Motivation

Dependency structures are often used for relation extraction as they are able to capture long-range relations that are only implicit in the surface form alone. However, dependency parsing accuracy is relatively low in the medical domain, which may downstream the relation extraction pipeline.

Figure (a) shows the 1-best dependency tree, (b) is the manually labeled tree.

- The 1-best dependency tree obtained by an off-the-shelf parser contains an error. Specifically, the entity phrase *phenylalanine hydroxylase* is broken since *hydroxylase* is mistakenly considered as the main verb.
- Unlike previous research efforts that rely on dependency parsers trained on newswire text, our model is able to generate task-specific dependency structures for capturing non-local interactions.

Model

- Our model consists of two modules: Forest Inducer and Forest Encoder. The forest inducer has two sub-modules, where the first one computes $N$ attention matrices and the second sub-module takes the $N$ attention matrices as inputs to obtain $N$ dependency forests. Then the forest encoder uses graph neural networks to encode the induced forests.

Latent Forests

Inspired by Liu and Lapata (2018), we use a variant of Kirchhoff's Matrix-Tree Theorem (MTT; Koo et al., 2007; Smith and Smith, 2007) to induce the latent structure of an input sentence. For the graph $G$, MTT takes the edge scores and root scores as inputs then generates a latent forest by computing the marginal probabilities for each edge:

$$P(z_{ij}^k = 1) = \sum_{y \in \mathbb{T}^*} P(y; \mathbf{\theta}^k)$$

Induced latent structure as shown below can be viewed as multiple full dependency forests, which efficiently represent all possible dependency trees within a compact and dense structure.

Experiments

- We evaluate our proposed model with four medical relation extraction datasets on two tasks, including cross-sentence n-ary relation extraction (Peng et al., 2017) and sentence-level relation extraction, CPR (Krallinger et al., 2017) and PGR (Sousa et al., 2019).

References

- Koo et al., Structured prediction models via the matrix-tree theorem. EMNLP07.
- Smith et al., Probabilistic models of nonprojective dependency trees. EMNLP07.