Robust Conversion of CCG Derivations
to Phrase Structure Trees

Jonathan K. Kummerfeld†, Dan Klein† and
James R. Curran‡
UC Berkeley† and U Sydney‡
Motivation
Motivation

Various parsing formalisms: HPSG, CCG, LTAG, etc.
Motivation

- Various parsing formalisms: HPSG, CCG, LTAG, etc.

- Scientific Motivation:
  Investigate the relationship between formalisms
Motivation

- Various parsing formalisms: HPSG, CCG, LTAG, etc.

- Scientific Motivation:
  Investigate the relationship between formalisms

- Engineering Motivation:
  Create a useful conversion tool
labeled his death a suicide
labeled his death a suicide
((S[dcl]\NP)/NP)/NP NP/N N NP/N N
labeled his death a suicide

$$((S[dcl]\NP)/NP)/NP\quad NP/N\quad N\quad NP/N\quad N\quad NP$$
labeled his death a suicide

((S[dcl]\NP)/NP)/NP NP/N N NP/N N

NP → NP
Italian magistrates labeled his death a suicide.
Italian magistrates labeled his death a suicide.
Italian magistrates labeled his death a suicide.
Naive Solution
Naive Solution

Categories → Labels
Naive Solution

Categories ➔ Labels

Combinator Applications ➔ Brackets
Italian magistrates labeled his death a suicide.
Italian magistrates labeled his death a suicide.
Clark and Curran (2009)

labeled his death a suicide
Clark and Curran (2009)

Italian magistrates labeled his death a suicide
Italian magistrates labeled his death a suicide.
Clark and Curran (2009)

more than doubled ...

\[
\frac{(S\backslash NP)/(S\backslash NP)}{S[dcl]\backslash NP} \rightarrow S[dcl]\backslash NP
\]
Clark and Curran (2009)

more than doubled ... \[(S\setminus NP)/(S\setminus NP) \quad S[dcl]\setminus NP\] \[\Rightarrow S[dcl]\setminus NP\]

also worried ... \[(S\setminus NP)/(S\setminus NP) \quad S[dcl]\setminus NP\] \[\Rightarrow S[dcl]\setminus NP\]
Clark and Curran (2009)

ADVP

more than

VBD
doubled ...

VP

(S\NP)/(S\NP) (S\NP)/(S\NP)

S[dcl]\NP

S[dcl]\NP

also

S[dcl]\NP

worried ...

S[dcl]\NP
Clark and Curran (2009)

more than doubled ...

also worried ...

(S\NP)/(S\NP) (S\NP)/(S\NP)

S[dcl]\NP S[dcl]\NP
Clark and Curran (2009)

S \[dcl\] \(\mapsto\) S \(\mapsto\) S \(\mapsto\) S

(\(S \NP\) \(\mapsto\) (\(S \NP\)) \(\mapsto\) S \[dcl\] \NP)

\(\mapsto\)

ADVP

VBP

more than

doubled ...

\(S \NP\) \(\mapsto\) (\(S \NP\)) \(\mapsto\) S \[dcl\] \NP

\(\mapsto\)

ADVP

RB

also

\(S \NP\) \(\mapsto\) (\(S \NP\)) \(\mapsto\) S \[dcl\] \NP

\(\mapsto\)

VP

VBN

worried ...

\(S \NP\) \(\mapsto\) (\(S \NP\)) \(\mapsto\) S \[dcl\] \NP
Clark and Curran (2009)

more than doubled ...

(S/np)/(S/np)  S[dcl]\np

also worried ...

(S/np)/(S/np)  S[dcl]\np
Resolving the Issue
Resolving the Issue

also

\((S\backslash NP)/(S\backslash NP)\)

than

\(((S\backslash NP)/(S\backslash NP))\backslash(S[adj]\backslash NP)\)
Conversion from Gold

| F-score | Clark and Curran (2009) | This Work |
|---------|--------------------------|------------|
|         | 94.6                     | 96.3       |

Clark and Curran (2009) compared to This Work, showing an improvement in F-score from 94.6 to 96.3.
Conversion from Parsed

F-score

Clark and Curran (2009) 84.6
This Work 86.2
Conversion from Parsed

F-score

| All Sentences | Clean Sentences |
|---------------|-----------------|
|               | 86.2            |
Conversion from Parsed

F-score

- All Sentences: 86.2
- Clean Sentences: 91.7
Conclusion
Conclusion

- A more robust and accurate conversion method
Conclusion

- A more robust and accurate conversion method
- Insight into treebank differences
Conclusion

- A more robust and accurate conversion method
- Insight into treebank differences
- Code is available at: nlp.cs.berkeley.edu/software
Conclusion

- A more robust and accurate conversion method
- Insight into treebank differences
- Code is available at: nlp.cs.berkeley.edu/software

Thank you!
Comparison with PTB Parsers

All Sentences

- CCG - Auli and Lopez (2011): 86.2
- PTB - Klein and Manning (2003): 85.8
- PTB - Petrov and Klein (2007): 90.1

Clean Sentences

- CCG - Auli and Lopez (2011): 91.7
- PTB - Klein and Manning (2003): 89.8
- PTB - Petrov and Klein (2007): 93.6
Parsing vs. Converting

Converted C&C, Parseval

 Converted Gold, Parseval
Parsing vs. Converting

![Graph showing the relationship between Converted C&C, Parseval and Converted Gold, Parseval. The points are scattered across the graph, indicating a correlation between the two variables.]
Parsing vs. Converting

![Graph showing the relationship between Native C&C (ideps) and Converted Gold (evalb)]