President Barack Obama received the Serve America Act after congress’ vote. **He signed the bill last Thursday.** The president said it would greatly increase service opportunities for the American people.
Question Answering

Q: Who signed the Serve America Act?
A: Barack Obama

Los Angeles Times

President Barack Obama received the Serve America Act after congress’ vote. He signed the bill last Thursday. The president said it would greatly increase service opportunities for the American people.

Machine Translation

امضاء کرد جوانان را

He signed the bill
Discourse Structure

President Barack Obama received the Serve America Act after congress’ vote. He signed the bill last Thursday. The president said it would greatly increase service opportunities for the American people.
Entity Reference Resolution

Input:

President Barack Obama received the Serve America Act after congress close vote. He signed the bill last Thursday. The president said it would greatly increase service opportunities for the American people.
Entity Reference Resolution

Output:

President Barack Obama received the Serve America Act after congress close vote.
He signed the bill last Thursday.
The president said it would greatly increase service opportunities for the American people.
Entity Resolution Model

Barack Obama ... the Serve America Act ... congress

He ... the bill

The president ... it ... American People

[Haghighi & Klein, ACL '07 ]
Finite Mixture Model

Entity Distribution

Z₁
W₁
“Obama”

Z₂
W₂
“Service Act”

Z₃
W₃
“he”

Barack Obama
“Barack Obama”: 0.30
“president”: 0.20
“he”: 0.15
“him”: 0.20

Infinite Mixture Model

Entity Distribution

Z₁
W₁
“Obama”

Z₂
W₂
“Service Act”

Z₃
W₃
“he”
Infinite Mixture Model

The Weir Group, whose headquarters is in the U.S, is a large specialized corporation. This power plant, which, will be situated in Jiangsu, has a large generation capacity.

Enriching the Mention Model

Mention Model

\[ P(W | \text{Weir Group}) \]

"Weir Group"=0.4, "whose"=0.2, .......
Enriching the Mention Model

Pronoun

Non-Pronoun

Entity Parameters

Pronoun Parameters

W | SING, MALE, PERS
P "he": 0.5, "him": 0.3, ...

W | PL, NEUT, ORG
P "they": 0.3, "it": 0.2, ...
Enriching the Mention Model

Pronoun

Non-Pronoun

Enriching Mention Model

Mention Type
Proper, Pronoun, Nominal

Non-pronoun
Pronoun
Enriching Mention Model
The Weir Group, whose headquarters is in the U.S, is a large specialized corporation. This power plant, which, will be situated in Jiangsu, has a large generation capacity.
What is Learned?

PROPER

Probability

<table>
<thead>
<tr>
<th>Level</th>
<th>NONE</th>
<th>LOW</th>
<th>MED</th>
<th>HIGH</th>
<th>TOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>1</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Less Salience More

What is Learned?

PRONOUN

Probability

<table>
<thead>
<tr>
<th>Level</th>
<th>NONE</th>
<th>LOW</th>
<th>MED</th>
<th>HIGH</th>
<th>TOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>1</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Less Salience More
What is Learned?

NOMINAL

Probability

0
NONE LOW MED HIGH TOP

Less Salience More

Discourse Salience
The Weir Group, whose headquarters is in the U.S, is a large specialized corporation. This power plant, which will be situated in Jiangsu, has a large generation capacity.
The Weir Group, whose headquarters is in the U.S, is a large specialized corporation. This power plant, which, will be situated in Jiangsu, has a large generation capacity.
Global Entity Resolution

- Bush
- Rice
- he
- she

Rice
- Bush
- she

Experiments

- **MUC6 English NWIRE (all mentions)**
  - 53.6 F1* [Cardie and Wagstaff 99]  
    *Unsupervised*
  - 70.3 F1  [Haghighi & Klein 07]  
    *Unsupervised*
  - 73.4 F1  [McCallum & Wellner 04]  
    *Supervised*
  - 81.3 F1  [Luo et al 04]  
    *Supervised+Oracle*

  - *MUC score, not all systems are completely comparable*

  - *Why not: some assume oracle gives boundaries, NER types, etc.*
Evaluation

Cluster Similarity: MUC F₁

* Now system mentions!
These are "real numbers"

Unsupervised  Supervised  Our Work, '07  [Denis et al., '07]

Larger is better

Errors

Input
America Online announced on Monday that the company plans to update its instant messaging service.

Correct
America Online the company its instant messaging service

Guess
America Online the company its instant messaging service
Adding Semantic Knowledge

**America Online** ← --- → **company**

**America Online, LLC** (commonly known as AOL) is an American global internet services and media company operated by Time Warner. It is headquartered at 770 Broadway in Midtown Manhattan, New York City. Founded in 1983 as Quantum Computer Services, it has franchised its services to companies in several nations around the world or set up international versions of its services.

**America Online**
- **Type**: Subsidiary of Time Warner
- **Founded**: 1983 as Quantum Computer Services

Adding Semantic Knowledge

**America Online, LLC** (commonly known as AOL) is an American global internet services and media company.

**George W. Bush**, the 43rd president, is already the subject of more than 100 books.

**trade union**: an organization whose members belong ...
Richer Semantic Modeling

The Federal Reserve announced that the dollar dropped this week. It fell ...

The Federal Reserve announced that the dollar dropped this week. It predicted ...
Prototype-Driven Learning

Unlabeled Data + Prototype List → Annotated Data

Target Label | Prototypes
---|---
COMPANY | company, unit, firm
PERSON | president, director, official
FINANCIAL CURRENCY | dollar, yen, euro
LOCATION | New York, Paris, London
FACILITY | hospital, airport, university

[Haghigi & Klein, NAACL ’06]

Evaluation

MUC F₁ - Cluster Similarity

Unsupervised Supervised Our Work, '07 [Denis et al., '07] [Bengston & Roth, '08] Our Work, Current

50 60.1 64.2 67.2 76.8 80.8

85

Unsupervised Supervised Our Work, '07 [Denis et al., '07] [Bengston & Roth, '08] Our Work, Current