Event Argument Extraction and Linking: Discovering and Characterizing Emerging Events (DISCERN)

Archna Bhatia, Adam Dalton, Bonnie Dorr,* Greg Dubbin, Kristy Hollingshead, Suriya Kandaswamy, and Ian Perera

Florida Institute for Human and Machine Cognition

11/17/2015
NIST TAC Workshop
Main Take-Away’s

• Symbolic (rule-based) and machine-learned approaches exhibit complementary advantages.
• Detection of nominal nuggets and merging nominals with support verbs improves recall.
• Automatic annotation of semantic role labels improves event argument extraction.
• Challenges of expanding rule-based systems are addressed through an interface for rapid iteration and immediate verification of rule changes.
The Tasks

• Event Nugget Detection (EN)

• Event Argument Extraction and Linking (EAL)
The Tasks

• Event Nugget Detection (EN)

  The **attack** by insurgents occurred on Saturday. Kennedy was **shot dead** by Oswald.

• Event Argument Extraction and Linking (EAL)
The Tasks

• Event Nugget Detection (EN)
  The attack by insurgents occurred on Saturday.
  Kennedy was shot dead by Oswald.

• Event Argument Extraction and Linking (EAL)
The Tasks

- Event Nugget Detection (EN)
  The **attack** by insurgents occurred on Saturday.
  Kennedy was **shot dead** by Oswald.

- Event Argument Extraction and Linking (EAL)
  The **attack** by **insurgents** occurred on **Saturday**.
  **Kennedy** was **shot dead** by **Oswald**.
The Tasks

• Event Nugget Detection (EN)

The **attack** by insurgents occurred on Saturday.
Kennedy was **shot dead** by Oswald.

• Event Argument Extraction and Linking (EAL)

The **attack** by **insurgents** occurred on **Saturday**.
**Kennedy** was **shot dead** by **Oswald**.
The Tasks

• Event Nugget Detection (EN)

The **attack** by insurgents occurred on Saturday.
Kennedy was **shot dead** by Oswald.

• Event Argument Extraction and Linking (EAL)

The **attack** by **insurgents** occurred on **Saturday**.
Kennedy was **shot dead** by **Oswald**.
Discovering and Characterizing Emerging Events (DISCERN)

Two Pipelines:
• Development time
• Evaluation time
DISCERN: Development time

Preprocessing training/development data
- Automatic annotations
- Support verb & event nominal Merger

Rule Creation/learning & development
- Hand crafting/ ML for rules
- Web-based front-end used for further development of hand-crafted rules

Implementation
- Detect event trigger
- Assign Realis
- Detect arguments
- Canonical Argument String resolution
DISCERN: Evaluation time

Preprocessing unseen data
- Automatic annotations
- Support verb & event nominal Merger

Implementation
- Detect event trigger
- Assign Realis
- Detect arguments
- Canonical Argument String resolution
DISCERN Preprocessing (both pipelines)

Stanford CoreNLP
- Stripping XML off
- Splitting sentences
- POS tagging, lemmatization, NER tagging, Coreference, Dependency tree

CatVar
- Word-POS pairs added

Support-verb & Event nominal merger
- New dependency tree generated with support verbs and nominals merged into a single unit

Senna
- Semantic Role Labeling (SRL) with PropBank labels
CatVar

- A database for categorial variations of English lexemes (Habash & Dorr, 2003)
- Connects derivationally-related words with different POS tags ➔ can help in identifying more trigger words (e.g., capturing non-verbal triggers)

Business.Merge-Org (before CatVar)
- Consolidate [V]
- Merge [V]
- Combine [V]

Business.Merge-Org (after CatVar)
- Consolidate [V], Consolidation [N], Consolidated [AJ], Merge [V], Merger [N]
- Combine [V], Combination [N]
Support-verb and Nominal Merger

• Support-verbs contain little semantic information but take the semantic arguments of the nominal as its own syntactic dependents.

<table>
<thead>
<tr>
<th>Support Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light Verbs:</td>
</tr>
<tr>
<td>Do, Give, Make, Have</td>
</tr>
<tr>
<td>Other:</td>
</tr>
<tr>
<td>Declare, Conduct, Stage</td>
</tr>
</tbody>
</table>

• Support verb and nominal are merged

Detroit *declared bankruptcy* on July 18, 2013.
Support-verb and Nominal Merger

• Support-verbs contain little semantic information but take the semantic arguments of the nominal as its own syntactic dependents.

