Effective Slot Filling Based on Shallow Distant Supervision Methods

Benjamin Roth, Tassilo Barth, Michael Wiegand, Mittul Singh, Dietrich Klakow

Spoken Language Systems (LSV), Saarland University

November 18, 2013

Outline

TAC KBP English Slot Filling

Corpus

… **Marc Bolland**, **50**, former **CEO** of **Morrison Supermarkets PLC**, is joining **M&S** …

Queries

```
...
 <query id="ID_002">
   <name>Marc Bolland</name>
   <enttype>PER</enttype>
 </query>
…
 <query id="ID_100">
   <name>Galleon Group</name>
   <enttype>ORG</enttype>
 </query>
...
```

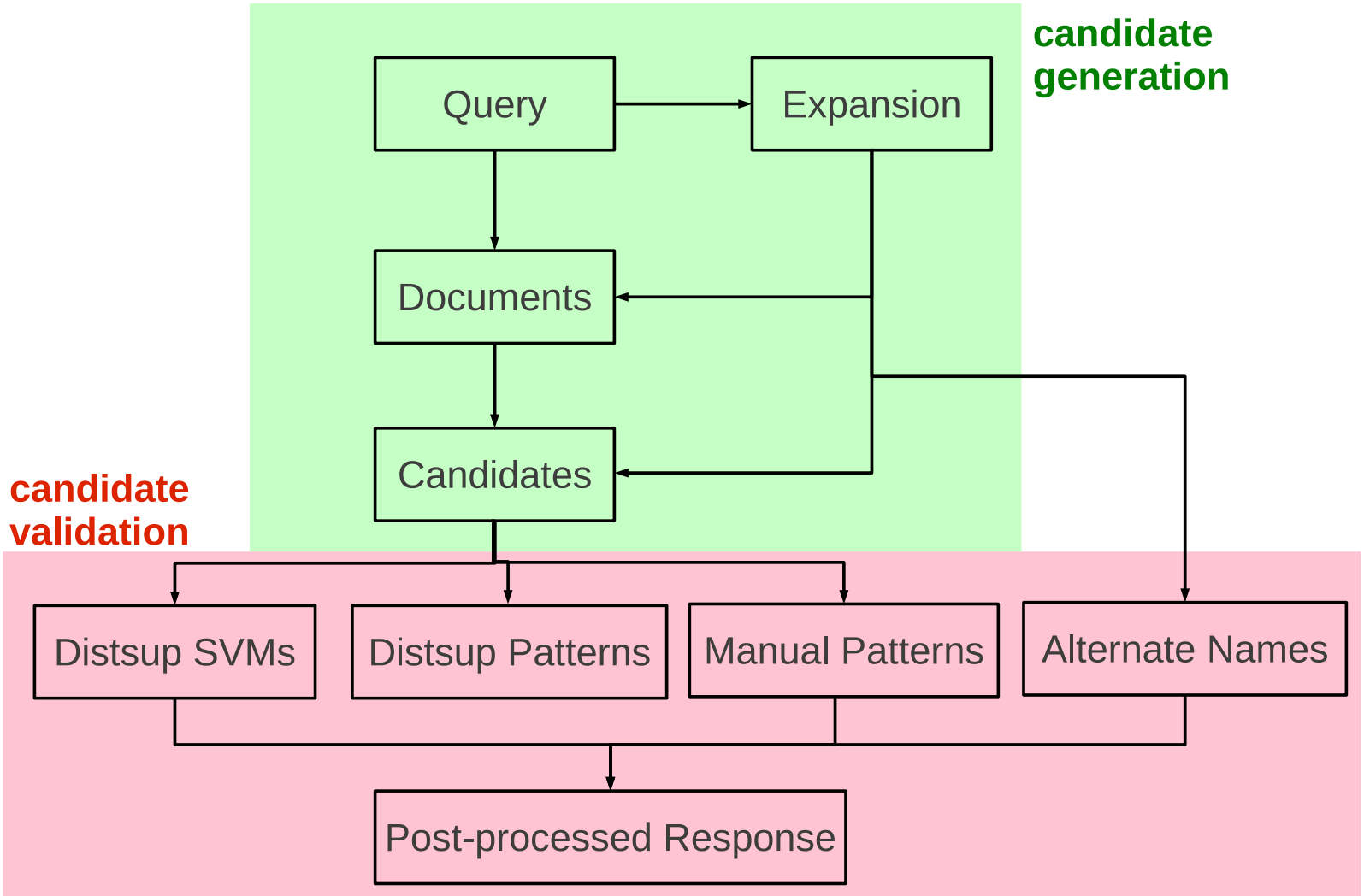
System Response

```
...
ID_002   per:age   50
ID_002   per:title   CEO
ID_002   per:employee_or_member_of   Morrison Supermarkets PLC
ID_002   per:employee_or_member_of   M&S
…
```

