A Simple Domain-Independent Approach to Generation

Gabor Angeli
gangeli@berkeley.edu

Percy Liang
pliang@eecs.berkeley.edu

Dan Klein
dan@eecs.berkeley.edu
Introduction

Task: Generate text from database records

s:

- temperature (time=5pm-6am, min=48, mean=53, max=61)
- windSpeed (time=5pm-6am, min=3, mean=6, max=11, mode=0-10)
- windDir (time=5pm-6am, mode=SSW)
- gust (time=5pm-6am, min=0, mean=0, max=0)
- skyCover (time=5pm-9pm, mode=0-25)
- skyCover (time=2am-6am, mode=75-100)
- precipPotential (time=5pm-6am, min=2, mean=14, max=20)
- rainChance (time=5pm-6am, mode=someChance)

w: A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.
Introduction

Task: Generate text from database records

- temperature (time=5pm-6am, min=48, mean=53, max=61)
- windSpeed (time=5pm-6am, min=3, mean=6, max=11, mode=0-10)
  - windDir (time=5pm-6am, mode=SSW)
- gust (time=5pm-6am, min=0, mean=0, max=0)
- skyCover (time=5pm-9pm, mode=0-25)
- skyCover (time=2am-6am, mode=75-100)
- precipPotential (time=5pm-6am, min=2, mean=14, max=20)
- rainChance (time=5pm-6am, mode=someChance)

Requires both content selection
Introduction

Task: Generate text from database records

s:

- temperature(time=5pm-6am, min=48, mean=53, max=61)
- windSpeed(time=5pm-6am, min=3, mean=6, max=11, mode=0-10)
- windDir(time=5pm-6am, mode=SSW)
- gust(time=5pm-6am, min=0, mean=0, max=0)
- skyCover(time=5pm-9pm, mode=0-25)
- skyCover(time=2am-6am, mode=75-100)
- precipPotential(time=5pm-6am, min=2, mean=14, max=20)
- rainChance(time=5pm-6am, mode=someChance)

w: A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.

Requires both content selection and surface realization
Main Ideas

Complete Pipeline:

Do content selection and surface realization in a unified framework
Main Ideas

Complete Pipeline:

Do content selection and surface realization in a unified framework

Content Selection: e.g. Barzilay and Lee (2004)

Surface Realization: e.g. Soricut and Marcu (2006), White, et al. (2007)
Main Ideas

Complete Pipeline:
Do content selection and surface realization in a unified framework
  Content Selection: e.g. Barzilay and Lee (2004)
  Surface Realization: e.g. Soricut and Marcu (2006), White, et al. (2007)

Domain Independent Approach:
Applicable to multiple domains with minimal tweaking
Main Ideas

Complete Pipeline:
Do content selection and surface realization in a unified framework
  Content Selection: e.g. Barzilay and Lee (2004)
  Surface Realization: e.g. Soricut and Marcu (2006), White, et al. (2007)

Domain Independent Approach:
Applicable to multiple domains with minimal tweaking
Contrast with domain-tuned systems,
e.g. Chen and Mooney (2008), Belz (2008)
A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. Southwest wind between 5 and 10 mph.
A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.

w: Purple3 made a bad pass that was picked off by pink9.
Domain Independent

\[
\begin{align*}
\text{wind10m} & (\text{time}=6am, \text{dir}=SW, \text{min}=16, \text{max}=20, \text{gust min}=0, \text{gust max}=0) \\
\text{wind10m} & (\text{time}=9pm, \text{dir}=SSW, \text{min}=28, \text{max}=32, \text{gust min}=40, \text{gust max}=0) \\
\text{wind10m} & (\text{time}=12am, \text{dir}=-, \text{min}=24, \text{max}=28, \text{gust min}=36, \text{gust max}=0)
\end{align*}
\]

\[w: \text{sw 16 - 20 backing ssw 28 - 32 gusts 40 by mid evening}
\text{ easing 24 - 28 gusts 36 late evening}\]
Main Ideas

Complete Pipeline:

Do content selection and surface realization in a unified framework
Content Selection: e.g. Barzilay and Lee (2004)
Surface Realization: e.g. Soricut and Marcu (2006), White, et al. (2007)

Domain Independent Approach:

Applicable to multiple domains with minimal tweaking
Contrast with domain-tuned systems,
  e.g. Chen and Mooney (2008), Belz (2008)
Main Ideas

Complete Pipeline:
Do content selection and surface realization in a unified framework
Content Selection: e.g. Barzilay and Lee (2004)
Surface Realization: e.g. Soricut and Marcu (2006), White, et al. (2007)

Domain Independent Approach:
Applicable to multiple domains with minimal tweaking
Contrast with domain-tuned systems,
e.g. Chen and Mooney (2008), Belz (2008)

Data Driven:
Learned from annotated training data
Learning Based
Learning Based

Train

- A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. Southwest wind between 5 and 10 mph.

...
Learning Based

Train

Test

s:  
- temperature(time=5pm-6am, min=48, mean=53, max=61)
- windSpeed(time=5pm-6am, min=3, mean=6, max=11, mode=0-10)
- windDir(time=5pm-6am, mode=SSW)
- gust(time=5pm-6am, min=0, mean=0, max=0)
- skyCover(time=5pm-9pm, mode=0-25)
- skyCover(time=2am-6am, mode=75-100)
- precipPotential(time=5pm-6am, min=2, mean=14, max=20)
- rainChance(time=5pm-6am, mode=someChance)

\[\text{w}: A 20 \text{ percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.} \]
A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.
Representations:

- **Temperature**: (time=5pm-6am, min=48, mean=53, max=61)
- **Wind Speed**: (time=5pm-6am, min=3, mean=6, max=11, mode=0-10)
- **Wind Direction**: (time=5pm-6am, mode=SSW)
- **Gust**: (time=5pm-6am, min=0, mean=0, max=0)
- **Sky Cover**: (time=5pm-9pm, mode=0-25)
- **Sky Cover**: (time=2am-6am, mode=75-100)
- **Precipitation Potential**: (time=5pm-6am, min=2, mean=14, max=20)
- **Rain Chance**: (time=5pm-6am, mode=some Chance)

Weather:

A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. Southwest wind between 5 and 10 mph.
A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. Southwest wind between 5 and 10 mph.

