Statistical Models for Frame-Semantic Parsing

Dipanjan Das
Google Inc.
76 9th Avenue,
New York, NY 10011
dipanjand@google.com

Abstract

We present a brief history and overview of statistical methods in frame-semantic parsing – the automatic analysis of text using the theory of frame semantics. We discuss how the FrameNet lexicon and frame-annotated datasets have been used by statistical NLP researchers to build usable, state-of-the-art systems. We also focus on future directions in frame-semantic parsing research, and discuss NLP applications that could benefit from this line of work.

1 Frame-Semantic Parsing

Frame-semantic parsing has been considered as the task of automatically finding semantically salient **targets** in text, disambiguating their semantic **frame** representing an event and scenario in discourse, and annotating arguments consisting of words or phrases in text with various frame elements (or **roles**). The FrameNet lexicon (Baker et al., 1998), an ontology inspired by the theory of frame semantics (Fillmore, 1982), serves as a repository of semantic frames and their roles. Figure 1 depicts a sentence with three evoked frames for the targets "million", "created" and "pushed" with FrameNet frames and roles.

Automatic analysis of text using frame-semantic structures can be traced back to the pioneering work of Gildea and Jurafsky (2002). Although their experimental setup relied on a primitive version of FrameNet and only made use of "exemplars" or example usages of semantic frames (containing one target per sentence) as opposed to a "corpus" of sentences, it resulted in a flurry of work in the area of automatic semantic role labeling (Màrquez et al., 2008). However, the focus of semantic role labeling (SRL) research has mostly been on PropBank (Palmer et al., 2005) conventions, where verbal targets could evoke a "sense" frame, which is not shared across targets, making the frame disambiguation setup different from the representation in FrameNet. Furthermore, it is fair to say that early research on PropBank focused primarily on argument structure prediction, and the interaction between frame and argument structure analysis has mostly been unaddressed (Màrquez et al., 2008). There are exceptions, where the verb frame has been taken into account during SRL (Meza-Ruiz and Riedel, 2009; Watanabe et al., 2010). Moreoever, the CoNLL 2008 and 2009 shared tasks also include the verb and noun frame identification task in their evaluations, although the overall goal was to predict semantic dependencies based on PropBank, and not full argument spans (Surdeanu et al., 2008; Hajič et al., 2009).

The SemEval 2007 shared task (Baker et al., 2007) attempted to revisit the frame-semantic analysis task based on FrameNet. It introduced a larger FrameNet lexicon (version 1.3), and also a larger corpus with full-text annotations compared to prior work, with multiple targets annotated per sentence. The corpus allowed words and phrases with noun, verb, adjective, adverb, number, determiner, conjunction and preposition syntactic categories to serve as targets and evoke frames, unlike any other single dataset; it also allowed targets from different syntactic categories share frames, and therefore roles. The repository of semantic role types was also much richer than PropBank-style lexicons, numbering in several hundreds.

Most systems participating in the task resorted to a cascade of classifiers and rule-based modules: identifying targets (a non-trivial subtask), disambiguating frames, identifying potential arguments, and then labeling them with roles. The system described by Johansson and Nugues (2007) performed the best in this shared task. Next, we focus on its performance, and subsequent improvements made by the research community on this task.

Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014), pages 26–29,
Baltimore, Maryland USA, June 27, 2014.

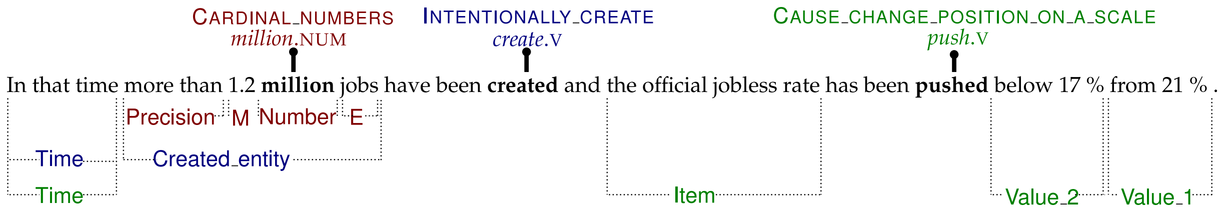

Figure 1: A partial depiction of frame-semantic structures taken from Das et al. (2014). The words in bold correspond to targets, which evoke semantic frames that are denoted in capital letters. Above each target is shown the corresponding lexical unit, which is a lemma appended by a coarse part-of-speech tag. Every frame is shown in a distinct color; each frame's arguments are annotated with the same color, and are marked below the sentence, at different levels. For the CARDINAL_NUMBERS frame, "M" denotes the role Multiplier and "E" denotes the role Entity.

