A Survey on
Label-Efficient Information Extraction with Natural Language Guidance

Qinyuan Ye
qinyuany@usc.edu

1 Introduction

Natural language explanations (NLE), as additional information provided by the annotator, are informative and effortless (*i.e.* requiring only tripled annotation time since the annotator already knows the instance well). Recent work in relation extraction, common-sense QA and visual QA utilize these explanations and achieve competitive performances efficiently (i.e., with fewer labeling effort). One way to leverage NLE is to construct labeling rules from explanations, which help relieve annotation burden, while such method usually adopt hard matching and suffer from low coverage issue. The proposed work aims to further improve the efficiency by constructing a joint bootstrapping framework for both instances and rules, so that rule coverage in unlabeled corpora is further expanded. In this survey we discuss related work on natural language explanations and machine executable forms.

2 Related Work

Natural Language Explanation. Srivastava et al. (2017) first introduced the usage of natural language explanation in concept learning. Each statement s in set $\mathcal{S}$ is first parsed into logical form with a CCG parser and acts like a binary feature function $z = f(x) \in \{0, 1\}$. The original representation of the instance, x, is augmented with binary feature outputs z, and is later fed into a classifier. The training objective $\log p(y|x, \mathcal{S})$ is roughly decomposed into classification part $\log p(y|x, z)$ and parsing part $\log p(z|x, \mathcal{S})$ with Jensen's Inequality. The proposed work demonstrates good generalization ability with a 30% relative F1 improvement on their crowd-sourced dataset, and the authors highlight the potential application for non machine-learning experts to guide model training with natural language. Though only a small number of instances are required to achieve competitive performance, the task is done in a purely supervised manner, and still much manual labor is spent on this task, including crowd-sourcing and domain-specific lexicon construction.

More recently, Hancock et al. (2018) proposed a BabbleLabble framework for training classifiers with NLE, and succeeded in three relation extraction tasks. One major difference from the prior work is that BabbleLabble abandoned trainable CCG-based parser and adopted a simpler and fixed rule-based parser with no domain-specific predicate dictionary. The drawback of this approach is that it iterates subspans of explanation sentence, construct a labeling function for each subspan, and then filter the invalid ones. The computation cost for such iteration and filtering may be huge. Another difference is that the z is not used to augment x but is used to pseudo-label x instead. This characteristic enables semi-supervised learning on unlabeled corpus. One highlight of this work is it focus on efficiency instead of pushing forward the state-of-the-art — how many instance-label pairs are equivalent to one instance-label pair with natural language explanation in terms of contribution to model performance? Is it still efficient to use explanations when annotation time is taken into account? The proposed work cast our attention to these practical and meaningful issues in model training.

There are also *generative* approaches to leverage natural language explanations. Li et al. (2018) framed visual question answering into multi-task learning by requiring the model to generate an explanation s with hidden representation $\mathbf{h}$ constructed from image and question, i.e. modeling $p(s|\mathbf{h})$. The multi-task learning paradigm enables the generation task to regularize hidden representation $\mathbf{h}$ and thus improve the QA task performance. Rajani et al. (2019) leveraged advanced pre-trained, OpenAI GPT (Radford et al., 2018), to generate an explanation sentence with the question and candidate answers as the previous sentence. This work assumes that pre-trained models have "common-sense reasoning ability" to provide surface explanations, while the reasoning process remains a black box.

Natural language explanations can be considered as additional information from annotators. Prior to natural language explanation, there are also notable attempts in acquiring extra information, but in other forms such as highlighting decisive phrases, or *rationale* (Zaidan et al., 2007). Highlighting is less expressive, but more formalized in representation (i.e., can be expressed with binary sequences) and more easy for annotators to use. Previously, such information is leveraged with generative annotation modeling (Zaidan and Eisner, 2008) or sub-graph features (Arora and Nyberg, 2009). One potential way of using rationale is to align highlight sequence with attention scores over tokens.

Machine Executable Forms. Transforming nature language into machine executable forms is usually done by semantic parser and is extensively studied. Liang (2016) provided an extensive survey on related methods. Surface text pattern is the simplest form, which provides a positive output if a string is exactly matched in candidate sentence. Another representative method is combinatory categorial grammar (CCG) (Zettlemoyer and Collins, 2012) which originates in linguistics and then borrowed to semantic parsing.

