# Example-Based Machine Translation with a Multi-Sentence Construction Transformer Architecture

**Haozhe Xiao** **Yifei Zhou** **Yves Lepage**
IPS, Waseda University
2-7 Hibikino, Wakamatsu-ku, Kitakyushu-shi, 808-0135 Fukuoka-ken, Japan
xiaohaozhe@moegi.waseda.jp, yifei.zhou@ruri.waseda.jp, yves.lepage@waseda.jp

## Abstract

Neural Machine Translation (NMT) has now attained state-of-art performance on large-scale data. However, it does not achieve the best translation results on small data sets. Example-Based Machine Translation (EBMT) is an approach to machine translation in which existing examples in a database are retrieved and modified to generate new translations. To combine EBMT with NMT, an architecture based on the Transformer model is proposed. We conduct two experiments respectively using limited amounts of data, one on an English-French bilingual dataset and the other one on a multilingual dataset with six languages (English, French, German, Chinese, Japanese and Russian). On the bilingual task, our method achieves an accuracy of 96.5 and a BLEU score of 98.8. On the multilingual task, it also outperforms OpenNMT in terms of BLEU scores.

## 1 Introduction

An analogy is a relationship between four objects, $A$ is to $B$ as $C$ is to $D$. Studies on analogies have investigated their utility in different applications, like machine translation. Solving analogies between sentences involves the task of generating an unknown $D$ that satisfies an analogical equation $A : B :: C : D$ , where $A$, $B$, and $C$ are given. Here is an example of a sentence analogy:

$$\textit{he 's coming .} : \textit{i am coming .} :: \begin{array}{l}\textit{he 's eating}\\ \textit{an apple .}\end{array} : x$$

$$\Rightarrow \quad x = \begin{array}{l}\textit{i am eating}\\ \textit{an apple .}\end{array}$$

EBMT extracts knowledge from a corpus in two languages to perform translation. Concretely, the process of EBMT by analogy involves extracting analogical relationships in the source language to find the corresponding sentences in the target language and solve a sentence analogy.

Formula (1) defines the notation of analogies between sentences in two languages. As instantiated in Formula (2), the translation result for "*i am eating an apple .*" is "*je manger une pomme .*", which can be obtained through the reasoning process.

$$\begin{array}{ccccccc} A & : & B & :: & C & : & D \\ \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\ A' & : & B' & :: & C' & : & D' \end{array} \qquad (1)$$

$$\begin{array}{ccccccc} \textit{he 's coming .} & : & \textit{i am coming .} & :: & \textit{he 's eating an apple .} & : & \textit{i am eating an apple .} \\ \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\ \textit{il est en train d' arriver .} & : & \textit{j' arrive .} & :: & \textit{il est en train de manger une pomme .} & : & ?? \end{array} \qquad (2)$$

EBMT by analogy is a translation method that involves generating a target translation by using multiple example sentences for reference and reasoning. However, the vanilla Transformer (Vaswani et al., 2017) model can only handle one input at a time. To address this limitation, we propose a multi-sentence construction Transformer architecture designed specifically for EBMT by analogy.

## 2 Previous work and proposal

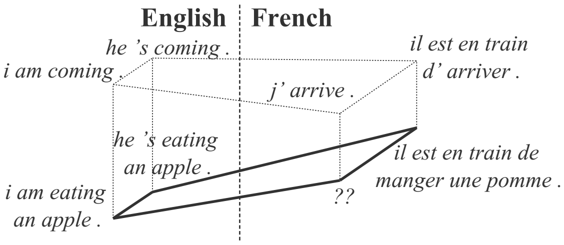
To perform translation, Nagao (1984) proposed an approach to EBMT that considers a bilingual analogy across two languages. Translations are made by transferring symbolic knowledge from the source language to the target language. In Figure 1(a), the translation of "*i am eating an apple .*" is achieved by solving a bilingual analogy:

$$\textit{i am coming .} : \textit{j' arrive .} :: \begin{array}{l}\textit{i am eating}\\ \textit{an apple .}\end{array} : ??$$

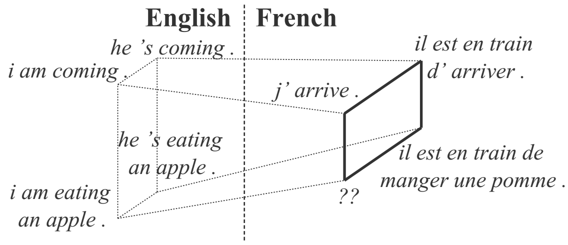
Figure 1(b) outlines the indirect approach to EBMT. As previously shown in Formula (2), previous research considered two monolingual analogies in two different languages that correspond to



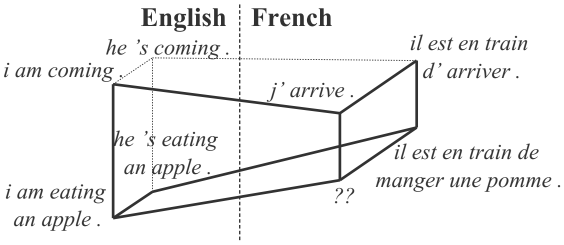
*Proceedings of the 2023 CLASP Conference on Learning with Small Data*, pages 72–80
September 11–12, 2023. 



