T.C.
İSTANBUL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
FELSEFE ANABİLİM DALI

YÜKSEK LİSANS TEZİ

YAPAY ZEKÂ FELSEFESİNDE FELSEFİ İŞLEVSELCİLİĞİN ELEŞTİRİSİ

Tarık TAZEGÜL

2501210163

TEZ DANIŞMANI

Doç. Dr. Necati Murad OMAY

İSTANBUL – 2023

ÖZ

YAPAY ZEKÂ FELSEFESİNDE FELSEFİ İŞLEVSELCİLİĞİN ELEŞTİRİSİ

TARIK TAZEGÜL

İşlevselciliğe göre bilişsel durumlar, bilişsel durumların gerçekleştiği materyalin yerine getirdiği işlevlerin sonucunda meydana gelir. Bu işlevleri yerine getirebilecek herhangi bir maddi yapının kendisi de yine bilişsel durumlara sahip olabilir. Yapay zekâ felsefesi bakımından işlevselcilik ise farklı bilgi ve bilim alanlarındaki kullanımlarıyla beraber düşünülmelidir. Yapay zekâ felsefesi içerisinde insan zihninin çözümlenmesi ve modellenmesi çalışmalarında farklı işlevselcilikler karşımıza çıkmaktadır. Bu eleştiride ele aldığımız bilgisayımsal-temsil işlevselciliği ise Ned Block'un işlevselcilik tasnifine dayanmaktadır. Bilgisayımsal-temsil işlevselciliği, zihin durumlarının işlevlerinin programlar vasıtasıyla temsil edildiği ve bu temsili yapının bilgisayarlar üzerinde gerçekleştirildiği bir düşünce üzerinde şekillenmektedir. Bilgisayımsal-temsil işlevselciliğinin eleştirimiz içerisindeki ele alınışı, onun yalnızca bir bilgisayar bilimi tartışması olarak değil ama aynı zamanda yapay zekâ felsefesinin asli tartışma unsurlarından biri olduğu düşüncemize dayanmaktadır. Bu vesileyle eleştirimizin "felsefi işlevselciliğin eleştirisi" olarak adlandırılmasını uygun gördük. Bu adlandırma yalnızca bilgisayımsal-temsil işlevselciliğini felsefi bir tartışma olarak yeniden ele almayı imlemez, aynı zamanda bu tartışmayı yapay zekâ felsefesine farklı alanlardan taşınan antropomorfik tasarımlarla birlikte ele almayı da imler. Böylece eleştirimizdeki amacımız ilk adımda yapay zekâ felsefesi tartışmalarının asli unsurlarının belirlenmesi amacını taşımaktadır. Çünkü mitsel ve edebi kurgular olarak kimi özne tasarımlarının bedensel veya davranışsal olarak antropomorfik bir tarzda düşünüldüğü ve bu düşünmenin felsefe öncesi/sağduyusal bir düşünme olarak yapay zekâ felsefesi içerisinde ortaya konmaya çalışıldığı görülmektedir. Eleştirimizdeki iddialarımızdan biri böyle bir düşünmenin yapay zekâ felsefesinde faillik, ahlakilik vb. gibi temel tartışmaların ilerletilmesinde belirleyici olamayacağıdır. Eleştirimizdeki diğer iddiamız ise, bilgisayımsal-temsil işlevselciliği temelinde inşa edilecek bir yapay

zekâlı sistemin bir zihne sahip olmak ve fail niteliği kazanmak bakımından kimi statüleri elde edemeyeceğidir.

Eleştirimizdeki temel iddialarımızı ortaya koyabilmek için ilk adımda felsefi işlevselciliği metafelsefi olarak inceledik. Bu tavırdaki antropomorfik ögeleri tasnifledik ve bu ögelerin felsefi bir tartışmanın belirleyici unsurları olamayacaklarını vurgulamaya çalıştık. Öte yandan yapay zekâ felsefesi içerisinde fizikalizmin aşırı kullanımlarına dikkat çekmek istedik. Böylece fizikalizmi, bu tavra getirilen eleştirilerle birlikte yeniden ele alarak yapay zekâ felsefesindeki örtük kabulleri yeniden değerlendirmeye çalıştık. Bilgisayımsal temsil işlevselciliğinin dayandığını düşündüğümüz temel bilimler olarak matematiği, mantığı ve kimi doğa bilimlerini ele aldık. Bu işlevselciliğin temelinde yer alan temel tavır olarak biçimselciliğin matematiksel ve mantıksal dayanaklarını bu alanların önde gelen tartışmalarına giderek yeniden incelemek istedik. Böylece sağduyu bilgisinin temsili, biçimselleştirmeye direnen nitelceler ve fenomenal deneyimin unsurları gibi birçok hususu da biçimsel sistemler üzerinden ele almaya çalıştık. Ayrıca matematik ve mantık bilimlerinin tam ve tutarlı alanlar olarak düşünüldüğü bilgisayımsal-temsil işlevselciliğine karşı olan kimi matematiksel itirazlar üzerinde de durmak istedik. İnsan zihninin nöral ateşlemelere indirgendiği bilgisayımsal temsil işlevselciliğinin doğa bilimlerine dayanan kabullerini de fizikalizme getirilen eleştiriler, determinizm-indeterminizm tartışmaları, kaos ve kuantum kuramları gibi hususlar üzerinden değerlendirmeye çalıştık. Tüm bu ele almalar neticesinde yapay zekâ felsefesi bakımından bilgisayımsal temsil işlevselciliğinin temelindeki tartışmaları açık kılmayı arzuladık. Ayrıca bu bölümdeki bir diğer amacımız felsefi işlevselciliğin yaygın felsefe yapma yöntemi olarak bu tavırda halihazırda kullanılan varsayımların sürmekte olan tartışmaların bir unsuru olarak değil, tamamlanmış tartışmaların bir ürünüymüş gibi kullanıldığını vurgulamak olmuştur.

Eleştirimizin ikinci bölümü “yapay zekâ” kavramının kavramsal incelemesine eğilmektedir. Literatürdeki yapay zekâ tanımlarını incelediğimiz çalışmamızda mevcut tanımların çoğunun zihnin algoritmalara indirgenebileceği düşüncesi etrafında temellendirildiği görülmektedir. Biz ise eleştirimizde zihnin, problem çözmeye ilişkin adımlar olarak algoritmalara indirgenebileceğine ilişkin destekleyici

ve karşı görüşleri ele aldık. Ayrıca algoritmaya indirgenemeyecek zihin yetilerinin neler olabileceği üzerinde durduk. Zihnin algoritmalara indirgenebileceği düşüncesinden hareketle sembolik ve nöral ağ sistemleri üzerinde gerçekleştirilecek yapay zekâ sistemlerinin zihin sahibi olabileceği ileri sürülmektedir. Çalışmamız içerisinde bu sistemleri semantik içeriğin temsili, gündelik bilginin kapsanması ve bu sistemlerin kullanımdaki başarıları üzerinden incelemeye çalıştık. "Yapay zekâ" kavramını ele almayı zorlaştıran "doğal zekâ" ve "yapay zekâ" ayrımlarına yer verdik. Böylece "doğal zekâ" kullanımına alternatif olarak "canlı zekâsı" kullanımını önerdik. Eleştirimiz zihin gibi birçok yetiyi içeren bir yapının zekâ gibi yalnızca hesaplama ve problem çözmeye ilişkin bir yetiye indirgenmesi çabasını da yeniden değerlendirmektedir. Kavramsal incelememizin son bölümünde ise yapay zekâlı varlıkların zihin sahibi olduklarının kabul edilmesi durumunda ahlaki özne olarak da kabul edilebileceklerine ilişkin iddianın, yapay zekâlı varlıkların zihin-beden ilişkisi bakımından olanakları üzerinden bir eleştirisini sunduk.

Eleştirimizin son bölümünde ise birinci ve ikinci bölümlerde yöntemsel ve kavramsal olarak ele aldığımız felsefi işlevselciliğin eleştirisini sunmayı amaçladık. Böylece işlevselcilik eleştirisi özelinde bilgisayımsal-temsil işlevselciliğinin felsefi bir açıdan eleştirisini sunmayı istediğimizi açık kıldık. Bilgisayımsal-temsil işlevselciliğine temel sağlayan bilgisayar bilimi üzerindeki kimi tartışmalara da çalışmamızda yer verdik. Matematikteki fonksiyonlar ile algoritmalar arasındaki nitelik farkından hareketle yapay zekâlı varlıklardaki yönelimselliğin kökenini sorguladık. Matematiksel-mantıksal temelde inşa edilen bir yapay zekâlı varlığın özsel sorunlarından biri olarak durma sorununu ele aldık. Durma sorununun yapay zekâlı bir varlığın insan zihnine denk veya onu aşan bir sahipliği elde etmek konusunda yapay zekâlı varlıklara bir engel teşkil ettiğini ileri sürdük. Simülasyonlar ile gerçeklik tasarımlarımız arasında bir ontolojik ayrım bulunduğu vurguladık. Transistörler ile sinir hücreleri arasında bir nitelik farkı olduğu üzerinde durduk. Son tahlilde ise felsefi işlevselci bir perspektiften yapay zekâ ve sanat ilişkisi üzerinde süregelen kimi tartışmalara ve itirazlara değinmek istedik.

Anahtar Kelimeler: Yapay Zekâ, İşlevselcilik, Zihin, Beden, Bilgisayar, Hesaplanabilirlik Kuramı, Turing, Putnam, Searle, McCarthy.

ABSTRACT

THE CRITIQUE OF THE PHILOSOPHICAL FUNCTIONALISM IN THE PHILOSOPHY OF ARTIFICIAL INTELLIGENCE

TARIK TAZEGÜL

Functionalism is the idea that objects are not defined by their physical existence but by the functions that they fulfill. Therefore, in the philosophy of artificial intelligence, there are many functionalist efforts to define human mental states, which is an aim to be achieved by constructing artificial intelligence models. In my critique, I do not focus on the basic functionalist approach, which is the functional analysis of a being. I am focusing on a specific type of functionalism, which is based on certain sciences such as computer science, mathematics, logic, and the natural sciences. This functionalism is called computational-representational functionalism by Ned Block in his paper "What is Functionalism?". But in order to inquire into the fundamental claims of this type of functionalism, I must also reveal the metaphysical and fictional aspects that are carried by artificial intelligence studies and literature into the philosophy of artificial intelligence. These metaphysical and fictional elements render the field of artificial intelligence and the fundamental concepts of this field, such as artificial intelligence, consciousness, agency, etc., unable to inquire philosophically into them. Therefore, I used the name philosophical functionalism to refer to both the computational-representational functionalist aspect and the metaphysical burdens of the field.

This critique contains three chapters. In the first chapter, I approached the field of artificial intelligence from a metaphilosophical perspective. In the first section, I focused on anthropomorphic elements and the physicalist presuppositions of the field. Physicalism, as one of the fields that are the foundations of the field of artificial intelligence, wants to reduce modes of human mental states to neural firings. But this idea has to deal with Levine's "explanatory gap" objection. On the other hand, it is not clear that the representation of information that is about reality is defined from the physicalist perspective. This idea leads us to the belief that

everything is represented by symbols that belong to formal systems such as mathematics or certain logical structures. However, this belief is not sound enough to construct an artificial intelligence system capable of reaching human level intelligence or agency. In the second section of our metaphilosophical inquiry, I focused on the fundamental sciences such as mathematics, logic and the natural sciences that created the field of artificial intelligence.

First, I analyzed the ontology and epistemology of the objects and operations of mathematics from the standpoint of the philosophy of mathematics. The ontological and epistemological status of mathematical objects and operations has many roots, and each of these roots causes different types of mathematics and reality designs. I had to reveal the exact mathematical method that artificial intelligence studies rest on. There are many ways, such as intuitionism, naturalism, logicism etc. The formalist method of doing mathematics is the main method that artificial intelligence researchers and thinkers in this field use. However, formalist mathematics has to answer a few problems. Even if Hilbert's program is to achieve a complete and consistent mathematical system, Gödel's first and second incompleteness theorems must be answered. We cannot claim that artificial intelligence systems, which are built on a formalist mathematical structure, can represent our reality before we solve these and further objections. Until these problems are solved, we can claim that artificial intelligence systems exist only in a mathematical ontology. **Atomos** of this mathematical ontology are numbers, and their laws are mathematical operations. Therefore, even if these systems seem to act in our reality, this is completely false. Through their sensory organs, these systems transform the objects of our reality into numbers and manipulate them in their mathematical ontology. Therefore, we can easily claim that the mathematicalization of reality cannot represent beings as they are.

Logic is the other area on which artificial intelligence systems are based. Most artificial intelligence researchers and thinkers thought that facts about reality could be represented with propositions. In order to do that, logic had to be accepted as a natural science. Because a being that is constructed on a propositional level must show certain qualities, such as inventiveness. However, logic cannot be considered a

natural science. Logic is both a method and the science of this method at the same time. But logic as a methodology does not carry the scientific qualities that the natural sciences do. We use logic to build knowledge in philosophy and science, but logic is a contentless area. That is why logic without sensory input does not provide inventiveness, which we humans have naturally. On the other hand, as we have shown in our inquiry into mathematics, a formalist system based on symbols and axioms cannot represent reality as well as we do.

In the last section of my metaphilosophical inquiry, I wanted to focus on the scientific basis of artificial intelligence studies. Natural sciences reduce human mental states to neural firings, and if these neural firings can be simulated in artificial neural networks, which are built on logic gates and symbolic structures, we can therefore claim that these artificial neural networks can provide the necessary and fundamental structure for the emergence of mental states. However, we claim that the realization of neural firings on artificial neural networks is necessary but not sufficient to simulate human mental states. Because human mental states cannot be reduced to neural firings or isolated from other mental states. Also, chaos and quantum theories are thought of as describing the unpredictable nature of the human mind, and these theories are suggested by artificial intelligence researchers and functionalist philosophers to realize the human mind in machines. But mental states are not random. Even if our mental states have a certain level of randomness, we can still control and decide our actions. Furthermore, robots, which are the inorganic realization of artificial intelligence systems, are expected to be fully determined systems. In order to solve this problem, also consistent with the physicalist basis of the sciences, we see many approaches to expanding and improving determinism to cover the indeterminist and probabilistic aspects of the human mind. But these attempts cannot be successful because even if we accept the existence of probabilistic events in a deterministic understanding, this acceptance would have ignored the fundamental qualities of beings, such as impenetrability. Because impenetrability must also have a continuum through time and space. This continuum requires a strict determination of the being.

My critique needed an analysis of the concept of artificial intelligence. The second chapter focuses on five points related to the concept. First, I analyzed the existing artificial intelligence definitions in the literature and came to the conclusion that most of these definitions are based on the presupposition that human mental states are reducible to one of the modes of mind, which is intelligence, and that intelligence can be represented by algorithms, which can also be executable on computers. Furthermore, these definitions provided the conditions for a philosophical functionalist understanding of human mental states. The technical processes and engineering basis of the artificial intelligence studies were also important. In these fields, there were several presuppositions that enabled the idea of mathematicalizing human mental states: representation of knowledge and artificial neural networks. On the scientific basis of these fields, human mental states have been considered as some neural firings and their relations. This idea led artificial intelligence researchers to the realization of the organizational structure of organic neural networks on inorganic material. If this inorganic material was designed properly, it could provide an appropriate basis for the emergence of mental states. The representation of knowledge in this system was easy: Every sensory input could be reduced to electrical voltages, and a bunch of zeros and ones could represent beings, facts, or concepts of our world. Through the conceptual analysis of artificial intelligence, I had to inquire into the concept of intelligence as the main and fundamental concept of this field. Because, without this inquiry, the most commonly used adjectives in the philosophy of artificial intelligence, such as natural or artificial would not be understandable in this critique. Furthermore, in the fourth section of the second chapter, I examined the relationship between the concept of intelligence and the concept of mind. Because in the field of artificial intelligence, human mental states are considered intelligence rather than modes and states of the mind. But I claim that this reductionist understanding of the human mind is not valid. First, even if we can reduce certain mental states to the fundamental level of calculation and computation, some mental states resist this formalization. And a mind, with all its mental states, is not something we can formalize with all of its phenomenal qualities and contents. That is why the efforts in artificial studies, which aim to realize human mental states in machines, cannot be successful because of this unjustified reduction. There were a

few topics that I must inquire into in order to further analyze the concept of artificial intelligence in the philosophical functionalist approach: Limits of a body from the perspective of mind-body problems, the concept of self, and agency. Because if we accept assigning agency status to artificial intelligence systems, we must answer several problems about these concepts.

The final chapter of my critique aims to analyze and discuss functionalism and the computational-representational functionalist approach. First, I focused on functionalism and its definitions and main arguments. Because, I had to clarify the distinction I made between functionalism and philosophical functionalism in order for my critique to be understood. After this clarification, the computational-representational aspects of philosophical functionalism had to be revealed. Therefore, I focused on the relationship between the philosophical functionalist approach and computer science. With this inquiry, I aimed to reveal the ontological and epistemological status of fundamental concepts from these fields, such as computers, programs, and algorithms. Because the computational-representational level of this approach can contain the main propositions and possibilities of artificial intelligence systems that could be used to build sentient and autonomous agents, according to certain philosophers and researchers in these fields. At the end of the chapter, I wanted to mention a few problems from the standpoint of the arts and the philosophy of art: What would happen to our understanding of art if we accepted the philosophical functionalist argument and built artificial intelligence systems that could act freely, feel, etc.? Could they be artists, or do they just simulate the processes?

Keywords: Artificial Intelligence, Functionalism, Mind, Body, Computer, Computability Theory, Turing, Putnam, Searle, McCarthy.

ÖNSÖZ

Yapay zekâ felsefesine ilgi duymamı sağlayan rahmetli babam İbrahim TAZEGÜL'ü anmak istiyorum. Kendisiyle canlı yaşamıyla ilgili tartışmalarda bulunur, insanlarla diğer canlıları birbirinden ayıran hususlar üzerinde tartışırdık. İnsanın neliği sorunu üzerine odaklanmamı sağlayan bu tartışmalar, kendisinin vefatından önce konuşmaya pek az fırsat bulduğumuz yapay zekâlı varlıklar üzerine eğilmemi mümkün kılmıştır.

Lisans bitirme tezimi yazarken kaybettiğim saygıdeğer babam, benim yapay zekâ felsefesi çalışmalarımın ve henüz kendisi hayattayken yeni yeni geliştirmekte olduğum fikirlerimin büyüyüp filizlenmesini görmeyi çok isterdi elbette. Dolayısıyla akademik kariyerimin bu aşamasında, bu çalışmada ortaya koyduğum düşüncelerim, beni bu yeni ve engin bir alan olan yapay zekâ felsefesine yönelten, beni düşünsel anlamda destekleyen ve kendi bilgisi ölçüsünde katkılar sunan saygıdeğer babamın hatırasına olan saygımın bir gerçekleşmesidir.

Saygıdeğer babamın farklılıklara olan saygısı, uyuma ve çalışmaya verdiği değer sahip olduğum felsefi tavır ve disiplinin oluşmasında temel olmuştur. Düşünce yaşamını ve felsefeciliği arzu ettiğimde kendisi bu arzumu koşulsuzca desteklemiştir. Nitekim ben de bu koşulsuz desteği ve özgürlük alanına olan saygıyı bir kor olarak taşımaktayım. Umarım bu kor, günün birinde miras bırakabileceğim bir değer olarak canlı kalır.

Saygıdeğer annem Sakine TAZEGÜL'ü de anmak istiyorum. Kendisinden öğrendiğim disiplin ve sıkı çalışma gibi kişilik özellikleri, okuyacak olduğunuz bu *ilk eleştirimin* meydana gelmesinde büyük rol oynamıştır.

İSTANBUL – 2023

TARIK TAZEGÜL

İÇİNDEKİLER

ÖZ ii

ABSTRACT v

ÖNSÖZ x

İÇİNDEKİLER xi

GİRİŞ 1

BİRİNCİ BÖLÜM

METAFELSEFİ İNCELEMELER

1.1. "Yapay Zekâ" Kavramının Ele Alınması Bakımından Yapay Zekâ Felsefesi Üzerine Saptamalar 7

1.1.1. Yapay Zekâ Felsefesindeki Antropomorfik Yükler 22

1.1.2. Yapay Zekâ Felsefesinde Fizikalist Savların Yeri 30

1.2. Yapay Zekâ Felsefesine Temel Oluşturan Bilimler Üzerine 40

1.2.1. Yapay Zekâ Çalışmalarının Matematiksel Temelleri 41

1.2.2. Yapay Zekâ Çalışmalarının Mantık Bilimine Dayanan Temelleri 59

1.2.3. Yapay Zekâ Çalışmalarının Doğa Bilimlerine Dayanan Temelleri 85

1.2.4. Bilimler Temelinde Ulaşılmaya Çalışılan Bir Metafizik Tasarım Olarak Yapay Zekâ Kavramı 104

İKİNCİ BÖLÜM

YAPAY ZEKÂ KAVRAMI ÜZERİNE

2.1. Mevcut Yapay Zekâ Tanımlarının İncelenmesi 107

2.2. Yapay Zekâ Çalışmalarındaki Teknik Unsurlar 116

2.3. Yapay Zekâ Çalışmalarında Zekâ Tanımları ve Yapay/Doğal Zekâ Ayrımı 123

2.4. Yapay Zekâ Çalışmalarında "Zihin" Kavramının Yadsınması 133

2.5. Zihin-Beden İlişkisi Bağlamında Yapay Zekâlı Varlıkların Sınırı 137

ÜÇÜNCÜ BÖLÜM

FELSEFİ İŞLEVSELCİLİĞİN ELEŞTİRİSİ

3.1. İşlevselcilik, Felsefi İşlevselcilik ve Yapay Zekâ İlişkisi 160

3.2. Felsefi İşlevselciliğin Bilgisayar Bilimine Dayanan Temelleri 174

3.3. Felsefi İşlevselciliğin Sonuçları Bakımından Yapay Zekâ ve Sanat İlişkisi Üzerine Bir Deneme..... 199

SONUÇ..... 204

KAYNAKÇA 213

GİRİŞ

Yapay zekâ felsefesi tartışmaları "yapay zekâ" kavramının neliği üzerinde şekillenen ve yapay zekâlı varlıkların zihin felsefesinden de devralınan problemleri karşısındaki statülerinin belirlenmeye çalışıldığı bir alandır. Yapay zekâlı varlıkların bilinç, özgür irade, faillik statüsü gibi birçok kavram ve durum karşısındaki konumu, yapay zekâ felsefesi tartışmaları içerisinde yalnızca felsefi temelde değil, aynı zamanda bilgisayar bilimi, matematik, mantık ve doğa bilimleri temelinde de çözümlenmeye çalışılmıştır. Ancak eleştirimizde konu edineceğimiz husus yapay zekâlı varlıkların burada dile getirdiğimiz zihnin kimi belli yetileri veya durumları bakımından statüleri üzerinde durmak değildir. Buradaki amacımız yapay zekâlı varlıkları zihnin tüm yetileri bakımından ele almaktır. Eleştirimiz, zihnin makinelerde gerçekleştirilmesinin koşullarının yapay zekâ çalışmaları içerisinde başarılabileceğini savunan felsefi işlevselcilik görüşü üzerine eğilmektedir. Ancak burada işlevselcilik var olanları, onları meydana getiren materyalden soyutlayarak işlevleri üzerinden analiz eden genel bir yaklaşım olarak ele alınmamaktadır. Tam aksine işlevselcilik, yapay zekâ çalışmalarına temel oluşturan bir bilim olarak bilgisayar bilimi bakımından özel bir işlevselcilik temelinde eleştirilmektedir: Bilgisayımsal-temsil işlevselciliği. İşlevselciliğin yapay zekâ felsefesi bakımından bu eleştiri içerisindeki tasnif edilmiş kullanımları, Ned Block'un işlevselcilik tasnifine dayanmaktadır. Bu vesileyle eleştirimizin hedefi, işlevselciliğin elenmesi değildir: İşlevselciliğin bilgisayımsal-temsil işlevselciliğine indirgenen ve problemli bir savunusunun eleştirilmesidir. Çünkü böyle bir indirgeme beyni bir bilgisayar ve zihinlerimizi ise bu bilgisayarlarda çalışmakta olan birer program olarak kabul etmektedir. Yapay zekâ çalışmalarının ise zihin sahibi bir varlık olmak için organik bir bedene sahip olmanın *ilineksel olduğunu* gösteren önemli bir başarı olduğu ileri sürülmektedir. İşlevlere indirgenen zihin, matematik ve mantık yoluyla bilgisayar biliminin sunduğu bir olanak olarak bilgisayarlarda gerçekleştirilmek istenmektedir. Bu yolla da etten ve kemikten kurtarılmış zihin, şimdiye gelesiye *insan zihni* olarak anılmışsa da artık zihin sahibi olmak bakımından insan olmanın bir önemi kalmamış bulunmaktadır. Çünkü zihnin işlevleri bir kere matematik ve mantık yoluyla biçimselleştirildiğinde, bu zihnin işlevlerini gerçekleştirebilecek *herhangi bir*

materyal organizasyonun zihin sahibi olabileceğini iddia edebilir duruma gelebilmekteyiz. Ancak bilgisayımsal-temsil temelinde şekillenen böyle bir işlevselcilik ne "insan" kavramının felsefece bir hesabını vermekte ne de üzerinde yükseldiği bilimler temelinde yer alan temel tartışmalara bir yanıt sunmaktadır. Dolayısıyla bu çalışmada üzerinde duracağımız bilgisayımsal-temsil işlevselciliği, sadece kendisi bakımından ele alınamaz. Böyle bir işlevselcilik, sadece bilgisayımsal temeldeki iddialarıyla değil, aynı zamanda antropomorfik ve fizikalist temellere dayanan kabulleriyle de yapay zekâ felsefesi tartışmalarına felsefece bir incelemenin sonucu olmayan kimi metafizik kabulleri dahil etmek istemektedir. Bu anlamda, eleştirimizde felsefi işlevselcilik olarak kullandığımız adlandırma, yalnızca yapay zekâ felsefesi içerisindeki bilgisayımsal-temsil işlevselciliğini değil aynı zamanda antropomorfik ve fizikalist temellere dayanan metafizik kabulleri de niteleyen bir adlandırmadır. Felsefi işlevselci yaklaşım, insanın zihinsel durumlarının matematik temelinde inşa edilebileceğini savunur, böylece matematiksel semboller ve operasyonların gerçekliğimizin temsil edilmesi konusunda sağlam bir temel oluşturabileceğini öne sürer. Matematiksel temelde inşa edilmiş bir yapay zekâlı varlığın, yalnızca onu oluşturan mantıksal yapı ve bu yapının güvencesini sağlayan mantıksal ilke ve nesnelere dayanarak sağduyu bilgisi de dahil olmak üzere gerçeklik içerisindeki tüm olgu ve durumları temsil edebileceği de savunulmaktadır. Felsefi işlevselciliğin burada dile getirdiğimiz matematik ve mantık temelinde inşa etmeye çalıştığı zihinler olarak yapay zekâlı varlıklar, insan zihinselliğinin doğa bilimleri içerisindeki bir ele alması olarak nöral ateşlemelerden ibaret sinir ağları üzerinde gerçekleştirilmek istenir. Böylece yapay nöral ağlar üzerinde matematik ve mantık temelinde biçimselleştirilmiş olan bir yapının, zihnin belirmesi için yeterli koşulları sağladığı düşünülür. İşte felsefi işlevselcilik olarak adlandırdığımız ve yapay zekâlı varlıkların sadece zihin durumlarını gerçekleştirebileceğini değil, aynı zamanda onlara sahip de olabileceklerini ileri süren görüş, bu varlıkların fail olmak bakımından da yetili olabileceklerini iddia etmektedir. Çünkü işlevselciliğin bir yapının işlevsel analizi hususunda önemli olanın o yapıyı meydana getiren materyali değil ama yapının işlevlerini özsel kabul ettiği düşünülürse, felsefi işlevselci yaklaşım açısından yapay zekâlı varlıkların bedenleşmeleri nasıl olursa olsun sahip oldukları bedenlerde, kendilik statüsüne de sahip olabileceği ilk adımda kabul

edilmektedir. Dolayısıyla felsefi işlevselciliğe göre mühim olan şey bir yapının *işlevsel analizi* olarak yalnızca insan bedeni ve zihninin onu oluşturan işlevlere ve yapılara ayrıştırılması ya da *metafizik işlevselciliğe* göre bu işlevlerin asıl anlamda ne olduklarının ortaya koyulması değil ama insan bedeninin yadsınması üzerinden insan zihninin bilgisayarlar ve bu bilgisayarlarda matematiksel ve mantıksal temelde temsil edilmesine dayanan bir yapıda gerçekleştirilebileceğinin savunulmasıdır. İşte eleştirimiz tam da böyle bir işlevselciliğe üç temel adımda bir eleştiri getirmektedir: Yapay zekâ felsefesi ve felsefi işlevselcilik üzerine yöntemsel inceleme, "yapay zekâ" kavramının kavramsal analizi ve işlevselcilik/felsefi işlevselcilik üzerine bir eleştiri.

Eleştirimizin yöntemsel incelemesi "metafelsefi incelemeler" olarak anılacaktır. Yöntemsel incelemenin "metafelsefi incelemeler" olarak anılmasının gerekçeleri açıktır: Bilgisayımsal-temsil işlevselciliği temelinde gelişen felsefi işlevselciliğin eleştirisini sunabilmek için bu alandaki metafizik kabullerin elenmesi gerekmektedir. İnsan ve yapay zekâ arasında kurulmak istenen ilişkinin felsefece yürütülebilmesi için, alanın edebi ve tarihsel kurgulardan arındırılması gerekmektedir. Çünkü tüm bu kurgular, felsefe dışı alanlardan alınıp yeniden felsefi tartışmaların içerisine konumlandırıldığında felsefi tartışmalar spekülatif tartışmalardan daha fazlası olmamaktadır. Ancak biz yapay zekâ felsefesi bakımından böyle bir felsefe yapmanın, felsefece olmadığını ortaya koyabilmek için yapay zekâ felsefesi içerisindeki antropomorfik ve fizikalist temellere dayanan ancak sıkıca incelendiklerinde güvenceleri bulunmayan kabullerin ve kullanımların okuyucu açısından da açık kılınması gerektiğini düşünüyoruz. Böylece bilgisayımsal-temsil işlevselciliğine dayanan ve "insan" kavramına alternatif bir kavram yaratma arzusunda bulunan böyle bir felsefe yapmanın felsefece olmayan kabullerinden kurtulabilirsek, "yapay zekâ" kavramının kavramsal incelemesinin de felsefece başarılabileceğine ilişkin olanağın bulunduğuna inanıyoruz. Yöntemsel incelememizin ikinci alt bölümü matematik, mantık ve doğa bilimleri temelinde yükselen felsefi işlevselciliğin bu bilimlerden gücünü alan ancak bu bilimler temelinde de güvencelerinin tartışmalı olduğu hususları hedef alacaktır. Böylece bilgisayımsal-temsil işlevselciliğinin sembollerin ve semboller arası ilişkilerin

gücüne dayanan güvencesi matematik ve mantık bilimleri temelinde yeniden değerlendirilecek ve bu bilimlerin olanakları ile felsefi tartışmaları bakımından yeniden ele alınacaktır. Doğa bilimlerinin incelenmesine duyulan ihtiyaç ise kuantum ve kaos kuramı başta olmak üzere kimi kuramların fizik evrenin sıkı determinist yapısı içerisinde indeterminist koşulları oluşturmak bakımından olanak sağladığının düşünülmesidir. Böylece bilgisayımsal-temsil işlevselciliğinin dayandığı bu ve benzeri kuramların, insan zihninin indeterminist yapısının makinelerde sağlanması bakımından ileri sürdüğü fikirleri incelemenin gerekli olduğunu düşünüyoruz.

Eleştirimizin ikinci bölümünü oluşturan kavramsal inceleme, felsefi işlevselciliğin dayandığı temel kavram olarak "yapay zekâ" kavramını çok yönlü olarak ele almayı amaçlamaktadır. Çünkü böyle bir işlevselciliğin başarıları güvence altına alınacaksa bu güvencenin sıkı bir kavramsal inceleme temelinde de sağlanması gerekmektedir. Kavramsal incelememiz literatürdeki yapay zekâ tanımlarının serimlenmesini ve bu tanımların karşılaştırmalı olarak incelemesini ortaya koyacaktır. "Yapay zekâ" kavramına ilişkin tanımların, yapay zekâ çalışmaları içerisindeki teknik unsurlarla olan ilişkisi, kavramsal incelememizin ikinci adımını oluşturacaktır. Böylece yapay zekâ tanımları ve yapay zekâ çalışmalarındaki teknik unsurlar arasındaki ilişkilerin ortaya konması ile bilgisayımsal-temsil işlevselciliği temelindeki felsefi işlevselciliğin dayandığı tanımların da açık kılınması amaçlanmaktadır. Bu vesileyle "zekâ" kavramına getirilen "doğal" ve "yapay" gibi sıfatların, zekâ tanımlarımız bakımından önemini de ortaya koyarak "doğal zekâ" ve "yapay zekâ" gibi literatürde zekâ kavramını incelememizi güçleştiren kullanımların işaret ettiği hususlar açık kılınacaktır. İncelememiz, yapay zekâ felsefesi ve yapay zekâ çalışmaları bakımından başarılmak istenen bir hedef olarak zihnin, "zekâ" kavramına indirgenmesini de içermektedir. Böylece zihnin makinelerde gerçekleştirilmesi için zekâ gibi zihnin yalnızca hesaplamaya ilişkin bir yetisini kendine temel edinen, dolayısıyla zihnin tüm yetilerini hesaplamaya dayanan zekâya indirgeyerek gerçekleştirmek isteyen felsefi işlevselciliğin, zihin kavramını yadsıması üzerinde durmak istiyoruz. Kavramsal incelememizin son alt bölümünde ise zihin-beden tartışmaları bakımından felsefi işlevselciliğin başarmak istediği bir kavram olarak yapay zekâlı varlıklar üzerinde duracağız. İşlevselleştirilen ve

bedenden soyutlanan zihin düşüncesinin kendilik, faillik, fenomenal deneyimler vb. gibi bedenli olmayla ilişkili bulunan birçok husus konusunda kimi tartışmalar barındırdığını düşündüğümüzden dolayı bu tartışmaların bir incelemesini sunarak felsefi işlevselciliğin zihin-beden probleminden devraldığı sorunlara ilişkin sıkı temellendirilmiş bir duruş ortaya koyup koyamadığını okuyucu açısından da açık kılmaya çalışacağız.

Eleştirimizin son bölümünde ise işlevselciliğin tanımı, türleri ve işlevselci argümanda öne çıkan düşünürlerin yapay zekâ felsefesi bakımından duruşlarını ele alacağız. Bu ele almanın okuyucu açısından birkaç hususu açığa çıkarmasını hedefliyoruz: İşlevselcilik yalnızca felsefe içerisinde sahip olunan bir tavır değildir. İşlevselcilik birçok bilgi ve bilim alanına sirayet etmiş olan bir tavırdır. Dolayısıyla felsefe içerisindeki izdüşümüyle işlevselcilik, diğer bilgi ve bilim alanlarına dayanan bir tarzla farklı alanlardaki işlevselcilikleri barındıran farklı alt yaklaşımlarla da karşımıza çıkabilmektedir. İşlevselciliğin tanım ve türlerine ilişkin ortaya koymayı arzuladığımız tasnif, işlevselciliğe bu çalışma içerisinde getirdiğimiz "felsefi" sıfatının da anlamını açık kılacaktır: Yalın olarak "işlevselcilik" adlandırmasının kullanılmasıyla ortaya çıkması muhtemel bir durum olarak işlevselcilik içerisindeki tüm işlevsel yaklaşımların eleştirildiği anlaşılabilir. Oysa felsefi işlevselcilik kullanımıyla, işlevselcilik içerisindeki kimi işlevselciliklerin işaret edilmesini arzuluyoruz. Bilgisayımsal-temsil işlevselciliği ve bu temele dayanan tüm işlevselcilikler, "felsefi işlevselcilik" olarak adlandırılmakta ve eleştirilmektedir. Dolayısıyla "felsefi işlevselcilik" adlandırmasının, yapay zekâ felsefesi içerisinde başta bilgisayar bilimi ve diğer ilintili bilim alanlarıyla da birlikte yapay zekâlı varlıkların, bilgisayımsal temellere dayanarak insansı failler olacakları, dünya tasarımlarının insanlardan farklı olmayacağı ve onların da haklar ve sorumluluklar bakımından insanlarla eşit olmaları gereken veya olacakları bir noktanın bulunacağı türünden iddiaları kapsayan genel bir adlandırma olduğunu ortaya koymak istiyoruz. Bu bakımdan yapay zekâ felsefesi içerisindeki işlevselcilik düşüncesinde, bazı işlevselciliklerin terk edilmesi ancak işlevsel analiz veya metafizik işlevselcilik gibi bazı işlevselciliklerin tercih edilmesi yoluyla sürdürülebilecek işlevselci tavırlar, eleştirimizin yalnızca işlevselciliğin eleştirisi olarak adlandırılması yoluyla

kapsanamaz. “Felsefi işlevselcilik” olarak adlandırdığımız tavrın, eleştirimizin merkezine aldığı işlevselciliklerin tespit ve işaret edilmesi konusunda aydınlatıcı olduğunu düşünüyoruz. “Felsefi işlevselcilik” olarak bu eleştiride özellikle tanımladığımız türdeki bilgisayımsal-temsil temelli işlevselciliğin yapay zekâ çalışmalarıyla olan ilişkisini de ortaya koymayı amaçlıyoruz. Bu anlamda bilgisayar bilimi tarihi bakımından bilgisayarlara ve kimi kavramlara ilişkin arka plan bilgisi sunulduktan sonra bu alana özgü olup felsefi işlevselci savlar bakımından tartışma yaratan durma sorunu, bilginin temsili, kuantum bilgi işleme gibi problemler üzerinde durulacaktır. Son kertede ise felsefi işlevselci görüşün bir sonucu olarak bir bilgisayarda işleyen bir programın tıpkı biz insanların sahip olduğu gibi bir zihin sahibi olduğu varsayıldığında, bu durumun sanat alanında doğuracağı problemlere ilişkin bir deneme sunma niyetindeyiz. Eleştirimizin yöntemsel, kavramsal ve tavır olarak ortaya koyduğu analiz ve eleştirilerin, yapay zekâ felsefesi bakımından etik ve politik sorunların çözümsüzlüğünün aşılması hususunda gösterdiği ontolojik ve epistemolojik mahiyetin de etik ve politik felsefenin yapay zekâ felsefesi bakımından sahip olduğu problemlere çözüm getirmek konusunda bir alternatif oluşturacağını düşünüyoruz.

BİRİNCİ BÖLÜM
METAFELSEFİ İNCELEMELER

1.1. "Yapay Zekâ" Kavramının Ele Alınması Bakımından Yapay Zekâ Felsefesi Üzerine Saptamalar

"... otomat hikayesi ruhlarının derinliklerinde kök salmış ve insan figürlerine karşı korkunç bir kuşku gerçekten de içlerine dolmuştu. Âşık oldukları kişinin bir tahta bebek olmadığından tamamen emin olabilmek için, pek çok aşık sevgilisinden biraz usule aykırı şarkı söylemesini ve dans etmesini, bir şey okunurken sürekli nakış işlemesini, küçük köpeklerle oynamasını vs. istiyordu; ama özellikle de yalnızca dinlememesini, arada bir konuşmasını ve bu konuşma biçiminin gerçekten de düşünmeyi ve hissetmeyi gerektirmesini bekliyordu."[1]

İnsanlık tarihini, insanın yapıp etmelerinin tarihi olarak ele almakla belki de o tarihe haksızlık ediyor olabiliriz. Bir toprak parçasına benim demekle başlayan süreçte sabanın, kılıcın, tüfeğin ve çarkların insanın yaşamına kattığı ve onda dönüştürdüğü anlamlarla birlikte düşünüldüğünde insanlık tarihini, devinen ve değişen şeylerin tarihi olarak görmeyi tercih ediyoruz. Peki yalnızca metalar mı değişir? Kesinlikle hayır. İnsanın yaşamındaki araç ve o araçla yöneldiği her türlü doğa nesnesi gibi insanın düşüncesi ve şeyleri algılayışı bakımından dünya kavramı da kendine bu değişimde yer bulur. Ancak insanın düşüncesi, taşıdığı yükler bakımından ne kökeninden ne de o düşünceyi doğurmuş ve doğurmakta olan nesnelerin organize oluşu bakımından insan yaşamından yalıtık bir şekilde ele alınabilir. O sebeple bir zamanlar çarklar ve kömürler, sahiplik ve sınıf için ne anlam ifade ettiyse, veri ve bilgi işleme çağı olan günümüzde de yapay zekâlar ancak ve ancak yine yaşamında yer bulduğu insan, üzerinde var olduğu bilgi işlem araçları ve yine ona hizmet ettiği insan yaşamı ile anlamlı ve açık kılınabilir. Böylece yapay zekânın açık kılınabileceği tek yol, ne onu matematik ve mantıkla donatan mühendis sınıfının icra ettiği mühendislik olabilir ne de onu eyleyebileceği araçlar ve antropomorfik bedenlerle yaşama nüfuz edebilmesini sağlayan teknikerlik zanaatı olabilir. Yapay zekânın açık ve seçik kılınabileceği yegâne alan felsefedir. Çünkü felsefe, kavramlarımızın yalnızca kendileri bakımından değil, eşzamanda diğer

[1] E. T. A. Hoffmann, **Kum Adam**, çev. Ersel Kayaoğlu ve Sami Türk, İstanbul, Can Yayınları, 2020, s. 45.

kavramlarla birliktelikleri bakımından da düşünüldüğü, ancak bu düşünmenin spekülatif kabuller ve tavırlardan arındırıldığı bir alandır.

Bilgi işleme çağında insanlığın yüzleştiği sorunlar yalnızca zanaat kollarını kapsamaz. "Yapay zekâ" kavramının 20. yy. ortası belirmesiyle beraber insanlık bir bilgi ve alan sorunu yaşamaktadır. Artık insan, bir **polis**'in öznesi olmaktan ziyade, aynı zamanda sanal bir var olma tarzının da nesnesidir. O artık *şeyleşmiştir.* Gerçek ve sanal arasındaki ayrım yok olmuş, insan kavramımız ise makineler tarafından silikleştirilmiş, böylece bilgiye erişemeyen aydınlanma insanından, bilgi bolluğu altında kafa karışıklığı yaşayan modern insana dönüştürülmüş durumdayız.[2] Bu bilgi bolluğu ve belirsizliğinin sonuçları kendisini "yapay zekâ" kavramının bulanıklaştığı, bu projede kendisine pay bulan her alan tarafından "yapay zekâ" kavramının modern zaman mitlerine konu edildiği, Orta Çağ insanının Tanrı karşısındaki anonimliğinden, modern zamanların alternatif insanına dönüştürüldüğü bir gündemi yaşamaktayız.

Peki yapay zekâ projesi nedir? Kuramsal ve deneysel yanlarla sürdürülen yapay zekâ çalışmaları bir yanıyla kendisini insan zekâsının biyolojik mahiyetlerinde – yani insanın psikolojik ve fizyolojisini taklit etmeyi hedefleyen – bulurken bir yanıyla da insan zihninin fenomenal yaşantısını matematik ve mantık yolu ile biçimselleştirme amacını taşımaktadır.[3] Dikkat edilirse yapay zekâ projesinin temel amaçlarından birinin insan zekâsının makineler üzerinde gerçekleştirilmesi çabası olduğu açıkça görülebilir. Bu çaba boşuna değildir. İnsanın Dünya üzerinde durmak bilmeyen *kendini yaratma* süreci, artık süreci sadece kendi yürütebileceği bir yük değildir. İnsan kendisine bir yardımcı aramaktadır. Böylece yapay zekâ projesinin bir diğer amacının da yeryüzünde zekice eyleyebilecek, insanları bu yaratım sürecinde destekleyebilecek ve bunu yine insanca, insan zekâsının işleyişinden pay alırcasına

[2] Luciano Floridi, "The Onlife Manifesto: Being Human in Hyperconnected Era", (Çevrimiçi) https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/Manifesto.pdf, 25 Aralık 2022.
[3] John McCarthy, "What is Artificial Intelligence?", (Çevrimiçi) http://www-formal.stanford.edu/jmc/whatisai.pdf, 25 Aralık 2022.

yapabilecek bilgisayar programları – yani yapay zekâlar – tasarlamak olduğunu görmek mümkündür.[4]

Ancak bu çalışmada ne bu çalışmanın ne de felsefenin – burada diğer bilim ve bilgi alanları bakımından düşünülürse – daha önemli bir kısmını oluşturan mühendislik ve iktisatla ilgili kaygıları değil, "yapay zekâ" kavramının düşünme, eyleme gibi temel insan edimleri karşısındaki rolünü ele alacağız. Nitekim bu kavramın felsefe ile olan ilintisini görmezden gelmek ve onu felsefe dışı olarak kabul etmek kavramın felsefi önemini anlamamaktan ileri gelmektedir. Çağdaş felsefenin bizatihi konularından biri olarak yapay zekâ, zihin felsefecilerinin olduğu kadar sağduyu bilgisinin söz konusu olduğu her felsefi tartışmanın da bir kavramıdır.[5] Çünkü, yapay zekâ kavramı, felsefenin değil, olsa olsa zanaat alanlarının kavramı olabilir demek, Dünya'da var olan, Dünya üzerinde eyleyen ve orada iş gören bir varlığın yine var olanlarla olan ilişkisinde ancak kişinin felsefenin rolünü unutmuş olmasıyla açıklanabilir. "Var olan nedir?", "Var olan nasıl bilinir?" ve "Var olan ne şekilde düşünülür ve var olan nasıl düşünülür kılınır?" gibi sorular, eğer yapay zekâlı varlıklara failliği atfetme hakkını kendimizde göreceksek yanıtlamamız gereken sorular olacaktır. Çünkü yapay zekâlara faillik atfeden McCarthy gibi isimler bile, yapay zekâlı varlıkların duyu verilerinin ancak ve ancak yine felsefi uslamlamalarla ortaya konan ontolojik ve epistemolojik görüşler vasıtasıyla yapay zekâlar tarafından çerçevelenebileceğini iddia etmektedir.[6]

İnsanın düşünmesinin tarihinde, canlı ile cansızı ayıranın ne olduğu, taş ve toprakla benzer materyalden güç bulan insan usunun nasıl oluyor da taş ve toprakta zuhur edemediğini – eğer bir animist değilseniz – felsefeciler ve teologlar binlerce yıldır tartışmıştır. Zihin felsefesinin de göbeğinde yer alan bilinç, özgür irade, nitelceler* vb. gibi fenomenler genel olarak düşünüldüğünde, insan zihninin yine ondan can bulduğu o cansız, durağan ve edilgen taş ve toprak ile aynı doğa yasaları

[4] Herbert A. Simon, "Artificial Intelligence: an empirical science", **Artificial Intelligence**, No: 77, 1995, s. 96.
[5] McCarthy, "What is Artificial Intelligence?".
[6] John McCarthy, "Mathematical Logic in Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 305.
* Lat. **Qualia**

tarafından yönetilmemesi için bir sebebin olmadığını da düşünebiliriz.[7] Nitekim yapay zekâ devrimi de kendisine buralardan güç bulur. Silikon ve elektrikle donatılmış ve algoritmalara – şu an bilgisayarınızın monitörünü kapatmanız için gereken adımları emir cümleleri ile satır satır bir kâğıt üzerine yazarsanız bu emirlerin tümüne algoritma denebilir – dökülmüş olan zekânın, insan zekâsı gibi düşünmesinin önünde bir engel olmadığını da söylemek bu görüşe göre oldukça tabiidir. Ve hatta insanlar gibi düşünüyorlarsa **polis**'te onlarla eş haklara sahip olmaları gerektiğini belki de haykırabiliriz![8]

Ancak böyle bir iddia – her ne kadar felsefi işlevselcilikle, materyalist bir zihin felsefeciliği ile kendine yankı bulmuşsa bile – bu çalışmanın kapsamı bakımından detaylıca incelenmek zorunda kalınmıştır. Çünkü kavramları dile getirme hakkımızın olmadığı, onlar üzerine düşünmemizle ilgili engellerin olduğu yollu değil ama bu kavramların ve düşüncelerin sıkıca dile getirilmesinin, o kavram ve düşüncelerin dayandığı bilimler ve alt kavramlar bakımından da hesabının verilmesi türünden bir disiplinin, felsefe yapmanın en temel ilkelerinden biri olduğu unutulmamalıdır. Dolayısıyla bir felsefecinin veya bir bilim insanının işi yalnızca düşünceleri dile getirmek değil, dile getirilen düşüncelerin yankıları bakımından da sorumluluğu ele almaktır.[*] Nitekim toplumun içinden toplumca yaşama ve toplumun yaşamı hakkındaki belirleyici rollere bürünmekle ayrıştığımız biz akademisyen ve bilim insanları, ilgi ve yargı alanlarımızla ilgili felsefi payı üstlenmek, bu payla birlikte felsefi kaygıları asla unutmamak ve bu sorgulamaların topluma yansıması işini yerine getirmeyi bir insani görev olarak daima yerine uygulamalıyız.[9]

Ancak 20. yy. biliminin, felsefeyi tarihte hiç olmadığı kadar spekülatif ilan ettiği, felsefece olanın "henüz bilimleşmemiş çabalar" olduğu yollu görüşle birlikte çalışmanın seyri boyunca okuyucu tarafından görülecektir ki, yapay zekâ çalışmalarında bilim insanları ve mühendislerin etki gücü, felsefecilerden oldukça

[7] Raymond Kurzweil, **The Age of Intelligent Machines**, Cambridge, The MIT Press, 1990, s. 13.
[8] Larry Hauser, "Artificial Intelligence", Internet Encyclopedia of Philosophy, (Çevrimiçi) https://iep.utm.edu/artificial-intelligence/#SSH4c.i, 15 Aralık 2022.
[*] Buradan lisans programımda felsefenin ve felsefeciliğin sistem filozoflarını hatim etmekten farklı olduğunu görmemi sağlayan ve metinde dile getirdiğim bu öğüdünü asla unutmayacağım Sayın Prof. Dr. Levent Yurdakul Kavas'ı saygı ile anarım.
[9] Utku Köse, **Yapay Zekâ Felsefesi**, İstanbul, Doğu Kitabevi, 2022.

yüksek olmuştur. Öyle ki, zihin felsefesinin kadim – ve eğer güçlü belirimci bir zihin felsefecisiyseniz çözümsüz kalacak – fenomenleri hakkında kendisine yapay zekâ projesini hedef koymuş olan bilim, bu yolla insan beyninin ve zihninin her türlü yeti ve fenomeni hakkında yapay zekâ projesi vasıtasıyla yanıtlar üretebileceği inancını kendinde taşımaktadır. Bu inanç, insan beynini algoritmalara indirgeyen ve bu algoritmalar sonucunda kendimize ve şeylere *ne olduklarını* sormamız yoluyla çözümsüz sorunun kalmayacağını kabul eden bir inançtır.[10] Eğer bu inanç takip edilirse karşımıza çıkacak bir sonraki kabulün, algoritmalara dökülmüş ve sayısallaştırılmış insan zihninde işleyen düzenlilik ve genelliklerin, tıpkı doğal nesnelerin doğa yasalarına uyarken gösterdikleri düzenlilik ve genellikler gibi nitelikleri göstermesi gerektiği, böylece insan eylemlerinin, insan yaşantısının tüm detaylarıyla bilimsel doktrine uyduğu, onun aracılığı ile boşluksuz bir şekilde ifade edilebileceği, dolayısıyla insanın geçmişi ve geleceği arasındaki eylem ve düşüncesindeki tüm boşlukların ortadan kaldırıldığı bir inanç olacağı açıktır.[11]

Burada okuyucu şu soruyu sorabilir: Burada beis nedir? Bilimin insan doğası üzerindeki bilinmezleri kaldırmasındaki korkutucu unsur nedir? Belki de bu yolla insan doğasının mükemmelleştirilmesi olanaklı kılınabilir. Bu doğru. Ancak bu çalışmanın da kendisine amaç veren asıl düşünceler şunlardır: Biz bu yolla insan kavramını yıkan, ona alternatif değil ama onunla artık fail olmayı iki kardeş gibi paylaşabileceği iddia edilen bir çabanın, sıkıca inşa edilmemiş bir modern zaman mitinin, yani yapay zekâların olanağını güncel sorularla yeniden tartışmaya açmış oluyoruz. Böylece şu sorular ortaya çıkıyor: Yapay zekâlı varlıklar fail olabilir mi? Yapay zekâlar düşünebilir mi? Yapay zekâlar bir midir çok mudur? Yapay zekâlar haklara sahip olabilir mi? Yapay zekâlı bir varlık hayal kurabilir mi? Yapay zekânın ortaya koyduğu eser sanat eseri niteliği kazanabilir mi? O halde biz bu çalışmayla insan doğasını ve insan kavramını muhafazakâr bir teknoloji karşıtlığının veya kimi inançların nesnesi yaparak değil ama felsefece bir kavram olarak, felsefenin bir nesnesi olarak tekrardan ele almaktayız. Onun, sahip olduğu faillik özelliğini taşırken

[10] Anthony J. Bell, "Levels and loops: the future of artificial intelligence and neuroscience", **Philosophical Transactions of the Royal Society of London**, No: 354, 1999, s. 2014.
[11] Mark Fisher, "A note on free will and artificial intelligence", **Philosophia**, C. I, No: 13, 1983, s. 76-77.

onda nelerin ona özgü, ona içkin olduğunu anlama çabasını güdüyoruz. Ancak bunu yapay zekâ felsefesinin, bilgisayar bilimcilerin ve daha sayısı ve kapsamı burada ele alınamayacak kadar çok alanın yapay zekâ hakkındaki felsefeden ve felsefi ciddiyetten uzak çabalarına giderek, oralarda o tanımların ve temellerin sıkı olmadığını göstererek yapma amacını taşıyoruz.

Böylece bilimin insan doğası üzerindeki bilinmezleri ortadan kaldırma ve onu yine kendisinin bir nesnesi kılma çabasının sonuçlarından biri olarak yapay zekâ çalışmaları, şimdiye kadar failliği yalnızca insana özgü kılmış olan felsefenin[*] özgürlük, otonomi, hayal kurma, duygular, bilinç, özgür irade, yurttaşlık, hak, yaratım vb. gibi her kavramını yeniden elden geçirmeyi ve antropomorfik tanımlamalara karşı çıkarak "insan" kavramını yıkmayı, yerine fail olan bedensiz bir zihni, bir **logos**'u koymayı hedeflemektedir. Bu düşüncenin ardında, yapay zekâ projesinin gerçekleştirilebileceği[**] inancı yatmakta, suni bir insanın, metalik bir bedende, bir zekâ olarak yaratılabileceği düşüncesi kabul edilmektedir.[12] Peki ama sormalıyız ki yeryüzüne indirilen, Tanrı ile değil ama yine kendi emeği ile baş başa bırakılan ve ona kendi gücü bahşedilen insan, sahip olduğu bilim ve bilimsel paradigmalar yoluyla tüm metafizik kabullerinden arınmış mıdır? O artık sadece refleksif değil ama aynı zamanda kendi yaratımları bakımından da insan bilgisinin sınırlılığını aşan üretimlerde bulunmuyor mu? İnsan artık mitlerden arınmış mıdır?

Uzay ve zamana iye bir varlık olarak insan ve onun yetileri ile zihinsel fenomenleri, bilimin nesnesidir. Ancak bilimin yanıtsız bıraktığı bilinç, özgür irade, buluşçuluk/yaratıcılık vb. gibi insansı olana atfedilen unsurlar, metafizik kurguların alanı olarak yine de kendine yer bulmaktadır. Bu alanlar, tıpkı Gözkan'ın da dediği gibi doğabilimlerinin yöntembilimsel çerçevelerinin çizdiği sınırların dışında alanlar olarak metafizik sorularla yüzleşen alanlardır. İnsan zihni deneysel psikoloji veya bilişsel bilimler gibi alanlara sıkıştırılamaz, bu doğru. Kendisinin çözümü, felsefe

[*] Örneğin bir Orta Çağ düşünürü için Tanrı karşısında onun muhatabı olabilecek olan varlık olarak, onun yaratımları arasındaki astlık ve üstlük ilişkisi bakımından düşünüldüğünde iradesiyle yeryüzünde en tepede bulunan varlık olarak insan…

[**] Yapay zekâların mevcut olmadığını iddia etmiyoruz. Sadece yapay zekâlı varlıkların insandan aşağı bir yanının kalmadığı/kalmayacağı bir günün geleceğini iddia eden savı hatırlatıyoruz.

[12] Çağlar Ersoy, **Robotlar, Yapay Zekâ ve Hukuk**, 2. bs., İstanbul, On İki Levha Yayıncılık, 2017, s. 10.

içerisinde, metafizik yaparak bir sonuca varmak olmuşsa da[13] döngüsel olarak bakıldığında, yapay zekâ felsefesindeki asıl sorunun, yapay zekâ tanımlanmaları ve o tanımlarla ulaşılmak istenen amaçlar bakımından "yapay zekâ" kavramının sahip olduğu metafizik boşluklar olduğu çok açıktır.

Peki bu metafizik boşluklar ve sorunlar nelerdir? Bilindiği gibi zihin felsefesi, konu edindiği kavramların ele alınışı ve açıklanışı bakımından düalizm ve materyalizm tartışmalarına sahne olmuştur. Ancak Descartes'ın düalizmi günümüzde yalnızca azınlık sayılabilecek felsefeci tarafından dillendirilmektedir ki bu yolla da literatürde düalist argümanların materyalist cephe karşısında terkedildiğini söylemek yanlış olmayacaktır. Peki ama beden ile zihni iki ayrı töz olarak kabul eden[14] düalist yaklaşımın *ruhu* terkedilmiş midir? Düalizmin ruhu, bedenin iye olduğu yasalara tabi değildir. Beden bozunur, devinir ve yiter. Ancak ruh farklı bir töz olarak hiçbir doğa yasasının onda işleyeceği zemin olamaz. Bu şekliyle Descartes'ın ruhu insan bilgisinin sınırlılığını aşan bir alana konumlandırılmış bir varlık olarak zihin felsefesinde yüzyıllar sürecek bir metafizik kabulün alanı olmuştur.[15] Öyle ki bu kabul, nasıl olur da ruh gibi ayrı türden bir tözün, beden gibi yeryüzü varlığı olan, yeryüzünde bir tözü etkileyebildiği, onu devindirebildiği, bu iki töz arasındaki iletişimin ne şekilde olacağı gibi metafizik sorularla tartışılıp durulmuştur. Ancak yanıt alınamamıştır. Aydınlanmanın ve felsefe tarihinin hiçbir zaman unutulmayacak filozofu olan Immanuel Kant bize bunun gerekçesini idealizmin çerçevesinden vermiş olsa da metafiziğin ve kimi metafizik kavram ve konuların insan bilgisinin sınırlılığı dışında yer aldığını göstererek Descartes'ın düalizminin insan bilgisinin sınırlarını aşan kimi yargılarını ayıklamamızı da mümkün kılmıştır.[16]

Çünkü Descartes'ın zihni, insansı bir yaratım ve ifade tarzı ile ortaya konamaz. Matematikselleştirilemez, resmi yapılamaz ve anlatılamaz. Böylece düalist argümanın karşısında üstün gelen materyalizm, insana ait olup izah edemediği,

---

[13] Bülent Gözkan, "Bilgi, Bilinç ve Yapay Zekâ", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs. İstanbul, Pan Yayıncılık, 2012, s. 164.
[14] S. Atakan Altınörs, "Zihin-Beden Problemine Searle'ün Yaklaşımı", **Temaşa Felsefe Dergisi**, No: 9, 2008, s. 48.
[15] Mehtap Doğan, "Biyolojik Doğalcılık Ekseninde John Searle ve Zihin-Beden Problemi", **MetaZihin**, C. I, No: 1, 2018, s. 228.
[16] David Vernon ve Dermot Furlong, "Philosophical Foundations of AI", **50 Years of Artificial Intelligence**, Berlin, Springer, 2007, s. 54.

matematik yoluyla dile getiremediği, üzerindeki sisi kaldıramadığı her türden fenomeni düalizmin iddialarına da referans vererek spekülatif addedecek, gerektiğinde onların birer metafizik kurgu olduğunu söyleyecek, elemeci bir materyalist tavırla bilinç, özgür irade vb. gibi zihinsel yeti ve olanaklılıkların bir zamanların *filojistonu* olduğunu söyleyerek onu bilimden ve insan düşünmesinden ihraç edecektir. Böylece örneğin Zambak'ın metodolojik argümanının materyalist bir evrende yaşadığımıza ve bu evrenin her köşesinde doğa yasalarının hüküm sürdüğüne, dolayısıyla fizik üstü varlıkları – örneğin Tanrı – kabul etmediğine ilişkin görüşlerini kabul ediyor olmamıza karşın aynı argümanın zihnin materyalist bir düzende bulunduğunu dolayısıyla ona ilişkin epistemolojik ve metodolojik tespitlerin materyalist kuramlara dayanması gerektiği ile ilgili görüşlerine katılmamaktayız.[17] Çünkü materyalizm, örneğin nedenlerin zorunluluğu ve var olma tarzı bakımından Spinozacı bir Tanrı anlayışına karşı[18] güçlü gözükse de bu görüş bize insanın da inşa edilebilir, parçaları bakımından anlaşılabilir bir var olan olduğunu iddia etmiyor mu? Bir adım daha ileri gidildiğinde parça ve bütün ilişkisi içerisinde olarak kabul edilen insanı, etinden ve kemiğinden soyutlayarak bir işlevler bütünü, bir "plan" sonucu olarak kabul etmemiz istenmiyor mu? Var olanı düalizm-materyalizm açmazı içerisinde görmek istemeyen Searle gibi filozoflar da vardır. Nitekim Searle zihinsel ve fiziksel, materyalizm ve düalizm, ruh ve et gibi zıt terminolojileri kabul etmemekle birlikte beyinde gerçekleşen fenomenal deneyimlerin de sıradan fiziksel haller olduğunu kabul etmez. Eş zamanda var olan gerçeklik kategorileri olarak zihin-beden problemini aşmanın yolu bu terminolojinin terk edilmesidir.[19] Ancak bizler, yapay zekâlı varlıkların tıpkı insanlardaki gibi çeşitli nöral unsurları taşıdıkları, onları tıpkı insanlarda olduğu gibi temsil edebildikleri, kısacası onlara sahip olabildiklerini saptayabileceğimiz araçlara sahip olmadığımız gibi[20], bir yapay zekânın – tıpkı materyalist ve işlevselci tavrın iddia ettiği gibi – nitelcelere sahip olabildiğini, bir fikre, örneğin satrancın ne olduğuna dair bir fikre sahip olup

[17] Aziz Fevzi Zambak, "Free Will and Artificial Intelligence", **MetaZihin**, C. I, No: 2, 2018, s. 169.
[18] Çetin Türkyılmaz, "Spinoza'da Özgürlük-Zorunluluk Bağlantısı", **Yeditepe'de Felsefe**, No: 6, 2007, s. 180.
[19] John R. Searle, "Free Will as a Problem in Neurobiology", **Philosophy**, C. LXXVI, No: 298, 2001, s. 493.
[20] David Gamez, "The Potential for Consciousness of Artificial Systems", **International Journal of Machine Consciousness**, C. I, No: 2, 2009, s. 214.

olmadığını, karşılaştığı rakibin üst düzey bir hamlesine karşılık farklı hislere bürünüp bürünmediği veya o esnada satranç oynuyor olduğunun bile farkında olup olmadığını tespit edemeyiz.[21]

Düalizmin metafizik yükleri belki de bilim yoluyla, bilimleşmiş bir felsefenin gündeminden düşmüş olabilir. Belki de biz düşünürler, doğa bilimleri ve matematik yoluyla varlık hakkında bilinmezlerin kalmayacağı inancını taşıyor, böylece bu inanca sıkı sıkıya sarılarak varlığın tek bir töze, maddi töze iye olan bir yapıda bulunduğunu kabul ediyor olabiliriz. Bu büyük ölçüde doğrudur. Nitekim insan bilgisinin sınırlılığını aşan kavram ve var olma tarzları hakkında yargılarda bulunmak olsa olsa Orta Çağcı bir metafiziğe özlem olabilir ancak. Fakat burada okuyucunun dikkatinden kaçmasını istemediğimiz bir husus bulunmaktadır: Matematikle ifade edilemediği, maddi tözle ilişkisinin izah edilemediği şey olarak düalizmin zihinsel tözü bilimler ve zanaat alanları tarafından reddedilmiş olabilir. Ancak yapay zekâ projesi ve bu projenin insan doğası hakkındaki amaçları bakımından düşünülecek olursa, böyle bir metafizik kabulden kurtulduğumuz söylenebilir mi? Kesinlikle söylenemez. Çünkü matematik ve mantık yoluyla sayı ve sembollere dökülmeye çalışılan insan zihni ile mekanizmin ve parça bütün ilişkisinin bir ürünü olarak anlaşılmak istenen insan doğası, filozofların yüzyıllardır hayalini kurdukları bir arzudur.[22] Hobbes'un felsefesinde doğa yasalarına sıkı sıkıya bağlı bir mekanizma olarak tasavvur edilen insan, Leibniz'in felsefesinde ise zihni bakımından sahip olduğu tüm felsefi ve metafizik bilgilerin ifade edilip biçimselleştirilebildiği bir tasarının, bir dilin ürünü haline gelecektir.[23] Doğaya sayılar ve geometri yoluyla – analitik geometri ile birlikte tamamıyla sayısallaşan bir durumdan da söz etmek mümkündür – yönelen ve ona çarklar ve buhar gücüyle hükmeden insanoğlu, yine de bu araçlar ve yöntemler yoluyla Descartes'ın felsefesindeki metafizik kabullere ulaşamaz. Böylece insanoğlunun önünde iki seçenek yer alır: Bu öğeleri kabul etmek veya bu öğeleri elimine etmek. Nitekim,

[21] Emir H. Özel, "Church-Turing Tezi ve Algoritmik Zihin Argümanı: Zihinsel Fenomenleri Matematiğe İndirgemek Mümkün müdür?", **Pasajlar**, No: 9, 2021, s. 160.
[22] Daniel C. Dennett, "When Philosophers Encounter Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 284.
[23] Nils J. Nilsson, **The Quest for Artificial Intelligence**, Cambridge, Cambridge University Press, 2010, s. 28.

bilimler yoluyla da kendine güç bulan materyalizm, bu öğelerden kurtulmak istemiştir. Fakat bu kurtulma, kendisine yapay zekâ projesi gibi bir projede yeni bir metafizik tasarım, yeni bir *modern zaman miti* yaratmıştır. İnsan doğasının metafizik kabullerinin, doğaya kesinlik ve güvence veren matematiksel ve mantıksal yöntemlerle buluştuğu bir yaratımdır bu. Böylece yapay zekâ projesi, insan doğasının daha önceleri matematikle ifade edilemeyen, mantıksal dizgeler vasıtasıyla ortaya konamayan zihinsel ve fenomenal durumlarını matematikselleştirme, mantıksallaştırma ve bu yolla işlemciler ve bellekler vasıtasıyla bilinmezleri ortadan kaldırma amacıyla insan doğasını ve "insan" kavramını hiçleştiren bir amacı taşımaktadır. Önceleri materyalist bir dünyada ifade edilemediği için ihraç edilen bilinç, özgür irade gibi fenomenler, yapay zekâ projesi vasıtasıyla örtük bir düalizmin, bir metafizik kabulün canlandığı fenomenler olarak yapay zekâlı varlıklarda kendine yer bulur, silikleştirilir ama onlar vasıtasıyla da yapay zekâlara faillik niteliği kazandırılmış olur. Nitekim Nagel'ın, fenomenal bilincin dile getirilmesi ve üzerine tartışılması hususunda dilin, önermelerin ve bunların doğruluk değerlerinin yetkinliğini sorguladığını biliyoruz.[24] Nitelceler hakkında ve daha birçok fenomenal olgu hakkındaki yargılarımızın, esasında bu nitelce ve olguların erişilmez olmaları karşısında olsa olsa metafizik yargılar olacağını söylememiz için bir engel bulunmamaktadır. Böylece aslında yapay zekâ projesi ile insan zihni hakkında, nitelcelere varasıya bir güvençle her türden fenomenal olguyu açıklama iddiasında bulunan bilim ve mühendislik, kendisini kırmızıyı görmenin, ağrı hissine sahip olmanın veya birine âşık olmanın – duygu sahibi olmanın – nasıl hissettirdiği ile ilgili haller hakkında bile söz sahibi görebiliyor, matematik ve mantık yoluyla, silikon mimarili duyargalar vasıtasıyla da bunu dışsallaştırabildiğini iddia etmiyor mu?[*]

Nagel gibi isimlerin itirazları değerli olsa da materyalist bir felsefenin karşısında bulunan bu ve benzeri sorunlar, türlü analojiler ve düşünce deneyleri vasıtasıyla elimine edilmek istenmiştir. Örneğin Dennett, insanı insan yapan

[24] Thomas Nagel, "What Is It Like to Be a Bat?", **The Philosophical Review**, C. LXXXIII, No: 4, 1974, s. 441.

[*] Okuyucu açısından matematik ve mantık başlıklarında değineceğimiz hususlarla da açık kılınacaktır ki bilimlerin ve mühendisliğin, araçları ve temelleri bakımından yanıtlamaları gereken sorunlar olduğu açıktır.

olguların davranışçı/dışsal gözlemler vasıtasıyla bile ortaya koyulamayacağını vurgulayarak zombi düşüncesinden bahseder ve zombilerin mükemmel bir şekilde doğal, uyanık, geveze ve neşeli tavırlar sergileyen bir insan veya insan olabilecek bir tür otomasyon olduğunu dile getirir.[25] Öte yandan Zeman ise zombi örneğinin, fiziksel olgular sabit kalırken deneyimin yok olabileceğini göstermeyi amaçladığını ortaya koymaktadır.[26] Açıkça görülebilir ki düşünce deneyleri vasıtasıyla "insan" kavramının dayandığı temeller sorgulanmakta ve insan olmanın yalnızca *insana özgü* olmadığı vurgulanmaktadır. Ancak insanın sahip olduğuna inandığımız[*] durumların *öyle görünen* durumlar olabileceğini iddia edebilir miyiz? Nitekim fizikalizmin her çabasını olumsuzlamak fayda sağlamaz. Fıçıdaki beyin türünden bir düşünce deneyinde bilincin içsel veya dışsal bir fenomen olup olmadığı düşünülür. Böylece bu düşünce deneyi ile bilincin zaman ve mekânda nerede bulunduğuna ilişkin tartışmaya bir yanıt aranır.[27] Ancak dikkat edilmelidir ki fıçıdaki beyin deneyinde de zombi deneyinde de düşünce deneyinin gerçekliği ile şu anda yaşadığımız gerçekliğin koşulları arasında bir uçurum bulunmaktadır. Olanaklılıklarıyla eleştirilen bu düşünce deneyleri için Dennett'in çözümü her ne kadar bu deneylerin birer sezgi pompaları[**] olduğunu iddia etmek olmuşsa da bu düşünce deneylerinin sonuçları bakımından geçerlikleri henüz sınanabilir bile değilken, yapay zekâ projesinde dillendirilen savların bu ve benzeri düşünce deneylerinin olasılıksal sonuçlarına dayanıyor olmaları da bir soru işaretidir.

İşte burada, şimdiye kadar sıkça dile getirdiğimiz, okuyucu açısından da önemli bulduğumuz bir izleği yeniden ve açıkça dile getiriyoruz: Metafiziksel

---

[25] Daniel Dennett, **Bilinç Açıklanıyor**, çev. Sibel Kibar, İstanbul, Alfa Basım Yayım, 2017, s. 91.
[26] Adam Zeman, **Bilinç Kullanım Kılavuzu**, çev. Gürol Koca, 3. bs., İstanbul, Metis Yayınları, 2017, s. 389.
[*] Esasen diğer zihinler problemi bakımından elbette ki bir solipsist yalnızca kendi varlığından emin olduğunu söyleyebilir ancak burada herkesin kendi deneyiminden emin olduğunu sağduyuya dayalı olarak söylemekte bir sakınca yoktur. Böylece felsefenin ve felsefeciliğin gündelikle, olup bitenle olan ilişkisini, herkesçe açık olunanlar hakkındaki konuların soru haline getirilmesini sağduyuya aykırı olduğunu söylemiyor muydu Aristoteles? Bkz.: Aristoteles, **Organon V -Topikler-**, çev. Prof. H. Ragıp Atademir, C. V, İstanbul, Milli Eğitim Basımevi, 1996, I, 10.
[27] Daniel C. Dennett, "Ben Neredeyim?", **Aklın Gözü: Benlik ve Ruh üzerine Hayaller ve Düşünceler**, ed. Daniel C. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 211.
[**] **Intuition Pumps** olarak anılan sezgi pompaları, Dennett'a göre, bir düşünürün bir problemle ilgili yanıt üretebilmesi için sezgilerini uyaran, sezgileri aracılığı ile bir yanıt üretebilmesine yardımcı olan düşünce deneyleridir.

işlevselcilik, zihnin doğasına ilişkin bir kuramdır. Metafiziksel işlevselciler zihin durumlarının davranışa ne şekilde yol açtığı ile ilgilenmemektedirler.* Onların davranışçılık ve fizikalizm ile ortak olan kaygıları bir noktada birleşir ve örneğin ağrının ne olduğunu ve tüm ağrı hissedilen durumlarda ortak olanın ne olduğunu soruşturur. Böylece bir ağrı hissinin veya bir zevk hissinin *ne* olduğu yanıtlanabilirse ve her zevkte ve ağrıda *ortak olanların ne olduğu* tespit edilebilirse böylece bu koşulları yerine getiren materyal ne türden olursa olsun, uygun şekilde organize olmuş o materyal ağrı veya zevk hissine sahip olabilecektir.[28] Artık fizikötesi alanlar ve var olma tarzları ile ilgilenilmemektedir. Düşüncenin cereyan ettiği, düşünmenin gerçekleştiği zihin artık bir modelleme biçimi sorununa dönüşmektedir.[29] Böylece bilinç, nitelceler, özgür irade gibi pürüzleri olan insan zihni, fizikalist düşünce deneyleri ve doğa bilimlerindeki paradigma değişimleri tarafından elenmekte, modelleme biçiminin bir nesnesi haline gelen insan zihni böylece sayısallaştırılıp silikonda var edilmekte, bu yolla da doğa bilimlerinin adım atamayacağı söylenen yerde "insan" kavramını yok edercesine bir adım atılmış olunmaktadır. Bir Pirus zaferine dönen bu arzudan hareketle, sterile edilmiş bir metafizik kabul kendisini kontrol edilemez bir yapay zekâ kaygısına bırakmakta, tasarım ürünü olan yapay zekâlar ve makinelerin insanların sınırlılıklarını aşabilecekleri ile ilgili kuruntular dillendirilmektedir.[30] Hatta bu kuruntular, bir süre sonra doğduğu tartışmaları unutup, yapay zekâlı varlıkların hangi koşullarda insansı olabileceği üzerine spekülatif tartışmalara doğru evrilmektedir. Böylece kökeni bakımından olanağı tartışmaya dahi tenezzül edilmeyen yapay zekâlı varlıkların hangi yetilere sahip olması gerektiği gündemimizde yer almaya başlamaktadır. Âdem ve Havva ile başlatılan yaratım süreci abiyogeneze indirgenmekte ancak abiyogenezden yapay zekâlı varlıklara kadar kesintisiz bir şekilde devam ediyor diye düşünülen, sterile bir inancın şahidi olmaktayız. Tegmark gibi isimler, bakterileri "Yaşam 1.0" olarak

* Nitelceler olarak örneğin kırmızıyı görüyor oluşumuz veya ağrı hissine sahip olduğumuz gibi durumlar metafiziksel işlevselcilerin gündeminde değildir.

[28] Ned Block, "What is Functionalism?", **Consciousness, Function, and Representation: Collected Papers**, C. I, Cambridge-Massachusetts, The MIT Press, 2007, s. 28.

[29] Aziz F. Zambak, "Düşünen Makineler ve Yeni Bir Metafizik Tasarım Olarak Yapay Zekâ", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 383.

[30] Norbert Wiener, "Some Moral and Technical Consequences of Automation", **American Association for the Advancement of Science**, C. CXXXI, No: 3410, 1960, s. 1355.

adlandırmakta, yazılım ve donanımlarının tasarlanmaktan çok evrimleşmiş olduğunu dile getirmektedir. Tegmark insanları ise "Yaşam 2.0" olarak adlandırmakta, donanımlarımızın* evrimleşmiş olmasına rağmen yazılımlarımızın büyük ölçüde duyularımızdan edindiğimiz verileri işleyip bilgi haline getirmemiz ve böylece ne yapacağımıza karar verdiğimiz bir süreç neticesinde tasarlanmakta olduğunu iddia etmektedir. Tegmark bir kademe daha yükselerek, yaşamın son bir güncelleme alması gerektiğini, yalnızca yazılımını değil, donanımını da tasarlayabilecek bir "Yaşam 3.0"ın gerçekleştirilmesi gerekliliğini vurgulamaktadır.[31] İşte "Yaşam 3.0" ile karşılanmaya çalışılan gelişme *yapay zekâ projesidir*. Yapay zekâ projesi ile insanın Tanrı'ya karşı olan savaşımı tamamlanmış olacaktır. Böylece kendi kendisini yaratan ve yaratmayı da elde edebilen bir varlığın arzusuna kavuşulmuş olacaktır. Bu metafizik yaratım öylesine güçlüdür ki, kimileri Nagel'ın düşüncesine de göndermelerde bulunarak makine olmak gibi bir şeyden bahsedilebilmesi için makinelerin dünyada sosyal bir varlık olması, dil yoluyla etkileşmesi ve ben/öteki ayrımını yapabilecek bilişsel yetilere sahip olması gerektiğini söyleyecek[32], bazıları ise daha da ileri giderek dil becerisi kazanan yapay zekâların resim yapabileceğini, şiir yazabileceğini iddia edecek[33], böylece insan eylem ve edimlerine denk yapay varlıklar üreteceğimizin göstergelerini bilgisayar programlarının o esnada yapıyor olduklarının ayırdında olup olmadığını sormadan – Searle ve Dreyfus'un itirazlarına yanıtlar getirmeden – matematiğin ve mantığın, bilgisayar bilimlerinin ve fiziğin varlık ve var olanlarla ilişkimiz bakımından taşıdığı sorunlara temas etmekten imtina edercesine, belki de tenezzül etmeyerek iddia edebilecektir.