The above semantic parsing approaches are explicit, interpretable, but can only be used with hard-matching. With the rise of large-scale pretrained models, soft and generalized execution can be achieved. Weber et al. (2019) softens first-order logic reasoning by maintaining an embedding for each entity, relation (*i.e.*, entity-masked sentence) and predicate. Entity unification is done by calculating the similarity between the embeddings of two candidate entities; and entailment can be grounded if the embedding of relation is close to that of a predicate. This approach is half explicit (interpretable rules, unification and entailment) and half implicit (hidden representation initialized with pre-trained models). One severe drawback of this method is differentiat-ability. Given supportive facts and rules, a *deterministic* solver is first used to solve the logical problem. As a workaround, the result from this solver guides the model to re-calculate a loss following its best solution. The trainable embeddings are updated with back propagation of this re-calculated loss. In practice, such deterministic solver are computationally expensive. Meanwhile, the meaning of (placeholder) predicates are still un-interpretable.

One on-going work within our group is to soften CCG-based labeling functions with soft neural execution tree (SNE-Tree). SNE-Trees split the operations into four categories: (1) logic calculation; (2) distance counting; (3) string matching and (4) deterministic. Operations in the first three categories can be softened. For example, distance requirement of less than 5 can be softened to a distance of 7 though the matching score is lowered; same thing happens to "CEO" matched with "chief executive". Ideally, such soft-matching scheme will expand rule coverage so that unlabeled corpus is better leveraged. However, CCG limits how the explanations are expressed and NLE become less 'natural' — In practice distance counting and string matching modules may still be insufficient to explain the reason, and it's hard to explain common sense with words in a pre-defined list.

3 Datasets

Srivastava et al. (2017) provided a crowd-sourced dataset with natural language statements to describe seven concepts in email texts. Hancock et al. (2018) gathered natural language explanations for three relation extraction tasks (Spouse, Disease and Protein). Rajani et al. (2019) constructed a Common Sense Explanations (CoS-E) dataset on top of Common sense Question Answering (CQA) dataset. In computer vision, image captions are good sources of natural language description, and can be potentially formed into natural language explanation for image-related classification tasks. Related dataset include VQA-E (Li et al., 2018) and MSCOCO (Lin et al., 2014). These datasets may be used in the research project.

4 Brief Conclusion

In this survey we broadly explored related research on natural language explanation and machine executable forms. Natural language explanation can be either leveraged in representation (augmenting representation of input), as labeling function (providing weak supervision for unlabeled corpus), or as regularization (as a second task in multitask learning). As for machine executable forms, we discuss simple surface text pattern and more complex CCG-based rules, as well as techniques to soften the rule grounding process. Pros and cons of each method are discussed so that a suitable approach may be chosen in fulfillment of the project.

References

Shilpa Arora and Eric Nyberg. 2009. Interactive annotation learning with indirect feature voting. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Student Research Workshop and Doctoral Consortium*, pages 55–60. Association for Computational Linguistics.

Braden Hancock, Martin Bringmann, Paroma Varma, Percy Liang, Stephanie Wang, and Christopher Ré. 2018. Training classifiers with natural language explanations. In *Proceedings of the conference. Association for Computational Linguistics. Meeting*, volume 2018, page 1884. NIH Public Access.

Qing Li, Qingyi Tao, Shafiq Joty, Jianfei Cai, and Jiebo Luo. 2018. Vqa-e: Explaining, elaborating, and enhancing your answers for visual questions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 552–567.

Percy Liang. 2016. Learning executable semantic parsers. *COMMUNICATIONS OF THE ACM*, 59(9).

Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer.

Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. Improving language understanding by generative pre-training. *URL https://s3-us-west-2. amazonaws. com/openai-assets/researchcovers/languageunsupervised/language understanding paper. pdf.*

Nazneen Fatema Rajani, Bryan McCann, Caiming Xiong, and Richard Socher. 2019. Explain yourself! leveraging language models for commonsense reasoning. *arXiv preprint arXiv:1906.02361*.

Shashank Srivastava, Igor Labutov, and Tom Mitchell. 2017. Joint concept learning and semantic parsing from natural language explanations. In *Proceedings of the 2017 conference on empirical methods in natural language processing*, pages 1527–1536.

L Weber, P Minervini, J Münchmeyer, U Leser, and T Rocktäschel. 2019. Nlprolog: Reasoning with weak unification for question answering in natural language. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, ACL 2019, Florence, Italy, Volume 1: Long Papers*, volume 57. ACL (Association for Computational Linguistics).

Omar Zaidan, Jason Eisner, and Christine Piatko. 2007. Using annotator rationales to improve machine learning for text categorization. In *Human language technologies 2007: The conference of the North American chapter of the association for computational linguistics; proceedings of the main conference*, pages 260–267.

Omar F Zaidan and Jason Eisner. 2008. Modeling annotators: A generative approach to learning from annotator rationales. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 31–40. Association for Computational Linguistics.

Luke S Zettlemoyer and Michael Collins. 2012. Learning to map sentences to logical form: Structured classification with probabilistic categorial grammars. *arXiv preprint arXiv:1207.1420*.