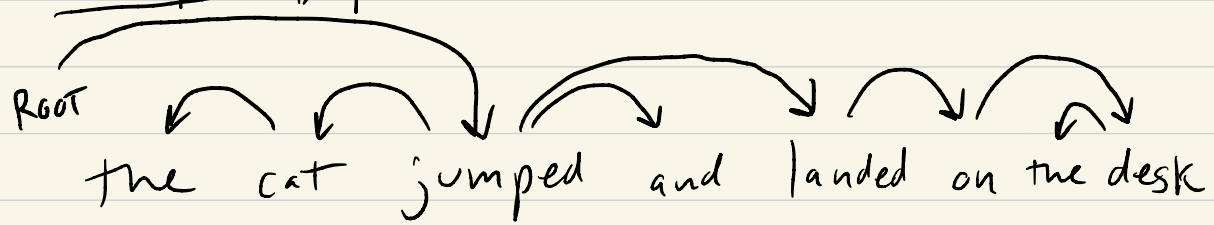
CS 378 Lecture 14: Shift-reduce parsing, review

Announcements

NO LECTURE THURSDAY

- A3 due
- Midterm
  - Qs via email or private ed STEM
  - open-book

Recap Dependencies



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

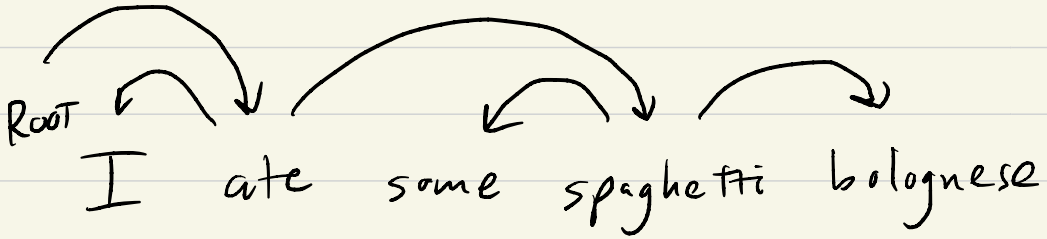
Today Shift-reduce dep parsers

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



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

(3) Right-arc

take top two from the stack, add arc, return to stack

What is the right sequence of ops to build this tree?

Stack Buffer

[ROOT] [I ate some spa bo]

Shift

[ROOT I] [ate ...]

Shift

[ROOT I ate] [some ...]

Left-arc

[① ROOT ② ate] [some ...]

↓

I

Shift, Shift

[ROOT ate some spa] [bo]

↓

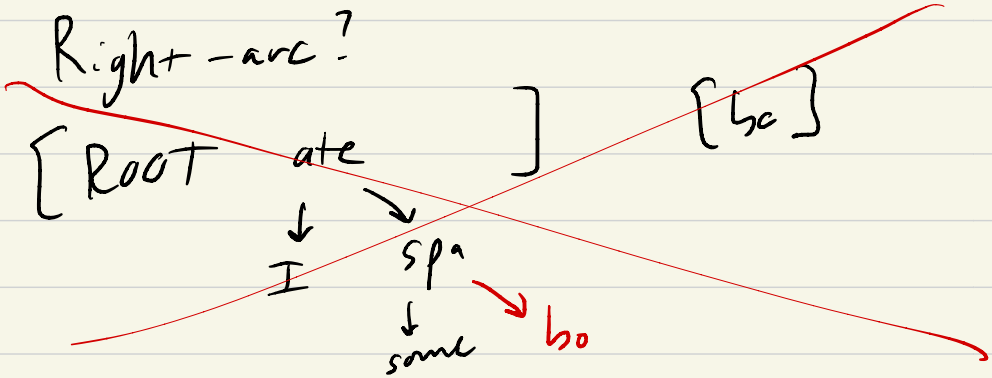
I

Left-arc

[ROOT ate spa] [bo]

↓ ↓

I some



PROBLEM: can't attach bolognese!

