Probabilistic Shift-Reduce Parsing Model Using Rich Contextual Information

Yong-Jae Kwak[1], So-Young Park[1], Joon-Ho Lim[1], Hae-Chang Rim[1]

[1] Natural Language Processing Lab., Dept. of CSE,
Korea University, Anam-dong 5-ga, Seongbuk-gu, 136-701, Seoul, Korea
{yjkwak,ssoya,jhlim,rim}@nlp.korea.ac.kr

Abstract. In this paper, we present a probabilistic shift-reduce parsing model which can overcome low context-sensitivity of previous LR parsing models. Since previous models are restricted by LR parsing framework, they can utilize only a lookahead and a LR state (stack). The proposed model is not restricted by LR parsing framework, and is able to add rich contextual information as needed. To show an example of contextual information designed for applying the proposed model to Korean, we devise a new context scheme named "*surface-context-types*" which uses syntactic structures, sentential forms, and selective lexicals. Experimental results show that rich contextual information used by our model can improve the parsing accuracy, and our model outperforms the previous models even when using a lookahead alone.

1 Probabilistic Shift-Reduce Parsing Model

Since the first approach [1] and [2] of integrating a probabilistic method with the LR parsing technique, some standard probabilistic LR parsing models have been implemented. [3] and [4] (or [5]) defined a parse tree candidate T as a transition sequence of LR state [3] or LR stack [4] that is driven by an action and a lookahead, as follows:

$$s_0 \overset{l_1,a_1}{\Rightarrow} s_1 \overset{l_1,a_1}{\Rightarrow} \ldots \Rightarrow s_{n-1} \overset{l_m,a_m}{\Rightarrow} s_m \text{ [3]} \quad \sigma_0 \overset{l_1,a_1}{\Rightarrow} \sigma_1 \overset{l_1,a_1}{\Rightarrow} \ldots \Rightarrow \sigma_{n-1} \overset{l_m,a_m}{\Rightarrow} \sigma_m \text{ [4]} \tag{1}$$

where s_i and σ_i are the i-th state [3] and stack [4], l_i is the i-th lookahead, a_i is the action that can be performed for the lookahead and the state (or the stack), and m is the number of actions to complete parsing procedure. A state/stack transition sequence gives the following probabilistic model:

$$P(T) = \prod_{i=1\ldots n} P(l_i, a_i, s_i \mid s_{i-i}) \text{ [3]} \quad P(T) = \prod_{i=1\ldots n} P(l_i, a_i, \sigma_i \mid \sigma_{i-i}) \text{ [4]} \tag{2}$$

These models are less context-sensitive, because the selection of action can be affected by information beyond the LR parsing framework, such as LR parsing table, LR stack, lookahead.

As actions are performed, not only the stack but also the input are changed. We propose a probabilistic shift-reduce parsing model considering both of the stack and the input word sequence $W = w_1 \ldots w_n$ as follows: