Illinois-Profiler: Knowledge Schemas at Scale

Zhiye Fei, Daniel Khashabi,
Haoruo Peng, Hao Wu & Dan Roth

Department of Computer Science
University of Illinois at Urbana-Champaign
Comprehension

(ENGLAND, June, 1989) - Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don’t know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.
Comprehension

(ENGLAND, June, 1989) – Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don’t know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

1. Christopher Robin was born in England.  
2. Winnie the Pooh is a title of a book.  
3. Christopher Robin’s dad was a magician.  
4. Christopher Robin must be at least 65 now.
Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don’t know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

1. Christopher Robin was born in England.  
2. Winnie the Pooh is a title of a book.  
3. Christopher Robin’s dad was a magician.  
4. Christopher Robin must be at least 65 now.
Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don’t know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

1. Christopher Robin was born in England. 2. Winnie the Pooh is a title of a book. 3. Christopher Robin’s dad was a magician. 4. Christopher Robin must be at least 65 now.
Comprehension

(ENGLAND, June, 1989) - (Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don’t know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

1. Christopher Robin was born in England.  
2. Winnie the Pooh is a title of a book.  
3. Christopher Robin’s dad was a magician.  
4. Christopher Robin must be at least 65 now.
Comprehension

(ENGLAND, June, 1989) - Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don’t know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

1. Christopher Robin was born in England.
2. Winnie the Pooh is a title of a book.
3. Christopher Robin’s dad was a magician.
4. Christopher Robin must be at least 65 now.
(ENGLAND, June, 1989) - Christopher Robin is alive and well. He lives in England. He is the same person that you read about in the book, Winnie the Pooh. As a boy, Chris lived in a pretty home called Cotchfield Farm. When Chris was three years old, his father wrote a poem about him. The poem was printed in a magazine for others to read. Mr. Robin then wrote a book. He made up a fairy tale land where Chris lived. His friends were animals. There was a bear called Winnie the Pooh. There was also an owl and a young pig, called a piglet. All the animals were stuffed toys that Chris owned. Mr. Robin made them come to life with his words. The places in the story were all near Cotchfield Farm. Winnie the Pooh was written in 1925. Children still love to read about Christopher Robin and his animal friends. Most people don't know he is a real person who is grown now. He has written two books of his own. They tell what it is like to be famous.

1. Christopher Robin was born in England.  
2. Winnie the Pooh is a title of a book.  
3. Christopher Robin's dad was a magician.  
4. Christopher Robin must be at least 65 now.  

This is an Inference Problem
What is being Repaired?

- The ball hit the window and Bill repaired it.
What is being Repaired?

- The ball hit the window and Bill repaired it.
What is being Repaired?

- The ball hit the window and Bill repaired it.
What is being Repaired?

- The ball hit the window and Bill repaired it.
What is being Repaired?

- The \textbf{ball} hit the \textbf{window} and Bill repaired \textbf{it}.

PERSON repaired \textbf{window} vs
PERSON repaired \textbf{ball}
What is being Repaired?

The ball hit the window and Bill repaired it.

PERSON repaired window
vs
PERSON repaired ball
What is being Repaired?

- The ball hit the window and Bill repaired it.

PERSON repaired window  
vs
PERSON repaired ball

N1
\textit{word(*)}

N2
\textit{word("repaired")}

N3
\textit{word(w)}
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant

Jimbo stole an elephant
vs.
Robert stole an elephant
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant

Jimbo stole an elephant

vs.

Robert stole an elephant

word(w)  word(“stole”)  word(“an elephant”)
Structured Knowledge

Jimbo arrested Robert because he stole an elephant

Jimbo stole an elephant vs. Robert stole an elephant
Structured Knowledge

- **Jimbo** arrested **Robert** because he stole an elephant

Diagram:

1. Jimbo stole an elephant
2. vs.
3. Robert stole an elephant

* arrested **SOMEONE** because **SOMEONE** stole *

vs.

**SOMEONE** arrested * because **SOMEONE** stole *
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant

* arrested SOMEONE because SOMEONE stole *

vs.

