Overview of the TAC2010 Knowledge Base Population (KBP) Track

Heng Ji, Ralph Grishman, Hoa Trang Dang and Kira Griffitt

November 15, 2010
Outline

- Motivation of KBP Tasks
- KBP2010 Task Overview
- Participants Overview
- Data Annotation and Analysis
- Evaluation Metrics
- Evaluation Results
- Discussions and Lessons
The Goal of KBP

- **General Goal**
  - Promote research in discovering facts about entities and expanding a knowledge source

- **What’s New**
  - Extraction at large scale (> 1 million documents);
  - Using a representative collection (not selected for relevance);
  - Cross-document entity resolution (extending the limited effort in ACE);
  - Linking the facts in text to a knowledge base;
  - Distant (and noisy) supervision through Infoboxes;
  - Rapid adaptation to new relations;
  - Support multi-lingual information fusion (KBP2011);
  - Capture temporal information (KBP2011)
Knowledge Base Population (KBP2010) Task Overview
KBP Setup

- Knowledge Base (KB)
  - Attributes (a.k.a., “slots”) derived from Wikipedia infoboxes are used to create the reference KB

- Source Collection
  - A large corpus of newswire and web documents (>1.3 million docs) is provided for systems to discover information to expand and populate KB

Destiny’s Child, Knowles released her debut solo album Dangerously in Love (2003), which spawned the number one and “Baby Boy” and became one of the most successful albums of that year, earning her a then record-tying five
Following the group’s disbandment in 2005, Knowles released B’Day in 2006. It debuted at number one on the

acting career in 2001, appearing in the musical film Carmen: A Hip Hopera. In 2006, she starred in the lead role in the

1981 Broadway musical Dreamgirls, for which she earned two Golden Globe nominations. Knowles launched her

House of Dereon, in 2004, and has endorsed such brands as Pepsi, Tommy Hilfiger, Armani and L’Oréal. In 2010,

bles at number two on its list of the 100 Most Powerful and Influential Celebrities in the world. Knowles was also listed

al and influential musician in the world. Time also included Knowles on its list of the “100 Most Influential People in the
Entity Linking: Create Wiki Entry?

Query = “James Parsons”
Query = “James Parsons”
Entity Linking Task Definition

- Involve Three Entity Types
  - Person, Geo-political, Organization

- Regular Entity Linking
  - Names must be aligned to entities in the KB; can use Wikipedia texts

- Optional Entity linking
  - Without using Wikipedia texts, can use Infobox values

- Query Example
  
  <query id="EL000304">
    <name>Jim Parsons</name>
    <docid>eng-NG-31-100578-11879229</docid>
  </query>
Jim Parsons, a graduate of the University of Houston School of Music, Theatre and Dance, won the Emmy on Sunday for Lead Actor in a Comedy Series for his work on The Big Bang Theory.

School Attended: University of Houston

Jim Parsons

Parsons in 2006

Born
James Joseph Parsons
March 24, 1973 (age 37)
Houston, Texas, U.S.

Occupation
Actor

Years active
2000–present
# Regular Slot Filling

<table>
<thead>
<tr>
<th>Person</th>
<th>Organization</th>
<th>Person</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>per:alternate_names</td>
<td>per:title</td>
<td>org:alternate_names</td>
<td></td>
</tr>
<tr>
<td>per:date_of_birth</td>
<td>per:member_of</td>
<td>org:political/religious_affiliation</td>
<td></td>
</tr>
<tr>
<td>per:age</td>
<td>per:employee_of</td>
<td>org:top_members/employees</td>
<td></td>
</tr>
<tr>
<td>per:country_of_birth</td>
<td>per:religion</td>
<td>org:number_of_employees/members</td>
<td></td>
</tr>
<tr>
<td>per:stateorprovince_of_birth</td>
<td>per:spouse</td>
<td>org:members</td>
<td></td>
</tr>
<tr>
<td>per:city_of_birth</td>
<td>per:children</td>
<td>org:member_of</td>
<td></td>
</tr>
<tr>
<td>per:origin</td>
<td>per:parents</td>
<td>org:subsidiaries</td>
<td></td>
</tr>
<tr>
<td>per:date_of_death</td>
<td>per:siblings</td>
<td>org:parents</td>
<td></td>
</tr>
<tr>
<td>per:country_of_death</td>
<td>per:other_family</td>
<td>org:founded_by</td>
<td></td>
</tr>
<tr>
<td>per:stateorprovince_of_death</td>
<td>per:charges</td>
<td>org:founded</td>
<td></td>
</tr>
<tr>
<td>per:city_of_death</td>
<td></td>
<td>org:dissolved</td>
<td></td>
</tr>
<tr>
<td>per:cause_of_death</td>
<td></td>
<td>org:country_of_headquarters</td>
<td></td>
</tr>
<tr>
<td>per:countries_of_residence</td>
<td></td>
<td>org:stateorprovince_of_headquarters</td>
<td></td>
</tr>
<tr>
<td>per:stateorprovinces_of_residence</td>
<td></td>
<td>org:city_of_headquarters</td>
<td></td>
</tr>
<tr>
<td>per:cities_of_residence</td>
<td></td>
<td>org:shareholders</td>
<td></td>
</tr>
<tr>
<td>per:schools_attended</td>
<td></td>
<td>org:website</td>
<td></td>
</tr>
</tbody>
</table>
Surprise Slot Filling