		P	R	F_1
SemEval'07 Data (automatic targets)	Johansson and Nugues (2007)	51.59	35.44	42.01
	Das et al. (2010)	**58.08**	**38.76**	**46.49**
FrameNet 1.5 Release (gold targets)	Das et al. (2014)	68.33	61.14	64.54
	Hermann et al. (2014)	**72.79**	**64.95**	**68.64**

Table 1: We show the current state of the art on the frame-semantic parsing task. The first section shows results on the SemEval 2007 shared task. The best system in the task, presented by Johansson and Nugues (2007) was later outperformed by SEMAFOR, a system described by Das et al. (2010). Both systems use a rule-based module to identify targets. On the FrameNet 1.5 data, Das et al. (2014) presented additional semi-supervised experiments using gold targets, which was recently outperformed by an approach presented by Hermann et al. (2014) that made use of distributed word representations.

2 Current State of the Art

Johansson and Nugues (2007) presented the system that resulted in the best F_1 score on the SemEval 2007 task of collectively identifying frame-evoking targets, a disambiguated frame for each target, and the set of role-labeled arguments for each frame. The system contained a set of rule-based heuristics to identify targets followed by a cascade of three learned models as mentioned in §1. Das et al. (2010) presented a tool called SEMAFOR,[1] which improved upon this system with a similar framework for target identification, but only used two probabilistic models, one for frame identification, and one for predicting the arguments. The frame identification subpart involved a latent-variable log-linear model, which intended to capture frames for unseen targets, many of which appeared in the test data. Moreover, the feature sets in both the models were sufficiently different from prior work, resulting in improvements. Table 1 shows results on the SemEval 2007 data for these two systems.

The FrameNet project released more annotations and a larger frame lexicon in 2010; Das et al. (2014) used this dataset, and presented a variety of experiments improving upon their prior work, setting the new state of the art. A few salient aspects of this updated version of SEMAFOR involved handling unseen targets using a graph-based semi-supervised learning approach and improved inference using a dual decomposition algorithm. Subsequently, Hermann et al. (2014) used a very similar framework but presented a novel method using distributed word representations for better frame identification, outperforming the aforementioned update to SEMAFOR. Table 1 shows the performance in terms of F_1 score for frames and arguments given gold targets. Recent work on the FrameNet corpora, including the aforementioned two papers have used gold targets to measure the performance of statistical methods because the distribution of annotated targets in the data varied significantly across documents and domains, making it difficult to build a learnable system for target identification.

The aforementioned papers focused on the task of sentence-internal frame-semantic analysis. There have been some investigation of finding implicit arguments of frames that may be present in other parts of a document, outside the sentential context. Although there has not been extensive research on this topic, a shared task at SemEval 2010 focused on this problem (Ruppenhofer et al., 2010).[2] Moreover, there has been significant effort

[1]See `http://www.ark.cs.cmu.edu/SEMAFOR`.

[2]Related work on the analysis of implicit arguments for

in developing unsupervised techniques for inducing frame-semantic structures (Modi et al., 2012), to induce FrameNet-like lexicons from weak supervision, such as syntactic parses.

3 Applications

Shallow semantic analysis based on FrameNet data has been recently utilized across various natural language processing applications with success. These include the generation of meeting summaries (Kleinbauer, 2012), the prediction of stock price movement using (Xie et al., 2013), inducing slots for domain-specific dialog systems (Chen et al., 2013), stance classification in debates (Hasan and Ng, 2013), modeling the clarity of student essays (Persing and Ng, 2013) to name a few.

There is strong potential in using frame-semantic structures in other applications such as question answering and machine translation, as demonstrated by prior work using PropBank-style SRL annotations (Shen and Lapata, 2007; Liu and Gildea, 2010).

4 Future Directions

Given the wide body of work in frame-semantic analysis of text, and recent interest in using frame-semantic parsers in NLP applications, the future directions of research look exciting.

First and foremost, to improve the quality of automatic frame-semantic parsers, the coverage of the FrameNet lexicon on free English text, and the number of annotated targets needs to increase. For example, the training dataset used for the state-of-the-art system of Hermann et al. (2014) contains only 4,458 labeled targets, which is approximately 40 times less than the number of annotated targets in Ontonotes 4.0 (Hovy et al., 2006), a standard NLP dataset, containing PropBank-style verb annotations. This comparison is important because FrameNet covers many more syntactic categories than the PropBank-style annotations, and features more than 1,000 semantic role labels compared to 51 in Ontonotes, but severely lacks annotations. A machine learned system would find it very hard to generalize to new data given such data sparsity. Increasing the quantity of such annotations requires exhaustive inter-annotator agreement studies (which has been rare in FrameNet corpora generation) and the development of annotation guidelines, such that these annotations can be produced outside the FrameNet project.

Other than increasing the amount of labeled data, there is a necessity of automatically aligning predicate-level semantic knowledge present in resources like FrameNet, PropBank, NomBank and VerbNet (Schuler, 2005). These lexicons share a lot of knowledge about predicates and current resources like Ontonotes do align some of the information, but a lot remains missing. For example, alignment between these lexicons could be done within a statistical model for frame-semantic parsing, such that correlations between the coarse semantic role labels in PropBank or NomBank and the finer labels in FrameNet could be discovered automatically.

Finally, the FrameNet data is an attractive test bed for semi-supervised learning techniques because of data sparsity; distributed word representations, which often capture more semantic information than surface-form features could be exploited in various aspects of the frame-semantic parsing task.

nominal targets in NomBank (Meyers et al., 2004) has been investigated recently (Gerber and Chai, 2012).