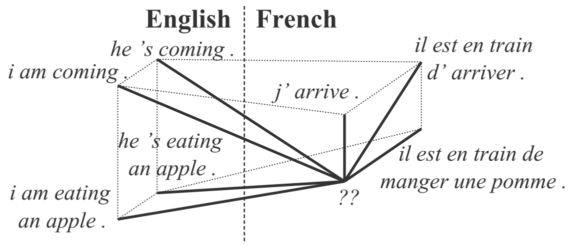

(a) direct approach (Nagao, 1984)




(b) indirect approach (Lepage and Denoual, 2005)




(c) combination of direct and indirect approaches (Taillandier et al., 2020)




(d) our proposed method

Figure 1: Different approaches to EBMT by analogy (adapted from (Taillandier et al., 2020)). In each sub-figure, the left half shows the embedding space for English sentences while the right half shows the embedding space for French sentences. Relationships between the sentences are represented by connecting lines.

generate translations (Lepage and Denoual, 2005; Langlais et al., 2008; Dandapat et al., 2010).

A step further, Taillandier et al. (2020) proposed to fuse the direct approach with the indirect approach. See Figure 1(c). The translation output can be obtained by solving the following three analogical equations.

*i am coming .* : *j' arrive .* :: *i am eating an apple .* : *??*

*he 's eating an apple .* : *il est en train de manger une pomme .* :: *i am eating an apple .* : *??*

*il est en train d' arriver .* : *j' arrive .* :: *il est en train de manger une pomme .* : *??*

Here, our proposal is to use the Transformer model to establish direct connections between each input sentence and the output sentence, in contrast to the fusion approach of using three quadrilateral relations to obtain the translation result as shown in Figure 1(c). With our approach, the input sentence information is better synthesized to generate the target translation as illustrated in Figure 1(d). The use of multiple attention is expected to enhance the translation accuracy of the results.

## 3 Multi-sentence construction Transformer architecture

We propose a novel Transformer structure that allows for multiple sentences to be inputted simultaneously, compared to the vanilla Transformer's single-sentence input. Concretely, this multi-sentence construction Transformer architecture takes seven sentences $A, B, C, D, A', B', C'$ as input to generate the target translation $D'$. Rather than concatenating them into a single input, we employ seven distinct inputs, which allows each individual input to compute attention with the output.

### 3.1 Structure of the decoder

The vanilla Transformer's decoder only receives two inputs to establish their connection: the sequence of vector representation of the source sentence from the encoder and the sequence of the target sentence. As an initial step towards building our multi-sentence construction Transformer architecture, a new decoder that can accommodate three inputs is designed in Figure 2.

To learn the relationship with the upper decoder, we add an extra layer of cross-attention after self-attention. This layer calculates the attention between the upper decoder's output and the target sentence, enabling the computation of attention to each input with the target output and establishing a connection. As a result, we create a decoder with three inputs for follow-up use.

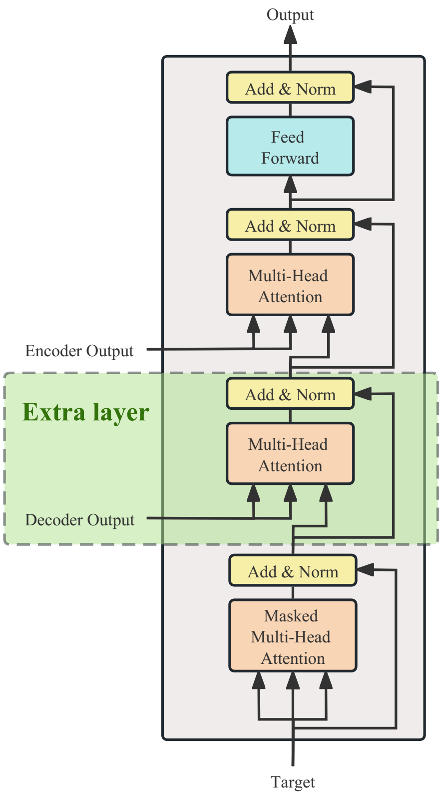



Figure 2: Structure of the decoder

## 3.2 Architecture for EBMT Transformer

Figure 3 illustrates the contrast between our proposed model architecture and the vanilla Transformer. Our proposal transforms the Transformer's single input into several independent encoder-decoder pairs. With multiple decoder layers overlaying each other, our EBMT Transformer can automatically encode the semantic information of the input sequence and use it to generate the appropriate target sequence.

