A Simple Domain-Independent Approach to Generation

Gabor Angeli
gangeli@berkeley.edu

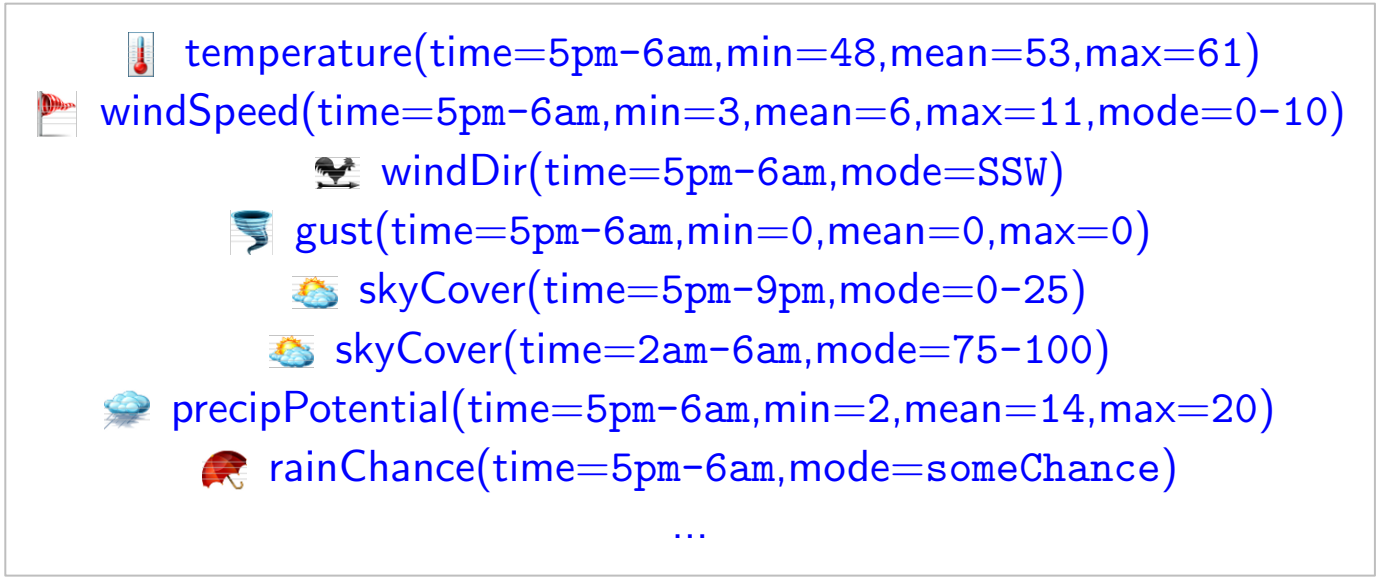
Percy Liang
pliang@eecs.berkeley.edu

Dan Klein
klein@eecs.berkeley.edu

Introduction

Task: Generate text from database records

s:

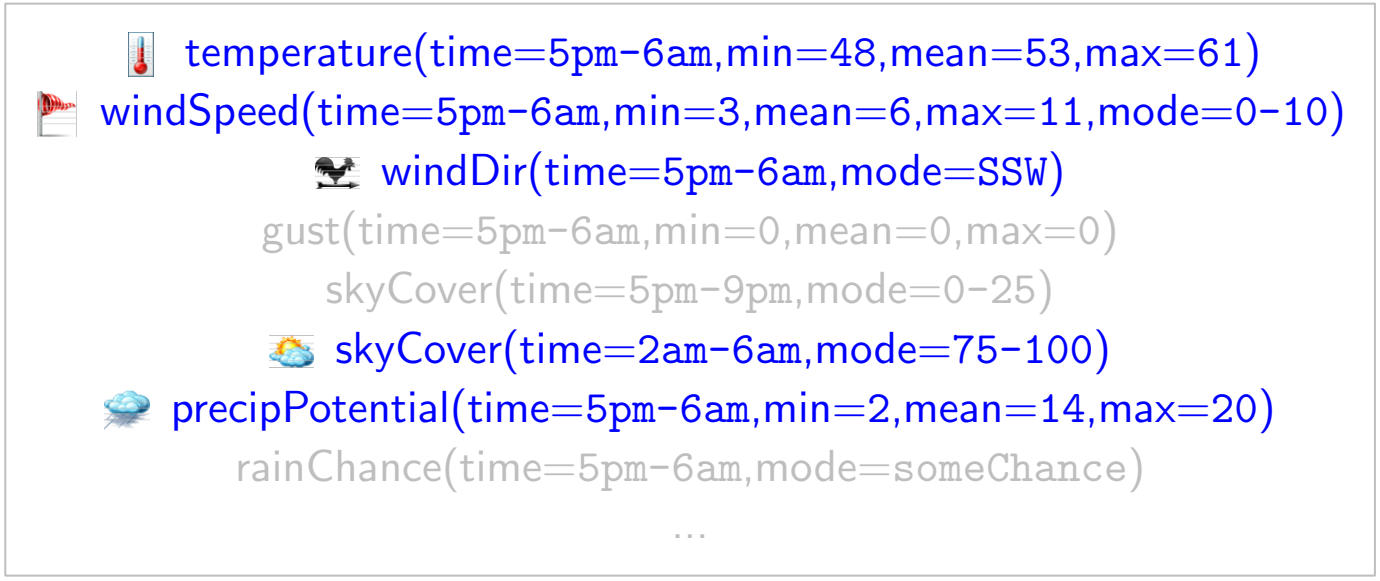


w: *A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.*

Introduction

Task: Generate text from database records

s:

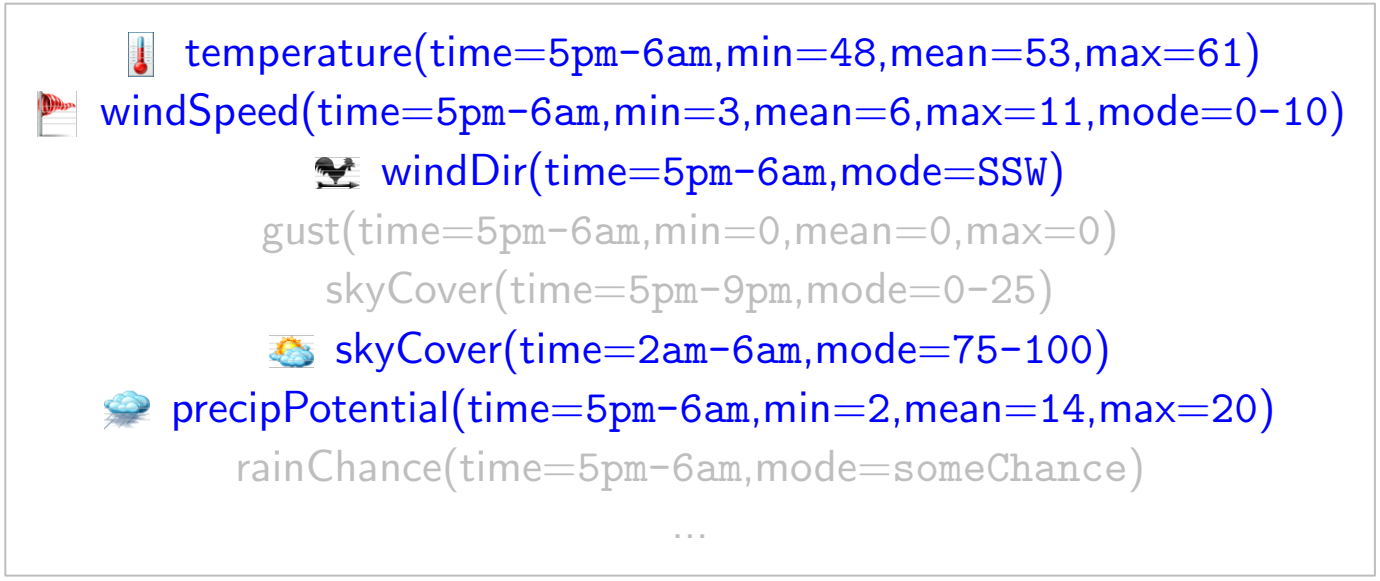


Requires both **content selection**

Introduction

Task: Generate text from database records

s:



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:

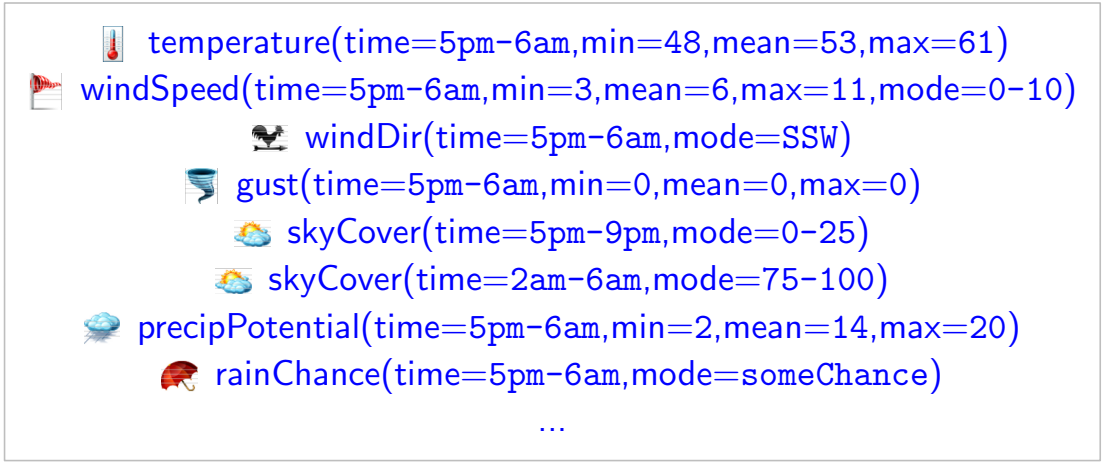
- Do content selection and surface realization in a unified framework
  - Content Selection: e.g. Barzilay and Lee (2004)
  - Surface Realization: e.g. Soricut and Marcu (2006), White, et al. (2007)

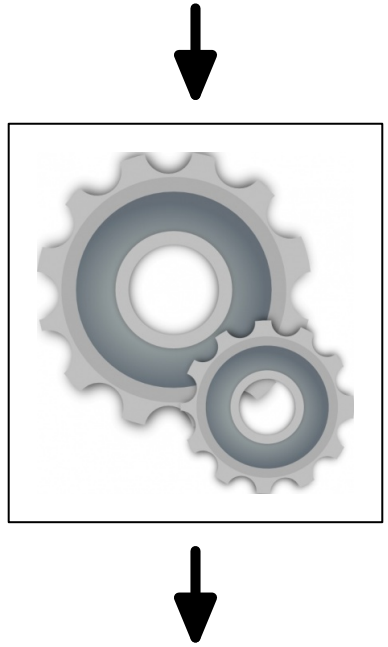
Domain Independent Approach:

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

Domain Independent

s:



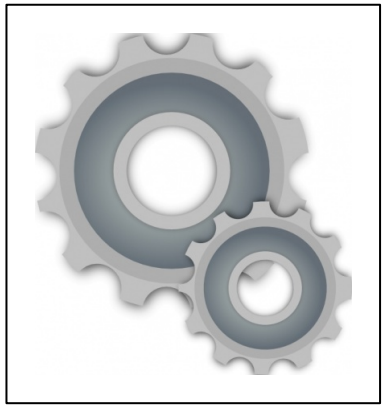


w: *A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.*

Domain Independent

s:

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

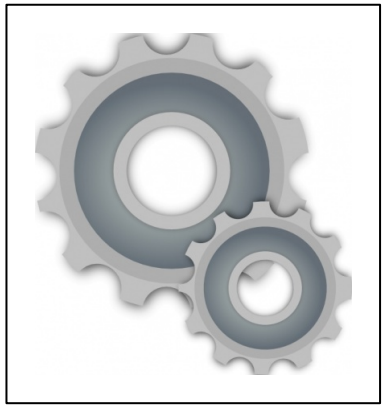


w: *Purple3 made a bad pass that was picked off by pink9.*

Domain Independent

s:

wind10m(time=6am,dir=SW,min=16,max=20,gust min=0,gust max=0)
wind10m(time=9pm,dir=SSW,min=28,max=32,gust min=40,gust max=0)
wind10m(time=12am,dir=-,min=24,max=28,gust min=36,gust max=0)



w: *sw 16 - 20 backing ssw 28 - 32 gusts 40 by mid evening 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

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.*

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.*

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.*

s:

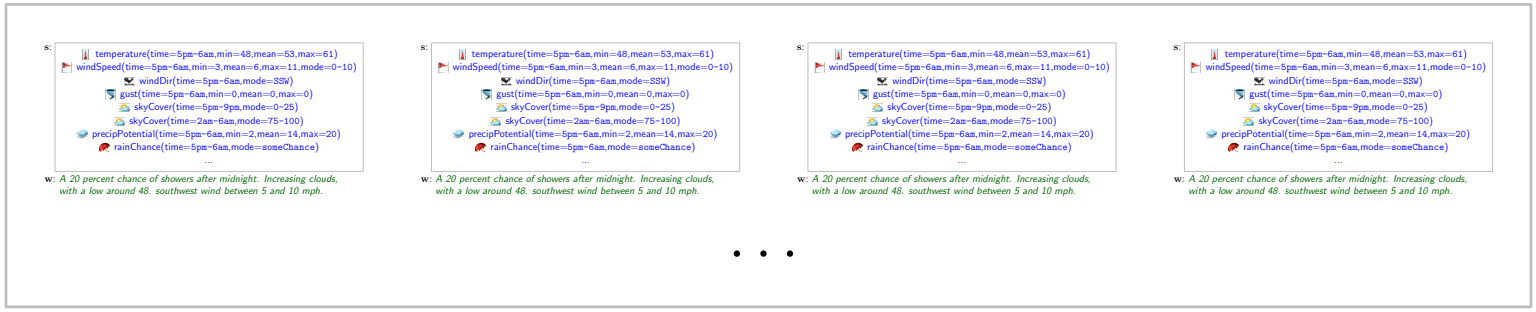
temperature(time=5pm-6am,min=48,mean=53,max=61)
windSpeed(time=5pm-6am,min=3,mean=6,max=11,mode=0-10)
windDir(time=5pm-6am,mode=SSW)
gust(time=5pm-6am,min=0,mean=0,max=0)
skyCover(time=5pm-9pm,mode=0-25)
skyCover(time=2am-6am,mode=75-100)
precipPotential(time=5pm-6am,min=2,mean=14,max=20)
rainChance(time=5pm-6am,mode=someChance)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

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

• • •

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)

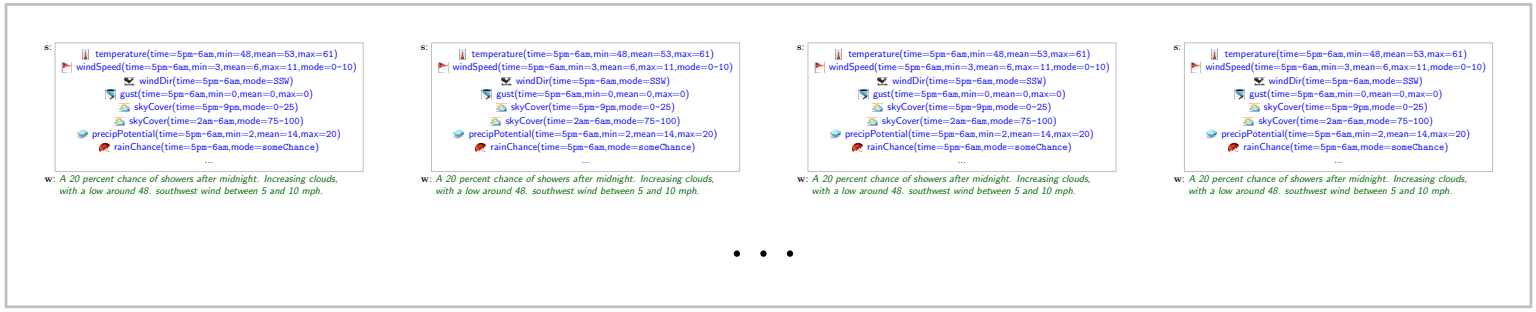
precipPotential(time=5pm-6am,min=2,mean=14,max=20)

rainChance(time=5pm-6am,mode=someChance)

...

