Learning Distributed Representations for Structured Output Prediction

Or: Are structured outputs discrete objects?

Vivek Srikumar

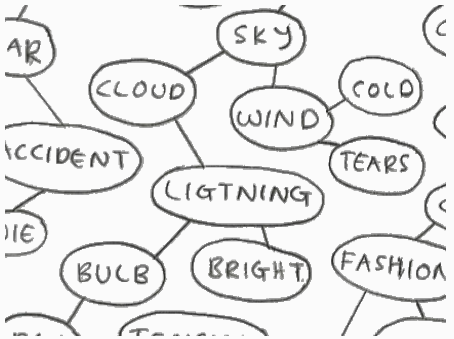
THE UNIVERSITY OF UTAH

Christopher Manning

The problem

Distributed representations for ***inputs***: A good idea

Inputs are not discrete units of meaning



Clear similarities between outputs (labeled sequences, trees, etc) exist

Yet, standard algorithms treat outputs as discrete objects!

Labels and sub-structures that compose **structured output** are semantically rich

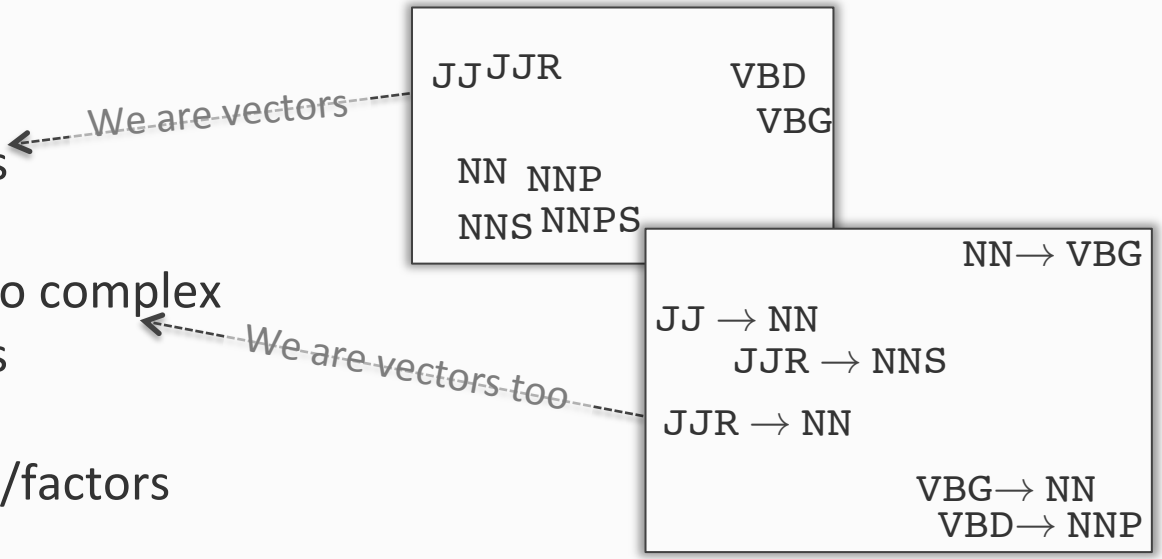
DT – **JJ – NN** – VB – DT – **JJ – NNP**
DT – **JJS – NNS** – VBD – DT – **JJR – NNPS**

All `Adjective` → `Noun` transitions

Why treat structures as discrete when the components aren't?

*DI*stributed *ST*ructured *O*utput: DISTRO

1. All labels are real vectors
2. Compose atomic labels to complex parts via tensor products
3. Define features for parts/factors using these tensors



CRF/structured SVM: special cases with one-hot label encoding

Learn label vectors and model parameters jointly

Use knowledge-driven inference algorithms with distributed representations!