Therefore, in the multi-sentence construction Transformer architecture for EBMT by analogy:

- All the encoders have the same structure as the vanilla Transformer's encoder, but each encoder has a weight specific to the corresponding input.

- Except for *decoder_A* which has the same structure as the vanilla Transformer's decoder (a two-input decoder), the other six decoders are the three-input decoders introduced in Section 3.1.

# 4 Datasets and metrics

## 4.1 Datasets

We use experimental data obtained directly from the bilingual analogy dataset developed by Taillandier et al. (2020) for comparison. For the task of multilingual machine translation, it will be necessary to create a multilingual analogy dataset.

### 4.1.1 Bilingual dataset

All sentences in the bilingual analogy dataset (Taillandier et al., 2020) are from Tatoeba[1]. The dataset is randomly divided into a training set (80%), validation set (10%), and test set (10%) by the number of analogies. As shown in Table 1, the average sentence length is approximately 5 words. Table 1 also counts the number of unique sentences contained in the dataset. Despite the fact that the whole dataset contains 239,594 analogies between sentences, it only contains 8,867 English sentences and 10,437 French sentences without repetition.

### 4.1.2 Multilingual dataset

To produce an analogy dataset for multiple languages, we extract analogies from the Tatoeba corpus using the Nlg package[2] (Fam and Lepage, 2018). Tatoeba is a collection of sentences in over 100 languages. In this work, we use English, French, German, Chinese, Japanese and Russian. Thus, we construct a multilingual dataset in six languages with 7,099 analogies and divide it into 80%, 10%, 10%.

Table 2 shows the statistics of the extracted multilingual dataset. In particular, each language has approximately 1,700 unique sentences. When considering the sentence length on the word level, Japanese has the longest average length and Russian has the shortest one.

## 4.2 Evaluation metrics

We automatically assess experimental results by comparing the translation output to the reference sentence in the test set. We use the three metrics listed below.

BLEU (Bilingual Evaluation Understudy) evaluates the similarity between the translated and reference sentences (Papineni et al., 2002). It features a 0 to 100 scale. The closer the translation output

[1] https://tatoeba.org
[2] http://lepage-lab.ips.waseda.ac.jp/media/filer_public/64/52/64528717-c3ce-4617-8208-c1fb70cf1442/nlg-v321.zip

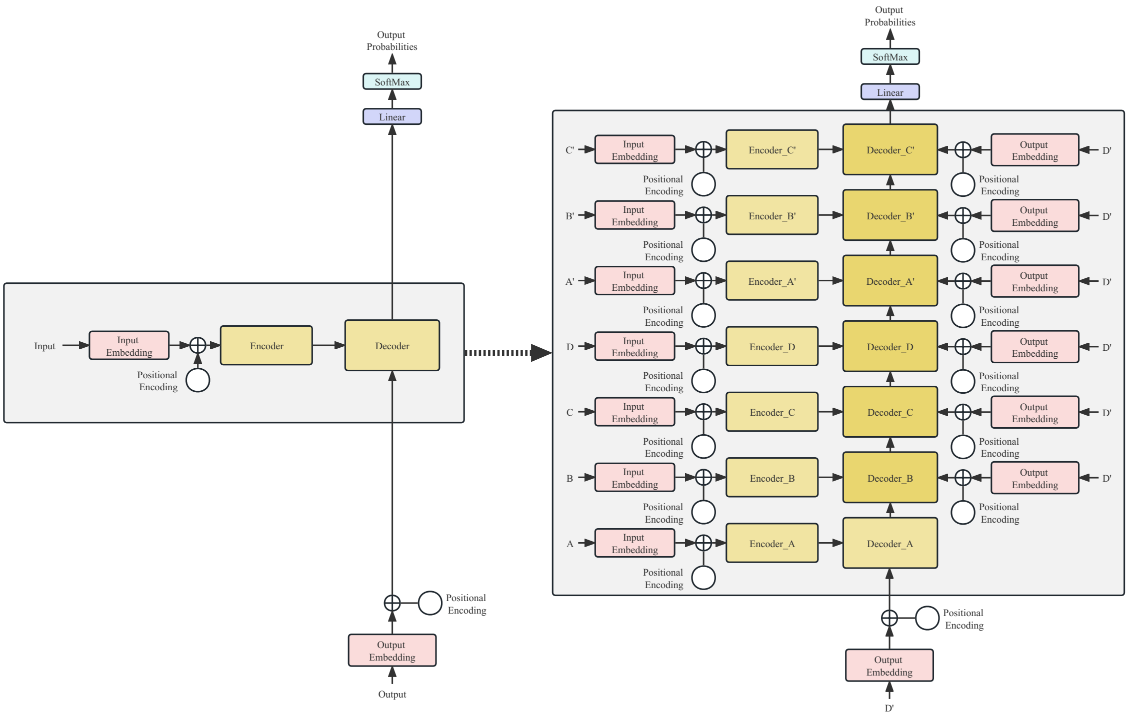



Figure 3: Model architecture: on the left, the vanilla Transformer, on the right, our EBMT Transformer.

