Effective Use of Target-side Context for Neural Machine Translation

美野 秀弥[†]

1 はじめに

本稿では、COLING2020 に採択された論文である “Effective Use of Target-side Context for Neural Machine Translation” (Mino et al. 2020) を解説する。採択論文の概要に加え、提案手法を思いつくまでの流れなどを交えて解説する。採択論文は文脈を考慮した機械翻訳に関するものである。従来の機械翻訳システムの多くは単文で独立に翻訳を行うことが多いが、単文が持つ情報だけでは適切な翻訳を行えない場合がある。例えば、単文翻訳では訳語を統一して翻訳することが難しくなる問題がある（訳語の統一の問題）。ある文で “東京デイズニーランド” と書かれているにも関わらず他の文で “TDR” と書かれている文書は良い文書ではないであろう（同じ表現を繰り返して使わない方が良い場合もあるので、一概には言えない問題である）。「ですます調」と「である調」が混在するような文末表現が統一していない文書も良い文書とは言えない。単文の持つ情報だけではこれらの表現を統一することは難しい。また、単文翻訳では曖昧性のある表現の翻訳が難しくなる問題がある（語義曖昧性解消の問題）。有名な例であるが、“I go to the bank.” という英文を日本語に翻訳する場合、“bank” が “銀行” なのか “土手” なのかこの文だけで判別して翻訳することは難しい。他にも、日本語などに代表されるゼロ代名詞が含まれる言語の翻訳ではゼロ照応解析が必要となるが、単文の情報だけでは文を跨いだゼロ照応解析は行えない（ゼロ照応解析の問題）。これらの問題は、文脈情報を使うことで解決できる可能性があり、多くの先行研究が存在する。文脈を考慮した機械翻訳に関する先行研究を知りたい方は Maruf et al. (2021) の論文を参照されたい。採択論文では、これらの既存手法の中から目的言語側の文脈情報を用いた機械翻訳 (Bawden et al. 2018) に着目した。Bawden et al. (2018) は文脈情報として目的言語側の前文を用いた手法では翻訳精度の向上が見られなかったと報告している。しかしながら、例えば訳語を統一させるには前文でどのように翻訳されているかが重要な情報となり、目的言語側の情報が有益な場合もあると考えられる。そこで、採択論文では、従来研究の問題点を洗い出し、目的言語側の文脈情報を用いて翻訳精度を向上させる手法を提案した。

[†] NHK 放送技術研究所

2 従来手法の問題点

目的言語側の文脈情報を用いた機械翻訳の従来手法は、以下のように学習、翻訳を行う。

学習 目的言語側の文脈情報の**参照訳**と原言語側の翻訳対象の文を入力し、翻訳対象の文の参照訳を出力するように翻訳モデルを学習する。

翻訳 目的言語側の文脈情報の**機械翻訳結果**と原言語側の翻訳対象の文を学習済みの機械翻訳器に入力し、機械翻訳結果を得る。

目的言語側の文脈情報として、学習時は参照訳を使うのに対して翻訳時には機械翻訳結果を使う。目的言語側の文脈情報が適切に翻訳されている参照訳に対して、機械翻訳結果は訳抜け、過剰訳、誤訳などの機械翻訳特有の誤りが含まれる。仮に誤りがない場合でも、機械翻訳結果は Translationese（翻訳調）と呼ばれる偏りのある文であることが知られている。Toral (2019) は、機械翻訳は人手翻訳よりもシンプルかつ標準化された結果になる傾向にあると示しており、参照訳と機械翻訳結果は多様性の側面からも異なる特徴を有していると考えられる。

著者らは、従来研究では学習時と翻訳時で異なる特徴を有するデータを文脈情報として用いたことで適切に翻訳モデルが学習されずに翻訳精度が低下すると仮定し、学習時と翻訳時の文脈情報の違い（ギャップ）の問題を低減させる手法を提案する。

