

Recurrent Neural Network Grammers

Chris Dyer, Adhiguna Kuncoro, Miguel Ballesteros,
Noah A. Smith

導入

Recurrent Neural Network Grammars (RNNGs)

- ▶ 文の確率的生成モデル
 - ▶ 単語や句の入れ子的・階層的構造を陽に表現
- ▶ 問題：構文解析，文生成
- ▶ 動機：Sequential な Recurrent Neural Networks (RNNs) は自然言語の潜在的な入れ子構造を考慮できていない
- ▶ 構文解析または文生成のアルゴリズム

提案手法

RNNG の定義

$$RNNG := (N, \Sigma, \Theta)$$

$$\left\{ \begin{array}{l} N : \text{非終端記号の有限集合} \\ \Sigma : \text{終端記号の有限集合} (N \cup \Sigma = \emptyset) \\ \Theta : \text{NN のパラメータ} \end{array} \right.$$

提案手法

構文解析のアルゴリズム

$$f : X \rightarrow Y$$

$$\left\{ \begin{array}{l} x : \text{終端記号（単語）の列（入力）} \\ y : \text{構文木（出力）} \\ S : \text{スタック} \\ B : \text{入力バッファ} \end{array} \right.$$

- ▶ スタックの要素：終端記号，open または closed な非終端記号
- ▶ 入力バッファの要素：終端記号

提案手法

構文解析のアルゴリズム

- ▶ 初期状態

$$S = \emptyset$$

$$B = [T_1, \dots, T_n]$$

提案手法

構文解析のアルゴリズム

▶ 遷移の制約

- ▶ n : スタック内の open な非終端記号の数

遷移	制約
nt(X)	$B \neq \emptyset \wedge n < 100$
SHIFT	$B \neq \emptyset \wedge n \geq 1$
REDUCE	スタック内の一番上の要素が open な非終端記号でない $\wedge (n \geq 2(1 \text{ じゃないの?}) \vee B = \emptyset)$

提案手法

文生成のアルゴリズム

$$f : X \rightarrow Y$$

$$\left\{ \begin{array}{l} x : ? \\ y : \text{終端記号 (単語) の列 (出力)} \\ S : \text{スタック} \\ \textcolor{red}{T} : \text{出力バッファ} \end{array} \right.$$

- ▶ スタックの要素：終端記号，open または closed な非終端記号
- ▶ 出力バッファの要素：終端記号

提案手法

文生成のアルゴリズム

- ▶ 初期状態

$$S = \emptyset$$

$$T = \emptyset$$

提案手法

文生成のアルゴリズム

▶ 遷移の制約

- ▶ n : スタック内の open な非終端記号の数

遷移	制約
GEN(X)	$n \geq 1$
REDUCE	スタック内の一番上の要素が open な非終端記号でない $\wedge n \geq 1$

提案手法

生成モデル

- ▶ 単語列 (x) と構文木 (y) の結合確率

$$\begin{aligned} p(x, y) &= \prod_{t=1}^{|a(x, y)|} p(a_t | a_{<t}) \\ &= \prod_{t=1}^{|a(x, y)|} \frac{\exp r_{a_t}^T u_t + b_{a_t}}{\sum_{a' \in A_G(T_t, S_t, n_t)} \exp r_{a'}^T u_t + b_{a'}} \\ u_t &= \tanh(W[o_t; s_t; h_t] + c) \end{aligned}$$

$$\begin{cases} u_t : \text{アルゴリズムの状態を表す埋め込み} \\ r_a : \text{各遷移の埋め込み (パラメータ)} \\ b_a : \text{各遷移のバイアス (パラメータ)} \end{cases}$$

提案手法

生成モデル

- ▶ u_t : アルゴリズムの状態を表す埋め込み

$$u_t = \tanh(W[o_t; s_t; h_t] + c)$$

$$\left\{ \begin{array}{l} o_t : \text{出力バッファの状態を表す埋め込み} \\ s_t : \text{スタックの状態を表す埋め込み} \\ h_t : \text{遷移歴を表す埋め込み} \\ W, c : \text{パラメータ} \end{array} \right.$$