Combined Distributional and Logical Semantics

Mike Lewis and Mark Steedman
University of Edinburgh
Semantics

Was Obama born in Kenya?

Obama's birthplace is not Kenya

No.
Distributional Semantics

- Induce the meanings of words from corpora
Distributional Semantics

- Induce the meanings of words from corpora

<table>
<thead>
<tr>
<th>X was born in Y</th>
<th>X's birthplace is Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Obama, Hawaii)</td>
<td>(Obama, Hawaii)</td>
</tr>
<tr>
<td>(Shakespeare, Stratford)</td>
<td>(Shakespeare, Stratford)</td>
</tr>
<tr>
<td>(Obama, 1961)</td>
<td>(Jesus, Bethlehem)</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

(e.g. Lin & Pantel, 2001)
\[ \text{sim}(X \text{ was born in } Y, \ X's \ birthplace \ is \ Y) > t \]

\[ \Rightarrow \text{Obama was born in Hawaii} \]

\[ \Rightarrow \text{Obama's birthplace is Hawaii} \]
Distributional Semantics

Was Obama born in Kenya?

Obama's birthplace is not Kenya

Yes!
Formal Semantics

Obama was born in Hawaii.
Formal Semantics

Obama
NP  obama

was
(S\NP)/(S\NP)
λpλx . p(x)

born in
(S\NP)/NP
λyλx . born_in'(x, y)

Hawaii
NP  hawaii

S\NP

S\NP

S
Formal Semantics

\[
\text{Obama} \quad \text{was} \quad \text{born in} \quad \text{Hawaii}
\]

\[
(S\NP)/(S\NP) \quad \lambda p \lambda x . p(x) \quad (S\NP)/NP \quad \lambda y \lambda x . \text{born_in'}(x, y) \quad \text{NP}
\]

\[
(\lambda x . \text{born_in'}(x, \text{hawaii}))
\]

\[
\lambda x . \text{born_in'}(x, \text{hawaii})
\]

\[
\text{born_in'}(\text{obama}, \text{hawaii})
\]
Formal Semantics

Obama NP
  obama wasn't
  (S\NP)/(S\NP)
    λpλx . ¬p(x)

  born in
  (S\NP)/NP
    λyλx . born_in'(x, y)

  Kenya NP
    kenya


S\NP
  λx . born_in'(x, kenya)

S\NP
  λx . ¬born_in'(x, kenya)

S
  ¬born_in'(obama, kenya)
Formal Semantics

Obama wasn't born in Kenya

⇒

Obama was born in Kenya
Every US president was born in the United States

⇒

Obama wasn't born in Kenya
Formal Semantics

Nicole Kidman and Barack Obama were born in Hawaii

⇔

Nicole Kidman was born in Hawaii and Barack Obama was born in Hawaii
Was Obama born in Kenya?

Obama's birthplace is not Kenya

Don't know.
$\neg \text{birthplace}'(\text{obama, kenya})$

$\Rightarrow ?$

$\text{born\_in}'(\text{obama, kenya})$
Formal Semantics

+ Models semantic operators ("not", "every", etc.)

+ Abstracts away from syntax (conjunctions, passive voice, relative clauses, etc)

- Doesn't model meaning of content words
Combining Formal and Distributional Semantics

Existing approaches:

- **Compositional vector space models**
  - (e.g. Coecke et al., 2011, Baroni et al., 2012, Grefenstette 2013)

- **Learn inference rules as logical axioms**
  - (e.g. Garrette et al., 2011, Beltagy et al., 2013)
Combining Formal and Distributional Semantics

Obama was born in Hawaii

\[ \text{born_in}'(\text{obama}, \text{hawaii}) \]

Hawaii is Obama's birthplace

\[ \text{birth_place}'(\text{obama}, \text{hawaii}) \]
Combining Formal and Distributional Semantics

