Synopsis

Translation of English to Mauritian Creole

By: Muhammad Arshad Goburdhun

ID Number: 1220377

MSc Software Engineering Projects and Management

Table of Contents

[1. Tentative Title 2](#_Toc377080364)

[2. Introduction 3](#_Toc377080365)

[2.1. Natural Language Processing 3](#_Toc377080366)

[2.2. Applications of Natural Language Processing 3](#_Toc377080367)

[3. Statement of the Problem 5](#_Toc377080368)

[4. Rationale of the Study 5](#_Toc377080369)

[5. Objectives of the Study 5](#_Toc377080370)

[6. Brief literature Review 5](#_Toc377080371)

[6.1. Machine Translation approaches 5](#_Toc377080372)

[7. Existing tools for Machine Translation 6](#_Toc377080373)

[7.1. Babelfish 6](#_Toc377080374)

[7.2. Bing Translator 6](#_Toc377080375)

[7.3. Google Translate 7](#_Toc377080376)

[8. Methodology 7](#_Toc377080377)

[9. Expected Output 7](#_Toc377080378)

[10. Research plan (Time Frame)/Activity (Gantt) Chart 8](#_Toc377080379)

[References 9](#_Toc377080380)

# Tentative Title

Translation of English to Mauritius Creole

# Introduction

The translation of English to Mauritian Creole is part of the field of Natural Language Processing. It is also known as Machine Translation which is the task of automatically converting one natural language into another, preserving the meaning of the input text, and producing fluent text in the output language. Machine Translation is one of the oldest subfields of artificial intelligence research.

## Natural Language Processing

Natural Language Processing (NLP) is an area of research and application that explores how computers can be used to understand and manipulate natural language text or speech to do useful things. NLP researchers aim to gather knowledge on how human beings understand and use language so that appropriate tools and techniques can be developed to make computer systems understand and manipulate natural languages to perform the desired tasks. The foundations of NLP lie in a number of disciplines, like computer and information sciences, linguistics, mathematics, electrical and electronic engineering, artificial intelligence and robotics, psychology, etc.

## Applications of Natural Language Processing

Applications of NLP include a number of fields of studies, such as machine translation, natural language text processing and summarization, user interfaces, multilingual and cross language information retrieval (CLIR), speech recognition, artificial intelligence and expert systems.

In our case we will be focusing on the Machine Translation part of NLP. Machine translation is the process of translating from source language text into the target language. The following diagram shows all the phases involved.

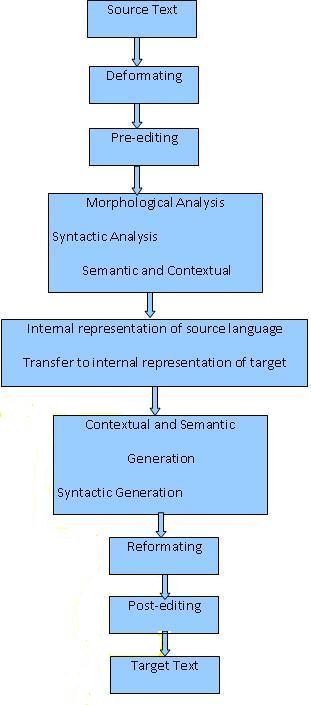


Figure 1. Phases of Machine Translation

# Statement of the Problem

On the web today, there are many tools for translating different languages. For example we have Google Translate. The latter tool has translations to many different language but however it does not have any options to translate English to Mauritian Creole.

# Rationale of the Study

The study will be undertaken to better understand the needs of having English to Mauritian Creole translation in the Mauritian society and also to better know the targeted audience for such a tool in Mauritius. A survey will be conducted amongst different people coming from different backgrounds and having different IT skills. The survey will help us know to what extent people are using Machine Translation systems and the amount of times they use such tools.

# Objectives of the Study

For the study, we will be performing research on existing translation tools already available. Also we will compare different algorithms for Natural Language Translation. Finally at the end of the project, I expect to have a tool that will help to perform translation of English language to Mauritian Creole language.

# Brief literature Review

## Machine Translation approaches

Generally, MT is classified into seven broad categories: rule-based, statistical-based, hybrid-based, example-based, knowledge-based, principle-based, and online interactive based methods. The first three MT approaches are the most widely used and earliest methods.

# Existing tools for Machine Translation

## Babelfish

Yahoo! Babel Fish was a free web-based multilingual translation application. In May 2012 it was replaced by Bing Translator, to which queries were redirected. Although Yahoo! has transitioned its Babel Fish translation services to Bing Translator, it did not sell its translation application to Microsoft outright. Babel Fish is considered the oldest web translation site, as it superseded AltaVista in 2003 and has remained active since its re-location in Bing Translator. As the oldest free online language translator, the service translated text or web pages between 38 languages, including English, Simplified Chinese, Traditional Chinese, Dutch, French, German, Greek, Italian, Japanese, Korean, Portuguese, Russian, Swedish, and Spanish.

The internet service derived its name from the "Babel fish", a fictional species in Douglas Adams's series The Hitchhiker's Guide to the Galaxy that could instantly translate languages. In turn, the name of the fictional creature refers to the biblical account of the confusion of languages that arose in the city of Babel.