- Research on Portability of IE/QA techniques
  - how easily and rapidly an IE system can be adapted to new types of relations and events

- Sites were given a maximum of 4 days to develop their systems

- KBP2010 Surprise Slot Types
  - Per: diseases
  - Per: awards-won
  - Per: charity-Supported
  - Org: products
Differences from KBP2009

- Entity Linking
  - Added Optional Entity Linking Task

- Slot Filling
  - Location slots were each divided into three slots (city, state/province, country)
  - Origin was changed from a single to a list value
  - Dropped Geo-political Entities as a query type
  - Added Surprise Slot Filling Task

- Query Selection
  - Independent selection processes for entity linking and slot filling

- Submission Rules
  - Alternative runs were encouraged to evaluate the impact of web access and external knowledge base and offline Wikipedia text mining
  - At least one submission had to run without web access
Participant Overview
## KBP2010 Participants

<table>
<thead>
<tr>
<th>Team Name</th>
<th>Reg. EL</th>
<th>Opt. EL</th>
<th>Reg. SF</th>
<th>Surp. SF</th>
<th>Team Name</th>
<th>Reg. EL</th>
<th>Opt. EL</th>
<th>Reg. SF</th>
<th>Surp. SF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARPANI</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>NUSchime</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BUDAPEST T-ACAD</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>NYU</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>BUPTPRIS</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>SIEL</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>CMCRC</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>SMU</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>CORTEX</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>STANFORD</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>CUNY</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>STANFORD _UBC</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>HLTCOE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>TCAR</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td>UBC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>ICL</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>UC3M</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIRG</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td>USFD</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>LCC</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>WebTLab</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSV</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
KBP2010 Participants

- over 45 teams registered for KBP 2010 (not including the RTE-KBP Pilot task)
- Each team can submit up to 3 submissions

<table>
<thead>
<tr>
<th>Tasks</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#Teams</td>
<td>#Submissions</td>
</tr>
<tr>
<td>Regular Entity Linking</td>
<td>13</td>
<td>35</td>
</tr>
<tr>
<td>Optional Entity Linking</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Regular Slot Filling</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Surprise Slot Filling</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Data Annotation and Analysis
## Data Annotation Overview

**Source collection:** about 1.3 million newswire docs and 500K web docs, a few speech transcribed docs

### Entity Linking Corpus

<table>
<thead>
<tr>
<th>Genre/Source</th>
<th>Size (entity mentions)</th>
<th>Person</th>
<th>Organization</th>
<th>GPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>2009 Training</td>
<td>627</td>
<td>2710</td>
<td>567</td>
</tr>
<tr>
<td></td>
<td>2010 Web data</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Newswire</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>Web data</td>
<td>250</td>
<td>250</td>
<td>250</td>
</tr>
</tbody>
</table>

### Slot Filling Corpus

<table>
<thead>
<tr>
<th>Task</th>
<th>Source</th>
<th>Size (entities)</th>
<th>Person</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>Regular Task</td>
<td>2009 Evaluation</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>2010 Participants</td>
<td></td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2010 LDC</td>
<td></td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Surprise Task</td>
<td>2010 LDC</td>
<td></td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Regular Task</td>
<td>LDC</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Surprise Task</td>
<td>LDC</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>
Data Selection
Data Selection: Inputs

