

# 双方向学習と再現学習を統合したニューラル機械翻訳



Innovative R&D by NTT

森下 睦<sup>1</sup>, 鈴木 潤<sup>1,2</sup>, 永田 昌明<sup>1</sup>

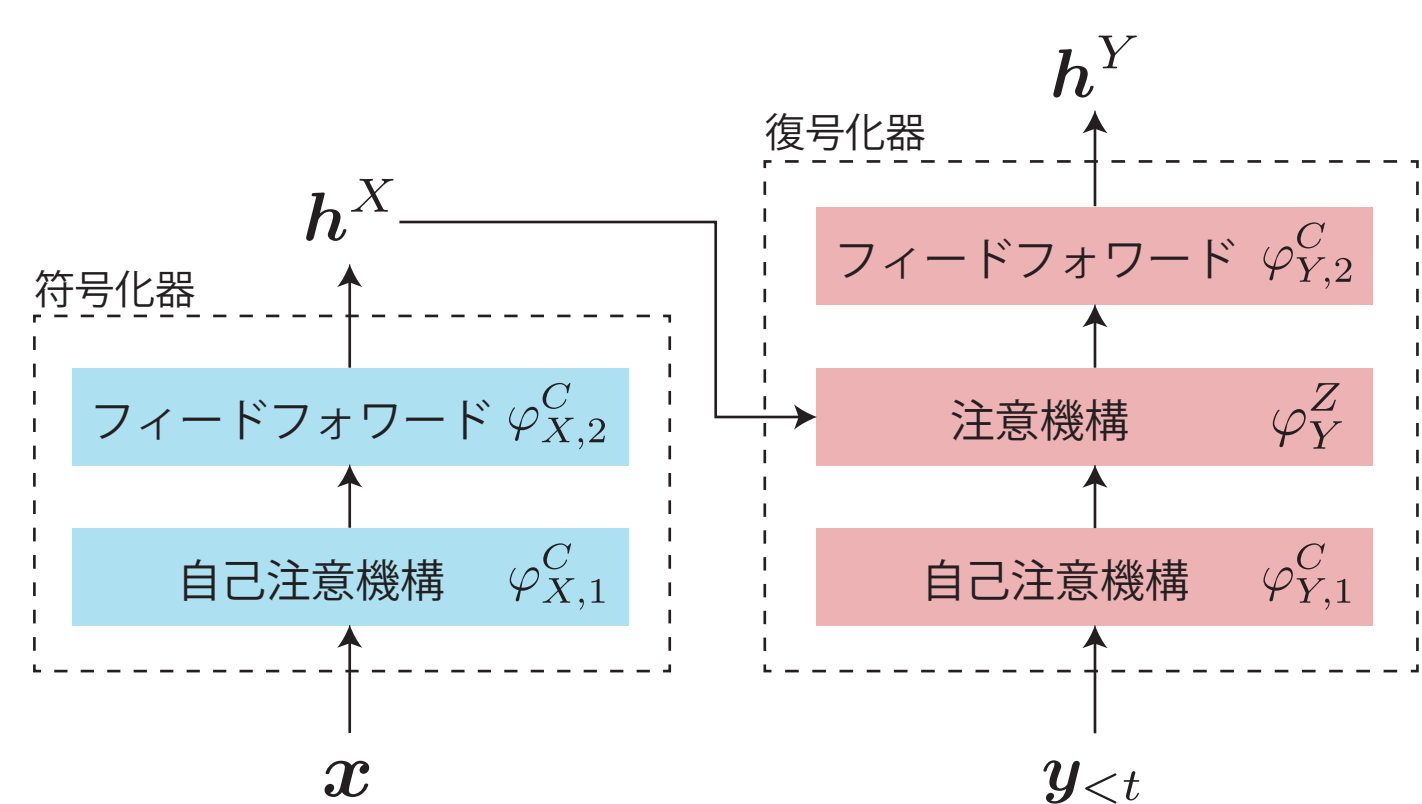
<sup>1</sup> NTT コミュニケーション科学基礎研究所, <sup>2</sup> 東北大学

## 概要

- ・ニューラル機械翻訳は時折**重要な単語を欠落する**など、**原文に忠実でない**ことがある。
- ・**再現学習**: 復号化器の隠れ層から原文を復元する際の誤差を用いてモデルを学習。  
→ ただし、再現器追加に伴う**パラメータ増加**
- ・**双方向学習**: 両翻訳方向を同時に学習することで**精度向上**かつ**パラメータ削減効果**あり
- ・提案法では、双方向学習と再現学習を統合することにより、**新たにパラメータを追加することなく**再現学習を可能に。実験により**統計的に有意な翻訳精度向上**が見られた。