Obama was born in Hawaii

Hawaii is Obama's birthplace

relation53(Obama, Hawaii)

relation53(Obama, Hawaii)
Learn lexical entries such that:

born \quad \vdash (S\backslash NP)/PP[\text{in}] : \lambda x\lambda y . relation53(x,y)
birthplace \quad \vdash N/PP[\text{of}] : \lambda x\lambda y . relation53(x,y)
Formal Semantics

Obama NP

was

(S\NP)/(S\NP)

\lambda p \lambda x . \ p(x)

born in

(S\NP)/NP

\lambda y \lambda x . \text{relation53}(x, y)

Hawaii NP

hawaii

(S\NP)

\lambda x . \text{relation53}(x, \text{hawaii})

S\NP

\lambda x . \text{relation53}(x, \text{hawaii})

S

\text{relation53}(\text{obama, hawaii})
Combining Formal and Distributional Semantics

Gather statistics on predicates in large corpus:

<table>
<thead>
<tr>
<th>born_in'(X, Y)</th>
<th>birthplace_of'(X, Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Obama, Hawaii)</td>
<td>(Obama, Hawaii)</td>
</tr>
<tr>
<td>(Shakespeare, Stratford)</td>
<td>(Shakespeare, Stratford)</td>
</tr>
<tr>
<td>(Obama, 1961)</td>
<td>(Jesus, Bethlehem)</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>
Combining Formal and Distributional Semantics

born_in'(X, Y)
birth_place(X,Y)

buy'(X, Y)
purchase'(X, Y)
acquisitionOf'(X, Y)

employeeOf'(X, Y)
workFor'(X, Y)
write'(X, Y)
authorOf'(X, Y)
Combining Formal and Distributional Semantics

- born_in'(X, Y)
- birth_place(X,Y)
- buy'(X, Y)
- purchase'(X, Y)
- acquisitionOf'(X, Y)
- employeeOf'(X, Y)
- workFor'(X, Y)
- write'(X, Y)
- authorOf'(X,Y)
Combining Formal and Distributional Semantics

Cluster using Chinese Whispers (Biemann, 2006)
Was Obama born in Kenya?

Obama's birthplace is not Kenya

No.
Ambiguity

Obama was born in Hawaii
Obama was born in 1961
Ambiguity

Obama was born in 1961

\[\not\Rightarrow\]

Obama's birthplace is 1961
Create multiple typed lexical entries:

\[
\begin{align*}
\text{born} & \vdash (S\backslash NP)/PP[\text{in}] : \lambda y \ \lambda x \quad . \ birthPlace(x,y) \\
\text{born} & \vdash (S\backslash NP)/PP[\text{in}] : \lambda y \quad \lambda x \quad . \ birthDate(x,y)
\end{align*}
\]

(similar to: Schoenmackers et al., 2010; Berant et al., 2011)
Clustering Typed Predicates

Only cluster predicates with the same type

- born_in(X:PER, Y:DAT)
- birth_date(X:PER, Y:DAT)

\[ \text{relation53}(X, Y) \]

- born_in(X:PER, Y:LOC)
- birth_place(X:PER, Y:LOC)
- native_of(X:PER, Y:LOC)

\[ \text{relation91}(X, Y) \]
### Topic Model

One 'document' per predicate
One 'word' per argument

<table>
<thead>
<tr>
<th>lives in X</th>
<th>year of X</th>
<th>born in X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>1564</td>
<td>Hawaii</td>
</tr>
<tr>
<td>London</td>
<td>2001</td>
<td>2001</td>
</tr>
<tr>
<td>France</td>
<td>1945</td>
<td>London</td>
</tr>
</tbody>
</table>

1564
Topic Model

One 'document' per predicate
One 'word' per argument

<table>
<thead>
<tr>
<th>lives in X</th>
<th>year of X</th>
<th>born in X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>1564</td>
<td>Hawaii</td>
</tr>
<tr>
<td>London</td>
<td>2001</td>
<td>2001</td>
</tr>
<tr>
<td>France</td>
<td>1945</td>
<td>London</td>
</tr>
</tbody>
</table>

