Combinatory Categorial Grammars
Lexicalized Semantically Guided Syntax

Yonatan Bisk
The People (a very abridged version)

Theory
- Mark Steedman (Edinburgh)
- Jason Baldridge (Google)
- Julia Hockenmaier (UIUC)

Data and Parsers

Weakly Supervised Parsing
- Me!
- Dan Garrette (Google)

Efficient Parsing
- Stephen Clark (Cambridge)
- James Curran (Sydney)
- Mike Lewis (FAIR)

Semantic Parsing
- Mirella Lapata
- Adam Lopez
- Johan Boss
- Cem Bozsahin
- Michael White
- Luke Zettlemoyer (UW)
- Yoav Artzi (Cornell)
- Siva Reddy (McGill)
Where is syntax?

- Hierarchical syntactic labels
- Where did the labels come from?
- Are all VPs the same? …

- Semantic labels
- How do you represent long-distance effect?
- Are all languages really tree-structured?
Where is syntax?

S

VP

VP

Sally, who ran home, ate dinner

Sally, who ran home, ate dinner
Where is syntax?

Sally, who ran home, ate dinner

ran( ?, home)  ate( ?, dinner)

...
Syntax as Functions

Noun ↓ Sentence ↓

Sally    ate
Sally    ate( X )
ate( Sally )

Function Application

N  S \ N
← S

S \ N is a function that if applied to a N on the left returns an S

Grammar:

Sally: N
ate:    S \ N

Every word is a function or an argument
Syntax as Functions

Sally ate dinner

Grammar:
Sally: N
dinner: N
ate: S\N, (S\N)/N

Functions returning functions
(S\N)/N N
S\N
S\N
S
Syntax as Functions

Sally: N
ate: (S\N)/N
the: N/N
(dinner: N

?: Takes in N (on right) and returns N

Grammar:
Sally: N
dinner: N
ate: S\N, (S\N)/N
the: N/N
Syntax as Semantics

\[
\begin{align*}
\text{Sally} & \quad \text{ate} \quad \text{the} \quad \text{dinner} \\
N & \quad (S\backslash N)/N \quad N/N \quad N \\
\end{align*}
\]

\[
\begin{align*}
\text{the}(\text{dinner}) \\
\text{ate}(X, \text{the}(\text{dinner})) \\
\text{ate}(\text{Sally}, \text{the}(\text{dinner}))
\end{align*}
\]
Function Composition

\[
S/S \quad (S\backslash N)/N \quad \rightarrow \quad (S\backslash N)/N
\]

Sally quickly ate the dinner

\[
\text{quickly(ate}(X,Y)) \quad \text{the(dinner)}
\]

\[
\text{quickly(ate}(X, \text{the(dinner)})
\]

\[
\text{quickly(ate}(Sally, \text{the(dinner)})
\]
Recap

Grammar

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sally</td>
<td>N</td>
</tr>
<tr>
<td>dinner</td>
<td>N</td>
</tr>
<tr>
<td>the</td>
<td>N/N</td>
</tr>
<tr>
<td>ate</td>
<td>S\N, (S\N)/N</td>
</tr>
<tr>
<td>quickly</td>
<td>S/S</td>
</tr>
</tbody>
</table>

Function Math

Application

\[
\begin{align*}
X\!/Y & \rightarrow X \\
Y & \rightarrow X\!/Y \\
X\!/Y & \rightarrow X
\end{align*}
\]

Composition

\[
\begin{align*}
X\!/Y & \rightarrow X\!/Z \\
X\!/Y & \rightarrow X\!/Z \\
Y\!/Z & \rightarrow X\!/Y \\
Y\!/Z & \rightarrow X\!/Z \\
Y\!/Z & \rightarrow X\!/Z
\end{align*}
\]
What is a relative clause?
A noun modifier

ran(Sally, home)    ate(Sally, dinner)

N
(N\N)?(S\N)
(S\N)/N
N
N\N
N\N
S

Puzzle Time

Sally who ran home ate dinner
Unification & Dependencies

Unification & Dependencies

(S\N)/N Arg1 (S\N)/N Arg2

Sally who ran home

N (N\N)/(S\N) (S\N)/N N

ran(Sally, home)

Note: These are more fine-grained labels than nsubj/dobj
Coordination

Sally ran home and ate dinner

Dependency Grammar

Sally ran home and ate dinner

CCG Dependency Grammar
Coordination

Sally heard but John heard and saw the explosion

Dependency Grammar

CCG Dependency Grammar
Coordination

\[ X \text{ conj } X[\text{conj}] \rightarrow X \]

\[ \text{conj } X[\text{conj}] \rightarrow X \]

apple and orange

\[
\begin{array}{c}
N \text{ conj } N \\
\rightarrow N[\text{conj}] \\
\rightarrow N \\
\end{array}
\]

heard and saw

\[
\begin{array}{c}
S\text{N conj } S\text{N} \\
S\text{N}[\text{conj}] \\
\rightarrow S\text{N} \\
\end{array}
\]
Side note: Crossing Dependencies Are Real

