Unsupervised Concept-to-text Generation with Hypergraphs

Ioannis Konstas, Mirella Lapata

Institute for Language, Cognition and Computation
University of Edinburgh

NAACL 2012, Montréal

Introduction

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

Introduction

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

Temperature			
Time	**Min**	**Mean**	**Max**
06:00-21:00	9	15	21

Cloud Sky Cover	
Time	**Percent (%)**
06:00-09:00	25-50
09:00-12:00	50-75

Wind Speed			
Time	**Min**	**Mean**	**Max**
06:00-21:00	15	20	30

Wind Direction	
Time	**Mode**
06:00-21:00	S

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

Introduction

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

Flight

direction	**from**	**to**
oneway	edinburgh	montreal

Day

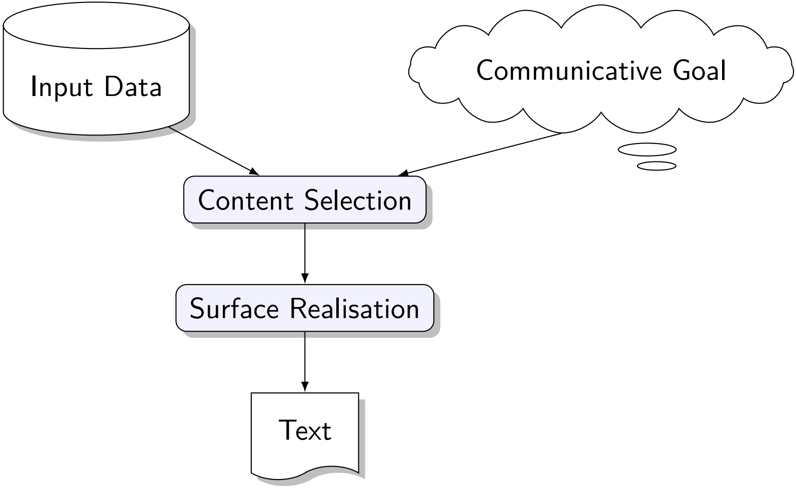
day	**dep/ar/ret**
saturday	departure

Search

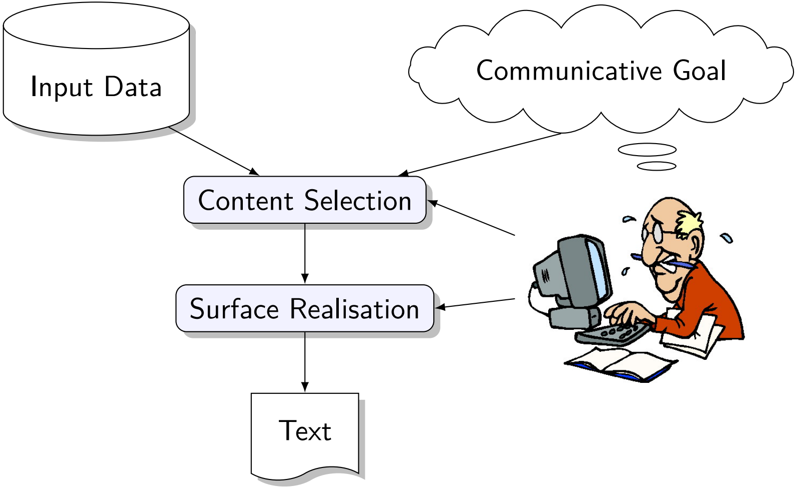
of	**type**	**what**
fare	argmin	flight

Show me the cheapest one way flights
from Edinburgh to Montreal leaving on Saturday

Traditional NLG Pipeline



Traditional NLG Pipeline



Our Approach

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

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

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

Our Approach

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

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

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

Our Approach

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

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

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

Our Approach

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

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

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

Our Approach

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

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

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

Our Approach

Temperature			
Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover	
Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

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

Wind Speed			
Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction	
Time	Mode
06:00-21:00	S

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

Our Approach

Temperature			
Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover	
Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

Wind Speed			
Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction	
Time	Mode
06:00-21:00	S

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

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

Key idea: recast generation as a parsing problem

1. Describe the structure of the input with a PCFG
2. Convert PCFG to a hypergraph
3. Goal: Find the most fluent and grammatical derivation

Related Work

Angeli et al., 2010

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

Related Work

Angeli et al., 2010

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

Our Approach

- Unsupervised generative model
- Joint content selection and surface realisation, breaks the traditional NLG pipeline
- Domain independent, trainable end-to-end system

