

# 高性能なオートエンコーダとのマルチタスク学習を利用した ニューラル機械翻訳

根石将人 (東京大学) 吉永直樹 (東京大学生産技術研究所)

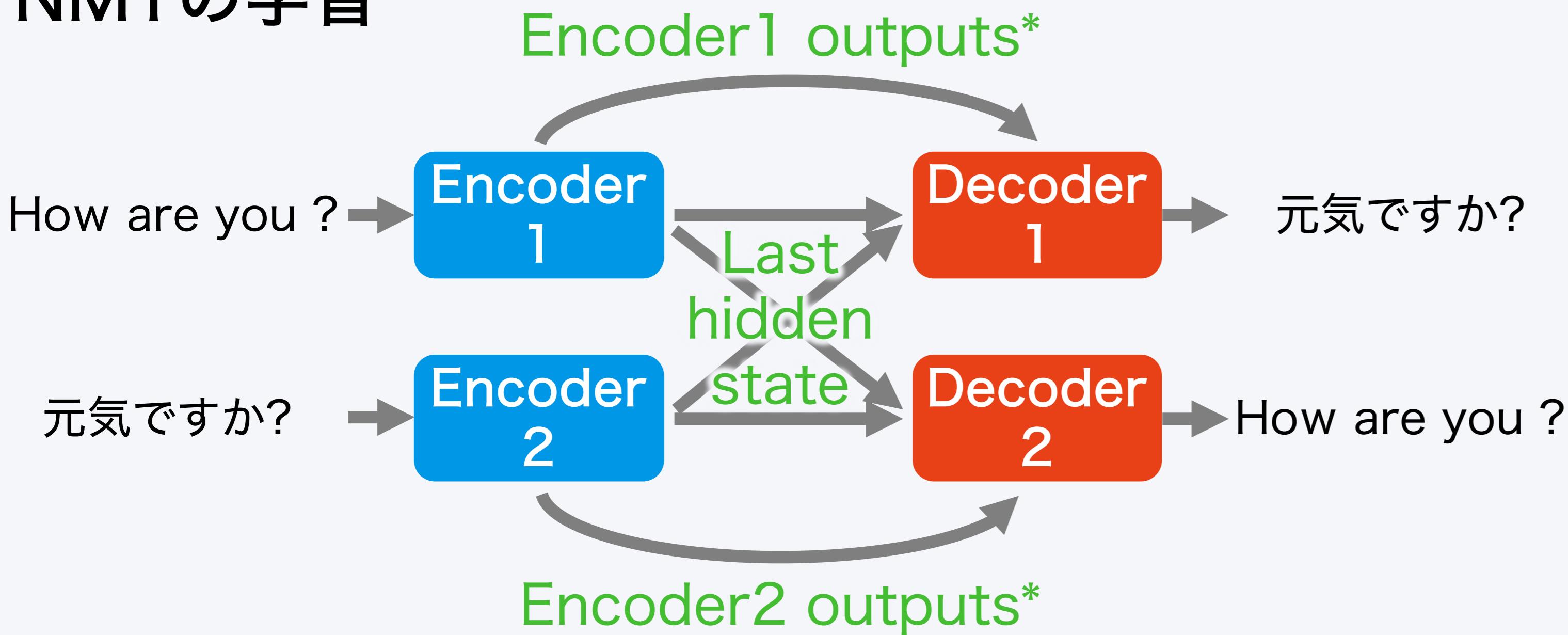
## 概要

- [研究背景]** • 深層学習において、関係する他のタスクとのマルチタスク学習は、学習データの増加やモデルの汎化性能の向上により目的タスクでの性能を高める  
• ニューラル機械翻訳 (NMT) は、言語モデル[1]、品詞解析、オートエンコーダ[2]などとのマルチタスク学習により翻訳性能が向上

- [研究目的]** NMTとオートエンコーダとのマルチタスク学習において、オートエンコーダの復元性能を高めることで、文のトークン列と分散表現間の変換能力を高め、同時学習するNMTの翻訳性能を高める

