that are learned with *TransE* (Bordes et al., 2013) on the structure of the FN hierarchy with the collection of (*frame*, *relation*, *frame*)-triples (Botschen et al., 2017). We also use *TransE* to pre-train entity embeddings (100-dim.) on the WD graph. The an-

³www.wikidata.org/wiki/Special:Statistics



Figure 2: Different embeddings from layers of annotations for a sentence: words, frames, entities.

notation of the ARC data leads to more frames per sentence (6.6 on avg.) than entities per sentence (0.7 on avg.).

2.6 Model - Enriching with Semantics

We extend the baseline model by Habernal et al. (2018) with embeddings for frames and entities (cf. Sec. 2.5 for frame embeddings and entity embeddings). The baseline model is an intra-warrant attention model that only uses conventional pre-trained word embeddings as an input. We apply a *late fusion* strategy: obtain the annotations separately and combine them afterwards by appending the frame and entity embeddings to the word vectors on the token level. Each input sentence is processed by a bi-directional long short-term memory (LSTM) network that reads not only word embeddings, but also frame embeddings for all event mentions and entity embeddings for all entity mentions (Figure 2). Now, the attention weights for the decision between the two warrants are adapted based on the semantically enriched representation of claim (b) and reason (c).⁴

We optimize hyperparameters on the development set with random search. All models are trained using the Adam optimizer (Kingma and Ba, 2014) with a batch size of 16. For our evaluation, we perform ten runs and report the mean and max accuracy together with the standard deviation.

3 Results

In Table 1 we report our results on the ARC task. Our extended approaches ‘+FN’ and ‘+WD’ for semantic enrichment with information about frames and entities increase the averaged performance by more than one percentage point against the baseline. For the best run, the advantage of ‘+FN’ and ‘+WD’ becomes even clearer (+2.2 pp.). On the other hand, the straightforward combination of the two external knowledge source, ‘+FN/WD’, does not lead to further improvements. This points out the

⁴We refer to Habernal et al. (2018) for more details.

Approach	Dev.	mean		max Test
		(\pm)	(\pm)	
Habernal et al. (2018) (reimpl.)	0.6712	0.0155	0.5570	0.5878
+WD	0.6623	0.0071	0.5680	0.6036
+FN	0.6741	0.0119	0.5676	0.6104
+FN/WD	0.6630	0.0088	0.5592	0.5946

Table 1: Mean and max accuracy over ten runs on the ARC dev. and test sets (best results highlighted in bold).

need for advanced models that are able to fuse annotations of different types. Albeit positive the results do not seem to be a strong support for the hypothesis of (Habernal et al., 2018) about external knowledge being beneficial for the defined task, as we observe only moderate improvements. Overall, the enriched models (+WD, +FN and +FN/WD) make mostly the same predictions as the baseline system. For instance, for +WD there is 79,5% overlap of the predictions with the baseline, and for +FN, it is 76.6%. In the following section, we try to identify the reasons why the structured knowledge in the form of FN and WD does not further improve the results.

3.1 Error analysis

The amount of semantic information that the model can utilize is dependent on the number of annotations for a test instance⁵. We analyze the performance of the enriched models by the number of annotations for +WD and for +FN.

Figure 3 shows the performance of +WD in comparison to the baseline against the number of WD entities per test instance. As expected, there is no difference in performance for the instances

⁵Each instance is four sentences: a claim, a reason, a debate title and a candidate warrant.

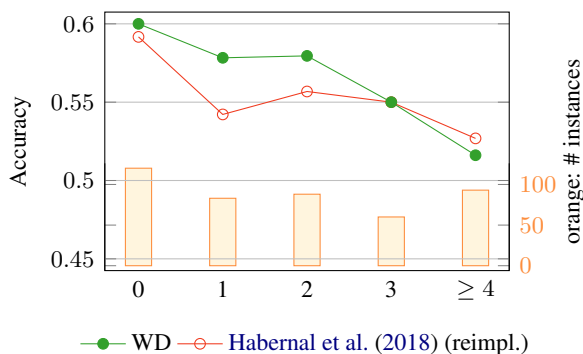


Figure 3: Performance for the +WD approach by the number of WD entities in a test set instance.

without WD annotations. We can see a clear improvement for the instances with one or two entities, which indicates that some semantic knowledge is helping to draw the decision between the two warrants. Contrary, +WD performs equal to the baseline for three or more annotations.