## Transformer

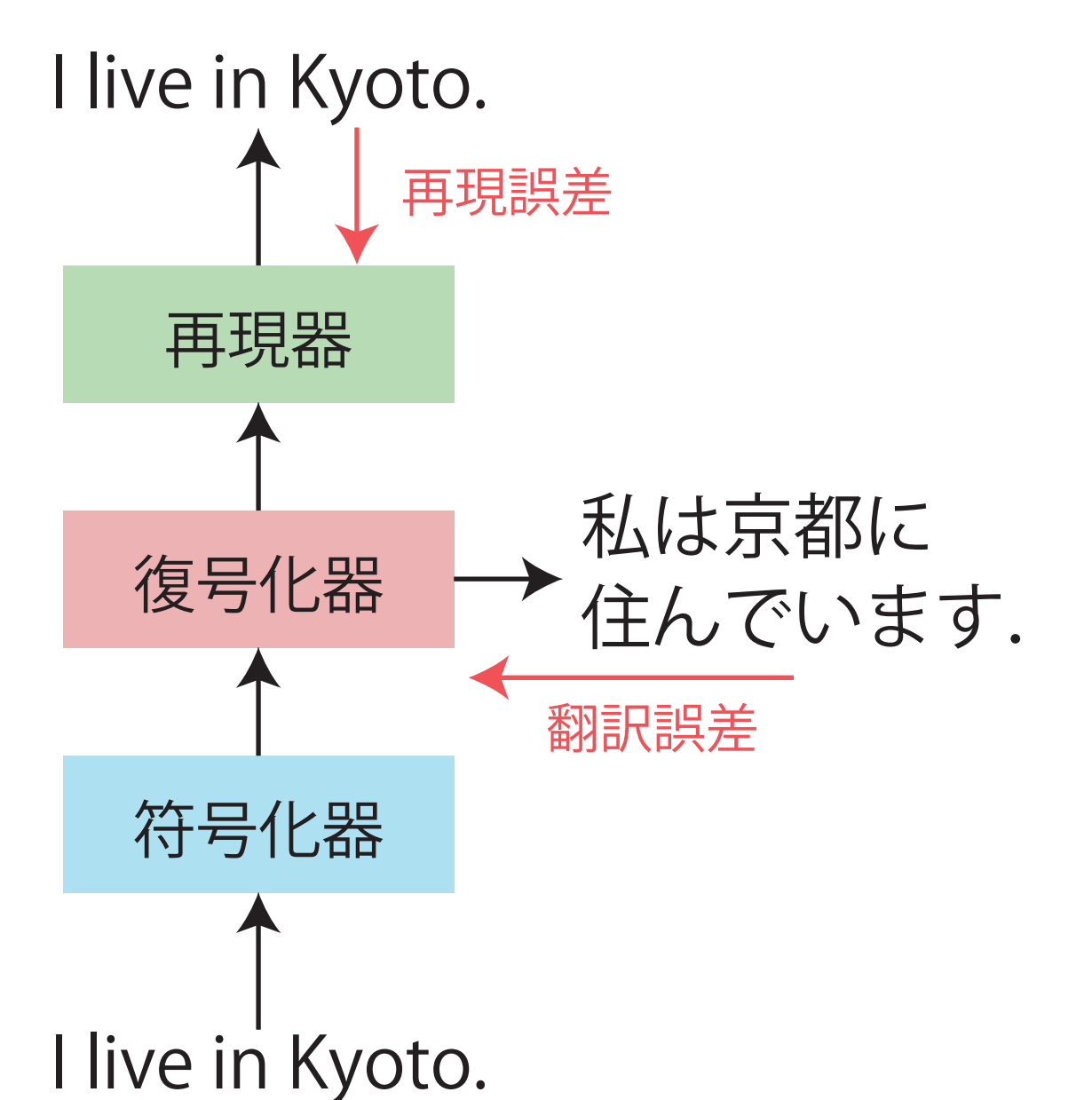
- ・2017年にGoogleによって提案された新たな系列変換モデル
- ・**高い精度が達成できる**ため、ニューラル機械翻訳では近年主流のモデル