(see also: Melamud et al., 2013)
<table>
<thead>
<tr>
<th>Example Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>suspect, assailant, fugitive, accomplice, ...</td>
</tr>
<tr>
<td>author, singer, actress, actor, dad, ...</td>
</tr>
<tr>
<td>city, area, country, region, town, capital, ...</td>
</tr>
<tr>
<td>subsidiary, automaker, airline, Co., GM, ...</td>
</tr>
<tr>
<td>musical, thriller, sequel, special, ...</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>
Combining types

\[
\lambda y: \begin{cases}
\text{LOC} & \sim 0.5 \\
\text{DATE} & \sim 0.4 \\
\text{...}
\end{cases}
\]

\[
\lambda x: \begin{cases}
\text{PER} & \sim 0.9 \\
\text{IDEA} & \sim 0.05 \\
\text{...}
\end{cases}
\]

\[
\text{born\_in}(x, y)
\]

\[
\begin{cases}
\text{2001} \\
\text{NP}
\end{cases}
\]

\[
\begin{cases}
\text{DATE} & \sim 0.8 \\
\text{FILM} & \sim 0.15 \\
\text{...}
\end{cases}
\]

\[
(S\backslash\text{NP})/\text{NP}
\]
Combining types

\[ \lambda y: \{ \text{LOC} \sim 0.5, \text{DATE} \sim 0.4, \ldots \} \quad \lambda x: \{ \text{PER} \sim 0.9, \text{IDEA} \sim 0.05, \ldots \} \quad \text{born}_\text{in}(x, y) \quad 2001: \{ \text{DATE} \sim 0.8, \text{FILM} \sim 0.15, \ldots \} \]

\[ (S\langle NP \rangle)/NP \]

\[ S\langle NP \rangle \]

\[ \lambda x: \{ \text{PER} \sim 0.9, \text{IDEA} \sim 0.05, \ldots \} \quad \text{born}_\text{in}(x, 2001: \{ \text{DATE} \sim 0.95, \text{LOC} \sim 0.02, \ldots \}) \]
Combining types

\( \lambda y: \{ \text{LOC} \sim 0.5, \text{DATE} \sim 0.4, \ldots \} \)

\( \lambda x: \{ \text{PER} \sim 0.9, \text{IDEA} \sim 0.05, \ldots \} \)

\( \text{born in} \)

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

\( \text{born}_in(x, y) \)

\( \text{Washington} \)

\( NP \)

\( \text{Wash}: \{ \text{LOC} \sim 0.8, \text{PER} \sim 0.15, \ldots \} \)
Combining types

\[ \lambda y: \begin{cases} \text{LOC } \sim 0.5 \\ \text{DATE } \sim 0.4 \\ \text{...} \end{cases} \quad \lambda x: \begin{cases} \text{PER } \sim 0.9 \\ \text{IDEA } \sim 0.05 \\ \text{...} \end{cases} \quad \text{born}_\text{in}(x, y) \]

\[ \text{Washington} \]

\[ \text{NP} \]

\[ \text{Wash}: \begin{cases} \text{LOC } \sim 0.8 \\ \text{PER } \sim 0.15 \\ \text{...} \end{cases} \]

\[ \text{S}\text{\backslash NP} \]

\[ \lambda x: \begin{cases} \text{PER } \sim 0.9 \\ \text{IDEA } \sim 0.05 \\ \text{...} \end{cases} \quad \text{born}_\text{in}(x, \text{Wash}: \begin{cases} \text{LOC } \sim 0.95 \\ \text{PER } \sim 0.01 \\ \text{...} \end{cases}) \]
Compositional Semantics