I ate the red and yellow, apple and banana, respectively
Sally heard and John saw the explosion

What should apply to what?
Puzzle Time

Sally heard and John saw the explosion

N (S\N)/N conj N (S\N)/N N

S/N S/N S/N
Puzzle Time

Fill Arg 1 before Arg 2

Sally heard

\[ \begin{array}{c}
N \ (S\backslash N)/N \\
S/N
\end{array} \]

Type-Raising

Sally heard

\[ \begin{array}{c}
N \ (S\backslash N)/N \\
S/(S\backslash N) \\
S/N
\end{array} \]
Lexicon & Rules

Grammar

<table>
<thead>
<tr>
<th>Word</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sally</td>
<td>N</td>
</tr>
<tr>
<td>dinner</td>
<td>N</td>
</tr>
<tr>
<td>the</td>
<td>N/N</td>
</tr>
<tr>
<td>ate</td>
<td>S\N, (S\N)/N</td>
</tr>
<tr>
<td>quickly</td>
<td>S/S</td>
</tr>
<tr>
<td>and</td>
<td>conj</td>
</tr>
</tbody>
</table>

Application

\[
\begin{align*}
X/Y & \rightarrow X \\
Y & \rightarrow X\\Y \\
\end{align*}
\]

Composition

\[
\begin{align*}
X/Y & \rightarrow X/Z \\
X/Y & \rightarrow X/Z \\
Y/Z & \rightarrow X/Z \\Y/Z & \rightarrow X/Z
\end{align*}
\]

Conjunction

\[
\begin{align*}
X & \rightarrow X[\text{conj}] \\
X[\text{conj}] & \rightarrow X \\
\end{align*}
\]

Type-Raising

\[
\begin{align*}
N & \rightarrow S/(S\N) \\
N & \rightarrow S/(S\N)
\end{align*}
\]
That’s it!
Just make up categories
Puzzle Time

Sally ate sushi with tuna

N (S\N)/N  N (N?N)/N  N

N\N

N

S\N

S
Puzzle Time

Sally ate with chopsticks

N S\N (S\S) /N N

S\S

S\N

S
Sweeping things under the rug

• It’s really just S, N, and conj? Well... no
  NP, PP

• Ok, so 5 categories? Sorta...
  S[adj], S[dcl], S[b],... NP[nb], ...

• OK, but we did learn all the rules right?
  Ugh,... “Yes” — ignore Type-Changing, it’s not really real...

• Are you lying to make this formalism sound prettier then it really is?
As I said, everything is perfect, there are no questions, everyone is happy with this result. CCG is beautiful and perfect.
Why CCG?

Sally, who ran home, ate dinner

$$\lambda y. \lambda x. f(x, y)$$

- ran: $(S\backslash N)/N$
- ate: $(S\backslash N)/N$
- who: $(N\backslash N)/(S\backslash N)$

$\lambda y. \lambda x. run(x, y)$
$\lambda y. \lambda x. eat(x, y)$
$\lambda f.f$

ran(Sally, home)    ate(Sally, dinner)

Could be SQL, SPARQL, python, etc
How should we define a probabilistic model?
Modeling Supertagging

Input:
Sally
ate
with chopsticks

Labels:
N S\N (S\S)\N N

S\S
S\N
S

Normal-form parsing for Combinatory Categorial Grammars with generalized composition and type-raising — Hockenmaier 2010
# Supertag LSTM Analysis

<table>
<thead>
<tr>
<th>Supertag</th>
<th>F-For</th>
<th>Forward</th>
<th>Backward</th>
<th>bi-LSTM</th>
<th>+LM(g–train)</th>
<th>ss–train–1</th>
<th>ss–train–5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(NP\NP)/NP</td>
<td>90.00</td>
<td>88.89</td>
<td>81.91</td>
<td>92.09</td>
<td>92.18</td>
<td>91.72</td>
<td><strong>92.31</strong></td>
</tr>
<tr>
<td>((S\NP)(S\NP))/NP</td>
<td>75.75</td>
<td>69.53</td>
<td>61.60</td>
<td><strong>80.38</strong></td>
<td>78.21</td>
<td>79.91</td>
<td>78.77</td>
</tr>
<tr>
<td>S[dcl]\NP</td>
<td>77.29</td>
<td>61.14</td>
<td>58.52</td>
<td><strong>84.28</strong></td>
<td>83.41</td>
<td>82.97</td>
<td>80.35</td>
</tr>
<tr>
<td>(S[dcl]\NP)/NP</td>
<td>91.39</td>
<td>56.58</td>
<td>69.86</td>
<td>92.34</td>
<td>92.46</td>
<td>92.46</td>
<td><strong>92.82</strong></td>
</tr>
<tr>
<td>((S[dcl]\NP)/PP)/NP</td>
<td>42.30</td>
<td>30.77</td>
<td>42.31</td>
<td>56.41</td>
<td><strong>64.10</strong></td>
<td>62.82</td>
<td>60.26</td>
</tr>
<tr>
<td>(S[dcl]\NP)/(S[adj]\NP)</td>
<td>86.80</td>
<td>22.84</td>
<td>83.25</td>
<td>87.31</td>
<td><strong>88.83</strong></td>
<td>87.82</td>
<td>86.80</td>
</tr>
<tr>
<td>((S[dcl]\NP)/(S[to]\NP))/NP</td>
<td>86.49</td>
<td>56.76</td>
<td>75.68</td>
<td><strong>94.59</strong></td>
<td>91.89</td>
<td>91.89</td>
<td>91.89</td>
</tr>
</tbody>
</table>