* Okuyucu açısından da bilinmelidir ki tematik olarak çoğu zaman "donanım" kelimesiyle insan bedeni ve onun biyolojik imkânları kastedilirken, "yazılım" kelimesi ile insanın bilişsel imkânları karşılanmaya çalışılır.

[31] Max Tegmark, **Yaşam 3.0**, çev. Ekin Can Göksoy, İstanbul, Pegasus Yayınları, 2019, s. 45-47.

[32] İbrahim Ethem Deveci v.d., "Makinenin Beni ve Ötekisi", **Pasajlar**, No: 9, 2021, s. 104.

[33] Erhan Demircioğlu, "Mehmet Hilmi Demir: 'Aklımızın Sınırlarını Deniyoruz Aslında, Nereye Kadar Gidebiliriz?'", **Pasajlar**, No: 9, 2021, s. 139.

Bu arzunun başarılabilmesi için türlü yöntemler önerilir. Öğrenmenin temel alındığı* McCarthy tarzı çözümler, John Locke'un **tabula rasa** görüşüne dayanan robot bir bebeğin dil üzerinden dünyayı öğreneceği ve bu yolla da insansılaşacağını iddia etmektedir. McCarthy nesnelerin ve nesneler arası ilişkilerin ise mantık nesneleri, mantıksal ilişkiler türünden temsil edileceğini vurgular.[34] Waltz gibi isimler ise metafiziksel işlevselciliğin en yalın örneklerinden birini sunmakta ve insanlar kadar zeki bir makinenin inşa edilmesinin şu üç problemin aşılmasına bağlı olduğunu savunmaktadır: Bir biliş biliminin inşa edilmesi, inşa edilecek sistem için yazılım ve sensörleri vs. bakımından mühendisliğin uygulanması, uygun bir donanımın bulunması.[35] Ancak tam da bu sebepledir ki etik ve politik zeminlerde karşımıza çıkan yapay zekâ ile ilintili problemler, çözümleri yine yalnızca etik ve politik zeminlere aitmiş gibi düşünülür. Oysa köken bakımından sorun buralardan doğmamaktadır. Problem, insanın doğaya hükmetme istencinin metafizik kabullerle buluştuğu yapay zekâ projesinin kendisidir. Böylece insansı bir yapay zekânın olanağına ilişkin ileri sürülen reçeteler olsa olsa böyle bir mitin gerçekleşmesine duyulan heyecanın sonucundan başka bir şey değildir. Öyle ki bu arzu, kendine etik ve politik sahada birçok problem yaratmış olur. Yapay zekâ etiği denen ve politik tartışmaları da içeren bu alt alanın kendisi, yine etik ve politik sahada çözümlenmeye çalışılır. Oysa bu çalışmanın da iddialarından biri, etik ve politik bir sorun olarak yapay zekâ ile ilgili sorunların, bu alanların kendilerinde sonuçlandırılamayacağıdır. Yapay zekânın, var olmasının koşulları bakımından ele alınması, konunun transandantal olarak incelenmesi zaruridir. Ancak ve ancak bu yolla bu tartışmalar bir yere evrilebilir. Oysa günümüzde gerek irade özgürlüğü bakımından yapay zekâlı

---

* McCarthy, yapay zekâ alanında çalışan çoğunluğun nihai hedefinin, insanların edindikleri tecrübelerden hareketle gösterdikleri öğrenme süreçleri gibi süreçleri makinelerde gerçekleştirebilmek ve bu yolla da makinelerin insan düzeyinde eylemesini sağlamak olduğunu aktarır. Bunun aşamalarını ise tüm davranışların sistem içerisinde temsil edilebilmesi, davranıştaki değişikliklerin basitçe ifade edilebilmesi, davranışın her özelliğinin geliştirilebilir olması, makine açısından başarmanın ve başarının tanımlanabilir olması ve sistemin eylemlerine ilişkin alt rutinlere sahip olması üzerinden açıklar. Bkz.: J. McCarthy, "Programs with Common Sense", Cambridge, **RLE and MIT Computation Center**, 1960, s. 78-79.

[34] John McCarthy, "The well-designed child", **Artificial Intelligence**, No: 172, 2008.

[35] David L. Waltz, "The Prospects for Building Truly Intelligent Machines", **Dædalus**, C. CXVII, No: 1, 1988, s. 191.

varlıkların cezai ya da hukuki sorumluluğunun olup olmadığına ilişkin[36] hukukta ya da robotların işleyebilecekleri/neden olabilecekleri suçlar bakımından insanların onların özgür iradeli varlıklar olup olmadıklarına ilişkin hislerinin kökenini soruşturan[37] psikoloji gibi alanlarda bu çözümsüzlük daha net hissedilir durumdadır. Yapay zekâ felsefesindeki tartışmaların felsefenin alanında, felsefi ciddiyetle ele alınması gereken konularda mühendislerce ve bilim insanlarınca gayriciddi şekilde sürdürülmeye çalışması bu çabaların yansıları bakımından kendilerini etik ve politik çabaların nesneleri olarak bulacaktır. Oysa biz, felsefenin varlığı kendi olarak ele alan, onu aydınlatma çabalarını başka alanların veya bir amacın yüzü suyu hürmetine değil, yine kendisi olarak kendisi için gerçekleştiren ve spekülatif/retorik tavırlara karşı olan titizliğini kendimize ilke ediniyoruz.

Daha net anlaşılacaktır ki, yapay zekâ felsefesindeki sorunların, yapay zekâlı varlıklara karşı duyduğumuz antropomorfik bir hissin, varlıklara insani özellikler atfetmeden duyduğumuz hazzın bir yansısı olması pek de uzak bir ihtimal değildir.[38] Ancak daha derinlemesine bakılırsa görülebilir ki, böylesine dikkatsizce sürdürülen bir felsefe yapma tarzının ve felsefeci olmayanlarca yürütülen spekülasyonların sonuçları bakımından yapay zekâ devrimine ilişkin çekinceler yalnızca onun yaşamımızda bir yenilik olması ve bu yenilikle yüzleşiyor olmamız değil, tam da bu devrimin kendisi yoluyla insan üzerinde kurulmak istenen hakimiyetin, hakikatle insan arasındaki binlerce yıllık bağın artık hakikat ve sayısallaştırılmış zekâlı herhangi türden bir varlıkla bağdaştırılmak istenmesi, belki de biraz daha ileri giderek, Baker'in de değindiği gibi, "insan" kavramının yanına değil ama kendisinin çeperlerini kırarak, "*insan" kavramını yok ederek* ona hükmetmeye çalışan bir iktidar ve güç arzusunun ussallaştığı bir devrimdir[39] yapay zekâ devrimi...

[36] Ahmet Haşim Alagüney, "Hukukta İrade Özgürlüğü Kavramının Yapay Zekâ Türleri Bakımından Tartışılması", **Pasajlar**, No: 9, 2021, s. 201.
[37] Eddy Nahmias v.d., "When Do Robots Have Free Will? Exploring the Relationships between (Attributions of) Consciousness and Free Will", **Free Will, Causality, and Neuroscience**, ed. Bernard Feltz, Marcus Missal ve Andrew Sims, Leiden, Brill Rodopi, 2020, s. 57.
[38] Mustafa Arslantunalı, "Yapay Zekâ ve İnsanbiçimcilik: 'Bacakları Olan Bir Araba'", **Pasajlar**, No: 9, 2021, s. 229.
[39] Ulus Baker, "Homo Homunculus: Kendi Kendini Yaratamayan İnsan", **Pasajlar**, No: 9, 2021, s. 225.

1.1.1. Yapay Zekâ Felsefesindeki Antropomorfik Yükler

"Domin: Böylece genç Rossum kendi kendine dedi ki:
'Bir insan, eğlenen, keman çalan, yürüyüşe çıkmayı seven,
gerçekten gereksiz pek çok şeyi yapan bir varlıktır.'"[40]

Felsefi işlevselcilik ve bilimler yoluyla yürütülen yapay zekâ projesi, "insan" kavramını metafizik bir yaratım uğruna yıkmakla kalmaz, yıktığı bu kavramın yerine yeni bir metafizik tasarım olarak yapay zekâlı varlığı inşa eder. Yalnızca bir **logos** olarak değil, eşzamanda bir *Dünya varlığı* olarak düşünülür yapay zekâlar. Böylece orada eyleyen, çevresini değiştiren, algılayan, düşünen bir varlık olarak yapay zekâ, doğallaştırılır ve daha önceleri taksonomik bir incelemenin konusu olan karbon bazlı hayatın mirasçısı kılınmaya çalışılır. İnsanın özgünlüğüne karşı yaratılan bu tehdit, zeki makinelerin potansiyel olarak düşünülmesinin tartışmaya açılması yoluyla ve zihinlerimizin ve inorganik faillerin olanaklılığın sorulması türünden bir yolla da insanın fail olmak bakımından biricikliğine karşı üzerinde durulması gereken bir hususu meydana getirir.[41] Öyle ki, bu tehditten hareketle, duvarda asılı duran ve tek işi oda sıcaklığını ölçmek olan termometrelerin bile *bir kanıya sahip olduğuna*, bununla da kalmayıp cep telefonlarının veya hesap makinelerinin bile *bir düşünceye sahip olduğuna* inanmamız beklenir.[42] Oysa okuyucu açısından düşünülmelidir ki şeylere tinler atfetme arzumuzun kökeni nedir? Düşüncelerimizin insan doğasına yöneldiği hangi anda bundan vazgeçtik ve kavramlarını sıkıca temellendirmediğimiz bir felsefe yapma tarzı ile yıkmayı arzuladığımız "insan" kavramının yerine inorganik insanlar yaratma arzusunu hangi anda taşır olduk? Felsefi işlevselci tavır bu soruların üstesinden gelemez. Çünkü tavrın ve problemin kökeni bakımından o artık kendi köküne yabancılaşmıştır. Onun gündemi artık yapay zekâlı varlıkların **polis**'teki yeridir. Dolayısıyla o, mümkün mertebe insanın **polis** yaşamı içerisindeki *Dünyalılaşmasının* reçetelerini aramakla meşgul olacaktır. Bu düşüncenin en açık örnekleri, felsefi işlevselci metinlerin kendi içerisinde de oksimoron düşünceler

[40] Karel Capek, **R.U.R. – Rossum'un Uluslararası Robotları**, çev. Bilge Kösebalaban, 2. bs, İstanbul, İthaki Yayınları, 2021, s. 30.
[41] Jerry Kaplan, **Artificial Intelligence: What Everyone Needs to Know**, New York, Oxford University Press, 2016, s. 67.
[42] John R. Searle, "Minds, brains, and programs", **The Behavioral and Brain Sciences**, No: 3, 1980, s. 420.

barındırmasıyla da açıktır. Örneğin Tegmark, yaşamı bir yandan yazılım ve donanım katmanları bakımından ele almayı tercih ederken, bir yandan da günümüzde ve bu çağda insan olmanın, bizi diğer yaşam formları ve makinelerden ayıranın ne olduğunu sormaktadır.[43] Bununla da kalmayıp yapay zekâların gerçek dünyada kullanıyor olduğumuz kelimelerin anlamlarının nelere tekabül ettiğini bilmeden, kelime ilişkilerini ve örüntülerini keşfettiğini söylemekte, böylece yapay zekâların erkek veya kadın olmanın ne anlama geldiğini ya da uzay, zaman ve maddeden oluşan fiziksel bir gerçekliğin bulunduğuna ilişkin en ufak fikrinin bile olmadığını açıkça öne sürmektedir.[44] Ancak fikrini yeniden değiştirerek zeki olarak atfedebileceğimiz makinelerin bir tür hakka sahip olmaları söz konusu olacaksa bunun çözümünün ancak ve ancak bu makinelerin bilinçli olup olmadığının ve acı ile mutluluk gibi hisleri yaşayıp yaşayamadıklarının tespiti yoluyla mümkün olabileceğini iddia etmektedir.[45] Görüleceği gibi, yapay zekâ felsefesindeki antropomorfik yüklerin, köken bakımından yapay zekâ alanını besleyen mühendislik ve bilim kollarındaki kafa karışıklıklarından doğduğu, "insan" kavramına karşı olan yabancılaşmanın sonucu bakımından ise inorganik bir varlığı insansılaştırmaya doğru döndüğü bir duruma şahitlik etmekteyiz. Öyle ki, her ne kadar Descartes'ın felsefesi, insan zihnini ve bu zihne ilişkin yetileri bilim dışı bir alana taşımayı denemiş olsa da insan davranışlarının biyolojik temellere indirgenmesi girişimlerinin tarihi bakımından Darwinci girişimleri veya devamı olarak çağdaş fizyoloji ve psikoloji çalışmalarını görmek mümkündür.[46]

İnsan doğasındaki bilinmezler, felsefe tarihi boyunca tartışıldığı ve bu tartışmaların zamansal olarak bakıldığında ezici bir kısmı, çözümü fizik üstü alan ve yöntemlerle elde etmeye çalışmışsa da bilimin insan yaşamında hiç olmadığı kadar güçlendiği son birkaç yüzyılda ibre yeniden yeryüzüne, fiziksel olana dönmüş durumdadır. Ve insan doğasındaki bu bilinmezlere karşı olan savaşımında bilim, yapay zekâ projesi vasıtasıyla insan doğasında yer alan bu bilinmezleri her ne kadar

[43] Tegmark, **Yaşam 3.0**, s. 113.
[44] **A.e.**, s. 124.
[45] **A.e.**, s. 362.
[46] Harold J. Morowitz, "Zihni Yeniden Keşfetmek", **Aklın Gözü: Benlik ve Ruh Üzerine Hayaller ve Düşünceler**, ed. Daniel C. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 42.

gerektiğinde elemeyi, gerektiğinde ise sadece ona özgü olmadığını söyleyerek değersizleştirmeyi, onu incelemekten ve açığa çıkarmaktan ziyade yıkmayı tercih etmişse de yine ondan kurtulamamıştır. Yapay zekâ projesi, doğası gereği insanı örnek alır, insansıdır ve insansı unsurlar içermek zorundadır. Çünkü McCarthy'nin de söylediği gibi yapay zekâ çalışmaları insan zekâsını her zaman eşit düzeyde değil ama çoğu durumda simüle etmek üzerine kurulmuştur. Bu çaba, bir yandan insanların dünya problemlerini nasıl çözdüklerini inceleyerek mesafe katetmeye çalışmakta, bir yandan da insansı ve hayvansı problem çözme çabalarını bir kenara bırakarak matematiksel mantık üzerinden steril bir "zekâ" kavramını hedeflemektedir.[47] Böylece aslında bir yandan evrimsel süreç sonunda şekillenmiş ancak bir mühendisin gözünden oldukça kötü ve kusurlu tasarlanmış olan insan zihni ve bedeni yeniden düşünülmekte, bir yandan da bu düşünme kendisini insan doğasındaki unsurlardan alıkoyamıyor gibi görünmektedir. Nitekim Wang, yapay zekâ çalışmalarının ayırıcı bir özelliğinin ise herhangi bir beyne ihtiyaç duyulmadan bir zihnin elde edilmeye çalışılması olduğunu söylemekte ve bu yolla zihnin herhangi bir aracı olmadan tanımlanma çabasının bulunduğunu iddia etmektedir.[48] Böylece yeryüzünde failliği insanın elinden almak isteyen yapay zekâ projesi kendine biyolojik kaygılar da edinmiş olur. **Artificial life, ALife** veya Türkçe "yapay yaşam" olarak ifade edebileceğimiz alan, öz organizasyon, çoğalma, adaptasyon ve evrim gibi biyolojik süreç ve işlevlerin bilgisayar simülasyonları üzerinden anlaşılmaya çalışıldığı bir alan olduğu gibi[49] bir yandan da yapay zekâ projesini bir alt alan olarak içeren ve yaşamın kendisi hakkındaki işlevselci yaklaşımların doğru olup olmadığı ya da simüle edilmiş bir yaşamın ne zaman canlı olabileceği yollu kimi felsefi sorulara da yanıtlar arayan bir alandır.[50]

Bedenin kopya edilmesi, bedeninin evrimsel geçmişinin inorganikleştirilmesi ve sahiplenilmesi yoluyla "insan" kavramının maddi yönünün üstesinden gelinmesi,

---

[47] John McCarthy, "What is Artificial Intelligence?", (Çevrimiçi) http://www-formal.stanford.edu/jmc/whatisai.pdf, 25 Aralık 2022.

[48] Pei Wang, "The Logic of Intelligence", **Artificial General Intelligence**, ed. Ben Goertzel ve Cassio Pennachin, C. II, New York, Springer, 2007, s. 36.

[49] Simon Blackburn, "Artificial life", **The Oxford Dictionary of Philosophy**, 2. bs., Oxford, Oxford University Press, 2005, s. 25.

[50] Robert Audi, "artificial life", **The Cambridge Dictionary of Philosophy**, 2. bs., New York, Cambridge University Press, 1999, s. 57-58.

kendisini bir sonraki aşamaya bırakır: İnsanın zihin süreçlerinin devralınması. "İnsan" kavramının bedensel ve zihinsel olarak yıkımıyla başlayan süreç kendisine modern bir **Frankenstein** yaratmaktadır. Eksik olan tek şey kalmıştır, o da yapay zekâ araştırmacılarının arzuladığı şey olarak insan gibi davranan, insansılaşan zihinler elde etmektir. Bu amaca ulaşmak kolay değildir. İnsanın düşüncesi psikoloji ve türlü bilimler yoluyla etüt edilmeli, bir kuram oluşturulmalı, daha sonrasında ise bu kurama dayanarak bilgisayar programları üretilmelidir. Bu programlar uygun girdi ve çıktı ilişkilerine uygun olarak çalıştırılır, böylece bir bilgisayarda insan zihninin ürünlerinin elde edileceği ümit edilir.[51] Bu ümit psikoloji ve fizyoloji gibi alanların insan zihni üzerine yapmış oldukları tanımlama ve tespit çalışmalarına sıkı bir şekilde dayanır. Çünkü Şengör'ün de dediği gibi, yapay zekâ çalışmaları canlıların ve özellikle insanların akıl yürütme, planlama, öğrenme, bellek, karar verme gibi bilişsel süreçlerine ilişkin elde edilen bilgilere ve bu bilgilerden türetilen önermelerden yola çıkarak akıl sahibi makineler ve sistemler kurgulamayı kendine amaç edinmektedir.[52] Makineler ve sistemler ile kastedilen şey bilgisayarlardır. İnsanların bir şeyleri nasıl anladığını anlayabildiğimiz ve ifade edebildiğimiz durumda bilgisayarların da formalize edilmiş bu anlama tarzına uygun inşa edilebileceğini söyleyen[53] Schank ve Abelson tarzı görüşler yaygındır. Ancak dikkat edilmelidir ki Tegmark'ın da vurguladığı noktalar Dennett açısından da gözden kaçmaz. Yapay zekâların bilgisayarların tıpkı bir failmiş gibi simüle edilmesini sağlayan programlar olduğunu kabul etsek bile onun dünya hakkındaki bilgisinin **nihil** olduğunu söylemek mümkündür. Çünkü Dennett açısından da bilgisayarın, ona bilgilerin yüklenmesi/sağlanması gerektiği bir **tabula rasa** olduğu, dolayısıyla yapay zekâyı programlayanlara bağlı bir durumun söz konusu olduğunu söyleyebiliriz.[54]

Şimdiye kadar gelinen noktada bilgisayarlara yapay zekâlar vasıtasıyla failsi niteliklerin kazandırılmaya çalışıldığını görmekteyiz. Ancak bilgisayar nedir?

---

51 H. Levent Akın, "Yapay Zekâda Vücut Beyin Problemi", (Çevrimiçi) https://www.cmpe.boun.edu.tr/~akin/papers/beyin.pdf, 19 Kasım 2022.

52 Neslihan Serap Şengör, "Bilinç, Farkında Olan…", **Özgür İrade / Hukuk, Nörobilim, Psikoloji ve Ötesi**, ed. Ozan Erözden, Salih Haydar Günler, Mustafa Ebrar Palteki ve Cem Uysal, İstanbul, Zoe Yayıncılık, 2021, s. 96.

53 Roger Schank ve Robert Abelson, **Scripts Plans Goals and Understanding: An Inquiry into Human Knowledge Structures**, New Jersey, Lawrance Erlbaum Associates, 1977, s. 8.

54 Daniel C. Dennett, "Cognitive Wheels: The Frame Problem of AI", (Çevrimiçi) http://cinc.uaem.mx/~bruno/material/dennet_93_cognitiveWheels.pdf, 14 Aralık 2022.

Bilgisayarlar, yapay zekâ projesinin arzuladığı amaçlar bakımından düşünüldüğünde bu amaçları taşıyabilecek nitelikte araçlar mıdır? Parçaları ve o parçaların bir bütünü olması bakımından bilgisayarın sınırlılıkları nedir? Nitekim yalnızca bizim monitör ve diğer birimleri barındıran kasa gibi ayrık bir formda değil ama dizüstü bilgisayar, cep telefonu vb. gibi birçok formda karşılaştığımız bir sistem olarak bilgisayarların, işaret ettikleri bakımından var olma tarzı, erimi ve gücü gibi birçok hususla ele alınmalıdır. Ancak biz bilgisayarların başlangıçta yalnızca birer hesap makinesi olduklarını biliyoruz. Bilgisayarların toplama, çıkarma, çarpma ve bölme işlemleri üzerindeki kabiliyetleriyle de hesaplama teknolojisinin bilgiyi sayılarla temsil etmekten gelen gücünü pekiştirdikleri açıktır. İşte tam da böyle bir güç, kendisini renklerin, tonların, müziğin, görsellerin ve daha nice sanallaşmış gerçekliğin temsilini **bit'**ler vasıtasıyla sağlayabildiğini iddia eder, böylece sayısallaştırılmış bir gerçeklik, masaüstünü gösteren bir ekranda dosyalara ve ikonlara dönüştürür.[55] Makinelerin gerçekliği sayısallaştırılmış bir tarzda bize yeniden sunabileceği düşüncesi bir kez elde edilince, makinelerin hangi koşulları yerine getirdiklerinde "İşte, düşünüyor!" diyebileceğimiz durumları sağladıklarını aramaya başlamamız oldukça doğaldır. Nitekim böyle de olmuştur! İnsan zekâsını kendine model alan yapay zekâ çalışmaları, bu zekâyı gerçekleştirmenin türlü yollarını denemektedir. Örneğin sonuçları bakımından düşünüldüğünde bir insanın eylemesi ve eyleminin neticesinde bir sonuca ulaşması, benzer sonuç bir yapay zekâ tarafından da elde edilebiliyorsa, yapay zekâya "zeki" dememiz için yeterli olurken, süreç bakımından düşünüldüğünde ise sonuç ne olursa olsun yapay zekâ ile insan zekâsı, zekânın çalışma ilke ve aşamaları bakımından benzerlik taşıyorsa yapay zekâlara "zeki" dememiz için yeterlidir.[56] Buradan yeni tartışmalar doğar. Eylemde bulunan bir varlığın fail olup olmadığı nasıl anlaşılır? Bir akıl yürütmenin aşamaları reçete edilebilir mi? Zekânın sınırları nasıl çizilir? Bir sınırı var mıdır? Bu ve benzeri soruları yönelten Russell'a göre rasyonel davranan yapay zekâların zeki olduğu yollu tavra ilişkin itirazlar çeşitlenebilir. Örneğin sayılabilirlik bakımından düşünülürse birlik olarak bir insan bedeni, bir hayvan bedeni veya bir böcek bedeni, failliğin onun

[55] Ethem Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, çev. Aylin Ağar, 2.bs, İstanbul, Tellekt, 2021, s. 18-19.
[56] Köse, **Yapay Zekâ Felsefesi**, s. 27.

sınırları içinde – örneğin bir bedende bir adet *ben* olur – olduğu bir durumu temsil edebiliyorsa, aileler, şirketler, kültürler veya devletler fail olabilir mi? Örneğin Russell, karıncalar gibi sosyal canlıları incelediğimizde tek bir karıncayı mı yoksa feromonlarla iletişim kuran tüm bir karınca kolonisini mi fail saymamız gerektiğini sorar, böylece beyin dediğimiz ve sayılabilirlik bakımından bir sayısıyla ifade ettiğimiz bir bütünü, karar alma bakımından çok sayıdaki küçük karar birimlerine yaymamızın ve bunu uzamsal olarak da birlik altında düşünmemizin türlü olasılıklarını dile getirir.[57] Manzotti ve Tagliasco ise, daha radikal bir şekilde, düşünen makinelerin insan yapması şeyler olduğunu da kabul ederek, bu makinelerin bir kez çalıştırılmaları durumunda, makinelerin sahip oldukları hallerin onların düşünüyor olduklarını gösteren durumlar olduğunu, bu düşünmenin biz insanların düşünmesi gibi bir düşünme de olmak zorunda olmadığını, düşünmenin farklı tarzları olabileceğini iddia eder.[58]

İnsan ve yapay zekâ arasındaki ilişkide bilgisayarlara yüklenilen rol, bir sonraki aşamada göründüğü üzere bilgisayarlara insansı niteliklerin kazandırılması ile pekiştirilmek istenmektedir. Bu rol yalnızca mühendislikte ve bilimde değil, aynı zamanda dilde ve söylemde de pekiştirilmek istenmektedir. Onların düşündüğü, özgür seçimler yaptıkları, duygu durumlarına sahip oldukları ve benlik sahibi olduklarının kurgu düzeyinden kuramsal bir ihtimale dönüştürüldüğü bir durumla karşı karşıyayız. Böylece seçimler yapabilen yapay zekâ sistemlerinin, onları kodlayan programcıların sınırlılıklarını da aşarak hedeflerini ve kendilerini güncelleyebilen – burada öğrenme vasıtasıyla insanların yaşam boyu kendilerini değiştirmelerine ilişkin bir olanağa da dayanarak – varlıklar olacakları kurgusu ön plana çıkmaktadır.[59] Önceleri çalışan çarkların ve araçların – burada ise konumuz gereği bilgisayarlar – insanın bir özelliğine benzetilmeye çalışıldığı durumdan şimdi ise insansı özelliklerde insanı aşan bir söyleme doğru evrilmekteyiz. Örneğin Moravec, hayvanların ve robotların öğrenmeleri arasında bir fark bulunduğunu,

[57] Stuart Russell, **İnsanlık için Yapay Zekâ: Yapay Zekâ ve Kontrol Problemi**, çev. Barış Satılmış, Ankara, Buzdağı Yayınevi, 2021, s. 41.
[58] Riccardo Manzotti ve Vincenzo Tagliasco, "Artificial consciousness: A discipline between technological and theoretical obstacles", **Artificial Intelligence in Medicine**, No: 44, 2008, s. 107.
[59] Mehtap Doğan, "Yapay Zekâ ve Özgür İrade: Yapay Özgür İradenin İmkânı", **TRT Akademi**, C. VI, No: 13, 2021, s. 791.

hayvanların tekil şekilde öğrenirken robotların öğrenmesinin bir makineden diğerine kopyalanabilecek bir tarzda olduğunu söyler[60], Işıklı ise doğal zekâlı* varlıklar olarak insanların bilgiyi kolayca unutabilirken, yapay zekâlı bilgisayarların kalıcı hafızaya sahip olduğunu, böylece bu bilgilerin kaybolmasının ve unutulmasının söz konusu olabileceği bir durumun olmadığını, bununla da kalmayarak bilgileri depolama yöntemlerimiz arasında da farklar bulunduğunu, bilgisayarların matematiksel yöntemleri kullanarak bilgileri 0'lar ve 1'ler ile işleyip kaydettiğini ancak biz insanların ise sesler, görüntüler vb. gibi duyu verilerini birbiri üzerine binmiş şekilde kaydettiğimizi aktarır.[61] Ancak burada daha ciddi bir sorun vardır: Arf bu sorunu iyi analiz eder ve bir makinenin düşünmesi ile bir makineye onun düşünüyor olduğunun *yüklenmesi* arasındaki farka dikkat çeker ve makinelerin düşünmediğini savunur. Çünkü örneğin düşüncenin elle tutulabilen ve gözle görülebilen meydana gelişine davranış diyeceksek bir insanın kendisine söylenen sözlere veya maruz kaldığı etkilere değişik davranışlarla yanıtlar üretmesi doğaldır. Ancak okuyucu şunu sorabilir: Makineler de böyle değil midir? Arf bu itiraza yanıt verir: Bir makine, örneğin bir saat, bizi saat kaçta uyandırmasını istediğimizi *ona kendi diliyle ifade ettiğimiz zaman* çoğu durumda bunu gerçekleştirir. Çarkları çevirir, bizi uyandırmasını istediğimiz saate onu ayarlarız ve ilgili saat geldiğinde alarm öter. Böylece makine bize kendi dilinde "Saat geldi, uyan artık!" der. Ancak Arf şunu ekler: Peki ama bunu biz bir bekçiye de yaptıramaz mıyız? Evvela bir bekçi de bizi uyandırmasını istediğimiz bir saatte gelip kapımızı çalabilir, uyanmazsak bizi dürtebilir ve uyandırabilir. Ancak Arf'a göre bir saat *bunu düşünemez*. Saat – veya herhangi bir makine – biz onu nasıl tasarlarsak ve ona ne gibi işlevler kazandırırsak ancak o tasarım ve işlevin olanaklılıklarını yerine getirebilir.[62]

Arf'ın makinelerin düşünmediği ve düşünemeyeceği yollu argümanının ortaya koyduğu iddialar bir kenara Penrose, Arf'tan farklı olarak, makinelerin

[60] Hans Moravec, "When will computer hardware match the human brain?", **Journal of Evolution and Technology**, C. I, No: 1, 1998, s. 12.

* Çalışmanın ikinci bölümünde zekâ kavramına getirilen "yapay" ve "doğal" sıfatlarını detaylıca ele aldık.

[61] Şevki Işıklı, "Fuzzy Mantık ve Sibernetik'in Siber Toplum ve Yapay Zekâ Üzerindeki Etkileri", Erzurum, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2004, s. 117.

[62] Cahit Arf, "Makine Düşünebilir mi ve Nasıl Düşünebilir?", **Atatürk Üniversitesi 1958-1959 Öğretim Yılı Halk Konferansları**, No: 1, s. 94-95.

davranışları üzerine yaptığımız antropomorfik yorumlamaların bir ciddiyete sahip olmadığını düşünür, onların ancak ve ancak makinelerin davranışlarını anlamamıza yardımcı olmaları türünden bir işlevi olduğunu savunur.[63] Yalnız unutulmamalıdır ki yapay zekâ felsefesindeki antropomorfik yükler yalnızca benzetim ve atfetmelerin tahliyesi yoluyla çözülemez. Bu sorun aynı zamanda bir dil sorunudur. Makinelerin yapıp etmelerini *davranış* olarak tanımlamakla, *hareket* olarak tanımlamak arasında bir fark bulunur. Eylemde bir *ben*, bir *kişi*, bir *fail*, bir *tin* bulunur. Böylece Penrose türü bir naiflikle olsa bile makinelerin *davranışlarını* anlamayı istemek, onda gizli bir *tin* görmeyi kabul etmek demeye gelir. Bunun en açık yansısı bilgisayar bilimcilerde görülür. Örneğin Say, hesaplama yapabilen her sisteme "bilgisayar" diyor oluşumuzun insanların da bilgisayar olduğu anlamına geldiğini, böylece bilgisayarların insanların söz ve davranışlar gibi bilişsel çıktılarını taklit etmeye muktedir olduklarını, birbirlerini taklit etmek konusunda insan ve makineler arasında birinin diğerine bir üstünlüğünün bulunmadığını, böylece bilgisayar olarak insan modelinde de insanın zihnindeki bilgi, fıkra, şarkı vb. gibi ürünlerin elektronik bilgisayarlardaki programlara denk geldiğini, bu programların ise makineye *can* katan unsurlardan biri olduğunu ve insanlarda da bunun böyle bir karşılığının bulunduğunu iddia eder.[64] Ancak Say unutur ki bilgisayarların zeki olmasını, bir hisse sahip olmasını ve örneğin bir benliğe iye olmasını istemekle, *bu durumları simüle etmesini istemek* ayrı şeylerdir.[65] Çünkü örneğin Lem'e göre, bir programcı, bir dijital makineyi adeta karşısında zeki bir insan varmış gibi konuşmak üzere programlayabilir. Gereksinim duyduğu her durumda bu makine "ben" zamirini kullanabilir. Ancak bu bir kandırmacadır! Bu makine en aptal insan kadar değil ama en fazla çok iyi eğitilmiş bir milyar papağan kadar başarılı olabilir. Dolayısıyla insan davranışını konuşma düzeyinde taklit etmekten daha fazlası söz konusu olamaz. Makinede bir *ben* bulunmaz. Dolayısıyla onu hiçbir şey eğlendiremeyecek, şaşırtmayacak, paniğe sokmayacak veya üzmeyecektir. O yalnızca ona girilen verileri seslendiren bir *ses* olacaktır. En iyi satranç oyuncularını yenen bir mantık olacaktır. İnsanların hiç anlayamayacakları kadar sonsuz bir umursamazlık içerisinde

[63] Roger Penrose, **Kralın Yeni Usu**, çev. Tekin Dereli, C. I, 10. bs., Ankara, TÜBİTAK, 2000, s. 17.
[64] Cem Say, **50 Soruda Yapay Zekâ**, 15. bs., İstanbul, 7 Renk, 2020, s. 44-46.
[65] Patrick Henry Winston, **Artificial Intelligence**, 3. bs, Reading-Massachusetts, Addison-Wesley Publishing Company, 1993, s. 8.

olacaktır. Ondan bir şeylere sempati veya antipati duyması beklenemeyecektir. Dolayısıyla o, her şeyi kusursuzca taklit eden bir nesne, bir *aktör* olacaktır.[66] Böylece Baker'in sözlerini daha iyi anlayabilir hale gelmekteyiz. O, insan biçimli bir yapay zekânın bulunmadığını söylemektedir. Yapay bir "ruh"un üretilemediğini, bilgisayarların tüm hatasızlıklarına karşın insana erişemediğini, program düşüncesinin yürümediğini dile getirmektedir. Ve eklemektedir: "… içinde bir cücenin saklandığı, 18. Yüzyılın şu sarıklı Satranç Oynayan Türk otomatı, yapımcısının ya da içindeki cücenin, bu oyunu ağzı açık izleyenleri kandırmasını mı, yoksa zihni ile eli arasındaki bağlantıyı, oyun için de olsa koparan cücenin kendi kendini kandırmasını mı simgelemektedir?"[67]

1.1.2. Yapay Zekâ Felsefesinde Fizikalist Savların Yeri

"Yaşlı Adam: Farenin aklıyla insanınkinin aynı makineler olduğunu düşünüyorum. Yalnızca farklı kapasitelere sahipler."[68]

Yapay zekâ projesinin, bilimler yoluyla gerçekleştirilmek istenen bir proje olduğunu bilmekteyiz. Ve bu projeyi, bir zamanların penisilini, buharlı makinesi veya barutunun insanın yaşamında ne gibi gelişmeleri sağladığının ya da ne gibi değişimleri dayattığının tespitini yapmak gibi bir yolla ele almak bizce oldukça hatalı bir ele alma olacaktır. Gerekçemiz açıktır: Yapay zekâ projesi, insanın yeryüzündeki statüsünü ve "insan" kavramının kapsamını sınır tanımazcasına yıkmaktadır. Ve daha önceki bölümlerde dile getirdiğimiz gibi, biz "insan" kavramını dokunulmaz bir kavram olarak ilan etmiyoruz, sadece bu kavramın yapım ve yıkım süreçlerinin yapay zekâ projesinde rol oynayan kimi bilim insanları ve felsefeciler tarafından sıkıca tartışılmadığını ve bunun olası sonuçları bakımından etik ve politik sorunların meydana geldiğini, dolayısıyla bu sorunlara yol açan yaklaşımlardan biri olarak felsefi işlevselciliğin tavır ve dayanakları bakımından kimi sorunlar aşılmadıkça yapay zekâ çalışmalarının meydana getirdiği ve getireceği olası problemlerin de

[66] Stanislaw Lem, "Non Serviam", **Aklın Gözü: Benlik ve Ruh Üzerine Hayaller ve Düşünceler**, ed. Daniel D. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 297.
[67] Baker, "Homo Homunculus: Kendi Kendini Yaramayan İnsan", s. 226.
[68] Mark Twain, **İnsan Nedir?**, çev. Esra Damla İpekçi, 11. bs., İstanbul, Dedalus, 2020, s. 96-97.

aşılamayacağını iddia ediyoruz. Felsefi işlevselciliği anlamanın, sorunlarını tahlil etmenin bir yolu da dayanakları üzerinde durmaktır. Çeşitli bilim alanları da bu dayanaklardan biridir. Biz bilimin kontrolsüz ve tutarsız bir alan olduğunu savunmuyoruz. Sadece gerçekliği inandırıcı bir şekilde yansıtan dünya modelleri inşa etme girişimi olarak bilimin, dünyada ne tür olguların bulunduğu, bunların nasıl düzenlendiği, birbirleriyle nasıl bir ilişki içerisinde olduğu ve nasıl işlediklerini anlama çabasında[69] yapay zekâ projesinin yeniden düşünülmesi gerektiğini vurguluyoruz. Çünkü Seymour Papert'a göre yakın bir zamana kadar insan zihnini anlama çabaları hayvanların zihinleri ile karşılaştırma yapmak veya Tanrı ya da mitolojik varlıklar üzerinden bir inceleme yapmakla sınırlıdır. Ancak şimdi ise bilgisayarlar vardır.[70] Böylece bir yönüyle insan zihnini anlama çabası olarak yapay zekâ projesi, meydana geldiği bilimler temelinde büyük bir öneme sahiptir. Çünkü bilimler düzeyinde sıkıca incelenmeyen ve spekülatif söylemlerle tartışılan kimi kavramların sosyal bilimler ve felsefe gibi alanlara da ihraç edildiği çok açıktır. Örneğin Demircioğlu, yapay zekâ disiplininin belirişini, evrenin en büyük gizemlerinden biri olan akılsallığın anlaşılmasında *soyut, felsefi ve spekülatif* tartışmaları, kapalı dil oyunlarını aşmanın ve uygulanabilirliğin, somut olan ile temasın, gözleme dayalı teori inşasının mümkün kılınması bakımından önemli kabul etmektedir. Ancak sorulabilir: Yapay zekâ projesi şimdiye kadar bize zihin hakkında önemli bir şey öğretmiş midir?[71] Yoksa o tam da somut ve uygulanabilir diye düşünülen ama metafizik tasarımlarımıza ilham olan, edebi arzularımızın ve gelecek rüyalarımızın bir sembolü müdür? İşte Demircioğlu'nun düşüncesinde de kendine yer bulan problem, bilimin materyalist/fizikalist[*] kabullerinde ve tam da var olmanın koşullarını kendisi olarak inceleyen felsefenin *spekülatif* ilan edilmesinde[72] yatar.[**]

[69] Antti Revonsuo, **Bilinç: Öznelliğin Bilimi**, çev. Selim Değirmenci, 2. bs., İstanbul, Küre Yayınları, 2017, s. 127.
[70] Heinz R. Pagels v.d., "Has Artificial Intelligence Research Illuminated Human Thinking?", **Annals of the New York Academy of Sciences**, No: 426, 1984, s. 140.
[71] Hilary Putnam, "Much Ado About Not Very Much", **Dædalus**, C. CXVII, No: 1, 1988, s. 269.
[*] Materyalizm ve fizikalizm birbirleri yerine kullanılabilir iki addır ve çalışmamızın buradan sonraki kısmında fizikalizm adını kullanacağız.
[72] Erhan Demircioğlu, "Yapay Zekâ: Açmaz ve İmkân", **Pasajlar**, No: 9, 2021, s. 11.

İnsan zihnine ve bedenine duyulan bu şüphe, yapay zekâ projesinin temel taşıdır. Tegmark, yapay zekâ araştırmacıları arasındaki genel görüşün zekânın en nihayetinde et, kan ya da karbon atomlarıyla değil bilgi ve hesaplamayla ilgili olduğunu söylemekte, temelde ise her şeyin etrafta dolaşan madde ve enerji olduğu mevcut durumda bilgi ve hesaplamanın ne olduğunu sormamakta, fiziksel varlığı bulunmayan bilgi ve hesaplamanın fiziksel varlıklarda nasıl vücut bulduğunu merak etmektedir.[73] Say ise, insanlar dahil evrendeki her şeyin fiziksel sistemler olduğunu, fiziksel sistemlerin makro düzeydeki davranışlarının onları oluşturan parçacıkların basit yasalarca yönetilen mikro düzeydeki davranışlarına bağlı olduğunu iddia etmekte, böylece bu mikro alemin yasalarının hesaplanabilir şeyler olduğu kabul edilirse hesap makinelerinin bile "ulvi" şeyler olduğunu söylemekten geri durmamaktadır.[74] Fizikalist temellere dayandırılan düalist arzular, kendilerini McCarthy'nin termometrelerinde veya Say'ın hesap makinelerinde bulur ve örneğin insan düzeyinde bir yapay zekânın epistemolojik yetkinliğe sahip olması gereken bir bilgisayar programına gereksindiği söylenir.[75] Böylece insansı olan tüm kavram ve yapılar mekanist bir tavrın gündemi olarak kabul edilir, bu yolla da beynin yaptıklarının benzerinin daha kesinlikli ve daha hızlı bir şekilde yapılabildiği bilgisayarların inşa edilebileceği savunulur. İnsanları satrançta, politikada veya aşk şiirleri yazmak konusunda geride bırakabilecek bilgisayar programlarının

[**] Öncelikle Demircioğlu'nun savında iki varsayım var: Bunlardan ilki, felsefenin bilimsiz, bilimden kopuk sürdürülen çabaların alanı olduğu yollu varsayım ve ikincisi ise felsefenin her istediğini söyleyebilenlerin alanı olduğu görüşüdür. Öncelikle, felsefe bilimden kopuk bir alan değildir. Felsefe tam da bilimsel paradigmalar yoluyla Thales'in ontolojisinden çağdaş ontolojilere gelebilmiştir. Felsefe tam da "İnsan nedir?" sorusunun renk ve cinsiyetle ilgisinin olmadığının anlaşıldığı bir yere antropoloji ile gelmiştir. Felsefe bilimsiz yürüyemez, bilimsiz felsefe Kant öncesi metafiziktir. Ancak felsefe bilimsiz yürüyemez demek felsefe bilimleşmemiş düşüncelerin alanıdır demek değildir. Bir bilim adamı, kendi alanı ne ise, o alanın temel varsayımlarının transandantal bir incelemesini sunamaz, sunmak için çabalamaz da. Sunmak için çabaladığı bir durumda felsefenin alanına geçmiş olur, örneğin biyoloji felsefesi yapar ve "Canlı nedir?" sorusunu sorar. Dolayısıyla Demircioğlu'nun ilk varsayımının kendisi tam da *spekülatif* olandır. İkinci varsayım ise konjonktüreldir. Felsefe ve felsefe akademisi hiç olmadığı kadar değersizleşmiş ve değersizleştirilmiştir. Ülkemizdeki felsefi hareketliliğin – veya Demircioğlu uluslararası felsefi hareketliliği referans veriyor olabilir – bilimleşmenin ve bilimsel paradigmanın etkisi altında kalması, bu hareketliliğin felsefeciler vasıtasıyla halka inmesi gerekliliği gibi bir kaygının bulunmuyor oluşu, belki de felsefe akademimizin hiç olmadığı kadar *spekülatif* ve *denetimden uzak* bir alan olmasına sebep olmuştur.

[73] Tegmark, **Yaşam 3.0**, s. 80.

[74] Say, **50 Soruda Yapay Zekâ**, s. 44.

[75] John McCarthy, "What has AI in Common with Philosophy?", (Çevrimiçi) https://www.cp.eng.chula.ac.th/~prabhas/teaching/cbs-it-seminar/2012/aiphil-mccarthy.pdf, 18 Kasım 2022.

yazılabileceği, bunun sadece bir sıkı çalışma meselesi olduğu iddia edilir.[76] Oysa Schwartz'a göre yapay zekâ alanındaki araştırmacıların kendi alanlarındaki bu ve benzeri söylemleri dayandırdıkları temel, fizikalist varsayımlardır. Bu varsayımlar, zihnin basitçe beyindeki bir bilgi işleme sürecinin adı olduğu, beynin biyokimyasal yasalara göre çalışan ve indirgenemez ruh veya diğer zihinsel bir fenomene, herhangi bir belirsiz varlığa dayanmayan fiziksel bir yapı olduğu varsayımlarıdır.[77]

Okuyucu şu soruları sorabilir: Fizikalizmden duyulan çekincenin sebebi nedir? Fizikalist savların yanlışlığı mı savunuluyor? Burada belirtmeliyiz ki amacımız fizikalizmin savlarını yanlışlamak değildir. Amacımız, fizikalizmin felsefi işlevselci tavır tarafından belli bir tarzda kullanılışına işaret etmektir. Fizikalizme yüklenilmeye çalışılan yükün, fizikalist savlara ilişkin problemler sebebiyle de taşınamaz olması, tam da bu sebeple bilgisayar bilimi ve doğa bilimleri tarafından ortaya konan yapay zekâ projesinin fizikalizme dayanan iddialarının sıkıca temellendirilmemiş olduğunu göstermek gayretindeyiz. Çünkü Descartes'ın düalizmi her ne kadar insan ruhu idesinin, uzunluk, genişlik ve derinlikçe yayılımı olmayan ve fizik nesnelere ait niteliklerin hiçbirinden pay almayan bir yapıda olduğunu[78] ve bizi biz yapan şey olarak bu ruhun bedensiz de var ve mevcut olabileceğini söylediyse[79] de fizikalizm bu savların yanlışlığını ortaya koymuştur. Örneğin bilincimiz gibi kimi fenomenal unsurların fiziksel olmayan bir varlık tarzına, örneğin bir ruha ait olduğu söyleniyorsa da bu ruhun tam olarak ne olduğunu sormak, nasıl davrandığını ve nerede bulunabileceğinin hesabının verilmesini bir şart koşmak[80] ya da fiziksel olmayan ve kütle, hareket, enerji, çekim, mekânsal uzam ve konum ile hayal edilebilecek başka tüm fiziksel unsurlardan yoksun bir şeyin nasıl fizik nesnelerle etkileştiğini sormak[81], fizikalizmin felsefi ve başka diğer bilgi alanları bakımından da düalist savlar karşısındaki önemli kazanımlarıdır. Fizikalizmin bu metinde *yadsınan* bir felsefi tavır *olmadığını* belirtmekle beraber okuyucunun dikkatini fizikalizme

[76] A. C. Cem Say, "Diller, Çözümsüz Problemler ve Sonsuzluk", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. Bs., İstanbul, Pan Yayıncılık, 2012, s. 184.
[77] Jacob T. Schwartz, "The New Connectionism: Developing Relationships Between Neuroscience and Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 123-124.
[78] Rene Descartes, **İlk Felsefe Hakkında Meditasyonlar**, çev. İsmet Birkan, 3. bs., Ankara, BilgeSu Yayıncılık, 2014, s. 53.
[79] **A.e.**, s. 77.
[80] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 37.
[81] **A.e.**, s. 40.

çekmek istememizin sebebi, tam da fizikalizmin bu kazanımlarının yapay zekâ projesi içerisinde yeni bir metafizik tasarımın doğurulması için kullanılmak istenilmesidir. Çünkü zihin çalışmalarının bilimsel ortamda yapılması gerektiğine ilişkin bir inanç söz konusudur ve bu inanç uğruna, bilimsellik için zihnin doğasındaki öznel ögelerin tahliye edildiğini, buradan hareketle de zihin felsefesi tartışmalarının bilimle iç içe olan fizikalistler tarafından indirgemeci bir tavırla sürdürülmeye çalışıldığını görmekteyiz.[82]

Realizm, algı nesnelerinin algılayandan bağımsız olarak bir gerçekliğinin bulunduğunu ve bu gerçekliğin kendinde şeyler olarak algılandığını söylemektedir. Dolayısıyla zihin ve bedenin bir aradalıkları, zihin-beden problemini meydana getirmiştir. Çünkü realist bir açıdan bakıldığında beden vardır ancak zihin gözlem ve deneyin konusu değildir. O yalnızca davranışlarla kendini göstermektedir.[83] Fizikalizm ise, realizmle tam bir uyumluluk göstererek, zihnin maddeyle eş türden açıklama modelleri ile ortaya konabileceğini, deneyimlerimizin ise beynimizdeki faaliyetlerden başka bir şey olmadığını, dolayısıyla zihinsel fenomenlerin alanının matematik ve fizik olduğunu iddia eder.[84] Dennett'a göre ise fizikalistler, ilkesel olarak, her bir zihinsel olguyu, radyoaktiviteyi, kıtaların sürüklenmesini, fotosentezi vb. açıklamakta kullandıkları fiziksel ilkelere ve materyale dayandırarak açıklayabileceklerini savunurlar.[85] Böylece organik veya inorganik materyalin doğası hakkında fizik, kimya ve biyolojideki en iyi kuramlar, fizikalizmde olduğu gibi kabul edilir ve dünya hakkındaki tüm yargılarımızın buralarda söylenebilecek şeylerden fazla olmadığı varsayılır.[86] Günümüz zihin felsefecilerinin çoğunluğu fizikalisttir ve zihin dediğimiz şeyin, beynin bazı kısımlarının davranışları ya da işlevlerinin sonucu olarak tanımlanabileceğini düşünür.[87] Ancak fizikalizmin de aşması gereken türlü problemler vardır. Örneğin Searle, zihinsel süreçlerin her ne kadar beyindeki birimlerin işleyişleri tarafından gerçekleştirildiğini ifade etse de beynin nasıl çalıştığı ve beyindeki fiziksel süreçlerle zihinsel süreçler arasındaki ilişkiler gibi konularda

[82] Doğan, "Biyolojik Doğalcılık Ekseninde John Searle ve Zihin-Beden Problemi", s. 228.
[83] Vernon ve Furlong, "Philosophical Foundations of AI", s. 54.
[84] Zeman, **Bilinç Kullanım Kılavuzu**, s. 393-398.
[85] Dennett, **Bilinç Açıklanıyor**, s. 46.
[86] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 54.
[87] Erdinç Sayan, "Analitik Zihin Felsefesinin Temel Problemlerine Bir Bakış", **Kaygı**, No: 19, 2012, s. 42.

herhangi bir bilgimizin henüz söz konusu olmadığını vurgular.[88] Veya bilinci ele alırsak bilince ilişkin fizikalist bir açıklamanın yaşayacağı zorluk, bilinç gibi fenomenal bir durumun madde ve onun ilişkileri türünden nasıl açıklanabileceğidir. Revonsuo'ya göre bu zorluk üç şekilde aşılmaya çalışılmıştır: Bilincin gerçekte var olmayan ve bizim bir yanılgımız olduğu yollu düşünce, bilincin var olan ancak tamamen fiziksel ve nörofizyolojik süreçlerin ürünü olduğu yollu düşünce ve bilincin var olan, gerçekliğe sahip, fiziksel bir fenomen olduğu ancak beynin üst düzey bir etkinliğinin sonucu olduğu yollu düşünce. İlk düşünce, elemeci fizikalizm olarak adlandırılan düşüncedir. İkinci düşünce indirgemeci fizikalizm olarak adlandırılır. Son düşünce ise belirimci fizikalizm olarak anılır.[89] Elemeci fizikalizmin, Dennett türü bir fizikalizmde temel bir rol oynadığını biliyoruz. Bilinç, özgür irade, nitelceler vb. hakkındaki açıklama girişimlerini spekülatif olarak gören bu tavır zihinsel olanı bir zamanların *filojistonu* olarak görür ve bilimsel gelişmeler sürdürüldüğü müddetçe, spekülasyonların döndüğü bu kavramların eleneceğini, bu fenomenlerin zaman içerisinde de ifade edilebilir formlar olarak belireceğini savlamaktadır. Ancak biz indirgemeci fizikalizme dikkat çekiyoruz. İndirgemeci fizikalist tavır, felsefi işlevselci savların konumlanışı bakımından önemlidir. Çünkü bu tavır zihinselliği, onun yetilerini ve fenomenlerini kabul eder ve bunların varlığını düşünmekle bir hata yapmadığımızı söyler. Ancak yine de insan zihninin ve beynin iki ayrı şey olduğunu düşünmekle bir hatada bulunduğumuzu belirtir. Bunların iki ayrı şey değil ama tek ve aynı şey olarak düşünülmesi gerekir. Nitekim de indirgemeci bir fizikalizmin amacı insanın zihinselliğinin beyinde gerçekleşen fiziksel süreçlerden farklı olmayan bir yapıda olduğunu savunmaktır.[90]

İndirgemeci fizikalizm, zihinsel fenomenlerin varlığını kabul ederek, fizikalizmin maddi olmayan fenomenlerin maddeden belirişine ilişkin yanıtlayamadığı sorunları aşamaz. Örneğin, kırmızı renkli bir elmayı duyumsuyor oluşumuz, belli bir anda bir acı hissine sahip olmamız ya da mutluluk hissinin kendisi, birer deneyim olarak fiziksel bir karşılığa sahip midir? Örneğin beynimizde

[88] John Searle, **Minds, Brains and Science**, 13. bs., Cambridge, Harvard University Press, 2003, s. 28.
[89] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 54-55.
[90] **A.e.**, s. 60.

kırmızının veya acı hissinin doğduğu yer işaret edilebilir mi? Fizikalizme göre, bu durumların fiziksel karşılıklarının olduğunu söylemek mümkündür. Ancak, Nagel'ın da belirttiği bu durum, fizikalizmin anlamını pekiştirir: Zihin durumları maddenin durumlarıdır ve zihin durumları fiziksel süreçlerdir.[91] Ancak burada büyük bir sorun bulunur. Örneğin, kırmızı bir elmayı duyumsadığımızda, duyumsuyor olduğumuz *o kırmızının*, beyinde biyokimyasal bir karşılığının bulunduğunu söylemekle, maddenin belli bir tarzının, belli bir örgütlenişinin ve düzenlenişinin, kırmızının ve kırmızı deneyiminin kendisini temsil edebileceğini, bu yolla da evrenin herhangi bir yerinde herhangi bir koşulda maddenin eş türdeki bir organizasyonun deneyimleyen açısından kırmızıyı temsil edebileceğini söyleyebilir miyiz? Zihin felsefesinin önemli sorunlarından biri olan izah gediği[*] bize maddi süreçler ile zihinsel deneyimler arasında, eş deyişle psiko-fiziksel yargılar arasında, bir izah gediğinin bulunduğunu söyler.[92] Dolayısıyla örneğin biz zihnimizdeki M türünden bir elektro-kimyasal sürecin mavi gibi bir nitelceyi meydana getirdiği bir durumu deneyimleriz ya da K türünden bir elektro-kimyasal sürecin neticesinde kırmızı gibi bir nitelceyi deneyimleriz. Ancak M veya K türünden bir elektro-kimyasal sürecin neden bu nitelcelere sebep olduğunu, bunların neden farklı nitelceleri meydana getirmediklerini izah edemeyiz.[93] İşte indirgemeci bir fizikalizmin aşması gereken önemli bir sorun olarak izah gediği sorunu, zihinsel olana ait anlam, öznellik vb. gibi üçüncü şahıs ontolojisine indirgenemez kimi hususları içerir[94], böylece indirgemeci fizikalizmin bu hususlara yanıt vermesi gerekir. Çünkü her ne kadar bilimler yoluyla fizikalizm, davranışların ruh gibi fizik üstü ögeler tarafından değil ama biyolojik mekanizmalar tarafından gerçekleştirildiğini göstermiş olsa da nörobilimsel sonuçlar bizim, fizikalizmde olduğu gibi, bir mekanizmadan başka bir şey olmadığımızı ispatlamamaktadır.[95]

---

[91] Nagel, "What Is It Like to Be a Bat?", s. 446.

[*] İng. **Explanatory Gap**

[92] Joseph Levine, "Materialism and Qualia: The Explanatory Gap", **Pacific Philosophical Quarterly**, No: 64, 1983, s. 354.

[93] Murat Arıcı ve Mehtap Doğan, "Prof. Dr. Erdinç SAYAN ile Söyleşi: Zihnin Doğası Üzerine", **MetaZihin**, C. I, No: 2, 2018, s. 252.

[94] Vedat Çelebi, "Searle ve Nagel'in Bilinç Kuramlarında Fizikalizm Eleştirisi Bağlamında Yapay Zekânın Sınırı Tartışması", **Beytülhikme**, C. IX, No: 9, 2019, s. 364.

[95] Adina Roskies, "Neuroscientific challenges to free will and responsibility", **Trends in Cognitive Sciences**, C. X, No: 9, 2006, s. 421.

İndirgemecilik, bilim felsefesinde eski kuramın ögeleri ile daha kapsamlı olan yeni bir kuramın ögeleri arasında kurulan özdeşlik önermeleri vasıtasıyla eski kuramın yeni kurama indirgenmesine verilen addır. Böylece felsefede indirgemeci fizikalizm, "tip fizikalizmi"* veya "özdeşlik kuramı"** gibi isimlerle anılır.[96] Şeylerin, belli başka şeyler türünden ifade edilebileceği düşüncesi olarak da tanımlanabilen indirgemeciliğin beş türü olduğu söylenebilir. Ontolojik indirgemecilik, bir varlığın kendisini oluşturan daha temel varlıklardan hareketle açıklanması girişimidir. Böylece bir sandalye, onu oluşturan moleküllerden başka bir şey değildir ve o moleküller üzerine etkiyen doğa yasaları ile açıklanabilir. Ontolojik indirgemeciliğin niteliksel bir türü de vardır. Bu yaklaşıma göre örneğin ısı veya gaz, moleküllerin kinetik enerjisi üzerinden ifade edilir. Kuramsal indirgemecilik ise bilimde olguların yasalarla ve yasalar arası ilişkilerle ifade edilmesidir. Böylece eski yasalar daha yeni ve kapsayıcı türden kuram ve yasalarla açıklanabilir. Mantıksal veya tanımsal indirgemecilik, felsefeciler arasında yaygın olan ve kelimeler ile yargıların aracısız bir şekilde dönüştürülebilmesi ve kullanılabilmesine denir. Örneğin sayılarla ilgili belirlemelerimizi kümeler hakkındaki yargılara indirgeyebiliyoruz. Nedensel indirgemecilik ise varlığın belli özelliklerinin veya varlıklar arası ilişkilerin belli durumlarının, o varlıkların sahip oldukları başka özelliklerin sonucu olduğu yollu görüştür.[97]

Yapay zekâ projesinde ele alınan ve bilgisayar bilimciler tarafından da üstlenilen zihin-program analojisi, niteliksel ve mantıksal indirgemeci tavırlara dayanır, indirgemeci bir fizikalizmle de buluşur. Böylece kişiliğin ve belleğin elle tutulamayacağı söylenir, yazılım katmanlarının da bu özelliğe sahip olduğu vurgulanır, ancak fizikalist tavrı da unutmadan, yazılımların 0'lar ve 1'ler ile ifade edilebilecek bir tarzda makine dili ile inşa edildiği söylenir[98], daha önceleri fizik üstü bir tavırla açıklanmaya çalışılan zihinsel fenomenler şimdi ise transistörlerin açık veya kapalı oldukları durumlar üzerinden biçimselleştirilir. Sembol sistem

* İng. **Type Physicalism**
** İng. **Identity Theory**
[96] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 61.
[97] Searle, **The Rediscovery of the Mind**, s. 112-115.
[98] Haluk Bingöl, "Bilgisayar Üzerine", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 30-31.

hipotezinin mimarları olan Allen Newell ve Herbert Simon'da insan davranışını açıklamakta özgün ve analitik, yeni bir araç olarak programların söz konusu olduğunu, uygun şekilde programlanmış bir bilgisayarın tıpkı insanların gerçekleştirdiği gibi karmaşık düşünme süreçlerini paralel bir şekilde gerçekleştirebileceğini düşünür.[99] Daha da temele indiğimizde ise Turing ile karşılaşırız. Ünlü makalesinde Turing, karmaşık bir işlem sürecinde bir insanın neler yaptığının bir bilgisayara taklit ettirilmek istenmesi durumunda insana yalnızca o karmaşık süreç esnasında ne yaptığının sorulmasının yeterli olacağını, böylece alınacak yanıtın bir yönerge tablosu şeklinde bilgisayarlar tarafından gerçekleştirilebileceğini, bu bilgisayarların ise dijital bilgisayarlar olarak soyut hal makineleri olarak düşünülmesi gerektiğini savunur[100], ancak bu soyut makinelerin fiziksel makineler olmaktan ziyade matematiksel kurgular olduğunu da belirtir.[101] Turing'in çabası fizikalist ögeler içerse de daha çok işlevselci bir yaklaşımdır. Çünkü Block'a göre, fizikalizm ve işlevselcilik arasında bir fark bulunur: İşlevselcilik zihinsel bir durumu beyinde gerçekleşen bir rol olarak görürken fizikalizm ise zihinsel durumları fiziksel ve biyolojik olarak bu rolü dolduran veya yerine getiren veya gerçekleştiren bir şey olarak düşünür.[102] Dolayısıyla Turing ile başlayan yapay zekâ fikri, savunucuları tarafından da fizikalist bir temele oturtulur ve işlevselci bir tavırla sürdürülür. Ancak yapay zekâ projesinin oturduğu fizikalist temeller düşünülürse içerdiği kimi boşlukların bulunduğu görülebilir. Örneğin Searle, ısının indirgenebildiği gibi bir durumda neden acının benzer şekilde indirgenemediğini sorar. Çünkü ısı öznel bir fenomen değildir, o yalnızca fiziksel olayların bir neticesidir.[103] Sadece niteliksel bir indirgemeciliğe değil, ontolojik bir indirgemeciliğe de itiraz söz konusudur. Örneğin Searle'ün felsefesinde bilinç, tıpkı katılık gibi belli bir sistem özelliğidir. Ancak bilinç, katılık gibi onu oluşturan mikro yapının kendisi üzerinden açıklanamaz.[104] Böylece yapay zekâ projesiyle

[99] Allen Newell ve Herbert A. Simon, "The Simulation of Human Thought", (Çevrimiçi) https://stacks.stanford.edu/file/druid:cn091ff4880/cn091ff4880.pdf, 7 Aralık 2022.
[100] A. M. Turing, "Computing Machinery and Intelligence", **Mind**, C. LIX, No: 236, 1950, s. 438-439.
[101] **A.e.**, s. 449.
[102] Ned Block, "Paradox and cross purposes in recent work on consciousness", **Cognition**, No: 79, 2001, s. 203.
[103] Searle, **The Rediscovery of the Mind**, s. 120.
[104] **A.e.**, s. 123.

biçimselleştirilmeye çalışılan insan zihni, ne onun üzerinde var olduğu yapı olarak nöronlarla ne de nöronlar arası ilişki ve örüntülerle açık kılınabilir.

Fizikalizmin yapay zekâ felsefesindeki yeri açıktır: İnsanın zihinsel süreçleri, bilimler yoluyla, sinirsel aktiviteden ibaret kılınır. Sinirsel aktivitenin incelenmesiyle erişilemeyecek nitelceler, bilinç, fenomenal deneyim vb. gibi unsurlar gerçekleştikleri maddi yapının kendisine indirgenir, bu yolla da onların ne olduklarına ilişkin metafizik bir sorgulamanın yükünden kurtulan bilim, zihin felsefesinin problemlerini *çözmüş* olur. Okuyucu merak edebilir: Zihin felsefesinin problemlerinin aşılmasında veya bu problemlerde bir aşamanın kaydedilmesinde bilimin ilerici rolünü yadsıyor muyuz? Kesinlikle hayır. Bilimsel yöntem ve paradigmalar yoluyla felsefenin sayısız problemi çözülmüştür. Bizler artık evrenin merkezinde yer aldığımızı düşünmüyor, düşüncelerimizin göğsümüzde durmadan atmakta olan kalpten cereyan ettiğine inanmıyor, bir zamanlar fizik üstü varlıklar tarafından bahşedildiğini düşündüğümüz yönetim erkini, **polis**'te ve yurttaşlarca var edilen bir erk olarak tanıyoruz. Ancak unutulmamalıdır ki bilimsel yöntem ve paradigmalar değiştikçe, felsefenin de yeni problemler edinmesi, gündelik olanla kuramsal olan arasındaki ilişkinin incelenmesi söz konusudur ve söz konusu olacaktır. Yapay zekâ projesi de bilimin insanlığın ve pek tabii felsefenin gündemine getirdiği bir gelişmedir. Bu gelişmenin felsefenin her alt alanı ve diğer bilgi alanları bakımından da sıkıca düşünülmesinde bir beis görülemez. Tam da bu yolla felsefe, yapay zekâyı ihraç etmek için değil, tam aksine onu insan yaşamında anlamlı kılmak için çabalar. Bu çalışmanın amacı bir yapay zekâ karşıtlığı üretmek değildir. Tam aksine, "yapay zekâ" kavramının her yönü ve temeli ile ele alınması, onun *insan yaşamında ve insanca yaşamda*, diğer bilgi ve bilim alanları bakımından ve **polis** açısından da hak ettiği yerin tesis edilmesi için zorunludur. Ancak görüyoruz ki, fizikalist temellere dayanan yapay zekâ projesi, insan zihnini ve beynini bir makine olarak tasarlar, onu simüle etmeye çalışır, matematik ve mekanik yasalar ölçüsünde biçimselleştirir fakat insan zihni Zambak ve Vergauwen'in de ifade ettiği üzere makinelerde gerçekleştirilemez, matematikselleştirilemez ve mekanik ilkelerle

açıklanamaz.[105] Aksi halde Twain'in Yaşlı Adam'ı gibi, aklımızın yalnızca bir makineden ibaret olduğu, üzerinde yetkimizin bulunmadığı ve beynimizin de kendi üzerinde yetkisinin olmadığı, dışarıdan çalıştırılan ve tüm makinelerin kanunu diyerek onun da kanunun bu olduğu[106] yollu düşünceyi savunmak için elimizde veri var mıdır ve bu düşünce *olan* ile örtüşür mü?

1.2. Yapay Zekâ Felsefesine Temel Oluşturan Bilimler Üzerine

Bu çalışmanın temelinde de yer alan incelemenin, felsefi işlevselci savların temelleri ve iddiaları bakımından bir incelemesinin, yalnızca bu tavrın felsefi yöntemini tahlil etmek yoluyla gerçekleştirilemeyeceğini, öte yandan bu tavrın dayandığı bilimleri, bu bilimlerin temel iddia ve problemlerini de inceleyerek ortaya koyabileceğimizi düşünüyoruz. İşte burada, yapay zekâ felsefesindeki temel bilimler olarak görülen matematik, mantık ve fizik, kimya ve biyoloji gibi doğa bilimlerini inceleme niyetindeyiz. Ancak burada okuyucunun zihninde bir soru canlanabilir: Neden yalnızca matematik, mantık ve doğa bilimleri incelemenin amacı gereği önde gelmektedir? Psikoloji, sinirbilim ve diğer alanlar bakımından da bu öncelik sıralaması değişemez mi? Okuyucu için açık kılmalıyız ki, bu çalışmada amacımız, felsefi işlevselci savların bir eleştirisini sunmaktır. Bu savlar çok yalın bir tarzla, zihni bir bilgisayar programı gibi görüp, bedeni ilinekselleştiren bir amacı beslemektedir. Bir başka çalışmanın veya çalışmaların konusu olarak okuyucunun zihnindeki diğer bilgi ve bilim alanları bakımından da incelemeler yapılabilir. Ancak bu çalışmada biz, zihni matematikselleştiren, transistörler ve koşullu ifadeler yoluyla, önermeler yoluyla onu mantıksallaştıran ve doğa bilimleri vasıtasıyla bu arzunun fizikalist beklentilerini bilgisayar bilimi ve bilgisayar teknolojileri vasıtasıyla karşılayan bir projenin, yapay zekâ projesinin, temellerini anlama çabasındayız. Okuyucu belki felsefi işlevselci tavrın ve bu tavra ait savların metnin daha önceki bölümlerinde aktarılması gerektiğini düşünebilir. Belki de çalışmayı, öncelik ve sonralık sıralaması bakımından sıkıca düşünülmemiş bir çalışma olarak niteleyebilir.

[105] Aziz Zambak ve Roger Vergauwen, "M-S Models: A New Approach to Models and Simulations in Artificial Intelligence", **Logique et Analyse**, C. L, No: 198, 2007, s. 187.
[106] Twain, **İnsan Nedir?**, s. 11.

Ancak çalışmanın bütünü düşünüldüğünde, yapay zekâ felsefesinin ve daha özelinde felsefi işlevselci savların, Turing ve sonrası dönemde şekillenen konumlanışı, geometrik olarak ifade edilirse, bir Möbius şeridine benzemektedir. Görülecektir ki Leibnizci bir arzu olmasaydı, Hilbert'in karar verme problemi adını verdiği sorun doğmayacaktı. Turing'in insanın matematik yaparken mekanik süreçleri işlettiği yollu fikri zihinsel olanın matematiksel ve mantıksal olarak ifade edilebileceği yollu arzunun sonucu gerçekleşmemiş midir? Rölelerden transistörlere giden süreç matematik, mantık ve doğa bilimleri bakımından bir bütün olarak ele alınamaz mı? Dolayısıyla bu çalışma bir Möbius şeridi gibidir. Bu çalışmanın sonu, bir başlangıçtır ve başlangıcı da bir sondur. İşte biz döngüsel bir hale gelmiş düşünce süreçlerini kırmayı ve felsefi işlevselci tavrı elemeyi değil ama bu tavra, savlarının yeniden ele geçirilmesi için bir sorunlar listesi sunmayı amaçlıyoruz. Dolayısıyla gelecek alt bölümler, bu tavrın dayandığı matematik bilimi, mantık bilimi ve doğa bilimleri bakımından kökleri ortaya koyacak, ancak bunları ortaya koyarken bu bilim alanlarının yapısını, içerdikleri önem arz eden ilintili tartışmaları ve soru işaretlerini dile getirecektir. Böylece okuyucu açısından da açık kılınacaktır ki yapay zekâ felsefesi ve daha özelinde felsefi işlevselci tavır, yalnızca savlarının doğruluk değerleri hesaba katılarak bir başarıya kavuşamaz. Bu tavır aynı zamanda, savlarını kurmak bakımından dayandığı temel bilimlerin sorunlarını aşmak durumundadır. Dolayısıyla bu problemlerin örtüldüğü, tartışılmadığı bir alanın kavramı olarak "yapay zekâ" kavramını güvenilir bulmuyoruz. Felsefi işlevselci tavır bakımından ulaşılmak istenen arzunun yöntem bakımından mevcut olan değil ama bu çalışmada dile getirilen sorunları da kapsayıcı, daha modern bir yöntemle başarılabileceğini iddia ediyoruz.

1.2.1. Yapay Zekâ Çalışmalarının Matematiksel Temelleri

Medeniyetimizi sağlamlaştıran, onun sınırlarını genişleten en temel bilgi alanlarından biri olarak matematik, insan yaşamının her unsurunu geliştirmede önemli bir rol oynar. Nesnelerin sayılması, sayılan nesnelerle temel işlemlerin – çarpma, bölme, toplama, çıkarma – gerçekleştirilmesi ve bu yolla da insan

yaşamındaki unsurlar, yaşam koşulları ve ihtiyaçlar değiştikçe matematiğin buna yanıtlar üretme döngüsü, günümüzde de devam etmektedir. Hatta matematik, şekillerin alanı olarak geometriyi, var olanların alanı olarak ise fiziği sayısallaştırma, onları kendisine indirgeme gücüne bile sahiptir. Peki ama matematik nedir? Matematik, sayılar, şekiller ve soyut yapıların özelliklerini incelemeyi ve bu nesneler hakkında zorunlu önermeler ortaya koymayı amaçlayan bir bilimdir. Konusu bakımından matematik, tek tek nesneleri değil ama nesne kümeleri veya nesne sistemlerini incelemektedir. Doğa bilimlerinde bu sistemlerle alakalı olumsal önermeler ortaya koyma amacı taşınırken, matematikte ise bu sistemlerin yapısal özellikleri hakkında zorunlu önermeler üretilmeye çalışılır. Matematiğin temel ilkelerini ise, kendisinden daha öte bir temellendirmenin söz konusu olmadığı yargılar olarak *aksiyomlar* oluşturur.[107]

Matematik felsefesini doğuran, matematik felsefecilerini ve matematikçileri meşgul eden sorunların büyük bir kısmı matematiğin aksiyomları hakkındaki tartışmalardan doğar. Sayı nedir? Sayıların ontolojik statüsü nedir? Sayılabilmek ne anlama gelir? Matematikte dile getirilen yargılar güçlerini ve geçerliliklerini nereden alırlar? Kümeler gerçeklikte karşılıkları bulunan kategoriler midir? Fizik matematiğe indirgenebilir mi? Burada matematik felsefesi bağlamında bu çalışmanın kapsamını oldukça aşabilecek pek çok soru sorulabilir. Bu sorular matematik alanında çalışan felsefeciler ve matematikçileri meşgul eder. Çünkü matematik alanı, bu alanda üretilen önermelerin, ortaya konan ifadelerin güvenilirliği ve sıkılığı iddiası ile ön plana çıkar. Eş deyişle, matematiksel bir ifade, bu ifadeyi dile getiren iki zihin tarafından iki ayrı önerme olarak algılanamaz, yorumlanamaz. Matematiksel ifadeler bu sıkılıklarını, sembollere ve matematiği var eden temel aksiyomlara borçludur. Matematik felsefesi gibi bir alanın var olmasının koşulları da tam da matematiğin var olanlara yöneltilmesi ve var olanlar hakkında matematikten hareketle yargılarda bulunulabileceği yollu düşüncedir. Matematik yapmanın ve matematikteki buluşçuluğun kökeni nedir? Matematik nesneler güvencelerini hangi ontolojik temele dayandırırlar? Bu sorular matematik felsefesinde sıkça tartışılmış ve felsefe tarihi boyunca türlü filozoflar ve matematikçiler tarafından matematik alanının ve onun

[107] Teo Grünberg, **Sembolik Mantık El Kitabı: Sembolik Mantığın Uygulamaları**, C. III, Ankara, METU PRESS, 2000, s. 212.

nesnelerinin güvencesi farklı ontolojik ve epistemolojik anlayışlarda inşa edilmeye çalışılmıştır.

Bir sistem olarak matematik, icra edilebilmesinin koşulları bakımından bazı aksiyomları ve bu aksiyomlardan çıkarımların ne şekilde elde edilebileceğini gösteren çıkarım kurallarını, bir ispata dayanmadan barındırmak zorundadır çünkü matematiksel ispat denen olgu, en nihayetinde herhangi bir ispata dayanmayan bir noktadan hareketle gerçekleşmek durumundadır.[108] Peki ama bu ispata dayanamayan aksiyomların ve bu aksiyomlarda istifade edilen sembollerin güvenceleri nasıl sağlanacaktır? Temel matematiğin ve temel geometrinin kavramlarının, gerçekliğimizde yer alan varlıkları tanımlamak ve varlıkları işlemin konusu yapmak için biçimsel hale getirilmiş doğruluklar olduğuna ilişkin bir görüş bulunsa da[109] bu temel kavramlar, tıpkı bir teorem gibi başlangıç aksiyomlarından hareketle geliştirilen, dolayısıyla onlara ve onların işletilmesini olanaklı kılan akıl yürütme kurallarına dayanan bir yapıda değildir. Çünkü teorem için böyle bir yapı sağlandığında teoremin sonunda **quod erat demonstrandum**[*] ifadesini kullanabiliyoruz ve teoremi ispat ettiğimizi söyleyebiliyoruz.[110] Ancak matematiğin temel aksiyomlarının ve onun temel kavramlarının bir ispat ürünü olduğunu söyleyemiyoruz. Öyle ki matematiği bir dil olarak kabul edersek, alfabesi ve o dildeki doğru önermeler kümesinden oluşan iki temel parçasıyla bu dil, temel aksiyom ve kavramlarından sonsuz adet önermenin türetilebildiği bir olanağı meydana getirebiliyor. Nitekim doğru önermeler kümesi ifadesiyle de kastedilen, bu sonsuz adet önerme içerisinde hem anlamlı olan – yani bu dilde yargıların üretilmesi için gereken kurallara uygun olarak kurulmuş – hem de doğruluk değeri alabilen önermelerdir.[111] Matematik tarihi incelendiğinde de bu önermelerin ve dört temel operasyondan farklı olarak yeni önermelerin ve mevcut sorunların çözümü için

---

[108] Kurt Gödel, "Some basic theorems on the foundations of mathematics and their implications", **Collected Works: Unpublished essays and lectures**, ed. Solomon Feferman v.d., C. III, New York, Oxford University Press, 1995, s. 305.

[109] Eugene P. Wigner, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", **Communications on Pure and Applied Mathematics**, C. XIII, 1960, s. 2.

[*] T. Gösterilmek istenen şey buydu.

[110] Torkel Franzen, **Gödel's Theorem: An Incomplete Guide to Its Use and Abuse**, Wellesley-Massachusetts, A. K. Peters, 2005, s. 16.

[111] Vladimir A. Uspensky, "Gödel's incompleteness theorem", **Theoretical Computer Science**, C. LXXX, No: 2, 1994, s. 242-243.

ihtiyaç duyulan yeni operasyonların türetildiğini – yani matematikte halihazırda olanaklı olan sonsuz önerme içerisindeki anlamlı ve doğruluk değeri alabilen önermelerin araştırılması çalışmalarını – görüyoruz. Nitekim matematikte yaşanan paradigma değişimlerinin de mevcut paradigma ile kapsanamayan, açıklanamayan matematiksel – veya fiziksel – olguların daha kapsayıcı veya açıklama gücü yüksek matematiksel teoremler veya operasyon türleri ile açıklama çabaları olduğunu söyleyebiliriz. Çünkü örneğin bir matematiksel ifadenin genel geçerliği, yani açıklama gücü, bu matematiksel ifade ile kapsanan sayısız durum ve nesne için geçerli olmasıyla ilgilidir[112] ve matematik tarihinden de bildiğimiz gibi diferansiyelin bulunması, analitik geometrinin – geometri alanında tam sayılı sonuçlar bulma çabalarına yanıt verilememesinden dolayı – geliştirilmesi vb. gibi paradigma değişimleri de bu duruma birer örnektir.

