Data-Driven Approaches to Concept-to-Text Generation

Ioannis Konstas, Mirella Lapata

Institute for Language, Cognition and Computation
University of Edinburgh

Heriot-Watt University
5 September 2012
Outline

1. Introduction

2. Generative Model
   - Konstas and Lapata, NAACL 2012

3. Discriminative Reranking Model
   - Konstas and Lapata, ACL 2012

4. Demos
**Concept-to-text** generation refers to the task of automatically producing textual output from nonlinguistic input (Reiter and Dale, 2000).
Concept-to-text generation refers to the task of automatically producing textual output from nonlinguistic input (Reiter and Dale, 2000).

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong></td>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cloud Sky Cover</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong></td>
<td>06:00-09:00</td>
</tr>
<tr>
<td></td>
<td>09:00-12:00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Speed</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong></td>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Direction</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong></td>
<td>06:00-21:00</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.
Concept-to-text generation refers to the task of automatically producing textual output from nonlinguistic input (Reiter and Dale, 2000).

<table>
<thead>
<tr>
<th>Flight</th>
<th>Day Number</th>
<th>Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>from</td>
<td>number</td>
<td>month</td>
</tr>
<tr>
<td>edinburgh</td>
<td>dep/ar</td>
<td>dep/ar</td>
</tr>
<tr>
<td>to</td>
<td>7</td>
<td>july</td>
</tr>
<tr>
<td>jeju</td>
<td>departure</td>
<td>departure</td>
</tr>
</tbody>
</table>

Condition:
- arg1 arg2 type
- arrival_time 1600 <

Search:
- type what
- query flight

Give me the flights leaving Edinburgh July seventh coming back to Jeju Island before 4pm.
Traditional NLG Pipeline

Input Data → Content Selection → Surface Realisation → Text → Communicative Goal
Traditional NLG Pipeline
Unsupervised Concept-to-text Generation with Hypergraphs
Joint Forest Rescoring with Hypergraphs
Joint Forest Rescoring with Hypergraphs

**Training**

**Temperature**

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

**Cloud Sky Cover**

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

**Wind Speed**

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

**Wind Direction**

<table>
<thead>
<tr>
<th>Time</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>S</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.
Joint Forest Rescoring with Hypergraphs

Training

Temperature

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

Cloud Sky Cover

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

Wind Speed

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

Wind Direction

<table>
<thead>
<tr>
<th>Time</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>S</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.

S → R(start)
R(r_i,t) → FS(r_j, start) R(r_j, r_i)
R(r_i,t) → FS(r_j, start)
FS(r, r.f_i) → F(r, r.f_j) FS(r, r.f_j)
FS(r, r.f_i) → F(r, r.f_i)
F(r, r.f) → W(r, r.f) F(r, r.f)
F(r, r.f) → W(r, r.f)
W(r, r.f) → \alpha
W(r, r.f) → g(f.v)
Joint Forest Rescoring with Hypergraphs

**Temperature**

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

**Wind Speed**

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

**Cloud Sky Cover**

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

**Wind Direction**

<table>
<thead>
<tr>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.

```
S → R(start)
R(r,t) → FS(r_j, start)R(r_j,t)
R(r,t) → FS(r_j, start)
FS(r,r.f) → F(r,r.f)FS(r,r.f)
FS(r,r.f) → F(r,r.f)
F(r,r.f) → W(r,r.f)F(r,r.f)
F(r,r.f) → W(r,r.f)
W(r,r.f) → α
W(r,r.f) → g(f,v)
```

Training

FS_0,1(temp_1,start) → F_0,1(temp_1,min) → F_0,1(temp_1,max)
FS_0,2(temp_1,start) → F_0,2(temp_1,min) → F_0,2(temp_1,max)
FS_1,2(temp_1,start) → W_4,5(skyCover_1,t)
Joint Forest Rescoring with Hypergraphs

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Wind Speed</td>
<td>Time</td>
<td>Min</td>
<td>Mean</td>
<td>Max</td>
</tr>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Cloud Sky Cover</td>
<td>Time</td>
<td>Percent (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wind Direction</td>
<td>Time</td>
<td>Mode</td>
<td></td>
<td></td>
</tr>
<tr>
<td>06:00-21:00</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.

S → \( R(\text{start}) \)
\( R(r_i, t) \rightarrow FS(r_j, \text{start})R(r_j, t) \)
\( R(r_i, t) \rightarrow FS(r_j, \text{start}) \)
\( FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j) \)
\( FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j) \)
\( F(r, r.f) \rightarrow W(r, r.f)F(r, r.f) \)
\( F(r, r.f) \rightarrow W(r, r.f)F(r, r.f) \)
\( W(r, r.f) \rightarrow \alpha \)
\( W(r, r.f) \rightarrow g(f.v) \)

Train using EM

FS\(_{0,1}\)(temp\(_1\),start)
FS\(_{0,2}\)(temp\(_1\),start)
FS\(_{1,2}\)(temp\(_1\),start)

k-best decoding via LM integration

\( F_0,1(\text{temp}_{1,\text{min}}) \)
\( F_0,1(\text{temp}_{1,\text{max}}) \)
\( F_0,2(\text{temp}_{1,\text{min}}) \)
\( F_0,2(\text{temp}_{1,\text{max}}) \)

Concept-to-Text Generation

Konstas, Lapata (ILCC)

5 September 2012
### Our Approach

#### Temperature

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

#### Cloud Sky Cover

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.

#### Wind Speed

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

#### Wind Direction

<table>
<thead>
<tr>
<th>Time</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>S</td>
</tr>
</tbody>
</table>
Our Approach

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cloud Sky Cover</th>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
<td></td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Speed</th>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Direction</th>
<th>Time</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>S</td>
<td></td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.

Partly cloudy, with a low around 9. Breezy, with a south wind between 15 and 30 mph.
Our Approach

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Speed</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cloud Sky Cover</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Direction</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>S</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.
Our Approach

<table>
<thead>
<tr>
<th>Temperature</th>
<th>06:00-21:00</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>9</td>
<td>15</td>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Speed</th>
<th>06:00-21:00</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>15</td>
<td>20</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cloud Sky Cover</th>
<th>06:00-09:00</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>25-50</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Direction</th>
<th>06:00-21:00</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>S</td>
<td></td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.
## Our Approach

### Temperature

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

### Wind Speed

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>

### Wind Direction

<table>
<thead>
<tr>
<th>Time</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>S</td>
</tr>
</tbody>
</table>

### Cloud Sky Cover

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.
Our Approach

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>06:00-21:00</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cloud Sky Cover</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>06:00-09:00</td>
</tr>
<tr>
<td></td>
<td>09:00-12:00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Speed</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>06:00-21:00</td>
<td>15</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wind Direction</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>06:00-21:00</td>
</tr>
</tbody>
</table>

Cloudy, with a low around 10. South wind between 15 and 30 mph.