Learning Based

Train



Test

s:

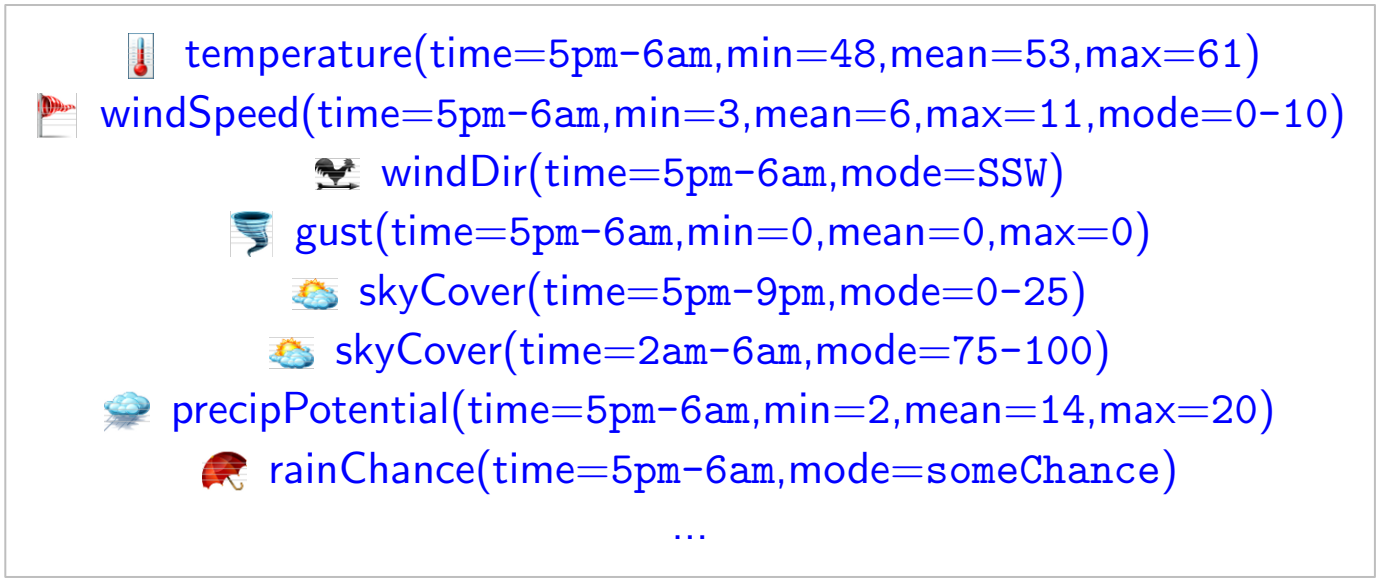
temperature(time=5pm-6am,min=48,mean=53,max=61)
windSpeed(time=5pm-6am,min=3,mean=6,max=11,mode=0-10)
windDir(time=5pm-6am,mode=SSW)
gust(time=5pm-6am,min=0,mean=0,max=0)
skyCover(time=5pm-9pm,mode=0-25)
skyCover(time=2am-6am,mode=75-100)
precipPotential(time=5pm-6am,min=2,mean=14,max=20)
rainChance(time=5pm-6am,mode=someChance)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

↓

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

Representation

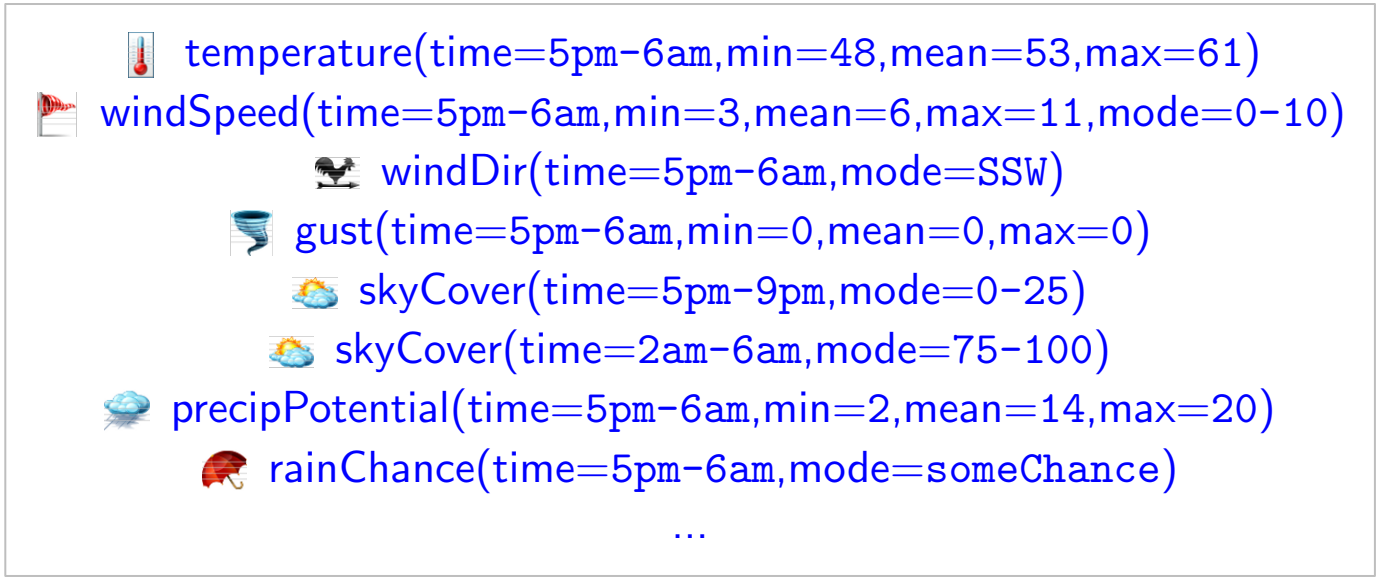
s:



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

Representation

s:

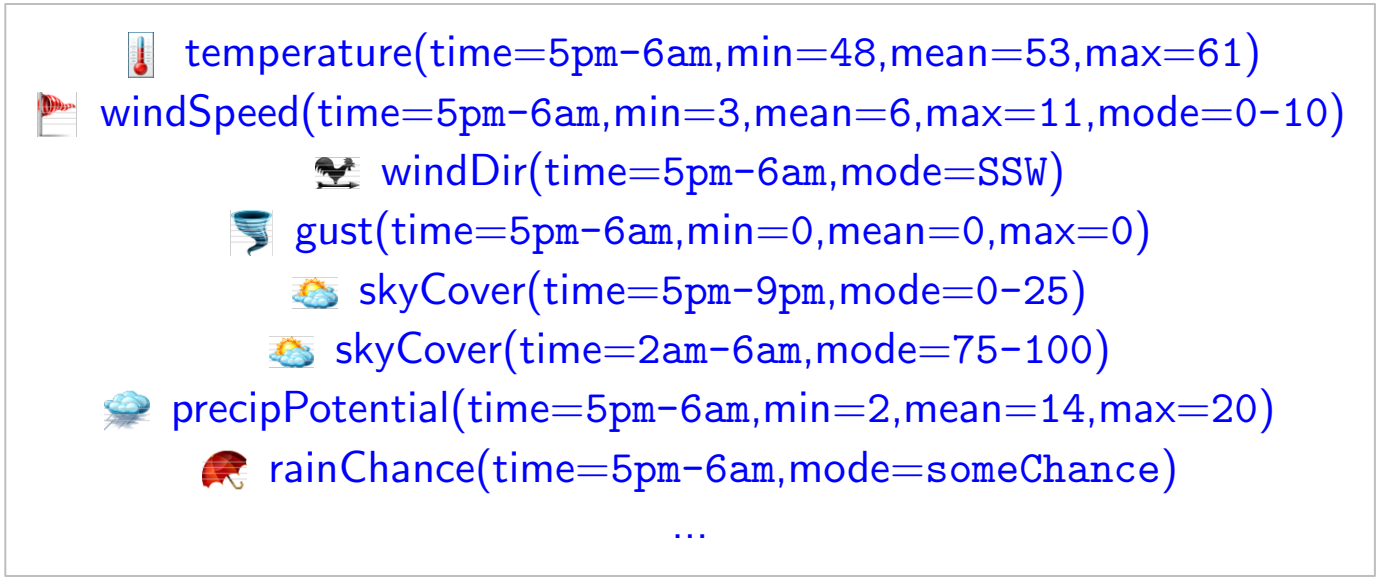


w: *A 20 percent chance of showers after midnight. Increasing clouds, with a low around 48. southwest wind between 5 and 10 mph.*

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

Representation

s:



w: *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**

Representation

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.*

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

World state: Set of database **records**

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,mode=0-10)

Representation

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,mode=0-10)

Representation

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,mode=0-10)

Each **record** is of a given **record type**

Representation

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,mode=0-10)

Each **record** is of a given **record type**

Each **record** has a set of **fields**

Representation

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,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

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,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

Representation

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,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

$\mathbf{r}_1 =$  windSpeed(time=5pm-6am,min=3,mean=6,max=11,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**

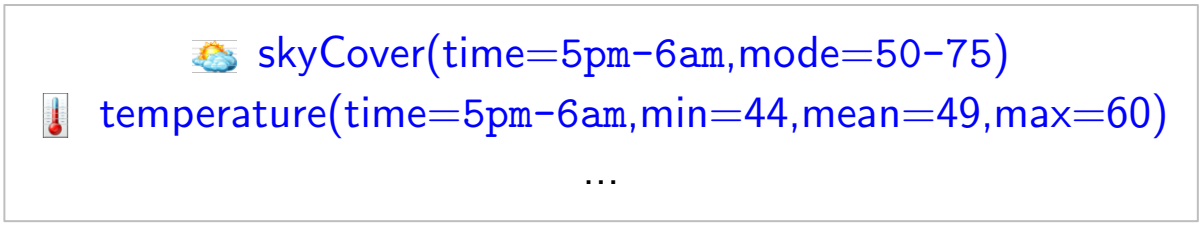
temperature(time=5pm-6am,min=48,mean=53,max=61)

⟨*with a low around* [min] .⟩

Templates extracted automatically from the training data (stay tuned)

Generation as Decisions

s:



Generation as Decisions

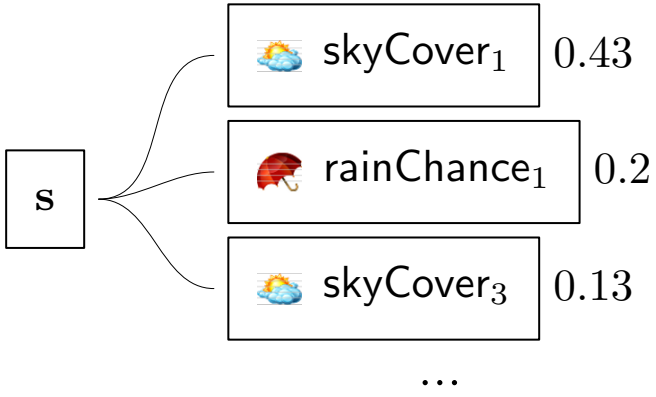
s:

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

Record

Field set

Template



Generation as Decisions

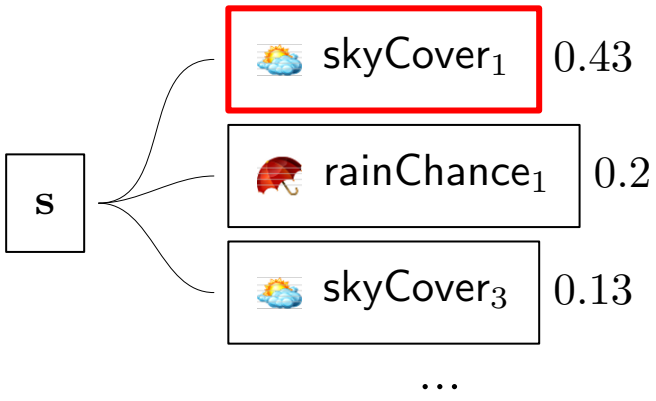
s:

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

Record

Field set

Template



Generation as Decisions

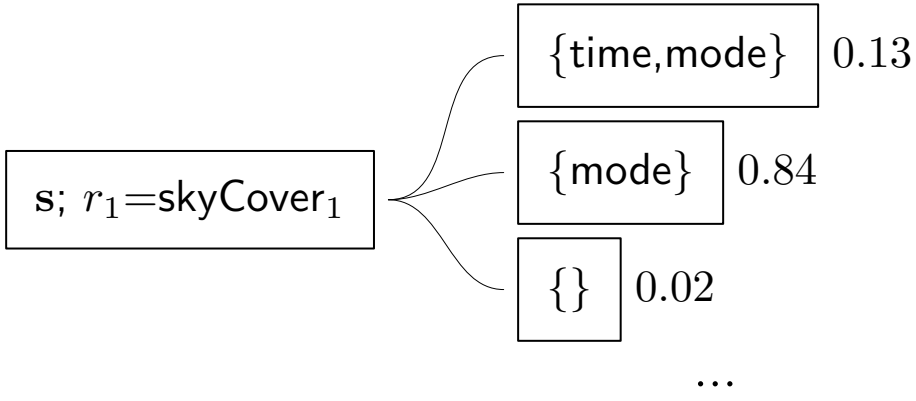
s:

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

Record r_1 = skyCover$_1$

Field set

Template



Generation as Decisions

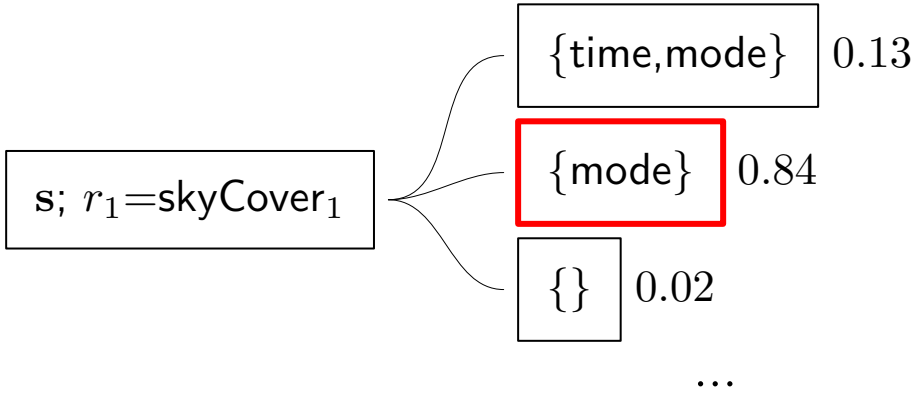
s:

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

Record r_1 = skyCover$_1$

Field set

Template



Generation as Decisions

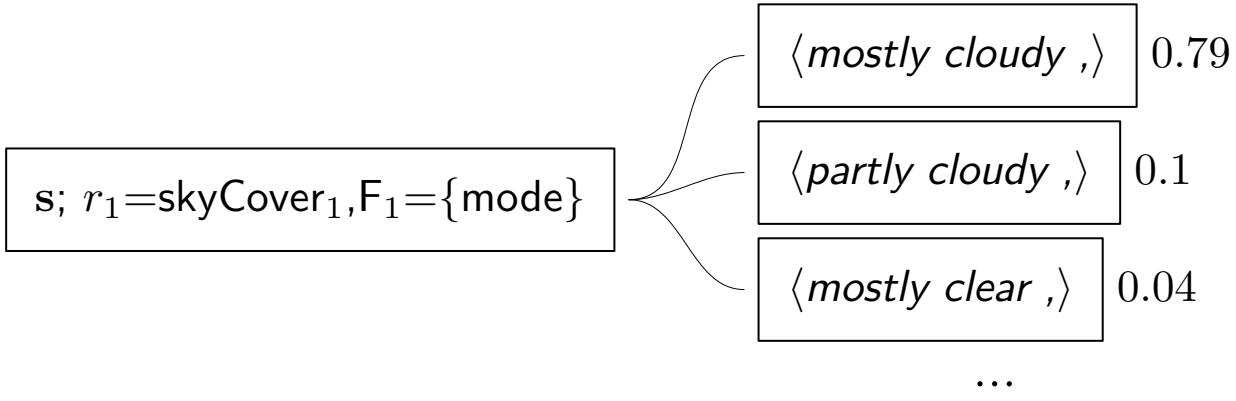
s:

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

Record $r_1 =$ skyCover$_1$

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

Template



Generation as Decisions

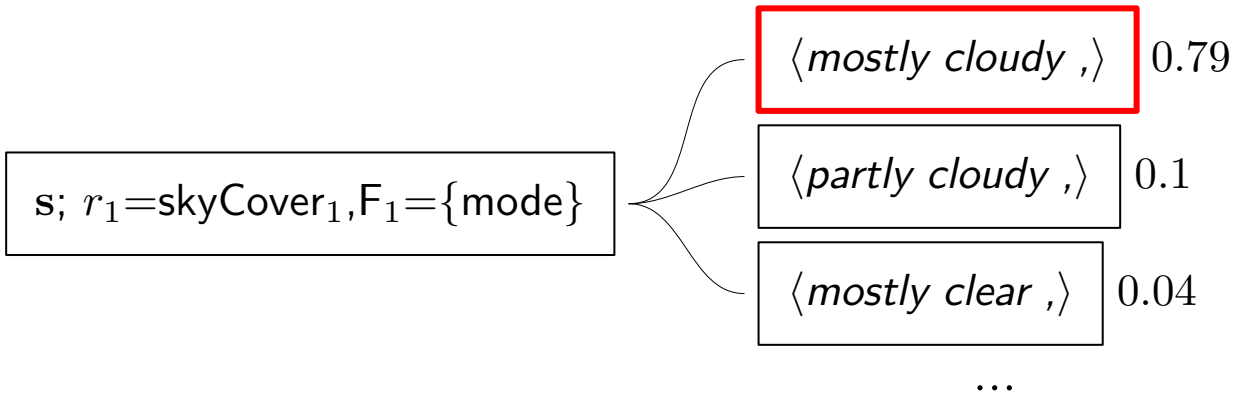
s:

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

Record $r_1 =$ skyCover$_1$

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

Template



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

Template $W_1 = \langle \textit{mostly cloudy} , \rangle$

s; r_1=skyCover$_1$,F$_1$={mode},T$_1$=$\langle$*mostly cloudy* ,$\rangle$

- rainChance$_1$ 0.02
- thunderChance$_3$ 0.01
- temperature$_1$ 0.95
- ...