Text: Generated from a sequence of templates (stay tuned)
representation

\[ s: \]
- \(\text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=48, \text{mean}=53, \text{max}=61)\)
- \(\text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10)\)
- \(\text{windDir}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=\text{SSW})\)
- \(\text{gust}(\text{time}=5\text{pm}-6\text{am}, \text{min}=0, \text{mean}=0, \text{max}=0)\)
- \(\text{skyCover}(\text{time}=5\text{pm}-9\text{pm}, \text{mode}=0-25)\)
- \(\text{skyCover}(\text{time}=2\text{am}-6\text{am}, \text{mode}=75-100)\)
- \(\text{precipPotential}(\text{time}=5\text{pm}-6\text{am}, \text{min}=2, \text{mean}=14, \text{max}=20)\)
- \(\text{rainChance}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=\text{someChance})\)

\[ w: \] A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.

\[ \text{Text:} \] Generated from a sequence of templates (stay tuned)

\[ \text{World state:} \] Set of database records
Representation

\[ s: \]
\[
\text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=48, \text{mean}=53, \text{max}=61) \\
\text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \\
\text{windDir}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=\text{SSW}) \\
\text{gust}(\text{time}=5\text{pm}-6\text{am}, \text{min}=0, \text{mean}=0, \text{max}=0) \\
\text{skyCover}(\text{time}=5\text{pm}-9\text{pm}, \text{mode}=0-25) \\
\text{skyCover}(\text{time}=2\text{am}-6\text{am}, \text{mode}=75-100) \\
\text{precipPotential}(\text{time}=5\text{pm}-6\text{am}, \text{min}=2, \text{mean}=14, \text{max}=20) \\
\text{rainChance}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=\text{someChance}) \\
\ldots
\]

\[ w: \text{ A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.} \]

**Text:** Generated from a sequence of templates (stay tuned)

**World state:** Set of database records

\[ r_1 = \text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \]
Representation

\[ r_1 = \text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \]
Representation

\[ r_1 = \text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \]

Each record is of a given record type
Representation

\( r_1 = \text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \)

Each record is of a given record type
Each record has a set of fields
**Representation**

\[ r_1 = \text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \]

Each **record** is of a given **record type**
Each **record** has a set of **fields**
Each **field** has an associated **value**
Representation

\[ r_1 = \text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \]

Each **record** is of a given **record type**
Each **record** has a set of **fields**
Each **field** has an associated **value**

- **Numeric**: \( \text{min}=3, \text{mean}=6, \text{max}=11 \)
- **Symbolic**: \( \text{time}=5\text{pm}-6\text{am}, \text{mode}=0-10 \)
Representation

\[ r_1 = \text{windSpeed}(time=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \]

Each record is of a given record type
Each record has a set of fields
Each field has an associated value
   - Numeric: min=3, mean=6, max=11
   - Symbolic: time=5pm-6am, mode=0-10

Note: No a priori correspondence between symbolic values and words
Representation

\[ r_1 = \text{windSpeed}(\text{time}=5\text{pm}-6\text{am}, \text{min}=3, \text{mean}=6, \text{max}=11, \text{mode}=0-10) \]

Each record is of a given record type

Each record has a set of fields

Each field has an associated value

- Numeric: min=3, mean=6, max=11
- Symbolic: time=5pm-6am, mode=0-10

Note: No a priori correspondence between symbolic values and words

Language independent
Representation

Text is generated from a sequence of templates
Representation

Text is generated from a sequence of templates

temperature(time=5pm-6am,min=48,mean=53,max=61)
Representation

Text is generated from a sequence of templates

temperature(time=5pm-6am,min=48,mean=53,max=61)

⟨with a low around [min] .⟩
Representation

Text is generated from a sequence of templates

```
\begin{itemize}
    \item temperature(time=5pm-6am,min=48,mean=53,max=61)
\end{itemize}
```

\begin{equation}
\langle \text{with a low around } \text{[min]} \rangle
\end{equation}

Templates extracted automatically from the training data (stay tuned)
Generation as Decisions

s:
  skyCover(time=5pm-6am, mode=50-75)
  temperature(time=5pm-6am, min=44, mean=49, max=60)
  ...
Generation as Decisions

\[ s: \]
\[ \text{skyCover}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75) \]
\[ \text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60) \]
\[ ... \]

Record

Field set

Template

\[ s \]
\[ \text{skyCover}_1 \ 0.43 \]
\[ \text{rainChance}_1 \ 0.2 \]
\[ \text{skyCover}_3 \ 0.13 \]
\[ ... \]
Generation as Decisions

$s$: 
- $\text{skyCover}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75)$
- $\text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60)$
- ...

Record

Field set

Template

$\text{skyCover}_1 = 0.43$
$\text{rainChance}_1 = 0.2$
$\text{skyCover}_3 = 0.13$
...

...
Generation as Decisions

\[ s: \]

\[ \text{skyCover} (\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75) \]
\[ \text{temperature} (\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60) \]
\[ \ldots \]

Record

\[ r_1 = \text{skyCover}_1 \]

Field set

Template

\[ \{ \text{time, mode} \} \quad 0.13 \]

\[ \{ \text{mode} \} \quad 0.84 \]

\[ \{ \} \quad 0.02 \]

\[ \ldots \]
Generation as Decisions

s:

- skyCover(time=5pm-6am, mode=50-75)
- temperature(time=5pm-6am, min=44, mean=49, max=60)

...
Generation as Decisions

s:

\[ \text{skyCover}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75) \]
\[ \text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60) \]

...
Generation as Decisions

\[
\text{s:} \quad \begin{align*}
\text{skyCover} & (\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75) \\
\text{temperature} & (\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60) \\
\text{...}
\end{align*}
\]

Record
\[
r_1 = \text{skyCover}_1
\]

Field set
\[
F_1 = \{\text{mode}\}
\]

Template
\[
\langle \text{mostly cloudy} , \rangle \quad 0.79
\]
\[
\langle \text{partly cloudy} , \rangle \quad 0.1
\]
\[
\langle \text{mostly clear} , \rangle \quad 0.04
\]
\[
\ldots
\]
Generation as Decisions

s:

```
skyCover(time=5pm-6am,mode=50-75)
temperature(time=5pm-6am,min=44,mean=49,max=60)
...```

Record

```
r_1 = skyCover_1
```

Field set

```
F_1 = \{ mode \}
```

Template

```
W_1 = \langle mostly cloudy , \rangle
```

s;  

```
r_1 = skyCover_1, F_1 = \{ mode \}, T_1 = \langle mostly cloudy , \rangle
```