SOMEONE arrested * because SOMEONE stole *
Structured Knowledge

- **Jimbo** arrested **Robert** because **he** stole an **elephant**

```
<table>
<thead>
<tr>
<th>N1</th>
<th>N2</th>
<th>N3</th>
<th>N4</th>
</tr>
</thead>
<tbody>
<tr>
<td>word(“arrest”)</td>
<td>word(*)</td>
<td>word(*)</td>
<td>word(“steal”)</td>
</tr>
<tr>
<td>objOf/subjOf?</td>
<td>Co-referred</td>
<td>subjOf</td>
<td></td>
</tr>
</tbody>
</table>
```

* arrested **SOMEONE** because **SOMEONE** stole *

vs.

**SOMEONE** arrested * because **SOMEONE** stole *
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant

* arrested SOMEONE because SOMEONE stole *

vs.

SOMEONE arrested * because SOMEONE stole *
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant

The subject of “stole” is more likely to be

* arrested SOMEONE because SOMEONE stole *

vs.

SOMEONE arrested * because SOMEONE stole *

* arrested SOMEONE because SOMEONE stole *
Structured Knowledge

- Jimbo arrested Robert because he stole an elephant

The subject of “stole” is more likely to be the object of “arrest” then the subject of “arrest”.

* arrested SOMEONE because SOMEONE stole * vs. SOMEONE arrested * because SOMEONE stole *
Knowledge for Many Tasks

- [Larry Robbins], founder of Glenview Capital Management, bought shares of [Endo International Plc] ...”
Knowledge for Many Tasks

- [Larry Robbins], founder of Glenview Capital Management, bought shares of [Endo International Plc] ...”

NER TAGS = {PERSON, LOCATION, ORGANIZATION, ...}
Knowledge for Many Tasks

- [Larry Robbins], founder of Glenview Capital Management, bought shares of [Endo International Plc] ...

NER TAGS = {PERSON, LOCATION, ORGANIZATION, ...}
Knowledge for Many Tasks

- [Larry Robbins], founder of Glenview Capital Management, bought shares of [Endo International Plc] ...

NER TAGS = {PERSON, LOCATION, ORGANIZATION, ...}

- [Seattle] fired John Doe after the team lost its 7th game in a row.
Knowledge for Many Tasks

- **[Larry Robbins]**, founder of Glenview Capital Management, bought shares of **[Endo International Plc]** ...

  NER TAGS = {PERSON, LOCATION, ORGANIZATION, ...}

- **[Seattle]** fired John Doe after the team lost its 7\textsuperscript{th} game in a row.

  NER TAGS = {PERSON, LOCATION, ORGANIZATION, ...}

- Organization is more likely than a location to be the subject of “fire”.

Knowledge is Essential
Knowledge is Essential

- Textual Inference requires additional knowledge.
  - More than just local features: need to know things.
Knowledge is Essential

- Textual Inference requires additional knowledge.
  - More than just local features: need to know things.
- There is a need to think about how to make use of this knowledge in textual inference.
  - Not here; see my talk at the Neuro-Symbolic Workshop
Knowledge is Essential

- Textual Inference requires additional knowledge.
  - More than just local features: need to know things.
- There is a need to think about how to make use of this knowledge in textual inference.
  - Not here; see my talk at the Neuro-Symbolic Workshop
- This work:
  - Graph-based formulation for modelling knowledge schemas
    - The necessity of disambiguation
    - The acquisition process
  - Profiler as a public resource
    - Contains pre-computed statistics
    - Many concepts/entities; many knowledge schemas
Terminology

- **Schema**: A way to define a structure and, consequently, semantics, specified by a template.
Terminology

- **Schema**: A way to define a structure and, consequently, semantics, specified by a template.
  - We use it Graphically to define pieces of knowledge

```
N1 → after → N2
  word("Obama", "Barack Obama") → POS(VB)

N1 → after → N2
  word("Obama", "Barack Obama") → POS(VB)

N3 → before → N1 → after → N2
  word("Obama", "Barack Obama") → POS(VB)
```
Terminology