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

Template $W_1 = \langle \textit{mostly cloudy} , \rangle$

s; r_1=skyCover$_1$,F$_1$={mode},T$_1$=$\langle$*mostly cloudy* ,$\rangle$

- rainChance$_1$ 0.02
- thunderChance$_3$ 0.01
- temperature$_1$ 0.95

...

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$_1$ $r_2 =$ temperature$_1$

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

Template $W_1 = \langle \textit{mostly cloudy} , \rangle$

s; ...,F_1={mode},T_1=$\langle$*mostly cloudy* ,$\rangle$,r_2=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 $r_1 =$ skyCover$_1$ $r_2 =$ temperature$_1$

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

Template $W_1 = \langle \textit{mostly cloudy} , \rangle$

s; ...,F_1={mode},T_1=⟨*mostly cloudy* ,⟩,r_2=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	$r_1 =$ skyCover$_1$	$r_2 =$ temperature$_1$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$
Template	$W_1 = \langle \textit{mostly cloudy} , \rangle$	

s; ...,T_1=⟨*mostly cloudy* ,⟩,r_2=temperature,F_2={time,min}

- ⟨*with a low around* [min .⟩ 0.83
- ⟨*with a high around* [max] .⟩ 0.01
- ⟨*with a high near* [max] .⟩ 0.05

...

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$_1$	$r_2 =$ temperature$_1$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$
Template	$W_1 = \langle \textit{mostly cloudy} , \rangle$	

s; ...,T_1=⟨*mostly cloudy* ,⟩,r_2=temperature,F_2={time,min}

- ⟨*with a low around* [min .⟩ 0.83
- ⟨*with a high around* [max] .⟩ 0.01
- ⟨*with a high near* [max] .⟩ 0.05

...

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$_1$	$r_2 =$ temperature$_1$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$
Template	$W_1 = \langle \textit{mostly cloudy} , \rangle$	$W_2 = \langle \textit{with a low around } [\min] \ . \rangle$

s; ...,r$_2$=temperature,F$_2$={time,min},T$_2$=⟨*with a low around* min .⟩

- windDir$_1$ 0.37
- temperature$_1$ 0.03
- STOP 0.43
- ...

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$_1$	$r_2 =$ temperature$_1$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$
Template	$W_1 = \langle \textit{mostly cloudy} , \rangle$	$W_2 = \langle \textit{with a low around } [\text{min}] \ . \rangle$

s; ...,r$_2$=temperature,F$_2$={time,min},T$_2$=⟨*with a low around* min .⟩

- windDir$_1$ 0.37
- temperature$_1$ 0.03
- STOP 0.43
- ...

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$_1$	$r_2 =$ temperature$_1$	$r_3 = \text{STOP}$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$	
Template	$W_1 = \langle$*mostly cloudy ,*$\rangle$	$W_2 = \langle$*with a low around* $[\text{min}]$.$\rangle$	

mostly cloudy , with a low around 44 .

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$_1$	$r_2 =$ temperature$_1$	$r_3 = \text{STOP}$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$	
Template	$W_1 = \langle$*mostly cloudy ,*$\rangle$	$W_2 = \langle$*with a low around* $[\text{min}]$.$\rangle$	

mostly cloudy , with a low around 44 .

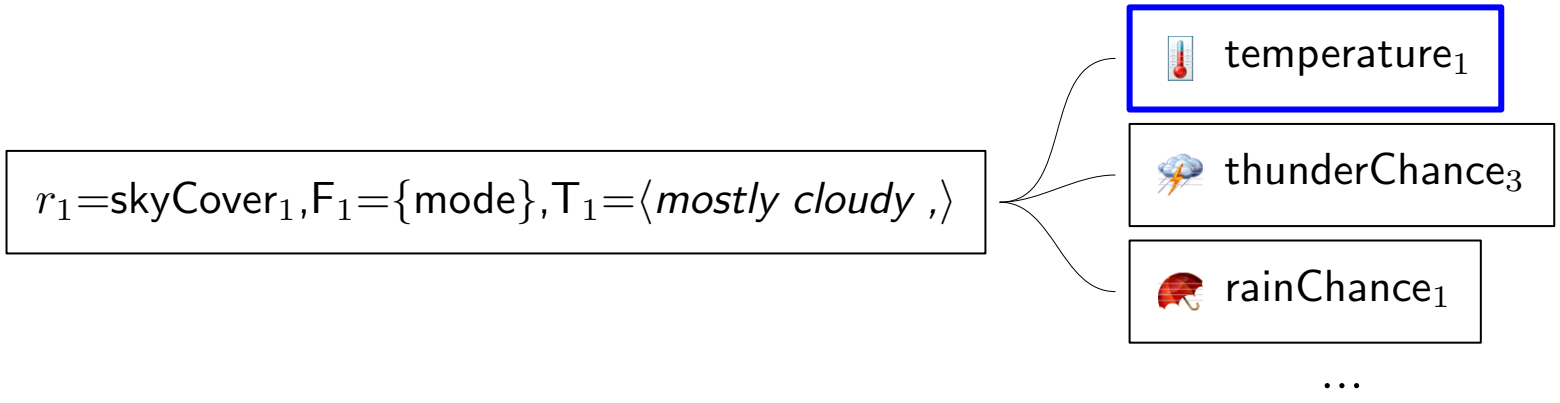
Features: Record

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

Features: Record

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

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



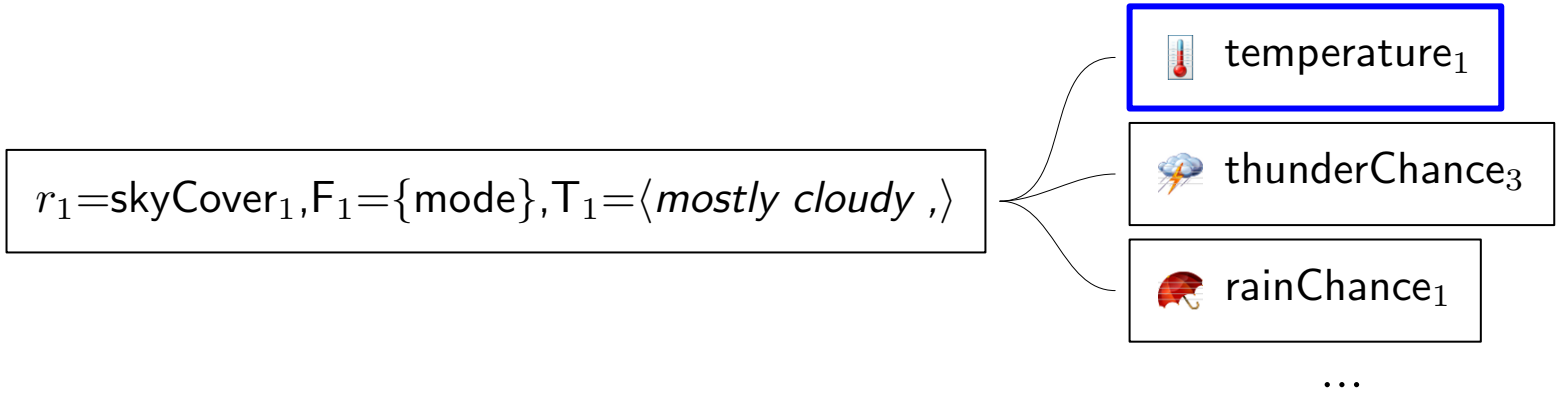
Features: Record

(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**

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



Features: Record

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

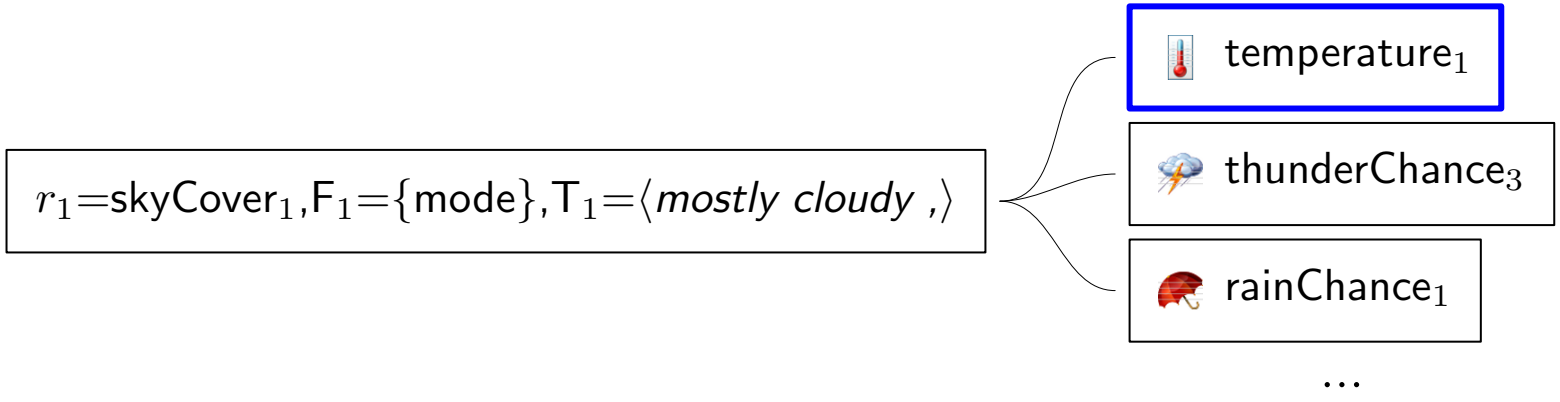
$[\![r_2.t = \text{temperature and } (r_1.t, r_0.t) = (\text{skyCover}, \text{START})]\!]$

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

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

(R3) **Record type already generated**

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



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**

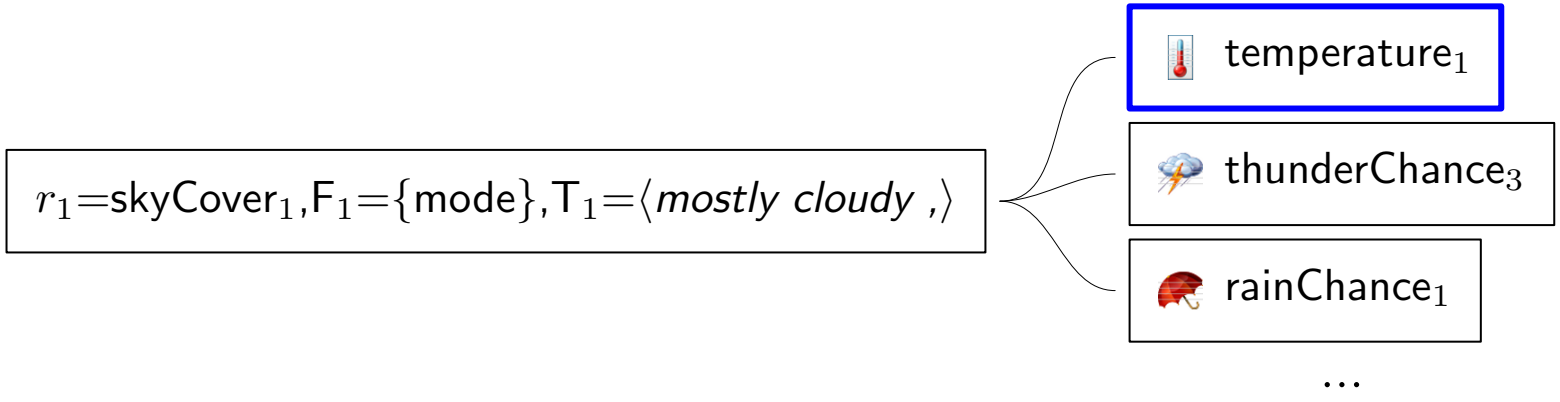
$[\![r_2.t = \text{temperature and } \{r_1.t\} = \{\text{skyCover}\}]\!]$

(R3) **Record type already generated**

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

(R4) **Field values**

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



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**

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

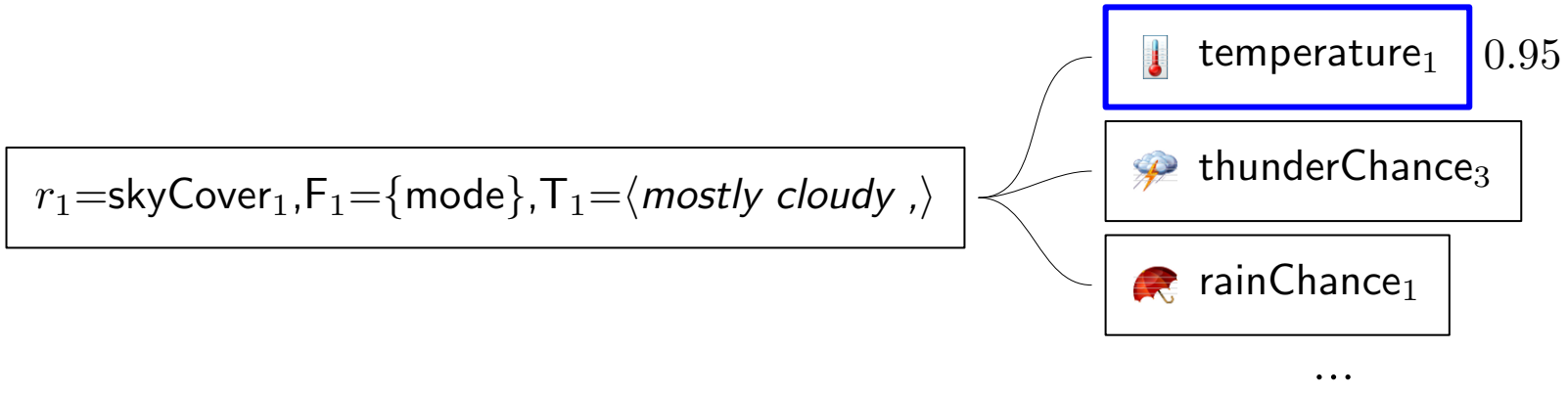
(R3) **Record type already generated**

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

(R4) **Field values**

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

(R5) **Stop under language model**



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**

$[\![r_2.t = \text{thunderChance and } \{r_1.t\} = \{\text{skyCover}\}]\!]$

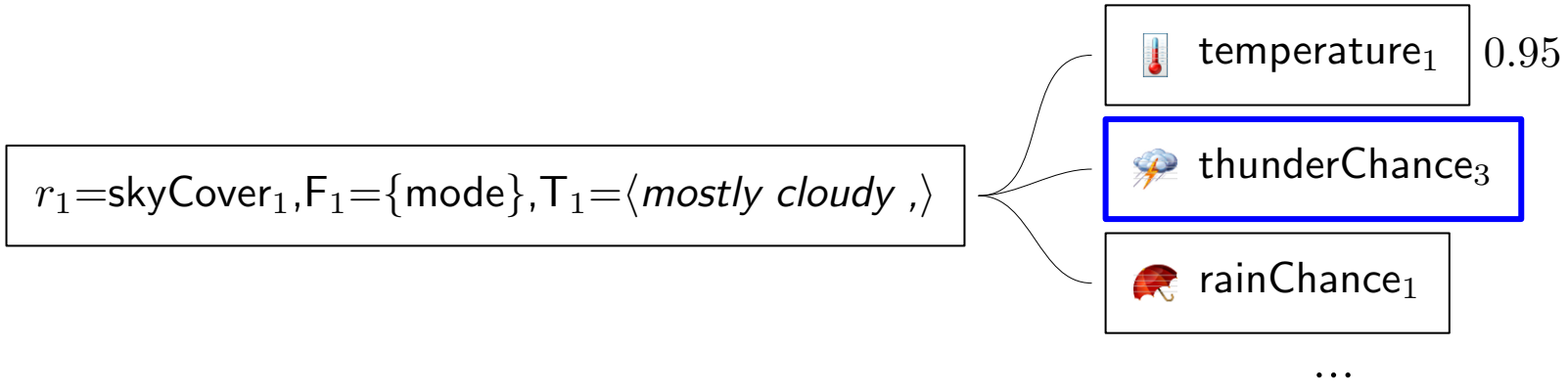
(R3) **Record type already generated**

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

(R4) **Field values**

$[\![r_2.t = \text{thunderChance and } r_2.v[\text{time}] = \texttt{2am-6am}]\!]$