LSV / Saarland University 2013 Slot Filling System

- Modular and easily extensible distant supervision relation extractor
- Using shallow textual representations and features
- Based on LSV 2012 system [Roth et al., 2012]
  - same training data
  - same architecture
  - improved algorithms & context modeling

Data Flow


candidate generation
Query
Expansion
Documents
Candidates
candidate validation
Distsup SVMs
Distsup Patterns
Manual Patterns
Alternate Names
Post-processed Response

Outline

Candidate Generation

- Entity expansion based on Wikipedia anchor text language models
  - Query: *"Badr Organization"*
  - Expansion: *"Badr Brigade"*, *"Badr Organisation"*, *"Badr Brigades"*, *"Badr"*, *"Badr Corps"*
  - Also used for removing redundant answers (postprocessing)

Candidate Generation

- Entity expansion based on Wikipedia anchor text language models
  - Query: *"Badr Organization"*
  - Expansion: *"Badr Brigade"*, *"Badr Organisation"*, *"Badr Brigades"*, *"Badr"*, *"Badr Corps"*
  - Also used for removing redundant answers (postprocessing)
- Document retrieval
  - Lucene index
  - Selection of expansion terms based on point-wise mutual information

Candidate Generation

- Entity expansion based on Wikipedia anchor text language models
  - Query: *"Badr Organization"*
  - Expansion: *"Badr Brigade"*, *"Badr Organisation"*, *"Badr Brigades"*, *"Badr"*, *"Badr Corps"*
  - Also used for removing redundant answers (postprocessing)
- Document retrieval
  - Lucene index
  - Selection of expansion terms based on point-wise mutual information
- Candidate matching
  - NE Tagger [Chrupała and Klakow, 2010]
  - NE types from Freebase: CAUSE-OF-DEATH, JOB-TITLE, CRIMINAL-CHARGES, RELIGION

Outline

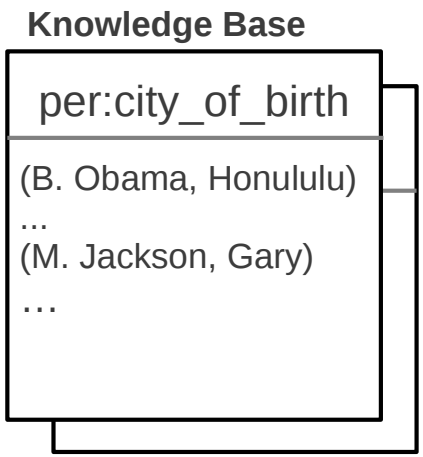
Candidate Validation Modules

- Distant Supervision SVM Classifiers
- Distant Supervision Patterns
- Manual Patterns
- Alternate Names from Query Expansion

