



Despite an identical intermediate representation in a recursive/tree-based model of grammar, baseline Encoder-Decoder Transformers (e.g. Wu et al baseline 2 layer) trained from scratch on generating logical forms for a corpus including the first type of sentence (a) but not the second (b), cannot reliably generalize to doing so for the second type of sentence (suggests cannot understand the type (b) sentences after learning type (a)).