- A3 due on Next Monday

Outline Dependency parsing, shift-reduce

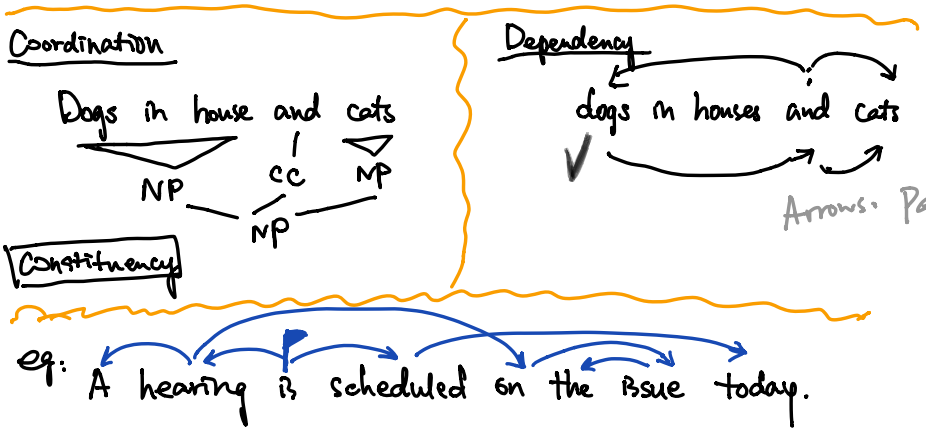
Dependency structure

Root DT NN VBD IN DT NN

the dog ran to the house

PP Attachment

eg. children ate the cake with a spoon



Non-projective

Project tree: has no crossing arcs

$\forall$ edges $i \to j$, all words in (i, j) have parents in $[i, j]$

Arabic : 88.8% proj
Czech : 76.8%
Danish : 84.4%

Shift-reduce transition-based

- Linear time (CKY/Eisner: $O(n^3)$)
- stack: partially built tree

Buffer: words left in the sentence

Shift : word from buffer, pushes it onto stack

Reduce: take partial trees from stack + combines + pushes back on stack

Initial state:

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

→ shift

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

↓ S

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

↓ L-Arc → RA:

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

↓ I ↓ ate

S, S, LA

Reduce: Left-arc / Right-arc } pop two from S, add arc, push on stack

S [ROOT ate] B [some spa bo]

↓ I ↓ S

S [ROOT ate some] B [spa bo]

↓ I ↓ S, L-A

S [ROOT ate spa] B [bo] S [some spa]

↓ I ↓ some ↓ bo

S [ROOT ate spa] B [] → R-A S [ROOT ate]

↓ I ↓ some ↓ bo ↙ spa ↘

Q: n words, how many ops do we need?

A: 2n

arc standard

Short Break

S, S, LA NN JJ

S [ROOT ate some spa] B [bo]

↓ I ↓ LA

Parser Fcn ($\widehat{S, B}$) → op I [stack last POS=NN ∧ buf first POS=JJ ∧ Dec=LA]

3-way classification

"Different features" approach . I [∧ Dec=LA]

- tags in stack - buf
- 2x tags in stack
- previously built arcs
- buffer feats : counts, "peek into future"
- action history
- words in stack and buff

Training

Trees T O(m) trees O(n) states/tree

T ⇒ ((S_1, B_1, a_1^*) , (S_2, B_2, a_2^*) ----)

↖ Oracle (1st part of lecture)

Train classifier over { (S, B, a^*)]

from all exs O(mn)

The horse raced past the barn .

The horse raced past the barn fell .

S [ROOT horse raced] B [put the barn (fell)]

↓ The

Solutions: (1) Beam search : maintain a beam of analyses

(2) Have really good features CKY ???