Dynamic Beam Search in Sequence-to-Sequence Models

Yu-Hsiang Lin, Shuxin Lin, Hai Pham
Language Technologies Institute, Carnegie Mellon University
{yuhsianl, shuxinl, htpham}@andrew.cmu.edu

Carnegie Mellon University Language Technologies Institute

1. Introduction

Sequence-to-sequence (seq2seq) models are typically trained by the maximum likelihood estimation with the "teacher forcing" technique, and beam search is used to accelerate the test-time decoding. While beam search has an advantage over greedy search for accuracy, it is time consuming as the beam size increases. We aim to address this problem and propose two methods of adapting the beam size for decoding: (1) **deterministic (heuristic) agent** with a fixed threshold and (2) **dynamic asynchronous actor-critic agent**.

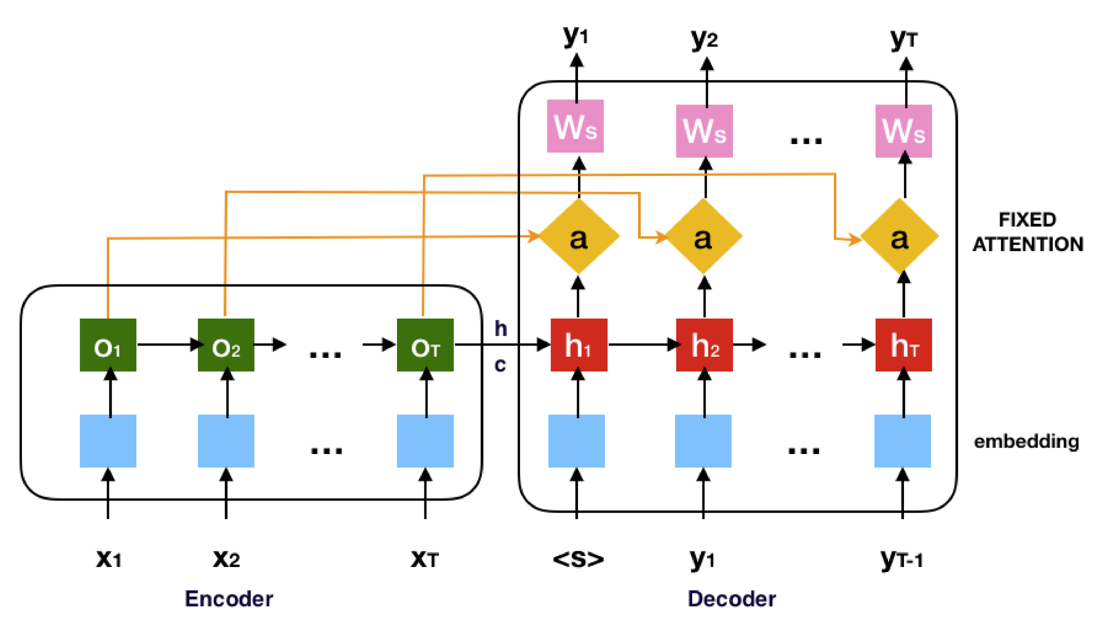
2. Seq2Seq with Fixed Attention

- Seq2seq model are both LSTM RNNs, except that the encoder is *bi-directional*
- Input augmentation with GLoVe word embedding
- We use the "fixed" attention, by which we attend at h_t^{enc} when decoding for h_t^{dec}.

$$p_{y_t} = softmax(W_s \cdot attn(h_t^{dec}, h^{enc}) + b_s)$$

where attention is:

$$attn(h_t^{dec}, h^{enc}) = h_t^{enc}$$



Fixed attention architecture

3. Search with Fixed Beam Size

- Beam search is the traditional way to handle search in decode time
- Trade processing time with the accuracy compared to greedy search; it always process $|B| * |V_{label}|$ tokens at each time step
- Time and resource hungry if the label size is big

8. Future work

- Refine reinforcement learning techniques
- Pre-train the Actor-Critic model with Deterministic Agent (using Imitation Learning)

9. References

[1] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International Conference on Machine Learning*, pages 1928–1937, 2016.

