CS224N Final Project

Neural Network Joint Language Model: An Investigation and An Extension with Global Source Context

Yuhao Zhang, Charles Ruizhongtai Qi
(Advised by Thang Luong)

Objective

- Implement and evaluate a state-of-the-art joint language model.
- Extend the model to make use of global source context information.

Example

Source: <s> 他 每天 早上 走路 去 学校 </s>

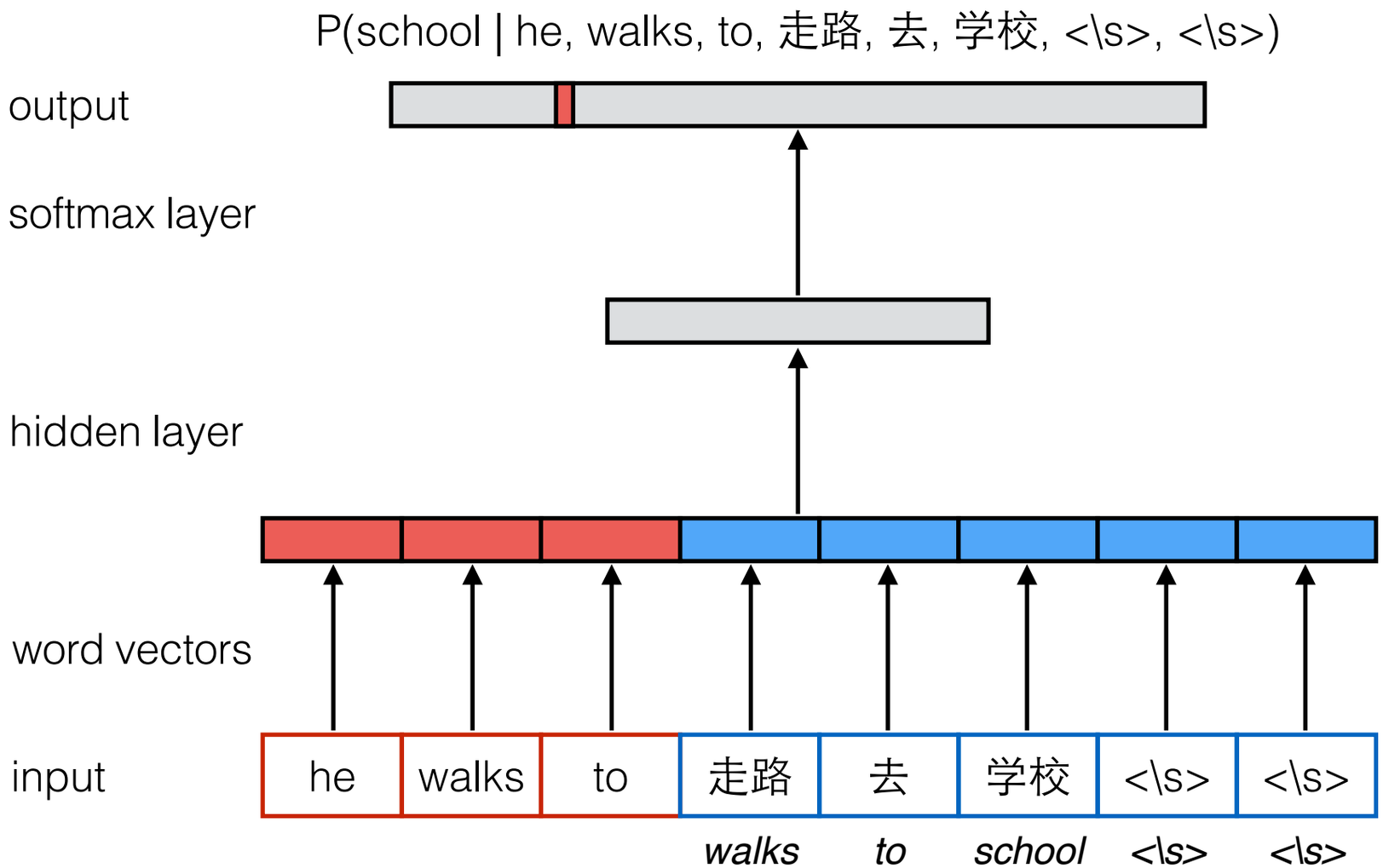
he every morning walk to school

Target: <s> he walks to school every morning </s>

P(morning I to, school, every, 他, 每天, 早上, 走路, 去)