Partly cloudy, with a low around 9. Breezy, with a south wind between 15 and 30 mph.
Related Work

Angeli et al., 2010

- Unified content selection and surface realisation
- Obtain alignments from Liang et al. (2009)
- Sequence of discriminative (log-linear) local decisions (records - fields - templates)
Related Work

Angeli et al., 2010
- Unified content selection and surface realisation
- Obtain alignments from Liang et al. (2009)
- Sequence of discriminative (log-linear) local decisions (records - fields - templates)

Our approach
- **Joint** model allows for more global decisions
- **Hypergraphs** are a compact representation and allow for efficient inference (k-best decoding via cube pruning)
- **Forest rescoring** allows for rescoring k-best trees at all internal nodes via LM integration
Input

- Input: database records \( d \)
- Output: words \( w \) corresponding to some records of \( d \)
- Each record \( r \in d \) has a type \( r.t \) and fields \( f \)
- Fields have values \( f.v \) and types \( f.t \) (integer, categorical)

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

mostly cloudy,
Input

- Input: database records $d$
- Output: words $w$ corresponding to some records of $d$
- Each record $r \in d$ has a type $r.t$ and fields $f$
- Fields have values $f.v$ and types $f.t$ (integer, categorical)

Cloud Sky Cover

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

mostly cloudy,
Input

- Input: database records $d$
- Output: words $w$ corresponding to some records of $d$
- Each record $r \in d$ has a type $r.t$ and fields $f$
- Fields have values $f.v$ and types $f.t$ (integer, categorical)

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

mostly cloudy,
Input

- Input: database records \( d \)
- Output: words \( w \) corresponding to some records of \( d \)
- Each record \( r \in d \) has a type \( r.t \) and fields \( f \)
- Fields have values \( f.v \) and types \( f.t \) (integer, categorical)

### Cloud Sky Cover

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-09:00</td>
<td>25-50</td>
</tr>
<tr>
<td>09:00-12:00</td>
<td>50-75</td>
</tr>
</tbody>
</table>

mostly cloudy,
Input

- Input: database records \( d \)
- Output: words \( w \) corresponding to some records of \( d \)
- Each record \( r \in d \) has a type \( r.t \) and fields \( f \)
- Fields have values \( f.v \) and types \( f.t \) (integer, categorical)
Grammar Definition

1. $S \rightarrow R(start)$
2. $R(r_i.t) \rightarrow FS(r_j, start)R(r_j.t)$
3. $R(r_i.t) \rightarrow FS(r_j, start)$
4. $FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j)$
5. $FS(r, r.f_i) \rightarrow F(r, r.f_j)$
6. $F(r, r.f) \rightarrow W(r, r.f)F(r, r.f)$
7. $F(r, r.f) \rightarrow W(r, r.f)$
8. $W(r, r.f) \rightarrow \alpha$
9. $W(r, r.f) \rightarrow g(f.v)$
Grammar Definition

\[ \text{R}(\text{skyCover}_1.t) \rightarrow \text{FS}(\text{temperature}_1, \text{start})\text{R}(\text{temperature}_1.t) \]
Grammar Definition

$$FS(wSpeed_1, \text{min}) \rightarrow F(wSpeed_1, \text{max})FS(wSpeed_1, \text{max})$$

1. $$S \rightarrow R(\text{start})$$
2. $$R(r_i.t) \rightarrow FS(r_j, \text{start})R(r_j.t)$$
3. $$R(r_i.t) \rightarrow FS(r_j, \text{start})$$
4. $$FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j)$$
5. $$FS(r, r.f_i) \rightarrow F(r, r.f_j)$$
6. $$F(r, r.f) \rightarrow W(r, r.f)F(r, r.f)$$
7. $$F(r, r.f) \rightarrow W(r, r.f)$$
8. $$W(r, r.f) \rightarrow \alpha$$
9. $$W(r, r.f) \rightarrow g(f.v)$$
Grammar Definition

F(\textit{gust}_1, \textit{mean}) \rightarrow W(\textit{gust}_1, \textit{mean})F(\textit{gust}_1, \textit{mean})

1. S \rightarrow R(\textit{start})
2. R(r_i.t) \rightarrow FS(r_j, \textit{start})R(r_j.t)
3. R(r_i.t) \rightarrow FS(r_j, \textit{start})
4. FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j)
5. FS(r, r.f_i) \rightarrow F(r, r.f_j)
6. F(r, r.f) \rightarrow W(r, r.f)F(r, r.f)
7. F(r, r.f) \rightarrow W(r, r.f)
8. W(r, r.f) \rightarrow \alpha
9. W(r, r.f) \rightarrow g(f.v)
Grammar Definition

\[
W(\text{skyCover}_1, \%) \rightarrow \text{cloudy} [\% . v = '75-100']
\]

1. \( S \rightarrow R(\text{start}) \)
2. \( R(r_j . t) \rightarrow \text{FS}(r_j, \text{start}) R(r_j . t) \)
3. \( R(r_j . t) \rightarrow \text{FS}(r_j, \text{start}) \)
4. \( \text{FS}(r, r.f_i) \rightarrow F(r, r.f_j) \text{FS}(r, r.f_j) \)
5. \( \text{FS}(r, r.f_i) \rightarrow F(r, r.f_j) \)
6. \( F(r, r.f) \rightarrow W(r, r.f) F(r, r.f) \)
7. \( F(r, r.f) \rightarrow W(r, r.f) \)
8. \( W(r, r.f) \rightarrow \alpha \)
9. \( W(r, r.f) \rightarrow g(f.v) \)
Hypergraphs

Definition

An ordered hypergraph $H$ is a tuple $⟨N, E, t, R⟩$, where $N$ is a finite set of nodes, $E$ is a finite set of hyperarcs, $t ∈ N$ is a target node and $R$ is the set of weights. Each hyperarc $e ∈ E$ is a triple $e = ⟨T(e), h(e), f(e)⟩$, where $h(e) ∈ N$ is its head node, $T(e) ∈ N^*$ is a set of tail nodes and $f(e)$ is a monotonic weight function $R_{|T(e)|}$ to $R$. 

![Diagram of a hypergraph with nodes t, a, and b, and a hyperarc f(e) connecting t to the other nodes.](attachment://hypergraph_diagram.png)
Hypergraphs

**Definition**

An ordered hypergraph $H$ is a tuple $\langle N, E, t, R \rangle$, where $N$ is a finite set of nodes, $E$ is a finite set of hyperarcs, $t \in N$ is a target node and $R$ is the set of weights. Each hyperarc $e \in E$ is a triple $e = \langle T(e), h(e), f(e) \rangle$, where $h(e) \in N$ is its head node, $T(e) \in N^*$ is a set of tail nodes and $f(e)$ is a monotonic weight function $R_{|T(e)|}$ to $R$. 

