

## ЛИНГВИСТИКА

Reshetniak M.V.  
Banking academy of the NBU

### COMPUTATIONAL LINGUISTICS AS A SUBFIELD OF AI

The article is devoted to the Artificial intelligence field of study basic understanding on the example of a kids' game. Such subfield of AI as computational linguistics is described in the second part of the article. And the practical use of the computational linguistics is highlighted with the help of the English module in a Robot AI Mind programming approach.

«It is not my aim to surprise or shock you – but the simplest way I can summarize is to say that there are now in the world machines that can think, that can learn and that can create. Moreover, their ability to do these things is going to increase rapidly until – in a visible future – the range of problems they can handle will be coextensive with the range to which the human mind has been applied.» -Herbert Simon [1].

Artificial intelligence (AI) is the intelligence of machines and the branch of computer science that aims to create it. AI textbooks define the field as «the study and design of intelligent agents» where an intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success. John McCarthy, who coined the term in 1956, defines it as «the science and engineering of making intelligent machines».

AI research is highly technical and specialized, deeply divided into subfields that often fail to communicate with each other. Subfields have grown up around particular institutions, the work of individual researchers, the solution of specific problems, longstanding differences of opinion about how AI should be done and the application of widely differing tools. The central problems of AI include such traits as reasoning, knowledge, planning, learning, communication, perception and the ability to move and manipulate objects. General intelligence (or «strong AI») is still among the field's long term goals [2].

In the following part of the article the demonstration of a very simple practical example of artificial Intelligence programming is described. A Nepali game named «GATTA TIPNE KHEL» (meaning pebble picking game) is used for this purpose. We can see small children playing this game in the playground. In this pebble picking game a pile of some pebbles is kept in the ground. One of the two players picks one, two or three pebbles at a time in his turn, leaving the pile for the other player to pick for his alternate turn. In this alternate picking process, the player who picks the last pebble(s) will be the loser and called to be a DOOM in Nepali.

The main logic of the game is to leave the pile of pebbles with 13, 9, 5 or 1 pebble(s) for the opponent to pick. In the program the starting number of pebbles are set to 17, 21, 25, 29 ... etc. so that computer could win always if it does not make a mistake. But in the real play computer seems to be gradually learning by correcting mistakes of the previously played games. At last it finds all its mistakes and corrects them to become an unbeatable champion.

It seems computer simulates the psychological learning process of animal, learning by correcting and not repeating the mistakes. A multidimensional array of elements (1.4.1.3) is chosen as the instruction book for the computer to pick the pebbles. The instruction book contains four pages with three lines of instructions to pick pebbles. The first line instructs to pick a single pebble, the second line instructs to pick 2 and the third line instructs to pick 3 pebbles. At the beginning, computer chooses a random page and a random line of instruction to pick the pebble. When the game finishes, if computer loses the game, the last instruction is red-marked (erased) and the instruction will not be read in the future. After playing many games, all the instructions leading to a lost game will be red marked and there will be left only the instructions those lead to a win. Well, it is enough for the description of the game [3].

Computational linguistics as a field predates artificial intelligence, a field under which it is often grouped. Computational linguistics originated with efforts in the United States in the 1950s to use computers to automatically translate texts from foreign languages, particularly Russian scientific journals, into English [4]. Since computers can make arithmetic calculations much faster and more accurately than humans, it was thought to be only a short matter of time before the technical details could be taken care of that would allow them the same remarkable capacity to process language [5].

When machine translation (also known as mechanical translation) failed to yield accurate translations right away, automated processing of human languages was recognized as far more complex than had originally been assumed. Computational linguistics was born as the name of the new field of study devoted to developing algorithms and software for intelligently processing language data. When artificial intelligence came into existence in the 1960s, the field of computational linguistics became that sub-division of artificial intelligence dealing with human-level comprehension and production of natural languages.

In order to translate one language into another, it was observed that one had to understand the grammar of both languages, including both morphology (the grammar of word forms) and syntax (the grammar of sentence structure). In order to understand syntax, one had to also understand the semantics and the lexicon (or 'vocabulary'), and even to understand something of the pragmatics of language use. Thus, what started as an effort to translate between languages evolved into an entire discipline devoted to understanding how to represent and process natural languages using computers.

