

## THE FIELD OF ARTIFICIAL INTELLIGENCE

Ashlı Aslan

### Öz

*Yapay zeka (Artificial Intelligence) bilgisayarların ne ölçüde zeka sahibi insanlar gibi davranabileceğini anlamayı hedefleyen bir bilim dalıdır (Flanagan, 1984; Erden, 2003). Yapay zeka alanında çalışan araştırmacılar insan zekasına benzer çalışan yapay zekaların olup olamayacağı, bir makinenin bütün insan davranışlarını, duyarlılığını ve duygularını taklit edip edemeyeceği gibi sorulara cevap ararlar. Bu yazı boyunca insan beyninin de bir çeşit bilgisayar olduğuna ya da en azından bir bilgisayar gibi çalıştığına inanan araştırmacıların bu tür düşünceleri kanıtlama veya çürütme teşebbüsleri üzerinde odaklanacağım. Bunu yaparken de bu konunun farklı yönlerini yapay zeka savunucularının ve karşıtlarının bakış açılarına göre karşılaştıracam.*

### Anahtar Sözcükler

*Yapay zeka, insan aklı, bilgisayar zekası  
Yapay Zeka Alanı*

### Abstract

*Artificial intelligence is a field of science aimed at understanding to what degree the computers can behave like intelligent people (Flanagan, 1984; Erden, 2003). Researchers who study in the field of artificial intelligence search for answers to questions like whether there can be artificial minds operating like the human mind, whether a machine can duplicate all forms of human behavior, sensitivity, and emotions. Throughout this paper, I will concentrate on the attempts to prove or disprove the claims made by researchers who believe that mind is a computer or at least works like a computer. While doing that, I will compare the different aspects of that topic from the view point of the proponents and opponents of artificial intelligence.*

### Key Words

*Artificial intelligence, Human, Computation*

### Artificial Intelligence

Researchers and philosophers who study artificial intelligence are concerned with questions such as whether there can be artificial minds analogous to human mind, whether a machine can duplicate all forms of human behavior, sensitivity, and emotions (Ergün, 2003). Artificial Intelligence especially aimed at understanding to what degree the computers can behave like intelligent people (Flanagan, 1984; Erden, 2003). In order to find answers for these questions, AI scientists attempt to map the human mind onto computers, generally by using two methods (Wagman, 1993): they either imitate cognitive processes of human mind with computer programs or apply formal processes of human logic to computer algorithms. Flanagan (1984) separates the artificial intelligence studies into four branches, depending on which method is used. One of them is non-psychological AI research. Non-psychological AI studies aim at developing an intelligent computer without any intention or claim to explain the way the human mind works. Another branch seeks to explain the cognitive processes of human mind by developing computer programs which implement certain aspects of the human mind. This second branch is called "weak-psychological AI" (Flanagan, 1984; Searle, 1990: 67-88). Weak AI researchers use the computer as a tool in studying the human intelligence; they attempt to test and probe the human brain indirectly, by studying computer algorithms which are modeled from the human mind. The third branch is called the "supra-psychological AI" (Flanagan, 1984). The researchers of this branch believe that AI studies must concentrate on all

intelligent life forms as models of intelligence, as well as human being. Thus AI programs must represent all intelligent entities. This idea has not yet occupied the minds of philosophers as much as the other three. The fourth kind is called the “strong AI” (Flanagan, 1984; Searle, 1990: 67-88). Strong AI researchers claim that human brain and digital computers are physical symbol systems and its operations can be performed by computer programs. Therefore, they see no reason why human mind could not be duplicated or completely imitated by computers. This idea is analogous to the computational theory of the mind which views the mind as a collection of automatic process (Wagman, 1991). According to this theory, all functions of the mind can be explained based on mathematical logic and certain symbolic rules. Thus, computers should be able to think and reason in the same way as the human mind, by using the same type of logic or the same symbolic rules.

#### **Can a machine duplicate human mind?**

Regarding the question of whether a machine can duplicate human mind, Cohen (1955: 36-41) concentrates on the dilemma of both proponents and opponents of strong AI claims. While he does not find it impossible for science to duplicate or imitate human mind, he also states that human mind seems to be unique with respect to some qualities that no machine can duplicate. Erden (2003) also states that, due to the current technological limitations it is not possible for machines to behave exactly like human beings, at least in the near future. This dilemma continues to occupy the minds of those who are interested in artificial intelligence studies. In addition to trying to explain the human mind or attempting to develop human-like computers, they also argue why it is possible or impossible to duplicate human mind.