The performance of the +FN model against the number of the frame annotations is plotted in Figure 4. Whilst the difference between +FN and baseline varies more, we can observe a similar tendency: once some semantic annotations are available the enriched model outperforms the baseline, whereas with the growing number of frames the difference in performance decreases.

Both annotation tools, the WD entity linker as well as the FN frame identifier, introduce some noise: for the entity linker, Sorokin and Gurevych (2018) report 0.73 F-score and the frame identifier has an accuracy of 0.89 (Botschen et al., 2018). We perform a manual error analysis on 50 instances of the test set to understand the effect of the noisy WD annotation.⁶ In 44% of errors, no WD annotation was available and in 52%, the annotations were (partially) incorrect. Only 4% of errors oc-

⁶Judging if a predicted frame is correct requires deep linguistic expertise and special training on the FrameNet guidelines. Therefore, we excluded FN from this first study.

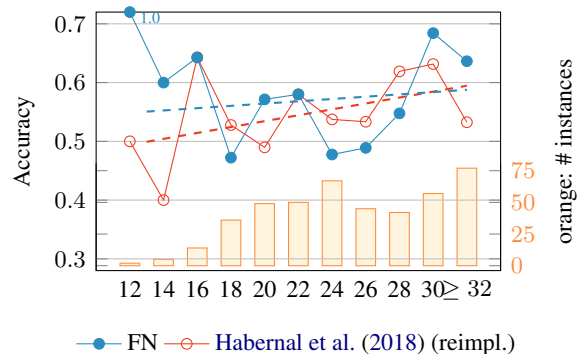


Figure 4: Performance for the +FN approach by the number of frames in a test set instance.

cur despite correct entities being available to the model. Notably, in 65% of the cases a correct entity annotation leads to a correct prediction.

Taken together, for instances with little context and therefore only some annotations with frames or entities, the semantic enrichment helps to capture a broader context of the argumentation chain which in turn benefits the classification. However, for instances with more context and therefore more annotations with frames or entities, the benefit is turned down by a worse precision of the annotation tools. Interestingly, the effect of improved performance only for shorter sequences with less annotations is in line with findings of research on information retrieval (Müller et al., 2008), where the trade-off between some annotations that increase the accuracy and more annotations that can hurt the performance is known as precision-recall balance (Riedel et al., 2010; Manning et al., 2008).

3.2 Qualitative analysis

When manually inspecting the annotations of frames and entities, it becomes questionable whether these actually have the potential of contributing to a clear distinction between the two warrants. Figure 5 shows two instances of the ARC corpus with FN and WD annotations. Both annotation layers contribute useful information about the world, which is not contained in the textual input. For instance, ‘companies’ and ‘corporations’ are correctly disambiguated and linked to the same frame and the phrase ‘use of force’ is mapped to the entity Q971119 for a legal concept. Nevertheless, when manually inspecting the annotations of frames and entities it becomes apparent that the provided background knowledge is not sufficient to draw the distinction between the two warrants. In the first example in Figure 5, the key difference between the two warrants is the negation (and similar in the second example). Even if our approach performs a semantic enrichment of the context, the crucial capability of performing reasoning is still missing. This means, our input representation is semantically enriched, but is not parsed into a logic-based representation.