![Diagram of a hypergraph with nodes t, a, b and a hyperarc f(e) connecting t to b]
Hypergraphs

Definition

An ordered hypergraph $H$ is a tuple $\langle N, E, t, R \rangle$, where $N$ is a finite set of nodes, $E$ is a finite set of hyperarcs, $t \in N$ is a target node and $R$ is the set of weights. Each hyperarc $e \in E$ is a triple $e = \langle T(e), h(e), f(e) \rangle$, where $h(e) \in N$ is its head node, $T(e) \in N^*$ is a set of tail nodes and $f(e)$ is a monotonic weight function $R_{|T(e)|}$ to $R$. 
A hypergraph $H$ is a tuple $\langle N, E, t, R \rangle$, where $N$ is a finite set of nodes, $E$ is a finite set of hyperarcs, $t \in N$ is a target node and $R$ is the set of weights. Each hyperarc $e \in E$ is a triple $e = \langle T(e), h(e), f(e) \rangle$, where $h(e) \in N$ is its head node, $T(e) \in N^*$ is a set of tail nodes and $f(e)$ is a monotonic weight function $R_{|T(e)|}$ to $R$. 

![Diagram of a hypergraph with nodes t, a, and b, and a hyperarc labeled f(e)]
Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph $H : \langle N, E, t, R \rangle$

$$f(e) = f(FS_{5,7}(flight_{1}.t, start)) \otimes f(R_{7,9}(flight_{1}.t)) \otimes w(R(search_{1}.t) \rightarrow FS(flight_{1}, start) \ R(flight_{1}.t))$$

$$R(r_i.t) \rightarrow FS(r_j, start) R(r_j.t)$$
Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph $H : \langle N, E, t, R \rangle$

$$\text{R}\text{(search}_1\text{.t)}$$

$$\text{FS}\text{(flight}_1\text{.t,start)} \quad \text{R}\text{(flight}_1\text{.t)}$$

$$f(e) = f(\text{FS}_{5,7}\text{(flight}_1\text{.t, start)}) \otimes f(\text{R}_{7,9}\text{(flight}_1\text{.t)}) \otimes w(\text{R}\text{(search}_1\text{.t)} \rightarrow \text{FS}\text{(flight}_1\text{, start) R}\text{(flight}_1\text{.t)})$$

$$\text{R}\text{(r}_1\text{.t)} \rightarrow \text{FS}\text{(r}_j\text{, start) R}\text{(r}_j\text{.t)}$$
Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph $H : \langle N, E, t, R \rangle$

$$f(e) = f(FS_{5,7}(flight_1.t, start)) \otimes f(R_{7,9}(flight_1.t)) \otimes w(R(search_1.t) \rightarrow FS(flight_1, start) R(flight_1.t))$$

$$R(r_j.t) \rightarrow FS(r_j, start) R(r_j.t)$$
Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph \( H : \langle N, E, t, R \rangle \)

\[
f(e) = f(FS_{5,7}(flight_1.t, start)) \otimes f(R_{7,9}(flight_1.t)) \otimes w(R(search_1.t) \rightarrow FS(flight_1, start) R(flight_1.t))
\]

\[
R(r_i.t) \rightarrow FS(r_j, start) R(r_j.t)
\]
Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph $H : \langle N, E, t, R \rangle$

$$f(e) = f(FS_{5,7}(flight_1.t, start)) \otimes f(R_{7,9}(flight_1.t)) \otimes w(R(search_1.t) \rightarrow FS(flight_1, start) R(flight_1.t))$$

$$R(r_i.t) \rightarrow FS(r_j, start) R(r_j.t)$$
Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph $H: \langle N, E, t, R \rangle$

$$f(e) = f(FS_{5,7}(flight_1.t, start)) \otimes f(R_{7,9}(flight_1.t)) \otimes w(R(search_1.t) \rightarrow FS(flight_1, start) \ R(flight_1.t))$$

$$R(r_i.t) \rightarrow FS(r_j, start) \ R(r_j.t)$$
Hypergraph Example

```
FS_{0,1}(skyCover_1, start)
FS_{0,2}(skyCover_1, start)
FS_{1,2}(skyCover_1, start)
FS_{1,2}(temp_1, start)
FS_{0,1}(temp_1, start)

F_{0,1}(skyCover_1, %)
F_{0,1}(skyCover_1, time)
F_{0,2}(skyCover_1, %)
F_{0,2}(skyCover_1, time)
W_{0,1}(skyCover_1, %)
W_{0,1}(skyCover_1, time)
W_{1,2}(skyCover_1, %)
W_{1,2}(skyCover_1, time)

R_{0,7}(start)
R_{1,7}(skyCover_1, t)
R_{1,7}(temp_1, t)
R_{2,7}(skyCover_1, t)
R_{2,7}(temp_1, t)

F_{1,2}(temp_1, min) W_{1,2}(temp_1, min) g_{1,2}(min, v=10)
F_{1,2}(temp_1, max) W_{1,2}(temp_1, max) g_{1,2}(max, v=20)
```

with sunny

EM Training: dynamic program similar to the inside-outside algorithm
EM Training: dynamic program similar to the inside-outside algorithm
Hypergraph Example

EM Training: dynamic program similar to the inside-outside algorithm
**k-best Decoding**

\[
\arg \max_w P(w | d) = \arg \max_w P(w) \cdot P(d | w)
\]
### k-best Decoding

\[
\arg \max_w P(w \mid d) = \arg \max_w P(w) \cdot P(d \mid w)
\]

- Bottom-up Viterbi search
**k-best Decoding**

\[
\arg\max_w P(w|d) = \arg\max_w P(w) \cdot P(d|w)
\]

- Bottom-up Viterbi search
- Keep k-best derivations at each node, cube pruning (Chiang, 2007)
k-best Decoding

\[ \arg\max_w P(w \mid d) = \arg\max_w P(w) \cdot P(d \mid w) \]

- Bottom-up Viterbi search
- Keep k-best derivations at each node, cube pruning (Chiang, 2007)
- Nodes in hypergraph → +LM items (Huang and Chiang, 2007)
  e.g. \( R_{2,8}(temp_1.t)^{a \text{ low*15 degrees}} \)
k-best Decoding

\[ \text{arg max}_w P(w | d) = \text{arg max}_w P(w) \cdot P(d | w) \]

- Bottom-up Viterbi search
- Keep k-best derivations at each node, cube pruning (Chiang, 2007)
- Nodes in hypergraph \( \rightarrow \) +LM items (Huang and Chiang, 2007)
  - e.g. \( R_{2,8}(\text{temp}_1.t)^a \) low*15 degrees
- Root node: most grammatical and semantically correct derivation
k-best Decoding