Before I discuss the viewpoints of certain philosophers and researchers, I would like to make a point about the different structures of human brain and the computer. First of all the building materials of the two are different: Brain is made up of a biological tissue intermingled with neurons, but the computer is built with metal and silicon. The inner workings of the two also differ: Information is transmitted by chemical substances known as neuro-transmitters in brain, while a computer relies solely on electrical signals (Haliloğlu, 2003). Whenever an analogy or a distinction will be made between the two entities, those obvious differences will be ignored.

Wagman (1993) puts down the fundamental differences between computers and human mind in certain aspects: The basic methods of information processing and the methods of communication between the information processing elements are different in brain and in the computers. Ignoring the obvious physical differences mentioned above, AI proponents note that the communication in both entities requires symbolic logic structures and operations. But two entities also differ in the way they represent and transmit information to other entities. Computers use digital programming languages, while humans use natural semantic languages. According to AI theories, these two types of languages differ only in appearance; the representation of information involves similar syntactic and semantic logical elements both for an artificial human entity and an AI entity.

Still another question is how the brain or the computer acquires information. Brain acquires by learning, problem solving and reasoning. These

processes may proceed differently for one person than for another due to differences in private characteristics (Wagman, 1991). Moreover, the information-acquiring mechanisms are not indifferent to the content. For an AI entity, on the other hand, information acquisition is based on a few standard theorems and is indifferent to the content (Wagman, 1991). AI theories, however, view the information acquisition as theorem proving operations for both humans and AI entities. (Wagman, 1993).

As seen from the Wagman's efforts to show the differences between the human mind and an AI entity, we sometimes differentiate similar phenomena by describing them with different concepts, while we sometimes establish an analogy between two different phenomena based on their different aspects. The problem is that, we do not yet have the capacity to clearly state which phenomena are really similar which are really different. It seems that discussions on whether a hand-made machine can have ability to duplicate human mind will not be solved in near future, unless we agree on certain definitions. Many of the philosophers who are interested in strong AI emphasize the different aspects of the problem and approach with different conceptualizations. These different aspects of the problem will be discussed in following sections.

### **Intentionality**

Perhaps the most fundamental dilemma faced by philosophers is whether the human mind is a formal or an informal system. Many of the strong AI opponents such as Searle (1990: 67-88) insists that human mind is an informal system in contrast with a computer program which is a formal system. Although he accepts the idea that, in many aspects, the cognitive processes occurring in brain involve formal, symbolic rules like a computer program does, he argues that brain has an undeniable capacity to behave informally. According to Searle, what makes brain an informal system is its causal capacity to produce intentionality. He views the existence of a biological entity a necessary and sufficient condition for intentionality, hence unpredictability. Because brain is a biological structure, it possesses the necessary casual power for intentional behavior.

Searle also claims that human intelligence or mind is more than a matter of formal symbol manipulation. It requires a set of causal understanding which is intentionality embodied in the biological brain tissue For Searle, intentionality is the essence of being a human; it cannot easily be explained. On the other hand, Searle agrees that it is possible to produce a machine that works exactly like the human brain, if it is constructed with a neural system that contains all the hallmarks of the human brain, such as neurons, neuro-transmitters, etc. He further claims that, without biological materials, no digital machine made up of inorganic material can imitate or duplicate human mind. Searle insists that a computer program reacting to certain inputs does not have intentionality, because it is based on formal and symbolic manipulations which are wholly predictable.

Dennett (1971: 87-106; 1990: 147-170) disagrees with Searle's definition of intentionality. According to him, ascribing intentionality to a system is related to ascribing beliefs, goals and rationality to this system. If a digital machine is assigned beliefs, goals and rationality that can be predicted and explained, it can also be considered as having intentionality. Moreover,

Turing (1990) claimed that any system capable of instantiating an appropriate formal information processing method can be said to be behaving intentionally. In order to prove his idea, Turing developed what has since been called the "Turing machine phenomenon" which is based on the idea that two systems are alike, if they behave sufficiently alike. Basically described, if we can not distinguish between a human being and a programmed machine based on their behavior, then the two must be identical (Rosen, 1992: 87-95). In other words, a computer program that appears to behave intelligently is, indeed, intelligent.

