

Statistical Machine Translation with Long Phrase Table and without Long Parallel Sentences

□ Jin'ichi Murakami, Masato Tokuhisa, Satoru Ikehara
(Tottori University, Japan)

The Strategy of Our Statistic Machine Translation

1) Long Phrase Tables (Adequacy)

Adequacy ~ translation model $P(E/C)$

long phrase tables = achieve high accuracy
20 words

English to Italian

Word position change is very small.

☐ short phrase table

Japanese to English

Some word are moved from their original position.

☐ long phrase tables.

The Strategy of Our Statistic Machine Translation

2) 5-gram Model (Fluency)

- Fluency \sim language model $P(E)$

Not use higher N-gram model.

(the reliability for each parameter becomes low)

normal 5-gram model

The Strategy of Our Statistic Machine Translation

3) Remove Long Parallel Sentences

- Long Parallel Sentences ☐ Wrong Phrase Table
- ☐ Low Blue Score
- Many Parallel Sentences ☐ High Translation?

The Strategy of Our Statistic Machine Translation

4) Standard Tools

GIZA++.2003-09-30.tar.gz

moses.2007-05-29.tgz

training-release-1.3.tgz(train-phrase-model.perl)

(Made only some small tools to build a temporal corpus.)

Wrong Phrase Table

図 3 及び 図 4 に 示す よう に As shown in FIGS . 6 and 7

0.047619 4.03037e-09 0.0243902 1.52103e-11

、それ 以外 4 , while all other

1 3.84583e-05 0.0217391 4.24439e-09

1 3 は 1 is a

0.0010582 0.000901274 0.000698568

1 3 は 1

4.97396e-06 7.91046e-05 0.000349284 0.0497444

コンデンサ a

0.333333 0.0001482 9.04732e-05 1.01046e-06

コンデンサ i

0.000857633 0.0001482 9.04732e-05 2.10078e-06

マイクロコンピュータ merely

0.00263852 0.000423 0.000176398 0.0001555

図30図に示す実施例は、路面のセンターに1本のガイド5を敷設し、そのガイド5を車体15に取り付けたガイドローラ3により挟み込んで支持するように構成したものであり、前方又は後方から観た状態を示している。

In the embodiment shown in FIG. 30, one guide 5 is laid at the center of the road surface, and the chassis 15 is supported by clamping the guide 5 by the guide rollers 3 attached to the chassis 15, the view being taken from the front or rear of the vehicle.

また、車軸9は、ダブルウィッシュボーン式のリンク14により車体15に支持されていると共に、コイルスプリング16及びダンパー(不図示)により衝撃が吸収緩和されるようになっている。

The axle 9 is supported on a chassis 15 by means of a double wish-bone type link 14, and shocks are absorbed and alleviated by a coil spring 16 and a damper (not shown).

3輪車の場合は、その構造から明らかなようにセンターに車輪があるため、図30に示す実施例は適用できず、車両の両外側に2本のガイド5を敷設した図34が最も適用しやすい構成である。

In the case of the three-wheeled vehicle, since one wheel is present at the center as is apparent from its structure, the embodiment shown in FIG. 30 is not applicable, and FIG. 34 in which two guides 5 are laid on the opposite outer sides of the vehicle is easiest to apply.

Remove Long Parallel Sentences

Japanese-English Parallel (> 128 char) : 736080 sentences

Decoder : Moses (moses.ini)

ttable-limit 40 0

weight-d 0.1

weight-l 1.0

weight-t 0.5 0.0 0.5 0.0 0.0

(Cross Entropy)

weight-w -1

distortion-limit -1

(The position of the verb changed significantly)

Results

- Effective for JE
- Not effective for EJ

task	method	bleu	Human Adequacy	evaluation Fluency
Intrinsic-JE	Proposed	22.29	2.58	3.44
	standard	21.62		
Intrinsic-EJ	Proposed	23.93		
	standard	25.33		
		MAP	Recall relax	
Extrinsic-EJ			0.77	

Proposed: Remove Long Parallel Sentences (> 128 char)

1062491 Japanese-English parallel sentences

Standard: Standard Moses:

1798571 Japanese-English parallel sentences

Outputs J-E (Proposed)

Input :図5は回転羽根2を駆動するモータの構成例を示す図である。

Output : FIG. 5 is a diagram showing an example of the configuration of a motor for driving the rotary blade 3.

Correct :G. 5 is a diagram showing a structural example of a motor for driving the rotating blade

Input :さらに、心線ワイヤ51の先端部がラミネートフィルム59により挟まれて保持され、その変形、ピッチの狂いを防止する。

Output :Further, the leading end portion of the core wire 51, and held in the modification of the pitch deviation between the laminated film adhering 59 can be prevented. \

Correct :Moreover, the front ends of the core wires 51 are sandwiched with laminated films 59 to prevent deformation of the core wires 51 for the purpose of maintaining their relative positions intact.

Input :この絶縁ハウジング10の外面に取り付けられるシールドカバー30を図6に示している。

Output :The insulating housing 10 is mounted to an outer surface of the shield cover 30 shown in FIG. 6.

Correct :FIG. 6 shows the shield cover 30, which is to be mounted on the insulative housing 10.

Outputs J-E (Standard)

Input :図5は回転羽根2を駆動するモータの構成例を示す図である。

Output :The rotator 2 driven by a motor structure shown in FIG. 5.

Correct :G. 5 is a diagram showing a structural example of a motor for driving the rotating blade 2.

Input :さらに、心線ワイヤ51の先端部がラミネートフィルム59により挟まれて保持され、その変形、ピッチの狂いを防止する。

Output :Further, a distal end portion of the core wire 51 is clamped and held by the pitch deviation of the laminated film 59, the deformation is prevented.

Correct :Moreover, the front ends of the core wires 51 are sandwiched with laminated films 59 to prevent deformation of the core wires 51 for the purpose of maintaining their relative positions intact.

Input :この絶縁ハウジング10の外面に取り付けられるシールドカバー30を図6に示している。

Output :The insulating housing 10 attached to the outer surface of the shield cover 30 is shown in FIG. 6.

Correct :FIG. 6 shows the shield cover 30, which is to be mounted on the insulative housing 10.

Conclusions

- Remove Long Parallel Sentences
 - Long phrase table
 - Standard tools
 - Statistical Example Based Translation
-
- Good results for J-E
 - 0.2229 BLEU score

Future study

- Optimize parameters
- Unknown word procedure
- More large database
- Not used parallel sentence
(If output likelihood is high,
use as parallel sentence)