Statistical

Machine Translation

A True Story (circa 1949)

Give me encoded text. I've got a new decoding algorithm.

81593
29712
03210
39711
21031
20103
32113
12011
21145
71141
11000

P: I'm German, but know Turkish. M does not know that.

But M got it right anyway (without knowing Turkish)

M

P

Sure!

(one day later) I failed! All I got was gibberish!

bu ne anlama geliyor

It is exactly right!

Statistical Machine Translation

Warren Weaver,
July 1949

one naturally wonders if the problem of translation could conceivably be treated as a problem in cryptography.

When I look at an article in Russian, I say "This is really written in English, but it has been coded in some strange symbols.

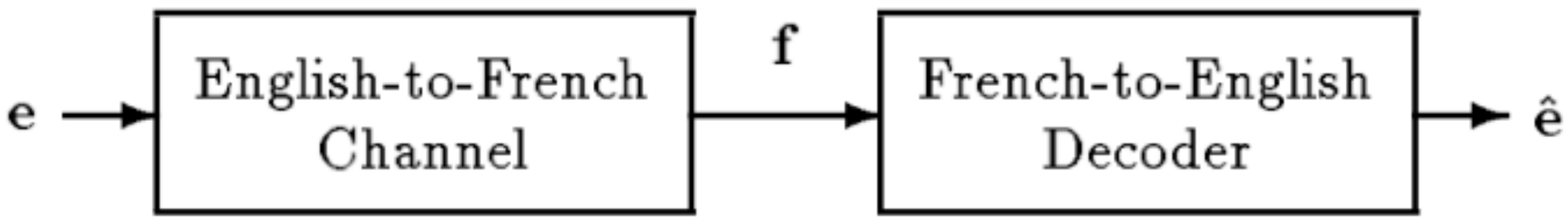
I will now proceed to decode.

(1988-1994)

The Candide System for Machine Translation

Adam L. Berger, Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra,*
John R. Gillett, John D. Lafferty, Robert L. Mercer, Harry Printz, Luboš Ureš*

IBM Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598



Bayes' Rule

$$\hat{\mathbf{e}} = \underset{\mathbf{e}}{\operatorname{argmax}} \Pr(\mathbf{e} \mid \mathbf{f}) = \underset{\mathbf{e}}{\operatorname{argmax}} \Pr(\mathbf{f} \mid \mathbf{e}) \Pr(\mathbf{e})$$

⏭ Present time

http://googleresearch.blogspot.ca/2006/04/statistical-machine-translation-live.html

Franz Och,
Apr 28 2006

Most state-of-the-art commercial machine translation systems in use today have been developed using a rules-based approach ...

Several research systems, including ours, take a different approach: **we feed the computer with billions of words of text**, both monolingual text in the target language, and aligned text consisting of examples of human translations between the languages.

We then apply statistical learning techniques to build a translation model.

Jean-François Champollion Deciphers Egyptian hieroglyphs using the Rosetta Stone, 1822

Machine Translation

澳洲　是　与　北韩　有　邦交　的　国家　之一

Australia is one of the countries that has diplomatic relations with North Korea

Google　Australia has diplomatic relations with North Korea is one of the countries

Bing　Australia is among the countries that have formal diplomatic relations with North Korea

SDL　Australia is still having diplomatic ties with North Korea , one of the countries

IBM Translation Server　costs USD 25,400

Discriminative Training for Statistical Machine Translation

Parallel Text
(Web, United Nations, European Parliament, Canadian Parliament, Wikipedia, ...)
Learn alignments between words in source and target languages (IBM Models)