Ancak bizler matematiğin temel aksiyomları ve matematiğin temel nesneleri ile kavramları hakkında benzer iddialarda bulunamıyoruz. Onlara bir dayanak noktası, bir son nokta bulamıyoruz. Kavramların arka planları bakımından daima onları önceleyen bir kavram bulunsa da "İşte burada durdum!" diyebileceğimiz bir noktanın bulunmuyor oluşu, düşünürleri türlü yanıtlar ve türlü durma noktaları bulmaya yöneltmiştir. Matematiğin temelleri bakımından ontolojik ve epistemolojik dayanaklar bulma ihtiyacını hissederiz. Nitekim matematik felsefesi tarihi ve matematik tarihi bakımından da felsefeciler ve matematikçiler, matematiğin ve geometrinin ögelerinin insanın zihninde birer ürün olabilmesinin koşullarını soruşturmuşlar, bazı filozoflar realist bir çizgide yanıtlar ararken, bazı filozoflar ise bu ögelerin güvencesini fizik üstü bir gerçeklik alanında konumlandırmayı tercih etmişlerdir. Platon'da ideal bir gerçekliğe sahip olan matematik ve geometrik ögeler[113], Aristoteles'in felsefesinde ideal bir alandan içinde yaşadığımız gerçekliğe indirgenmiş[114], böylece matematik felsefesi bakımından iki ana düşünce kampı meydana gelmiştir. Çağdaş matematik felsefesi tartışmalarında mantıkçılık, doğalcılık, sezgicilik, biçimselcilik vb. gibi birçok çözüm yolu bulunur, bu yolların

[112] David Hilbert ve W. Ackermann, **Principles of Mathematical Logic**, çev. Lewis M. Hammond, George G. Leckie ve F. Steinhardt, ed. Robert E. Luce, New York, Chelsea Publishing Company, 1950, s. 112.
[113] Platon, **Menon**, çev. Furkan Akderin, 2. bs, Ankara, Say Yayınları, 2013.
[114] Aristoteles, **Metafizik**, çev. Y. Gurur Sev, 4. bs., İstanbul, Pinhan Yayıncılık, 2018.

Antiklere gidildiğinde Platoncu veya Aristotelesçi bir kaynaktan doğdukları görülebilir. Burada dile getirdiğimiz eğilimler, onları meydana getiren farklı felsefi anlayışların bulunuyor olmasıyla da matematik ve geometri alanlarının güvencesini sağlama yolları bakımından birbirlerinden ayrılırlar.

Sezgiciler, temel aksiyom ve kavramların sezgi vasıtasıyla dolaysızca bilindiğine, bir önermenin doğruluğu ya da dolaysızca bilindiği söylenen bir kavram hakkındaki felsefi akıl yürütmenin bir noktada durduğuna, matematik ve geometri gibi alanlardaki buluşçu durumların ancak ve ancak sezgi vasıtasıyla gerçekleşebileceğine işaret ederler.[115] Mantıkçılar ise her türlü psikolojizmden uzak, dolaysızca türetilen kavramlara izin vermeyen, sıkı bir inşa sürecini savunurlar. Matematiksel kavramların inşa edilmesinde ve anlaşılmasında temel kavramların analiz yoluyla incelenmesi, böylece alt kavramlara ve alt kavramlar vasıtasıyla da temel kavramlara anlamlarının kazandırılması amaçlanır. Çünkü mantıkçılara göre matematikteki doğruluklar özleri itibarıyla mantıksal doğruluklar olarak kabul edilir. Dolayısıyla bu doğruluklar mantıktan pay aldıkları için yine mantığın ilkeleri ve yasaları ile türetilebilir, ifade edilebilirler. Matematikteki önermeler bu yolla mantıksal bir ifade formunda yazılabilir durumdadır, bu bakımından matematik, mantık temelinde inşa edilebilir bir alandır.[116] Mantıkçı bir matematik anlayışının tam karşında bulunan sezgiciler, mantıkçıların analiz yöntemini eleştirir, böylece mantıkçı bir matematik yapmanın kavramların sonu gelmez analizleriyle veya mantıksal ilkelerin farklı yorumlanmalarıyla elde edilecek farklı matematikler gibi bir sorunlarla yüzleşmesi gerektiğini düşünürler. Sezginin bulunmadığı bir matematiksel sistemin ise buluşçuluk özelliğinin nasıl sağlanacağı mantıkçı bir matematik yapma tarzında açık değildir.

Matematik felsefesinde doğalcılık, matematiğin doğalcı bir felsefi anlayışla inşası düşüncesinde dayanır. Doğalcılık, metodolojik ve ontolojik doğalcılık olarak iki farklı yoruma sahiptir. Metodolojik doğalcılık, fizik, kimya, biyoloji gibi doğa bilimlerini oluşturan bilimsel metotları kapsar. Böylece doğru metotların uygulanmasıyla dünya hakkındaki bilgimiz ve anlayışımız genişler. Ontolojik

[115] Blackburn, "intuition", **The Oxford Dictionary of Philosophy**.
[116] Blackburn, "logicism", **The Oxford Dictionary of Philosophy**.

doğalcılık ise doğa bilimlerinin ontolojisini yani her türden fizik objeleri ve hatta fiziki nitelikleri ve ilişkileri dikkate alır.[117] Dolayısıyla doğalcı bir matematik anlayışı, matematik alanını doğalcı bir metodoloji ve ontoloji ile inşa edecektir. Örneğin Mill'e göre sayılar nesnelerin yalnızca bir özelliğidir. Nesnelerin özellikleri olmalarından ötürü sayılar, nesnelerin bir araya gelme ilkeleri olarak, belli bir yığışımdaki nesnelerin sayısıdır. O halde Mill bakımından sayı kavramı şeylerin sayısı yani sayılabilmenin bir sonucu olarak karşımıza çıkmaktadır. Ancak Mill bakımından her sayının türetilebilmesi için doğada duyumsanması gerektiği düşünülmemelidir. Temel sayılar bir kez elde edildiğinde, sayılar arası ilişkiler vasıtasıyla diğer sayıların da türetilmesi olanaklıdır. Doğalcı bir matematiğin yanıtlaması gereken önemli hususlardan biri ise türev, logaritma vb. gibi matematiksel işlemlerin doğada gözlemlenemiyor olmalarına rağmen matematikte var olabilmelerinin doğalcı perspektiften açıklamasının nasıl yapılacağıdır. Mill bakımından bu işlemler, gündelik deneyimin kendisinde yer alan ve sayılara ulaşmanın örüntüleri olarak karşımıza çıkarlar. Dolayısıyla bu işlemler de doğada bulunabilir, doğadan elde edilebilir niteliktedir.[118] Matematiksel doğalcılık, Philip Kitcher gibi günümüz düşünürleri tarafından da savunulmaktadır. Kitcher'a göre matematiksel ilerleme nesiller arası bir pratik aktarımıdır. Böylece matematik tarihinin, matematiksel paradigmaların ve ilerlemelerin, onların nesiller arası pratiksel geçişkenliğinden ve onların devinmesinden meydana geldiğini, bu pratiklerin yaşamın soyut ihtiyaçlarına birer yanıt olduğunu savunur. Bu anlamda doğalcı bir açıdan matematik, pratikçe belirlenen ve pragmatik kaygılara hitap eden entelektüel ve zihinsel bir etkinlik alanıdır.[119]

Görülebileceği gibi sezgicilik, mantıkçılık ve doğalcılık, matematiğin ve geometrinin güvencesini farklı ontolojik yaklaşımlarda ve bunlara dayanan epistemolojik dayanaklarda bulmak ister. Biçimselcilik ise burada dile getirdiğimiz yaklaşımlardan farklı olarak matematiği semboller ve operasyonlar arası bir ilişki

[117] Alan Weir, "Naturalism Reconsidered", **Oxford Handbook of Philosophy of Mathematics and Logic**, ed. Stewart Shapiro, New York, Oxford University Press, 2005, s. 460.
[118] John Stuart Mill, "Of the Remaining Laws of Nature", **A System of Logic**, 8. bs., New York, Harper&Brothers, 1882, iii, xxiv, § 5, § 6.
[119] Philip Kitcher, "Mathematical Naturalism", **From Kant to Hilbert Volume 1: A Source Book in the Foundations of Mathematics**, C. I., ed. William Bragg Ewald, Oxford, Oxford University Press, 1996.

olarak kabul eder. Böylece bu alandaki önermeler bir dizi sembolün türlü ilke ve kurallara uygun şekilde sıralanması işi olarak görülür.[120] Biçimselcilik yaklaşımında en akla gelen isimlerden biri David Hilbert'tir. Hilbert'e gelesiye matematik içerisinde çözülmeyen türlü sorunlar söz konusudur. Sonsuz kavramı, sayının türetilme çabaları vb. gibi sorunlar mevcuttur. Dolayısıyla Hilbert, matematiğin biçimselci bir yeniden inşası ile, matematiğin her ihtiyacının yine matematiğin içerisinde karşılanabileceğini, matematikte çözümsüz herhangi bir problemin kalmayacağını, böylece matematiksel ifadelerin, yine kendilerine dayanarak sıkı bir sistemin inşasında rol oynayabileceğini savunmuştur. Ancak bu düşüncesinin gerçekleştirilebilmesi için matematik gibi olanaklı sonsuz önermenin içerisinde anlamlı ve doğruluk değeri alabilen önermelerin bulunmaya çalışıldığı bir alanın sınırlanması gerekliliği açıktır. "Sonluculuk" olarak anılan yaklaşım, bir matematiksel metodun veya sistemin, artık ona başka bir nesne veya ifadeyi kabul etmeyecek bir tarzda olmasıdır. Böylece tam hale gelmiş olan bir sistem, dışarıdan başka bir öge veya ifadeye ihtiyaç duymaksızın sistem içerisinde sınırlı sayıda adım ve elemanla, sezgicilikten farklı olarak, matematiksel ihtiyaçlara yanıt üretebilir.[121]

17. yy. düşünürü olan Leibniz, evrensel ve yapay bir matematiksel dil üzerinden bilginin ifade edildiği, hesaplama kurallarının onun üzerinde işletildiği ve buradan hareketle biçimselleştirilecek bilgilerin tümü arasındaki mantıksal ilişkilerin yürütülebileceği bir dil hayali kurmuştur. Öyle ki Leibniz, seslerle değil ama kavramlarla temsil edilen özel bir alfabe aramaktadır. Bu alfabe sayesinde bir dilde dile getirilebilen ifadelerin sembolik hesapları yapılabilecek, böylece bunlar arasındaki mantıksal ilişkiler tayin edilebilecektir.[122] Görüyoruz ki matematiğin ve mantığın psikolojizme izin vermeyen sıkılığı, dünya hakkındaki olguların dile getirilebilmesi bakımından matematik ve mantık alanlarının en uygun alanlar olarak görülmesini sağlamaktadır. Hilbert'in sonlucu ve tam bir matematik sistem tasarlama programı, Leibniz'in hayallerinden birinin gerçekleştirilmesi olarak nitelenebilir.

---

[120] Blackburn, "logicism", **The Oxford Dictionary of Philosophy**.
[121] Michael Proudfoot ve A. R. Lacey, "Finitism", **The Routledge Dictionary of Philosophy**, 4. bs., New York, Routledge, 2010.
[122] Martin Davis, **The Universal Computer: The Road From Leibniz to Turing**, New York, W. W. Norton & Company, 2000, s. 4-5.

Bir matematiksel sistem incelenirse, bu sistemi meydana getiren ve bu sistemde istifade edilen aritmetik ögeleri olanaklı kılan savların, sistemi inşa eden düşünürler tarafından da analitik olarak görüldüğünü, dolayısıyla temel aritmetik önermelerin doğruluklarının mantıksal olarak temel aksiyomlardan dedüktif olarak çıkarılabildiğini ve tanımı olarak doğru olduğunu görürüz. Mantıkçılık olarak anılan bu yaklaşıma göre matematiği meydana getiren her unsur zorunlu doğruluklardır. Ancak bir sorun bulunur: Tüm bu temel önermelerin doğruluk değerlerinin sabit ve zorunlu olduğu yollu düşüncenin dayanağı nedir?[123] Aksiyomlar belirlenirken herhangi bir dayanağa sahip değildir ve örneğin mantıkçılıkta, bunun temellerinin mantık ilkelerine dayandırılıyor olması bize nihai bir güvence veremez. Çünkü mantık ilkelerinin farklı yorumlanmaları olanaklıdır ve olanaklı dünyalarda mantık ilkelerinin farklı yorumlanmaları söz konusu olabilir. Böylece farklı aksiyomlara dayanan farklı matematikler gibi bir sorunla karşılaşmış oluruz, nitekim farklı matematiklerin olanaklı olması, mevcut matematiğin kurgusal olduğunu, dolayısıyla bize bu kurgusal yapı içerisinde türetilen önermelerin sağın olmadığını göstermez mi? İşte Hilbert, bir sistem içerisinde hiçbir yargının karar verilemez olmadığı bir sistemi düşlemiş, böylece bir sistemde doğruluk değeri hakkında karar veremediğimiz hiçbir önermenin kalmadığı tarzda, tamamlanmış bir matematik sistem tasarlamayı istemiştir. Çünkü bir sistemde bir önermenin doğruluk değeri ile ilgili bir yargıya varılamazsa bu sistem tamamlanmamıştır, yani eksiktir.[124] Bu problemi bir ana problem olarak ele alan Hilbert[125], 1928 yılında üç adet önemli soru yöneltmiştir: Matematik tamamlanmış mıdır yani her matematik ifade bir sistem içerisinde doğrulanabilir veya yanlışlanabilir mi? Matematik tutarlı mıdır yani 1=2 türünden çelişkili ifadeler ispatlanamaz bir yapıda mıdır? Matematiksel bir sistem karar verilebilir bir yapıda mıdır yani s gibi bir önermenin S gibi bir mantıksal matematiksel sistem içerisinde ispatlanıp ispatlanamayacağına karar verebilen bir mekanik metot var mıdır? Biz son soruyu karar verme problemi[*] olarak tanıyoruz.[126] Böylece Hilbert'in matematik projesinin, içerisinde birbiriyle çelişkili ifadelerin

[123] Peter Smith, **An Introduction to Gödel's Theorems**, C. II, Logic Matters, 2020, s. 2-3.
[124] Franzen, **Gödel's Theorem: An Incomplete Guide to Its Use and Abuse**, s. 18.
[125] Hilbert ve Ackermann, **Principles of Mathematical Logic**, s. 113.
[*] Alm. **Entscheidungsproblem**.
[126] Luciano Floridi, **Philosophy and Computing: An Introduction**, New York, Routledge, 1999, s. 26-27.

türetilemeyeceği, tam ve tutarlı, sonlucu bir matematik sistemin inşa edilmesi olduğunu anlıyoruz.[127]

Tam ve tutarlı bir matematik sistem, "a" ve "a-olmayan" türünden iki önermenin aynı anda ispatlanamadığı dedüktif, sentaktik bir sistemdir. Bu dedüktif sistemin tam ve tutarlı olabilmesi için ya "a" ya da "a-olmayan" önermelerinden sadece birinin ilgili sistem içerisinde ispatlanabilir olması gerekir.[128] Hilbert'in programının amacı ise her türden matematiksel doğruluğa ulaşmamızı sağlayan uslamlama yöntemlerini de kapsayacak kadar geniş kapsamlı aksiyomlar ve ilkeler listesine sahip sıkı tanımlanmış bir matematik sistemin yaratılmasıdır. Örneğin bu programa uygun olarak ancak ve ancak geniş kapsamlı bir aksiyomlar ve ilkeler sistemine sahip olunursa herhangi bir aritmetik önerme ile ilgili matematiksel ispatın doğruluğuna ait kesinlikli ilkelere sahip olunabilir. Fakat önermelere doğruluklarını sağlayan bu sistemdeki her bir öge – örneğin aksiyomlar – sağlamlıklarını sistemin bütünü tarafından kazanır.[129]

Matematiksel ifadelere, biçimsel bir matematik sisteminde yer alan sembol dizilerinden başka bir şey olmadıkları şeklinde yaklaştığımızda ve bu yolla da onların anlamlarından kurtulmak istediğimizde matematiği biçimsel bir şekilde inşa etmeye başlamış oluruz. Hilbert'in sistemi ise, biçimsel bir sistem olarak, önermelerin anlamı hakkındaki tartışmaları sonlandırmaya yöneliktir. Örneğin "T", uygun bir semboller dizisi olarak aynı zamanda ya "t" ya da "t-olmayan" önermelerinden yalnızca birinin ispatlanabildiği bir teoremdir. Bu teoremin bu koşulları sağlayabilmesi için tam ve tutarlı olması gerekir. Ancak Penrose'un da dile getirdiği gibi, matematiğe özünü kazandıran şey anlamdır.[130] Anlam en nihayetinde matematiğin var olanlarla, insan yaşamıyla, insanın gerçekliği ile ilişkisinde ortaya çıkar. Biçimsel bir sistem ise, bir semboller dizisi ve bu dizide yer alan semboller

[127] Davis, **The Universal Computer: The Road From Leibniz to Turing**, s. 99.
[128] Uspensky, "Gödel's incompleteness theorem", s. 269.
[129] Penrose, **Kralın Yeni Usu**, s. 124.
[130] **A.e.**, s. 125.

arası ilişkilerden ibaret bir sistem olarak*, bu anlamdan yoksundur, anlamı dikkate almaz.

Hilbert'in matematik projesi, Kurt Gödel tarafından kimi eleştirilere maruz kalır. Gödel'e göre bizler Hilbert'in düşünceleri hakkında kendimizi yalnızca doğal sayılar kuramı ile sınırlasak bile "A" gibi dizge içerisinde hem "a" hem de "a-olmayan" önermelerinin her zaman türetilebileceği bir aksiyom ve biçimsel kurallar sisteminin inşa edilmesinin olanağını bulamayız.[131] Gödel bu itirazını, tamamlanamazlık kuramları üzerinden güçlendirmektedir. Tamamlanamazlık kuramları, herhangi bir biçimsel dizge içerisinde belli koşullar altında doğru olan ama doğruluğu veya yanlışlığı içinde bulunduğu sistem içerisinde ispatlanamayan önermelerin bulunduğunu ortaya koyar.[132] Gödel'in birinci ve ikinci tamamlanamazlık kuramları, Hilbert'in matematik anlayışına getirilmiş önemli eleştirilerdir. Birinci tamamlanamazlık kuramına göre bir matematiksel sistemin dili içerisinde yine o sistem içerisinde ne ispatlanabilir ne de yanlışlanabilir durumda olan en az bir yargı bulunur.[133] Böyle bir yargı sistem içerisinde karar verilebilir bir durumda değildir.[134] Dolayısıyla birinci tamamlanamazlık kuramına göre, örneğin matematikte sayının inşası söz konusu olduğunda temel bir aritmetik sistemin inşası, biçimsel bir anlayışla mümkün değildir.[135] İkinci tamamlanamazlık kuramı ise bir biçimsel dizgenin, bir matematiksel sistemin tutarlılığının sistemin kendisinin içerisinde sağlanamayacağını, bir sistemin yine kendisine dayanarak kendisinin tutarlılığı hakkında bir şey söylenemeyeceğini ortaya koyar.[136] Dolayısıyla bir biçimsel kuram, bu biçimsel sembol sistemini ögeleri bakımından ne kadar

* Ancak vurgulamalıyız ki Güven'in Hilbert'in biçimselciliğinin salt semboller arası ilişkilerden ibaret bir biçimselcilik olmadığı, Hilbert açısından sezginin de dikkate alındığını vurguladığı görüşleri, matematik felsefesi literatürümüz açısından önemlidir. Ancak Hilbert'in biçimselciliğinde sezginin oynadığı rolü yadsımasak bile, karar verme problemi bakımından Hilbert'in matematik anlayışına ciddi itirazlar söz konusudur. Bkz.: Özgüç Güven, "Hilbert, Matematiğin Temelleri ve Görü", **Felsefe Arkivi**, No: 52, 2020, 113-149.

[131] Kurt Gödel, "The modern development of the Foundations of mathematics in the light of philosophy", **Collected Works: Unpublished essays and lectures**, ed. Solomon Feferman ve diğerleri, C. III, New York-Oxford, Oxford University Press, 1995, s. 383.

[132] Uspensky, "Gödel's incompleteness theorem", s. 241.

[133] Cristian S. Calude, "Incompleteness and the Halting Problem", **Studia Logica**, No: 129, s. 1159.

[134] Franzen, **Gödel's Theorem: An Incomplete Guide to Its Use and Abuse**, s. 3.

[135] Smith, **An Introduction to Gödel's Theorems**, s. 3.

[136] Franzen, **Gödel's Theorem: An Incomplete Guide to Its Use and Abuse**, s. 3.

detaylandırırsak detaylandıralım kendisinin tutarlılığını ispatlayamaz.[137] İkinci tamamlanamazlık kuramının önemi, Hilbert'in sonlucu bir anlayışla inşa etmeye çalıştığı, tamlığı ve tutarlılığı hedefleyen bir matematiksel sistemin temellerine ilişkin ciddi eleştiriler[*] getirmesidir.[138] Çünkü biçimselci bir matematiksel sistemde, matematiksel olgu gibi bir şeyin bulunmadığını, matematik olguları ifade ettiğine inandığımız önermelerin doğruluğunun yalnızca ilgili önermenin hangi koşullar altında doğru olduğunu tanımlayan diğer ifadelerle birlikte bir anlam ifade ettiğini, böylece Gödel'in ifadesiyle *boş bir dil akışının* meydana geldiğini söyleyebiliyoruz. Dolayısıyla biçimsel bir matematik dizge oluşturmak için gereken olanaklı önermelerin bu anlamda içerikten yoksun olduğunu söylemek mümkündür.[139] Buradan da açıktır ki Gödel'in kuramları, tutarlı ve temel bir aritmetik için uygun olan yani doğal sayıları ve toplama veya çarpma gibi operasyonları içeren sistemlerde ispatlanamayan ancak oldukça anlamlı olan, bu sistemin dışında olsa bile doğru olduğunu görebildiğimiz önermelerin bulunduğunu ortaya koymaktadır.[140] Gödel bize matematiksel doğrulukların biçimsel dizgeler içerisine sıkıştırılamayacağını göstermiştir. Yalın bir biçimselciliğin çok ötesinde olan matematiksel doğruluklar, matematik yapmada ve matematiksel bir sistemi inşa etmede sezginin önemini hatırlatır. Böylece yalın biçimsel sistemler kendilerinin tamlığını ve tutarlılığını ortaya koymak veya anlamlı olup olmadıklarını tayin etmek konusunda sezgisel olarak hangi durumların uygun olup olmadığına karar veremez.[141]

Şimdiye kadar çizilmeye çalışılan çerçeve, matematiğin ve matematik felsefesinin temel problemlerini ve bizi yapay zekâ projesinde karşılayacak temel problemleri serimlemeye ilişkin bir çabadır. Ancak okuyucu sorabilir: Matematik

---

[137] Calude, "Incompleteness and the Halting Problem", s. 1159.

[*] Burada okuyucu açısından da akla gelebilir ki, halihazırda bir matematiksel ve mantıksal sistemler olarak yapay zekâlı varlıklar, ikinci tamamlanamazlık teoreminin eleştirilerinin odağında varlıklar olarak, kendileri üzerine yeniden düşünmesi gereken durumlarda, ikinci tamamlanamazlık teoreminin sonuçları dikkate alınırsa, herhangi bir yanıt üretemeyecek böylece bir sonuç üretemediği durumda yanıtsız kalması beklenecektir. Ayrıca yine birinci tamamlanamazlık teoremi sonuçları bakımından dikkate alınırsa, yapay zekâlı varlıkların birbirine zıt iki önermeye eş zamanda sahip olamayacağını da söyleyebilir durumda oluruz.

[138] Franzen, **Gödel's Theorem: An Incomplete Guide to Its Use and Abuse**, s. 39.

[139] Gödel, "Some basic theorems on the foundations of mathematics and their implications", s. 319.

[140] J. R. Lucas, "Minds, Machines and Gödel", **Philosophy**, C. XXXVI, No: 137, 2009, s. 112-113.

[141] Penrose, **Kralın Yeni Usu**, s. 133.

felsefesine veya matematik alanına ilişkin bir problemin, bu çalışmanın yöneldiği problemler bakımından ilintisi nedir? Açıktır ki yapay zekâ çalışmaları ve bilgisayar teknolojisi, matematiğe sıkı sıkıya bağlı gelişmelerdir. Çünkü Gödel, Hilbert'in sınırlı sayıda prosedürle sonuca ulaşmayı da içeren matematik projesinin bir sonucunun, büyük bir gelişme olarak Turing'in sınırlı sayıda parçaya sahip makinesini olanaklı kılan bir girişim olduğunu savunur.[142] Buradan, Gödel'in Hilbert'in sonlucu ve biçimsel matematik anlayışına getirdiği eleştirilerin yapay zekâ ve bilgisayar bilimi bakımından yansılarının olabileceğini söylemenin de mümkün hale geldiğini görüyoruz. Nitekim ünlü makalesinde Turing, tasarladığı makinelerin gücüne ilişkin kimi matematiksel ve mantıksal sınırlılıkların söz konusu olduğunu, bunlardan en iyi bilineninin Gödel'in kuramlarının ortaya koyduğu, doğruluğu veya yanlışlığı ispatlanamayacak kimi mantıksal sistem önermelerinin bulunduğu yollu görüşü dile getirir.[143]

Yapay zekâ projesi içerisinde, matematiksel ve mantıksal sistemlere ilişkin Gödelci itiraza bir yanıt üretme çabası, özellikle bilgisayar bilimcilerin üstlendiği bir uğraş olarak belirir. Çünkü Gödel'in tamamlanamazlık kuramlarının insanlarla makineler arasında bir hiyerarşi oluşturduğuna, böylece makinelerin hiçbir zaman insansı niteliklere sahip olamayacağına ilişkin hatalı bir yorumlama söz konusudur. Örneğin McCarthy, programların insan düzeyinde matematik yapamayacağına ilişkin hatalı bir sezginin olduğunu, bu varsayımın, makinelerin matematiği zorunlulukla tek bir aksiyomatik sistem içerisinde ve önceden tanımlanmış yorumlardan hareketle yapmak zorunda olduğu varsayımı olduğunu dile getirir.[144] Gödel'in tamamlanamazlık kuramlarının makinelere uygulanıyor olması doğaldır çünkü Lucas'a göre makine olmanın özü, makinenin tamamen biçimsel bir sistemin somutlaştırılması, onun gerçekleşmesi olmasıdır.[145] Ancak bu, Gödel'in kuramlarının hangi zekânın *insan zekâsı* veya *başka bir zekâ* olduğunu tayin eden bir turnusol olduğunu göstermez. Söz konusu kuramlar bize yalnızca bir biçimsel sistemin belli

[142] Gödel, "Some basic theorems on the foundations of mathematics and their implications", s. 304-305.
[143] Turing, "Computing Machinery and Intelligence", s. 444.
[144] John McCarthy, "Review of the Emperor's New Mind by Roger Penrose", (Çevrimiçi) http://jmc.stanford.edu/artificial-intelligence/reviews/penrose1.pdf, 7 Ocak 2023.
[145] Lucas, "Minds, Machines and Gödel", s. 113.

kabiliyetlere sahip olup olmadığına ilişkin bir itirazdan daha fazlasını sunmaz. İnsan ve makineler arasındaki karşılaştırmanın kökeninin çok daha eskilere dayandığını biliyor olsak da bu çalışmanın kapsamı ve iddiaları düşünülürse, yapay zekâ projesinin de güncel bir çaba olduğunu söylemek doğaldır. Yapay zekâlar ve insan zekâsı arasında kurulmaya çalışılan eşitlik ilişkisi, matematiksel bir engel olarak Gödel'in kuramlarını aşamaz. Nitekim Canbeyli'nin de dile getirdiği gibi, öğrenen ve temelinde yatan aksiyomatik sistemdeki belirsizliklerin Gödel'in kuramları tarafından ortaya koyulduğu bir durumda bilgisayarlar ve insan arasında yapılmaya çalışılan karşılaştırmaların temelindeki varsayımları sorgulamamız olanaklı kılınmış olur.[146]

Yeni bir metafizik tasarım olarak yapay zekâ projesinin, insanın dünyadaki yerini, insanın failsi özelliğini ve insanın biricikliğini, onun var oluşunun koşullarını tanımazcasına ve bu çabada rol oynayan felsefe ve diğer bilgi alanlarındaki tüm ciddi tavırları spekülatif ilan ederek büyüdüğünü iddia ettiğimizde, böyle bir çabanın en önemli örneklerinin bilgisayar bilimciler tarafından ortaya koyulduğunu görebilir hale geliyoruz. Örneğin Say, yapay zekâlar, makineler ve insanlar arasındaki ilişki üzerine yapılan tartışmaların, bu var olanlar arasında kurulmak istenen bir hiyerarşiye ilişkin kaygıları taşıdığını düşünüyor olsa gerek ki[*] makinelerin asla insanlar gibi düşünemeyeceği gibi bir sava itiraz eder, Gödel'in tamamlanamazlık kuramlarının burada öne sürüldüğünü belirler, böylece yaratıcılık hususunda insanlarla yarışamayacakları söylenen makinelerin şimdi insanlarla matematik alanında da yarışamayacağı iddiasının bulunduğunu dile getirir.[147] Gödel'in kuramlarının insanların makinelerden üstün olduklarını gösteren bir kuram olduğu

[146] Reşit Canbeyli, "Psikoloji Açısından Beyin-Bilgisayar Karşılaştırması", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 116.

[*] Say burada kısmen haksız sayılmaz. Çünkü yapay zekâ felsefesi, tarih boyunca var olan insansı yaratıklar ve insan arasında kurulmuş olan ilişkilerin de bir sonucu olarak insansı niteliklere sahip var olanlarla insanlar arasındaki varlık farkları üzerine yapılmış tartışmalarla doludur. Bu durum bir yanıyla da bu daha düşük kademeden varlıkların hangi koşulları sağladıklarında insansı olacakları, bu koşulları sağlamaları için hangi bilimsel ve teknik unsurların yerine getirilmesi gerektiği gibi tartışmaları da barındırır. Ancak tartışmanın bu özelliklerinin, yapay zekâ felsefesinin felsefece icrası mümkün olacaksa, yapay zekâ felsefesinden arındırılması gerektiğini ve yapay zekâ felsefesinin modern bir çizgiye oturtulması gerektiğini bu çalışmanın seyrinde açıkça ve bütünlüklü olarak ortaya koyuyoruz.

[147] Cem Say, "Yapay Zekâ ve Gödel Teoremi", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 254.

yollu hatalı yorumlamayı kabul eden Say, bu fikre sahip olanların iddialarını şu şekilde özetler:

> "Bakın, çelişkisiz biçimsel sistemlerde belitlerden çıkarım kurallarıyla 'mekanik' bir şekilde elde edilemeyen, ama biz insanların doğru olduğunu şıp diye görebildiğimiz önermeler var. Demek ki bizler çelişkisiz biçimsel sistemler değiliz. İnsan matematikçiler çelişkisiz olduklarına göre demek ki insanlar biçimsel sistem değildir. Bilgisayarlar ne yaparlarsa yapsınlar, o yetersiz 'mekanik' şekilde yapacaklardır. Demek ki her zaman bilgisayarların bilemeyeceği ama insanların görebileceği kimi önermeler olacaktır."[148]

Say özetlediği varsayımda iki yanlış basamağın olduğunu söyler, insanların çelişkisiz oldukları yollu savın kökenini anlamadığını, örneğin matematikçilerin hata yapabildiğini savunur, bilgisayarların ise uslamlama kuralları önceden belirlenmiş, asla şaşmayan ve çelişkisiz sistemleri olduğu yollu görüşü dile getirir. Buradan hareketle de insan beyni gibi sonlu sayıda parçacığın birbirleriyle fizik ilkelerine göre etkileşmeleri sonucu ortaya çıkan, örneğin bir matematikçinin zihninde gerçekleşen matematiksel işlemler, sonuçlardan hareketle beynin yapısı yeterli incelenir ve bu yapıda işleyen süreçler bir bilgisayar programına hesaplatılabilirse bilgisayarın da matematik yapabileceği, bu durumda matematik kurallarını işleten bir zihin olarak insan zihni ile makine zihni arasında herhangi bir farkın bulunmayacağını iddia eder.[149]

Say'ın iddiaları iki adımda incelenebilir. İlkin, insanların hata yapmayan varlıklar oldukları savunulmamaktadır. Hatta insanlar çelişkiler üretebilen varlıklardır da. Ancak makinelerdeki çelişkiler ile insanların içine düştükleri çelişkiler arasında bir nitelik farkı bulunur. Örneğin bir insan özdeşlik veya çelişmezlik ilkesini ihlal edebilir. Ancak insanın içinde bulunduğu mantıksal çelişkiler, insanın özünde değil, insanın yöneldiği şeyle insanın o andaki zihinsel durumu arasındaki ilişkide gerçekleşir. Böylece bir insan hiçbir zaman bir makine gibi yanıt üretemeyeceği sonsuz bir döngünün içine giremez ya da bir hata mesajı iletip o anda doğumundan beri sürdürmekte olduğu var oluşundan, kimliğinden, içinde bulunduğu zamandan ve deneyimlerinden kopuk bir biçimde kendini yeniden

[148] **A.e.**, s. 263.
[149] **A.e.**, s. 264.

başlatmaya çalışmaz. İnsanın içine düştüğü hatalar mantık ilkelerine tabi olmamasından kaynaklanmaz, *mantık ilkelerini içinde bulunduğu durumda işletememesinden kaynaklanır*. Örneğin her insan özdeşlik veya çelişmezlik ilkesinin ne olduğunu sezgisel olarak bilir. Ancak bir insana özdeşlik veya çelişmezlik ilkesinin ne olduğu sorulursa bu soruya bu kavramların mantık bilimindeki tanımları türünden bir yanıt alınamayabilir. Bu durum insanlarda mantık ilkelerinin işlemediğini ya da insanların mantık ilkelerini aşabildiğini göstermez, insanların mantık ilkelerini ayrık birer bilgi olarak zihinlerinde tutmadıklarını gösterir. Böylece örneğin zihin felsefesinde yer alan benlik problemi düşünülürse, insanlar hiçbir zaman beni duyumsadığımı, duyumsadığımı, duyumsadığımı, duyumsadığımı… duyumsarım türünden sonsuz bir döngünün içine girmez. Sonsuz bir düşünmeden ziyade ben varım, benliğim vardır derler. Çünkü insanların yaptıkları hatalar, içinde düştükleri çelişkiler vb. türünden durumlar, yani tutarsızlıklarımız, yukarıda da dile getirdiğimiz gibi bir durumun yerine getirilmesinde ve o zaman ile mekânda kendini gösterir, yalındır.[150] Ancak bir makinenin – burada ilkece yapay zekâ kavramı da kullanılabilir ama makine ve yapay zekâ gibi kavramlar artık iç içe geçmiştir – yaşadığı çelişki insanlardaki gibi ilineksel değil, özüne ilişkindir. Makinenin çelişkisi bizzat onu o yapan her ne ise orada gerçekleşir. Bir makine, içine düştüğü çelişkide bir yanıt üretemeyecek duruma gelir. Bu durumda makine, yapmakta olduğu şeyi sonlandıramaz dolayısıyla bir program olmanın koşullarından biri olarak sonsuz döngülerden kaçınmak durumundadır. Böylece bir matematikçi hata yapar, ancak matematikçinin hatası matematikçiyi o matematikçi yapan zihin durumlarına ilişkin değil, matematik bilgisinde ve matematik yaparken, onu icra ederken var olan bir hatadır. Oysa bir makinenin hatası, bizzat makineyi o makine yapan şey ne ise – bilgisayar bilimciler bunun program olduğunu çok açık bir şekilde dile getiriyorlar – bizzat orada gerçekleşir. İnsan zihni bu anlamda makine zihninden üstündür veya aşağıdır gibi hiyerarşik bir yaklaşımı kabul etmediğimizi söylemeliyiz. Ancak biz makine zekâsı ile insan zekâsı arasındaki varlık farklarını dikkate alıyoruz. Makine zekâsının mutlak, belirlenmiş bir zekâ olarak insan zekâsının baş edebildiği sonsuz döngülerle ve belirsiz durumlarla mücadele edemeyeceğini söyleyen ve bu durumda

[150] Lucas, "Minds, Machines and Gödel" s. 121.

hiçbir makinenin insan zihni için uygun ve tam bir model olamayacağını iddia eden, bunun sebebinin de onların özsel olarak farklı olduğunu savunan Lucas'ın düşüncelerini vurguluyoruz.[151]

Say'ın ikinci iddiası için bilgisayarların varlık koşullarını incelememiz gerekir. Bir bilgisayar, bir matematiksel ve mantıksal sistem olarak, insan zekâsının ve insan beyninin kimi işlevlerini başarıyla yerine getirebilir. Çünkü insan zihninin hesaplamayla ve hesaplamayla bağlantılı bazı yetileri çok daha güçlü ve isabetli araçlar olarak bilgisayarlar üzerinde gerçekleştirilebilir.[152] Ancak, bilgisayarların insan zihninin bir program üzerinde, o zihni oluşturan parçacıkların davranışları da dahil olmak üzere o zihinde gerçekleşmekte olan fenomenal yaşantıyı gerçekleştirebileceğini savlamak için elimizde yeterince temellendirme bulunmaz. Burada Say, bu iddiasını güçlü kılmak için iki soruya yanıt vermelidir: Makine zekâsı ile insan zekâsı aynı gerçekliği mi deneyimlemektedir? Katı bir şekilde matematik ve mantık ilkeleri üzerinde gerçekleştirilen makine zekâsının buluşçu, yaratıcı unsurları nasıl sağlanacaktır?

Öncelikle makine zekâsı ve insan zekâsının var olduğu gerçekliği incelemek gerekiyor. İnsan zekâsı, *bu dünya üzerinde* ve *bu fizik evrende* vardır. Bu zekâ, duyargaları vasıtasıyla duyumsar ve bu duyu verilerini işler. Böylece içinde yaşadığı gerçeklikle ilgili kimi yargılara sahip olur. Ancak insan zekâsı, içinde yaşadığı evrenin sınırlılıklarına iye bir zekâ olarak örneğin matematik evrenler gibi *sanal* bir evrende bulunmaz. Dolayısıyla *bu fizik evrende* ve *bu yaşantıda*, *bu fizik evren ve yaşantı içerisinde* bulunan fizik nesnelerle ve ilişkide bulunan bir fail olarak bizim olanaklılıklarımız bir düzenliliğe sahiptir. Örneğin hiçbir zaman gökten bir deve inemez, bir insan Dünya'nın yer çekimine tabi olan bir canlı olarak 92 metre zıplayamaz veya çok daha temel olarak, bir tepkimeye girenlerle bir tepkimeden çıkanlar arasındaki eşitlik asla bozulamaz. Ancak makine zekâsının içerisinde var olduğu matematik evren böyle değildir.

Bu çalışma içerisinde yapay zekâ felsefesinde yer alan kimi antropolojik yüklere yer vermiş olmamızın ve bu yükleri dayanakları bakımından incelemiş

[151] **A.e.**, s. 113-114.
[152] **A.e.**, s. 115.

olmamızın bir önemi de kendisini burada göstermektedir. Yapay zekâlı makinelere insansı özellikler atfetme eğilimimiz bizi, herhangi bir yapay zekânın da bizim gibi şu anda ve burada olan bir kırmızı elmayı deneyimlediğini düşünmemiz gerektiği inancına sevk eder. Ancak yanılıyoruz. Yapay zekâlı bir varlık, metalik bir bedenle gelip, şu anda ve burada bulunan kırmızı bir elmayı avuçlayabilir, hatta ona gereken özeni göstererek çok da fazla sıkmayabilir, böylece o kırmızı elmayı herhangi bir zarar vermeden örneğin bir B noktasına doğru götürüp bırakabilir. Biz bu yapay zekâlı varlığın elmayı algıladığını, onunla bir temas kurduğunu, bu temasın derecesine ilişkin bir kaygısının bulunduğunu, elmanın fizik yasalarına göre davranışlarını tıpkı bizim bir içeceği dökmeden taşımak için akışkan dinamiklerini hesaba katıyor olduğumuz gibi dikkate aldığını düşünüyor olabiliriz ancak burada Lem'e göre tüm bunlar şu anda ve burada bu metni okumakta olan okuyucular da dahil olmak üzere insanların içinde yaşadıkları, bakmakta ve duyumsamakta oldukları *bu gerçekliğin* kendisinde değil, aksine, *bir matematik evren, bir matematik sistem* içerisinde gerçekleşmektedir. O yapay zekâ veya herhangi bir yapay zekâ, hiçbir zaman içinde bulunduğu matematik evrenden dışarı çıkmamıştır ve çıkamaz da.[153] Elmanın kendisi, elmanın rengi, elmanın bulunduğu mekân ve elmanın zamansallığı vb. gibi bizce dışsal olan ama bir yapay zekâ için de dışsal olması gerektiğini düşündüğümüz her şey, bir yapay zekâ bakımından matematiksel ifadelerden, formüllerden ve matematik evrende yeniden tasarlanan ve sayılarla ifade edilen – çünkü analitik geometri ile şekillere de gerek kalmadı – bir varlık tarzından başka bir şey değildir. Makinenin kırmızı elmayı t miktarda zaman boyunca elinde tuttuğunu söylediğimizde makineye dışsal bir var oluşa iyeymiş gibi yaklaşırız. Oysa olan şey makine açısından yalnızca sayısal süreçlerdir. Makine için elma yoktur, mekân yoktur, bizim zamanımız orada geçerli değildir. Makine kuru, sabit, sonsuz boyutlu, kesin, belli aksiyomlara dayanan, içerisinde inşa edilebilecek şeylerin olanağı bakımından da sonsuz olanaklara sahip bir matematik evrende yaşar. Bu makineler matematik evrenin içinde varlar ve onların havası, dünyası, bulutları, suyu ve hatta *o kırmızı elmaları* bile birer matematiksel ifadedir. Dolayısıyla bu evrenden dışarı çıkamazlar, çıkamadıkları gibi, bizim fizik evrenimizden de bu evrene fizik

[153] Lem, "Non Serviam", s. 289.

nesnelerin ve fizik nesneler arası ilişkilerin *kendinde şeyler olarak* aktarılması olanaklı değildir.[154] Bu durumda makinenin matematik evrende yaşayan bir varlık olarak fizik evrenimizle ve bizim gerçekliğimizle ilgili olguları anladığını söylemek ya da Say'ın bilgisayarların insan zihninden aşağı kalır bir yanının olmadığını iddia ettiği savı savunmak için elimizde metafizik kabullerimizden başka bir neden bulunmaz.

Tüm bu sonuçlarla beraber okuyucu açısından açık kılınmıştır ki matematik, köklü bir bilim olarak, evrene yönelişimizde güçlü bir dayanaktır. Ancak matematik de – tıpkı ileride göreceğimiz diğer bilim alanları gibi – sorunlar taşır. Matematiğin alanının bilgisayar bilimcilerin ve yapay zekâ projesinde rol alan diğer bilgi alanlarında çalışan araştırmacıların yapmaya çalıştığı gibi, sorunlarını görmezden gelerek genişletilmeye çalışılması oldukça sorunludur. Gödel'in tamamlanamazlık kuramlarının, bilgisayar bilimlerine gücünü veren Hilbertçi bir matematik anlayışına getirdiği eleştirilerin önemi açıktır: Matematik önermeler uzayda ve zamanda var olan fiziksel gerçeklik hakkında bir şey söylemez. Çünkü bu önermeler gerçek var olanların bulunduğu bu evrenden bağımsız olarak, terimlerin anlamlarından ötürü doğruluk değerlerini alan bir yapıdadırlar. Bu vesileyle Gödelci itiraz matematiği yıkmaz, matematiğin güvenilirliğini, onun gücünü ve nesnelliğini yadsımaz. Tam aksine onu kimi sembollerin ilişkileri olmaktan kurtarıp yeniden Platoncu[*] bir çizgiye oturtur. Gödel bize matematik kavramların bizim yaratamayacağımız veya değiştiremeyeceğimiz, sadece algılayıp tanımlayabileceğimiz, kendilerine ait nesnel gerçeklikler oluşturduklarını söylemektedir.[155] Burada şimdiye kadar matematiğin yapay zekâ felsefesine temel bir bilim olarak hatalı kullanılışlarına ilişkin sorunlara dikkat çekmeye çalıştık. Böylece önemli bir yere sahip olan matematik biliminin, felsefi işlevselci savlar bakımından oynadığı rol, sorunu betimleme ve yanıtlar üretme çabamız bakımından da okuyucu için açık kılınmıştır.

---

[154] **A.e.**, s. 290-292.

[*] Matematik felsefesi bakımından Platoncu bir çizgide bulunmasak da Kurt Gödel'in matematiğin biçimselleştirilmesi girişimlerine karşı ortaya koyduğu tavrı, matematik felsefesi bakımından da olması gereken olarak niteliyor, böylece bu tavırla beraber matematikle ilgili tartışmaların yeniden felsefeleşerek, felsefeye dahil edilerek sürüdürüleceği bir zeminin olanağının açıldığını kabul ediyoruz.

[155] Gödel, "Some basic theorems on the foundations of mathematics and their implications", s. 320.

1.2.2. Yapay Zekâ Çalışmalarının Mantık Bilimine Dayanan Temelleri

"Cutie: ... Doğru bir mantık zinciri beni gerçeğe ulaştıracaktır."[156]

Evreni veya çevremizde olup bitenleri incelemek istediğimizde, fizik nesneler olarak var olanların kimi düzenliliklere iye olduğunu görürüz. Yer çekimi, Doppler etkisi vb. gibi birçok düzenlilik fizik nesneler üzerinde etkir. Doğa yasası diyebileceğimiz kimi düzenlilikler bulunduğu ve bunların fizik nesneler üzerinde etkidiği bilinir ancak insanın düşünmesinin ve düşüncelerinin de ona tabi olduğu kimi ilkelerin bulunması gerektiğini düşünemez miyiz? Bir bilim olarak mantık, düşünmenin ve düşünceler arası ilişkilerin ilkeleri olarak mantık ilkelerini soruşturan ve bu ilkelerin akıl yürütmelerimizde işletilmesini inceleyen bir alandır.[157] Ancak mantığın ihtiyaç duyduğu zemin, kesin ve anlamlı cümlelerden oluşan bir dildir çünkü böyle bir zemin bulunamazsa mantık ilkelerinin ve bu ilkelerin etkidiği önermelerin doğrulukları veya yanlışlıkları hakkında karar verilmesi olanaklı olamaz. Bu olanak sağlandığı sürece mantık, kesinlikli bir alan olarak, yargılarımızın güvencesini sağlayabilir, böylece bir dil içerisinde dile getirilen ve mantığın konusu olabilen yargılar olarak önermelerin tamamı üzerinde söz sahibi olabilir.[158] Dolayısıyla Boole'un da dile getirdiği gibi mantığın dilinin, insanın usu için bir araç veya enstrüman olduğunu söylemek oldukça doğaldır.[159] Ancak mantığın dilinin, sembolik-simgesel mahiyeti unutulmamalıdır. Mantığın, var olanlarla ilişkimizde meydana çıkan önermeler üzerindeki gücü yadsınamaz. Ancak mantığın odağı, önermelerimizi içerikleri bakımından ele almak ve içerikleri üzerinde kimi sonuçlara varmak değildir. Tam aksine mantık, önermelerin içerikleri olarak var olanlarla

[156] Isaac Asimov, **Ben, Robot**, çev. Ekin Odabaş, 4. bs., İstanbul, İthaki Yayınları, 2018, s. 60.
[157] Şafak Ural, **Temel Mantık**, 4. bs., İstanbul Çantay Kitabevi, 2017, s. 2.
[158] Russell, **İnsanlık için Yapay Zekâ: Yapay Zekâ ve Kontrol Problemi**, s. 309-310.
[159] George Boole, **An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities**, Cambridge, Macmillan and Co., 1854, s. 24.

ilişkisiyle ilgilenmez, yalnızca düşünmenin formlarını veya bu formların ifade edilme biçimlerini ele alan bir bilim olarak iş görür.[160]

Anlaşılacağı üzere mantık hem bir yöntem hem de bir bilim olarak görülebilir. Yöntem olarak mantığın, doğru düşünmenin – yani geçerli akıl yürütmeler yapmanın – bir yöntemi olduğu söylenebilir. Öte yandan bir bilim olarak mantık, yöntem olarak görülebilen mantığın bilimidir. Burada bu yöntemde yer alabilecek ve akıl yürütmelerin geçerliliğini sağlayan kurallar incelenir, metodolojik bir tavır ortaya konur. Ancak unutulmamalıdır ki mantıkta incelenen önermeler olarak yargılar, gerçeklik hakkındaki iddialarımız olduğu gibi gerçekliğin kendisi de yargılarımızdan farklı olarak, var olanı yani somutu ve soyut var olanları imler.[161] Buradan mantığın direkt olarak gerçeklikle ilgili değil ama gerçeklikle ilgili yargılarımızla ilgilenen bir bilim olduğunu görebiliyoruz. Ancak yargılarımızı incelerken bir araç olarak mantığın, yargılarımızın doğruluklarını gösterebilmesi ve bu vesileyle de önermelerin kabul edilmesini haklı gösteren kimi gerekçeleri ortaya koyabilmesi için söz konusu yargıyı başka yargılara veya bu yargıların uydukları düzenlilikleri, yani mantık ilkelerini, başka ilkelere dayandırabilmesi gerekmektedir.[162] Ancak biliyoruz ki matematik alanında olduğu gibi, mantık alanı da kimi aksiyomlara dayanır. Aksiyomlar ve mantık ilkeleri, doğrulukları dolaysızca kabul edilen ögelerdir.[163] Dolayısıyla doğru düşünmenin yöntemi veya doğru düşünmenin kurallarını içeren bir bilim olarak mantık, düşünmeyi – yani akıl yürütmeyi – ele alır. Böylece akıl yürütmede dile getirilen bir öncül veya öncüllerden sonuç mahiyetinde bir yargı elde edilir. Akıl yürütmenin doğru olması ise, öncüllerin doğru olması durumunda sonucun da doğru olmasının bir sonucudur.[164]

Akıl yürütme bir düşünme sürecidir ve bir düşünme süreci, düşünceden ayrık olarak ele alınamaz. Düşünme ve düşünce arasındaki ilişki, dil aracılığı ile dışsallaşır. Düşünülmüş şey olarak düşünce ile dile getirilmiş şey olarak dilsel ifade arasında tam bir örtüşme veya çakışma bulunmaz. İşte düşünülmüş şey olarak

[160] Işıklı, "Fuzzy Mantık ve Sibernetik'in Siber Toplum ve Yapay Zekâ Üzerindeki Etkileri", s. 30.
[161] Teo Grünberg, **Felsefe ve Felsefi Mantık Yazıları**, İstanbul, YKY, 2005, s. 367-368.
[162] Teo Grünberg, **Sembolik Mantık El Kitabı: Temel Mantık**, C. I, 2. bs., Ankara, ODTÜ Yayıncılık, 2017, s. 1.
[163] Işıklı, "Fuzzy Mantık ve Sibernetik'in Siber Toplum ve Yapay Zekâ Üzerindeki Etkileri", s. 34.
[164] Grünberg**, Sembolik Mantık El Kitabı: Temel Mantık**, s. 1.

düşünce ile onun dışsallaşmasını sağlayan dil arasındaki ilişkide mantığın oynadığı rolün, yalnızca düzgün düşünmenin veya mantıksal düşünme denen bir tür düşünme tarzının ilke ve formlarının incelendiği bir çabadan daha fazlası olmadığını söylememiz gerekir. Çünkü düşünme ve düşünülen şey olarak düşünce farklılık gösterir. "Elma kırmızıdır." cümlesi bir önermedir ve düşünülmüş bir şey olarak bir düşüncedir, dil ile dışsallaşır. Öte yandan "Hava ne kadar da sıcak!" cümlesi de bir tür düşüncedir. Ancak mantık, her türden düşünme üzerinde ve bu her türden düşünme sonucunda ortaya çıkan her türden düşünce ile ilgilenmez, bazı yargılar mantığın alanında bulunmaz. Nitekim algılama, hayal kurma, anımsama, sezgi, tasarlama, hesap yapma vb. gibi birçok kipe sahip olan düşünme, kiplerinin yalnızca bazıları bakımından mantığın konusu olarak ele alınır. Çünkü mantığın ilgisi, doğrulanabilir veya yanlışlanabilir ve bu değerleri alabilir bir ürün olarak, düşüncenin ve bu düşünceler arasındaki ilişkinin ilkeleri üzerinde şekillenir. Dolayısıyla mantık, bir olgu olarak ne düşüncenin kendisini ele alır ne de mantıksal düşünmenin fizyolojisi veya beyindeki nedenleri ve işlevleri gibi konularla ilgilenir. Mantık yalnızca düşünmenin biçimi ile ilgilenir.[165] O halde bir bilim olarak mantığın konusu, yöneldiği çıkarımları herhangi bir akıl yürütmeden veya üzerinde durduğu yargıları herhangi bir yargıdan ayıran kimi ayrımları şart koşuyorsa, bunun "objektif olarak denetlenebilme" gibi bir nitelik olması gerektiği açıktır.[166] "Elma kırmızıdır." önermesi, doğruluk değeri bakımından doğru veya yanlış değer alabilen, denetlenebilen bir önermedir. Elma kırmızı değil ama yeşil olabilir. Böylece elmanın kırmızı olduğu yollu önerme yanlıştır denebilir. Ancak havanın ne kadar da sıcak olduğunu dile getiren bir ifade doğrulanamaz veya yanlışlanamaz. O yalnızca fenomenal bir deneyimi dile getiren bir düşüncedir, fazlası değil. Dolayısıyla mantığın konusu olabilecek önerme ve akıl yürütmelerle, mantığın konusu olamayacak olanlar arasındaki ayrım, bir biçimsel dil içerisinde bu önerme ve akıl yürütmelerin ifade edilebiliyor olmaları ve bu ifade edilenlerin denetlenebilir bir yapıda olmalarıyla belirlenebilir. Nitekim mantıksal olarak denetlenemeyen çıkarımlar saçma olarak da tanımlanamaz, ancak denetlenemeyen bu çıkarımlar nesnel bir incelemenin konusu olamadıkları için de herkesin aynı sonucu elde

[165] Doğan Özlem, **Mantık**, 7. bs., İstanbul, İnkılap Kitabevi, 2004, s. 28-29.
[166] Ural, **Temel Mantık**, s. 118.

edemediği kimi durumlar olabilir. Burada herkesin üzerinde mutabık kalabildiği sonuçların üretilebileceği bir dilin gereksindiği düşünülebilir, nitekim mantık bilimi tıpkı matematik bilimi gibi gücünü sıkı bir biçimsel dilden alır. Önermelerimiz biçimsel bir dilde ifade edilir, böylece mantıksal ilkeler bu biçimsel ifadeler üzerinde işletilir. Çıkarım, bu yolla mantığın güvencesi altına alınır. Biçimsel dilin yapısı ise çıkarımın türünü belirleyen önemli etmenlerden biridir.[167]

Mantığa gücünü kazandıran biçimselliğin, yani biçimsel bir dilin, kurulabilmesi için iki koşul sağlanmalıdır: Akıl yürütmelerin ifade edilebileceği sembolik ögeler ve sembollerin nasıl kullanılacağını belirleyen ilkeler. Bu iki koşul sağlandığında, mantıksal her türden ilişki bu dil içerisinde tıpkı matematikte olduğu gibi denetlenebilir ve incelenebilir olur. Mantık içerisindeki ilişkilerin, çıkarım türlerinin vb. ifade edilebilmesine imkân veren dillerden biri Boole cebridir. Önermelerdeki terimlerin, kaplamları açısından düşünüldüğü Boole cebrinde her terim, bu terimin ifade ettiği nesneler sınıfına işaret etmektedir.[168] Boole cebri, Aristoteles mantığının birçok sorununu aşmak bakımından – örneğin dile getirilen önermelerdeki anlam bulanıklıkları – önemlidir. Ünlü metninde Boole, soruşturmasının dilin kuruluşu ve yapısı üzerinde şekilleneceğini, şeylerin ve onların ilişkilerinin bir temsili olarak sembolleri ve onlar üzerinde etkiyen yasaları soruşturmanın ise bilimlerin işi olduğunu, bu yasa ve semboller hakkında sahip olunacak bilginin, sembollerin insanın anlama yetisinin kavramları ve operasyonlarının bir temsili olması bakımından önemli olduğunu iddia etmiştir.[169] Ona göre bir dil, bir işaretler yani semboller sistemidir, dolayısıyla kullanıyor olduğumuz kelimeler de birer işarettir. Bir işaret olarak kelimeler, kimi zaman bir şeyleri temsil ederler, bazen daha basit kavramlar olarak bir araya gelip karmaşık kavramları oluştururlar ya da fenomenal deneyimlerimiz hakkında – örneğin arzu, acı vb. – kimi durumları dile getirirler böylece algılayan zihnin kimi duygulanımları gibi birçok durum kelimelerle karşılanır. Ancak Boole'e göre kullandığımız semboller ve işaretler yalnızca kelimeler değildir. Sembol – Bir sembolün ne olacağı ilinekseldir.

[167] **A.e.**, s. 118.
[168] **A.e.**, s. 119-124.
[169] Boole, **An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities**, s. 24.

Örneğin x, [, * vb. gibi semboller kullanılabilir. – kavramı şeylerin sayılarını yani niceliklerini ifade etmekten çok onlar arasındaki operasyonları ve ilişkileri temsil etmektedir. Böylece semboller, bir akıl yürütme sürecinde zihindeki kavramların ve operasyonların yerini doldurmaktadır ancak yeri doldurulan kavramlar ve operasyonlar da bir şeyleri temsil etmektedir. Dolayısıyla sembollerin de örtük olarak yerine durdukları kavram ve operasyonların temsil ettikleri şeyleri temsil ettiklerini, bu vesileyle de zihinde yer alan kavram ve operasyonların tabi oldukları yasalara, sembollerin de iye olduklarını söylemek mümkün olmaktadır.[170] Düşüncenin yasalarının semboller için de geçerli olduğunu savunan Boole, her türden özel ismin veya tanımın ona uygulanacağı şeyler olarak – örneğin beyaz ve boynuzlu olma tanımlarının beraberce uygulanabildiği bir yığın olarak koyun – "sınıf" kavramını ortaya atmakta, böylece bir sınıfa ait olmayı 1, ait olmamayı ise 0 değeri[*] üzerinden tanımlamayı önermektedir.[171] Ancak unutulmamalıdır ki mantığı içerikli kılmak ve mantığı içerikle ilgili bir alan gibi kabul etme girişimleri, mantığın içeriksiz, biçimsel bir dil olması sebebiyle olanaklı değildir. Mantık tıpkı matematik gibi, çıkarımların geçerli olup olmadığına karar verebileceğimiz bir alanı bize sunmaktadır.[172]

Matematiğin inşasında, matematik aksiyomların ve kavramların dayandırıldıkları ontolojik ve epistemolojik dayanaklardan bahsedilmiştir. Benzer şekilde mantık bilimi ve mantık alanının inşa sürecinde de benzer dayanakların olduğunu görmek mümkündür. Çünkü mantık ilkelerinin kökeni, yani bu ilkelerin dayandıkları temeller, bulunmak durumundadır. Örneğin dış dünya hakkındaki deneyimlerimizden elde ettiğimiz ilkeler olarak mantık ilkeleriyle, halihazırda zihnimizde bulunan, doğuştan elde ettiğimiz, **a priori** ilkeler olarak mantık ilkeleri iki farklı temellendirmeyi bize sunmaktadır. Nitekim mantık ilkelerinin yalnızca bize ait, öznel düşünme ilkeleri olup olmadığı ya da bu ilkelerin aynı zamanda birer ontolojik yasalar yani gerçeklik yasaları olup olmadığı konusu da ayrı bir

[170] **A.e.**, s. 25-26.
[*] Çünkü hiç kimse evrensel küme gibi her şeyi içeren bir kümeye, o kümenin içermediği başka bir şey ile göndermede bulunamazken yine hiç kimse evrensel küme dışında kalan herhangi bir şeye referans veremez. Bu sebeple $x^2 = x$ eşitliğini sağlayabilecek değerler düşünülürse – çünkü Boole'e göre x^2 ile karşılanan şey bir şeyin kendisine iki kere referans vermektir örneğin koyun koyun demekle koyun demek eş anlamlıdır – bunların matematikte 0 ve 1 sayıları olduğu görülür. Kurulan önermelerin doğruluk değerleri de 0 ve 1 sayıları ile karşılanmış olur. Bkz.: **A.e.**, s. 47-48.
[171] **A.e.**, s. 28-41.
[172] Ural, **Temel Mantık**, s. 154.

tartışmadır.[173] Biz bu çalışmanın kapsamı içerisinde mantık ilkeleri ile ontoloji ve epistemoloji arasındaki ilişkiyi kapsayan tartışmaları çözümleme ve buradaki kimi görüşleri ele alma çabasını taşımıyoruz. Ancak çalışmanın bu bölümünde, yapay zekâ kavramı ve mantık arasındaki ilişkinin açığa çıkarılması uğruna görülecektir ki mantık biliminin rolü, ontoloji ve epistemoloji alanları karşısındaki konumu ve bu alanlarla mantık biliminin ilişkisi, yapay zekâ çalışmalarında oldukça yanlış anlaşılmıştır. Dolayısıyla biz bu bilim ve alanlar arasındaki ilişkinin tahlili bakımından bu tartışmaları aktarmak istiyoruz. Çünkü yapay zekâ projesinin bir kavramı olarak yaratılmaya çalışılan yapay zekâ kavramı, mantığın rasyonalist kabullerine dayanır. Zihnin yapısıyla varlığın yapısı arasında tam bir uygunluk varsayan ve bu yolla da mantıksal olanın reel olduğunu kabul eden rasyonalizme dayanan yapay zekâ projesi, bu yolla mantık ve varlığı, mantık ilkelerini ve varlık alanını, yani gerçekliği, uzlaştırır böylece mantık ilkeleri üzerinden gerçeklikle bir bağ kurulmuş olur. Nitekim mantık ilkelerinin kullanılmadığı bir ontoloji ve epistemoloji söz konusu olamaz. Bu çalışmada mantığa dayanmayan bir ontoloji ve epistemolojinin mevcut olduğu veya olacağı savunulmamaktadır. Ancak unutulmamalıdır ki burada ifade edilen durum, mantık ilkelerinin ontolojik yasalar olduğuna ilişkin görüşün mantığın veya mantıkçının iddia edeceği veya bu alanların konusu olan bir çaba olduğu değil tam aksine bir varlık felsefecisinin veya alan olarak ontolojinin sorunu olduğunu görmemiz gerektiğidir. Gerçeklik tasarımlarımız mantığa uyabilir, mantığa uygun olarak kurulabilir. Tüm insanlar ölümlüdür ve insan hem ölümlü hem de ölümsüz olamaz, bu doğru. Ancak bu durum, gerçekliğin mantıkça belirlendiği anlamına da gelmez. Bu durumda mantıkta ontolojik kabuller görmek veya mantık ilkelerinin ontolojik yasalar olduğunu varsaymak için bir gerekçemiz bulunmaz. Epistemoloji için de durum benzerdir. Gerçeklik hakkında bilgi elde etme sürecimizde mantıktan istifade etmememiz mümkün değildir ancak gerçeklik hakkındaki çabamızda mantığın kullanılması bir zorunluluk değildir nitekim insanlar hayal edebilirler, uydurma veya kurgusal, gerçekdışı olan şeylerle ilgili de mantıksal düşünebilirler.[174] Görülebilir ki mantık bilimi ile ontoloji ve epistemoloji alanları arasında bir ayrım vardır. Mantık ilkelerinin ontolojik ve

[173] Özlem, **Mantık**, s. 56.
[174] **A.e.**, s. 57-60.

epistemolojik yasa ve ilkeler oldukları yollu kabul, mantık biliminin sınırlılıklarını aşmaktadır. Düşünmenin ilkelerinin işletilmesi ve bu ilkelerin bilimi olarak mantık ile bu ilkelerin birer ontolojik veya epistemolojik ilke olduğunu söyleyen çaba iki ayrı şeydir.

Yapay zekâ projesi içerisinde mantığa kazandırılmaya çalışılan rol, bu proje içerisindeki metafizik kabullere ilişkindir ve düalizmden kurtulan ancak düalizmdeki metafizik kabullerden kurtulamayan bu girişim, tıpkı matematikte olduğu gibi, mantıkla ve mantıkça da izah edilmiş *bir düalist ruhun,* yani yapay zekânın, gerçekleştirilmesi uğruna bilimleri, bilimlerin problemlerini görmezden gelircesine ve hatta metafizik kabulleri uğruna onların *varlık koşullarını yıkarcasına* yoluna devam etmektedir. Okuyucu açısından da açık kılınmalıdır ki, matematiksel ve mantıksal bir biçimcilikle inşa edilen yapay zekâlar ne matematik ne de mantık bilimlerinin olanaklılıkları bakımından gerçeklikle bir ilişki içerisinde olabilirler. Yapay zekâ projesi içerisinde, yapay zekâlı varlıkların böyle bir olanağı matematik ve mantık vasıtasıyla elde edebileceğini düşünen düşünürler ve bilim insanları bulunur. Ve bu tavır, matematik ve mantık bilimlerinin tanımlarını ve sınırlılıklarını değiştirircesine bir çabanın neticesinde mümkündür.[175] Ancak bu düşünürler, gerçekliğin ne mantığın ne de matematiğin bizzat nesnesi olmadığı ve nesne olarak yalnızca bilimleri ve felsefeyi ilgilendirdiği bir durumda mantığın bizzat gerçeklikle ilgili bir bilim olduğunu söylemek için detaylı bir temellendirme sürecini de bize sunmamaktadırlar. Görüyoruz ki gerçekliğe deney ve gözlem ile eğilen bilim insanları ve gerçeklik hakkındaki bilgimizin kaynağı ve değeri hakkında çalışan felsefeciler bir kenara, mantıkçıların gerçeklikle olan ilişkisi, yalnızca bilim insanları ve felsefecilerin gerçeklikle olan ilişkilerinde, bu ilişkilerin düşünme ilkelerine olan uygunluğunun tesisi konusunda ortaya çıkar. Böylece mantığın bilim olarak anılması, onun bir doğa bilimi veya bir felsefe etkinliği olduğunu söylemeye gelmez.[176]

Mantıkta bir önermenin tutarlı olması ile doğru olması ayrı durumlardır. Bir önermeye, doğru *olabilen* bir yapıdaysa tutarlıdır denebilir. Keza bir grup önerme

[175] **A.e.**, s. 61.
[176] **A.e.**, s. 63.

için de tutarlılık, bu önermelerin bir arada *doğru olabilmeleri*[*] durumunda sağlanmış olur. Nitekim bizim mantıkta doğrular olarak ele aldığımız doğru önermeler, **a priori** yani akıl doğuları ve **a posteriori** yani olgu doğruları olmak üzere birbirlerinden ayrılırlar. Akıl doğruları, tecrübeden bağımsız olarak salt akılla belirlenebilen önermelerdir. "Bir şey ya elmadır ya da elma olmayandır." yollu bir önermenin biçimi "Bir şey ya A'dır ya da A olmayandır." şeklinde ortaya konduğunda görülebilir ki bu biçimdeki bir önerme **a priori** olarak, deneyime gidilmeden de doğruluğu görülebilecek bir yapıdadır. Olgu doğruları ise doğruluğu bizzat deneyime bağlı olarak belirlenebilen önermelerdir. Burada bilimler ve felsefe önemli bir rol oynar.[177] Örneğin "Su deniz seviyesinde 100 santigrat derecede kaynar." önermesi bir **a posteriori** ifadedir. Deney ve gözlemle suyun deniz seviyesinde 100 santigrat derecede kaynayıp kaynamadığına ilişkin çalışmalar yapılır. Böylece bu önermenin bir düzenlilik gösterdiği, yani *doğru olduğu* söylenir.

Bir akıl yürütmenin geçerli olması ile o akıl yürütmede istifade edilen önermelerin doğrulukları arasında bir fark bulunur. Bir akıl yürütme geçerli olabilir çünkü mantık biliminde ortaya konan düşünmenin yasaları, biçimleri vardır. Bir akıl yürütme mantıkça ortaya konan yasalara ve biçimlere uygunsa geçerlilik kazanır. Ancak bir akıl yürütmenin geçerli olması, yani biçimce uygunluk göstermesi, o akıl yürütmede kullanılan önermelere doğruluk kazandırmaz.[178] Çünkü önermelerin doğrulukları, gerçeklikle ilgili bir durumdur. Gerçeklikte yer alan durumlarla ilgili yargılarımızın doğrulukları biçimce belirlenemez, deneyimce ortaya konur.[**] Fakat deneyime yani algılarımıza dayanan bilgi de güvenilir değildir. Çünkü algılarımız doğru ya da yanlış olabilir. Kanılarımızın, beklentilerimizin, alışkanlıklarımızın vb.

[*] Bir olanaktan bahsediliyor, bir olgudan değil. Bir önermenin *doğru olması* ile bir önermenin tutarlı olması ayrıdır. "Elma kırmızıdır." önermesi *doğru olabilen* – deneyimimizce de *doğru olan* – ve dolayısıyla tutarlı bir önermedir. Deney ve gözlemle söz konusu elmanın kırmızı olup olmadığına ilişkin doğruluk **a posteriori** bir bilgi olarak ortaya konabilir. Ancak "Rengarenk, boynuzlu ve uçabilen bir at vardır." önermesi mantık açısından *doğru olabilecek* ve bu yolla da tutarlı denen bir yapıda olsa da önermede dile getirilen *atın var olup olmadığına ilişkin bir soruşturma* mantığın ilgilendiği bir çaba değildir. Bir mantıkçı için böyle bir atın *var olup olmadığı önemsizdir*. Mantıkça hesaba katılan şey önermenin biçimsel olarak uygunluğudur.

[177] Grünberg, **Felsefe ve Felsefi Mantık Yazıları**, s. 369-370.

[178] Erik J. Larson, **Yapay Zekâ Miti -Bilgisayarlar Neden Bizim Gibi Düşünemez-**, çev. Kadir Yiğit Us, Ankara, Fol Kitap, 2022, s. 146.

[**] Elbette burada deneyimce belirlenen doğrulukların olasılıksal mahiyetleri ya da güvenilirlikleri hakkında derin bir epistemolojik tartışma sürdürülebilir. Ancak bu çalışmanın amaçları düşünüldüğünde bu tartışmalar bu çalışmanın bağlamını aşar.

etkisinde olarak algılarımız aldatıcı olabilir. Bu noktada bilimler, kuramlar yoluyla algıyı işler ve onları deneyle arındırır. Nitekim deneyim bilgisinin bu mahiyeti, mantığa kendi mahiyetini aşan bir kimlik kazandırmakta kullanılamaz. Mantıkla yalnızca akıl doğruları, yani **a priori** doğruluklar elde edilebilir. Mantık bizi, somut gerçeklik alanının – yani kırmızı bir elmanın veya bir kayanın vb. – bilgisi hakkında, onların doğrulukları ile ilgili olarak bir yargıya ulaştıramaz. Mantık gerçeklikle ilgili doğrulukların üretilebileceği bir alan olamaz.[179] Dolayısıyla içerik bakımından doğruluk değeri taşıyan önermelerin elde edilmesi işi mantığa değil bilimlere ait bir çabadır. Bilimler, bilimsel yöntemlerle var olanlara yani gerçekliğe gider ve oradan kimi yargılar elde eder. Bu anlamda bilimlerin işi, yöneldiği gerçeklikle ilgili doğru önermeler elde etmektir. O halde denebilir ki bilimler içeriksel, yani nesneye ilişkin, bilgiye erişme çabasına sahiptir ve felsefe etkinliği ise bilgimizin neliği ile bu bilginin elde edilmesi sürecine ilişkin bir çaba olarak karşımıza çıkar. Mantık ise bu iki alan üzerinde yer alır ve gerçekleştirilen çabalara biçimsel bir açıdan yaklaşır, denetleyici bir rol oynar.[180]

Önermelerin doğru olmaları, semantik açıdan önemlidir. Semantik, dile getirilen şey ile o şeyin işaret ettiği başka şeyler arasındaki ilişkide ortaya çıkar. Sentaks ise dile getirilen şeyin dile getirilme kurallarını, örneğin Türkçe dilinde kurulacak bir cümlenin kurallarını, imler. Görülür ki mantık, semantikten çok sentaks temelli bir bilimdir. Mantık, önermelerin ve akıl yürütmelerin alanı olarak bu önerme ve akıl yürütmelerin biçimsel olarak sahip olması gereken kurallarını içerir, böylece mantığın sentaks temelli bir bilim olduğu açıktır.[181] Düşüncenin içeriği olarak semantiğin, mantığın bir konusu olmadığı vurgulanır ancak mantık açısından önemli olanın, bir önermenin biçimce kuruluşunun, yani sentaktik koşullarının, yerine getirilmesi olduğu da söylenir. Aksi halde sentaks üzerinde şekillenen mantık biliminin "Alev yeşilleşti." gibi bir önermenin doğruluğu bakımından da buluşçu kimi mahiyetlerinin olduğunu söyleyebiliyor oluruz. Oysa bu önermenin doğruluğu, yani ifade ettiği bilgi olarak "alevin yeşilleşmesi" olgusu, bilginin kazanılması süreci

[179] Grünberg, **Felsefe ve Felsefi Mantık Yazıları**, s. 370-373.
[180] Özlem, **Mantık**, s. 36-37.
[181] Işıklı, "Fuzzy Mantık ve Sibernetik'in Siber Toplum ve Yapay Zekâ Üzerindeki Etkileri", s. 62.

bakımından mantıkla değil deneyimle sabitleşir.[182] Ancak mantıkça belirlenen sentaktik kurallar ölçüsünde dile getirilir böylece mantıkla denetlenen ama mantıkça türetilmeyen bir ürün olarak "Alev yeşilleşti." önermesi kurulur.

Okuyucu açısından bir dizi soru düşünülebilir: Mantık biliminin yapay zekâ çalışmalarında merkezi bir rol oynamasının gerekçeleri nedir ve mantık biliminin, diğer bilimlerden ve felsefe etkinliğinden farkına neden bu kadar vurgu yapıldı? Yapay zekâ projesinin özünde, insansı olanın insan yapması bir varlık olarak yapay zekâlı varlıklarda gerçekleştirilmeye çalışılması ve bu yolla da insan zihnine ve bedenine ilişkin bilinmezlerin ortadan kaldırılması veya bilinmezlerin anlamsızlaştırılıp elenmesi türünden birkaç arzunun bulunduğu üzerinde durmuştuk. İşte bu arzunun, yapay zekâların inşa edilmesi bakımından, dayandığı temellerden biri de mantıktır. Ancak yapay zekâ çalışmalarında sıkça karşılaşılan tutum, mantık temelinde inşa edilecek yapay zekâların gerçeklikle olan ilişkisinde mantığın gücüne dayanarak, yaratıcı unsurları da elde edebileceği böylece sıkıca inşa edilmiş bir mantık sistemi olarak yapay zekâların insansı olandan ve insan zihninden aşağı kalır bir yanının olmayacağı inancıdır. Bu inanç hatalıdır. Sıkça dile getirdiğimiz gibi mantığın gerçeklikle bağı örtüktür. Gerçekle ilişki içerisindeki bilimler ve felsefe üzerinde sahip olduğu metodolojik bir denetleme rolü bakımından mantık yalnızca biçimsel bir sistem ve bu sistemde yer alan ilkeleri soruşturan bir metodolojiden başka bir bilim değildir. Ancak karmaşık ve birbiriyle koordine şekilde çalışan moleküler durumların gerçekleştiği ve bunlar neticesinde algı, düşünce ve eylemin sinirsel olarak meydana geldiği beyinlerimizi[183], mantıksal yapılar olarak gören bir görüş bulunur.[184] Nitekim simgeci yaklaşımda zihnin kural ve mantık tabanlı, tıpkı bir bilgisayar gibi düşündüğü esas alınır[185], bu yolla da Turing türü bir akıl yürütme – karmaşık bir işlemde bir insana ne yaptığı sorulur ve verilen cevap bir bilgisayarda yönerge tablosu şeklinde işlenir[186] – pekiştirilmeye çalışılır. Yapay zekâ çalışmalarında çözülmesi gereken problemlerin tam da mantık yoluyla başarılabileceği ve böylece mantık tarafından biçimselleştirilen her türden problemin

[182] Grünberg, **Sembolik Mantık El Kitabı: Temel Mantık**, s. 2.
[183] Bell, "Levels and loops: the future of artificial intelligence and neuroscience", s. 2015.
[184] Hilary Putnam, "Beyinler ve Davranış", çev. İdil Gemici, **Phainomena**, No: 1, 2020, s. 154.
[185] Say, **50 Soruda Yapay Zekâ**, s. 99.
[186] Turing, "Computing Machinery and Intelligence", s. 438.