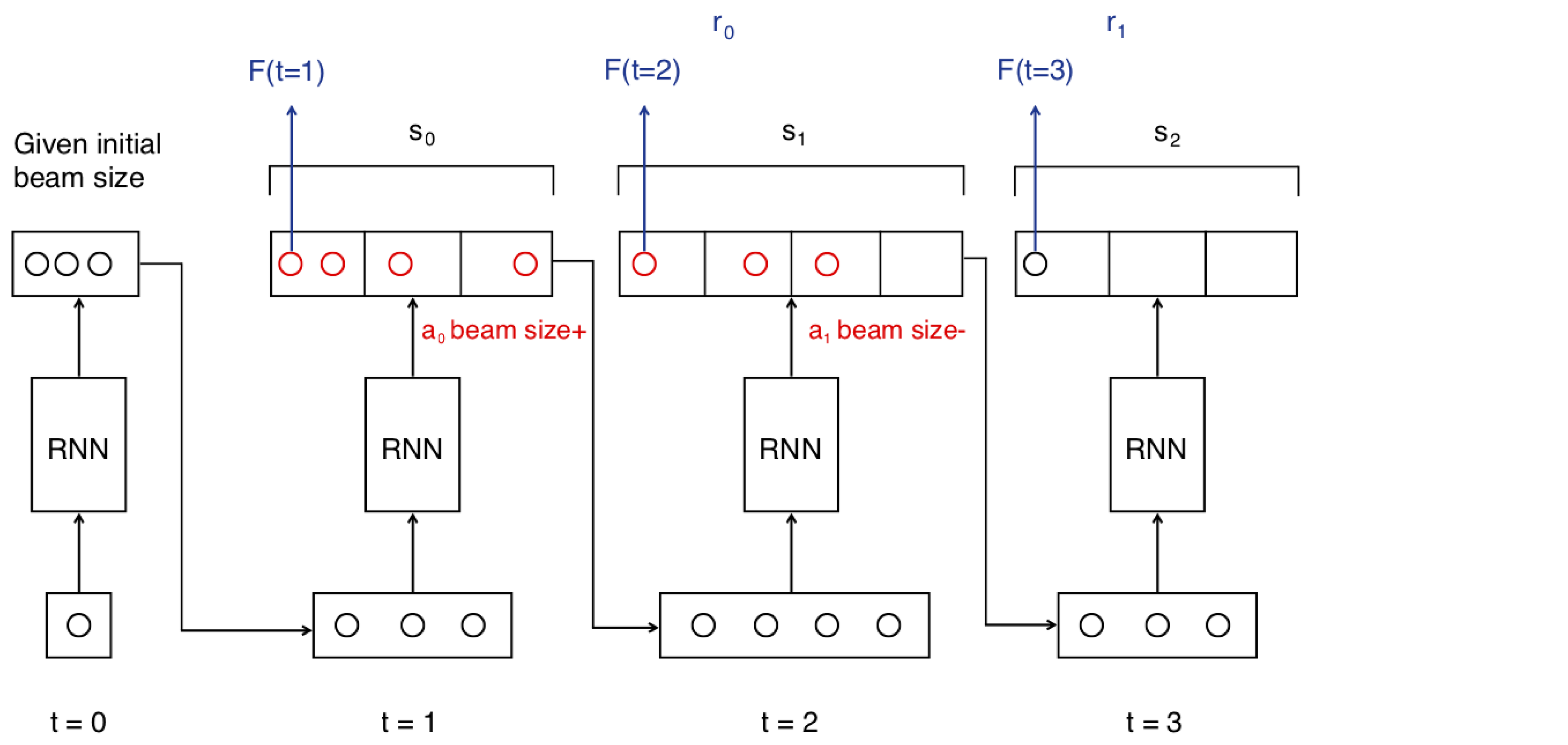
[2] Kartik Goyal, Graham Neubig, Chris Dyer, and Taylor Berg-Kirkpatrick. A continuous relaxation of beam search for end-to-end training of neural sequence models. *arXiv preprint arXiv:1708.00111*, 2017.

