

Ordered Memory

Yikang Shen, Shawn Tan, Arian Hosseini, Zhouhan Lin, Alessandro Sordoni, Aaron Courville

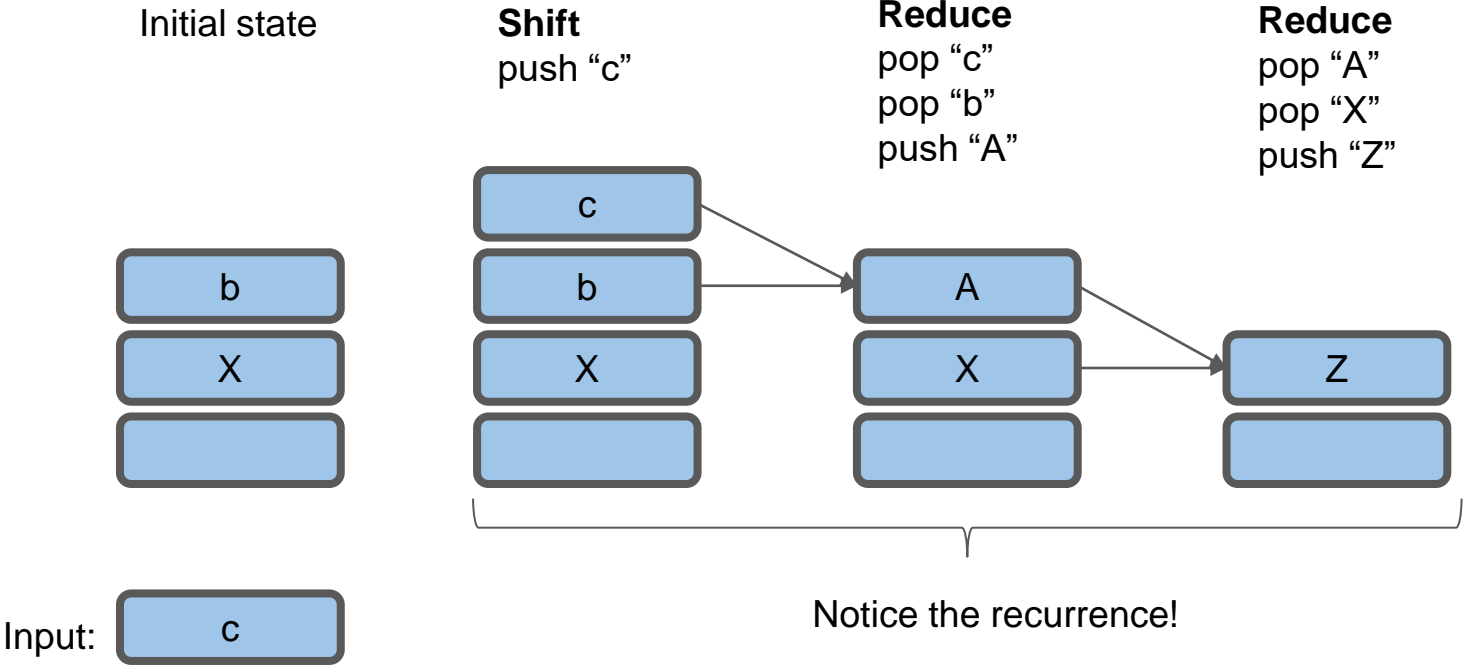
Background

- Recent work has shown that recursive composition of sentences is crucial to *systematic generalisation* (Bowman et al., 2015; Bahdanau et al., 2018)
- There has been work that shows that a tree-like composition structure helps with certain types of tasks (Socher et al., 2015; Kuncoro et al., 2018; Dyer et al., 2016)
- We propose a stack-based method that performs well on syntax-reliant tasks, and that is able to recover the tree structure that is induced.