```
\begin{array}{ll}
\text{rainChance}_1 & 0.02 \\
\text{thunderChance}_3 & 0.01 \\
\text{temperature}_1 & 0.95 \\
\end{array}
```

...
Generation as Decisions

\[ s: \]

\[ \text{skyCover}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75) \]
\[ \text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60) \]
\[ \ldots \]

Record

\[ r_1 = \text{skyCover}_1 \]

Field set

\[ F_1 = \{ \text{mode} \} \]

Template

\[ W_1 = \langle \text{mostly cloudy} , \rangle \]

\[ s; r_1 = \text{skyCover}_1, F_1 = \{ \text{mode} \}, T_1 = \langle \text{mostly cloudy} , \rangle \]

\[ \begin{align*}
\text{rainChance}_1 & \quad 0.02 \\
\text{thunderChance}_3 & \quad 0.01 \\
\text{temperature}_1 & \quad 0.95 \\
\ldots 
\end{align*} \]
Generation as Decisions

s:

skyCover(time=5pm-6am,mode=50-75)
temperature(time=5pm-6am,min=44,mean=49,max=60)
...

Record

$ r_1 = \text{skyCover}_1 $  

$ r_2 = \text{temperature}_1 $  

Field set

$ F_1 = \{\text{mode}\} $  

Template

$ W_1 = \langle mostly\ cloudy, \rangle $  

s; ..., $ F_1 = \{\text{mode}\}, T_1 = \langle mostly\ cloudy, \rangle, r_2 = \text{temperature} $  

{mean,min,max} 0.12  

{time,min} 0.6  

{min} 0.13  

...
Generation as Decisions

s:
- skyCover\((time=5pm-6am, mode=50-75)\)
- temperature\((time=5pm-6am, min=44, mean=49, max=60)\)
...

Record
- \(r_1 = \text{skyCover}_1\)
- \(r_2 = \text{temperature}_1\)

Field set
- \(F_1 = \{\text{mode}\}\)

Template
- \(W_1 = \langle \text{mostly cloudy} ,\rangle\)

s; ...,\(F_1 = \{\text{mode}\}\), \(T_1 = \langle \text{mostly cloudy} ,\rangle\), \(r_2 = \text{temperature}\)

- \(\{\text{mean, min, max}\} \) 0.12
- \(\{\text{time, min}\} \) 0.6
- \(\{\text{min}\} \) 0.13
...
### Generation as Decisions

$$s: \quad \text{skyCover}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75)$$

$$\quad \text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60)$$

...
Generation as Decisions

\[
\begin{align*}
\text{Record} & \\
\quad r_1 &= \text{skyCover}_1 & \quad r_2 &= \text{temperature}_1 \\
\text{Field set} & \\
\quad F_1 &= \{\text{mode}\} & \quad F_2 &= \{\text{time, min}\} \\
\text{Template} & \\
\quad W_1 &= \langle \text{mostly cloudy} , \rangle &
\end{align*}
\]

\[
\begin{align*}
\text{s}; \ldots, T_1 &= \langle \text{mostly cloudy} , \rangle, r_2 = \text{temperature}, F_2 = \{\text{time, min}\} \\
\langle \text{with a low around} [\text{min} .] \rangle &= 0.83 & \langle \text{with a high around} [\text{max} .] \rangle &= 0.01 & \langle \text{with a high near} [\text{max} .] \rangle &= 0.05 \\
\ldots
\end{align*}
\]
Generation as Decisions

s: skyCover(time=5pm-6am, mode=50-75)

\[ \text{temperature(time=5pm-6am, min=44, mean=49, max=60)} \]

...
Generation as Decisions

s:
- skyCover\(\text{time}=5\text{pm}-6\text{am},\text{mode}=50-75\)
- temperature\(\text{time}=5\text{pm}-6\text{am},\text{min}=44,\text{mean}=49,\text{max}=60\)

Record
- \(r_1 = \text{skyCover}_1\)
- \(r_2 = \text{temperature}_1\)

Field set
- \(F_1 = \{\text{mode}\}\)
- \(F_2 = \{\text{time}, \text{min}\}\)

Template
- \(W_1 = \langle \text{mostly cloudy } \rangle\)
- \(W_2 = \langle \text{with a low around } [\text{min}] \rangle\)

s; \(\dots, r_2 = \text{temperature}, F_2 = \{\text{time}, \text{min}\}, T_2 = \langle \text{with a low around } \text{min} \rangle\)

\begin{align*}
\text{windDir}_1 & \quad 0.37 \\
\text{temperature}_1 & \quad 0.03 \\
\text{STOP} & \quad 0.43
\end{align*}
Generation as Decisions

s:

\[
\text{skyCover}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75)
\]
\[
\text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60)
\]
...

Record

\begin{align*}
\mathbf{r_1} &= \text{skyCover}_1 \\
\mathbf{r_2} &= \text{temperature}_1 \\
\mathbf{r_3} &= \text{STOP}
\end{align*}

Field set

\begin{align*}
F_1 &= \{\text{mode}\} \\
F_2 &= \{\text{time}, \text{min}\}
\end{align*}

Template

\begin{align*}
W_1 &= \langle\text{mostly cloudy ,}\rangle \\
W_2 &= \langle\text{with a low around } [\text{min}] .\rangle
\end{align*}

mostly cloudy , with a low around 44 .
Generation as Decisions

\[ s: \]
\[ \text{\( \checkmark \)} \text{skyCover}(\text{time}=5\text{pm}-6\text{am}, \text{mode}=50-75) \]
\[ \text{\( \odot \)} \text{temperature}(\text{time}=5\text{pm}-6\text{am}, \text{min}=44, \text{mean}=49, \text{max}=60) \]
\[ \ldots \]

<table>
<thead>
<tr>
<th>Record</th>
<th>Field set</th>
<th>Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_1 = \text{( \checkmark )} \text{skyCover}_1 )</td>
<td>( F_1 = {\text{mode}} )</td>
<td>( W_1 = \langle \text{mostly cloudy} , \rangle )</td>
</tr>
<tr>
<td>( r_2 = \text{( \odot )} \text{temperature}_1 )</td>
<td>( F_2 = {\text{time, min}} )</td>
<td>( W_2 = \langle \text{with a low around } [\text{min}] \rangle )</td>
</tr>
<tr>
<td>( r_3 = \text{STOP} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

mostly cloudy, with a low around 44.
Features: Record

| Record | $r_1 = \text{skyCover}_1$ | $r_2 = \text{temperature}_1$ |
| Field set | $F_1 = \{\text{mode}\}$ |
| Template | $W_1 = \langle \text{mostly cloudy}, \rangle$ |
Features: Record

\[(R1)\] List of last two record types
\[
[r_2.t = \text{temperature and } (r_1.t, r_0.t) = (\text{skyCover}, \text{START})]
\]

\[
r_1 = \text{skyCover}_1, F_1 = \{\text{mode}\}, T_1 = \langle \text{mostly cloudy} \rangle
\]

\[
\text{temperature}_1
\]

\[
\text{thunderChance}_3
\]

\[
\text{rainChance}_1
\]

...
Features: Record

(R1) List of last two record types
\[ r_2.t = \text{temperature} \text{ and } (r_1.t, r_0.t) = (\text{skyCover}, \text{START}) \]\n
(R2) Set of previous record types
\[ r_2.t = \text{temperature} \text{ and } \{r_1.t\} = \{\text{skyCover}\} \]
Features: Record

(R1) List of last two record types
\[ r_{2,t} = \text{temperature and } (r_{1,t}, r_{0,t}) = (\text{skyCover}, \text{START}) \]

(R2) Set of previous record types
\[ r_{2,t} = \text{temperature and } \{r_{1,t}\} = \{\text{skyCover}\} \]