3 提案手法

学習時と翻訳時の文脈情報のギャップをなくす最も簡単な手法は同じ情報を使うことであり、学習時と翻訳時ともに参照訳を使うか、機械翻訳結果を使うかの選択肢がある。翻訳誤り（ノイズ）のない参照訳を用いれば翻訳モデルはノイズを考慮した学習をする必要がないために学習が効率的に進むことが期待できるが、翻訳時は参照訳を使うことが困難であるため、機械翻訳結果を使う選択肢のみが残る。しかし、ノイズが多く含まれる機械翻訳結果を用いることで文脈情報から適切な情報を抽出することができなくなる可能性がある。機械翻訳器の精度が低いほど、この問題は顕著になると考えられる。すなわち、翻訳時と同じ特徴を有する目的言語側の文脈情報を使うという観点では機械翻訳結果を用いることが好ましいが、効率的な学習という観点では参照訳を用いることが好ましい。そこで、特徴の異なる参照訳と機械翻訳結果の両方を用いて効率的に学習する、カリキュラム学習 (Bengio et al. 2009) とスケジューリング (Bengio et al. 2015) のアプローチを参考にしたサンプリング制御の手法を提案した。この手法は、学習エポックごとに学習データ内に含まれる参照訳の割合（サンプリングレート p ）と機械翻訳の割合（サンプリングレート $1 - p$ ）を変更することで、翻訳モデルを段階的にノイズに対応させるものである。

(1) 1 エポック目の学習はサンプリングレート $p_1 = 1$ とし、目的言語側の文脈情報には全て

参照訳を用い、機械翻訳結果は用いない。

- (2) 2 エポック目以降の学習はサンプリングレート p_e を下記の逆シグモイド関数により減少させる。

$$p_e = \frac{k}{k + \exp(e - 2/k)}, \quad (1)$$

$e = 2, 3, \dots$ はエポック数であり、 $k (\geq 1)$ は p の減少スピードを制御するハイパーパラメータである。1 エポック終了時の翻訳モデルの性能が高いのであれば k の値を 1 に近づけ、2 エポック目以降の学習データ内の機械翻訳の割合を増やす。上記の手順により、ノイズに頑健な翻訳モデルの学習を実現する。実装する機械翻訳器のアーキテクチャには、Concatenation-based Context-aware NMT (Tiedemann and Scherrer 2017) と Multi-encoder Context-aware NMT (Bawden et al. 2018; Kim et al. 2019) を用いた。どちらもエンコーダとデコーダにより構成されるが、文脈情報の入力方法が異なる。Concatenation-based Context-aware NMT は、文脈情報と翻訳対象の文とを特殊記号で結合させた 1 つのデータを 1 種類のエンコーダに入力するアーキテクチャであり、Multi-encoder Context-aware NMT は、文脈情報と翻訳対象の文を別々のエンコーダに入力するアーキテクチャである。

4 実験

提案手法の効果を、時事通信社ニュースの英日・日英機械翻訳実験と、TED トークの英日・日英、および英独・独英機械翻訳実験で確認した。ベースライン手法として、文脈情報を用いない手法 (Single-Sentence)、学習時の文脈情報として目的言語側の参照訳のみを用いた手法 (Trg-Context GT) を用意した。また、学習時と翻訳時のギャップをなくした、学習時の文脈情報として目的言語側の機械翻訳結果のみを用いた手法 (Trg-Context Pred.) も実験に加えた。

実験結果を表 1 に示す。Concatenation-based Context-aware NMT と Multi-encoder Context-aware NMT との比較では、タスクによって優劣が異なる結果となった。Trg-Context GT、及び Trg-Context Pred. と提案手法との比較では、全タスクにおいて提案手法が BLEU スコアで上回っており、参照訳と機械翻訳結果の両方のデータを制御して用いた効果があったといえる。文脈情報を用いていない Single-Sentence の結果と提案手法との比較では、Multi-encoder Context-aware NMT では全タスクにおいて BLEU スコアで上回っており文脈を用いる効果を確認できた。採択論文には記載していない余談になるが、当初は学習時に機械翻訳結果を使うことで効果が出るものと考えており、これを提案手法とする予定でいた。しかし、表 1 の Trg-Context Pred. の結果から分かるように効果のあるタスクとないタスクが混在するという結果となってしまった。そこで、Trg-Context Pred. よりも Trg-Context GT の方が翻訳精度向上が大きいタスク（時事通信社ニュースの JA-EN、TED トークタスクの DE-EN）があることに着目し、両方を使う手

手法	アーキテクチャ	時事通信社ニュース		TED トーク			
		EN-JA	JA-EN	EN-JA	JA-EN	EN-DE	DE-EN
Single-Sentence	vanilla	41.99	24.23	15.69	8.48	24.04	28.46
Trg-Context GT	concat.	41.45	24.40	14.72	7.55	25.09	29.11
Trg-Context Pred.	concat.	42.07	23.83	15.32	7.37	23.99	29.05
提案手法	concat.	42.89	24.62	15.33	8.36	25.09	29.74
Trg-Context GT	multi-enc.	42.40	24.80	16.15	8.98	25.96	30.25
Trg-Context Pred.	multi-enc.	42.45	24.31	16.27	9.15	25.98	30.23
提案手法	multi-enc.	42.79	24.86	16.37	9.66	26.01	30.60