Input

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

Cloud Sky Cover

Time	**Percent (%)**
06:00-09:00	25-50
09:00-12:00	50-75

mostly cloudy,

Input

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

Cloud Sky Cover

Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

mostly cloudy,

Input

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

Cloud Sky Cover	
Time	**Percent (%)**
06:00-09:00	25-50
09:00-12:00	50-75

mostly cloudy,

Input

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

Cloud Sky Cover	
Time	**Percent (%)**
06:00-09:00	25-50
09:00-12:00	50-75

mostly cloudy,

Input

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

Cloud Sky Cover

Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

mostly cloudy,

Grammar Definition

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

Grammar Definition

$\mathrm{R}(skyCover_1.t) \rightarrow \mathrm{FS}(temperature_1, start)\mathrm{R}(temperature_1.t)$

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

Grammar Definition

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

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

Grammar Definition

$\mathrm{F}(gust_1, mean) \rightarrow \mathrm{W}(gust_1, mean)\mathrm{F}(gust_1, mean)$

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

Grammar Definition

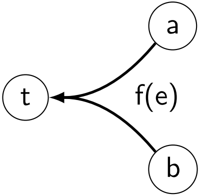
$\mathrm{W}(windDir_1, mode) \rightarrow \text{southeast}$

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

Hypergraphs

Definition

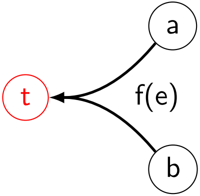
An ordered hypergraph H is a tuple $\langle N, E, t, \mathbf{R} \rangle$, where N is a finite set of nodes, E is a finite set of hyperarcs, $t \in N$ is a target node and $\mathbf{R}$ is the set of weights. Each hyperarc $e \in E$ is a triple $e = \langle T(e), h(e), f(e) \rangle$, where $h(e) \in N$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



Hypergraphs

Definition

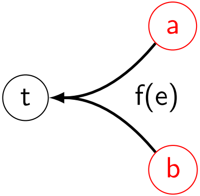
An ordered hypergraph H is a tuple $\langle N, E, t, \mathbf{R} \rangle$, where N is a finite set of nodes, E is a finite set of hyperarcs, $t \in N$ is a target node and $\mathbf{R}$ is the set of weights. Each hyperarc $e \in E$ is a triple $e = \langle T(e), h(e), f(e) \rangle$, where $h(e) \in N$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



Hypergraphs

Definition

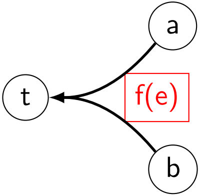
An ordered hypergraph H is a tuple $\langle N, E, t, \mathbf{R} \rangle$, where N is a finite set of nodes, E is a finite set of hyperarcs, $t \in N$ is a target node and $\mathbf{R}$ is the set of weights. Each hyperarc $e \in E$ is a triple $e = \langle T(e), h(e), f(e) \rangle$, where $h(e) \in N$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



Hypergraphs

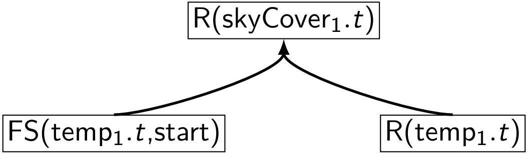
Definition

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



Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph

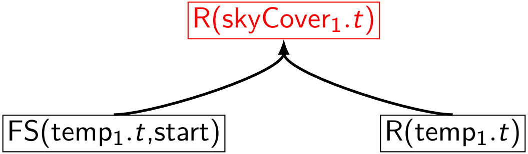


$$f(e) = f(\mathrm{FS}_{1,2}(\mathrm{temp}_1.t, \mathit{start})) \otimes f(\mathrm{R}_{2,3}(\mathrm{temp}_1.t)) \otimes \\ w(\mathrm{R}(\mathrm{skyCover}_1.t) \rightarrow \mathrm{FS}(\mathrm{temp}_1, \mathit{start})\ \mathrm{R}(\mathrm{temp}_1.t))$$

$$\boxed{\mathrm{R}(r_i.t) \rightarrow \mathrm{FS}(r_j, \mathit{start})\mathrm{R}(r_j.t)}$$