Leaf nodes $\epsilon$ emit a k-best list of words

$W_{0,1}(\text{skyCover}_1.t,\%)$

$( mostly$

$( cloudy$

$( sunny$

$( ... )$
k-best Decoding

\[
\begin{align*}
&\left( \text{mostly cloudy} \ast \text{the morning} \right) \\
&\left( \text{mostly cloudy} \ast \text{after 11am} \right) \\
&\left( \text{mostly cloudy} \ast \text{then becoming} \right) \\
&\left( \text{mostly cloudy} \ast \text{cloudy} \right) \\
&\left( \text{most} \ast \text{cloudy} \right) \\
&\left( \text{mostly cloudy} \ast \text{sunny} \right)
\end{align*}
\]

\[
\begin{align*}
&\text{FS}_{0,5}(\text{skyCover}_1.t,\text{start}) \\
&\text{F}_{0,2}(\text{skyCover}_1.t,\%) \\
&\text{W}_{4,5}(\text{skyCover}_1.t,\text{time}) \\
&\text{W}_{0,1}(\text{skyCover}_1.t,\%) \\
&\text{W}_{1,2}(\text{skyCover}_1.t,\%)
\end{align*}
\]
k-best Decoding

\[
\begin{align*}
&\left( \text{mostly cloudy } \ast \text{ the morning} \right) \\
&\left( \text{mostly cloudy } \ast \text{ after 11am} \right) \\
&\left( \text{mostly cloudy } \ast \text{ then becoming} \right)
\end{align*}
\]

\[
\begin{align*}
&\left( \text{mostly cloudy} \right) \\
&\left( \text{mostly clouds} \right) \\
&\left( \text{cloudy} \right)
\end{align*}
\]

\[
\begin{align*}
&\left( \text{mostly cloudy} \right) \\
&\left( \text{sunny} \right)
\end{align*}
\]

\[
\begin{align*}
&\text{FS}_{0,5}(\text{skyCover}_1.t, \text{start}) \\
&\text{F}_{0,2}(\text{skyCover}_1.t, \%) \\
&\text{W}_{0,1}(\text{skyCover}_1.t, \%) \\
&\text{W}_{1,2}(\text{skyCover}_1.t, \%)
\end{align*}
\]

\[
\begin{align*}
&\text{W}_{4,5}(\text{skyCover}_1.t, \text{time}) \\
&\left( \text{morning} \right) \\
&\left( \text{11am} \right) \\
&\left( \text{after} \right) \\
&\left( \text{...} \right)
\end{align*}
\]
k-best Decoding

\[
\begin{align*}
(F_{0,2}(skyCover_{1.t},\%)) & \\
(W_{0,1}(skyCover_{1.t},\%)) & \\
(W_{1,2}(skyCover_{1.t},\%)) & \\
(W_{4,5}(skyCover_{1.t},\text{time})) & \\
(F_{0,5}(skyCover_{1.t},\text{start})) & \\
\end{align*}
\]
Experimental Setup

- **RoboCup**: simulated sportscasting [214 words] (Chen and Mooney, 2008)
- **WeatherGov**: weather reports [345 words] (Liang et al., 2009)
- **ATIS**: mapping from $\lambda$-version [927 words] (Zettlemoyer and Collins, 2007)
Experimental Setup

- **RoboCup**: simulated sportscasting [214 words] (Chen and Mooney, 2008)
- **WeatherGov**: weather reports [345 words] (Liang et al., 2009)
- **ATIS**: mapping from $\lambda$-version [927 words] (Zettlemoyer and Collins, 2007)
- Model parameters estimated on dev set ($k$-best, n-grams)
Experimental Setup

- **RoboCup**: simulated sportscasting [214 words] (Chen and Mooney, 2008)
- **WeatherGov**: weather reports [345 words] (Liang et al., 2009)
- **ATIS**: mapping from $\lambda$-version [927 words] (Zettlemoyer and Collins, 2007)
- Model parameters estimated on dev set ($k$-best, n-grams)
- Automatic evaluation: BLEU-4
- Human evaluation (MTurk): fluency, semantic correctness
Results: Automatic Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>RoboCup</th>
<th>WeatherGov</th>
<th>Atis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Best</td>
<td>10.79</td>
<td>8.64</td>
<td>11.85</td>
</tr>
<tr>
<td>k-Best</td>
<td>30.90</td>
<td>33.70</td>
<td>29.30</td>
</tr>
<tr>
<td>Angeli</td>
<td>28.70</td>
<td>38.40</td>
<td>26.77</td>
</tr>
</tbody>
</table>

- **RoboCup** results with fixed content selection;
- **WeatherGov** and **Atis** results with content selection and surface realization.
### Results: Automatic Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>RoboCup</th>
<th>WeatherGov</th>
<th>Atis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Best</td>
<td>10.79</td>
<td>8.64</td>
<td>11.85</td>
</tr>
<tr>
<td>k-Best</td>
<td>30.90</td>
<td>33.70</td>
<td>29.30</td>
</tr>
<tr>
<td>Angeli</td>
<td>28.70</td>
<td>38.40</td>
<td>26.77</td>
</tr>
</tbody>
</table>

