Journal of Computational Intelligence and Secure Systems of Artificial Intelligence
Volume 01, Issue 03, July 2025
www.jcissai.com

Algorithm for Arabic-English Hybrid Machine Translation Optimisation

P Rehaman Shaik
Department of Computer Science Engineering, KL University, Hyderabad
Corresponding Author: rehamanshaikcse@gmail.com

To Cite this Article
Rehaman Shaik, "Algorithm for Arabic-English Hybrid Machine Translation Optimisation", Journal of Computational Intelligence and Secure Systems of Artificial Intelligence, Vol. 01, Issue 03, July 2025, pp:19-21.

Abstract*: The distinct designs of Rule Based Machine Translation and Statistical Machine Translation make it difficult for a pure machine translation solution to handle all translation obstacles. Once lexical and syntactic analyzers are dealt with using Rule Based, the Expectation–Maximization algorithm is run to resolve certain uncertainties with maximum likelihood. We have designed and explained an integrated hybrid machine translation system in this paper. Fusing the most effective features from different strategies is what is pursued. In the first step, RBMT converts Arabic text to English and then EM is run on the output to improve it and create the final version. To do well and become better, the EM algorithm must be able to accurately translate a frequency from one language to another. The results prove that the new approach works better than the common Rule Based approach and the EM algorithm, as proven by the BLEU system. The survey suggests that scores from the HMT system are regularly greater than those from SMT.*

Keywords*: Machine Translation, Arabic-English Machine, Expectation–Maximization Translation, Hybrid Machine Translation, (EM) Algorithm*

I. Introduction

There are numerous languages in the world, and finding a machine translator that can translate them while yet meeting human standards is challenging. The translation process also encounters a number of linguistic issues, particularly when translating from Arabic to English, such as confusing words, grammatical discrepancies, and word structure. Arabic machine translation issues and evaluation have been approached from a variety of angles [5]. [1] provides a detailed discussion of these methods.

One of these methods is the Rule Based machine translation system, which indicates linguistic rules between the source and target languages. Rule-based systems are able to fully analyse linguistic information at both the semantic and syntactic levels. It takes a lot of linguistic expertise to produce the Rule Based language rules, which makes it expensive to develop. The output of a rule-based system is less fluid than that of a statistical system and is dependent on the accuracy of each level. MT for Romance Languages to Spain, Arabic to English, Indonesian to Malaysian, Bulgarian to Macedonian, English to Sankrit, and many other scholars have employed rule-based approaches as their translation techniques.

II. Hybrid Translation System Between English and Arabic

The emphasis on hybrid machine translation systems has been a significant trend in recent years. Combining tools and methods from various technological backgrounds—such as rule-based and statistical approaches—is the goal of these systems. This study presents an Arabic-English HMT system. In general, this system consists of two parts—a statistical and a rule-based section. Medical terminology is converted from Arabic to English using the RBMT Parser, while SMT copes with confusion in the language by accessing huge corpora. This approach is focused on linguistic aspects of the two languages from dictionaries and looks at rule usage just like in transfer system machine translation.

This type of machine translation works from translation rules, not from a dictionary, so it is considered a knowing system. A sentence is translated literally from the dictionary if its form meets one of the rules. The new sentence type is based on the examination of the morphology as well as the syntax of the original sentence. Then, the

translation is switched to the target language according to its rules and then a better translation is made for the final result.

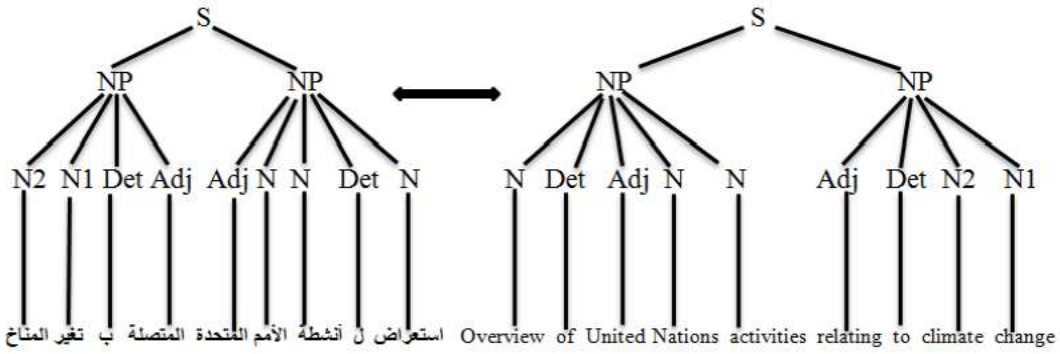

Fig 1: Structure tree

III. Implementation

Arabic to English has been the subject of the experiments. The hybrid approach depends on the structures produced by the statistical and rule-based systems. For every input sentence, a rule-based machine translation technique produces a candidate sentence. It is crucial to choose which Arabic word is appropriate for the source sentence because a single word can have multiple meanings in English. As a result, this produces candidate sentences for every word meaning.

It generates grammatically correct translated English candidate sentences, although some sentences may be useless because of the unclear word meanings. Out of those possible sentences, the EM algorithm will select the best one. SMT techniques are applied in this situation. Find the number of fuzzy words and the number of multiple meanings in the rule side of the language rule, then use this to create potential versions of the source text. A source sentence can be created in several ways when the words' meanings are not certain.