Hypergraph Construction

Map standard weighted CYK algorithm to hypergraph

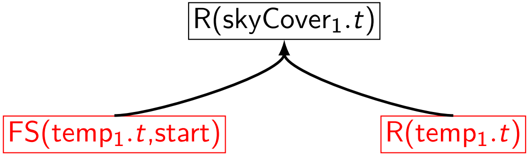


$$f(e) = f(\mathrm{FS}_{1,2}(\mathrm{temp}_1.t, start)) \otimes f(\mathrm{R}_{2,3}(\mathrm{temp}_1.t)) \otimes$$
$$w(\mathrm{R}(\mathrm{skyCover}_1.t) \rightarrow \mathrm{FS}(\mathrm{temp}_1, start)\ \mathrm{R}(\mathrm{temp}_1.t))$$

$$\boxed{\mathrm{R}(r_i.t) \rightarrow \mathrm{FS}(r_j, start)\mathrm{R}(r_j.t)}$$

Hypergraph Construction

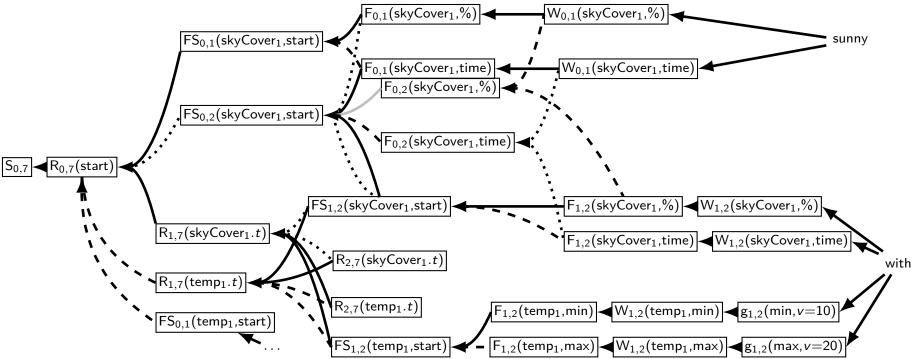
Map standard weighted CYK algorithm to hypergraph



$$
\begin{aligned}
f(e) = & f(\mathrm{FS}_{1,2}(\mathrm{temp}_1.t, start)) \otimes f(\mathrm{R}_{2,3}(\mathrm{temp}_1.t)) \otimes \\
& w(\mathrm{R}(\mathrm{skyCover}_1.t) \rightarrow \mathrm{FS}(\mathrm{temp}_1, start)\ \mathrm{R}(\mathrm{temp}_1.t))
\end{aligned}
$$

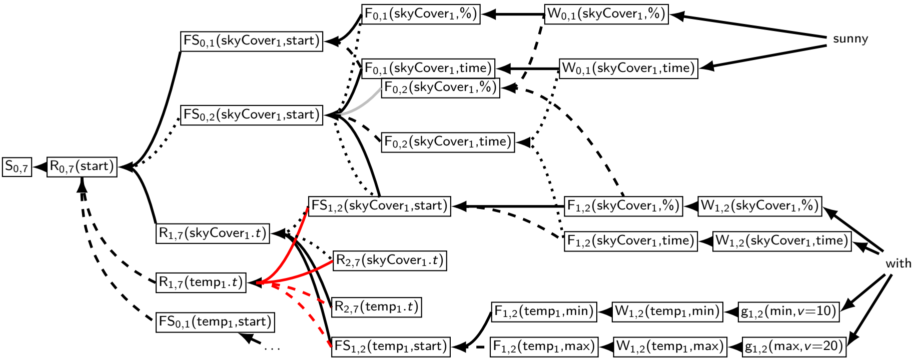
$$\mathrm{R}(r_i.t) \rightarrow \mathrm{FS}(r_j, start)\mathrm{R}(r_j.t)$$

Hypergraph Example



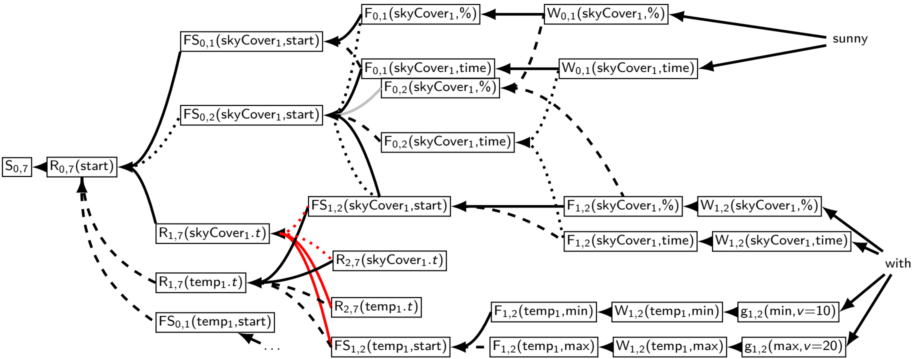
EM Training: dynamic program similar to the inside-outside algorithm