- **RoboCup** results with fixed content selection;
- **WeatherGov** and **Atis** results with content selection and surface realization.
## Results: Human Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-Best</strong></td>
<td>2.47</td>
<td>2.33</td>
</tr>
<tr>
<td><strong>k-Best</strong></td>
<td>4.31</td>
<td>3.96</td>
</tr>
<tr>
<td><strong>ANGELI</strong></td>
<td>4.03</td>
<td>3.70</td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>4.47</td>
<td>4.37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-Best</strong></td>
<td>1.82</td>
<td>2.05</td>
</tr>
<tr>
<td><strong>k-Best</strong></td>
<td>3.92</td>
<td>3.30</td>
</tr>
<tr>
<td><strong>ANGELI</strong></td>
<td>4.26</td>
<td>3.60</td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>4.61</td>
<td>4.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-Best</strong></td>
<td>2.40</td>
<td>2.46</td>
</tr>
<tr>
<td><strong>k-Best</strong></td>
<td>4.01</td>
<td>3.87</td>
</tr>
<tr>
<td><strong>ANGELI</strong></td>
<td>3.56</td>
<td>3.33</td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>4.10</td>
<td>4.01</td>
</tr>
</tbody>
</table>
## Results: Human Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-BEST</strong></td>
<td>2.47</td>
<td>2.33</td>
</tr>
<tr>
<td><strong>k-BEST</strong></td>
<td>4.31</td>
<td>3.96</td>
</tr>
<tr>
<td><strong>ANGELI</strong></td>
<td>4.03</td>
<td>3.70</td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>4.47</td>
<td>4.37</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-BEST</strong></td>
<td>1.82</td>
<td>2.05</td>
</tr>
<tr>
<td><strong>k-BEST</strong></td>
<td>3.92</td>
<td>3.30</td>
</tr>
<tr>
<td><strong>ANGELI</strong></td>
<td>4.26</td>
<td>3.60</td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>4.61</td>
<td>4.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-BEST</strong></td>
<td>2.40</td>
<td>2.46</td>
</tr>
<tr>
<td><strong>k-BEST</strong></td>
<td>4.01</td>
<td>3.87</td>
</tr>
<tr>
<td><strong>ANGELI</strong></td>
<td>3.56</td>
<td>3.33</td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>4.10</td>
<td>4.01</td>
</tr>
</tbody>
</table>
### Results: Human Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ROBOCUP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-BEST</td>
<td>2.47</td>
<td>2.33</td>
</tr>
<tr>
<td>k-BEST</td>
<td>4.31</td>
<td>3.96</td>
</tr>
<tr>
<td>ANGELI</td>
<td>4.03</td>
<td>3.70</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.47</td>
<td>4.37</td>
</tr>
<tr>
<td><strong>WEATHERGOV</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-BEST</td>
<td>1.82</td>
<td>2.05</td>
</tr>
<tr>
<td>k-BEST</td>
<td>3.92</td>
<td>3.30</td>
</tr>
<tr>
<td>ANGELI</td>
<td>4.26</td>
<td>3.60</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.61</td>
<td>4.03</td>
</tr>
<tr>
<td><strong>ATIS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-BEST</td>
<td>2.40</td>
<td>2.46</td>
</tr>
<tr>
<td>k-BEST</td>
<td>4.01</td>
<td>3.87</td>
</tr>
<tr>
<td>ANGELI</td>
<td>3.56</td>
<td>3.33</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.10</td>
<td>4.01</td>
</tr>
</tbody>
</table>
## Results: Human Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ROBOCUP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-BEST</td>
<td>2.47</td>
<td>2.33</td>
</tr>
<tr>
<td>k-BEST</td>
<td>4.31</td>
<td>3.96</td>
</tr>
<tr>
<td>ANGELI</td>
<td>4.03</td>
<td>3.70</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.47</td>
<td>4.37</td>
</tr>
<tr>
<td><strong>WEATHERGOV</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-BEST</td>
<td>1.82</td>
<td>2.05</td>
</tr>
<tr>
<td>k-BEST</td>
<td>3.92</td>
<td>3.30</td>
</tr>
<tr>
<td>ANGELI</td>
<td>4.26</td>
<td>3.60</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.61</td>
<td>4.03</td>
</tr>
<tr>
<td><strong>ATIS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-BEST</td>
<td>2.40</td>
<td>2.46</td>
</tr>
<tr>
<td>k-BEST</td>
<td>4.01</td>
<td>3.87</td>
</tr>
<tr>
<td>ANGELI</td>
<td>3.56</td>
<td>3.33</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.10</td>
<td>4.01</td>
</tr>
</tbody>
</table>
**WEATHERGOV**

### Temperature

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>30</td>
<td>38</td>
<td>44</td>
</tr>
</tbody>
</table>

### Cloud Sky Cover

<table>
<thead>
<tr>
<th>Time</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>75-100</td>
</tr>
</tbody>
</table>

### Wind Speed

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

### Wind Direction

<table>
<thead>
<tr>
<th>Time</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>ENE</td>
</tr>
</tbody>
</table>

### Precipitation Potential (%)

<table>
<thead>
<tr>
<th>Time</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:00-21:00</td>
<td>9</td>
<td>20</td>
<td>35</td>
</tr>
</tbody>
</table>

**k-Best:** A chance of rain showers before 11am. Mostly cloudy, with a high near 44. East wind between 6 and 7 mph.

**Angeli:** A chance of showers. Patchy fog before noon. Mostly cloudy, with a high near 44. East wind between 6 and 7 mph. Chance of precipitation is 35%

**Human:** A 40 percent chance of showers before 10am. Mostly cloudy, with a high near 44. East northeast wind around 7 mph.
**ATIS**

**Input:**

<table>
<thead>
<tr>
<th>Flight</th>
<th>Day</th>
<th>Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>from</td>
<td>day</td>
<td>type</td>
</tr>
<tr>
<td>milwaukee</td>
<td>dep/ar/ret</td>
<td>what</td>
</tr>
<tr>
<td>to</td>
<td>saturday departure</td>
<td>query flight</td>
</tr>
</tbody>
</table>

**k-Best:**

*What are the flights from Milwaukee to Phoenix on Saturday*

**Angeli:**

*Show me the flights between Milwaukee and Phoenix on Saturday*

**Human:**

*Milwaukee to Phoenix on Saturday*
Conclusions

- Generation as parsing problem using the hypergraph framework
- Unsupervised end-to-end generation system
- Performance comparable to state-of-the-art
Conclusions

- Generation as parsing problem using the hypergraph framework
- Unsupervised end-to-end generation system
- Performance comparable to state-of-the-art
- How about introducing features on hyperarcs?
Discriminative Reranking Model

Konstas and Lapata, ACL 2012

Concept-to-text Generation via Discriminative Reranking
Joint Discriminative Reranking with Hypergraphs
Joint Discriminative Reranking with Hypergraphs

Training

Flight
from to
edinburgh jeju

Day Number
number dep/ar
7 departure

Month
month dep/ar
july departure

Search
type what
query flight

Give me the flights leaving Edinburgh July seventh coming back to Jeju Island.
Joint Discriminative Reranking with Hypergraphs

Training

Flight
from to
edinburgh jeju

Day Number
number dep/ar
7 departure

Month
month dep/ar
july departure

Search
type what
query flight

Give me the flights leaving Edinburgh July seventh coming back to Jeju Island

S → R(start)
R(r_i.t) → FS(r_j, start) R(r_j.t)
R(r_i.t) → FS(r_j, start)
FS(r, r.f_i) → F(r, r.f_i) FS(r, r.f_i)
FS(r, r.f_i) → F(r, r.f_i)
F(r, r.f) → W(r, r.f) F(r, r.f)
F(r, r.f) → W(r, r.f)
W(r, r.f) → α
W(r, r.f) → g(f.v)
Joint Discriminative Reranking with Hypergraphs

Training

- **Flight**
  - from to
  - edinburgh jeju

- **Day Number**
  - number dep/ar
  - 7 departure

- **Month**
  - month dep/ar
  - july departure

- **Search**
  - type what query flight

\[ f(e) : \Phi = (\Phi_1, \ldots, \Phi_m) \]