çözüme kavuşturulabileceği McCarthy tarafından da bir araştırma metodolojisi olarak savunulmuştur.[187]

Yapay zekâ çalışmalarının temelinde sembolik mantığın, insan ve diğer canlıların zihin süreçlerini modelleyebileceği düşüncesi vardır.[188] Nitekim mantıksal yapay zekâ, amaçları bakımından yapay zekânın, matematiksel mantığın dilinde yer alan yargılarla ifade edildiği ve yapay zekâya dünyadaki genel olgulara ilişkin bilginin ve türlü özel durumlarda nasıl davranacağı ile ilgili koşulların bu yolla sağlandığı bir yaklaşımdır.[189] Yapay zekâ araştırmalarının ilk amaçlarından biri de makinelere zekâ kazandırmanın sembolik mantık ile nasıl başarılabileceğini bulmaktır.[190] Sembolik mantıkla çalışan sistemlerin özünde ise insan dilinin sembolik bir yapının ürünü olduğu, insan zihninin ise bu sembolleri işleyerek çalıştığı görüşü bulunur.[191] Ancak sembolik mantık temelli yapay zekâ projesi türlü sorunlarla karşılaşır. Örneğin Boden açısından sembolik mantık temelli inşa denemelerinin, insanın kavramlaştırma ve düşünme süreçlerini açıklamak, semantik içerikleri ve insanların sahip oldukları yönelimselliği bilgisayarlarda gerçekleştirebilmek için yeterli olup olmadığı şüphelidir. Çünkü insan zihni her zaman mantıksal bir katılığa sahip değildir, esnek ve değişken bir düşünce tarzına da sahiptir. Sağduyuya dayalı birçok akıl yürütmeye sahip bir varlık olarak insanın zihni, yapay zekâlarda gerçekleştirilecekse mantığın sağduyu bilgisini de kapsayacak şekilde geliştirilmesi gerekmektedir. Bu problemi aşmak amacıyla monoton-olmayan mantık, semantik ağlar vb. gibi çözümler bulunur.[192] Çünkü sağduyu bilgisi, tamamlanmamış bir bilgi türü olarak, bizim fizikte veya mühendislikte kullandığımız türde bir kesinliğe sahip olmayan ve nesneler ile onlar arasındaki ilişkileri, nesnelerin değişimlerini ifade eden bir yapıya sahiptir. Dolayısıyla sağduyu bilgisinin ifade edilebilmesi için mantığın kesin olmayan ifadelerin de yer alabildiği bir dil

[187]Richmond Thomason, "Logic and Artificial Intelligence", The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta, (Çevrimiçi) https://plato.stanford.edu/entries/logic-ai/, 2 Ocak 2023.
[188] Margeret A. Boden, "Artificial Intelligence", Routledge Encyclopedia of Philosophy, (Çevrimiçi) https://www.rep.routledge.com/articles/thematic/artificial-intelligence/v-1#, 18 Kasım 2022
[189] McCarthy, "What is Artificial Intelligence?".
[190] Simon, "Artificial Intelligence: an empirical science", s. 104.
[191] Ayşe Çağıl, "Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekânın Olasılığı", Ankara, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2022, s. 38.
[192] Boden, "Artificial Intelligence".

tasarlaması gerekmektedir.[193] Böyle bir dilin üretilmesi çalışmaları yapay zekâ araştırmacılarını türlü düşünce kamplarına ayırmıştır. Bazı araştırmacılar böyle bir zihnin insan gibi düşünen sistemler hedeflenerek gerçekleştirilebileceğini savunurken bir grup araştırmacı ise rasyonel karar verebilen, sıkı bir mantıksal temele dayanan sistemlerle bu hedefin gerçekleştirilebileceğini savunmuştur. Rasyonel düşünen sistemler tasarlama çözümünün temelinde mantık yer alır. Buradaki amaç, bir problemin çözümü için gereken mantıksal adımların saptanması ve çıkarım kurallarının uygulanmasıyla bir problemin mantıksal adımlar yoluyla çözülmesidir. Dolayısıyla problemlerin çözülmesi için mantık temelli inşa edilmiş programların bulunduğunu görürüz.[194]

Sağduyu bilgisini mantıksal bir sistem içerisinde temsil etme sorunu, yapay zekâ çalışmalarının insan seviyesi zekâdan en uzak olduğu alandır ve monoton-olmayan mantık gibi kimi çözümler yer alsa da yeni fikirlerin ortaya atılması gerektiği vurgulanır.[195] Ancak Holland'a göre erken dönem yapay zekâ geleneğini benimseyenlerin varsaydığı mühim sorun, yüksek seviye sembol işlemenin – yani bilgiyi bir dil içerisinde yargılarla temsil etmeye çalışmak ve bu temsil üzerinde akıl yürütmeler yapmak – insanın tüm davranışları için temel olabileceğinin düşünülmesidir.[196] Çünkü bir yapay zekâ, bir faili *simüle etmek*[*] için programlanmıştır ve dünya hakkında hiçbir bilgisi bulunmamaktadır. Makine zekâsının, ona bilgilerin yüklenmesi gereken bir **tabula rasa** olduğu açıktır ve programcının veya öğrenen bir sistemin bu boş zihni işlemesi gerekir. Ancak

[193] John McCarthy, "What has AI in Common with Philosophy?"
[194] Akın, "Yapay Zekâda Vücut Beyin Problemi".
[195] McCarthy, "What is Artificial Intelligence?".
[196] Owen Holland, "The future of embodied artificial intelligence: Machine consciousness?", **Embodied Artificial Intelligence**, Ed. Fumiya Iida, Rold Pfeifer, Luc Steels, Ysuo Kuniyoshi, Berlin, Springer, 2004, s. 38.
[*] Bir şeyi simüle etmek ile o şeyi kopyalamak farklıdır. Örneğin Bennu'nun zihnini simüle etmemizle Bennu'nun zihnini kopyalamamız bize farklı sonuçlar verir. Çünkü ilk durumda Bennu'nun zihninin koşullarını, o koşullara en yakın alternatif bir zeminde, o zihnin çalışma ilkelerini de dikkate alarak gerçekleştiririz. Böyle bir zihin Bennu'nun zihnine benzer nitelcelere ya da düşünme yetisine sahipmiş gibi görülebilir ancak bu bir yanılgıdır çünkü söz konusu simülasyonun bir nitelceye sahip olabilmesinin koşullarının, simülasyonun bir *bene* sahip olması gerektiği açıktır. Ayrıca bir simülasyonda gerçekleşen acı hissi ile Bennu'nun hissedebileceği herhangi bir acının gerçekliği birinin matematiksel bir ontolojiye ait olması, diğerinin ise asıl anlamda gerçekliğe ait olmasıdır. Öte yandan Bennu'yu kopyalamak ise Bennu ile aynı fenomenal deneyime ve benliğe sahip olmasa da kendi fenomenal deneyimlerine sahip olabilen, asıl anlamda acı hissine sahip olan bir varlığın, bir failin gerçekleşmesidir.

mantıksal bir sistem olduğu için de bir yapay zekânın geliştirilmesi için bilginin yapay zekâ tarafından kullanılabilir bir formatta – yani dünya hakkındaki her türlü bilgimizin, fenomenal yargılarımız da dahil olmak üzere, mantıksallaştırılarak makineye sağlanması – sunulması gerekmektedir.[197] Fakat mantıksal yapay zekânın epistemolojik ve problem çözmeye ilişkin kimi problemleri bulunur. Epistemolojik açıdan bir yapay zekânın herhangi bir durumda ihtiyaç duyacağı bilginin ne olacağı ve bu bilginin nasıl temsil edileceği sorunu yer alır. Öte yandan bir problemin çözümü için ise problemlerin çözümünde hangi bilginin ve nasıl bir yolun tercih edilmesi gerektiğine ilişkin sorunlar bulunur.[198] Nitekim bu noktada sorulabilir: Tüm bu problemlere rağmen yapay zekâ çalışmalarında mantığın kullanılmasındaki gerekçeler nelerdir? Minsky'ye göre burada iki neden bulunur. İlkin mantık, anlam bulanıklıklarının bulunmadığı kesinlikli bir alandır. Diğer neden ise mantık üzerine birçok çalışma bulunuyor olmasıdır. Ancak Minsky, mantığın dezavantajlarının da bulunduğunu belirtir: Mantıksal bir dizgede dile getirilen önermeler, onları tanımlayan, onları olanaklı kılan daha önceki önermelere dayanırlar ve mantıksal sistemler esneklik özellikleri oldukça düşük sistemlerdir dolayısıyla mantık sistemleri varsayımsal ifadeleri, analojileri ve metaforları dışlamaktadır.[199] Öte yandan Thomason'a göre ise mantığı kullanma motivasyonu, teoremlerin kanıtlanması gibi doğrudan ve temel bir şekilde mantıksal çıkarım türleri kullanılmasa bile, mantıksal biçimleştirmenin, akıl yürütmeye ilişkin problemlerin kendilerini anlamamıza oldukça yardımcı oluyor olmasıdır.[200] Yapay zekâ gibi insan zihnini ve failliğini kendine model seçen bir alanda mantık gibi sıkı ve insan zihninin çalışma ilkelerini anlamamızda bize yardımcı olmuş bir bilimin kullanılması oldukça doğaldır. Nitekim McCarthy'nin de işaret ettiği gibi, yapay zekâ çalışmalarında insan düzeyi bir zekânın yakalanması için türlü aşamaların başarılmaya çalışılacağı görülebilir: Tüm eylemlerin mantıksal bir sistem içerisinde temsil edilebilmesi, davranıştaki değişikliklerin temel olarak ifade edilebilmesi, eylemin her özelliğinin

[197] Daniel C. Dennett, "Cognitive Wheels: The Frame Problem of AI".
[198] John McCarthy, "Concepts of Logical AI", (Çevrimiçi) http://www-formal.stanford.edu/jmc/concepts-ai.pdf, 28 Aralık 2022.
[199] Marvin Minsky, "Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy", **AI Magazine**, C. XII, No: 2, 1991, s. 41.
[200] Thomason, "Logic and Artificial Intelligence".

geliştirilebilir olması, eylemin gerçekleşmesine ilişkin makinenin başarmaya ilişkin kavramlara sahip olması vb.[201]

Biz insanlar, dünya hakkında çeşitli yargılara sahibiz. Elmalar kırmızı olabilir, savaş uçakları oldukça hızlı olabilir ve bir insana sevgi besleyebiliriz. Görüleceği gibi dünyada olan kimi olgulara ilişkin bazı yargılarımız diğer yargılarımızdan biçimsel ve içeriksel olarak ayrılır. Mantıksal bir sistem içerisinde birinci ve ikinci önermenin ifade edilmesi işi oldukça kolaydır. Ancak bir insana duyulan sevginin veya herhangi bir fenomenal deneyimin, bir mantıksal sistem içerisinde ifade edilmesi işi, sağduyu bilgisinin mantık sistemi içerisinde ifade edilmeye çalışılması işidir. Sağduyu bilgisine – olaylar ve onların sonuçları hakkındaki temel yargılarımız olarak örneğin kanılarımız, arzularımız vb. – sahip varlıklar olarak dünya hakkında birçok yargıya sahibiz ancak bunlar kimi kuramlarla açıklanamamaktadır.[202] Ancak yapay zekâ çalışmaları incelenirse sağduyu bilgisinin de temsil edilebileceği girişimlerden biri olarak, McCulloch ve Pitts'in çalışmaları bizi karşılar. McCulloch ve Pitts, yapay sinir hücrelerine dayanan ve bu hücrelerde "ve" ve "veya" gibi mantıksal operatörler ile çalışan bir mantıksal sistem ile hesaplanabilir herhangi bir fonksiyonun mantıksal kurallara göre inşa edilmiş bir sinir ağı üzerinde hesaplanabileceğini savunmaktadır. Böylece bu mantıksal sistemler uygun şekilde tasarlanabilirse öğrenme becerisi de elde edebilmektedir.[203]

Mantık temelli yapay zekâ çalışmalarında bir diğer çözüm ise Allen Newell ve Herbert Simon'un fiziksel sembol sistemi hipotezidir. Bu hipoteze göre fiziksel bir sembol sistemi, sembolleri saklayabilen, bunları girdi ve çıktı ilişkisi içerisinde organize edebilen, koşullu çalışan sistemlerdir.[204] Burada dile getirilen semboller nedir? Bir sembol, bir temsil nesnesi olarak başka bir varlığın ifadesi olarak karşımıza çıkar[205] ve tekil nesneler, nitelikler, kavramlar, arzular, duygular, milletler, renkler, moleküller gibi birçok şey sembol olarak kullanılabilir durumdadır ancak bir

[201] McCarthy, "Programs with Common Sense", s. 79.
[202] John McCarthy, "Artificial Intelligence, Logic and Formalizing Common Sense", (Çevrimiçi) http://www-formal.stanford.edu/jmc/ailogic.pdf, 27 Aralık 2022.
[203] Akın, "Yapay Zekâda Vücut ve Beyin Problemi", s. 148.
[204] Simon, "Artificial Intelligence: an empirical science", s. 104.
[205] Vernon ve Furlong, "Philosophical Foundations of AI", s. 57.

şeyin sembol olarak kullanılabilmesi için tanımlanmış olması gerekmektedir.[206] Newell, fiziksel sembol sistem hipotezine göre inşa edilmiş bir sistemde istifade edilecek sembollerin, insanların gündelik yaşamlarında kullandığı sembollerden farklı olmadığını, insanların da bir fiziksel sembol sistemi olarak kabul edilebileceğini iddia etmektedir.[207] Newell'ın bu görüşü iki gerekçeye dayanır: İlkin, bir fiziksel sembol sistem fizik evrenimiz içerisinde gerçekleştirilebilir bir durumdadır ve bu sistemin dayandığı sembol kavramı ise **a priori** bir kavramdır. Öte yandan bir sembol sistemi, bir bellekten, bir grup operatörden, bir kontrolcüden, bir girdi ve çıktıdan oluşur. Girdiler bir tür sembol olup, çıktılar ise türlü operatörlerce işlenmiş olan semboller olarak karşımıza çıkar. Böylece makine girdilerin yani sembollerin işlevleri sonucu bir dışsal davranış üretmiş olur.[208] Nitekim Newell, dijital bilgisayarların fiziksel sembol sistemlerinin gerçekleştirilmesi konusunda en iyi aday olduğunu savunur çünkü dijital bilgisayarların devre düzeyi, mantık düzeyi ve program düzeyi gibi hiyerarşik bir düzene sahip olmasının bilgisayarın – veya bir makine zekâsının – dışsal davranışının oluşmasında sembollerin önemini gösterdiğini vurgular. Semboller burada sıkı bir mantıksal mimari içerisinde yer alarak türlü ilişki ve nesneleri temsil ederler. Ancak Newell, bir mantıksal sistemin sonsuz sayıda farklı gerçekleşmesinin söz konusu olabileceğini dolayısıyla da tüm mimari türlerini kapsayan genel bir mimari sınıfından söz edilemeyeceğini belirtir, insan zihninin nöral yapısının da sembol sistemi olarak kabul edilebilecek bir tür mimari olduğunu savunur.[209] Nilsson'a göre de fiziksel sembol sistemleri, bilgisayarlara uygun sembol işleme programları sağlandığında zekice eyleyebilecek, zekâyı ortaya koyabilecek olanaklara sahiptir.[210]

Fiziksel sembol sistemi hipotezi bize güçlü bir açıklama modeli sunuyor gibi görünse de bu hipoteze ilişkin kimi itirazlar bulunur. Her ne kadar fiziksel sembolleri

---

[206] Schank ve Abelson, **Scripts Plans Goals and Understanding: An Inquiry into Human Knowledge Structures**, s. 4.
[207] Allen Newell, "Physical Symbol Systems", La Jolla Conference on Cognitive Science, California, 15 Ağustos 1979, (Çevrimiçi) https://scholar.archive.org/work/skfkjrtc3nftnacjc7berhd7vq/access/wayback/https://s3-eu-west-1.amazonaws.com/pstorage-cmu-348901238291901/12099818/file.pdf, 6 Ocak 2023, s. 1.
[208] **A.e.**, s. 7-8.
[209] **A.e.**, s. 44-45.
[210] Nils J. Nilsson, "The Physical Symbol System Hypothesis: Status and Prospects", **50 Years of Artificial Intelligence: Essays Dedicated to the 50th Anniversary of Artificial Intelligence**, ed. Max Lungarella, Fumiya Iida, Josh Bongard ve Rolf Pfeifer, Springer, 2007, s. 9.

manipüle edebilen herhangi bir sistemin insanların sahip olduğu zekâya sahip olabileceği iddia edilse de[211] bunun, sembollerin manipülasyonları neticesinde gerçekleşemeyeceği Turing'in bir makinenin zihin durumunun o anda işlediği/ayırdında olduğu sembolden ve bu sembolden hareketle ortaya çıkan eylemden başka bir şey olmadığı yollu görüşünden[212] de açıktır. Bir insanın düşünmesinin semboller arası ilişkilerden ibaret olduğunu kabul etsek bile zihin durumlarımız ve zihinsel yaşantımız yalnızca şimdide bulunmaz, şimdide dışsallaşan o davranışla da sınırlanamaz. Örneğin ben önümde duran bir kırmızı elmaya uzanabilirim, onu kavrayabilir ve 1 metre yukarı kaldırabilirim. Daha sonrasında bir gözlemci benim bu elmayı yemek istediğime ilişkin davranışsal kimi tespitlerde bulunabilir. Turing'in de açıkça ifade ettiği gibi, bir makine – fiziksel sembol sistemler de bir makine olacağı için – ancak ve ancak o anda işliyor olduğu bilgiden – ki bu bilgi sembollerle, semboller arası ilişkilerle temsil ediliyor – başka bir bilginin ayırdında değildir. Belki okuyucu açısından şöyle bir soru açığa çıkabilir: Bir bilgisayarın rastgele erişilebilir belleğinde geçici olarak kimi bilgiler tutulmamakta mıdır? Böylece bir bilgisayarın da hiç olmazsa tıpkı insanlarda olduğu gibi sınırlı da olsa bir hafızası olduğunu, semboller arasında geçmiş ve şimdi ilişkisini sağlayabildiğini söylemek olanaklı değil midir? Biz bu gelişmeleri yadsımıyor, modern bilgisayarların bu olanaklara sahip olduğunu kabul ediyoruz. Ancak iddia ediyoruz ki bir bilgisayarın rastgele erişimli belleğinin bulunması ona kendi var oluşunu bir bütün olarak ya da tamamlanmış ve tamamlanmakta olan bir birlik olarak kavramasını sağlamaz, Turing'in de belirttiği gibi, yalnızca makinenin o andalığının erimini arttırır. Makine belki şu anda işliyor olduğu 4 sembolün ayırdında değildir ancak örneğin 8^9 sembolün eş zamanda ayırdında olması demek de onun şu andalığının ancak ve ancak genişlediğini gösterir bize, makinenin bir *kavrayışa* eriştiğini değil. Anlıyoruz ki fiziksel sembol hipotezini temel alan bazı yapay zekâ araştırmacıları, bilgisayarlara kapatılmış, duyargaları veya beden uzantıları olmayan elektronik beyinler üretmeyi arzulamaktadırlar. Ancak bu hipoteze göre inşa edilecek bir makinenin bir dünya modeli olması gerekli değil

[211] Searle, **Minds, Brains and Science**, s. 29.
[212] A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem", **Proceedings of the London Mathematical Society**, C. XLII, 1936, s. 250.

midir? Böylece Akın'ın da dile getirdiği gibi, dünya modelinin de ve bu model içerisinde var olan nesnelerle makine arasındaki ilişkilerin de bir insan yani bir programcı tarafından makineye sağlanması gerekir.[213]

Mantık temelinde şekillenen yapay zekâ çalışmalarının üstesinden gelmesi gereken bir problem ise çerçeve aksiyomlarıdır.[*] Çerçeve problemi, ilk olarak John McCarthy ve Pat Heyes tarafından tanımlanmış bir problemdir ve bu problemin özü bir zekânın tanımlanmış herhangi bir problemin çözümü ile ilgili olan kanılarla, ilgili olmayan kanılar arasında ortaya koyduğu bir ayrımı bir makine zekâsının da gerçekleştirip gerçekleştiremeyeceğidir.[214] Burada temel düşünce, makinenin bir problemin çözümünde, problemle ilgili olan, problemin çözülmesine yardımcı olacak yargılarla problemle uzaktan ilgili olan veya problemle ilgili olmayan yargıların ayrılmasına ilişkin bir beceriye sahip olup olamayacağı ile ilgili bir tartışmadır.[215] Örneğin bir makine, bir odada bulunan bir elmanın bulunması, bu elmanın odadan çıkarılması ve X gibi bir konuma ulaştırılması probleminin üstesinden gelmek için, sahip olduğu diğer sayısız yargıları bir kenara bırakabilmeli, bu problemin çözümüyle ilgili olarak hareket, kavrama, araştırma, konum tespiti vb. türünden birçok alt başlıkla analiz edilebilecek bu problemle alakalı yargıları tespit edip bunları kullanabilmeli ve problemi çözebilmelidir. Dolayısıyla çerçeve aksiyomları sisteme tanımlanan her eylem tipi için genel koşulları tanımlamaktadır.[216] Çerçeve aksiyomlarının önemi şudur: Yapay zekâ çalışmaları, doğal dil işlemenin sentaktik süreçlere çok duyarlı olduğu ve dile getirilmemiş semantik varsayımların öneminin bulunduğu bir durumda programın daha önce dile getirilmemiş kimi durumları yanıtlaması gerektiğini bize söylemektedir.[217] Örneğin yapay zekânın belli konu ve problemlerle ilgili özelleştirildiği ve bu yolla gündelik yaşamın türlü sorunlarıyla ilgili tek bir zekânın değil ama özelleşmiş birçok zekâ ile mücadele edilebileceği düşüncesinden hareketle ortaya çıkan uzman sistemler anlayışının başarılı olamamasının sebebi de budur. Çünkü gündelik deneyim durağan değildir,

[213] Akın, "Yapay Zekâda Vücut Beyin Problemi".
[*] İng. **Frame Problem**
[214] Simon Blackburn, "frame problem", **The Oxford Dictionary of Philosophy**, 2. bs., Oxford, Oxford University Press, 2005, s. 140.
[215] Dennett, "Cognitive Wheels: The Frame Problem of AI".
[216] **A.e.**
[217] Boden, "Artificial Intelligence".

değişkendir. Yargılarımız farklı zamanlarda farklı doğruluk değerleri alabilir. Böylece verilerden öğrenmenin olmadığı ve bilgi tabanının uzun bir süreçle uzman insanlar tarafından elle oluşturulması gereken uzman sistemler, mantığın da dünyayı temsil etmekteki başarısızlığından ötürü başarılı olamamışlardır.[218] Nitekim, her olgunun bir insan tarafından programa dahil edilmesi gerektiği bir durum insanları olguları öğrenebilen programlar tasarlamak konusunda bir çözüm bulmaya yöneltmiştir.[219]

Mantığın ve mantık temelli yapay zekâ çalışmalarının burada dile getirilen ve benzer birçok problemi olduğu açıktır. Ancak Boole'un çalışmalarından da sonra mantığın cebir dilinde, yani makine dili de olan 0 ve 1'ler ile, ifade edilmesi[220], yapay zekâ projesi içerisindeki gerçekçi olmayan, özellikle mantığın varlık koşullarını aşan kimi düşüncelerin meydana gelmesini sağlamıştır. Örneğin McCarthy, matematiksel mantığın, sağduyu bilgisi veya bilimsel bilginin her türü için ifade olanağı sağlayan bir temel olabileceğini savunmaktadır.[221] Böylece 0 ve 1 ile cebirleştirilmiş olan mantığın, gündelik insan yaşamının her kavramının makine dilinde temsil edilebileceğine ilişkin bir kanıyı destekleyen bir gelişme olduğu görülmektedir. Ancak Arf, insan beyni ve makine zekâsı arasında bulunan varlık farklarına dikkat çeker, insanların daha önce karşılaşmadıkları problemleri çözebildiklerini böylece insan beyninin karakteristik bir özelliğinin ise yeni olan durumlara adapte olabilmesi olduğunu dile getirir. Makine zekâsı konusunda ise Arf, makine zekâsına yapay bir beyin olarak bakmanın bu koşullar altında mümkün olmadığını dolayısıyla makine zekâsının, makine inşa edilirken düşünülmemiş olan problemleri de çözebilip çözemeyeceği ile ilgili bir sorunun bulunduğuna işaret etmektedir.[222] Gerçi Arf'ın bu itirazlarını dile getirdiği yıllarda henüz daha kuramsal kimi arka planlara sahip olan nöral ağlar temelli öğrenen yapay zekâlar, gerçekleşmeleri bakımından yaygın olmadığı için, okuyucu açısından bu düşüncenin, güncelliği yitmiş bir düşünce olduğu kabul edilebilir. Ancak unutulmamalıdır ki Arf düşüncelerinin devamında, kendi kendisini geliştirebilecek, öğrenebilen makinelerin

[218] Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, s. 51.
[219] Bell, "Levels and loops: the future of Artificial intelligence and neuroscience", s. 2014.
[220] Say, **50 Soruda Yapay Zekâ**, s. 19.
[221] McCarthy, "What has AI in Common with Philosophy?".
[222] Arf, "Makineler Düşünebilir mi ve Nasıl Düşünebilir?", s. 99.

yapılabilme olanaklarını yadsımamaktadır. Tam aksine kendi kendini geliştirebilen, öğrenebilen makinelerin yapılması durumunda bile insan beyni ve makineler arasında estetik yönden – muhtemelen yaratıcılık konusuna bir vurguyla beraber – ve bir işi yapmak veya yapmamak konusunda bir karar alıcı unsurun onda mevcut olup olmaması üzerinden bir dizi varlık farkına dikkat çekmektedir. Ancak unutulmamalıdır ki kendisinin de ifade ettiği üzere, bu unsurlar derin düzeylerde belirsizlik içerir. Nitekim Arf, belirsizlik unsurunun yalnızca insan beynine has bir durum olmadığını da kabul eder. Kuantum gelişmelerin, atom altı düzeyde açıkladığı belirsizlik durumlarının insan beyni dışında da belirsizliğin üstesinden gelebilecek makinelerin inşasında bir olanak olarak görülebileceğinin mümkün olduğunu ancak bu gelişmenin hiçbir zaman başarılamayacağını iddia eder.[223]

Herbert Simon'un bir fiziksel sembol sisteminin zekâya sahip olabilmek ve zekice eyleyebilmek bakımından gerekli ve yeterli araçlara sahip olduğunu söylemesi[224], yapay zekâ çalışmalarında tıpkı insanların sahip olduğu gibi bir zekâya, insanlarla eş bir varlığa bizi ulaştıracak yöntemlerin araştırılması konusunda bir hedef göstermiştir. Yapay zekâ çalışmalarını insan düzeyinde zekâya ulaştıracak yöntem ve koşulların soruşturulması sürecinde fiziksel sembol sistemi hipotezinin önemi büyüktür: Bir fiziksel sembol sistemi olarak kabul edilebilecek olan insan zihninin *kablolama diyagramı* ilkece anlaşılabilirse, o halde insanın duyumsamaya ilişkin örüntüleri tespit edilebilir ve insan zihnin ürünleri yeniden inşa edilebilir şekilde sunulabilir. Böylece elektrik sinyallere indirgenen zihin, dijital bilgisayarlarda yeniden gerçekleştirilebilir.[225] Nitekim Churchland ve Sejnowski'ye göre de sembolik yapılar olarak temsiller, hesaplamanın bunlar üzerinde manipülasyonlar yoluyla gerçekleştiği yapılardır.[226] Transistörlere ve mantığa indirgenen zihin, bu yolla semboller arası ilişkinin, hesaplamanın bir sonucu olarak karşımıza çıkmaktadır. Bu indirgeme, sentaks ve semantik üzerinden devam etmekte olan tartışmaları da semantiğin sentaksa indirgenmesi yoluyla aşmak istemektedir.

---

[223] **A.e.**, s. 103.
[224] Simon, "Artificial Intelligence: an empirical science", s. 104.
[225] F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", **Psychological Review**, C. LXV, No: 6, 1958, s. 386.
[226] Patricia Smith Churchland ve Terrence J. Sejnowski, "Neural Representation and Neural Computation", **Philosophical Perspectives**, C. IV, 1990, s. 352.

Örneğin Demir, her ne kadar semantiğin, terimlerin, işaret ettiği şeylerle olan ilişkisinde ortaya çıktığını söylüyor olsa da asıl meselenin bir *işaret etme meselesi* olduğunu dile getirir.[227] Ancak burada bir fark vardır: Sentaks, sembollerin ve semboller arası ilişkilerin düzenlenmesini imliyor ise semantik, sentakstan farklı olarak, sembollerle sembollerin eşlendiği şeyler olarak var olanlar arasındaki ilişkiyi ifade etmektedir. Mantığın içeriksiz bir alan olduğu düşünülürse, mantığa içeriğini kazandıran şeyin deneyim olduğunu, sembollerin var olanlarla eşlenmesinin tek yolunun ise deneyim olacağını görmemiz oldukça doğaldır. Ancak her ne kadar sentaks sembollerin bir şeye işaret etmelerinin ilkelerini içeriyor olsa da bir *şeye* işaret etmek ile bir şeye *işaret edebilmek* iki ayrı olgudur. İlkinde bir şeyin *duyumsanması* ve bir sembolle *eşlenmesi* söz konusudur, ikinci durumda ise duyumsanan bir şeyin bir sembolle *eşlenmesinin kuralları* söz konusudur. Dolayısıyla sembollerin bir şeye işaret edebilmeleri, kendi başlarına semantiğin oluşturulması için gereklidir ancak yeterli değildir. Ancak ve ancak bir sembol bir *deneyim süreciyle beraber* ve *bir deneyimleyenle* semantiğin belirmesinde katkı sunabilir. Demir'in sentaksın semantik için yeterli olacağı görüşünü Schank ve Abelson da paylaşır. Onlara göre insan zihnindeki semantik işleyiş, tıpkı Boole'un sınıf/aitlik ilişkisi görüşü gibi, hiyerarşik biçimde organize edilmiş ve sınıf aitliği üzerinde temel ilişkilerin kurulduğu bir işleyiştir.[228]

Mantık temelli yapay zekâ çalışmalarında John Searle'ün bir problem olarak sıkıca tanımladığı semantik sorun, sentaks üzerinden aşılmaya çalışılan bir sorundur. Sentaksın semantiği üretebileceğini düşünen Zambak ve Vergauwen gibi düşünürler de bulunur.[229] Bu düşünürler, makine zekâsının fail niteliği kazanmasının koşullarını duyu verileri, algılama, öğrenme ve bellek gibi bilişsel yetilerin bu zekâya kazandırılması gerekliliği üzerinden kurgular ancak makine zihni ve insan zihni arasında yapısal ve organizasyonel kimi farklar bulunduğundan dolayı da makinelerde böyle bir bilişsel altyapının oluşturulabilmesi için alternatif yolların

---

[227] Erhan Demircioğlu, "Mehmet Hilmi Demir: 'Aklımızın Sınırlarını Deniyoruz Aslında, Nereye Kadar Gidebiliriz?'", s. 150-151.
[228] Schank ve Abelson, **Scripts Plans Goals and Understanding: An Inquiry into Human Knowledge Structures**, s. 18.
[229] Aziz Zambak ve Roger Vergauwen, "Artificial Intelligence and Agentive Cognition: A Logico-Linguistic Approach", **Logique et Analyse**, C. LII, No: 205, 2009, s. 70.

denenmesi gerektiğini ileri sürerler, mantıksal-dilsel bir modelin ise bu alternatif yollar arasında önde geldiğini çünkü bir makine zekâsının dışsal dünyada failsi bir niteliği elde edebilmesinin tek yolunun dil olduğunu iddia etmektedirler. Çünkü faillik açısından ihtiyaç duyulan soyut kavramların inşası mantıksal-dilsel bir çerçeve tarafından üretilebilir. Bu çerçeve, makine zekâsındaki deneyimi işlevsel hale getirir yani faillik için gereken bilişsel yetiler, bu çerçeve sayesinde deneyim üzerine yansıtılır, böylece deneyimde yer alan kavramlar dilsel bir çerçevede ifade edilmiş olur.[230] Zambak ve Vergauwen her ne kadar öne sürdükleri mantıksal-dilsel modellerinin, dil öncesi düşünmenin olmadığına ilişkin bir anlamı ifade etmediğini vurgulasalar da faillik niteliği kazanabilecek bir zekâ olarak bir makine zekâsının, tüm bilişsel süreçlerinin bu mantıksal-dilsel yapıya dayanması gerektiğini savunurlar. Çünkü bu düşünürler, dilin makine zekâsı ve çevre arasındaki bilişsel bir araç olduğunu iddia ederler, böylece makine zekâsı dünyayı, yani çevreyi, ancak ve ancak dil üzerinden deneyimleyebilir. Deneyim kavramı ise makinede linguistik bir yapı üzerinden düşünülebilir. İnsan zihni bakımından düşünülürse deneyim direkt olarak linguistik kategorilerde karşılık bulmayabilir ancak linguistik süreç, insan zihninde anlamın üretilmesinde bir rol oynar.[231] Burada sorulabilir: Dil de bir mantıksal sistem olarak kurallardan ve sembollerden oluşan bir sistem değil midir? O halde Zambak ve Vergauwen'in düşüncelerinde kendine yer bulan sentaksın aşılması ve semantiğin basitçe bir sentaks işi olduğu yollu görüşler, mantıksal sistemlere geliştirilen eleştiriler bakımından nasıl ele alınabilir? Açıkça görülebilir ki dil de deneyimce zenginleşen ve esasında içeriksiz, sembolik bir yapıdır. Dil, makine zekâsının failsi niteliği kazanması için gerekli olan bilişsel yetilerin kazanılmasında rol oynar demek ile bir mantıksal sistem insan zihnini onda gerçekleştirebileceğimiz zemini bize sunar demek arasında bir fark yoktur. Dolayısıyla Zambak ve Vergauwen'in öne sürdüğü dilsel yapı da mantık sistemlere getirilen eleştirileri üstlenmek zorundadır. Çünkü mantıksal-dilsel yapının iddiaların gerçekleştirilmesi açısından, herhangi bir mantıksal sistemin de yanıtlaması gereken noktalar olarak, duyargaların rolü ve dünyanın duyumsanmasının tarzı bakımından kimi hususlar yadsınmaktadır. Mantıksal sistemler, sistemin inşa edilmesi ve anlamın bu sistemde var edilmesi için

[230] **A.e.**, s. 58-59.
[231] **A.e.**, s. 59-60.

duyargaları ve duyumsama yetisini yadsımaktadır. Sembolik mantık esaslarına göre inşa edilmiş bir makinenin, sahip olduğu program üzerinden birtakım çıkarımlar yapacağı açıktır ancak Işıklı tarafından da sorulur ki sağduyuya sahip olmayan bu makineler, mantıksal saçmalığı veya gerçek dışılığı nasıl tespit edebilir?[232] Sağduyu bilgisinin elde edilmesinin koşullarından biri de duyargalara sahip olmaktır. Ancak matematiksel temelleri incelediğimiz alt bölümde de gördüğümüz gibi bir makinenin duyargaları da Boole cebrine, elektrik sinyallerine dayanır. Böylece bir makinenin duyargaları da fizik evrenimizden ona gelen verilerden ancak ve ancak matematiksel bir ontolojinin izin verdiği, onun uygun kıldığı bir tarzda etkilenebilir.

John Searle, mantıksal sistemlerin dayandığı temel kavramlardan biri olan "sembol" kavramına ciddi eleştiriler getirmiştir. Searle, sembollerin semantik içerikten yoksun olan ögeler olarak, kendi başlarına herhangi bir şeyi bile dile getirmediğini, dolayısıyla da bir zihne sahip olmanın, biçimsel ve sentaktik süreçlere sahip olmaktan fazlası olduğunu iddia eder. Görüyoruz ki örneğin Zambak ve Vergauwen'in dil üzerinden ileri sürdükleri görüşler, Searle bakımından kabul görmez çünkü Searle, bir dili anlamanın ya da zihinsel süreçlere sahip olmanın biçimsel sembollerden daha fazlasına sahip olmak anlamına geldiğini savunur.[233] Uzman sistemler anlayışının da temelini oluşturan, uzman insanlar tarafından yargılarla inşa edilen bir mantıksal sistemin, örneğin bir elmayı yetiştirecek uzman bir çiftçi sistemin, başarısızlıklara uğramış ve uğrayacak olmasının gerekçeleri de budur: Programlar yalnızca sentaktik şekilde çalışır ve sağduyu bilgisini, semantiği kullanacak kabiliyette değildir.[234] McCarthy'nin de dile getirdiği gibi insan düzeyindeki zekâ ile makine zekâsı arasında birçok problem bulunur: Dünyanın sağduyu bilgisi, epistemolojik olarak uygun diller, nesneler olarak kavramlar, monoton-olmayan mantık, bağlamlar olarak nesneler, iç gözlem vb.[235] Dikkat edilirse tüm bu problemler, bir metodoloji olarak mantığın ilgilenmediği, bilimlerin ve felsefenin ilgilendiği alanlarla ilgili konuların mantıkla çözümlenmek istenmesi

[232] Işıklı, "Fuzzy Mantık v e Sibernetik'in Siber Toplum ve Yapay Zekâ Üzerindeki Etkileri", s. 62.
[233] Searle, **Minds, Brains and Science**, s. 31-33.
[234] Altınörs, "Zihin-Beden Problemine Searle'ün Yaklaşımı", s. 48.
[235] John McCarthy, "From here to human-level AI", **Artificial Intelligence**, No: 171, 2007, s. 1180-1181.

ile mantığı bilimler ve felsefeyle denk, **a posteriori** mahiyetteki bilginin üretildiği bir alan kılmak istencinin bir sonucu olarak ortaya çıkan sorunlardır.

Bir mantıksal sistem yaklaşımı olarak fiziksel sembol sistem hipotezine getirilen kimi eleştiriler bulunur. Nilsson bu eleştirileri dört ana grupta özetler: Bilgisayarlar anlamsız sembolleri manipüle etmektedir, algı sembolik olmayan bir tarzdadır ve analog sinyallere dayanmaktadır, hesaplama kuramı zekâ için uygun bir model değildir, zeki davranış tıpkı böceklerde ve bitkilerde olduğu gibi zihinsiz de gerçekleşebilir.[236] Nilsson bu eleştirilerin fiziksel sembol sistem hipotezinin önemini ve geçerliliğini ortadan kaldırdığını düşünmemektedir. Çünkü ilk eleştirinin, hipotezin sembolleri temellendiremediği gibi hatalı bir varsayımdan hareket ettiğini, ikinci eleştirinin ise sembolik olmayan süreçlerden bahsettiğini ve ihtiyaç duyulursa onların da sisteme dahil edilebileceğini düşünmektedir. Eleştirilere yanıtını sürdüren Nilsson, üçüncü itirazın geçersiz olduğunu savunur çünkü beyin üzerine yapılan araştırmalarda araştırmacılar büyük çoğunlukla beyindeki işlevleri bilgisayar biliminin kavramlarını kullanarak tanımlamaktadır. Son eleştiri ise zihin sahibi olmadığını düşündüğümüz varlıkların davranışlarının zihne ihtiyaç duymayan türden davranışlar olduğunu kabul ettiğimizde ortadan kalkmaktadır.[237] Bu yanıtlar yeterli midir? İlk olarak, sembollerin temellendirilmesi meselesi sentaks-semantik tartışmasını çözümlemek için gerekli ve yeterli bir gelişme olarak fiziksel sembol sistemi hipotezi tarafından çözümlenmiş değildir. Bir hipotez olarak düşünüldüğünde sembollerin statüleri hakkında kimi yargıları öne sürüyor olsa, fiziksel sembol sistemlerinin sembollere ve semboller arası ilişkilere dayanarak semantik içeriği üretebileceği yalnızca bir varsayımdır. Öte yandan insan zihnine ilişkin onun koşulları bakımından sembolik olmayan kimi durumlar söz konusudur. Örneğin nöronlar analog sinyallerle çalışmaktadır. Dolayısıyla sembol sistem hipotezinde sembolik olmayan varlık koşullarının da dahil edilebileceği bir olanak bulunmamaktadır. Aksi halde semboller üzerine kurulmuş olan bu hipotez nitelik değiştirmek durumundadır. Nilsson'un üçüncü eleştiriye ilişkin itirazı da yalın bir **ad populum** örneği olarak karşımıza çıkar. Beyindeki süreçlerin ve türlü işleve sahip bölgelerin birtakım bilgisayar bilimi kavramıyla ifade edilmesi ayrıdır – ki bu

[236] Nilsson, "The Physical Symbol System Hypothesis: Status and Prospects", s. 9-10.
[237] **A.e.**, s. 16.

çalışmanın kapsamı bakımından hem şimdiye kadar ele alındığı kadarıyla hem de ayrıntıda üçüncü bölümde görülecektir ki – bu süreç ve işleve sahip bölgelerin bilgisayar kavramlarıyla ifade edilmeye çalışılması işini pek çoklarının da yapıyor olduğu ayrıdır. Birinci durumda bu kavramları kullanma hakkımızın, kavramların karşılamak istedikleri varlığın koşulları bakımından ele alınması gerekmektedir. İkinci durumda ise bir kavramın bir varlığın belli bir koşulunu ya da kendisini karşılaması bakımından birçokları tarafından kullanılıyor olması söz konusudur. Nitekim bu durum o kavrama bir meşruiyet kazandırmaz. Sayıca çokluk akıl yürütmelerimizin gücünü arttırmaz. Nilsson'un zihnin bulunmadığı varlıkların davranışlarının zihin gerektirmeyen davranışlar olarak adlandırılması gerekliliğine ilişkin görüşü de incelenebilir. Burada birkaç sorun bulunur: Zihin durumu bir midir yoksa zihin birçok kipe sahip bir yapı mıdır? Örneğin insan zihni ile Nilsson'un dile getirdiği gibi bir örümcek veya bitki zihni aynı şeyin iki farklı görünüşü müdür? Yoksa zihne sahip olmak zihne sahip olan her varlık tarafından aynı şekilde mi deneyimlenir? Öte yandan bir canlının zihne sahip olup olmadığı nasıl tayin edilebilir? Tüm bu sorunlar çözülse bile bir zihne ilineksel olan – zihin olmasa da yerine getirilebilecek zekâ gerektiren işler – ile bir zihne özsel olarak bağlı olan – yani zihinsiz gerçekleştirilemeyecek türden zekâ gerektiren işler – eylemleri nasıl tespit edebiliriz? Bu bitimli bir iş midir? Nilsson'un özellikle bu son itirazı bize elemeci fizikalist tavrı hatırlatır. Görüyoruz ki sembol sistem hipotezinin revize edilmesi gereken durumlarda bile bilgisayar bilimciler ve daha genelinde yapay zekâ projesi bakımından rol oynayan düşünür ve bilim insanları, *kuramı varlığın koşullarına uygun olarak tasarlamak yerine varlığın koşullarını öne sürülen kuramın koşullarına uydurmak* için çabalamaktadırlar.

Fiziksel sembol sistem hipoteziyle dile getirilen – veya daha genel olarak ele almamız gerekirse kuralları ve sembolleriyle bir mantıksal sistem bakımından – iddiaları yadsıyor değiliz. Ancak tıpkı Doğan'ın da dile getirdiği gibi, mantıksal olasılıkla ontolojik olasılık arasında bir fark bulunur. Bir iddia mantıksal olarak olanaklı olabilir ancak mühim olan şey mantıksal olarak olanaklı olan durumun ontolojik olarak da bir gerçekliğinin bulunmasıdır.[238] Çünkü Allen Newell ve Herbert

[238] Doğan, "Biyolojik Doğalcılık Ekseninde John Searle ve Zihin-Beden Problemi", s. 234.

Simon'un makineler için düşündükleri bilişsel yapının insanların bilişsel yapılarıyla aynı olduğunu kabul ettiğimiz durumda Searle'ün de işaret ettiği gibi bir otomatik kapının fotoelektrik hücresinden geçen emri anlaması* ile bizim örneğin İngilizceyi anlıyor oluşumuz arasında herhangi bir fark bulunmadığını kabul etmiş oluyoruz.[239] Ancak Choi, sezgisel çıkarımların fizikten halk psikolojisine kadar dünyanın nasıl işlediğine ilişkin geniş bir arka plan bilgisine dayandığını, sosyal dünyanın nasıl çalıştığına dair geniş kapsamlı sağduyu modellerine ihtiyacımız olduğunu dile getirir, böylece dilin yalnızca kelimeler ve kelimelerin imleri değil aynı zamanda karşılaşılan durumun bütünü bakımından bir bağlamı da kapsadığını ifade eder. Dolayısıyla türlü sembolcü yaklaşımlarla doğal dil arasında daima bir boşluk bulunacaktır çünkü insan düzeyindeki sezgisel çıkarımlar, nesneler, eylemler, konumlar, nitelikler ve duygular dahil olmak üzere çeşitli kavramların bir bütün olarak algılanmasını, bir bütün olarak kavranmasını gerektirmektedir. Bu anlayışa göre, biz insanlar bir cümleyi dile getirirken sadece cümlenin sonu bakımından anlamı düşünmeyiz, cümleyi dile getirme sürecinde kullanılan her bir kelime, cümlenin bitişi ve cümlenin sahip olması gereken anlam bakımından da değerlendirilir, böylece kelimeler cümlenin olası sezgisel anlamı bakımından da seçilmiş olur.[240] Dolayısıyla bizim sezgisel kavrayışımız önceden tanımlanmış tanımlar kümesinin sınırlılığını aşan, çıkarım yapmak bakımından tahmine dayalı, doğal dilin kapsamını aşan bir tarzdadır. Bu bakımdan örneğin bir makinenin bir görseli algılaması ile bir insanın aynı görseli algılaması arasında bir fark vardır. Çünkü bir görselin insan zihninde anlaşılır olması piksellerin algılanmasından öte bir durumdur. Bir görselin algılanması ve semantik içeriğin kavranması süreci algı bakımından olanaklı kılınan ancak algının da üzerinde gerçekleşen, görselin statik

* "Bennu anlıyor." cümlesi ile "Makine anlıyor." arasında biçimsel olarak bir fark bulunması gerektiğini ısrarla vurgulamak istiyoruz. Her iki cümlenin biçiminin "O anlıyor." şeklinde çıkarılabildiği açıktır. Okuyucu açısından da dikkat çekmek isteriz ki O zamirinin yerine koyulabilecek Bennu ve makine adları, yerine durdukları şeyler bakımından benzer varlık koşullarını karşılar mı? Bennu'nun anlıyor olması tıpkı Searle'ün de öne sürdüğü gibi yönelimsel ve failsi bir durumdur. Burada Bennu'nun deneyimine eşlik eden bir *ben* bulunur. Peki makinede gerçekleşen durumları makinede bir ben *varmış gibi*, onda bir yönelimsellik *bulunuyormuş gibi* ifade ediyor olmamız sorunlu değil mi? Makinenin emri anladığını söylediğimizde ona bir benlik, bir failsi nitelik kazandırmıyor muyuz ister istemez? Bu durumun nasıl önleneceği elbette bir dilbilim sorunudur da.

[239] John R. Searle, "Minds, brains, and programs", s. 419.

[240] Yejin Choi, "The Curious Case of Commonsense Intelligence", **Daedalus**, C. CLI, No: 2, 2022, s. 142-143.

durumunu aşan ve görseldeki ilişkilerin kavrandığı bir durumdur.[241] Dolayısıyla makine zekâsı gibi bir zekâda olan durum ancak ve ancak bir programcının kimi semboller ve mantıksal ifadeler ile dünyada kimi olgular arasında yaptığı eşleştirmelerden fazlası olamaz.[242]

Son tahlilde mantık temelli yapay zekâ çalışmaları nasıl yorumlanabilir? Mantık bilimi ve mantık biliminin yapay zekâ çalışmaları bağlamında bilgisayar bilimindeki kullanılışları olarak incelenirse bu çalışmalar iki şekilde yorumlanabilir. İlkin, mantığın gerçekliğimizdeki nesneleri ve onların ilişkilerini temsil etmek konusundaki gücü yadsınamaz. Ancak mantığın temsil gücünün, kullandığı temsilin, yerine durduğu gerçekliğin veya ilişkinin bizzat kendisi ile eş olmasını sağlayacak bir güç olduğunu kabul etmek de hatalıdır. Çünkü mantığın ilke ve ögelerinin gerçekliğimize sirayet eden bir rolünün bulunduğunu söylemek, mantık ilkelerinin ve ögelerinin değişkenliği karşısında gerçeklikteki nesnelerin de bu değişmeden etkileneceğini söylemeye gelir. Oysa tam aksine, nesneler ve nesneler arası ilişkiler gerçekliği, önermelerin doğruluk değerlerini vb. mantığa dayatmaktadır.[243] Üstelik mantık, gerçeklikte yer alan kimi olguların ve nesnelerin ifade edilmesi bakımından da sınırlılıklara sahiptir. Winston'a göre mantığın notasyonu bazı şeylerin mantıkta ifade edilmesine izin vermemektedir çünkü dünyadaki bazı olgular aksiyomlaştırılmaya direnirler.[244] Şimdiye kadar da dile getirdiğimiz gibi, sağduyu bilgisi veya sezgiye dayalı bilgi, mantıkça kapsanabilmenin henüz uzağındadır. Sağduyuya dayalı bilginin üstesinden gelebilmek için monoton-olmayan mantık gibi birtakım çözüm önerileri sağlanmış olsa da ne monoton-olmayan mantık bakımından sağduyu bilgisinin dile getirilmesi tarifi kolay ve güvenilirdir ne de herhangi bir insan, bir programın ihtiyaç duyacağı herhangi bir durum için gereken sağduyu bilgisini içeren bir veri havuzu oluşturmayı bilmektedir.[245] Dolayısıyla denebilir ki yapay zekâ projesi kapsamında makinelerin mantık temeline dayanarak inşa

---

[241] **A.e.**, s. 141-142.
[242] John R. Searle, "Is the Brain's Mind a Computer Program?", **Scientific American**, C. CCLXII, No: 1, 1990, s. 27.
[243] Steve Hanks ve Drew McDermott, "Default Reasoning, Nonmonotonic Logics, and the Frame Problem", **AAAI**, C. LXXXVI, 1986, s. 330.
[244] Winston, **Artificial Intelligence**, s. 303.
[245] John McCarthy, "Generality in Artificial Intelligence", **Communications of the ACM**, C. XXX, No: 12, 1987, s. 1030-1034.

edildiğinde insansı olan herhangi bir alanda tıpkı insanlar gibi eyleyebileceği yollu iddia yanlıştır. Çünkü bir program olarak yapay zekâlar, yeni ilkeler türetemezler. Yeni ilkeler – veya buluşçu mahiyet – bir programcı tarafından sağlanmak durumundadır.[246] Bu durumda, en insansı görünen yapay zekâ eylemi ne ontolojik açıdan *bu gerçeklikte* ve *şu andadır*, ne de epistemolojik açıdan *mantıkça ifade edilenden* daha fazlasıdır. Pylyshyn'in de ifade ettiği gibi, hesaplama bağlamında düşünüldüğünde mantıksal bir sistem içerisinde dile getirilenlerin, yani sembollerin, temsil ettikleri şeyler olan gerçeklikteki nesne ve nesneler arası ilişkilerin alanına bir erişimi bulunmaz.[247] O halde mantıkla başarılmaya çalışılan yapay zekâ projesi, mantıkla başarılabileceğini düşündüğü iddialarını yeniden gözden geçirmek durumundadır. Aksi halde yapay zekâ felsefesinde sağlamlaştırılmak istenen bu metafizik kabullerin sonucu, metafelsefi açıdan da açıktır ki salt bu kabullerin başarılabilmesi için yapay zekâ projesinin dayandığı temellerin sorunlarını yadsımak ve bu alanları varlık koşullarının ötesine taşımaya çalışmak olur. Biz bu tavrın felsefece olduğunu kabul etmiyoruz. Çünkü bir felsefi tavır, dayandığı temellerin sorunlarıyla yüzleşmek ve bu sorunlara ilişkin açıklama modelleri sunmak zorundadır. Aksi halde felsefenin içine düşeceği durum, tam da bilimler karşısında felsefenin getirilmek istendiği konum olarak, spekülatif tartışma ve konuşmaların alanı olmaktan başka bir şey olamaz.

1.2.3. Yapay Zekâ Çalışmalarının Doğa Bilimlerine Dayanan Temelleri

Çalışmanın bu alt bölümüne kadar, yapay zekâ çalışmalarının dayandığı matematiksel ve mantıksal temeller incelenmiştir. Bu incelemelerdeki amaç, insan zekâsının makinelerde gerçekleştirilebilmesinin koşullarının biçimselci temeller bakımından ele alınışlarını okuyucu açısından da görünür kılmaktır. Ancak bu alt bölümde görülecektir ki, yapay zekâ çalışmalarının yalnızca biçimsel bir ayağı bulunmamaktadır. Biçimsel bir zeminde inşa edilen yapay zekâlar, hem çıkış

[246] McCarthy, "Artificial Intelligence, Logic and Formalizing Common Sense".
[247] Zenon W. Pylyshyn, "Computation and cognition: issues in the foundations of cognitive science", **The Behavioral and Brain Sciences**, C. III, No: 1, 1980, s. 113.

noktaları olan insan zihninin ve insan fizyolojisinin etüt edilmesi bakımından hem de bu zihin ve fizyolojinin olanaklarının makinelerde, somut bir tarzda gerçekleştirilebilmesi bakımından da doğa bilimleriyle sıkı bir ilişki içerisindedir. Çünkü felsefe tarihi bakımından düşünülürse yapay zekâ çalışmaları, fizyolojik ve psikolojik olarak insan zihnini anlama çabalarımızdan ve zihin felsefesinin tarihi bakımından düşünüldüğünde ise insan zihnine ilişkin felsefi yaklaşımlarımızdan bir miras devralır. Böylece bilgisayar bilimi – bu bilimin dayandığı temeller bakımından da doğa bilimleri sayılabilir – ve felsefe arasındaki bağ sağlanmış olur.[248] İşte tam da burada açıktır ki felsefi çalışmaların içeriği ne sadece felsefi incelemelerin kendisi olabilir ne de doğa bilimleri ve diğer bilimler ya da zanaat alanları bakımından mevcut felsefi incelemenin konusu olan kavramla ilintili kazanımların, bir felsefi çalışmanın içerisinde dile getirilmesi felsefe dışı olarak kabul edilebilir. Çünkü her ne kadar yapay zekâ felsefesi, zihin felsefesinin bir alt alanı olarak kabul edilebilse de insan bilgisi, zihin üzerine yalnızca felsefe ile değil deney ve gözlem yoluyla da eğilir. Böylece sinirbilim, bilişsel bilim, psikoloji gibi alanlar da yapay zekâ çalışmalarında rol oynar.[249] İçeriğini psikoloji ve yapay zekâ gibi alanların oluşturduğu bilişsel bilim, makineler ve insanlarda zekâ üzerine araştırmaların yapıldığı bir alan olarak öne çıksa da[250], biz bu alt bölümde bilişsel bilimin bizzat kendisini değil ama bu bilimi de kapsayan bir tartışma konusu olarak beynin biyolojik mimarisi ile makinelerde bu mimarinin gerçekleştirilmesinin koşulları bakımından kimi tartışmaları ele almayı tercih edeceğiz.

Yapay zekâ çalışmalarında insan beyninden ilham alınır. Öyle ki insan beyninin işleyişinin anlaşılabilmesi için sinirbilimciler ve bilişsel bilimciler, insan beyninin en temel birimleri olan sinir hücrelerini ve onların yapısını modelleyerek bir benzetim süreci tasarlar.[251] Nitekim insan beyninin ve genel olarak insan varoluşunun ilham verici olması oldukça anlaşılabilir bir durumdur: İnsan, bedeni ve zihni ile bir sistem olarak kabul edilir böylece dış dünya ile iletişim kurabilen, dış

[248] Floridi, **Philosophy and Computing: An Introduction**, s. 18.
[249] Manzotti ve Tagliasco, "Artificial consciousness: A discipline between technological and theoretical obstacles", s. 110.
[250] Robert Audi, "cognitive science", **The Cambridge Dictionary of Philosophy**, 2. bs., New York, Cambridge University Press, 1999, s. 148.
[251] Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, s. 29.

dünyadan gelen uyarıları değerlendirebilen ve bunlara anlamlı cevaplar verebilen bir sistem olarak[252] onun, duyu verileri, anıları, duyguları ve eylemleriyle bir bütün olan deneyimi anlaşılmak istenir.[253] Ancak unutulmamalıdır ki insan belleğinde bu deneyimlerin nasıl depolandığı, henüz ilkelerini bildiğimiz bir durum değildir.[254] Bu belirsizliğe rağmen sistemleştirilen insan varoluşu ve yapay zekâlı varlıklar arasında bir koşutluk ilişkisi kurulmak istenir. Gözkan'ın da dile getirdiği gibi tüm insansı özelliklerin fiziğe tabi olan bedensel işlevlere indirgenebileceği yollu bir inanca sahipsek, buradan hareketle beynin, etten meydana gelmiş bir bilgisayar olduğunu ve yapay zekâların ise bu durumda henüz yeterince gelişmemiş bir insan prototipi olduğunu kabul ediyoruz demektir.[255] Gözkan'ın yapay zekâ felsefesi bakımından da değerli olan bu tespiti, bilgisayar biliminde de onaylanır. Nitekim Say, insan beyninin devamlı bir şekilde duyu organlarından gelen verileri alan, fizik kurallarına tabi bir mekanizmayla bu verileri işleyen ve hesapladığı çıktıyı da bedene sinirsel sinyaller yoluyla dağıtan bir *bilgisayar* olduğunu, bu bilgisayarın ise et ve kemikten yapılmış bir *otonom robot* olarak kabul edilebileceğini savunur.[256] Buradan hareketle Say, sinir hücrelerinin yaptığı girdi ve çıktı süreçlerinin, sinyallerin ağırlıklı toplamlarının hesaplanması ve bu toplamın büyüklüğü üzerinden etkileşmeleri yollu bir işlevi yerine getiren milyarlarca sinirsel birimin beyindeki gibi birbirlerine bağlanması yoluyla doğadakine benzer şekilde gerçekleştirilebilmesi sonucunda düşünen bir sistemin elde edilebileceğini iddia etmektedir.[257] Fakat buradaki asıl sorun, beyin ne türden bir bilimsel incelemenin konusu olursa olsun, onda herhangi bir failin, herhangi bir kırmızının, herhangi bir *Bennu'nun* olduğu görülemediğidir. Keza beyin, hiçbir parçası bakımından da düşünmeyi gerçekleştiren veya hissetmeyi gerçekleştiren kısımlar olarak düşünülemez.[258] İndirgemeci olan bu yaklaşım, bir sistemin daha iyi anlaşılabilmesi bakımından sistemin daha basit alt sistem ve parçalara bölünmesi yoluyla karşımıza çıkar, ancak Bingöl'ün de dile getirdiği gibi, bütün alt sistemlerin anlaşılması, bir sistemin kendisini anlamamız için yeterli

[252] Bingöl, "Bilgisayar Üzerine", s. 28.
[253] Zeman, **Bilinç Kullanım Kılavuzu**, s. 39.
[254] Bingöl, "Bilgisayar Üzerine", s. 38.
[255] Gözkan, "Bilgi, Bilinç ve Yapay Zekâ", s. 158.
[256] Say, **50 Soruda Yapay Zekâ**, s. 45.
[257] **A.e.**, s. 98.
[258] Dennett, **Bilinç Açıklanıyor**, s. 41.

olamaz. Özellikle doğrusal olmayan sistemlerde indirgemeci yaklaşım oldukça başarısızdır.[259]

Newell ve Simon, bir fenomenin açıklanmasının, o fenomeni kaçınılmaz bir şekilde meydana getiren ve o fenomenden daha basit olan mekanizmaların ifade edilmesi olduğunu söyler, böylece örneğin kimyasal reaksiyonların atomların fiziksel özellikleri ve davranışları ile açıklanabildiği bir durumda sinirsel süreçlerin, bu süreçleri oluşturan nöral ve sinaptik süreçlere indirgenmesi yoluyla açıklanabileceğini iddia eder.[260] Ancak okuyucu açısından da anımsanmalıdır ki bu çalışmanın fizikalizm üzerinde durduğumuz alt bölümünde, indirgemeci bir fizikalizmin sorunlarına değinilmiştir. Özellikle Nagel'in vurguladığı husus, fizikalizmin iddialarının ortaya konabilmesinin koşulu olarak zihinsel fenomenlerin, örneğin nitelceler gibi, fiziksel karşılıklarının söz konusu olabilmesi gerektiğidir.[261] Çünkü biz her ne kadar beynin çalışma ilkelerini ve fiziksel koşullarını bir makinede gerçekleştirebilsek ve davranışsal olarak da örneğin Bennu'nun ağlamasını bir makinenin ağlaması yoluyla mümkün kılabilsek bile bir makinenin Bennu'nun ağlarken sahip olduğu *fenomenal deneyime* sahip olup olmadığını ortaya koyamayız. Böylece Lamme'nin de ifade ettiği gibi, bir zihne birçok duyu verisi ulaşır ancak zihne gelen bu duyu verilerinden yalnızca bazıları, dilsel veya davranışsal unsurlarla dışsallaşır.[262] Dolayısıyla Jackson'un Fred'i veya Mary'sinin de esasen fizikalist tutuma karşı ortaya koymaya çalıştığı fikir, bir sisteme ilişkin fiziksel tüm koşulların bilindiği bir durumun bile o sistemin fenomenal durumları hakkında bir kavrayışa erişmemiz anlamına gelemeyeceğidir.[263] Ancak Walter Pitts ve Warren McCulloch'ın düşünceleri bu yönde değildir. Onlara göre belli girdilere göre belli çıktıları üretebilecek sinir ağlarının inşa edilmesi için genel metotlar aranmalıdır.[264] Nitekim Rosenblatt'a göre de yüksek düzeyli varlıkların zihinleri hakkında – örneğin

[259] Haluk Bingöl, "Karmaşık Ağlar Beyni Anlamamızda Yardımcı Olabilir mi?", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 304.
[260] Newell ve Simon, "The Simulation of Human Thought", s. 155.
[261] Nagel, "What Is It Like to Be a Bat?", s. 437.
[262] Victor A. F. Lamme, "Why visual attention and awareness are different", **Trends in Cognitive Sciences**, C. VII, No: 1, 2003, s. 12.
[263] Frank Jackson, "Epiphenomenal Qualia", **The Philosophical Quarterly**, C. XXXII, No: 127, 1982, s. 129-130.
[264] Walter Pitts ve Warren S. McCulloch, "How We Know Universals the Perception of Auditory and Visual Forms", **Bulletin of Mathematical Biophysics**, C. IX, 1947, s. 128.

insanlar – algısal tanıma, genelleme, düşünme vb. gibi yetilerin anlaşılmasının koşulları kimi soruların yanıtlanmasına bağlıdır: Fizik dünyadaki bilgi biyolojik bir sistemde nasıl duyumsanır ve kavranır? Bilgi hangi biçimde saklanır ve hatırlanır? Bellekte depolanan bilgi tanımayı ve davranışı nasıl etkiler?[265]

İnsan zihninin anlaşılabilmesi ve yukarıda dile getirilen soruların yanıtlanabilmesi için insan zihnin fizyolojisi incelenmelidir. Beyinlerimiz, bilgi iletimi sağlayan sinir hücreleri olarak nöronlardan oluşmaktadır. Nöronlar, akson adı verilen uzantıdan ve diğer nöronun aksonundan gelen sinyallerin gövdeye iletilmesini sağlayan dendritlerden meydana gelir. Bir nöronun aksonunun diğer nöronun dendritine temas etmesine ise sinaps denir. Nöronlar arasındaki bilgi transferi ise nöronal ateşlemeler yoluyla gerçekleşir[266], belli bir yer ve zamanda kurulan bu nöronal bağlantılar kullanım ve deneyimle sabitleşir, güçlenir.[267] İşte tüm deneyimlerimiz bu fizyolojik yapı üzerinde gerçekleşir. Kırmızı bir elmanın duyumsanması, güneş ışınlarının cildimizi ısıtmasına ilişkin his vb. gibi birçok fenomenal durum, beynimizdeki nöronal ateşlemeler ve ilişkiler vasıtasıyla olanaklı kılınır, fenomenal bilincimizde deneyimlenir. İnsan zihninin fizyolojik yapısına odaklanan yapay zekâ çalışmaları, ateşlemeler ve bağlantılar yoluyla zihinsel yaşantıyı olanaklı kılan nöronları, mantık kapıları şeklinde yüksek düzeyde idealize edilmiş ve basit kombinasyonları yerine getiren birimler olarak kabul eden bir yaklaşıma sahiptir.[268] Akman'a göre de sinir sistemimiz içerisinde bulunan sinirsel ateşlemeler, bir elektrik sisteminde olduğu gibi, evet-hayır ya da devre mantığı ile açık-kapalı biçimde gerçekleşmektedir.[269] Böylece insan zihninin fizyolojisi ve psikolojisi üzerinde şekillenen bu yaklaşım, zihin felsefesinde işlevselciliğin temel savlarından biridir. Ancak sinir hücrelerinin mantıksal kapılar olduğu ve mantıksal işlemlerin onlar üzerinde gerçekleştirilebileceği yapılar olarak kabul edilmeleri gerektiğine ilişkin bu görüş, bazı itirazlarla karşılaşmaktadır. Örneğin bazı düşünürler, insan davranışlarının ve insan zihninde gerçekleşmekte olan durumların,

[265] Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", s. 386.
[266] Toygar Akman, **Sibernetik: Dünü, Bugünü, Yarını**, İstanbul, Kaknüs Yayınları, 2003, s. 27.
[267] Barış Korkmaz, "Çocuklarda Nöronal ve Zihinsel Gelişim", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 66.
[268] Boden, "Artificial Intelligence".
[269] Akman, **Sibernetik: Dünü, Bugünü**, Yarını, s. 23.

bunların arkasında bulunan evrimsel kimi mekanizmalar kavranmadan gerçekleştirilmeye çalışılmasının uygun bir yöntem olmadığını savunmaktadırlar. Bu anlamda evrimsel yapay zekâ çalışmaları, yapay zekâları bir evrimsel süreçle geliştirmek için genetik algoritmaların kullanıldığı bir alandır. Genetik algoritmalar, her jenerasyonda programda rastgele değişiklikler gerçekleştirir ve en uyumlu sonuç korunur. Böylece insan zihninin makinelerde gerçekleştirilmesine ilişkin çalışmalara getirilen evrim yönünden eleştiriler aşılmak istenir.[270] Ancak unutulmamalıdır ki evrimsel süreç birçok unsurla gerçekleşir. Programlar üzerinde yalnızca rastgele unsurların farklılaşması yoluyla ilgili programın en uygun hale gelmesini sağlamak zekânın elde edilmesi için yeterli olmayabilir. Doğal ve yapay seçilim gibi birçok unsur evrimsel süreçte rol oynar. Dolayısıyla insan zekâsının, antropolojik verilerle de sabit olmak üzere, insanın çevresi, buluşçuluğu ve yaşama tarzı gibi birçok yapay seçilim unsuruyla olan ilişkisi ve tarihselliği yadsınamaz. Ancak evrimsel yapay zekâ çalışmalarıyla da bu unsurların nasıl kapsanacağı açıklanamaz.

İnsan beyninin fizyolojik yapısıyla, bu yapının makinelerde gerçekleştirilmesi için öne sürülen çözümler arasında kurulmaya çalışılan koşutluk ilişkisi, sıkça kategori hatasına düşmektedir. Örneğin insanın zihinselliğinin, tıpkı aerodinamik yasalarına benzer şekilde uymakta olan serçe ve uçak örneğindeki gibi farklı donanımlarda gerçekleştirilebileceğine ilişkin bir inanç bulunur. İnsanın zihinselliği, onun biyolojik doğasından soyutlanır, böylece beyin, zihinselliğin ortaya çıktığı *bir tür* donanım olarak kabul edilir. Zekânın hesaplama kuramı bir kere keşfedildiğinde ise serçelerin sahip olduğu tüylerin uçmayı gerçekleştirebilmek için ilineksel donanımlar olması gibi nöronlar ve sinapsların da zekâ için ilineksel olan donanımlar olduğunu söylememiz mümkün olabilir.[271] Mühim olan şey aerodinamik yasalarına uyan herhangi bir donanımın uçmayı gerçekleştirebildiği bir durumda zihnin hesaplama kuramını gerçekleştirebilen herhangi bir donanım mimarisinin zihni gerçekleştirebileceğine olan inançtır. Ancak burada şöyle bir sorun bulunur: Aerodinamiğin iki ayrı donanımda gerçekleşmesi ile var olduğu düşünülen *bir zihin*

[270] Boden, "Artificial Intelligence".
[271] Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, s. 31.

kuramının[*] iki ayrı donanımda gerçekleştirilebileceğine ilişkin koşutluk ilişkisi bir kategori hatasıdır. Çünkü aerodinamik yasalara uymak, yani uçabilmek, serçenin varlık koşullarından sadece bir tanesidir. Öte yandan uçak olmanın varlık koşullarından biri de uçabilmektir, bu doğru. Oysa bu iki var olanın varlık koşulları bakımından birçok farklı unsurdan da söz etmek mümkündür. Örneğin serçe memeli bir hayvan değildir, ötebilir vb. Uçak ise türlü teknolojik aksamları, metalik bedeni ve diğer türlü özellikleriyle apayrı varlık koşullarına sahiptir. Serçe ile uçağın aerodinamik yasalarına uygun varlık koşuluna sahip iki ayrı varlık olması, serçenin ve uçağın varlık koşulları bakımından paylaştıkları yalnızca bir özelliktir. Bu durum serçe hakkında veya uçak hakkındaki yargılarımızı, yine serçe veya uçak için genişletebileceğimiz anlamına gelmez. Dolayısıyla tıpkı aerodinamik yasalar gibi, insan zihninin de bir zihnin bir canlıda belirmesinin koşulları olarak *henüz keşfedilmeyi bekleyen* türlü yasalara sahip olduğunu söylemek, buradan hareketle de nöron ve sinapsların zihnin belirmesi için ilineksel olduklarını iddia etmek doğru değildir. Çünkü ne zihinde zihnin ortaya çıkmasının koşulları bakımından türlü yasaların olduğu ve işlediğine ilişkin bir durum söz konusudur ne de aerodinamik yasalarla zihinde işleyen düzenliliklerin eş düzeyde düzenlilikler olduğunu söylemek mümkündür. Çünkü aerodinamik yasalar bir varlığın yalnızca birkaç varlık koşulunu kapsayan mahiyetteyken bir zihne sahip olabilmek, zihne sahip var olanlar içerisinde o var olanların her türden varlık koşullarını olanaklı kılan ve en önde gelen koşullardan biridir. Uçamayan bir serçe, serçe olmaklığını kaybetmez ancak bir zihne sahip olmamak insan olmamayı sağlamak için yeterlidir. O halde insan zihninde düzenliliklerin olduğunu kabul etmek demek insan zihninin *kendisinin* bir düzenliliğin sonucunda belirdiğini söylemek demeye gelmez. Bir yerleşkenin fakülte olduğunu söylemekle, bu yerleşkede yer alan binalar incelendiğinde "Fakülte nerede?" türünden bir soru sormak, bir insanın bir zihne sahip olduğunu söylemekle onun fizyolojik yapısını ve bu yapının birimlerini inceleyip "İnsan zihni nerede?" diye sormak eş türden sorulardır. Bu durumda bir zihne sahip olmanın ne anlama

[*] Bu çalışmada şimdiye kadar anmakta olduğumuz "metafizik kabuller" ifadesiyle de sık sık dile getirmek istediğimiz durum tam olarak budur: Yapay zekâ felsefesindeki anlam bulanıklıklarının ve yapay zekâ çalışmalarındaki antropomorfik eğilimlerin tam da kökünde, zihnin bir program olduğu, bu programın şaşmaz bir biçimde işleyen ve keşfedilmeyi bekleyen *bir doğa yasası* olduğu sanılır. Çokça mühendisler tarafından taşınan bu metafizik kabulden arındırıldığı zaman felsefi işlevselci tavrın ilerleme kaydedebileceğine inanıyoruz.

geldiği ve nasıl tespit edildiği açıkça ifade edilmedikçe fizyolojik yapının işlevselci bir tavırla başka materyalde inşa edilmesi sonucu ortaya çıkacak bir yapının, davranış gösterse bile bir zihne sahip olacağını, düşüneceğini söylemek için bir sebebimiz bulunmaz.

McCulloch ve Pitts'in düşüncesinde nöronlar, nöral aktivitenin onlar üzerinde var olduğu birimler olarak, etkinlikleri önermelerle temsil edilebilecek ve nöronal reaksiyonların her birinin yalın önermelere karşılık geldiği yapılardır.[272] Nöral ağlar düşüncesinde de gördüğümüz bu yapı, nöronları temsil eden birbirine bağlı birimler arasında sayısallaştırılmış ağırlıkların bulunduğu, her bir birimin aktivasyon düzeylerinin girdinin gücüne göre belirlendiği, böylece çıktının girdi ve birim ağırlığına göre kararlaştırıldığı bir yapıdır.[273] Sayısallaştırılmış ve biçimselleştirilmiş nöral ağlar fikri, zihnin fenomenal unsurlarının fizikalist tavırla tahliyesini de mümkün kılmaktadır. Örneğin Manzotti ve Tagliasco, türlü işlevlerin yerine getirilebildiği bir yapıda fenomenal bir yanın bulunması için herhangi bir gerekçenin bulunmadığını, dış dünyada başlayıp beyinde biten bir algılama sürecinin bir birlik olarak kabul edilmesi gerektiğini, dolayısıyla fiziksel bir bakış açısından yaklaşıldığında algılanan ile beyindeki aktiviteler arasında kesintisiz bir fiziksel zincirin bulunduğunu iddia etmektedir. Bu durumda örneğin bir insan, bir çikolata tattığında, beyinde gerçekleşen şey yalnızca farklı özellikler taşıyan nöral örüntüler olmaktadır.[274] Burada Levine'in fizikalist açıklamalara ilişkin itirazlarını hatırlayabiliriz. İlk olarak Levine, izah gediği olarak da adlandırılan bir itirazı bize sunmaktadır: Tıpkı fizik ve kimya bilgimizin moleküllerin hareketi ve ısı arasındaki nedensel ilişkiyle alakalı olarak sahip olduğu açıklama gücü gibi zihnimizde gerçekleşen deneyimlerin, fenomenal unsurların, nöral ateşlemelerle olan nedensel ilişkileri de benzer şekilde açıklanabilir mi?[275] Öte yandan Levine fizikalist açıklamanın kabul edildiği bir durumda, örneğin ağrı hissi belli nöral ateşleme

[272] Warren S. McCulloch ve Walter Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity", **Bulletin of Mathematical Biology**, C. LII, No: ½, 1990, s. 100-101.
[273] Selmer Bringsjord ve Naveen Sundar Govindarajulu, "Artificial Intelligence", The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta ve Uri Nodelman, (Çevrimiçi) https://plato.stanford.edu/entries/artificial-intelligence/, 2 Ocak 2023.
[274] Manzotti ve Tagliasco, "Artificial consciousness: A discipline between technological and theoretical obstacles", s. 111-113.
[275] Levine, "Materialism and Qualia: The Explanatory Gap", s. 357.

örüntüleri ile eşlenirse ve bu örüntülerle açıklanmaya çalışılırsa, sinirsel bir yapıya sahip insan olmayan bir varlığın, kendi ifadesiyle bir uzaylı, bu belli sinir örüntülerine sahip olmadığını tespit ettiğimizde ise onun ağrı hissine sahip olup olamayacağını sorma hakkımızın bulunduğunu ortaya atmaktadır.[276] Gerçi her ne kadar Dennett türü kimi çözümler – örneğin nitelceleri reddetmek[277] gibi çözümler literatürde fizikalistlerin sıkça dile getirdiği itirazlardan biridir – önerilse de, Arıcı'nın da dile getirdiği gibi, var olan tek şey sadece nitelceler değildir, aynı zamanda farkındalık, benlik hissi vb. gibi diğer fenomenal unsurlar da zihinsel yaşantının bir parçasıdır.[278] Peki bu durumda doğa bilimleri açısından veya genel olarak fizikalist bir açıdan zihin nasıl ele alınabilir? Çünkü Penrose'un da dile getirdiği gibi düşünmenin, hissetmenin ya da usun ne olduğunu sorduğumuzda ve bu yapılarla fizyolojik durumlar arasındaki ilişkinin mahiyetini sorguladığımızda, bu yapıların fizyolojik durumlardan bağımsız olup olamayacağı ya da bu yapıların herhangi bir fizik yasaya bağımlı olup olmadığı türünden sorularla karşılaşmaktayız.[279]

Bilindiği gibi koşullu önermeler olarak tüm doğa yasaları, gelecekte yaşanması olası olaylar hakkında şimdinin bilgisine dayanarak kimi tahminlerde bulunmamıza izin vermektedir. Bu durumda doğa yasaları, şimdide yer alan tüm belirleyici etkenler tamamen bilindiğinde geleceği tahmin etmek için kullanılabilir.[280] Ancak sorulabilir: Şimdinin bilgisine dayanarak gelecek hakkında kimi tahminlerde bulunabilmemizi sağlayan bu düzenlilikler olarak doğa yasalarının, düzenli olduklarını söylemek ne anlama gelir? Doğanın düzenli olması, aynı nedenlerin aynı sonuçları sağlaması anlamına gelir.[281] Dolayısıyla mikro düzeyde nöronlara, sinapslara ve nörotransmiterlere sahip olduğumuz ve bu mikro düzey üzerinde işleyen zihinsel olanaklılıklar olarak bilinç, yönelimsellik, karar verme vb. gibi

[276] **A.e.**, s. 357-360.
[277] Çağlan Çınar Dilek, "Zor Problem ve Yüksek Düzey Düşünce Kuramı", **Phainomena**, No: 1, 2020, s. 144-145.
[278] Murat Arıcı, "Materyalizm, Fenomenal Özne ve Ontolojik Statüsü", **Beytülhikme**, C. IV, No: 2, 2014, s. 6.
[279] Penrose, **Kralın Yeni Usu**, s. 2.
[280] Eugene P. Wigner, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", s. 5.
[281] Oğuz İnel, **Özgür İrade Yanılsaması**, İstanbul, LEJAND Kitap, 2022.

fenomenal durumların da bulunduğu[282] hesaba katılırsa, aynı nedenlerin aynı sonuçları doğurması gerektiğiyle ilgili düşüncenin, yani determinizmin, sıkıca incelenmesi gerekir. Determinizm, her olayın daha önceki nedenlerin zorunlu olarak[*] bir sonucu olduğuna ilişkin bir görüştür.[283] Dolayısıyla geçmişte yaşanmış olaylar, deterministik bir evren anlayışına göre, gelecekteki olayları belirlemekte[284], böylece koşullar ve nedenler sabit tutulduğu sürece bir olayın sonucunun değişme olasılığı bulunmamaktadır.[285] Determinizmin dört türü olduğu söylenebilir: Mantıksal determinizm, teolojik determinizm, psikolojik determinizm, fiziksel determinizm. Mantıksal determinizme göre gelecek, zaten var olan bir geçmiş gibi halihazırda sabittir ve değişmeyecektir. "Kader" kavramıyla karşılamaya çalıştığımız anlam, mantıksal determinizmde bütünüyle karşılanır. Teolojik determinizm ise Tanrı'nın her şeye gücü yeten[**] mahiyetinden güç alan ve Tanrı'nın bilgisinin evrende her zamanı ve her mekânı kuşatması[***] zorunluluğuna dayanan bir anlayıştır. Psikolojik determinizm, belli psikolojik düzenliliklerin, yani yasalılıkların, var olduğunu, bunlardan hareketle ise failin gelecekteki yaşamında geçmişinde yaşamış olduğu deneyimlerinden hareketle verebileceği tepkilerin belirlenmesine ilişkin bir yaklaşımdır. Fiziksel determinizm ise kimi fiziksel doğa yasalarının var olduğunu, var olanların ise bu doğa yasalarına tam bir uyumlulukla devindiklerini ve bunlara uyum gösterdiklerini ortaya koyar. Determinizmin geçmişin tam bilgisi verildiğinde gelecekle ilgili bilgiye sahip olunabileceği yollu yorumunun temeli determinizmin fiziksel yaklaşımıdır.[286] İndeterminizm ise, determinizmin karşıtı olarak, olayların bir düzenlilik göstermediği, neden ve sonuç ilişkilerinin rastgele yaşandığına ilişkin görüştür. Roskies'e göre iki tür indeterminizmden söz edilebilir: Görünürde

---

[282] Searle, "Free Will as a Problem in Neurobiology", s. 503.

[*] Okuyucu açısından açık kılmalıyız ki çalışmanın bu bölümünden sonra kullanacağımız "determinizm", "indeterminizm" vb. gibi kavramları Türkçeleştirerek de – belirlenimcilik, belirlenmezcilik vs. – kullanabilirdik. Ancak amacımız, oldukça yalın bir arzu olarak, çalışmamızın okuma kolaylığını akıcılık açısından sağlamaktır. Böylece çalışmamızın buradan sonraki kısmında bu kavramlar da dahil olmak üzere kimi kavramları Türkçe karşılıklarını kullanmadan metin içerisinde kullanmayı tercih edeceğiz.