- **Schema**: A way to define a structure and, consequently, semantics, specified by a template.
  - We use it Graphically to define pieces of knowledge
- **Pivot**: a key node in each schema
Terminology

- **Schema**: A way to define a structure and, consequently, semantics, specified by a template.
  - We use it Graphically to define pieces of knowledge
- **Pivot**: a key node in each schema

Profile: Entity: Obama (President)
Terminology

- **Schema**: A way to define a structure and, consequently, semantics, specified by a template.
  - We use it Graphically to define pieces of knowledge
- **Pivot**: a key node in each schema
- **Profile(pivot)**: a set of schemas with a common pivot

Profile: Entity: Obama (President)
**Terminology**

- **Schema**: A way to define a structure and, consequently, semantics, specified by a template.
  - We use it Graphically to define pieces of knowledge
- **Pivot**: a key node in each schema
- **Profile(pivot)**: a set of schemas with a common pivot
  - Instantiated schema, with statistics
Disambiguation is Important
Disambiguation is Important

The airport is located south of Seattle.

Seattle played well today!
Disambiguation is Important

The airport is located south of Seattle.

Seattle played well today!
Disambiguation is Important

The airport is located south of Seattle.

Seattle played well today!
Disambiguation is Important

The airport is located south of Seattle.

Seattle played well today!
Disambiguation is Important

The airport is located south of Seattle.

Seattle played well today!
Disambiguation is Important

The airport is located south of Seattle.

Seattle played well today!

A pivot is a pair: (mention, Wikipedia URL)
Knowledge Schema as a Graph

- Knowledge as graph:
  - Generalization of Feature Description Logic (Cumby & Roth, ‘02,’03)
  - Assumes a structured (relational) representation of the data.

- Definition:
  - **Attribute:** type of value on each node \( \mathcal{A} = \{a_1, a_2, a_3, \ldots\} \)
    - In the example: \( \mathcal{A} = \{\text{POS, raw – text}\} \)
  - **Values:** possible values each attribute take \( \mathcal{V} = \{v_1, v_2, v_3, \ldots\} \)
    - In the example, the values of POS are \( \mathcal{V} = \{\text{N, VP, NP, ...}\} \)
  - **Roles** (Relation): connection between nodes \( \mathcal{R} = \{r_1, r_2, r_3, \ldots\} \)
    - In the example, we have \( \mathcal{R} = \{\text{subjOf, objOf, ...}\} \)
Knowledge Schema Descriptions

- A Description:
  - A schema (template) defined in the FDL language, which corresponds to a set of grounded elements matching the definition.
Knowledge Schema Descriptions

- A Description:
  - A schema (template) defined in the FDL language, which corresponds to a set of grounded elements matching the definition.

- Description of a schema graph
  - The set of all instances matching the schema graph.
  - Descriptions are defined Recursively
Knowledge Schema Descriptions (2)

- **A Description:**
  - A schema (template) defined in the FDL language, which corresponds to a set of grounded elements matching the definition.

- **Description of a schema graph**
  - The set of all instances matching the schema graph.
  - Descriptions are defined Recursively

- **Basic rules:**
  - For an attribute $a \in \mathcal{A}$ and a value $v \in \mathcal{V}$, $a(v)$ is a description, and it represents the set of $x \in \mathcal{X}$ for which $a(x, v)$ is True.
  - Example:
    
    $\text{word(“defeat”)}$
    
    \[ \text{POS(N)} \]
Knowledge Schema Descriptions (3)

- Basic rules (continued):
  - For a description $D$ and a role $r \in \mathcal{R}$, $(r \ D)$ is a description. Such description represents the set $x \in \mathcal{X}$ such that $r(x, y)$ is True, where $y \in \mathcal{X}$ is described by $D$.

\[
\begin{align*}
\text{(subjectOf word("defeat"))} \\
\text{(objectOf word("defeat"))}
\end{align*}
\]