Table 1: Statistics of the bilingual dataset

| Dataset | Analogies | # of unique sentences | | words/sentence | | characters/sentence | |
|---|---|---|---|---|---|---|---|
| | | English | French | English | French | English | French |
| train | 191,676 | 8,867 | 10,437 | 5.50±1.45 | 5.66±1.57 | 21.31±6.04 | 24.54±7.33 |
| valid | 23,959 | 7,734 | 8,868 | 5.49±1.45 | 5.65±1.56 | 21.27±6.01 | 24.48±7.30 |
| test | 23,959 | 7,768 | 8,955 | 5.51±1.46 | 5.66±1.57 | 21.35±6.06 | 24.53±7.33 |

Table 2: Statistics of the multilingual dataset

| Dataset | Analogies | Language | # of unique sentences | words/sentence | characters/sentence |
|---|---|---|---|---|---|
| | | English | 1,746 | 5.65±1.50 | 21.39±6.41 |
| | | French | 1,676 | 5.75±1.82 | 24.78±8.44 |
| train | 5,679 | German | 1,665 | 5.17±1.34 | 24.14±7.68 |
| | | Chinese | 1,628 | 4.95±1.33 | 11.35±3.08 |
| | | Japanese | 1,662 | 6.90±2.42 | 17.17±5.82 |
| | | Russian | 1,664 | 4.50±1.40 | 20.25±7.46 |
| | | English | 956 | 5.61±1.45 | 21.13±6.20 |
| | | French | 927 | 5.73±1.75 | 24.51±8.14 |
| valid | 710 | German | 928 | 5.14±1.29 | 23.88±7.51 |
| | | Chinese | 915 | 4.91±1.30 | 11.22±2.99 |
| | | Japanese | 916 | 6.91±2.35 | 17.17±5.69 |
| | | Russian | 925 | 4.50±1.37 | 20.31±7.39 |
| | | English | 946 | 5.67±1.53 | 21.40±6.37 |
| | | French | 915 | 5.74±1.82 | 24.63±8.43 |
| test | 710 | German | 917 | 5.18±1.40 | 24.08±7.83 |
| | | Chinese | 900 | 4.92±1.33 | 11.30±3.04 |
| | | Japanese | 912 | 6.92±2.42 | 17.18±5.84 |
| | | Russian | 916 | 4.49±1.38 | 20.11±7.40 |

is to the reference sentence, the higher the BLEU score is. We use SacreBLEU[3] (Post, 2018).

Accuracy refers to the percentage of translation results where the model outputs are identical to the reference sentences. This value can be expressed as the ratio of the number of identical results, denoted by $n$, to the total number of references, denoted by $m$. Mathematically, it can be represented as Accuracy $= n/m$.

Levenshtein edit distance (Levenshtein, 1966) is defined as the minimum number of edit operations (insertions, deletions or substitutions) required to transform one string into another. We evaluate the results using two units: word and character. A smaller edit distance indicates better results.

## 5 Experiments and analysis

To evaluate the performance of our proposed model, we compare its translation results to those of other methods. For the bilingual translation task from English to French, we use OpenNMT[4] (Klein et al., 2017) and the method proposed by Taillandier et al. (2020) as baselines. For the multilingual translation task across six languages, we use OpenNMT only. The parameter settings for OpenNMT and our proposal are detailed in Appendix A.

### 5.1 Bilingual translation task

Table 3 shows the translation results of various methods on the bilingual dataset mentioned in Section 4.1.1. Our proposed EBMT Transformer achieved a BLEU score of 98.8, outperforming OpenNMT's 90.3 and Taillandier et al. (2020)'s 94.7. Additionally, our model outperformed the baselines in terms of accuracy and edit distance metrics, demonstrating the stability of the results. Therefore, the multi-sentence construction Transformer architecture clearly provides a substantial improvement on this task.

Appendix B provides an error case for translation into French. Our proposed method faces challenges when it comes to accurately incorporating punctuation marks during the inference process.

### 5.2 Multilingual translation task

For multilingual translation across six languages, a total of $C_2^6 \times 2 = 30$ models need to be trained for each translation direction. The complete results are attached in Appendix C. Figures 4 and 5 present the BLEU score and accuracy of multilingual translation across six languages, respectively.