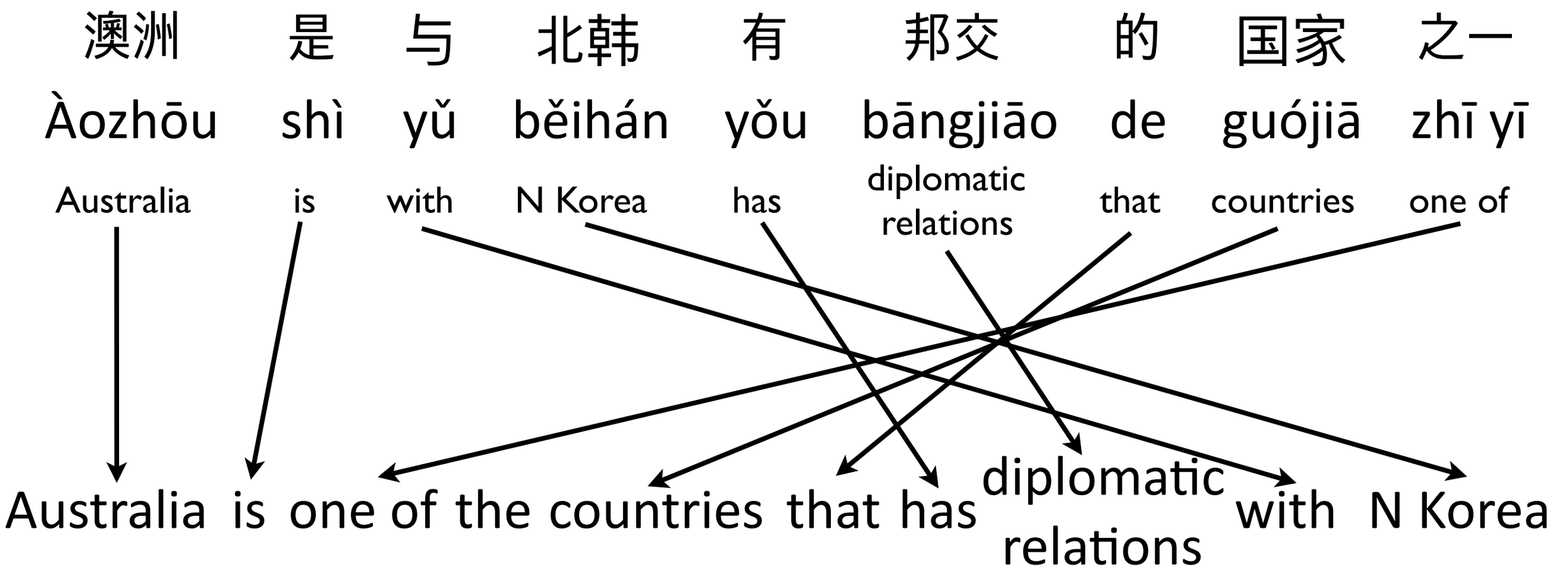
封面說明
百葉窗後
編者的話
八月號光
近幾個月
五月，是
重開雙橡
兩邊看」
七月四日
一九八七
國建言，
近來，美
本刊編輯
闢專欄「
在這個專
「香港的
至於美國
海外採訪
會面。
不期而遇
除了僑胞
我們來自
「台灣很
我們當然
中美關係
「他兒子
「這批貨
「電影你

Cover
Is the view of Sino-American relations from Taipei and Washington the same? (design by Walter Shih, photo by Vincent Chang)
Editor's Note
This issue has a particularly large number of overseas reports.
Two groups of Sinorama reporters recently went to the U. S. May is the pleasantest month in Washington, D. C., and the time when Twin Oaks is loveliest

ved by
yone
e 1,300
ent.
rted on
carried
rseas
rt is the
ger
m
n in
reign"
our
lented
oment
an
ctor of

North American affairs, stated in a public speech that U. S.-R. O. C. reltions are "friendly and close," the best they have been in the past ten years.
Just what is the current state of relations?
What does Fredrick Chien, on the front line of contacts between the two countries as the ROC's representative to the U. S., have to say about them?
And how has the public felt about them over the years?