Train using Perceptron

- **FS**
  - \[ FS_{0,1}(query_1, start) \]
  - \[ FS_{0,2}(query_1, start) \]
  - \[ FS_{1,2}(query_1, start) \]

- **W**
  - \[ W_{0,5}(search_1, type) \]
  - \[ W_{0,1}(search_1, type) \]
  - \[ W_{1,2}(search_1, type) \]

- **K-best decoding** via reranking

Give me the flights leaving Edinburgh July seventh coming back to Jeju Island.
Joint Discriminative Reranking with Hypergraphs

Discriminative Reranking Model

Training
- Flight: from to
- Day Number: number dep/ar
- Month: month dep/ar
- Search: type what

Month
- july departure

Search
- query flight

S → R(start)
R(r₁,t) → FS(r₂, start)R(r₃,t)
R(r₁,t) → FS(r₄, start)
FS(r₃,r₄) → F(r₅,r₄)FS(r₆,r₄)
FS(r₃,r₄) → F(r₅,r₄)
F(r₅,r₄) → W(r₆,r₄)F(r₇,r₄)
F(r₅,r₄) → W(r₆,r₄)
W(r₆,r₄) → α
W(r₆,r₄) → g(f.v)

Testing
- k-best decoding via reranking

Show me * the flights [type what]
Show me * what flights [type what]
Show me * all flights [type what]

f(e) : Φ = (Φ₁, ..., Φₘ)
Train using Perceptron

FS₀₁(query₁,start)
FS₀₂(query₁,start)
FS₁₂(query₁,start)

FS₀₁(query₁,type)
FS₀₂(query₁,type)
FS₀₂(query₁,what)
FS₁₂(query₁,start)

f(e) : Φ = (Φ₁, ..., Φₘ)
Train using Perceptron

f(e) : Φ = (Φ₁, ..., Φₘ)
Train using Perceptron
Generative vs Discriminative Reranking Model

Generative Model

- **Joint** model allows for more global decisions
- **Hypergraphs** are a compact representation and allow for efficient inference (k-best decoding via cube pruning)
- **Forest rescoring** allows for rescoring k-best trees at all internal nodes via LM integration
Generative vs Discriminative Reranking Model

Generative Model

- **Joint** model allows for more global decisions
- **Hypergraphs** are a compact representation and allow for efficient inference (k-best decoding via cube pruning)
- **Forest rescoring** allows for rescoring k-best trees at all internal nodes via LM integration

Discriminative Reranking Model
Generative vs Discriminative Reranking Model

Generative Model

- **Joint** model allows for more global decisions
- **Hypergraphs** are a compact representation and allow for efficient inference (k-best decoding via cube pruning)
- Forest rescoring allows for rescoring k-best trees at all internal nodes via LM integration

Discriminative Reranking Model

- Use generative decoder as a baseline
- Introduce lexical and structural features at each hyperarc
- **Discriminative reranking** reranks k-best trees at all internal nodes
Hypergraph Reranking

Hidden correspondence $h$ between database $d$ and words $w$
Hidden correspondence $\mathbf{h}$ between database $\mathbf{d}$ and words $\mathbf{w}$

$$(\mathbf{\hat{w}}, \mathbf{\hat{h}}) = \arg \max_{w,h} \alpha \cdot \Phi(\mathbf{d}, \mathbf{w}, \mathbf{h})$$

- $\Phi = (\Phi_1, \ldots, \Phi_m)$: high dimensional feature representation
- $\alpha$: weight vector
- Learn $\alpha$ with averaged structured perceptron (Collins, 2002)
Hypergraph Reranking

Give me the flights leaving Edinburgh July seventh coming back to Jeju Island

Hidden correspondence $h$ between database $d$ and words $w$

$$(\hat{w}, \hat{h}) = \arg \max_{w, h} \alpha \cdot \Phi(d, w, h)$$

- $\Phi = (\Phi_1, \ldots, \Phi_m)$: high dimensional feature representation
- $\alpha$: weight vector
- Learn $\alpha$ with averaged structured perceptron (Collins, 2002)
Hidden correspondence $h$ between database $d$ and words $w$

$$(\hat{w}, \hat{h}) = \arg \max_{w,h} \alpha \cdot \Phi(d, w, h)$$

- $\Phi = (\Phi_1, \ldots, \Phi_m)$: high dimensional feature representation
- $\alpha$: weight vector
- Learn $\alpha$ with averaged structured perceptron (Collins, 2002)
Oracle Derivation

Oracle derivation \((w^*, h^+))

- Use the existing decoder but observe the training text.
- \(w^*\): gold standard text
- \(h^+\): best latent configuration
Baseline Features

- Baseline Model Feature (local): Log score of unsupervised generative decoder (Konstas and Lapata, 2012a)
Baseline Features

- **Baseline Model Feature (local)**: Log score of unsupervised generative decoder (Konstas and Lapata, 2012a)
- **Alignment Features (local)**: Count of each PCFG rule

![Diagram](image)
Lexical Features

- Word Bigrams/Trigrams (non-local)
- Number of Words per Field (local)
- Consecutive Word/Bigram/Trigram (non-local)
Lexical Features

- Word Bigrams/Trigrams (non-local)
- Number of Words per Field (local)
- Consecutive Word/Bigram/Trigram (non-local)

\[
\begin{align*}
& FS_{0,3}(\text{search}_{1}.t, \text{start}) \\
& W_{0,1}(\text{search}_{1}.t, \text{type}) \\
& \hspace{1cm} \ldots \\
& W_{1,3}(\text{search}_{1}.t, \text{what}) \\
& \begin{pmatrix}
show \\
me \\
what \\
\ldots
\end{pmatrix} \\
& \begin{pmatrix}
me \\
the \\
me \\
flights \\
the \\
flights \\
\ldots
\end{pmatrix}
\end{align*}
\]

\text{<show me the>, <show me flights>, etc.}
Lexical Features

- Word Bigrams/Trigrams (non-local)
- **Number of Words per Field** (local)
- Consecutive Word/Bigram/Trigram (non-local)

```
FS_{0,3}(\text{search}_1.t,\text{start})
```

```
W_{0,1}(\text{search}_1.t,\text{type})
```

```
W_{1,3}(\text{search}_1.t,\text{what})
```

```
<show me the>, <show me flights>, etc.
```

```
F_{2,4}(\text{flight}_1.t,\text{from})
```

```
| 2 words |
```

```
<2 | from>
```
Structural Features

- Field Bigrams/Trigrams (non-local)
- Number of Fields per Record (local)
- Fields with no Value (local)
Structural Features

- Field Bigrams/Trigrams (non-local)
- Number of Fields per Record (local)
- Fields with no Value (local)

FS\(_{2,6}(flight_1.t,start)\)

F\(_{2,4}(flight_1.t,from)\)  \(\text{FS}_{4,6}(flight_1.t,from)\)

