A Systematic Comparison of Alignment and Attention Models

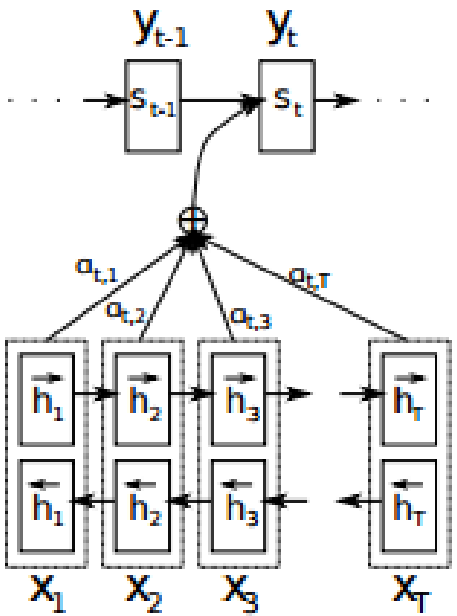
Anahita Mansouri Bigvand
Senior Supervisor: Dr. Anoop Sarkar

Simon Fraser University

14 November, 2016

Motivation and Objective

- Correspondence between attention model and the alignment model
- HMM-based alignment mode (Vogel et al., 1996)
  - a_j depends on a_{j-1}
- Idea: incorporate Markov conditioning to the attention model (Bahdanau et al., 2014)
  - c_t depends on c_{t-1}



$$c_t = \sum_{j=1}^{T_x} \alpha_{tj} h_j$$

Methodology

- Compare word alignment quality between three models:
  - NMT coverage (Tu et al., 2016)
    - $C_{t,j} = f(C_{t-1,j}, \alpha_{tj}, h_j, s_{t-1})$
  - Attention model (Bahdanau et al, 2014)
  - HMM-based alignment model (Vogel et al., 1996)

Experiments

- German-English data
- Train:100K sentences
- Test: 150 sentences
- Evaluation using the AER measure