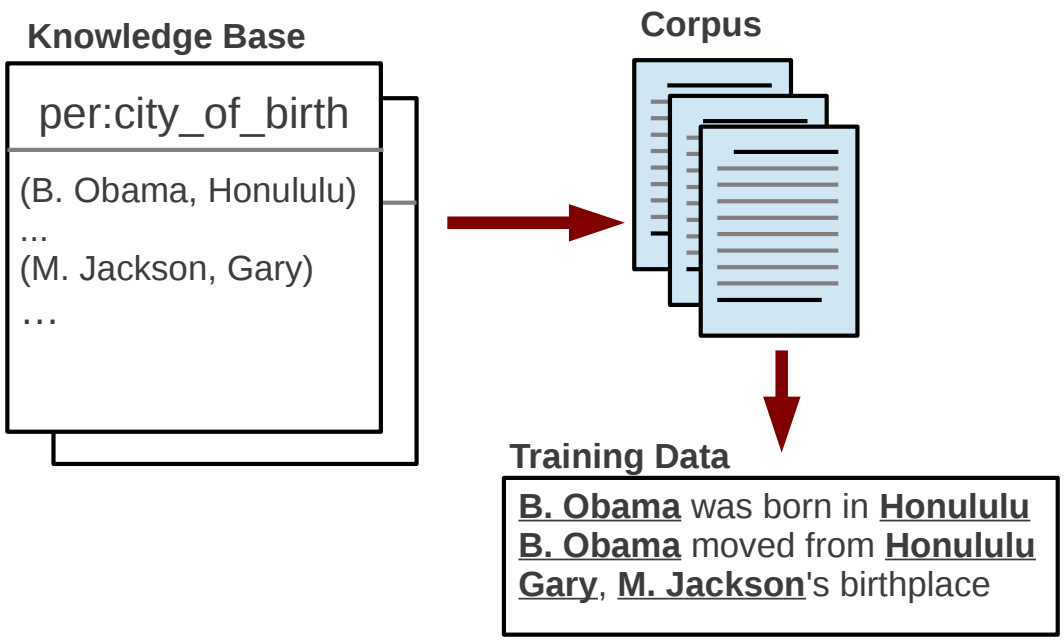
Candidate Validation Modules

- Distant Supervision SVM Classifiers
- Distant Supervision Patterns
- Manual Patterns
- Alternate Names from Query Expansion