(R5) **Stop under language model**



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**

$[\![r_2.t = \text{thunderChance and } \{r_1.t\} = \{\text{skyCover}\}]\!]$

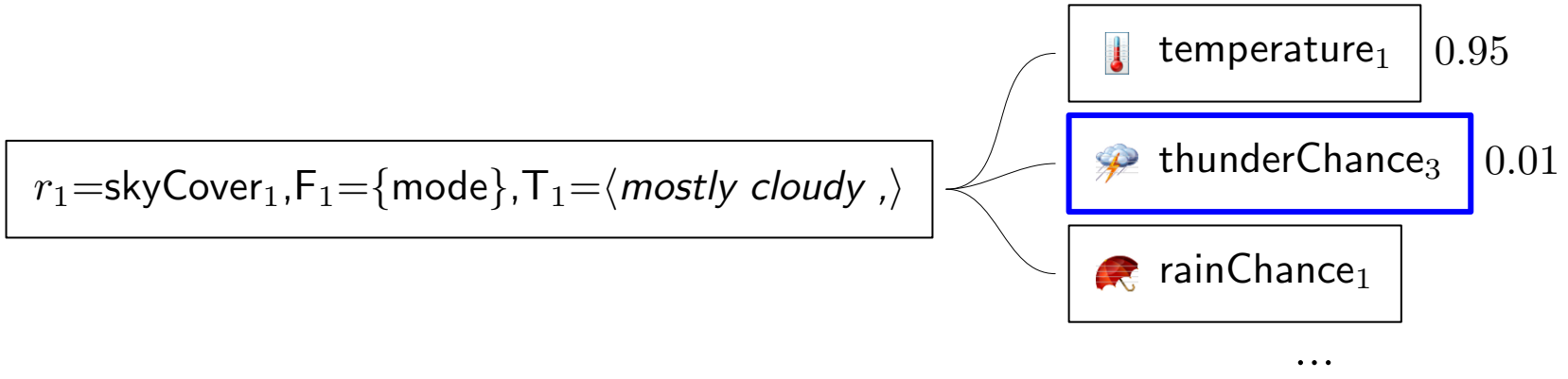
(R3) **Record type already generated**

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

(R4) **Field values**

$[\![r_2.t = \text{thunderChance and } r_2.v[\text{time}] = \texttt{2am-6am}]\!]$

(R5) **Stop under language model**



Features: Record

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

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

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

$[\![r_2.t = \text{rainChance and } \{r_1.t\} = \{\text{skyCover}\}]\!]$

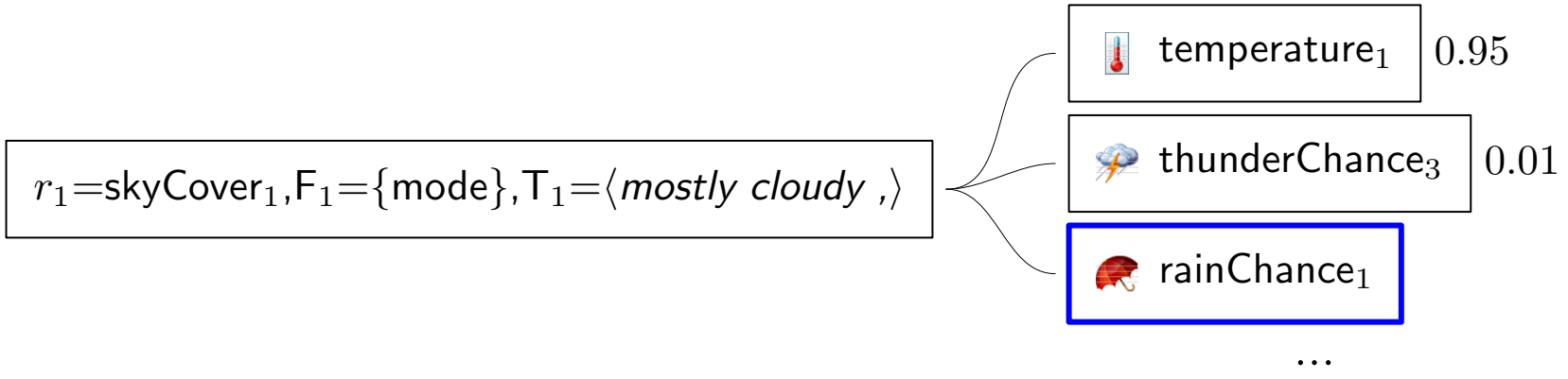
(R3) **Record type already generated**

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

(R4) **Field values**

$[\![r_2.t = \text{rainChance and } r_2.v[\text{time}] = \texttt{5pm-6am}]\!]$

(R5) **Stop under language model**



Features: Record

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

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

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

$[\![r_2.t = \text{rainChance and } \{r_1.t\} = \{\text{skyCover}\}]\!]$

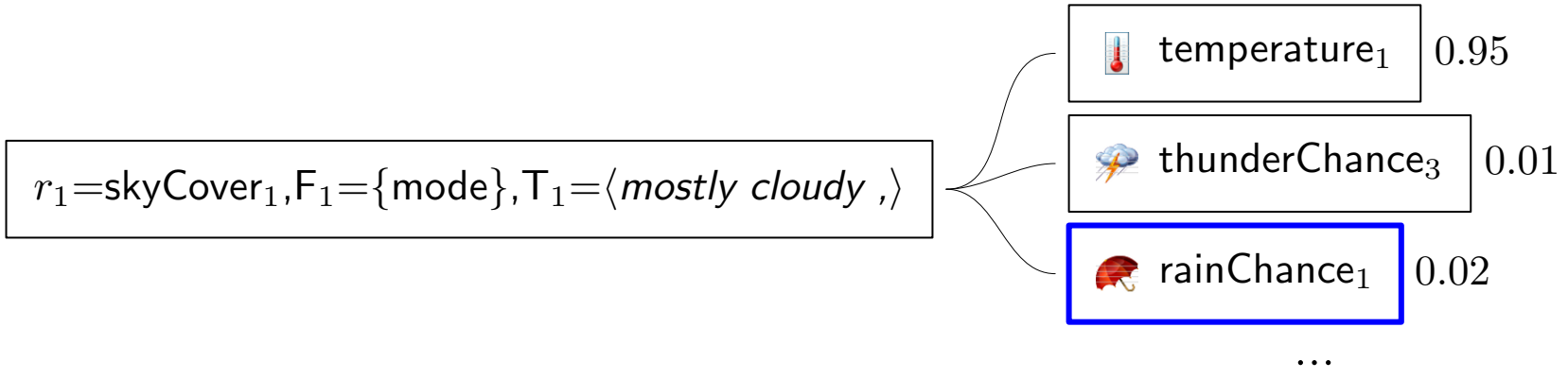
(R3) **Record type already generated**

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

(R4) **Field values**

$[\![r_2.t = \text{rainChance and } r_2.v[\text{time}] = \texttt{5pm-6am}]\!]$

(R5) **Stop under language model**



Features: Record

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

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

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

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

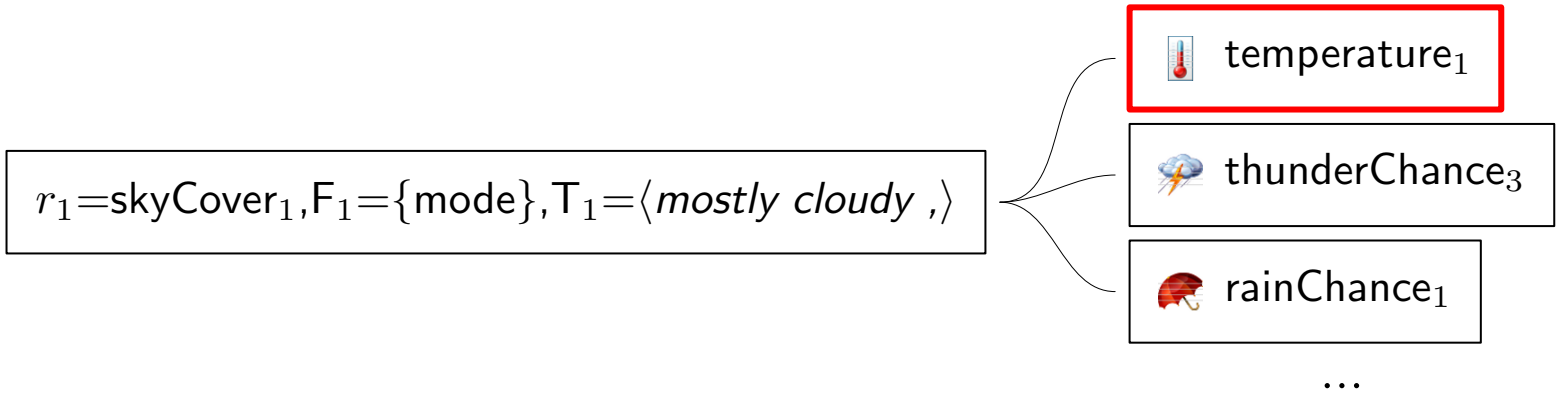
(R3) **Record type already generated**

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

(R4) **Field values**

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

(R5) **Stop under language model**



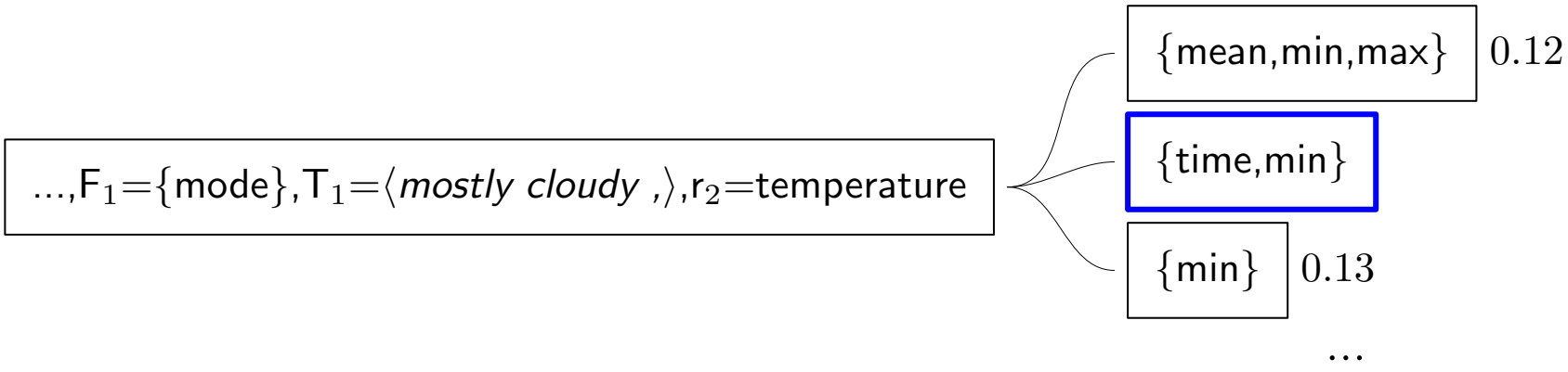
Features: Field Set

Record	$r_1 =$ 🌥️ skyCover$_1$	$r_2 =$ 🌡️ temperature$_1$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$
Template	$W_1 = \langle \textit{mostly cloudy} , \rangle$	

Features: Field Set

(R1) **Field Set**

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



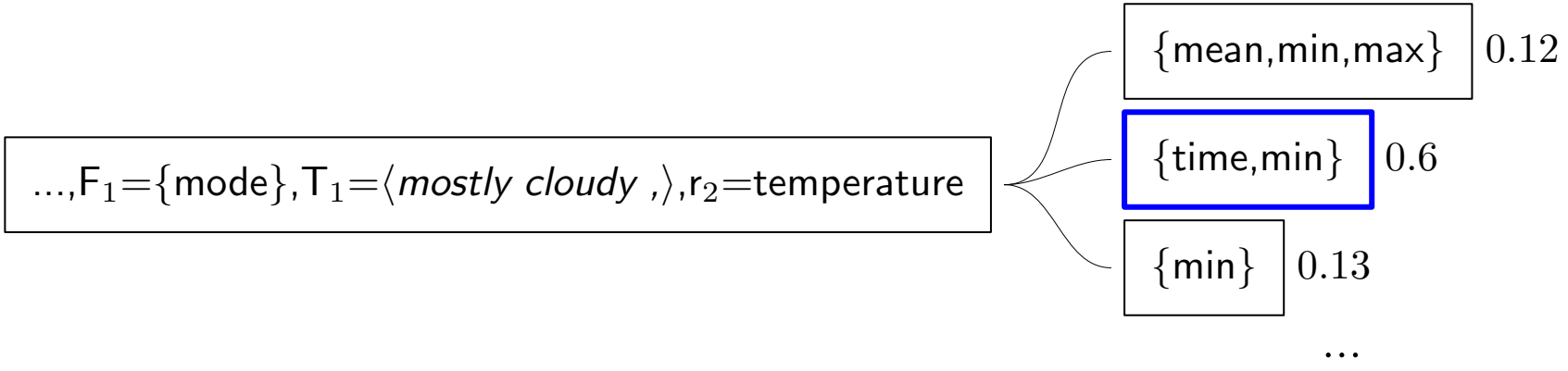
Features: Field Set

(R1) **Field Set**

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

(R2) **Field Set Values**

$[\![F_2 = \{\text{time}, \text{min}\} \text{ and } r_2.v[\text{time}] = \texttt{5pm-6am}]\!]$



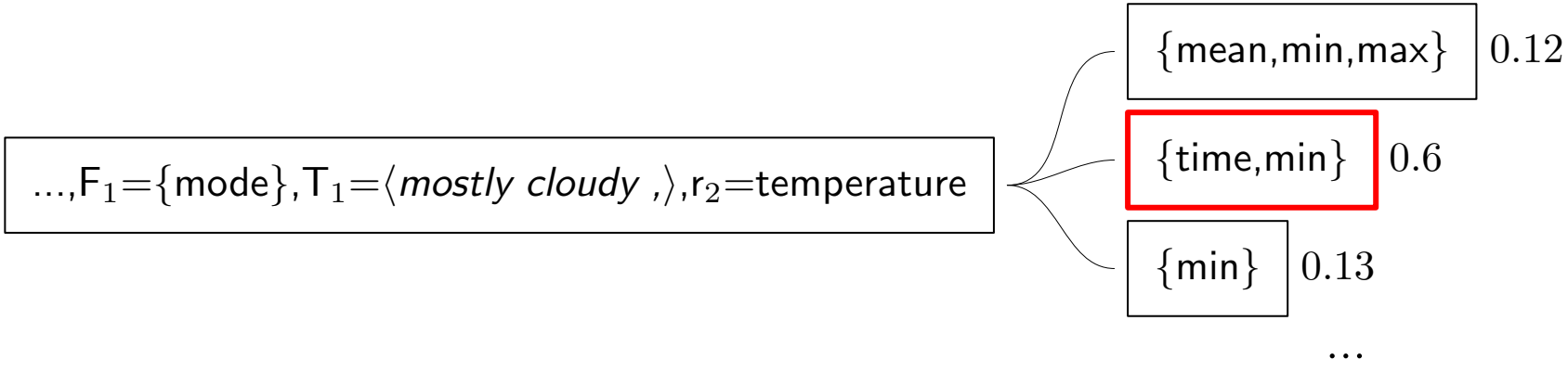
Features: Field Set

(R1) **Field Set**

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

(R2) **Field Set Values**

$[\![F_2 = \{\text{time}, \text{min}\} \text{ and } r_2.v[\text{time}] = \texttt{5pm-6am}]\!]$

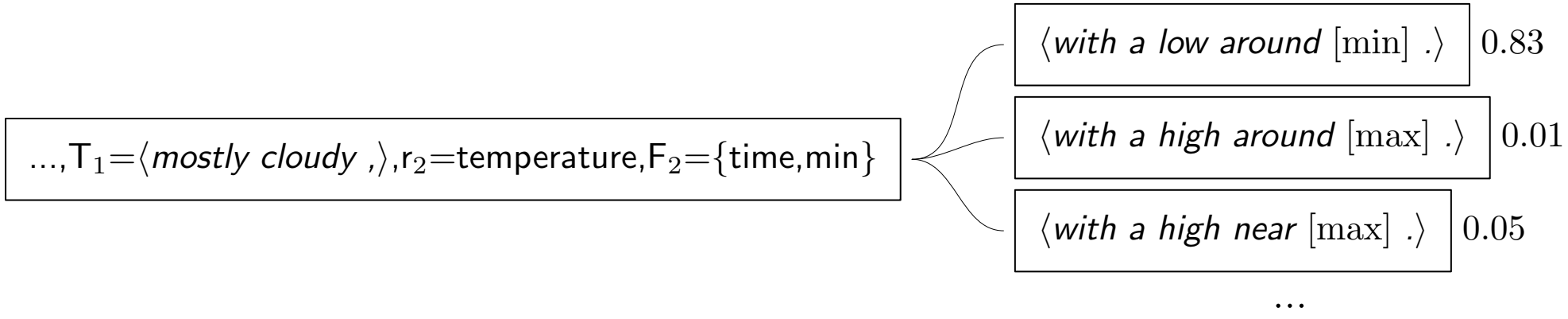


Features: Template

Record	$r_1 =$ skyCover$_1$	$r_2 =$ temperature$_1$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$
Template	$W_1 = \langle \textit{mostly cloudy} \ , \rangle$	$W_2 = \langle \textit{with a low around} \ [\text{min}] \ . \rangle$

Features: Template

Record	$r_1 =$ skyCover$_1$	$r_2 =$ temperature$_1$
Field set	$F_1 = \{\text{mode}\}$	$F_2 = \{\text{time}, \text{min}\}$
Template	$W_1 = \langle \textit{mostly cloudy} , \rangle$	$W_2 = \langle \textit{with a low around } [\text{min}] \ . \rangle$



Template Extraction

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

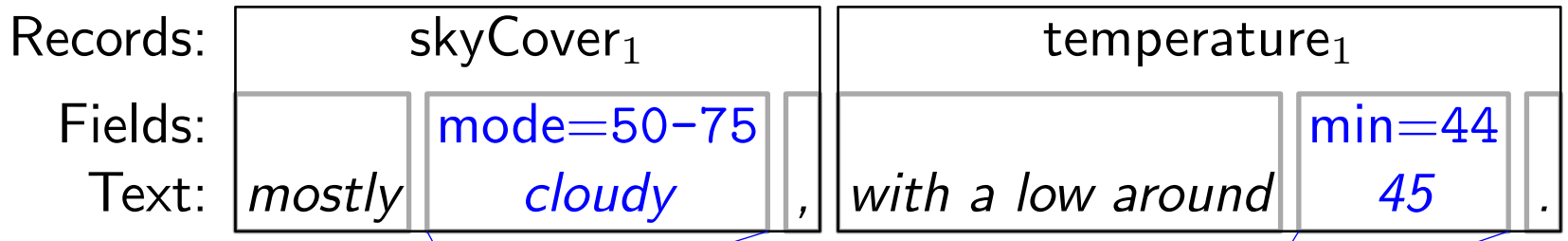
Template Extraction

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

Records:	skyCover$_1$			temperature$_1$		
Fields:		mode=50-75			min=44	
Text:	*mostly*	*cloudy*	,	*with a low around*	*45*	.