- For given descriptions $D^{(1)}, \ldots, D^{(k)}$ then $(\text{AND } D^{(1)}, \ldots, D^{(k)})$ is a description, which represents a conjunction of all elements described by each description.

\[
\begin{align*}
\text{(AND (POS(N))(subjectOf word("defeat")))} \\
\text{(AND (POS(N))(objectOf word("defeat")))}
\end{align*}
\]
Knowledge Schema Descriptions (4)

- Basic rules (continued):
  - Denote the description of node $i$ with $D_i$.
  - Denote the description of nodes $i_1, \ldots, i_k$ with $D_{i_1, \ldots, i_k}$.
  - The description of the whole graph can be found with:
    \[
    D_{1,2,3} = D_1 \otimes D_2 \otimes D_3
    \]
  - This is the set of all instances matching the pattern defined by the schema graph.

\[
D_1 = (\text{AND} (\text{POS}(N)) (\text{subjOf} \ \text{word}(\text{"defeat"})))
\]

\[
D_2 = \{\text{word}(\text{"defeat"})\}
\]

\[
D_3 = (\text{AND} (\text{POS}(N)) (\text{objOf} \ \text{word}(\text{"defeat"})))
\]
Knowledge Schemas

- The definition can be generalized for any graph.
  - See the general definition in the paper.
Knowledge Schemas

- The definition can be generalized for any graph.
  - See the general definition in the paper.

- Why important?
  - A way of formalizing general knowledge over relational structures
  - A systematic way to represent and acquire knowledge
  - Compatible with functional programming languages
Knowledge Schemas

- The definition can be generalized for any graph.
  - See the general definition in the paper.

- Why important?
  - A way of formalizing general knowledge over relational structures
  - A systematic way to represent and acquire knowledge
  - Compatible with functional programming languages

- See a talk on:
  **Saul: Towards Declarative Learning Based Programming**
  Parisa Kordjamshidi, Hao Wu, Dan Roth
  Presented on Tuesday, July 28; 9:40 Relational Learning Session
The Acquisition Procedure
The Acquisition Procedure

4 million
The Acquisition Procedure

4 million

NLP
The Acquisition Procedure

4 million → 1,455 GB

Cognitive Computation Group
The Acquisition Procedure

Illinois CloudNLP: a suite of state-of-the-art NLP tools. Made available also on AWS.
The Acquisition Procedure

- 4 million documents
- 1,455 GB of data
- Illinois CloudNLP: a suite of state-of-the-art NLP tools. Made available also on AWS.
The Acquisition Procedure

- 4 million
- 1,455 GB

Illinois CloudNLP: a suite of state-of-the-art NLP tools. Made available also on AWS.

200 mid-end EC2 nodes,
3 hours, 420$
The Acquisition Procedure

4 million documents → NLP → 1,455 GB data → S3

200 mid-end EC2 nodes, 3 hours, 420$ to prepare 200GB

Illinois CloudNLP: a suite of state-of-the-art NLP tools. Made available also on AWS.
The Acquisition Procedure

4 million documents are processed through Illinois CloudNLP, a suite of state-of-the-art NLP tools. The processed data is then stored in S3. Each node requires 200GB of storage and the entire process takes 3 hours at a cost of 420$. The processed data is then made available on AWS.
The Acquisition Procedure

Illinois CloudNLP: a suite of state-of-the-art NLP tools. Made available also on AWS.