[283] İnel, **Özgür İrade Yanılsaması**, s. 11.

[284] Seth Lloyd, "A Turing test for free will", **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, No: 370, 2012, s. 3598.

[285] Doğan, "Yapay Zekâ ve Özgür İrade: Yapay Özgür İradenin İmkânı", s. 794.

[**] Lat. **Omnipotent**

[***] Lat. **Providentia**

[286] J. R. Lucas, **The Freedom of the Will**, Oxford, Clarendon Press, 1970, s. 65-66.

indeterminist sistemler, metafizik olarak indeterminist sistemler. Görünürde indeterminist sistemler ilk izlenimde bir kurallılığa sahip olmadığı düşünülebilecek ancak daha detayda incelenirse determinist bir şekilde çalışan, karmaşıklık düzeyleri yüksek sistemlerdir.[*] Metafizik olarak indeterminist sistemler ise tam anlamıyla indeterministtir. Bu sistemdeki olaylar tamamen rastgele yaşanmaktadır ve olaylar arasında hiçbir düzenlilik ve kurallılık bulunmamaktadır.[287]

İnsan beyninin fizyolojik yapısı doğa bilimleri tarafından incelendiğinde, doğa bilimleri söz konusu olduğunda tartışmalı bir husus olan determinizm-indeterminizm tartışması bizi karşılar. Klasik fiziğin deterministik olduğu görülebilir. Böylece klasik fiziğe göre materyal evren, zaman bakımından evrenin daha önceki durumu tarafından fizik yasalarına göre belirlenir.[288] Klasik fizik, görüleceği gibi sıkı bir determinist anlayışı içerir. Bu durumda beynin fizyolojisinin klasik fizikle incelenmesi ve bunun makinelerde gerçekleştirilebilmesi uğruna, makinelerin katı kuralların birbirini izlediği sistemler olduğu yollu görüşün elenmesi gerekliliği ortaya çıkar. Determinizm içerisinde insan zihnini ve pek tabi olarak bu zihnin olanaklarını kendisine örnek kabul eden yapay zekâ çalışmalarının, kimi belirsizlikleri içeren yapısının muhafaza edilebilmesi için bağdaşırcı[**] kimi girişimler söz konusu olmuştur. Bağdaşırcılık, insan zihnindeki bilinç, özgür irade vb. gibi deterministik şekilde işlemeyen unsurların da güvencesini sağlayan bir görüş olarak, deterministik yasalara aykırı davranma özgürlüğümüzün bulunduğunu iddia eder.[289] Dolayısıyla ılımlı determinizm olarak da adlandırılabilen bağdaşırcı anlayışa göre, insanlar her zaman neden-sonuç ilişkisine uygun olarak eylemezler. İnsanda, fizik dünya için geçerli olan nedenselliğin işlemediği bir alanın mevcut olduğunu söyleyebiliriz. Bu

[*] Roskies, yapay zekâ sistemlerinin görünürde indeterminist olan ama ilkece determinist sistemler olduğunu savunur. Örneğin rastgele sayı üreten bilgisayar programları, bu sayıları üreten algoritmalar bilinmediği sürece indeterminist bir tarzda bu sayıları üretir denebilir. Ancak algoritmalar bir kere bilindiğinde, bu programın sayıları üretirken rastgelelik unsuru içermediği ve determinist bir tarzda bu sayıları ürettiği görülebilir. Bkz.: Adina L. Roskies, "Can Neuroscience Resolve Issues about Free Will?", **Moral Psychology: Free Will and Moral Responsibility**, ed. Walter Sinnott-Armstrong, C. IV, Cambridge, The MIT Press, 2014, s. 105.

[287] **A.e.**

[288] David Hodgson, "Quantum Physics, Consciousness, and Free Will", **The Oxford Handbook of Free Will**, ed. Robert Kane, New York, Oxford University Press, 2002, s. 85-86.

[**] İng. **Compatibilist**

[289] **A.e.**, s. 86.

görüşe göre insan fizik dünyada, fiziksel yasalara tam bir uygunluk gösteren kimi yanlarına ek olarak bu yasalara uymayabilen kimi yanları da kendinde taşır.[290]

Determinizm-indeterminizm tartışmasının determinizm lehine çözümlenebileceğini savunan görüşler bulunur. Örneğin Denkel, determinizmin katı deterministik mahiyetinden arındırılmış bir yorumunu – determinizmin olasılıkçı yaklaşımı kuşatabilecek ve fiziksel doğa içerisinde bu durumu dışlamayacak mahiyetlerinin bulunduğunu – savunur.[291] Denkel'e göre bir genellemeye dayanarak şimdiki bir olayın gelecekteki belirli türden olayları doğuracağını iddia etmenin, ereksel açıklama girişimlerinden hiçbir farkı bulunmaz. Denkel açısından "Her olayın belirli bir etkisi vardır." önermesi ile "Her olayın belirli bir ereği vardır." önermesi arasında bir ayrım bulunmaz. Buradaki ayrım her iki önermenin bir duruma ilişkin genellikten hareketle kimi açıklamalar getiriyor olmalarına karşın ilgili olayı gelecekteki mi yoksa geçmişteki eşleşiğe mi dayandırmalarında yatar. Açıktır ki gelecekteki eşleşiğe dayandırılan olay, henüz olmamış, henüz yaşanmamış ve gözlemlenebilir olmayan bir gerçekliğe dayandırılmak istenmektedir. Gelecekteki eşleşik olaya ilişkin durum bir açıklama değil ancak ve ancak bir tahmin olabilir ve burada dile getirilen verinin değeri ontik değil epistemik bir değere sahiptir.[292]

Determinist kabul içerisinde yer alan hem erekselci hem olasılıksal girişimler, determinizmin kabul edebileceği bir tavır olamaz. Dolayısıyla olasılıksal nedenselliği, determinizmin çürütülmesi olarak değil ama erekselci bir determinizmin çürütülmesi olarak kabul edebiliriz. Çünkü ontik anlamdaki olaylar için belirli olmak başkadır, bu olayların henüz var olmadıkları bir durumda nasıl gerçekleşeceklerini düşüncede ortaya koyabilmek başkadır. Ontik anlamda durumların gelecekte ne olacaklarının belirli olduğunu iddia etmek, kaderciliktir.[293] Denkel'in epistemik belirlenimcilik olarak adlandırdığı bu yaklaşım, ontik sonuçlar içermemekle birlikte determinizmi tahmin edilebilirlik düzeyine indirgemektedir. Böylece epistemik belirlenimciliğe göre olan tek şey gelecekteki henüz

[290] E. Funda Neslioğlu, "İrade Özgürlüğü Sorununun Çözümsüzlüğü Üzerine", **Felsefe Dünyası**, C.I, No: 47, 2008, s. 182.
[291] Arda Denkel, "Belirlenim, Neden ve Değişim", **Düşünceler ve Gerekçeler: Felsefe Yazıları**, ed. Ahmet Haluk Atalay, C. II, İstanbul, Göçebe Yayınları, 1997, s. 61.
[292] **A.e.**, s. 66-68.
[293] **A.e.**, s. 71-73.

gerçekleşmemiş olayların hesaplanması ya da geçmişte gerçekleşmiş olaylar dizisinden, yine geçmişte gerçekleşmiş başka olayların türetilmesi süreçleri bakımından bir tahminden fazlası değildir.[294] Ancak Denkel'in bağdaşırcı yaklaşımı iki şekilde eleştirilebilir. İlkin Denkel'in, determinizmin olasılıkçı yaklaşımı kapsadığı yollu görüşü kabul edilirse, insan zihninin bilinç, özgür irade gibi kimi fenomenal unsurları bakımından determinizmin bu unsurlarla bir uyum sağladığı görülebilir. Ancak Inwagen'e göre, ne determinizm özgür irade ile bağdaşabilir – çünkü determinist bir evrende bir fizik var olan olarak insanın da bu evrendeki kurallılığa uyması gerekiyor – ne de özgür irade olasılıkçı bir determinizm ya da bir indeterminizm – görüleceği üzere kaos ve kuantum kuramlarının da yardımıyla – ile bağdaşabilir.[295] Çünkü zihin argümanı[*] bakımından düşünülürse örneğin özgür irade için indeterminist koşullar gerekiyorsa, o halde fail nasıl kendi eylemlerini yine kendisine dayanarak belirleyebilir? Belirleyemez ise o halde tüm eylemlerin şans eseri gerçekleşiyor olması gerekir.[296] Ancak açıktır ki eylemlerimiz bir düzenlilik gösterir. Bu durumda zihin durumlarının deterministik veya indeterministik şekilde çalışıp çalışmadığı problemi, ne determinizmin olasılıkçı bir mahiyete kavuşturulmasıyla çözülebilir ne de mevcut durumda indeterminist açıklamalar zihin durumlarıyla bağdaşır. Öte yandan Denkel'in epistemik belirlenimciliği, olasılıkçı mahiyetinden ötürü, şeylerin *neden olabilmek* bakımından özelliklerini de yadsır.

Neden ve etki arasındaki ilişkinin zorunlu olmadığı Hume'dan beri bilinmektedir. Nitekim Denkel'in epistemik belirlenimciliğinin itirazlarından biri de neden ve etki arasında zorunlu bir bağlantı bulunduğu, bu zorunlu bağlantıdan ötürü determinizmin şimdinin veya geçmişin herhangi bir bilgisi sağlandığında geleceğin de bilgisine bizi kavuşturabileceğine ilişkin inançtır. Denkel'in bu itirazı yerinde bir itiraz olsa da epistemik belirlenimciliğin, determinizmi gelecek konusunda olasılıksal sonuçlar üreten bir konuma getirmesi, nedensellik düşüncesinin kendisini ve fizik nesneler arasındaki neden olabilmeye ilişkin mahiyetleri ihlal etmektedir. Bir örnek üzerinden düşünelim: Bennu, şimdiki yaşına kadar sayısız elmanın 4 metre

[294] **A.e.**, s. 85-86.
[295] Peter Van Inwagen, "Free Will Remains a Mystery", **The Oxford Handbook of Free Will**, ed. Robert Kane, New York, Oxford University Press, 2002, s. 158.
[*] İng. **Mind Argument**
[296] **A.e.**, s. 168.

yükseklikten sert bir zemine bırakıldığında hasar gördüğünü deneyimlemiş olsun. Böylece Bennu, bir elmanın 4 metre gibi bir yükseklikten bırakıldığında sert bir zeminle temas etmesi durumunda elmada hasar gerçekleşeceğine ilişkin bir neden etki ilişkisi kurmuş olsun. Epistemik belirlenimci argümanın varsayımlarını hatırlarsak, nedenler ve etkiler arasındaki bağın zorunlu olduğunu söylemek olanaklı değildir, bu doğru. Nitekim bu durum David Hume'un nedensellik anlayışından beri bilinmektedir. Ancak Denkel'in epistemik belirlenimciliğinin bu varsayımının aksine Hume, nedenselliğin neden ve etki arasındaki zorunlu bağlantısının kurulmasının insan zihninde gerçekleştiğini, neden ve etki arasındaki zorunluluğun ne nedende ne de etkide içkin olduğunu söylemesine karşın *nedenselliğin kendisini yadsımaz* ancak neden ve etki arasındaki zorunlu olduğu düşünülen ilişkinin mahiyeti psikolojik olanla açıklar. Oysa Denkel'in epistemik belirlenimciliğinde, determinizmin gelecekle ilgili yargılarının olasılıksal olması neden ve etki arasındaki bağlantının zorunluluğuna bir şerh düşüyor olsa da *neden olabilmeyi yadsımak* bakımından bir problem taşımaktadır. Örneğin, Bennu'nun belli bir yükseklikten bırakılan bir elmanın zemine temas ettiğinde zarar göreceğine ilişkin kurduğu neden-etki ilişkisini ele alalım. Zeminin elmaya bir direnç göstermesi, elmanın gösterilen bu direnç karşısında formunu koruyamayacak kadar az dirence sahip olması gibi birçok durum iki fizik nesnenin de fizik nesne olmaklık bakımından birtakım ortak özellikleri taşıdığını bize açıkça gösterir. Böylece örneğin bir zeminin geçirmezlik özelliğini bıraktığını ama o esnada bir elmanın geçirmezlik özelliğine sahip olmayı sürdürebileceğini söyleyebilmemiz için bir gerekçemiz bulunmaz. Çünkü hem elma hem de zemin bir fizik nesnedir. Aynı türden varlıkların, en azından fizik nesne olmaklıkları bakımından, farklı yasalılıklara ve farklı varlık koşullarına sahip olacaklarını söylemek doğru olamaz. Fizikalizmin varlık anlayışını hatırlarsak var olanların tümü, var olmak bakımından maddi bir temele sahiptir. O halde maddi temele sahip her varlığın paylaştığı genel ilkeler ve yüklemler olduğu açıktır. Üstelik aynı türden varlıkların, aynı mekân ve zamanda zıt özelliklere – örneğin bir kayanın geçirmezlik özelliğini yitirmesi ama aynı esnada elmanın, içinden geçmek isteyen nesnelere karşı direnç uyguluyor olması – sahip olabileceğini söylemek mümkün değildir. Böylece neden olabilmenin de fizik nesne olmanın varlık koşullarından biri olarak bir özellik olduğunu iddia ediyoruz. Bu iddiamız neden ve etki arasındaki

ilişkiyi zorunlu kılmak uğruna değil, aksine bu ilişkinin zorunlu olmadığını da kapsar bir tarzda bulunur böylece örneğin Denkel'in epistemik belirimciliğinin olasılıksal mahiyeti bir kere kabul edildiğinde elmanın zemine teması sonucunda elmada yaşanacak bir deformasyon – olasılıksal olarak yaklaşıldığı için – yaşanmayabilecektir, bu doğru. Ancak sorulabilir: Tam da burada, bir zemin ve bir elma arasında gerçekleşen neden ve etki ilişkisinin bir gerçekleşmesi olarak elmada yaşanacak deformasyon gerçekleşmezse, her iki fizik nesnenin *neden olabilmek* gibi taşıdığı bir özelliği nasıl gerçekleşecektir? Çünkü bir elma ve bir zemin arasında yaşanacak neden ve etki ilişkisinin olasılıksal bir sonucu olarak elmanın zarar görmediği bir senaryo düşünülebilir. O halde beklenen – elmanın zarar görmesi – neden ve etki ilişkisi gözlemlenememiş olur. Ancak elmanın tam da zemine temas ettiği sırada elma ya da zeminin birbirleri üzerinde en az bir neden ve etki ilişkisi kurmaları, onların iki aynı türden varlık olarak düşünüldüğünde bunun zorunlu bir sonucu olamaz mı? Elma zemine temas ettiğinde deforme olmayabilir ancak ısınabilir ya da ısınmayabilir. Yine aynı elma, bırakıldığı durumdan daha da yükseğe erişecek şekilde zeminden sekebilir. Bu gibi sonsuz durum arasından yalnızca çok sık gerçekleşiyor diye elmanın deforme olmasını beklemek, ama bunu olasılıksal olarak ummak, bu yüksek olasılık bir kere gerçekleşmediğinde ise neden olabilmenin kendisini fizik nesne olmanın doğasından yadsımak için herhangi bir gerekçemiz de bulunmuyor. Dolayısıyla determinizm içerisindeki olasılıksal sonuçları kapsayan çözümler, zihinsel fenomenlerin olasılıksal yaklaşımlar tarafından açıklanamaması ve neden olabilmenin yadsınması bakımından kimi sorunlar taşımaktadır. Böylece insan beyninin fizyolojik yapısı hakkındaki determinizm-indeterminizm tartışması henüz çözüme kavuşmuş bir tartışma değildir. Doğa bilimlerinin insan beyninin fizyolojik koşulları üzerine araştırmalarında ise determinizm, ne insan zihninin fenomenal unsurlarının elenmesi bakımından ne de insan zihninin çalışma ilkeleri bakımından çözüme kavuşturucu bir tavır olarak ele alınabilir. Bu durumda nöral ağlar üzerinden insan beyninin fizyolojik yapısının taklit edilmesi, bu yolla da nöral ağlarda bir zihin olarak düşünen bir benin gerçekleştirilebileceğini savunmak amacıyla bilimin taşıyıcı unsurlarından biri olan determinizmin, insansı unsurların sahip olduğu indeterminist olanakları kapsamak bakımından yetersiz olduğunu görüyoruz.

Nöral ağlar düşüncesinde insansı olan zihinsel fenomenlerin gerçekleşebileceğine ilişkin düşünceler kuantum ve kaos kuramları tarafından da desteklenir. Penrose zihnin çalışma mekanizması ile kuantum mekanik sistemler arasında bir analojinin kurulma olasılığından bahsetmektedir.[297] Penrose bu önerisinde haksız değildir. Çünkü kuantum kuramı, Newton fiziğinin ortaya koyduğu katı bir kurallılıkla işleyen evren anlayışımızı sarsmış, atom ve atomaltı parçacıkların davranışları üzerindeki yargılarımızın Newton fiziğinin olanakları ile açıklanamayacağını göstermiştir. Ancak kuantum kuramının bu itirazı, doğanın deterministik mi yoksa rastlantısal mı davrandığına ilişkin bir sorunun çözümü olmamıştır.[298] Çünkü kuantum kuramına göre, bir sistemin atomaltı davranışı artık farklı davranma durumları arasındaki olasılıklarla ifade edilir. Böylece daha önceleri hangi durumların deterministik olarak yaşanacağına ilişkin hesaplama anlayışından, bu durumlardan hangilerinin gerçekleşme olasılıklarının daha yüksek olduğunu tespit etmeye doğru bir anlayışa geçmiş bulunuruz.[299] Bunun sebebi, klasik mekanikte ölçülen dinamik değişkenin belirli değerini kullanarak ölçme öncesindeki belirli değerini bulabiliyor olmamıza karşın, kuantum mekaniğinde ise ölçme öncesi ile sonrasını veren tasvirler arasında kategorik bir ayrım yapılması gerekmesidir. Çünkü kuantum mekaniği içerisinde düşünülen ölçme bakımından, ölçülen dinamik değişkenin ölçme sonrasındaki belirli bir değerinin kullanılmasıyla, ölçme öncesine dair belirli ve kesin bir değeri bulmak mümkün değildir. Burada ölçme öncesi ve sonrası üzerinde olasılıklar bulunduğu söylemek mümkündür. Dolayısıyla bir durumun ölçülmesi, kuantum kuramında ilgili durumun potansiyel durumdan gerçek duruma dönüşmesini sağlar, kuantum kuramında gözlem öncesi sonuç kestirilemez.[300] Bu özellikleriyle kuantum kuramı, insan zihnine ilişkin problemlerden biri olan özgür irade probleminde de sıkça kullanılmakta olan bir kuramdır. Sebebi açıktır: Kuantum kuramı, evrenin katı kurallılığına karşı çıkmakta ve olasılıksal, rastlantısal kimi durumların olanağının bulunduğunu öne sürmektedir.

[297] R. Ömür Akyüz, "Zihnin Fiziği?", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 133.
[298] Yağmur Denizhan, "Beyin Bağlamında Kaotik Sistemlere Bakış", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 168-169.
[299] Ömür Akyüz, "Beyin, Zihin, Zaman, Fizik (Kaos, Kuantum Mekaniği vs.)", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, Pan Yayıncılık, İstanbul, 2012, s. 208-211.
[300] Yalçın Koç, **Determinizm ve Mekân**, İstanbul, Boğaziçi Üniversitesi Yayınları, 1984, s. 12-13.

İnsanın iradesinin özgür olabilmesinin koşullarından biri de insan üzerinde etkiyen fizik yasaların, insanda mevcut bir irade tarafından denetlenebilmesi, ona aykırı şekilde davranılabilmesidir. Ancak felsefeciler ve nörobilimcilere göre özgür iradenin rastgele gerçekleşen olaylar neticesinde belirdiğini söylemek iradeyi özgür kılmak için yeterli değildir.[301] Kuantum kuramının nöral ağlar düşüncesi için de bir temel olduğu düşünülebilir. Çünkü nöral ağlar düşüncesine getirilen eleştirilerden biri de onun sıkı bir deterministik sistem olduğudur. Çoğu zaman bilgisayar bilimciler tarafından itiraz edilen bu eleştiriye yanıt olarak, sisteme esneklik kazandıracak unsurlar öne sürülür. Bunlardan biri olarak kuantum kuramının kimi savları dile getirilir. Oysa kuantum kuramının özgür irade problemi bakımından yüzleştiği sorunlar, nöral ağlar düşüncesi bakımından da geçerlidir.

Kuantum kuramına ek olarak kaos kuramı da öne sürülmektedir. Kaos kuramının insan zihni bakımından önemi, benzer fizyolojik mimariye sahip olan beyinlerimizin benzer etkilenimler karşısında gösterdiği birbirinden oldukça uzak davranışları açıklamaya çalışmasıdır. Kaotik sistemlerin yaklaşımı şudur: Klasik yaklaşıma göre eğer bir sistemin deterministik ilkelere göre çalıştığı biliniyorsa, sistemin ne türden girdiler karşısında hangi çıktıları verebileceğini hesaplamak mümkündür. Ancak kaotik sistemlerde durum değişir. Kaos kuramı kestirilemez determinizm yaklaşımını sunmaktadır. Bu yaklaşıma göre kaotik sistemler içerisinde gerçekleşen durumlar birbiriyle nedensel olarak ilintili olmalarına karşın, her bir olayın nedenselliğinin sistemin karmaşıklığından dolayı izlenememesinden ötürü tespit edilemeyeceğini, böylece sistem içerisinde işleyen kurallılıkları saptayamayacağımızı söylememiz mümkündür.[302] Beyinlerimizin de bir fizik nesne olarak doğada işleyen deterministik ilkelere iye olmak zorunda olduğu düşünülür ancak davranışsal olarak bakıldığında beyinlerimizin determinizme tam bir uygunluk göstermediği açıktır. Böylece kaos kuramı üzerinden beyinlerimiz, kaotik sistemler olarak yorumlanmak istenir. Beynimizde kuralsızlıktan ziyade örtük bir kurallılığın, Denizhan'ın ifadesiyle kestirilemez bir determinizmin, olduğu varsayılır. Kaos kuramının ilgi çekiciliği açıktır: Kaos denklemlerindeki en ufak değişiklikler, sonuç

[301] Hodgson, "Quantum Physics, Consciousness, and Free Will", s. 86.
[302] Denizhan, "Beyin Bağlamında Kaotik Sistemlere Bakış", s. 169.

bakımından oldukça farklı çıktıların ortaya çıkmasını sağlamaktadır.[303] Beyinlerimizin de benzer girdiler olarak düşünülseler bile oldukça ufak farklara sahip duyu verileri bakımından, sonuçları oldukça farklı olan çıktılar üreten ve kaos denklemleriyle ifade edilen kaos sistemlerine benzediği düşünülür.

Nöral ağlar düşüncesiyle inorganik materyaller üzerinde gerçekleştirilmeye çalışılan insan zihninin fizyolojik yapısı, maddenin doğasından ötürü kimi problemlerle karşılaşır. Determinizm problemi üzerinde durulmuştur. Kuantum ve kaos kuramları ile de açıklanmaya çalışılan, böylece silikon mimaride gerçekleştirilecek insan beyni benzeri yapının çeşitli karmaşıklık ve atomaltı şartları sağladığı sürece bir zihne sahip olabileceğine ilişkin inanç da birtakım problemlerle yüzleşmek durumundadır. Örneğin Roskies'e göre kaotik sistemler – daha önce dile getirdiğimiz görünürde veya metafizik olarak determinist veya indeterminist sistemler ayrımı hatırlanırsa – görünürde indeterminist ancak metafizik olarak determinist sistemlerdir. Kaotik sistemler dışardan gözlemlendiğinde onların indeterminist bir şekilde çalıştığı düşünülebilir çünkü gözlemcinin mevcut sistemde gerçekleşen her olaya ilişkin bilgisi bulunmamaktadır. Ancak Roskies, herhangi bir kaotik sistem içerisinde de biz her ne kadar gerçekleşen olayların tümünün ayırdında olamıyor olsak bile bu olayların deterministik ilkelere göre işlediğini vurgular. Dolayısıyla kaotik bir sistemin herhangi bir zamanında gerçekleşmekte olan tüm olaylar bilinirse, kaotik sistemlerin deterministik sistemler olduğu görülebilir.[304] Kuantum kuramı ise insan zihnindeki kimi fenomenlerin doğası üzerinden eleştirilir. Searle'e göre kuantum kuramı özgür irade problemine bir çözüm sunamamaktadır çünkü kuantum kuramı rastgelelilikler sunmaktadır. Ancak dikkat edilirse eylemlerimizin özgür olduğu savı ile eylemlerimizin rastgele meydana geldiği savları iki ayrı savdır.[305]

Yapay zekâ çalışmalarının doğa bilimlerine dayanan temelleriyle alakalı şimdiye kadar üzerinde durduğumuz noktalardan da anlaşılabilir ki, insan fizyolojisinin bilimler tarafından etüt edilmesi ve yeniden inşasına ilişkin süreçler, metodolojik ve felsefi sorunlar taşır. Burada amacımız bilimlere yön veya tanım

[303] Akyüz, "Beyin, Zihin, Zaman, Fizik (Kaos, Kuantum Mekaniği vs.)", s. 211.
[304] Roskies, "Can Neuroscience Resolve Issues About Free Will?", s. 106.
[305] Searle, "Free Will as a Problem in Neurobiology", s. 495.

verme girişimi değildir. Ancak Denkel'in de ifade ettiği gibi, nasıl ki felsefe içerisinde doğruluk değerlerini felsefeden bağımsız bir bilmeye göre onayladığımız birçok ifade varsa, benzer şekilde bilimler içerisinde de doğruluk değerleri yine bilimlere dayanarak ortaya konamayacak kimi ifadeler bulunur. İşte bilimlerin dayandığı bu ifadeler felsefenin konu ve nesneleri olarak sayılabilirler.[306] Dolayısıyla, bu alt bölümde de ele alındığı üzere, bilimler tarafından ele alınan ancak felsefi yükleri bakımından bilimlerce çözüme kavuşturulamayan kimi problemlerin olduğunu söylemek felsefe üzerinden bilimsel çalışmaların yadsınması değil, tam aksine onlara felsefi kıymetinin verilmesi ve insan fizyolojisinin incelenmesi uğruna bu çabanın felsefi statüsünün hatırlatılmasıdır. Nöral ağlar üzerinden gerçekleştirilmek istenen zihin ve zihinsel fenomenlere ilişkin determinizm, kuantum ve kaos kuramları üzerinden taşınan problemleri dile getirmemizin sebebi de tam olarak budur: Bilimler ve onların kavramsal çerçevesi içerisinde oluşturulan tasvirler, doğadaki nesnelerin *temel tasvirleri* değildir.[307] Nöral ağlar düşüncesinin bilimler temelinde de savunulması sürecinde sıklıkla yapılan hatalardan biri de budur. İnsan zihninin nöral süreçlerinin modellenmesi ayrı bir şeydir, bu modellenen ağın maddi zeminde yeniden gerçekleştirilmesi farklı bir şeydir. Belirtmeliyiz ki örneğin maddenin fizik kurallarına uygun bir şekilde düzenlendiğinde hatırlama, hesaplama ve öğrenebilme gibi kimi yetileri elde edebileceğini savunmak[308] türünden açıklama girişimleri bu çalışmanın kapsamında yadsınmaz. Yalnızca bu açıklama girişimlerinin, tıpkı Koç'un da dile getirdiği gibi, nesnelerin veya olguların temel tasvirleri olduğunu kabul etmek, ama bu kabulü ederken bu tavrın şimdiye gelesiye bu çalışma içerisinde aktardığımız temeller bakımından, örneğin yapay zekâ projesini besleyen unsurların problemlerini görmezden gelmenin, felsefece bir tavır olmadığına işaret edilir.

[306] Denkel, "Belirlenim, Neden ve Değişim", s. 83-84.
[307] Koç, **Determinizm ve Mekân**, s. 17.
[308] Tegmark, **Yaşam 3.0**, s. 109.

1.2.4. Bilimler Temelinde Ulaşılmaya Çalışılan Bir Metafizik Tasarım Olarak Yapay Zekâ Kavramı

Yapay zekâ çalışmalarının ve bu çalışmalar içerisinde ulaşılmak istenen bir metafizik tasarım olarak "yapay zekâ" kavramının antropolojik ve bilimsel temelleri ortaya koyulmuştur. Okuyucu açısından daha açık kılınmıştır ki, bu çalışmanın da arzularından biri olarak, "yapay zekâ" kavramının felsefece analizini güçleştiren ve onu felsefe ya da bilimler uğruna değil ama pek çoklarının – kültür endüstrisinin – metafizik kabullerini tatmin edercesine inşa eden bir tavrın analizini sunmak asıl arzumuzdur. Bu uğurda antropolojik ve fizikalist temelleri bakımından incelediğimiz "yapay zekâ" kavramını ve bu kavramı ortaya çıkaran tavrı yer yer yapay zekâ projesi olarak adlandırmış olmamız okuyucu açısından anlaşılmaz bulunmuş olabilir. Ancak burada amacımız, yapay zekâ çalışmalarının tümünü değil ama bu çalışmanın da konusu olarak, felsefi işlevselci tavrın yükleri bakımından yapay zekâ alanında yaratılmak istenen bir metafizik tasarım olarak "yapay zekâ" kavramını ve bu kavram uğruna çalışan düşünür ve bilim insanlarını nitelemektir. Böylece insan karşısına konmak istenen ve insanla denk olması için çalışılan, ancak bu uğurda insanın ve yapay zekânın varlık koşullarının onlara yeterli dikkat verilmeyerek ele alındığı, dolayısıyla felsefece bir incelemenin değil ama bilimler düzeyinde onaylanmaya çalışılan steril kimi inançların literatürde yer aldığını vurguluyoruz. "Yapay zekâ" kavramının, bir metafizik tasarım olması işte tam da bu sebepledir: Bilimleştirilen kanıların, hakikat olduğunu savunmak için temellendirmelerimiz mevcut mudur? Kanılarımız, üzerinde durduğumuz var olanların varlık koşullarıyla örtüşür mü? Yoksa biz bu örtüşmenin sağlanması uğruna yine o varlıkların varlık koşullarını *yontmaya*, onları *kanılarımıza uymaya zorlamaya* mı çalışıyoruz? İşte okuyucu açısından şimdiye kadar bu çalışmada üzerinde durulan hususlarla da açık kılınmıştır ki felsefi işlevselci girişimlerin mevcut durumu, beyin-bilgisayar koşutluğu gibi sıkıca incelenmemiş ve üzerinde durulma dikkati gösterilmemiş bir kanıya dayanan bir tavırdan daha fazlası değildir. Bu kavramı bir metafizik tasarım olarak adlandırmamız, bu kavramın felsefece bir incelemenin ve bilimler düzeyinde bu incelemenin sınanması yoluyla değil ama bilimlerden hareketle felsefeye dayatılmak istenen bir tavırdan, *düalist ruha* ilişkin arzularını fizikalist temeller

bakımından gerçekleştirmek isteyen bir anlayıştan hareketle *doğurtulmaya çalışılmasından* ötürüdür.

Bu çalışmanın amaçlarından biri bilimlerin yadsınması veya bilimlerin sahip oldukları problemlerden ötürü bir şeyler ortaya koyamayacak durumda olduklarını iddia etmek değildir. Bilimler, medeniyetimizi giderek hızlanmakta olan bir şekilde geliştirmekte, ileri götürmektedir. Bir asır önceki insanın koşullarıyla şu anda yaşamakta olan bir insanın koşullarının birbirinden oldukça farklı olması tam da bilimler sayesindedir. Ancak bilimlerin bu ilerletici ve aydınlatıcı mahiyetlerinin, bu çalışmanın da odağı olarak, yapay zekâ çalışmaları içerisinde *aşırı kullanımlarını*[*] okuyucu açısından da açık kılmak istemekte bir beis bulunamaz. İşte tam da metafelsefi incelemeler adını verdiğimiz bu bölümde matematik, mantık ve doğa bilimi üzerine kimi tespitlerde bulunmuş olmamız, bu bilimlerin değersizleştirilmesi ya da bu bilimlerin yapay zekâ çalışmaları için bir temel olamayacağının gösterilmesi uğruna değil, aksine bu bilimlerden istifade eden yapay zekâ çalışmalarının bu bilimler düzeyinde de kimi sorumlulukları üstlenmek durumunda olduğunu ortaya koymaktan daha fazlası değildir.

Yapay zekâ felsefesi içerisinde yapay zekâ projesi olarak adlandırdığımız ve insanın varlık koşullarını dikkate almayan bir felsefe yapma anlayışının bulunduğunu ve bu anlayışın bir sonucu olarak ele aldığımız "yapay zekâ" kavramının antropolojik ve fizikalist temeller bakımından, üzerinde felsefece bir tartışmanın yürütülemeyeceği kadar kimi yüklere sahip olduğunu dile getirmiştik. Bizim bu saptamalardaki amacımız, tıpkı bu çalışma içerisinde bilimleri ele alırken sahip olduğumuz tutum gibi, *benzetim* ve *örnekleme* yoluyla kavramın aydınlatılmasına ilişkin bir düşünce yöntemini elemek veya yadsımak değil, tam aksine bu benzetim ve örnekleme süreçlerinin artık felsefece bir sorgulamaya tabi olmadan bizzat ilgili kavramların yerine geçmeleri yollu bir felsefe yapma tarzına dikkat çekmek içindir.

[*] "Aşırı kullanım" ile ifade etmek istediğimiz durum, bilimlere sıkça başvurulması değil tam aksine bilimlerin varlık koşullarının esnetilmesi, aşılması ve bu yolla bilimlere, bilimlerin sınırlılıklarını aşan yargıların taşınmak istenmesidir. Bu arzu kendisini en yalın haliyle mantık bilimde gösterir. Mantık bilimi tartışmaya oldukça kapalı olan yapısıyla yargılarımızın onda inşa edilmek istendiği ve dünya tasarımımızın bu yolla mantıksal sıkılıkla güvenceye alınmak istendiği bir alandır. Yapay zekâ projesinde de sıkça mantığın, dünya tasarımımız için en uygun dil olduğu, yapay zekâların da bu dile dayanarak inşa edilebileceği savunulur. Oysa bu çalışma içerisinde mantığa buluşçu, yaratıcı mahiyetlerin yüklenip yüklenemeyeceğiyle ilgili husus tartışılmıştır.

Böylece bu çalışmada ortaya konan tavır, ne felsefi işlevselci tavrın elenmesi uğrunadır ne de felsefi işlevselci tavrın yapay zekâ felsefesi bakımından konumunu yadsımaktır. Bu çalışmadaki anlayışımız, metafelsefi incelemeler bölümü bakımından, felsefi işlevselci tavrın felsefe yapma tarzı bakımından oldukça *felsefece olmayan bir yöntemi* benimsemiş olduğu, bu yöntemin dayanaklarının antropomorfik ve fizikalist temellere dayanan ve bilimler düzeyinde de pekiştirilen bir yöntem olduğunu vurgulamaktadır. Biz tam aksine, felsefenin tüm bu kurgusal ve gündelik kaygılardan arınmış, duru ve sıkı bir disiplin olduğunu savunuyoruz. Felsefi işlevselciliğin tüm bu şimdiye kadar sayılan yüklerden arınmış bir felsefi yöntemle yeniden ele alınması ve bu tavrın bilimler düzeyinde de kimi başarılar elde etmesi oldukça mümkündür. Ancak bizim yöntem ve içerik bakımından işaret ettiğimiz hususlar üzerinde yeni ve felsefece bir yöntem tercih edilmedikçe felsefi işlevselci tavrın literatürdeki kavramsal yansılarının, felsefenin bir modern zaman miti olarak *golemlerin* ya da bir *Frankenstein'ın* türetilmesi uğruna koşulduğu bir tavrın ürünlerinden daha fazlası olmadığını söylememiz gerekir. Felsefe etkinliğinin doğası ve bir felsefecinin işi, bir kavramın ele alınışında ne o kavramı ele alırken tutku ve duygulanımların işin içine katılmasına izin vermektir ne de kavramın gündelik insanlar bakımından ilgi çekici kılınması uğruna kimi girişimleri onaylamaktır. Yapay zekâ projesi olarak adlandırdığımız tavrın tam olarak ihlal ettiği hususlar bu hususlardır. Ne yapay zekâ projesi içerisinde yapay zekâ kavramı, insanın varlık koşulları bakımından sıkıca incelenmektedir ne de "yapay zekâ" kavramının mitleştirilip popülarize edilmesinden ve bu çabanın ürünlerinin yeniden felsefenin içine ithal edilmesinden bir rahatsızlık duyulmaktadır. Çalışmanın bir sonraki bölümünde okuyucu açısından daha da açık kılınacaktır ki "yapay zekâ" kavramının literatürde yer alan ve bu çalışma içerisinde üzerinde durmaya çalışacağımız tanımları birçok sorunla yüzleşmek durumundadır.

İKİNCİ BÖLÜM
YAPAY ZEKÂ KAVRAMI ÜZERİNE

2.1. Mevcut Yapay Zekâ Tanımlarının İncelenmesi

Genel anlamıyla yapay zekâ felsefesi, yapay zekânın ne olduğunun soruşturulduğu ve bu kavramın ilintili olduğu tartışmalar üzerinde durulan bir felsefe alanıdır. Bu alana adını veren yapay zekâ disiplini, adını John McCarthy'nin çalışmalarından almaktadır. İnönü'ye göre yapay zekâ felsefesinin oluşması Alan Turing'e ve onun çalışmalarına dayandırılabilmektedir.[1] McCarthy'ye göre ise yapay zekâ felsefesi bakımından değil, ama yapay zekâ çalışmalarının başlaması bir grup insan tarafından bağımsız yürütülen çalışmalarla ortaya çıkmıştır[*], ancak Alan Turing'in çalışmalarının başlangıç olarak sayıldığı kabul edilmektedir.[2] "Yapay zekâ" kavramının oluşmasının tarihi incelenirse Dartmouth Konferansı ve bu konferansta tartışılan fikirler önem kazanır. Nitekim bu konferansın önemi, bilgisayarların yalnızca sayıları işlediği anlayışının terk edilmesi ve sembolleri işleyebilmek amacıyla veri işleyen işlevlerin artırılması ve geliştirilmesi bakımından bir yol gösterici olmasıdır.[3] Ancak Say'ın da dile getirdiği gibi yapay zekâ çalışmaları sıradan bir mühendislik projesi olarak ele alınamaz, çünkü inşa edilmek istenen şey olarak yapay zekâlar, bir bilgisayar programı olarak, biçimsel bir dilde oluşturulmuş bir metinden meydana gelir.[4] Bu anlamda çalışmanın kapsamı içerisinde "yapay zekâ" kavramının ve alanının tanımları üzerinde dikkatlice durmak

---

[1] Mahbube Nazlı İnönü, "Yapay Zekâ Felsefesi", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 136-137.

[*] McCarthy'nin de ifadesine dayanarak, yapay zekâ çalışmalarının değil ama yapay zekâ felsefesinin fikren Turing'in çalışmalarına değil, David Hilbert'in matematik felsefesi bakımından değerli olan karar verme problemine dayandırılmasını öneriyoruz. Çünkü Hilbert'in karar verme problemi bir önermenin doğruluğu veya yanlışlığı hakkında genel bir algoritmanın olanaklı olup olmadığı üzerinde şekilleniyordu. Yapay zekâ kavramının varlık koşullarından biri tam olarak bu algoritma vasfının taşınıp taşınmadığıdır. Hilbert'in karar verme probleminin ortaya atılması, yapay zekâ alanının olanaklılığını sağlamak bakımından hem bilgisayar bilimini mümkün kılmış hem de yazılım/programlama üzerinden böyle bir problemin üstesinden gelecek bir zekâ bakımından da yapay zekâ düşüncesinin olanaklılığını sağlamıştır. Dolayısıyla Turing'in başlangıç olarak kabul edilmesi Hilbert'in karar verme probleminin öneminin yadsınması olarak nitelenebilir.

[2] McCarthy, "What is Artificial Intelligence?".

[3] Kaplan, **Artificial Intelligence: What Everyone Needs to Know**, s. 13-14.

[4] Say, "Yapay Zekâ ve Gödel Teoremi", s. 254.

ve bu kavram ve alanla ilgili tanımları farklı tartışma sahalarındaki başarıları bakımından ele almak durumundayız.

Her felsefi alanın temel sorusu olarak nelik sorusu yapay zekâ felsefesi bakımından da üzerinde durulan bir sorudur. Yapay zekâ felsefesi içerisinde yapay zekâlı varlıkların ne olduğu, hangi yetilere sahip olması gerektiği, hangi koşullar sağlandığında bir varlığın yapay zekâ sahibi olduğunun tespit edileceği yollu birçok soru mevcuttur. Say'a göre yapay zekâ alanı, doğal sistemlerin gerçekleştirebildiği her türden bilişsel kabiliyeti yapay sistemlere yaptırabilmenin araştırıldığı bir alandır.[5] Dolayısıyla yapay zekâ alanı ve yapay zekâ felsefesi içerisinde de yapay sistemlere yaptırılmak istenen bilişsel yetilerin ne olması gerektiği, bunların hangilerinin yapay zekâlar bakımından kaçınılmaz bir şekilde içerilmesi gerektiği türünden birçok soruya yanıt verebilmek amacıyla farklı tanımlama çabaları bulunur. Ancak bu tanımlama çabalarının gerçekleştirilebilmesi için yapay zekâ alanının belirli temellere dayandırılması gerekir. McCarthy açısından bu temel, yapay zekâ alanının varlık koşullarından biri olarak, insan zekâsının anlaşılmasıdır. Çünkü insan zekâsı anlaşılamazsa, bilgisayar programlarına zekâ kazandırılamaz. Dolayısıyla yapay zekâ alanı içerisinde üretilecek çözümler, çözümü bakımından zekâ gerektiren problemlerin analiz edilmesi üzerinden elde edilmeye çalışılır.[6] Nitekim 1960'larda yapılan yapay zekâ çalışmalarının öncelikli hedeflerinden biri yapay zekânın insan zekâsıyla olan ilişkisidir. Bu anlamda bir mühendislik hedefi olarak insan zekâsının işlevsel kabiliyetlerinin yeniden üretilebilmesi için sezgiye dayanan programların yazılması amaçlanmıştır.[7] Çünkü Darthmouth Konferansı düşünülürse, makinelerin yalnızca sayıları işleyen cihazlar olmadıkları kabul edildiğinde insan zekâsının sezgisel unsurlarını da içeren bir yapay zekânın üretilmesi gerektiği oldukça açıktır. Bu anlamda yapay zekâ alanının somutlaşması, düşünen makinelerin gerçekleştirilmesi için yürütülen çalışmalardır.[8]

Makinelere düşünme yeteneğinin kazandırılması amacının başarılabilmesi için yapay zekâlı varlıkların düşünebilme koşullarının ortaya konması gereklidir.

[5] Say, **50 Soruda Yapay Zekâ**, s. 83.
[6] McCarthy, "Review of the Emperor's New Mind by Roger Penrose".
[7] Terrence J. Sejnowski, "The unreasonable effectiveness of deep learning in artificial intelligence", **Proceedings of the National Academy of Sciences**, C. CXVII, No: 48, 2020, s. 30036.
[8] Demircioğlu, "Yapay Zekâ: Açmaz ve İmkân", s. 12.

Zeki makineler ya da daha özelinde zeki bilgisayar programlarını gerçekleştirebilmenin bilimi ve mühendisliği olarak yapay zekâ alanı, bilgisayarların ve teknik imkânların kullanılması yoluyla insan zekâsını incelemektedir. Nitekim bu inceleme, yapay zekânın insan zekâsını simüle etmek uğruna geliştirilip geliştirilmediği ile ilgili bir soruyu doğurur. McCarthy'ye göre insan zekâsını inceleme çalışmaları kimi zaman insan zekâsının belli yetilerinin bir yapay zekâya kazandırılmak istenmesi, bazen de tüm yetilerin bir yapay zekâya kazandırılmak istenmesi şeklinde farklı çabaları içerir. Bu anlamda kimi zaman bir problemi nasıl çözdüğümüzle ilgili akıl yürütürken kimi zaman ise insanların ve hayvanların dünyadaki bir problemi çözme süreçlerinde zekânın rolünü ortaya koymaya çalışırız. Ancak yapay zekâ alanının, insan seviyesi bir zekâyı hedefleyen kimi arzulara sahip olduğunu söylememiz de mümkündür.[9] Peki yapay zekâ alanının bu ve benzeri amaçlarını gerçekleştirmek için dayandığı temel varsayımlar nelerdir? Boden'a göre yapay zekâ araştırmacılarının iki temel varsayımı bulunur: Zekâ gerektiren süreçler algoritmalarla ifade edilebilir ve tüm bu algoritmalar bir bilgisayara entegre edilebilir.[10] Öte yandan Şengör'e göre ise yapay zekâ alanında çalışan insanlar, insan zihninin akıl yürütme, planlama, karar verme, öğrenme vb. gibi bilişsel süreçlerine odaklanmaktadır ve bu çalışmalardan elde ettikleri bilgilere dayanarak zeki makineleri gerçekleştirmeyi hedeflemektedir.[11] Fakat Şengör, yapay zekâ alanında ele alınan problemlerin ve bu problemlere önerilen çözümlerin, zekânın modellenmesi sürecinde ne kadar başarılı olduğunun, önerilen yöntemlerin zekâyı tam anlamıyla temsil edip edemediğinin ve zeki bir sistemin oluşturulup oluşturulamadığının tartışmalı başlıklar olduğunu dile getirir.[12]

Yapay zekâ alanının araştırma odaklarının, başarı unsurlarının ve hedeflerinin ne olması gerektiği ya da "yapay zekâ" kavramının ne olduğu ve insansı hangi yetilere sahip olması gerektiği, kısacası onun varlık koşulları ile ilgili bir uzlaşı bulunmamaktadır. "Yapay zekâ" kavramının koşulları hem yapay zekâ alanının farklı tanımlamaları hem de "yapay zekâ" kavramının farklı bilimler ve alanlar

[9] McCarthy, "What is Artificial Intelligence?".
[10] Boden, "Artificial Intelligence".
[11] Şengör, "Bilinç, Farkında Olan…", s. 96.
[12] **A.e.**, s. 101.

bakımından ele alınması karşısında değişiklikler gösterir. Bu anlamda literatür üzerinde birçok tanımın bulunduğu görülür. Yapay zekâ alanı bakımından temel birkaç tanımın bulunduğundan söz edilebilir. Örneğin yapay zekâ alanı, insanların yapabildiği türden işleri yapabilen makinelerin bilimi olarak eylem ve faillik üzerinden tanımlanabilirken[13] öte yandan Audi'ye göre insan yapması zeki varlıklar tasarlamaya yönelik bilimsel bir çaba olarak da nitelendirilebilir. Burada Audi, "zekâ" kavramının önemini vurgular. Yapay zekâ çalışmalarında zekânın ele alınışı, sistemin hangi materyalden yapıldığının öneminin olmadığı işlevselci bir tarzla, parçalar ve parçalar arası ilişkilerin zeki davranışı meydana getirmesi üzerinden gerçekleşir. Burada zekânın kendini gösterdiği eylemler olarak sıklıkla mantık problemlerin çözülmesi, satranç oynamak vb. gibi işler karşımıza çıkmaktadır.[14] Klasik yapay zekâ ya da Boden'in ifadesiyle geleneksel yapay zekâ, koşullu ifadelerin bulunduğu sıralı bir programlama anlayışına dayanmaktadır. Burada listeler, semantik ağlar ve diğer bilgi işleme yapıları üzerine eğilen bir anlayışla karşılaşılır. Nitekim Boden, temsil yapıları ve bileşenleri olarak kavramların, kanılarımız, önermelerimiz vb. için bir temel oluşturduğu sembolik temsillerin, sembolik yapay zekâ yaklaşımını oluşturduğunu öne sürer.[15] Boden'in bu yaklaşımı bakımından ele alınırsa yapay zekâ ile ilgili tanımlamaların insan zekâsının ve bu zekânın yetilerinin gerçekleştirilmesi konusunda yoğunlaştığını söylemek mümkündür. Bu anlamda yüz veya fotoğraf tanıma, dili kullanma, öğrenme vb. gibi kabiliyetler ele alınabildiği gibi[16] algı, akıl yürütme, öğrenme, iletişim kurma ve karmaşık bir çevrede eyleyebilme gibi birçok olanaklılığın[17] da makinelere kazandırılmak istendiği açıktır. Dolayısıyla Penrose'un yapay zekânın amaçlarından birinin insanın ussal etkinliğinin olabildiğince taklit edilmesi olduğu yollu yaklaşımı bu temellere dayandırılabilir.[18]

---

[13] Simon Blackburn, "artificial intelligence", **The Oxford Dictionary of Philosophy**, 2. bs., Oxford, Oxford University Press, 2005, s. 25.
[14] Robert Audi, "artificial intelligence", **The Cambridge Dictionary of Philosophy**, 2. bs., New York, Cambridge University Press, 1999, s. 53-56.
[15] Boden, "Artificial Intelligence".
[16] **A.e.**
[17] Nils J. Nilsson, **Artificial Intelligence: A New Synthesis**, San Francisco, Morgan Kaufmann Publishers, 2009, s. 1.
[18] Penrose, **Kralın Yeni Usu**, s. 11.

Yapay zekâ çalışmalarının insanın ussallığı ve insan zihninin yetileri üzerinde yoğunlaşmış olması, bu çabanın başarılabilmesi amacıyla türlü çözüm önerilerinin ortaya çıkmasını sağlamıştır. Bu çalışmanın ilk bölümünde de üzerinde durduğumuz hususlardan biri olarak mantık bilimine dayanan yapay zekâ önerileri mevcuttur. Allen Newell ve Herbert Simon'un sembol temelli yapay zekâ düşüncesi, yapay zekâ alanı ve kavramı bakımından yukarıda dile getirdiğimiz bazı hedeflerin gerçekleştirilebileceği bir zemin olarak önerilir. Nitekim Newell ve Simon'a göre insan davranışını yakın bir şekilde simüle edebilen bir bilgisayar programı bir bilgisayarda yazılabilirse o halde insanın herhangi bir davranışını ortaya çıkaran yeterli mekanizmalar keşfedilebilecektir.[19] Simon, bu keşfin sağlanabilmesi için yapay zekâ çalışmalarının ilk görevinin zeki davranışın sembolik süreçlerle elde edilip edilemeyeceğinin tayin edilmesi gerektiğini düşünmektedir.[20] İnsan zihni ve onun yetileri sembolik yapay zekâ anlayışı bakımından bir kere mantıksallaştırıldığında, bu düşüncenin bilgisayarlar ve robotlar bakımından da bağı kurulmuş olur. Böylece Akın'ın yapay zekâ araştırmacılarının baştan beri ulaşmak istediği noktanın insan gibi davranan sistemler üretmek olduğu yaklaşımından hareketle[21], Simon'un yapay zekânın, insanlar tarafından yerine getirildiğinde zekâ ve düşünce gerektiğini düşündüğümüz işlerin bilgisayarlar tarafından yerine getirilmesi işiyle ilgilendiği yollu düşüncesi oldukça açık kılınmış olur.[22] Ancak "yapay zekâ" olarak adlandırdığımız bu fenomeni tanımlama, ona bir nelik kazandırma konusunda bazı sorunlarla karşılaştığımızı söyleyebiliriz. Örneğin Köse'ye göre biyolojik olanla ifade edilen doğal zekânın sayısallaştırılarak bir bilgisayara aktarılması, bu yolla da problem çözme yönünde olan doğal zekâların bilgisayarda gerçekleştirilmesinin sonucunda makinenin sahip olacağı zekâ, yapay zekâ olarak tanımlanabilir.[23] Nitekim bu tanım, Bell'in yapay zekâ çalışmalarının nihai amacının dünyada beyin yerine bir bilgisayar kullanarak yaşayacak robotlar yaratmak olduğu görüşüyle de örtüşür.[24] Ancak çalışmamızda matematik ve mantık bilimleri bakımından üzerinde durduğumuz kimi sorunlarla da açıktır ki insan

[19] Newell ve Simon, "The Simulation of Human Thought".
[20] Simon, "Artificial Intelligence: an empirical science", s. 104.
[21] Akın, "Yapay Zekâda Vücut Beyin Problemi".
[22] Simon, "Artificial Intelligence: an empirical science", s. 95.
[23] Köse, **Yapay Zekâ Felsefesi**, s. 17-18.
[24] Bell, "Levels and loops: the future of artificial intelligence and neuroscience", s. 2014.

zekâsının veya herhangi karbon temelli bir zekânın sayısallaştırılabilmesi, sembolleştirilebilmesi ve biçimsel bir sistem içerisinde kayıpsızca temsil edilebilmesinin koşulları tartışmalıdır. Ayrıca, tamamen matematiksel ve sembolik temellere dayanan bir biçimsel dizge olarak bir yapay zekânın, matematiksel bir ontolojiyi aşıp aşamayacağı, bizim gerçekliğimizde eyliyor ve burada bulunuyor gibi gözükse bile matematiksel ontolojinin sınırlılığı bakımından şeylere anlam yükleyemeyeceği, bu bakımdan da semantiğin makineler için söz konusu olmadığı bir durumda makine zekâsının asıl anlamda bir zekâ olarak tanımlanabilmesinin olanaksız olduğu bu çalışma içerisinde sıkıca tartışılmıştır.

Öte yandan yapay zekâlı varlığın ne olduğu ve bir makinede bulunan zekânın yapay zekâ olup olmadığının tespit edilebilmesi için kimi tanımlama girişimlerinin de olduğunu söylemek mümkündür. Bir sosyal bilim olarak hukukun bu ve benzeri tanımlara ihtiyacı, günümüz bakımından yadsınamaz. Dolayısıyla hukukçular açısından önem arz eden bu ve benzeri tanımlama girişimlerinden biri olarak Taşdemir, Özbay ve Kireçtepe açısından bir zekâ, bir başka varlığın yönlendirmesine ihtiyaç duymuyorsa, yani otonomsa, karbon bazlı yani organik değilse ve üretilmesini insan veya bir başka yapay zekâ sağlıyorsa yapay zekâ olarak adlandırılabilir.[25] Öncelikle bu tanımın birkaç sorunu bulunur: İlk olarak tanım döngüseldir. Yapay zekânın bir başka zekâ yarattığı durumda o zekânın yapay zekâ olacağını söylemek, ilk yapay zekânın nasıl yapay zekâ tayin edilebileceğini söylemeye gelmez. Öte yandan bir insan tarafından üretilen zekânın, asıl anlamda bir yapay zekâ olarak tanımlanabilmesinin koşulları sunulmadığı için de insan tarafından üretilen her zekânın yapay zekâ olacağını söylememiz için bir zorunluluk bulunmaz. Tanımla ilgili bir diğer sorun ise bir yapay zekânın yönettiği birden çok robotun söz konusu olabileceği bir durumda tüm bu robotların aynı anda yalnızca bir yapay zekâyı mı yoksa birden çok yapay zekâyı mı temsil edip etmediğine ilişkin bir açıklama verilmedikçe tanımda vurgulanan otonomluk ifadesi anlamını yitirmektedir. Çünkü yapay zekâ çalışmalarında örnek alınan insan zihni de tüm yetileri ve olanaklılıkları bakımından otonom bir şekilde çalışmamaktadır. Otomatik ve kontrolümüz dışında gerçekleşen duyumsama, algılama süreçlerimiz gibi bilinç

[25] Özgür Taşdemir v.d., "Robotların Hukuki ve Cezai Sorumluluğu Üzerine Bir Deneme", **Ankara Üniversitesi Hukuk Fakültesi Dergisi**, C. LXIX, No: 2, 2020, s. 797-798.

dışında kalan diğer süreçlerimiz de zekâmızın onda otonomca eyleyemediği fenomenlerdir. Dolayısıyla bir zekânın "yapay zekâ" olarak tanımlanabilmesinin bir koşulu olarak mutlak anlamda otonomluğun beklenmesi* tartışmalıdır.

İnsan zihninin birçok yetiyi içermesi ve insan zekâsının dünyada birçok problemi çözmek bakımından gösterdiği başarı, yapay zekâ çalışmalarının farklılaşmasını sağlamıştır. Çünkü yapay zekâ çalışmaları, bu çalışma içerisinde de sıkça dile getirildiği gibi, insan zihninden ve özellikle insan zekâsından ilham almaktadır. Arslantunalı'ya göre "yapay zekâ" kavramını kullanışımızda üç tür anlamı karşılıyor olabiliriz: Otomasyon ve sibernetik karşılık olarak elektronik aygıtların ve makinelerin otonom olarak çalışmalarını sağlayan bir olanak olarak yapay zekâ, insan zihni ve zekâsına ilişkin yetilerin taklit ve simülasyonunu sağlayan programlar olarak özel yapay zekâ, insanlar gibi düşünebilen, hissedebilen ve insan benzeri bilinçli bir var olan olarak genel yapay zekâ düşüncesi.[26] İlk anlamıyla yapay zekâlar, artık gündelik yaşamın her noktasında kendine yer bulabilmektedir. Cihazlar akıllı hale gelmektedir böylece kullanım alışkanlıklarımızı etüt eden ve bu sayede tercihlerimize uygun seçenekleri bize sunan yapay zekâlar olduğu gibi kullandığımız görsel materyalleri zenginleştiren yapay zekâlar da söz konusudur. Öte yandan Arslantunalı'nın özel yapay zekâ olarak adlandırdığı ancak literatürdeki genel kullanım olarak zayıf** yapay zekâlar, bir insanın zihinsel kabiliyetlerinden bir veya birkaçını taklit etmek üzerine geliştirilmiş programlar olarak bulunmaktadır. Günümüzde zayıf yapay zekâlar rastgele verilen kelimelerden bir kompozisyon oluşturmakta ve bir şarkı veya resim üretebilmektedir.*** Öte yandan işi insanın araştırma ve sohbet kabiliyetlerini taklit etmek ve kullanıcı ile sohbet edip

* Örneğin bir iş yerinde bir yapay zekânın eş zamanlı olarak yönettiği 100 adet robotun var olduğunu düşünelim. Bu durumda yapay zekânın bu robotların tümü üzerinde eş zamanda farkında olarak yani tümünün fenomenal deneyimlerine t_1 gibi bir zamanda sahip olabileceğini – bu aynı anda 100 kişinin size bakarken sizi nasıl gördüğünü eş zamanda deneyimleyebilmekle aynı durumdur – ve olması gerektiğini iddia ederek otonomluğu – otonom olabilmenin bir koşulu olarak deneyimlerimizin tümüne tüm zamanlarda hâkim olmamız gerekir – sağlayabileceğini ileri sürmemiz gerekir. Ancak tıpkı insan zihni gibi makine zekâsı da oldukça otomatikleşmiş süreçlere sahip olmalıdır. Dolayısıyla otonom olmayan ama yapay zekâ olarak kabul edilebilecek bir varlık da tasavvur edilebilir.

[26] Arslantunalı, "Yapay Zekâ ve İnsanbiçimcilik: 'Bacakları Olan Bir Araba'", s. 227-228.

** İng. **Weak Artificial Intelligence**

*** Üretilen bu şarkıların veya resimlerin sanat eseri olmak bakımından mahiyetleri, sanat felsefesinin asli tartışma konularından biri olmuştur. Bu anlamda yapay zekâ devriminin ve genel anlamda bilimsel ve toplumsal yeniliklerin felsefenin alt alanları ve tartışma konuları bakımından dönüştürücülüklerini canlı olarak deneyimlemekteyiz.

kullanıcının arzularına göre internet üzerinde arama yapmak olan yapay zekâ botları da söz konusudur.[*] Ancak biz bu çalışmanın amaçları bakımından Arslantunalı'nın genel yapay zekâ olarak belirttiği ancak John Searle tarafından güçlü yapay zekâ olarak tasnif edilmiş bir tür yapay zekâ anlayışı üzerinde durma niyetindeyiz.

Güçlü yapay zekâ kuramına göre yapay zekâ çalışmaları insan zihninin doğası hakkındaki araştırmalarda yalnızca bir araç olamaz. Eğer bir yapay zekâ uygun bir şekilde programlanabilirse o halde bu programın yüklendiği ve çalıştırıldığı bir bilgisayarda bir zihnin belirmesi mümkündür. Bu zihin dilsel ifadeleri anlayabilir ve başka zihinsel yetilere de sahip olabilir. Searle'e göre bu kuramın iki iddiası bulunur: Programlanmış bir bilgisayar dilsel ifadeleri anlayabilir ve bilgisayar programları insanın anlama yetisi üzerinde bir açıklama gücü bulunur. Dolayısıyla güçlü yapay zekâlar yalnızca bir insanın psikolojik durumunun test edildiği programlar değil aynı zamanda o zihin durumlarının kendisine de sahip durumdadır. Oysa zayıf yapay zekâlar bize yalnızca formüllerimizi ya da hipotezlerimizi sınamak konusunda olanak sağlayan daha dar anlamdaki yapay zekâlar olarak karşımıza çıkar.[27] Doğan'ın da belirttiği gibi zayıf yapay zekâ çalışmaları zihnin hesaplama yetisini amaçlarken, güçlü yapay zekâ çalışmaları ise insan zihninin tüm yetilerini bir makinede gerçekleştirmeyi amaçlamaktadır çünkü güçlü yapay zekâ alanında çalışan insanlar tasarlamayı arzu ettikleri güçlü yapay zekâların insan zihnine ilişkin fenomenleri simüle etmekten çok onlara asıl anlamda sahip de olabileceğini savunmaktadır.[28]

Yapay zekâ çalışmalarında bir dönem yaygın olarak kullanılan makine öğrenmesi, güçlü yapay zekâya yol açan genel yapay zekâ anlayışı gibi zihnin tüm yetilerini modellemeyi değil ama daha mütevazı olarak algılama, dil, motor kontrol, çıkarım yapma vb. gibi pratik kimi problemler üzerine eğilmektedir.[29] Ancak McCarthy'ye göre Allen Newell ve Herbert Simon'un Genel Problem Çözücü[**] programı genel zekânın başarılması ve belki de güçlü zekâya giden yolun açılması

[*] ChatGPT bu anlamda en yalın ve en güncel örnektir.
[27] Searle, "Minds, brains, and programs", s. 417.
[28] Mehtap Doğan, "Güçlü Yapay Zekâ Mümkün müdür?", **Pasajlar**, No: 9, 2021, s. 87.
[29] Sejnowski, "The unreasonable effectiveness of deep learning in artificial intelligence", s. 30033.
[**] İng. **General Problem Solver**

uğruna çok optimist bir girişimdir.[30] Görülebilir ki genel yapay zekâ üzerinden gerçekleştirilmeye çalışılan güçlü yapay zekâ düşüncesinin asıl yöntemi, karmaşık problemleri çözebilen, kendisini otonom olarak kontrol edebilen, kendi düşünceleri, kaygıları, hisleri, güçleri ve zayıflıkları olan bir yazılım inşa etmektir.[31] Güçlü yapay zekâ düşüncesi, bizim bu çalışmada yapay zekâ projesi olarak adlandırdığımız bir anlayışa hizmet eder ve daha önceleri de dile getirdiğimiz yapay zekâ projesi ifademiz, yapay zekâ çalışmaları içerisinde ağırlıklı olarak benimsenen zayıf yapay zekâ tanımını[32] ve bu tanımı benimseyen araştırmacı ve düşünürleri kapsamamaktadır.

Güçlü yapay zekâ fikrinin ortaya atılmasıyla bu tarz bir zekânın gerçekleşmesinin koşulları üzerinde birtakım tartışmalar ortaya çıkmıştır. Bir grup düşünür makinelerin insanlar gibi eyleyebilmelerinin olanaklı olmadığını, insanlar gibi duygu sahibi varlıklar olamayacaklarını, iyi ve kötü arasında bir ayrım yapamayacaklarını savunurken, bir grup düşünür ise insan beyninin de birçok işlevi yerine getiren çeşitli parçalardan meydana gelen yapısından ötürü makinelerde güçlü yapay zekânın gerçekleştirilmesinin olanaklı olduğunu savunmaktadır.[33] Güçlü yapay zekâ savunucularına göre makinenin davranışları yalnızca bir simüle etme süreci değildir tam aksine makinenin söylenilen şeyleri anladığını, kimi hislere sahip olduğunu ortaya koyan önemli bir göstergedir.[34] Çünkü bu anlayışa göre düşünme süreçleri semboliktir ve bu süreçler sembollerin biçimsel olarak manipülasyonu olarak kabul edilebilir. Bu anlamda bir bilgisayarın yaptığı iş olarak sembollerin manipülasyonu, bir yapay zekânın güçlü yapay zekâ olarak kabul edilebilmesi ve insana denk bir şekilde kimi fenomenal deneyimlere sahip olduğu anlamına gelebilir.[35] Ancak İnönü'nün de ifade ettiği gibi Searle, bilgisayar programlarının biçimsel yapısı içerisinde semantiğin ortaya konabileceği düşüncesine

[30] McCarthy, "From here to human-level AI", s. 1174.
[31] Cassio Pennachin ve Ben Goertzel, "Contemporary Approaches to Artificial General Intelligence", **Artificial General Intelligence**, ed. Ben Goertzel ve Cassio Pennachin, C. II, New York, Springer, 2007, s. 1.
[32] Antonio Chella ve Riccardo Manzotti, "Artificial intelligence and consciousness", **Association for the Advancement of Artificial Intelligence Fall Symposium**, 2007, s. 1.
[33] Kadir Öztürk ve Mustafa Ergin Şahin, "Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış", **Takvim-i Vekayi**, C. VI, No: 2, 2018, s. 25-26.
[34] Searle, "Minds, brains, and programs", s. 417.
[35] Searle, "Is the Brain's Mind a Computer Program?", s. 26.

katılmamaktadır.[36] Nitekim Holland'da erken dönem yapay zekâ çalışmalarına katılan insanların varsaydığı önemli bir hatanın, yüksek seviye sembol manipülasyonunun – bilgiyi bir dil içerisinde yargılar olarak temsil etmek ve bu temsil üzerinde işlem yapmak – tüm eylemler için temel sağlayabileceğinin düşünülmesi olduğunu belirtmektedir.[37] Bu anlamda sembol işleme üzerinde yükselen güçlü yapay zekâ düşüncesinin, bu çalışma içerisinde de dile getirdiğimiz, mantıksal ve matematiksel kimi problemleri yanıtlamak zorunda olduğunu söyleyebiliriz. Dolayısıyla mantıksal ve matematiksel problemlerin çözülmesinin olanağı olarak yapay zekâ çalışmaları içerisinde öne sürülen kimi teknik çözüm önerileri bir sonraki bölümde ele alacağımız hususlar olacaktır.

2.2. Yapay Zekâ Çalışmalarındaki Teknik Unsurlar

Yapay zekâ çalışmaları dünya üzerinde eyleyebilen, bu yolla da tıpkı insanlar gibi problemler çözebilen varlıklar olarak yapay zekâların olanaklarının araştırıldığı bir alandır. Şimdiye kadar görüldüğü üzere matematik ve mantık gibi biçimsel bilimler temelinde inşa edilmek istenen yapay zekâlar için insanın sinirsel yapılarının örnek alındığı görülür. Bu anlamda insan beyni, bu beyni oluşturan en küçük birimler olarak nöronlar ve nöronlar arası organizasyonlar bakımından incelenmekte ve yapay zekâların başarısı, bu nöronal yapıyı en uygun şekilde modelleyebilecek sistemler üzerinde pekiştirilmek istenmektedir. Yapay zekâ çalışmalarının bir kolunun insanın zihinsel süreçlerini anlamak olduğu üzerinde durmuştuk. Nitekim bu anlamda bilişsel bilimciler ve sinirbilimciler, insan beynini anlamak üzere çeşitli sinirsel modeller inşa etmektedir.[38] Bu sinirsel modellerin başarısı ise modellenen sistem olarak insan beyninin yapısının detaylıca anlaşılmasında yatar. Yapay zekâ çalışmaları tarihine bakıldığında ise insan beyninin fizyolojik yapısının çeşitli açıklama modelleri üzerinden anlaşılmaya çalışıldığı görülür. Örneğin McCulloch ve Pitts, yapay sinir hücrelerinin kullanıldığı bir hesaplama modelini öne sürmüştür. Hesaplanabilir fonksiyonların sinir hücrelerinden oluşan ve mantıksal – "ve" ve "veya" – kapılara

[36] İnönü, "Yapay Zekâ Felsefesi", s. 142-143.
[37] Holland, "The future of embodied artificial intelligence: Machine consciousness?", s. 38.
[38] Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, s. 29.

sahip ağlarla gerçekleştirilebileceğini savunurlar. Dolayısıyla bu ağların uygun şekilde tasarlandıklarında öğrenme becerisini de elde edebileceğini ileri sürmüşlerdir.[39] McCulloch ve Pitts'in yapay zekâ üzerine çalışmaları, yapay zekâ alanında iki temel yaklaşımın ortaya çıkmasını sağlamıştır: Mantık temelli yapay zekâ anlayışı, bağlantıcı[*] yapay zekâ anlayışı.[40] Newell ve Simon'a göre de uygun şekilde programlanmış bir bilgisayarın karmaşık işlem örüntülerini, tıpkı insanların sahip oldukları zihin süreçlerine paralel bir şekilde gerçekleştirebileceğine ilişkin bir inanç söz konusudur.[41] Ancak yapay zekâ çalışmaları içerisinde bu başarının nasıl elde edileceğine ilişkin bir yöntem bulunmalıdır.

