Don’t ‘have a clue’?

Unsupervised co-learning of downward-entailing operators

Cristian Danescu-Niculescu-Mizil
Lillian Lee

Department of Computer Science, Cornell University

ACL 2010
Why my poster?

Tackles a complex semantic phenomenon.

Uses only raw text + one “seed word”.

The only solution that works for Klingon.
Monotonicity

“I know I’ll buy a Mac” → “I know I’ll buy a computer”
Monotonicity

“I know I’ll buy a Mac” $\Rightarrow$ “I know I’ll buy a computer”

“I doubt I’ll buy a Mac” $\nRightarrow$ “I doubt I’ll buy a computer”
Monotonicity

“I know I’ll buy a Mac” $\implies$ “I know I’ll buy a computer”

“I doubt I’ll buy a Mac” $\impliedby$ “I doubt I’ll buy a computer”
Monotonicity

Upward monotone:
“\( \text{I know I’ll buy a Mac} \) ⇒ “\( \text{I know I’ll buy a computer} \)”

Downward monotone:
“\( \text{I doubt I’ll buy a Mac} \)⇐ “\( \text{I doubt I’ll buy a computer} \)”
Monotonicity

Upward monotone:

“I know I’ll buy a Mac” $\implies$ “I know I’ll buy a computer”

<subset> <set>

Downward monotone:

“I doubt I’ll buy a Mac” $\iff$ “I doubt I’ll buy a computer”

<subset> <set>
Downward-entailing operators

Downward-entailing operators invert the default monotonicity, allowing one to “reason from sets to subsets”.

Examples:

- "I doubt I'll buy a computer" = ⇒ "I doubt I'll buy a Mac"
- "He came without cash or cards" = ⇒ "He came without cash"
- "She is too lazy to run" = ⇒ "She is too lazy to run a 10k"
Downward-entailing operators

Downward-entailing operators invert the default monotonicity, allowing one to “reason from sets to subsets”.

Examples:

\[ \text{<set>} \quad \text{<superset>} \]

“I doubt I’ll buy a computer” \(\implies\) “I doubt I’ll buy a Mac”

“He came without cash or cards” \(\implies\) “He came without cash”

“She is too lazy to run” \(\implies\) “She is too lazy to run a 10k”
Downward-entailing operators

Task:
Automatically discover DE operators.

Challenges:
No monotonicity-annotated corpora.
Not deducible from any public lexical database.
[Nairn et al., 2006]
Downward-entailing operators

Why?

Linguistic importance:
DE operators play “an extremely important role in natural language”
[van der Wouden, 1997; van Benthem, 1986; Hoeksema, 1986; Dowty, 1994;
Sánchez Valencia, 1991]

Textual Entailment:
TE systems that approach monotonicity rely on relatively small
hand-annotated lists of English DE operators.
[Nairn et al.; 2006, MacCartney and Manning, 2008; Bar-Haim et al., 2008.]

Natural Language Generation:
DM inferences induce greater cognitive load than UM inferences.
[Geurts and van der Slik, 2005]

Prevalence:
At least 6% of newswire sentences contain a non-trivial DE operator
[Danescu-Niculescu-Mizil et al., 2009]
How to find DE ops.?

Before:

[Danescu-Niculescu-Mizil et al., 2009]

Sprinkle some linguistic magic powder over the raw text: “NPIs” are noisy clues for DE operators.

Examples:

anymore, have a clue, budge, give a damn, ...
How to find DE ops.?

Before:

[Danescu-Niculescu-Mizil et al., 2009]
Sprinkle some linguistic magic powder over the raw text: “NPIs” are noisy clues for DE operators.

Examples:

“They do not listen anymore.” vs. “*They do listen anymore.”

“I doubt they have a clue.” vs. “*They have a clue.”
How to find DE ops.?

Before:

[Danescu-Niculescu-Mizil et al., 2009]
Sprinkle some linguistic magic powder over the raw text: “NPIs” are noisy clues for DE operators.

Examples:
“They do not listen anymore.” vs. “*They do listen anymore.”
“I doubt they have a clue.” vs. “*They have a clue.”

Where’s the green?
“It is wise to try compensating for any excess.”
How to find DE ops.?

Before [Danescu-Niculescu-Mizil et al., 2009]:

**Input:**

<table>
<thead>
<tr>
<th>English Text + English NPIs:</th>
</tr>
</thead>
<tbody>
<tr>
<td>have a clue</td>
</tr>
<tr>
<td>give a damn</td>
</tr>
<tr>
<td>anymore</td>
</tr>
<tr>
<td>a red cent</td>
</tr>
<tr>
<td>budge</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

**Output:**

DE ops.
How to find DE ops.?

Before [Danescu-Niculescu-Mizil et al., 2009]:

<table>
<thead>
<tr>
<th>English Text + English NPIs:</th>
<th>Romanian Text + Romanian NPIs:</th>
</tr>
</thead>
<tbody>
<tr>
<td>have a clue</td>
<td></td>
</tr>
<tr>
<td>give a damn</td>
<td></td>
</tr>
<tr>
<td>anymore</td>
<td></td>
</tr>
<tr>
<td>a red cent</td>
<td></td>
</tr>
<tr>
<td>budge</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

Input:

Output:
DE ops.
How to find DE ops.?

Before [Danescu-Niculescu-Mizil et al., 2009]:

<table>
<thead>
<tr>
<th>English Text + English NPIs:</th>
</tr>
</thead>
<tbody>
<tr>
<td>have a clue</td>
</tr>
<tr>
<td>give a damn</td>
</tr>
<tr>
<td>anymore</td>
</tr>
<tr>
<td>a red cent</td>
</tr>
<tr>
<td>budge</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Romanian Text + Romanian NPIs:</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
</tr>
<tr>
<td>?</td>
</tr>
<tr>
<td>?</td>
</tr>
<tr>
<td>?</td>
</tr>
<tr>
<td>...</td>
</tr>
</tbody>
</table>

Input:  

Output: DE ops.
How to find DE ops.?

Before [Danescu-Niculescu-Mizil et al., 2009]:

**Input:**
- English Text + English NPIs: have a clue, give a damn, anymore, a red cent, budge...

**Output:**
- DE ops.

![Diagram with examples of English and Romanian NPIs and their evaluation]
How to find DE ops.?

This work: a co-learning approach.

**Input:**

<table>
<thead>
<tr>
<th>EN Text +</th>
<th>RO Text +</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>One EN NPI</td>
<td>One RO NPI</td>
<td>...</td>
</tr>
</tbody>
</table>

**Byprod.:**

<table>
<thead>
<tr>
<th>pseudo-NPIs</th>
<th>pseudo-NPIs</th>
<th>...</th>
</tr>
</thead>
</table>

**Output:**

<table>
<thead>
<tr>
<th>DE ops.</th>
<th>...</th>
</tr>
</thead>
</table>
Come to poster #4!

Results:
First time DE operators are learned for a language other than English!

How to chose the seed word?
Spoiler: seed word for Klingon is vay’

Does it really work for Klingon?
Connections to linguistic typology.