

Document-Level Machine Translation with Hierarchical Attention

*Experiments with rule-based machine translation*

Presented to Prof. Chris Pollett  
Department of Computer Science  
San Jose State University

In Partial Fulfillment  
Of the Requirements of the Class  
Fall 2022: CS 297

By  
Yu-Tang Shen  
September 2022

## TABLE OF CONTENTS

<b>I. INTRODUCTION ON RULE-BASED MACHINE TRANSLATION.....</b>	<b>1</b>
<b>II. MACHINE TRANSLATION TYPES.....</b>	<b>1</b>
<b>III. EXPERIMENT .....</b>	<b>3</b>
<b>IV. CONCLUSION .....</b>	<b>4</b>
<b>REFERENCES.....</b>	<b>5</b>

## I. INTRODUCTION ON RULE-BASED MACHINE TRANSLATION

Rule-based machine translation (RBMT) is one of the more apparent ways of implementing machine translation (MT) technologies. RBMT follows a set of rules to perform translation, and the difference between swapping every word in the source language (SL) into target language (TL) is more complex rules, such as re-ordering, can be specified in RBMT.

Although the concept of MT was brought up in the seventeenth century, it was until 1933 that George Artsrouni and Petr Sminrnov-Troyanskii published a concrete proposal utilizing a paper tape machine to translate [1]. Most of the pioneers in the MT field, including the one published by George Artsrouni and Petr Sminrnov-Troyanskii, can be categorized as RBMT.

RBMT requires delicate preparation on the rules, where a huge number of rules consist of translating Russian polysemes to English terms in a MT system developed by Erwin Reifler [1]. Even though a great amount of mappings of Russian-English were done, the results were unsatisfactory. Therefore, the research split into two directions, trial-and-error approaches and theoretical approaches, where the former tried to get the immediate working MT systems and the latter aimed to improve MT by scrutinizing the linguistics.

## II. MACHINE TRANSLATION TYPES

Three MT strategies, direct translation, indirect translation, and transfer translation will be discussed in this section.

### *A. Direct translation*

Direct translation utilizes a set of rules that directly translate SL to TL, so developers can easily revise the rules and quickly identify the errors. For example, RAND Corporation project started with basic rules and keep on revising the rules by examining the translation generated by the rules.

However, the disadvantage of this strategy is the exponential growth on the set of rules: for translation between  $N$  languages,  $N(N - 1)$  sets of rules are required, which grows exponentially.

### ***B. Indirect translation***

Indirect translation adds an additional interlingua (IL) layer between two languages so that less set of rules are required for multi-language MT. Instead of translating the source language to the target language, SL is first mapped to IL, which preserves the semantic information, and TL is generated from IL information. With this approach, merely  $2N$  sets of rules are required for translation between  $N$  languages.

However, an universal interlingua might fail to abstract the transition between every pair of languages, because two TLs might be inherently different and it would be challenging to synthesize two divergent languages with a same piece of IL.

### ***C. Transfer translation***

Transfer translation attaches an extra layer of IL from the indirect translation schema, so the two layers of IL act as abstractions of SL and TL respectively. In this approach, it resolves the challenge of synthesizing different TLs with a same piece of IL. Instead of synthesizing TL from the abstraction of SL, which is the first IL, an additional transition to convert abstracted information from SL to TL is deployed. Therefore, the second IL will act as seed information to generate the TL.

### III. EXPERIMENT

Simple experiments are conducted to evaluate the strengths and weaknesses of RBMT. The experiments are done with Universal Rule-Based Machine Translation toolkit [2].

#### A. Experiments with one-to-one translations

In this configuration, the relationships between SL to TL is one-to-one without reordering, and this experiment shows satisfying results.

('I love good dogs', '我愛好狗')  
( 'I hate bad dogs', '我討厭壞狗' )

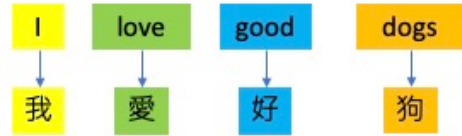


Fig 1. Tuples of translations from English to Traditional Chinese

Fig 2. One-to-one relationships without reordering

The translations from the experiment are correct. Due to its simplicity, once the rules are set correctly, the translation will be satisfactory.

#### B. Experiments with one-to-many translations

Knowing simple-structured sentences can be correctly translated, a more sophisticated test is shown in this section. One of the shortcomings of RBMT is the inability to provide more than one translation from a same word.