4 million documents → 1,455 GB → S3 → Profiles

Profiles:
- ~3.5 M Wiki profiles
- ~300,000 verb sense profiles

EMR → hadoop → S3 → 200 mid-end EC2 nodes, 3 hours, 420$

Try our demo:
http://cogcomp.cs.illinois.edu/profile
Knowledge Schema as a Graph

- The annotations used in the current system:
  - Attribute & values; Roles

<table>
<thead>
<tr>
<th>Attributes (𝒜)</th>
<th>Values (𝑉)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Raw text</td>
</tr>
<tr>
<td>Lemma</td>
<td>Raw text</td>
</tr>
<tr>
<td>POS</td>
<td>labels form Penn Treebank</td>
</tr>
<tr>
<td>NER</td>
<td>{ PER, ORG, LOC, MISC }</td>
</tr>
<tr>
<td>Wikifier</td>
<td>Wikipedia urls</td>
</tr>
<tr>
<td>Verbsense</td>
<td>Verb sense from Verbnet</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Roles (ℛ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
</tr>
<tr>
<td>After</td>
</tr>
<tr>
<td>NearestBefore</td>
</tr>
<tr>
<td>NearestAfter</td>
</tr>
<tr>
<td>AdjacentToBefore</td>
</tr>
<tr>
<td>AdjacentToAfter</td>
</tr>
<tr>
<td>ExclusiveContaining</td>
</tr>
<tr>
<td>HasOverlap</td>
</tr>
<tr>
<td>DependencyPath(l)</td>
</tr>
<tr>
<td>Co-referred</td>
</tr>
<tr>
<td>SubjectOf</td>
</tr>
<tr>
<td>ObjectOf</td>
</tr>
</tbody>
</table>
Knowledge Schema as a Graph

The annotations used in the current system:

- Attribute & values; Roles

<table>
<thead>
<tr>
<th>Attributes (𝒜)</th>
<th>Values (𝑉)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Raw text</td>
</tr>
<tr>
<td>Lemma</td>
<td>Raw text</td>
</tr>
<tr>
<td>POS</td>
<td>labels form Penn Treebank</td>
</tr>
<tr>
<td>NER</td>
<td>{ PER, ORG, LOC, MISC }</td>
</tr>
<tr>
<td>Wikifier</td>
<td>Wikipedia urls</td>
</tr>
<tr>
<td>Verbsense</td>
<td>Verb sense from Verbnet</td>
</tr>
</tbody>
</table>

Roles (𝑅)

- Before
- After
- NearestBefore
- NearestAfter
- AdjacentToBefore
- AdjacentToAfter
- ExclusiveContaining
- HasOverlap
- DependencyPath(l)
- Co-referred
- SubjectOf
- ObjectOf
Blumenthal (D) is a candidate for the U.S. Senate seat now held by Christopher Dodd (D), and he has held a commanding lead in the race since he entered it. But the Times report has the potential to fundamentally reshape the contest in the Nutmeg State.
Blumenthal (D) is a candidate for the U.S. Senate seat now held by Christopher Dodd (D), and he has held a commanding lead in the race since he entered it. But the Times report has the potential to fundamentally reshape the contest in the Nutmeg State.
Who is Alex Smith?

The Chiefs didn't trade for Alex Smith this **offseason** solely because they wanted a smart game manager who wouldn’t kill their offense with turnovers. They acquired him because they needed a quarterback who knows how to win. Sometimes that requires him to do what he’s done for most of this season: throw the safe pass, make the key play when necessary and use his feet to keep the chains moving when his arm can’t get the job done. These days it means **Smith** has to show people more of what he revealed in Sunday’s 41-38 loss to San Diego -- that he can elevate his game when his team is in dire straits.

From the sounds of it, **Bengals** tight end Alex Smith might be gone for the season. The **veteran tight end** suffered a wrist injury in the third quarter during the regular season finale against Baltimore. Bengal**s head coach Marvin Lewis described the injury as a “wrist dislocation”, also said during the postgame radio interview on 700 WLW with Dave Lapham that “It looks like we lost” Smith, all but confirming an eventualty. More will be known this week, so hold off on declaring him done. On the other hand, Lewis confirmed that the Bengals should have tight ends **Tyler Eifert** and Jermaine Gresham for the next game. Both were out Sunday against the **Ravens**.
Who is Alex Smith?

The Chiefs traded for Alex Smith during the season solely because they wanted a smart game manager who wouldn't kill their offense with turnovers. They acquired him because they needed a quarterback who knows how to win. Sometimes that requires him to do what he's done for most of this season: throw the safe pass, make the key play when necessary and use his feet to get the chains moving when his arm can't get the job done. These days it's the little things people more of what he revealed in Sunday's 41-38 loss to San Diego -- that he can elevate his game when his team is in dire straits.

