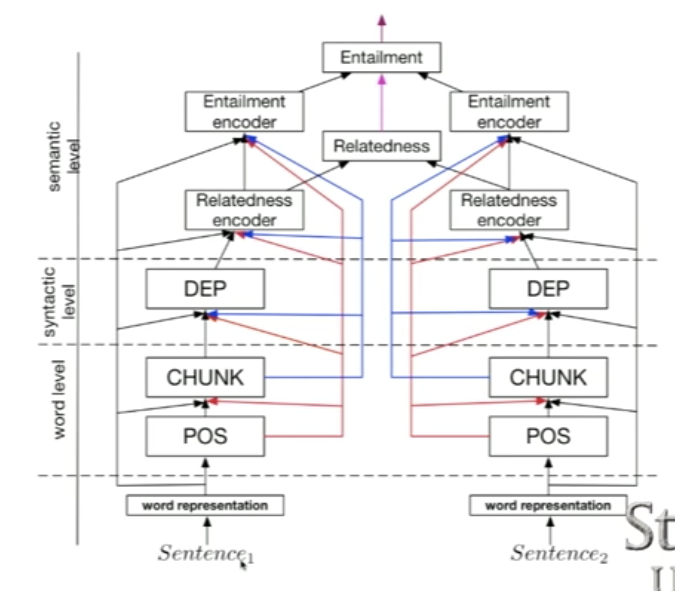
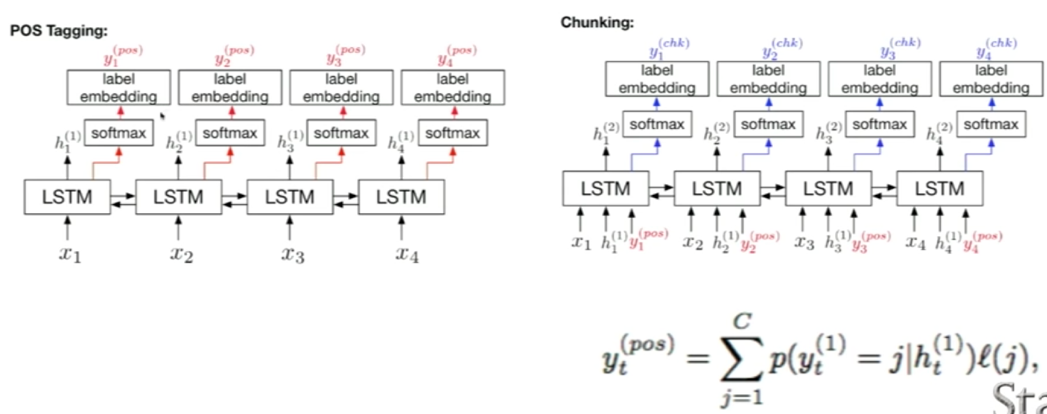
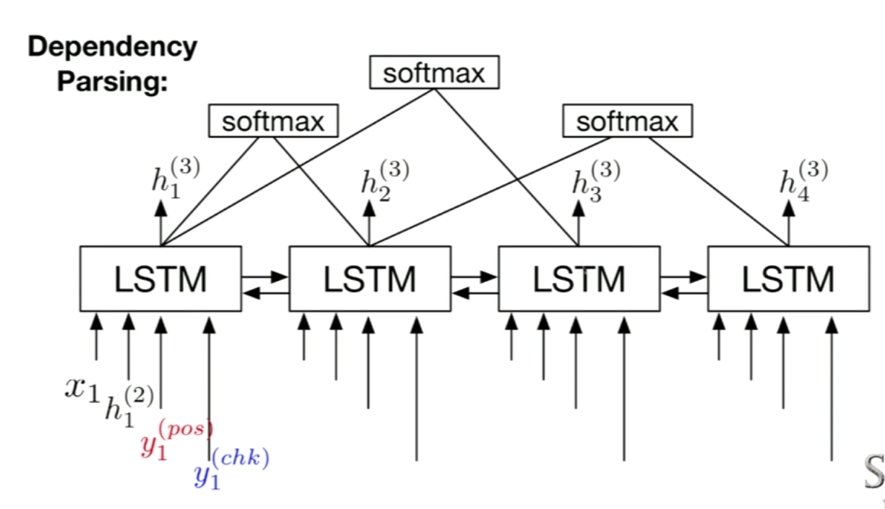
Lecture 18 | Tackling the limits of Deep learning for NLP