As shown in Figure 4(a), all OpenNMT models achieved a BLEU score of over 75. This is impressive given that OpenNMT typically requires a large amount of training data to achieve good results. However, as discussed in Section 4.1.2, the multilingual dataset used in this experiment only contains a total of 7,099 analogies, indicating that the dataset is very particular. We further observe that when English, French and Russian are the target language, the results are generally better than for other languages.

After comparing the BLEU score and accuracy in Figures 4 and 5, it can be concluded that the EBMT Transformer outperforms OpenNMT for all six languages. Although both methods have high translation performance, this is likely due to the fact that the languages involved do not have a large vocabulary and the sentences are short. The BLEU scores for Chinese as the target language are lower than those of other languages. This is mainly because Chinese has the lowest average number of characters per sentence, which results in a lower BLEU score calculation.

## 6 Conclusion

We proposed a multi-sentence construction Transformer architecture model to implement EBMT by analogy. Our proposal outperformed the two baselines on the bilingual dataset, achieving a BLEU score of 98.8 and an accuracy of 96.5. Additionally, for the multilingual translation task across six languages, our proposed method produced significantly better results than OpenNMT.

## Limitations

Note that the used datasets are relatively easy ones. This raises questions about the generalizability of our proposed model when used in a real EBMT by analogy setting where retrieval of analogies from an input sentence should be taken into consideration. Future research will explore this issue using more complex datasets.

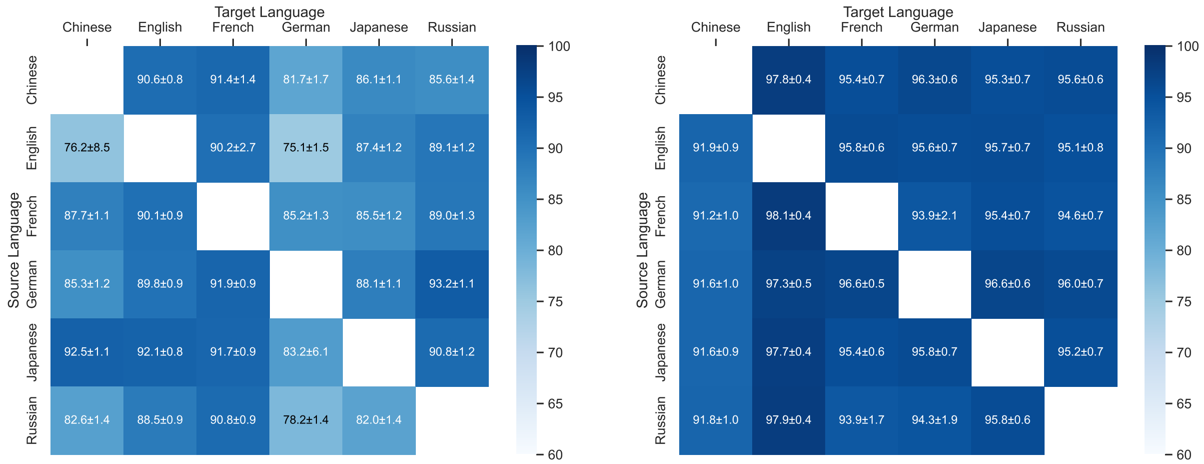
## Acknowledgements

The results reported here were obtained in part under a JSPS grant, Kakenhi Kiban C, n° 18K11447 entitled "Self-explainable and fast-to-train example-based machine translation using neural networks."

[3] `https://github.com/mjpost/sacrebleu`
[4] `https://opennmt.net/`

Table 3: Translation results of different methods on the bilingual dataset (en → fr)