• Support verb and nominal are merged

Detroit **declared** bankruptcy on July 18, 2013.
Support-verb and Nominal Merger

- Support-verbs contain little semantic information but take the semantic arguments of the nominal as its own syntactic dependents.

- Support verb and nominal are merged

Detroit *declared* **bankruptcy** on July 18, 2013.

**Support Verbs**

<table>
<thead>
<tr>
<th>Light Verbs:</th>
<th>Other:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do, Give, Make, Have</td>
<td>Declare, Conduct, Stage</td>
</tr>
</tbody>
</table>

**Diagram:**

- *declared* (nsubj)
- **bankruptcy** (nmod:on)
DISCERN: Development time

- Preprocessing training/development data
  - Automatic annotations
  - Support verb & event nominal Merger

- Rule Creation/learning & development
  - Hand crafting/ML for rules
  - Web-based front-end used for further development of hand-crafted rules

- Implementation
  - Detect event trigger
  - Assign Realis
  - Detect arguments
  - Canonical Argument String resolution
How are rules created for DISCERN?

- Manually created linguistically-informed rules (DISCERN-R)
- Machine learned rules (DISCERN-ML)
- A combination of the manually created rules and the machine learned rules (DISCERN-C)

Three variants of DISCERN submitted by IHMC
**DISCERN-R:**

- DISCERN-R uses handcrafted rules for determining nuggets and arguments
- Event sub-types are assigned representative lemmas

<table>
<thead>
<tr>
<th>Event Sub-type</th>
<th>Justice.Arrest-Jail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemmas</td>
<td>arrest, capture, jail, imprison</td>
</tr>
<tr>
<td>Roles</td>
<td>Person</td>
</tr>
<tr>
<td>Features</td>
<td>Person</td>
</tr>
<tr>
<td>Values</td>
<td>A1</td>
</tr>
</tbody>
</table>
DISCERN-R:

- Rules map roles for each event sub-type to semantic and syntactic features
- Lexical resources inform rules: OntoNotes, Thesaurus, CatVar, VerbNet, Senna/PropBank (SRL)

**Event Sub-type**

- Justice.Arrest-Jail

**Lemmas**

- arrest, capture, jail, imprison

**Roles**

- Person
- Agent[1]

**Features**

- Dependency Type
- Senna/PropBank
- VerbNet
- Senna/PropBank
- VerbNet

**Values**

- dobj
- nmod:of
- A1
- Patient
- A0
- Agent
DISCERN-ML

- Decision trees trained using ID3 algorithm
- Every event sub-type has a binary decision tree
  - Every word is classified by that decision tree.
  - A word that is labeled as a yes is trigger of that sub-type
- Each role belonging to an event sub-type has a binary decision tree
  - This example classifies the Entity role Contact.Meet
  - Tested against dependents of Contact.Meet triggers in dependency tree
DISCERN-C

- Combines DISCERN-R with DISCERN-ML, where DISCERN-R rules act like a set of decision trees
- DISCERN-R rules are compared to DISCERN-ML rules and considered five times as strong
Web-based Front-End for Rule Development