Searle disagrees with the Turing method. He states his counter-claims by using the Schanks program which is a Turing machine analogous. Briefly described, the Schanks program aims at simulating the human ability to understand stories (Searle, 1990: 67-88). Searle tests the Shanks program in an imaginary game which he calls a Gedanken experiment. He asks us to imagine a person locked in a room with a batch of Chinese writing. We are also asked to assume that this person neither speaks nor understands Chinese. Therefore, the Chinese writing that he is given is completely meaningless for this imaginary prisoner. We now imagine that the prisoner is given a second batch of Chinese script, along with an extra script which contains a set of grammatical rules in English. Provided that the prisoner knows English, this extra script will enable him/her to correlate the second Chinese script with the first Chinese script. We now introduce a third batch of Chinese script together with some instructions in English. The instructional script enables the man to correlate all three Chinese writings with each other. The instructions are arranged to give information about how to correlate certain Chinese symbols with certain shapes in response to certain other shapes given in the third batch. At that point, the prisoner is able to answer certain questions in Chinese, without knowing any Chinese at all. All that he/she does is to establish analogies between the shapes in Chinese writings and form new groupings according to orderings that he/she has observed.

In Searle's imitation game, the first Chinese script is a story written in Chinese, and the third Chinese script consists of questions on the first story. The instructional English scripts are programs that are fed to a computer to help it to form new word groupings like those found in the story. He goes on to suggest that new scripts produced by the prisoner according to the given rules are analogous to new stories produced by the computer. In other words, he simply gives an example of how a computer program works. According to Searle (1990: 67-88), the man locked in the room is able to answer the questions by manipulating certain symbols using certain formal rules, but without understanding, and so does a computer. On the other hand, from the viewpoint of strong AI proponents, the Chinese room illustration explains just the way human beings understand a story. As highlighted in Searle's illustration, a programmed computer can also understand the story the way a human being understands the same story. A machine running a Schanks program, or any Turing machine analogous, does not merely simulate the human ability of understanding stories, but it also actually understand the stories. AI proponents insist that, when we listen to and speculate on a story in our native language, what we are doing is actually what the prisoner of the Chinese room does.

Turing (1990) claims that, by modifying a computer so that it will have an adequate storage capacity, speed and an appropriate program, it will act in the

same way a human does. According to Turing (1990), the two systems which behave sufficiently alike are actually alike. He sees no reason to think otherwise. Finally, he argues that, once a machine is able to write a sonnet which sounds emotional and thoughtful, it will be hard to claim that the machine is merely responding to electric signals and producing material output lacking emotional content. On the other hand, even strong AI proponents argue that computer system cannot be considered to be identical to human mind simply because the computer happens to pass the Turing machine tests. Regarding the Turing machine analogous, Searle insists that a computer understands and thinks differently from a human being. He finds no reason to believe that human understanding requires a formal system like the computer operations do. In the Chinese room illustration, there is no understanding at all. The reason for our viewing computers as agents capable of understanding is extending our own intentionality to them. According to Searle, many different levels of intelligence found among human beings are another evidence that computer intelligence may really be different from that of humans.

Newell and Simon (1990: 105-132) disagree with the idea that intelligence is something more than a physical, symbolic information processing system. They believe that intelligence is inherent in all physical symbol systems. They concentrate more on the question of what the structural requirements are for the intelligence. According to them, at least some of the requirements for intelligence involve ability to store and manipulate symbols. Differing from some other strong AI proponents such as Turing (1990: 40-66) and Boden (1990: 89-104), they do not claim that their system offers a complete model of human intelligence. They are satisfied with searching for the general characteristics of intelligence. Indeed, they offer two laws of qualitative structure for AI (Newell and Simon, 1990: 105-132). These are physical symbol systems based on heuristic search. Newell and Simon insist that human beings and computers are the two most significant classes of symbol systems that has finite resources. According to Newell and Simon's (1990: 105-132), research on artificial intelligence must be concerned with how symbol systems must be organized in order for the system to behave intelligently. For instance, symbol systems cannot reveal intelligent behavior when they are surrounded by chaos. Machines behave intelligently by extracting information from a problem domain and by using that information to guide their search. This method helps an AI entity avoid wrong turns and mistakes. As a matter of fact, as we mentioned above, Searle agrees that a machine using an information-processing approach can possess a genuine understanding but it will still lack mind's intrinsic intentionality.

#### **Informality**

The informality aspect of the mind is an important point for the opponents of strong AI; it supports their idea that no formal system can represent the human mind. Emotionality also seems to be a good hallmark for an informal system (Flanagan, 1984). In addition to that, the way that humans think is always associated with human emotions. Consequently, philosophers are also concentrated on emotions. It seems that emotions are mysterious and it may be said that we can never explain the rationality behind at least some of the emotions (Pylyshyn, 1983: 57-77). Proponents of strong AI object to the

informality attributed to emotions. Since they view emotions to be under the control of a bio-chemical system that employ hormones and neuro-transmitters, they insist that emotions also abide certain formal laws.