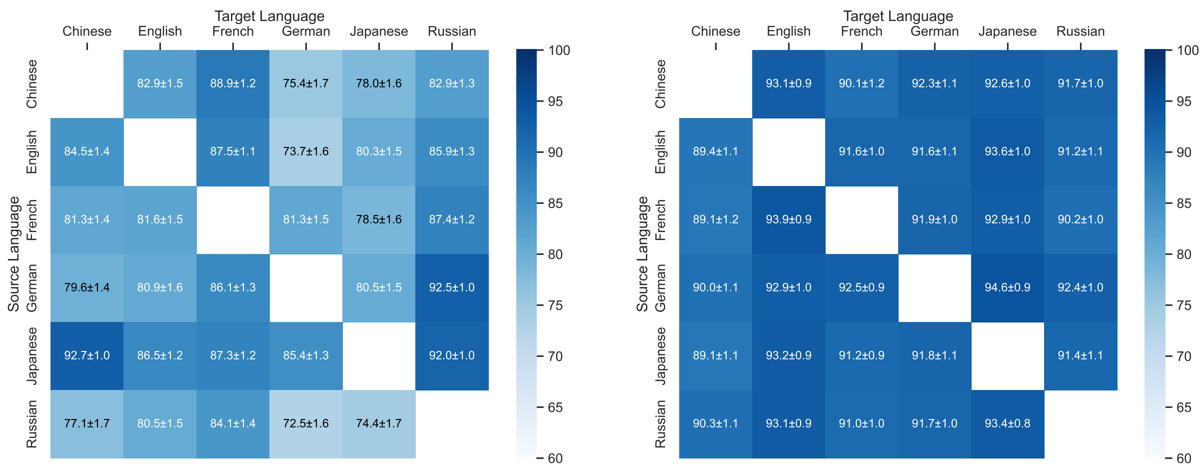
| **Method** | **BLEU** | **Accuracy** | **Edit distance** | |
|---|---|---|---|---|
| | | | **in word** | **in char.** |
| OpenNMT | 90.3 | 82.7 | 0.5 | 1.0 |
| (Taillandier et al., 2020) | 94.7 | 90.2 | 0.2 | 0.6 |
| EBMT Transformer | **98.8** | **96.5** | **0.1** | **0.2** |




(a) OpenNMT

(b) EBMT Transformer

Figure 4: BLEU scores of multilingual translation across six languages




(a) OpenNMT

(b) EBMT Transformer

Figure 5: Accuracy of multilingual translation across six languages

## References

Sandipan Dandapat, Sara Morrissey, Sudip Kumar Naskar, and Harold Somers. 2010. Mitigating problems in analogy-based ebmt with smt and vice versa: a case study with named entity transliteration. In *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation*, pages 365–372.

Rashel Fam and Yves Lepage. 2018. Tools for the production of analogical grids and a resource of n-gram analogical grids in 11 languages. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Guillaume Klein, Yoon Kim, Yuntian Deng, Jean Senellart, and Alexander Rush. 2017. OpenNMT: Open-source toolkit for neural machine translation. In *Proceedings of ACL 2017, System Demonstrations*, pages 67–72, Vancouver, Canada. Association for Computational Linguistics.

Philippe Langlais, François Yvon, and Pierre Zweigenbaum. 2008. Analogical translation of medical words in different languages. In *Advances in Natural Language Processing: 6th International Conference, GoTAL 2008 Gothenburg, Sweden, August 25-27, 2008 Proceedings*, pages 284–295. Springer.

Yves Lepage and Etienne Denoual. 2005. Purest ever example-based machine translation: Detailed presentation and assessment. *Machine translation*, 19(3/4):251–282.

V.I. Levenshtein. 1966. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics-doklady*, 10(8):707–710.

Makoto Nagao. 1984. A framework of a mechanical translation between japanese and english by analogy principle. *Artificial and human intelligence*, pages 351–354.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.

Matt Post. 2018. A call for clarity in reporting BLEU scores. In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 186–191, Brussels, Belgium. Association for Computational Linguistics.

Valentin Taillandier, Liyan Wang, and Yves Lepage. 2020. Réseaux de neurones pour la résolution d'analogies entre phrases en traduction automatique par l'exemple. In *6e conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 2 : Traitement Automatique des Langues Naturelles*, pages 108–121, Nancy, France. ATALA.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.


## A Experimental setup

| Encoder&Decoder | |
|---|---|
| Type | Transformer |
| Embedding dimension | 512 |
| Number of layers | 6 |
| Number of heads | 8 |
| Size of feedforward layer | 2048 |
| Optimizer | Adam |
| Learning rate | 1.0 |

Table 4: Parameter settings for OpenNMT

| Encoder&Decoder | |
|---|---|
| Embedding dimension | 512 |
| Number of layers | 1 |
| Number of heads | 8 |
| Size of feedforward layer | 2048 |
| Optimizer | Adam |
| Learning rate | 0.0001 |
| Dropout | 0.1 |
| Max length | 80 |

Table 5: Parameter settings for our proposal

## B Error case in bilingual translation task

Table 6: Error case in bilingual translation task

| IDs | Sentences |
|---|---|
| $A$ | you're the love of my life . |
| $B$ | you're such a jerk . |
| $C$ | he's the love of my life . |
| $D$ | he's such a jerk . |
| $A'$ | tu es l'amour de ma vie . |
| $B'$ | tu es un de ces pauvres types ! |
| $C'$ | c'est l'amour de ma vie . |
| Output | c'est un de ces pauvres . |
| Reference | c'est un de ces pauvres types ! |

## C Results of multilingual translation task

See next pages.