- KBP 2010 source data
  - KBP 2009 newswire data (~1M documents)
  - KBP 2010 weblog data (~500k documents)
- Named Entity (NE) tagger output from NYU:
  - Namestrings selected from source data, labeled with type information (i.e. PER, ORG, or GPE)
  - Number of possible KB matches for a namestring
  - Number of KBP 2010 source data documents containing a namestring
Data Selection: Entity Linking Stage 1

- Namestrings selected for the Entity-Linking task were confusable:
  - Quantitatively - No or many possible KB matches:
    - 0 or 7+ possible KB matches
  - Qualitatively – Entities difficult to determine:
    - Incorrect or alternative spelling (e.g. ‘Jon’ vs. ‘John’)
    - Abbreviated (e.g. ‘CDC’)
    - Ambiguous or more common name (e.g. ‘John Smith’)

20 documents were selected from the source data for each of the Entity Linking namestrings

- If more than 20 documents were found, the selected set equally represented the unique entities possibly referred to by the namestring

Example: ‘John W’ namestring

- Source data search produced 50 documents containing ‘John W.’
- In these documents, three unique entities were referred to by the namestring ‘John W.’ (e.g. John Williams, John Wagner, John Wilson)
- The 20 documents selected would have equally represented these three unique entities
Data Selection: Slot-Filling
Stage 1

Namestrings selected for the Slot-Filling task:

- Were not confusable:
  - Quantitatively - Few possible KB matches
    - 1-6 possible KB matches
  - Qualitatively - Easy to determine entity namestrings:
    - Correct spelling (e.g. ‘Bill Clinton’)
    - Not abbreviated (e.g. ‘American Medical Association’)
    - Unambiguous or less common name (e.g. ‘Tony Blair’)

- Had slot-filling answers in source data
- Had an incomplete or non-existent KB entry
Data Selection: Slot-Filling
Stage 2

- Searched source data to ensure there were slot-filling answers for each selected namestring
  - 2-3 slot-filling answers were needed for a SF entity to remain in pipeline

- Also searched source data for a reference document for each SF entity
  - Selection of the reference document created namestring-document pair, which was used to determine if SF entity had a KB entry
  - If no reference document was available, the namestring was excluded from the pipeline
Data Selection:
Excluded Namestrings

- Excluded Namestrings did not meet:
  - The confusability requirements for the Entity Linking or Slot Filling tasks
    - Namestrings had to meet both the quantitative AND qualitative criteria
  - General data-selection standards
    - incorrectly typed by NE tagger
    - Nonsensical
    - objectionable content
**KB-Linking:**
Stage 3 for both EL and SF entities

- Searched to determine whether entities in namestring-document pairs had KB entries
  - If it was unclear whether an entity had a KB entry, it was removed from the pipeline
  - Entity Linking entities marked “KB entry” or “No KB entry” were released as Entity Linking queries
  - Slot-Filling entities marked “KB entry” or “No KB entry” were reviewed and used in slot-filling annotation
Slot-Filling Annotation
Slot-Filling Annotation Preparations

- KB entries for all SF entities were reviewed to prevent redundant annotations.
  - Already-filled slots were made visible to annotators but could not be altered.

- Slot-Filling Guidelines Developed
  - Used 2009 Assessment guidelines as a basis
  - Added descriptions of slot-filling task and annotation categories

- In-House Annotator Training
  - Done in-house so that potential annotators could complete a practice kit with supervision.
Slot-Filling: Annotation

- Annotators were presented with an entity and all available slots for that entity
  - If an entity was NOT linked to the KB, annotators were presented with
    - All single-value slots
    - All list-value slots
  - If an entity was linked to the KB, annotators were presented with
    - Empty single-value slots
    - All list-value slots
- Slot-filling annotation performed within a two-hour time limit for each SF entity
Slot-Filling Assessment
Updates to Assessment Guidelines

- In the interim between Slot-Filling annotation and Assessment, continued discussions between LDC staff and TAC coordinators on questions raised by annotators produced clarifications on slot descriptions
  - Classification of specific GPEs (Emirates, Capital Districts)
  - Employers of actors
  - Residence of Universities
  - Headquarters of sports teams
  - Non-recognized states and areas of contention
  - Alternate names
Slot-Filling: Assessment