From the sounds of it, Bengals tight end Alex Smith might be gone for the season. The veteran tight end suffered a wrist injury in the third quarter during the regular season finale against Baltimore. Bengals head coach Marvin Lewis described the injury as a "wrist dislocation," also said during the postgame radio interview on 700 WLW with Dave Lapham that "It looks like we lost" Smith, all but confirming the eventuality. More will be known this week, so hold off on declaring him done. On the other hand, Lewis confirmed that the Bengals should have tight ends Tyler Eifert and Jermaine Gresham for the next game. Both were out Sunday against the Ravens.
Who is Alex Smith?

The Chiefs signed Alex Smith to a two-year, $5 million deal with $1 million guaranteed, making the move in his fifth year as a free agent. They acquired him because they needed a quarterback who knows how to win. Sometimes that requires him to do what he's done for most of this season: throw the safe pass, make the key play when necessary and use his feet to keep the chains moving when his arm can't get the job done.

These days it appears people more of what he revealed in Sunday's 41-38 loss to San Diego -- that he can elevate his game when his team is in dire straits.

From the sounds of it, Bengals tight end Jermaine Gresham is in for the season. The veteran tight end suffered a wrist injury in the third quarter during the regular season finale against Baltimore. Bengals head coach Marvin Lewis described the injury as a "wrist dislocation," also said during the post-game interview on 700 WLW with Dave Lapham that "It looks like we're dealing with an eventuality. More will be known this week, so hold off on declaring him done." On the other hand, Lewis confirmed that the Bengals should have tight ends Tyler Eifert and Jermaine Gresham for the next game. Both were out Sunday against the Ravens.
Who is Alex Smith?

Alex Smith

Quarterback of the Kansas City Chief

Alex Smith

Smith

Quarterback of the Kansas City Chief
Who is Alex Smith?

Quarterback of the Kansas City Chiefs

Tight End of the Cincinnati Bengals
Who is Alex Smith?

Alex Smith
Quarterback of the Kansas City Chiefs
Tight End of the Cincinnati Bengals
San Diego: The San Diego Chargers (A Football team)
Who is Alex Smith?

Alex Smith

Quarterback of the Kansas City Chiefs

Tight End of the Cincinnati Bengals

San Diego: The San Diego Chargers (A Football team)

Ravens: The Baltimore Ravens (A Football team)
Who is Alex Smith?

Alex Smith

Quarterback of the Kansas City Chiefs

Tight End of the Cincinnati Bengals

San Diego: The San Diego Chargers (A Football team)

Ravens: The Baltimore Ravens (A Football team)

Contextual Disambiguation
Middle Eastern Politics

Over and over again I’d heard these perorations from certain Jewish circles arguing that there is no difference between Fatah and Hamas, or between Mahmoud Abbas and Khaled Maashal. I would cringe at such comments, while knowing full well that Abu Mazen was hardly the perfect interlocutor. I’m a strong believer in identifying the threats to Israel without pulling any punches. But I also believe that it is important to give peace a chance, to search for signs that the Palestinians are open to change from the destructive and self-destructive path they have pursued for decades. Hamas was and is a hopeless proposition. It not only rejects Israel’s existence on extremist religious grounds but it is anti-Semitic to the extreme. Its charter sounds like the "Protocols of the Learned Elders of Zion," blaming Jews for all the world’s ills since the French Revolution. Its leaders have denied the Holocaust and blamed the financial crisis on Jewish control.
Middle Eastern Politics

Over and over again I'd hear these perorations from certain sources that there is no difference between Fatah and Hamas, or between Mahmoud Abbas and Khaled Maashal. I would cringe at such comments, while knowing full well that the Palestinians are open to change from the destructive and self-destructive path they have pursued for decades. Hamas was and is a hopeless proposition. It not only rejects Israel's existence on extremist religious grounds but it is anti-Semitic to the extreme. Its charter sounds like the "Protocols of the Learned Elders of Zion," blaming Jews for all the world's ills since the French Revolution. Its leaders have denied the Holocaust and blamed the financial crisis on Jewish control.
Middle Eastern Politics

