CS 378 Lecture 14: Shift-reduce parsing, review

Announcements
- NO LECTURE THURSDAY
- A3 due
- Midterm Qs via email or private open-book ed STEM

Recap Dependencies

- Verbs are heads of sentences
- Verbs have nouns/prepositions as children usually
- Nouns are modified by other stuff in the NP

The cat jumped and landed on the desk
Today Shift-reduce dep parsers

Move through a sent word-by-word, make decisions greedily

Input: sentence
Output: dep parse

Two data structures
Stack: partial parse trees
Buffer: remaining words of the sent
I ate some spaghetti bolognese

Initial state:
stack: [Root]
Buffer: [I ate some spa bo]

Operations: (arc-standard)
① Shift: first word from buf → end of stack
② Left-arc: take top two from the stack, add arc, return to stack
③ Right-arc

What is the right sequence of ops to build this tree?
[Root] [I ate some spa bo]
Shift
[Root I] [ate ...]
Shift
[Root I ate] [some ...]
Left-arc

[Root ate] [some ...]

I
Shift, Shift

[Root ate some spa] [bo]

Left-arc

[Root ate spa] [bo]

I

I

Left-arc
PROBLEM: can't attach bolognese!

Shift

[ROOT ate spa bo]  []
I \  \  I some

Right-arc

[ROOT ate spa]  []
I some bo ROOT
ate
spa
some bo

R-A, R-A
Arc-standard shift-reduce can build any projective tree

\[ \rightarrow \]

\[ a \quad b \quad c \quad d \quad \times \quad \text{can't build} \]

nonprojective

Building S-R parsers

Parser is a classifier

Maps from \((\text{stack, buffer})\) \rightarrow \text{action} \(\langle S, LA, RA \rangle\)

Multiclass classification

Features \(f(S, B)\)
OPTIONAL

Feats are really complex

S [ Root ate spa ] [ bo ]
  ↓  ↓
  I  some

Shift

if instead

[ on Monday ]

Right-arc

feats conjoin info from buf + stack

Ind [ top of stack = N & first in buffer = bolognese ]

× 100s of feature templates like this

⇒ 100K features
Midterm Review

\[ p_1, p_2 = [0, 0] \]

\[
\frac{e^{p_1}}{e^{p_1} + e^{p_2}} = \frac{1}{1 + 1} = \frac{1}{2}
\]

cats and dogs

\[
\begin{bmatrix}
0.3 \\
0.6
\end{bmatrix}
+ 
\begin{bmatrix}
-0.4 \\
0.1
\end{bmatrix}
+ 
\begin{bmatrix}
0.1 \\
-0.7
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
0
\end{bmatrix}
\]

dogs and dogs and dogs

\[
\begin{bmatrix}
0.1 \\
-0.7
\end{bmatrix}
\times 3
+ 
\begin{bmatrix}
-0.4 \\
0.1
\end{bmatrix}
\times 2
= 
\begin{bmatrix}
-0.5 \\
-1.9
\end{bmatrix}
\]

\[ W = \begin{bmatrix} a \\ b \end{bmatrix} \]

\[
a = -100 \\
b = -100
\]

\[ Ws_2 \text{ is large} \]

Ans: E

\[
\text{Softmax } \rightarrow 0.5 \ 0.5
\]
NP: -2 -1 from role

dT: O N: -1

the ring

ROOT - NP role

ROOT log prob: -3

1/2 1

prob: 1/8

NP

1/2 /

DT N 1/2

the ring

7. NP \rightarrow NPP

3. NPP \rightarrow \text{rings}