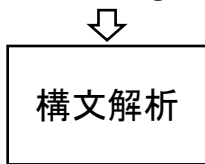
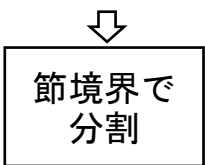


## 処理概略図

I think it is good .



(S I think (S it is good )S . )S



I think

It is good

.

学習済み  
Transformer1

学習済み  
Transformer1

学習済み  
Transformer1

私は思う。

それは良い。

。

I think @@@ 私は思う。 ||| it is good.  
@@@それは良い ||| . @@@。

学習済み  
Transformer2

私は良いと思う。

## 問題

既存のニューラル機械翻訳では、長文翻訳が難しい。

## 結論

- コーパスBLEUにおいては顕著な差は見られないが、RATIOは改善している。
- 分割の仕方や、並べ替えの手法のさらなる検討が必要。

## 解決策

短く区切って翻訳し、それらを並べ替えて繋げる。

## 実験

### コーパス

- ASPEC 英日

train	valid	test
1,000,000	1790	1812

### サブワード分割

- SentencePiece
- 語彙サイズ:16,000
- 言語間で語彙共有

表1 入力トークン長ごとの BLEU

モデル	all	1-20	21-40	41-60	61-
Transformer	<b>41.7</b>	40.8	41.5	<b>43.5</b>	39.1
Linearized tree	41.6	41.0	<b>41.8</b>	42.7	39.1
Not reordering	28.8	32.4	28.2	28.8	26.4
Proposed	<b>41.7</b>	<b>41.3</b>	41.4	<b>43.5</b>	<b>39.9</b>

表2 入力トークン長ごとの ratio

モデル	all	1-20	21-40	41-60	61-
Transformer	0.960	<b>0.944</b>	0.956	0.971	0.975
Linearized tree	0.958	0.939	0.956	0.970	0.974
Not reordering	1.272	1.136	1.276	1.343	1.301
Proposed	<b>0.962</b>	0.943	<b>0.958</b>	<b>0.977</b>	<b>0.979</b>

表3 モデルごとの出力例

モデル	出力例
英語の節分割	The measures for regional improvement aiming at expansion of the interchanging zone area /, and / promotion and creation of the industry in the area were examined , / taking advantage of the time that / National Route 361 connecting Inadani and Kisodani was approved in the business as a high standard road / .
参照訳	伊那谷と木曾谷を結ぶ国道 361 号線が高規格道路として事業認可されたのを機に、交流圏域の拡大と域内産業の振興・創出をめざす地域整備の方策を検討した。
節の翻訳結果	交流圏域の拡大を目指した地域整備の方策。 / 又、及び、また、そのためのものである。 / 地域における産業の振興と創出について検討した。 / 時間の利点を活かした。 / インダニと木曾谷を結ぶ国道 361 号が高規格道路として事業で認可された。 / その他。
Transformer	インターチェンジ・ゾーンエリアの拡大を目指した地域整備の方策と、インターチェンジ・ゾーンエリアを結ぶ国道 361 号を高規格道路として、ビジネスで認可された時期を背景に、エリアの産業の振興・創造について検討した。
Proposed	インダニと木曾谷を結ぶ国道 361 号が高規格道路として事業に認可された時の利点を活かし、交流圏域の拡大、産業の振興と創出を目指した地域整備の方策を検討した。