Efficient Extraction of Oracle-best Translations from Hypergraphs

Zhifei Li and Sanjeev Khudanpur

Center for Language and Speech Processing

Computer Science Department

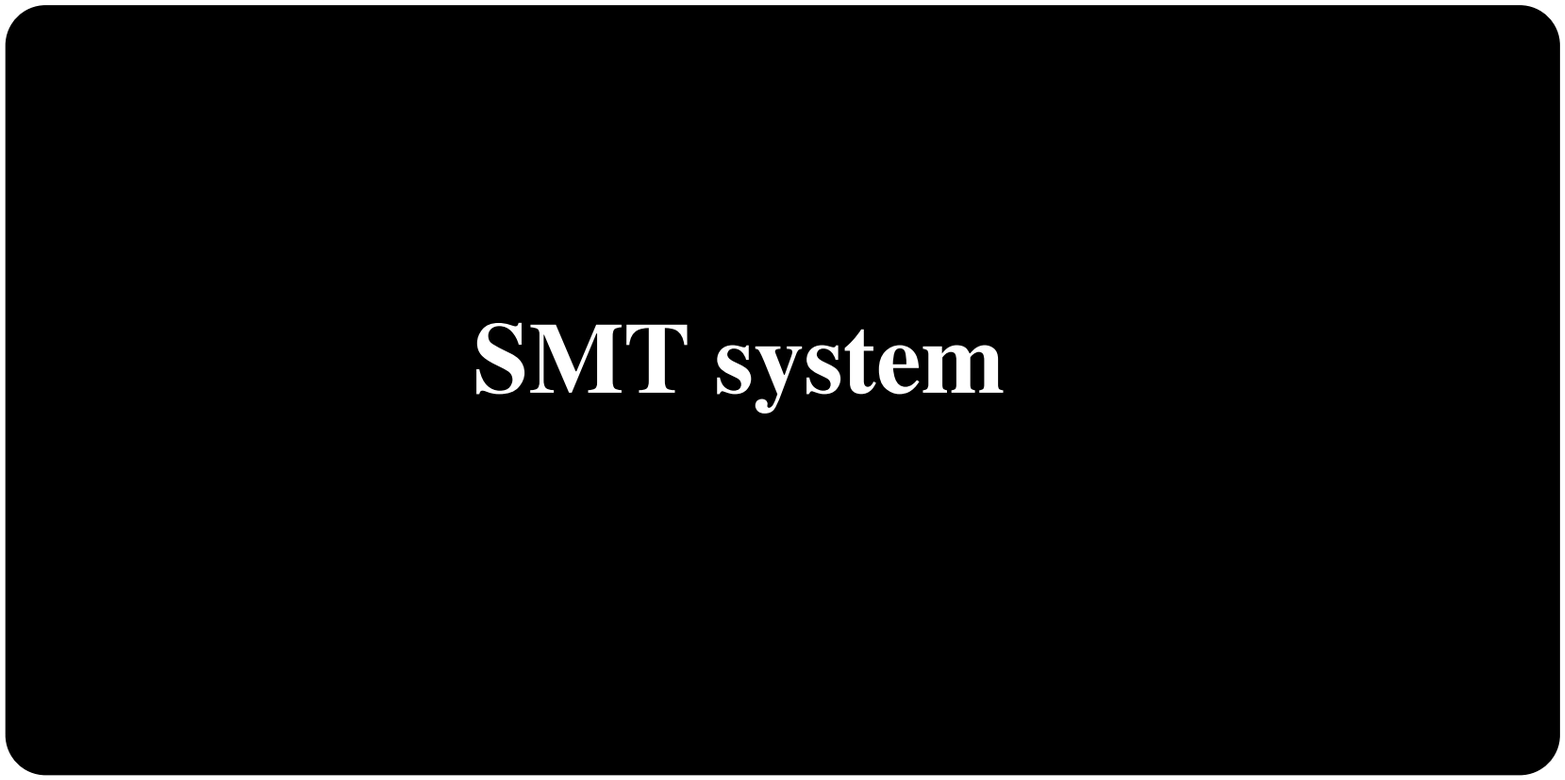
Johns Hopkins University

我垫子上的猫 $\Rightarrow$ **my cat on the mat**

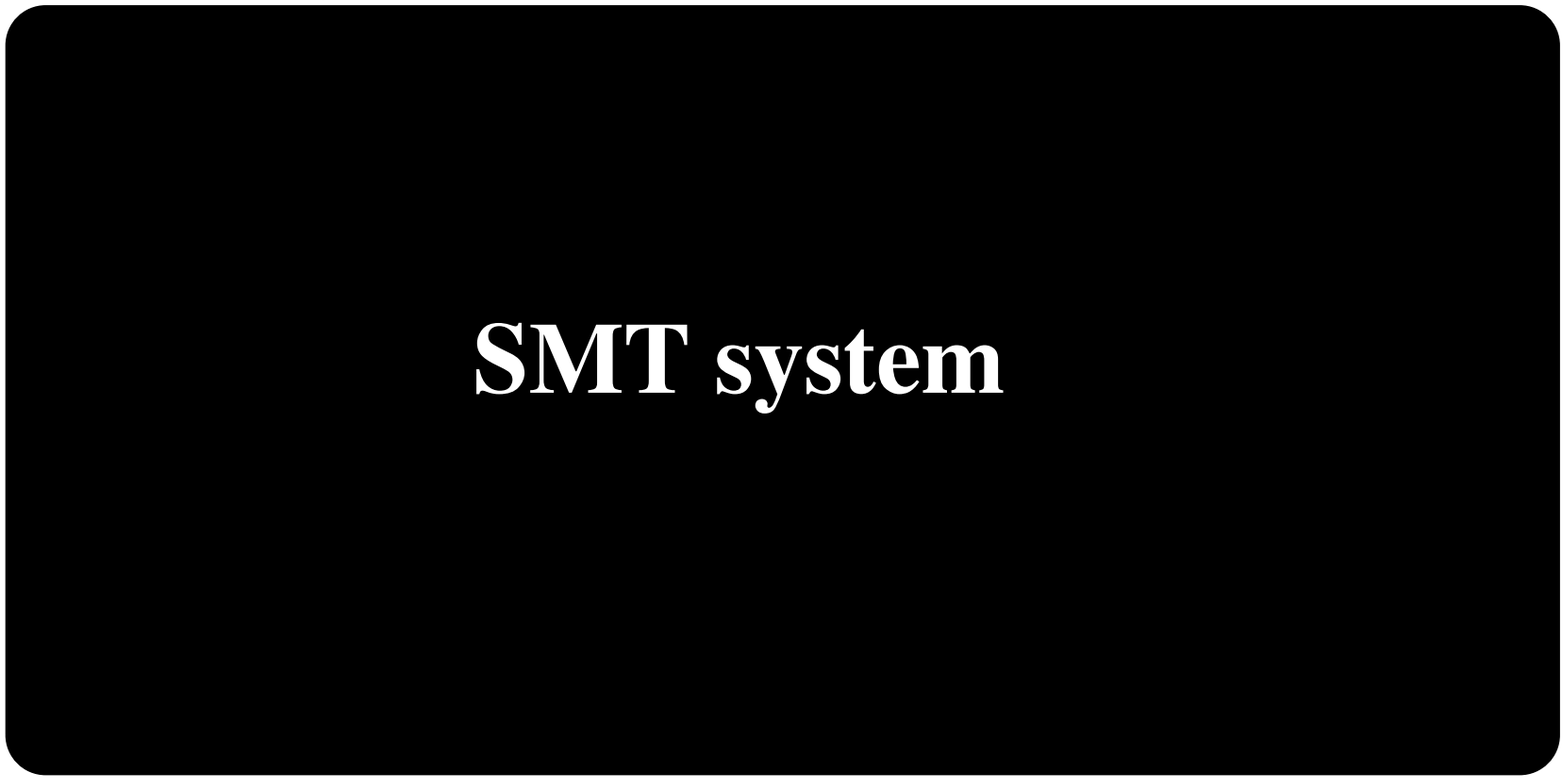
我垫子上的猫 ⇒ **my cat on the mat**

↓

我垫子上的猫 ⇒ **my cat on the mat**



我垫子上的猫 ⇒ **my cat on the mat**



the mat a cat

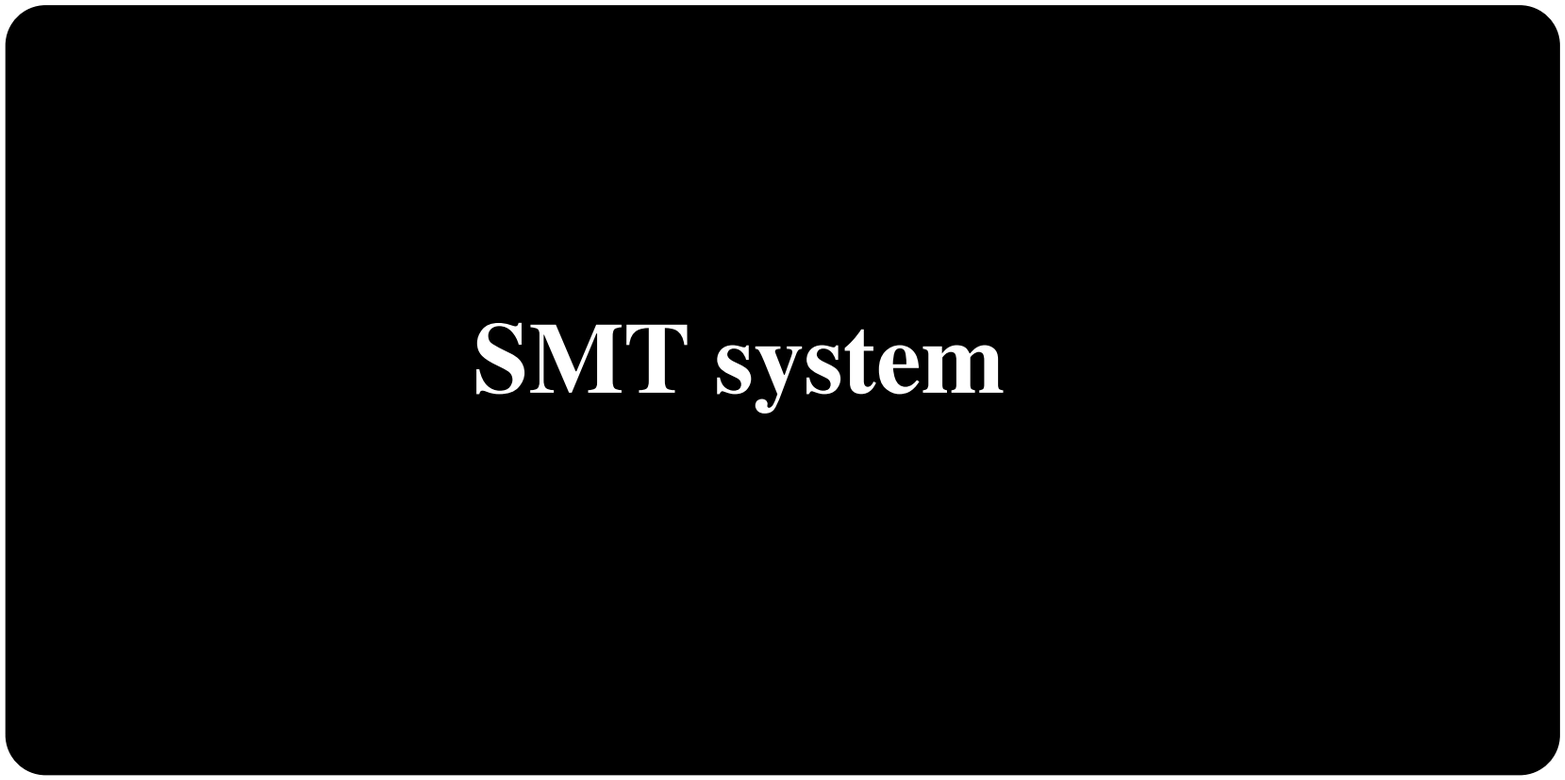
我垫子上的猫 ⇒ **my cat on the mat**

SMT system

the mat a cat

BLEU = 0.75

我垫子上的猫 ⇒ **my cat on the mat**

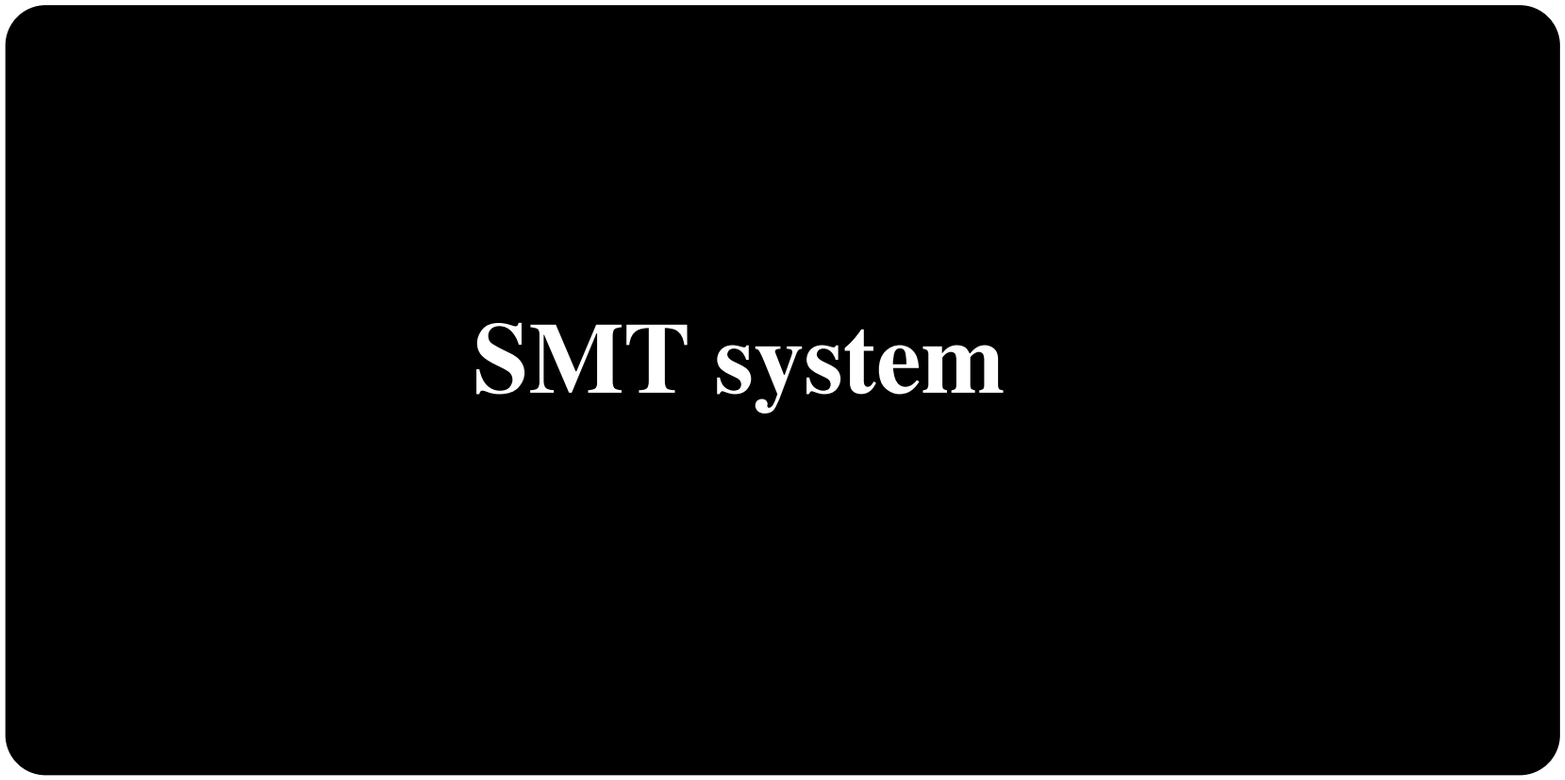


the mat a cat

BLEU = 0.75

the mat ‘s a cat

我垫子上的猫 ⇒ **my cat on the mat**



the mat a cat

BLEU = 0.75

the mat ‘s a cat

BLEU = 0.76

我垫子上的猫 ⇒ **my cat on the mat**

SMT system

the mat a cat

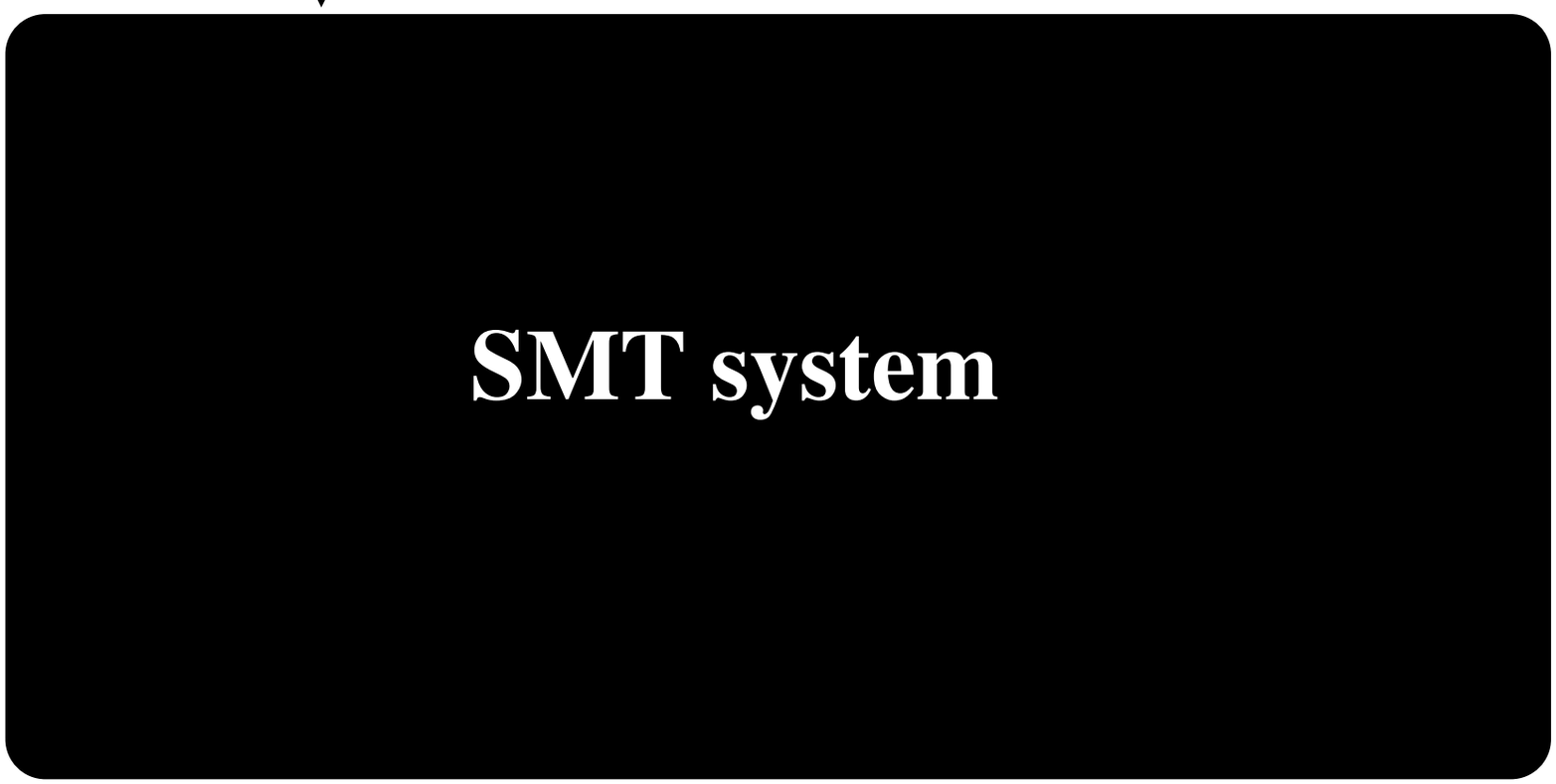
BLEU = 0.75

the mat ‘s a cat

BLEU = 0.76

a cat of the mat

我垫子上的猫 ⇒ **my cat on the mat**



the mat a cat

BLEU = 0.75

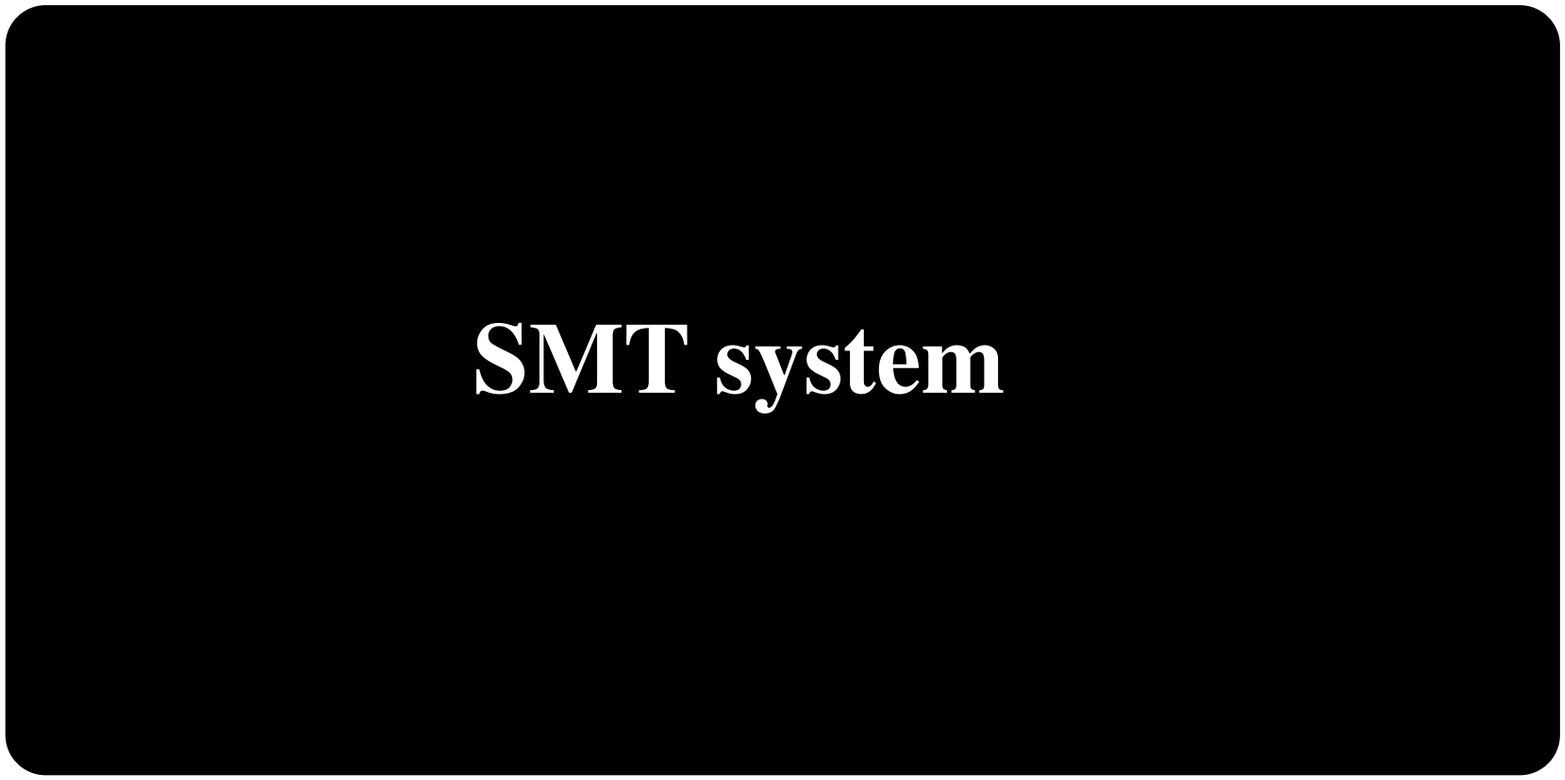
the mat ‘s a cat

BLEU = 0.76

a cat of the mat

BLEU = 0.77

我垫子上的猫 ⇒ **my cat on the mat**



the mat a cat

BLEU = 0.75

the mat ‘s a cat

BLEU = 0.76

a cat of the mat

BLEU = 0.77

a cat on the mat

我垫子上的猫 ⇒ **my cat on the mat**

SMT system

the mat a cat

BLEU = 0.75

the mat ‘s a cat

BLEU = 0.76

a cat of the mat

BLEU = 0.77

a cat on the mat

BLEU = 0.85

我垫子上的猫 ⇒ **my cat on the mat**

SMT system

the mat a cat

BLEU = 0.75

the mat ‘s a cat

BLEU = 0.76

a cat of the mat

BLEU = 0.77

a cat on the mat

BLEU = 0.85

我垫子上的猫 ⇒ **my cat on the mat**

SMT system

the mat a cat

BLEU = 0.75

the mat ‘s a cat

BLEU = 0.76

a cat of the mat

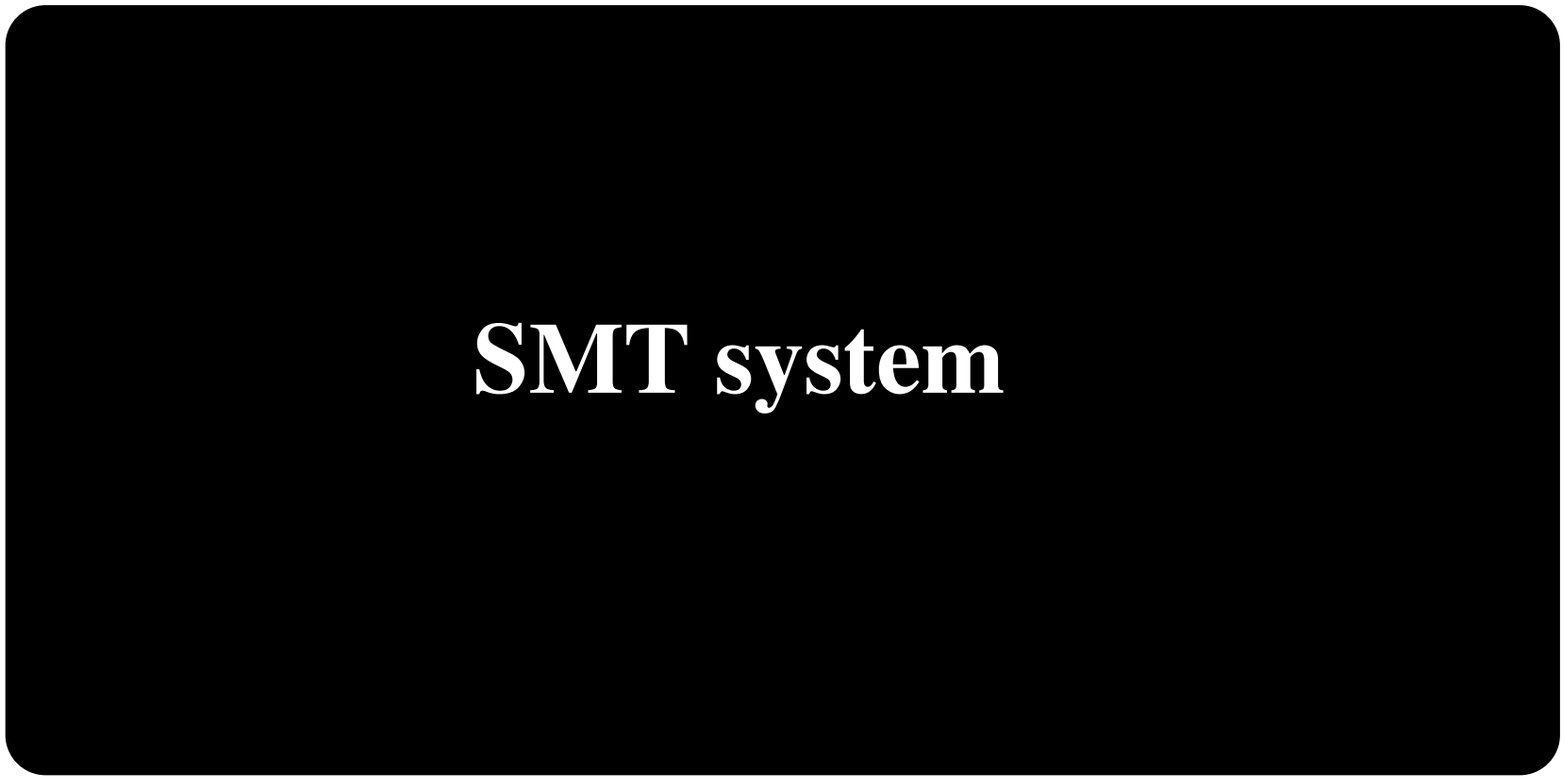
BLEU = 0.77

a cat on the mat

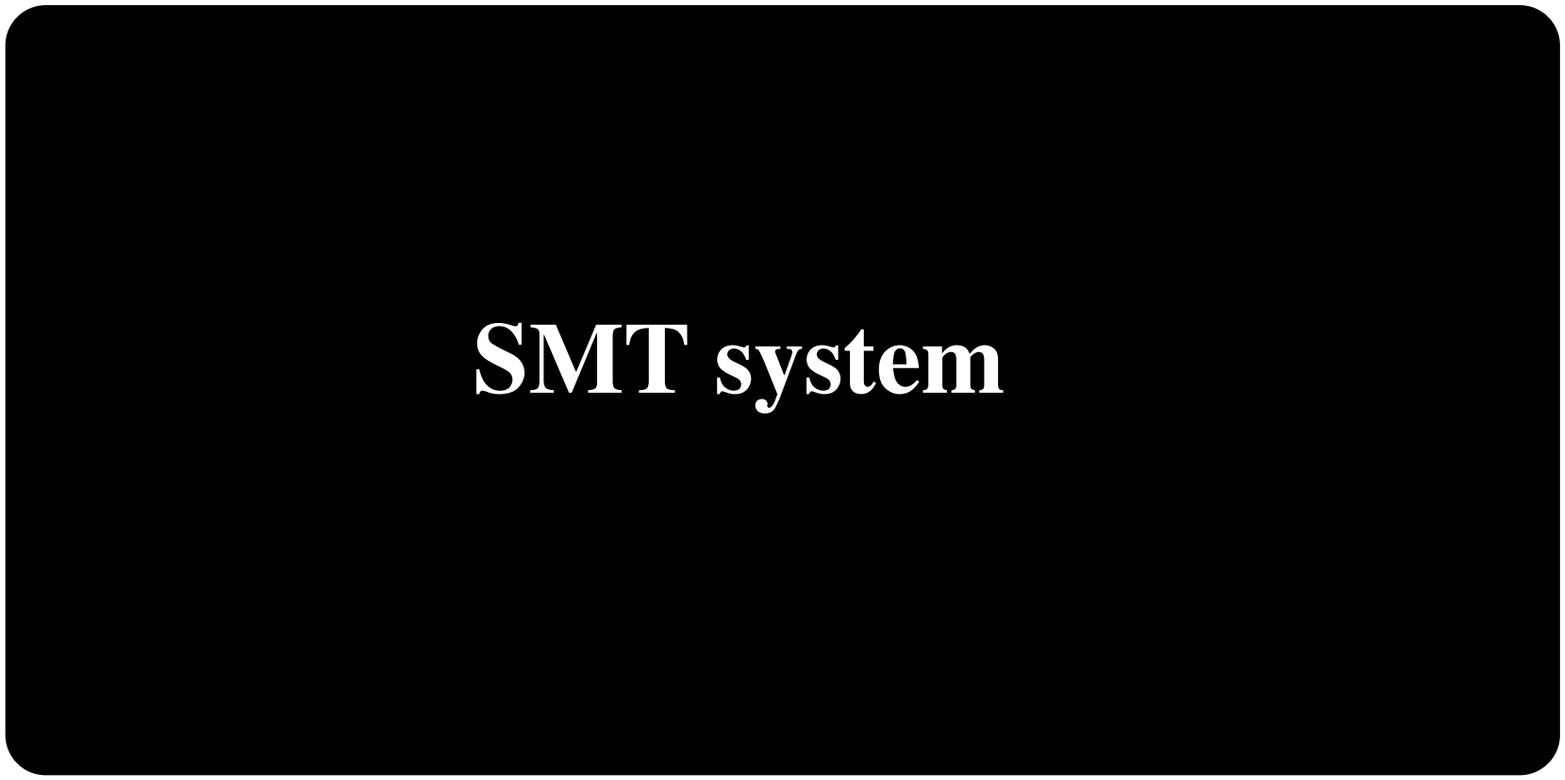
oracle translation

BLEU = 0.85

我垫子上的猫 ⇒ **my cat on the mat**

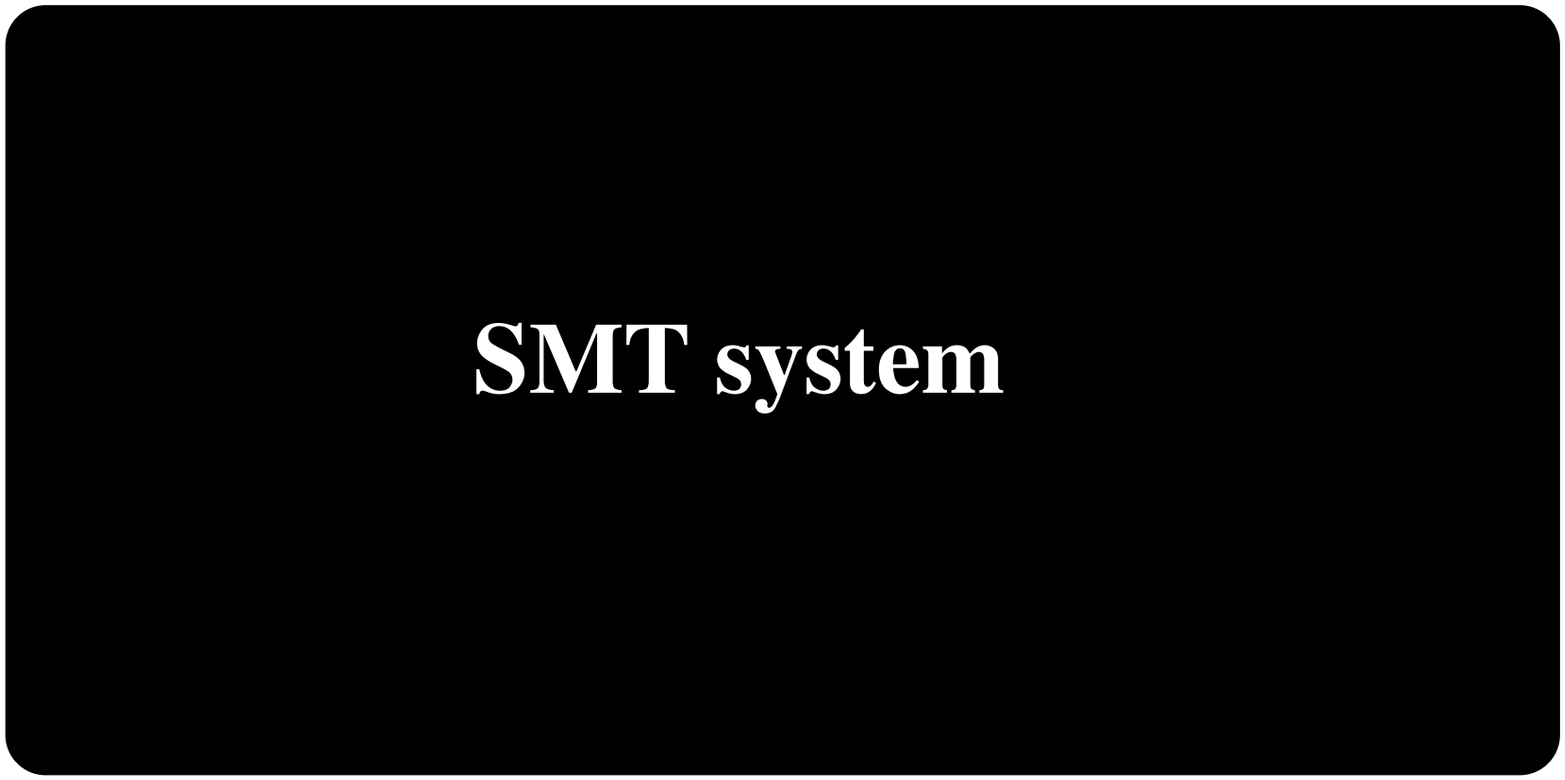


我垫子上的猫 ⇒ **my cat on the mat**

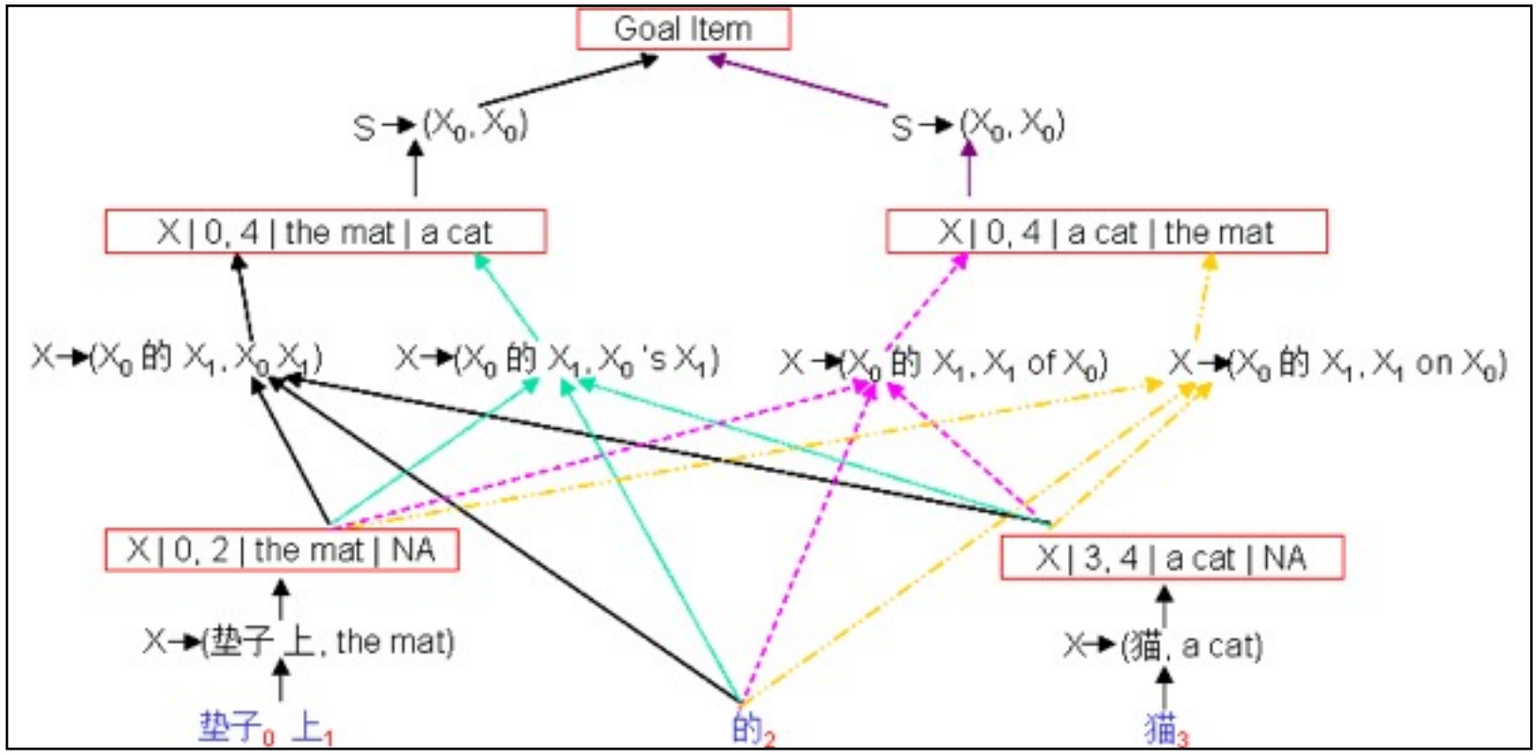


- Applications of oracle extraction:
  - discriminative training
  - measuring pruning quality
  - system combination
  - multi-source translation

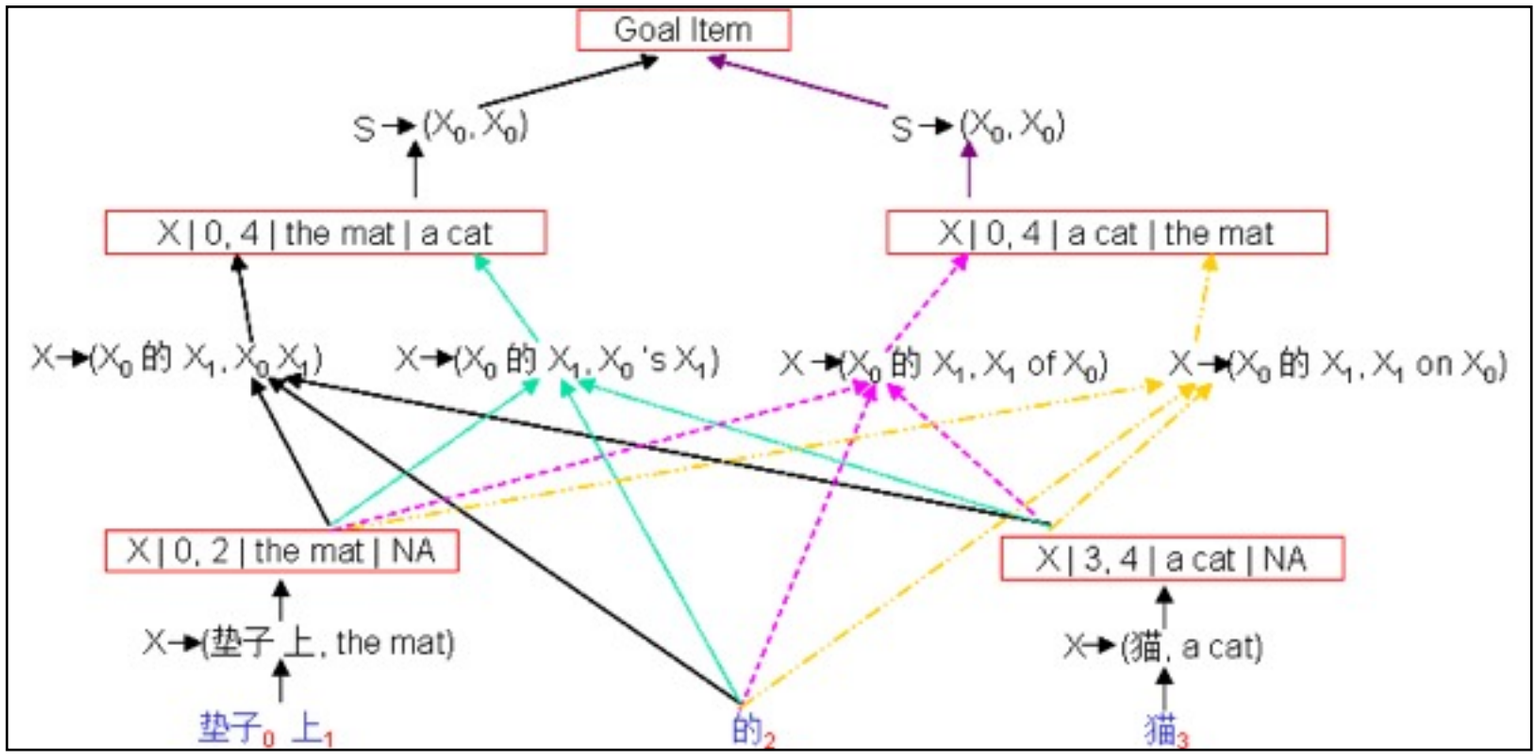
我垫子上的猫 ⇒ **my cat on the mat**

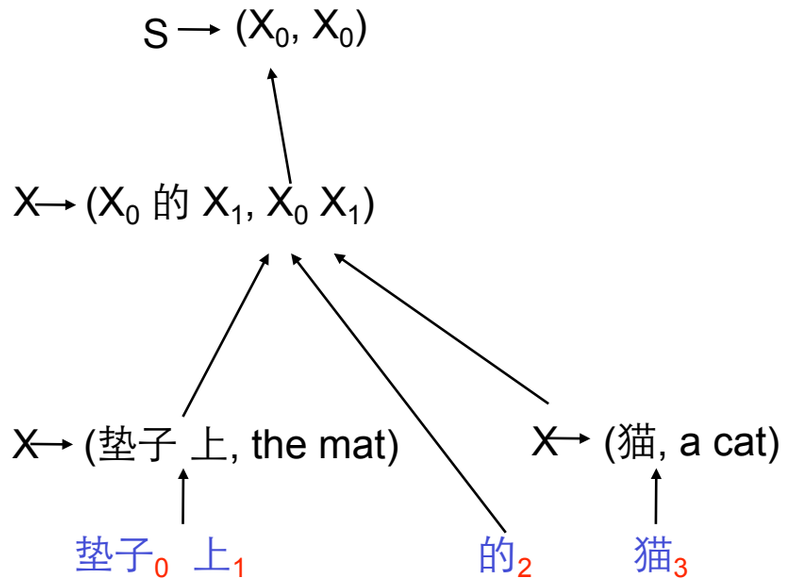


我垫子上的猫 ⇒ **my cat on the mat**

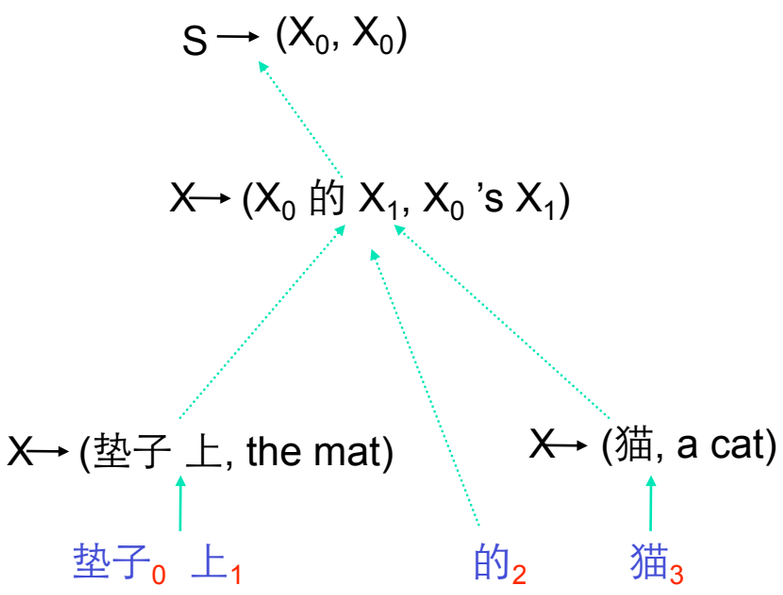


我垫子上的猫 ⇒ **my cat on the mat**

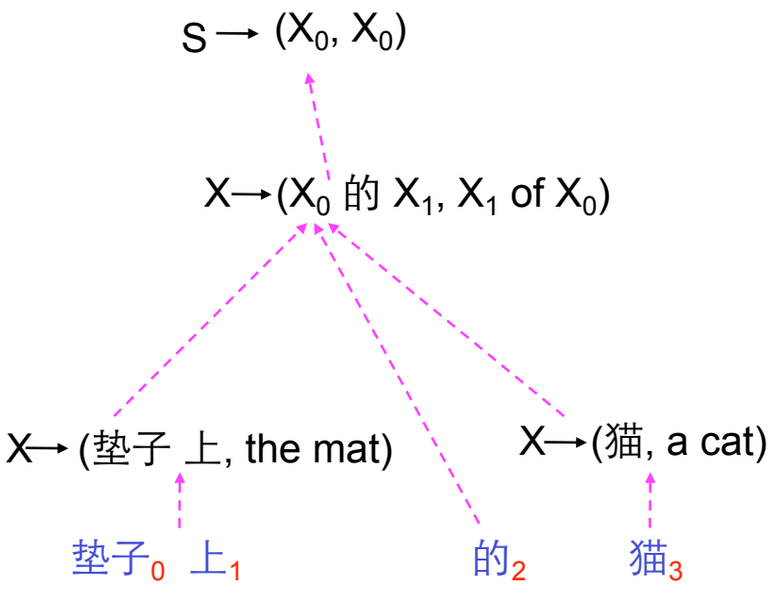




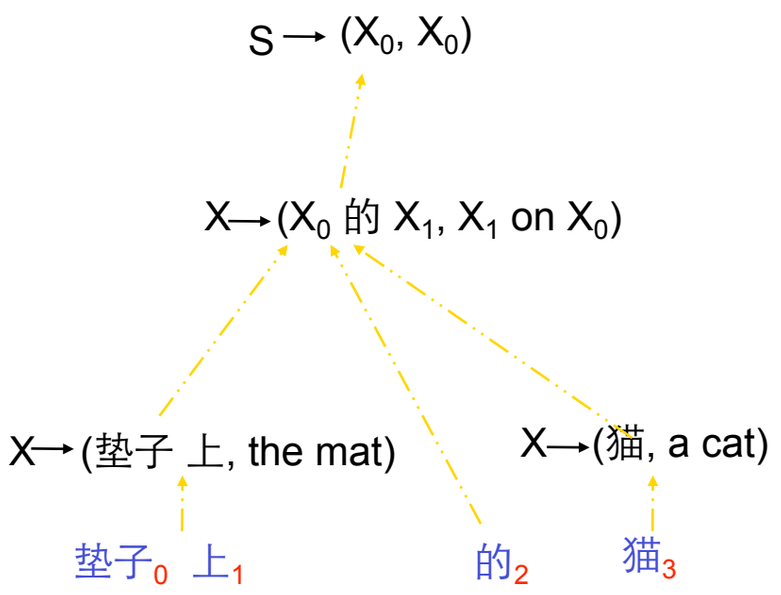
the mat a cat



the mat 's a cat



a cat of the mat



a cat on the mat

我垫子上的猫 ⇒ **my cat on the mat**

the mat a cat

BLEU = 0.75

the mat ‘s a cat

BLEU = 0.76

a cat of the mat

BLEU = 0.77

a cat on the mat

BLEU = 0.85

我垫子上的猫 ⇒ **my cat on the mat**

the mat a cat

the mat ‘s a cat **a cat of the mat** **a cat on the mat**

我垫子上的猫 ⇒ **my cat on the mat**

a cat on the floor

the mat a cat

a mouse on the mat

my cat of the mat

a pig on the mat

his cat of the mat

a dog on the mat

the mat ‘s a cat

a cat of the mat

a cat on the mat

我垫子上的猫 ⇒ my cat on the mat

a cat on the floor

the mat a cat

a mouse on the mat

my cat of the mat

his cat of the mat

the mat ‘s a cat

a cat of the mat

a pig on the mat

a dog on the mat

a cat on the mat

exponentially many translations!

Oracle extraction from a hypergraph

- Idea 1: oracle extraction as model scoring
- Idea 2: equivalent oracle state maintenance

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	**count**
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

Hypothesis:

n-gram	**count**
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

Hypothesis:

your cat on the mat

n-gram	**count**
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

Hypothesis:

your cat on the mat

$$p_1 = 4/5$$

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

Hypothesis:

your cat on the mat

n-gram	**count**
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

$$p_1 = 4/5$$

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction =

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction =

n-gram language model scoring on a hypergraph

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction $\approx$

n-gram language model scoring on a hypergraph

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction $\approx$

n-gram language model scoring on a hypergraph

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction $\approx$

n-gram language model scoring on a hypergraph

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction $\approx$

n-gram language model scoring on a hypergraph

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction $\approx$

n-gram language model scoring on a hypergraph

$$\log \mathrm{BLEU}(r,h) = \min\left[1 - \frac{|r|}{|h|}, 0\right] + \sum_{n=1}^{4} \frac{1}{4} \log p_n$$

Reference:

my cat on the mat

n-gram	count
my	1
cat	1
on	1
the	1
mat	1
my cat	1
cat on	1
on the	1
the mat	1
my cat on	1
cat on the	1
on the mat	1
my cat on the	1
cat on the mat	1

Hypothesis:

your cat on the mat

$p_1 = 4/5$ $p_2 = 3/4$

$p_3 = 2/3$ $p_4 = 1/2$

$\mathrm{BLEU} = 0.669$

???

Oracle extraction $\approx$

n-gram language model scoring on a hypergraph

BLEU score

Idea 1: oracle extraction as model scoring

- Similar to regular n-gram LM integration
- At each node, we maintain:
  - Oracle states
    - left- and right n-gram states
    - hypothesis length
  - Scores
    - # of n-gram matches
    - log BLEU value

Idea 1: oracle extraction as model scoring

- Similar to regular n-gram LM integration
- At each node, we maintain:
  - Oracle states
    - left- and right n-gram states
    - hypothesis length
  - Scores
    - # of n-gram matches
    - log BLEU value

- two approximations:
  - ignore clipping
  - approximated reference length

Idea 1: oracle extraction as model scoring

- Similar to regular n-gram LM integration
- At each node, we maintain:
  - Oracle states
    - left- and right n-gram states
    - hypothesis length
  - Scores
    - # of n-gram matches
    - log BLEU value

- two approximations:
  - ignore clipping
  - approximated reference length

Node	$\|\mathbf{h}\|$	$\|\widetilde{\mathbf{r}}\|$	**matches**	$\log$ **BLEU**
Node A	5	6.2	(3, 2, 2, 1)	-0.82
Node B	10	9.8	(8, 7, 6, 5)	-0.27
New	2	NA	(1, 1, 1, 0)	NA
Node C	17	18.3	(12, 10, 9, 6)	-0.62

Idea 1: oracle extraction as model scoring

- Similar to regular n-gram LM integration
- At each node, we maintain:
  - Oracle states
    - left- and right n-gram states
    - hypothesis length
  - Scores
    - # of n-gram matches
    - log BLEU value

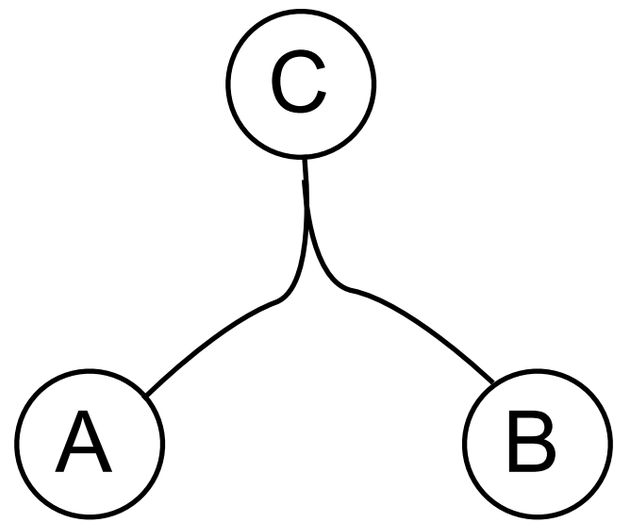
- two approximations:
  - ignore clipping
  - approximated reference length

Node	$\|\mathbf{h}\|$	$\|\widetilde{\mathbf{r}}\|$	**matches**	$\log$ **BLEU**
Node A	5	6.2	(3, 2, 2, 1)	-0.82
Node B	10	9.8	(8, 7, 6, 5)	-0.27
New	2	NA	(1, 1, 1, 0)	NA
Node C	17	18.3	(12, 10, 9, 6)	-0.62

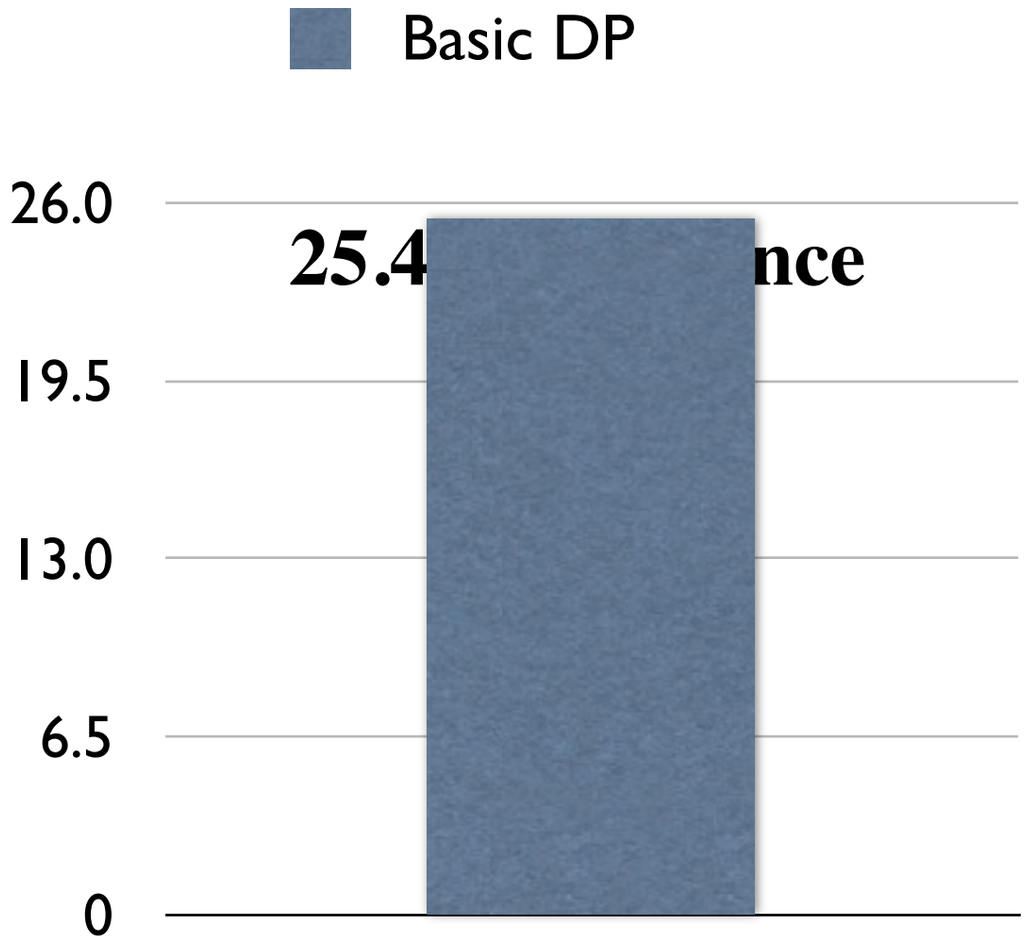
Idea 1: oracle extraction as model scoring

- Similar to regular n-gram LM integration
- At each node, we maintain:
  - Oracle states
    - left- and right n-gram states
    - hypothesis length
  - Scores
    - # of n-gram matches
    - log BLEU value

- two approximations:
  - ignore clipping
  - approximated reference length



Node	$\|\mathbf{h}\|$	$\|\widetilde{\mathbf{r}}\|$	**matches**	$\log$ **BLEU**
Node A	5	6.2	(3, 2, 2, 1)	-0.82
Node B	10	9.8	(8, 7, 6, 5)	-0.27
New	2	NA	(1, 1, 1, 0)	NA
Node C	17	18.3	(12, 10, 9, 6)	-0.62


Basic DP
26.0
25.4
nce
19.5
13.0
6.5
0

Equivalent oracle state maintenance

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

  left-state ... right-state

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

  left-state ... right-state

  my cat on ... on the mat

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

my cat on ... on the mat

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

my cat on ... on the mat

my cat in ... on the mat

my cat at ... on the mat

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

my cat on ... on the mat

my cat in ... on the mat

my cat at ... on the mat

my dog in ... on the mat

my mouse at ... on the mat

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

my cat on ... on the mat

my cat in ... on the mat

my cat at ... on the mat

my dog in ... on the mat

my mouse at ... on the mat

his cat on ... on the mat

her cat on ... on the mat

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

my cat on ... on the mat

my cat in ... on the mat

my cat at ... on the mat

my dog in ... on the mat

my mouse at ... on the mat

his cat on ... on the mat

her cat on ... on the mat

my cat on ... on the bed

my cat on ... on the floor

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

Reference: watch my cat on the mat , please !

my cat on ... on the mat

my cat in ... on the mat

my cat at ... on the mat

my dog in ... on the mat

my mouse at ... on the mat

his cat on ... on the mat

her cat on ... on the mat

my cat on ... on the bed

my cat on ... on the floor

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

Reference: watch my cat on the mat , please !

my cat on ... on the mat $\Rightarrow$ my cat on ... on the mat

my cat in ... on the mat

my cat at ... on the mat

my dog in ... on the mat

my mouse at ... on the mat

his cat on ... on the mat

her cat on ... on the mat

my cat on ... on the bed

my cat on ... on the floor

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

Reference: watch my cat on the mat , please !

my cat on ... on the mat $\Rightarrow$ my cat on ... on the mat

my cat in ... on the mat
my cat at ... on the mat
} $\Rightarrow$ my cat * ... on the mat

my dog in ... on the mat
my mouse at ... on the mat

his cat on ... on the mat
her cat on ... on the mat

my cat on ... on the bed
my cat on ... on the floor

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

Reference: watch my cat on the mat , please !

my cat on ... on the mat $\Rightarrow$ my cat on ... on the mat

my cat in ... on the mat
my cat at ... on the mat
} $\Rightarrow$ my cat * ... on the mat

my dog in ... on the mat
my mouse at ... on the mat
} $\Rightarrow$ my * * ... on the mat

his cat on ... on the mat
her cat on ... on the mat

my cat on ... on the bed
my cat on ... on the floor

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

Reference: watch my cat on the mat , please !

my cat on ... on the mat $\Rightarrow$ my cat on ... on the mat

my cat in ... on the mat
my cat at ... on the mat
} $\Rightarrow$ my cat * ... on the mat

my dog in ... on the mat
my mouse at ... on the mat
} $\Rightarrow$ my * * ... on the mat

his cat on ... on the mat
her cat on ... on the mat
} $\Rightarrow$ * * * ... on the mat

my cat on ... on the bed
my cat on ... on the floor

Equivalent oracle state maintenance

- Need to maintain n-gram states at each node

left-state ... right-state

my cat on ... on the mat

Reference: watch my cat on the mat , please !

my cat on ... on the mat $\Rightarrow$ my cat on ... on the mat

my cat in ... on the mat
my cat at ... on the mat
} $\Rightarrow$ my cat * ... on the mat

my dog in ... on the mat
my mouse at ... on the mat
} $\Rightarrow$ my * * ... on the mat

his cat on ... on the mat
her cat on ... on the mat
} $\Rightarrow$ * * * ... on the mat

my cat on ... on the bed
my cat on ... on the floor
} $\Rightarrow$ my cat on ... * * *

Idea 2: equivalent oracle states

Idea 2: equivalent oracle states

- reference
  - a b c d e

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

State	Is-a-suffix	Reduced state
c f a	no	c f *
c f *	no	c * *
c * *	yes	c * *

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

State	Is-a-suffix	Reduced state
c f a	no	c f *
c f *	no	c * *
c * *	yes	c * *

- right LM state: f a b => * a b

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

State	Is-a-suffix	Reduced state
c f a	no	c f *
c f *	no	c * *
c * *	yes	c * *

- right LM state: f a b => * a b

State	Is-a-prefix	Reduced state
f a b	no	* a b
* a b	yes	* a b

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

State	Is-a-suffix	Reduced state
c f a	no	c f *
c f *	no	c * *
c * *	yes	c * *

- right LM state: f a b => * a b

State	Is-a-prefix	Reduced state
f a b	no	* a b
* a b	yes	* a b

c f a ... f a b

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

State	Is-a-suffix	Reduced state
c f a	no	c f *
c f *	no	c * *
c * *	yes	c * *

- right LM state: f a b => * a b

State	Is-a-prefix	Reduced state
f a b	no	* a b
* a b	yes	* a b

c f a ... f a b

=> c * * ... * a b

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

State	Is-a-suffix	Reduced state
c f a	no	c f *
c f *	no	c * *
c * *	yes	c * *

- right LM state: f a b => * a b

State	Is-a-prefix	Reduced state
f a b	no	* a b
* a b	yes	* a b

c f a ... f a b
c e a ... f a b => c * * ... * a b

Idea 2: equivalent oracle states

- reference
  - a b c d e

n-gram	count
a	1
b	1
c	1
d	1
e	1
a b	1
b c	1
c d	1
d e	1
a b c	1
b c d	1
c d e	1
a b c d	1
b c d e	1

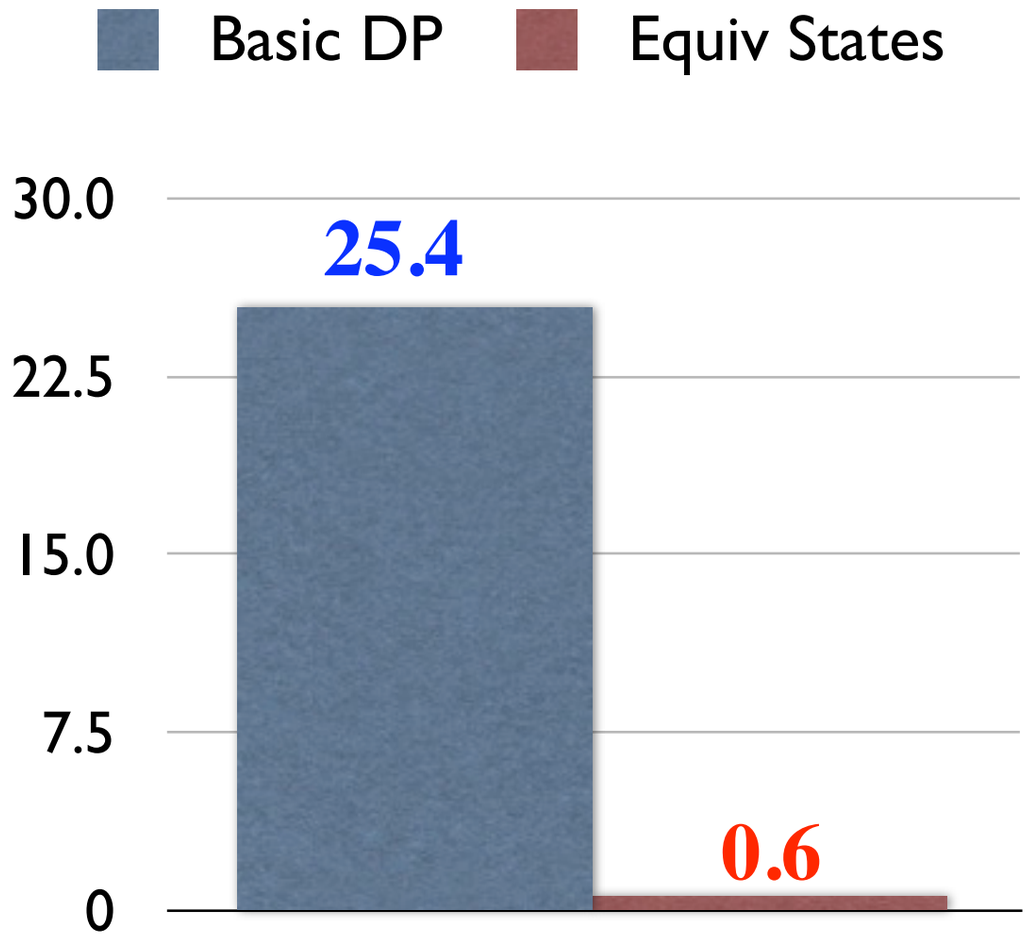
- regular LM state: c f a ... f a b
- left LM state: c f a => c * *

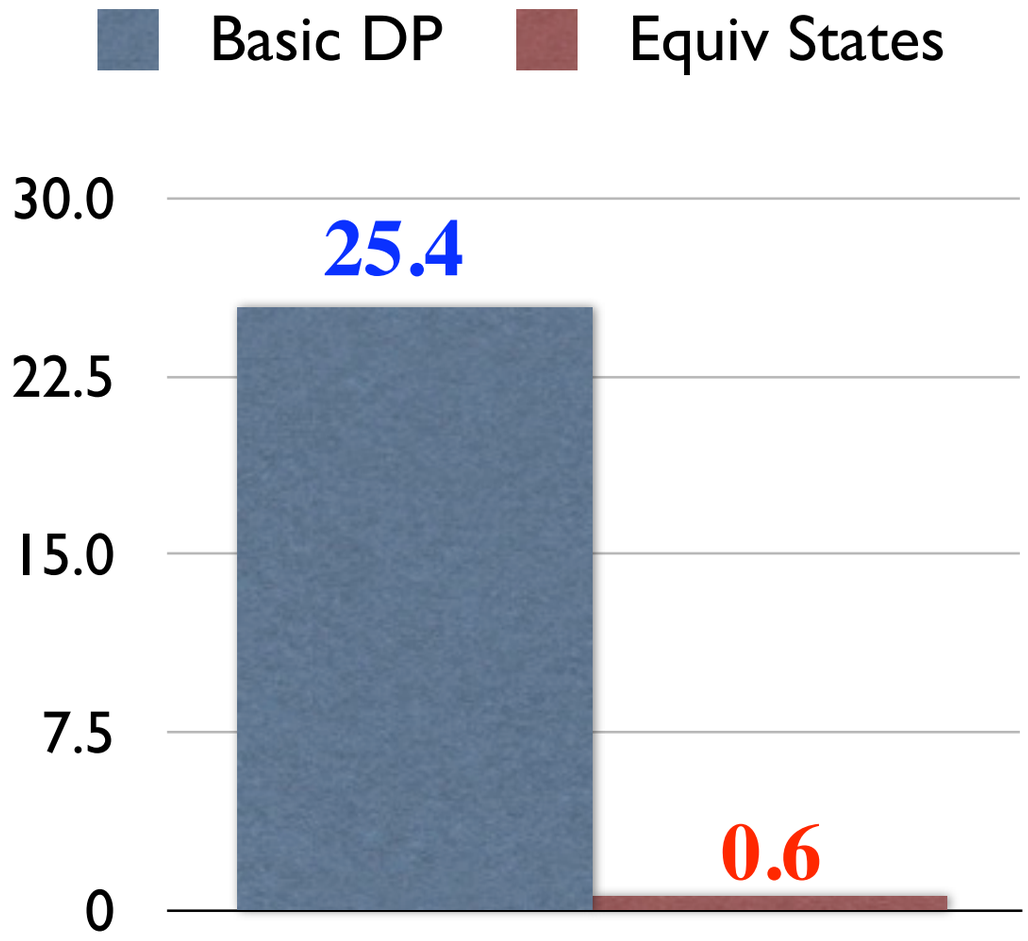
State	Is-a-suffix	Reduced state
c f a	no	c f *
c f *	no	c * *
c * *	yes	c * *

- right LM state: f a b => * a b

State	Is-a-prefix	Reduced state
f a b	no	* a b
* a b	yes	* a b

c f a ... f a b
c e a ... f a b => c * * ... * a b
c f a ... e a b


Basic DP
Equiv States
30.0
22.5
15.0
7.5
0
25.4
0.6



42 times speed up!

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

- Oracle BLEU is much better than the model's 1-best BLEU

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

- Oracle BLEU is much better than the model's 1-best BLEU

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

- Oracle BLEU is much better than the model's 1-best BLEU
- Oracle BLEU on hypergraphs is much better than on n-bests

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

- Oracle BLEU is much better than the model's 1-best BLEU
- Oracle BLEU on hypergraphs is much better than on n-bests

Results: oracle BLEU scores

- On NIST Chinese to English MT sets

Hypothesis space	MT'04	MT'05	MT'06
1-best (Baseline)	35.7	32.6	28.3
500-unique-best	44.0	41.2	35.1
Hypergraph	52.8	51.8	37.8
500-best oracles	**53.2**	**52.2**	**38.0**

- Oracle BLEU is much better than the model's 1-best BLEU
- Oracle BLEU on hypergraphs is much better than on n-bests
- Approximations (i.e., no clipping and approximated reference length) do not affect the oracle quality very much

Discriminative hypergraph reranking

- Similar to forest reranking in monolingual parsing (Huang 2008)
- Trained on about one million hypergraphs/oracles using the averaged perceptron algorithm

Test Set	MT'04	MT'05	MT'06
Baseline	35.7	32.6	28.3
Discrim. LM	35.9	33.0	28.2
Discrim. TM	**36.1**	**33.2**	**28.7**
Discrim. TM+LM	36.0	33.1	28.6

- Oracle extraction is useful for training a discriminative translation or language model
- Discriminative hypergraph reranking improves BLEU scores


Joshua

Joshua project

- An open-source parsing-based MT toolkit (Li et al. 2009)
- Team members
  - Zhifei Li, Chris Callison-Burch, Chris Dyer, Sanjeev Khudanpur, Wren Thornton, Jonathan Weese, Juri Ganitkevitch, Lane Schwartz, and Omar Zaidan
- Functions
  - Chart-parsing, pruning, language model integration, kbest extraction, distributed and parallel decoding
  - Suffix-array based grammar extraction
  - Minimum error rate training
  - More to add: variational decoding, minimum risk annealing, semiring parsing, hypergraph rescoring, and so on

Thank you!

谢谢!