Hypergraph Example



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

Hypergraph Example



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

k-best Decoding

$$\arg\max_{w} P(\mathbf{w} \mid \mathbf{d}) = \arg\max_{w} P(\mathbf{w}) \cdot P(\mathbf{d} \mid \mathbf{w})$$

k-best Decoding

$$\arg\max_w P(\mathbf{w} \mid \mathbf{d}) = \arg\max_w P(\mathbf{w}) \cdot P(\mathbf{d} \mid \mathbf{w})$$

- Motivation: fluency and grammaticality
- Nodes in hypergraph $\rightarrow$ +LM items (Huang and Chiang, 2007)
  e.g. $\mathrm{R}_{2,8}(temp_1.t)^{\text{a low}\star\text{15 degrees}}$
- k-best Viterbi, cube pruning (Chiang, 2007)

k-best Decoding

Leaf nodes ϵ emit a k-best list of words

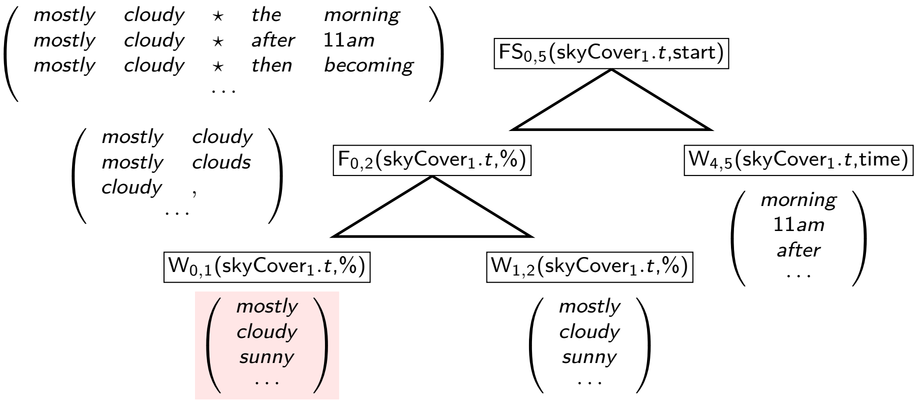
$$\boxed{\mathrm{W}_{0,1}(\mathrm{skyCover}_1.t,\%)}$$

$$\uparrow$$

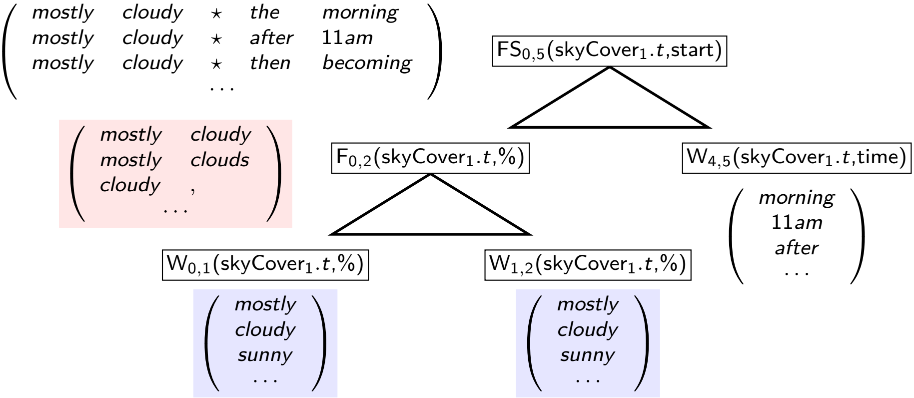
$$\epsilon$$

$$\begin{pmatrix} \textit{mostly} \\ \textit{cloudy} \\ \textit{sunny} \\ \ldots \end{pmatrix}$$

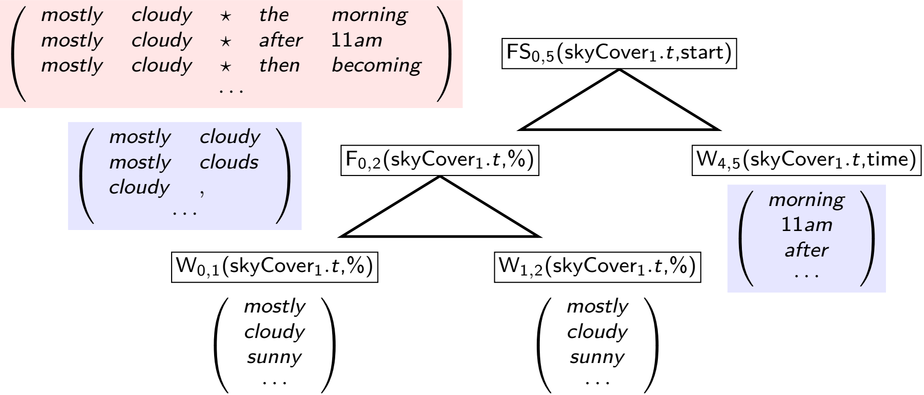
k-best Decoding



k-best Decoding



k-best Decoding



Experimental Setup

- ROBOCUP : simulated sportscasting [214 words] (Chen and Mooney, 2008)
- WEATHERGOV : weather reports [345 words] (Liang et al., 2009)
- ATIS : mapping from λ-version [927 words] (Zettlemoyer and Collins, 2007)

Experimental Setup

- ROBOCUP : simulated sportscasting [214 words] (Chen and Mooney, 2008)
- WEATHERGOV : weather reports [345 words] (Liang et al., 2009)
- ATIS : mapping from λ-version [927 words] (Zettlemoyer and Collins, 2007)
- Model parameters estimated on dev set (k-best, n-grams)

Experimental Setup

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

Results: Automatic Evaluation

System	RoboCup	WeatherGov	Atis
1-Best	10.79	8.64	11.85
k-Best	30.90	33.70	29.30
Angeli	28.70	38.40	26.77

- RoboCup results with fixed content selection;
- WeatherGov and Atis results with content selection and surface realization.

Results: Automatic Evaluation

System	RoboCup	WeatherGov	Atis
1-Best	10.79	8.64	11.85
k-Best	30.90	33.70	29.30
Angeli	28.70	38.40	26.77

- RoboCup results with fixed content selection;
- WeatherGov and Atis results with content selection and surface realization.

Results: Human Evaluation

RoboCup

System	Fluency	SemCor
1-Best	2.47	2.33
k-Best	4.31	3.96
Angeli	4.03	3.70
Human	4.47	4.37

WeatherGov

System	Fluency	SemCor
1-Best	1.82	2.05
k-Best	3.92	3.30
Angeli	4.26	3.60
Human	4.61	4.03

Atis

System	Fluency	SemCor
1-Best	2.40	2.46
k-Best	4.01	3.87
Angeli	3.56	3.33
Human	4.10	4.01

Results: Human Evaluation

RoboCup System	Fluency	SemCor
1-Best	2.47	2.33
k-Best	4.31	3.96
Angeli	4.03	3.70
Human	4.47	4.37

WeatherGov System	Fluency	SemCor
1-Best	1.82	2.05
k-Best	3.92	3.30
Angeli	4.26	3.60
Human	4.61	4.03

Atis System	Fluency	SemCor
1-Best	2.40	2.46
k-Best	4.01	3.87
Angeli	3.56	3.33
Human	4.10	4.01

Results: Human Evaluation

RoboCup

System	Fluency	SemCor
1-Best	2.47	2.33
k-Best	4.31	3.96
Angeli	4.03	3.70
Human	4.47	4.37

WeatherGov

System	Fluency	SemCor
1-Best	1.82	2.05
k-Best	3.92	3.30
Angeli	4.26	3.60
Human	4.61	4.03

Atis

System	Fluency	SemCor
1-Best	2.40	2.46
k-Best	4.01	3.87
Angeli	3.56	3.33
Human	4.10	4.01

Results: Human Evaluation

RoboCup

System	Fluency	SemCor
1-Best	2.47	2.33
k-Best	4.31	3.96
Angeli	4.03	3.70
Human	4.47	4.37

