Task: Frame-semantic parsing
The goal is to parse sentences into FrameNet-style semantic graphs (Baker et. al., 1998).

Segmental Recurrent Neural Networks (Kong et. al., 2016)
- Variant of a semi-Markov conditional random field (Sarawagi and Kohen, 2004)
- Span representations are computed using bidirectional RNNs.
- Provide a generalization of BIO tagging schemes.
- Directly model an entire variable-length segment (rather than fixed-length label n-grams).
- Exact inference takes O(nd), n being the length of sentence, d maximum length of spans, and ℓ the number of labels.

Recall-oriented Softmax-Margin Segmental RNNs
A modified logloss objective that encourages recall over precision, by applying a cost function which penalizes false negatives by a factor α is used:

\[
\text{loss}(x, s^*) = -\log \frac{\exp(s^*, x)}{\sum_s \exp \left\{ (s, x) + \text{cost}(s, s^*) \right\}},
\]

\[
\text{cost}(s, s^*) = \alpha \text{FN}(s, s^*) + \text{FP}(s, s^*),
\]

This objective results in a boost in F1, primarily due to increase in recall.

Learning with a syntactic scaffold
A binary logistic regression loss is used to predict if a text span could be a constituent.

\[
\text{loss}_{\text{scaffold}}(i, j, r^*, x) = -\log \frac{\exp \psi(i, j, r^*, x)}{\sum_{r = (0, 1)} \exp(\psi(i, j, r, x))},
\]

The joint multi-task loss for a single sentence is:

\[
\text{loss}(x, s^*) + \delta \sum_{i < j < k < D} \text{loss}_{\text{scaffold}}(i, j, r^*, x),
\]

Argument Identification
Performance of argument identification only, using gold frames, on the FrameNet 1.5 test set.

Incorporating Syntax I: Pipelining Syntactic Features
Constituency Features
- dep_phrase
- lca_type
- constit_path_lstm

Dependency Features
- head_word
- head_label
- out_#heads
- dep_path_lstm

Incorporating Syntax II: Syntactic Scaffolding
- Frame-semantic arguments are also syntactic constituents.
- Multi-task learning setup: simultaneously learn to predict syntactic constituents and frame-semantic arguments.
- Can exploit constituent span annotations from Penn Treebank.
- Bidirectional RNN parameters are shared between tasks.
- Scaffold is only needed at train time; usual test setup is followed.