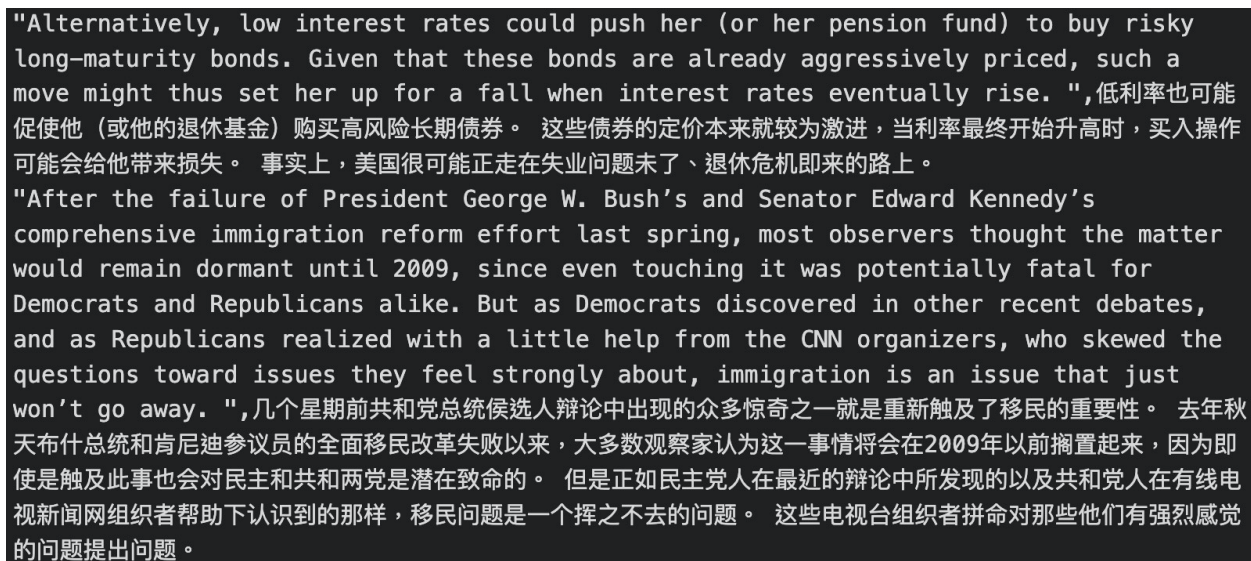
('They make perfect steaks', '他們做完美的牛排')  
( 'They make a cute couple', '他們做一可愛的情侶' )

Fig 3. Tuples of translations from English to Traditional Chinese

In the two source sentences in Fig3., the two 'make's mean differently. Since the rules must be deterministic, RBMT failed to resolve different meanings in these two sentences.

### C. Attempts on complex sentence structures

With unsatisfactory results from one-to-many configuration, another experiment is configured to observe the results from more complicated sentence structures such as clauses, conjunctions, etc.



"Alternatively, low interest rates could push her (or her pension fund) to buy risky long-maturity bonds. Given that these bonds are already aggressively priced, such a move might thus set her up for a fall when interest rates eventually rise. ",低利率也可能促使他（或他的退休基金）购买高风险长期债券。这些债券的定价本来就较为激进，当利率最终开始升高时，买入操作可能会给他带来损失。事实上，美国很可能正走在失业问题未了、退休危机即来的路上。

"After the failure of President George W. Bush's and Senator Edward Kennedy's comprehensive immigration reform effort last spring, most observers thought the matter would remain dormant until 2009, since even touching it was potentially fatal for Democrats and Republicans alike. But as Democrats discovered in other recent debates, and as Republicans realized with a little help from the CNN organizers, who skewed the questions toward issues they feel strongly about, immigration is an issue that just won't go away. ",几个星期前共和党总统候选人辩论中出现的众多惊奇之一就是重新触及了移民的重要性。去年秋天布什总统和肯尼迪参议员的全面移民改革失败以来，大多数观察家认为这一事情将会在2009年以前搁置起来，因为即使是触及此事也会对民主和共和两党是潜在致命的。但是正如民主党人在最近的辩论中所发现的以及共和党人在有线电视新闻网组织者帮助下认识到的那样，移民问题是一个挥之不去的问题。这些电视台组织者拼命对那些他们有强烈感觉的问题提出问题。

Fig 4. Complicated sentence structure corpus

A sub-corpus is derived from [3], where 10 sentence pairs are sampled. Due to the complexity of the grammatical transition, the attempts to set rules fail. But by observing the relationship between SL and TL, it is obvious that an in-depth understanding on the linguistic knowledge in both languages is required to create a RBMT system.

## IV. CONCLUSION

As the pioneer of machine translation, RBMT showcases the practicality of translating with machines. Despite the poor translation quality, it provides users a general idea of the information written in languages ones don't understand.

RBMT generates deterministic translations once the rules are set, and that can be a beneficial property in MT systems, where developers can easily adjust the system for wanted effects. On the other hand, the determinism also hinders the ability for the system to provide different translations under different context.

## REFERENCES

- [1] W. J. Hutchins, "Machine translation: A brief history," in *Concise history of the language sciences*, Elsevier, 1995, pp. 431-445.
- [2] T.-P. Nguyen, *URBANS: Universal Rule-Based Machine Translation NLP toolkit*, GitHub, 2021.
- [3] J. Tiedemann, "Parallel Data, Tools and Interfaces in OPUS," in *Proceedings of the 8th International Conference on Language Resources and Evaluation*, 2012.