(R3) Record type already generated
\[ r_{2,t} = \text{temperature and } r_{j,t} \neq \text{temperature} \ \forall j < 2 \]

\[ r_1 = \text{skyCover}_1, F_1 = \{\text{mode}\}, T_1 = \langle \text{mostly cloudy} , \rangle \]

\[ \begin{align*}
\text{temperature}_1 \\
\text{thunderChance}_3 \\
\text{rainChance}_1
\end{align*} \]

...
(R1) **List of last two record types**

\[ r_2.t = \text{temperature and } (r_1.t, r_0.t) = (\text{skyCover}, \text{START}) \]

(R2) **Set of previous record types**

\[ r_2.t = \text{temperature and } \{ r_1.t \} = \{ \text{skyCover} \} \]

(R3) **Record type already generated**

\[ r_2.t = \text{temperature and } r_j.t \neq \text{temperature } \forall j < 2 \]

(R4) **Field values**

\[ r_2.t = \text{temperature and } r_2.v[\text{time}] = 5\text{pm}-6\text{am} \]

\[ r_1 = \text{skyCover}_1, F_1 = \{ \text{mode} \}, T_1 = \langle \text{mostly cloudy} \rangle \]

\[ \text{temperature}_1 \]

\[ \text{thunderChance}_3 \]

\[ \text{rainChance}_1 \]

\[ \ldots \]
Features: Record

(R1) List of last two record types
\[ r_{2.t} = \text{temperature and } (r_{1.t}, r_{0.t}) = (\text{skyCover}, \text{START}) \]

(R2) Set of previous record types
\[ r_{2.t} = \text{temperature and } \{r_{1.t}\} = \{\text{skyCover}\} \]

(R3) Record type already generated
\[ r_{2.t} = \text{temperature and } r_{j.t} \neq \text{temperature } \forall j < 2 \]

(R4) Field values
\[ r_{2.t} = \text{temperature and } r_{2.v}[\text{time}] = 5\text{pm} - 6\text{am} \]

(R5) Stop under language model

\[ r_1 = \text{skyCover}, F_1 = \{\text{mode}\}, T_1 = \langle \text{mostly cloudy} \rangle \]

\[ \text{temperature}_1 = 0.95 \]

\[ \text{thunderChance}_3 \]

\[ \text{rainChance}_1 \]

...
### Features: Record

**R1)** List of last two record types

\[ r_{2.t} = \text{thunderChance} \text{ and } (r_{1.t}, r_{0.t}) = (\text{skyCover}, \text{START}) \]

**R2)** Set of previous record types

\[ r_{2.t} = \text{thunderChance} \text{ and } \{r_{1.t}\} = \{\text{skyCover}\} \]

**R3)** Record type already generated

\[ r_{2.t} = \text{thunderChance} \text{ and } r_{j.t} \neq \text{thunderChance} \forall j < 2 \]

**R4)** Field values

\[ r_{2.t} = \text{thunderChance} \text{ and } r_{2.v}[\text{time}] = 2\text{am}-6\text{am} \]

**R5)** Stop under language model

\[ r_1 = \text{skyCover}_1, F_1 = \{\text{mode}\}, T_1 = \langle \text{mostly cloudy} , \rangle \]

\[ \text{temperature}_1 = 0.95 \]

\[ \text{thunderChance}_3 \]

\[ \text{rainChance}_1 \]

...
Features: Record

(R1) List of last two record types
\[ r_{2.t} = \text{thunderChance and } (r_{1.t}, r_{0.t}) = (\text{skyCover}, \text{START}) \]

(R2) Set of previous record types
\[ r_{2.t} = \text{thunderChance and } \{r_{1.t}\} = \{\text{skyCover}\} \]

(R3) Record type already generated
\[ r_{2.t} = \text{thunderChance and } r_j.t \neq \text{thunderChance} \quad \forall j < 2 \]

(R4) Field values
\[ r_{2.t} = \text{thunderChance and } r_{2.v}[\text{time}] = 2\text{am–6am} \]

(R5) Stop under language model

\[ r_1 = \text{skyCover}_1, F_1 = \{\text{mode}\}, T_1 = \langle \text{mostly cloudy} , \rangle \]

\[ \text{temperature}_1 = 0.95 \]

\[ \text{thunderChance}_3 = 0.01 \]

\[ \text{rainChance}_1 \]

...
Features: Record

(R1) List of last two record types
\[ r_2.t = \text{rainChance} \text{ and } (r_1.t, r_0.t) = (\text{skyCover}, \text{START}) \]

(R2) Set of previous record types
\[ r_2.t = \text{rainChance} \text{ and } \{r_1.t\} = \{\text{skyCover}\} \]

(R3) Record type already generated
\[ r_2.t = \text{rainChance} \text{ and } r_j.t \neq \text{rainChance} \forall j < 2 \]

(R4) Field values
\[ r_2.t = \text{rainChance} \text{ and } r_2.v[\text{time}] = 5\text{pm}-6\text{am} \]

(R5) Stop under language model

\[ r_1 = \text{skyCover}_1, F_1 = \{\text{mode}\}, T_1 = \langle \text{mostly cloudy} , \rangle \]

\begin{itemize}
  \item temperature_1 \quad 0.95
  \item thunderChance_3 \quad 0.01
  \item rainChance_1
  \item ...
\end{itemize}
Features: Record

(R1) **List of last two record types**
\[ r_{2.t} = \text{rainChance} \text{ and } (r_{1.t}, r_{0.t}) = (\text{skyCover}, \text{START}) \]

(R2) **Set of previous record types**
\[ r_{2.t} = \text{rainChance} \text{ and } \{r_{1.t}\} = \{\text{skyCover}\} \]

(R3) **Record type already generated**
\[ r_{2.t} = \text{rainChance} \text{ and } r_{j.t} \neq \text{rainChance} \forall j < 2 \]

(R4) **Field values**
\[ r_{2.t} = \text{rainChance} \text{ and } r_{2.v}[\text{time}] = 5\text{pm}-6\text{am} \]

(R5) **Stop under language model**
Features: Record

(R1) List of last two record types
\[ r_2.t = \text{temperature and } (r_1.t, r_0.t) = (\text{skyCover, START}) \]

(R2) Set of previous record types
\[ r_2.t = \text{temperature and } \{r_1.t\} = \{\text{skyCover}\} \]

(R3) Record type already generated
\[ r_2.t = \text{temperature and } r_j.t \neq \text{temperature } \forall j < 2 \]

(R4) Field values
\[ r_2.t = \text{temperature and } r_2.v[\text{time}] = 5\text{pm}-6\text{am} \]

(R5) Stop under language model

\[ r_1=\text{skyCover}_1, F_1=\{\text{mode}\}, T_1=\langle \text{mostly cloudy ,} \rangle \]

\[ \text{temperature}_1 \]

\[ \text{thunderChance}_3 \]

\[ \text{rainChance}_1 \]

\[ ... \]
**Features: Field Set**

<table>
<thead>
<tr>
<th>Record</th>
<th>$r_1 = \text{skyCover}_1$</th>
<th>$r_2 = \text{temperature}_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field set</td>
<td>$F_1 = {\text{mode}}$</td>
<td>$F_2 = {\text{time}, \text{min}}$</td>
</tr>
<tr>
<td>Template</td>
<td>$W_1 = \langle \text{mostly cloudy} , \rangle$</td>
<td></td>
</tr>
</tbody>
</table>
Features: Field Set

(R1) **Field Set**

\[ F_2 = \{\text{time, min}\} \]

..., \( F_1 = \{\text{mode}\} \), \( T_1 = \langle \text{mostly cloudy} \rangle \), \( r_2 = \text{temperature} \)

\{mean, min, max\} 0.12

\{time, min\}

\{min\} 0.13

...
Features: Field Set

(R1) Field Set
\[ F_2 = \{\text{time, min}\} \]