<table>
<thead>
<tr>
<th>ID</th>
<th>True Positives</th>
<th>False Positives</th>
<th>FP Wrong Basefiller</th>
<th>FP Wrong Event Type</th>
<th>FP Wrong Role</th>
<th>FP Wrong Realisa</th>
<th>False Negatives</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>130.0</td>
<td>896.0</td>
<td>537</td>
<td>175</td>
<td>139</td>
<td>35</td>
<td>1530.0</td>
<td>0.1267</td>
<td>0.0663</td>
<td>0.0871</td>
</tr>
<tr>
<td>4</td>
<td>80.0</td>
<td>410.0</td>
<td>250</td>
<td>85</td>
<td>59</td>
<td>7</td>
<td>1880.0</td>
<td>0.1633</td>
<td>0.0408</td>
<td>0.0653</td>
</tr>
<tr>
<td>5</td>
<td>103.0</td>
<td>681.0</td>
<td>391</td>
<td>135</td>
<td>117</td>
<td>28</td>
<td>1857.0</td>
<td>0.1314</td>
<td>0.0526</td>
<td>0.0751</td>
</tr>
<tr>
<td>6</td>
<td>96.0</td>
<td>519.0</td>
<td>291</td>
<td>105</td>
<td>95</td>
<td>18</td>
<td>1862.0</td>
<td>0.1588</td>
<td>0.0500</td>
<td>0.0761</td>
</tr>
<tr>
<td>7</td>
<td>130.0</td>
<td>896.0</td>
<td>537</td>
<td>175</td>
<td>139</td>
<td>35</td>
<td>1830.0</td>
<td>0.1267</td>
<td>0.0663</td>
<td>0.0871</td>
</tr>
<tr>
<td>8</td>
<td>124.0</td>
<td>683.0</td>
<td>408</td>
<td>136</td>
<td>107</td>
<td>22</td>
<td>1836.0</td>
<td>0.1537</td>
<td>0.0633</td>
<td>0.0896</td>
</tr>
<tr>
<td>9</td>
<td>86.0</td>
<td>590.0</td>
<td>337</td>
<td>118</td>
<td>90</td>
<td>36</td>
<td>1894.0</td>
<td>0.1006</td>
<td>0.0337</td>
<td>0.0505</td>
</tr>
<tr>
<td>10</td>
<td>45.0</td>
<td>636.0</td>
<td>512</td>
<td>83</td>
<td>12</td>
<td>4</td>
<td>1917.0</td>
<td>0.0661</td>
<td>0.0229</td>
<td>0.0341</td>
</tr>
<tr>
<td>11</td>
<td>218.0</td>
<td>2475.0</td>
<td>1980</td>
<td>303</td>
<td>106</td>
<td>30</td>
<td>1745.0</td>
<td>0.0810</td>
<td>0.1111</td>
<td>0.0936</td>
</tr>
<tr>
<td>12</td>
<td>221.0</td>
<td>3051.0</td>
<td>2418</td>
<td>434</td>
<td>108</td>
<td>38</td>
<td>1742.0</td>
<td>0.0675</td>
<td>0.1126</td>
<td>0.0844</td>
</tr>
<tr>
<td>13</td>
<td>222.0</td>
<td>7150.0</td>
<td>5775</td>
<td>1126</td>
<td>142</td>
<td>35</td>
<td>1741.0</td>
<td>0.0301</td>
<td>0.1131</td>
<td>0.0476</td>
</tr>
<tr>
<td>14</td>
<td>225.0</td>
<td>2119.0</td>
<td>1618</td>
<td>326</td>
<td>79</td>
<td>40</td>
<td>1738.0</td>
<td>0.0960</td>
<td>0.1146</td>
<td>0.1045</td>
</tr>
<tr>
<td>15</td>
<td>165.0</td>
<td>1127.0</td>
<td>674</td>
<td>236</td>
<td>143</td>
<td>45</td>
<td>1798.0</td>
<td>0.1277</td>
<td>0.0841</td>
<td>0.1014</td>
</tr>
<tr>
<td>16</td>
<td>154.0</td>
<td>1143.0</td>
<td>680</td>
<td>254</td>
<td>141</td>
<td>36</td>
<td>1809.0</td>
<td>0.1187</td>
<td>0.0785</td>
<td>0.0945</td>
</tr>
<tr>
<td>17</td>
<td>118.0</td>
<td>890.0</td>
<td>562</td>
<td>184</td>
<td>95</td>
<td>29</td>
<td>1845.0</td>
<td>0.1171</td>
<td>0.0601</td>
<td>0.0794</td>
</tr>
<tr>
<td>18</td>
<td>163.0</td>
<td>1185.0</td>
<td>701</td>
<td>263</td>
<td>154</td>
<td>38</td>
<td>1800.0</td>
<td>0.1209</td>
<td>0.0830</td>
<td>0.0985</td>
</tr>
<tr>
<td>19</td>
<td>14.0</td>
<td>251.0</td>
<td>128</td>
<td>38</td>
<td>3</td>
<td>3</td>
<td>1947.0</td>
<td>0.0528</td>
<td>0.0071</td>
<td>0.0126</td>
</tr>
<tr>
<td>20</td>
<td>144.0</td>
<td>1145.0</td>
<td>657</td>
<td>258</td>
<td>131</td>
<td>31</td>
<td>1821.0</td>
<td>0.1115</td>
<td>0.0733</td>
<td>0.0884</td>
</tr>
<tr>
<td>21</td>
<td>154.0</td>
<td>1161.0</td>
<td>528</td>
<td>279</td>
<td>135</td>
<td>42</td>
<td>1817.0</td>
<td>0.1171</td>
<td>0.0781</td>
<td>0.0907</td>
</tr>
<tr>
<td>22</td>
<td>136.0</td>
<td>964.0</td>
<td>526</td>
<td>230</td>
<td>97</td>
<td>40</td>
<td>1831.0</td>
<td>0.1236</td>
<td>0.0691</td>
<td>0.0887</td>
</tr>
<tr>
<td>23</td>
<td>76.0</td>
<td>1164.0</td>
<td>872</td>
<td>177</td>
<td>64</td>
<td>10</td>
<td>1891.0</td>
<td>0.0613</td>
<td>0.0386</td>
<td>0.0474</td>
</tr>
<tr>
<td>24</td>
<td>181.0</td>
<td>1727.0</td>
<td>1156</td>
<td>317</td>
<td>105</td>
<td>32</td>
<td>1799.0</td>
<td>0.0949</td>
<td>0.0914</td>
<td>0.0931</td>
</tr>
<tr>
<td>25</td>
<td>166.0</td>
<td>1350.0</td>
<td>801</td>
<td>313</td>
<td>82</td>
<td>50</td>
<td>1806.0</td>
<td>0.1095</td>
<td>0.0842</td>
<td>0.0952</td>
</tr>
<tr>
<td>Event Type</td>
<td>Role</td>
<td>True Positives</td>
<td>False Negatives</td>
<td>Basefiller</td>
<td>Type</td>
<td>Role</td>
<td>Realis</td>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>----------</td>
<td>----------------</td>
<td>-----------------</td>
<td>------------</td>
<td>------</td>
<td>------</td>
<td>--------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life.Die</td>
<td>Victim</td>
<td>12</td>
<td>112</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact.Meet</td>
<td>Entity</td>
<td>11</td>
<td>95</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice.Arrest-Jail</td>
<td>Person</td>
<td>9</td>
<td>54</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice.Charge-Indict</td>
<td>Crime</td>
<td>9</td>
<td>46</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflict.Attack</td>
<td>Attacker</td>
<td>8</td>
<td>44</td>
<td>11</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice.Charge-Indict</td>
<td>Defendant</td>
<td>7</td>
<td>69</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflict.Attack</td>
<td>Target</td>
<td>6</td>
<td>64</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice.Convict</td>
<td>Defendant</td>
<td>6</td>
<td>68</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life.Die</td>
<td>Place</td>
<td>5</td>
<td>47</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life.Divorce</td>
<td>Person</td>
<td>5</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
26

