A Markov Logic Approach to Bio-Molecular Event Extraction

Sebastian Riedel, Hong-Woo Chun, Toshihisa Takagi, Jun'ichi Tsujii
Task 1

- Find event **clues** and **arguments**

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
Task 2

- Find and attach cellular locations

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
Observations

- At least one Theme

Phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain.
Observations

- Regulation transitively involves proteins

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

- Site2 arguments require Theme2
Approach
Approach

• Due to **global interactions**: 
Approach

• Due to **global interactions**:
  • Learn distribution over **full event structures**
Approach

• Due to *global interactions*:
  • Learn distribution over *full event structures*
  • Joint inference is *difficult*
Approach

• Due to **global interactions**:
  • Learn distribution over **full event structures**
  • Joint inference is **difficult**

• Use **Markov Logic** & interpreter (Richardson & Domingos, 2006)
Approach

- Due to **global interactions:**
  - Learn distribution over **full event structures**
- Joint inference is **difficult**
- Use **Markov Logic** & interpreter (Richardson & Domingos, 2006)
- Markov Logic **likes small domains**
Approach

- Due to global interactions:
  - Learn distribution over full event structures
- Joint inference is difficult
- Use Markov Logic & interpreter (Richardson & Domingos, 2006)
- Markov Logic likes small domains
- Map to link structure over tokens
Event Prediction

- Event structure where events are entities

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

- Event structure where events are entities
phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
Link Prediction

- Events are projected on their clues

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
Link Prediction

- Events are projected on their clues
- Close to **Semantic Role Labelling** : cf Meza-Ruiz & Riedel, 2009)
Markov Logic

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
Markov Logic

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

- Describe solution with **predicates** and **possible worlds**:
  \{ event(1, Phosph.), role(1, 3, Theme), site(9)... \}
Phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain.

- Describe solution with **predicates** and **possible worlds**:
  \{\text{event}(1, \text{Phosph.}), \text{role}(1, 3, \text{Theme}), \text{site}(9)\ldots\}\
Markov Logic

- Describe solution with **predicates** and **possible worlds**:
  \{\text{event}(1, \text{Phosph.}), \text{role}(1, 3, \text{Theme}), \text{site}(9)\ldots\}
Markov Logic

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

- Describe solution with predicates and possible worlds:
  \{event(1,Phosph.), role(1,3,Theme), site(9)\}
<table>
<thead>
<tr>
<th>Protein</th>
<th>Event Clue</th>
<th>Location</th>
</tr>
</thead>
</table>

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
Clue Detection

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

- Distinguish between good and bad worlds
  \( \text{word}(i, \text{inhibits}) \Rightarrow \text{event}(i, \text{Neg. Regulation}) <1.2> \)
phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain.
Compaction of Representation

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

- Compact representation:
  \[ \text{word}(i, w) \Rightarrow \text{event}(i, t) \ < \text{weight}(w, t) > \]
Argument Extraction

phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

• Subjects are themes:
  \[ dep(i,j,subj) \Rightarrow role(i,j,Theme) <1.23> \]
phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain
Global Correlations

Phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain

• There need to be Themes:
  \( \text{event}(e,t) \Rightarrow \exists a. \text{role}(e,a,\text{Theme}) <999.9> \)
Joint Distribution
Joint Distribution

- **Ground Feature**: e.g.

\[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]
Joint Distribution

• **Ground Feature**: e.g.

\[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]

\[
f_{\phi,j/4,i/1}(y) = \begin{cases} 
1 & \text{if } \text{dep}(4,1,\text{subj}) \Rightarrow \text{role}(4,1,\text{Theme}) \\
0 & \text{otherwise}
\end{cases}
\]
Joint Distribution

- **Ground Feature**: e.g.

\[ \text{dep}(i, j, \text{subj}) \Rightarrow \text{role}(i, j, \text{Theme}) \]

\[ f_{j/4, i/1}^\phi (y) = \begin{cases} 
1 & \text{if } \text{dep} (4, 1, \text{subj}) \Rightarrow \text{role} (4, 1, \text{Theme}) \\
0 & \text{otherwise}
\end{cases} \]