Supertagging with LSTMs Vaswani 2016
Modeling the Arguments

\[
S \rightarrow Y = N \\
\text{Combinator} = >B_0
\]
### Induced Lexicons: Adjectives

<table>
<thead>
<tr>
<th>English</th>
<th>Arabic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Big</strong></td>
<td><strong>كرة</strong></td>
</tr>
<tr>
<td><strong>N/N</strong></td>
<td><strong>N</strong></td>
</tr>
<tr>
<td><strong>Ball</strong></td>
<td><strong>كبيرة</strong></td>
</tr>
<tr>
<td><strong>N</strong></td>
<td><strong>N/N</strong></td>
</tr>
</tbody>
</table>

**Examples:**
- English: Big Ball
- Arabic: كرة كبيرة

**Induced Lexicons:**
- English: *Obj Adj*
- Arabic: *Obj Adj*
## Induced Lexicons: Verbs

<table>
<thead>
<tr>
<th>Language</th>
<th>Sentence</th>
<th>English Word Form</th>
<th>Arabic Word Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>The man wrote a letter</td>
<td>S V O</td>
<td>فيكتتب الرجل رسالة</td>
</tr>
<tr>
<td>Child Directed Speech</td>
<td>write a letter</td>
<td>Ø V O</td>
<td>فيكتتب الرجل رسالة</td>
</tr>
<tr>
<td>Arabic</td>
<td>كتاب الرجال رسالة</td>
<td>V S O</td>
<td>فيكتتب الرجل رسالة</td>
</tr>
</tbody>
</table>

The symbols (S/N) and (S\N) represent different word forms, and the symbols N, S, and V correspond to noun, subject, and verb in the English and Arabic sentence structures.
### Induced Lexicons: Adpositions

<table>
<thead>
<tr>
<th>English</th>
<th>Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>ran</td>
<td>浜を走った</td>
</tr>
<tr>
<td>on</td>
<td>を (S/S)</td>
</tr>
<tr>
<td>beach</td>
<td>走ったN</td>
</tr>
<tr>
<td>(S/S)/N</td>
<td>(on)</td>
</tr>
<tr>
<td>N</td>
<td>S/N</td>
</tr>
</tbody>
</table>

*(beach)*
Prepositions can be tricky

Is “to Boston” a modifier?

Go
S/NP

S

PP

to

PP

N

Boston

S
\[
\begin{array}{cccc}
\text{show me} & \text{flights} & \text{to} & \text{Boston} \\
S/N \quad \lambda f. f & N \quad \lambda x. \text{flight}(x) & PP/NP \quad \lambda y. \lambda x. \text{to}(x, y) & NP \quad BOSTON \\
\hline
\hline
\end{array}
\]

\[
\frac{PP}{\lambda x. \text{to}(x, BOSTON)} >
\]

\[
\frac{N \setminus N}{\lambda f. \lambda x. f(x) \land \text{to}(x, BOSTON)} <
\]

\[
\frac{N}{\lambda x. \text{flight}(x) \land \text{to}(x, BOSTON)} >
\]

\[
\frac{S}{\lambda x. \text{flight}(x) \land \text{to}(x, BOSTON)}
\]
You parse so you can do something

How does a robot check if it’s at the right location?

\[
\begin{align*}
\text{chair} & \quad \text{in} & \quad \text{the} & \quad \text{corner} \\
\lambda x.\text{chair}(x) & \quad \lambda x.\lambda y.\text{intersect}(x, y) & \quad \lambda f.\lambda x.f(x) & \quad \lambda x.\text{corner}(x) \\
N & \quad PP/NP & \quad NP/N & \quad NP \\
\text{corner}(x) & \quad \lambda y.\text{intersect}(x, y) & \quad \lambda f.\lambda y.f(y) \land \text{intersect}(x, \text{chair}(x), y) & \quad \lambda y.\text{chair}(y) \land \text{intersect}(x, \text{chair}(x), y) \\
\end{align*}
\]

Weakly Supervised Learning of Semantic Parsers for Mapping Instructions to Actions — Artzi 2013

Query a knowledge base

(a) Who directed The Nutty Professor?

(b) Austin is the state capital of Texas.

Large-scale Semantic Parsing without Question-Answer Pairs — Reddy 2014
Where to learn more?

Semantic Parsing and Modeling

https://yoavartzi.com/tutorial/

Linguistics

Mark Steedman

Jazz

http://jazzparser.granroth-wilding.co.uk/Parser.html