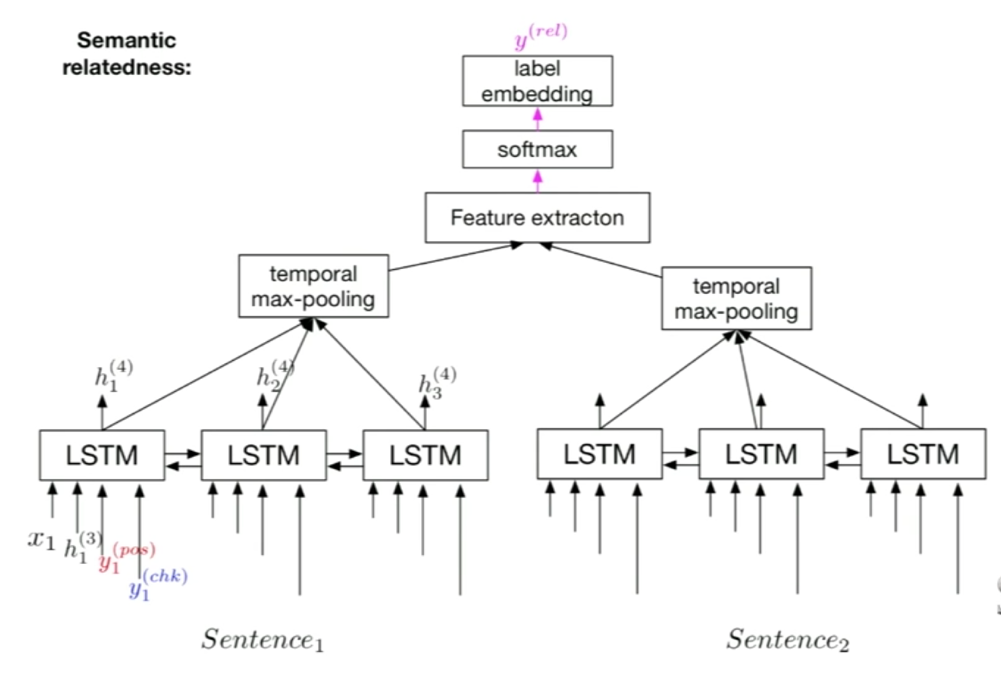
* The limits of single task learning
  + Obstacle 1: For NLP no single model architecture with consistent state of the art results across tasks
    - Attempt Solution: Dynamic Memory Network
  + Obstacle 2: Joint Many-task learning
    - Attempt solution: A joint many-task model: growing a neural network for multiple NLP tasks paper



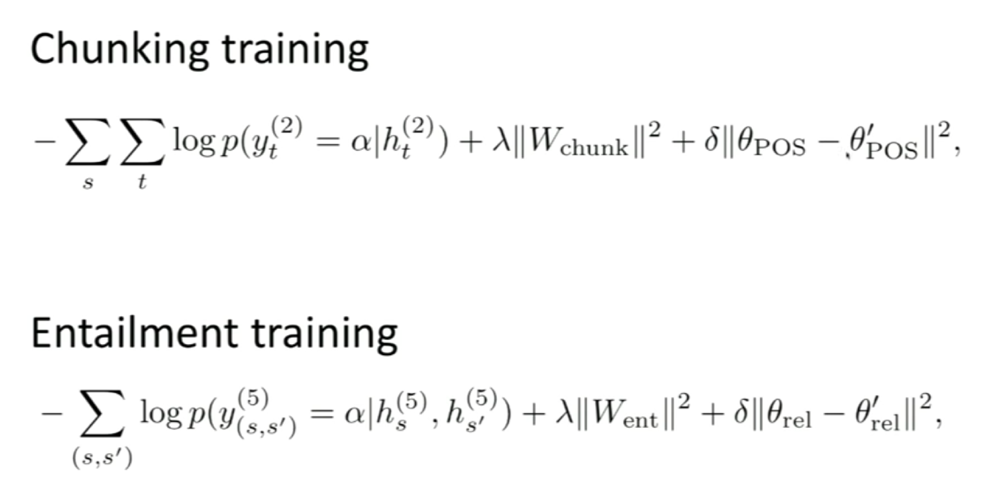
* + - Include character n-grams and short-circuits



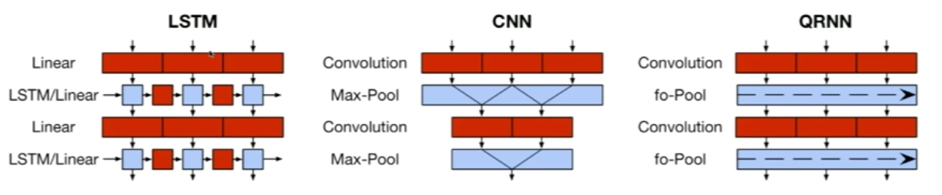




* + - Training details: regularisation



* + Obstacle 3: No zero shot word predictions - Answers can only be predicted if they were seen during training and part of the softmax but it’s natural to learn new words in an active conversation and systems should be able to pick them up
    - Attempt solution: Mixture model of softmax and pointers
  + Obstacle 4: Duplicate word representations – different encodings for encoder (Word2Vec and Glove word vectors) and decoder (softmax classification weights for words), therefore, duplicate parameters/meaning
    - Attempt solution: Tying word vectors and train single weights jointly
  + Obstacle 5: Questions have input independent representations – interdependence needed for a comprehensive QA model
    - Attempt solution: Dynamic Coattention Networks
  + Obstacle 6: RNNs are slow! RNNs are the basic building block for deepNLP
    - Attempt solution: Take the best and parallelisable parts of RNNs and CNNs – Quasi-recurrent neural networks (can be applied to sentiment analysis)



* + **Obstacle 7: Architecture search is slow – manual process that requires a lot of expertise. What if we could use AI to find the right architecture for any problem?**
    - Attempt solution: Neural architecture search with reinforcement learning by Zoph and Le, 2016
* Stanford Question Answering Dataset (SquAD)