Alignments



Word/Phrase Pairs

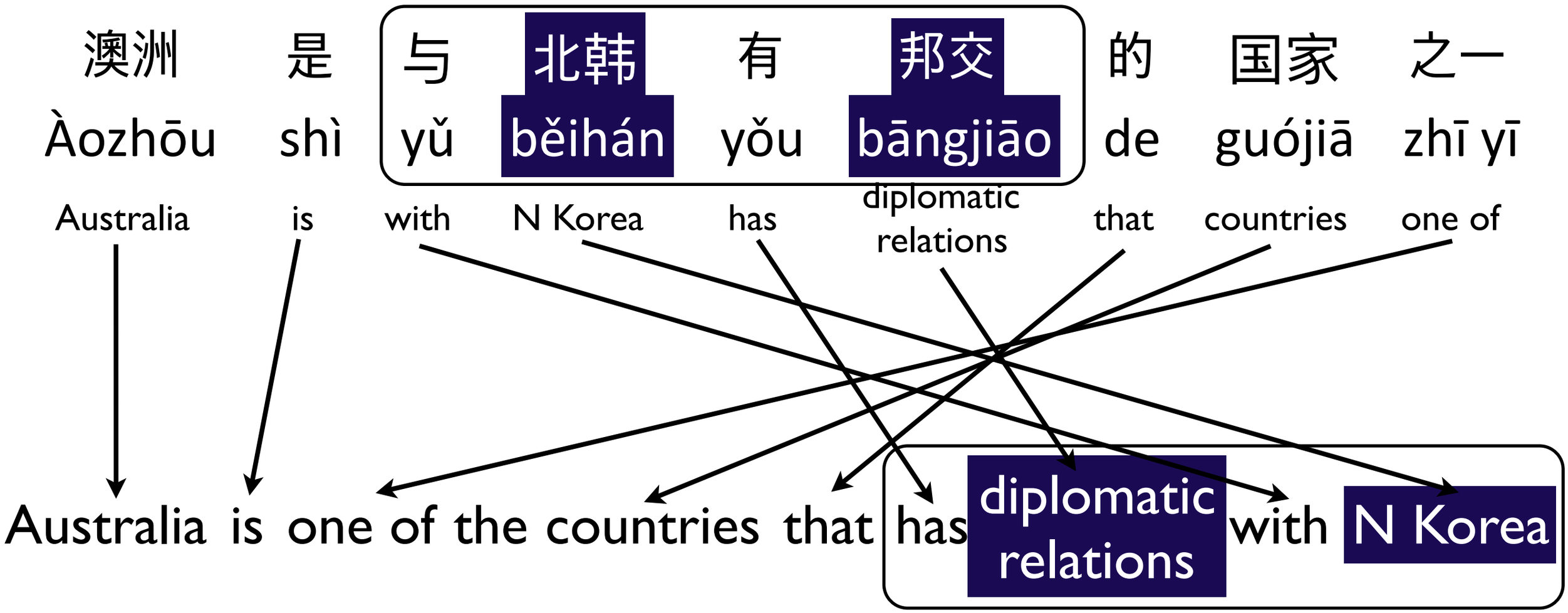
yǒu :: has

yǔ :: with

běihán :: N Korea

bāngjiāo :: diplomatic relations

Hierarchical Phrase Pairs



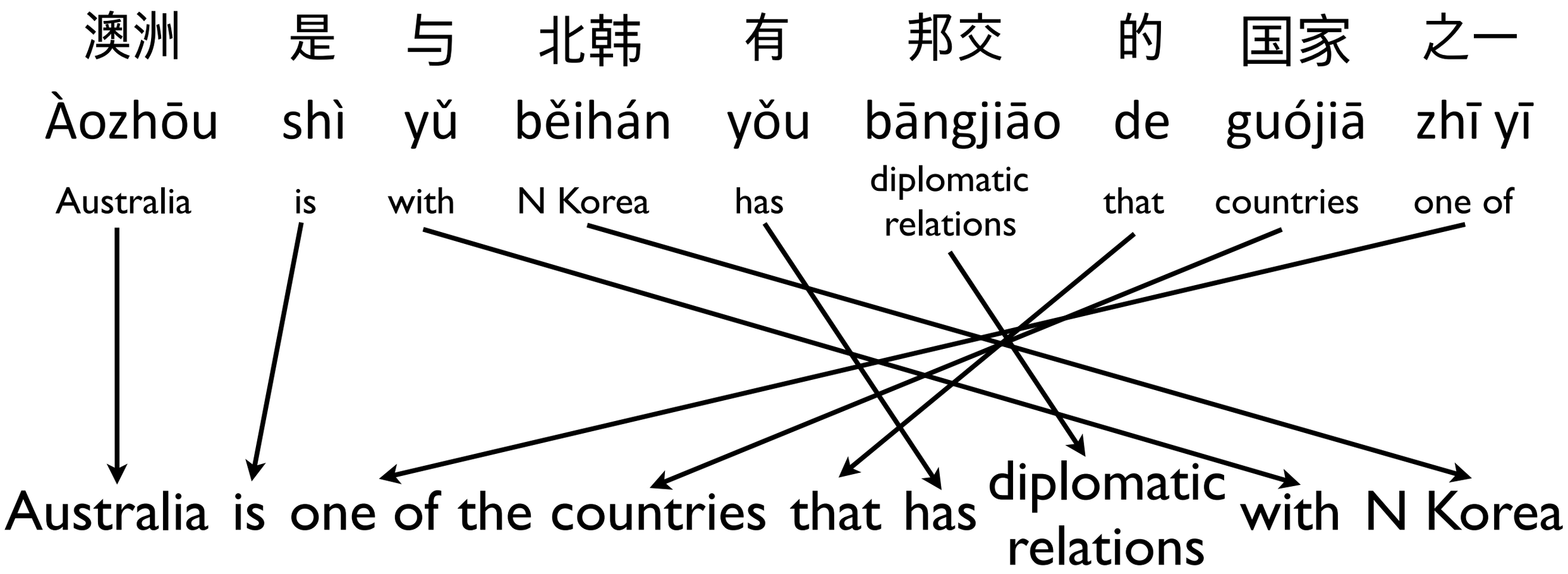
yǔ X1 yǒu X2 :: has X2 with X1

běihán :: N Korea

bāngjiāo :: diplomatic relations

yǔ běihán yǒu bāngjiāo, has diplomatic relations with N Korea

Hierarchical Phrase Pairs



yǔ X1 yǒu X2 :: has X2 with X1

běihán :: N Korea

bāngjiāo :: diplomatic relations

yǔ běihán yǒu bāngjiāo, has diplomatic relations with N Korea

Hierarchical Phrase Pairs

yǔ X1 yǒu X2 :: has X2 with X1 0.5

běihán :: N Korea 1.5

bāngjiāo :: diplomatic relations 3.0

yǔ běihán yǒu bāngjiāo, has diplomatic relations with N Korea

yǔ X1 yǒu X2 :: with X1 having X2 1.5

yǔ běihán yǒu bāngjiāo, with N Korea having diplomatic relations

Lower scores are better!

Next Problem: Efficient **Decoding**.

Given source language sentence find the best target language sentence.

Input: yǔ běihán yǒu bāngjiāo.

Output: ?