Template Extraction

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



$<$ *mostly* [mode] , $>$ $<$ *with a low around* [min] . $>$

Ideal: Extract alignments directly from the alignment

Template Extraction

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

Records:	skyCover$_1$	temperature$_1$			
Fields:	mode=50-75		time=17-30	min=44	mean=49
Text:	*mostly cloudy ,*	*with a*	*low around*	*45*	*.*

Ideal: Extract alignments directly from the alignment

Challenge: Alignments are often noisy

Template Extraction

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

Records:	skyCover$_1$	temperature$_1$			
Fields:	mode=50-75		time=17-30	min=44	mean=49
Text:	*mostly cloudy ,*	*with a*	*low around*	*45*	*.*

$\textsc{Coarse}$ $<$ [mode] $>$ $<$ *with a* [time] [min] [mean] $>$

$\textsc{Base}$ $<$ *mostly cloudy ,* $>$ $<$ *with a low around* [min] *.* $>$

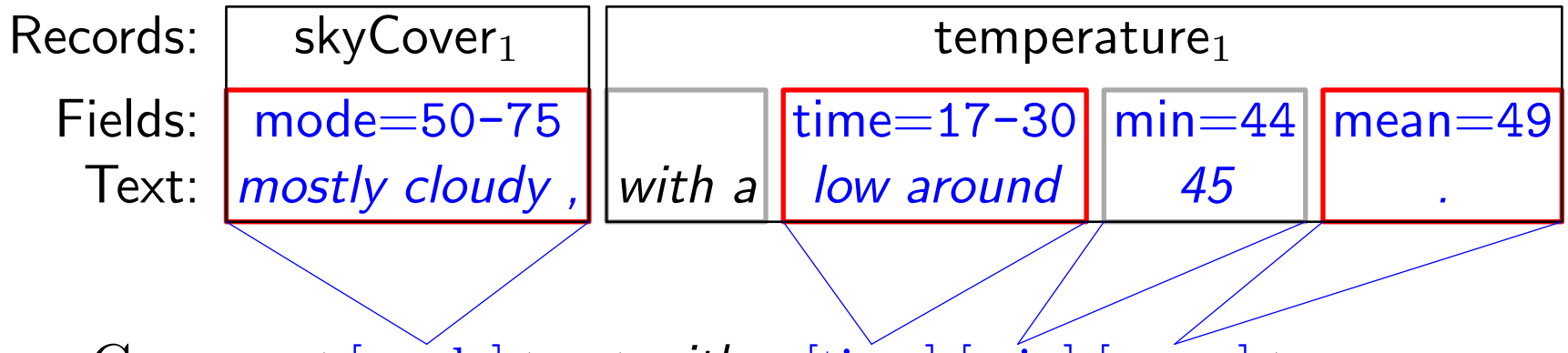
Ideal: Extract alignments directly from the alignment

Challenge: Alignments are often noisy

Approach: Extract templates of different granularities:

Template Extraction

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



COARSE< [mode] > < *with a* [time] [min] [mean] >

BASE < *mostly cloudy ,* > < *with a low around* [min] *.* >

Ideal: Extract alignments directly from the alignment

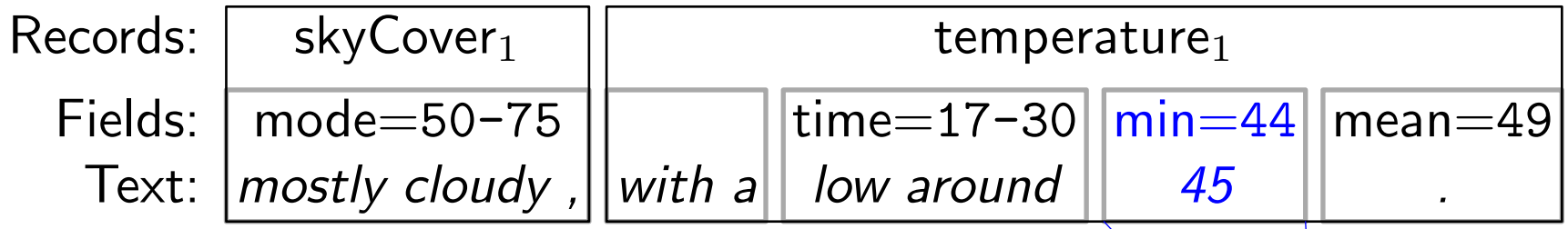
Challenge: Alignments are often noisy

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)



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

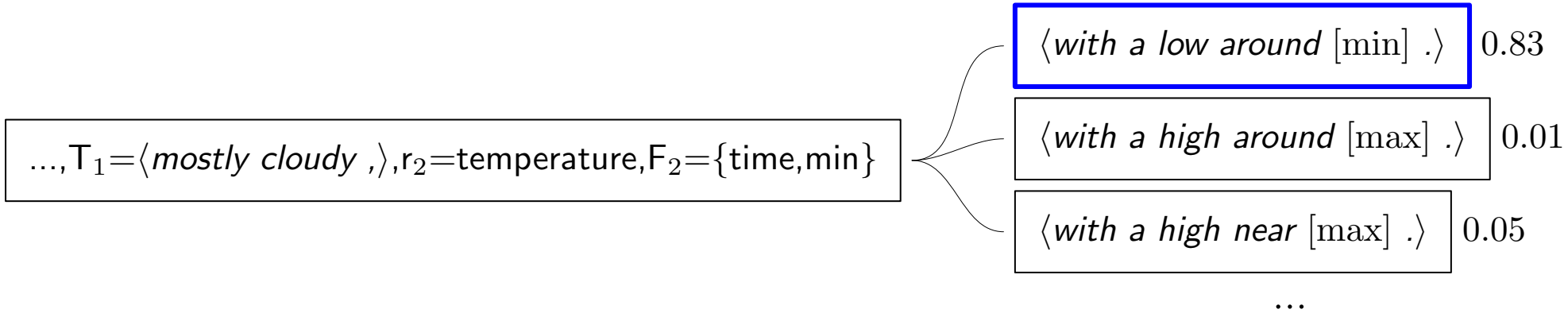
BASE: created by abstracting field matching regexp (e.g. *[0-9]+*)

Features: Template

Note: Features defined over both template granularities

(R1) **Template**

$[\![\textsc{Base}(T_2) = \langle \textit{with a low around } [\min]\rangle]\!]$



Features: Template

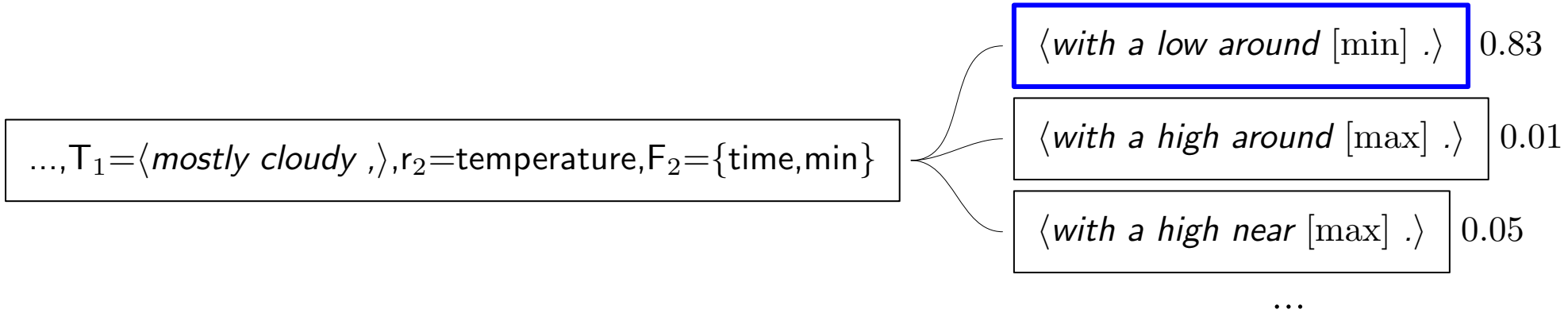
Note: Features defined over both template granularities

(R1) **Template**

$[\![\textsc{Base}(T_2) = \langle \textit{with a low around } [\min]\rangle]\!]$

(R2) **Field Values**

$[\![\textsc{Base}(T_2) = \langle \textit{with a low around } [\min]\rangle \text{ and } r_2.v[\text{time}] = \texttt{5pm-6am}]\!]$



Features: Template

Note: Features defined over both template granularities

(R1) **Template**

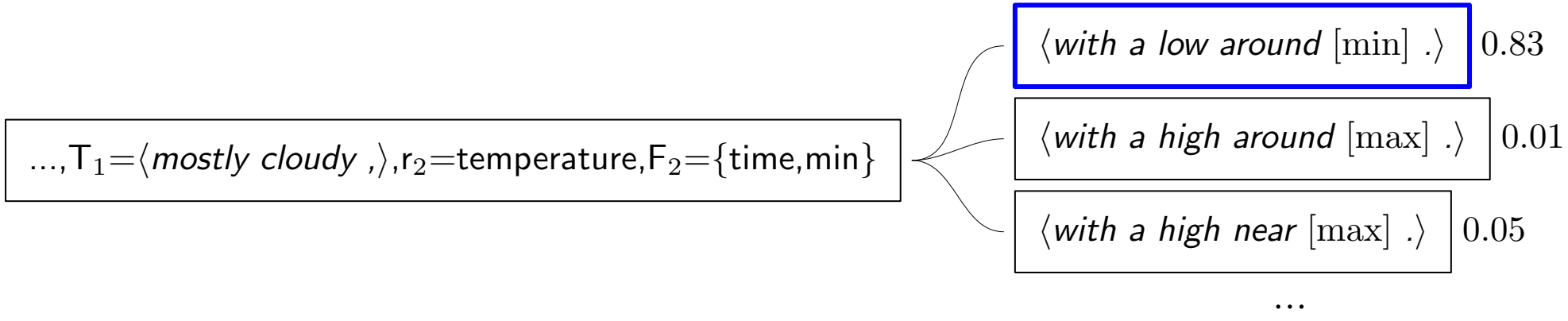
$[\![\textsc{Base}(T_2) = \langle \textit{with a low around } [\min]\rangle]\!]$

(R2) **Field Values**

$[\![\textsc{Base}(T_2) = \langle \textit{with a low around } [\min]\rangle \text{ and } r_2.v[\texttt{time}] = \texttt{5pm-6am}]\!]$

(R3) **First word in template under LM**

$\log p_{\text{LM}}(\textit{with} \mid \textit{cloudy ,})$



Features: Template

Note: Features defined over both template granularities

(R1) **Template**

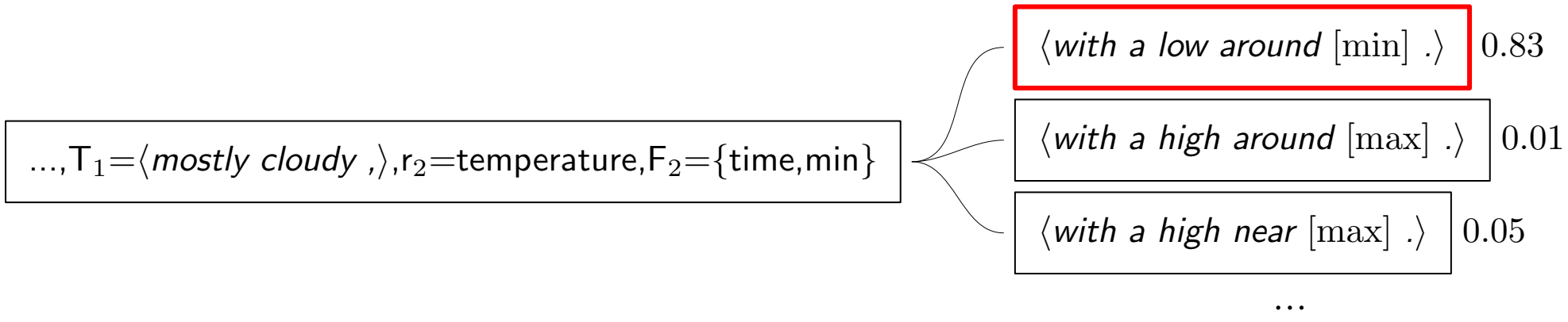
$[\![\textsc{Base}(T_2) = \langle \textit{with a low around } [\min]\rangle]\!]$

(R2) **Field Values**

$[\![\textsc{Base}(T_2) = \langle \textit{with a low around } [\min]\rangle \text{ and } r_2.v[\text{time}] = \texttt{5pm-6am}]\!]$

(R3) **First word in template under LM**

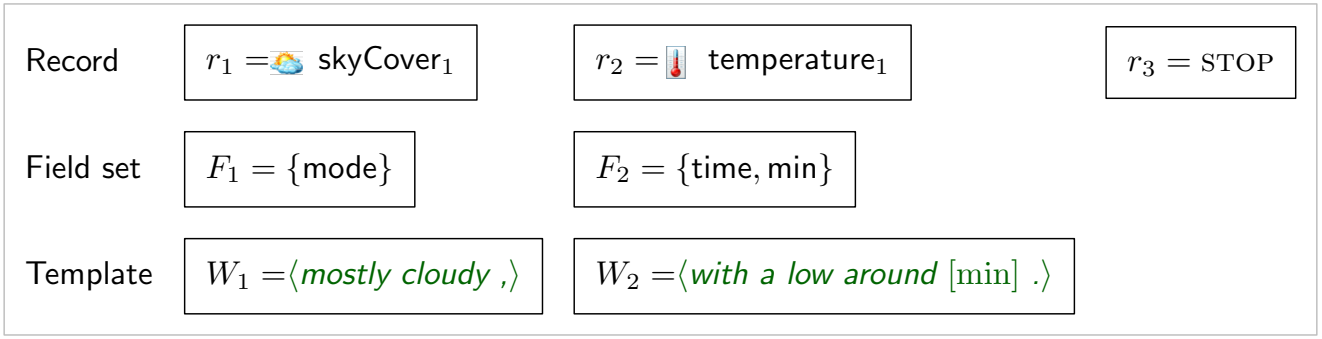
$\log p_{\text{LM}}(\textit{with} \mid \textit{cloudy ,})$



Training: Log-Linear Model

Generation treated as a sequence of local decisions

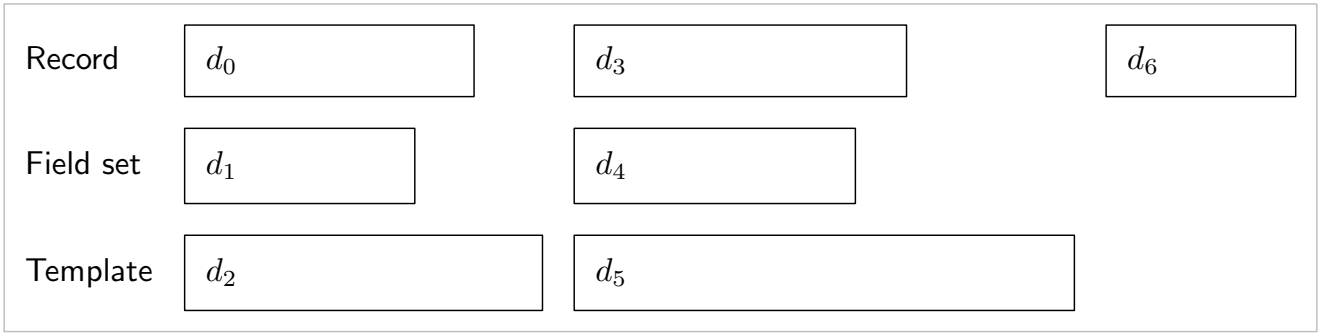
Similar in spirit to Ratnaparkhi (1996)