- Potential assessors were rigorously tested
  - Required to complete a full-size assessment kit
    - 12 slots filled with slot-filling answers for an actual SF entity
  - 90% accuracy or greater required to pass
    - A number of candidates did not make the cut
  - Questions raised during the test resulted in further clarifications to the guidelines:
    - Acceptability of past and future relationships
    - PER:Age, PER:Alternate Names, PER:Country of Birth, PER:Employee Of, PER:Origin
    - ORG:Alternate Names and ORG:Number of Employees/Members

- Those who passed went on to assess the validity of slot-filling answers from both humans and machines
Slot-Filling: Assessment

- Quality Control
  - After assessment was completed, systematic and spot-checking QC was performed on 50% of the assessment data.
  - These QC passes revealed only incidental errors, which were corrected prior to release of assessment results.

- Errors identified in assessment results will be incorporated into the 2011 slot-filling annotation and assessment guidelines.
### Entity Linking Inter-Annotator Agreement

#### Graphical Representation

![Venn Diagram showing agreement rates between three annotators for different entity types and genres.](image)

#### Table

<table>
<thead>
<tr>
<th>Entity Type</th>
<th>#Total Queries</th>
<th>Agreement Rate</th>
<th>Genre</th>
<th>#Disagreed Queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>59</td>
<td>91.53%</td>
<td>Newswire</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Web Text</td>
<td>1</td>
</tr>
<tr>
<td>Geo-political</td>
<td>64</td>
<td>87.5%</td>
<td>Newswire</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Web Text</td>
<td>5</td>
</tr>
<tr>
<td>Organization</td>
<td>57</td>
<td>92.98%</td>
<td>Newswire</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Web Text</td>
<td>1</td>
</tr>
</tbody>
</table>
Slot Filling Human Annotation Performance

Evaluation assessment of LDC Hand Annotation

<table>
<thead>
<tr>
<th>Performance</th>
<th>P(%)</th>
<th>R(%)</th>
<th>F(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Slots</td>
<td>70.14</td>
<td>54.06</td>
<td>61.06</td>
</tr>
<tr>
<td>All except per:top-employee, per:member_of, per:title</td>
<td>71.63</td>
<td>57.6</td>
<td>63.86</td>
</tr>
</tbody>
</table>

Why is the precision only 70%?

- 32 responses were judged as inexact and 200 as wrong answers
- A third annotator’s assessment on 20 answers marked as wrong: 65% incorrect; 15% correct; 20% uncertain
- Some annotated answers are not explicitly stated in the document
- … some require a little world knowledge and reasoning
- Ambiguities and underspecification in the annotation guideline
- Confusion about acceptable answers
- Updates to KBP2010 annotation guideline for assessment
Evaluation Metrics
Entity Linking Scoring Metric

- **Micro-averaged Accuracy (official metric)**
  - Mean accuracy across all queries

- **Macro-averaged Accuracy**
  - Mean accuracy across all KB entries
Slot Filling Scoring Metric

- Each response is rated as correct, inexact, redundant, or wrong (credit only given for correct responses)
  - Redundancy: (1) response vs. KB; (2) among responses: build equivalence class, credit only for one member of each class

- Correct = # (non-NIL system output slots judged correct)
- System = # (non-NIL system output slots)
- Reference =
  # (single-valued slots with a correct non-NIL response) +
  # (equivalence classes for all list-valued slots)

- Standard Precision, Recall, F-measure
Evaluation Results
Top-10 Regular Entity Linking Systems

<0.8 correlation between overall vs. Non-NIL performance
Human/System Entity Linking Comparison (subset of 200 queries)

Average among three annotators
Impact of Wikipedia Texts on Entity Linking

Not statistically significant
(<89% confidence level using the Wilcoxon Matched-Pairs Signed-Ranks Test)
Top-10 Regular Slot Filling Systems

F-Measure
Surprise Slot Filling Evaluation Results

- Only 5 sites participated in the task, possibly because it was scheduled immediately after the regular task.
- 4 sites fielded automatic systems.
- Only one site, LCC, exceeded 10% F score.