4. Heuristic Pruning and Growing for Dynamic Beam Search

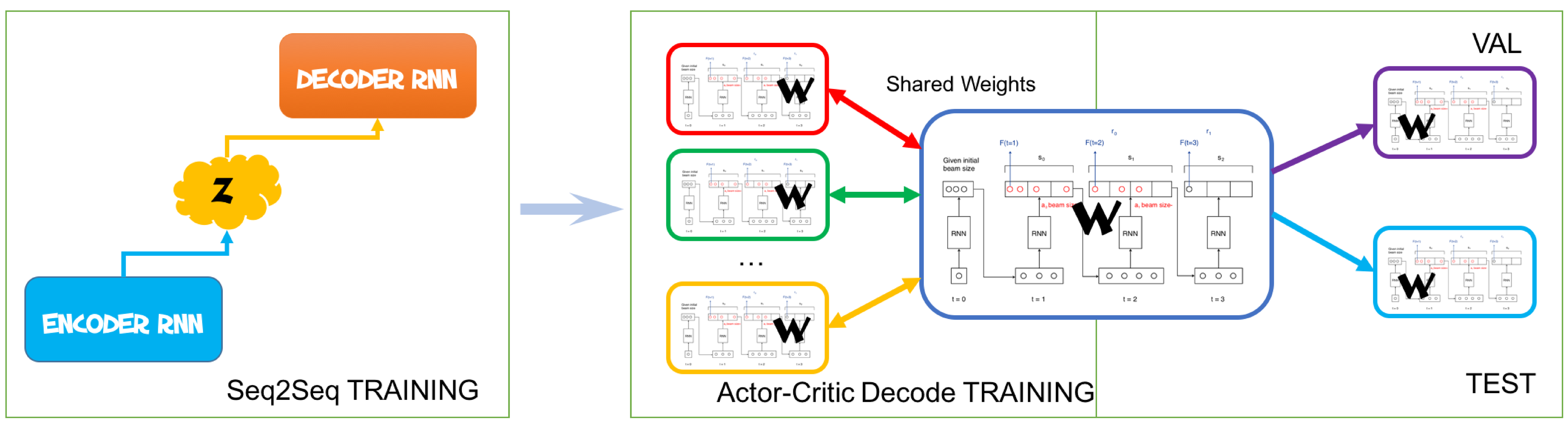
- Current beam size B at step t
- $B \to B-1$ if $\sum_t \log P(y_t^B) - \sum_t \log P(y_t^0) \le \log r_{low}^{sen}$ or $\log P(y_t^B) - \log P(y_t^0) \le \log r_{low}^{word}$
- $B \to B+1$ if $\sum_t \log P(y_t^{B+1}) - \sum_t \log P(y_t^0) > \log r_{low}^{sen}$ or $\log P(y_t^{B+1}) - \log P(y_t^0) > \log r_{low}^{word}$

5. Reinforcement Learning for Dynamic Beam Search

- Design the states, actions, and rewards for dynamic beam decoding



6. Training Decoder with Asynchronous Actor-Critic



Asynchronous Actor-Critic (A3C) REINFORCE [1] with many shared-weights training agents

7. Experiment Results

- Named Entity Recognition tagging task with CoNLL-2003 dataset (German)
- Compare with baseline Goyal 2017 [2]
- Results of Heuristic Pruning and Growing (adapt from given initial beam size):

	Greedy	Beam 3	Beam 3 Adaptive	Beam 6	Beam 6 Adaptive	Beam 9	Beam 9 Adaptive	Soft Beam
F-score	58.09	57.69	57.71	57.76	57.71	57.76	57.71	
Total beam #	48,571	145,713	92,727	291,426	126,759	437,139	182,785	
Avg. beam #	1	3	1.95	6	3.16	9	4.86	
Time (sec)	22	76	61	132	73	178	92	
Goyal F-score	54.92	51.34						56.38

- Results of Asynchronous Actor-Critic (adapt from given initial beam size):

	RL Beam 3	RL Beam 6	RL Beam 9
F-score	57.66	57.61	57.52
Avg. beam #	1.17	2.00	3.06