Bir bilgisayarın yeterince iyi bir şekilde programlanabilmesinin koşulları düşünüldüğünde, insan beyninden ilham alan yapay zekâ araştırmacıları, insan beyninin fizyolojik yapısına odaklanmışlardır. Bu fizyolojik yapı, yoğun bir şekilde nöronlardan oluşmaktadır ve nöronlar arası bağlantılar bir mühendislik projesinin ürünü olarak oldukça düzenli duran elektrik devreleri olarak inşa edilmemiştir. Bu fizyolojik yapı karmaşıktır ve zihnin işlevleri ile nöronlar arası ilişkilerin eşleştirilmesi mümkün değildir. Bu anlamda hangi nöronun hangi fenomenal duruma yol açtığı tespit edilemese bile insan beyninin fizyolojik yapısının örnek alınması, insanın fenomenal unsurlarının da bu fizyolojik yapı neticesinde başka var olanlar üzerinde belirebileceğine ilişkin bir beklentiyi doğurabilmektedir. Nitekim McCulloch ve Pitts, sinirsel aktiviteye ve sinirsel materyale yoğunlaşmış, böylece bir nöronlar ağı olarak sinir sistemini düşüncelerinin merkezi haline getirmişlerdir.[42] Fiziksel sembol sistemi hipotezi ise tam da böyle bir düşünceyle örtüşür tarzda bilgisayarlarda uygun sembol işleme programlarının gerçekleştirilmesi neticesinde programların tıpkı insanlar gibi zeki davranışı gerçekleştirebilecek olanaklara sahip olabileceğini iddia etmektedir.[43] Böylece Minsky'nin de belirttiği gibi bir makinenin öğrenebilmesinin koşulu olarak makinenin, öğrendiği şeyi temsil edebilmesi gereklidir. Bilginin o halde bir çeşit veri yapısı olarak ya da temsil edilerek

[39] Akın, "Yapay Zekâda Vücut ve Beyin Problemi", s. 148.
[*] İng. **Connectionist**
[40] **A.e.**, s. 149.
[41] Newell ve Simon, "The Simulation of Human Thought".
[42] McCulloch ve Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity", s. 99.
[43] Nilsson, "The Physical Symbol System Hypothesis: Status and Prospects", s. 9.

somutlaştırılması gerekmektedir. Yapay zekâ araştırmacıları bilginin somutlaştırılması için kural temelli sistemler, semantik ağlar, nesne temelli veri yapıları gibi birçok yöntem denemişlerdir.[44] Bu anlamda yapay zekâ araştırmacılarının iki tür yaklaşım etrafında toplandığı söylenebilir: Klasik yapay zekâ yaklaşımı, bağlantıcı yapay zekâ yaklaşımı. Klasik yapay zekâ yaklaşımına göre bir yapay zekâ seri şekilde sıralı olarak çalışan ve karmaşık kuralları işleyen bir yapıdadır. Bağlantıcı sistemler ise paralel ve eş zamanlı olarak çalışan, sayıca çok olan daha basit birimlerden meydana gelmektedir. Klasik yapay zekâ yaklaşımı semboliktir ve biçimsel kurallar içerir. Bu biçimsel kurallara göre sembolleri sıralı bir şekilde işler. Ancak bağlantıcı sistemlerde ise süreçler paralel mimari üzerinde eş zamanlı olarak gerçekleştirilir. Newell ve Simon'un Genel Problem Çözücü programı, klasik yapay zekânın bağlantıcı yapay zekâ anlayışından önce gerçekleştirdiği gelişmelerden biridir. Bu program bir problemi hedefler ve problemi alt hedefler hiyerarşisi üzerinden analiz eder. Böylece arzulanan durumla mevcut durum arasındaki farkı düşürecek en uygun yöntem keşfedilmeye çalışılır.[45] Genel Problem Çözücü, insanların zihinsel süreçlerinin çalışma prensipleri üzerine doğrudan eğilen ve bu prensipleri makinelerde gerçekleştirmek isteyen bir programdır. Ancak yapay zekâ alanında her problemi çözebilecek programların gerçekleştirilmek istenmesinden, problem türlerine göre daha etkili çalışabilecek uzman sistemler anlayışına geçilen bir süreç yaşanmıştır. Uzman sistemler, problemlerin çözüleceği alanlardaki uzman kişiler tarafından sahip olunan bilgilerin derlenmesi üzerinden geliştirilen yapay zekâlardır. Ancak bu sistemler başarısız olmuşlardır. Çünkü bu programlar yalnızca sentaktik bir şekilde çalışmaktadır.[46] Öte yandan uzman sistemler, verilerden öğrenmeyi de gerçekleştirememektedir. Çok geniş bir bilgi tabanının el ile doldurulması, verilerin bu bilgi tabanına işlenmesi gerekmektedir. Ayrıca bu sistemler dünya tasvirleri bakımından çift değerli mantığa bağımlıdır ve dünyadaki olguları 0 ve 1 arasındaki bir değeri de dikkate alarak

[44] Minsky, "Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy", s. 37.
[45] Boden, "Artificial Intelligence".
[46] Akın, "Yapay Zekâda Vücut Beyin Problemi".

hesaplayamamaktadır.[47] Bu anlamda simgelere dayanan bu anlayışın bilgisayar gibi, mantık temeline oturtulmuş bir düşünmeyi temel aldığını söyleyebiliriz.[48]

Yapay zekâ çalışmalarındaki bir diğer yaklaşım ise bağlantıcılıktır. Bağlantıcı anlayış, yapay sinir ağları düşüncesine dayanır. Bu düşünceye göre nöronlardan oluşan insan beyni ve bu nöronlar arası ilişkiler örnek alınırsa, yapay olarak kurulacak sinir ağları insan beyninde gerçekleşen süreçleri taklit edebilir ve makinelerimiz insanlar gibi düşünebilir.[49] Nitekim Sejnowski'ye göre de derin öğrenme ağları, dijital bilgisayarlar ile gerçek dünya arasında köprü oluşturabilir ve böylece bu durum bilgisayarlar ile kendi kavramlarımız aracılığı ile iletişim kurmamızı sağlayan[*] bir gelişme olarak kabul edilebilir.[50] Bu anlamda, McCulloch ve Pitts'in düşüncesi de tam olarak bu kavramsal anlayışa uygundur: Belli bir şekle ait olan her girdi için benzer çıktıları üretebilen ve şekilleri tanıyabilen sinirsel ağlar inşa etmenin genel metotları aranmalıdır.[51] Nitekim bir sinir ağı, birbirine bağlı olan çok sayıdaki basit işlem biriminden meydana gelir. Bu birimler birbirine bağlı diğer birimler tarafından gerçekleştirilen işlemler tarafından etkilenerek çıktı mahiyetinde değerler üretir. Birimlerin birbirlerini etkilemeleri, bu bağlantılar arasındaki ağırlıklarla belirlenir. Dolayısıyla bir birimin ne zaman aktif olup olmayacağı, diğer birimlerden gelen etkilenimlerin toplam değerine bağlı olarak bir değerin aşılması veya aşılmaması üzerinden belirlenir.[52] Birbiriyle bağlantılı nöronlar olarak da ifade edilebilecek yapay sinir ağları, girdiler tarafından etkilenir böylece tüm bu basit birimler tarafından gelen girdilerin bir ortalaması alınır, bu ortalama üzerinde türlü işlemler yapılır ve aktivasyon fonksiyonu uygulanır.[53] Görüleceği üzere yapay sinir ağları tıpkı beynimizi oluşturan nöronlar gibi bağlantılara sahiptir ancak nöronal bağlantılardan ziyade yapay sinir ağları, sayısallaştırılmış bağlantılara sahiptir. Bu

[47] Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, s. 51.
[48] Say, **50 Soruda Yapay Zekâ**, s. 99.
[49] Çağıl, "Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekânın Olasılığı", s. 34.
[*] Makineler ve insanlar görünürde aynı kavramı kullanıyor olabilirler ancak semantik açıdan bu kavramlar aynı semantik içeriğe sahiptir denebilir mi?
[50] Sejnowski, "The unreasonable effectiveness of deep learning in Artificial intelligence", s. 30038.
[51] Pitts ve McCulloch, "How We Know Universals the Perception of Auditory and Visual Forms", s. 128.
[52] Ethem Alpaydın, "Yapay Zekâ/Doğal Bilgisayarlar", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 99.
[53] Tegmark, **Yaşam 3.0**, s. 100-101.

anlamda her birimin aktivasyon düzeyi bulunur, bu düzey girdilerin gücüne göre belirlenir. Çıktı ise girdilere ve birimlerin ağırlıklarına göre seçilir.[54]

Yapay sinir ağları, öğrenme algoritmasından ve aktivasyon fonksiyonundan oluşur. Aktivasyon fonksiyonu, sistemi oluşturan her birim tarafından, bu birimler üzerinde etkiyen girdi kuvvetlerinin hesaplanması ve belirli bir çıktının oluşturulmasında rol oynar.[55] Bağlantıcılık düşüncesi tam da bu düşünce üzerinde belirir. Bağlantıcılık, nöral ağ modelleme veya paralel dağıtımlı işleme gibi kimi isimlerle de anılan bir yaklaşım olarak, bilişsel sistemleri modellemede sinir sisteminin temel yapısının örnek alındığı bir yaklaşımdır. Bağlantıcı bir sistem aktivasyon değerleri olan birimlerden meydana gelir ve aktivasyon değerine göre çıktı üreten bu birimler bağlı bulundukları diğer birimleri etkilemektedir.[56] Bağlantıcı sistemler bakımından temsil, yerel bir temsilden ziyade tüm ağ tarafından gerçekleştirilen bir temsil türüdür. Dolayısıyla bağlantıcı bir sistemi oluşturan farklı birimler, örneğin bir kavramın temsil edilmesi sürecinde farklı katkılar sağlamaktadır. Minsky, öğrenme becerisine sahip bir bağlantıcı sistemi tasarlayan ilk kişidir ve aynı zamanda, bağlantıcı basit ağların hesaplama ve bilgi işleme sınırlılıklarını da ortaya koyan bir araştırmacıdır.[57] Klasik yapay zekâların mantık temelli düşündüğü durumda bağlantıcı sistemlerin beyin gibi düşünen sistemler tasarlamayı arzu ettiklerini söylememiz mümkündür.[58] Ancak bu durumun, Say'ın iddia ettiği gibi, sinir hücrelerinin yaptığı girdi ve çıktı ilişkilerini inceleyerek bu ilişkileri yapay sinir ağları üzerinde milyarlarca birimle gerçekleştirmenin bir sonucu olarak, makinelerde düşünmenin gerçekleştirilebilmesini mümkün kıldığı söylenebilir mi?[59]

Yapay sinir ağlarının olanakları incelendiğinde bazı tartışmalı hususların olduğu görülebilir. Akın'a göre nöral ağlar içeriksizdir. Öte yandan Minsky ve Papert'ın tek katmanlı algılayıcıların bazı problemleri çözmek konusundaki

[54] Bringsjord ve Govindarajulu, "Artificial Intelligence".
[55] Öztürk ve Şahin, "Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış", s. 32-33.
[56] Robert Audi, "connectionism", **The Cambridge Dictionary of Philosophy**, 2. bs., New York, Cambridge University Press, 1999, s. 175-176.
[57] Boden, "Artificial Intelligence".
[58] Say, **50 Soruda Yapay Zekâ**, s. 99.
[59] **A.e.**, s. 98.

başarısızlığının çok katmanlı algılayıcılara sahip yapay sinir ağları bakımından da söz konusu olduğunu iddia etmeleri[60], yapay sinir ağlarının bazı problemleri çözmek konusundaki kabiliyetini sorgulamamıza neden olmuştur. Yapay zekâ çalışmalarının hem klasik hem de bağlantıcı yaklaşımlar bakımından detaylı bir eleştirisini Minsky'de görmek mümkündür. Minsky'ye göre çağdaş sembolik yapay zekâ sistemleri bir problemin çözümü konusunda muğlak bilgiyi ele almakta çok kısıtlıdır. Bağlantıcı sistemler ise her ne kadar daha esnek sistemler olsalar da öğrenebilme yetilerine rağmen daha düşük düzeyde zekâ belirtisi gösterebilmektedirler. Sembolik sistemlerin güçlü sistemler olmalarına rağmen sağduyu bilgisini temsil edemedikleri bir durumda, bağlantıcı sistemler de iyi akıl yürütmeden yoksun olarak karşımıza çıkmaktadır.[61] Öte yandan yapay zekâların makinelerde inşa edildiği ve inşa edilme sürecinde daha genel ögelerden daha özele doğru ya da daha özel ögelerden daha genele doğru bir inşa modeli izlenebileceği düşünülürse Minsky açısından durum daha da açmaz hale gelir: Yukarıdan aşağı olarak inşa edilen yapay zekâ sistemlerindeki sorun çeşitli temel ilişkileri ele almakta eksiklik çekmeleriyken aşağıdan yukarı doğru inşa edilen yapay zekâ sistemleri ise organizasyonel limitlere sahiptir. Dolayısıyla her iki inşa modelinde de farklı bilgi türlerini ele almak bakımından çeşitli sorunlar bulunur. Bu anlamda bir yapay zekâ sisteminin inşası bakımından insan zihnine ilişkin kimi fenomenlerin matematiksel ve mantıksal olarak aksiyomlaştırılmaları, birimleştirilmeleri mümkün değildir. Çünkü Minsky'ye göre beyindeki işlevsel süreçler binlerce alt sistemden oluşmaktadır ve bu sistemler milyonlarca yıllık evrim sürecinin bir ürünüdür. O halde bir bütün olarak beyin, fizikçilerin veya biyologların bir sistemi ve parçalarını anlarken kullandığı teknikler gibi teknikler uygulanarak anlaşılamaz.[62]

Yapay zekâ çalışmalarının Darthmouth Konferansından sonra izlediği seyir düşünülürse, insan düzeyi zekânın yakalanması ve makinelerde gerçekleştirilmesi için farklı metotların denendiği görülür. Makinelerin yalnızca sayılar üzerinde işlem yapan basit cihazlar oldukları yollu düşüncenin terk edilmesiyle başlayan süreç,

[60] Akın, "Yapay Zekâda Vücut Beyin Problemi".
[61] Minsky, "Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy", s. 35-36.
[62] **A.e.**, s. 37-38.

makine zekâsının olanağı bakımından makinelerin zihinsel fenomenleri donanım üzerinde gerçekleştirebilmelerinin koşullarının soruşturulmasını sağlamıştır. Bu anlamda klasik ve sembollere dayanan yapay zekâ anlayışından günümüzde yaygın olan bağlantıcı sistemlere – derin öğrenme, derin nöral ağlar vb. – kadar uzanan çeşitli metotlar denenmektedir. Yapay zekâ sistemlerinin başarısı, bağlantıcı sistemlerin öğrenmeye dayanan gücü ile açıklanabilir. Ancak öğrenmenin kendisi düşünülürse, makineler bakımından bilginin elde edilmesi sürecinin kendisi asıl anlamda öğrenme olarak nitelenebilir mi? Burada Kurzweil, öğrenmenin yalnızca olguların edinilmesi olduğu yollu görüşe karşı çıkmaktadır çünkü bilgi, yalnızca önermesel değil aynı zamanda olguları, olgular arası ilişkileri ve bu ilişkilerin gerçekleşmelerini de içeren bir üründür. Kurzweil bu düşüncesini bilgisayarlar ve insan zihinlerinin yapısal farklılıklarına dayandırır: Bilgisayarlar sembolik ilişkiler ve olguları hatırlama konusunda sınır tanımazken bir insan zihni ise örneğin bir romanı okumak ve romandaki karakterler arasındaki ilişkiler üzerine akıl yürütmek konusunda önde gelmektedir.[63] Anlıyoruz ki bilgisayarları nasıl tasarlarsak tasarlayalım, onlara bizzat sağlayacağımız bilgilerle ya da onların dünyadaki olguları keşfetmelerini sağlayarak kendilerini geliştirmeleri yoluyla bilgisayarların, tıpkı biz insanların dünyadaki olguları anlama yetimiz vasıtasıyla kavrayabildiği gibi kavrayabileceklerini savunmak için henüz bilgisayar biliminin ve yapay zekâ çalışmalarının çok yol alması gerektiği açıktır. Çünkü en ümit verici gelişme sayılabilecek bağlantıcı sistemlerin bile, açıklama güçleri sınırlıdır. Bağlantıcı sistemler her ne kadar beynin fizyolojik yapısı hakkında önemli bilgiler sağlıyor ve beynin fizyolojik yapısının inorganik materyalde gerçekleştirilmesinin koşullarını büyük ölçüde başarıyor olsalar da bir bağlantıcı sistemin en ufak birimleri ve bu birimler arası ilişkilerden neden zihinsel fenomenlerin belirmesi gerektiği, bu zorunluluğun kökeninin ne olduğu henüz açıklanabilmiş bir fenomen değildir. Çünkü Papert'a göre de bir evrimsel biyolog, bir kaplanın neden çizgi desenlere sahip olduğunu araştırabilir. Keza bir moleküler biyolog ise yaşamın moleküler süreçten nasıl meydana geldiğini araştırabilir ancak moleküler biyoloğun yaşamın nasıl başladığı konusunda ulaştığı sonuçlar, bir kaplanın neden çizgi desenlere sahip

[63] Kurzweil, **The Age of Intelligent Machines**, s. 17.

olduğunu açıklayamaz. Buradaki benzetim açıktır: Bağlantıcı sistemler ve bağlantıcı araştırmalar yoluyla ortaya konan nöronal deneyler, düşünmenin ne anlama geldiğini açıklayamaz.[64]

2.3. Yapay Zekâ Çalışmalarında Zekâ Tanımları ve Yapay/Doğal Zekâ Ayrımı

Yapay zekâ çalışmaları içerisinde farklılaşan tanımlar ve yöntemler bulunur. Önceki alt bölümlerde de görüldüğü gibi tanımların ve yöntemlerin farklılaşması, insanın varlık koşullarının – fizyolojik ya da fenomenal unsurlar bakımından – yorumlanmaları arasındaki ayrımdan ve yapay zekâ çalışmalarındaki teknik unsurlardan kaynaklanmaktadır. İnsanın nöral yapısını kendine örnek alan ve insan zihninin fenomenal unsurlarının da bu yapıdan hareketle gerçekleştirilebileceğini savunan görüşleri görebildiğimiz bir alan olarak yapay zekâ çalışmalarında, zekânın neliği üzerine farklı tanımlamalar bulunur. Zekânın ne olduğu ve hangi davranışların zekâ gerektirdiği ile ilgili bir tartışma halihazırda zihin felsefesinin ve insan zihnini çalışan diğer bilgi alanlarının bir konusudur. Ancak yapay zekâ felsefesi ve yapay zekâ çalışmaları açısından düşünülürse görülebilir ki zekânın ne olduğu sorusu, yapay zekânın gerçekleştirilebilmesi için kilit rol oynamaktadır. "Zekâ" kavramına getirilen "yapay" ve "doğal" sıfatlarının anlaşılmasının bir koşulu olarak "zekâ" kavramının sıkıca incelenmesi gerektiğini vurguluyoruz. Çünkü bu kavramın sıkıca ele alınması, insan ve makine zekâsı arasındaki varlık farklarının aydınlatılması bakımından önemlidir. Ancak ve ancak bu yolla yapay zekâ felsefesi bakımından da pek çoklarının zihnine gelen bir soru olarak hangi zekânın "yapay" hangi zekânın ise "doğal zekâ" olduğuna ilişkin sorun çözüme kavuşturulabilir.

Yapay zekâ çalışmalarının bir hedefi zekâyı ve zeki davranışı gösterebilen bilgisayar programları inşa etmek iken, bir hedefi ise benzer görevlerde tıpkı insanlar gibi akıl yürütebilen bilgisayar programlarının olanağının soruşturulmasıdır. Bu anlamda yapay zekâ araştırmacılarının paylaştığı temel çabalardan biri zekâ taşıyan

[64] Seymour Papert, "One AI or Many?", **Dædalus**, C. CXVII, No: 1, 1988, s. 2.

ve zeki davranışı sergileyebilen sistemlerin araştırılması ve inşa edilmesidir.[65] Ancak burada sorulabilir: Zekâ nedir? Bilgisayarlar zekâ gerektiren görevleri yerine getirdikleri zaman onların asıl anlamda düşünüyor olduklarını söyleyebilir miyiz?[66] Bu soruya yanıt verebilmek için zekânın ne olduğuna ve zekânın farklı canlılar bakımından farklı mı yoksa bir ve aynı şey mi olduğuna ilişkin bir incelemede bulunmamız gerekir. Çünkü Warwick'e göre insanlar, hayvanlar ve makinelerdeki zekânın ne olduğu yanıtlanmalıdır: Zekâ bir midir? Örneğin bir örümcek için ya da bir makine için zekâ farklı mıdır? İnsan üzerinden tanımlanmış bir zekâ kavramı günün birinde uzaylı başka bir canlı formu için uygulanabilir mi?[67] Warwick, insanların dünyadaki tek zeki varlık olduklarını düşünmez. Bir varlığın başarısının zekâsını çevreye uyum sağlayabilmek ve çevreye uygun davranabilmek bakımından kullanabilmesi olarak düşünülmesi gerektiğini, bu anlamda farklı varlıklar ve makinelerin çevreye uyum sağlayabilmek için kendilerine has yöntemlere sahip olduklarını ileri sürer. Dolayısıyla dünyadaki tüm zeki varlıkları kapsayacak bir zekâ tanımına ihtiyacımız olduğunu, kapsayıcı bir tanım olarak zekânın, varlığın hayatta kalma mücadelesini olanaklı kılan farklı türdeki bilgi işleme süreçlerinin kolektif bir çabası olarak anlaşılması gerektiğini önerir.[68]

Yapay zekâ araştırmacıları bakımından ve zekâ üzerine çalışan diğer bilgi alanlarındaki araştırmacılar ve düşünürler bakımından da zekânın mutlak bir tanımının olmadığını söylememiz gerekir. Tegmark'ın da ifade ettiği gibi birbirleriyle rekabet içerisinde olan ve zekâyı farklı fenomenler üzerinden tanımlamaya çalışan birçok girişim mevcuttur. Bu girişimler mantıksal beceri, anlama yetisi, planlama gücü, öz farkındalık, yaratıcılık, problem çözme, öğrenme vb. gibi olanaklar bakımından zekâyı tanımlamaya çalışmaktadır.[69] Örneğin Alpaydın'a göre bir sisteme zeki denebilmesinin koşulu, ilgili sistemin değişen bir çevrede bulunan bir sistem olarak, öğrenme yeteneğine sahip olup olmadığı üzerinden belirlenebilir.[70] Russell açısından ise zekâ eylemler ve hedefler ile ilgilidir.

[65] Simon, "Artificial Intelligence: an empirical science", s. 96-101.
[66] Hauser, "Artificial Intelligence".
[67] Kevin Warwick, **Artificial Intelligence: The Basics**, New York, Routledge, 2012, s. 13.
[68] **A.e.**, s. 17.
[69] Tegmark, **Yaşam 3.0**, s. 74.
[70] Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, s. 28.

Eğer insanlar, türlü hedeflere ulaşmak bakımından en uygun eylemleri tercih edebiliyorsa zekidir denebilir. En uygun eylemin seçilebilmesi ve başarılı şekilde icra edilebilmesi için zekânın ortaya çıktığı süreçler olarak algılama, düşünme, öğrenme, icat etme vb. gibi hususların önemli olduğu vurgulanır.[71] Köse açısından zekâ ise bir durumun analizinde, problemin çözülmesinde ve çözümün nesiller arası aktarımında ortaya çıkar. Bu anlamda hayatta kalabilmeye ilişkin eylemsel bir yana da sahip olan zekâ, Köse'ye göre yaşamın koşulu olarak hayvanlardan bitkilere kadar problem çözebilme kabiliyetine ilişkindir.[72] Görüleceği üzere zekânın tanımlanması çabaları canlılıkla bağı bakımından bir yandan insansı bir yandan da insansı olmayan mekanik bir tarzda şekillenmektedir. Ancak bu tanımlar kimi zafiyetler barındırmaktadır: Örneğin Köse'nin yaklaşımı ele alınırsa, hayatta kalmak ve bu uğurda mücadele etmek zorunda olmayan muhtemel inorganik yaşam formlarının bulunduğu bir evrende zekâ nasıl tanımlanacaktır? Öte yandan zekâ yalnızca problem çözebilmeye indirgenirse, akla gelebilir ki insan zihninin kimi yetileri, kendileri veya daha alt birimleri bakımından problem çözme yetisi ile ilişkilendirilemez. Örneğin hayal kurabilmek, tasarlayabilmek veya bir sanat eserini üretebilmek de bu işleri ortaya koyabilmek için öne çıkan bir zihin gücü olarak zekâyı gerektirmektedir. Oysa bu işlerin başarılabilmesi problem çözme ile açıklanamaz çünkü örneğin bir sanat eserinin başarılabilmesi bir problem ya da bir problemin sonucu olarak tanımlanamaz.

Zekâ, matematiksel ve mantıksal bir temelde de tanımlanabilir. McCarthy'nin zekâ tanımı böyle bir temelde şekillenir: Bir varlık, kendi amaçlarını ve diğer zihinsel süreçlerini matematik temelde inşa edebileceği bir dünya modeline sahipse ve bu modele dayanarak çeşitli problemlere yanıtlar türetebiliyorsa zekidir denebilir. Bu anlayışa göre zekânın iki parçası bulunur: Epistemolojik temel ve problem çözme. Bir zekâ, içinde bulunduğu dünyayı ve bu dünyada yer alan nesneleri temsil edebilmeli ve bu temsillere dayanarak problemleri çözebilmeli, problem çözme yeteneğine dayanarak da problemin çözümüne ilişkin atılacak adımların ne olacağına

[71] Russell, **İnsanlık için Yapay Zekâ: Yapay Zekâ ve Kontrol Problemi**, s. 22.
[72] Köse, **Yapay Zekâ Felsefesi**, s. 14.

karar verebilmelidir.[73] Fakat burada da bir sorun bulunur: Zekânın problem çözmeyle olan ilişkisi kabul edilse bile, inşa edilen bir zekânın matematiksel ve mantıksal kimi problemler taşıyacağını söylememiz mümkündür. Matematiksel ontolojinin bizim gerçekliğimizle olan ilişkisindeki kapalılığına vurgu yapılmıştır. Nitekim okuyucu açısından da bu vurgu anımsanırsa, matematiksel ve mantıksal temele dayanan bir zekâ, asıl anlamda zekâ değil ancak ve ancak simüle bir zekâ olacaktır. Çünkü gerçeklik farkından ötürü, matematiksel ontolojiye ait olan bir zekâ, gerçekliğimizdeki nesneleri ve nesneler arası ilişkileri temsil edebiliyor olsa bile hiçbir zaman temsil ile temsil edilen şey arasındaki ilişkiyi kavrayamayacak, semantik gediği aşamayacaktır. Bu anlamda Zambak ve Vergauwen'in zekânın gözlemlenebilir örneklerinin yalnızca canlı organizmaların zihinleri olarak özellikle insanlar olduğu yollu vurgusunun[74], Doğan'ın "zekâ" kavramının hesap yapma, analiz etme, akıl yürütme vb. gibi kimi yetilerden daha geniş bir anlama sahip olduğu yollu tanımlamasıyla[75] beraber düşünüldüğünde zekânın tanımlanması sürecinin ve zekânın insanın varlık koşullarıyla olan ilişkisinin önemini vurguluyoruz.

Zekâyı tanımlama girişimlerimiz, zekânın atfedildiği varlıkların koşullarından bağımsız olamaz. Çünkü zekâ, var olmaklık bakımından ilinekseldir. Zekânın belirmesinin koşulu bir varlık tarafından sahip olunması ve kullanılmasıdır.[*] O sebeple zekânın öğrenme, anlama, düşünme, planlama vb. gibi yetilerle ilişkilendirilmesi[76] ilk adımda olağan görünebilir: Zekâ üzerine düşüncemizde önümüzde güçlü bir örnek olarak insan bulunmaktadır. Şimdiye kadar zekâ ile ilişkili olduğu düşünülen birçok yetinin insan doğasıyla olan ilişkisi düşünülürse zekânın antropomorfik kimi tanımlama girişimlerinin odağında olduğu söylenebilir. Öte yandan failliği matematiksel ve mantıksal bir temelde görmek yollu yapılmaya çalışılan zekâ tanımları, fail ve çevre arasındaki ilişkide inşa edilmeye çalışılır. Wang'a göre zekâ, bir sistemin yetersiz bilgi ve kaynaklara rağmen çevresine uyum

[73] John McCarthy ve Patrick J. Hayes, "Some Philosophical Problems from the Standpoint of Artificial Intelligence", (Çevrimiçi) http://www-formal.stanford.edu/jmc/mcchay69.pdf, 8 Aralık 2022.
[74] Zambak ve Vergauwen, "Artificial Intelligence and Agentive Cognition: A Logico-Linguistic Approach", s. 95.
[75] Doğan, "Yapay Zekâ ve Özgür İrade: Yapay Özgür İradenin İmkânı", s. 790.
[*] Eğer bir düalist değilseniz…
[76] Alagüney, "Hukukta İrade Özgürlüğü Kavramının Yapay Zekâ Türleri Bakımından Tartışılması" s. 203.

sağlayabilmesi olarak tanımlanır.[77] Dikkat edilirse antropomorfik yüklerden arındırılmış bir tanım olarak Wang'ın tanımı, çevrenin veya failliğin söz konusu olmadığı durumda gerçekleşebilecek eylemleri kapsamamaktadır. Böylece zekâ, eylemin olmadığı bir yerde kendi başına bir olanak olarak düşünülememektedir.

Zekâyı tanımlama girişimlerimizin miktarı, bize zekânın ne olduğu üzerine bir uzlaşının bulunmadığını gösterir. Bilimler düzeyinde de soruşturulan zekâ, çoğu zaman zekâ gerektirdiği düşünülen işler üzerinden tanımlanmaktadır. Görüntü tanıma, problem çözme vb. işler söz konusu olduğunda biz insanlar bu işleri nasıl gerçekleştirdiğimizi bilmiyoruz ancak bu işlerde oldukça başarılıyız. Başarılı olduğumuz ve zekâ gerektirdiğini düşündüğümüz işlerin ise süreçlerine hâkim olamadığımız için zekânın tanımını ve ne olduğunu ortaya koyamayız.[78] Zekânın birçok tanımı söz konusuyken bu durumda yapay zekâ kavramımızı yeniden incelersek, zekânın yapaylığı ne anlama gelmektedir? Sokolowski'ye göre yapay zekâ derken "yapaylık" sıfatının iki kullanımı söz konusu olabilir: İlk kullanımda yapaylık, örneğin bir çiçeğin yapay olması gibi bir kullanımdır. Yapay bir çiçek, gerçek bir çiçeğin yalnızca *görünüşte* yerini tutabilir. İkinci kullanımında yapaylık ise bir ışığın yapay olduğunu söylediğimizdeki kullanımı ile eştir.[79] Doğal bir ışık kaynağı olarak Güneş ile yapay bir ışık kaynağı olarak herhangi bir ampul, üretilen ışık bakımından bir fark ortaya koymaz. Güneşin ürettiği ışık ile ampulün ürettiği ışık arasında hiçbir nitelik farkı bulunmaz. İşte yapay zekâ felsefesinde, bu çalışmanın da odağında olan böyle bir yapaylık yorumu söz konusudur. Yapay zekâ çalışmalarında bazı araştırmacılar, yapaylığı tıpkı yapay ışık kullanımındaki gibi bir kullanımla kabul etmektedirler[80] ve yapay zekâların, canlı zekâsı[*] ile herhangi bir farkının bulunmadığını ileri sürmektedirler. Yapay zekâ ve canlı zekâsı – literatürdeki kullanımı bakımından doğal zekâ – arasında kurulmak istenen bu

[77] Wang, "The Logic of Intelligence", s. 33.
[78] Alpaydın, "Yapay Zekâ/Doğal Bilgisayarlar", s. 98.
[79] Robert Sokolowski, "Natural and Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 45.
[80] Manzotti ve Tagliasco, "Artificial consciousness: A discipline between technological and theoretical obstacles", s. 106.
[*] Yapay zekâ/doğal zekâ karşıtlığının türetilmesinin, zekânın anlaşılması ve yapay zekânın konumlandırılması çabalarını anlaşılmaz ve çözümsüz kıldığını düşünüyoruz. Dolayısıyla doğal zekâ kullanımından ziyade canlı zekâsı kullanımını öneriyoruz. Yapay zekâ felsefesindeki felsefi işlevselci tartışmaların ve güçlü yapay zekâ argümanının odak noktalarından biri olarak biyolojik olanla inorganik olan üzerine söylenilenler düşünülürse, meselenin canlılıkla olan ilintisi oldukça açıktır.

geçişken ilişkinin temelini incelediğimizde, zekânın tanımlanması sürecinde makineler üzerinden gerçekleştirilmeye çalışılan çabalarla karşılaşırız. McCarthy ve Hayes'e göre bir makine, insanların belli problemleri çözmeleri için zekâya gereksindiği problemler sınıfını çözebiliyorsa ve çevresine uyum sağlayarak hayatta kalabiliyorsa zeki olarak tanımlanabilir.[81] Acar'a göre ise hedeflere ulaşabilmek bakımından belirli bir irade ortaya koyabilen makinelere zeki denebilir.[82] Nilsson ise bir sistemin zeki olabilmesinin koşulunu, sistemin dünya hakkındaki bilgilere sahip olabilmesi ve bunlardan sonuçlar çıkaracak araçlara sahip olabilmesi üzerinden tanımlar. Bu anlamda bilişsel bilimciler ve yapay zekâ araştırmacıları bilginin temsil edilmesine ilişkin iki türü dile getirirler: Prosedürel ve dekleratif. Bir davranışın yerine getirilmesi için ihtiyaç duyulan bilgi türüne prosedürel bilgi denirken örneğin hava kaç derece gibi bir sorunun yanıtı dekleratif bir bilgi türüdür.[83] Fakat Doğan'ın da belirttiği gibi zekâ, dünyadaki hedeflere ulaşmak bakımından bu yeteneğin bilgisayımsal bir parçasıdır. Bu anlamda biz hangi bilgisayımsal süreçlerin zekâ gerektirdiğini ve zeki olduğunu ifade edebileceğimizi henüz bilmiyoruz.[84] Dolayısıyla ne makine zekâsının bilgisayara aktarılmış bir canlı zekâsı – Köse'nin kullanımıyla *insan taraflı yorumlanan bir zekâ* anlayışı olarak doğal zekâ kullanımı – olduğu savunulabilir[85] ne de canlı zekâsının akla gereksinmeyen, bunun yerine algoritmalara ve çözüm adımlarına indirgenebilen bir makine zekâsı ile herhangi bir farkının bulunmadığı savunulabilir.[86] Hatta Alpaydın gibi bazı bilgisayar bilimciler bir adım daha ileri giderek, yapay zekâ ve canlı zekâsı arasındaki ayrım hakkında yürütülen tartışmaları, *biyolojik olanın şoven duygulara dayanarak savunulduğu tartışmalar* olarak niteleyecek, böylece beyni, belirli ilkelere dayanan bir makine olarak görmenin aşağılayıcı bir yanının bulunmadığını ileri sürecektir.[87] Bu çalışmanın metafelsefi incelemeler bölümü altında da sıkça dile getirilmiştir ki,

[81] McCarthy ve Hayes, "Some Philosophical Problems from the Standpoint of Artificial Intelligence".
[82] Ayşe Acar, "Autonomy in Human and Artificial Intelligence and Its Relationship with the Concepts of Dignity and Human Rights", İstanbul, İstanbul Bilgi Üniversitesi Lisansüstü Programlar Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2020, s. 17.
[83] Nilsson, **The Quest for Artificial Intelligence**, s. 199.
[84] Mehtap Doğan, "Yapay Zekâ Felsefesinde Bilinç Problemi", Ankara, Ankara Yıldırım Beyazıt Üniversitesi Sosyal Bilimler Enstitüsü, Yayımlanmamış Doktora Tezi, 2020, s. 2-3.
[85] Köse, **Yapay Zekâ Felsefesi**, s. 17-18.
[86] **A.e.**, s. 23.
[87] Alpaydın, "Yapay Zekâ/Doğal Bilgisayarlar", s. 102.

yapay zekâ felsefesi içerisine sık sık bilgisayar bilimciler tarafından taşınmak istenen bu tavır, biz felsefecilere ve felsefenin doğasına hiç ait olmayan bir tavırdır. Felsefe içerisinde bir kavramın ele alınışı, duygulanımlardan yalıtıktır. Felsefecinin işi retorik yapmak değildir. Felsefecinin işi *kavramların aydınlatılmasıdır*. Beyni bir makine gibi görmemek ve onun doğasının sayısallaştırılamayacağını ileri sürmek, ne onun doğasını anlamak istemiyor olduğumuz anlamına gelir ne de beynin dokunulmaz ve gizemli bir yapı olduğunu düşündüğümüzü gösterir. Çünkü felsefe alanında kavramlar bu yollu ele alınmaz. Felsefe etkinliği içerisinde kavramlar kendileri olarak ele alınırlar. Biz, kavramları onurlu veya aşağılayıcı olmaları bakımından değil, tutarlı olmaları ve var olanla örtüşmeleri bakımından ele alırız. Metafelsefi incelemeler bölümünü bu çalışmanın planlanması bakımından önde görmemizin sebeplerinden biri tam da çoğunlukla bilgisayar bilimciler tarafından felsefeye taşınmak, kabul ettirilmek istenen bu ve benzeri kimi varsayımlardır. Yapay zekâ devrimi, sıkıca tartışılması gereken bir devrimdir. Yapay zekâlı varlıklar örneğin yalnızca emek meselesi üzerinden ele alınamaz. Onlar yaşamın her sahasında yaşayışımızı ve düzenimizi değiştirmeye aday varlıklardır. Dolayısıyla bilgisayar bilimcilerin metafizik kabullerinden önce felsefenin kavramsal incelemesinin gelmesi gerektiğini savunmak doğaldır. Çünkü yaşamın yalnızca birkaç sahası bakımından kabul ettirilmek istenen bir kavramsal anlayışın, yaşamın diğer sahaları bakımından düşünülürse, çelişkiler doğurmayacağını söylemek olanaksızdır. Bu durumun hesaba katılması bakımından biz felsefeciler bir kavramı *yalnızca kendisi olarak değil,* diğer kavramlarla ve alanlarla olan ilişkisi bakımından da ele alırız. Böyle bir yöntemin, felsefece olmayan kimi yöntemlerle olan farkı Aristoteles'in *Sofistçe Çürütmeler* adlı metninde de açık bir şekilde ortaya konmuştur.

Yapay zekâ felsefesinde "zekâ" kavramının ele alınma tarzı, sıklıkla zekânın tüm canlılarca ve hatta makineler tarafından da sahip olunan lineer ve tek bir şey olduğu kabulü etrafında şekillenir. Makine zekâsının insan zekâsını aştığı, insan zekâsının hiçbir zaman bazı alanlarda makine zekâsını yakalayamayacağı yollu görüşler, zekânın lineerliği görüşünden hareketle makinelerin donanım anlamında daha güçlü ve hızlı olmalarına dayandırılır. Ancak Warwick, lineer bir zekâ anlayışını kabul etmez, farklı varlıkların zekâları arasında bir hiyerarşi olduğunu

reddeder. Ona göre bir varlığın zekâsı, o varlığın var olma tarzı bakımından – örneğin duyargalarının sayısı – belirlenir. Bu durumda makine zekâsı ve insan zekâsı arasında ya da bir hayvan zekâsı ile insan zekâsı arasında hiyerarşik bir ilişki değil ama bir farklılık ilişkisi söz konusudur.[88] Üstelik çevrede gerçekleşen bir olgunun anlaşılması, o olguya ilişkin bilginin edinilmesi sürecinde bir makinenin o olguyla ilişkisi ile bir insanın veya bir hayvanın ilişkisi farklıdır. Dolayısıyla bir olguya ilişkin epistemolojik süreçlerin farklılık gösteriyor olması, makine, insan veya hayvan zekâsı arasında birinin diğerine göre daha zeki olduğunu değil yalnızca zekâların farklı olabileceğini gösteren önemli bir durumdur.[89] Ancak Sokolowski'nin yaklaşımı düşünülürse, makinelerin zekâ sahibi olabileceği ve makinelerin dünyaya ilişkin kimi olguları edinebileceği düşüncesine imkân tanıyan Warwick'in modern yaklaşımı kabul edilemez. Çünkü "yapay" sıfatının kullanımları hatırlanırsa, örneğin her ne kadar yapay ışığın aydınlattığını ya da uçağın uçabildiğini söyleyebiliyor, bunların yapay şeyler olarak gerçek ilişkilerin ve varlıkların yerini tutabiliyor olduklarını kabul etsek bile, düşünmenin kendisi bu ilişkiler veya varlıklarla benzer statüde ele alınamaz. Düşünmenin kendisi biçimsel bir dizge içerisinde tüm unsurlarıyla serimlenemez. Bu anlamda, insan düşünmesinin yerine düşünüyor gibi davranan bir makine koymamız makinenin düşündüğü anlamına gelmez. Örneğin bir hesap makinesi, kendi düşünmesini değil, onu tasarlayan ve programlayan kişinin veya kişilerin düşüncelerini ortaya koymaktadır.[90] Ancak burada okuyucu itiraz edebilir: Sokolowski'nin bu itirazı neticesinde zayıf yapay zekâ düşüncesinin başarıları yadsınmış olmuyor mu? Bugün birçok özelleşmiş yapay zekâ programı, tasarlama, planlama ve uygulama konusunda birçok alanda insan düşünmesini modelleyemiyor mu? Sokolowski, makine düşünmesinin doğal düşünmeyi – Sokolowski doğal düşünme ile kendisinin de doğal zekâ olarak nitelediği insan ve hayvan zekâsına ait düşünmeyi imliyor – kısmen gerçekleştirebileceğini kabul etmektedir. Ancak bu kabiliyetler gerçekleşirken olan şeyin sinir sisteminin veya beynin taklit edilmesi olduğu kabul edilmemektir. Bir makinenin zekâsı kodlardan ve onların dönüşümlerinden başka bir şey değildir. Oysa doğal zekâ, bir makine zekâsı

[88] Warwick, **Artificial Intelligence: The Basics**, s. 17.
[89] **A.e.**, s. 64.
[90] Sokolowski, "Natural and Artificial Intelligence", s. 46-47.

gibi sadece mekanik süreçlerin gerçekleştiği organik yapıda bir beynin ürünü değil aynı zamanda zekânın etkileştiği çevre ile, çevredeki fizik nesnelerle bir bütün olarak düşünülmelidir.[91] Bu anlamda doğal zekânın alıntılama, ayrım yapabilme ve arzu gibi kimi yetileri bir yapay zekâ tarafından gerçekleştirilmedikçe, bir yapay zekânın bir doğal zekâ olarak insan zekâsını gerçekleştirdiği söylenemez. Çünkü yapay zekâlar *çıkarım makineleridir. Mantıksal çıkarımlar yapmak ise zekânın yalnızca bir kısmını oluşturur.* O halde bir zekâyı doğal zekâ olarak niteleyeceksek, o zekânın ilk olarak başka zekâların dile getirdiği şeyleri alıntılayabilmesi gerektiğini söyleyebiliriz. Alıntılamak, yani bir başka zekânın dilsel ifadelerini dile getirebilmek, yalnızca kelimelerin taklit edilmesi değildir. Doğal zekâlı bir varlık aynı zamanda bir başka varlığın deneyiminden yeni bir perspektif kazanır. Çünkü bir şeyin bir başka varlığa nasıl göründüğü, onun fenomenal deneyimine ait bir durumdur. O halde doğal zekâlı varlıklar dil aracılığı ile deneyimlerini paylaşabilir ve bir fenomene ilişkin deneyimlerini zenginleştirebilirler.[92] Ancak sorulabilir: Makineler de kendi aralarında bilgi paylaşmıyorlar mı? Makinelerin birbirleri arasındaki deneyim paylaşımları yoluyla kendilerini eğittikleri gerçeğini yadsıyor muyuz? Öncelikle deneyim ve onun kurulması bakımından makine ve insan zekâsı arasındaki epistemolojik koşullar incelenirse, insan zihninin deneyimi kurması, bir yazılım dili ya da genel olarak biçimsel bir dil üzerinden değil, henüz yapısını ve çalışma ilkelerini bilmediğimiz bir tarzda gerçekleşmektedir. Çünkü insanlarda deneyim, birbirleriyle ne kadar benzer kurulursa kurulsun hiçbir zaman aynı şekilde kurulamaz. Örneğin Bennu ve Nazım, kırmızı bir elmanın düşmesine şahitlik etsinler. Bennu ve Nazım'ın söz konusu kırmızı elmanın düşmesi olayını rapor etmeleri istendiğinde verecekleri yanıt ancak ve ancak benzer olabilir, aynı olamaz. Dolayısıyla Bennu ve Nazım, bu olaydan daha sonra kendi aralarında söz konusu kırmızı elmanın düşmesi olayı üzerine bir konuşma gerçekleştirdiklerinde tam da Sokolowski'nin alıntılamaya ilişkin vurgusu gerçekleşmiş olur: Hem Bennu hem de Nazım, birbirlerinin elmanın düşmesi ile ilgili deneyimlerini elde ederler, bunları gerekirse başkalarına dilsel olarak aktarırlar ve kendi deneyimlerini zenginleştirirler. Oysa bir makine için durum böyle değildir. Makine zekâsı biçimsel bir dizgenin,

[91] **A.e.**, s. 47-48.
[92] **A.e.**, s. 57-58.

yani bir programın, ürünüdür. Dolayısıyla kırmızı bir elmanın düşmesi olayına şahitlik eden birbirinden ayrık birden çok makine zekâsını görüyor ve sayabiliyor olsak bile bu zekâların hiçbiri biçimsel açıdan birbirinden ayrık olarak kabul edilemez. Dolayısıyla sayıca birden çok zekâ olduğu görülüyor olsa bile asıl durum yalnızca *bir* ve biçimsel olan bir makine zekâsının, yani bir programın farklı açılardan aynı duruma ilişkin deneyimlerinin depolanmasıdır. Tüm bu biçimsel makine zekâları birbirleri arasında sentaktik kurallara uygun olarak kurulmuş matematiksel/mantıksal yargılar yoluyla bilgi paylaşıyor gibi görünebilirler. Oysa olan tek şey, makine zekâsının failsi niteliği kabul edilse bile, bir ve aynı fenomenal deneyime sahip bir zekânın bir olguyu farklı açılardan deneyimlemesinden daha fazlası değildir. Dolayısıyla bir makine zekâsı, hiçbir zaman bir olgu hakkında tıpkı insanların birbirleri arasında deneyim paylaşımı yoluyla birbirilerinin deneyimlerini zenginleştirmeleri gibi bir yaratıcı/buluşçu unsurlara sahip olamaz. Bu durumda bir makine zekâsı, kuralları ona daha önceden verilmedikçe iki şey arasında da ayrım yapamaz.[93] Oysa bir doğal zekâ, şeyleri bir ve aynı biçimsel dizge içerisinde deneyimleştirmediği için daima iki şey arasında bir ayrım yapmaya doğru yönelmektedir. Ayrım yapamayan bir zekâ için, doğal zekânın koşullarından biri olarak Sokolowski açısından, arzunun da sağlanamayacağı açıktır. Çünkü arzu, bir fail için en az iki şey arasında ortaya çıkan ve düşünmemize amaçlarını veren bir yönelimdir. Bu yönelimin kendisi rasyonel herhangi bir tartışmanın konusu olamadığı gibi makine zekâsı hiçbir zaman doymak ile yemek yemek arzuları arasındaki ya da barınmak ile bir ev istemek arzuları arasındaki nitelik farkını anlayamaz, böyle bir arzuya sahip de olamaz. Çünkü arzularımız, zekânın alanında değildir ve zekâ tarafından belirlenmez.[94] O halde denebilir ki yapay zekâlar ne kadar zeki kılınırsa kılınsın, sahip oldukları *zekâ,* insansı problemlere ve insansı kaygılara daima yabancı kalacaktır.[95]

[93] **A.e.**, s. 60.
[94] **A.e.**, s. 61.
[95] Joseph Weizenbaum, **Computer Power and Human Reason**, New York, W. H. Freeman and Company, 1976, s. 213.

2.4. Yapay Zekâ Çalışmalarında "Zihin" Kavramının Yadsınması

Yapay zekâ felsefesi içerisindeki felsefi işlevselci tavrın eleştirisini sunmak bakımından bu çalışmada üzerinde durmamız gereken hususlardan biri, yapay zekâ alanının kendisinin de onun üzerinde inşa edildiği kavram olarak "zekâ" kavramıdır. Bu anlamda "zekâ" kavramını, yapay zekâ felsefesi içerisinde "zihin" kavramı karşısında konumlandırılmak istendiği yükleriyle ele almak ve buradan hareketle yapay zekâ felsefesi ve yapay zekâ çalışmaları bakımından "zihin" ve "zekâ" kavramları arasındaki nitelik ve anlam farklarına vurgu yaparak, bu kavramlar arasındaki kavramsal karmaşanın ortadan kaldırılabileceğini düşünüyoruz.

"Zihin" ve "zekâ" kavramlarının koşullarının belirlenmesi, insan yaşamından ve insanın zihinselliğinden ayrı düşünülemez. Ancak bu kavramların okuyucu açısından da açık kılınabilmesi ve bu alt bölümde gerçekleştirmeyi amaçladığımız kavramsal inceleme uğruna, söz konusu kavramların etimolojik arka planlarının serimlenmesini bir öncelik olarak kabul ediyoruz. "Zihin"[*] kavramı, bir çatı kavram olarak, insan zihinselliğinin tüm kiplerini kapsar. İnsanın zihnin ilkelerine[**] uygun olarak düşünmesi, bir şeylerin farkında olabilme yetisine sahip olabilmesi, bir şeyleri hatırlayabilmesi, kimi düşüncelere odaklanabilmesi, bir yargıda bulunabiliyor olması ya da bir kanıya sahip olabilmesi gibi sayısız olanaklılık, "zihin" kavramının altına yerleştirilebilir. İnsan zihni, tam da bu kipler yoluyla kendini ortaya koyar. Bir insanın zihin sahibi olduğunu ya da zihnini iyi kullandığını dile getirdiğimizde onun zihninin gerçekleşmeleri bakımından bir veya birkaç kipten birini başarı ile yerine *getirdiğini* veya en azından *yerine getirebildiğini*, zihnin kendini *bu kipler aracılığı ile* gösterebileceği azami iradeyi ortaya koyabildiğini söylemiş oluruz. Örneğin Bennu, kırmızı bir elmayı yemek arzusuna sahip olsun. Karşısında duran elmaya erişebilmek için bir planlama yapmak, bu planı uygulamaya koymak ve bu planı gerçekleştirirken bir irade sergilemek durumundadır. İşte Bennu'nun zihninin kendini bu ve benzeri kipler aracılığı ile ortaya koyduğunu söyleyebiliyor böylece Bennu'nun, *zihninin dışsallaştığı* bu ve benzeri kipler bakımından gösterdiği başarı

[*] Lat. **Mens**
[**] Mantık ilkelerine uygun olarak düşünmek.

ölçüsünde de bir zihne sahip olduğunu dile getirebiliyoruz. Peki zekâ nedir? "Zekâ" kavramı neden yapay zekâ çalışmalarının temelindedir? Neden yapay zekâ çalışmalarında düşünen bir zihin bir zekâ olarak kurgulanmıştır da bir zihin olarak kurgulanmamıştır? İnsan zihninden uzaklaşılması ve "zekâ" kavramında ısrar edilmesi herhangi bir kavramsal karmaşanın sonucu mudur?

Yapay zekâ çalışmalarında "zekâ" kavramı, Latince karşılığının[*] isimleşmesi neticesinde İngilizce diline yerleşmiş olan karşılığına[**] dayanılarak düşünülür. İsimleşen karşılığı ile "zekâ" kavramı, Latincede ayırdına varmak, anlamak anlamındaki içeriğinden – ki bu anlamlarıyla fiil olarak zekâ, "zihin" kavramının bir kipi olarak kabul edilebilir – koparılmış ve İngilizce dilindeki isim karşılığı bakımından olguları, doğruları veya bilgileri *anlayabilme* ve *edinebilme* gibi bir anlamda zihnin kapasitesi, zihinsel çabukluk veya anlama yetisinin derinliği bakımından bir nitelik olarak da bu yetilerin kullanılabilmesi ile ilgili bir içeriksel zemine oturtulmuştur. Okuyucu açısından dikkat edilirse "zekâ" kavramının modern kullanımı, zihnin kendisi veya yetilerinin *içeriği bakımından değil,* zihnin ve yetilerinin *kullanımındaki başarılara ilişkin* bir konuma getirilmiştir. Böylece bir insanın *düşünmesinin gücü ve isabetliliği* ile *düşünmesinin kendisi ve zihnin diğer yetileri* bakımından önemi, "zekâ" kavramımızın kullanımı bakımından iki ayrı alandır. Yapay zekâ çalışmaları içerisinde yer alan metafizik kabuller bu çalışmada detaylıca ele alınmıştır. İşte bu kabullerin ve genel olarak yapay zekâ alanının oturduğu zemin, "zekâ" kavramının isimleşen ve zihin ile onun yetilerini kendileri olarak değil *kullanımdaki başarılarıyla,* yani *nicelikleriyle* ele alan anlamıdır. Şimdiye kadar matematiksel, mantıksal ve doğa bilimleri bakımından ortaya koyduğumuz incelemelerin hizmet ettiği ve okuyucu açısından da şimdiye kadar sezgisel olarak işaret ettiği husus tam da budur: *Zekâ, niceliklerle ve niceliklerin başarılmasıyla ilgilidir. Oysa zihin, nicelikleri de kapsayan ancak niceliklerden ibaret olmayan bir çatı kavramdır.* İnsanın insan olmaklığını zihnin uygun bir modeli üzerinden düşünebiliriz, bu anlamda bu çalışma işlevselci savların değerini ve işlevselci girişimlerin önemini yadsımamaktadır. Ancak bu çalışmada karşı çıkılan görüş felsefi işlevselci savların "zekâ" kavramı üzerinden, zihnin yalnızca

[*] Lat. **Intelligo**, **Intelligere**
[**] İng. **Intelligence**

niceliklerin elde edilmesi yoluyla modellenebileceği ve gerçekleştirilebileceğine ilişkin sahip olduğu metafizik kabullerdir. Örneğin Bennu, yine söz konusu kırmızı elmayı yemek arzusuna sahip olsun. Yine bu elmaya ulaşabilmek için planlar yapsın. Zihni yadsıyan ve yerine "zekâ" kavramını yeterli gören çoğu yapay zekâ araştırmacılarının savı şudur: Bennu'nun kırmızı elmayı yemek istemesi bir problemdir, bu problemin başarılabilmesi için çevre ve koşullar matematiksel ve mantıksal modellemeler yoluyla ortaya konmalıdır ve Bennu'nun elmayı yiyebilmesi sonucuna en yakın çözümler en başarılı çözümler olarak kabul edilmelidir. Bu anlamda elmanın yenmesi ile sonuçlanan eylem dizileri, zekânın en açık göstergesidir. Zihin sahibi varlık olarak Bennu ile biçimselleşen ve sonuca götüren niceliklerin tespit edilmesi yoluyla benzer başarıyı sağlayan bir sistemin benzer fenomenal unsurlara sahip olması gerektiği de kolayca öne sürülür. Ancak zihin sahibi bir varlık olarak Bennu'nun elmayı arzulaması ve elmaya yönelmesi olayı ile salt zekâ olarak Bennu'nun bunu planlaması ve planlarken biçimsel bir epistemolojik düzeyde kalması arasında herhangi bir nitelik ve derece farkı görmemek için herhangi bir nedenimiz bulunmaz. Çünkü ilk durumda Bennu için olan şey yalnızca elmanın arzulanması ve elmaya erişmek değildir. Burada hesaplama ve zekâyı gerektiren kimi süreçler söz konusudur. Bennu'nun elmaya erişmek için uzanması gereken mesafenin hesaplanması ve kuvvetin nicelikçe belirlenmesi zaruridir. Ancak tüm bu olaylar devam ederken niceliğe indirgenemeyecek kimi fenomenal unsurların, duygulanımların ve arzuların vb. olduğunu unutmamamız gerekir. Dolayısıyla zekâ, bir zihinsel durumun tümünü değil ancak ve ancak bir kısmını kapsayabilir.

Yapay zekâlı varlıkların güçlü örneklerini gördüğümüz satranç oyununu düşünelim. Okuyucu bu noktada sorabilir: İşte, 64 kareden ve 32 taştan oluşan satranç oyununda olan şey, zihin sahibi varlıklar olsalar bile örneğin Bennu ve Nazım'ın yalnızca zekâlarının – yani hesaplama güçlerinin – ön plana çıktığı bir oyun değil mi? Bu alanda başarılı olan bir yapay zekânın, zihni zekâ üzerinden karşılayabileceğini söyleyemez miyiz? Biz bu itiraza iki gerekçe ile karşı çıkıyoruz. İlkin, satranç oyununun kendisi ya da herhangi bir oyunun kendisi yalnızca *o oyunun oyun olmasını sağlayan ilkelere, kurallara ve notasyonlara indirgenemez.* Satranç

oyunu veya herhangi bir oyun, yalnızca onu oluşturan kurallarla neliklerini kazanıyor gibi görünseler de herhangi bir oyunun, insansılıkla ve insansı olanla birlikte *oyunlaştığını ve neliğini kazandığını* savunuyoruz. Böylece satranç oynayan iki insan ile iki yapay zekâ arasında bir fark bulunur: Satranç oynayan iki insan, yalnızca satranç oyununun kurallarını işletmez. Aynı zamanda rakibin bir insan olduğunu bilir, en uygun hamleyi bulmak için üzerine etkiyen tüm duyu verileri ile mücadele etmek zorundadır ve olasılıkları hesaplamak bakımından sınırlı işlem gücüne sahip bir varlık olarak en uygun kararı vermeye çalışır. Satranç oyunu basitçe bir tarafın diğerini mağlup etmeye çalıştığı ve bunu yalnızca hamlelerle başardığı değil aynı zamanda tarafların kendi kusurlarıyla da mücadele ederek sonuç elde etmeye çalıştığı bir oyundur. Dolayısıyla bir insan bir başka insanı – salt düşünen varlık olarak değil, duyumsayan ve kendiliği olan bir varlık olarak bir faili – mağlup etmeye çalışır. Bu anlamda genel olarak bir oyun, insansı bir doğaya sahiptir. Bu doğa kusurludur, eksiktir, duyguları ve fenomenal yaşantıyı içerir. Bunun sonucu olarak bir oyunun oynanması ve oyunun sonucunda bir galibin ve bir mağlubun bulunuyor olması sadece o oyun açısından değil, insan yaşantısı bakımından da bir sonucu, bir anlamı kendinde taşır. Biz bu anlama ve çıktıya dayanarak bir oyuna "oyun" diyebiliyoruz. Oysa iki makine arasında gerçekleşen *görünürdeki satranç oyunu*, bir satranç oyunu olarak nitelenemez: Makineler ne satranç oynadıklarının ayırdındadır ne de matematiksel ontolojileri içerisinde satranç tahtası, masası, ortamı veya taşı bir anlam taşır. *Makineler için olan tek şey biçimsel bir dizgenin, bir başka biçimsel dizge ile olan mücadelesidir.* Burada matematiksel bir denklem eşitlenmeye çalışılır. Bir taraf bir tarafa üstün gelir. Oysa biz insanlar, yorumlayıcı varlıklar olarak, makineler arasında gerçekleşen ve *satranca benzeyen bu etkileşimi* bir satranç müsabakası olarak görürüz, öyle kabul ederiz ve *kabul etmek isteriz*. Bu çalışmanın antropomorfik yüklerle ilgili alt bölümünün okuyucu açısından da açık kılmaya çalıştığı hususlardan biri de buydu: Makineler arası etkileşimleri insansı görmeye olan eğilimimiz bize makinelerin de satranç oynadıklarını düşündürmektedir. Oysa bu düşünce için sezgilerimizden başka hiçbir gerekçemiz yoktur. Satranç oyuncularının *makine hamlesi* ve *insan hamlesi* ya da *insan satrancı* gibi ifadeleri kullanmasının arkasında da bu anlayış yatmaktadır: Makineler satranç oynamamaktadır, yalnızca sembolleri manipüle etmektedir. Makineler insansı

olamayacak düzeyde anlamdan yoksun hamlelerde bulunabilmekte ve bunun iki taraf arasındaki dengeyi bozmadığını kabul etmektedirler.

İkinci olarak, satranç oyunu veya herhangi bir oyun *kapalı bir sistemdir*. Bir kapalı sistemde başarılı olan zekânın, yaşamın diğer sınırsız zenginliği, problem çeşitliliği ve kapalı sistemleri bakımından da başarılı olabileceğini iddia edebilir miyiz? Böyle bir arzu güçlü yapay zekâ düşüncesinin temelinde yer alır. Ancak zihin gibi çok çeşitli yetileri kendisinde barındıran bir yapının, gerçekliğimizdeki sayısızca problem bakımından başarısı ve esnekliği karşısında makine zekâsının özelleşmiş biçimsel işleyişinin, yalnızca hesaplamaya dayanan yapısından hareketle zihni gerçekleştirebileceğini savunamayız. Böylece yapay zekâ felsefesinin ve yapay zekâ çalışmalarının odağındaki bir kavram olarak "zekâ" kavramının, insanın zihinselliği bakımından odak olarak kabul edilmesindeki pratik kaygıları anlayabiliyor ancak bu pratik kaygıların insan zihinselliği bakımından üstesinden gelemeyeceği fenomenal unsurların ve zihnin diğer yetilerinin mevcut olduğunu burada ileri vurguluyoruz.

2.5. Zihin-Beden İlişkisi Bağlamında Yapay Zekâlı Varlıkların Sınırı

Yapay zekâlı varlığın sınırı nedir? Bu varlıkların sınırı yalnızca içerisinde bulundukları bilgisayarların sınırları mıdır yoksa bilgisayarlara bağlı olan tüm donanımlarla birlikte düşünüldüğünde ağa bağlı olan tüm cihazların oluşturduğu bir sınır mıdır? Yapay zekâ felsefesinde felsefi işlevselci savların eleştirisini ele aldığımız bu çalışmanın böyle bir tartışmayı kapsıyor ve bu tartışmalar bakımından kimi incelemeleri yerine getirebiliyor olması gerekmektedir. Çünkü zihni bir programa, beyinlerimizi de bir donanıma eşlemek konusunda aceleci olduğunu savunduğumuz felsefi işlevselci savların, zihin-beden ikiliği üzerinde bu ve benzeri kimi sorunlu alanlarla ilgili sorulara yanıt üretebilmesi zaruridir.

Turing, düşünen bir makinenin inşa edilmesi amacını gerçekleştirmenin bir yolunun, insanı bir bütün olarak ele alıp onun tüm uzuvlarını – kol, göz, kulak vb. – kameralar, mikrofonlar, hoparlörler veya tekerlekler gibi insan yapması teknolojik araçlarla ve bir çeşit elektronik beyinle değiştirme üzerinden gerçekleştirilebileceğini

ileri sürmektedir.[96] Peki ama insansı bir zihnin örnek alındığı ve başarı kriteri olarak böyle bir zihnin gerçekleştirilmesine ilişkin arzunun bulunduğu hesaba katıldığında, insan zihninin varlık koşulları Turing türü bir kurgunun doğasına uymakta mıdır? Böyle bir arzu, insan zihninin bedeni ile birlikte düşünüldüğünde varlık koşullarını kapsayabilir mi? Burada tartışmayı sıkı bir antropomorfik yükle ele almak istemiyoruz. Keza bu çalışma içerisinde de antropomorfik bakma tarzlarının sorunlarını sıkça dile getirdik. Dolayısıyla Frankfurt'un da vurguladığı gibi özne, kişi kavramımız olarak biz insanlar, bize özgü kavramları ele alırken kimi nitelikleri yalnızca bize özgü, türümüze özgü olarak kabul etmemeliyiz. İnsan olmayan varlıklarda da faillik, kişi olmaklık bakımından kimi niteliklerin bulunabileceğini kavramsal olarak düşünebiliriz. İnsan olmayan ama kişilik özelliği gösterebilecek varlıklar düşünülebilir. Ancak bazı özelliklerimiz ise yalnızca bize özgü olabilir.[97] Bu anlamda, çalışmamızın amaçları bakımından da bu amaçlardan birinin işlevselci savların başarılarını ve başarma arzusunu güttüğü düşünsel kazanımların imkânını yadsımak olmadığını bir kere daha vurgulamamız gerekir. O halde yapay zekâlı varlıkların zihin-beden ikiliği bakımından statülerini inceleyebilmek için hareket noktası olarak insanı seçmemiz ve insanla ilgili kimi hususları vurgulamamız gerektiği açıktır.

Biz insanlar bir zihne ve bir bedene sahip varlıklarız. Zihin ve bedenin ne oldukları sorusu, bedenin ne olduğu yönünden bilimler nezdinde pek tartışma götürmez durumda olsa da zihnin ne olduğu sorusu felsefe tarihi boyunca tartışmalı olmuştur. Beyin, omurilik ve sinirlerden oluşan sinir sistemimiz bedenimiz dediğimiz tüm uzuv ve organlarımız boyunca yayılır böylece zihnimiz bedenimiz aracılığı ile dışsallaşır ancak asıl soru zihin gibi maddi olmayan bir yapının bedenle olan ilişkisi nasıldır? Bu ilişkinin doğası nedir? İşte bu sorular "zihin-beden problemi" denilen bir problemi doğurur. Bu iki kavram arasındaki ilişkinin incelenmesine ve zihnin doğasına ilişkin sorular ise zihin felsefesi denilen alanın ortaya çıkmasını sağlar.[98] Zihin-beden problemine ilişkin zihin felsefesi içerisinde kimi yanıtlama girişimleri

[96] A. M. Turing, "Intelligent Machinery", **Mechanical Intelligence: Collected Works of A. M. Turing**, ed. D. C. Ince, C.I, North Holland, 1992, s. 117.
[97] Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person", **The Journal of Philosophy**, C. LXVIII, No: 1, 1971, s. 6.
[98] Sayan, "Analitik Zihin Felsefesinin Temel Problemlerine Bir Bakış", s. 38-39.

bulunur. Örneğin zihnin ne olduğu ile ilgili soruya zihnin beyin ne ise o olduğu yollu özdeşlik ilişkisi bakımından yanıtlar verilebilirken, kimi yanıtlar ise zihnin tıpkı bir gölge gibi olduğu, dolayısıyla zihnin beynin ürettiği sanal bir yapı olarak kabul edilirse salt beynin devinimlerine indirgenemeyeceğini ileri sürer. Bu yanıtlar dilde birçok farklı şekilde kullanımlara karşılık gelir: "Ben bir bedenim" mi deriz yoksa "Benim bir bedenim var" mı deriz? Ya da "Ben bir zihinim" mi deriz yoksa "Benim bir zihnim var" mı deriz?[99] Bir başka tartışmalı konu ise zihnin uzaydaki konumunun ne olduğudur. Zihnin konumu neresidir? Yanıt verilebilir ki beyin nerede ise zihnin konumu ise uzay-zaman olarak orasıdır. Bu kabul edilebilir bir yanıttır ancak o halde beyin devindiğinde zihin de devinir diyebilir miyiz? Zihnin bir ivmesinden ya da hızından söz edilebilir mi?[100]

Zihnin yalnızca kendisinden değil, kiplerinden de söz etmek mümkündür. Bu kipler yönelimsel ve nitel olarak iki grupta incelenebilir. Yönelimsel zihin durumları arzuları, beklentileri, düşünceleri vb. kapsar ve bu durumlar daima bir *şey* hakkındadır. Nitel zihin durumları ise bir şey hakkında olmayan durumlar olarak sadece zihnin duyumsamaya ilişkin hallerini kapsar. Acı duymak, sevinmek, bir rengi algılamak, bir kokuyu duyumsamak vb. gibi durumlar nitel durumlardır. Bu durumların duyumsanmış hallerine ise nitelce denir.[101] Nitel zihin durumları zihin felsefesi alanında "fenomenal bilinç" kavramının merkezinde yer alır. Fenomenal durumlar deneyimlenmiş durumlar olarak öznel psikolojik bir gerçekliğe aittirler. Bu gerçekliğin içerisinde bulunmak bakımından canlı, hissedebilen ve duyumsayabilen bir varlık olarak bu hissedilebilir ve duyumsanabilir olanların ayırdına varır. Böylece acı hissi nesnede olmasa bile o salt var olma tarzı bakımından şeyler, fenomenal olarak *bir şey gibi* – örneğin acı olarak –hissedilir. Fenomenal bilinç ise tüm bu öznel zihin durumlarımız olarak nitelceleri barındırmaktadır.[102] Zihin felsefesi tartışmalarında zihin-beden probleminin ele alınması bakımından kimi düşünce deneyleri öne sürülür. Bunlardan biri zombi problemidir. Zombi problemi, bir insandan ayırt edilemeyecek düzeyde dışsallaşabilen ancak herhangi bir fenomenal

[99] **A.e.**, s. 40.
[100] **A.e.**, s. 42.
[101] **A.e.**, s. 49.
[102] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 129.

yaşantıya sahip olmayan mantıksal olarak olası varlıklar üzerinden insan zihninin doğasına ilişkin bir inceleme girişimidir.[103] Zombi probleminin Sayan ve Aytekin'in ortaya attıkları "fenombi" kavramı üzerinden bir yorumuna göre, fenombiler – mantıksal olarak olası varlıklar olarak – fenomenal zihin durumlarına sahip olabilirler ama bu onların bilinç sahibi oldukları anlamına gelmez. Dolayısıyla zihne sahip olmak bilinç sahibi olmak için gerek koşullardan birisiyken tek başına yeter koşul olarak kabul edilemez.[104] Öte yandan Hurlbert ve Poggio'ya göre de fenomenal bir durumun ayırdına varmak duyumsamanın herhangi bir kipi – örneğin görmek – bir zekâdan daha fazlasını ifade etmektedir. Bir nesneyi duyumsamamızı sağlayan ışık demetlerinin retinamızdan içeri girmesi ve bir görsel imgeye dönüşmesi süreci zekânın göstergesi olarak kabul edilemez. Görme olayı – veya herhangi bir duyumsama olayı – biz insanlar ve fail niteliğine sahip olabilecek herhangi bir varlık bakımından da ışık ışınlarının çözülmesi olayından fazla bir durum olarak dünyaya ilişkin dilsel ifadelerin kavranması ile birlikte düşünülmelidir.[105] Peki tüm bu düşünceler yapay zekâ felsefesi ve yapay zekâlı varlıklar bakımından bize ne gibi bir temel sağlar? Yapay zekâlı varlıkların zeki olduklarını kabul edebilmemiz için önkoşullardan birisi onların zihin sahibi de olabildiklerini ortaya koyabilmektir. Bu anlamda Sayan, robotların asıl anlamda mı yoksa bir taklit sürecinin sonucu olarak görünürde mi zihin sahibi olduklarının tespit edilmesi gerekliliğini öne sürer, bu ayrımı yapabilmek için bu zihin durumlarının görünürde mi yoksa gerçek zihin durumları mı olduğunu nasıl anlayacağımızı sorar. Çünkü zihin felsefesinin temel sorularından biri olarak diğer canlılarda zihnin var olup olmadığının tespitine ilişkin genel soru, yapay zekâlı varlıklar için de geçerlidir. Gerekçe açıktır: Her birimiz zihin durumlarına sahip olduğumuzu biliriz ancak başka insanların zihin durumlarına sahip olduğunu nasıl tespit edebiliriz? Bu anlamda biz henüz başka insanların zihin durumlarının tıpkı bizlerin sahip olduğu zihin durumlarıyla özdeş olup olmadığını

[103] Erdinç Sayan ve Tevfik Aytekin, "Being Mindful of Zombies and Phenombies", **MetaZihin**, C. I, No: 2, 2018, s. 160.
[104] **A.e.**, s. 164.
[105] Anya Hurlbert ve Tomaso Poggio, "Making Machines (and Artificial Intelligence) See", **Dædalus**, C. CXVII, No: 1, 1988, s. 213.

ortaya koyamıyorken yapay zekâlı varlıkların zihin durumlarını yüksek bir kesinlikle teşhis edebilir miyiz?[106]

Yapay zekâlı varlıklarda zihnin teşhisi konusunun açığa kavuşturulabilmesi için biz insanlarda zihin-beden ikiliğinin ne anlama geldiği, bu problem içerisinde deneyimin merkezi, zihnin içselliği ve dışsallığı gibi birkaç hususun aydınlatılması gerektiğini düşünüyoruz. Gündelik düşünme ile çok açıktır ki bizler bedenlerimizi, zihnimizi barındıran bir ev olarak görmeye ve zihinlerimizi ise iki yuvarlak pencereden dış dünyayı izleyen bir gözlemci olarak kabul etmeye meyilliyiz. Hatta bu zihnin nasıl meydana gelebileceği, zihnin dışsallaşmasını sağlayan beden parçaları, duyu organları vb. arasındaki ilişki gibi birçok konuyu hesaba kattığımızda, bizim gibi başka insanların elleri, ayakları olduğunu duyumsar, bizim de elimiz ve ayağımız olduğunu, bunun diğer insanlardan ayrı olarak bize ait olduğunu kabul ederiz. Böylece bizim de kendimize ait bir zihnimizin olduğunu düşünürüz ve bedenimiz üzerinden zihnimizin bize ait olduğunu hissetmek isteriz.[107] Ancak zihin her ne kadar beden üzerinden anlaşılıyor olsa bile bedenden, maddi olandan bağımsız mıdır? Duyumsamaların zihin üzerindeki etkisi *sanal* mıdır? Canbeyli'ye göre beynimizi uyaran duyumsama verileri bağımsız olarak ele alınamaz. Zihnimizi uyaranların etkileri zaman ve mekândadır ve birbirleriyle etkileşmektedir. Hatta bu uyaranların bazıları zihnimizin de onun üzerinde gerçekleştiği nöronal yapı üzerinde kalıcı kimi etkilere sebep olabilmektedir.[108] Buradan anlıyoruz ki zihin ve beden arasında sıkı bir ilişkinin bulunduğunu, zihnin belirebilmesi ve nöronal herhangi bir sürecin *zihinleşebilmesi* için bedenin zaruri olduğunu savunan kimi görüşler söz konusu olabilir. Örneğin Dilman'a göre bizler birey olarak özdeş olduğumuz bir zihinden ve nedensel açıdan bağlı bulunduğumuz bir bedenden oluşmaktayız. Yaşamlarımız boyunca bu bedene bağlı olarak kalmak zorundayız. Dolayısıyla bedenimizin koşulları, biz insanların varlık tarzlarını ve

[106] Sayan, "Analitik Zihin Felsefesinin Temel Problemlerine Bir Bakış", s. 47.
[107] D. E. Harding, "Bir Kafaya Sahip Olmamak Hakkında", **Aklın Gözü: Benlik ve Ruh Üzerine Hayaller ve Düşünceler**, ed. Daniel C. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 32-39.
[108] Canbeyli, "Psikoloji Açısından Beyin-Bilgisayar Karşılaştırması", s. 123-124.

zihinsel olarak var olmasının koşullarını sıkı bir şekilde belirlemektedir.[109] Bedenleşme yaklaşımı olarak anılabilecek bu yaklaşımın temel düşüncesi, zekânın fiziksel bir bedende ve fiziksel bir çevrede yer aldığının savunulmasıdır. Zekânın anlaşılacağı tek yolun çevre ile etkileşen ve zekânın çevre ile etkileşmesini onu bedenleştirerek sağlayan bir beden olduğu düşünülür.[110] Yapay zekâ felsefesindeki önde gelen eleştirilerden biri Dreyfus'un bedenleşme üzerinden gerçekleştirdiği eleştirilerdir. Çünkü ona göre insan dünyası, insanların bedenleşen ihtiyaçlarının karşılandığı ve dünyanın buna göre organize edildiği bir alandır. Dolayısıyla bedene göre örgütlenmiş bir dünya, başka yollarla erişilebilir kılınamaz.[111]

Zihin-beden tartışmaları bakımından zihnin sınırı ve bedenin rolü düşünülürse bu tartışmalı hususların içselcilik* ve dışsalcılık** olmak üzere iki ana bakma tarzı bakımından ele alındığı görülebilir. İçselciliğe göre bilinç gibi zihinsel durumlarımız beynimizin içinde gerçekleşir. Dışsalcılığa göre ise zihinsel durumlarımız kısmen veya tamamen beynimizin dışında veya vücudumuzun dışında kimi unsurlara dayanarak gerçekleşir.[112] Manzotti ve Tagliasco'ya göre zihinsel süreçler dışsalcı bir yaklaşıma uygun olarak işlemektedir. Çünkü onlara göre zihin yalnızca kafatası*** içerisinde yerleşik bir şekilde dış dünyanın temsilini oluşturamaz. Radikal dışsalcı bir yaklaşımdan zihin, beyin ve bedenin sınırlarından bağımsız olarak zaman ve mekânda özdeş fiziksel işlemler üzerinde gerçekleşmektedir. Böylece gözlemci ve gözlenen varlık arasındaki ayrımın bulunmadığı, bunların beraberce deneyimi meydana getirdiği düşünülürse zihin, failin ayırdında olduğu her şeyle özdeş bir yapı olarak karşımıza çıkar. Bu durumda örneğin bir gökkuşağı bir zihin tarafından algılanmazsa var olamaz. Gökkuşağı varlığını hem gözlemci olarak bir zihinle hem de gözlenen şey olarak kendisiyle olanaklı kılabilir.[113] Andy Clark ve David

---

[109] Ilham Dilman, **Free Will: An Historical and Philosophical Introduction**, New York, Routledge, 2001, s. 113-114.
[110] Ben Goertzel, "Artificial General Intelligence: Concept, State of the Art, and Future Prospects", **Journal of Artificial General Intelligence**, C. V, No: 1, 2014, s. 13.
[111] Hubert L. Dreyfus, "Bilgisayarların Akıllı Olması İçin Neden Bedenleri Olması Gerekir?", çev. Berk Celayir ve Selenay Yakın, **Pasajlar**, No: 9, 2021, s. 61.
[*] İng. **Internalism**
[**] İng. **Externalism**
[112] Manzotti ve Tagliasco, "Artificial consciousness: A discipline between technological and theoretical obstacles", s. 112.
[***] İng. **Cranium**
[113] **A.e.**, s. 112-113.

Chalmers'ın aktif dışsalcılık[*] olarak adlandırdığı bir dışsalcı tavır daha söz konusudur. Bu tavra göre zihnin içsel ve dışsal durumları, zihnin kafatası içerisinde bulunan ve dış dünya ile kendisini ayıran biyolojik yapılara sahip olduğu içselci yaklaşım ve anlamın dışsallığı üzerinden zihnin yalnızca kafatası içerisine sıkıştırılamayacağı yollu dışsalcı yaklaşım gibi iki yaklaşıma indirgenemez. Aktif dışsalcılık olarak adlandırılan üçüncü bir tavra göre çevrenin bilişsel süreçler üzerindeki aktif rolü ön plana çıkarılır. Clark ve Chalmers, aktif dışsalcı tavrın farkını insanın problem çözmesine ilişkin verdikleri üç örnek üzerinden açık kılar: İlk durumda bir insan bir bilgisayar karşısına oturur ve ilgili soketlere iki boyutlu cisimler yerleştirmesi istenir. Bir insan bu şekilleri zihinsel olarak döndürür ve soketlere yerleştirir. İkinci durumda ise bir insan benzer bir bilgisayar karşısına oturtulur ve bu sefer bu insana ekranda gördüğü şekli isterse fiziksel olarak bir butona basarak döndürebileceğini isterse yine zihinsel olarak bu döndürme işlemini yapabileceği söylenir. Burada fiziksel olarak butona basmanın bir avantajı bulunacaktır. Son durumda ise **cyberpunk**[**] bir gelecekte bir insan yine benzer bir bilgisayar ekranının karşısına oturtulur. Failin bu sefer nöral implantlar vasıtasıyla bir önceki geometrik şekilleri bir bilgisayar kadar hızlı şekilde döndürebileceği bir durum söz konusudur. Clark ve Chalmers'a göre bu durumlar arasında herhangi bir bilişsel fark bulunmaz ancak bilişsel süreçlerin bazıları dışsal nesnelere ve süreçlere *delege edilirler*. Bu anlamda dış dünya bilişsel süreçlerin aktif bir parçasıdır ve bilişsel süreçler içselciliğin iddia ettiği gibi kafatasının içerisine yerleştirilemez.[114]

Clark ve Chalmers, aktif dışsalcı pozisyonlarını bir düşünce deneyi ile de pekiştirirler. Alzheimer hastası olan Otto isimli bir insan düşünelim. Otto, Alzheimer hastalığının bir sonucu olarak, her gün birçok şeyi unutmaktadır. Otto bu durumun önüne geçebilmek için unutmak istemediği bilgileri bir not defterine kaydetmektedir ve not defterine kaydettiği bilgileri mütemadiyen kontrol etmektedir. Böylece hastalığından ötürü unuttuğu bilgileri hatırlayabilmektedir. Clark ve Chalmers'a göre zihinlerimiz de bu yollu çalışır: Zihinsel ürünler olarak bilgilerimiz, dış dünyaya

[*] İng. **Active Externalism**

[**] Popüler olan bu kelimenin isabetli bir Türkçe karşılığını bulamadığımız için bu kelimeyi metin içerisinde kullanılageldiği şekilde kullanmayı tercih ettik.

[114] Andy Clark ve David Chalmers, "The extended mind", **Analysis**, C. LVIII, No: 1, 1998, s. 7-8.

yayılır ve dışsal nesnelerde yer alır, onlar aracılığı ile düşünülür ve kurulur. Dolayısıyla bilgi derimizin ve kafatasımızın ötesinde dışsal bir durumdadır.[115] Otto açısından incelenirse Otto'nun içsel süreçleri ile not defterinin iki ayrık alan ve sistem olduğu düşünülebilir ancak Clark ve Chalmers'a göre Otto'nun sinir sistemi ve not defteri bir bütündür, tek bir bilişsel sistemi temsil ederler. Tek bir bilişsel sistem olmasından ötürü de not defteri ve Otto'nun zihni arasındaki iletişim sistemin dışından herhangi bir şeyin bu sürece katılmasına gereksinmez. O halde Alzheimer Otto düşünce deneyinin önemi nedir? Deneyin en önemli unsuru olan not defteri, tıpkı bizlerin hafızası gibi, Otto'nun hayatında bir sabittir. Not defterindeki bilgiler Otto'ya bir uygunluk içerisinde sunulmaktadır. Otto, gereksindiği bilgileri not defterinden elde ederken bilgiler onun epistemolojik koşullarına uygun olarak hazır bulunmaktadır. Not defterindeki bilgiler tıpkı biz sağlıklı insanların yaptığı gibi hafızamızda daha sonradan hatırlanmaya uygun bir biçimsel formata getirilmiştir ve saklanmaktadır.[116] Radikal ve aktif dışsalcı yaklaşımların içselcilik ve dışsalcılık tartışması bakımından taşıdığı önem, zihnin yayılımı üzerine kimi belirlemeler sunuyor olmalarıdır. Ancak yapay zekâlı varlıklar hakkındaki bu incelemenin amaçları düşünülürse zihnin yayılımı kadar önemli olan hususlardan biri zihnin yeridir. Zihin uzamsal ve zamansal olarak nerededir? Zihnin yeri konusunda birçok düşünce deneyi bulunmakla beraber Dennett'ın *fıçıdaki beyin* düşünce deneyinin çalışmamızın amaçları bakımından önemli olduğunu düşünüyoruz.

Dennett'ın fıçıdaki beyin kurgusu şu şekildedir: Bir insanın beyni ameliyatla bedeninden çıkarılsın ve beyindeki her girdi/çıktı yoluna bir ucu beyinde bir ucu ise boş kafatasında bulunmak üzere bir radyo alıcısı yerleştirilsin. Beyin bir fıçıda hayatta tutulsun ancak radyo alıcısı ile beyin sahibi olmayan bir beden ile iletişim kurabilsin. Bu sayede bu beden tıpkı sağlıklı bir insan gibi yaşamına devam edebilsin.[117] Dennett'a göre "İşte buradayım!" gibi bir düşünce belirdiğinde, düşüncenin gerçekleştiği yerin fıçıdaki beyin olması gerekirken *burada olan şey olarak* da bedenden söz etmemiz gerekmektedir. Böylece Dennett'ın Yorick adını verdiği beyin ve Hamlet adını verdiği beden düşünüldüğünde, "Ben neredeyim?"

[115] **A.e.**, s. 12-13.
[116] **A.e.**, s. 16-17.
[117] Dennett, "Ben Neredeyim?", s. 210.

sorusu sorulduğunda burada "ben" ile karşılanan öznenin konumu hakkında net bir yanıt veremiyoruz.[118] Çünkü sinirsel ateşlemeler beyinde gerçekleşmektedir, *burada olan* şey ise bedendir. Oysa sinirsel ateşlemeler fıçıdaki beyinde gerçekleşmekte olup, fenomenal deneyim radyo alıcıları vasıtasıyla beyin bulunmayan kafatasında yaşanıyor gibi görünmektedir. Dolayısıyla "Ben neredeyim?" sorusunu soran benliğin uzam-zamansal konumu tartışmalıdır. Ancak tartışmaya kapalı olan durum sayılabilirlik bakımından bir olan bir zihin ve yine bir olan bir bedenin söz konusu olmasıdır. Dolayısıyla fıçıdaki beyin deneyi her ne kadar zihnin konumu ile ilgili sezgilerimizi sarsıyor olsa da zihin ve bedenin sayılabilirlik bakımından birliğini yıkmaz, aksine bunu onaylar. Bu anlamda yapay zekâ felsefesi açısından da tartışmak istediğimiz husus yapay zekâlı varlıkların zihin-beden problemi bağlamındaki konumunun tespit edilmesinin tartışmalı birçok durumu barındırdığıdır. Çünkü yapay zekâlı varlıklar insanlardan farklı olarak birçok aygıtı, bedeni ve parçayı kontrol edebilmekte böylece "Onların bedeni nedir?" gibi bir soruya kolayca yanıt verilememektedir. Bedenini tespit edemediğimiz bir varlığın fenomenal durumlara sahip olmasının koşulları ise ayrı bir tartışma konusudur.

Zihin-beden problemini çözümsüz ve tartışmalı bırakan durumlardan biri zihnin elle tutulamayan, görülemeyen, maddi olmayan bir yapı olmasıdır. Bu durum yapay zekâ felsefesi bakımından da söz konusudur: Yazılım katmanları, yani programlar, elle tutulamazlar.[119] Bir program olarak yapay zekâlar, kodlanabilirler ve bu anlamda Descartes türü bir ruhtan bahsedilmemektedir. Ancak programlar daima bir donanımla beraber var olmak durumunda olan biçimsel dizgelerdir. Ve yapay zekâlar, benzer makine diline göre inşa edilmiş her türlü donanımda kendilerini kopyalayabilme ve yine bu donanımların tamamını bedeni kılabilme olanağına sahiptirler çünkü programlar da en nihayetinde makine diline dayanmaktadır ve birer yazılımdır.[120] Bu durumda okuyucu sorabilir: İşte yapay zekâların bedeni bu gördüğümüz monitörler, hidrolik kollar ve kameralar değil mi? Yapay zekâların bedeni konusunda daha fazlasını arzu etmek için bir gerekçemiz var mı? Burada

[118] **A.e.**, s. 211-212.
[119] Bingöl, "Bilgisayar Üzerine", s. 30.
[120] Nick Bostrom, "Ethical Issues in Advanced Artificial Intelligence", (Çevrimiçi) https://nickbostrom.com/ethics/ai, 17 Aralık 2022.

okuyucu açısından önemli bir itirazı vurgulamak isteriz: Dreyfus, bedenlenmiş olmanın salt harekete ve çevre ile etkileşmeyi sağlayan manipülatörleri uygulamaya indirgenemeyeceğini savunmaktadır.[121] Ancak bu bedenlenme herhangi bir bedenlenme olarak düşünülmemelidir. Sloman'ın beden vurgusu Dreyfus'un vurguladığı bedeni bize tanımlar niteliktedir: Beden aynı anda iki ayrı yerde olamaz, gözler aynı anda iki zıt yöne bakamaz.[122] Böylece yalnızca "beden" kavramını karşılar gibi görünen maddi parçalara sahip olmanın bedenleşme için yeter koşulları sağladığını söyleyemiyoruz. Çünkü Sloman'ın da vurguladığı noktadan hareketle bedenin *sayılabilirlik bakımından birliği* büyük önem arz eder. Bu birlik ki bedene ve bu bedendeki zihne birlik, düşünme ve dünya tasarımı bakımından bir olanaklılık sağlar. Dreyfus'un bilgisayarların düşünemeyeceği yollu itirazının bilgisayarların bedene sahip olmamaları olduğunu[123] hesaba katarsak, bedenin birliği ve zihnin bir yerdeliği bakımından kimi problemlerin olduğunu söylemek mümkündür.