## 提案手法

### オートエンコーダとのマルチタスク学習に基づく NMTの学習



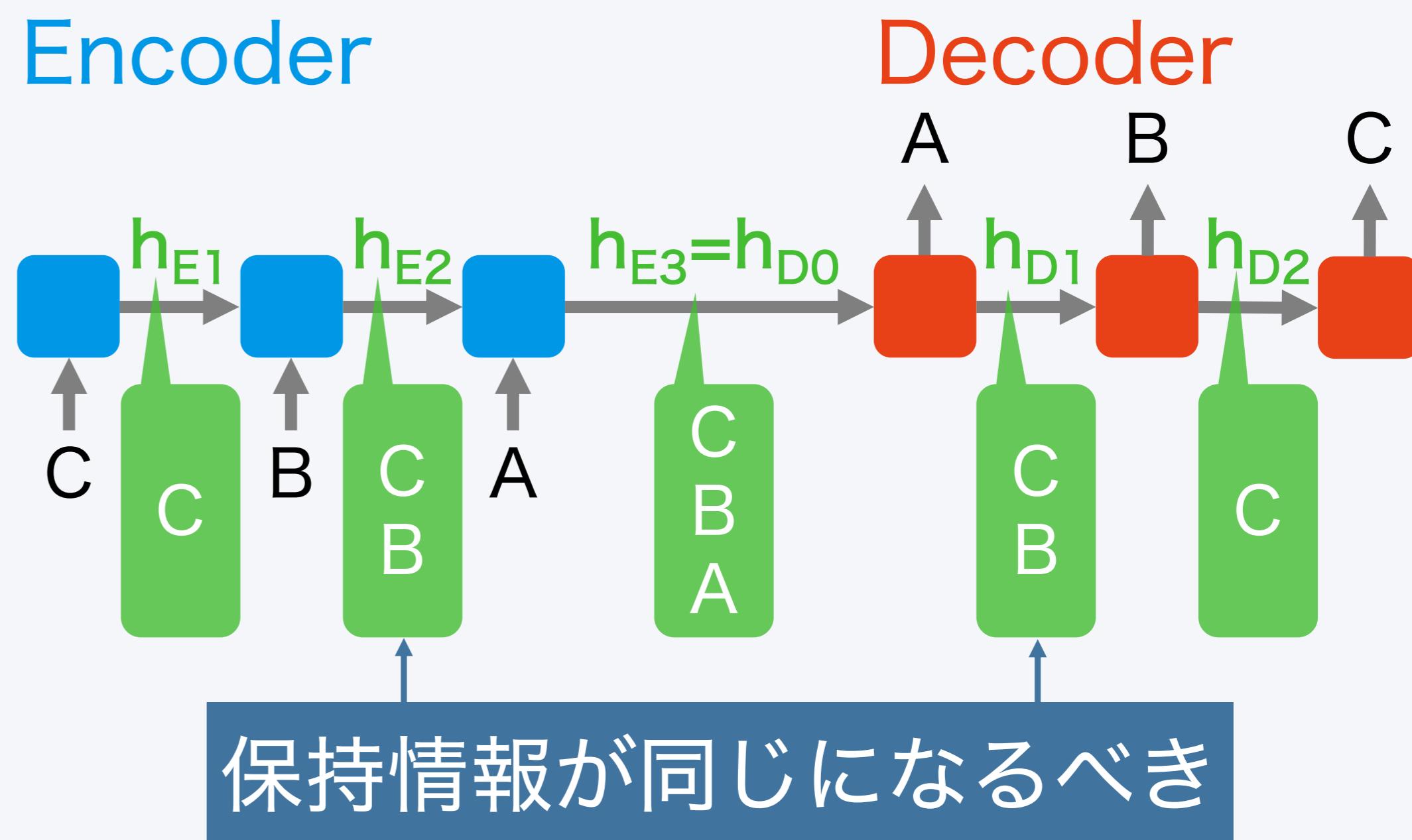
\*アテンション以外の部分の学習を目的としたため、オートエンコーダにはアテンションは不使用

- エンコーダ、デコーダの組み合わせにより、双方向の翻訳、両言語のオートエンコーダを内包し、これらを全て同時に学習する  
- 損失関数

cross entropy loss(trans1, trans2, autoenc1, autoenc2)  
+ stepwise autoenc loss(autoenc1, autoenc2)

