

Механизмы внимания для машинного перевода

Практикум на ЭВМ 2017/2018

Фоминская Галина Евгеньевна

МГУ имени М. В. Ломоносова, факультет ВМК, кафедра ММП

10 мая 2018 г.

- Дано: предложение на одном языке (исходное)
- Выход: предложение на другом языке (целевое)
- Для одного предложения много правильных переводов
- Метрика качества: BLEU (bilingual evaluation understudy)

Метрика BLEU (bilingual evaluation understudy)

$$p_n = \frac{\sum_{C \in \text{Candidates}} \sum_{ngram \in C} Count_{clip}(ngram)}{\sum_{C' \in \text{Candidates}} \sum_{ngram' \in C'} Count(ngram')},$$

Метрика BLEU (bilingual evaluation understudy)

$$p_n = \frac{\sum_{C \in \text{Candidates}} \sum_{ngram \in C} Count_{clip}(ngram)}{\sum_{C' \in \text{Candidates}} \sum_{ngram' \in C'} Count(ngram')},$$

$$Count_{clip}(ngram) = \max \left(Count(ngram), \max_{true_translations} Count(ngram) \right)$$

Метрика BLEU (bilingual evaluation understudy)

$$p_n = \frac{\sum_{C \in \text{Candidates}} \sum_{ngram \in C} Count_{clip}(ngram)}{\sum_{C' \in \text{Candidates}} \sum_{ngram' \in C'} Count(ngram')},$$

$$Count_{clip}(ngram) = \max \left(Count(ngram), \max_{true_translations} Count(ngram) \right)$$

$$r = \sum_{C \in \text{Candidates}} best_match_length(C),$$

Метрика BLEU (bilingual evaluation understudy)

$$p_n = \frac{\sum_{C \in \text{Candidates}} \sum_{ngram \in C} \text{Count}_{clip}(ngram)}{\sum_{C' \in \text{Candidates}} \sum_{ngram' \in C'} \text{Count}(ngram')},$$

$$\text{Count}_{clip}(ngram) = \max \left(\text{Count}(ngram), \max_{\text{true_translations}} \text{Count}(ngram) \right)$$

$$r = \sum_{C \in \text{Candidates}} \text{best_match_length}(C),$$

$$c = \sum_{C \in \text{Candidates}} \text{len}(C),$$

Метрика BLEU (bilingual evaluation understudy)

$$p_n = \frac{\sum_{C \in \text{Candidates}} \sum_{ngram \in C} \text{Count}_{clip}(ngram)}{\sum_{C' \in \text{Candidates}} \sum_{ngram' \in C'} \text{Count}(ngram')},$$

$$\text{Count}_{clip}(ngram) = \max \left(\text{Count}(ngram), \max_{\text{true_translations}} \text{Count}(ngram) \right)$$

$$r = \sum_{C \in \text{Candidates}} \text{best_match_length}(C),$$

$$c = \sum_{C \in \text{Candidates}} \text{len}(C),$$

$$\log BLEU_N = \min \left(1 - \frac{r}{c}, 0 \right) + \sum_{n=1}^N w_n \log p_n.$$

Обычно $N = 4$, $w_n = \frac{1}{N}$.

Метрика BLEU (bilingual evaluation understudy)

Достоинства:

- + Быстро и легко считается
- + Было показано, что есть корреляция с человеческими оценками

Недостатки:

- Не учитывает важность слов
- Не учитывает грамматическую корректность, т. к. работает на уровне фраз из 4-х слов
- BLEU человеческих переводов не сильно выше автоматических, хотя сам перевод лучше

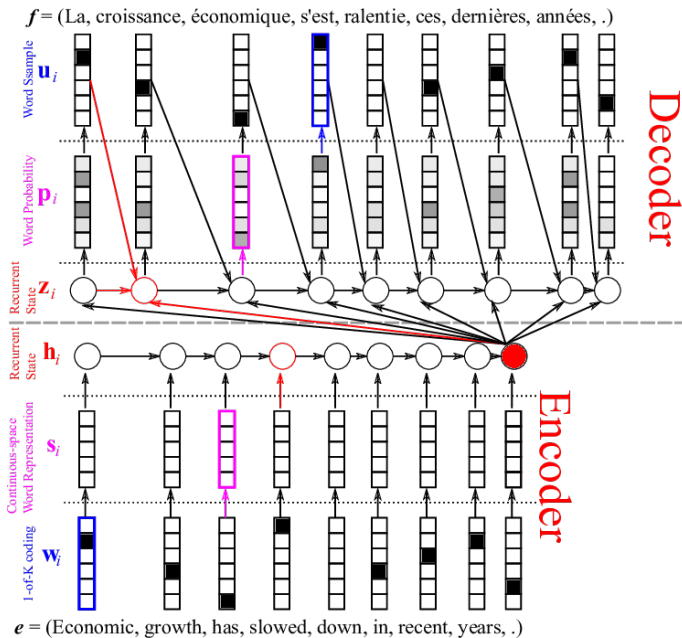
Классические подходы:

- Решения, основанные на задании правил и признаков, использующие предположения о структуре языка

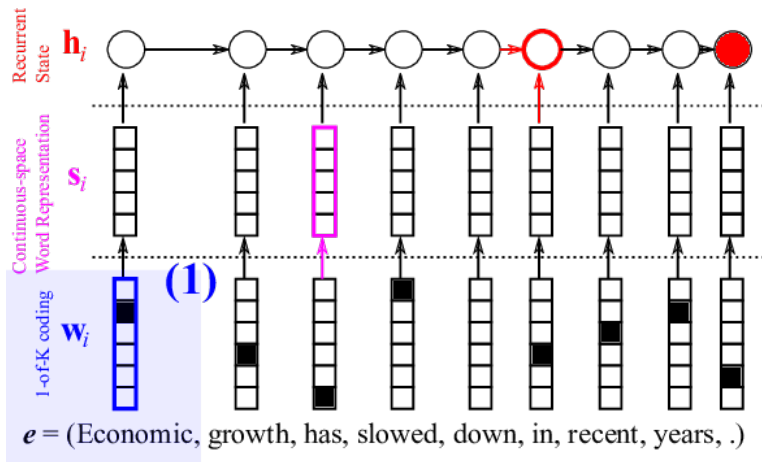
С использованием нейросетей:

- Энкодер-декодер
- Модель с механизмом внимания

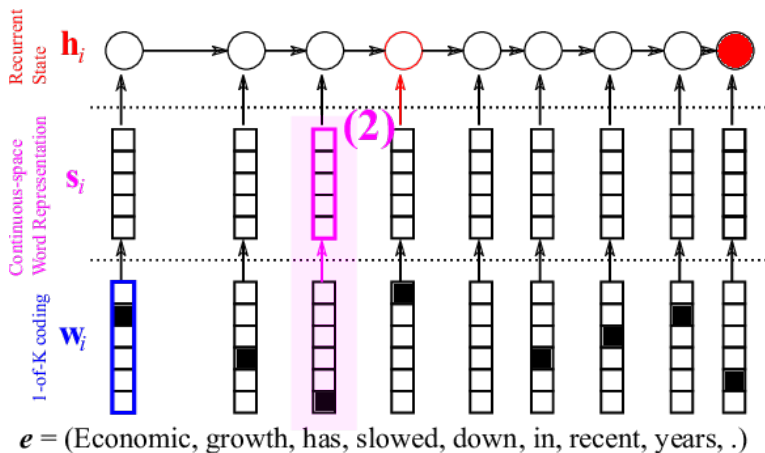
Encoder-decoder



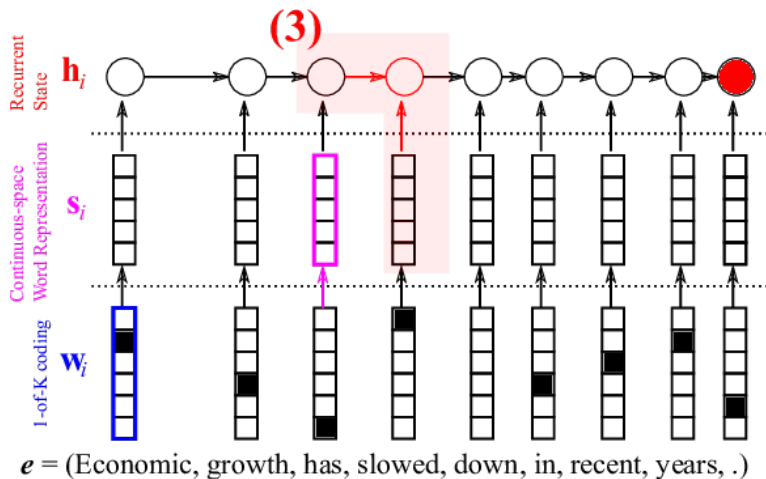
Encoder-decoder



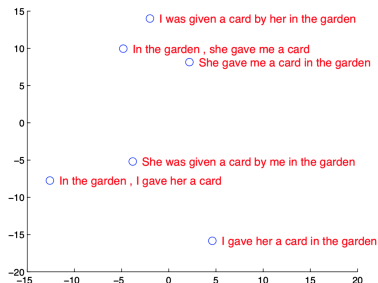
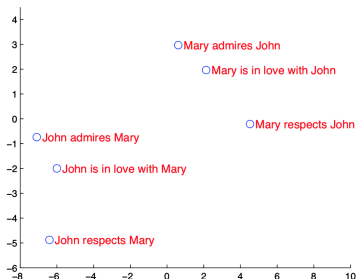
Encoder-decoder



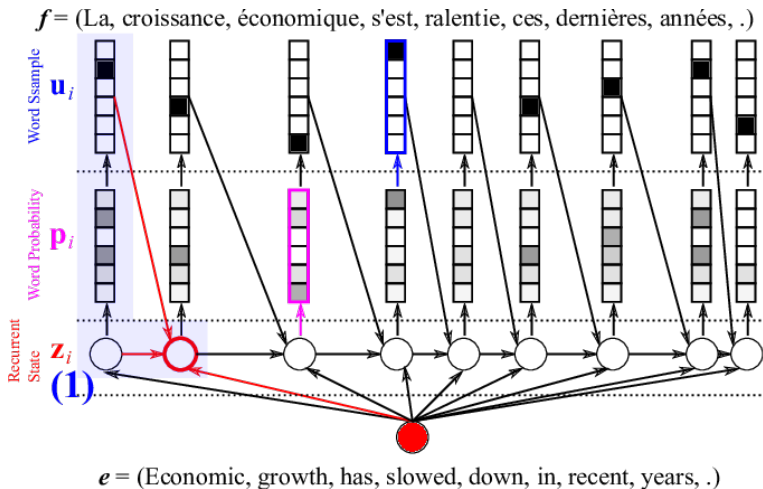
Encoder-decoder



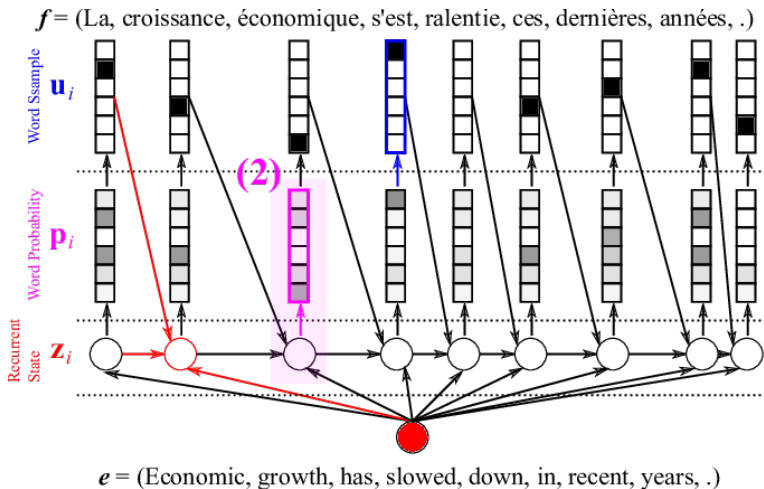
Внутренние векторные представления предложений