Vaswani et al., Attention Is All You Need, NIPS 2017

## 再現学習

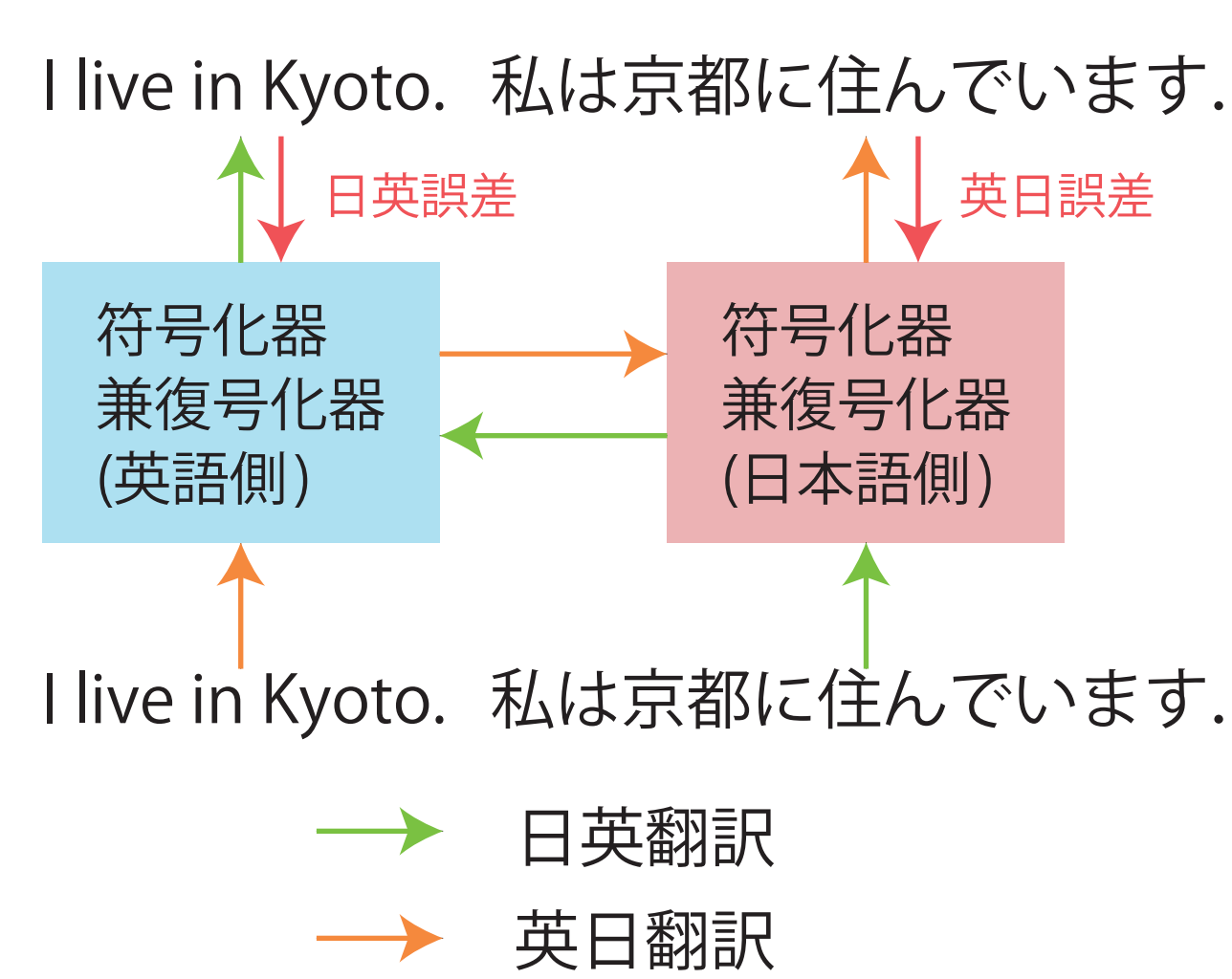
- ・復号化器の隠れ層を基に**原文を復元した際の誤差**をモデル学習に利用  
→ **より原文に忠実な訳**が生成されることを期待  
→ ただし、再現器に関する**新たなパラメータが必要**