Scores

Ground Truth Events: 05
Anchors Hits: 18
Anchors inside predicate: 15
Rule Hits: 18
Correct: 3
DISCERN: Evaluation time

Preprocessing
unseen
data

Automatic annotations
Support verb & event
nominal Merger

Implementation

Detect event trigger
Assign Realis
Detect arguments

Canonical Argument String resolution
DISCERN Implementation

- Detect event triggers (nuggets)
- Assign Realis
- Detect arguments from trigger’s dependents
- Canonical Argument String (CAS) Resolution
Detecting Triggers

• Each event subtype has a classifier to locate triggers of that subtype

• Main features:
  – Lemmas
  – CatVar
  – Part-of-Speech
Assigning Realis

• Each event trigger is assigned Realis
• Series of straightforward linguistic rules
• Examples:
  – Non-verbal trigger with no support verb or copula
    -> ACTUAL
    • “The AP reported an attack this morning.”
  – Verbal trigger with “MD” dependent -> OTHER
    • “The military may attack the city.”
Argument Detection

• Determine arguments from among the trigger’s dependents
• Support-verb collapsing includes dependents of the support verb
• Experimented with three variants
## Event Nuggets Results

<table>
<thead>
<tr>
<th>System</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCERN-R</td>
<td>32%</td>
<td>26%</td>
<td>29%</td>
</tr>
<tr>
<td>DISCERN-ML</td>
<td>9%</td>
<td>26%</td>
<td>14%</td>
</tr>
<tr>
<td>DISCERN-C</td>
<td>9%</td>
<td>31%</td>
<td>14%</td>
</tr>
</tbody>
</table>
# Event Argument Results

<table>
<thead>
<tr>
<th>System</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCERN-R</td>
<td>12.83%</td>
<td>14.13%</td>
<td>13.45%</td>
</tr>
<tr>
<td>DISCERN-ML</td>
<td>7.39%</td>
<td>9.19%</td>
<td>8.19%</td>
</tr>
<tr>
<td>DISCERN-C</td>
<td>8.18%</td>
<td>15.02%</td>
<td>10.59%</td>
</tr>
<tr>
<td>Median</td>
<td>30.65%</td>
<td>11.66%</td>
<td>16.89%</td>
</tr>
<tr>
<td>Human</td>
<td>73.62%</td>
<td>39.43%</td>
<td>51.35%</td>
</tr>
</tbody>
</table>
Ablation Experiments