Yapay zekâlı varlıkların bedeni ve statüleri konusunda kimi kavramsal tanımlama girişimleri bulunur. Bu tanımlama girişimleri, Ersoy'un da haklı olarak sorduğu yapay zekâlı varlıkların o an bir makinenin içinde olup olmadığı ya da yapay zekâlı bir varlığın aynı anda birkaç yerde olup olamayacağına[124] ilişkin kimi sorulara yanıt vermek bakımından çabaların bir sonucudur. Yapay zekâ çalışmalarının önde gelen isimlerinden biri olan McCarthy, bedenli yapay zekâlar olarak robotları fiziksel kabiliyetleri olan – örneğin hareket etme, objeleri manipüle edebilme ve algılayabilme – ve bir bilgisayar programı tarafından kontrol edilen bir kişi olarak tanımlar. Robotlar, bir varlık olarak var oluşları süreklilik arz etmek durumunda olan varlıklardır.[125] Ancak biz yapay zekâları yalnızca robot formunda görmemekteyiz. Yapay zekâlar **bot** formunda da karşımıza çıkmaktadır. Yapay zekâların robotlar gibi fiziki bedenlere sahip olmayan ama yine de kimi işlevleri yerine getiren yazılım formuna **bot** denmektedir. Bu anlamda Köken'in kullanımına göre yapay zekâ soyut

[121] Dreyfus, "Bilgisayarların Akıllı Olması İçin Neden Bedenleri Olması Gerekir?", s. 48.
[122] Aaron Sloman, **The Computer Revolution in Philosophy: Philosophy Science and Models of Mind**, Sussex, Harvester Press, 1978, s. 167.
[123] Doğan, "Güçlü Yapay Zekâ Mümkün müdür?", s. 90.
[124] Ersoy, **Robotlar, Yapay Zekâ ve Hukuk**, s. 138.
[125] McCarthy, "Concepts of Logical AI".

veya somut olabilen, robot ise somut olmak durumunda olan varlıktır.[*] Bir yazılım olarak yapay zekâ, robotlarda kullanıldığında somutlaşır.[126] Ersoy ise robotları fiziksel ve zihinsel olarak fail nitelikleri gösteren ancak biyolojik olarak canlılık özelliğine sahip olmayan varlıklar olarak tanımlamaktadır. Ersoy'un bu tanımının literatürdeki kullanımlar karşısında güçlü olduğunu savunuyoruz.

Robot ve **bot** arasında çizilen robotların somut varlıklar oldukları ve yapay zekâların bir program olarak soyut ve bedensiz varlıklar oldukları, dolayısıyla yapay zekâların dünyada eyleyemeyecek, çevre ile etkileşemeyecek varlıklar oldukları yollu hatalı ayrım hukukçularımız arasında yaygındır. Taşdemir'e göre robot olabilmenin koşulu fiziksel dünyada bir varlık olarak bulunmak ve dünyada kimi etkiler yerine getirebilmektir. Yapay zekâlar ise salt kendileri olarak sadece elektronik ortamda var olabilecek ve eyleyebilecek varlıklar olabilirler.[127] Bedenin ne olduğu yollu soruya yanıt vermediğimiz ve bu soruyu incelemediğimiz bir durumda yapay zekâlı varlıkların bedeninin program düzeyinde kalsalar bile mevcut olmadığını iddia etmenin önünde bir engel bulunmaz. Burada iki durum üzerinden düşünmek istiyoruz: İlk durumda Bennu, sahip olduğu arazisini evinde kurmuş olduğu veri merkezinde çalışmakta olan yapay zekâya bağlı 100 adet robot bedeni ile işliyor ve tarım ürünleri yetiştiriyor olsun. Bu ürünlerin ekilmesi, hasat edilmesi, iklim ve toprağın durumu gibi birçok hususta planlayan ve hesaplayan zekâ bu veri merkezinde çalışmakta olan yapay zekâ olsun. İkinci durumda ise veri merkezindeki yapay zekâ kalsın ancak 100 adet robot bulunmasın. Her eylem için Bennu bu yapay

[*] Sorulabilir: Herhangi bir bilgisayar programının bedensiz olduğu, herhangi bir donanımda barındırılmadan da erişilebilir olduğu ve işleyebildiği görülmüş müdür? Böyle bir durumun olanağı var mıdır? Bugün bize en soyut görünen bulut teknolojileri bile bizim göremediğimiz veri merkezlerindeki sabit diskler, işlemciler ve grafik birimleri tarafından saklanmakta ve çalıştırılmaktadır. Ancak bizler son kullanıcılar olarak bu programların donanımdan bağımsız çalıştıklarını düşünüyor olabiliriz. O halde ne tür bir program olursa olsun bir bilgisayar programı mutlaka bir hafızada saklanmalı, bir işlem birimi tarafından işlenmelidir ve bunlar fiziksel varlıklar olmak durumundadır. Biz bir yapay zekâyı antropomorfik bir tarzda bir bedenin içinde görmüyoruz diye yapay zekâların tıpkı Descartes'ın maddi olmayan tözü gibi maddeden bağımsız şekilde var olduğunu düşünemeyiz. Yine bu yapay zekâlar donanıma sahip programlar olarak ama bizim alışageldiğimiz beden algısının dışında bir bedende vardırlar. O halde yapay zekâların bedeni konusunda soyut veya somut gibi ayrımları yapmak için bir temelimiz bulunmuyor. Yapay zekâların bedensiz var olabileceğinin ileri sürüldüğü bir başka çalışma için Bkz.: Ersoy, **Robotlar, Yapay Zekâ ve Hukuk**, s. 7.

[126] Enes Köken, "Yapay Zekânın Cezai Sorumluluğu", **Türkiye Adalet Akademisi Dergisi**, No: 47, 2021, s. 257-259.

[127] Taşdemir v.d., "Robotların Hukuki ve Cezai Sorumluluğu Üzerine Bir Deneme", s. 799.

zekâya danışsın ve kendi bedeni ile bu yapay zekânın komutlarını yerine getirsin. Ortaya çıkan iş bakımından belki 100 robotun çalışma gücünden oldukça aşağı bir çalışma gücüyle muazzam bir zaman farkı oluşsun ancak yine benzer hasat elde edilsin. Buradaki fark nedir? İlk durumda hesaplayan ve planlayan bir zekâ olarak yapay zekânın iradesi, metalik bedenlere dayanıyordu ve onlar aracılığıyla kendine dayanarak gerçekleşiyordu. İkinci durumda ise bu yapay zekânın iradesi Bennu'nun bedenine dayanarak gerçekleşmiştir. Bennu, robotların vazgeçememesine rağmen bu iradeyi işletmekten vazgeçme olanağına sahiptir. Ancak sonuç bakımından düşünülürse, yapay zekâ bir program olarak bir veri merkezinde kaldığı halde, ki burada hem Köken, Ersoy ve Taşdemir'in iddialarının tam aksi olarak düşünülürse, yapay zekânın iradesi bir programken bile bir beden tarafından gerçekleştirilmedi mi? Böyle bir program olsalar bile yapay zekâlar, onların ortaya koyduğu[*] ve bir planlama ve hesaplama sürecinin sonucu olan kimi yargıları gerçekleştirmek bakımından bizler tarafından da bedenleşmiyorlar mı? Bu durumun süreklilik mi arz etmesi gerekiyor yapay zekâların somutlaştıklarını söylemek için? Bu noktada Sloman'ın öne sürdüğü görüşler daha da önem kazanır ve tartışmayı bir üst noktaya taşır: Ona göre radyo sinyalleriyle iletişim kuran bir merkezi sistem olarak bir yapay zekâya bağlı birden çok robotun olduğunu düşündüğümüzde, bu sistem bu birden çok robotun kabaca nerede olduğunu, çevresinde neler olup bittiğini veya tam olarak ne yaptığının farkında olabilecek bir sistem olarak düşünülebilir. Hatta bu sistemdeki robotlar, ilgili yapay zekânın bilincinde olmadığı alt süreçleri temsil etmek ve bilgileri saklamak konusunda işlevlere sahip olabilirler. Dolayısıyla bu robotların tamamının herhangi bir andaki karmaşık ayrıntılarıyla değil ama kimi görevleri ve bilgileri ile ilgilenen ancak istediği an tüm robotların o anda gördüğü ve hissettiği şeyleri hissedebilen değil ama herhangi bir anda yalnızca bir robotun tüm duyularına vakıf olabilen bir yapay zekâ düşünülebilir. Fakat bu durumda tek tek robotların ayırdında olduğu ancak sistemin kendisinin ayırdında olmadığı şeylerden söz etmek mümkündür.[128] Sloman'ın bu son vurgusu yapay zekâlı bir varlık ile onun bedeni

[*] Okuyucu açısından belirtmeliyiz ki burada yapay zekâlar üzerine kurduğumuz yargılarda "O" zamirini kullanmak zorunluluğum dilimizin yapısından kaynaklanmaktadır, onları fail saydığımız için değil. Bu sebeple bu cümle yapısını bir zorunluluk sonucu kullandığımız unutulmamalıdır.

[128] Sloman, **The Computer Revolution in Philosophy: Philosophy Science and Models of Mind**, s. 171.

olarak düşünebileceğimiz erimleri arasındaki uyuşmazlığı ve yapay zekâlı bir varlığın gerçekliğimizle olan ilişkisindeki problemleri gözler önüne sermektedir. Nitekim Putnam gibi işlevselciler, bu gibi çelişkili ve problemli durumlara rağmen robotun içsel durumları olduğunu kabul ederler, bu içsel durumların ayırdına varma tarzlarının her ne kadar insanlardaki gibi olmadığını kabul etseler bile işlevleri bakımından birebir oldukları için örneğin insanın duyumsamaya sahip olduğu durumu düşünürsek robotun da mantıksal ve semantik karşılık olarak bir duyumsamaya sahip olduğunun ayırdında olduğunu ileri sürerler.[129]

Yapay zekâlı varlıkların zihin-beden problemi bakımından taşıdığı ve şimdiye kadar üzerinde durduğumuz sorunlar söz konusuyken, onların kişilik statülerinin bulunduğunu, özgür iradeye sahip olduklarını vs. söyleyebilmemiz için gerekçelerimiz nelerdir? Biz muazzam donanımsal güce dayanarak ortaya konacak bir zekânın kişi olabilmesini, bir fail olabilmesini yadsımıyoruz ancak şimdiye kadar üzerinde durduğumuz problemlerin aşılması gerektiğini de vurguluyoruz. Çünkü biz insanlar için dünyayı etkilemek, seçimler yapabilmek ve eyleyebilmek olağandır, bizim var oluşumuza ve onun koşullarına içkindir. Dolayısıyla biz bu *bize göreliği* deneyimleriz ki bu durum farklı türlü de eyleyebileceğimize ilişkin en derin sezgilerden biridir ve özgür iradeli varlıklar olduğumuzu kabul ederiz. Dolayısıyla bu *bize görelik* – çoğumuzun insan olmak olarak vurguladığı sezgi olarak – yaşantımızın ve dünyada eyleyen bir varlık olma koşulumuzun temelinde yer almaktadır.[130] Oysa işlevselci girişimlerin, yapay zekâlı varlıkların insanlarla zihinsel anlamda eşit statüde varlıklar olduğu yollu iddiası şekillenecekse bu iddianın faillik ve ahlakilik üzerinden kimi eleştirilerle yüzleşmesi gerektiği açıktır. Çünkü insanla benzer bir zihnin failliğe aday olması gerektiğini düşünmek çok doğalken böyle bir zihnin ahlaki özne olmayacağını söylemek çelişkidir. Bu anlamda ahlaki failin, kendilik, ahlak vb. gibi kavramlara ve bu kavramların temsil ettiği unsurlara sahip olması gerektiğini söylememiz gerekir. Aksi halde kendiliğe sahip olmayan bir varlığın fail olarak kabul edilmesi olanaklı olmadığı için ahlaki eylemlerinden de sorumlu olması

[129] Hilary Putnam, **Mind, Language and Reality: Philosophical Papers**, C. II, Cambridge, Cambridge University Press, 1997, s. 391.
[130] Robert Kane, **A Contemporary Introduction to Free Will**, New York, Oxford University Press, 2005, s. 6.

söz konusu olamaz.[131] Bu çalışmanın yapay zekâ felsefesi bakımından hedeflediği tartışma açısından vurgulamak istediği ve önem arz ettiği düşünülen yargılarından biri tam da budur: Yapay zekâ felsefesi ve yapay zekâ çalışmaları içerisinde çözümlenmeye çalışılan etik ve politik sorunlar, her şeyden önce bir ontolojik ve epistemolojik sorun olarak, zihin felsefesinin merkezindedir ve *bu temellerden hareketle* çözüme kavuşturulabilir. Dolayısıyla günümüzde bu etik ve politik sorunların çözümsüz kalmış olmasının temel sebebini yöntem ve hareket noktasına ilişkin hatalar olarak kabul edebiliriz.

Yapay zekâlı varlıkların ahlaki özne olabilmeleri, onların kendilik düşüncesine sahip olmalarını gerektirir[132] çünkü kendilik – Nichols'un ifadesiyle faillik – düşüncesine sahip olmayan bir yapay zekâ, eyleminin özgür olduğunu hissedemez, iradesi özgür olamaz. Dolayısıyla özgür irade sahibi olmayan bir yapay zekânın, Zambak'ın da vurguladığı gibi bilinç veya düşünme yetisine sahip bir fail olarak, bir varlık olması olanaklı değildir.[133] Peki yapay zekâlı varlıkların özgür irade sahibi olabilmelerinin koşulu nedir? Doğan'a göre bir yapay zekâlı varlığın özgür iradeye sahip olabilmesinin koşulu onların bir özne olarak kabul edilebilmeleri ve deneyimleyen bir özne olarak yapay zekâların olanağının incelenmesidir.[134] Özgür iradeden söz edilebilmesi için öznenin, kararlarının ayırdında olması ve istemesi üzerinde kontrol sahibi olması gerekir. İsteme, özgür iradenin temel kavramlarından biri olarak bir varlığın bir nesne ya da duruma yönelmesine ilişkin arzudur. İsteme ediminin herhangi bir yönelmeden farkı, bir iradeye dayanması ve usumuzun burada rol oynamasıdır. Dolayısıyla istemeye sahip olabilmek, karar vermeye yetili olabilmektir ve kişi eylem esnasında kendi usuna dayanarak bir seçim yapmaktadır.[135] Kendilik sahibi insanlar ile kendilik sahibi olmayan yapay zekâlar arasındaki ayrım, kişi olmaklık bakımından isteme üzerinden belirlenebilir.

---

[131] Joel Parthemore ve Blay Whitby, "What Makes Any Agent a Moral Agent?: Reflections on Machine Consciousness and Moral Agency", **International Journal of Machine Consciousness**, C. V, No: 2, 2013, s. 106-108.
[132] Shaun Nichols, "Experimental Philosophy and the Problem of Free Will", **Science**, No: 331, 2011, s. 1401.
[133] Zambak, "Free Will and Artificial Intelligence", s. 167.
[134] Doğan, "Yapay Zekâ ve Özgür İrade: Yapay Özgür İradenin İmkânı", s. 809.
[135] Thomas Pink, **Free Will: A Very Short Introduction**, New York, Oxford University Press, 2004, s. 26-28.

Frankfurt'un istemeye ilişkin ayrımı insanlar, hayvanlar ve isteme sahibi olabilecek türden herhangi bir varlık arasında bir ayrım yapmamıza olanak sağlamaktadır.

Frankfurt'a göre kişilerle diğer varlıklar arasındaki ayrım, bu varlıkların istemelerinin yapısında bulunabilir. Örneğin insanlar yalnızca arzuları ve sezgileri olan varlıklar değildir. Ancak birçok hayvan yalnızca temel düzeyde arzulara ve sezgilere sahiptir. Frankfurt arzular arasında birinci ve ikinci dereceden arzular olmak üzere iki kademeli bir ayrım yapar. Hayvanlar, bir şeyi yapma veya yapmama anlamında temel bir düzey olarak birinci dereceden arzulara sahiptirler. Biz insanlar ise ikinci dereceden arzulara sahip varlıklar olarak kendi birinci dereceden arzularımızı nesne edinebiliyor ve onlar üzerine düşünebiliyoruz. Dolayısıyla farklı türden birinci dereceden arzulara sahip varlıklar olarak biz insanlar bir durum karşısında çeşitli tercihlerde bulunma imkânına sahibiz.[136] Frankfurt, ikinci dereceden arzulara sahip olmak yanında ikinci dereceden istemlere sahip olmayı da ön plana çıkaracak, böylece ikinci dereceden istemlere sahip olabilmeyi kişi olmaya özgü bir olanak olarak kabul edecektir. Ona göre ikinci dereceden isteme sahip olmayıp, birinci veya ikinci dereceden arzulara sahip olabilecek varlıklar ancak mantıksal olarak olanaklı olabilir. Frankfurt'un **wanton**[*] adını verdiği bu mantıksal olarak olası varlıklar bakımından iradenin bir önemi yoktur. İradenin bir önemi olmadığı için **wanton**'ları bir şey yapmaya yönelten şey herhangi bir arzu tarafından harekete geçirilmek veya başka bir arzu tarafından harekete geçirilmeye ilişkin bir seçimin doğru olduğu önemsiz bir şekilde bir arzu tarafından davranışın belirlenmesidir. **Wanton** olarak adlandırılabilecek varlıkları tüm insan olmayan hayvanlar ve çok genç çocuklar olarak tanımlayan Frankfurt, en rasyonel **wanton**'un diğer rasyonel faillerden olan farkının, bir **wanton**'un arzularının arzu edilebilirliği ile ilgilenmiyor olması, dolayısıyla kendi arzusunun ne olması gerektiğiyle ilgili soruyu görmezden gelmesi olduğunu ifade eder.[137] Frankfurt'un **wanton** kavramı bize yapay zekâlar hakkında önemli bir durumu gösterir: Yapay zekâlar kendilik

[136] Frankfurt, "Freedom of the Will and the Concept of a Person" s. 6-7.

[*] Frankfurt'un kendi adlandırması olarak **wanton** kelimesi, İngilizce karşılıkları bakımından "istemek" ve "otomat" kelimelerini çağrıştıran bir adlandırmadır. Böylece sürekli arzulayan ama bunu kendi üzerinde kontrolü kaybetmişçesine veya bir kontrol unsuru olmadan gerçekleştiren varlıkları nitelemek için gayet uygun bir adlandırma gibi görünüyor.

[137] **A.e.**, s. 10-11.

sahibi olmayan ve matematiksel ontolojiye gömülü varlıklar olarak ikinci dereceden istemeye sahip varlıklar olamazlar. Dolayısıyla arzularının içeriği ve bu içeriğe yönelme gerekliliği ile ilgili bir akıl yürütme yapamazlar. Çünkü arzularının içeriği ve bu içeriğe yönelmeye ilişkin kararlar, kendi istemelerinin nesnesi değildir. Böylece yapay zekâlı varlıkların ahlaki özne olmalarına ilişkin bir hususun kendilik koşulunu sağlamak olduğunu görüyoruz.

Kendiliğin edinilmesi sürecinde ilk adım olarak ben düşüncesinin sağlanması gerektiği savunulur. Deveci, bir canlının bilinç sahibi olmasının önkoşulunu, canlının kendisini bir *ben* olarak ayırt edebilmesine dayandırır. Bir ben olarak ayrışabilmek için gereken niteliğin ise birinci şahıs bakış açısı ve üçüncü şahıs bakış açısı arasındaki farkın sezilebilmesi olduğunu vurgular. Böylece bir canlı için ben olma hissinin ne olduğu sorusu, canlının *kendini kendi olmayan şeylerden ayırt edebildiği* durumlarda olanaklı hale gelmektedir.[138] Makineler hususunda ise – Nagel'a atıf yaparak – makine olmak gibi bir şeyin anlamlı kılınabilmesi için makinenin sosyal bir varlık olması, dil yoluyla etkileşmesi ve etkileştiği varlıklarla benzer zihinsel yapılara sahip olması gerektiği öne sürülür.[139] Peki benlik nasıl inşa edilir? Benlik, Hume türü bir yaklaşım tercih edilirse, algı demetlerinin meydana getirdiği birlik olarak düşünülebilir. Benlik, algılanan şeylerin bir arada düşünülmesi, onların birliğinin hesaba katılması yoluyla elde edilebilecek bir ide olarak kabul edilebilir. Ancak Hume'un yaklaşımı, ben idesinin tüm deneyimlerden bağımsız şekilde fenomenal bilinçte devamlı olarak deneyimlenmesini açıklayamaz.[140] Burada sorulabilir: O halde insanlara *kim*liklerini veren nedir? Onların herhangi bir varlıktan kendilerini ayırmalarını sağlayan benlik, kendilik gibi ideleri elde ettiği süreçleri nasıl açıklayabiliriz? Bu ideler nedir? İnsanların kim olduklarını, benliklerini vb. sahip oldukları materyalin organize olmasına indirgediğimizde insanın atom düzeyindeki yapısı, sürekli bir değişim içerisindedir. O halde insanların bu gibi

[138] Deveci v.d., "Makinenin Beni ve Ötekisi", s. 108-109.
[139] **A.e.**, s. 104.
[140] Murat Arıcı, "Zihinsel Alanda "Ben Hissi" Neye Karşılık Gelir?", **Dört Öge**, No: 7, 2015, s. 66-67.

idelerini onların sahip oldukları, bedenlerini oluşturan biyolojik materyallere indirgeyemiyor, onlar üzerinden tanımlayamıyoruz.[141]

Makine zekâsının fail niteliğine sahip olduğu ve ahlaki özne olmak için yeter koşullara sahip olduğu savunulacaksa "kendilik" kavramının hesabının verilmesi gerektiği açıktır. Lewis'e göre kendilik problemi üç soru üzerinden tartışılmalıdır: Kendilik nedir? Kendiliği nasıl ölçeriz? Kendilik hangi işlevleri yerine getirir? [142] Öncelikle biz insanlar kendimizi tanımak konusunda hiç zorluk çekmeyiz: Nerede olduğumuzu, neden burada olduğumuzu vs. biliriz. Bu anlamda kendimizi evrimsel olarak adapte olmuş ve karmaşık süreçlerin işlediği biyolojik makineler olarak düşünebiliriz: Eyleriz, hissederiz, düşünürüz, planlarız ve öğreniriz. Bu makinenin – biz insanların – bir özelliği de ben idesidir. Ancak ben idesi ile kendilik arasında epistemolojik bir ayrım bulunur: Bilmek ile bildiğimi bilmek ideleri arasında bir ayrım vardır. Bu ayrımdan ötürü ben idesi, kendilik idesinin bir özelliğidir. Çünkü kendilik tıpkı bildiğimizi biliyor olduğumuz gibi ben idesinin ayırdında olan bir yapıdır. Kendilik ve ben ideleri arasındaki hiyerarşi ilişkisi kendiliğin iki özelliğini oluşturur: Kendi-öteki ayrımı, kendiliğin zaman ve mekân boyunca korunumu.[143] Varlığın çevresi ile etkileşime girebilmesi için kendisi ve kendisi olmayan diğer varlıklar arasında bir ayrım yapması gerekmektedir. Bu süreçler içerisinde ben idesine ihtiyaç duyulmaz. Ancak varlığın karşılaştığı problemleri kurabilmesi için kendiliğin bir özelliği olan ben idesinin bulunması gerekmektedir. Ben idesi problemleri kurarken, kendilik idesi bu problemlerin çözülmesinde rol oynamaktadır. Bir varlık, ben idesine sahip olmadan önce kendisi üzerine düşünme yetisine sahip olamaz. Ben idesinin belirebilmesi için ise varlığın dil ve öz referanslı davranışlara sahip olabilmesi gerekmektedir.[144] Strawson'a göre ise biz insanlar, kendimizi ve kendi durumlarımızı, kendimiz olmayan ve kendimize ait olmayan durumlardan ayırmaktayız. Kendimizi eylemler, yönelimler, duyumsamalar, düşünceler, hisler, algılar ve bellek üzerinden tanımlarız. Kendiliğimizin – Lewis'in de vurguladığı

[141] Penrose, **Kralın Yeni Usu**, s. 27.
[142] Michael Lewis, "Myself and me", **Self-awareness in Animals and Humans: Developmental Perspectives**, ed. Sue Taylor Parker ve Robert W. Mitchell, New York, Cambridge University Press, 2006, s. 20.
[143] **A.e.**, s. 20-22.
[144] **A.e.**, s. 23-27.

üzere – bir konumu ve yönelimi olduğunu da biliriz. Ancak Lewis'den farklı olarak Strawson'a göre kendiliğe yalnızca zamansallık değil aynı zamanda boy, renk, şekil, kilo vb. gibi kalıcı fiziksel özellikler de atfederiz. Yani başkalarına atfetmediğimiz ancak fiziksel olarak maddi bedenlerimize atfettiğimiz şeyler de kendilik açısından söz konusudur. Kendilik gibi bir ideye fiziksel özellikler atamak ilk aşamada anlaşılmaz gelebilir ancak Strawson, öznenin ancak ve ancak sahip olduğu bedende gerçekleşen bir fenomenal durumu hissedebildiğini vurgular. Bu anlamda öznenin bedeni, aynı zamanda öznenin algısal deneyimi ile özel bir ilişki içerisindedir ki bu da yer kaplayan bir bedenle olanaklıdır.[145] Kendiliğin gerek koşullarından biri olarak kendi-öteki ayrımının yapılabilmesi için deneyimlerimizin onunla beraber uzamsal bir ilişki içerisinde olduğu bir bedenin bulunması zorunludur. Özne deneyimimiz olarak ben idesine sahip oluyor olmamız veya herhangi bir deneyimi bize ait olarak düşünmemizin koşulu diğer öznelerin tam da beden yoluyla ayırt edilebiliyor olmasıdır. Beden, tüm özel deneyimlerimizin başkalarına ait değil ama kendimize, bize ait olmasını sağlar. Dolayısıyla bedenden ayrık bir şey olarak saf kendilik kavramı, Strawson'a göre mevcut değildir.[146] Ayer'in beden vurgusu da Strawson'un beden vurgusu ile örtüşür. Ona göre "kişi" kavramının özü, herhangi bir zamandaki belli bir bedene yapılan referanstır. Burada belli bir bedene sahip olduğumuz ön kabulü temelde bulunmaktadır. Buradan hareketle deneyimlerimizin bizim deneyimlerimiz olduğunu kabul eder ve bu deneyimlerin başka bir bedende gerçekleşmediğini düşünürüz. Ancak bu bedene yapılan referanslar tüm zamanlar bakımından söz konusu olan bir referans mıdır? Bu referans sürekli midir?[147] Bilindiği üzere Descartes, insanı zihin ve beden olarak iki ayrı tözden meydana gelen bir varlık olarak tasarlamıştır. Hobbes ise Descartes'ın bu ayrımına karşıt bir biçimde, zihinlerin varlığını da maddeye indirgemiştir. Böylece zihnin tüm durumları ve süreçleri maddenin yani bedenin kendisine indirgenmiş olur. Her ne kadar Hobbes türü bir fizikalist yaklaşım güçlü duruyor olsa da – bu çalışma içerisinde de vurgulanan fizikalizme ilişkin sorunlar anımsanırsa – fizikalist yaklaşımın zihne ilişkin açıklama girişimlerinde kimi sorunların olduğunu vurguluyor olmak, Ayer'e

[145] P. F. Strawson, **Individuals: An Essay in Descriptive Metaphysics**, New York, Routledge, 1996, s. 89-92.
[146] **A.e.**, s. 102.
[147] A. J. Ayer, **The Concept of a Person and Other Essays**, London, Macmillan Press, 1973, s. 126.

göre Descartes'ın düalizmini kabul etmek anlamına da gelmez. Çünkü zihinsel olanın maddi var olanların nedensel ilişkilerinden ve onların fiziksel gerçekleşmelerinden bağımsız olabildiği kimi durumlar olabilir. Ancak bu bizi yine de bir zihinsel tözün bulunduğunu düşünmeye sevk etmez.[148] Hume'un benliğe ilişkin açıklaması anımsanırsa Ayer, Hume'un belleğe dayanan ve deneyimlerin bir ben idesine ait olduğunu savunan argümantasyonunun döngüsel olduğunu belirtir. Çünkü böyle bir açıklama girişiminin sonucu olarak ilk deneyimin bizim deneyimimiz olarak işaretlenebilmesinin koşulu, bir ilk deneyimin mevcut olmasıdır. Ancak ilk deneyimin bir "ben" idesine ihtiyaç duyacağı da açıktır. Bu anlamda Hume'un yaklaşımı, benlik tartışmasını aydınlatamaz.[149]

Kendilik ve benlik üzerine şimdiye kadar vurguladığımız hususlardan sonra yapay zekâ felsefesi ve yapay zekâ çalışmaları bakımından yapay zekâlı varlıkların faillik statüleri hakkında neler söyleyebiliriz? Yapay zekâlı varlıkları fail olarak atfedebilmek için yeterli gerekçemiz bulunur mu? Her şeyden önce yapay zekâlı varlığın kişi ve ahlaki bir özne olabilmesinin koşulunun bir kendiliğe ve benliğe sahip olmak olduğu açıktır. Kendiliğin ve benliğin, kısacası yapay zekâlı varlığın işaret edilebilmesi, uzay ve zamanda konumunun gösterilebilmesi gerekmektedir. Şimdiye kadar Lewis, Strawson ve Ayer'in de ortak olarak vurguladığı husus, kendilik sahibi bir varlığın bedenli olması, bu bedenin uzay ve zamanda süreklilik gösteren bir nitelikte bulunması zorunluluğudur. Bu açıdan bakılırsa Doğan'ın, insan bilinci gibi bilişsel fenomenlerin yapay bir sistemde oluşturulabileceğine ilişkin iddiası geçersizdir. Çünkü Doğan, her ne kadar bu iddiasını çevresi ile etkileşebilecek ve deneyim elde edebilecek bir robot üzerinden kurguluyor olsa da ilgili yapay zekânın o robota özgü, o robotta bedenlenmiş ve süreklilik gösteren bir varlık olması gerektiğine ilişkin eleştirilere bir yanıt sunmamaktadır.[150] Çünkü Ayer, kişi olmanın karakteristik koşullarının, fiziksel özelliklere sahip olmak yoluyla belirli bir süre boyunca belirli bir uzamsal pozisyonda devamlılığa sahip olmak ve çeşitli biçimlerde farkındalığa sahip olmak olduğunu vurgulamaktadır.[151] Dolayısıyla biz

[148] **A.e.**, s. 82-84.
[149] **A.e.**, s. 112-114.
[150] Doğan, "Yapay Zekâ Felsefesinde Bilinç Problemi", s. 185.
[151] Ayer, **The Concept of a Person and Other Essays**, s. 82.

bir yapay zekânın bir beden ile olan ilişkisinde bedenin birliğini ve uzay zamandaki sürekliliğini tesis edemedikçe Doğan'ın başka nesnelerle ve çevresiyle etkileşebilecek yapay zekâlı varlıkların sosyal bilince sahip olabileceğine ilişkin iddiasını kabul edemeyiz.[152] Gerekçelerimiz açıktır: Ayer, bir bedende birden çok zihnin olup olamayacağını, bir zihnin ise birden çok bedende bulunup bulunamayacağını sorar. Eğer bir zihin birden çok bedende bulunabilecekse bunun eş zamanda mı yoksa farklı zamanlarda mı gerçekleşeceği sorusunu da ekler. Zihnin bedenler arasında geçişken bir yapıda olabileceğini kabul etsek bile zihnin hangi bedende var olmak isteyeceğini, beden değişikliğini ne şekilde isteyebileceğini belirleyebilir miyiz?[153] Yapay zekâlı varlıkların zihin-beden problemi bakımından taşıdıkları ve en temelde yer alan sorun tam olarak Ayer'in burada dile getirdiğimiz türden sorularıyla özetlenebilir: Yapay zekâlı bir varlığın, herhangi bir t_1 anında, sayılabilirlik bakımından *bir* ile ifade edilebilecek bir bedende bulunduğunu nasıl iddia edebiliriz? Bir robot bedeninin birden çok yapay zekâya ev sahipliği yapamayacağını söylemek için bir gerekçemiz bulunuyor mu? Bir yapay zekâya bağlı ve dünyanın farklı kıtalarında bulunan yüzlerce robotun bulunabileceği ihtimalini düşündüğümüzde ilgili yapay zekânın herhangi bir anda *belirli bir bedende bulunduğunu* söyleyebilir miyiz? Bu ve benzeri soruları bir örnek üzerinden düşünelim ve felsefi işlevselciliğin zihin-beden problemi bakımından yapay zekâlı varlıkların bedeni ve zihni üzerindeki aşması gereken problemlerin neler olabileceğini ortaya koyalım.

Y isimli bir teknoloji şirketinin bulunduğunu varsayalım. Bu şirketin, kendi veri merkezlerinde bulunan bir yapay zekâ tarafından yönetilecek ve satın alanların ev işlerini yapmak konusunda onlara yardımcı olacak robotlar tasarladıklarını ve sattıklarını düşünelim. Bennu Y firmasından bir robot satın alsın ve robotu evinde aktifleştirsin böylece burada düşünce deneyimiz iki senaryoya ayrılsın: İlk durumda Bennu robotu çalıştırsın ve robot internet erişimi kazandığı andan itibaren veri merkezindeki yapay zekâ tarafından yönetilmeye başlansın. Binlerce robotun tek bir yapay zekâ tarafından internet bağlantısı üzerinden yönetildiği bir senaryodur bu. Diğer senaryomuz ise Bennu'nun, robotunu çalıştırıp internete bağladığı bir durumda

[152] Doğan, "Yapay Zekâ Felsefesinde Bilinç Problemi", s. 185.
[153] Ayer, **The Concept of a Person and Other Essays**, s. 84.

veri merkezindeki yapay zekânın internete erişiminin ya da robotların veri merkezindeki yapay zekâya erişimlerinin bilinmeyen bir sebeple sonlandığı bir senaryo olsun. Burada iki sorun vardır: İlk senaryomuz açısından akla gelebilir ki tüm bu robotlar veri merkezindeki yapay zekânın *kendisi* miydi? Yoksa bu robotların her biri veri merkezindeki zekânın yalnızca hesaplama gibi beynin nöral kaba kuvvetinden besleniyor, böylece bu robotlar, veri merkezindeki yapay zekâdan farklı olarak *kendi fenomenal unsurlarına sahip*, bu anlamda *kendilik sahibi varlıklar* olarak mı işliyorlardı? Eğer böyle ise kendi fenomenal unsurlarına sahip olan robotların kendi zekâlarına da sahip olmaları gerekmiyor mu? Kendi zekâlarına sahip olan robotlar ile bu veri merkezindeki yapay zekâ arasındaki ilişki, robotların fenomenal unsura sahip oldukları ilk anda bu iki varlığın *iki ayrı varlık olmalarını* gerektirmiyor mu? Dolayısıyla ilk senaryoda her bir robotun, çalıştırıldıkları anda fail statüsü kazanan ve veri merkezinde çalıştırılmakta olan yapay zekâdan ayrışması gereken özneler olması gerektiğini düşünemez miyiz? Eğer böyle değil ise ve tüm robotlar veri merkezindeki yapay zekâya bağlı varlıklar olarak bulunuyorlarsa ve veri merkezindeki yapay zekâ, sayılabilirlik bakımından *bir* ise, o halde bu veri merkezindeki yapay zekâ, kontrol ettiği her bir robotun duyu verileri vasıtasıyla elde edeceği fenomenal unsurları yine sayılabilirlik bakımından *bir* deneyim olarak mı deneyimler – böylece her bir robotun fenomenal deneyimi veri merkezindeki yapay zekânın t_1 zamanında tek bir fenomenal deneyimi olabileceği için birbiri üzerine çöker ve *birbiriyle tutarsız bir duyumsama yığışımı* olarak duyumsanır – yoksa bu deneyimler arasında geçiş yapabilir ve t_1 anında yalnızca *bir robotun* fenomenal deneyimlerini mi deneyimleyebilir? Böyle bir durumda veri merkezindeki bir yapay zekâya bağlı ve t_1 anında kontrol edilebilen, sayılabilirlik bakımından yalnızca bir robotun olması zorunlu değil mi? Aksi halde veri merkezindeki *bir* yapay zekânın yönettiği ve dolayısıyla ancak ve ancak *bir adet kendiliğe* sahip olabilecek binlerce robottan söz etmiş olmuyor muyuz? Dolayısıyla sayıca *bir benliğin* binlerce bedene yayılabileceğini söylemek için en ufak bir gerekçemizin bulunmadığını bu aşamada vurgulamamız mümkündür.

İkinci senaryomuzun sonuçları ise daha açıktır: Felsefi işlevselcilerin iddia ettiği gibi zihinsel işlevlerin yerine getirildiği bir varlık olarak veri merkezindeki

yapay zekânın kendilik ve dolayısıyla bir benlik sahibi olduğunu kabul edelim ve yönettiği robotlarla olan bağlantısının herhangi bir zamanda kesildiği ikinci senaryomuzu yeniden ele alalım. Bağlantılar kesildiği anda, yapay zekânın *kendiliği* robotlar tarafından da paylaşılır mı? Böylece ilk durumda veri merkezinde bulunan ve *bir kendiliği* olan yapay zekâya bağlı binlerce robotun olduğu düşünülürse, bağlantının kesildiği durumda bu kendilik – çünkü kendiliğin *bir bedende* bulunma ve uzay-zaman bakımından bu bedenin süreklilik göstermesi gerektiği farklı düşünürler tarafından da bu çalışmada vurgulanmıştır – içkin olduğu bu bedenlerin tamamınca sahiplenilir ve her robot bu andan itibaren kendi fenomenal deneyimlerinin bu tek yapay zekâdan ayrılması vesilesiyle kendi *kendiliklerini* mi zenginleştirmeye başlarlar? Böylece kendilik sahibi olan ve birbirinden ayrı, fail niteliği kazanmış olan binlerce robottan bahsetmemiz mümkün müdür? Eğer böyle değilse ve bağlantı koptuğunda veri merkezindeki yapay zekânın kendiliği yine kendisi tarafından sahip olunan bir şey olarak kalacaksa, o halde kendiliğin bedenli olmayla ilişkili sıkı bağı düşünülürse, veri merkezindeki yapay zekânın bedeni olarak kabul edilebilecek bu robotlar bağlantının kopması vesilesiyle bu yapay zekâdan ayrıldığında yapay zekânın kendiliğinin statüsü ne olur? Onun kendiliğinde bir azalma, eksilme vb. gerçekleşir mi? Yoksa tıpkı biz insanların kol veya bacak gibi uzuvlarımızı kaybettiğimizdeki gibi bir durum mu söz konusu olur? Bu ve benzeri sorular yapay zekâlı bir varlığın insan düzeyinde bir zihinsel yaşantıya sahip olabileceği, böylece dünyada düşünen, fenomenal deneyimlere sahip olan bir varlık konumunda bulunabileceği iddiası ileri sürülürse karşılaşılacak senaryolardan doğmaktadır. Bir yapay zekânın *basitçe işleyen biçimsel bir sistem* olduğu kabul edilirse pek tabii burada kurmuş olduğumuz düşünce deneyi bir anlam ifade etmez.

Burada ortaya koyduğumuz düşünce deneyinin içerdiği ve her iki senaryo bakımından sorulabilecek birçok soru bulunduğu görülebilir. Bu soruların varlığı robotların fail olabilmeleri bakımından kaçınılmaz bir şekilde sahip olmaları gereken kendilik, benlik gibi unsurları mevcut robot ve **bot** tanımlarımızla kapsayamadığımızı, onlara ilişkin olası problemler bakımından zengin bir kavramsal temel sunamadığımızı gösterir. Dolayısıyla zihin-beden probleminde yola çıkan ve felsefi işlevselci bir tavırla bu problemi yapay zekâ lehine çözülebilir olarak

yanıtlayan girişimlerin, burada şimdiye kadar dile getirdiğimiz problemleri aşmaları gerektiğini vurguluyoruz. Ancak açıktır ki yapay zekâ felsefesi içerisinde insansı nitelikleri, zihinsel fenomenleri vb. bilgisayar-beyin analojisi üzerinden kurmak ve bu yolla işlevler üzerinden insansı olanın makinelerde gerçekleştirilebileceğini savunmak için yanıtlanması gereken birçok sorunun bulunduğunu da kabul etmek gerekir.

ÜÇÜNCÜ BÖLÜM
FELSEFİ İŞLEVSELCİLİĞİN ELEŞTİRİSİ

3.1. İşlevselcilik, Felsefi İşlevselcilik ve Yapay Zekâ İlişkisi

Çalışmamızda şimdiye kadar aldığımız mesafeyi okuyucu açısından açık kılmak isteriz. Çalışmanın metafelsefi incelemelerle başlayan ve "yapay zekâ" kavramı üzerine kavramsal bir incelemenin ele alınmasıyla devam eden, şimdi ise bir felsefi yaklaşımın ele alınmasıyla devam edecek olan bu sıralamaya dayanarak inşa edilmesinde bir amaç vardır: Yapay zekâ çalışmalarını felsefece ele alabilmek ve yapay zekâ felsefesi bakımından bir inceleme ve eleştiri ortaya koyabilmek için her şeyden önce bu alanlarda hakim olan *felsefece-olmayan* akıl yürütmelerin ve kimi *metafizik kabullerin* tahlilini sunmalı, onları okuyucu açısından da açık kılmalıyız. Ancak ve ancak bu yolla, bu alanlarda çalışanların çoğu ve geniş halk kitleleri tarafından arzulanan **mitosları** eleyebilir, yapay zekâ felsefesini değerli ve saygın bir tartışma alanı olarak sunabiliriz. Felsefeciler arasında bile sıkça, yapay zekâ felsefesinin kuruntular ve kurgular alanı olduğu, bu sebepten ötürü de bu alana felsefece bir dikkatin layık görülmemesi yönünde görüşlere rastlamak mümkündür. Ancak yapay zekâ felsefesi tartışmalarının metafelsefi analizlerini sunduktan sonra yapmamız gerekenin, bu felsefi alt-alanın doğabilimleri ve kimi bilimler bakımından kökenlerini ortaya koymak, medeniyetimizin olanaklarını sunan bu bilimlerin ciddiye alınan ve alınması pek doğal olan sorularının bir sonucu olarak yapay zekâ felsefesi alanının mümkün olduğu düşüncesini serimlemek olduğunu biliyoruz. Böylece matematik, mantık ve kimi doğa bilimleri bakımından yapay zekâ felsefesinin ve yapay zekâ çalışmalarını meydana getiren temel soruları aktarmak amacıyla tarihsel çerçeve ve bu çerçeve içerisindeki tartışmaları okuyucuya aktarmak istiyoruz. Ancak okuyucu açısından da yapay zekâ felsefesi tartışmalarının bu bilimlere gidilmediği vakit anlamlı kılınamayacağını, dolayısıyla yapay zekâ felsefesine atılan bir ilk bakışın oldukça anlamsız gelmesi ile bu alana ve onun sorularına gerekli değerin verilmemesi yönündeki eğilimin bu kavramsal arka plan düşüncesini kavramamış olmaktan ileri geldiğini belirtmeliyiz. Böylece metafelsefi incelemeler altında ele aldığımız bu bilimleri yalnızca temel olmak bakımından önemleriyle değil, aynı

zamanda bu temel olmak bakımından önemlerinin *felsefe yapma tarzı bakımından hatalı kullanımlarının* yapay zekâ felsefesi içerisindeki izdüşümlerini de ele alarak gerçekleştirmiş olduk. Böyle bir ele alma, yapay zekâ felsefesi ve yapay zekâ kavramı bakımından yüklerin yeniden-incelemesini oldukça kolaylaştırmış ve kavramın felsefi işlevselci ele alınmaları tarafından görmezden gelinen problemlerinin serimlenmesinde oldukça yol gösterici ve aydınlatıcı olmuştur. Çalışmamızın ikinci bölümünde "yapay zekâ" kavramının incelemesi hedeflenmektedir. Kavramsal incelemenin felsefi işlevselciliği eleştirdiğimiz bu bölümden önce konumlandırılmasının kimi gerekçeleri bulunur: İlkin, kavramın analizinin, bir felsefi tavrı eleştirmek bakımından önde geldiğini düşünüyoruz. Görülecektir ki felsefi işlevselciliğin eleştirisinin gerçekleştirilebilmesi için yükleri bakımından ele alındığında "yapay zekâ" kavramının bir ön incelemesini ve onun tarihsel çerçevesini sunmak zorunluluğunu hissetmekteyiz. Bu anlamda ikinci bölümde "yapay zekâ" kavramını, üzerindeki felsefi tavırlardan ayrık olarak çok yönlü bir biçimde ele almak ve birinci bölümde sunduğumuz yöntem ve bilimler düzeyindeki temeller bakımından yeniden değerlendirmek amacını gerçekleştirdik. Birinci ve ikinci bölümde ortaya koyduğumuz bu düşünce planı, tam da bu bölümde nihayetlenecek olan eleştirimizin temellerini sağlamıştır: Bir tavır olarak felsefi işlevselciliğin yöntemsel ve kavramsal dayanakları açık kılınmıştır.

İşlevselciliğin incelemesini sunmadan önce zihin-beden tartışmaları bakımından işlevselciliğe temel sağlayan felsefi tavırları ele almanın, işlevselciliğin kazandırdığı yeniliği anlamak bakımından önemli olduğunu düşünüyoruz. İnsanın zihinselliği ve bu zihinselliğin bedenle olan ilişkisi felsefe tarihi içerisinde ve psikoloji gibi diğer bilimler bakımından farklı yaklaşımlarla açıklanmaya çalışılmıştır. Zihin durumlarının beyindeki fiziksel-biyolojik süreçlerle özdeş olduğunu düşünen indirgemeci bir tavır, zihnin beyindeki elektrokimyasal süreçlerden ibaret olduğu şeklinde bir açıklama sunmak isterken elemeci bir fizikalizm ise bilinç, özgür irade, nitelceler vb. gibi zihin durumlarının halk psikolojisine dayanan kavramlar olduğunu ve bilimsel gelişmeler neticesinde bir zaman sonra eleneceklerini öne sürmektedir. Ancak işlevselciliğe yakın bir tavır söz konusudur: Felsefi davranışçılık, zihin-beyin problemini zihin durumlarından ziyade

davranışların analizi üzerinden açıklamak istemektedir. Ancak işlevselciliğin felsefi davranışçılıktan kimi farkları bulunur: İşlevselcilik beynin durumlarına karşılık gelen zihin durumlarının *işlevlerinden ötürü* zihinsel durumlar olduklarını, bu durumların fiziksel-biyolojik devinimlerle bir ilintilerinin bulunmadığını savunur. Dolayısıyla zihinsel durumlar fiziksel koşulları bakımından değil, hangi işlevlerin yerine getirilmesinde rol oynadıkları yönünden ele alınırlar.[1] Ancak davranışçılıkta ne zihinsel durumlar dikkate alınır ne de fiziksel koşullar dikkate alınmayacak kadar önemsizdir. Davranışçı yaklaşıma göre bir canlıya uygun girdi ürünleri – örneğin duyu verileri veya maddi kimi unsurlar vs. – sağlanırsa canlının benzer davranış örüntülerini sergileyeceği kabul edilir. Örneğin insanlar uygun davranışsal girdiler ve temeller sağlanırsa bir avukat veya doktor olarak yetiştirilebilirler.[2] Zihin durumlarının kendi başına ne oldukları, davranışçılar için önem arz etmez. Ancak işlevselcilik Block'a göre iki kabul bakımından davranışçılıktan ayrılır: Davranışçılar zihin durumlarını uyarılma ve bu uyarılmaya ilişkin yanıt üzerinden ele alırken zihinsel durumların kendilerinin bu yanıtların nedeni olduğunu ve uyarılmaların bir etkisi olduklarını dikkate almamaktadırlar. Diğer husus ise işlevselciler, davranışçılar gibi belli girdilerle belli çıktılar arasında lineer bir ilişkiyi düşünmezler. İşlevselciler kimi zihinsel durumlarla başka zihinsel durumlar arasındaki bağlantıları da hesaba katmaktadırlar.[3] Peki ama bir tavır olarak işlevselcilik nedir? Temelleri nelerdir?

İnsan, zihin ve bedeniyle bir bütündür. Bu bütünü ele almak bilimler bakımından ve felsefe açısından onu oluşturan yapıların analizi ile açık kılınmaya çalışılmaktadır. Nitekim nasıl ki insan bedeninde solunum, dolaşım ve sindirim gibi birçok süreci yerine getiren sistemlerin bulunduğunu biliyorsak, insan zihninde de kimi süreçler olarak işlevlerin yerine getirilmesini sağlayan yapıların olduğunu ileri sürebiliyoruz. Örneğin Gamez'e göre bir zihinsel fenomen olarak bilinç, insan beynindeki birbiri ile ilişkili çok sayıdaki unsurun işlevi olarak düşünülebilir.[4] İnsan beynine ve insanın zihinselliğine ilişkin böyle bir yaklaşım, insan bilinci gibi diğer zihinsel fenomenlerin de işlevsel olarak tanımlanabileceği ve işlevsel olarak

[1] James A. Reggia, "The rise of machine consciousness: Studying consciousness with computational models", **Neural Networks**, No: 44, 2013, s. 114.
[2] Canbeyli, "Psikoloji Açısından Beyin-Bilgisayar Karşılaştırması", s. 119.
[3] Block, What is Functionalism?", s. 33.
[4] Gamez, "The Potential for Consciousness of Artificial Systems", s. 218.

tanımlanmış olan bu zihinsel fenomenlerin maddi olarak gerçekleşmeleri ne olursa olsun herhangi bir gerçekleşmede belirebileceği düşüncesine imkân sağlamaktadır.[5] İşlevselci görüşün gücü buradan gelmektedir: İşlevselciler herhangi bir sistemin biyolojik materyalden ya da inorganik – örneğin silikon mimari – bir materyalden yapılmasının önemli olmadığını öne sürerler ve önemli olanın materyal yapının ne olduğu değil bu materyal yapının yerine getirdiği işlevler olduğunu iddia ederler. Örneğin şarap tadan bir insanın beyni, silikon bir mimaride yeniden gerçekleştirilebilirse şaraptan alınan zevk, zorunlu bir sonuç olarak bu silikon mimari üzerinde yeniden üretilebilecektir.[6] İşlevselciler zihnin doğasının materyal olandan ayrık olarak zihinde çeşitli bölgeler arasında gerçekleşmekte olan karmaşık nedensel ilişkilerde bulunduğunu öne sürerler. Zihinsel bir durum ise beynin temsil ettiği bilgi işleme sisteminin bir işlevi olarak karşımıza çıkar. Bu işlev, sistemdeki girdi ve çıktı ilişkileri üzerinden tanımlanır.[7] Dolayısıyla işlevselci yaklaşıma göre bir bilgi işleme sistemi olarak mekanik veya hidrolik bir sistem bile zihin sahibi olabilecektir. Çünkü kaçınılmaz olan şey beyinlerimizdeki gibi nöronal yapılara sahip olmak değildir.[8] Herhangi bir bilgi işleme yeteneğine sahip yapı, zihin sahibi olmak için uygun koşulları taşımaktadır. İşlevselcilik yaklaşımında Putnam önde gelen düşünürlerden biridir. Putnam, ağrı hissi gibi kimi fenomenal durumların beyinde gerçekleşen fiziksel-kimyasal bir durum olmadığını, bu durumların sistemin tümü bakımından işlevsel kimi rolleri üstlendiğini ileri sürmektedir.[9] Dolayısıyla ona göre zihinsel süreçler olarak problem çözme ya da düşünme, bir insanın veya bir makinenin zihninin işlevsel organizasyonu olarak kabul edilebilir ve zihinsel ya da mantıksal durumların aşamaları olarak bu durumların fiziksel gerçekleşmelerinin doğasına referans vermeden açıklanabilmektedir.[10] Nitekim nöral ağlar ve bağlantıcı sistemler, Barrow gibi yapay zekâ araştırmacılarının öne sürdüğü gibi, biyolojik olana ait olan içsel temsilleri ve işlevsel organizasyonları meydana getirebilecek olanağı sunuyor

[5] Zeman, **Bilinç Kullanım Kılavuzu**, s. 397-398.
[6] Dennett, **Bilinç Açıklanıyor**, s. 43.
[7] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 83-84.
[8] Arıcı ve Doğan, "Prof. Dr. Erdinç SAYAN ile Söyleşi: Zihnin Doğası Üzerine", s. 244.
[9] Hilary Putnam, "Psychological Predicates", **Art, Mind and Religion**, C. I, London, University of Pittsburgh Press, 1967.
[10] Putnam, **Mind, Language and Reality: Philosophical Papers**, s. 373.

olabilirler.[11] Bu anlamda Putnam'ın işlevselciliğini genel olarak beş temel düşüncede ortaya koyabiliriz: İnsanlar ve diğer organizmalar sınırlı varlıklardır. Sınırlı varlıklar olarak onların duyumsama ve eylemleri Turing makinelerindeki girdi ve çıktı ilişkileri üzerinden ifade edilebilirler. Burada fiziksel yapılar değil işlevsel durumlar rol oynar böylece insanın soyut zihin durumlarına odaklanılır. Öte yandan verili bir mantıksal ifade olarak zihin durumları sonsuz sayıdaki farklı fiziksel yapı tarafından gerçekleştirilebilir. Bir insan yapması varlığın – örneğin bir robotun – mantıksal ve yapısal durumu, bir insanın psikolojik ve fizyolojik özellikleri ile örtüşebilir.[12] Ancak işlevselciliği Putnam ile başlatmak konusunda emin olamayız. Boden'e göre işlevselci yaklaşımın kökeni Turing'in ve McCulloch ve Pitts'in çalışmalarına götürülebilir.[13] Boden'in bu yorumu üzerinden *işlevselciliğin bir çatı kavram olarak ele alınması gerekliliğini ve farklı işlevselciliklerin söz konusu olabileceğini* düşünebiliriz.

Ned Block, zihin felsefesi ve psikoloji alanlarında etkili olan üç tür işlevselciliğin bulunduğunu öne sürer: İşlevsel analiz[*], bilgisayımsal-temsil işlevselciliği[**] ve metafiziksel işlevselcilik[***]. İşlevsel analiz, bir sistemin kapasitesini ve özelliklerini anlamak için onu parçalarına ayıran ve parçalar ile bu parçalar arası ilişkilere eğilen bir yaklaşımdır. Örneğin bir fabrikanın bir arabayı üretmesi sürecini anlamak için üretim bantlarını, işçileri ya da makineleri ve bunlar arasındaki organizasyonel süreçleri incelemek işlevsel analizin uygulanmasıdır. Bilgisayımsal-temsil işlevselciliğinin temelinde ise psikolojik açıklamaların zihnin yapısına göre tasarlanmış bilgisayar programları vasıtasıyla sağlanabileceğine ilişkin bir görüş bulunur. Bu anlamda zihinsel süreçlerin işlevsel analizleri yoluyla çözümlenen süreçler, dijital bilgisayarların bilgisayımsal/mekanik operasyonları ile oluşturulmak istenir. Psikolojik durumlar düşüncenin dili ile sistematik olarak dünyayı temsil etmeye ve psikolojik süreçler ise bilgisayarların dahil olduğu temsil

[11] Harry Barrow, "Connectionism and Neural Networks", **Artificial Intelligence**, ed. Margaret A. Boden, 2. bs., San Diego, Academic Press, 1996, s. 135.
[12] Margaret A. Boden, **Minds as Machine: A History of Cognitive Science**, C. II, Oxford, Clarendon Press, 2006, s. 1359.
[13] **A.e.**, s. 1357-1358.
[*] İng. **Functional Analysis**
[**] İng. **Computational-Representational Functionalism**
[***] İng. **Metaphysical Functionalism**

süreçlerine indirgenmiş durumdadır. Metafiziksel işlevselciliğin temelinde ise psikolojik açıklamalardan ziyade zihnin doğasının ne olduğu sorusu yer almaktadır. Bu anlamda önemli olan zihin durumları ve davranış arasındaki ilişkinin nasıl gerçekleştiği değil bu zihin durumlarının ne olduğudur. Bir metafiziksel işlevselci zihin durumları nedir sorusuna bu zihin durumlarının işlevsel durumlar olduğu yanıtını verecektir. Metafiziksel işlevselcilik bazen özdeşlik kuramının savları ile ya da davranışçılık ve fizikalizmin savları ile ifade edilir. Ancak metafiziksel işlevselcilikte zihin durumlarının statüleri önemlidir. Bir ağrı, herhangi bir tekil ağrının ne olduğu yollu değil ama bir genel kavram olarak ağrı şeklinde ele alınır. Böylece ağrı olmak bakımından ağrı olarak nitelenen tüm zihin durumlarında ortak olanın ne olduğu sorusu üzerinden bir işlevsel inceleme gerçekleştirilmeye çalışılır. Şimdiye kadar andığımız üç tür işlevselciliğin kesiştiği bir nokta, bu işlevselciliklerin Turing makinelerine dayanan yaklaşımlar olmalarıdır. Metafizik işlevselciler zihin durumlarını Turing makinelerindeki tablo durumları olarak tanımlarlar. Bilgisayımsal-temsil işlevselciliğine göre psikolojik açıklamalar zihin için bir bilgisayar programının sağlanması yöntemine dayanırken, zihin durumlarının işlevsel analizi ise süreçlerin mekanik süreçlere indirgenmesidir ki eğer bu mekanik süreçler algoritmik ise o halde bu süreçler Turing makinesinde hesaplanabilir bir niteliğe kavuşmuş olarak kabul edilebilirler.[14] Block'un sunduğu bu kavramsal çerçevenin, bu çalışmanın adlandırılmasında ve yapay zekâ felsefesinde belirli bir tutumun eleştirilmesi amacıyla işlevselciliğin "felsefi işlevselcilik" olarak kullanımı bakımından bir temel sunduğu görülebilir. Okuyucu sorabilir: Yalnızca "işlevselcilik" kelimesinin kullanılmamasının sebebi nedir? Tek başına bu kelime, eleştirinin odağını temsil edebilmek bakımından yetersiz midir? Bu sorulara ilişkin yanıt, her şeyden önce işlevselcilik ve felsefi işlevselcilik kullanımlarının eleştirimiz bakımından izdüşümlerini serimlememizi zorunlu kılmaktadır.

İşlevselcilik, Block'un ortaya koyduğu ve şimdiye kadar farklı düşünür ve bilim insanları tarafından işlevselciliğin ele alınışlarının bu çalışmada dile getirilen farklılıkları bakımından da yalnızca felsefe içerisinde bir tutum olarak ele alınamaz. Farklı bilgi alanlarında da söz konusu olabilen işlevselcilik, görülebileceği üzere

[14] Block, "What is Functionalism?", s. 27-29.

felsefi bir tutum olması bakımından farklı bilgi ve bilim alanlarına dayanarak da farklı işlevselcilikleri kendi içerisinde barındırabilmektedir. Dolayısıyla bu çalışmada "felsefi" sıfatını kullanmak zorunluluğumuz buradan gelmektedir: İşlevselciliğin eleştirisi olarak genel bir adlandırmanın sonucu, işlevselcilik adlandırması altındaki tüm işlevselci tavırların bir eleştirisini sunmayı arzuladığımızı ortaya koyabilir. Ancak eleştirimizin odağında böyle bir arzu bulunmamaktadır. Ayrıca bu çalışma içerisinde sıkça dile getirdiğimiz ve burada da dile getirmek zorunluluğunu hissettiğimiz bir husus bulunur: Bu eleştiri, işlevselciliğe karşı değildir. Güçlü bir tutum olarak işlevselciliğin diğer tüm bilgi alanları açısından değeri yadsınmamaktadır. Ancak "felsefi işlevselcilik" olarak adlandırmanın bir gerekçesi olarak bu eleştiri, felsefe içerisinde yer alan *kimi işlevselciliklere karşı* yazılmıştır. Dolayısıyla okuyucunun, "felsefi" sıfatını kullanmadığımız bir eleştirinin işlevselciliğe karşı bir reddiye yazma amacıyla ortaya konduğunu düşünmesini istemiyoruz. Nitekim bu çalışma felsefi bir metin olarak, bir reddiye olmak kaygısını gütmemektedir. Ancak bu çalışma, Block'un kavramsal ayrımına dayanarak, felsefe içerisinde özellikle yapay zekâ felsefesine dayanak sağlayan ve *bilgisayımsal temellere dayanan işlevselciliklerin tümünü* niteler ve bu işlevselcilikleri hedef alır. Fakat bu hedef alma, sadece bu işlevselciliklerin yöntem ve başarılar olarak bir eleştirisini sunmaktadır, işlevselciliği bir tutum olarak yadsımamaktadır. Bu eleştiride şimdiye kadar ortaya koyduğumuz, işaret ettiğimiz hususlar işlevselciliğin başarı olarak lanse ettiği hususların üzerinde sıkıca durulmamış ve eleştiriye oldukça açık hususlar olduğunu göstermekten öte değildir. Böylece "felsefi işlevselcilik" adlandırması, yapay zekâ felsefesi içerisinde başta bilgisayar bilimi ve diğer ilintili bilim alanlarıyla da birlikte yapay zekâlı varlıkların, bilgisayımsal temellere dayanarak insansı failler olacakları, dünya tasarımlarının insanlardan farklı olmayacağı ve onlarında haklar ve sorumluluklar bakımından insanlarla eşit olmaları gereken veya olacakları bir noktanın bulunacağı türünden iddiaları kapsayan genel bir adlandırmadır. Böyle bir adlandırmanın kullanılması bir amacı daha taşımaktadır: Yapay zekâ felsefesi bakımından bazı işlevselcilikler terk edilebilir ve farklı işlevselcilikler savunulabilir. Örneğin Block'un bilgisayımsal-temsil işlevselciliği

olarak adlandırdığı işlevselcilik savunulmayabilir – nitekim örneğin Özel, eleştirisini[15] *klasik fonksiyonalizme ilişkin bir eleştiri* olarak sunsa da biz Block'un ayrımına binaen Özel'in eleştirisinin bilgisayımsal-temsil işlevselciliğine ilişkin bir itiraz olduğunu niteleyebiliriz – ancak işlevsel analizci bir yaklaşım sürdürülebilir. Böylece işlevselciliğin eleştirisi olarak sunacağımız bir eleştiri bakımından eleştirimizin odağı anlaşılamaz.

İşlevselcilik ve yapay zekâ çalışmaları arasındaki bağlantının okuyucu açısından daha açık kılınması gerekmektedir. İnsanın zihinsel fenomenleri olarak örneğin bilinç ve bilinçle ilişkili yapıları düşündüğümüzde bu yapılar bir makinede temsil edilebilirse, söz konusu makinenin bu zihin durumlarına sahip olabileceğine ilişkin inanca güçlü bir temel bulunmuş olabilir. Bunun koşulu, materyal olarak nöral ağların herhangi bir materyalden inşası ve bilişsel yapıların ise işlevsel yapılar olduğuna ilişkin düşüncedir.[16] Turing, insanların gerçekleştirebildiği eylemleri ve zihinsel süreçleri bir hesaplama süreci olarak ileri sürmüştür. Böylece insanların zihin süreçlerini hesaplama yapabilen makineler tarafından gerçekleştirmenin olanağının bulunduğunu düşünmüştür. Turing'in bu düşüncesi, Çağıl'ın da ifade ettiği gibi bir tür işlevselcilik olarak anılabilir. Çünkü Turing, zihinsel süreçlerin makinelerde gerçekleştirilebileceğini savunduğu bu yaklaşımıyla örtük olarak bu zihinsel süreçlerin organik bir beyinde gerçekleşmesinin zorunluluğunun bulunmadığını da savunmuş olmaktadır.[17] Nitekim McCulloch ve Pitts'in amaçları da önermeler olarak ifade edilebilen her durumun nöral ağlar üzerinde gerçekleştirilebileceğini göstermeye ilişkindir ve bu nöral ağlar, Turing makineleri tarafından işlenebilecek ve her hesaplanabilir fonksiyon bir nöral ağ tarafından hesaplanmış ve işlenmiş olacaktır. McCulloch ve Pitts'in görüşleri psikoloji ve fizyoloji arasındaki ilişkinin gücüne dayanmaktadır ve Boden'a göre bu durum, yapay zekâ felsefesinde işlevselciliğin de temelini hazırlayacaktır.[18] Nitekim işlevselciliğin varsayımlarından birine göre de zihin beyne bağlı değildir. Zihin, beynin işlevlerinin bir sonucudur. İşlevlerin bir sonucu olarak zihin ise bilgisayarda

[15] Özel, "Church-Turing Tezi ve Algoritmik Zihin Argümanı: Zihinsel Fenomenleri Matematiğe İndirgemek Mümkün müdür?".
[16] Gamez, "The Potential for Consciousness of Artificial Systems", s. 214.
[17] Çağıl, "Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekânın Olasılığı", s. 15.
[18] Boden, "Artificial Intelligence".

programlanabilen bir yapı olarak düşünülebilir. Yapay zekâ çalışmalarının temelinde de bu anlayış yer alır.[19] Güçlü yapay zekâ yaklaşımı tam da bu anlayış üzerinde şekillenir. Çünkü güçlü yapay zekâ düşüncesi bakımından zihnin hangi materyal yapıda gerçekleşeceği önemsizdir.[20] Dolayısıyla güçlü yapay zekâ yaklaşımını savunanlar zihnin beyinde çalışan bir program gibi olduğunu düşünürler. Bu bakımdan bir robota, zihni gerçekleştirebileceği bir program sağlanırsa robotun zihin sahibi olduğunu söylemek olanaklı olabilir.[21] Bu anlamda güçlü yapay zekâ çalışmalarında Dilek'in de savunduğu gibi sıkça tercih edilen yöntemlerden biri, insan zihni ve onun nöral temeli bakımından bilgisayımsal terimlerin kullanılması ve bu süreçler ile yapıların işlevsel açıklamalar ile açık kılınmaya çalışılmasıdır.[22] Ancak bağlantıcı anlayış bakımından nöral ağlar, tek fikirlilikten ve aşırı işlevsel yerelleşmeden dolayı sorun yaşar. Burada değinilen nokta, bağlantıcı sistemlerin bir veya birkaç belirli problem ve problem sahası üzerinde yoğunlaşmalarıdır. Böyle bir izole yapının gerçek bir "nöral ağ" olarak kabul edilebilecek insan beyninin işlevlerini yerine getirmesi güçtür. Çünkü beyin birçok görevi eşzamanlı olarak yerine getirebilmekte ve farklı ağlar aynı problemler için katkılarda bulunabilmektedir. Bağlantıcı sistemler ile beyinlerimiz arasında bu bakımdan bir tasarım farkı bulunduğu açıktır.[23] Matematik ve mantık bilimleri bakımından bu çalışma içerisinde ortaya koyduğumuz incelemeler de anımsanırsa yapay zekâ sistemleri içerisinde evrene ilişkin olguların temsil edilebilmesine ilişkin sınırlılıklardan söz etmek mümkündür. Nitekim Fodor ve Pylyshyn de zihin süreçlerinin semantik mahiyetlerinin yapa zekâ programları içerisinde sentaktik yapılar ve formüller vasıtasıyla temsil edilmek istendiğini vurgular. Böylece bir yargının sentaktik anlamı, içinde bulunduğu dizge ve çıkarımdaki rolü bakımından bir değer taşır.[24] Dolayısıyla olgulara ilişkin yargıların anlamları hiçbir zaman

[19] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 84.
[20] Searle, "Minds, brains, and programs", s. 423.
[21] Arıcı ve Doğan, "Prof. Dr. Erdinç SAYAN ile Söyleşi: Zihnin Doğası Üzerine", s. 245.
[22] Dilek, "Zor Problem ve Yüksek Düzey Düşünce Kuramı", s. 140.
[23] Andy Clark, "Philosophical Foundations", **Artificial Intelligence**, ed. Margaret A. Boden, 2. bs., San Diego, Academic Press, 1996, s. 12.
[24] Jerry A. Fodor ve Zenon W. Pylyshyn, "Connectionism and cognitive architecture: A critical analysis", Cognition, C. XXVIII, No: 1-2, 1988, s. 28-30.

semantik düzeye çıkamaz. Nitekim Block'un Çin halkı[*] düşünce deneyi de işlevselciliğin sezgisel olmayan sonuçlarını eleştirmektedir.[25] Çin halkı düşünce deneyine göre oldukça yüksek nüfuslu olan Çin halkının bir saatliğine bir insan zihnini gerçekleştirmesini isteyelim. Bunu başarabilmek için her Çinliye, özel olarak tasarlanmış çift yönlü bir radyo alıcısı sağlayalım. Bu radyo alıcısı hem bir Çinlinin başka Çinli vatandaşlarla iletişim kurabilmesini hem de tüm Çinlilerin bağlı bulunduğu ve beyindeki bir nöron gibi işlev görebilecekleri bir yapay bedene bağlanmasını sağlasın. Gökyüzündeki uydular vasıtasıyla da tüm Çinlilerin eşzamanda görebileceği şekilde türlü harfler onlara gösterilsin. Birbiriyle iletişim halinde olan bir milyon Çinlinin ve uyduların bir bütün olarak çalıştığı bu sistem radyo sinyalleri ile yapay bir bedene bağlı olan dışsal bir beynin rolünü oynamaktadır.[26] Ancak Block'a göre böyle bir işlevsel yapı yalnızca belli bir işlevsel açıklama türünün gerekliliklerini sağlar.[27] Nitekim farklı işlevselciliklerin bulunduğundan ve bu işlevselciliklerin varlığın belli özellik ve durumlarını farklı temellerde açıklamayı arzuladıklarından bu çalışma içerisinde söz edilmiştir.

İşlevselci tavrın kimi savunucuları işlevsel açıklama girişimlerine getirilen, işlevsel sistemlerin fenomenal zihin durumlarına sahip olamayacakları yollu eleştirilere yanıtlar vermek durumunda kalmıştır. Bu yanıtlar fenomenal durumların yadsınması ya da fenomenal durumların işlevsel sistemlerde gerçekleşmeye devam edebileceği yollu iki zıt tavrı içerir. Manzotti ve Tagliasco'ya göre işlevsel bir yapının fenomenal bir yanının olması gerektiğini düşünmek için bir gerekçemiz bulunmamaktadır. Böylece nitelcelerin statüleri üzerinden işlevsel sistemlere getirilen eleştiriler, Manzotti ve Tagliasco'nun işlevsel sistemlerde gerçekleşen haz ve acı gibi fenomenal durumların yalnızca birkaç nöral ateşlemeden daha fazlası olmadığı yollu düşüncesi ile elenmeye çalışılır.[28] Öte yandan Clark tam aksi bir pozisyonu savunur ve Putnam'ın işlevselcilik hususundaki görüşlerinden sonra, bu

[*] **China Brain**, **Chinese Nation** ya da **Chinese Gym** olarak anılan bu düşünce deneyinin Türkçeleştirilmesinin bu yolla daha uygun olduğunu düşündüğüm için **China Brain**'i değil **Chinese Nation** adlandırmasını Türkçeleştirip kullandım.

[25] Games, "The Potential for Consciousness of Artificial Systems", s. 214-215.

[26] Ned Block, "Troubles with Functionalism", **Consciousness, Function, and Representation: Collected Papers**, C. I, Cambridge-Massachusetts, The MIT Press, 2007, s. 71.

[27] **A.e.**, s. 72.

[28] Manzotti ve Tagliasco, "Artificial consciousness: A discipline between technological and theoretical obstacles", s. 111.

görüşleri eleştiren tarafların bir *tür şovenizmi* yapmakta olduğunu iddia eder böylece nöronlardan yoksun olan bir silikon beynin acı hissine ya da kimi kanılara sahip olamayacağını iddia edebilmemiz için bir gerekçemiz bulunmadığını ileri sürer.[29] Ancak her ne kadar işlevselciliğe ve işlevsel sistemlere ilişkin kimi savunma girişimleri bulunsa da eleştirel durumlar da söz konusudur. Levine'in izah gediği düşüncesi anımsanırsa, Chrisley'in vurguladığı gibi işlevselciliğe ilişkin karşı argümanlar da çoğunlukla işlevsel durumların aynılığının nitelcelerin ya da zihin durumlarının aynılığı anlamına gelmediğini ortaya koymayı amaçlar.[30] Nitekim Searle'ün Çin Odası argümanına ilişkin Yale Robot Yanıtına[31] karşı Searle'ün robotların herhangi bir yönelimsel durumlarının olmadığı dolayısıyla onların hareketlerinin yalnızca elektrik kablolamalarının ve programların bir sonucunu olduğu yollu itirazı[32] da Chrisley'in vurguladığı gibi işlevselciliğin, varlığın zihinsel durumlarına ilişkin kuşatamayacağı kimi durumların söz konusu olduğu yollu düşünceyle paraleldir. Block'un vurgusuyla da açıktır ki deneyimin temsili içeriği ile onun fenomenal karakteri arasında bir gedik bulunur. Bu gedik, fenomenal karakterin temsili içeriğe indirgenemeyeceğini gösterir.[33] Görüyoruz ki her ne kadar bir yapının veya birimin işlevinin yerine getirilmesi fikri bakımından işlevselcilik güçlü bir yaklaşım olarak karşımıza çıksa da işlevsel durumların makinelerde gerçekleştirilmesi, fenomenal durumların temsili konusunda birçok problem taşımaktadır. Çünkü bir yapıya ilişkin işlevin bir başka yapıda gerçekleştirilebilir olduğunu düşünsek bile, gerçekleştirilen işlev ile duyumsanan fenomenal içerik arasında bir örtüşme zorunlu olarak bulunmaz. Bunu bir örnek üzerinden düşünelim. Bennu güzel ve güneşli bir günde bir parkta otursun ve bir ağacı seyretsin. O esnada ağacın dalları ve yaprakları arasından gelmekte olan ışık ışınları Bennu'nun cildini

---

[29] Clark, "Philosophical Foundations", s. 2.

[30] Ron Chrisley, "Philosophical Foundations of artificial consciousness", **Artificial Intelligence in Medicine**, No: 44, 2008, s. 134.

[31] Yale Robot Yanıtı'na göre bir robotun içine bilgisayar konur ve bu bilgisayar konvansiyonel bir şekilde sadece girdi-çıktı ilişkilerine göre çalışmasın aynı zamanda hareket eden ve çevresini duyargalar vasıtasıyla algılayabilen bir yapıya bürünmüş olsun. Bu yanıt, burada dile getirilen olanaklara sahip bir robotun semantik içeriği kurabilecek bir yapıya sahip olduğunu ileri sürer. Bu yanıt John Searle'ün Roger Schank'ın görüşlerine karşı ileri sürdüğü kimi argümanlara karşı ortaya atılmıştır. Bkz.: Searle, "Minds, brains, and programs", s. 420.

[32] **A.e.**

[33] Ned Block, "Is Experiencing Just Representing?", (Çevrimiçi) https://www.nedblock.us/papers/1998.ppr.representing.pdf, 9 Aralık 2022.

