

















language into facts. The experiments reveal the superiority of Logician in various open-domain information extraction tasks to the state-of-the-art algorithms.

Regarding future work, there are at least three promising directions. Firstly, one can investigate knowledge expression methods to extend SAOKE to express more complex knowledge, for tasks such as event extraction. Secondly, one can develop novel learning strategies to improve the performance of Logician and adapt the algorithm to the extended future version of SAOKE. Thirdly, one can extend SAOKE format and Logician algorithm in other languages.

## References

- [1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural Machine Translation By Jointly Learning To Align and Translate. In *Proceedings of ICLR*.
- [2] Michele Banko, Mj Cafarella, and Stephen Soderland. 2007. Open information extraction for the web. In *IJCAI*. 2670–2676.
- [3] Qingqing Cai and Alexander Yates. 2013. Large-scale Semantic Parsing via Schema Matching and Lexicon Extension. In *Proceedings of the 51st Annual Meeting of ACL*. 423–433.
- [4] Kaushik Chakrabarti, Surajit Chaudhuri, Tao Cheng, and Dong Xin. 2011. Entity-tagger: automatically tagging entities with descriptive phrases. In *Proceedings of the 20th International Conference Companion on WWW*. 19–20.
- [5] Wanxiang Che, Zhenghua Li, and Ting Liu. 2010. LTP: A Chinese Language Technology Platform. In *Proceedings of COLING*. 13–16.
- [6] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on EMNLP*. 1724–1734.
- [7] Janara Christensen, Mausam, Stephen Soderland, and Oren Etzioni. 2011. An analysis of open information extraction based on semantic role labeling. In *Proceedings of the sixth International Conference on Knowledge Capture*. 113–120.
- [8] Janara Christensen, Mausam, Stephen Soderland, Oren Etzioni, Mausam, Stephen Soderland, and Oren Etzioni. 2013. Towards Coherent Multi-Document Summarization. In *Proceedings of the 2013 Conference of NAACL: HLT*. 1163–1173.
- [9] Janara Christensen, Stephen Soderland, and Gagan Bansal. 2014. Hierarchical Summarization: Scaling Up Multi-Document Summarization. In *Proceedings of the 52nd Annual Meeting of ACL*. 902–912.
- [10] Luciano Del Corro and Rainer Gemulla. 2013. ClausIE: Clause-Based Open Information Extraction. In *Proceedings of the 22nd International Conference on WWW*. 355–366.
- [11] Li Dong and Mirella Lapata. 2016. Language to Logical Form with Neural Attention. In *In Proceedings of the Annual Meeting of ACL*. 33–43. arXiv:1601.01280
- [12] Oren Etzioni, Anthony Fader, Janara Christensen, Stephen Soderland, and Mausam. 2011. Open information extraction: The second generation. In *Proceedings of IJCAI* 3–10.
- [13] Anthony Fader, Stephen Soderland, and Oren Etzioni. 2011. Identifying Relations for Open Information Extraction. In *Proceedings of the Conference on EMNLP*. 1535–1545.
- [14] Anthony Fader, Luke S Zettlemoyer, and Oren Etzioni. 2014. Open Question Answering Over Curated and Extracted Knowledge Bases. In *Proceedings of the 20th ACM SIGKDD*. 1156–1165.
- [15] Jiatao Gu, Zhengdong Lu, Hang Li, and Victor O.K. Li. 2016. Incorporating Copying Mechanism in Sequence-to-Sequence Learning. In *Proceedings of the 54th Annual Meeting of ACL*. 1631–1640.
- [16] Rahul Gupta and A Halevy. 2014. Biperpedia: An Ontology for Search Applications. In *Proceedings of the VLDB Endowment*. 505–516.
- [17] Luheng He, Kenton Lee, Mike Lewis, and Luke Zettlemoyer. 2017. Deep Semantic Role Labeling: What Works and What's Next. In *Proceedings of the 55th Annual Meeting of the ACL*. 473–483.
- [18] Luheng He, Mike Lewis, and Luke Zettlemoyer. 2015. Question-Answer Driven Semantic Role Labeling: Using Natural Language to Annotate Natural Language. In *Proceedings of the 2015 Conference on EMNLP*. 643–653.
- [19] Marti A. Hearst. 1992. Automatic Acquisition of Hyponyms from Large Text Corpora. In *Proceedings of the 14th conference on Computational Linguistics*. 23–28.
- [20] Geoffrey Hinton, Nitish Srivastava, and Kevin Swersky. 2012. *Overview of mini-batch gradient descent*. Technical Report.
- [21] Nanda Kambhatla. 2004. Combining lexical, syntactic, and semantic features with maximum entropy models for extracting relations. In *Proceedings of the ACL: Interactive Poster and Demonstration Sessions*.
- [22] Rohit J Kate, Yuk Wah, and Wong Raymond. 2005. Learning to Transform Natural to Formal Languages. In *Proceedings of the 20th AAAI*. 1062–1068.
- [23] Tushar Khot, Ashish Sabharwal, and Peter Clark. 2017. Answering Complex Questions Using Open Information Extraction. In *Proceedings of the 55th Annual Meeting of the ACL*. 311–316. arXiv:1704.05572
- [24] Tom Kwiatkowski, Eunsol Choi, Yoav Artzi, and Luke Zettlemoyer. 2013. Scaling Semantic Parsers with On-the-fly Ontology Matching. In *Proceedings of the 2013 Conference on EMNLP*. 1545–1556.