\[
\begin{align*}
\text{born} \vdash (S\backslash NP)/PP[\text{in}] &: \lambda y_{\text{LOC}} \lambda x_{\text{PER}}. \text{birthPlace}(x,y) \\
\text{born} \vdash (S\backslash NP)/PP[\text{in}] &: \lambda y_{\text{DAT}} \lambda x_{\text{PER}}. \text{birthDate}(x,y)
\end{align*}
\]
born ⊢ (S\NP)/PP[<in>] : λy_{LOC}λx_{PER}. birthPlace(x,y)
born ⊢ (S\NP)/PP[<in>] : λy_{DAT}λx_{PER}. birthDate(x,y)

\[
\begin{align*}
\text{born} & \vdash (\text{S}\backslash\text{NP})/\text{PP[<in>]} : \lambda y_{\text{LOC}} \lambda x_{\text{PER}}. \text{birthPlace}(x,y) \\
\text{born} & \vdash (\text{S}\backslash\text{NP})/\text{PP[<in>]} : \lambda y_{\text{DAT}} \lambda x_{\text{PER}}. \text{birthDate}(x,y)
\end{align*}
\]

\[
\begin{align*}
\text{born} & \vdash (\text{S}\backslash\text{NP})/\text{PP[<in>]} : \lambda y \lambda x. \left\{ (x:\text{LOC}, y:\text{PER}) \Rightarrow \text{birthPlace}(x, y) \right. \\
& \left. (x:\text{DAT}, y:\text{PER}) \Rightarrow \text{birthDate}(x, y) \right\}
\end{align*}
\]
Output packed logical form capturing joint distribution:

Obama was born in Hawaii in 1961

\[
\begin{align*}
\{ & \text{birthPlace } \sim 0.9 \\
& \text{birthDate } \sim 0.1 \\
\} & (\text{Obama, Hawaii})
\end{align*}
\]

\[
\wedge \begin{align*}
\{ & \text{birthPlace } \sim 0.1 \\
& \text{birthDate } \sim 0.9 \\
\} & (\text{Obama, 1961})
\end{align*}
\]
Training Details

- Train on English Gigaword
- Use C&C parser for syntax (Clark and Curran 2004)
- 15 entity types
- Lexicon includes manual entries for some function words ("every", "not", etc.)
Evaluation

Automatically generate question set from corpus

Google bought YouTube

Who bought YouTube?
What did Google buy?
Evaluation

Automatically generate question set from corpus

Google bought YouTube

Who bought YouTube?

What did Google buy?

Generate 1000 questions

Manual evaluation
Comparison Systems

- Reverb (Fader et al., 2011)
- Relational LDA (Yao et al., 2011)
- CCG Baseline
- CCG + WordNet
- CCG using Distributional Clusters
Question: What did Delta merge with?
Sentence: The 747 freighters came with Delta's acquisition of Northwest
Answer: Northwest
Problems involving inference using quantifiers from the FraCaS Suite (Section 1):

Premises:
Every European has the right to live in Europe.
Every European is a person.
Every person who has the right to live in Europe can travel freely within Europe.

Hypothesis:
Every European can travel freely within Europe
## Evaluating Formal Semantics

<table>
<thead>
<tr>
<th></th>
<th>Single Premise Sentence</th>
<th>Multiple Premise Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Natural Logic 2007</strong></td>
<td>84%</td>
<td>-</td>
</tr>
<tr>
<td><strong>Natural Logic 2008</strong></td>
<td>98%</td>
<td>-</td>
</tr>
<tr>
<td><strong>CCG-Dist (parser)</strong></td>
<td>70%</td>
<td>50%</td>
</tr>
<tr>
<td><strong>CCG-Dist (gold syntax)</strong></td>
<td>89%</td>
<td>80%</td>
</tr>
</tbody>
</table>
Conclusions

- Formal and distributional semantics complement each other
- Modelling content words using cluster identifiers gives the benefits of both
Any questions?