F\(_{4,6}(flight_1.t,to)\)  \(\epsilon\)

\(<from\ to\) | flight\)
k-best Decoding

- Bottom-up Viterbi search
k-best Decoding

- Bottom-up Viterbi search
- Keep k-best derivations at each node, cube pruning (Chiang, 2007)
k-best Decoding

- Bottom-up Viterbi search
- Keep k-best derivations at each node, cube pruning (Chiang, 2007)
- Score of j-th derivation: $\alpha \cdot \Phi_L(e) + \alpha \cdot \Phi_N(<e,j>)$
k-best Decoding

- Bottom-up Viterbi search
- Keep k-best derivations at each node, cube pruning (Chiang, 2007)
- Score of j-th derivation: $\alpha \cdot \Phi_L(e) + \alpha \cdot \Phi_N(<e,j>)$
- Nodes in hypergraph augmented with lexical and structural sub-strings (Huang and Chiang, 2007)

  e.g. $R_{2,8}(flight_1.t)<\text{one way} \ast \text{to Seoul}; \text{direction from} \ast \text{stop-over to}>$
k-best Decoding

- Bottom-up Viterbi search
- Keep k-best derivations at each node, cube pruning (Chiang, 2007)
- Score of \( j \)-th derivation: \( \alpha \cdot \Phi_L(e) + \alpha \cdot \Phi_N(<e,j>) \)
- Nodes in hypergraph augmented with lexical and structural sub-strings (Huang and Chiang, 2007)
  e.g. \( R_{2,8}(flight_1.t)<\text{one way to Seoul; direction from stop-over to}> \)
- Root node: most grammatical and semantically correct derivation
k-best Decoding

\[
\begin{align*}
&\left(\text{show me } \ast \text{ the flights } [\text{type what}]\right) \\
&\left(\text{show me } \ast \text{ what flights } [\text{type what}]\right) \\
&\left(\text{show me } \ast \text{ all flights } [\text{type what}]\right) \\
&\cdots \\
&\left(\text{show me } [\text{type}]\right) \\
&\left(\text{show the } [\text{type}]\right) \\
&\left(\text{what are } [\text{type}]\right) \\
&\cdots \\
&\text{FS}_{0,5}(\text{search}_1.t,\text{start}) \\
&\text{F}_{0,2}(\text{search}_1.t,\text{type}) \\
&\text{W}_{4,5}(\text{search}_1.t,\text{what}) \\
&\text{W}_{0,1}(\text{search}_1.t,\text{type}) \\
&\text{W}_{1,2}(\text{search}_1.t,\text{type}) \\
&\text{W}_{0,1}(\text{search}_1.t,\text{type}) \\
&\text{W}_{1,2}(\text{search}_1.t,\text{type}) \\
&\cdots
\end{align*}
\]
k-best Decoding

\[
\begin{align*}
&\left( \text{show me } \star \text{ the flights } \right) [\text{type what}] \\
&\left( \text{show me } \star \text{ what flights } \right) [\text{type what}] \\
&\left( \text{show me } \star \text{ all flights } \right) [\text{type what}] \\
&\quad \cdots \\
&\left( \text{show me } \left[ \text{type} \right] \right) \\
&\left( \text{show the } \left[ \text{type} \right] \right) \\
&\left( \text{what are } \left[ \text{type} \right] \right) \\
&\quad \cdots
\end{align*}
\]

\[
\begin{align*}
&\text{FS}_{0,5}(\text{search}_1.t, \text{start}) \\
&\text{F}_{0,2}(\text{search}_1.t, \text{type}) \\
&\text{W}_{0,1}(\text{search}_1.t, \text{type}) \\
&\text{W}_{1,2}(\text{search}_1.t, \text{type}) \\
&\text{W}_{4,5}(\text{search}_1.t, \text{what}) \\
&\left( \text{flights } [\emptyset] \right) \\
&\left( \text{flight } [\emptyset] \right) \\
&\left( \text{airline } [\emptyset] \right) \\
&\quad \cdots
\end{align*}
\]
**k-best Decoding**

\[
\begin{align*}
& (\text{show me} \star \text{the flights} \ [\text{type what}]) \\
& (\text{show me} \star \text{what flights} \ [\text{type what}]) \\
& (\text{show me} \star \text{all flights} \ [\text{type what}]) \\
& \ldots
\end{align*}
\]

\[
\begin{align*}
& (\text{show me} \ [\text{type}]) \\
& (\text{show the} \ [\text{type}]) \\
& (\text{what are} \ [\text{type}]) \\
& \ldots
\end{align*}
\]

\[
\begin{align*}
& \text{FS}_{0,5}(\text{search}_1.t,\text{start}) \\
& \text{F}_{0,2}(\text{search}_1.t,\text{type}) \\
& \text{W}_{4,5}(\text{search}_1.t,\text{what}) \\
& \text{W}_{0,1}(\text{search}_1.t,\text{type}) \\
& \text{W}_{1,2}(\text{search}_1.t,\text{type})
\end{align*}
\]

\[
\begin{align*}
& (\text{flights} \ [\emptyset]) \\
& (\text{flight} \ [\emptyset]) \\
& (\text{airline} \ [\emptyset]) \\
& \ldots
\end{align*}
\]
Experimental Setup

Data

- **ATIS**: mapping from $\lambda$-version (Zettlemoyer and Collins, 2007)
- Model parameters estimated on dev set ($k$-best)
Experimental Setup

Data
- **ATIS**: mapping from λ-version (Zettlemoyer and Collins, 2007)
- Model parameters estimated on dev set (*k*-best)

Evaluation
- Automatic evaluation: BLEU-4, METEOR
- Human evaluation (MTurk): fluency, semantic correctness
Experimental Setup

Data
- ATIS: mapping from $\lambda$-version (Zettlemoyer and Collins, 2007)
- Model parameters estimated on dev set ($k$-best)

Evaluation
- Automatic evaluation: BLEU-4, METEOR
- Human evaluation (MTurk): fluency, semantic correctness