(R2) Field Set Values
\[ F_2 = \{\text{time, min}\} \text{ and } r_2.v[\text{time}] = 5\text{pm}-6\text{am} \]

..., \( F_1 = \{\text{mode}\}, T_1 = \langle \text{mostly cloudy} \rangle, r_2 = \text{temperature} \)

\{\text{mean, min, max}\} \quad 0.12

\{\text{time, min}\} \quad 0.6

\{\text{min}\} \quad 0.13

...
Features: Field Set

(R1) Field Set
\[ F_2 = \{ \text{time, min} \} \]

(R2) Field Set Values
\[ F_2 = \{ \text{time, min} \} \text{ and } r_2.v[\text{time}] = 5\text{pm}-6\text{am} \]

..., \text{F}_1=\{\text{mode}\}, \text{T}_1=\langle \text{mostly cloudy} , \rangle, r_2=\text{temperature}

\{\text{mean, min, max}\} 0.12
\{\text{time, min}\} 0.6
\{\text{min}\} 0.13
...

14
## Features: Template

<table>
<thead>
<tr>
<th>Record</th>
<th>$r_1 = \text{skyCover}_1$</th>
<th>$r_2 = \text{temperature}_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field set</td>
<td>$F_1 = {\text{mode}}$</td>
<td>$F_2 = {\text{time, min}}$</td>
</tr>
<tr>
<td>Template</td>
<td>$W_1 = \langle\text{mostly cloudy} ,\rangle$</td>
<td>$W_2 = \langle\text{with a low around} \ [\text{min}] \ .\rangle$</td>
</tr>
</tbody>
</table>
### Features: Template

<table>
<thead>
<tr>
<th>Record</th>
<th>$r_1 = \text{skyCover}_1$</th>
<th>$r_2 = \text{temperature}_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field set</td>
<td>$F_1 = {\text{mode}}$</td>
<td>$F_2 = {\text{time, min}}$</td>
</tr>
<tr>
<td>Template</td>
<td>$W_1 = \langle \text{mostly cloudy} , \rangle$</td>
<td>$W_2 = \langle \text{with a low around [min]} . \rangle$</td>
</tr>
</tbody>
</table>

$...T_1 = \langle \text{mostly cloudy} , \rangle, r_2 = \text{temperature}, F_2 = \{\text{time, min}\}$

- $\langle \text{with a low around [min]} . \rangle$ : 0.83
- $\langle \text{with a high around [max]} . \rangle$ : 0.01
- $\langle \text{with a high near [max]} . \rangle$ : 0.05

...
Template Extraction

Templates extracted from induced alignments of Liang, et. al (2009)
**Template Extraction**

Templates extracted from induced alignments of Liang, et. al (2009)

<table>
<thead>
<tr>
<th>Records:</th>
<th>skyCover&lt;sub&gt;1&lt;/sub&gt;</th>
<th>temperature&lt;sub&gt;1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields:</td>
<td>mostly cloudy, mode=50–75</td>
<td>with a low around min=44, 45</td>
</tr>
<tr>
<td>Text:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Template Extraction

Templates extracted from induced alignments of Liang, et. al (2009)

<table>
<thead>
<tr>
<th>Records:</th>
<th>skyCover&lt;sub&gt;1&lt;/sub&gt;</th>
<th>temperature&lt;sub&gt;1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields:</td>
<td>mostly &lt;&lt; mode=50-75 cloudy &gt; &gt;</td>
<td>with a low around &lt;&lt; min=44 45 &gt; &gt;</td>
</tr>
<tr>
<td>Text:</td>
<td>mostly cloudy, with a low around</td>
<td></td>
</tr>
</tbody>
</table>

Ideal: Extract alignments directly from the alignment
Template Extraction

Templates extracted from induced alignments of Liang, et. al (2009)

<table>
<thead>
<tr>
<th>Records:</th>
<th>skyCover₁</th>
<th>temperature₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields:</td>
<td>mode=50-75</td>
<td>time=17-30</td>
</tr>
<tr>
<td>Text:</td>
<td>mostly cloudy ,</td>
<td>low around</td>
</tr>
<tr>
<td></td>
<td></td>
<td>min=44 45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mean=49 .</td>
</tr>
</tbody>
</table>

**Ideal**: Extract alignments directly from the alignment

**Challenge**: Alignments are often noisy
## Template Extraction

Templates extracted from induced alignments of Liang, et. al (2009)

<table>
<thead>
<tr>
<th>Records:</th>
<th>skyCover&lt;sub&gt;1&lt;/sub&gt;</th>
<th>temperature&lt;sub&gt;1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields:</td>
<td>mode=50–75</td>
<td>time=17–30</td>
</tr>
<tr>
<td>Text:</td>
<td>mostly cloudy ,</td>
<td>low around</td>
</tr>
</tbody>
</table>

\[
\text{COARSE} < \text{[mode]} > < \text{with a [time] [min] [mean]} > \\
\text{BASE} \quad < \text{mostly cloudy ,} > < \text{with a low around} \text{[min]} >
\]

**Ideal:** Extract alignments directly from the alignment  
**Challenge:** Alignments are often noisy  
**Approach:** Extract templates of different granularities:
### Template Extraction

Templates extracted from induced alignments of Liang, et. al (2009)

<table>
<thead>
<tr>
<th>Records:</th>
<th>skyCover&lt;sub&gt;1&lt;/sub&gt;</th>
<th>temperature&lt;sub&gt;1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields:</td>
<td>mode=50–75, mostly cloudy.</td>
<td>time=17–30 &lt;br&gt; low around &lt;br&gt; min=44–45 &lt;br&gt; mean=49.</td>
</tr>
<tr>
<td>Text:</td>
<td>mostly cloudy.</td>
<td>with a low around [min].</td>
</tr>
</tbody>
</table>

**Coarse**: < [mode] > < with a [time] [min] [mean] >

**Base**: < mostly cloudy > < with a low around [min] >

**Ideal**: Extract alignments directly from the alignment

**Challenge**: Alignments are often noisy

**Approach**: Extract templates of different granularities:

**Coarse**: taken verbatim from the [noisy] alignment
Template Extraction

Templates extracted from induced alignments of Liang, et. al (2009)

<table>
<thead>
<tr>
<th>Records:</th>
<th>skyCover&lt;sub&gt;1&lt;/sub&gt;</th>
<th>temperature&lt;sub&gt;1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields:</td>
<td>mode=50-75</td>
<td>time=17-30</td>
</tr>
<tr>
<td>Text:</td>
<td>mostly cloudy ,</td>
<td>low around</td>
</tr>
</tbody>
</table>

Coarse < [mode] > < with a [time] [min] [mean] >

Base < mostly cloudy , > < with a low around [min] >

**Ideal:** Extract alignments directly from the alignment

**Challenge:** Alignments are often noisy

**Approach:** Extract templates of different granularities:

Coarse: taken verbatim from the [noisy] alignment

Base: created by abstracting field matching regexp (e.g. [0-9]+)
Features: Template

Note: Features defined over both template granularities

(R1) Template
