Neural networks

Natural language processing - recursive network
Topics: word phrase representations

• We’ve seen how to learn representations for single words

• How could we learn representations for phrases of arbitrary length?
  ‣ can we model relationships between words and multiword expressions
    - ex.: “consider” $\approx$ “take into account”
  ‣ can we extract a representation of full sentences that preserves some of its semantic meaning
    - ex.: “word representations were learned from a corpus” $\approx$ “we trained word representations on a text data set”
**Topics:** recursive neural network (RNN)

- Idea: recursively merge pairs of word/phrase representations

- We need 2 things
  - a model that merges pairs of representations
  - a model that determines the tree structure

---

**Abstract**

Recursive structure is commonly found in the inputs of different modalities such as natural scene images or natural language sentences. Discovering this recursive structure helps us to not only identify the units that an image or sentence contains but also how they interact to form a whole. We introduce a max-margin structure prediction architecture based on recursive neural networks that can successfully recover such structure both in complex scene images as well as sentences. The same algorithm can be used both to provide a competitive syntactic parser for natural language sentences from the Penn Treebank and to outperform alternative approaches for semantic scene segmentation, annotation and classification. For segmentation and annotation our algorithm obtains a new level of state-of-the-art performance on the Stanford background dataset (78.1%). The features from the image parse tree outperform Gist descriptors for scene classification by 4%.

---

**1. Introduction**

Recursive structure is commonly found in different modalities, as shown in Fig. 1. The syntactic rules of natural language are known to be recursive, with noun phrases containing relative clauses that themselves contain noun phrases, e.g., . . . the church which has nice windows . . . Similarly, one finds nested hierarchical structuring in scene images that capture both part-of and proximity relationships. For instance, cars are often on top of street regions. A large car region appearing in . . .

---

**Figure 1.** Illustration of our recursive neural network architecture which parses images and natural language sentences. Segment features and word indices (orange) are first mapped into semantic feature space (blue) and then recursively merged by the same neural network until they represent the entire image or sentence. Both mappings and mergings are learned.

---

*Socher, Lin, Ng and Manning, 2011*