0.5 yǔ X_1 yǒu X_2 :: have X_2 with X_1

⊗

1.5 běihán :: N Korea

⊗

3.0 bāngjiāo :: diplomatic relations

⊕

1.5 yǔ X_1 yǒu X_2 :: with X_1 having X_2

⊗ 1.5 ⊗ 3.0

Semi-ring

⊕ ⊗

$(\mathbb{R} \cup \{\infty\}, \min, +, \infty, 0)$

1.have dipl relns with N Kor = 5.0

2.with N Kor having dipl relns = 6.0

- min(5.0, 6.0) option #1 wins

Àozhōu	shì	yǔ	běihán	yǒu	bāngjiāo	de	guójiā	zhī yī
澳洲	是	与	北韩	有	邦交	的	国家	之一

Which output is best?

yǔ běihán yǒu bāngjiāo

has diplomatic relations with N Korea — 11.5 ✓

with N Korea having diplomatic relations — 21.5

N Korea has diplomatic links — 60.0

A good scoring function can tell us ...

How to ensure proper scores?

yǔ běihán yǒu bāngjiāo

has diplomatic relations with N Korea ??

with N Korea having diplomatic relations ??

N Korea has diplomatic links ??

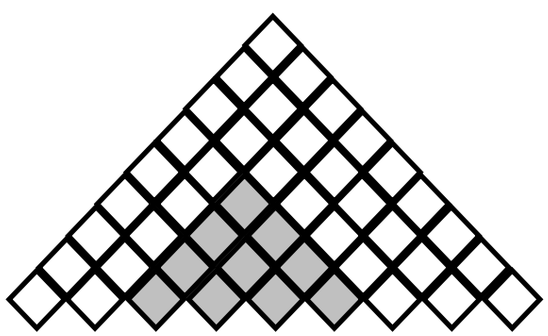
Use properties of the output translation ...

Using properties of the translation

yǔ X1 yǒu X2	::	has X2 with X1
běihán	::	N Korea
bāngjiāo	::	diplomatic relations

Features!

literal meaning?	2.0
fluency?	5.0
alignment scores	17.5
phrase count	3.0
…	



with N Korea having diplomatic relations

Using properties of the translation

Features!

f1: 2.0

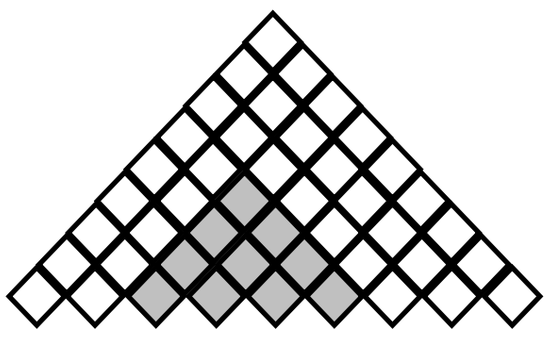
f2: 4.0

...

Each translation is scored with a weighted sum:

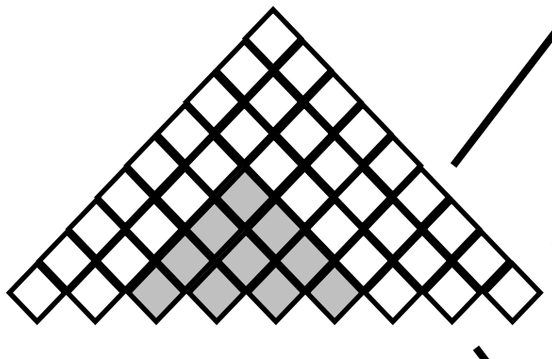
-2 * + 3 * = 8

Weights



with N Korea having diplomatic relations

Translations are feature vectors



has diplomatic relations with N Korea

3 8

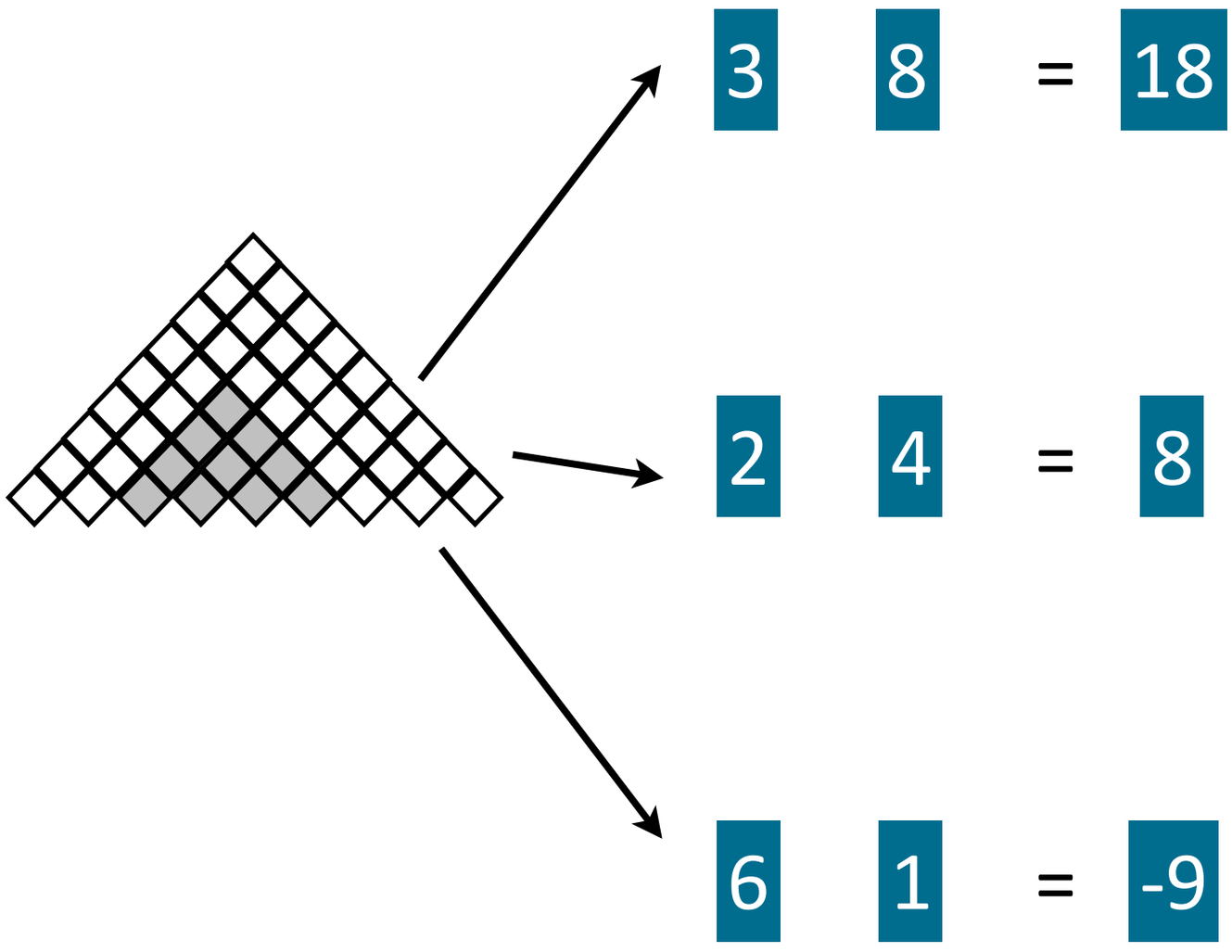
with N Korea having diplomatic relations

2 4

N Korea has diplomatic links

6 1

Weight vector determines the score



Model score

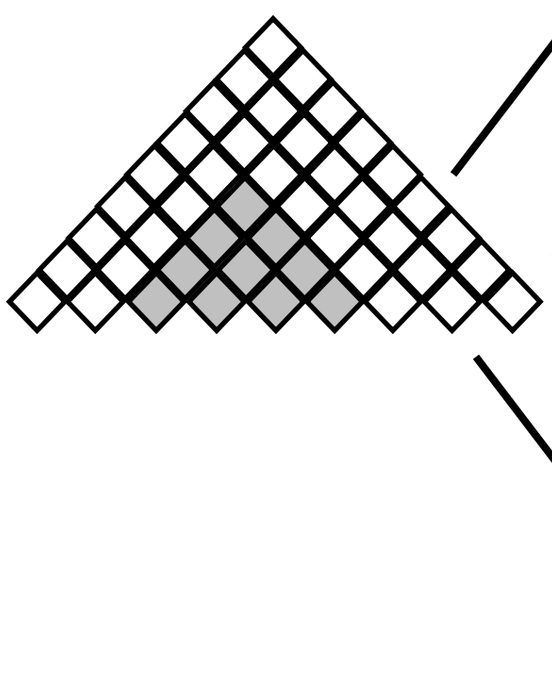
$$\mathbf{h} = \mathbf{f} \circ \mathbf{w}$$