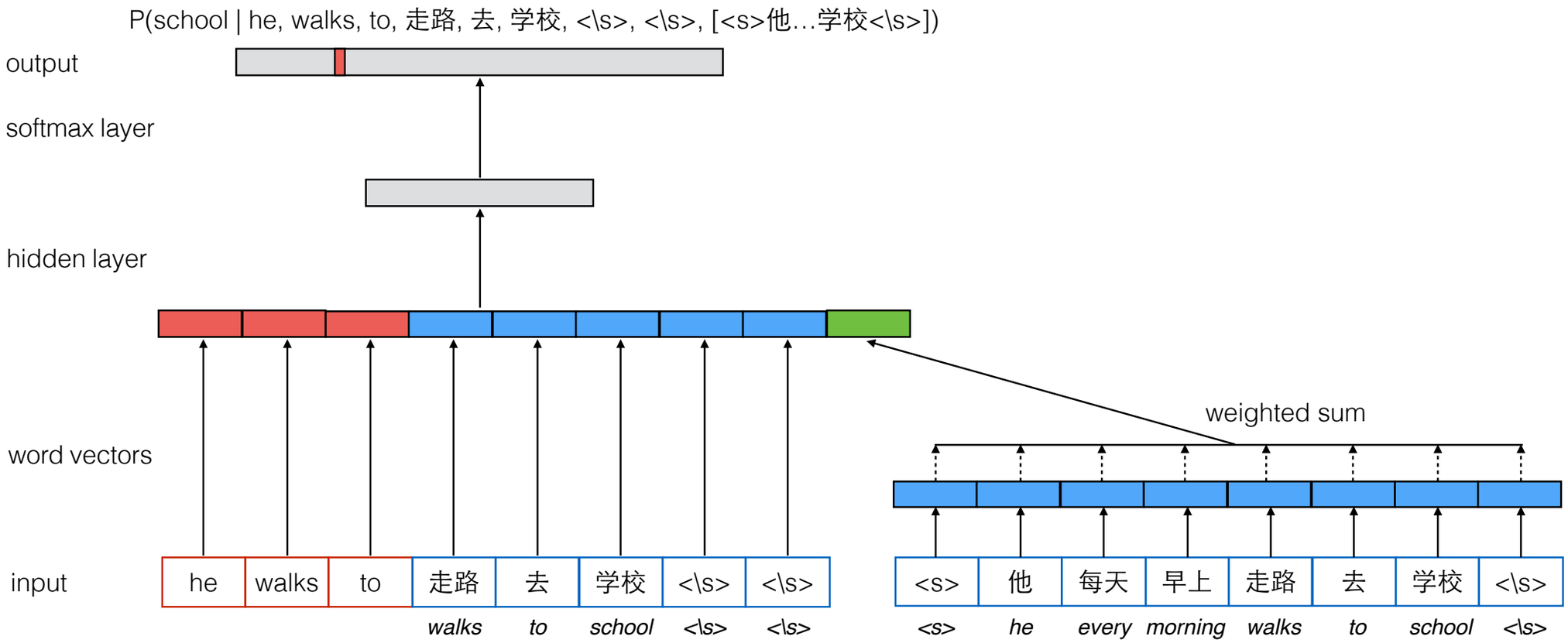
Model

Neural Network Joint Model (NNJM)



