A Structural Probe for Finding Syntax in Word Representations

Alireza Moradi

- Understanding what deep neural models learn is important.
- **Probe:** a supervised model for finding info in a representation.
- Learning whether a model's representation encode syntax is of interest.
- **Structural probe:** is a model, tests whether syntax trees are consistently embedded in a linear transformation of a networks word representation space.
- How to test?
    - Tree structure is embedded **if** squared L2 distance between 2 words' vectors corresponds to the number of edges in the parse tree.
    - Edge direction ⇒ depth of the word in the parse tree.
- **Embed a graph:** learn a vector representation of nodes such that geometry (distances and norms) in vector space approximates geometry in graph.
- Why do parse tree distances and depths matter to syntax?
    - distance is the length of the path between 2 nodes. d(u,v) = 1 tells us that nodes *u, v* are neighbors.
    - The node with greater norm (depth) is the child of the other.
    - Distance also explains hierarchical behaviour.
- This model encodes 2 things:
    - Which word is governed by which other word
    - Each word's proximity to all other words in the syntax tree
- **Parse depth:** number of edges between the word w and the root. Represented as a norm.
- **Distance evaluation:** how well predicted distances between all pairs of words reconstruct gold parse trees and correlate with the parse tree's distance metric. Report spearman correlation between true and predicted distances for each word in each sentence. Averaged between all sentences of the same length (distance spearman).

- **Tree depth evaluation:** ability of the model to create order of words based on their depth in the parse tree. Report spearman correlation between true depth ordering and predicted ordering. Averaged between sentences of the same length and then sentences of length 5 to 50 (norm spearman).