To sum up, in the previous Section 3.1, we show that our approach is of help by successfully enriching the context with semantics for shorter instances; and in this Section 3.2, we analyze why our approach is too limited to solve some key challenges of the ARC task. We conclude with the

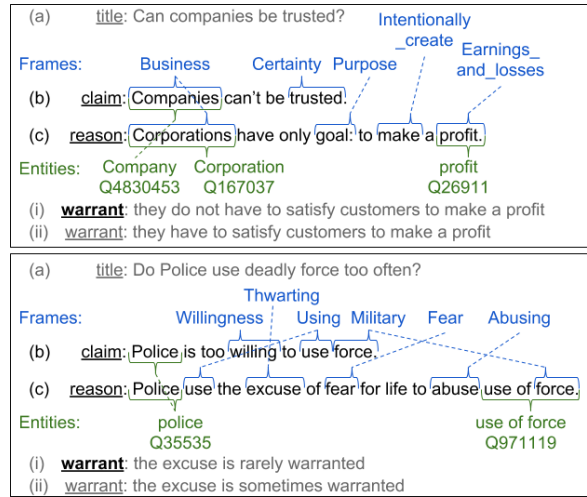


Figure 5: Instances of ARC corpus, with claim (b) and reason (c) annotated with frames and entities.

key challenge of ARC not being lexical-semantic gaps between warrants but rather different phenomena such as negation and that this challenge is to be resolved with logical analysis rather than with integrating world knowledge.

4 Conclusion

We integrate world knowledge from FrameNet and Wikidata into the task of common-sense argumentative reasoning and achieve an improvement in performance compared to the baseline approach. Based on the experiments and the manual analysis, we conclude that external world knowledge might not be enough to gain significant improvements in argumentative reasoning, and we rather point towards logical analysis.

Starting from the hypothesis of the evaluators of the shared task about world knowledge being essential for the Argument Reasoning Comprehension task, we show the potential of semantic enrichment of the context for shorter instances. Our results offer a first perspective on using external resources for the Argument Reasoning Comprehension task. More broadly, our approaches ‘+FN’ (events) and ‘+WD’ (facts) showcase the contribution of semantic enrichment to high-level tasks requiring common sense knowledge.

FrameNet and Wikidata are open-domain resources and our enrichment approach is task-independent. Consequently, we encourage utilizing event- and fact-knowledge in further language understanding tasks, e.g., Story Cloze (Mostafazadeh et al., 2016) or Semantic Textual Similarity (Agirre et al., 2012), using our freely available model.

Acknowledgments

This work has been supported by the DFG-funded research training group “Adaptive Preparation of Information from Heterogeneous Sources” (AIPHES, GRK 1994/1).

References

- Eneko Agirre, Mona Diab, Daniel Cer, and Aitor Gonzalez-Agirre. 2012. Semeval-2012 task 6: A pilot on semantic textual similarity. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pages 385–393. Association for Computational Linguistics.
- Simon Ahrendt and Vera Demberg. 2016. Improving event prediction by representing script participants. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 546–551, San Diego, USA. Association for Computational Linguistics.
- Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 1*, pages 86–90, Stroudsburg, PA, USA.
- Antoine Bordes, Nicolas Usunier, Alberto Garcia-Durán, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In *Proceedings of the 26th International Conference on Advances in Neural Information Processing Systems (NIPS)*, pages 2787–2795.
- Teresa Botschen, Iryna Gurevych, Jan-Christoph Klie, Hatem Mousselly Sergieh, and Stefan Roth. 2018. Multimodal Frame Identification with Multilingual Evaluation. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 1481–1491. Association for Computational Linguistics.
- Teresa Botschen, Hatem Mousselly-Sergieh, and Iryna Gurevych. 2017. Prediction of Frame-to-Frame Relations in the FrameNet Hierarchy with Frame Embeddings. In *Proceedings of the 2nd Workshop on Representation Learning for NLP (RepL4NLP, held in conjunction with ACL)*, pages 146–156.
- Samuel R. Bowman, Gabor Angeli, Christopher Potts, and Christopher D. Manning. 2015. A large annotated corpus for learning natural language inference. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 632–642. Association for Computational Linguistics.
- Qian Chen, Xiaodan Zhu, Zhen-Hua Ling, Diana Inkpen, and Si Wei. 2018. Neural Natural Language Inference Models Enhanced with External Knowledge. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 2406–2417, Melbourne, Australia. Association for Computational Linguistics.
- Pengxiang Cheng and Katrin Erk. 2018. Implicit argument prediction with event knowledge. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 831–840. Association for Computational Linguistics.
- HongSeok Choi and HyunJu Lee. 2018. Gist at semeval-2018 task 12: A network transferring inference knowledge to argument reasoning comprehension task. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (SemEval-Workshop)*, page (to appear), New Orleans, LA, USA. Association for Computational Linguistics.
- Michael Färber, Basil Ell, Carsten Menne, and Achim Rettinger. 2015. A Comparative Survey of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. *Semantic Web Journal*, 1:1–5.
- Charles J. Fillmore. 1976. Frame semantics and the nature of language. *Annals of the New York Academy of Sciences*, 280:20–32.
- Max Glockner, Vered Shwartz, and Yoav Goldberg. 2018. Breaking NLI Systems with Sentences that Require Simple Lexical Inferences. In *The 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, Melbourne, Australia.
- Jiang Guo, Wanxiang Che, Haifeng Wang, Ting Liu, and Jun Xu. 2016. A unified architecture for semantic role labeling and relation classification. In *Proceedings of the 26th International Conference on Computational Linguistics: Technical Papers (COLING)*, pages 1264–1274.
- Ivan Habernal, Henning Wachsmuth, Iryna Gurevych, and Benno Stein. 2018. The Argument Reasoning Comprehension Task: Identification and Reconstruction of Implicit Warrants. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 1930–1940, New Orleans, LA, USA. Association for Computational Linguistics.
- Daniel Khashabi, Tushar Khot, Ashish Sabharwal, and Dan Roth. 2018. Question Answering as Global Reasoning over Semantic Abstractions. In *Proceedings of the 32nd Conference of Association for the Advancement of Artificial Intelligence (AAAI)*, pages 1905–1914. AAAI Press.