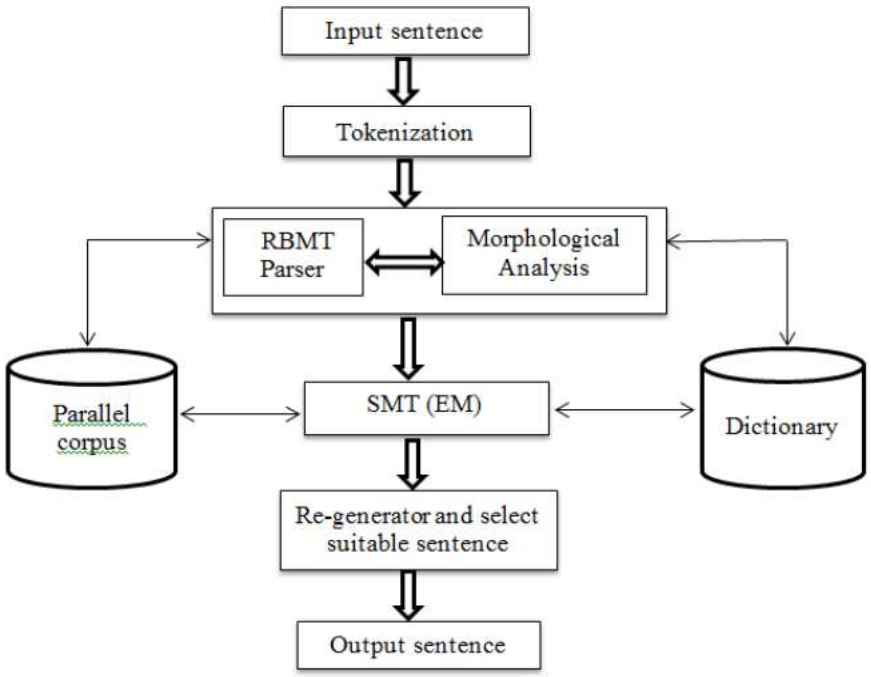

Fig 2: Architecture of HMT system

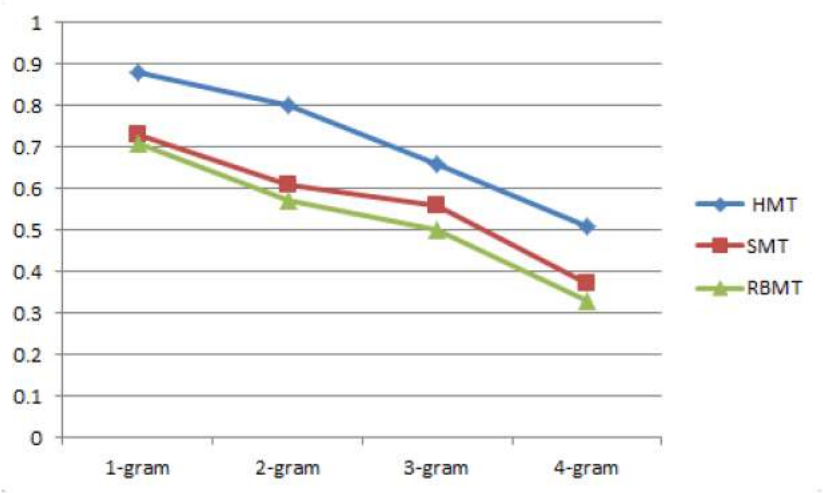

Fig 3: score of SMT, HMT and RBMT

IV. Conclusion

This research integrates rule-based and statistical methods into one framework which has allowed us to design a new way to link these methods in HMT. The approach tries to make our scientific outcomes better than with any method in use today. It allows the native strength of statistics to be combined with rules related to language used in rule-based translation models. The model addresses the lexical and syntactic analysis requirements in SMT as well as the language ambiguity issue that is the greatest challenge in RBMT. Because of this, we presented how our hybrid system was built, using the EM algorithm for statistics and drawing inspiration from rules. There are more languages, a bigger scope and multiple layers of trials within the system. The focus of this study was the United Nations parallel corpus from Arabic to English. This research aims to improve how translations are produced by joining the best parts of every MT system.

References

[1] Sokamso Tayang, Hubert and Deepak , Fuzzy Neural Networks: A Review from UPQC and Fuzzy Logic Statistical Perspective. Statistical science. 1998; 2(7): 12-18.

[2] Raju, Venkatesh, Hubert and Alwyn " Multilayer neural networks and PQ theory decision theory. *Science Direct*. 1991: 156-165.

[3] Guoqiang Peter Zhang. Neural Networks for Classification: A Survey. *IEEE Trans.On Man, Systems, and Cybernetics*. 2000; 30(4).

[4] G Rothermel, R Untch, C Chu, MJ Harrold. Prioritizing Test Cases for Regression Testing. *IEEE Trans. Software Eng*. 2001; 27(10): 929-948.

[5] P. C. Chang and C. H. Liu, "A TSK type fuzzy rule based system for stock price prediction," *Expert Syst. Appl*., vol.34, pp. 135–144, Jan 2008.

[6] SK Abdul Rehaman, John Vangli and Shanli, "Neural network system combined with Fuzzy-rough data reduction with ant colony optimization," *Fuzzy Set Syst*., vol. 231, pp. 56-65, March 2010.

[7] Chen Chen-Hung "A unctional-Link-Based Neurofuzzy Network for Nonlinear System Control"- *IEEE Transaction on Fuzzy Systems*, Vol 16 No 5, October 2008.

[8] Manju Bargavi, Siddharda Roy and Mohit Reddy, "ANN and FLANN Based Forecasting for Conceptual S&P 500 Index"*Information Technology Journal*, 6 (1): 121-132, 2010 Asian Network System for Artificial Scientific Information.