0.1 Semantic Role Labeling

Semantic role labeling (SRL), also known as shallow semantic parsing, is the task of semantically annotating natural language text. Conventionally, a syntactically parsed sentence is taken as input, and semantic arguments associated with predicate of the sentence are identified first, and then classified to particular semantic classes. The identification subtask is commonly known as Semantic Role Identification (SRI), while the classification subtask is known as Semantic Role Classification (SRC). For a better understanding, reconsider the previously given sentence: "The waiter served fresh water to the guests". In this sentence, "serve" is the verbal predicate, which has three associated semantic arguments (i.e. the noun phrases 'The waiter', 'fresh water', and 'the guests'). In a perfect scenario, SRI will identify these three arguments (among others), which will be classified to be ARG0, ARG1, and ARG2 respectively by the SRC part. ARG0, ARG1, and ARG2 are semantic role labels defined by the Propbank labeling scheme. See [1] for definitions and details about these and other semantic role labels.

The first automatic SRL systems was reported by Gildea and Jurafsky in 2002 [3], and since then, their ideas have been dominating the field. In their approach, they emphasized on selection of appropriate lexical and syntactical features for SRL, use of statistical classifiers and their combinations, and ways to handle the data sparseness issue. Many researchers have tried to build on that by augmenting and/or altering the feature set [8], by experimenting with various classification approaches [6, 5], and by attempting different ways to handle data sparseness [10]. Others have tried to extend it with new ways e.g. [2] tried a hierarchical feature selection strategy, while [4] proposed to exploit argument interdependence (i.e. the semantic role of one argument depends on the semantic roles of other arguments).

For this study, we developed a SRL system using previously explored features, statistical classifiers, and exploiting argument interdependence. In the following paragraphs, we will first briefly describe a baseline SRL system and explain how argument interdependence has been previously used. Later, we will propose a two layer architecture to better exploit the argument interdependence. Due to space limitations, we will concentrate on argument interdependence part, and will not go into details of the baseline SRL system.

A base line system was developed using Maximum Entropy classifiers, and the feature set given in Table 1. PenTreebank and Propbank data was used for building and training the models, which were then used for SRI and SRL. Let's consider an abstract example to see what argument interdependence is and how it has been used previously?

Take Fig. 1 and suppose for any given predicate P in a sentence, the SRI subtask has identified three potential arguments (i.e. A1, A2, and A3) of the predicate. Next, the task in the SRC is to predict semantic role labels of those three arguments. For that, a critical observation made by [4] is that the semantic roles of arguments may depend on each other. This phenomena is known as argument interdependence. A common way to escapade argument interdependence is to turn the SRC task into a sequence labeling problem, and use the features extracted from arguments around the current argument, in addition to the features of the current one, to predict label for the current argument. For example, if while predicting label of argument A2, the features extracted

Feature	**Explanation**
From [9]	
t-word	Target Word (i.e. Predicate)
t-w-pos	Target Word Part of Speech Tag
h-word	Head Word
h-w-pos	Head Word Part of Speech Tag
position	Position of constituent in focus w.r.t predicate
pt	Phrase Type
subcat	Subcate Frame
path	Parse tree path from constituent in focus to the predicate
path-to-BA	Parse tree path from constituent in focus to BA node
path-to-BEI	Parse tree path from constituent in focus to BEI node
verbClass	The rule that expands the constituent in focus
verbClass-pls-pt	Verb class + phrase type
verbClass-pls-h-word	Verb class + head word
first-word	First word of constituent in focus
last-word	Last word of constituent in focus
t-word-pls-pt	Target word + phrase type
From [2]	
voicePosition	
subcatAt	The rule that expands the constituent in focus
subcatStar	The rule that expands the parent node of constituent in focus
l-sib-pt	Phrase type of the left sibling of the constituent in focus
r-sib-pt	Phrase type of the right sibling of the constituent in focus
l-c-focus	Layer of constituent in focus
allFrameSet	The combination of all the framesets of a predicate
allFrameSet-pls-verbClass	allFrameSet + verb class
allFrameSet-pls-h-word	allFrameSet + head word
allFrameSet-pls-pt	allFrameSet + phrase type
semType-h-word	Semantic type of head word
semType-t-word	Semantic type of target word
semType-first-word	Semantic type of of first word
semType-last-word	Semantic type of of last word
semType-t-pls-l-word	Semantic type of target + last word
From [7]	
voice	Voice of the sentence
From [3]	
gov	Governing Category

Table 1: Feature Used for Baseline SRL

from arguments A1 and A3 are also used, this can assist the system to correctly predict the label for A2. As far the scope of the interdependence, usually, the concept of a window size is used for this purpose. For example, a wind size of [-m,n] means that the features of m preceding and n following arguments will be used to predict label for the current node. The window-size strategy have some practical limits, since the feature set is normally very big, resulting in the use of smaller window size (the window size of [-1,1] is common). The smaller window size means argument interdependence is not used to its full potential.

