

AMBIGUITY IN MACHINE TRANSLATION

Anuradha Pandey *Jawaharlal Nehru University*

Abstract

Machine translation is a process of translation in which a machine translates any input that we provide it. In this process of translation, we create a machine language to understand the machine. Through this language the translation process becomes complete. Some examples of such machine languages are COBOL, C+, C++, C# etc. The machine translation process involves some important issues that we need to understand. In machine translation, literary translation is a great problem. The machine cannot translate a literary text very properly. In a literary text translation, there can be many problems, amongst them, Ambiguity is a great problem. In this paper, we try to understand Ambiguity in machine translation.

Keywords:

Machine
Translation
Kinds of Machine
Translation
Ambiguity and
Kinds of
ambiguity

Machine translation

Translation is not just a process of the word to word translation; it is based on meaning and thought. In machine translation, we have to collect the proper and complete information about a single word or sentence. In languages like English, Spanish, etc. machine translation is completely possible only in some fields like science, technology, law, and medicine. It is possible because in these fields literal translation is accepted. But when it comes to literature, machine translation is extremely difficult irrespective of the language. There is not a single country which has managed to develop any machine language software that can perform literary translation with hundred-percent proficiency.

In literature, machine translation becomes difficult because literature's language is a collection of indirect meanings, satirical sentences, multi-word expressions, and ambiguities. And a machine cannot understand such types of words, sentences, phrases, and meanings. There should be all information about a word at the level of synthesis, analysis, and processing in machine translation. A word dictionary plays a very important role in machine translation.

For example, take the English word 'Boy'. Its Hindi meaning is- 'लड़का'. It is a

noun. It is a living thing, human being, masculine, singular person, etc. All this information should be given then only in the translation process of 'Boy' the machine can understand it properly. This type of dictionary can be very helpful in machine translation.

In Indian languages such dictionaries are is not yet available. But engineers and linguistics are working in this field to get hundred-percent purity in machine translation's output. There are four kinds of machine translation, which are assisted by humans.

1. Human translation
2. Machine-aided human translation
3. Human-aided machine translation &
4. Machine translation.

Human translation is fully processed by a human; there is no role of the machine. In this type of translation, the human translates a natural language text. Human translator plays an important role in the translation of a literary text.

Machine-aided Human Translation is a process of translation in which a human uses a machine as a helping tool. For example- in the translation of non-literature text like technology, medical, official and scientific text's, human uses a glossary of such fields and literal translation is possible in these fields.

Human-aided Machine Translation is a process of translation in which a human helps a machine in the translation process. For example- when a machine cannot understand the meaning of any word, human helps the machine to understand it through simple explanation and analysis, and the machine translation becomes complete. This type of translation is aided by humans. Human plays a key role in it.

Machine translation is fully processed by a Machine. The machine plays a key role and there is no presence of humans in it. This type of translation is not possible - yet. Experts are still working on it and they may become successful in some areas. In future, it might be possible in all natural languages.

The difficulties in machine translation can arise at many levels, such as grammatical, word, sentences, phrases, meanings, and communication level. Accordingly, problems associated with machine translation can be divided into two categories.

First one is the translation between two natural languages of the same culture, like- Hindi to Marathi, Marathi to Bengali and Gujarati etc. These natural languages belong to the same cultural areas, these are the Indian Languages.

The second one is the translation between two natural languages belonging to two different language families or cultural areas, like English to Hindi or Hindi to French, Spanish, etc. These languages are very different in meaning both socially and culturally. This is a great form of difficulty in the process of machine translation.

Ambiguity in Machine translation

Ambiguity is a great problem in the translation process as it creates lots of error in translation at various levels, like the word to phrases and in the level of social and cultural meaning. As we all know that a natural language is a collection of human languages which are in use in communication between human beings. Throughout the world, there are many natural languages that are in daily use. Each one is very different from others. When we start translation between such languages Ambiguity creates many problems in the translation process.

