

## MODELING AN INTERLINGUA FOR SEMANTIC MACHINE TRANSLATION

M.J. KUMARA AND S.R. KODITUWAKKU

*Department of Statistics and Computer Science, Faculty of Science,  
University of Peradeniya*

Machine translation, referred to by the acronym MT, is a subfield of computational linguistics that investigates the use of computer software to translate text or speech between natural languages. At its basic level, MT performs simple substitution of atomic words from one natural language to words in another. Using corpus techniques, more complex translations can be performed, allowing for better handling of differences in linguistic typology, phrase recognition, and translation of idioms, as well as the isolation of anomalies. However, current systems are unable to produce a same quality output as a human translator, particularly where the text to be translated uses casual language.

When translation involves more than two languages it is necessary to repeat the source language analysis process for each translation. It is time consuming and inefficient in terms of resources and using the same knowledge base repeatedly. The Interlingua is the best solution for these problems. Interlingua first analyses the source sentence and then builds the intermediate representation which can be used to translate to more than two languages without analysing the source sentence more than once. The following figure gives a summary of the translation process.

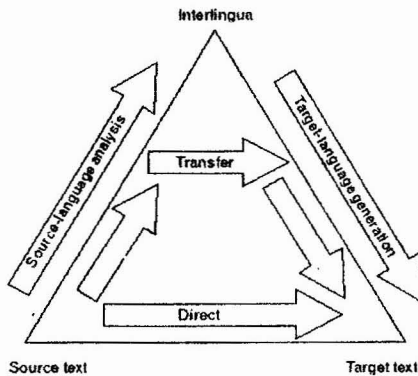


Figure: Interlingua representation