Evaluating Open Relation Extraction Over Conversational Texts

Mahsa Imani, Yashar Mehdad, Giuseppe Carenini, Raymond T. Ng
Department of Computer Science
University of British Columbia
Vancouver, Canada
mimani, mehdad, carenini, rng@cs.ubc.ca

1 Introduction

With the invention of internet and prevalence of email systems, blogs, fora discussions and social networking, now people around the world can participate in conversations and discuss their thoughts, feelings and opinions about new products and services. As a result, conversational data is growing in an exponential rate. Relation extraction (RE) is the task of finding relationships between entities in text and is an effective way to convert unstructured text data into structured knowledge which is required to effectively summarize, search and find relevant information. State of-the-art Open IE systems such as ReVerb [8], WOE [15], OLLIE [11], SONEX [12], TreeKernel [16], EXEMPLAR [7] extract web-scale information in the format of relational tuples (*arg1; rel; arg2*) in which the relation phrase *rel* expresses a relation between *arg1* and *arg2*.

Sentences in conversational data such as social streams, chat logs, blogs and Email threads are complex and noise-prone[2]. They often have an ungrammatical colloquial language, more abbreviations, and may not state the full relation which is often assumed in relation extraction task. Furthermore, performance of these techniques depends on the output of preprocessing steps such as POS tagging, NP chunking, NE tagging and dependency parsing whose accuracy degrade while running over conversational text. In order to partially tackle such challenges, we propose to exploit text simplification as a preprocessing step prior to relation extraction. There are several algorithms for text simplification including lexical, syntactic rules [5, 6, 13] and log-linear models [1]. Our hypothesis is that text simplification can improve the accuracy of preprocessing steps [4, 5, 13] as well as relation extraction by breaking down each complex sentence into semantically equivalent shorter sentences. To evaluate Open IE over conversational data, we created a dataset sampled from a wide range of conversational corpora. This work highlights the following contributions: i) we sample a conversational dataset from various conversational modalities; ii) we evaluate the performance of a state-of-the-art Open IE system over our sampled dataset; and iii) we evaluate the performance of the Open IE system after text simplification.

2 Method

To evaluate Open RE over conversational data, we create a dataset sampled from a wide range of conversational corpora including synchronous conversations (AMI [3] and ICSI corpus [9]), microblogs (tweets), threaded or asynchronous conversations (Email [14] and blog threads), and reviews on products and services (Opinosis Dataset[10]). To obtain a representative sample of sentences, we use two-stage stratified sampling. In the first stage, to capture key characteristics of each corpus, each corpus plays the role of one stratum independently. In the second stage, 100 sentences were sampled from each corpus. We extracte a set of syntactic and conversational features from each sentence and group them based on the resulting feature vectors.

For Open IE, we use OLLIE as one of the state-of-the-art available systems which does not need any training or tuning. OLLIE learns general open extraction patterns using high confidence extractions of ReVerb as input seeds of its bootsrapper. For text simplification, we use TriS [1] which uses syntactic rules in order to transform text input to a simpler version. We evaluate the extracted relation by OLLIE, before and after running the text simplification (i.e., TriS), based on the number of extracted relations, the accuracy of the extracted argument and relation phrases and the confidence score of the extracted relations. Experimental results are presented in the next section.

3 Experimental Results and Discussion

We first evaluate the performance of OLLIE on our created dataset before text simplification and report the accuracy when only relation phrase is correct and when both arguments and relation phrase are correct. In the next experiment, we run the text simplification over the created dataset and then we extract the relations using OLLIE. This step is referred as OLLIE-Simplified in our results. The result of these experiments has been shown in Table 1 and Table 2. The columns, from left to right, show number of extractions, accuracy of the relation phrase and accuracy when both arguments and relation phrase are correct. Tables 3 and 4 show average confidence scores assigned to the extracted relations. From left to right, the columns show average confidence score of all extractions, average confidence score of correct relations, average confidence score of the incorrect relations, and average confidence score when both arguments and relation phrase are correct. In all tables, the bold numbers show the cases OLIIE-Simplified outperforms OLIIE. TriS fails to simplify most of the sentences in AMI and ICICS corpus due to lack of punctuations or wrong punctuations which made sentences too long to be simplified by TriS.

Experimental results show that OLLIE-Simplified outperforms OLLIE in terms of accuracy and informativeness of the confidence score for all datasets except Slashdot. Text simplification is more promising in increasing the accuracy of OLLIE for Twitter dataset which has much more cryptic content and abbreviations than other conversations due to its length limit (140 characters). Overall, the most difficult conversational modality for relation extraction for both systems is reviews. The easiest ones

Dataset	#Extracts	Rel. phrase	All correct
ICSI	292	47.9%	45.2%
AMI	650	71.5%	43.2%
BC3	**148**	73.0%	48.6%
Slashdot	**301**	76.4%	**54.1%**
Reviews	**372**	64.5%	40.9%
Twitter	90	62.2%	45.6%

Table 1: Accuracy before simplification.