\[ \text{BASE}(T_2) = \langle with \ a \ low \ around \ [\min] \rangle \]

\[ ..., T_1 = \langle mostly \ cloudy, \rangle, r_2 = \text{temperature}, F_2 = \{ \text{time}, \text{min} \} \]

\[ \langle with \ a \ low \ around \ [\min] \rangle \ 0.83 \]
\[ \langle with \ a \ high \ around \ [\max] \rangle \ 0.01 \]
\[ \langle with \ a \ high \ near \ [\max] \rangle \ 0.05 \]

...
Features: Template

Note: Features defined over both template granularities

(R1) Template
\[ \text{BASE}(T_2) = \langle \text{with a low around [min]} \rangle \]

(R2) Field Values
\[ \text{BASE}(T_2) = \langle \text{with a low around [min]} \rangle \text{ and } r_2.v[\text{time}] = 5\text{pm–6am} \]

..., T_1 = \langle mostly cloudy , \rangle, r_2 = \text{temperature}, F_2 = \{ \text{time, min} \}

\begin{align*}
\langle \text{with a low around [min]} \rangle & \quad 0.83 \\
\langle \text{with a high around [max]} \rangle & \quad 0.01 \\
\langle \text{with a high near [max]} \rangle & \quad 0.05 \\
... & 
\end{align*}
Features: Template

Note: Features defined over both template granularities

(R1) Template

\[ \text{BASE}(T_2) = \langle \text{with a low around} \ [\text{min}] \rangle \]

(R2) Field Values

\[ \text{BASE}(T_2) = \langle \text{with a low around} \ [\text{min}] \rangle \text{ and } r_2.v[\text{time}] = 5\text{pm}-6\text{am} \]

(R3) First word in template under LM

\[ \log p_{LM}(\text{with} \mid \text{cloudy},) \]

... \[ T_1 = \langle \text{mostly cloudy}, \rangle, r_2 = \text{temperature}, F_2 = \{\text{time, min}\} \]

\[ \langle \text{with a low around} \ [\text{min}] \rangle \ 0.83 \]

\[ \langle \text{with a high around} \ [\text{max}] \rangle \ 0.01 \]

\[ \langle \text{with a high near} \ [\text{max}] \rangle \ 0.05 \]

...
Features: Template

Note: Features defined over both template granularities

(R1) Template
\[ \text{BASE}(T_2) = \langle \text{with a low around} [\text{min}] \rangle \]

(R2) Field Values
\[ \text{BASE}(T_2) = \langle \text{with a low around} [\text{min}] \rangle \text{ and } r_2.v[\text{time}] = 5\text{pm}-6\text{am} \]

(R3) First word in template under LM
\[ \log p_{LM}(with | cloudy,) \]

..., \text{T}_1=\langle mostly cloudy ,\rangle, r_2=\text{temperature}, F_2=\{\text{time, min}\}

\[ \langle \text{with a low around} [\text{min}] . \rangle \quad 0.83 \]
\[ \langle \text{with a high around} [\text{max}] . \rangle \quad 0.01 \]
\[ \langle \text{with a high near} [\text{max}] . \rangle \quad 0.05 \]

...
Training: Log-Linear Model

Generation treated as a sequence of local decisions
Similar in spirit to Ratnaparkhi (1996)

<table>
<thead>
<tr>
<th>Record</th>
<th>$r_1 = \text{skyCover}_1$</th>
<th>$r_2 = \text{temperature}_1$</th>
<th>$r_3 = \text{STOP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field set</td>
<td>$F_1 = {\text{mode}}$</td>
<td>$F_2 = {\text{time, min}}$</td>
<td></td>
</tr>
<tr>
<td>Template</td>
<td>$W_1 = \langle \text{mostly cloudy} , \rangle$</td>
<td>$W_2 = \langle \text{with a low around [min]} . \rangle$</td>
<td></td>
</tr>
</tbody>
</table>
Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)

<table>
<thead>
<tr>
<th>Record</th>
<th>$d_0$</th>
<th>$d_3$</th>
<th>$d_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field set</td>
<td>$d_1$</td>
<td>$d_4$</td>
<td></td>
</tr>
<tr>
<td>Template</td>
<td>$d_2$</td>
<td>$d_5$</td>
<td></td>
</tr>
</tbody>
</table>
## Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)

<table>
<thead>
<tr>
<th>Record</th>
<th>$d_0$</th>
<th>$d_3$</th>
<th>$d_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field set</td>
<td>$d_1$</td>
<td>$d_4$</td>
<td></td>
</tr>
<tr>
<td>Template</td>
<td>$d_2$</td>
<td>$d_5$</td>
<td></td>
</tr>
</tbody>
</table>

### Log-linear model: Each decision trained using a log-linear model

\[
p(d_j \mid d_{j-1}, d_{j-2}, \ldots, s; \theta) \propto \exp\{\phi_j(d_j, d_{j-1}, d_{j-2}, \ldots, s) \top \theta\}
\]
Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)

<table>
<thead>
<tr>
<th>Record</th>
<th>Field set</th>
<th>Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_0$</td>
<td>$d_1$</td>
<td>$d_2$</td>
</tr>
<tr>
<td>$d_3$</td>
<td>$d_4$</td>
<td>$d_5$</td>
</tr>
<tr>
<td>$d_6$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Log-linear model**: Each decision trained using a log-linear model

$$p(d_j | d_{j-1}, d_{j-2}, \ldots, s; \theta) \propto \exp\{\phi_j(d_j, d_{j-1}, d_{j-2}, \ldots, s)^\top \theta\}$$

Log-likelihood maximized using L-BFGS
Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)

<table>
<thead>
<tr>
<th>Record</th>
<th>Field set</th>
<th>Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_0$</td>
<td>$d_1$</td>
<td>$d_2$</td>
</tr>
<tr>
<td>$d_3$</td>
<td>$d_4$</td>
<td>$d_5$</td>
</tr>
<tr>
<td>$d_6$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Log-linear model: Each decision trained using a log-linear model

$$p(d_j \mid d_{j-1}, d_{j-2}, \ldots, s; \theta) \propto \exp\{\phi_j(d_j, d_{j-1}, d_{j-2}, \ldots, s)^\top \theta\}$$

Log-likelihood maximized using L-BFGS

No dynamic program since decisions have long-range dependencies
Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)

Log-linear model: Each decision trained using a log-linear model

\[
p(d_j \mid d_{j-1}, d_{j-2}, \ldots, s; \theta) \propto \exp\{\phi_j(d_j, d_{j-1}, d_{j-2}, \ldots, s)^\top \theta\}
\]

Log-likelihood maximized using L-BFGS

No dynamic program since decisions have long-range dependencies

Training problem is fully supervised
Training: Log-Linear Model

Generation treated as a sequence of local decisions
Similar in spirit to Ratnaparkhi (1996)

<table>
<thead>
<tr>
<th>Record</th>
<th>$d_0$</th>
<th>$d_3$</th>
<th>$d_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field set</td>
<td>$d_1$</td>
<td>$d_4$</td>
<td></td>
</tr>
<tr>
<td>Template</td>
<td>$d_2$</td>
<td>$d_5$</td>
<td></td>
</tr>
</tbody>
</table>