Tu et al., Neural Machine Translation with Reconstruction, AAAI 2017

## 双方向学習

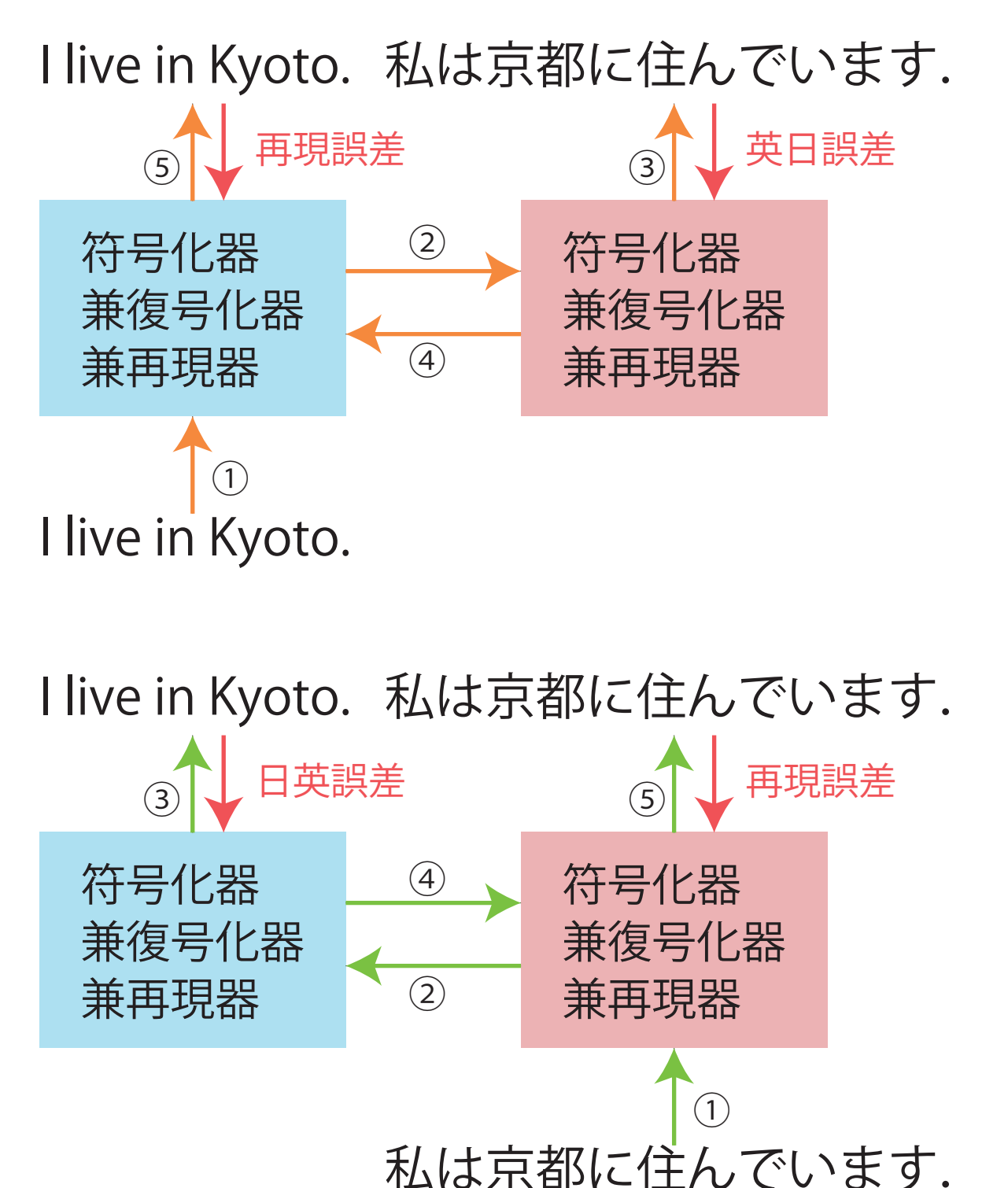
- ・**両翻訳方向を同時に学習**することで、翻訳精度向上を狙う
- ・Xiaらの手法では同一言語の**符号化器と復号化器のパラメータを共有**
- ・通常モデルより**パラメータ数は削減**かつ正則化の効果があり**翻訳精度は向上**



Xia et al., Model-Level Dual Learning, ICML 2018

## 提案法

- ・**双方向学習と再現学習を統合**
- ・入力言語側の**符号化器兼復号化器を再現器として使用**
- ・**パラメータ数を増加させずに再現誤差を導入可能**



## 実験

### 実験設定

- ・NMTモデル: Transformer
- ・TEDトークコーパス(独↔英, 仏↔英)

	独↔英	仏↔英
Train	189,318	208,323
Dev	888	890
Test	3,998	3,455

コーパスに含まれる対訳文数

- ・再現器無しで15万ステップ学習  
→ その後再現器を追加し合計25万ステップ学習

### 翻訳精度

	英→独	独→英	英→仏	仏→英
片方向学習 (ベースライン)	25.57	30.77	39.79	37.74
+ 再現器	28.04 **	33.72 **	41.99 **	40.62 **
双方向学習	29.38	34.78	42.86	42.14
+ 再現器 (提案法)	<b>29.55</b>	<b>34.83</b>	<b>43.46 **</b>	<b>42.42 *</b>

ベースラインおよび提案法の BLEU スコア. 再現器なしの結果と比べて,

\* は 5% 水準で, \*\* は 1% 水準で統計的に有意な差があることを示す.

### パラメータ数

提案法では逆側言語の復号化器を再現器として使用するため、再現器を追加することによる**パラメータ増加は無い**

	パラメータ数
片方向学習 (ベースライン)	46.45M * 2方向 = 92.9M
+ 再現器	65.35M * 2方向 = 130.7M
双方向学習	52.7M
+ 再現器 (提案法)	52.7M