Dataset	#Extracts	Rel. phrase	All correct
BC3	141	**77.3%**	**58.2%**
Slashdot	211	**76.8%**	51.5%
Reviews	233	**68.2%**	**44.2%**
Twitter	**99**	**72.7%**	**55.6%**

Table 2: Accuracy after simplification.

Dataset	All	Corr. rel.	Incorr. rel	All corr.
ICSI	0.6	0.43	0.15	0.29
AMI	0.56	0.4	0.15	0.26
BC3	0.61	0.46	0.14	0.31
Slashdot	0.66	0.49	0.14	**0.35**
Reviews	0.64	0.42	0.21	0.27
Twitter	0.7	0.44	0.26	0.34

Table 3: Average confidence score before simplification.

Dataset	All	Corr. rel.	Incorr. rel	All corr.
BC3	**0.69**	**0.54**	**0.11**	**0.4**
Slashdot	**0.7**	0.49	**0.12**	0.33
Reviews	**0.66**	**0.47**	**0.17**	**0.31**
Twitter	**0.74**	**0.54**	**0.19**	**0.42**

Table 4: Average confidence score after simplification.

for OLIIE are blogs and emails and for OLLIE-Simplified are emails and microblogs (tweets). In reviews, the use of incomplete sentences and informal text effects the performance of OLLIE. Emails written in a corporation usually have more formal and grammatical sentences and technical blogs like Slashdot have more technical and grammatical sentences which makes relation extraction less challenging. Based on the above finding, as a future work, we propose the use of different Open IE and text simplification systems and aggregation of the extracted relations in order to improve the performance of our current approach. Both find new relations not found by the other one. So, a new system utilizing the union of extracted relations of two systems will outperform both systems in term of recall.

References

[1] Nguyen Bach, Qin Gao, Stephan Vogel, and Alex Waibel. Tris: A statistical sentence simplifier with log-linear models and margin-based discriminative training. In *IJCNLP*, pages 474–482, 2011.

[2] Giuseppe Carenini, Gabriel Murray, and Raymond Ng. *Methods for Mining and Summarizing Text Conversations*. Morgan & Claypool Publishers, 1st edition, 2011.

[3] Jean Carletta, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, Guillaume Lathoud, Mike Lincoln, Agnes Lisowska, Iain McCowan, Wilfried Post, Dennis Reidsma, and Pierre Wellner. The ami meeting corpus: A pre-announcement. In *Proceedings of the Second International Conference on Machine Learning for Multimodal Interaction*, MLMI'05, pages 28–39, Berlin, Heidelberg, 2006. Springer-Verlag.

[4] R. Chandrasekar, Christine Doran, and B. Srinivas. Motivations and methods for text simplification. In *PROCEEDINGS OF THE SIXTEENTH INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS (COLING '96)*, 1996.

[5] R. Chandrasekar and B. Srinivas. Automatic induction of rules for text simplification, 1997.

[6] Jan De Belder and Marie-Francine Moens. Text simplification for children. In *Prroceedings of the SIGIR Workshop on Accessible Search Systems,*, pages 19–26. ACM, 2010.

[7] Filipe de S Mesquita, Jordan Schmidek, and Denilson Barbosa. Effectiveness and efficiency of open relation extraction. In *EMNLP*, pages 447–457. ACL, 2013.

[8] Anthony Fader, Stephen Soderland, and Oren Etzioni. Identifying relations for open information extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11, pages 1535–1545, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.

[9] Adam Janin, Don Baron, Jane Edwards, Dan Ellis, David Gelbart, Nelson Morgan, Barbara Peskin, Thilo Pfau, Elizabeth Shriberg, Andreas Stolcke, and Chuck Wooters. The icsi meeting corpus. pages 364–367, 2003.

[10] Won Kim, Ron Kohavi, Johannes Gehrke, and William DuMouchel, editors. *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, Washington, USA, August 22-25, 2004*. ACM, 2004.

[11] Mausam, Michael Schmitz, Robert Bart, Stephen Soderland, and Oren Etzioni. Open language learning for information extraction. EMNLP-CoNLL '12, pages 523–534, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.

[12] Yuval Merhav, Filipe Mesquita, Denilson Barbosa, Wai Gen Yee, and Ophir Frieder. Extracting information networks from the blogosphere. *ACM Trans. Web*, 6(3):11:1–11:33, October 2012.

[13] Advaith Siddharthan. Syntactic simplification and text cohesion. Technical report, Research on Language and Computation, 2003.

[14] J. Ulrich, G. Murray, and G. Carenini. A publicly available annotated corpus for supervised email summarization. In *AAAI08 EMAIL Workshop*, Chicago, USA, 2008. AAAI.

[15] Fei Wu and Daniel S. Weld. Open information extraction using wikipedia. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, ACL '10, pages 118–127, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[16] Ying Xu, Mi-Young Kim, Kevin Quinn, Randy Goebel, and Denilson Barbosa. Open information extraction with tree kernels. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 868–877, Atlanta, Georgia, June 2013. Association for Computational Linguistics.