**Log-linear model**: Each decision trained using a log-linear model

$$p(d_j \mid d_{j-1}, d_{j-2}, \ldots, s; \theta) \propto \exp\{\phi_j(d_j, d_{j-1}, d_{j-2}, \ldots, s)^\top \theta\}$$

Log-likelihood maximized using L-BFGS

No dynamic program since decisions have long-range dependencies

Training problem is fully supervised

Generating a new text is done using greedy search
Evaluation: Domains

**Robocup**: RoboCup sportscasting data
Evaluation: Domains

**Robocup**: RoboCup sportscasting data

**Content Selection**: Chose one record

\[
\begin{align*}
    s: & \quad \text{kick}(\text{arg1}=\text{purple3}) \\
    & \quad \text{badPass}(\text{arg1}=\text{purple3}, \text{arg2}=\text{pink9}) \\
    & \quad \text{turnover}(\text{arg1}=\text{purple3}, \text{arg2}=\text{pink9})
\end{align*}
\]
Evaluation: Domains

Robocup: RoboCup sportscasting data

Content Selection: Chose one record

s:

- kick(arg1=purple3)
- badPass(arg1=purple3, arg2=pink9)
- turnover(arg1=purple3, arg2=pink9)

Surface Realization:

w: Purple3 made a bad pass that was picked off by pink9.
Evaluation: Domains

Robocup: RoboCup sportscasting data
SumTime: Weather forecasts for oil rigs
Evaluation: Domains

**RoboCup**: RoboCup sportscasting data

**SumTime**: Weather forecasts for oil rigs

**Content Selection**: No content selection

\[
\begin{align*}
\text{s:} & \quad \text{wind10m}(\text{time}=6\text{am}, \text{dir}=\text{SW}, \text{min}=16, \text{max}=20, \text{gust min}=0, \text{gust max}=0) \\
& \quad \text{wind10m}(\text{time}=9\text{pm}, \text{dir}=\text{SSW}, \text{min}=28, \text{max}=32, \text{gust min}=40, \text{gust max}=0) \\
& \quad \text{wind10m}(\text{time}=12\text{am}, \text{dir}=\text{-}, \text{min}=24, \text{max}=28, \text{gust min}=36, \text{gust max}=0)
\end{align*}
\]
Evaluation: Domains

**Robocup**: RoboCup sportscasting data

**SumTime**: Weather forecasts for oil rigs

**Content Selection**: No content selection

s:

\[
\begin{align*}
\text{wind10m}(\text{time}=6\text{am}, \text{dir}=\text{SW}, \text{min}=16, \text{max}=20, \text{gust min}=0, \text{gust max}=0) \\
\text{wind10m}(\text{time}=9\text{pm}, \text{dir}=\text{SSW}, \text{min}=28, \text{max}=32, \text{gust min}=40, \text{gust max}=0) \\
\text{wind10m}(\text{time}=12\text{am}, \text{dir}=-, \text{min}=24, \text{max}=28, \text{gust min}=36, \text{gust max}=0)
\end{align*}
\]

**Surface Realization**:

w: *sw 16 - 20 backing ssw 28 - 32 gusts 40 by mid evening easing 24 - 28 gusts 36 late evening*
Evaluation: Domains

\textbf{Robocup}: RoboCup sportscasting data
\textbf{SumTime}: Weather forecasts for oil rigs
\textbf{WeatherGov}: Weather forecasts from \texttt{weather.gov}
Evaluation: Domains

**Robocup**: RoboCup sportscasting data

**SumTime**: Weather forecasts for oil rigs

**WeatherGov**: Weather forecasts from weather.gov

**Content Selection**: Complex; chose a few out of around 35 records

<table>
<thead>
<tr>
<th>s:</th>
</tr>
</thead>
<tbody>
<tr>
<td>temperature(time=5pm-6am,min=48,mean=53,max=61)</td>
</tr>
<tr>
<td>windSpeed(time=5pm-6am,min=3,mean=6,max=11,mode=0-10)</td>
</tr>
<tr>
<td>windDir(time=5pm-6am,mode=SSW)</td>
</tr>
<tr>
<td>gust(time=5pm-6am,min=0,mean=0,max=0)</td>
</tr>
<tr>
<td>skyCover(time=5pm-9pm,mode=0-25)</td>
</tr>
<tr>
<td>skyCover(time=2am-6am,mode=75-100)</td>
</tr>
<tr>
<td>precipPotential(time=5pm-6am,min=2,mean=14,max=20)</td>
</tr>
<tr>
<td>rainChance(time=5pm-6am,mode=someChance)</td>
</tr>
</tbody>
</table>
Evaluation: Domains

**Robocup**: RoboCup sportscasting data

**SumTime**: Weather forecasts for oil rigs

**WeatherGov**: Weather forecasts from weather.gov

**Content Selection**: Complex; chose a few out of around 35 records

| s: | temperature(time=5pm-6am, min=48, mean=53, max=61) |
|    | windSpeed(time=5pm-6am, min=3, mean=6, max=11, mode=0-10) |
|    | windDir(time=5pm-6am, mode=SSW) |
|    | gust(time=5pm-6am, min=0, mean=0, max=0) |
|    | skyCover(time=5pm-9pm, mode=0-25) |
|    | skyCover(time=2am-6am, mode=75-100) |
|    | precipPotential(time=5pm-6am, min=2, mean=14, max=20) |
|    | rainChance(time=5pm-6am, mode=someChance) |

**Surface Realization:**

**w**: A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.
Evaluation: Metrics

Automatic metrics:

   Content selection: $F_1$ score (record precision and recall)
   Surface realization: BLEU score (system output versus human annotated)
Evaluation: Metrics

Automatic metrics:

Content selection: $F_1$ score (record precision and recall)
Surface realization: BLEU score (system output versus human annotated)
Objective metric, but may not agree with human judgement
Evaluation: Metrics

Automatic metrics:

Content selection: $F_1$ score (record precision and recall)
Surface realization: BLEU score (system output versus human annotated)
Objective metric, but may not agree with human judgement

Human evaluation: Mechanical turk
Evaluation: Metrics

Automatic metrics:

- Content selection: $F_1$ score (record precision and recall)
- Surface realization: BLEU score (system output versus human annotated)

Objective metric, but may not agree with human judgement

Human evaluation: Mechanical turk

100 random shuffled scenarios shown to 10 Turkers
Evaluation: Metrics

Automatic metrics:

Content selection: $F_1$ score (record precision and recall)

Surface realization: BLEU score (system output versus human annotated)

Objective metric, but may not agree with human judgement

Human evaluation: Mechanical turk

100 random shuffled scenarios shown to 10 Turkers

Rank *english fluency* and *semantic correctness* on 1-5 scale

<table>
<thead>
<tr>
<th>Score</th>
<th>English Fluency</th>
<th>Semantic Correctness</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Flawless</td>
<td>Perfect</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
<td>Near Perfect</td>
</tr>
<tr>
<td>3</td>
<td>Non-native</td>
<td>Minor Errors</td>
</tr>
<tr>
<td>2</td>