Table 7: Result of multilingual translation task with OpenNMT

| Source Language | Target Language | BLEU | Accuracy | Edit distance in word | Edit distance in char. |
|---|---|---|---|---|---|
| English | French | 90.2 ± 2.7 | 87.5 ± 1.1 | 0.5 ± 0.2 | 1.9 ± 0.6 |
| English | German | 75.1 ± 1.5 | 73.7 ± 1.6 | 1.0 ± 0.1 | 4.3 ± 0.3 |
| English | Chinese | 76.2 ± 8.5 | 84.5 ± 1.4 | 1.1 ± 0.5 | 2.1 ± 1.0 |
| English | Japanese | 87.4 ± 1.2 | 80.3 ± 1.5 | 0.9 ± 0.1 | 1.8 ± 0.2 |
| English | Russian | 89.1 ± 1.2 | 85.9 ± 1.3 | 0.3 ± 0.1 | 1.6 ± 0.2 |
| French | English | 90.1 ± 0.9 | 81.6 ± 1.5 | 0.5 ± 0.1 | 1.7 ± 0.2 |
| French | German | 85.2 ± 1.3 | 81.3 ± 1.5 | 0.6 ± 0.1 | 2.8 ± 0.3 |
| French | Chinese | 87.7 ± 1.1 | 81.3 ± 1.4 | 0.4 ± 0.1 | 1.0 ± 0.1 |
| French | Japanese | 85.5 ± 1.2 | 78.5 ± 1.6 | 1.0 ± 0.1 | 2.1 ± 0.2 |
| French | Russian | 89.0 ± 1.3 | 87.4 ± 1.2 | 0.3 ± 0.0 | 1.5 ± 0.2 |
| German | English | 89.8 ± 0.9 | 80.9 ± 1.6 | 0.5 ± 0.1 | 1.6 ± 0.2 |
| German | French | 91.9 ± 0.9 | 86.1 ± 1.3 | 0.5 ± 0.1 | 1.8 ± 0.2 |
| German | Chinese | 85.3 ± 1.2 | 79.6 ± 1.4 | 0.5 ± 0.1 | 1.1 ± 0.1 |
| German | Japanese | 88.1 ± 1.1 | 80.5 ± 1.5 | 0.8 ± 0.1 | 1.7 ± 0.2 |
| German | Russian | 93.2 ± 1.1 | 92.5 ± 1.0 | 0.2 ± 0.0 | 0.9 ± 0.1 |
| Chinese | English | 90.6 ± 0.8 | 82.9 ± 1.5 | 0.5 ± 0.1 | 1.6 ± 0.2 |
| Chinese | French | 91.4 ± 1.4 | 88.9 ± 1.2 | 0.4 ± 0.1 | 1.6 ± 0.2 |
| Chinese | German | 81.7 ± 1.7 | 75.4 ± 1.7 | 0.8 ± 0.1 | 3.8 ± 0.4 |
| Chinese | Japanese | 86.1 ± 1.1 | 78.0 ± 1.6 | 0.9 ± 0.1 | 1.9 ± 0.2 |
| Chinese | Russian | 85.6 ± 1.4 | 82.9 ± 1.3 | 0.5 ± 0.1 | 2.0 ± 0.2 |
| Japanese | English | 92.1 ± 0.8 | 86.5 ± 1.2 | 0.4 ± 0.1 | 1.3 ± 0.2 |
| Japanese | French | 91.7 ± 0.9 | 87.3 ± 1.2 | 0.4 ± 0.1 | 1.7 ± 0.2 |
| Japanese | German | 83.2 ± 6.1 | 85.4 ± 1.3 | 0.8 ± 0.4 | 3.3 ± 1.6 |
| Japanese | Chinese | 92.5 ± 1.1 | 92.7 ± 1.0 | 0.2 ± 0.0 | 0.5 ± 0.1 |
| Japanese | Russian | 90.8 ± 1.2 | 92.0 ± 1.0 | 0.2 ± 0.0 | 1.1 ± 0.1 |
| Russian | English | 88.5 ± 0.9 | 80.5 ± 1.5 | 0.6 ± 0.1 | 1.9 ± 0.2 |
| Russian | French | 90.8 ± 0.9 | 84.1 ± 1.4 | 0.5 ± 0.1 | 2.0 ± 0.2 |
| Russian | German | 78.2 ± 1.4 | 72.5 ± 1.6 | 0.8 ± 0.1 | 3.8 ± 0.3 |
| Russian | Chinese | 82.6 ± 1.4 | 77.1 ± 1.7 | 0.7 ± 0.1 | 1.4 ± 0.1 |
| Russian | Japanese | 82.0 ± 1.4 | 74.4 ± 1.7 | 1.2 ± 0.1 | 2.4 ± 0.2 |