Features

Weights

Features

Model score



3 8 = -18

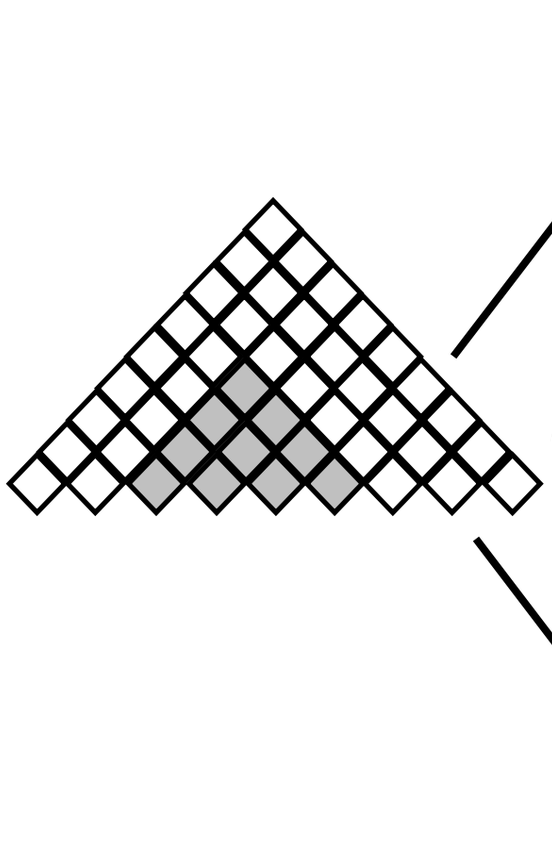
2 4 = -8

6 1 = 9

Weights **w**: 2 -3

Features

Model score



3 8 = 9

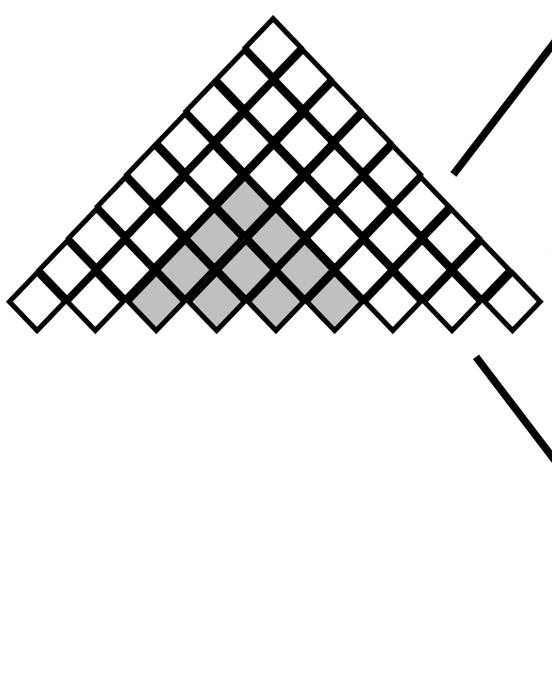
2 4 = 2

6 1 = -27

Weights **w**: -5 3

Features

Model score



3 8 = 9

2 4 = 2

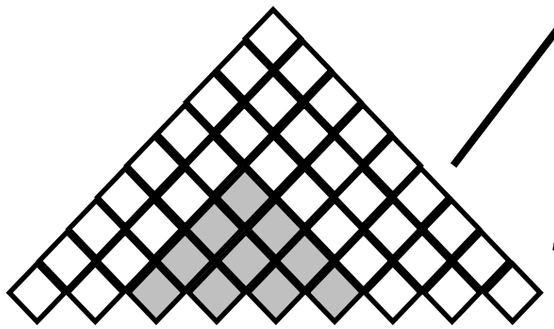
6 1 = -27

Weights

w: -5 3

Discriminative training is all about choosing this weight vector

Features | Model score | Extrinsic score



3 8 = 9 (3) .44 (1)

2 4 = 2 (2) .28 (2)

6 1 = -27 (1) .12 (3)

Bad Match!

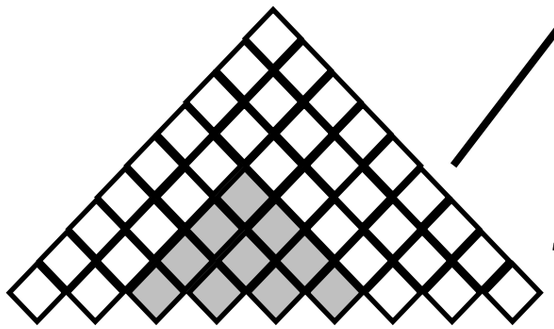
Ranking

Weights

w: -5 3

Choose a weight vector that matches the extrinsic score

Features | Model score | Extrinsic score



3 8 = 9 (3) .44 (1)

2 4 = 2 (2) .28 (2)

6 1 = -27 (1) .12 (3)

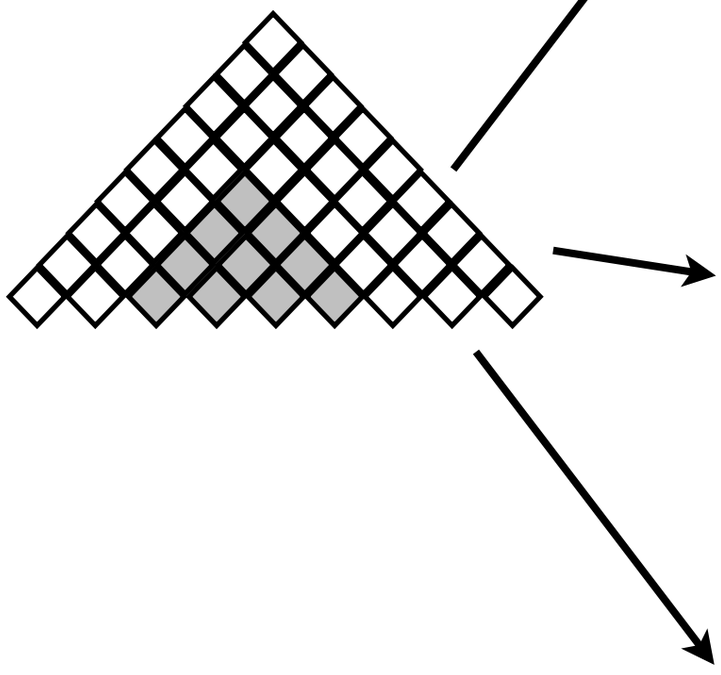
Bad Match!