<table>
<thead>
<tr>
<th>Time (#hours)</th>
<th>P(%)</th>
<th>R(%)</th>
<th>F(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>50.33</td>
<td>15.45</td>
<td>23.64</td>
</tr>
<tr>
<td>34</td>
<td>52.36</td>
<td>24.16</td>
<td>33.06</td>
</tr>
</tbody>
</table>
Discussion of Entity Linking
Unsupervised/Minimally-Supervised vs. Supervised Learning for Entity Linking
Impact of Semantic Features on Entity Linking

- CUNY-BLENDER: Use Slot Filling results as features (entity profile)

<table>
<thead>
<tr>
<th>System</th>
<th>Person</th>
<th>Organization</th>
<th>Geo-Political</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without SF feedback</td>
<td>84.6%</td>
<td>63.1%</td>
<td>57.5%</td>
<td>59.9%</td>
</tr>
<tr>
<td>With SF feedback</td>
<td>92.8%</td>
<td>65.7%</td>
<td>84.1%</td>
<td>69.3%</td>
</tr>
</tbody>
</table>

- BuptPris: Use name tagging, infoboxes etc. as features

<table>
<thead>
<tr>
<th>System</th>
<th>Person</th>
<th>Organization</th>
<th>Geo-Political</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without semantic Features</td>
<td>83.9%</td>
<td>59.5%</td>
<td>33.4%</td>
<td>58.9%</td>
</tr>
<tr>
<td>With semantic features</td>
<td>79.1%</td>
<td>74.1%</td>
<td>66.6%</td>
<td>73.3%</td>
</tr>
</tbody>
</table>
Impact of Data Genre (All queries)
Impact of Data Genre (Non-NIL queries)
Discussion of Slot Filling
Slot-Specific Analysis

- A few slots account for a large fraction of the answers:
  - `per:title`, `per:employee_of`, `per:member_of`, and `org:top_members/employees` account for 37% of correct responses.

- For a few slots, delimiting exact answer is difficult … result is ‘inexact’ slot fills
  - `per:charges`, `per:title` (“rookie driver”; “record producer”)

- For a few slots, equivalent-answer detection is important to avoid redundant answers
  - `per:title` again accounts for the largest number of cases. e.g., “defense minister” and “defense chief” are equivalent.
How much Inference is Needed?

- ACE event mentions (7.5%)
- ACE relation mentions (26.4%)
- Cross-slot Inference (1.8%)
- Cross-sentence Inference (18.8%)
- Cross-sentence Coreference (22.5%)
- Other (26.4%)

Within-Sentence
Why KBP is more difficult than ACE

- Cross-sentence Inference – non-identity coreference(per:children)
  - Lahoud is married to an Armenian and the couple have three children. Eldest son Emile Emile Lahoud was a member of parliament between 2000 and 2005.

- Cross-slot Inference (per:children)
  - People Magazine has confirmed that actress Julia Roberts has given birth to her third child a boy named Henry Daniel Moder. Henry was born Monday in Los Angeles and weighed 8 lbs. Roberts, 39, and husband Danny Moder, 38, are already parents to twins Hazel and Phinnaeus who were born in November 2006.
## Impact of Using External Knowledge/Wiki Text Mining for Answer Validation

<table>
<thead>
<tr>
<th>System</th>
<th>Use Answer Validation?</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCC</td>
<td>No</td>
<td>45.33</td>
<td>18.76</td>
<td>26.54</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>44.87</td>
<td>19.44</td>
<td>27.13</td>
</tr>
<tr>
<td>CUNY</td>
<td>No</td>
<td>27.99</td>
<td>26.02</td>
<td>26.97</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>28.74</td>
<td>27.85</td>
<td>28.29</td>
</tr>
</tbody>
</table>
Answer-Key Preparation

Is it better to fill slots by hand or review system-generated fills?
Impact of Web Access

Micro-Average Accuracy/F-Measure

- SMU_SIS(EL)
- Siel(EL)
- Icc(EL)
- Stanford(SF)
Conclusions

- KBP2010 was marked by a growing variety of tasks and a growing pool of participants

- Entity Linking
  - General improvement in performance over last year’s results
  - System performance approached and in some cases (and for some entity types) exceeded the human benchmark
  - Performance on new web data was close to newswire
  - Optional task was slightly but not significantly worse

- Slot Filling
  - A wide variety of approaches were represented
  - Substantial training corpora produced
  - A better understanding of the shortcoming of these approaches will lead to continued progress