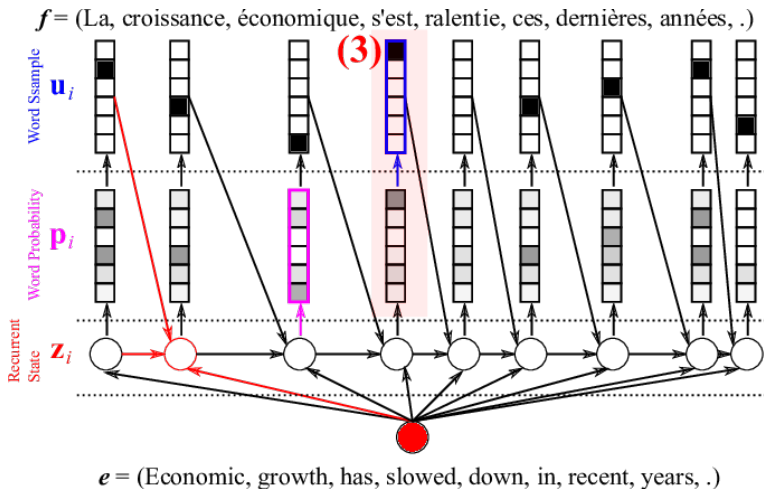
Encoder-decoder



Encoder-decoder



Encoder-decoder

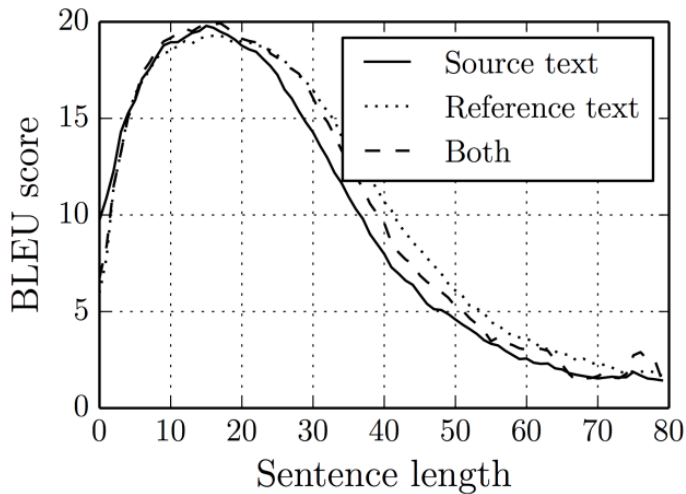


$$p(y|x) = \prod_{t=1}^{T_y} p(y_t|y_1, \dots, y_{t-1}, c),$$

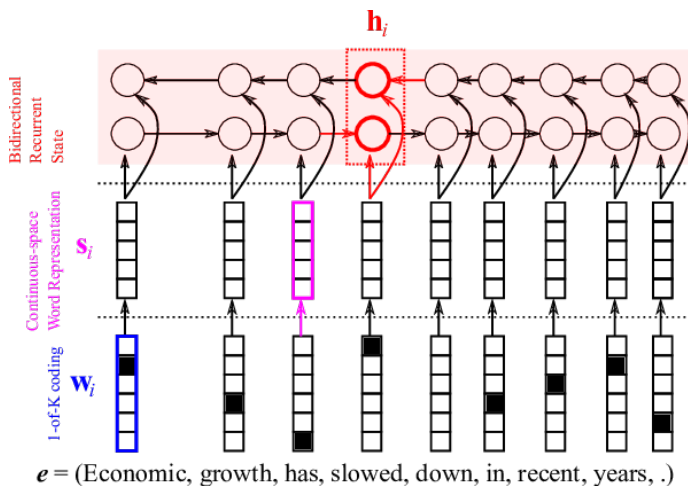
$$p(y_t|y_1, \dots, y_{t-1}, c) = g(y_{t-1}, s_t, c).$$

$$\hat{\Theta} = \arg \max_{\Theta} \frac{1}{N} \sum_{n=1}^N \log(p_{\Theta}(y_n|x_n))$$

