A Knowledge-Based Machine Translation Using AI Technique

Sahar A. El and Rahman, Benha University, Cairo, Egypt & Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia
Tarek A. El-Shishtawy, Benha University, Banha, Egypt
Raafat A. El-Kammar, Benha University, Banha, Egypt

ABSTRACT

This article presents a realistic technique for the machine aided translation system. In this technique, the system dictionary is partitioned into a multi-module structure for fast retrieval of Arabic features of English words. Each module is accessed through an interface that includes the necessary morphological rules, which directs the search toward the proper sub-dictionary. Another factor that aids fast retrieval of Arabic features of words is the prediction of the word category, and accesses its sub-dictionary to retrieve the corresponding attributes. The system consists of three main parts, which are the source language analysis, the transfer rules between source language (English) and target language (Arabic), and the generation of the target language. The proposed system is able to translate, some negative forms, demonstrations, and conjunctions, and also adjust nouns, verbs, and adjectives according their attributes. Then, it adds the symptom of Arabic words to generate a correct sentence.

KEYWORDS

Arabic Machine Translation, Artificial intelligence, Knowledge-Based Approach, Knowledge-Based Machine Translation, Machine Translation

1. INTRODUCTION

Machine Translation (MT) is a translation between languages, but not with a human tongue, but by computer programming with a satisfactory accuracy in a minimum possible time (Allen & Trucker, 1992). It is undoubtedly an important topic socially, politically, commercially, scientifically, and intellectually or philosophically (Arnold, Balkan, SMeijer, Humphreys, & Sadler, 1994).

MT is widely spread over the world between different language such as (French, English, and Arabic), but the Arabic language is not considered efficiently, because of the complexity of the Arabic language (Al-Khonaizi, Al-Ali, & Al Zobaidie, 1994). There are Very few systems translate from and to Arabic language, but not perfect and with the lack especially in the generation of Arabic language. They lack an open structure of the underlying theoretical basics of Arabic language. This lack does not give a good picture of the state of MT that involves Arabic language. It was for this reason that we decided to begin a system for an opened structure MT involving Arabic.
2. APPROACHES TO MACHINE TRANSLATION

Historically, the strategies for Machine Aided Translation fall into the following categories (Hutchins & Somers, 1992):

2.1. Direct Approach

The translation system was reduced to word-for word substitution (El-Hamayed, 1992). The direct approach works best when the two languages are both lexically and structurally similar.

2.2. Indirect Approach (Thunes, 1993)

It can be summarized as:

- Interlingua approach (Blekhman, Kursin, & Fagradiants, 1998; Bel et al., 1999): Source text --[analysis]-- Interlingua --[synthesis]-- Target text
- Transfer approach (Gimenez & Forcada, 1998; El-Shishtawy, 1997) Source text –[analysis]—Intermediate Structure (source)–[transfer]—Intermediate Structure (target)–[synthesis]—Target text

2.3. Knowledge-Based Approach

There is a fairly new approach called knowledge-based machine translation (KBMT) (Carbonell Cullingford, & Gershman, 1981). Successful translation by machine requires the use of various types of knowledge (hence the term "Knowledge Based Machine Translation") for each language (Leavit, Londale, & Franz, 1994). This includes spelling, contraction, and formatting rules; morphological rules; lexical knowledge, including syntactic features, semantic concepts, collocation and terminological information; knowledge about grammatical structure; and semantic rules (Elkateb & Black, 2001; El-Azim, 1994).

3. THE PROPOSED SYSTEM

In system implementation, the” knowledge-based “and “transfer” approaches are used. Where, The knowledge-based approaches discards the syntax of the original text, and the output is an interpretation rather than translation (El-Shishtawy, 1997). The current system relies on the syntax of the input text, and provides it with semantic information as necessary to solve specific ambiguity (Choi, Kim, Lee, & Park, 1997). I.e. we used an approach which makes integration between two approaches to get an Arabic sentence with correct syntax and semantics. This approach is used to achieve the stated requirements of the proposed translation system. The proposed system consists of two parts that are system resources and the processing stages as illustrated in Figure 1.

3.1. System Resources

In this section, we review the resources used through our Machine Translation System. The main resources are:

3.1.1. Online Dictionary

The proposed system relies on a rich on-line dictionary in all phases of the translation system. The used dictionary was designed and prepared previously for ACLP (Arabic Computational Linguistic Project) machine translation (2000). The dictionary contains about 40,000 English words and their corresponding Arabic words.

In The dictionary, each English word has a grammatical category (as noun, verb, etc.) and its Arabic translation. The dictionary also includes <Previous words> and <next words> that
Differentiated Process Support for Large Software Projects
[www.igi-global.com/chapter/differentiated-process-support-large-software/29511?camid=4v1a](www.igi-global.com/chapter/differentiated-process-support-large-software/29511?camid=4v1a)

To Prevent Reverse-Engineering Tools by Shuffling the Stack Status with Hook Mechanism