In conclusion, strong AI proponents insist that human mind is a formal system as a whole. They claim that anything that is natural involves lawful relations. Since mind is also part of the nature, the mind also involves lawful relations. Thus, there is no reason why any formal system should not duplicate any other formal system.

#### **Originality**

Another leading concern among researchers and philosophers who are interested in strong AI is finding a good answer to the no-originality objection. The opponents of strong AI argue that a computer does what is instructed by its programmer. It adds nothing to them on its own. Hence a computer can behave like anything, but only if it is provided with the right program (Rosen, 1992: 87-95). The conclusion of the strong-AI opponents is that the computer never does anything creative or unpredictable. Even when it seems to be behaving on its own, this illusion lasts until its current program is known to us. They contrast these limitations with human behavior whose hallmarks are novelty, creativity and unpredictability.

Cohen (1955: 36-41) also states that some of the most prominent features of human beings are unpredictability and creativity. Even though they may seem to be controlled by their governments, parties, families, and certain ideological belief systems, it would be wrong to say that they behaved with no minds of their own. According to Cohen, there are always reasons to believe that they think for themselves. A person carries, in himself/herself, a very unique subjective meaning. Flanagan (1990) also attributes a set of unit aspects to each person's personalities that cannot be explained by outside powers.

Proponents of strong AI answer to that by denying the notion of unpredictable human beings. They claim that human behavior is determined by the operational principles of biology, the functions of the cognitive equipment, life experiences, and the level of education. A computer can also acquire a set of such law-like rules if it is provided with appropriate settings. Flanagan, on the other hand, brings up the issue of human creativity of pervading many different areas such as music, literature, science, and poetry, because there seem to be no law-like rules governing the creative actions in such fields. However the actual question stands: Is creative behavior a mere spontaneous natural act, or is it necessitated by some social, law-like rules? (Flanagan, 1990)

#### **Consciousness**

Another point made by the proponents of strong-AI is that the predictability of a computer's actions depends on the size and the complexity of the system, and the level of randomness (Turing, 1990: 40-66). Turing reminds that there are already many programs whose outcomes can not be predicted easily such as the programs designed to play chess. These are programs that often surprise even their programmers. According to Turing (1990: 40-66), a machine can have much diversity of behavior depending on the storage capacity, and when the machine has adequate storage capacity, its behavior can not easily be predicted. Finally, Turing notes that the main reason for the criticisms of the strong-AI opponents must be their preconception that a computer cannot have

much storage capacity. Otherwise, even they would not have any reason to think that computers could not produce unpredicted and novel results. He argues that these criticisms are just disguised forms of consciousness problem. The strong-AI components do not foresee intelligent computers, because they think computers will never have consciousness. However, it is not yet clear that consciousness is an exclusive characteristic of a human being. We do not yet have an adequate definition of consciousness.

### **Conclusions**

Throughout the paper, I have focused on the attempts to prove or disprove the claims made by strong-AI researchers. Briefly stated, the strong-AI claims center on the idea that mind is a computer or, at least, works like a computer. Thus, there is no reason why a computer cannot duplicate the human mind. According to Turing (1990: 40-66) and some other proponents of strong-AI such as Newell and Simon (1990: 105-132), intelligence requires nothing more than a formal symbolic system. Turing aims at proving that if any program behaves sufficiently intelligent, it is intelligent. Based on this claim, he suggests a criterion to verify the success of strong-AI claims: if we cannot discriminate between the intelligent behaviors of a human being from the actions of a computer program, the program is an adequate prototype of the human intelligence. On the other hand, regarding the Turing's claims, strong-AI opponents argue that what really matter is whether the program behaves intelligently the way that a human being does. They bring up the Chinese room analogy in which a computer represented by an imaginary prisoner answers questions on Chinese stories, but still does not understand Chinese. Another objection made by Cohen (1955: 36-41), being able to instantiate a formal symbolic system is not a sufficient condition for intelligent behavior. Having intelligence requires consciousness, but machines lack consciousness. For Searle (1990: 67-88), the mind is an intentional system, rather than a representational system. According to Searle, intentionality is the most important indicator of an intelligent mind and it requires biological structures like a brain. Hence a silicon machine will never acquire intentionality and, therefore, intelligence.