Ranking

Weights

w: -5 3

Choose a weight vector that matches the extrinsic score



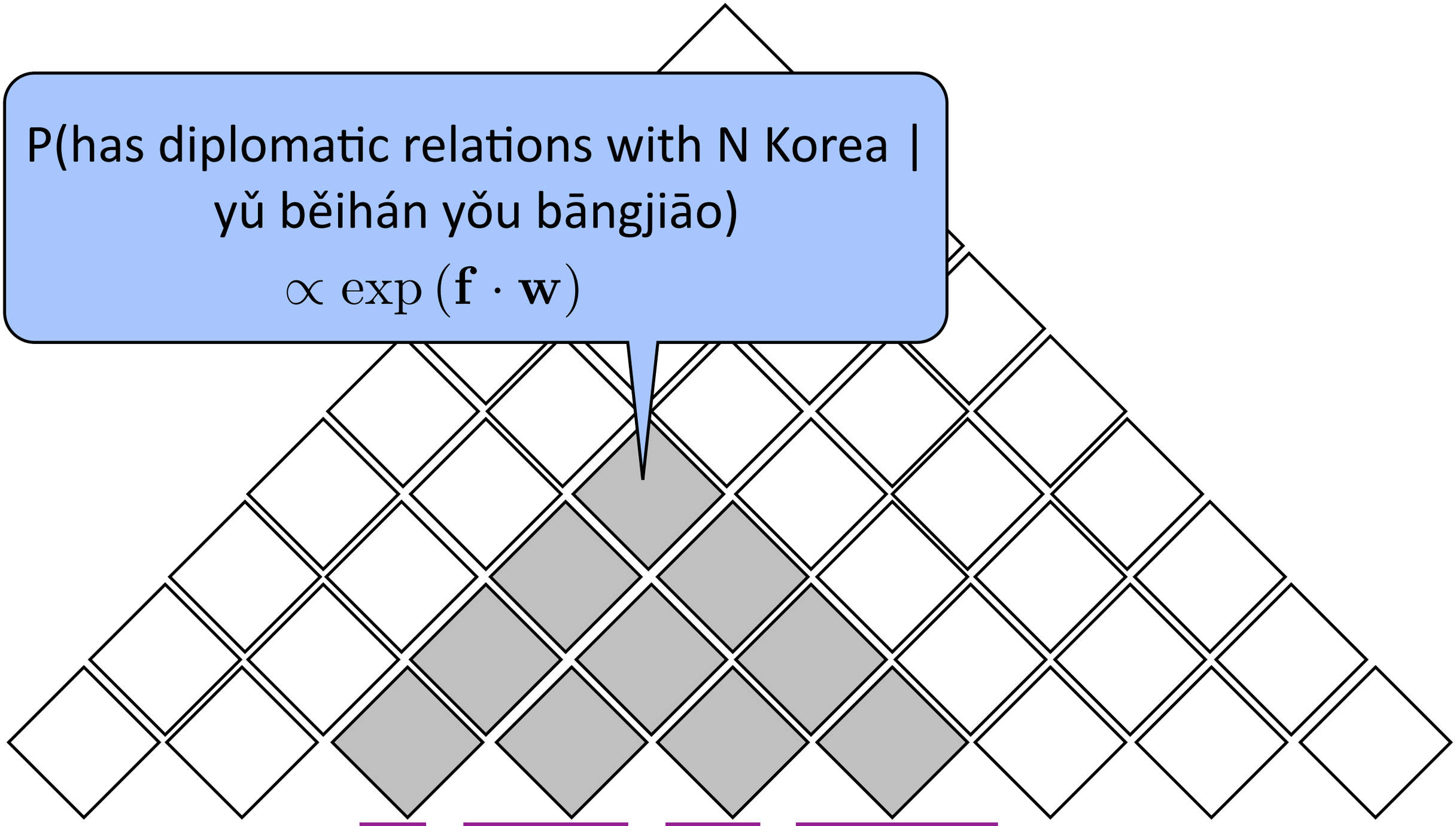
	Features	Model score	Extrinsic score	
(1)	3 8	= -2	.44	(1)
(2)	2 4	= 3	.28	(2)
(3)	6 1	= 11	.12	(3)

Good Match!