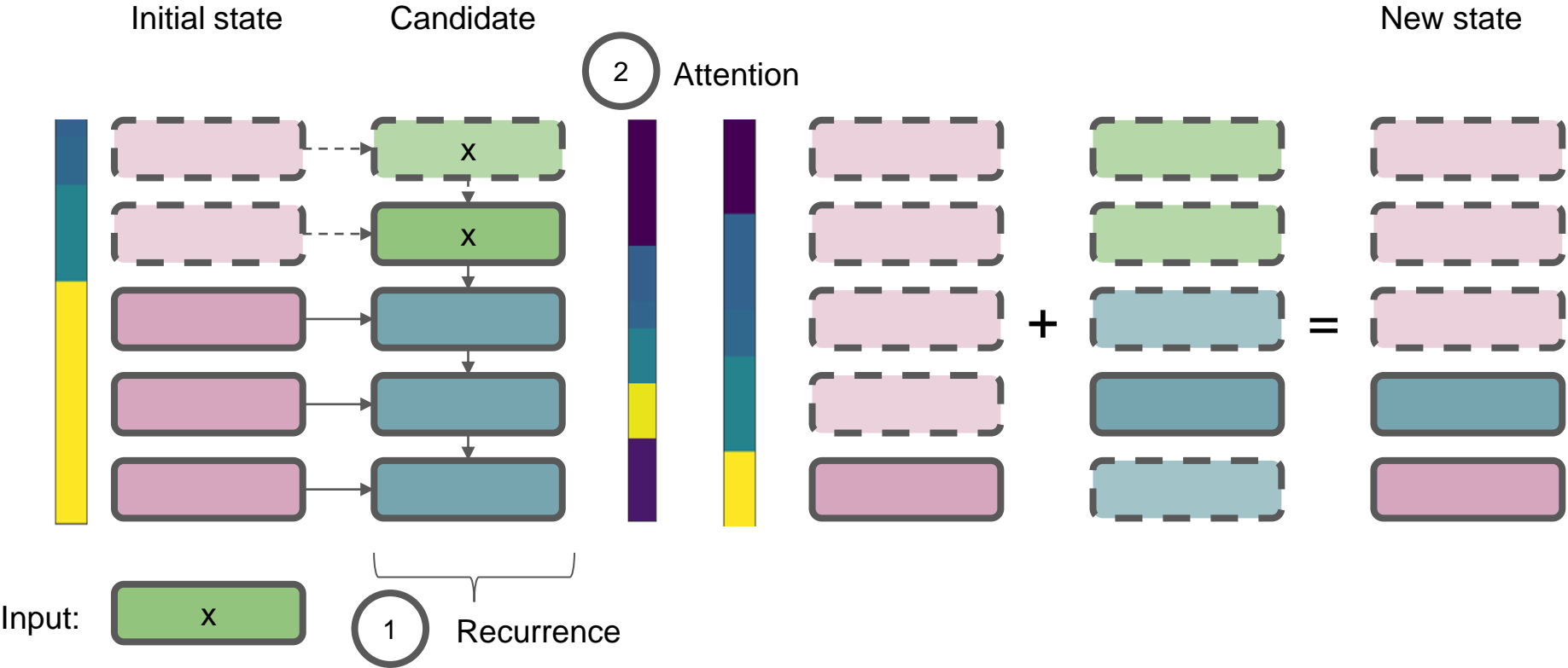
Shift-reduce parser

Generative rules:

$A \rightarrow b\ c$
 $Z \rightarrow X\ A$



Ordered Memory



Logical Inference (Bowman et al. 2015) and Systematic Generalisation Tests

- We perform well compared to sequential sentence representation
- We perform further tests similar to the ones in the SCAN dataset (Loula et al., 2018)
 - We remove training examples that have as constituents certain subtrees.
 - We do this three times, in increasing difficulty
 - Our models' performance on these generalisation test match closely with tree-based models that have the true tree structure as input.

ListOps (Nangia and Bowman, 2018)

- Task was constructed so that if the right composition structure is used, it is easily solvable.
- We achieve a 99.97 % accuracy, while being able to retrieve the induced structure completely.

Stanford Sentiment Treebank (Socher et al. 2013)

- Should be noted that our model is a sentence embedding model
- Performs comparatively with other learned structure models
- Interestingly, the ground-truth tree structure performs worse than the learned models.

Summary

- Stack-augmented model designed to mimic behaviour of shift-reduce parsing.
- Performs well on synthetic tasks:
 - Able to generalise on the logical inference tasks over
 - Operator lengths
 - Unseen sub-trees
 - Able to achieve 99.97% on ListOps, and correctly retrieve the compositional tree structure.
- Performance on real tasks is comparable to existing work.