WeatherGov

System	Fluency	SemCor
1-Best	1.82	2.05
k-Best	3.92	3.30
Angeli	4.26	3.60
Human	4.61	4.03

Atis

System	Fluency	SemCor
1-Best	2.40	2.46
k-Best	4.01	3.87
Angeli	3.56	3.33
Human	4.10	4.01

Output

WeatherGov

Temperature

Time	Min	Mean	Max
06:00-21:00	30	38	44

Cloud Sky Cover

Time	Percent (%)
06:00-21:00	75-100

Chance of Rain

Time	Mode
06:00-21:00	Slight Chance

Wind Speed

Time	Min	Mean	Max
06:00-21:00	6	6	7

Wind Direction

Time	Mode
06:00-21:00	ENE

Precipitation Potential (%)

Time	Min	Mean	Max
06:00-21:00	9	20	35

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

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

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

Output

ATIS

Input:

Flight	
from	**to**
milwaukee	phoenix

Day	
day	**dep/ar/ret**
saturday	departure

Search	
type	**what**
query	flight

k-BEST: **What are the flights from Milwuakee to Phoenix on Saturday**

ANGELI : Show me the flights between Milwuakee and Phoenix on Saturday

HUMAN: Milwuakee to Phoenix on Saturday

Conclusions

- Generation as parsing problem using the hypergraph framework
- Unsupervised end-to-end generation system
- Performance comparable to state-of-the-art
- Future work: discriminative reranking (Konstas and Lapata, 2012b)

Demo

Live Weather Forecast Generator

Cross domain

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

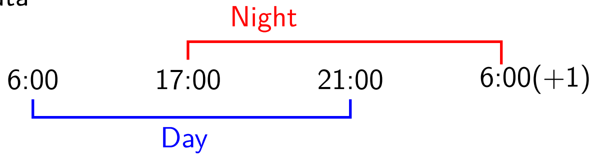
Demo

Live Weather Forecast Generator

Cross domain

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

Live Data



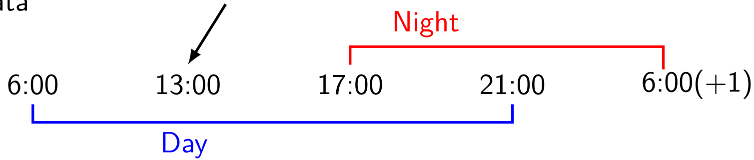
Demo

Live Weather Forecast Generator

Cross domain

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

Live Data



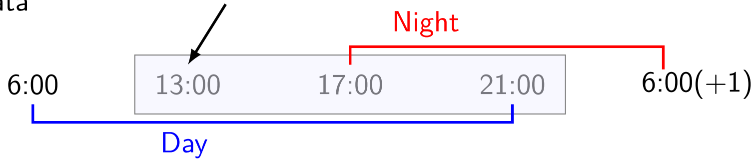
Demo

Live Weather Forecast Generator

Cross domain

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

Live Data



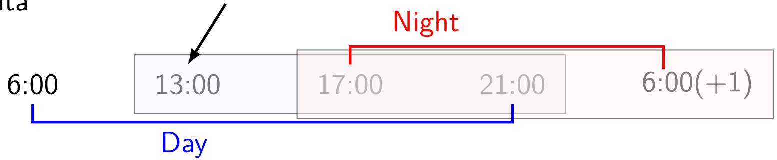
Demo

Live Weather Forecast Generator

Cross domain

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

Live Data



Thank you

Questions ?

Human Evaluation (Mturk)

Table 2

Category	Fields - Values			
Temperature (F)	**time:** 06.00 - 21.00	**min:** 31	**mean:** 38	**max:** 45
Wind Speed (mph)	**time:** 06.00 - 21.00	**min:** 6	**mean:** 7	**max:** 8
Wind Direction	**time:** 06.00 - 21.00	**mode:** SE		
Cloud Sky Cover (%)	**time:** 06.00 - 21.00	**percent:** 75-100		

Translation 2

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

Fluency ○ 1 ○ 2 ○ 3 ○ 4 ◉ 5
Semantic Correctness ○ 1 ○ 2 ○ 3 ◉ 4 ○ 5

Determining Text Length

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

Output

ROBOCUP

Input:

Pass	
From	**To**
purple10	purple11

k-BEST: **purple10 passes back to purple11**

ANGELI : purple10 passes to purple11

HUMAN: purple10 immediately passes to purple11