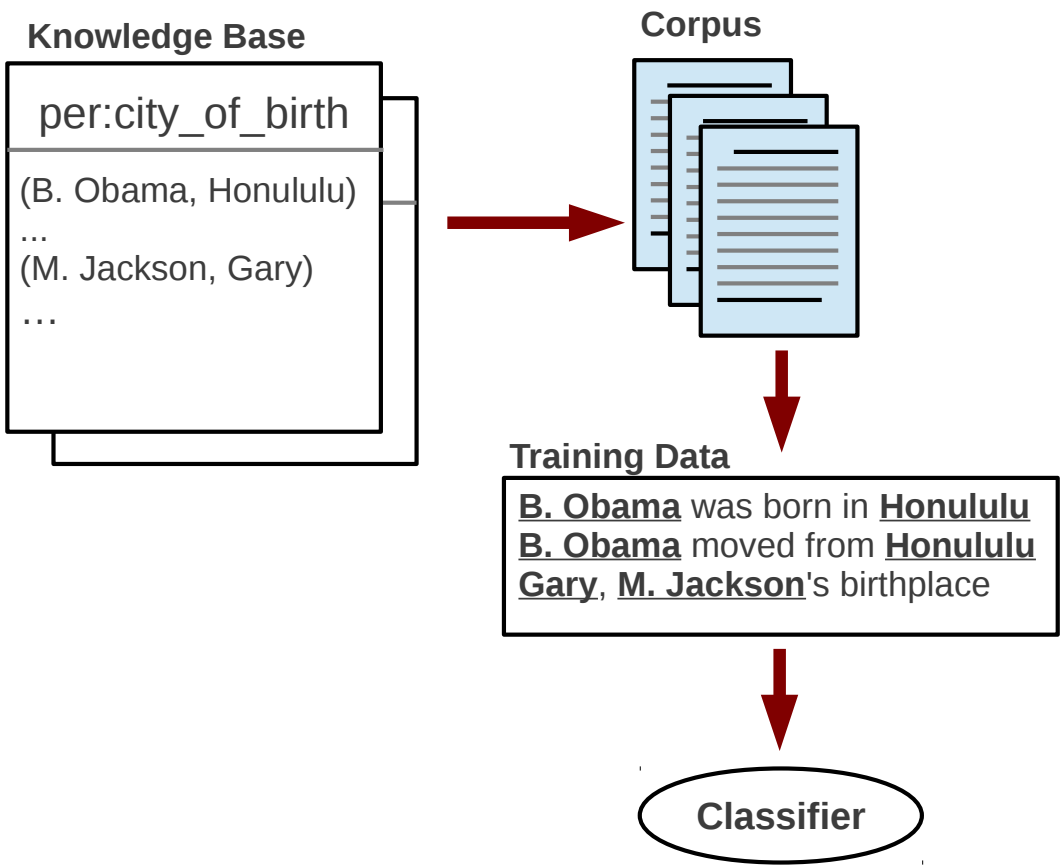
Distant Supervision



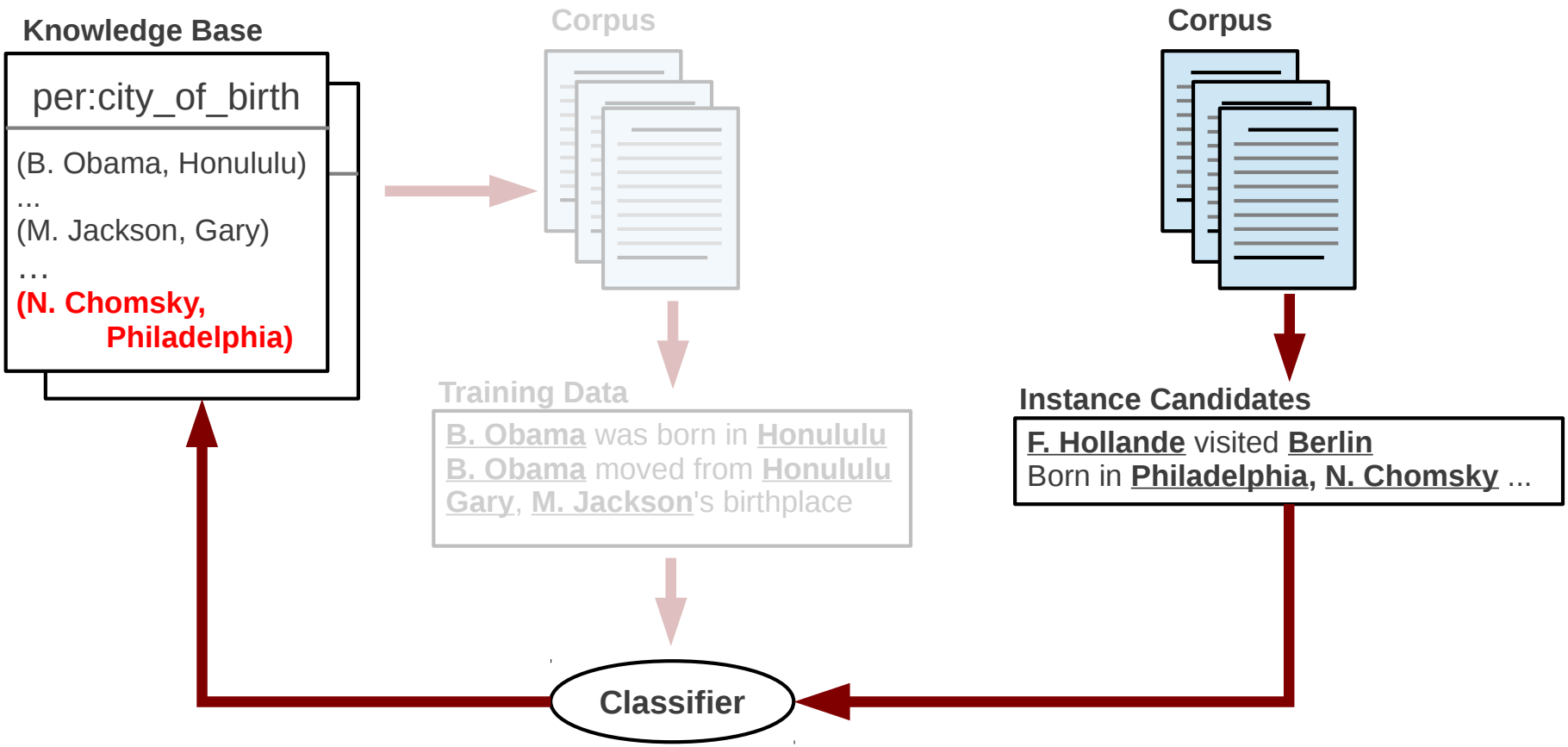
Distant Supervision



Distant Supervision



Distant Supervision



Distant Supervision (DS) SVM Classifiers

- "Workhorse" for candidate validation.

Distant Supervision (DS) SVM Classifiers

- "Workhorse" for candidate validation.
- Argument pairs for training data
  - Freebase
  - Pattern matches

Distant Supervision (DS) SVM Classifiers

- "Workhorse" for candidate validation.
- Argument pairs for training data
  - Freebase
  - Pattern matches
- Minimalistic feature set
  - n-grams between relation arguments
  - n-grams outside relation arguments
  - *sparse* (or *skip*) n-grams
  - marking of argument order for every feature

Distant Supervision (DS) SVM Classifiers

- "Workhorse" for candidate validation.
- Argument pairs for training data
  - Freebase
  - Pattern matches
- Minimalistic feature set
  - n-grams between relation arguments
  - n-grams outside relation arguments
  - *sparse* (or *skip*) n-grams
  - marking of argument order for every feature
- Training scheme:
  - *aggregate* training
  - *global* parameter tuning

DS SVMs: Training

- One binary SVM per relation

DS SVMs: Training

- One binary SVM per relation
- *Aggregate training*
  - Training sentences are aggregegaged per argument pair
  - Feature weights averaged
  - Better generalization than single-sentence training

DS SVMs: Training

- One binary SVM per relation