- Diederik Kingma and Jimmy Ba. 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980*.
- Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press, New York, NY, USA.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. *arXiv preprint arXiv:1301.3781*.
- George A Miller. 1995. WordNet: a lexical database for English. *Communications of the ACM*, 38(11):39–41.
- Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. 2016. A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 839–849, San Diego, California. Association for Computational Linguistics.
- Hatem Mousselly-Sergieh and Iryna Gurevych. 2016. Enriching Wikidata with Frame Semantics. In *Proceedings of the 5th Workshop on Automated Knowledge Base Construction (AKBC, held in conjunction with NAACL)*, pages 29–34.
- Christof Müller, Iryna Gurevych, and Max Mühlhäuser. 2008. Closing the Vocabulary Gap for Computing Text Similarity and Information Retrieval. *International Journal of Semantic Computing*, 2(02):253–272.
- Hao Peng, Sam Thomson, Swabha Swayamdipta, and Noah A Smith. 2018. Learning joint semantic parsers from disjoint data. *arXiv preprint arXiv:1804.05990*.
- Sebastian Riedel, Limin Yao, and Andrew McCallum. 2010. Modeling Relations and Their Mentions without Labeled Text. In *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 148–163, Barcelona, Spain.
- Josef Ruppenhofer, Michael Ellsworth, Miriam R.L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. 2016. *FrameNet II: Extended Theory and Practice*, revised november 1, 2016 edition. International Computer Science Institute, Berkeley, USA.
- Daniil Sorokin and Iryna Gurevych. 2018. Mixing Context Granularities for Improved Entity Linking on Question Answering Data across Entity Categories. In *Proceedings of the 7th Joint Conference on Lexical and Computational Semantics (*SEM, held in conjunction with NAACL)*, pages 65–75. Association for Computational Linguistics.
- Su Wang, Greg Durrett, and Katrin Erk. 2018. Modeling Semantic Plausibility by Injecting World Knowledge. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, pages 303–308. Association for Computational Linguistics.
- Michael Zhai, Johnny Tan, and Jinho D Choi. 2016. Intrinsic and extrinsic evaluations of word embeddings. In *Proceedings of the 30th Conference of Association for the Advancement of Artificial Intelligence (AAAI)*, pages 4282–4283. AAAI Press.