8th. World Congress on Computational Mechanics (WCCM8)
5th. European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2008)
June 30 – July 5, 2008
Venice, Italy

On Intelligent Conversations with a Computer

***Pedro V. Marcal[1] and Nobuki Yamagata[2]**

[1] MPACT Corp.
3606 Calico Ranch Gate
Julian, CA 92036, USA
marcalpv@cox.net

[2] Nippon ESI K.K.
45-1 Oyama-cho
Shibuya-ku,Tokyo,Japan
nym@esi.co.jp

Key Words: Natural Language, Conceptual Dependency, AI, Meaning.

ABSTRACT

The goal of Artificial Intelligence has been to imbue the Computer with enough understanding that it can respond and act with intelligence. Much progress has been made in teaching the Computer to act in restricted and specific domains. Most of this has been made through the development and application of Expert Systems, Neural Networks and GUIs. Yet all this progress has been made without teaching the computer to communicate in man's natural language. The ability to use natural language has made much progress in recent years [1]. This field of study is known as Natural Language Programming (NLP). Progress is such that there are many NLP with context free grammars. Schank[2] developed a different approach to NLP through the direct application of semantics known as Conceptual Dependency. (CD). The main thesis in CD is that the meaning of a word can be expressed in a few basic concepts with its own grammar. Thus a simple sentence can be constructed as

PA PP, ACT PA, PA PP **(1)**

Where **ACT** are **ACT**ions or verbs, **PP** are **P**icture **P**ainters or nouns, and **PA** is **P**icture **A**iders or adjectives and/or adverbs

Though the expression above resembles the traditional grammar, it is very different because each of the **PP**, **PA** and **ACT** terms represent concepts that are noun, adjectival and verbal in nature at the sentential level respectively. Independent of any grammatical parsing, it is the responsibility of the CD to search out their meaning irrespective of where the terms may be placed in an actual sentence. Each term can in turn be written in a primitive form. [2] Adopted a notation that is two dimensional in nature. For the sake of computational convenience, the writers have developed their own notation in linear form. The semantic parsing of a sentence is based on a number of rules. These are listed below for the sentence " The tall John hit Mary on her head with a bat."

C-Rule	Example	
PP ⇔ ACT	John ⇔ hit	
ACT <-o PP	hit <- o Mary	
PP T<- PA	John T<- tall	where T transposes <-

PP <-o-<- R <- PP ->PP Mary <-o-<-R <-(with) bat ->(on) head

Where each of the operational symbols represents a transaction (TRANS)

and the symbol <-o-<-R represents a case belonging to an ACT with an input <- and an output ->. In all there are some 16 basic rules and cases. They are supplemented by about 100 special rules and cases. The following shows the semantics of the word

John PP HUMAN MALE UPPER-CASE

hit ACT PROPEL <-o-<-R <- TOOL -> LOCATION or PURPOSE

The following describes the development of the **S**emantic **CAN**vasser (SCAN) program. The implementation of CD in a computer program is reasonably straightforward. The main task was to build a semantic dictionary on which the program could operate. The Japanese have a set of characters named 'Toyo Kanji' that consists of 1850 characters. These characters are the only characters allowed to be used in their Newspapers. It seemed reasonable to use words based on these charactersin order to obtain a working vocabulary. This basic dictionary resulted in 4,500 words. In the original CD [2] it was suggested that the ACT could be described by about 18 verbs such as give, propel, ingest etc. This would be comprised of the capitalised words in our dictionary. As it turned out our dictionary resulted in 1,500 such root words.

Our work has been simplified by the use of the Natural Language Toolkit (NLTK) [1]. NLTK includes Wordnet, a lexical semantic dictionary that is constructed as sets of hierarchic semantic trees. It can measure the similarity distance that two words have when measured against these trees. Finally the similarity measure of the root words were used as a measure of the similarity in meaning between two sentences. The NLP described here was used to parse documents that would respond to questions from a user. Examples are provided of applications of this NLP in FEA.

It is apparent that continuing improvement will result in improved search in data mining, automatic building of expert systems, automatic computer programming in that order. It is also apparent that at a certain stage this effort will be self-evolving and further progress will be accelerated.

REFERENCES

[1] S. Bird, E Klein, and E. Loper, Introduction to Natural Language Processing, http://nltk.org/book/ 2007.

[2] R. Schank, "Identification of Conceptualizations Underlying Natural Language", Computer Models of Thought and Language, W.H. Freeman and Co., Chapter 5, 1973.