Training: Log-Linear Model

Generation treated as a sequence of local decisions

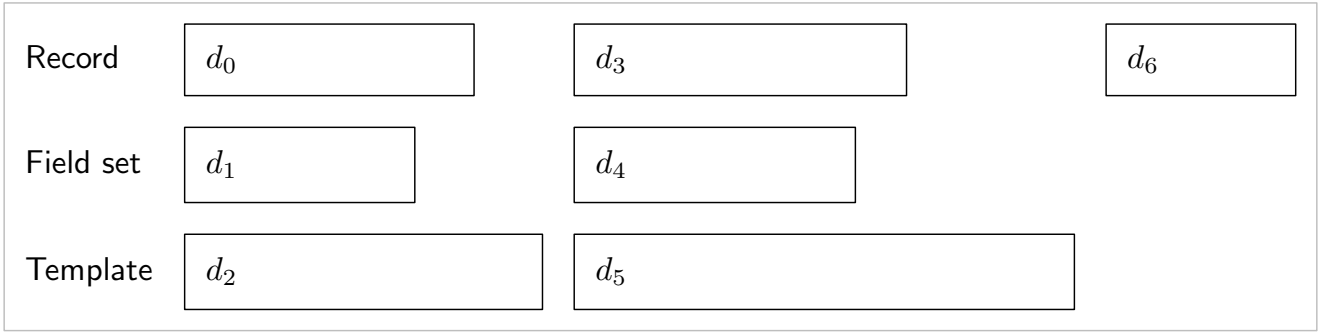
Similar in spirit to Ratnaparkhi (1996)



Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)



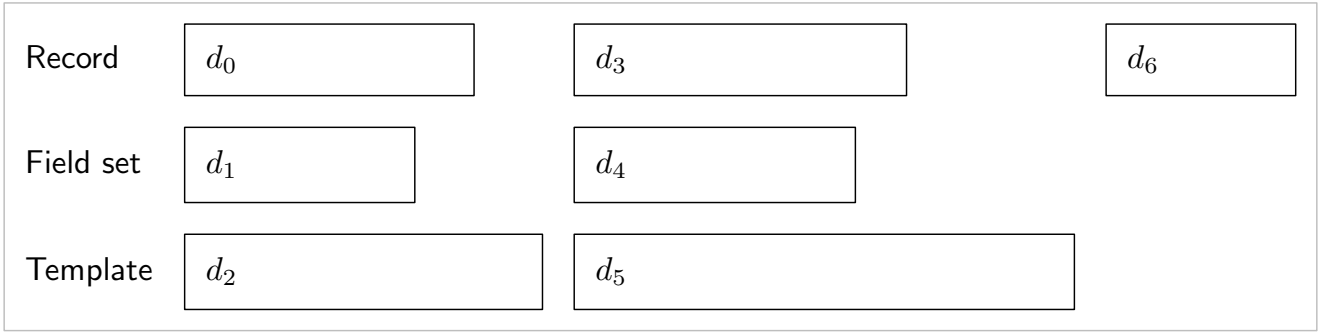
Log-linear model: Each decision trained using a log-linear model

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

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

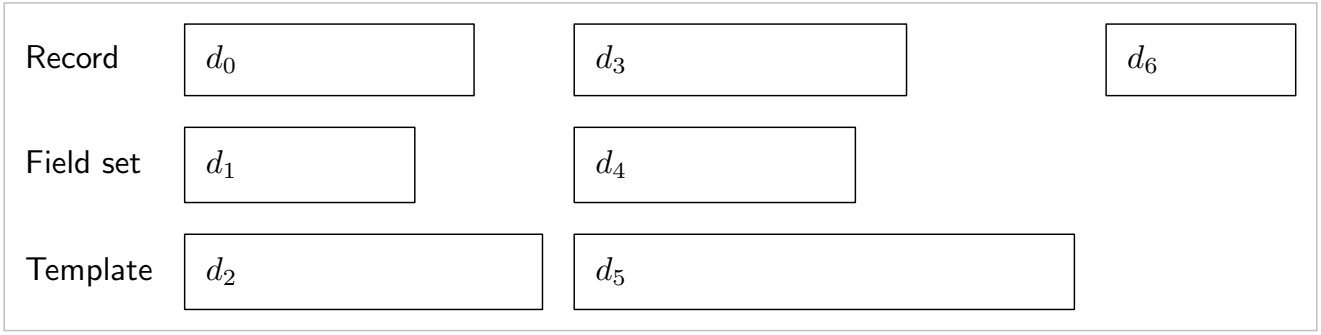
$$p(d_j \mid d_{j-1}, d_{j-2}, \ldots, \mathbf{s}; \theta) \propto \exp\{\phi_j(d_j, d_{j-1}, d_{j-2}, \ldots, \mathbf{s})^\top \theta\}$$

Log-likelihood maximized using L-BFGS

Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)



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

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

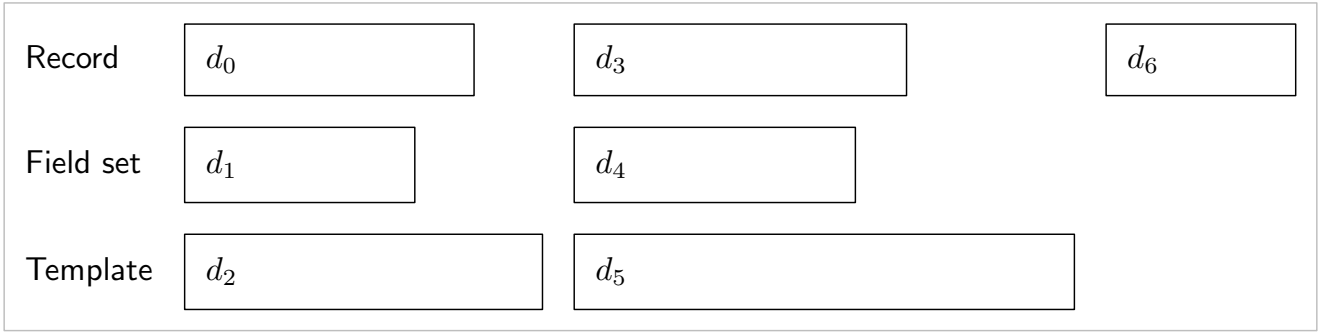
Log-likelihood maximized using L-BFGS

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

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

Log-likelihood maximized using L-BFGS

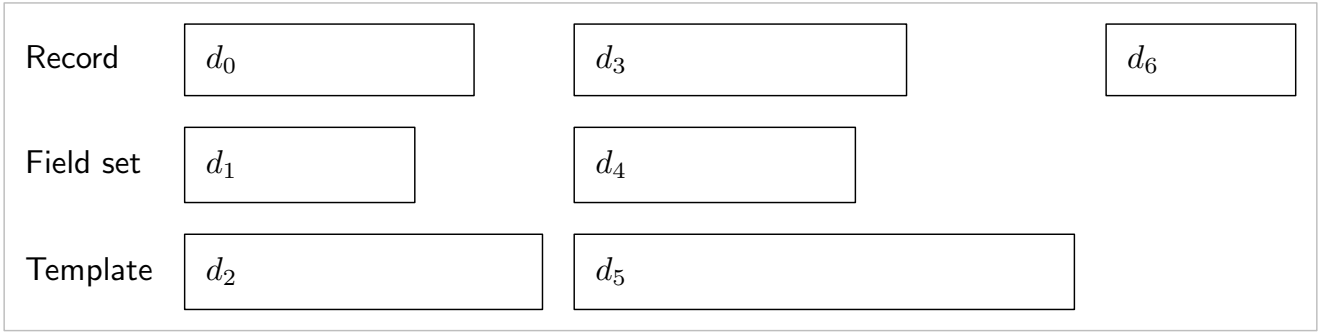
No dynamic program since decisions have long-range dependencies

Training problem is fully supervised

Training: Log-Linear Model

Generation treated as a sequence of local decisions

Similar in spirit to Ratnaparkhi (1996)



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

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

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

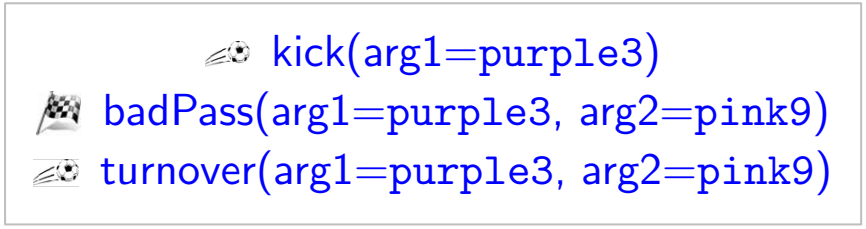
ROBOCUP: RoboCup sportscasting data

Evaluation: Domains

ROBOCUP: RoboCup sportscasting data

Content Selection: Chose one record

s:

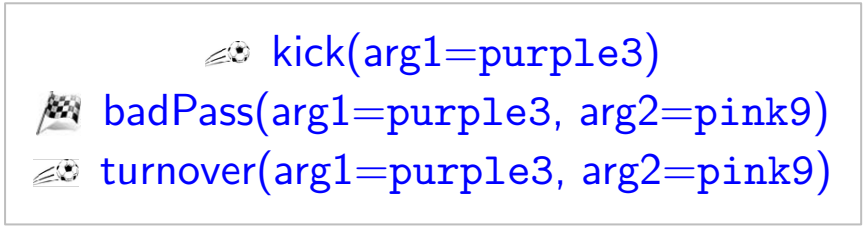


Evaluation: Domains

ROBOCUP: RoboCup sportscasting data

Content Selection: Chose one record

s:



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

s:

wind10m(time=6am,dir=SW,min=16,max=20,gust min=0,gust max=0)
wind10m(time=9pm,dir=SSW,min=28,max=32,gust min=40,gust max=0)
wind10m(time=12am,dir=-,min=24,max=28,gust min=36,gust max=0)

Evaluation: Domains

ROBOCUP: RoboCup sportscasting data

SUMTIME: Weather forecasts for oil rigs

Content Selection: No content selection

s:

wind10m(time=6am,dir=SW,min=16,max=20,gust min=0,gust max=0)
wind10m(time=9pm,dir=SSW,min=28,max=32,gust min=40,gust max=0)
wind10m(time=12am,dir=-,min=24,max=28,gust min=36,gust max=0)

Surface Realization:

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

Evaluation: Domains

ROBOCUP: RoboCup sportscasting data

SUMTIME: Weather forecasts for oil rigs

WEATHERGOV: Weather forecasts from `weather.gov`

Evaluation: Domains

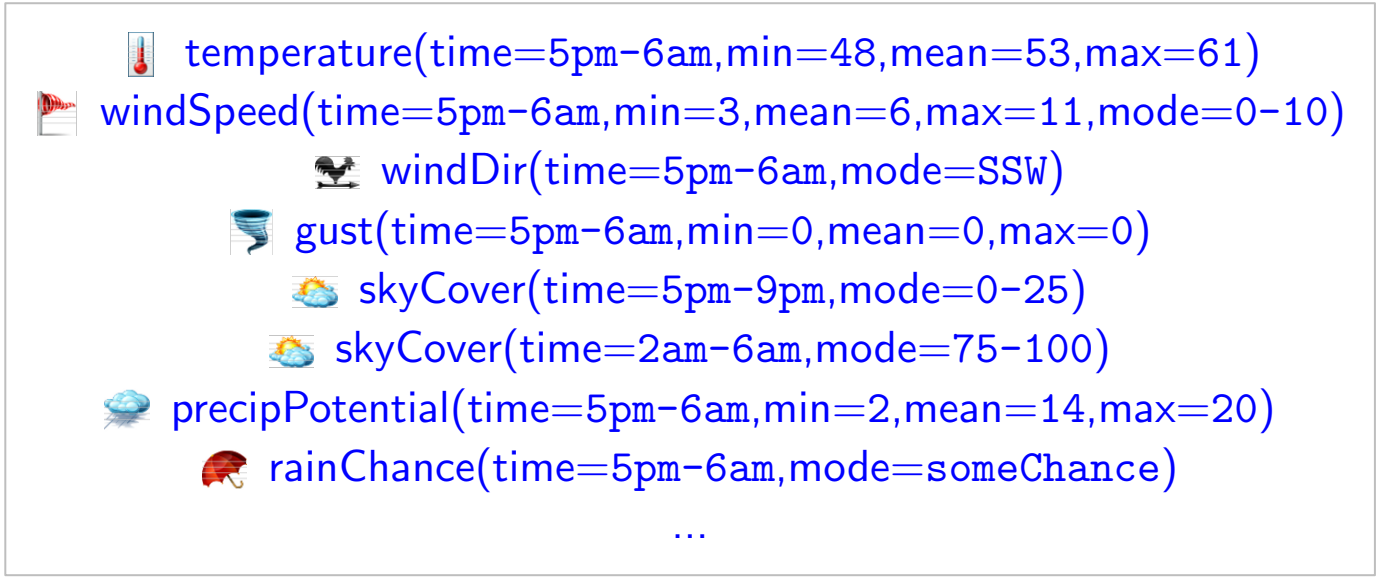
ROBOCUP: RoboCup sportscasting data

SUMTIME: Weather forecasts for oil rigs

WEATHERGOV: Weather forecasts from `weather.gov`

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

s:



Evaluation: Domains

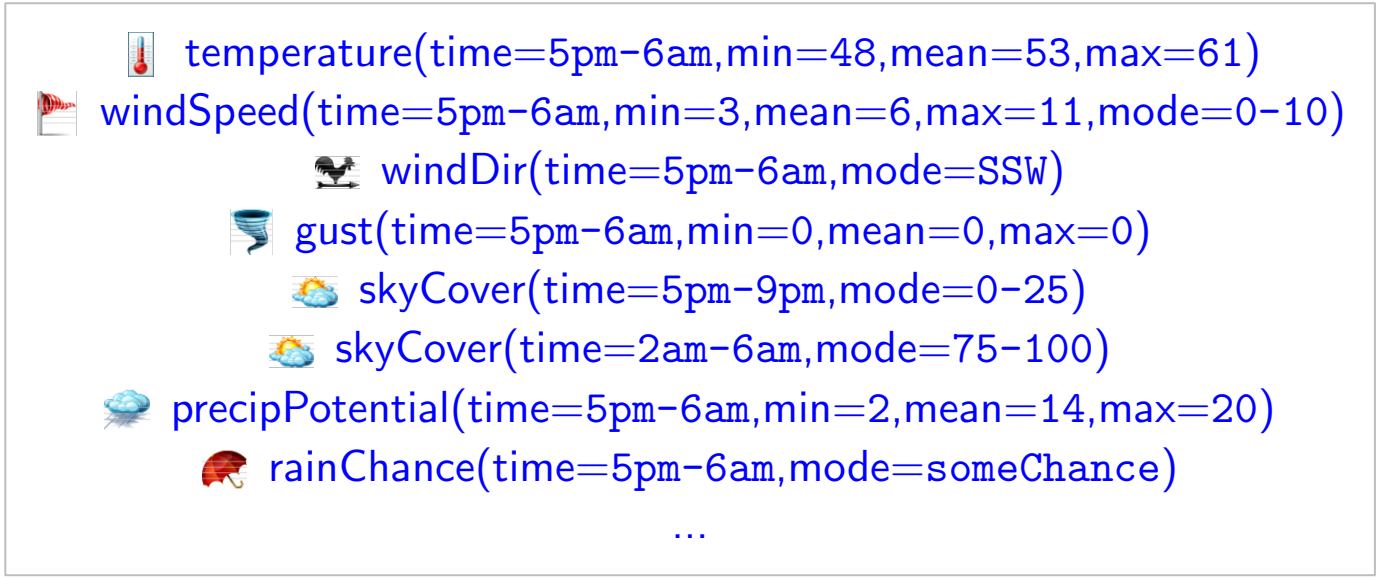
ROBOCUP: RoboCup sportscasting data

SUMTIME: Weather forecasts for oil rigs

WEATHERGOV: Weather forecasts from `weather.gov`

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

s:



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

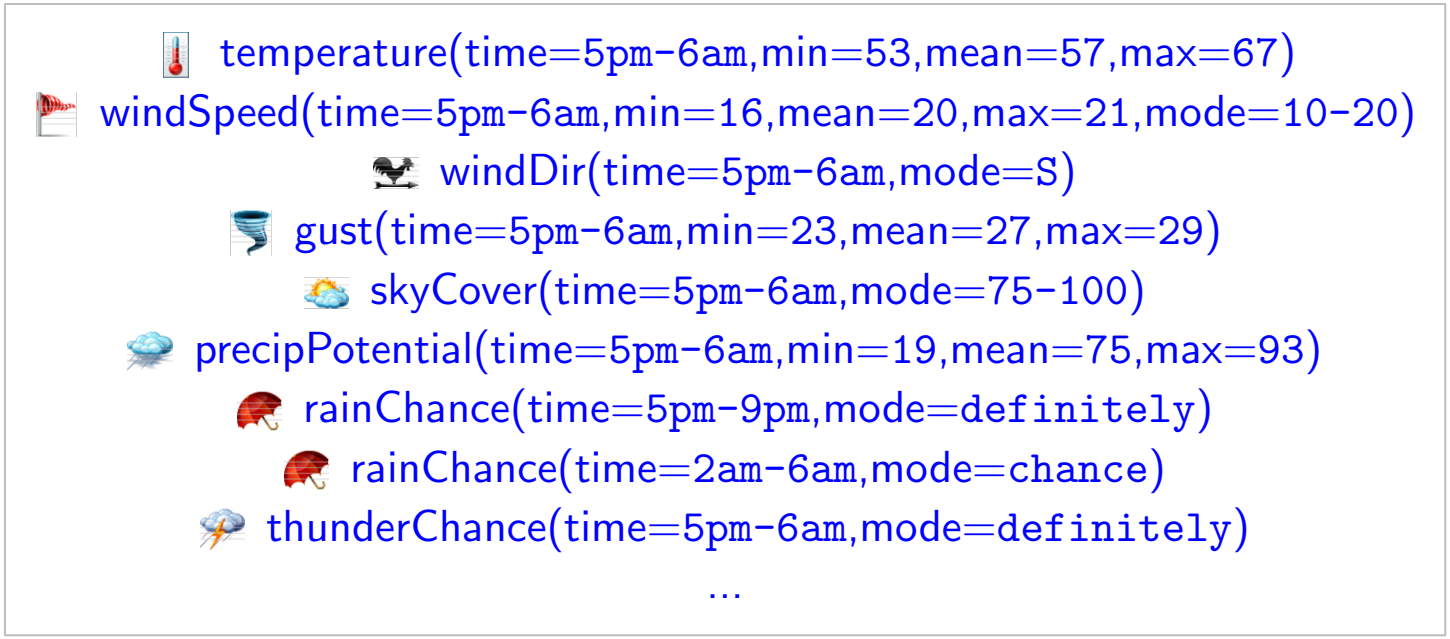
100 random shuffled scenarios shown to 10 Turkers

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

Score	**English Fluency**	**Semantic Correctness**
5	Flawless	Perfect
4	Good	Near Perfect
3	Non-native	Minor Errors
2	Disfluent	Major Errors
1	Gibberish	Completely Wrong

Evaluation: Metrics

s:



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

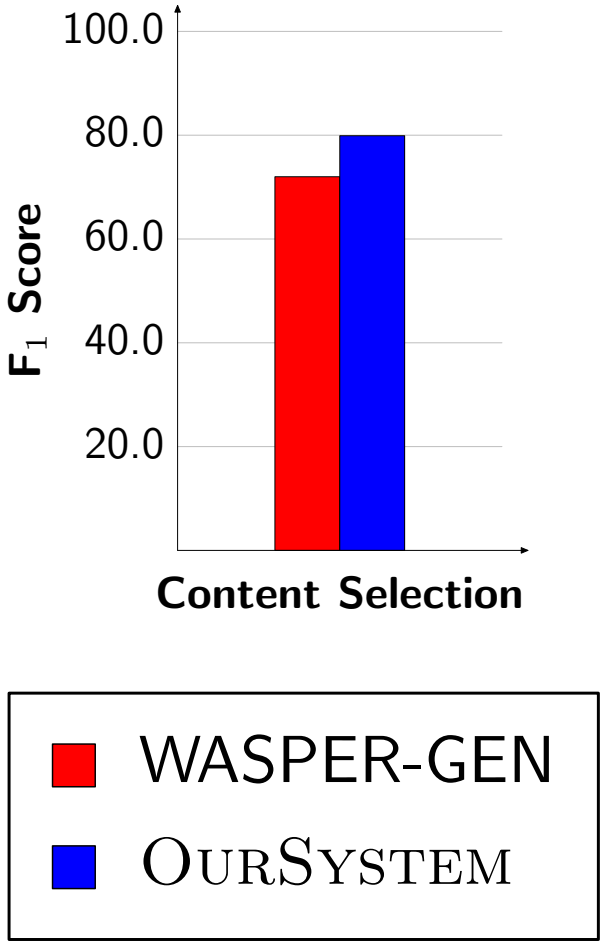
OURSYSTEM-CUSTOM: 8 domain-dependent features for SUMTIME

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

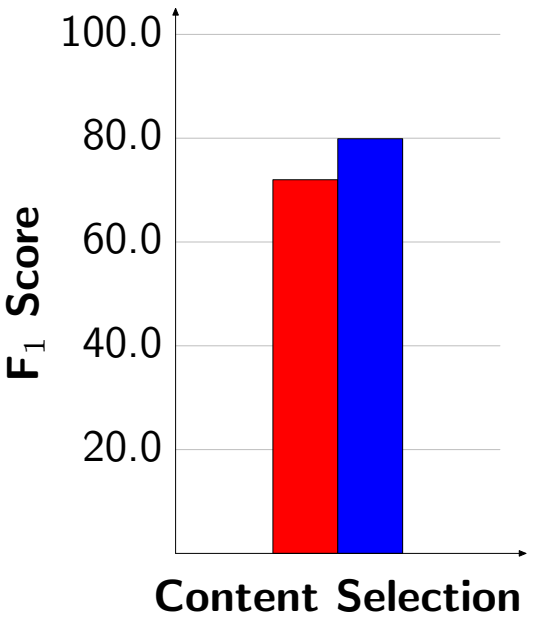
SUMTIME-Hybrid: SUMTIME hand crafted system of Reiter, et al. (2005)

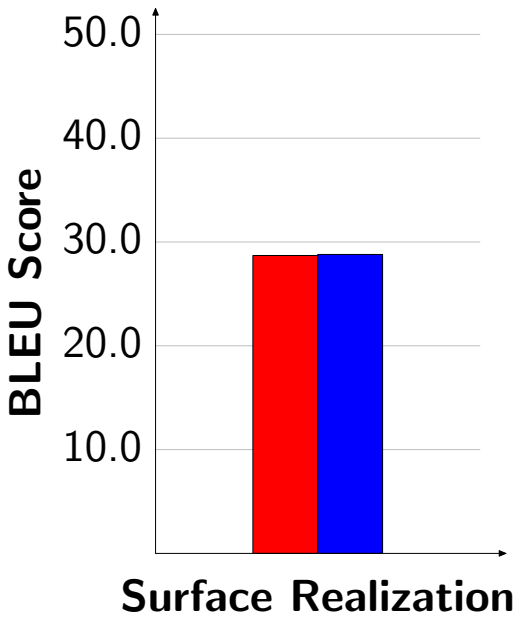
*p*CRU-greedy: SUMTIME state of the art by Belz (2008)

