Assessment of Multi-Engine Machine Translation for English to Hindi Language (MEMTEHiL): Using F&A and iBLEU Metrics

Pankaj K. Goswami, Amity University, Noida, India
Sanjay K. Dwivedi, Babasaheb Bhimrao Ambedkar University, Lucknow, India
C. K. Jha, Banasthali University, Rajasthan, India

ABSTRACT

English to Hindi translation of the computer-science related e-content, generated through an online freely available machine translation engine may not be technically correct. The expected target translation should be as fluent as intended for the native learners and the meaning of a source e-content should be conveyed properly. A Multi-Engine Machine Translation for English to Hindi Language (MEMTEHiL) framework has been designed and integrated by the authors as a translation solution for the computer science domain e-content. It was possible by enabling the use of well-tested approaches of machine translation. The humanly evaluated and acceptable metrics like fluency and adequacy (F&A) were used to assess the best translation quality for English to Hindi language pair. Besides humanly-judged metrics, another well-tested and existing interactive version of Bi-Lingual Evaluation Understudy (iBLEU) was used for evaluation. Authors have incorporated both parameters (F&A and iBLEU) for assessing the quality of translation as regenerated by the designed MEMTEHiL.

KEYWORDS

Evaluation of Machine Translation, Fluency & Adequacy (F&A), Interactive Bi Lingual Evaluation Understudy (iBLEU), Multi Engine Machine Translation (MEMT)

INTRODUCTION

Machine Translation (MT) is the solution which jointly developed by the computer and linguistic scientists. It translates the user’s text from one natural language to the other without human intervention.

DOI: 10.4018/IJALR.2016010102

Copyright © 2016, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
It has been observed that Machine Translation (MT) of technical e-contents of computer science domain, which is acquired from the web does not produce adequate quality of translation from English to Hindi. It is very generic that the best-quality of machine translated e-contents can only be evaluated by the native evaluator. The evaluator has to be conversant with both source and target languages. The assessment of any machine translated e-contents, performed by a human evaluator have been truly menial, time taking and not a cost-effective approach. It may be stretched up to the couple of days or months to finish the task. The pivotal disadvantage of such assessment is that the human efforts made during this process cannot be re-used (Papineni, 2002).

A Multi-Engine Machine Translation for English to Hindi (MEMTEHiL) is a translation and evaluation framework designed for the native learner of Hindi. It has been designed and developed by us, to translate technical e-contents of computer science domain from English to Hindi. This framework is influenced by some cardinal features of the existing Multi-Engine Machine Translation (MEMT) for other language pairs. MEMTEHiL has inbuilt key features of multiple contributor engines. It has been incorporated three Machine Translation (MT) engines and one evaluator engine, which is used as a validation engine. The evaluator engine validates the result by using a (iBLEU) metric. MEMTEHiL works on generated outputs of each contributor machine translation engine. Firstly, each output examined by a human evaluator by using fluency and adequacy metrics and then post-editing plays an important role to achieve the best final result.

The designed framework and algorithm of MEMTEHiL were empirically tested for the computer science e-content. The encouraging results in this domain proved the suitability of the designed framework. (Goswami et al., in press).

The assessment of Multi Engine Machine Translation for English to Hindi (MEMTEHiL) for the computer science domain e-content has been done by Fluency and Adequacy (F&A). It is a humanly evaluated metric which contains user’s acceptance. The same translation has been further validated interactively by an online engine, i.e. interactive Bi-Lingual Evaluation Understudy (iBLEU). It’s an automatic statistical evaluation metric. Through this method, the validation of previous human evaluation has also been done.

The subsequent sections are organized as to cover all necessary aspects of this research paper. It covers the challenges for machine translation evaluation, Fluency and Adequacy (F&A) metrics, brief details about the evaluators of fluency and adequacy. Furthermore, an evaluation metric, Bi Lingual Evaluation Understudy (BLEU), an interactive BLEU (iBLEU) and its scoring mechanism are also explained. The details of participating source e-contents in assessment as well as about the contributor engines of MEMTEHiL have been covered. An evaluation of MEMTEHiL with test result analysis, using (F&A) and iBLEU metrics has been represented. Finally, conclusion and prospect have been briefly discussed.
Related Content

A Review on Time Series Motif Discovery Techniques an Application to ECG Signal Classification: ECG Signal Classification Using Time Series Motif Discovery Techniques
[www.irma-international.org/article/a-review-on-time-series-motif-discovery-techniques-an-application-to-ecg-signal-classification/238127](www.irma-international.org/article/a-review-on-time-series-motif-discovery-techniques-an-application-to-ecg-signal-classification/238127)

Symmetric Encryption Algorithm Inspired by Randomness and Non-Linearity of Immune Systems
[www.irma-international.org/article/symmetric-encryption-algorithm-inspired-randomness/72872](www.irma-international.org/article/symmetric-encryption-algorithm-inspired-randomness/72872)

Aroma Chip Using a Functional Polymer Gel
Dong Wook Kim (2013). *Human Olfactory Displays and Interfaces: Odor Sensing and Presentation* (pp. 384-400).
[www.irma-international.org/chapter/aroma-chip-using-functional-polymer/71935](www.irma-international.org/chapter/aroma-chip-using-functional-polymer/71935)

Multi-Objective Higher Order Polynomial Networks to Model Insertion Force of Bevel-Tip Needles
A Computational Comparison of Swarm Optimization Techniques for Optimal Load Shedding Under the Presence of FACTS Devices to Avoid Voltage Instability