[substitute free variables wrt. binding]
Joint Distribution

- **Ground Feature**: e.g.

\[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]

\[
f_{j/4,i/1}^{\phi}(y) = \begin{cases} 
1 & \text{if } \text{dep}(4,1,\text{subj}) \Rightarrow \text{role}(4,1,\text{Theme}) \\
0 & \text{otherwise}
\end{cases}
\]

- Weighted formulae **loglinear distribution**:

\[ \text{substitute free variables wrt. binding} \]
Joint Distribution

- **Ground Feature**: e.g.

\[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]

\[ f_{j/4,i/1}^\phi(y) = \begin{cases} 
1 & \text{if } \text{dep}(4,1,\text{subj}) \Rightarrow \text{role}(4,1,\text{Theme}) \\
0 & \text{otherwise}
\end{cases} \]

- **Weighted formulae** loglinear distribution:

\[ p(y) = \frac{1}{Z} \exp \left( \sum_{(\phi,w) \in M} w \sum_{e \in C^\phi} f_e^\phi(y) \right) \]
Joint Distribution

- **Ground Feature**: e.g.

  \[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]

  \[ f_{j/4,i/1}^{\phi}(y) = \begin{cases} 1 & \text{if } \text{dep}(4,1,\text{subj}) \Rightarrow \text{role}(4,1,\text{Theme}) \\ 0 & \text{otherwise} \end{cases} \]

  substitute free variables wrt. binding

- **Weighted formulae loglinear distribution**:

  \[ p(y) = \frac{1}{Z} \exp \left( \sum_{(\phi,w) \in M} w \sum_{c \in C^\phi} f_c^{\phi}(y) \right) \]

  possible world
Joint Distribution

- **Ground Feature**: e.g.

  \[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]

  \[ f_{j/4,i/1}^\phi(y) = \begin{cases} 
  1 & \text{if } \text{dep}(4,1,\text{subj}) \Rightarrow \text{role}(4,1,\text{Theme}) \\
  0 & \text{otherwise} 
\end{cases} \]

- **Weighted formulae loglinear distribution**:

  \[ p(y) = \frac{1}{Z} \exp \left( \sum_{(\phi,w) \in M} w \sum_{c \in C^\phi} f_c^\phi(y) \right) \]
Joint Distribution

• **Ground Feature**: e.g.

\[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]

\[ f_{j/4,i/1}^\phi (y) = \begin{cases} 
1 & \text{if dep}(4, 1, \text{subj}) \Rightarrow \text{role}(4, 1, \text{Theme}) \\
0 & \text{otherwise} \end{cases} \]

• Weighted formulae **loglinear distribution**:

\[
 p(y) = \frac{1}{Z} \exp \left( \sum_{(\phi,w) \in M} w \sum_{c \in C^\phi} f_c^\phi (y) \right)
\]

possible world formula weight
Joint Distribution

- **Ground Feature**: e.g.

\[
dep(i,j,\text{subj}) \Rightarrow role(i,j,\text{Theme})
\]

\[
f_{j/4,i/1}^\phi (y) = \begin{cases} 
1 & \text{if } dep(4,1,\text{subj}) \Rightarrow role(4,1,\text{Theme}) \\
0 & \text{otherwise}
\end{cases}
\]

- **Weighted formulae loglinear distribution**:

\[
p(y) = \frac{1}{Z} \exp \left( \sum_{(\phi,w) \in M} w \sum_{c \in C^\phi} f_c^\phi (y) \right)
\]
Joint Distribution

- **Ground Feature**: e.g.

\[ \text{dep}(i,j,\text{subj}) \Rightarrow \text{role}(i,j,\text{Theme}) \]

\[ f_{j/4,i/1}^\phi(y) = \begin{cases} 
1 & \text{if } \text{dep}(4,1,\text{subj}) \Rightarrow \text{role}(4,1,\text{Theme}) \\
0 & \text{otherwise}
\end{cases} \]

- **Weighted formulae loglinear distribution**:

\[ p(y) = \frac{1}{Z} \exp \left( \sum_{(\phi,w) \in M} w \sum_{c \in C^\phi} f_c^\phi(y) \right) \]
Learning and Inference
Learning and Inference