- *Aggregate training*
  - Training sentences are aggregaged per argument pair
  - Feature weights averaged
  - Better generalization than single-sentence training
- *Parameter tuning*
  - Misclassification cost tuning is essential
  - Optimizing per-relation cost parameter does not lead to global optimum
    $\Rightarrow$ Greedy parameter tuning algorithm for global $F1$ optimization

Distant Supervision Patterns

- Surface patterns from DS data
  - with *"goodness"* scores
  - `org:alternate_names`
    `0.9784 [ARG1] , abbreviated [ARG2]`
    `0.4023 [ARG2] is the core division of [ARG1]`

Distant Supervision Patterns

- Surface patterns from DS data
  - with *"goodness"* scores
  - `org:alternate_names`
    `0.9784 [ARG1] , abbreviated [ARG2]`
    `0.4023 [ARG2] is the core division of [ARG1]`
- Combination of DS noise reduction models [Roth and Klakow, 2013]
  - discriminative *at-least-one* perceptron model:

$$P(relation|pattern, \theta)$$

  - generative hierarchical topic model:

$$n(pattern, topic(relation))$$

  - relative frequency of pattern:

$$\frac{n(pattern, relation)}{n(pattern)}$$

Outline

Effect of Removing Single Components (one at a time)

Component	P	R	F1	F1 gain
LSV main run	42.5	33.2	37.3	
−Query expansion	41.1	17.5	24.5	+12.8
−Distsup SVM classifier	53.3	21.8	30.9	+6.4
−Distsup patterns	39.6	28.6	33.2	+4.1
−Manual patterns	38.2	29.5	33.2	+4.1
−Alternate names	41.1	31.0	35.4	+1.9

Effect of Removing Single Components (one at a time)