Table 8: Result of multilingual translation task with EBMT Transformer

| Source Language | Target Language | BLEU | Accuracy | Edit distance in word | Edit distance in char. |
|---|---|---|---|---|---|
| English | French | 95.8 ± 0.6 | 91.6 ± 1.0 | 0.2 ± 0.0 | 0.8 ± 0.1 |
| English | German | 95.6 ± 0.7 | 91.6 ± 1.1 | 0.2 ± 0.0 | 0.8 ± 0.1 |
| English | Chinese | 91.9 ± 0.9 | 89.4 ± 1.1 | 0.3 ± 0.0 | 0.4 ± 0.1 |
| English | Japanese | 95.7 ± 0.7 | 93.6 ± 1.0 | 0.3 ± 0.1 | 0.6 ± 0.1 |
| English | Russian | 95.1 ± 0.8 | 91.2 ± 1.1 | 0.2 ± 0.0 | 0.8 ± 0.1 |
| French | English | 98.1 ± 0.4 | 93.9 ± 0.9 | 0.1 ± 0.0 | 0.3 ± 0.1 |
| French | German | 93.9 ± 2.1 | 91.9 ± 1.0 | 0.3 ± 0.1 | 1.2 ± 0.4 |
| French | Chinese | 91.2 ± 1.0 | 89.1 ± 1.2 | 0.3 ± 0.0 | 0.4 ± 0.1 |
| French | Japanese | 95.4 ± 0.7 | 92.9 ± 1.0 | 0.3 ± 0.1 | 0.6 ± 0.1 |
| French | Russian | 94.6 ± 0.7 | 90.2 ± 1.0 | 0.2 ± 0.0 | 0.9 ± 0.1 |
| German | English | 97.3 ± 0.5 | 92.9 ± 1.0 | 0.1 ± 0.0 | 0.5 ± 0.1 |
| German | French | 96.6 ± 0.5 | 92.5 ± 0.9 | 0.2 ± 0.0 | 0.6 ± 0.1 |
| German | Chinese | 91.6 ± 1.0 | 90.0 ± 1.1 | 0.3 ± 0.0 | 0.3 ± 0.1 |
| German | Japanese | 96.6 ± 0.6 | 94.6 ± 0.9 | 0.2 ± 0.1 | 0.5 ± 0.1 |
| German | Russian | 96.0 ± 0.7 | 92.4 ± 1.0 | 0.2 ± 0.0 | 0.7 ± 0.1 |
| Chinese | English | 97.8 ± 0.4 | 93.1 ± 0.9 | 0.1 ± 0.0 | 0.4 ± 0.1 |
| Chinese | French | 95.4 ± 0.7 | 90.1 ± 1.2 | 0.3 ± 0.1 | 0.9 ± 0.1 |
| Chinese | German | 96.3 ± 0.6 | 92.3 ± 1.1 | 0.2 ± 0.0 | 0.6 ± 0.1 |
| Chinese | Japanese | 95.3 ± 0.7 | 92.6 ± 1.0 | 0.3 ± 0.1 | 0.7 ± 0.1 |
| Chinese | Russian | 95.6 ± 0.6 | 91.7 ± 1.0 | 0.2 ± 0.0 | 0.7 ± 0.1 |
| Japanese | English | 97.7 ± 0.4 | 93.2 ± 0.9 | 0.1 ± 0.0 | 0.4 ± 0.1 |
| Japanese | French | 95.4 ± 0.6 | 91.2 ± 0.9 | 0.3 ± 0.1 | 0.9 ± 0.1 |
| Japanese | German | 95.8 ± 0.7 | 91.8 ± 1.1 | 0.2 ± 0.0 | 0.8 ± 0.1 |
| Japanese | Chinese | 91.6 ± 0.9 | 89.1 ± 1.1 | 0.3 ± 0.0 | 0.4 ± 0.1 |
| Japanese | Russian | 95.2 ± 0.7 | 91.4 ± 1.1 | 0.2 ± 0.0 | 0.8 ± 0.1 |
| Russian | English | 97.9 ± 0.4 | 93.1 ± 0.9 | 0.1 ± 0.0 | 0.4 ± 0.1 |
| Russian | French | 93.9 ± 1.7 | 91.0 ± 1.0 | 0.3 ± 0.1 | 1.1 ± 0.3 |
| Russian | German | 94.3 ± 1.9 | 91.7 ± 1.0 | 0.3 ± 0.1 | 1.1 ± 0.4 |
| Russian | Chinese | 91.8 ± 1.0 | 90.3 ± 1.1 | 0.3 ± 0.0 | 0.3 ± 0.1 |
| Russian | Japanese | 95.8 ± 0.6 | 93.4 ± 0.8 | 0.3 ± 0.1 | 0.5 ± 0.1 |