УДК 004=111

# PROBLEMS OF INTELLIGENCE SYSTEMS DESIGN

**Iermak A.M., Kyrbatova I.A.**
**Scientific supervisor – Associate professor Redkina A.V.**

***Siberian Federal University***

**"The question of whether a computer can think is no more interesting than the question of whether a submarine can swim", – Edsger Dijkstra**

At present artificial intelligence (AI) is defined as the scientific branch of learning which includes hardware and software development for modeling of intelligent human activity. The main problem of this definition consists in the fact that it is difficult to explain and understand the concept of intelligence. Is intelligence a single system? How is information presented in the living cells of organisms and how can it be incarnated in a computer? Is it possible to create AI using a computer-oriented approach? The science has found no answers to these questions yet, but they are the questions which favoured the creation of the modern categorical-conceptual framework that is the basis for the science about AI.

The middle of the XX century, when computers had such computational power that it could be used for complex mathematical calculations, can be considered the birth of artificial intelligence. In that very time the first premises about possibilities to model AI in computer systems appeared. Since that time AI science has passed some stages: brain structure and function simulation, heuristic method to limit a search range by discarding preliminary illegal variants, modeling of specialists' specific knowledge in the field of chemistry and medicine, and so-called 5th generation intelligence systems based on self-learning. The example of the last-named is an extension of the given heuristic patterns on the basis of the data being processed.

The main problem of the concepts stated above consists in the fact that such computer systems perform their actions only with the help of a human being, who sets starting conditions at the least.

There are two theories in the philosophy of artificial intelligence: the theory of strong artificial intelligence and the theory of weak artificial intelligence. Strong AI theory assumes the creation of some machine which has self-consciousness, reasoning and feelings that are not distinguishable from human ones. This theory supposes that machinery will finally be able to overshadow a human being. Weak AI theory assumes the opposite. Each theory has its supports and opponents.

The creation of strong AI has been the most challenging task for the science over a period of some decades. Unfortunately, even if it is developed the task of its testing is rather hard today.

The assessment method of machine intelligence was first offered in 1950 by an English mathematician Alan Turing. He did not create a collection of rules, conditions and regulations but offered the heuristic test which would with certainty answer if a computer has intelligence or not. The author's condition of the experiment is the following: "A man is communicating with a human being and a computer. Relying on the answers given to the questions he must define who he is talking to – a human being or a computer program. The computer program task is to mislead the man, get him make the wrong choice." The classical

test excludes any physical communication. Alan Turing supposed that by 2000 year a computer with the storage of 1 billion bits (about 119MB) would have been able to deceive a man in 30% of cases.

60 years has passed since the day of test publishing but even with exponential growth of computer performance not a single machine has come nearer to the solution of this task. Certainly, there were cases when judges were cheated by intelligence simulation. The most obvious ones are misprints while answering in chats and some mistakes in computations. Therefore, the judges' expertise is one of the weak points in the Turing intelligence assessment method.

The Chinese Room is one more thought experiment demonstrating non-applicability of the Turing test while assessing system intelligence. Imagine that there is a man in the closed room. He knows English but he does not know the Chinese language, which is only a set of hieroglyphs for him. We also assume that in the room there are some baskets with cut pieces of paper depicting hieroglyphs and that the man has a dictionary in English where the rules of hieroglyphs relation are stated and it is absolutely unnecessary to know what is written on the paper. Now we suppose that someone outside the room passes a set of Chinese symbols with a question and the man has to give an intelligible answer to it. Of course, to do it the man uses the dictionary. For example, imagine the following dialogue: "What's your favourite colour?", - the man outside the room asks. "Green", - the man being tested answers having absolutely no idea what he has been asked about.

The man has answered the question but has not understood anything. Considering this case the Turing test would say that the man knows the mater at least. Is it intelligence? Certainly, not, it is simulation. In this example the dictionary for the man means the same thing a computer program means for a machine as it is guided by the program when choosing an answer. Therefore, if a computer passes the Turing test it does not mean that this machine is intelligent. Thereby, the developer of the Chinese Room – John Searle – rejects the existing formal system approaches to the creation of strong AI.

There is no problem in the science that would remain unsolved, so science-fiction writers and researchers are already thinking of future solutions and developments. It is really necessary to consider the possible problems of global risk that can appear if AI is not programmed to be human-friendly.

The Three Laws of Robotics were first introduced by an American writer Isaac Asimov:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

One should be really a genius to put these laws into practice. Nevertheless, it should be done to provide safety measure as robots are becoming more and more advanced. This task is one of the cornerstones and the future of robots depends on its solution.

Thus, research in the field of AI is one of the most complicated, essential and challenging. This field of science is being developed very quickly: building of a new pattern brings the introduction of new ideas and technologies. All-round appearance of robots can solve any problem of the modern world but bring more other troubles. A real breakthrough in the creation of strong AI will take place only with the introduction of a new approach to computing systems design. At present techniques capable to completely replace usual transistor circuits by more advanced ones with absolutely different approach to computational

problems solution are being developed. In some years one of such systems might tell us “Hello!”