Advanced Artificial Intelligence

Part II. Statistical NLP

Probabilistic Context Free Grammars

Wolfram Burgard, Luc De Raedt, Bernhard Nebel, Lars Schmidt-Thieme

Slides taken from the Stanford Group (Manning and Schuetze)

Viterbi Algorithm and PCFGs

- This is like the inside algorithm but we find the maximum instead of the sum and then record it

  $\delta_i(p,p)$ = highest probability parse of a subtree N^i_{pq}

1. Initialization: $\delta_i(p,p) = \mathrm{P}(N^i \rightarrow w_p)$
2. Induction:

   $\delta_i(p,q) = \max \mathrm{P}(N^i \rightarrow N^j N^k)\ \delta_j(p,r)\ \delta_k(r+1,q)$
3. Store backtrace:

   $\psi_i(p,q) = \mathrm{argmax}\ \mathrm{P}(N^i \rightarrow N^j N^k)\ \delta_j(p,r)\ \delta_k(r+1,q)$
4. From start symbol N^1, most likely parse t is:

   $\mathrm{P}(t) = \delta_1(1,m)$

Calculation of Viterbi probabilities (CKY algorithm)

	1	2	3	4	5
1	$\delta_{NP} =$ 0.1		$\delta_{S} =$ 0.0126		$\delta_{S} =$ 0.000907
2		$\delta_{NP} =$ 0.04 $\delta_{V} =$ 1.0	$\delta_{VP} =$ 0.126		$\delta_{VP} =$ 0.00907
3			$\delta_{NP} =$ 0.18		$\delta_{NP} =$ 0.01296
4				$\delta_{P} =$ 1.0	$\delta_{PP} =$ 0.18
5					$\delta_{NP} =$ 0.18
	astronomers	*saw*	*stars*	*with*	*ears*

Calculation of inside probabilities (CKY algorithm)

	1	2	3	4	5
1	$\beta_{NP} =$ 0.1		$\beta_S =$ 0.0126		$\beta_S =$ 0.001587
2		$\beta_{NP} =$ 0.04 $\beta_V =$ 1.0	$\beta_{VP} =$ 0.126		$\beta_{VP} =$ 0.01587
3			$\beta_{NP} =$ 0.18		$\beta_{NP} =$ 0.01296
4				$\beta_P =$ 1.0	$\beta_{PP} =$ 0.18
5					$\beta_{NP} =$ 0.18
	astronomers	*saw*	*stars*	*with*	*ears*

Two settings for learning

- Learning from complete data
  - Everything is “observed” “visible”, examples are parse trees
  - Cf. POS-tagging from tagged corpora
  - PCFGs : learning from tree banks,
  - Easy : just counting
- Learning from incomplete data
  - Harder : The EM approach
  - The inside-outside algorithm
  - Learning from the sentences (no parse trees given)

A Penn Treebank tree (POS tags not shown)

```
( (S (NP-SBJ The move)
     (VP followed
         (NP (NP a round)
             (PP of
                 (NP (NP similar increases)
                     (PP by
                         (NP other lenders))
                     (PP against
                         (NP Arizona real estate loans)))))
         ,
         (S-ADV (NP-SBJ *)
                (VP reflecting
                    (NP (NP a continuing decline)
                        (PP-LOC in
                                (NP that market))))))
     .))
```