Conversions between dependencies & constituencies

L715: Seminar on: Data manipulation for parser improvement
Dept. of Linguistics, Indiana University
Fall 2011

Introduction

We are going to focus on dependency & constituency conversions

- Mostly in the direction of: constituency $\rightarrow$ dependency
  - Parser evaluation: dependencies are argued to better capture important aspects of evaluation (e.g., Lin 1995)
  - Dependency parsing has become increasingly popular (linear time algorithms, closeness to semantics, multi-linguality, ...)
    - But constituency treebanks were developed first (at least for English)
  - Even for PCFGs, dependency information can help the parsing model (e.g., Collins 1999)

We'll focus on English ...

Earlier work on conversions

- Magerman (1994); Collins (1999) developed rules to identify the head of a constituent in the PTB
  1. NP $\rightarrow$ DT NN*
  2. VP $\rightarrow$ VBD* NP
  3. S $\rightarrow$ NP VP*
    - head percolation table: priority lists to identify the head in each type of constituent
- Yamada and Matsumoto (2003): modified the table
  - plus: defined arc labeling rules
  - See figures 1 & 2 in Johansson and Nugues (2007)

Extending conversions to be more semantic

Johansson and Nugues (2007)

Goal: improve upon previous methods by making dependencies which interface better with semantics

- richer set of labels
- better treatment of long-distance phenomena
  - PTB-II contains information on wh-movement, topicalization, it-clefts, expletives, & gapping
  - older conversions do not use such information

Make use of extended structure in PTB-II

- The conversion procedure will illustrate a number of issues facing any parser or any annotation scheme

New procedure

Penn2Malt:

Why, they wonder, should it belong to the EC?

New conversion (LTH):

Why, they wonder, should it belong to the EC?

1. Modify dependency links
   - e.g., Penn2Malt misses relation between belong and Why

2. Richer set of dependency labels
   - e.g., Penn2Malt only used SBJ and PRD from PTB-II grammatical function labels
### Heuristically deepening NPs

Need to add internal structure to NPs: PTB has flat structure
- e.g., flat NP of other small apparel makers
- ... but not every word is truly dependent on the head noun (makers)

Heuristics:
- certain adverbs (e.g., quite) are joined with consecutive adjective to form ADJP
- certain words in coordinated NPs (e.g., and Sons) provide clues as to bracketing
- words with identical POS around a conjunction assumed to be coordinated (e.g., small and venomous snake)

nb: see also Vadas and Curran (2007) for NP deepening

### Head rule modifications

Head rules from before are adapted
- make use of the context of phrases
- make use of grammatical functions
  - e.g., SQ ← VBZ VBD VBP VB MD VP *-PRD VP SQ

See table 1 in the paper

### Modification of arc labeling rules

Used 17 of 21 grammatical function labels to label dependency relations
- properties may be combined (e.g., LOC-PRD-TPC)
- excluded ones reflecting structural properties & not grammatical functions (e.g., HLN (headline))

### Structural labels

EXP (expletive) and CLF (cleft) are structural labels, but represent complex constructions
- result in a fronted it
- handled different in PTB, but similarly after conversion
Secondary edges used for a variety of purposes in PTB
- When they represent a “deep governor”, they are useful as dependency arcs (close to semantics)
  - e.g., “T” (trace of wh & topicalization), “ICH” (discontinuous constituent)
  - Such cases are relinked (unless cyclicity is introduced)
- “RNR” (right node raising), e.g., *a U.S. and a Soviet naval vessel*
  - Of the two secondary edges, only the first one is used for conversion

Conversion introduces nonprojectivity: 6.17% of the sentences
Small clauses

Small clause treatment in the PTB:

```
S
 NP  VP
 He  VBP S-1
 made NP-SBJ  ADJP-PRD
   us happy
```

Function tags

14 function tags are used to create dependency labels:
- LTH converts joined tags (e.g., LOC-TMP) into unique tags
- Choi & Palmer select one tag from a joined pair
  - e.g., LOC-TMP ⊃ LOC
  - based on the notion that parsers do not often get joined tags correct (cf. external criteria)

Precendence table:

```
DTV|EXT|LGS|SBJ > LOC > BNF|DIR|MNR|PRP|TMP > SEZ|VOC > PRD > ADV
IGNORE ::= CLF|CLR|ETC|HLN|IMP|NOM|PUT|TPC|TTL|UNF
```

Gapping relations

Parsers perform poorly on gapping constructions:
- LTH tends to give flat structures with long-distance dependencies
- ... which parsers generally get wrong

```
LTH:

SBJ  ROOT  SBJ  OBJ  TMP  PMOD  P  GAP-SBJ  DEP  GAP-PMOD
root Some said. Putin visited in April, some said2 May
```

Coordination

Take a right-branching approach for coordination:
- Difficulty: does a phrase contain coordination?
  - contains UCP, a child annotated with a function tag (ETC), or at least one conjunction (CC) or CONJP
  - “Even if there is a conjunction, if either the left or the right conjunct does not appear within the same phrase, we do not consider there to be a coordination”

```
root We sold old books and then bought new books
```

Note in the algorithm (p. 59) that SKIP defines POS tags which are skipped to find the correct conjuncts

Conversions between dependencies & constituencies

Introduction
Johansson and Nugues (2007)
Choi and Palmer (2010)

References
Choi and Palmer
Nugues (2007)
Johansson and
Nugues (2007)
Empty category mappings

Right node raising is treated slightly differently by Choi & Palmer

- remove link between first conjunct and object
- eliminates non-projective dependencies, but keeps semantic interpretation recoverable

root: I know his admiration for and trust in you

References


Evaluation

- Non-projective dependencies go from 0.82% (LTH) to 0.73% (table 4)
  - largely due to “RNR” treatment
- Unclassified dependencies go from 2.20% (LTH) to 0.60% (table 5)

Parsing accuracy also increases (tables 6 & 7), as does accuracy on semantic dependencies (tables 8 & 9)