The Development of Language Processing Products and its Implications

Neetu Seth

Assistant professor in English

RR Bawa dav college for Girls, Batala.

Introduction

Mathematics has impacted upon each and every discipline of social sciences and language being one of the social sciences, is no exception. Language products have occupied sizeable chunk of the global economy. They have integrated inseparably into social-commercial milieu. Be it Machine Translation System, Speech Recognition System Dictionaries, or Optical Recognizers – they are omnipresent in every home and office. Interestingly, each and every language produce is actuated by some mathematical formulas running at backend. While mathematicians develop mathematical algorithms and formulas to support backend, linguists support the development of corpora for use in developing language products.

Global demand for Language Products

Language products have global demand in corporate world and in homes as well. For example, Microsoft Word is the most widely used word processing software. Microsoft estimates that over 500,000,000 people use the Microsoft Office suite.[1] Similarly, Translation generates by far the largest share of revenues in the industry. It represents more than 80% of all the revenues of the firms. Translation generates by far the largest share of revenues in the industry. It represents more than 80% of all the revenues of the firms.[2]

---

1 <http://en.wikipedia.org/wiki/Word_processor>

2 <http://www.uottawa.ca/associations/csict/represum.pdf.>

Definitional outline and Nature of Language Products

Language products are those products which have linguistic features either in terms of speech or text, reprography or translation. They are developed by the interdisciplinary collaboration of mathematicians and linguists. Whereas mathematicians extend mathematical algorithms

and formulas to support backend, linguists creates corpora for use in developing language products.

Nature of a language products varies from product to product. Some are translator in nature, while other are reprographic in nature. Likewise, some have text-processing properties while other have speech-processing properties.

Scope of Language Products

There is a wide spectrum of language products ranging from word processors to dictionaries, machine language translation to speech recognition system to Optical Character Recognition which can be expatiated upon in the manner specified below:

Dictionary

Dictionary is the most important language product. It is packaged in software. It is also available online. It has monolingual or bilingual version. It is highly demanded product. Its speech-oriented properties make it even more popular. Behind its textual & speech-oriented properties lies the mathematics which, not only, lends it pronunciation but make search possible through its database as well.

Machine Translation

Machine Translation is yet another language product. It translates text or speech from one natural language to another. Complex mathematical formulas work behind machine language translation system. It is highly demanded product in corporate world where communication is to be effected globally in multilingual milieu.

On a basic level, MT performs simple substitution of words in one natural language for words in another, but that alone usually cannot produce a good translation of a text, because recognition of whole phrase and their closest counterparts in the target language is needed. Solving this problem with corpus and statistical techniques is a rapidly growing field that is leading to better translations, handling differences in linguistic typology, translation of idioms, and the isolation of anomalies.[3]

Speech Recognition System

Speech Recognition System is yet another language product which is highly demanded in corporate and academic world. **Speech recognition** is the translation of spoken words into text. It is also known as "automatic speech recognition", "ASR", "computer speech recognition", "speech to text", or just "STT".[4]

_____

3 <http://en.wikipedia.org/wiki/Machine_translation>
4 <http://en.wikipedia.org/wiki/Speech_recognition>

Automatic speech recognition is a process by which a computer takes a speech signal (recorded using a microphone) and converts it into words.[5] It helps auto-type texts. It has application in voice recognition and password protection.

Speech recognition applications include voice user interfaces such as voice dialing (e.g., "Call home"), call routing (e.g., "I would like to make collect call"), demotic appliance control, search (e.g., find a podcast where particular words were spoken), simple data entry (e.g., entering a credit card number), preparation of structured documents (e.g., a radiology report), speech-to-text processing (e.g., word processors or emails), and aircraft (usually termed Direct Voice Input).[6]

The University of Edinburgh has a large group of researchers working on speech recognition. The work ranges from very basic research (building mathematical models of how speech works), through research into recognizing the speech of elderly users, to recognizing speech recorded using distant microphones (e.g. on a table top).[7] English or for that matter some other language speech corpus for large vocabulary is built by linguists while speech algorithms are built by mathematicians.

_____

Optical Character Recognition

It is yet another language product which operates by virtue of application of network of mathematical formulas. It optically recognizes characters and reduces typological effort through scanners.

Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic translation of scanned images of handwritten, typewritten or printed text into machine-encoded text. It is widely used to convert books and documents into electronic files, to computerize a record-keeping system in an office, or to publish the text on a website. OCR makes it possible to edit the text, search for a word or phrase, store it more compactly, display or print a copy free of scanning artifacts, and apply techniques such as machine translation, text-to-speech and text mining to it.[8]

Word Processors

Word Processors have become synonymous with modern office life. Millions of people around the world use Microsoft Office in the course of their day-to-day work. It might come as a surprise, though, that Office also is the most popular software in the world for personal use – whether for writing school reports, designing party invitations, scheduling carpools, keeping grocery lists or organizing budgets.[9] its pure language-based word processing functions include spell checking (actually checks against wordlists), "grammar checking" (checks for what seem to be simple grammar errors), and a "thesaurus" function (finds words with similar or opposite meanings).

_____

Conclusion

To encapsulate, an innovative and interdisciplinary approach applied in the field of linguistics and mathematics has led to development of language processing products and can go a long way if both work in collaboration to churn out even sophisticated language processing products. They should even resort to multidisciplinary approaches, if this helps them bring home goal of developing even better language processing products.