表 1 英語 $\leftrightarrow$ 日本語（時事通信社ニュース, TED トーク）, 及び英語 $\leftrightarrow$ ドイツ語（TED トーク）機械翻訳実験の BLEU スコア. vanilla は Transformer を用いた NMT, concat. は Concatenation-based Context-aware NMT, multi-enc. は Multi-encoder Context-aware NMT を表す.

法が適切と考え、提案手法ができあがった。さらに、提案手法を用いたことによる具体的な改善を見るために人手評価を行い、訳語の統一やゼロ照応解析で改善が見られる事例を確認した。より詳細な改善事例については、採択論文を参照されたい。

5 おわりに

近年、機械翻訳の精度向上のために翻訳対象文以外の情報を使う研究が盛んに行われ、多くの成果が見られている。本稿で解説した文脈情報だけでなく、マルチモーダル翻訳でよく用いられる画像情報（マルチモーダル翻訳とは視覚や聴覚、触覚など複数のモダリティを用いた機械翻訳を指す）や、BERT を始めとする事前学習された言語モデルなども含まれるであろう。扱う情報の種類が増えるほど翻訳精度向上の可能性は高まるが、その分だけそれぞれの情報の特徴を熟知しなくてはならない。研究者としてはうれしい悲鳴であろう。また、採択論文は数多くある文脈を考慮した機械翻訳手法の中の 1 手法に着目しただけである。冒頭で述べたとおり、文脈情報を使う利点は多くある。今後も機械翻訳のための文脈情報の利用に関する研究を進め、さらなる成果を出していきたい。

参考文献

- Bawden, R., Sennrich, R., Birch, A., and Haddow, B. (2018). “Evaluating Discourse Phenomena in Neural Machine Translation.” In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pp. 1304–1313, New Orleans, Louisiana. Association for

Computational Linguistics.

- Bengio, S., Vinyals, O., Jaitly, N., and Shazeer, N. (2015). “Scheduled Sampling for Sequence Prediction with Recurrent Neural Networks.” In Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R. (Eds.), *Advances in Neural Information Processing Systems* 28, pp. 1171–1179. Curran Associates, Inc.
- Bengio, Y., Louradour, J., Collobert, R., and Weston, J. (2009). “Curriculum Learning.” In Danyluk, A. P., Bottou, L., and Littman, M. L. (Eds.), *ICML*, Vol. 382 of *ACM International Conference Proceeding Series*, pp. 41–48. ACM.
- Kim, Y., Tran, D. T., and Ney, H. (2019). “When and Why is Document-level Context Useful in Neural Machine Translation?” In *Proceedings of the 4th Workshop on Discourse in Machine Translation, DiscoMT@EMNLP 2019, Hong Kong, China, November 3, 2019*, pp. 24–34.
- Maruf, S., Saleh, F., and Haffari, G. (2021). “A Survey on Document-level Neural Machine Translation: Methods and Evaluation.”
- Mino, H., Ito, H., Goto, I., Yamada, I., and Tokunaga, T. (2020). “Effective Use of Target-side Context for Neural Machine Translation.” In *Proceedings of the 28th International Conference on Computational Linguistics*, pp. 4483–4494, Barcelona, Spain (Online). International Committee on Computational Linguistics.
- Tiedemann, J. and Scherrer, Y. (2017). “Neural Machine Translation with Extended Context.” In *Proceedings of the Third Workshop on Discourse in Machine Translation*, pp. 82–92, Copenhagen, Denmark. Association for Computational Linguistics.
- Toral, A. (2019). “Post-edits: an Exacerbated Translationese.” In *Proceedings of Machine Translation Summit XVII Volume 1: Research Track*, pp. 273–281, Dublin, Ireland. European Association for Machine Translation.

略歴

美野 秀弥：2002 年東京工業大学情報理工学部計算工学科卒業。2004 年同大学院情報理工学研究科修士課程修了。同年 NHK 入局。北見放送局、国立研究開発法人情報通信研究機構を経て、現在、放送技術研究所スマートプロダクション研究部研究員。自然言語処理の研究に従事。