

Short description of BOWser

Matthias Lindemann

November 1, 2019

1 Modeling

BOWser models a distribution of trees (without variables as they occur in e.g. SQL) given sentences w by reading off a CFG, from which we generate using a straightforward factorization of the conditional distribution:

$$P(t|w) = \prod_{i=1}^N P(\text{Children}(v_i)|T_i, w_1, \dots, w_n)$$

where T is a sequence of subtrees of t such that T_i contains exactly those nodes of t that have been visited in a pre-order traversal of t up to step i . Similarly, v_i is the node of t that is visited at step i in the pre-order traversal. In practice, the distribution $P(\text{Children}(v_i)|T_i, w_1, \dots, w_n)$ is estimated with a small MLP for each non-terminal of the CFG that doesn't take into account the entire subtree T_i but only certain features:

- the words, *disregarding their order!*
- the depth of v_i
- number of left sisters of v_i
- the identity of the left and right sisters of v_i (disregarding the order among them)
- the leaves in T_i , without order
- parent and grandparent of v_i

That is, this model is in spirit very similar to many modern deep semantic parsers but does this in the style of log-linear models.

2 Inference

Inference is performed greedily in a pre-order traversal, following the factorization. A cutoff of depth 50 is used to ensure termination.

3 Conclusion

It is surprising that this works at all beyond tiny sentences/questions. Ignoring the syntax should have a noticeable price in accuracy and perhaps corpora with more complex syntactical structures are required to train and evaluate semantic parsers.