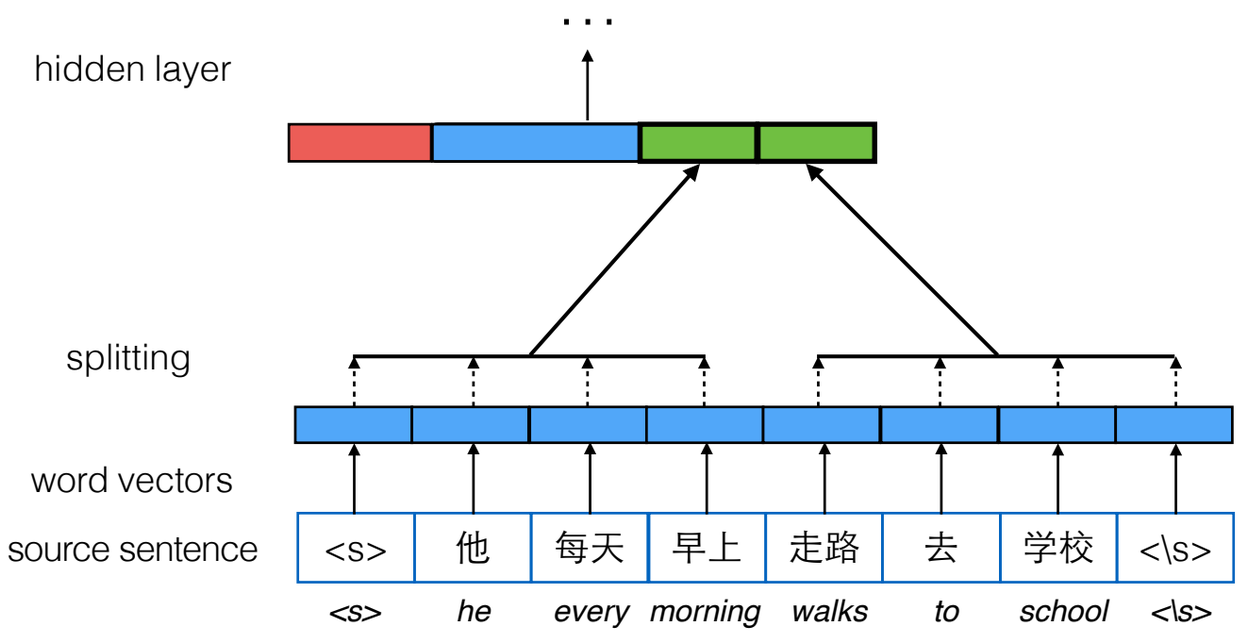
Model (cont.)

Our extension to the basic model:
NNJM with global source context (NNJM-Global)

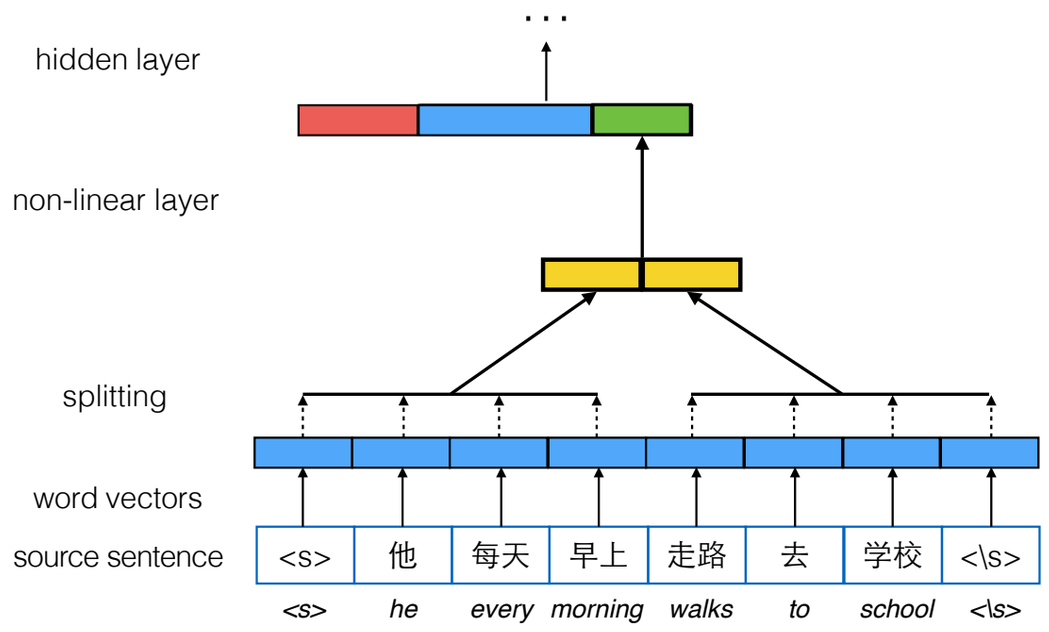


Model (cont.)