※[Luong+ 2016]では片方の翻訳と両言語のオートエンコーダの3つを同時学習  
また、オートエンコーダにもアテンションを使用

### 隠れ状態の対応関係に着目したオートエンコーダ



- 入力の逆順で出力することで、各時刻のエンコーダとデコーダの隠れ状態  $h$  が保持すべき情報に対応関係
- 保持情報が同じになるべき隠れ状態同士を近づけることで目的の単語を出力しやすくする

#### ◆ 対応する隠れ状態を用いた平均二乗誤差 (MSE)

$$\text{stepwise autoenc loss} = \frac{1}{L_x - 1} \sum_{i=1}^{i=1} (h_{Ei} - h_{D(L_x-i)})^2$$

## 実験

### 目的

提案する損失関数の以下の2つへの影響の確認

- オートエンコーダの復元性能
- NMTとのマルチタスク学習時の翻訳性能

### 実験設定

- データセット
  - ASPEC英日対訳 (Train: first 1.5M, Dev: 1790, Test: 1812 sents)
  - 語彙数: 16,000 (英日共通語彙、SentencePiece(unigram)によるトークン分割)
- モデル
  - 埋め込み層: 512次元
  - エンコーダ・デコーダ: 2層512次元LSTM
  - アテンション: Luong(2015) General (マルチレイヤー)
- 学習設定
  - 最適化アルゴリズム: Adam (初期学習率0.005)
  - ミニバッチサイズ: 32

#### ◆ オートエンコーダ実験 (日本語→日本語)

| モデル                              | BLEUスコア | 完全復元した文の割合 |
|----------------------------------|---------|------------|
| Forward                          | 77.53   | 48.1%      |
| Backward                         | 83.61   | 56.8%      |
| Bi-directional                   | 82.78   | 67.5%      |
| Backward + stepwise autoenc loss | 92.19   | 76.8%      |

| モデル                              | 出力例                                                    |
|----------------------------------|--------------------------------------------------------|
| [Reference]                      | 科学論文の不正行為には、捏造、改ざん、データやアイデアの盗用のほかに、様々な形態、程度の行為がある。     |
| Bi-directional                   | 科学論文の不正行為には、UNK造、改ざん、データやアイデアのUNK用のほかに、様々な形態の程度、行為がある。 |
| Backward + stepwise autoenc loss | 科学論文の不正行為には、UNK造、改ざん、データやアイデアのUNK用のほかに、様々な形態、程度の行為がある。 |

- 提案手法のエラーは主に未知語や低頻度
- また、途中の不正解出力が及ぼす後方への悪影響が少ない傾向
- その他のモデルは、尤度は高そうだが元と微妙に異なる文を生成

#### ◆ 機械翻訳実験 (マルチタスク学習)

| モデル                                | 英日翻訳<br>BLEUスコア | 日日オートエンコーダ<br>BLEUスコア |
|------------------------------------|-----------------|-----------------------|
| Backward (Non multi-task)          | 26.50           | (無し)                  |
| Multi-task                         | 22.73           | 77.60                 |
| Multi-task + stepwise autoenc loss | 28.13           | 80.47                 |

| モデル                                | 出力例                                                                                                                             |
|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| [Source]                           | a database for development of superconducting material was constructed, and deduction system for material design was developed. |
| [Reference]                        | 超伝導材料開発のためのデータベースを構築し、材料設計用演算システムの開発を行った。                                                                                       |
| Backward (Non multi-task)          | 超電導材料開発用データベースを構築し、材料設計用のシステムを開発した。                                                                                             |
| Multi-task                         | 超電導材料の開発に資するためのデータベースを構築し、設計のためのシステムを開発した。                                                                                      |
| Multi-task + stepwise autoenc loss | 超電導材料の開発のためのデータベースを構築し、材料設計用のデカッシュシステムを開発した。                                                                                    |

- マルチタスクにより、尤度が高そうな誤った文を生成
- 提案手法は比較的、情報の欠落が抑えられている

## 今後の予定

- 大規模単言語コーパスを用いたオートエンコーダの事前学習
- 高性能なオートエンコーダの他への応用

## 参考文献

- [1] Ramachandran+, Unsupervised Pretraining for Sequence to Sequence Learning, EMNLP2017  
[2] Luong+, Multi-task Sequence to Sequence Learning, ICLR2016