- [25] Jinyang Li, Chengyu Wang, Xiaofeng He, Rong Zhang, and Ming Gao. 2015. User Generated Content Oriented Chinese Taxonomy Construction. In *Lecture Notes in Computer Science*. Vol. 9313. 623–634.
- [26] Yankai Lin, Shiqi Shen, Zhiyuan Liu, Huanbo Luan, and Maosong Sun. 2016. Neural Relation Extraction with Selective Attention over Instances. In *Proceedings of the 54th Annual Meeting of ACL*. 2124–2133.
- [27] Christopher D Manning, John Bauer, Jenny Finkel, Steven J Bethard, Mihai Surdeanu, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of 52nd Annual Meeting of ACL: System Demonstrations*. 55–60. arXiv:arXiv:1011.1669v3
- [28] Mausam. 2016. Open Information Extraction Systems and Downstream Applications. In *Proceedings of the 25th IJCAI*. 4074–4077.
- [29] Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th IJCNLP*, Vol. 2. 1003.
- [30] Makoto Miwa and Mohit Bansal. 2016. End-to-End Relation Extraction using LSTMs on Sequences and Tree Structures. In *Proceedings of the 54th Annual Meeting of ACL*. 1105–1116.
- [31] Harinder Pal and Mausam. 2016. Demonyms and Compound Relational Nouns in Nominal Open IE. In *Proceedings of the 5th Workshop on AKBC*. 35–39.
- [32] Likun Qiu and Yue Zhang. 2014. ZORE: A Syntax-based System for Chinese Open Relation Extraction. In *Proceedings of the 2014 Conference on EMNLP*. 1870–1880.
- [33] John W Ratcliff and David E Metzener. 1988. Pattern Matching: The Gestalt Approach. *Dr Dobbs's* 13, 7 (1988).
- [34] Sebastian Riedel, Limin Yao, Andrew McCallum, and Benjamin M Marlin. 2013. Relation Extraction with Matrix Factorization and Universal Schemas. *Proceedings of the 2013 Conference of NAACL: HLT June (2013)*, 74–84.
- [35] Sebastian Riedel, Limin Yao, Andrew McCallum, and Benjamin M Marlin. 2013. Relation Extraction with Matrix Factorization and Universal Schemas. In *Proceedings of the 2013 Conference of NAACL: HLT*. 74–84.
- [36] Michael Schmitz, Robert Bart, Stephen Soderland, and Oren Etzioni. 2012. Open language learning for information extraction. In *Proceedings of the 2012 Joint Conference on EMNLP and CoNLL*. 523–534.
- [37] Stephen Soderland, Brendan Roof, Bo Qin, and Shi Xu. 2010. Adapting Open Information Extraction to Domain-Specific Relations. *AI Magazine* 31, 3 (2010), 93–102.
- [38] Gabriel Stanovsky and Ido Dagan. 2016. Creating a Large Benchmark for Open Information Extraction. In *Proceedings of the 2016 Conference on EMNLP*. 2300–2305.
- [39] Gabriel Stanovsky, Ido Dagan, and Mausam. 2015. Open IE as an Intermediate Structure for Semantic Tasks. In *Proceedings of the 53rd Annual Meeting of ACL and the 7th IJCNLP*. 303–308.
- [40] Zhaopeng Tu, Zhengdong Lu, Yang Liu, Xiaohua Liu, and Hang Li. 2016. Modeling Coverage for Neural Machine Translation. In *Proceedings of the Annual Meeting of ACL (2016)*. 76–85.
- [41] Vered Schwartz, Yoav Goldberg, Ido Dagan, Vered Schwartz, Yoav Goldberg, and Ido Dagan. 2016. Improving hypernymy detection with an integrated path-based and distributional method. In *Proceedings of the 54th Annual Meeting of ACL*. 2389–2398. arXiv:1603.06076
- [42] Chengyu Wang and Xiaofeng He. 2017. A Short Survey on Taxonomy Learning from Text Corpora: Issues, Resources and Recent Advances. In *Proceedings of the Conference on EMNLP*.
- [43] Wikipedia. 2017. Assignment problem—Wikipedia, The Free Encyclopedia. (2017).
- [44] Fei Wu and Daniel S Weld. 2010. Open Information Extraction using Wikipedia. In *Proceedings of the 48th Annual Meeting of ACL*. 118–127.
- [45] Wentao Wu, Hongsong Li, Haixun Wang, and Kenny Q Zhu. 2012. Probase: A probabilistic taxonomy for text understanding. In *Proceedings of the 2012 ACM SIGMOD*. 481–492.
- [46] Mohamed Yahya, Steven Euijong Whang, Rahul Gupta, and Alon Halevy. 2014. ReNoun: Fact Extraction for Nominal Attributes. In *Proceedings of the Conference on EMNLP 2014, Doha, Qatar*. 325–335.
- [47] Pengcheng Yin, Zhengdong Lu, Hang Li, and Ben Kao. 2016. Neural Enquirer: Learning to Query Tables. In *In Proceedings of the Annual Meeting of ACL*. 29–35.
- [48] Dmitry Zelenko, Chinatsu Aone, Anthony Richardella, Jaz Kandola, Thomas Hofmann, Tomaso Poggio, and John Shawe-Taylor. 2003. Kernel Methods for Relation Extraction. *Journal of Machine Learning Research* 3 (2003), 1083–1106.
- [49] Luke S Zettlemoyer and Michael Collins. 2005. Learning to Map Sentences to Logical Form: Structured Classification with Probabilistic Categorical Grammars. In *Proceedings of the 21st Conference on UAI*. 658–666.
- [50] Suncong Zheng, Feng Wang, Hongyun Bao, Yuexing Hao, Peng Zhou, and Bo Xu. 2017. Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme. *Proceedings of the 55th Annual Meeting of the ACL (2017)*, 1227–1236.