Here, a question arises as to whether the mind is a representational system or an intentional system. Regarding this question, it is known that intelligence can be explained by formal cognitive processes. But how much of intelligence is necessary to be considered human is the right question to be asked. As Flanagan (1984) points out, maybe intentionality, consciousness and emotions are, in some essential way, tied only to a specific type of organic machine, a human. At that point, it will not be wrong to say that intelligence is tied to law-like rules of nature. Thus, the claims made by the proponents of strong AI cannot be contradicted or disproved easily. Furthermore are right in pointing out that we have not searched enough to say that there are no such laws. Unless otherwise is proved, the opponents of strong AI will continue to object to the concept of a completely formal, symbolic intelligence.

It is obvious that researchers who are interested in artificial intelligence need to consider the knowledge and experience gained in other areas such as cognitive science, biology and other physical sciences, as well as social sciences such as developmental psychology, cultural sciences (Erden, 2003). For now, as

Flanagan points out, it is really difficult to find because mind still seems to be extremely complex. First, we have to find out what intelligence is. In order to understand the nature of mind and intelligence, we will have to know what is innate at the birth, what develops during maturing, and how the environment affects the process. In conclusion, we need much information from many diverse areas such as developmental psychology, cultural sciences as well as physical sciences. In addition, we should not forget that findings of the studies on artificial intelligence can result in expert systems that can solve complicated and tedious problems just as a human expert does. For example, in many areas such as geography, a well-designed AI system makes decisions like a human and it is much more superior to an ordinary computer program that works by manipulating binary data according to the rules of boolean logic (Ölgen, 2003).

#### References

- BODEN, M.A.(1978), "Artificial intelligence and Piagetian theory", **Synthese** 38, 389-414.
- BODEN, M.A.(1990), "Escaping from the Chinese room", In M.A.Boden (Ed.), **The philosophy of Artificial Intelligence** (pp.89-104), Oxford University Press.
- COHEN, J. (1955), "Can there be artificial minds", **Analysis**.Vol.15-16, 36-41.
- DENNETT, D. C.(1971), Intentional systems. **Journal of Philosophy** 68, 87-106.
- DENNETT, D.C.(1990),Cognitive Wheels: the frame problem of AI. In M.A.Boden (Ed.),**The philosophy of Artificial Intelligence** (pp. 147-170), Oxford University Press.
- ERDEN, A., (2003), "Robotik", **Bilim ve Teknik Dergisi**, Tubitak. <http://www.biltek.tubitak.gov.tr>
- ERGUN, M. Yapay zekanın gelişimi ve tarihçesi. <http://www.egitim.aku.edu.tr/yz1.htm>
- FLANAGAN, O.J.(1984), The Science of the Mind. Cambridge , Massachusetts, London, England: The MIT Press.
- HALILOĞLU, N. (2003), Yapay zeka ve Kapsamı. <http://www.geocites.com/nurayhaliloglu/yapayzeka/yapay3.htm>
- NEWELL, A. & Simon, H.A.(1990), "Computer science as empirical enquiry : symbols and search" In M.A.Boden (Ed.),**The philosophy of Artificial Intelligence** (pp. 105-132), Oxford University Press.
- ÖLGEN, M. K. (2003), Yapay zeka ve coğrafya. International 12. Turkish Symposium on Artificial Intelligence and Networks-TAINN.
- PRESTON, B.(1993), "Heidegger and Artificial Intelligence". **Philosophy and Phenomenological Research** No:1.
- PYLYSHYN, Z.W.(1983), "Minds, machines and phenomenology:some reflections on Dreyfus: what computers cannot do", **Cognition**.3(1), 57-77.
- ROSEN, R. (1992), On psychomimesis" *Idealistic Studies*. 22-23.87-95.
- SEARLE, J.R. (1990), "Minds, Brains and Programs", In M.A.Boden (Ed.), **The Philosophy of Artificial Intelligence**, (pp.67-88). Oxford University Press.

- TURNIG, A.M. (1990), "Computing machinery and intelligence", In M.A.Boden (Ed.), **The Philosophy of Artificial Intelligence** (pp.40-66). Oxford University Press.
- JACQUETTE, D. (1989), **Adventures in the Chinese room. Philosophy and Phenomenological Research**. Vol.XLIX. 605-623.
- WAGNIAN, M.(1993), **Cognitive Psychology And Artificial Psychology**, Praeger: Westport, Connecticut, London.
- WAGMAN, M.(1991), **Artificial Intelligence And Human Cognition**, Praeger: Westport, Connecticut, London.