Shift

[ROOT ate spa bo] [ ]

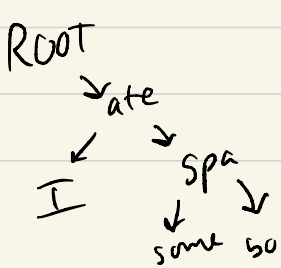
I some

Right-arc

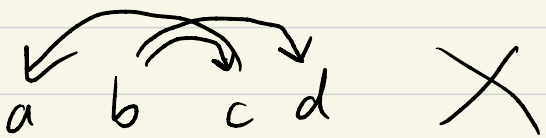
[ROOT ate spa] [ ]

I some bo

R-A, R-A



Arc-standard shift-reduce can build any projective tree



can't build

nonprojective

Building S-R parsers

Parser is a classifier

Maps from (stack, buffer) → action
(S, LA, RA)

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

$p_1\ p_2 = [0\ \ 0]$

Midterm Review

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

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

$$\text{dogs } \begin{bmatrix} 0.1 \\ -0.7 \end{bmatrix} \times 3 + \begin{bmatrix} -0.4 \\ 0.1 \end{bmatrix} \times 2 = \begin{bmatrix} -0.5 \\ -1.9 \end{bmatrix}$$

$$W = \begin{bmatrix} a & \\ & b \end{bmatrix}$$

$a = -100$

$b = -100$

$W s_2$ is large

Ans: E

$W s_1 = 0$

↓ Softmax

→ 0.5 0.5

NP: -2 -1 from rule

DT: 0 N: -1

the ring

ROOT → NP rule

ROOT log prob: -3

1/2 1 prob: $\frac{1}{8}$

NP

1/2

DT N 1/2

1.0 the ring

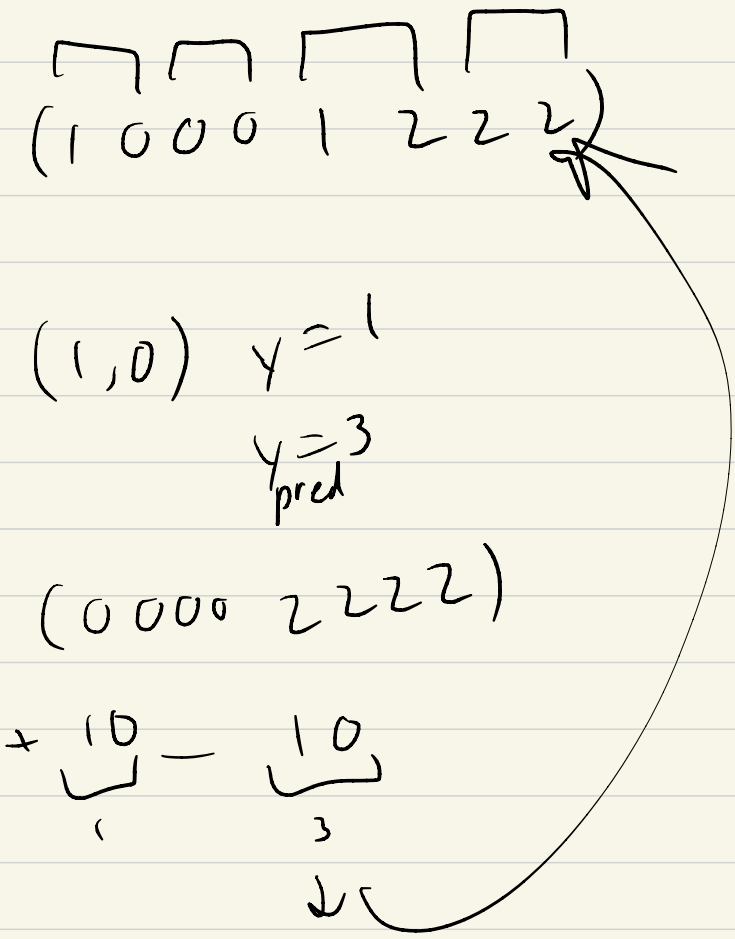
7. NP → NPP NS

3. NPP → rings

OR

7

+ NP → rings


(1 0 0 0 1 2 2 2)
(1,0) y=1
$y_{pred}=3$
(0 0 0 0 2 2 2 2)
+ 1 0 − 1 0
1
3