DISCERN-R with varying features

- Support verbs
- Semantic role labeling (SRL)
- Named entity recognition (NER)
- CatVar
- Dependency types
### Ablation Results Table

<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support Verbs</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SRL</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>NER</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CatVar</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Dependencies</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>10.88%</th>
<th>10.89%</th>
<th>11.99%</th>
<th>11.00%</th>
<th>11.71%</th>
<th>12.08%</th>
<th>10.93%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall</td>
<td>5.49%</td>
<td>5.39%</td>
<td>3.76%</td>
<td>3.76%</td>
<td>3.66%</td>
<td>3.66%</td>
<td>4.99%</td>
</tr>
<tr>
<td>F-Score</td>
<td>7.30%</td>
<td>7.21%</td>
<td>5.73%</td>
<td>5.61%</td>
<td>5.58%</td>
<td>5.62%</td>
<td>6.85%</td>
</tr>
</tbody>
</table>

CatVar and support verbs boosts recall but lowers precision.
CatVar and Support-verbs Merging

• CatVar detects nominal triggers:

In Switzerland... the real estate owner... remained in detention.
CatVar/Support-verb improves recall

- Support verbs are located:

In Switzerland... the real estate owner... *remained* in *detention.*
CatVar/Support-verb improves recall

- Support verb and nominal are merged:

In Switzerland... the real estate owner... \textcolor{red}{remained} in \textcolor{red}{detention}.
Where does CatVar hurt?

• “Catvariation” can be overly aggressive

Even within the confines of `pure country’, Jones did not stand still...

The case was transferred ... to the State Security prosecutor for further investigation.

South African Leader cites `progress’ in Mandela’s condition
### Ablation Results Table

<table>
<thead>
<tr>
<th>Support Verbs</th>
<th>+</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRL</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>NER</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CatVar</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Dependencies</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Precision</td>
<td>10.88%</td>
<td>10.89%</td>
<td>11.99%</td>
<td>11.00%</td>
<td>11.71%</td>
<td>12.08%</td>
<td>10.93%</td>
<td></td>
</tr>
<tr>
<td>Recall</td>
<td>5.49%</td>
<td>5.39%</td>
<td>3.76%</td>
<td>3.76%</td>
<td>3.66%</td>
<td>3.66%</td>
<td>4.99%</td>
<td></td>
</tr>
<tr>
<td>F-Score</td>
<td>7.30%</td>
<td>7.21%</td>
<td>5.73%</td>
<td>5.61%</td>
<td>5.58%</td>
<td>5.62%</td>
<td>6.85%</td>
<td></td>
</tr>
</tbody>
</table>

SRL boosts recall, but lowers precision
SRL improves recall

• Helps with general dependency types:
  the Iraqi car bombing ... that **killed** 50 +
  **xcomp**

• Helps with mislabelled dependencies:
  **NEW YORK** ... A pedestrian was killed ...
  **rcmod**
Where does SRL hurt?

• Mislabeled semantic roles:
  
  $4.6 \text{ million} \ldots \text{to be } \textbf{distributed} \text{ among the victims' relatives}^*.$

• Heterogeneous semantic role labels:

  1. The New York investor didn’t demand the company also \textbf{pay} a \textit{premium} to other shareholders.

  2. He wouldn’t \textbf{accept} anything of value \textit{from those he was writing about}.

  AM-LOC
  nmod:among

  A2

Where does SRL hurt?

• Overly general semantic roles:

  ... the second Catholic ever* nominated...

  ... nominated for 3 MAMAs*...
Future Work

• Implementation of semantic role constraints to ensure each role assigned to at most argument for potential precision improvement of 5%.

• Joint learning of event trigger and argument extraction (e.g. Li et al, 2013) for improvements in event/argument detection

• Improving semantic role labeller precision to compensate for mislabeling and incorrect parses
  – Adapting roles to individual domain
  – Deep semantic parsing e.g. TRIPS (Allen, 2008)
Conclusions

• Web-interface enables rapid iteration and improvement
• Support-verb merging in conjunction with CatVar improves recall, surpassing median
• Semantic roles can help in cases where dependencies fall short, but they must be used with care due to inaccurate or overly general assignments.
• Combining linguistic knowledge with machine learning methods can improve over either method alone
THANKS!

This work was supported, in part, by the Defense Advanced Research Projects Agency (DARPA) under Contract No. FA8750-12-2-0348, The Office of Naval Research (N000141210547), and the Nuance Foundation.