Component	P	R	F1	F1 gain
LSV main run	42.5	33.2	37.3	
−Query expansion	41.1	17.5	24.5	+12.8
−Distsup SVM classifier	53.3	21.8	30.9	+6.4
−Distsup patterns	39.6	28.6	33.2	+4.1
−Manual patterns	38.2	29.5	33.2	+4.1
−Alternate names	41.1	31.0	35.4	+1.9

Effect of Removing Single Components (one at a time)

Component	P	R	F1	F1 gain
LSV main run	42.5	33.2	37.3	
−Query expansion	41.1	17.5	24.5	+12.8
−Distsup SVM classifier	53.3	21.8	30.9	+6.4
−Distsup patterns	39.6	28.6	33.2	+4.1
−Manual patterns	38.2	29.5	33.2	+4.1
−Alternate names	41.1	31.0	35.4	+1.9

Effect of Removing Single Components (one at a time)

Component	P	R	F1	F1 gain
LSV main run	42.5	33.2	37.3	
−Query expansion	41.1	17.5	24.5	+12.8
−Distsup SVM classifier	53.3	21.8	30.9	+6.4
−Distsup patterns	39.6	28.6	33.2	+4.1
−Manual patterns	38.2	29.5	33.2	+4.1
−Alternate names	41.1	31.0	35.4	+1.9

Effect of Removing Single Components (one at a time)

Component	P	R	F1	F1 gain
LSV main run	42.5	33.2	37.3	
−Query expansion	41.1	17.5	24.5	+12.8
−Distsup SVM classifier	53.3	21.8	30.9	+6.4
−Distsup patterns	39.6	28.6	33.2	+4.1
−Manual patterns	38.2	29.5	33.2	+4.1
−Alternate names	41.1	31.0	35.4	+1.9

Effect of Removing Single Components (one at a time)

Component	P	R	F1	F1 gain
LSV main run	42.5	33.2	37.3	
−Query expansion	41.1	17.5	24.5	+12.8
−Distsup SVM classifier	53.3	21.8	30.9	+6.4
−Distsup patterns	39.6	28.6	33.2	+4.1
−Manual patterns	38.2	29.5	33.2	+4.1
−Alternate names	41.1	31.0	35.4	+1.9

Single Component Performance

Component	P	R	F1
Distsup SVM classifier	34.7	23.6	28.1
Distsup Patterns	42.7	15.6	22.9
Manual patterns	50.2	10.3	17.1
Alternate names	54.2	1.8	3.4

Bottleneck: Candidate Generation

- Lost recall on candidate level cannot be undone by validation modules.
- Query and argument matching is of crucial importance.
- Recall analysis (on 2012 queries):
  - good recall on document level
  - big potential on candidate sentence extraction

Query expansion	document recall	candidate recall	end-to-end F1
yes	90.2	58.8	32.1
no	87.7	34.4	23.0

Outline

Conclusion

- TAC KBP English Slot-Filling

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching
    - Relation modeling

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching
    - Relation modeling
- LSV system

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching
    - Relation modeling
- LSV system
  - Modular, shallow approach

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching
    - Relation modeling
- LSV system
  - Modular, shallow approach
  - Query expansion: Wikipedia anchor text

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching
    - Relation modeling
- LSV system
  - Modular, shallow approach
  - Query expansion: Wikipedia anchor text
  - N-gram-based distant supervision SVMs

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching
    - Relation modeling
- LSV system
  - Modular, shallow approach
  - Query expansion: Wikipedia anchor text
  - N-gram-based distant supervision SVMs
  - Scored distant supervision surface patterns

Conclusion

- TAC KBP English Slot-Filling
  - Query-driven relation extraction
  - Subtasks
    - Candidate retrieval and matching
    - Relation modeling
- LSV system
  - Modular, shallow approach
  - Query expansion: Wikipedia anchor text
  - N-gram-based distant supervision SVMs
  - Scored distant supervision surface patterns
- More details and analysis in our workshop paper!