Mahmoud Abbas: http://en.wikipedia.org/wiki/Mahmoud_Abbas

Over and over again I'd hear these perorations from my interlocutor. I'm a strong believer in ideas without pulling any punches. But I also believe that it is important to give peace a chance, to search for signs that the Palestinians are open to change from the destructive and self-destructive paths they have pursued for decades. Hamas was and is a hopeless proposition. It is on extremist religious grounds but it is anti-Semitic to blame the "Protocols of the Learned Elders of Zion," blaming the French Revolution. Its leaders have denied the Holocaust and blamed the financial crisis on Jewish control.
Over and over again I’d hear these perorations from my interlocutor. I’m a strong believer in idea without pulling any punches. But I also believe that it is important to give peace a chance, to search for signs that the Palestinians are open to change from the destructive and self-destructive paths they have pursued for decades. Hamas was and is a hopeless proposition. Founded on extremist religious grounds by anti-Semitic to blame the "Protocols of the Learned Elders of Zion," blaming the French Revolution. Its leader denied the Holocaust and blamed the financial crisis on Jewish control.
Middle Eastern Politics


Mahmoud Abbas:
http://en.wikipedia.org/wiki/Mahmoud_Abbas

Abu Mazen:
http://en.wikipedia.org/wiki/Mahmoud_Abbas
The Profiler DB

- Each entry corresponds to a disambiguated entity/Concept
- Mapping to Wikipedia grounds entities in the “world” and allow us to profile unique entities, rather than “mentions” in text.

- In particular, we have distinct entries for:
- Clinton (Bill), Clinton (Hilary), Clinton (lake), Clinton (Illinois),.....
The Acquisition Procedure

4 million

1,455 GB

Profiles

~3,5 M Wiki profiles
~300,000 verb sense profiles

200 mid-end EC2 nodes,
3 hours, 420$

Illinois CloudNLP: a suite of state-of-the art NLP tools. Made available also on AWS.

Try our demo:

http://cogcomp.cs.illinois.edu/profiler
Experimental Evidence

- We are at early stages of experimental validation (and refinement) of the acquisition and inference with the profiler.

- Co-reference Resolution

- Identifying Attributes of Entities
  - Profession
Experiment 1: Co-reference Resolution
Experiment 1: Co-reference Resolution

- We build upon our previous work (Peng et al, 2015).
Experiment 1: Co-reference Resolution

- We build upon our previous work (Peng et al, 2015).
- Extended the Winograd data (Rahman & Ng)
Experiment 1: Co-reference Resolution

- We build upon our previous work (Peng et al, 2015).
- Extended the Winograd data (Rahman & Ng)
  - to general co-reference instances: WinoCoref data
Experiment 1: Co-reference Resolution

- We build upon our previous work (Peng et al, 2015).
- Extended the Winograd data (Rahman & Ng)
  - to general co-reference instances: WinoCoref data

Jack threw the bags of John into the water since he mistakenly asked him to carry his bags.
Experiment 1: Co-reference Resolution

- We build upon our previous work (Peng et al, 2015).
- Extended the Winograd data (Rahman & Ng)
  - to general co-reference instances: WinoCoref data

Schemas are converted automatically, given an instance, into constraints that are used in an Integer Linear Programming formulation.
Experiment 1: Co-reference Resolution

- **Metrics:**
  - Precision for Winograd dataset
  - AntePre for WinoCored dataset:
    - Consider all the binary decisions of connecting pronouns to nominal mentions
    - AntePre is the ratio of correct binary decisions to the total decisions

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Winograd</th>
<th>WinoCoref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metric</td>
<td>Precision</td>
<td>AntePre</td>
</tr>
<tr>
<td>(Rahman &amp; Ng, 2012)</td>
<td>73.05</td>
<td>----</td>
</tr>
<tr>
<td>(Peng et al, 2015)</td>
<td>76.41</td>
<td>89.32</td>
</tr>
<tr>
<td>Our paper</td>
<td>77.16</td>
<td>89.77</td>
</tr>
</tbody>
</table>
Experiment 2: Classifying occupations

- Observations: Profiles of people contain information about their occupation.