<td>Disfluent</td>
<td>Major Errors</td>
</tr>
<tr>
<td>1</td>
<td>Gibberish</td>
<td>Completely Wrong</td>
</tr>
</tbody>
</table>
Evaluation: Metrics

s:
- temperature (time=5pm-6am, min=53, mean=57, max=67)
- windSpeed (time=5pm-6am, min=16, mean=20, max=21, mode=10-20)
- windDir (time=5pm-6am, mode=S)
- gust (time=5pm-6am, min=23, mean=27, max=29)
- skyCover (time=5pm-6am, mode=75-100)
- precipPotential (time=5pm-6am, min=19, mean=75, max=93)
- rainChance (time=5pm-9pm, mode=definitely)
- rainChance (time=2am-6am, mode=chance)
- thunderChance (time=5pm-6am, mode=definitely)

w: Rain. Some of the storms could be severe and 53. South wind between 16 and 21 mph chance of precipitation is 95%.

English Fluency: 3.45    Semantic Correctness: 4.00
Evaluation: Systems

Baseline

Subset of features

R1  List of last record type
R5  Stop under language model
F1  Field set
T1  Template
T3  First word of template under LM
Evaluation: Systems

Baseline

Subset of features

- **R1**: List of last record type
- **R5**: Stop under language model
- **F1**: Field set
- **T1**: Template
- **T3**: First word of template under LM

Bigram record model; most common field set; language model
Evaluation: Systems

Baseline

Subset of features

R1 List of last record type
R5 Stop under language model
F1 Field set
T1 Template
T3 First word of template under LM

Bigram record model; most common field set; language model

Oursystem: All features
Evaluation: Systems

Baseline
Subset of features

- **R1** List of last record type
- **R5** Stop under language model
- **F1** Field set
- **T1** Template
- **T3** First word of template under LM

Bigram record model; most common field set; language model

**OurSystem**: All features

**OurSystem-Custom**: 8 domain-dependent features for SumTime
Evaluation: Systems

Baseline

Subset of features

R1 List of last record type
R5 Stop under language model
F1 Field set
T1 Template
T3 First word of template under LM

Bigram record model; most common field set; language model

OurSystem: All features

OurSystem-Custom: 8 domain-dependent features for SumTime

WASPER-GEN: Robocup state of the art by Chen and Mooney (2008)
Evaluation: Systems

Baseline

Subset of features

R1 List of last record type
R5 Stop under language model
F1 Field set
T1 Template
T3 First word of template under LM

Bigram record model; most common field set; language model

OurSystem: All features

OurSystem-Custom: 8 domain-dependent features for SumTime

WASPER-GEN: Robocup state of the art by Chen and Mooney (2008)

SumTime-Hybrid: SumTime hand crafted system of Reiter, et al. (2005)
Evaluation: Systems

Baseline
Subset of features

R1 List of last record type
R5 Stop under language model
F1 Field set
T1 Template
T3 First word of template under LM

Bigram record model; most common field set; language model

OurSystem: All features
OurSystem-Custom: 8 domain-dependent features for SumTime

WASPER-GEN: Robocup state of the art by Chen and Mooney (2008)
SumTime-Hybrid: SumTime hand crafted system of Reiter, et al. (2005)
pCRU-greedy: SumTime state of the art by Belz (2008)
Evaluation: Robocup Results

![Bar chart showing F1 scores for different systems in content selection. The chart compares WASPER-GEN and OurSystem.]
Evaluation: Robocup Results

- F1 Score
- BLEU Score

- WASPER-GEN
- OURSYSTEM
Evaluation: Robocup Results

![Chart showing F1 Score for Content Selection and BLEU Score for Surface Realization.]

- **WASPER-GEN**
- **OursSystem**

![Bar chart showing Average Turker Score for Fluency and Correctness Human Evaluation.]

- Fluency Human Evaluation
- Correctness Human Evaluation
Evaluation: **SumTime** Results

![Graph showing BLEU Score vs. Surface Realization for SumTime-Hybrid and OurSystem]
Evaluation: \texttt{SumTime} Results

![Bar chart showing BLEU scores for different systems.

- \texttt{SumTime-Hybrid}
- \texttt{OurSystem}
- \texttt{pCRU-greedy}
- \texttt{OurSystem-Custom}]}
Evaluation: **SumTime** Results

- **BLEU Score**
- **Correctness**
- **Human Evaluation**

### Surface Realization

- **SumTime-Hybrid**
- **OurSystem**
- **pCRU-greedy**
- **OurSystem-Custom**

### Average Turker Score

- **Fluency**
- **Correctness**

---

*Note: The chart shows the comparison of different systems in terms of BLEU score, correctness, and human evaluation.*
Evaluation: WeatherGov Results

<table>
<thead>
<tr>
<th>F1 Score</th>
<th>Baseline</th>
<th>OurSystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Content Selection

- **Baseline**
- **OurSystem**
Evaluation: WeatherGov Results

F_1 Score

Baseline

Our System

Surface Realization

Baseline

Our System
Evaluation: WeatherGov Results

- **Content Selection**
  - F1 Score
  - Baseline: 20.0
  - OurSystem: 60.0

- **Surface Realization**
  - BLEU Score
  - Baseline: 20.0
  - OurSystem: 60.0

- **Human Evaluation**
  - Fluency
    - Baseline: 4.0
    - OurSystem: 4.0
  - Correctness
    - Baseline: 3.5
    - OurSystem: 3.5
  - Average Turker Score: 4.2

Legend:
- Red: Baseline
- Blue: OurSystem
Evaluation: \textbf{WEATHERGOV} Results

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{chart.png}
\end{figure}
Summary

Complete pipeline

Same framework for content selection and surface realization
Summary

Complete pipeline
  Same framework for content selection and surface realization

Domain independent Approach
  Minimal tweaking between domains
Summary

Complete pipeline
  Same framework for content selection and surface realization

Domain independent Approach
  Minimal tweaking between domains

Data driven
  Weighted tuned to human generated output
Summary

Complete pipeline
  Same framework for content selection and surface realization

Domain independent Approach
  Minimal tweaking between domains

Data driven
  Weighted tuned to human generated output

Comparable results
  Results comparable to state of the art
Thank You!

Berkeley

NLP