Introduction

**Goal:** leverage behavioral web search and browse sessions to bootstrap multi-turn spoken dialog systems (SDS)

---

### Entity Mention Identification
- Linear-chain Conditional Random Fields (CRF): identify all the entity mentions in each query.

### Entity Type Determination
- CRFs trained with missing labels: replace the turns with no entity matched with missing labels.

### Relation Extraction
- Relation Templates
- Entity-Relation-Entity (E-R-E)
- Type-Relation-Entity (T-R-E)
- Multiclass logistic regression classifiers (one vs. all strategy)

---

### Web Search Query Interpretation

**Data**
- **Web Search Session Log**
- 12 weeks of IE search/browse sessions from March 18 to June 9, 2013
- **Domain:** Film
- 297,352 sessions in total.
- **Knowledge Base:** extension from Freebase.
- **Training via Distant Supervision**

#### Table: # Entity type # Entity # Relation Type # Relation

| Training | 44 | 215,053 | 33 | 20,197 |
| Test     | 44 | 104,657 | 27 | 5,255  |
| Test Only| 0  | 29,557  | 0  | 1,104  |

---

### Pilot Study: Entity-based Search Session Modeling

**Results**
- Session = sequence of entities.
- Session likelihoods: Markov models with different history.
- 2 smoothing techniques: additive smoothing & linear interpolation.

---

### Experimental Results

- Baseline 1 (Gazetteer Only): string matching with the gazetteer.
- Baseline 2 (No Session): log-linear model using the same features.
- CRF (mention + type): jointly learn the entity mention and entity type.
- Session model (CRF): removed turns without entity match.
- Session model (CRF w/ missing label): replace turns without entity match with missing labels.

#### Table: Recall on Entity Extraction

| Baseline 1 (Gazetteer Only) | 84.00 |
| Baseline 2 (No Session)     | 86.34 |
| CRF (mention + type)        | 87.33 |
| Session model (CRF)         | 88.06 |
| Session model (CRF w/ missing label) | 90.38 |

---

### Discussion

- Our session-based models trained on web search logs outperform two nontrivial baselines and state-of-the-art methods in most scenarios.
- Our models effectively leverage behavioral web search/browse patterns to improve the entity extraction performance in spoken dialog systems.
- Genre detection is relatively poor. Due to different context people use for genres, e.g., "Disney" is a genre in "Disney movie", but is a place name in "Disney world".
- Future Work: add co-reference resolution

---

### Contributions

- 1st successful method to directly leverage web search & browse sessions to bootstrap statistical multi-turn models for SDS
- Significantly improved multi-turn SDS with a scalable method