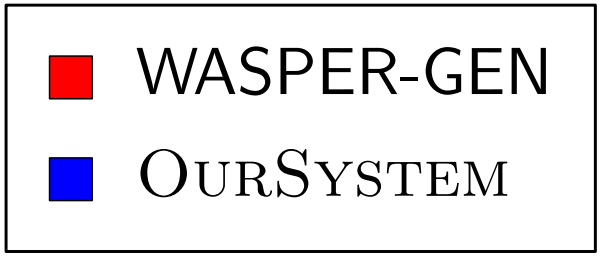
Evaluation: ROBOCUP Results



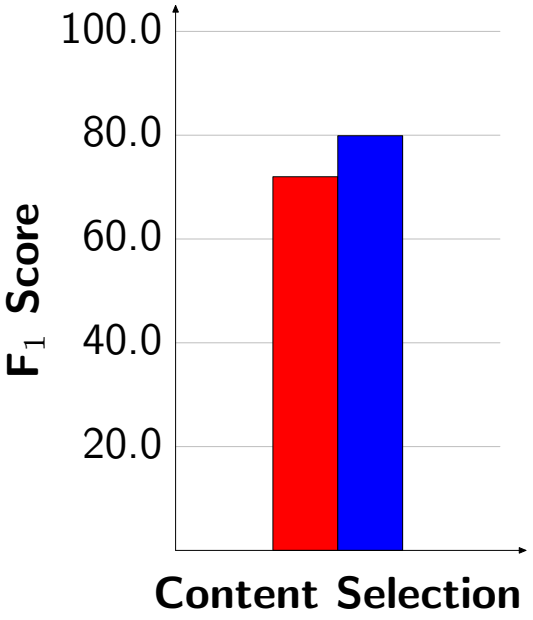
Evaluation: ROBOCUP Results

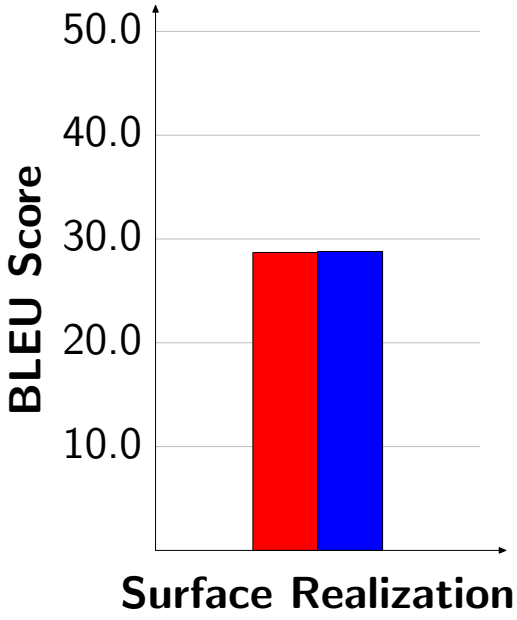


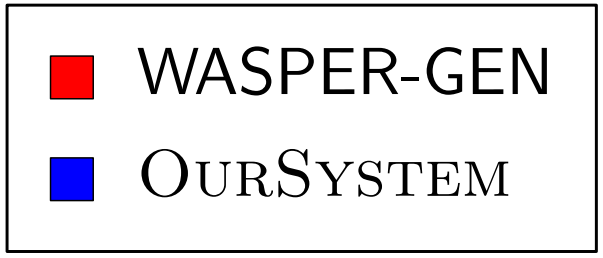


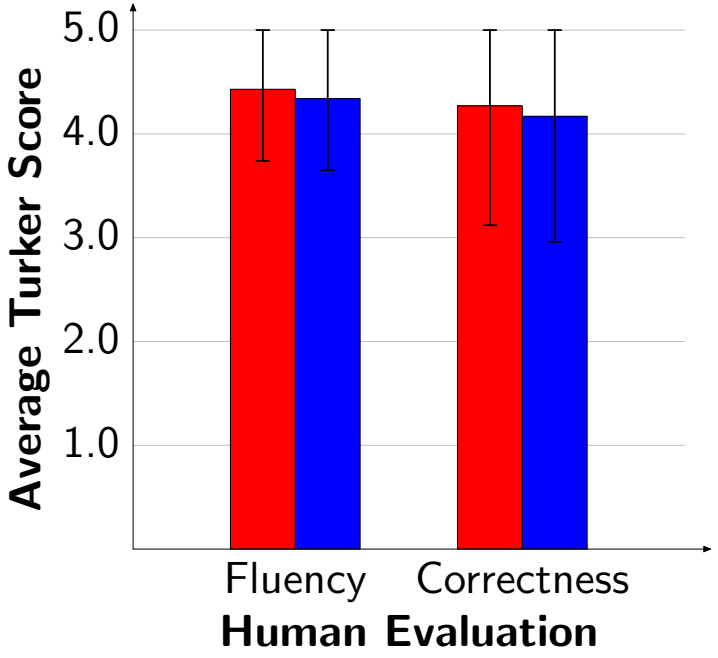


Evaluation: ROBOCUP Results

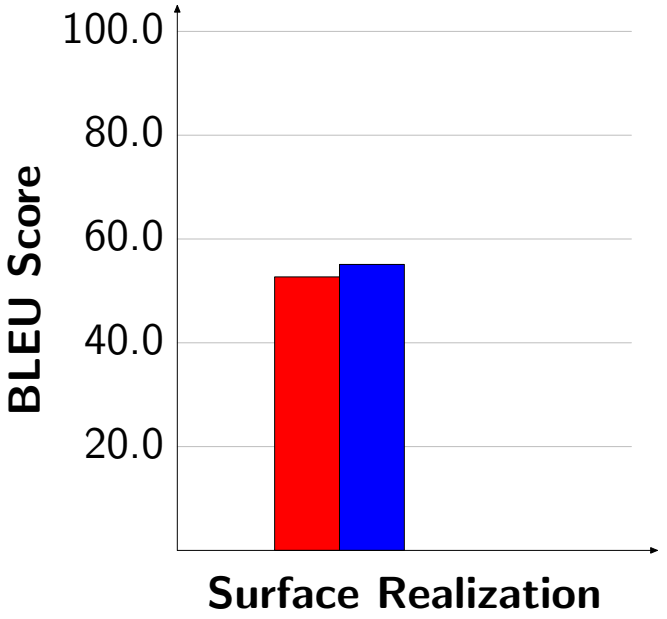








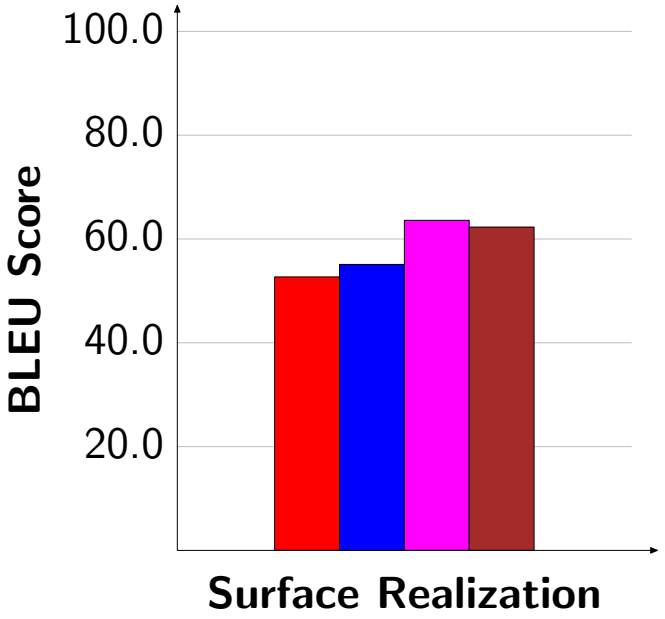
Evaluation: SumTime Results



SumTime-Hybrid

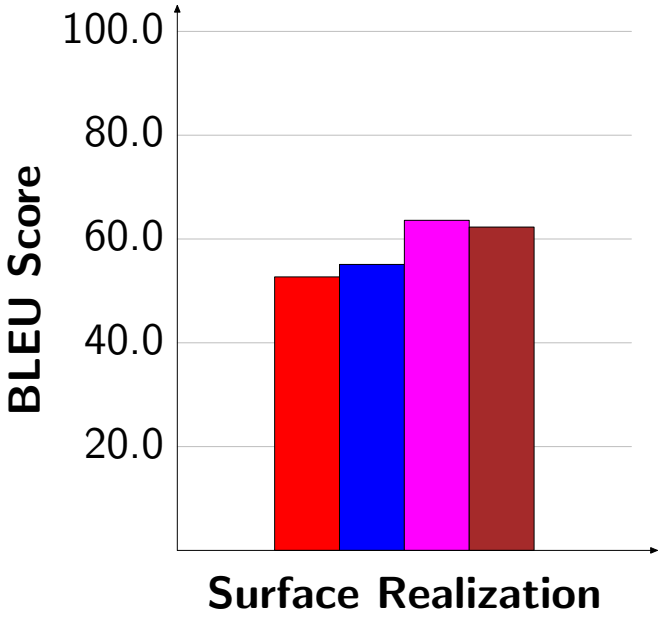
OurSystem

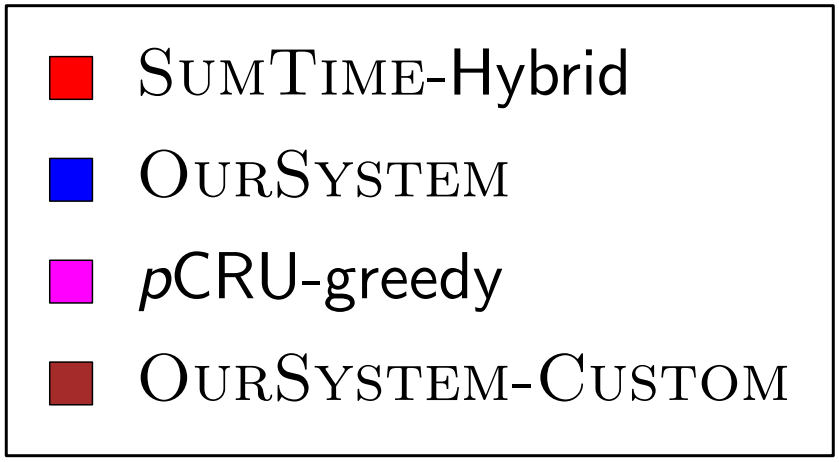
Evaluation: SumTime Results

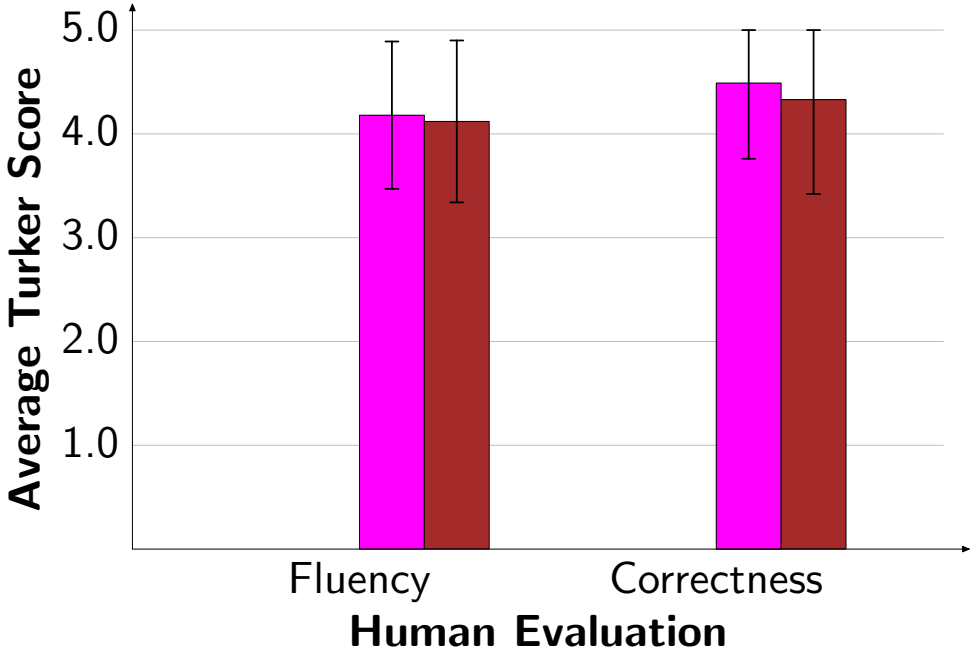


- SumTime-Hybrid
- OurSystem
- *p*CRU-greedy
- OurSystem-Custom

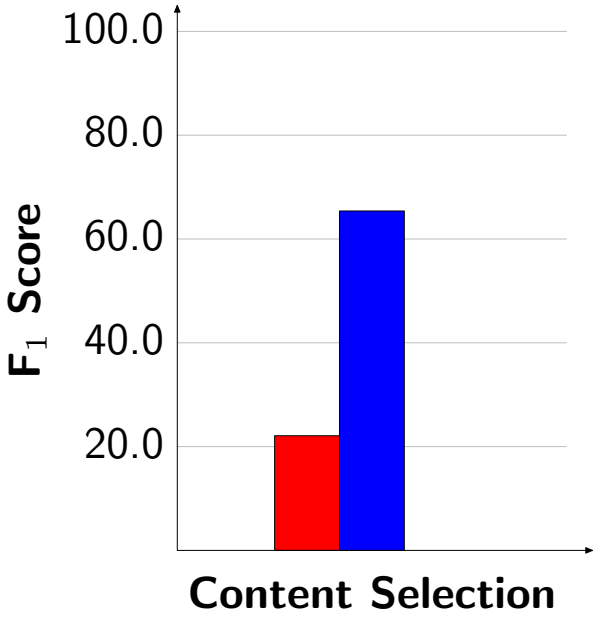
Evaluation: SumTime Results

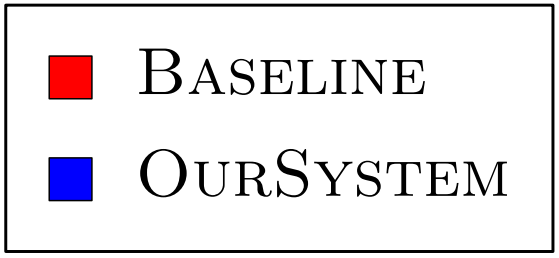




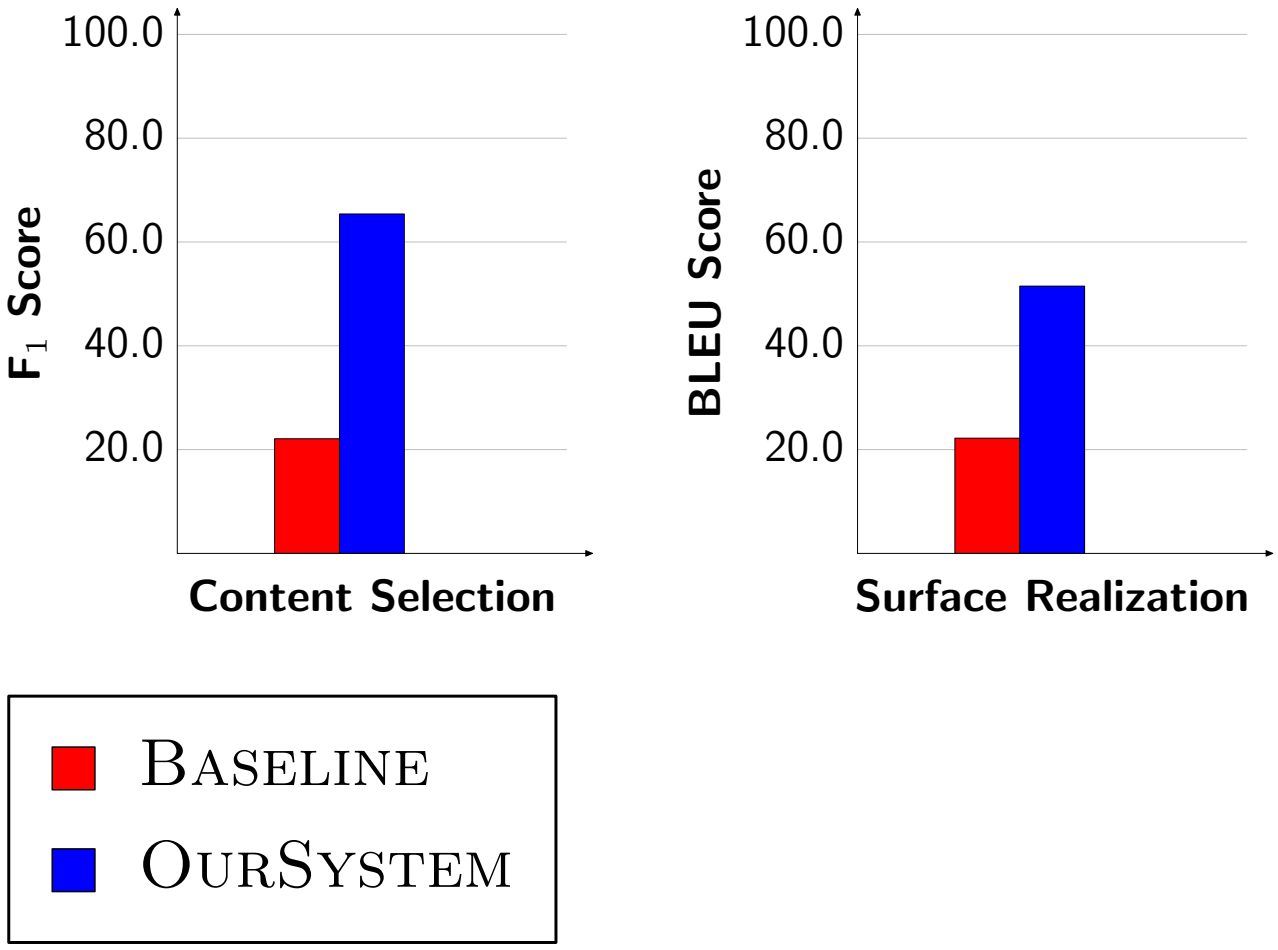


Evaluation: WeatherGov Results

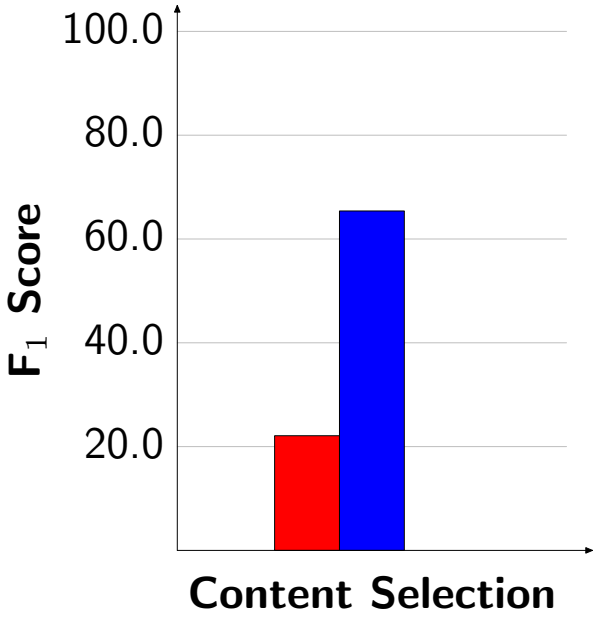


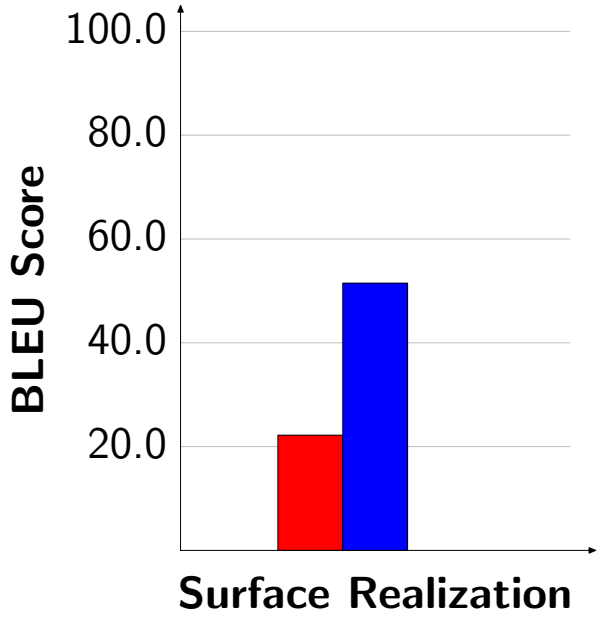


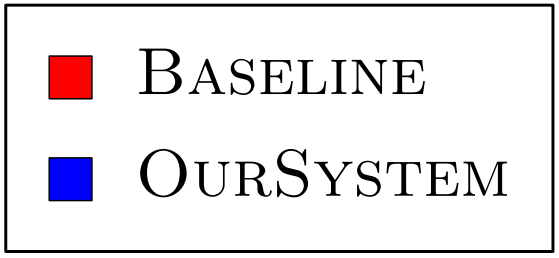
Evaluation: WeatherGov Results

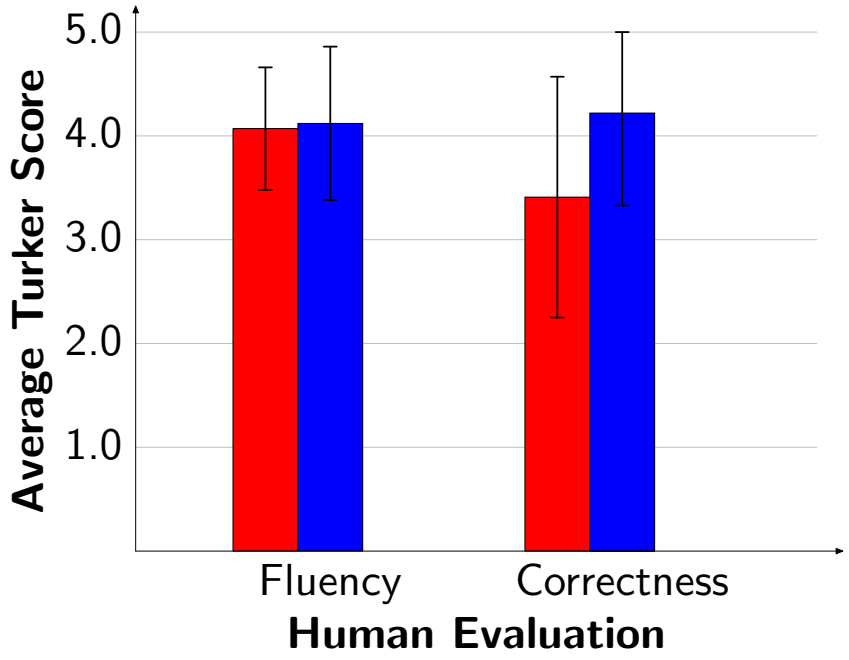


Evaluation: WeatherGov Results

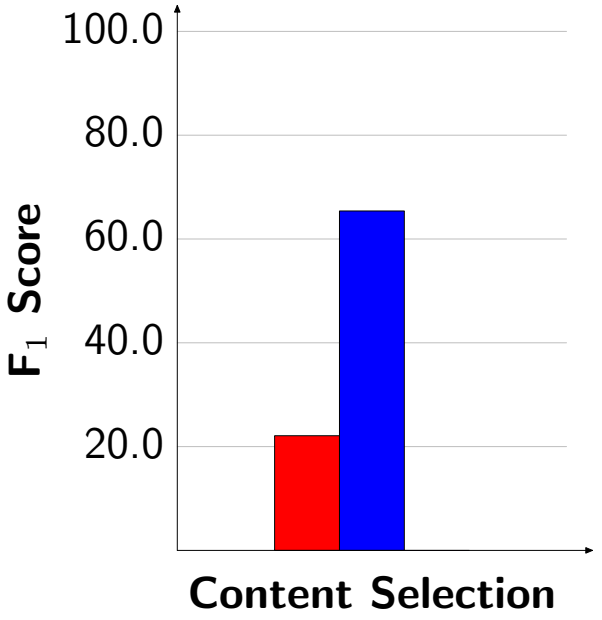


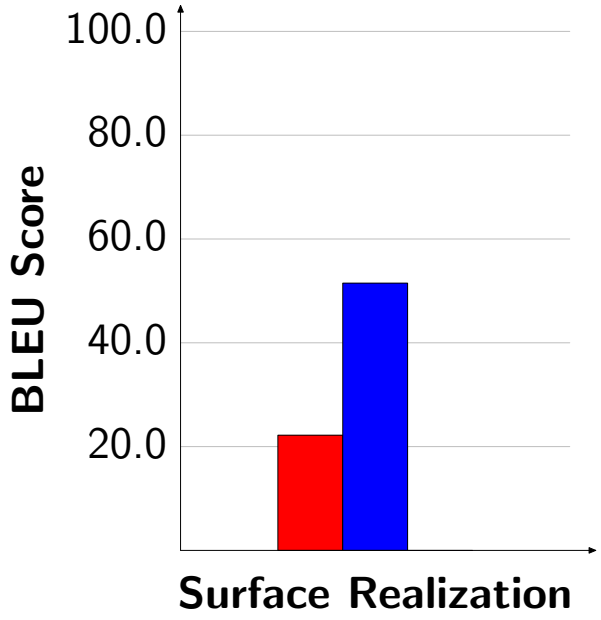


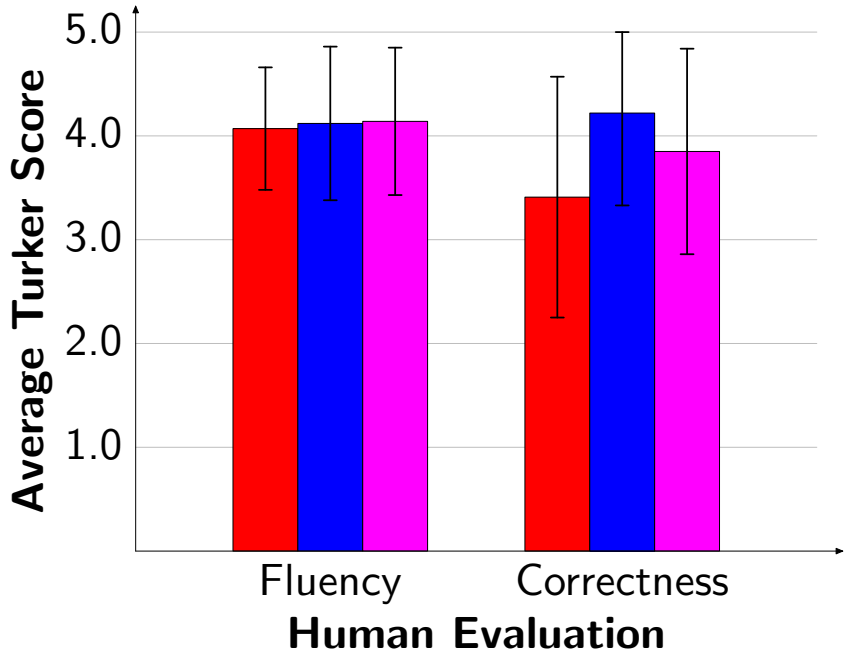




Evaluation: WeatherGov Results







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!