Variants of NNJM-Global:



Splitting the source context



Add non-linear layer

Experiment

- **Evaluation Metric**
  - Perplexity on test set

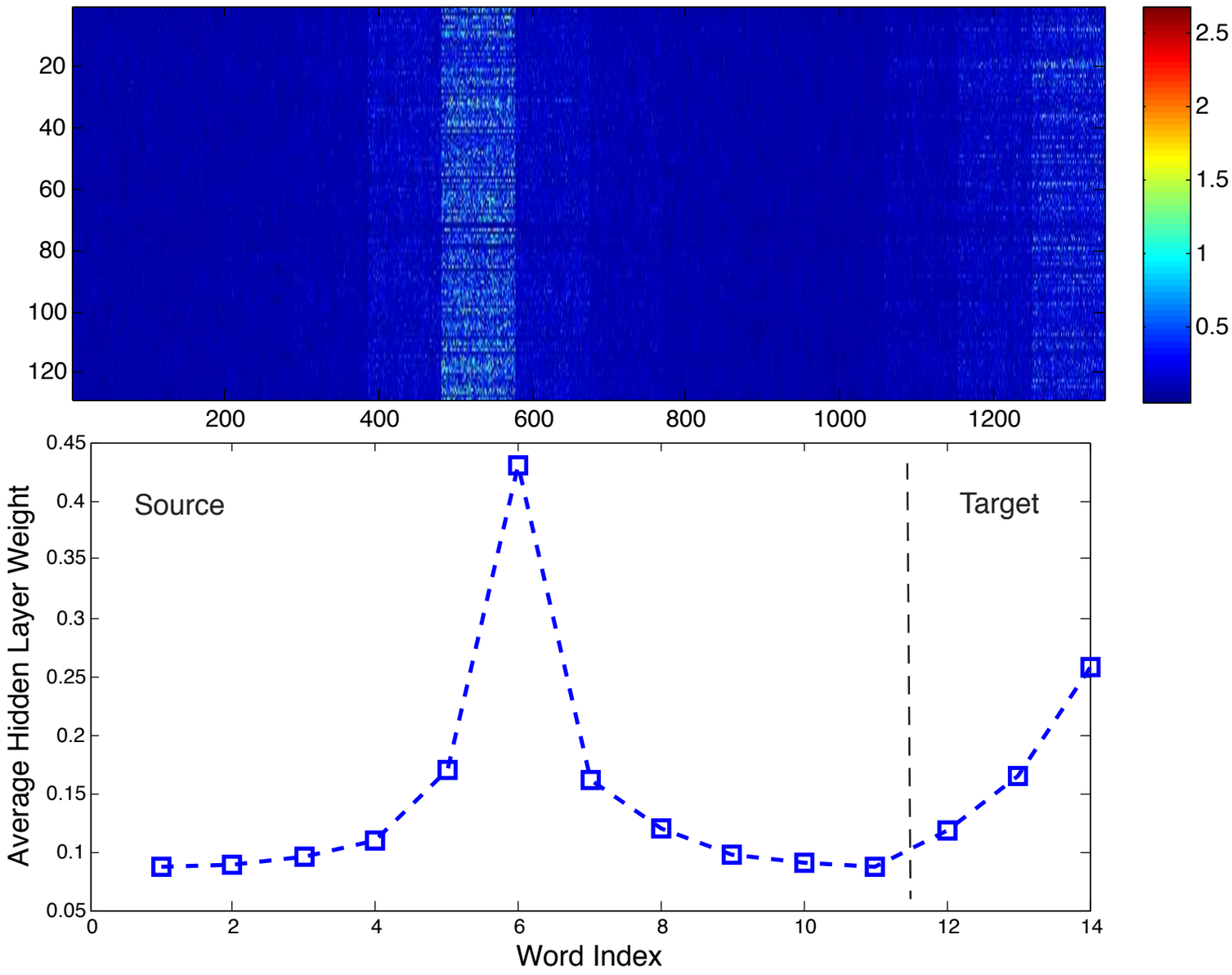
$$PP(W) = p(w_1, w_2, ..., w_N)^{\frac{-1}{N}}$$

- **Data**
  - French-English parallel text (25K pairs) with alignments from European Parallel Corpus.
- **Training**
  - Mini-batch gradient descent
  - Theano, GPU

Results

Model	Perplexity
NNJM	9.51
NNJM-Global	9.45
NNJM-Global+Split	9.38
NNJM-Global+Split+NonLinear	**9.33**

Visualization



Average value of hidden layer weights for each word

Thank You!