Computational linguistics can be divided into major areas depending upon the medium of the language being processed, whether spoken or textual; and upon the task being performed, whether analyzing language (recognition) or synthesizing language (generation).

Speech recognition and speech synthesis deal with how spoken language can be understood or created using computers. Parsing and generation are sub-divisions of computational linguistics dealing respectively with taking language apart and putting it together. Machine translation remains the sub-division of computational linguistics dealing with having computers translate between languages.

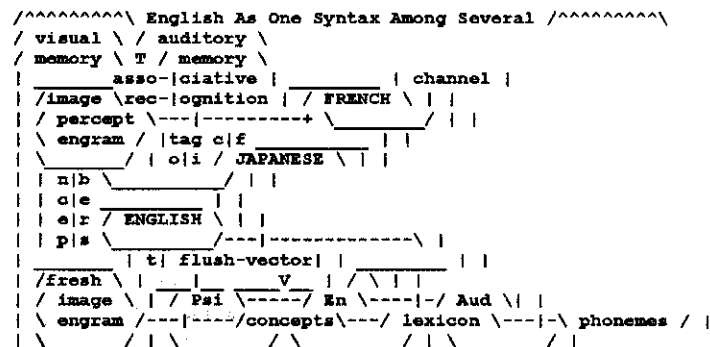
Some of the areas of research that are studied by computational linguistics include:

- Computational complexity of natural language, largely modeled on automata theory, with the application of context-sensitive grammar and linearly-bounded Turing machines.
- Computational semantics comprises defining suitable logics for linguistic meaning representation, automatically constructing them and reasoning with them.
- Computer-aided corpus linguistics.
- Design of parsers or chunkers for natural languages.
- Design of taggers like POS-taggers (part-of-speech taggers).

Machine translation as one of the earliest and least successful applications of computational linguistics draws on many subfields [6].

The Association for Computational Linguistics defines computational linguistics as: «...the scientific study of language from a computational perspective. Computational linguists are interested in providing computational models of various kinds of linguistic phenomena.» [7]

The brain-mind diagram below shows how the English module may easily co-exist with several other human languages in a Robot AI Mind.



The brain-mind diagram above shows how the English module may easily co-exist with several other human languages in a Robot AI Mind. Once the Think module has chosen which language to think in (perhaps because it is listening to input in a certain language), the English or other selected module generates and comprehends sentences of thought in the particular language.

Machine translation (MT) is achievable in a Robot AI Mind that specializes in a subject area in particular human languages [8].

Artificial Intelligence is a common topic in both science fiction and projections about the future of technology and society. The existence of an artificial intelligence that rivals human intelligence raises difficult ethical issues, and the potential power of the technology inspires both hopes and fears.

Informational Sources:

1. <http://library.thinkquest.org/2705/>.
  2. [http://en.wikipedia.org/wiki/Artificial\\_intelligence](http://en.wikipedia.org/wiki/Artificial_intelligence).
  3. <http://delphi.about.com/od/gameprogramming/a/aigamesample.htm>.
  4. John Hutchins: Retrospect and prospect in computer-based translation. Proceedings of MT Summit VII, 1999, pp. 30-44.
  5. Arnold B. Barach: Translating Machine 1975: And the Changes To Come.
  6. [http://en.wikipedia.org/wiki/Computational\\_linguistics](http://en.wikipedia.org/wiki/Computational_linguistics)
  7. The Association for Computational Linguistics / What is Computational Linguistics? Published online, Feb, 2005.
  8. <http://visitware.com/AI4U/english.html>.
- Additional source:
9. <http://www.informatics.sussex.ac.uk/research/groups/nlp/gazdar/nlp-in-prolog/ch04/chapter-04-sh-1.html#sh-1>.

Reshetniak M. V. Computational linguistics as a subfield of AI [Text] / M. V. Reshetniak // Поглед върху световната наука.- Бялград: Publishing house Education and Science s.r.o., 2010.- P. 64-67.