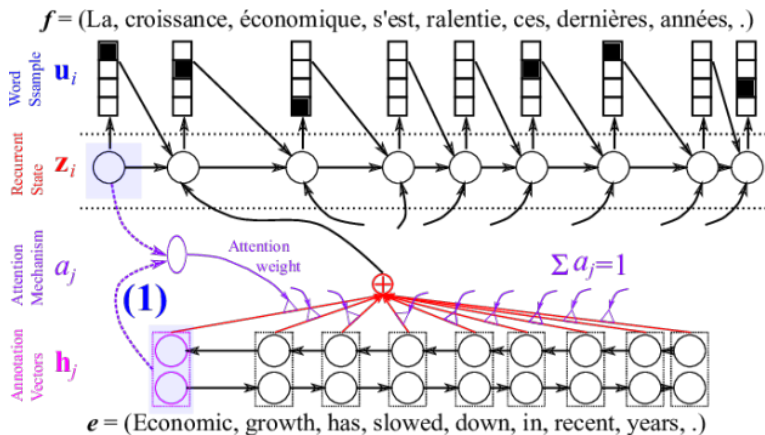
Проблема рассмотренного подхода



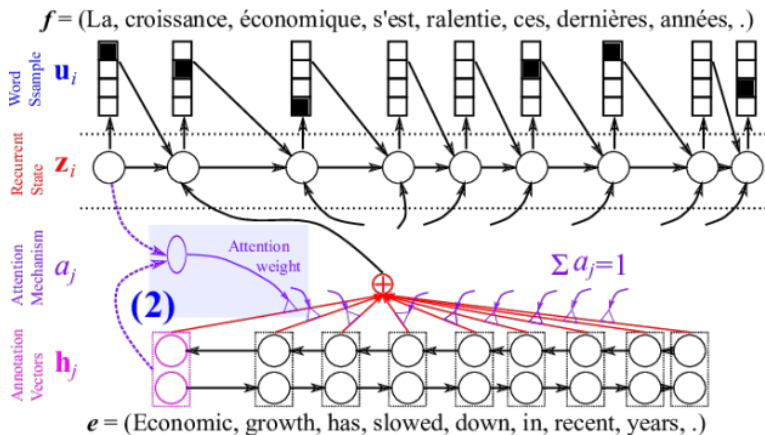
Подход с attention: bidirectional RNN



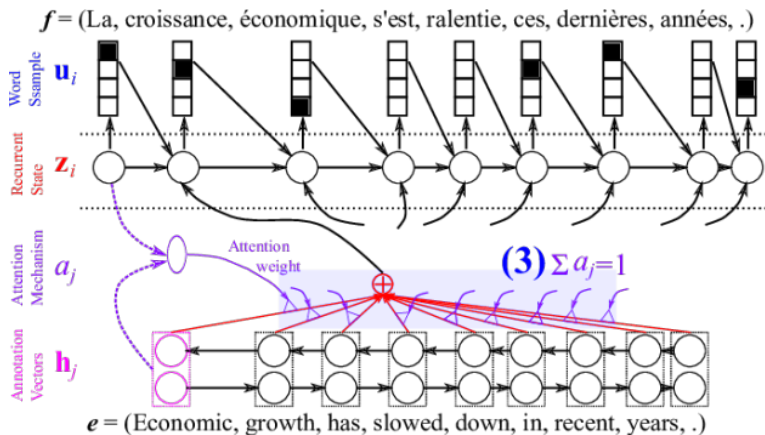
Подход с attention



Подход с attention



Подход с attention



$$p(y|x) = \prod_{t=1}^{T_y} p(y_t|y_1, \dots, y_{t-1}, c),$$

$$p(y_t|y_1, \dots, y_{t-1}, c) = g(y_{t-1}, s_t, c).$$

$$\hat{\Theta} = \arg \max_{\Theta} \frac{1}{N} \sum_{n=1}^N \log(p_{\Theta}(y_n|x_n))$$

$$p(y|x) = \prod_{t=1}^{T_y} p(y_t|y_1, \dots, y_{t-1}, \mathbf{c}_t),$$

$$p(y_t|y_1, \dots, y_{t-1}, \mathbf{c}_t) = g(y_{t-1}, s_t, \mathbf{c}_t).$$

$$\hat{\Theta} = \arg \max_{\Theta} \frac{1}{N} \sum_{n=1}^N \log(p_{\Theta}(y_n|x_n))$$

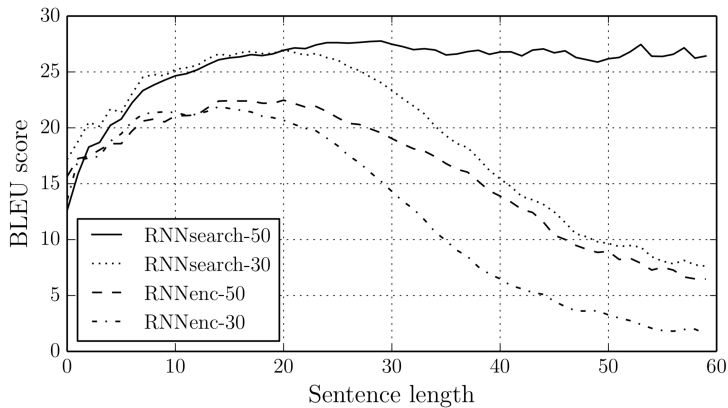
Economic growth has slowed down in recent years .

Das Wirtschaftswachstum hat sich in den letzten Jahren verlangsamt .

Economic growth has slowed down in recent years .

La croissance économique s' est ralentie ces dernières années .

Сравнение моделей



- Papineni, Roukos, Ward, Zhu. BLEU: a method for automatic evaluation of machine translation
- Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation
- Bahdanau, Dzmitry, Cho, Kyunghyun, and Bengio, Yoshua. Neural machine translation by jointly learning to align and translate
- Kyunghyun Cho. Нейронный машинный перевод с применением GPU. Вводный курс.