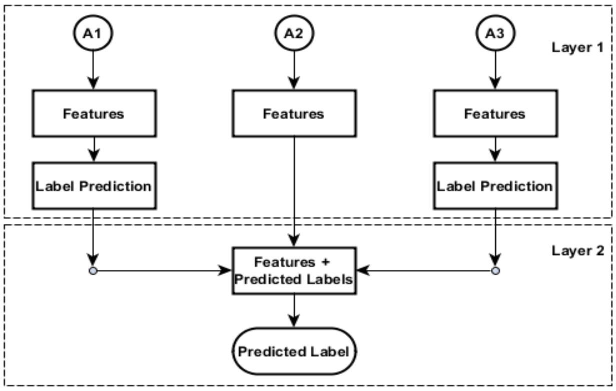

Figure 1: Two Layered SRL Architecture

To overcome the limitations of window-size strategy and to better exploit argument interdependence, we suggest to use the predicted labels of the surrounding arguments, instead of their features, to predict the label of the current node. For this purpose, we propose a two-layered SRC architecture shown in Fig. 1. As can be seen, a baseline system is used at layer 1 to predict the labels of surroundings arguments (i.e. A1 and A3), and then at layer 2, these predicted labels together with other features of the current node (i.e. A2) are used to build and train models which are then used to predict the label of the current node. Experimental results show that this strategy works better than the window-size strategy, and hence, results in a better system performance. However, further details are beyond the scope of this paper, and are planned to given in another article.

References

[1] Claire Bonial, Olga Babko-Malaya, Jinho D. Choi, Jena Hwang, and Martha Palmer. Propbank annotation guidelines. Center for Computational Language and Education Research Institute of Cognitive Science University of Colorad at Boulder, 2010. http://clear.colorado.edu/compsem/documents/ propbank_guidelines.pdf.

[2] Weiwei Ding and Baobao Chang. Improving chinese semantic role classification with hierarchical feature selection strategy. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP

'08, pages 324–333, Stroudsburg, PA, USA, 2008. Association for Computational Linguistics.

[3] Daniel Gildea and Daniel Jurafsky. Automatic labeling of semantic roles. *Comput. Linguist.*, 28(3):245–288, September 2002.

[4] Zheng Ping Jiang, Jia Li, and Hwee Tou Ng. Semantic argument classification exploiting argument interdependence. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, IJCAI'05, pages 1067–1072, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc.

[5] Kyung-Mi Park and Hae-Chang Rim. Maximum entropy based semantic role labeling. In *Proceedings of the Ninth Conference on Computational Natural Language Learning*, CONLL '05, pages 209–212, Stroudsburg, PA, USA, 2005. Association for Computational Linguistics.

[6] Sameer Pradhan, Wayne Ward, Kadri Hacioglu, and James H. Martin. Shallow semantic parsing using support vector machines. 2004.

[7] Honglin Sun and Daniel Jurafsky. Shallow semantc parsing of chinese. In Daniel Marcu Susan Dumais and Salim Roukos, editors, *HLT-NAACL 2004: Main Proceedings*, pages 249–256, Boston, Massachusetts, USA, May 2 - May 7 2004. Association for Computational Linguistics.

[8] Nianwen Xue. Calibrating features for semantic role labeling. In *In Proceedings of EMNLP 2004*, pages 88–94, 2004.

[9] Nianwen Xue. Labeling chinese predicates with semantic roles. *Comput. Linguist.*, 34(2):225–255, June 2008.

[10] Beñat Zapirain, Eneko Agirre, and Lluís Màrquez. Ubc-upc: Sequential srl using selectional preferences: An aproach with maximum entropy markov models. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, SemEval '07, pages 354–357, Stroudsburg, PA, USA, 2007. Association for Computational Linguistics.