



### 3. 文長制御のための条件付き言語モデル

#### 3.1. 条件付き言語モデル

条件付き言語モデルでは、学習時にモデルの入力の先頭に所与の制御トークンを挿入する必要がある。文長制御において、制御トークンは出力させたい文長の表現  $l$  である。よって、出力するトークン列を  $x$ 、文長の制御トークンを  $l$  とすると、条件付き言語モデル  $p(x|l)$  は以下のように表される。

$$p(x|l) = \prod_{i=1}^L p(x_i|x_{<i}, l) \quad (1)$$

#### 3.2. 文長符号化

文長を制御するために、本研究では、文長を離散値として埋め込むかわりに、位置符号化[4]を用いて文長の埋め込み表現を獲得する。位置符号化は、機械翻訳で提案されたモデルである Transformer において、各入力トークンの位置情報をベクトルに変換する手法である。本研究では、位置符号化を文長の分散表現獲得に応用した文長符号化を提案する。以下に文長符号化の式を示す。

$$LE_{(length, 2i)} = \sin\left(\frac{length}{10000^{\frac{2i}{d}}}\right) \quad (2)$$

$$LE_{(length, 2i+1)} = \cos\left(\frac{length}{10000^{\frac{2i}{d}}}\right) \quad (3)$$

## 4. 実験

### 4.1. 実験設定

提案手法の精度を評価するため、wikitext-103をデータセットとしてモデルの訓練し、実験を行う。実験では、事前学習済みの言語モデルを文長の条件付き言語モデルとしてファインチューニングすることにより、指定した長さで自然な文が生成できるかを測定する。

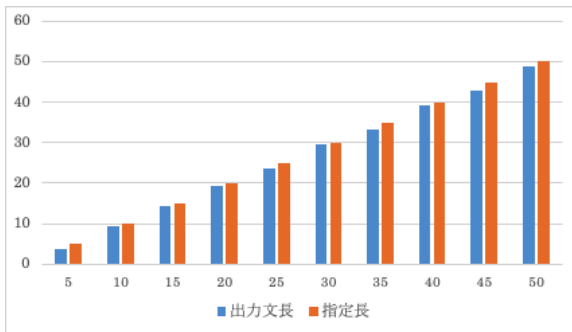


図2 実際の出力文長と指定長

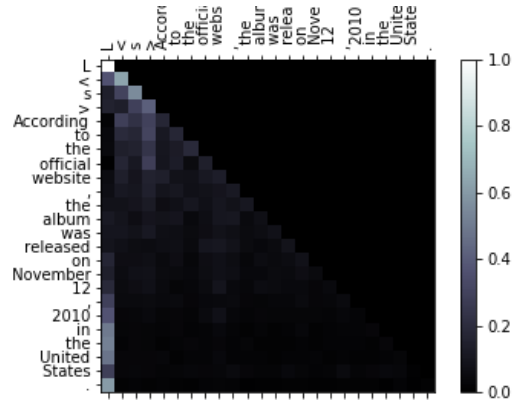


図3 文長制御時の最終層におけるアテンション

### 4.2. 実験結果

図2に文長制御の実験結果を示す。実験結果から提案した条件付き言語モデルが指定の長さに応じて文長を制御できていることがわかる。また、図3に示すように、文生成のプロセスの後半では、制御トークン  $l$  の列に対し強くアテンションが払われており、条件付き言語モデルが文長制御を学習していることがわかる。また、生成した文の質的評価では、各指定長に応じて文法的かつ意味的に一貫性のある文を生成できることが確認された(表1)。

表1 指定長ごとの生成文の例

指定長*	入力単語(太字)+生成文(トークン数)*
5*	<b>He</b> died in 1885.(6)*
10*	<b>Gore</b> was born in New York City.(9)*
15*	"The song was released on the album as a single on September 26.(15)*
20*	<b>These</b> were the first of the three ships to be built for the Royal Navy in the 1930s.(20)*
25*	<b>As</b> a result, the city was forced to close its schools and to close its schools to the public in the summer.(24)*

## 5. まとめ

本研究では、条件付き言語モデルによる文長制御を提案し、それらの有効性を評価した。今後の課題として、文長条件付き言語モデルを事前学習モデルとして自然言語処理タスクに応用することが挙げられる。

### 参考文献

- [1]A. Radford, J. Wu, R. Child, D. Luan, D. Amodei and I. Sutskever, "Language Models are Unsupervised Multitask Learners," CoRR, 2019.
- [2]N. S. Keskar, B. McCann, L. R. Varshney, C. Xiong and R. Socher, "CTRL: A CONDITIONAL TRANSFORMER LANGUAGE MODEL FOR CONTROLLABLE GENERATION," CoRR, 2019.
- [3]A. Fan, D. Grangier and M. Auli, "Controllable Abstractive Summarization," ACL, 2018.
- [4]Y. Liu, Z. Luo and Kenny Zhu, "Controlling Length in Abstractive Summarization Using a Convolutional Neural Network," EMNLP, 2018.
- [5]高瀬翔, 岡崎直観, "位置エンコーディングを用いた出力長制御," 言語処理学会 第25回年次大会, 2019.