<table>
<thead>
<tr>
<th>Tom Brady (football player)</th>
<th>Nikola Tesla (Inventor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>say 0.254</td>
<td>say 0.224</td>
</tr>
<tr>
<td>throw 0.16</td>
<td>develop 0.104</td>
</tr>
<tr>
<td>pass 0.06</td>
<td>sell 0.06</td>
</tr>
<tr>
<td>man 0.052</td>
<td>buy 0.06</td>
</tr>
<tr>
<td>play 0.048</td>
<td>invent 0.06</td>
</tr>
<tr>
<td>go 0.045</td>
<td>build 0.06</td>
</tr>
<tr>
<td>take 0.037</td>
<td>continue 0.045</td>
</tr>
<tr>
<td>spend 0.037</td>
<td>upgrade 0.03</td>
</tr>
<tr>
<td>look 0.034</td>
<td>play 0.03</td>
</tr>
<tr>
<td>win 0.034</td>
<td>help 0.03</td>
</tr>
<tr>
<td>spot 0.034</td>
<td>magnify 0.03</td>
</tr>
<tr>
<td>complete 0.031</td>
<td>suppress 0.03</td>
</tr>
<tr>
<td>order 0.029</td>
<td>alternate 0.03</td>
</tr>
<tr>
<td>come 0.028</td>
<td>hire 0.03</td>
</tr>
<tr>
<td>start 0.023</td>
<td>remove 0.03</td>
</tr>
<tr>
<td>lead 0.021</td>
<td>intend 0.03</td>
</tr>
<tr>
<td>know 0.021</td>
<td>waste 0.03</td>
</tr>
<tr>
<td>launch 0.017</td>
<td>seek 0.03</td>
</tr>
</tbody>
</table>
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia
- Steps:
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia

Steps:

- Pick a bunch of profiles (of people) with known jobs
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia

Steps:
- Pick a bunch of profiles (of people) with known jobs
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia
- Steps:
  - Pick a bunch of profiles (of people) with known jobs
  - Merge the profiles of the ones with the same job (occupation profiles)
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia

Steps:

- Pick a bunch of profiles (of people) with known jobs
- Merge the profiles of the ones with the same job (occupation profiles)
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia
- Steps:
  - Pick a bunch of profiles (of people) with known jobs
  - Merge the profiles of the ones with the same job (occupation profiles)
  - Given a new profile, decide which occupation profile it is closest:

John F. Kennedy
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia

Steps:

- Pick a bunch of profiles (of people) with known jobs
- Merge the profiles of the ones with the same job (occupation profiles)
- Given a new profile, decide which occupation profile it is closest:

John F. Kennedy
Experiment 2: Classifying occupations

- Created a dataset of People-Profession based on Wikipedia

**Steps:**

- Pick a bunch of profiles (of people) with known jobs
- Merge the profiles of the ones with the same job (occupation profiles)
- Given a new profile, decide which occupation profile it is closest:

72.1% of the test cases, the correct answer is among the top-5

John F. Kennedy
Future Directions

- Extensions of the Profiler
  - Richer set of schemas
  - Richer annotations
  - More data
  - Incorporating the profiler as a part of feature extraction system, within a learning framework
    - Profiler, beyond a resource, but as a tool to engineer knowledge.

- Inference
  - How to best use the profiles

- Experiments
  - Different tasks need to be explored.
Future Directions

- Extensions of the Profiler
  - Richer set of schemas
  - Richer annotations
  - More data
  - Incorporating the profiler as a part of feature extraction system, within a learning framework
    - Profiler, beyond a resource, but as a tool to engineer knowledge.

- Inference
  - How to best use the profiles

- Experiments
  - Different tasks need to be explored.
*Haoruo Peng*, *Daniel Khashabi* and *Dan Roth*. NAACL 2015.