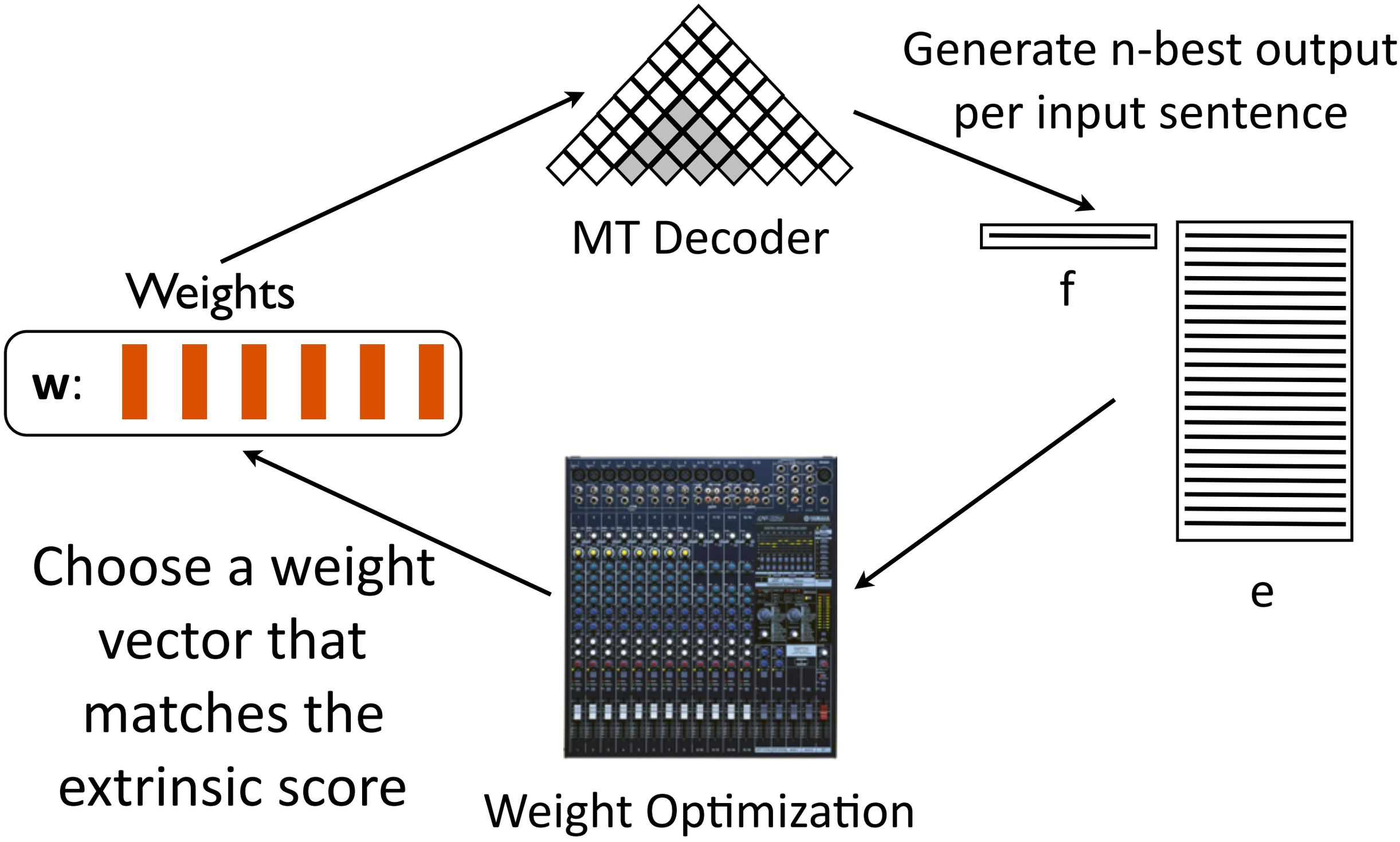
Weights **w**: 2 -1

Choose a weight vector that matches the extrinsic score

MT Decoder


P(has diplomatic relations with N Korea | yǔ běihán yǒu bāngjiāo)
$\propto \exp(\mathbf{f} \cdot \mathbf{w})$
Àozhōu shì yǔ běihán yǒu bāngjiāo de guójiā zhī yī
澳洲 是 与 北韩 有 邦交 的 国家 之一

Discriminative Training Framework



Automatic Evaluation of Translation

Automatic Evaluation of Translation

虽然	北	风	呼啸	,	但	天空	依然	十分	清澈
Suīrán	běi	fēng	hūxiào	,	dàn	tiānkōng	yīrán	shífēn	qīngchè
Although	north	wind	howls	,	but	sky	still	very	clear

Although a north wind was howling , but the sky remained clear

However , the sky remained clear under the strong north wind

Despite the strong northernly winds , the sky remains clear

The sky was crystal clear , though the north wind was howling

Although a north wind was howling , the sky remained clear and blue

Automatic Evaluation of Translation

- BLEU (B)
  - Scores precision of n-gram matches to reference
  - Also penalizes system output that is shorter than reference
- TER (T)
  - # edits / # of reference words
  - edits are indels and block shifts (NP-hard)

Automatic Evaluation of Translation

- RIBES (R)

  - Aligns system output to reference to find which words/phrases are out of order.

  - Reports the Spearman's correlation coefficient.

- METEOR (M)

  - Like BLEU, but

  - Computes precision and recall

  - Also matches synonyms and stems