Acknowledgments

The author thanks Desai Chen, André Martins, Nathan Schneider and Noah Smith for numerous discussions and several years of collaboration on this topic at Carnegie Mellon. He is also grateful to Kuzman Ganchev, Oscar Täckström, Karl Moritz Hermann, Ryan McDonald, Slav Petrov, Fernando Pereira and the greater Research at Google community for helping research in this area thrive over the past couple of years.

References

C. F. Baker, C. J. Fillmore, and J. B. Lowe. 1998. The Berkeley Framenet project. In *Proceedings of COLING-ACL.*

C. F. Baker, M. Ellsworth, and K. Erk. 2007. SemEval-2007 Task 19: Frame semantic structure extraction. In *Proceedings of SemEval.*

Y.-N. Chen, W. Y. Wang, and A. Rudnicky. 2013. Unsupervised induction and filling of semantic slots for spoken dialogue systems using frame-semantic parsing. In *Proceedings of ASRU.*

D. Das, N. Schneider, D. Chen, and N. A. Smith. 2010. Probabilistic frame-semantic parsing. In *Proceedings of NAACL-HLT.*

D. Das, D. Chen, A. F. T. Martins, N. Schneider, and N. A. Smith. 2014. Frame-semantic parsing. *Computational Linguistics*, 40(1):9–56.

C. J. Fillmore. 1982. Frame Semantics. In *Linguistics in the Morning Calm*, pages 111–137. Hanshin Publishing Co., Seoul, South Korea.

M. Gerber and J. Y. Chai. 2012. Semantic role labeling of implicit arguments for nominal predicates. *Computational Linguistics*, 38(4):755–798.

D. Gildea and D. Jurafsky. 2002. Automatic labeling of semantic roles. *Computational Linguistics*, 28(3):245–288.

J. Hajič, M. Ciaramita, R. Johansson, D. Kawahara, M. A. Martí, L. Màrquez, A. Meyers, J. Nivre, S. Padó, J. Štěpánek, P. Straňák, M. Surdeanu, N. Xue, and Y. Zhang. 2009. The CoNLL-2009 shared task: Syntactic and semantic dependencies in multiple languages. In *Proceedings of CoNLL*.

K. S. Hasan and V. Ng. 2013. Frame semantics for stance classification. In *Proceedings of CoNLL*.

K. M. Hermann, D. Das, J. Weston, and K. Ganchev. 2014. Semantic frame identification with distributed word representations. In *Proceedings of ACL*.

E. Hovy, M. Marcus, M. Palmer, L. Ramshaw, and R. Weischedel. 2006. Ontonotes: The 90% solution. In *Proceedings of NAACL-HLT*.

R. Johansson and P. Nugues. 2007. LTH: semantic structure extraction using nonprojective dependency trees. In *Proceedings of SemEval*.

T. Kleinbauer. 2012. *Generating automated meeting summaries*. Ph.D. thesis, Saarland University.

D. Liu and D. Gildea. 2010. Semantic role features for machine translation. In *Proceedings of COLING*.

L. Màrquez, X. Carreras, K. C. Litkowski, and S. Stevenson. 2008. Semantic role labeling: an introduction to the special issue. *Computational Linguistics*, 34(2):145–159.

A. Meyers, R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, and R. Grishman. 2004. The NomBank project: An interim report. In *Proceedings of NAACL/HLT Workshop on Frontiers in Corpus Annotation*.

I. Meza-Ruiz and S. Riedel. 2009. Jointly identifying predicates, arguments and senses using Markov logic. In *Proceedings of NAACL-HLT*.

A. Modi, I. Titov, and A. Klementiev. 2012. Unsupervised induction of frame-semantic representations. In *Proceedings of the NAACL-HLT Workshop on the Induction of Linguistic Structure*.

M. Palmer, D. Gildea, and P. Kingsbury. 2005. The Proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106.

I. Persing and V. Ng. 2013. Modeling thesis clarity in student essays. In *Proceedings of ACL*.

J. Ruppenhofer, C. Sporleder, R. Morante, C. Baker, and M. Palmer. 2010. SemEval-2010 task 10: Linking events and their participants in discourse. In *Proceedings of SemEval*.

K. K. Schuler. 2005. *Verbnet: A Broad-coverage, Comprehensive Verb Lexicon*. Ph.D. thesis, University of Pennsylvania.

D. Shen and M. Lapata. 2007. Using semantic roles to improve question answering. In *Proceedings of EMNLP-CoNLL*.

M. Surdeanu, R. Johansson, A. Meyers, L. Màrquez, and J. Nivre. 2008. The CoNLL 2008 shared task on joint parsing of syntactic and semantic dependencies. In *Proceedings of CoNLL*.

Y. Watanabe, M. Asahara, and Y. Matsumoto. 2010. A structured model for joint learning of argument roles and predicate senses. In *Proceedings of ACL*.

B. Xie, R. J. Passonneau, L. Wu, and G. G. Creamer. 2013. Semantic frames to predict stock price movement. In *Proceedings of ACL*.