## Bing Translator

Bing Translator (previously Live Search Translator and Windows Live Translator) is a user facing translation portal provided by Microsoft as part of its Bing services to translate texts or entire web pages into different languages. In addition to standard text and web page translation, Bing Translator includes several additional features:

* When translating an entire web page, or when the user selects "Translate this page" in Bing search results, the Bilingual Viewer will be shown, which allows users to browse the original web page text and translation in parallel, supported by synchronized highlights, scrolling, and navigation
* Four Bilingual Viewer layouts are available:
  + Side by side
  + Top and bottom
  + Original with hover translation
  + Translation with hover original
* Website owners can add a translation widget to their website for translating it into other languages supported by Bing Translator; this is done by inserting an HTML code snippet on the web page
* Any-to-Any language translation pairs
* Automatically detect the language of the text or website being translated
* Ability to easily reverse the translation direction
* The user can play back a spoken version of the translation through text-to-speech (not supported in every language)

## Google Translate

Google Translate is a free translation service that provides instant translations between dozens of different languages. It can translate words, sentences and web pages between any combination of our supported languages. With Google Translate, we hope to make information universally accessible and useful, regardless of the language in which it’s written.

When Google Translate generates a translation, it looks for patterns in hundreds of millions of documents to help decide on the best translation for you. By detecting patterns in documents that have already been translated by human translators, Google Translate can make intelligent guesses as to what an appropriate translation should be. This process of seeking patterns in large amounts of text is called "statistical machine translation". Since the translations are generated by machines, not all translation will be perfect. The more human-translated documents that Google Translate can analyze in a specific language, the better the translation quality will be. This is why translation accuracy will sometimes vary across languages.

* More tools and applications will be discussed in details in the Project documentation.

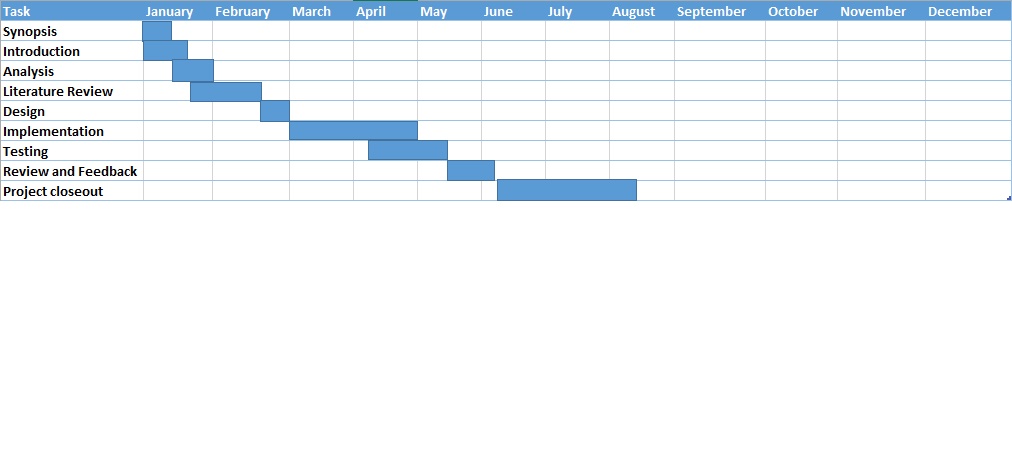
# Methodology

For the methodology, we will be conducting research on existing tools and algorithms. Based on the findings, we will be working on an algorithm to implement the translation of English to Mauritian Creole and try to improve on the existing algorithms.

# Expected Output

At the end of the project, we will be expecting to have a working tool to perform translation of English to Mauritian Creole.

# Research plan (Time Frame)/Activity (Gantt) Chart



# References

Machine Translation: The Stanford Natural Language Processing Group [Online] Available from: <http://nlp.stanford.edu/projects/mt.shtml> [accessed on 8th January 2014]

Gobinda G. Chowdhury. Natural Language Processing. *Dept. of Computer and Information Sciences*

*University of Strathclyde, Glasgow G1 1XH, UK*

Sneha Tripathi and Juran Krishna Sarkhel. Approaches to Machine Translation. *Annals of Library and Information Services. Vol. 57, December 2010, pp 388-393*

Yahoo! Babel Fish: From Wikipedia, the free encyclopedia [Online] Available from: <http://en.wikipedia.org/wiki/Yahoo!_Babel_Fish> [Accessed on 9th January 2014]

Bing Translator: From Wikipedia, the free encyclopedia [Online] Available from: <http://en.wikipedia.org/wiki/Bing_Translator> [Accessed on 9th January 2014]

Find out how our translations are created. Inside Google Translate [Online] Available from: <http://translate.google.mu/about/intl/en_ALL/> [Accessed on 9th January 2014]

University of Pennsylvania. Department of Computer and Information Science. RESEARCH IN NATURAL LANGUAGE PROCESSING Available from: <http://acl.ldc.upenn.edu/H/H86/H86-1005.pdf> [Accessed on 9th January 2014]

MIRELLA LAPATA and FRANK KELLER. Web-based Models for Natural Language Processing.

*University of Edinburgh.* *ACM Transactions on Speech and Language Processing, Vol. 2, No. 1, February 2005, Pages 1-30*

Antony P. J. Machine Translation Approaches and Survey for Indian Languages.

*Computational Linguistics and Chinese Language Processing Vol. 18, No. 1, March 2013, pp. 47-78*