System Comparison
- Baseline: 1-BEST+BASE+ALIGN
- $k$-best (+Lexical): $k$-BEST+BASE+ALIGN+LEX
- $k$-best (+Structural): $k$-BEST+BASE+ALIGN+LEX+STR
- Angeli et al. (2010)
### Results: Automatic Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>BLEU</th>
<th>METEOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-best + base + align</td>
<td>21.93</td>
<td>34.01</td>
</tr>
<tr>
<td>k-best + base + align + lex</td>
<td>28.66</td>
<td>45.18</td>
</tr>
<tr>
<td>k-best + base + align + lex + str</td>
<td>30.62</td>
<td>46.07</td>
</tr>
<tr>
<td>Angeli</td>
<td>26.77</td>
<td>42.41</td>
</tr>
</tbody>
</table>
## Results: Automatic Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>BLEU</th>
<th>METEOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-best + base + align</td>
<td>21.93</td>
<td>34.01</td>
</tr>
<tr>
<td>k-best + base + align + lex</td>
<td>28.66</td>
<td>45.18</td>
</tr>
<tr>
<td>k-best + base + align + lex + str</td>
<td>30.62</td>
<td>46.07</td>
</tr>
<tr>
<td>Angeli</td>
<td>26.77</td>
<td>42.41</td>
</tr>
</tbody>
</table>
Results: Human Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-BEST</td>
<td>2.70</td>
<td>3.05</td>
</tr>
<tr>
<td>k-BEST</td>
<td>4.02</td>
<td>4.04</td>
</tr>
<tr>
<td>ANGELI</td>
<td>3.74</td>
<td>3.17</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.18</td>
<td>4.02</td>
</tr>
</tbody>
</table>

1-BEST significantly better than 1-best and Angeli (α < 0.01).

k-BEST and Human are not significantly different.
Results: Human Evaluation

<table>
<thead>
<tr>
<th>System</th>
<th>Fluency</th>
<th>SemCor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-BEST</td>
<td>2.70</td>
<td>3.05</td>
</tr>
<tr>
<td>$k$-BEST</td>
<td>4.02</td>
<td>4.04</td>
</tr>
<tr>
<td>ANGELI</td>
<td>3.74</td>
<td>3.17</td>
</tr>
<tr>
<td>HUMAN</td>
<td>4.18</td>
<td>4.02</td>
</tr>
</tbody>
</table>

- $k$-BEST significantly better than 1-BEST and ANGELI ($\alpha < 0.01$)
- $k$-BEST and HUMAN are not significantly different
Output

<table>
<thead>
<tr>
<th>Flight</th>
<th>Time</th>
<th>Day</th>
<th>Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>from</td>
<td>to</td>
<td>when dep/ar</td>
<td>day dep/ar</td>
</tr>
<tr>
<td>phoenix</td>
<td>milwaukee</td>
<td>evening departure</td>
<td>wednesday departure</td>
</tr>
</tbody>
</table>

1-BEST: On Wednesday evening from Phoenix to Milwaukee on Wednesday evening

k-BEST: Please list the flights from Phoenix to Milwaukee on Wednesday evening

ANGELI: Show me the flights from Phoenix to Milwaukee on Wednesday evening flights from Phoenix to Milwaukee

HUMAN: Give me the flights from Phoenix to Milwaukee on Wednesday evening
Conclusions

- Discriminative reranking using the structured perceptron
- Introduced local and non-local features
- Performance better than state-of-the-art
- Future work: dependency relations, discourse
Demo

Live Weather Forecast Generator

Cross domain

- Model trained on: weather.gov
- Demo runs on: wunderground.com
- Discrepancies (e.g. no gust information, inferred fields)
Demo

Live Weather Forecast Generator

Cross domain
- Model trained on: weather.gov
- Demo runs on: wunderground.com
- Discrepancies (e.g. no gust information, inferred fields)

Live Data

Night
6:00  17:00  21:00  6:00(+1)

Day

Konstas, Lapata (ILCC)  Concept-to-Text Generation  5 September 2012
Demo

Live Weather Forecast Generator

Cross domain
- Model trained on: weather.gov
- Demo runs on: wunderground.com
- Discrepancies (e.g. no gust information, inferred fields)

Live Data

6:00  13:00  17:00  21:00  6:00(+1)

Day

Night
Demo

Live Weather Forecast Generator

Cross domain
- Model trained on: weather.gov
- Demo runs on: wunderground.com
- Discrepancies (e.g. no gust information, inferred fields)

Live Data

Day

6:00 13:00 17:00 21:00

Night

6:00(+1)
Demo

Live Weather Forecast Generator

Cross domain

- Model trained on: weather.gov
- Demo runs on: wunderground.com
- Discrepancies (e.g. no gust information, inferred fields)

Live Data

6:00 13:00 17:00 21:00 6:00(+1)

Day

Night
Demos

Flight Booking System - lowJet

Scenario 1
Imagine you are in Los Angeles for the weekend and you would like to fly directly back home to Denver the following Monday. Try to find the earliest flight.

Scenario 2
You have won a free one-way vacation from Miami to Washington DC. Book the cheapest tickets for next Thursday morning. Alternatively, just find the earliest itinerary.
Thank you
### Table 2

<table>
<thead>
<tr>
<th>Category</th>
<th>Fields - Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature (F)</td>
<td>time: 06.00 - 21.00, min: 31, mean: 38, max: 45</td>
</tr>
<tr>
<td>Wind Speed (mph)</td>
<td>time: 06.00 - 21.00, min: 6, mean: 7, max: 8</td>
</tr>
<tr>
<td>Wind Direction</td>
<td>time: 06.00 - 21.00, mode: SE</td>
</tr>
<tr>
<td>Cloud Sky Cover (%)</td>
<td>time: 06.00 - 21.00, percent: 75-100</td>
</tr>
</tbody>
</table>

### Translation 2

MOSTLY CLOUDY, WITH A HIGH NEAR 45. SOUTH SOUTHEAST WIND BETWEEN 6 AND 8 MPH.

<table>
<thead>
<tr>
<th>Fluency</th>
<th>Semantic Correctness</th>
</tr>
</thead>
<tbody>
<tr>
<td>○ 1</td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5</td>
</tr>
<tr>
<td>○ 1 ○ 2</td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5</td>
</tr>
</tbody>
</table>
Table 3

<table>
<thead>
<tr>
<th>Category</th>
<th>Fields - Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flight Info</td>
<td>from: milwaukee</td>
</tr>
<tr>
<td></td>
<td>to: phoenix</td>
</tr>
<tr>
<td>Query</td>
<td>type: show</td>
</tr>
<tr>
<td></td>
<td>what: flight</td>
</tr>
<tr>
<td>Day</td>
<td>day: saturday</td>
</tr>
<tr>
<td></td>
<td>dep/ar/ret: departure</td>
</tr>
</tbody>
</table>

Translation 3

SHOW ME THE FLIGHTS BETWEEN MILWAUKEE AND PHOENIX ON SATURDAY

Fluency

Semantic Correctness
Train a linear regression model
Idea: The more records and fields that have values in the database → the more facts need to be uttered
Input to the model: Flattened version of the database input, i.e. each feature is a record-field pair
Feature values: Values vs Counts of Fields
RoboCup

Input:

<table>
<thead>
<tr>
<th>Pass</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>purple10</td>
<td>purple11</td>
</tr>
</tbody>
</table>

$k$-Best: purple10 passes back to purple11

ANGELI: purple10 passes to purple11

HUMAN: purple10 immediately passes to purple11