Hindi is an Indian language and English is a western language. Both languages are different from each other, socio-culturally and grammatically. Both of them have many forms of uses like- official, trade, educational, political, medical and literary too. So there are many forms of languages. But there come a lot of problems at the level of literary form. Every language has social and cultural meanings too.

Kinds of Ambiguity

There is no single kind of Ambiguity. Linguists are defining it in different ways. In the field of linguistics, there are two kinds of Ambiguity.

1. Lexical Ambiguity.
2. Structural Ambiguity.

In the field of Natural Language processing, there are five kinds of Ambiguity-

1. Lexical Ambiguity.
2. Structural Semantic Ambiguity.
3. Syntactic Ambiguity.
4. Discourse Ambiguity.
5. Missing Information Ambiguity.

In the field of machine translation or in the process of machine translation ambiguity creates problems at the level of Phoneme, Morpheme, Word, Sentences, Phrases, Semiotics, Pitch, Stress, Juncture, Prefix, Suffix & etc.

Phoneme-level Ambiguity- (दिन और दीन) in this word Ambiguity is based on Phoneme level. The meaning of this word has been different from each other just because of the phone. The first one 'Din' is 'Day' and the second one 'Deen' is 'Poor'.

Morpheme level Ambiguity- (खाना और खाना) in this word Ambiguity is based on Morpheme level. Both word formations are the same and its sounds same, but it represents different meanings, like the Food and Counter.

Word level Ambiguity- (सोना और सोना) in this example, both are same but represents different meaning. The first ones meaning is 'sleeping' and the second is 'Gold'. In the English language this type of problem is also very common. Example- Back means (पीछे और पीठ) and Bank means (बैंक और किनारा). The problem is that only a human can understand it while the machine can never solve it alone.

Sentence level Ambiguity- (उसने जाते हुए घोड़े को देखा) in these sentences there comes an ambiguity problem at the sentence level. In this sentence this is not clear that which one is moving, the horse or the observer.

Stress level Ambiguity- The style of pronunciation also creates ambiguity in different ways. For example-

You say Sorry to her.

You say sorry to her.

You say **sorry** to her.

Pitch level Ambiguity- in the pronunciation of any sentences when we put emphasis on words in different ways in the same sentence, as it represents different meanings.

Juncture level Ambiguity- when we pronounce any word in very fast or very roughly. In this way, a sentence represents a different meaning. It creates confusion. For example-

यह गा ली है. यह गाली है.

सूरज की धूप पीली है. धूप पी ली है.

Conclusion

This present time is known as the time of technology. It is possible just because of computers. Today, computers play a very important role in our daily life. We are very deeply related to it. Today we can't survive without a mobile phone or social networking. These are our daily needs. Just like such things, translation is also very important in our life. When we go out from our country to another country we need translators in case we do not understand that country's language. If we are very fond of literature and if we want to know about other country's literature we have to take the help of a

translation.

Today, in the field of translation new aspects of research are emerging daily. People want to know about more things in a very short time. Researchers and language engineers are working in this field to solve lots of problems. Ambiguity is a great problem in machine translation, but linguists and engineers are working to solve it. In future very soon there will be a success.

References:

1. Blaheta, Don and Johnson, Mark. (2001). Unsupervised learning of multi-word verbs. Brown University.
2. Dyre, Christopher James. (2012). A Formal Model of Ambiguity and its Applications in Machine Translation. Bibliobazaar Publisher.
3. Munday, Jeremy, and Basil Hatim. (2004). Translation: An Advanced Resource Book. Routledge: London and New York.
4. मल्होत्रा, विजय कुमार. (2013). कंप्यूटर के भाषिक अनुप्रयोग. वाणी प्रकाशन: दिल्ली.
5. सिंह, सूरजभान. (2006). अंग्रेजी -हिंदी अनुवाद व्याकरण. नई दिल्ली: प्रभात प्रकाशन.

Contributor details :

Anuradha Pandey

Ph.D. Research Scholar

Hindi Translation

Jawaharlal Nehru University

Email- anu.pantran@gmail.com

