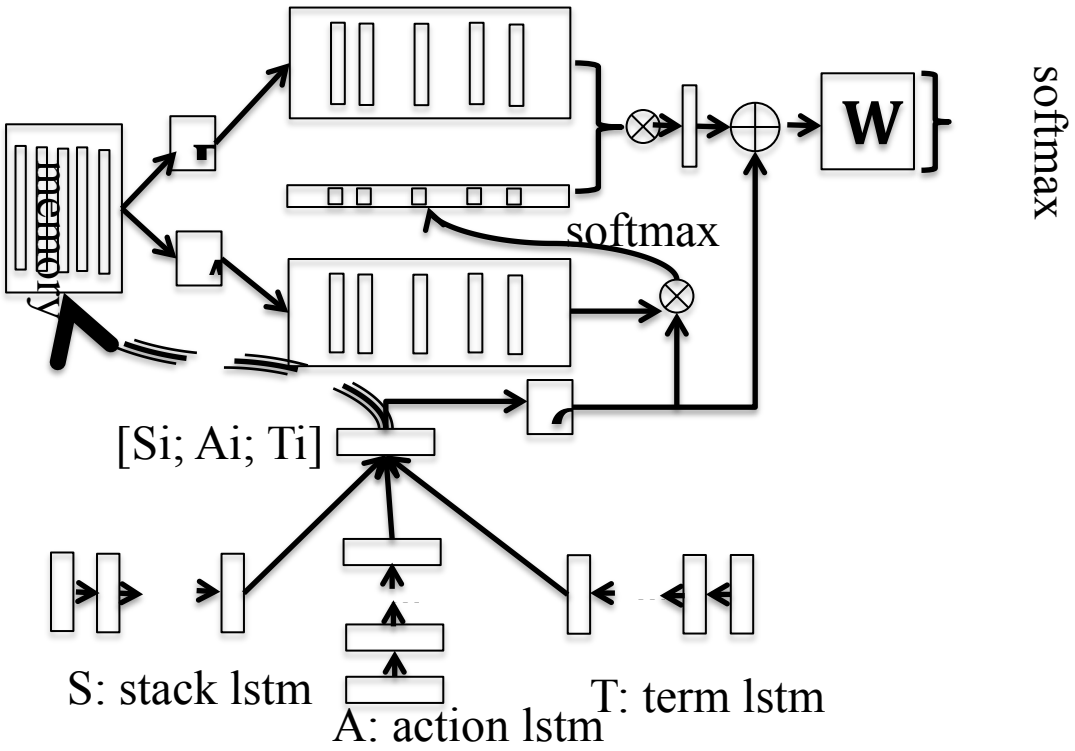
Progress of RNNG with memory network

Yang Feng

RELEASE ON 11-07-2016

The current version

Cascade training of lstm and memory network, which means first train lstm as shown in the paper (Dyer et al. 2016) and then train memory network with the parameters of lstm fixed.



Experiment configuration

Memory size	Memory dimension	nhops	Graph memory	Whether succeed	Reason
100000	768	2	512M	No	Graph is out of memory
10000	768	2	2048M	No	Ditto
5000	768	2	10240M	No	The cluster node is out of memory
5000	768	2	4096M	Under training	---

Also, I optimized my code to reduce memory use, and now the running status is

System	Total memory	Memory for lstm states	Speed
Rnng-gen-mem	20G	5~8G	4.5s/sentence
Rnng-gen	2G	--	3ms/sentence

There are 39832 sentences in the training set and it takes about 50 hours to run a round and it takes at least 10 rounds to converge, so the training time is at least 500 hours. We need to speed up!

To do list:

- To speed up. Change Concatenation lstm states to linear transform. In this way, the memory can reduce from 5000*768 to 5000*256.
- To try different strategies to select memory contents;
- To try different memory size;
- To implement batch training;
- To jointly train memory networks and lstm models.