- How do we find most likely structure?
Learning and Inference

• How do we find most likely structure?
• Cutting Plane Inference+ILP (Riedel, 2008)
Learning and Inference

• How do we find most likely structure?
  • Cutting Plane Inference+ILP (Riedel, 2008)

• How do we learn weights?
Learning and Inference

- How do we find most likely structure?
  - Cutting Plane Inference+ILP (Riedel, 2008)

- How do we learn weights?
  - Single best MIRA (Krammer, 2006)
System
System

- Local Formulae
System

- Local Formulae
- Dependency paths (labelled, unlabelled)
System

• Local Formulae
• Dependency paths (labelled, unlabelled)
• Words, POS tags, Stems, 2 Dictionaries
System

• Local Formulae
  • Dependency paths (labelled, unlabelled)
  • Words, POS tags, Stems, 2 Dictionaries
• Global Formulae
System

• Local Formulae
  • Dependency paths (labelled, unlabelled)
  • Words, POS tags, Stems, 2 Dictionaries
• Global Formulae
  • (Inspired by the shared task validator)
Results
Results

• **44.4** F-score for Task 1 (4th rank)
Results

- **44.4** F-score for Task 1 (4th rank)
- Using dev set for training: 45.1 (3th rank)
Results

- **44.4** F-score for Task 1 (4th rank)
- Using dev set for training: **45.1** (3th rank)
- **43.1** F-score for Task 2 (1st rank)
Results

- **44.4** F-score for Task 1 (4th rank)
- Using dev set for training: **45.1** (3th rank)
- **43.1** F-score for Task 2 (1st rank)
- Poor results for **Bindings**
Results

- **44.4** F-score for Task 1 (4th rank)
- Using dev set for training: **45.1** (3rd rank)
- **43.1** F-score for Task 2 (1st rank)
- Poor results for **Bindings**
- Fast to train: **3 hours** on MacBook Pro
Impact of Joint Inference?
Impact of Joint Inference?

- Formula sets:
Impact of Joint Inference?

- Formula sets:
  - **CORE**: not more than one type, role
Impact of Joint Inference?

- Formula sets:
  - **CORE**: not more than one type, role
  - **VALID**: consistent arguments and events
Impact of Joint Inference?

- Formula sets:
  - **CORE**: not more than one type, role
  - **VALID**: consistent arguments and events
  - **FULL**: “additional” formulae
Impact of Joint Inference (on Atom F1)

<table>
<thead>
<tr>
<th>Predicate</th>
<th>CORE</th>
<th>VALID</th>
<th>FULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>50.7</td>
<td>60.1</td>
<td>61.9</td>
</tr>
<tr>
<td>event</td>
<td>52.8</td>
<td>63.2</td>
<td>64.3</td>
</tr>
<tr>
<td>role</td>
<td>44.0</td>
<td>53.5</td>
<td>55.7</td>
</tr>
<tr>
<td>site</td>
<td>42.0</td>
<td>46.0</td>
<td>51.5</td>
</tr>
</tbody>
</table>
Impact of Joint Inference (on Atom F1)

<table>
<thead>
<tr>
<th>Predicate</th>
<th>CORE</th>
<th>VALID</th>
<th>FULL</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>50.7</td>
<td>60.1</td>
<td>61.9</td>
</tr>
<tr>
<td>event</td>
<td>52.8</td>
<td>63.2</td>
<td>64.3</td>
</tr>
<tr>
<td>role</td>
<td>44.0</td>
<td>53.5</td>
<td>55.7</td>
</tr>
<tr>
<td>site</td>
<td>42.0</td>
<td>46.0</td>
<td>51.5</td>
</tr>
</tbody>
</table>

- FULL does not explicitly consider site
- But Turku still did better w/ a local model
Conclusion
Conclusion

• **Joint** and **Declarative** Approach
Conclusion

• **Joint** and **Declarative** Approach

• Global Formulae help **across the board**
Conclusion

• **Joint** and **Declarative** Approach
• Global Formulae help **across the board**
• Location extraction **easy to implement**
Conclusion

• **Joint** and **Declarative** Approach
• Global Formulae help **across the board**
• Location extraction **easy to implement**
• **Compact** representation of model