ısıtsın. Hava ise polen ve bitki kokularıyla tipik bir bahar havası olsun. Bir işlevselci, Bennu'nun güneşten gelen ışık ışınları vesilesiyle ısınmakta olan cildinde gerçekleşen sıcaklık hissinin ve havadaki kokular vesilesiyle burnunda gerçekleşmekte olan kimyasal tepkimelerden dolayı fenomenal bir deneyime dönüşen o güzel kokuların zorunlu olarak herhangi bir güneş ışınına veya bitkiye gerek duyulmadan da gerçekleştirilebileceğini savunabilir. Nitekim bu büyük oranda doğrudur: İşlevselcilik, yapıların işlevlerinin güvencesini sağlamak konusunda güçlüdür. Ancak bir sorun bulunur: İşlevselcilik, işlevlerin güvencesini sağlarken benzer işlevleri yerine getiren yapıların farklılık göstermelerinden ötürü ilk durumda deneyimlemekte olduğumuz ve kurduğumuz gerçeklik tasarımımızı ikinci durumda güvence altına almakta mıdır? Örneğin Bennu'nun cildindeki sıcaklık hissi, işlevselci bir anlayışla, güneş ışınları tarafından değil onun metalik bedenindeki ekzotermik tepkimeler vesilesiyle gerçekleşsin böylece metalik bedenindeki Bennu, güneş ışınlarının ilk durumdaki gibi cildini ısıttığını ve bir sıcaklık hissini duyumsasın. İlk ve ikinci durum hesaba katılırsa Bennu'nun gerçeklik tasarımı değişmemiş midir? Evet ekzotermik tepkimeler vesilesiyle ilk durumdaki gibi ciltteki ısınma işlevi bakımından yerine getirilmiş olabilir ancak güneş ışınının cildi ısıtması ile ekzotermik bir tepkime sonucu cildin ısınması ama bunun görünürde bir güneş ışını tarafından gerçekleştirildiğinin düşünülmesi arasında bir fark bulunmuyor mu? İşlevselcilik deneyimin içeriğinin güvencesini sağlamak konusunda başarı gösteriyor gibi durmamaktadır. Çünkü işlevselciler Bennu'nun cildindeki sıcaklık hissinin güvencesini sağlıyor olsalar bile Bennu'nun *o andalığını* ve fenomenal deneyimlerinin içeriğini temin edememektedirler. Revonsuo'ya göre de işlevselcilik, deneyimimizdeki fenomenal gerçekliğe ilişkin herhangi bir açıklama sunmamaktadır. Çünkü fenomenal deneyimin içeriği, girdi ve çıktı ilişkileri olarak betimlenemez ve işlevsel durumlar olarak nitelenemez.[34] Burada dile getirilen itirazdan kaçmak ve Çin halkı gibi argümanlara karşı çıkmak amacıyla yukarıda da dile getirildiği üzere işlevselcilik savunucuları, deneyimin öznel içeriğine/nitelcelere karşı çıkmakta ve onların varlığını yadsımaktadırlar. Dennett, zihinsel içeriğin sinir sisteminin yapısı ve süreçleri hakkında yeterli bilgi sahibi olunduğunda tam anlamıyla anlayabileceğimiz

[34] Revonsuo, **Bilinç: Öznelliğin Bilimi**, s. 85-86.

durumda olduğunu savunur böylece açıklama girişimlerimizin dışında kalabilecek bir nitelceden söz etmek mümkün değildir.[35]

İşlevselciliğe zihnin ve bedenin ontolojisinden hareketle bir eleştiri sunmak da mümkündür. Searle'ün ünlü düşünce deneyi, beynin işlevsel rollerinin, nedensel ilişkilerinin veya devinimlerinin bilinçli zihinsel durumların varlığı için önemsiz olduğunu ortaya koymaktadır. Bu düşünce deneyine göre beynimizdeki sinir hücrelerinin aşama aşama silikon çiplerle değiştirildiğini düşündüğümüzde, üç seçenek ile karşılaşırız. İlk durumda mantıksal olarak düşünüldüğünde tüm hücrelerimiz silikon olsa bile biz *biz* olmaya devam ederiz. Böylece yalnızca girdi çıktı süreçlerinin işlevsel olarak yerine getirildiği bir durum değil ama aynı zamanda zihinsel fenomenlerin de kopyalandığı bir durum gerçekleşebilir. İkinci bir olasılık, silikon çiplerin sayısı arttıkça bilinçli deneyimimiz küçülecek ancak hiçbir dışsal davranışımızda bir değişiklik yaşanmayacaktır. Bu durumda zamanla bilincimiz yitecektir ama dışsal olarak tıpkı zombi analojisinde olduğu gibi dışsal, davranışsal olarak ilk durumumuzdan ayırt edilemeyen bir özne söz konusu olacaktır. Üçüncü durumda ise dışsal davranışlarımız kaybolacaktır ancak zihin durumlarımız var olmaya devam edecektir. Böylece zihnimiz gelen duyu verilerini işleyip algılayabilecek bir durumda bulunurken bedenimizle olan bağı kopuk kalacaktır. Dolayısıyla bitkisel hayatta olan bir insandan farkımız kalmayacaktır.[36] Bu mantıksal olasılıklar gösterir ki insan bilgisinin, veri yapıları gibi karmaşık şekillerde kodlanabileceğine ve zihinsel süreçlerin işlevler ve fonksiyonlara indirgenebileceğine ilişkin kanı kesinlik taşımaz. Çünkü bir hissin *gerçekleşmesiyle* o hissin *duyumsanması* ve özne bakımından bir bilgi niteliği kazanması iki ayrı durumdur.[37] Nitekim Özel gibi bilgisayımsal-temsil işlevselciliğinin sorunlarını gören ancak bu işlevselciliğin psikoloji ve nörobilim alanları bakımından ayakta kalan bir tarzda savunulabileceğini iddia eden isimler de vardır. Özel'in Algoritmik Zihin Argümanına göre insan zihni, matematik vasıtasıyla makinelere indirgenemez. Ancak makinelere indirgenemeyen kimi zihin işlevlerinin matematikselleştirilmesinin olanağı da yadsınmaz. Böylece psikoloji ve nörobilim

[35] Dennett, **Bilinç Açıklanıyor**, s. 537.
[36] Searle, **The Rediscovery of the Mind**, s. 65-69.
[37] Weizenbaum, **Computer Power and Human Reason**, s. 208.

gibi alanların zihinsel olanlara ilişkin inceleme sahalarının güvencesi matematik vasıtasıyla sağlanmış olur. Zihne ait bir işlev olarak matematik, işlevler bütünü olan zihnin matematiksel bir gösteriminin olanağını sağlar ancak bu olanağın makinelere indirgenemediği savunulur.[38] Özel'e göre organik bir yapı olan beynin fiziksel süreçlerini ve dinamiklerini bilmiyor oluşumuz, beyni bir makine gibi düşünmek için bir engel oluşturmaktadır. Bu anlamda kendi ifadesiyle klasik anlamdaki işlevselci – bilgisayımsal-temsilci işlevselciliğin işaret edildiğini düşünüyoruz – görüş geçerli değildir. Ancak Özel, alternatif ve kendi ifadesiyle "organik bir fonksiyonalizm" olarak tanımladığı Algoritmik Zihin Argümanını öne sürer ve böyle bir işlevselciliğin zihnin matematiksel olarak incelemesini mümkün kıldığını iddia eder.[39] Makine olmayan ve makine gibi çalışmayan ancak *makineye benzer* şekilde çalıştığı savunulan bu tarz bir işlevselcilikteki beyin, matematiksel incelemeler neticesinde üzerindeki bilinmezlerin kaldırılabileceği ve anlaşılmaz olan hiçbir unsurun bulunmayacağı bir yapı olarak tasvir edilir.[40] Ancak Özel'in ortaya koyduğu organik işlevselcilik anlayışının kimi sorunları bulunur: Her şeyden önce makine olmayan ama makine gibi işlemekte olan beyin ifadesi çelişki barındırmaktadır. Nitekim bu çalışmanın matematik bilimi ile ilgili alt bölümünde dile getirilen yargılarımız anımsanırsa ve Hilbert'in matematik projesi bakımından düşünülürse makine denilen insan yapması nesnenin, karar verme problemi bakımından bir Turing makinesinin veya günümüz bilgisayarlarının bir sonucu olarak bulunduğu görülebilir. Böylece matematik biliminin bir sonucu olarak makine denilen yapı, matematik ya da ürünü olduğu biçimsel herhangi bir dizgenin kurallarına sıkı sıkıya bağlı bir yapıdır. Beynin tümü üzerinde değil ama bir bölümü üzerinde "Makine gibi çalışır." demek, beynin bazı bölümlerinin makine gibi çalışmadığını, yani *matematikselleştirilemeyeceğini* kabul etmek anlamına gelir. Böylece her ne kadar beyinde işleyen bir süreçler bütünü olarak zihnin kimi kipleri matematik vasıtasıyla incelenebilir kılınsa bile Özel'in organik işlevselciliğinin kapsayamadığı husus, zihnin tüm yetileri bakımından düşünülmesi gereken şey olarak insanın zihinsel yaşantısıdır. İnsanın zihinsel yaşantısının matematikselleştirilemeyen bölümlerinin

[38] Özel, "Church-Turing Tezi ve Algoritmik Zihin Argümanı: Zihinsel Fenomenleri Matematiğe İndirgemek Mümkün müdür?", s. 170.
[39] **A.e.**, s. 174.
[40] **A.e.**, s. 166.

olduğunu kabul etmek onun var oluşunun, gerçeklik tasarımlarının ve nitelcelerinin matematiğin yani makinelerin ve algoritmaların ötesinde, *transandantal* bir alanda bulunduğu kabul etmektir. Minsky'nin de vurguladığı gibi, zihinsel aktiviteler birimsel veya temel fenomenler olarak matematik veya mantıksal aksiyomlarla ifade edilemezler. Çünkü beyindeki işlevsel süreçler, binlerce alt sistemin ve evrimin milyonlarca yıllık sürecinin bir ürünüdür. Bu sistem, herhangi bir bilimin bir varlığı ele alma tarzı gibi ele alınamaz. Örneğin fizikçiler gibi fiziğin teknik ve imkânları kullanılarak beyin, bu tartışma bakımından bizi bir noktaya getiremez.[41] Çünkü beyin matematik ve mantık temelinde sembolik bir tarzda ifade edilemez.

3.2. Felsefi İşlevselciliğin Bilgisayar Bilimine Dayanan Temelleri

Ünlü makalesinde Alan Turing, yapay zekâ felsefesi bakımından önemli bir soruyu dile getirmiştir: *Makineler düşünebilir mi?* Bu sorunun yanıtlanabilmesinin ise her şeyden önce makine ve düşünme kavramlarının ele alınması üzerinden başarılabileceğini öne sürmüştür.[42] Nitekim çalışmamız içerisinde de ortaya konduğu üzere Turing'in çalışmaları ve devamında gelişmekte olan bilgisayar bilimi ve yapay zekâ çalışmaları bakımından bu alanlardaki araştırmacılar ve düşünürlerin de üzerinde durduğu soru, hesap yapan makinelerin düşünmek bakımından insan gibi düşünmeye sahip olup olmadıklarıdır.[43] Yapay zekâ çalışmalarındaki bağlantıcı düşüncenin temelleri anımsanırsa, bu düşüncenin hedefleri yalnızca beynin nöral yapısının gerçekleştirilmesi değildir. Bağlantıcı düşünce, aynı zamanda insanın nöral yapısının onun üzerinde gerçekleştiği şey olarak beynin fiziksel yapısının da bilgisayar bilimi ve donanım kavramları ile ifade edilmesi ve karşılanması üzerinden bir yönelimi de barındırmaktadır. Öztürk ve Şahin'in de ifade ettiği gibi nöronlara, dendritlere, aksonlara veya sinapslara işlemci elemanı, nöral ağdaki ağırlıklar, yapay nöron çıkışları ya da toplama fonksiyonu gibi işlevler atayan, böylece bu işlevlerin

[41] Minsky, "Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy", s. 38.
[42] Turing, "Computing Machinery and Intelligence", s. 433.
[43] Şengör, "Bilinç, Farkında Olan…", s. 97.

yazılımsal ve donanımsal olarak bilgisayarlarda karşılanmaya çalışıldığı bir çabayı görmekteyiz.[44] İşte bu alt bölümdeki incelememize de konu olan bilgi alanı olarak bilgisayar bilimini ele alıyor olmamız, tam da işlevselci tavır vasıtasıyla bilgisayar bilimine taşınan insan zihinselliğinin bilgisayar bilimi, onun olanakları ve kavramları bakımından çözüme kavuşturulmak istenmesine ilişkin bir tavır üzerine eğilmektedir.

Bilgisayar bilimi içerisinde sıkça karşılaşılan yaklaşım, beynin bir bilgi işleme makinesi olduğu, bu bilgi işleme makinesinin en temel özelliğinin ise hesaplama yapma kabiliyeti olduğu görüşüdür. Tegmark tam da işlevselci bir temele dayanarak hesaplamanın parçacıkların uzay ve zaman düzeninde bulunan bir örüntü olduğunu, mühim olanın bu örüntüyü dolduran fiziksel parçacıkların değil örüntünün kendisi olduğunu iddia etmektedir.[45] Böylece bilgisayardaki birimler olarak bitler, *bilgi atomları* olarak bilgiyi oluşturan ve ondan daha küçüğü düşünülemeyecek yapılar gibi – bilgisayar bilimi içerisindeki Sokrates öncesi eğilimleri de görebiliyoruz çünkü böyle bir bit yaklaşımının, Antik Yunan'ın **atomos**'undan bir farkı bulunmuyor – kabul edilirlerse[46] o halde herhangi bir bilgi işleme yapısının da tıpkı matematikteki gibi fonksiyonların uygulanmasıyla bilginin alınması ve dönüştürülmesini işini gerçekleştirebileceğini kabul edebilmeliyiz.[47] Peki zihinlerimiz ve beyinlerimiz bit düzeyine indirgenebilir ve bu düzeyde temsil edilebilir mi? Zihin durumlarımız ve beyinlerimiz bilgisayar biliminin kavramları kullanılarak açıklanabilir mi? Beyinlerimiz, karmaşıklık ve koordine olmak bakımından moleküler süreçlerin gerçekleştiği ve bunun sonucunda algı, düşünme ve eylem gibi birçok durumun gerçekleştiği bir yapıdır. İşlevselcilik, tam da bu yapının kendisini bilgisayar üzerinden açıklamaya ilişkin bir tavrı da barındırır. Zihinlerimizin ise bilgisayarlarda çalıştırılan programlardan bir farkının bulunmadığı düşünülür. Ancak Bell'in de dile getirdiği gibi, bir bilgisayar modelinin hesaplama yapabileceğini biliyor olsak bile onun *düşünebileceğini ya da canlı bir unsur olduğunu* iddia edebilir miyiz?[48] Böyle bir soruya duyulan ihtiyaç, felsefi işlevselciliğin temelinde yer alan bir varsayım olarak beynin *bilgi işleyen bir makine*

[44] Öztürk ve Şahin, "Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış", s. 28.
[45] Tegmark, **Yaşam 3.0**, s. 94.
[46] **A.e.**, s. 83.
[47] **A.e.**, s. 87.
[48] Bell, "Levels and loops: the future of artificial intelligence and neuroscience", s. 2015.

olduğu yollu kabuldür.[49] Güçlü yapay zekâ düşüncesi ile ulaşılmak istenen hedeflerden biri, bir bilgi işleme makinesi olarak beynin inorganik materyalde ve yapay zekâ üzerinden gerçekleştirilmek istenmesidir. Searle, güçlü yapay zekâ düşüncesinin temelinde beynin işlevleri ile dijital bilgisayarlar arasında kurulan bir analojinin bulunduğunu saptar. Beynin dijital bir bilgisayar olduğu, zihnin ise bu dijital bilgisayar üzerinde çalışmakta olan bir programdan fazlası olmadığına ilişkin radikal bir tavrın söz konusu olduğunu dile getirir. Bu tavra göre insan zihni, herhangi bir biyolojik yapıya veya temele bağlı değildir. Dolayısıyla biyolojik temel olarak beyin, insan zihninin gerçekleştirilmesi için zorunlu bir yapı olarak düşünülemez. Bu anlamda doğru program ve doğru girdi çıktıların sağlandığı bir dijital bilgisayar – veya fiziksel herhangi bir sistem – tıpkı biz insanlar gibi zihin durumlarına sahip olabilir.[50] Hatta böyle bir makinenin gerçekleşmesi düşüncesinden hareketle hesap makineleri ve insan beyni arasında bir fark ortaya konur: Bu iki yapı türce aynı olmayan ancak ölçü, hız, otonomi vb. gibi durumlar bakımından benzer özelliklere sahip yapılar olarak düşünülür.[51]

Bilgisayarlara duyulan ihtiyacın temeli ve medeniyetimiz açısından bilgisayarların oynadığı rol nedir? Modern bilgisayarlar, medeniyetimiz açısından oldukça yeni araçlardır. Antropolojik açıdan düşünülürse insanlık tarihi bakımından modern bilgisayarları ortaya çıkaran ihtiyacın en temelde bir *hesaplama ihtiyacı* olduğu görülebilir. Hesaplama işi için en uygun ve bilinen araç ise uzunca bir süre abaküstür.[52] Abaküs, insanın el ve eklemleri ile yaptığı saymaya göre daha etkili bir sayma ve hesaplama olanağı sağlamaktadır. Bu olanağı sağlayan iki durum bulunmaktadır: Veri girdisi ve verinin işlenmesi süreci. Sayma ve hesaplama süreçlerinde gerçekleşen şey elimizdeki işlenmemiş verilerin makineye girilmesi ve makineye girilmiş bu verilerin makine tarafından işlenmesidir. Makinenin son durumu ise çıktıyı yani sonucu temsil etmektedir. İşte analog ve dijital bilgisayar kavramlarını ortaya çıkaran şey de tam olarak budur: Fiziksel parçalardan oluşan ve

[49] Gözkan, "Bilgi, Bilinç ve Yapay Zekâ", s. 158.
[50] Searle, **Minds, Brains and Science**, s. 28.
[51] Peter Stone v.d., "Artificial Intelligence and Life in 2030", One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, (Çevrimiçi) https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/ai100report10032016fnl_singles.pdf, 9 Aralık 2022.
[52] Floridi, **Philosophy and Computing: An Introduction**, s. 21.

yapmakta oldukları hesaplamaları zamanda devamlılık gösteren, girdi mahiyetindeki fiziksel fenomenler vasıtasıyla gerçekleştiren – örneğin güneş saatinin çalışması için gölge, kum saatindeki kum vb. – araçlar analog bilgisayar olarak anılır. Analog bilgisayarlar, gerçek zamanlı makinelerdir. Yani şu anda ve burada çalışan makineler olarak bulunurlar. Bu anlamda veri bozunmasına ve fiziksel sorunlarla karşılaşmaya oldukça uygundurlar. Bu yapıları onların, günümüzde kullanıyor olduğumuz gibi, genel amaçlı makineler olarak bilgisayarlarımız gibi makineler olamayacaklarını göstermektedir.[53] Dijital bilgisayarlar ise 0 ve 1 gibi iki sembole dayanarak hesaplamaları gerçekleştiren ve çıktıyı ise yine bu semboller cinsinden sağlayan makinelerdir. Burada kullanılan her 0 ve 1 bir *bit* olarak tanımlanmıştır. Bitler mantıksal olarak doğru veya yanlışı, matematiksel olarak 0 veya 1'i, fiziksel olarak transistörlerin açık veya kapalı olmasını, elektrik devresinde ise yüksek veya alçak voltaj durumlarını temsil etmektedirler. Analog sinyallerden farklı olarak dijital sinyaller vardır ya da yoktur. Ara bir durum söz konusu olamaz. Böylece dijital bilgisayarlar yüksek bir kesinlik ile çalışır.[54] 0 ve 1'ler olarak bitler, makine dilinin en temel birimleridir ve bu kesinlik matematiksel bir kesinliğe, bir tek anlamlılığa dayanmaktadır. Programlar ve yazılım dilleri farklılık gösterse bile tüm bu farklı program ve yazılım dilleri, en temelde 0 ve 1'lerden oluşmakta olan makine diline çevrilmektedir.[55] Dijital bilgisayarın temel birimleri düşünülürse, çalışma mantığının sayılara karşılık gelen soyut bilgi birimlerinin manipüle edilmesi üzerinde şekillendiği görülebilir. Yapay zekâ çalışmaları düşünüldüğünde bitler, şeyleri temsil eden semboller ve bu sembollerin sıkı mantık kuralları ile işletildiği bir yapıyı oluşturur.[56] Günümüz bilgisayarlarına gelen süreç ise tamamen dijital bilgisayarların sağladığı olanaklar ve bu olanakların ortaya çıkmasını sağlayan medeniyetimizin ihtiyaçları ölçüsünde ortaya çıkmıştır. Abaküs gibi hesaplama bakımından sınırlı ve bellek bakımından dış etkilere oldukça açık bir aracın medeniyetimizin yüksek hesaplama ve bellek gücü isteyen sayısal süreçleri karşısında, ihtiyacı karşılayacak bir niteliği bulunmamaktadır. Nitekim biz insanlar büyük miktardaki veri ve büyük

[53] **A.e.**, s. 22.
[54] **A.e.**, s. 23-24.
[55] Bingöl, "Bilgisayar Üzerine", s. 31-32.
[56] Daniel Crevier, **AI: The Tumultuous History of the Search for Artificial Intelligence**, New York, Basic Books, 1993, s. 3.

sayılar üzerindeki hesaplamaları yürütmek için dijital bilgisayarları kullanmaktayız. Dijital bilgisayarların, medeniyetimizin yüksek hesaplama gücü ihtiyacını karşılamak uğruna sahip olması gereken ve şimdiye kadar bu ihtiyacı karşılamak konusunda gösterdiği başarıyı ortaya koyabilmesini sağlayan temel özelliği, George Boole'un cebrine dayanan ve karşılaşılan tüm durum ve varlıkların 0'lar ve 1'ler ile temsil edildiği yapıdır. Bu yapı, Claude Shannon'un *bir yüksek lisans tezinde ortaya attığı düşüncelere* dayanmaktadır. Ona göre dönen diskler üzerinden yapılan hesaplamadan ziyade, birbirlerine bağlı olan ve birbirlerini açıp kapatma kabiliyetine sahip olan anahtarların oluşturduğu dijital bir hesap makinesi tasarlanabilir. Bu durum Boole cebrinin rölelere entegre edilmesi ve iki durumun – açık/kapalı – temsil edilmesini olanaklı kılmıştır. Sayılar bu yolla ikili sistem üzerinde temsil edilebilecektir.[57] Yüksek lisans tezinde Shannon, verili herhangi bir zamanda bir devrenin açık veya kapalı olduğunu vurgular, açık veya kapalı olan devre durumlarının kapalı olduğunda 0, açık olduğunda ise 1 sembolü ile temsil edilebileceğini ileri sürer.[58] Böylece 0 ve 1 sembolleri üzerine kurulmuş olan devre sisteminin durumu, Shannon'a göre Boole'un sistemindeki eşitliklere uygulanabilmektedir[*] ve Boole'un ortaya koyduğu mantıksal ilkeler bu devreler üzerinde gerçekleştirilmektedir.[59] Dolayısıyla Shannon'un bu önerisi, bir dizi anahtarın bir araya geldiğinde belirli önermeler ve bu önermeler arası ilişkilerin analiz ve temsil edilebileceği bir olanağı da bize sağlamaktadır.[60]

Bir hesaplama makinesi olarak düşünülen ve dijital bilgisayarlar vasıtasıyla inorganik mimaride gerçekleştirilebileceği savunulan insan beyninin zihin

[57] **A.e.**, s. 10-12.
[58] Claude Elwood Shannon, "A Symbolic Analysis of Relay and Switching Circuits", Cambridge, Massachusetts Institute of Technology, Yayımlanmamış Yüksek Lisans Tezi, 1940, s. 4.
[*] Medeniyetimizin tarihini ve seyrini etkileyen böyle bir önerinin *bir yüksek lisans tezinde* ortaya atılmış olması, akademimizde alışılagelmiş olan yüksek lisans tezi ve doktora tezi kabullerinin ne kadar da *temelsiz* ve *ezbere düşünceler* olduğunu bize göstermektedir. Yüksek lisans tezlerinin belli bir sorumluluk olarak değil, *kendi başına arzulanan ve başarılmak istenen, özen ve emeğin bulunduğu, özgün ve yaratıcı bir çalışma* olması beklenirken akademimizdeki yaygın anlayışın, burada ortaya koyduğumuz anlayışın *tam aksi bir istikamette* bulunduğunu saptayabiliriz. Yüksek lisans ve hatta kimi doktora tezlerinin yeni ve başarılı bir girişim olmaktan ziyade belirli bir *bürokratik engelin aşılması uğruna* ortaya konan çalışmalar olduğunu akademimizde sıkça görmekteyiz. Bu çalışmalar akademimizdeki "felsefe" ve "felsefeci" algısını aşağı çeken, onu değersizleştiren bir boyuta ulaşmıştır. *Endüstrileşen* akademimizdeki bu anlayışın terk edilmesinin, felsefe akademimizin çağdaşlaşması yolundaki önemli adımlardan biri olacağını düşünüyoruz.
[59] **A.e.**, s. 5-6.
[60] Crevier, **AI: The Tumultuous History of the Search for Artificial Intelligence**, s. 18.

durumlarının da gerçekleştirilmesi gerekmektedir. Zihin durumları için düşünülen karşılık ise programlardır. Programların bir zihin olabilmesi için her şeyden önce ona – epistemolojik bir temelin sağlanması amacıyla – bilginin ne olduğu ve bilginin nasıl elde edildiği ile ilgili yetilerin kazandırılması gerektiği düşünülmüştür.[61] Ancak bir program nedir? Programların burada dile getirilen yetileri elde edebilecek yapılar olduğunu iddia edebilir miyiz? Programlar, bir işin yerine getirilmesi için hazırlanmış komut dizileridir, dolayısıyla sonlu olmak durumundadırlar. Sonlu bir biçimsel yapı olarak programların çalıştırılması ise, programın içerisinde bulunan yargıların sıralı bir biçimde gerçekleştirilmesidir.[62] Programların bu mahiyeti, bilgisayarların yapabileceklerinin sınırı düşünüldüğünde bize ucu açık bir olanak alanı sağlamaktadır. Çünkü Alpaydın'a göre bir zamana kadar yalnızca hesap makineleri olarak düşünülen bilgisayarların, bilginin en ufak parçalarının sayılar olarak temsil edilebileceği düşüncesi bir kere ortaya atıldığında, bitler üzerinden yalnızca sayılar veya metinleri değil, aynı zamanda görseller veya sesler gibi diğer unsurları da temsil edebileceği ve onları işleyebileceği anlaşılır.[63] Böylece tüm bu süreçleri yürütebilecek bir biçimsel yapı olarak programlar daha da önem kazanmış olur. Algoritmalar olarak da ifade edilebilecek programlar, belirli bir girdinin belirli bir çıktıya dönüştürülmesi işi için tasarlanmıştır.[64] Böylece bilgisayar bilimciler ve mühendisler için algoritmalara dökülmüş olan gerçeklik, yine bu algoritmaların uygun şekilde tasarlanması yoluyla bilgisayarlar üzerinde gerçekleştirilebilecektir.[65] Algoritmaların bir problemin çözümüne yönelik tasarlanmış biçimsel yapılar olarak düşünüldüklerinde, zekâ sahibi olmalarının koşulları üzerinde durulmuştur. Örneğin Köse gibi isimler bir algoritmanın zekâ sahibi olma koşullarını onun sonlu ve tasarlanmış bir yapı olması ve bir problemin çözümü bakımından insan müdahalesinin dışında eyleyebilecek bir otonomluğa sahip olabilmesi gibi birçok özellik üzerinden düşünmektedirler.[66] Ancak programların yapay zekâ felsefesi bakımından ele alınışları kimi sorunlarla yüzleşmek durumundadır. Örneğin Hehner'e göre, katı bir biçimsel notasyonla – yani bir programlama dili ile –

[61] McCarthy, "What has AI in Common with Philosophy?".
[62] Say, "Yapay Zekâ ve Gödel Teoremi", s. 255.
[63] Alpaydın, **Yapay Öğrenme: Yeni Yapay Zekâ**, s. 18.
[64] **A.e.**, s. 27.
[65] **A.e.**, s. 77.
[66] Köse, **Yapay Zekâ Felsefesi**, s. 19-21.

belirlenmiş olan bilgisayar programları ile bilgisayar biliminin de temelini oluşturan ve gerçeklik tasarımlarının onun notasyonu üzerinden kurulduğu matematik arasında bir ayrım bulunmaktadır. Matematiğin, programların gücünü ve güvencesini sağladığı ilk adımda düşünülse bile Hehner, programların tıpkı matematikteki fonksiyonlar gibi ifade edilemeyeceğini savunmaktadır. Çünkü programlar fiziksel süreçleri vurgular ve bu süreçleri belirli adımlar şeklinde düzenler. Dolayısıyla program belirlenmiş ve süresiz bir şekilde bulunur. Oysa bir matematik fonksiyonu böyle değildir: Bir matematik fonksiyonu için varılan sonucun bir önemi bulunmaz. Burada yalnızca bir nesneden bir başka nesneye doğru bir işaret etme söz konusudur. Dolayısıyla bilgisayar programları her ne kadar katı bir matematiksel ve mantıksal notasyona dayanıyor olurlarsa olsunlar programlardaki *yönelimsellik*, bu notasyonun kendisine ait değildir.[67] Burada açıktır ki programlar vasıtasıyla bilgisayar mimarileri üzerinde gerçekleştirilmek istenen insan zihninin yönelimselliği, notasyonun kendisinde sağlanamaz. Sağlanamadığı için de bu yönelimselliğin yine varlığın kendine dayanarak ortaya koyduğu bir yönelimsellik mi yoksa ona dışarıdan yine bir başka zihin tarafından sağlanan bir yönelimsellik mi olduğu sorusu bilgisayar bilimi tarafından ve genel olarak yapay zekâ felsefesi içerisinde yanıtlanması gereken önemli bir soru olarak durmaktadır.

Programların ne olduğu ve programların insan zihinselliğini temsil etmek bakımından becerileri ile ilgili tartışmalar hesaba katıldığında, bir bilgisayarın ne olduğu sorusuna verilen yanıtların da değiştiği görülebilir. Dreyfus'lar, bilgisayarların ne olduğuna ilişkin iki karşıt görüşün bulunduğunu saptar. Bir görüşe göre bilgisayarlar zihinsel sembolleri manipüle eden ve bu yolla dünyanın biçimsel temsillerinin onun üzerinde işlendiği bir yapı olarak anlaşılır. Bu yapının gerçekleştirilmesi için ise mantık temel alınır. Bir diğer görüşe göre ise bilgisayarlar beyni modellemek için bir araç olarak kabul edilir. Böylece önemli olanın biçimsel sembollerin manipülasyonu değil ama nöronlar arası ilişkilerin simüle edilmesi olduğu öne sürülür. Bu simülasyon sürecinin başarılabilmesi için ise mantık bilimi

[67] Eric C. R. Hehner, "Problems with the Halting Problem", **Advances in Computer Science and Engineering**, C.I, No:10, 2013, s. 5-7.

değil ama istatistik temel olarak düşünülür.[68] Dreyfus'ların yaptığı bu saptama, eleştirimiz açısından da önem arz etmektedir. Çünkü şimdiye kadar yapay zekâ felsefesi bakımından eleştirisini sunmaya çalıştığımız felsefi işlevselciliğin dayandığı bilgisayımsal-temsilci işlevselciliğin burada da dile getirildiği üzere hem mantık hem de istatistik bilimlerine dayanan anlayışlarının tamamı, insan ve makine zekâsı arasındaki ayrımı önemsememekte ve mühim olanın, Moravec'in de vurguladığı gibi insan performansına erişmek için daha fazla hesaplama gücüne sahip olmak olduğunu savunmaktadır. Bu savununun çıkış noktasının ise hesaplama gücünün miktarı olduğu dile getirilir.[69] Peki ama Turing'in düşüncelerine, ortaya attığı ve kendi adıyla anılan Turing makineleri fikrine geri dönülürse, bilgisayarların ne olduğuna ilişkin burada aktardığımız iki farklı yaklaşımın temelleri sağlanabilir mi? Çünkü Floridi'nin de ifade ettiği üzere Turing makineleri katı deterministik sistemlerdir ve girdilerle çıktılar arasındaki ilişki sıkıdır. Ve bu sıkılık, Turing makinelerinin herhangi bir durumda açık bir şekilde yorumlanıp anlaşılabilmelerine olanak sağlamaktadır.[70] Ancak insan zihni ve makineler arasında kurulmaya çalışılan ilişkide, insan zihninin doğası gereği Turing makinelerinin açıklığı ve sıkılığı karşısında insan zihninin henüz açıklayamadığımız bir doğasının bulunduğunu da hatırlamamız gerekir. Ancak Turing'in görüşlerine ve Turing makinelerinin ortaya çıkmasını sağlayan temel fikre eğilmemiz gerektiği de açıktır.

Turing'in bilgisayar bilimi bakımından önemli olan kimi görüşlerini anlamak için bir örnek üzerinden düşünelim: Bennu bir elma bahçesi işletsin ve işlettiği bu elma bahçesindeki ürün girdi çıktılarını kâğıt ve kalem ile hesaplasın. Haftanın her günü kaç kilogram elma ihraç edildiği ile ilgili toplama işlemleri, Bennu tarafından isabetli bir şekilde yerine getirilsin. Örneğin sıkı çalışılan bir haftada gün başına 100 kilogram elma toplansın ve Bennu, kalem vasıtasıyla kâğıt üzerinde haftanın yedi günü için toplamda 700 kilogram elma toplandığını basit bir toplama işlemini yerine getirerek bulsun. Sorumuz şudur: Bennu'nun doğru sonuç olan 700 kilogram sonucuna ulaşması için sahip olması gereken *zorunlu yeti ve haller* nelerdir? Örneğin

[68] Hubert L. Dreyfus ve Stuart E. Dreyfus, "Making a Mind Versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint", **Dædalus**, C. CXVII, No: 1, 1988, s. 15-16.
[69] Moravec, "When will computer hardware match the human brain?", s. 11.
[70] Floridi, **Philosophy and Computing: An Introduction**, s. 28.

Bennu toplama işlemini yaparken kırmızı bir elbise giymeseydi bu sonucu elde edemez miydi? Ya da bu toplama işlemini yaparken ortam sıcaklığı 24 santigrat derece olmasaydı bu sonuca ulaşılamaz mıydı? Bennu somurtursa 700 kilogram yanıtına ulaşılamaz mı? Bu ve benzeri sonsuz sayıda soru sorulabilir. İşte Turing'e ilham sağlayan en temel saptamalardan biri de bu durumdur: Bir insan hesap yaparken hesap yapma sürecinde *içkin olan*, *özsel unsurlar ile ilineksel unsurlar* arasında bir ayrım yapılmalıdır. Bennu kâğıt ve kalem ile toplama işlemini yaparken Turing'in ifadesiyle bir *kâğıt makinesi* olarak bulunmaktadır. Çünkü her şeyden önce Turing'e göre hesaplama, yazılı belirli yönergelerin işletilmesidir. Dolayısıyla bir program tam anlamıyla saf bir kâğıt işidir. Kâğıt, kalem ve silgi sağlanmış bir insan sıkı bir şekilde çalışarak Turing'in evrensel makine dediği makine gibi, yani genel amaçlı makine gibi eyleyebilir.[71] İşte Turing'in Turing makinesi olarak anılan fikri, insanların problem çözerken veya hesaplama yaparken sahip oldukları ilineksel durumların bu hesaplama ve problem çözme durumlarından ayrı tutulması ve bu durumlardaki özsel aşama ve yetilerin makinelerde gerçekleştirilmesi fikrine dayanmaktadır. Evrensel Turing makineleri ise belirli problem ve hesaplamalara özgü değil ama tüm problem ve hesaplamalarda içkin olan ortaklıkların doğasına dayanarak ortaya atılmış bir fikir olarak, her bir özel durum için programlanabilecek genel amaçlı makineler olarak düşünülmüştür. Böylece farklı işleri yapmak için sınırsız sayıda makineye sahip olmaktansa tek bir makineye sahip olmanın makinenin farklı problemleri çözmek için programlanabilmesi yoluyla yeterli olabileceğine inanılır.[72] Turing'e göre düşünen makineler, elektronik bilgisayarlar veya dijital bilgisayarlar olarak anılmalıdır. Bu bilgisayarların arkasında yatan düşünce, *bir insan bilgisayar* tarafından yapılabilecek herhangi bir operasyonu yerine getirebilmektir. Dolayısıyla dijital bir bilgisayar üç birimden oluşmalıdır: Depolama, işlem birimi ve kontrol birimi. Turing'in kalem ve kâğıt örneği üzerinden yaklaşımı düşünülürse depolama birimi, bilginin depolandığı yer olarak insan bilgisayarın kağıdına denk gelmektedir. Kâğıt, insan bilgisayarın işlemleri onun üzerinde yaptığı veya kuralların mevcut olduğu materyaldir. İşlem birimi ise hesaplama süreçlerindeki her tekil işlemi yerine getirmektedir. Kontrol biriminin

[71] Turing, "Intelligent Machinery", s. 112-113.
[72] **A.e.**, s. 111.

görevi ise kuralların ve işlemlerin uygun şekilde işletip işletilmediğinin kontrolünü sağlamaktır.[73] Turing, dijital bilgisayarların inşa edilebileceği fikrine inanmaktadır ve dijital bilgisayarlar *insan bilgisayar* olarak bizlerin eylemlerini oldukça yakın bir şekilde taklit edebilecektir. Çünkü karmaşık bir işlem sürecinde bir insanın eylemleri bir makinede gerçekleştirilmek istenirse basit bir şekilde insana bu işlem sürecinde ne yaptığı ve bunları nasıl gerçekleştirdiği sorulmalı ve yanıtlar bir yönerge tablosuna çevrilmeli, makineye sunulmalıdır.[74] Turing'e göre bu makinelere ilişkin kimi matematiksel itirazlar dile getirilebilir: Gödel'in tamamlanamazlık kuramları[*] bir aksiyomatik sistemde doğruluğu veya yanlışlığı ispatlanamayacak kimi önermelerin bulunduğunu öne sürmektedir. Böylece Turing, burada dile getirilmiş olan dijital makinelerin gücüne ilişkin kimi kısıtlamaların söz konusu olduğunu kabul etmektedir.[75] Ancak Gödel, her ne kadar kuramının Turing makinelerinin olanaklılığı bakımından sonuçlarının ayırdında olsa bile "sınırlı prosedür" kavramının ortaya koyduğu gelişmenin gerçekleşmeleri olarak bu kavramın sınırlı sayıda parçaya sahip makinelere, yani Turing makinelerine, indirgenmesini olumlu karşılamaktadır.[76] Çünkü Turing makinesinin hesaplama açısından önemi, insanın hesaplama yaparken gerçekleştirdiği sürecin sembollerin kağıda yazılmasından ibaret olduğu düşüncesinden hareketle tek boyutlu ve karelere bölünmüş bir kâğıt olarak *bir bant* üzerinde gerçekleştiği bir tasarım olmasıdır.[77] Bu tasarım, tam da insan bilgisayarın hesaplama yaparken sahip olduğu ancak hesaplamanın sonucuna etki etmeyen durumların elenmesi sonucu ortaya çıkmış bir düşüncedir. Çünkü hesaplama yaparken yalnızca temel belli kurallar işletilmektedir. Böylece yukarıda da dile getirdiğimiz gibi hesaplamanın doğasına ait olmayan ancak insan olmanın doğasına özgü durumlar ile hesaplama ile direkt ilişkili durumlar arasında bir ayrım yapılmaktadır. Turing'in bir sonraki fikri ise bir insanın bir makine ile değiştirilebilecek türden bir varlık olduğunu göstermektir. Çünkü zihinde

[73] Turing, "Computing Machinery and Intelligence", s. 436-437.
[74] **A.e.**, s. 438.
[*] Eleştirimizin ilk bölümünde matematik bilimi üzerinden yapay zekâ felsefesine ilişkin problemler dile getirilmiş ve kimi itirazlar sunulmuştur.
[75] **A.e.**, s. 444.
[76] Gödel, "Some basic theorems on the foundations of mathematics and their implications", s. 304-305.
[77] Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem", s. 249.

gerçekleşen süreçler hesaplamanın bir sonucudur ve makineler hesaplama yapmak konusunda insanlardan geri kalmamaktadır.[78] Nitekim bu yaklaşımdaki ana fikirlerden biri, "hesaplama" kavramının belirli bir yorumuna dayanmaktadır: Algoritmaya dökülebilen ve özyinelemeli her durum bir Turing makinesi tarafından hesaplanabilmektedir. Dolayısıyla bazı spesifik eylemlerin Turing makinesi tarafından gerçekleştirilemeyeceği gösterilirse o halde bu eylemlerin algoritmik bir temelde ifade edilemeyeceğini de ortaya koyulmuş olur.[79] Nitekim bir problemin çözülebileceği mümkünse o halde bu çözümü gerçekleştirebilecek en az bir Turing makinesinin bulunduğundan da söz etmek mümkündür.[80] Alonzo Church, Turing ile yakın zamanlarda ortaya attığı hesaplama kuramında da benzer bir fikri savunur. Church'ün 1935 yılında ortaya attığı teze göre doğal sayılar üzerinde etkili bir şekilde hesaplanabilecek bir fonksiyon yalnızca bir Turing makinesi tarafından hesaplanabilir. Hesaplanabilir olan bir fonksiyon ifadesi, etkili prosedür veya algoritma olarak düşünülebilir. Böylece döngüye düşmeyen, öz-yinelemeli algoritmalarla ifade edilebilecek olan her fonksiyon hesaplanabilir olarak düşünülebilir.[81] Turing'e göre de Church, kendisinin hesaplanabilirlik düşüncesine eşdeğer fakat farklı tanımlanmış etkin bir hesaplanabilirlik fikrini ortaya atmıştır.[82] Peki tüm bu gelişmeler insan zihni ve makineler arasında kurulmak istenen ilişki bakımından neleri ifade eder?

Hesaplama araçlarının ve kuramlarının gelişmesi ile beraber hesaplamanın erimi, yalnızca sayılabilirliğin söz konusu olduğu durumlarla sınırlı değil ama neredeyse tüm fenomenal durumlara doğru genişletilmek istenmiştir. İnsan zihninin sayısallaştırılabileceği düşüncesi, hesaplama kuramının nesnesi olarak insan zihni ve onun yetilerini yeniden ele almaya doğru bir düşünceye imkân sağlamıştır. Böylece bilgisayarlar ve programlar, insan zihnin yapıp etmelerini algoritmik bir düzeyde ve transistörler ile gelişmiş diğer donanımlar vasıtasıyla gerçekleştirmek konusunda bir aday olarak görülmüşlerdir. Turing ve Church'ün hesaplama kuramları ve Turing

[78] Davis, **The Universal Computer: The Road From Leibniz to Turing**, s. 147-148.
[79] **A.e.**, s. 151.
[80] Say, "Diller, Çözümsüz Problemler ve Sonsuzluk", s. 183.
[81] Robert Audi, "Church's thesis", **The Cambridge Dictionary of Philosophy**, 2. bs., New York, Cambridge University Press, 1999, s. 141-142.
[82] Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem", s. 231.

makinesi fikri birlikte düşünülürse, insan zihninin makinelerde gerçekleştirilebilmesine ilişkin arzunun başarı ölçütüne ilişkin bir fikrin de bulunması gerektiği düşünülebilir. Taklit oyunu* olarak anılan oyunda Turing, bir bilgisayarın insan zekâsına denk bir zekâya sahip olup olmadığının belirlenebileceğini savunur.[83] Turing bu oyunu cinsiyeti önemsiz bir sorgulayıcı, bir kadın ve bir erkek üzerinden tasarlamıştır. Bu oyunda sorgulayıcı, bir arayüz üzerinden kendisini insan olduğuna ikna etmeye çalışan bu deneklerle konuşmaktadır. Turing bu esnada kadın denek yerine bir makinenin geçmesi durumunda ne olacağını sorar.[84] Burada sorgulayıcının, konuşmakta olduğu deneklerden hangisinin insan olduğunu hangisinin ise yapay zekâ olduğunu tespit etmesi beklenmektedir. Ancak sorgulayıcı, iletişimde olduğu deneklerden hangisinin insan hangisinin ise yapay zekâ olduğunu tespit edemezse ne olur? Turing, ortaya konan bu yeni problemin insanın fiziksel ve entelektüel kapasiteleri arasına keskin bir çizgi çekmekte olduğunu düşünür.[85] Nitekim de öyledir: Bu eleştirinin antropomorfik bir felsefe yapma tarzı ile ilgili sunmuş olduğu alt incelemede de vurgulandığı üzere, yapay zekâ felsefesinde bir felsefi düşünmenin sonucu olarak değil ama bir metafizik kabulün hedefi anlamında türetilmeye çalışılan yapay zekâlı varlık tasarımlarının meydana gelmesine sebep olan şey, insan tasarımımızın yapay zekâ çalışmaları karşısında güncelliğini yitirmesidir. "İnsan nedir?" sorusunu yeniden yanıtlamaya her zamankinden daha fazla ihtiyaç duymaktayız. İnsanın bedeniyle olan ilişkisini ve bedeninin insanın var olmaklığı bakımından temsil ettiği kaçınılmaz önemi vurgulamak, bu durumlar kaçınılmaz değilse bile *neden öyle olmadığını anlamak* zorundayız. Aksi halde Turing'in taklit oyunundan beri taşınmakta olan bu düşünsel mirasla daima yüzleşmek durumundayız. Bedenin yadsınması ve insanın bedeniyle olan sıkı ilişkisi açığa çıkarılmadıkça yapay zekâlı varlıkların insan olmak bakımından insanların yalnızca zihinsel varlığını taklit etmesi üzerinden bir yeterlilik elde edebileceklerini düşünmek için bir gerekçemiz bulunmamaktadır. Ancak unutulmamalıdır ki Turing'in ortaya attığı bu test, yapay

* İng. **Imitation Game**
[83] Simon Blackburn, "Turing test", **The Oxford Dictionary of Philosophy**, 2. bs., Oxford, Oxford University Press, 2005, s. 371-372.
[84] Turing, "Computing Machinery and Intelligence", s. 433-434.
[85] **A.e.**

zekâ çalışmaları içerisinde insan zihninin yapısal olarak tamamen kopyalanabileceğini ve makinelerde başarılabileceğini savunanların tümü tarafından kabul görmez.[86] Dolayısıyla beden üzerine yapmış olduğumuz vurgular taklit oyunu bakımından bir önem taşır.

Yapay zekâ felsefesi ve bilgisayar bilimi bakımından yapay zekâ üzerine yürütülen tartışmalarda taklit oyunu, üzerinde sıkça durulan ve tartışılan bir düşünce deneyidir. Taklit oyunu ile ilgili eleştiriler, makinenin sentaks üzerinden ve yönelimsel durumlara sahip olmadan arzulanan başarıları elde edip edemeyeceği üzerinden geliştirilir. Nitekim bu eleştiriler Searle'ün yapay zekâ eleştirisinin de merkezinde yer alır. Çünkü ona göre makinenin içinde çalışan programlar olarak yapay zekâlar, tamamen sentaktik bir düzende inşa edilmektedir. Oysa anlam ve yönelim ancak ve ancak bir semantik yapı içerisinde bulunabilir.[87] Searle'ün eleştirilerine vesile olan iddia, Roger Schank'ın ileri sürdüğü bir programdır. Bu programın amacı insanın anlatıları ve bir bağlam içerisinde gerçekleşen belirli diyalogları anlama yeteneğinin simüle edilmesidir. İnsanların bu alanlardaki anlama yeteneğinin, bazı bilgiler eksik olsa bile yine de bu gibi durumlarda yanıt üretebilecek kadar gelişmiş olduğu unutulmamalıdır. Buradaki asıl sorun güçlü yapay zekâ düşüncesini savunan düşünürlerin, makinenin soru ve cevap süresince insanın yeteneğini yalnızca simüle etmediğini, aynı zamanda makinenin söyleneni gerçekten de anladığını, kendisinin de bu yolla insanın anlama süreçlerini açıkladığını iddia etmeleridir.[88] Searle, ünlü Çin Odası düşünce deneyi ile Schank'ın programının ulaşmak istediği hedeflere ilişkin bir eleştiri sunmuştur. Çin Odası düşünce deneyini şu şekilde özetlenebilir: Bennu'nun bir odaya kapatıldığını düşünelim. Odanın herhangi bir camı veya dışarı ile temas kurulabilecek herhangi bir unsuru bulunmasın. Yalnızca odanın kapısında odanın içerisine zarf bırakılabilecek kadar bir boşluk bulunsun. Bennu ile bu boşluktan iletişim kurulsun. İlk durumda Bennu'ya içerisinde Çince bir metin bulunan bir zarf iletelim. Bennu'dan bu zarfı Türkçeye çevirmesini isteyelim. Bennu ilk durumda zarfın içerisindeki Çince metni anlayamayacaktır, dolayısıyla metni Türkçeye çeviremeyecektir. İkinci durumda

[86] Boden, "Artificial Intelligence".
[87] **A.e.**
[88] Searle, "Minds, brains and programs", s. 417.

Bennu'ya iki zarf verelim: İlk zarfta bir Çince metin bulunsun, ikinci zarfta ise metni çevirmek için genel kurallar Türkçe olarak yer alsın. Üçüncü durumda ise ilk iki yazma ile ilgili türlü semboller ve bu sembollere ilişkin yanıtların ne olacağını söyleyen üçüncü bir yazma Bennu'ya sunulsun. Burada Searle'e göre ilk yazma bir metni, ikinci yazma bir anlatıyı – yani bağlamı –, üçüncü yazma ise sorulara ilişkin yanıtları temsil etmektedir. Ve bunu sağlayan Türkçe* kurallara ise program denmektedir. Dolayısıyla bu örnekte Bennu tıpkı bir bilgisayar gibi davranmakta ve yalnızca biçimsel sembolleri manipüle ederek yine bir başka biçimsel semboller yığını olan Türkçeye dönüştürmektedir. Güçlü yapay zekâ düşüncesini savunan düşünürlere göre burada makinenin Çinceyi tıpkı insanların anladığı gibi anladığını söyleyebilmeliyiz. Oysa burada da olduğu gibi Schank'ın programı da hiçbir şeyi anlamamaktadır. Bilgisayara sağlanan biçimsel ilkeler her ne ise, onlar anlama yetisi için yeterli koşulları sağlamamaktadırlar. Çünkü insanlar her ne kadar anlama yetisine sahip olsalar bile bir biçimsel ilkeler dizisini, tıpkı bu düşünce deneyinde olduğu gibi, anlamadan da işletebilme olanağına sahiptirler. Searle'e göre anlamanın farklı seviyeleri söz konusudur. Bu anlamda bir otomatik kapının fotoelektrik hücresinden gelen emri anlaması ile bizim İngilizceyi veya Türkçeyi anlamamız aynı anlamda anlama olarak ifade edilemez. Dolayısıyla programlanmış bir bilgisayar, bir araba veya bir hesap makinesi için anlama ne ise o ölçüde bir anlama gerçekleşecektir. Searle bir adım daha ileri giderek bir makinenin kısmen ya tamamlanmamış bir biçimde değil hiçbir biçimde bir anlama yetisine sahip olmadığını savunmuştur.[89]

Searle'ün ortaya koyduğu Çin Odası argümanına karşı türlü eleştiriler dile getirilmiştir. Örneğin Yale Robot Yanıtına göre robotun içerisine bir bilgisayar yerleştirdiğimizi ve bu bilgisayarın yalnızca girdi çıktı olarak sembolleri işlemediğini, aynı zamanda bir robotu da algılama, yürüme, beslenme vb. işler için kontrol ettiğini düşünelim. Bu robotlar kolları ve bacakları olan ve bu sayede dünyamızda eyleyebilecek bir varlıklar olarak tasarlansın. Bu durumda Schank'ın

* Eleştirimiz boyunca yaptığımız gibi, düşünürlerin argümanlarını metinde düşünsel bütünlük sağlanması amacıyla "Bennu" veya "Nazım" gibi yarattığımız karakterler ve yerelleştirme üzerinden kurgulamaktayız.

[89] **A.e.**, s. 418-419.

programı anlama yetisine sahip kılınmaz mı? Searle'e göre robotun herhangi bir yönelimsel durumu bulunmamaktadır. Robotun eylemleri onun yalnızca elektriksel süreçlerinden, kablolamalarından ve sahip olduğu yazılım olan programlardan ileri gelmektedir. Berkeley/MIT Beyin Simülatörü Yanıtına göre ise Çince anlatıyı anlayabilen bir Çinlinin beyninde gerçekleşmekte olan nöral aktivitelere sahip olan paralel işlemeli, çok programlı ve nöral ateşlemelere sahip sinapslardan oluşan bir beyin tasarlanırsa bu durumda Çin Odası argümanının sonuçlarının ne olacağı sorulur. Ancak Searle, güçlü yapay zekâ yaklaşımının beynin nasıl çalıştığını merak eden ve anlamaya çalışan bir çaba olmaktan çok zihnin hangi fiziksel yapılarda gerçekleşebileceğine odaklanan bir yaklaşım olduğunu hatırlatır. Dolayısıyla zihnin beynin programı olduğu, beynin ise bu programın donanımı olduğu yolla bilgisayımsal-temsil işlevselciliğine dayanan güçlü yapay zekâ düşüncesi bakımından zihni nörofizyolojiye başvurmadan anlayabilmekteyiz. Üstelik beyin simülatörü yanıtının asıl sorunu, beyindeki nöral ateşlemelerin simüle edilmesi yoluyla beyindeki yönelimsel durumların da simüle edilebileceğini savunmaya çalışmasıdır. Oysa nöral ateşlemelerin simüle edilmesi zorunlu olarak yönelimsel durumların da simüle edileceği anlamına gelmez.[90] Yale Diğer Zihinler Yanıtına göre ise Çin Odası argümanının sonuçlarından biri, başka insanların Çinceyi veya başka bir şeyi anladıklarına ve anlayabildiklerine ilişkin güvencemizi ortadan kaldırmasıdır. Her şeyden önce Searle, burada tartışılan konunun başkalarının bilişsel durumlara nasıl sahip oldukları veya olmadıkları değil, biliş durumlarına sahip olmayı onlara ne zaman atfedebileceğimize ilişkin bir sorun üzerinde durduğunu belirtir. Bir Marslı varlık da beyninin yapısı veya materyali farklı olsa bile yönelimsel durumlara sahip olabilir. Bu durum, tıpkı fotosentez olayının klorofille veya farklı bir kimyasal tepkime ile gerçekleştirilebilmesine ilişkin bir durumdan çok farksızdır. Searle'ün vurguladığı ve eleştirdiği asıl nokta, saf biçimsel bir modelin yönelimsellik için asla yeterli olmayacağıdır. Çünkü biçimsel özelliklerin yönelimsel güçleri yoktur ve olan tek şey çalışmaya bir kere başladıklarında bir sonraki durum için biçimsel durumların üretilmesidir. Dolayısıyla *bir makine olarak bizler* de düşünürüz ancak bizlerin düşünmesi, yönelimsel durumlara sahip olabilmesi ve

[90] **A.e.**, s. 420-421.

anlamayı gerçekleştirebilmesi için organik materyalden meydana gelen nöronlara, dendritlere ve aksonlara sahip olmaya dayandırılmaktadır. Bir makine de tıpkı biz insanların sahip olduğu biyolojik materyale sahip olarak düşünmeyi gerçekleştirebilir.[91] O halde zihnin bir program olduğu yollu görüş üç açıdan eleştirilebilir: İlk olarak, bir program ile bu programın gerçekleşmesi arasındaki fark bir programın çok farklı şekilde ve hiçbir yönelimselliğe sahip olmayan sayısız biçimde gerçekleştirilebilmesidir. Örneğin Çinceyi anlamak üzere tasarlanmış bir program, bir bilgisayarda, bir dizi su borusu ya da rüzgârlı makinelerde çalıştırılabilir ancak bu yapıların da Çinceyi anlayacağının savunulması sağduyuya aykırı olur. Çünkü yalnızca beynin sahip olduğu nedensel güçlere sahip bir yapının yönelimselliğe sahip olacağını ileri sürmek oldukça doğaldır. İkinci olarak, bir program tamamen biçimseldir ancak yönelimsel durumlar benzer şekilde biçimsel bir yapıda değildir. Onlar biçimleri bakımından değil içerikleri bakımından ele alınırlar. Yağmurun yağdığı düşüncesi sembollerle değil kimi zihinsel içerikler vs. ile ifade edilir. Üçüncü olarak, zihinsel haller ve olaylar beynin ürünüdür ancak bir program bir bilgisayarın ürünü değildir. Dolayısıyla doğru girdi-çıktı ilişkilerini sağladığımızda gerçekleşen tek şey *simülasyondur*.[92] Searle'ün bu pozisyonu, bir yandan zihnin beyin ve merkezi sinir sistemimiz ile mevcut olan biyolojik ve fiziksel ilişkisini yadsımamaktayken bir yandan da fizikalizmin indirgemeci ve elemeci yaklaşımlarına karşı çıkmamasıyla uzlaşmacı bir duruş sergiler.[93]

Makine zekâsı ve insan zekâsı arasında kurulmak istenen ilişkinin sahip olduğu problemler yalnızca onların yönelimsel durumlara sahip olmamalarından ibaret değildir. Turing, bir yandan hesaplama kuramının da gücünü kullanarak ortaya attığı bilgisayar fikrini savunur ama bilgisayarların çözüm üretemeyeceği, yanıtsız kalacağı kimi durumların da mevcut olduğunu ortaya koyar. Bu durum, durma problemi[*] olarak anılmaktadır. Durma probleminin özü, bazı problemler ne kadar sıkı tanımlanmış ve tasarlanmış olsalar bile, bu problemlere yanıt üretebilecek bir algoritmanın mevcut olmamasıdır. Durma probleminde bir algoritmanın, belli bir

[91] **A.e.**, s. 422.
[92] **A.e.**, s. 422-423.
[93] Altınörs, "Zihin-Beden Problemine Searle'ün Yaklaşımı", s. 59.
[*] İng. **Halting Problem**

programın sonlanmasını engelleyen ve sonsuza kadar sürecek olan bir döngüde olup olmadığına karar verip veremeyeceğine ilişkin soru ele alınır.[94] Bilindiği üzere algoritmalar, bir problemi veya problemler sınıfını çözmek için oluşturulmuş komutlar kümesidir. Ve bir algoritma bir kez çalıştırıldığında, bu algoritmanın sonlanıp sonlanmayacağını, durup durmayacağını tayin edebilmek oldukça önemli bir husustur. Çünkü sonsuza kadar çalışan bir algoritma asla durmayacak, böylece bir algoritmanın durmasını çıktı olarak tanımladığımız durumda, sonsuza kadar çalışmakta olan bir algoritma asla çıktı üretemeyecektir. Burkholder'a göre durma problemi algoritmik olarak çözümsüzdür.[95] Durma probleminde, problemin çözümü için kurgulanan fonksiyona durma fonksiyonu denir. Durma probleminin farklı kurgulamaları söz konusu olsa bile çoğu durumda durma fonksiyonunun uygulandığı iki terim bulunur: Bir program ve bir girdi. Bu fonksiyon, bir program çalıştırıldığında ilgili girdiden hareketle o programın sonlanıp sonlanmayacağına karar vermelidir.[96] Durma problemini anlamak için bu problemi bir örnek üzerinden düşünelim: Bennu'nun önünde iki adet makine olsun. Birinci makine sayılar ona sağlandığında uzunluk birimlerini birbirleri arasında dönüştürsün. Bu makineyi U(x) gibi bir fonksiyon olarak analım. İkinci makine ise doğum tarihi bilgisi girildiğinde kişinin yaşını bir çıktı olarak bize sağlasın. Bu makinemizi ise D(x) fonksiyonu olarak analım. Burada açıktır ki U makinesine, D makinesine girdiğimiz girdileri verirsek U makinesi bir çıktı üretemeyecektir. Çünkü U makinesi yalnızca uzunluk dönüşümlerini sağlamak için tasarlanmıştır. Benzer şekilde yaş hesabı için tasarlanan D makinesine uzunluk bilgisi girersek, makinemiz bir çıktı üretemeyecektir çünkü makine doğum tarihi girdileri ile çalışmak üzere tasarlanmıştır. Durma problemine göre burada ikinci adım olarak H(x,y) gibi bir fonksiyonu düşünelim. Bu fonksiyon, burada dile getirdiğimiz U ve D fonksiyonlarımızın, doğru girdiler verildiğinde sonlanıp sonlanmayacağına karar veren fonksiyonumuz olsun. Böylece H(x,y) fonksiyonumuzdaki (x) makine tipini, (y) ise girdi tipini temsil etsin. Bu durumda H makinemizin olası dört durumu olacaktır. H(U, uzunluk girdisi) durumunda makinemiz "durur" yanıtını versin. H(U, doğum tarihi bilgisi) durumunda makinemiz

[94] Russell, **İnsanlık için Yapay Zekâ: Yapay Zekâ ve Kontrol Problemi**, s. 54.
[95] Leslie Burkholder, "The Halting Problem", **ACM SIGACT News**, C. XVII, No: 3, 1987, s. 48.
[96] Hehner, "Problems with the Halting Problem", s. 31-32.

“durmaz” yanıtını versin. H(D, uzunluk bilgisi) durumunda makinemiz “durmaz” yanıtını versin. H(D, doğum tarihi bilgisi) durumunda ise makinemiz “durur” yanıtını versin. Görüleceği üzere uygun fonksiyona uygun girdilerin verildiği durumda fonksiyonlar, yani makinelerimiz, duracak ve bir yanıt üretmiş olacaktır. Aksi durumda ise fonksiyonlarımızın, yani makinelerimizin, durmayacağını ve sonsuza kadar çalışacağını söyleyebiliyoruz. H fonksiyonumuz bize makinelerin sonsuza kadar çalışıp çalışmayacağını söyleyen, bunları tayin eden bir makinedir. Durma probleminin özü, Turing ilkelerine göre çalışan bir H makinesinin bulunmadığıdır. Bu durumun ispatı açıktır. Diyelim ki H gibi bir makine mevcuttur ve gerçekleştirilebilir. Şimdi N gibi bir makinemiz olsun ve bu makine, ona verilen girdilerin tersi yönünde çıktılar türetsin. Bu durumda H makinemize kendisini girdi olarak sağlarsak ve H makinemizin sonuçlarını N makinesinde işletirsek bu durumda sonuç ne olur? İlk durumda H(H,H) fonksiyonuna göre makinemiz çalıştırılırsa H makinemizin “durmaz” yanıtına göre N makinemiz “:)” yanıtını döndürsün. Böylece H makinemizin yanıtı makinenin sonsuza dek çalışması gerektiği olmuşsa da N makinesinin “:)” yanıtı H makinesinin yanıldığını göstermektedir. İkindi durumda ise H makinesinin durur yanıtı ürettiğini kabul edelim. Böylece N makinesi yanıtı tersine çevirir ve N makinesi H makinesinin durmayacağı sonucunu verir. Bu durumda H makinesi yine yanılmış olur. Bu durum bize gösterir ki bir programın verilen girdilere göre çalışıp çalışamayacağını tayin edebilecek bir program yazılamaz. Turing makinelerinde, Turing’in ilkelerine göre çalışmakta olan bir durma fonksiyonu tasarlanamaz.[97]

Durma probleminin sonuçları, eleştirimizin amaçlarına hizmet etmek bakımından yapay zekâ felsefesi ve yapay zekâlı varlıkların fail statülerinin tayin edilmesi ve zihin durumlarının tespiti konusunda büyük önem taşır. Açıktır ki biz insanların sahip olduğu kendimizi kendi nesnemiz edinebilme ve böylece kendimiz üzerine düşünebilme konusundaki becerilerimiz olarak adlandırabileceğimiz diğer tüm benzer zihinsel süreçlerimiz bir makine tarafından gerçekleştirilebilir gibi durmamaktadır. Çünkü H fonksiyonu üzerinden görüldü ki bir makine kendi düşünmesini kendi nesnesi yapamamaktadır. Böyle bir fonksiyon, makinenin var

[97] Talha Zekeriya Durmuş, Sonlanma Problemi, (Çevrimiçi) https://talhadurmus.com/bilgisayar-bilimi/sonlanma-problemi, 26.06.2022.

oluşu açısından bir paradoksun doğmasına yol açmaktadır. Bu açmazın nasıl aşılabileceği ise bu aşamada felsefenin değil ama mühendisliğin alanına girmektedir. Durma probleminin bir sonucu ise Ayral'ın da ifade ettiği üzere doğal sayılar kuramımızın tamlığı ve tutarlılığı bakımından da önemli olan, doğal sayılar üzerine olan yargılarımızın doğruluğunu sınayacak genel bir algoritmanın olmadığını göstermesidir.[98] Böylece aslında bu eleştiri içerisinde de dile getirdiğimiz ve yapay zekâ felsefesi ile yapay zekâ çalışmalarının başlangıcı olarak Turing'in değil ama matematik bilimi bakımından düşünülürse David Hilbert'in çalışmaları olarak düşünülmesi gerektiği ile ilgili önerinin bir gerekçesine ulaşmış bulunuyoruz: Hilbert'in tam ve tutarlı bir matematik sistem inşa etme arzusunun bir sonucu olarak böyle bir fonksiyonun arzulanması, yapay zekâ çalışmalarının temel bulması ve yol göstericilik açısından önem taşır. Ancak ilginçtir ki Turing, böyle bir problemin çözümsüz olacağını göstermiştir. Dolayısıyla durma probleminin varlığı matematiğin temelleri açısından da önem taşır. Çünkü matematik de tam ve tutarlı olması arzulanan, sıkı bir güvencenin inşa edilmeye çalışıldığı bir sistemdir ve matematiğin durma problemi gibi bir problemi aşmak konusundaki yetersizliği matematiğin temellerinin sorgulanmasına neden olmuştur. Fakat tüm bu durumlara rağmen Russell gibi bilgisayar bilimciler, durma probleminin yapay zekâlı varlıkların ya da bilgisayarların zeki olmaları konusunda herhangi bir engel yaratmadığını savunmaktadır. Buradaki gerekçe, benzer sınırlamaların insan beyni için de söz konusu olmasıdır.[99] Nitekim burada insan doğasının sezgisel mahiyeti yadsınmıyor mu? Matematik felsefesi bakımından sezgicilik önemli bir rol oynamaktadır ve insan doğasındaki yaratıcı ve buluşçu unsurların sezgi söz konusu olmasa da ortaya çıkabileceğini savunmak için güçlü gerekçelerimiz bulunmamaktadır. Böylece bir makine zekâsı ile bir insan zekâsı arasındaki nitelik farklarını yadsımamak gerektiği açıktır. Çünkü bir makine için kendi *birliğini* sağlamak onun için özsel bir durumsa, insan için böyle bir zorunluluk bulunmamaktadır. Makine, doğası gereği *bir şey için* iken bir insan bir şey için değil ama amaçsız bir varlık olarak duygularıyla, sezgileriyle ve diğer tüm zihinsel olanaklılıklarıyla bir bütündür. Bu bütünlüğün

[98] Hakan Ayral, "Halting Prediction on Busy Beaver Type Turing Machines Based on Information Entropy", İstanbul, Galatasaray Üniversitesi Fen Bilimleri Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2008, s. 6.
[99] Russell, **İnsanlık için Yapay Zekâ: Yapay Zekâ ve Kontrol Problemi**, s. 54-55.

nöronal temelde nasıl sağlandığına ilişkin insan bilgisinin sınırları gün geçtikçe gelişmektedir fakat bu eleştirinin amaçlarından biri böyle bir bilginin sağlanması değildir. Ancak bu durum yapay zekâ felsefesi açısından önemli bir itirazın ortaya çıkmasını sağlar: Durma problemi yapay zekâlı varlıkların fail olabileceğine ilişkin sava güçlü bir yanıt niteliği taşımaktadır.

Bilgisayar bilimi ve yapay zekâ çalışmaları bakımından bilgisayarların oynadığı rol ve bu yapıların temelleri düşünüldüğünde, şimdiye kadar ortaya koyduğumuz kavramsal çerçeve ne ifade eder? Her şeyden önce bu eleştirinin böyle bir alt incelemeye sahip olmasının gerekçeleri açıktır: Zihin ve beynin işlevlerini anlamak için en iyi modelin Turing veya von Neumann mimarilerine dayanan mantıksal temellere göre inşa edilmiş makineler olarak dijital bilgisayarlar olduğunu savunan, dolayısıyla bu metafora göre inşa edilmiş bir bilişsel psikolojinin amacının beyinde çalışan bir program olarak zihni anlama çabası olduğunu ileri süren Churchland türü bir bilgisayımsal-temsilci işlevselcilik söz konusudur.[100] Putnam'ın bir Turing makinesinin soyut bir makine olarak sonsuz sayıda farklı yolla fiziksel olarak gerçekleştirilebilir bir makine olduğu yollu vurgusu da böyle bir işlevselcilikle örtüşmektedir.[101] Bu tür bir işlevselcilik, sonsuz sayıda gerçekleştirilebilirlik düşüncesinin de verdiği güvenceyle birlikte, insan gibi düşünme kapasitesine sahip *bir makinenin* ne tür özelliklere sahip olması gerektiği yollu kurgulara da olanak sağlamıştır: Nilsson'a göre makineler insan gibi düşünmeyi gerçekleştireceklerse ikili değil bulanık mantığa sahip olmalı, paralel işlem gücüne sahip yapılar olarak gerçekçi yapay nöronlarla desteklenmelidir.[102] Say'a göre ise hesaplama yapabilen her sistemi bilgisayar olarak adlandırmanın insanların da bilgisayar olduğu yollu görüşü dile getirmekten bir farkı olmadığı açıktır. Bilgisayarlarımız insanları birebir taklit edebilecek en uygun adaylardır. Ancak Say, taklidin yalnızca söz ve davranış gibi davranışçılık açısından önemli görülen durumların taklit edilmesini hedeflediğini dile getirmektedir. Böylece davranışsal olarak birbirini taklit edebilen iki farklı varlık olarak bilgisayarlar ve insanlar arasında bir astlık ve üstlük ilişkisi de kurulamaz. Çünkü insanlar da tıpkı bilgisayarlar gibi, duyargalarından gelen verileri

[100] Churchland ve Sejnowski, "Neural Representation and Neural Computation", s. 352.
[101] Putnam, **Mind, Language and Reality: Philosophical Papers**, s. 371.
[102] Nilsson, **Artificial Intelligence: A New Synthesis**, s. 4.

işleyen, fizik kurallarına göre çalışan ve hesaplama yapan ancak et ve kemikten oluşan bir bilgisayardan fazlası değildir. Dolayısıyla insan zihninde bulunan düşünceler, bilgiler, şarkılar vb. gibi fenomenal tüm unsurlar da bilgisayarlarımıza çalıştırılacak programlar vasıtasıyla *makineye can katan* bir unsur olarak gerçekleştirilebilir.[103] Bu anlamda Say için beyin bir bilgi işleme makinesinden daha fazlası değildir: Beyin tam anlamıyla bir bilgisayardır. Ona göre beyinlerimizin mimarilerinin mevcut dijital bilgisayarlarımızın mimarileriyle karşılaştırılması yoluyla *bu gerçeğin sindirilmesi*[*] gerekmektedir. Öte yandan Turing için ise insan zekâsının potansiyeli uygun bir eğitimle sağlanabilir.[104] Çünkü insan zekâsının makinelerde gerçekleştirilmesi çabası ilk adımda yetişkin bir zihnin gerçekleştirilmesi üzerinden değil ama bir çocuğun zihnini simüle eden bir programın eğitilmesi üzerinden başarılabilir. Eğitim ile kazandırılacak olan deneyimlerden hareketle çocuk zihnini simüle edebilecek bir program zamanla yetişkin zihni düzeyine erişecek ve insan zekâsı makinelerde gerçekleştirilmiş olacaktır.[105] Churchland'lara göre ise Turing'in "Makineler düşünebilir mi?" sorusu fiziksel sembol sistemler düşüncesinin ortaya atılmasıyla birlikte bu mimariye dayanan makinelerin düşünüp düşünmediği ile ilgili bir soruya evrilmiştir. Dolayısıyla makineler düşünebilir mi sorusuna "evet" demek için Church'un

[103] Say, **50 Soruda Yapay Zekâ**, s. 43-46.

[*] Felsefi bir konuşmayı spekülatif tartışmalardan ayıran niteliklerden biri, felsefi bir konuşmada tarafların kavramların kullanımı ve argümanın dile getirilmesi bakımından öne sürülen ifadelerin seçimi konusunda bir hassasiyete sahip olmalarıdır. Felsefi bir konuşmada taraflardan birinin, dile getirdiği argümanın veya savın kuvveti konusundaki ısrarı ancak ve ancak sıkı bir temellendirme ile pekişir. Dolayısıyla Say'ın hem bu eleştiri boyunca yer yer dile getirdiğimiz kullanımları bakımından hem de yapay zekâ felsefesi özelinde tartışmaları ele alış tarzının hiç de felsefece olmadığını burada vurgulamak istiyoruz. Eleştirimizin de odağında olan ve bizce yapay zekâ felsefesindeki açmazların ortaya çıkması bakımından oldukça hatalı bulduğumuz bu ve benzeri *mühendisçe bakmalar*, eleştirimizde şimdiye kadar ortaya koyduğumuz tartışmaları her şeyden önce bir felsefi tartışma olarak kabul etmiyorlar, tam aksine bir mühendislik sorunu olarak kabul ediyorlar. Böylece metafizik kurgular, mühendisçe bir temele oturtularak *sindirilmesi gereken gerçekler* olarak ileri sürülebiliyor. Üstelik, *sindirme* gibi bir ifadeyi bir kenara koymayı ve Say'ın kullanımını naif bir şekilde ele almayı seçsek bile gerçekliğin kabul edilmesi olarak düşünüldüğünde burada Say'a basitçe şu sorular yöneltilebilir: Hangi gerçeklik? Kimin gerçeklik tasarımı? Hangi düşünürün gerçeklik anlayışına dayanılıyor? Burada örtük bir gerçeklik anlayışı varsa bunun ontolojik ve epistemolojik temelleri nedir? Buradaki sorular gibi birçok soru Say'a yöneltilebilir. Gerekçe açıktır: Biz felsefede kavramları gelişigüzel kullanmıyoruz. Felsefeciler kavramların hesabını vererek argümantasyon yapmak durumundalar. "Gerçeklik" gibi felsefe alanının temel kavramlarından birini, sırf bir görüşü dile getirmenin yüzü suyu hürmetine kullanma alışkanlığının ülkemiz akademik tartışmalarından çıkarılması gerektiğini ısrarla vurguluyoruz.

[104] Turing, "Intelligent Machinery", s. 107.

[105] Turing, "Computing Machinery and Intelligence", s. 455-456.

hesaplanabilirlik ile ilgili kuramı ve Turing'in evrensel Turing makineleri ile ilgili görüşleri gibi birçok unsur bulunur.[106] Böylece bilgisayarların insan zihnini modelleyebileceğine ilişkin tüm bu düşünceler, insan zihninin basit hesaplama cihazları olarak hiçbir *büyülü* özelliğe sahip olmadığı yollu bir görüşe temel sağlamıştır.[107]

Bilgisayımsal temelde şekillenen işlevselci yaklaşıma ilişkin eleştirilerin elenmesi amacıyla kimi güncel bilimsel gelişmelerin, bilgisayar bilimi içerisinde işlevselci savların gücünü ve başarısını pekiştireceğine ilişkin bir görüş bulunur. Kuantum kuramının, insan zihinselliğinin makinelerde başarılması konusundaki engellerin aşılmasında önemli bir kuram olduğu savunulur. Bit kavramımız, kuantum işleme süreçlerinde "**qubit**" kavramıyla karşılanabilir. **Qubit**, klasik bitlerin 0 ve 1 gibi durumlarından ziyade süper pozisyonel olarak aynı anda hem 0 hem de 1 değerlerini alabilmektedir. **Qubit**'in herhangi bir andaki durumu iki boyutlu bir Hilbert uzayında bir vektör olarak tanımlanabilir. Klasik bitlerin gözlemden bağımsız durumlarından ziyade bir **qubit**'in gözlemlenmesi aynı anda 0 ve 1 durumlarında bulunan söz konusu **qubit**'in, onu oluşturan bu durumlardan birine çökmesine sebep olmaktadır. Böylece gözlemlendiği anda bir **qubit** 0 ve 1 gibi değerlerden birini temsil etmektedir. Bitlerden ziyade **qubit'**lerin bu olanağı, bilgi işlemenin gücü ve esnekliği bakımından büyük avantajlar sağlamaktadır.[108] Kuantum hesaplama süreçlerinin işlevselci savlar bakımından önemi mantıksal kapılara, yalnızca 0 ve 1 değerlerine sahip olunabilen bir mimari gibi katı ve kesin bir yapıdan ziyade, insan zihni gibi esnek ve çok değerli bir yapının koşullarını sağlamak konusunda bir imkânı bize sunmasıdır. Çünkü insan zihninin herhangi bir anda yalnızca 0 veya 1 gibi değerleri alabilen işlem birimlerinden ibaret olduğunu söylemek, herhangi bir girdi için sürekli aynı çıktının alınabileceğinin iddia edilmesinden farksızdır. Böylece kuantum bilgisayarların sunduğu durum, verili bir girdiye karşılık aynı çıktının

---

[106] Paul M. Churchland ve Patricia Smith Churchland, "Could a Machine Think?", **Scientific American**, C. CCLXXII, No: 1, 1990, s. 32.
[107] Kevin O'Regan, "How to Build a Robot that is Conscious and Feels", **Minds & Machines**, No: 22, 2012, s. 120.
[108] Amit Hagar ve Michael Cuffaro, "Quantum Computing", The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta ve Uri Nodelman, (Çevrimiçi) https://plato.stanford.edu/archives/fall2022/entries/qt-quantcomp/, 20 Aralık 2022.

alınamayacağı karmaşık durumların sağlanmasıdır.[109] Kuantum bilgisayarların bu olanaklılıkları, insan zihninin makinelerde gerçekleştirilebileceğine ilişkin düşünceleri güçlendirmişse de kuantum bilgisayarların zihinsel durumların ya da gerçekliğin simülasyonu konusundaki başarılarına ilişkin kimi itirazlar söz konusudur. Örneğin Feynman'a göre bilgisayarlar, fiziksel olayları tıpkı doğada yaşandıkları gibi simüle edemezler.[110] Çünkü bir kuantum sistem, verili bir durumdaki tüm sayısal değerlerin hepsine vakıf olamaz. Olamadığı gibi de böyle bir sistemin sonuçları bir evrensel Turing makinesinde temsil edilemez. Dolayısıyla fiziksel bir fenomen, bir bilgisayar tarafından simüle edilemez.[111] Bu durumda insan beyninde gerçekleşmekte olan ve insan zihnini olanaklı kılan fiziksel süreçler olarak nöral ateşlemeler bir bilgisayarda, *doğada ve şu anda gerçekleştikleri gibi* tam anlamıyla simüle edilemezler. Ancak bu durum, bilgisayarlar ve simülasyonlar yoluyla insan zihninin yapısını anlamayı ve insanın ruh durumlarına ilişkin çalışmaların önemini yadsımak anlamına gelmez. Çünkü Dennett'a göre, her ne kadar bir kasırga simülasyonunun bizi ıslatacağını düşünmemiz için bir gerekçe bulunmuyor olsa da iyi bir simülasyonun oynadığı rol, bir fenomene ilişkin sahip olduğumuz bir kuramın sonuçlarını görebilmemizi ve kuramımızdaki tutarsızlık ve eksiklikleri tespit etmemizi kolaylaştırmalarıdır.[112] Dolayısıyla her ne kadar bilgisayar programları, fenomenlerin simüle edilmesi bakımından imkânlar sağlasa da Gödel ve Turing'in de ileri sürdüğü ve bu eleştirinin de merkezindeki savlardan biri olarak, yalnızca insanların çözebileceği ve tasarlayabileceği kimi problemlerin ve yalnızca insanların sahip olabileceği kimi zihinsel durum ve duyguların bulunduğunu kabul etmek – her ne kadar Turing, Gödel ile bu noktada uzlaşsa da yine de bir bilgisayarın düşünebileceğine ilişkin inancını terk etmemiştir – oldukça doğaldır.[113] Çünkü beyinlerimiz ve onların temel birimleri olarak nöronlar, birbirinden ayrık işlevlere sahip değildir. Onlar hem depolama hem de işlem birimi

[109] Doğan, "Yapay Zekâ ve Özgür İrade: Yapay Özgür İradenin İmkânı", s. 799.
[110] Richard P. Feynman, "Simulating Physics with Computers", **International Journal of Theoretical Physics**, C. XXI, No: 6/7, 1982, s. 468.
[111] **A.e.**, s. 476-477.
[112] Daniel C. Dennett, "Why You Can't Make a Computer That Feels Pain", **Synthese**, No: 38, 1978, s. 421-422.
[113] Crevier, **AI: The Tumultuous History of the Search for Artificial Intelligence**, s. 23.

olarak çalışmaktadır.[114] Beyinlerimizle bilgisayarlar arasındaki bu mimari farklılık kimi zihin durumlarının yalnızca bize özgü olarak düşünülmesi için önemli durumlardan biri olarak kabul edilebilir.

İnsan zihninin makinelerde gerçekleştirilmesine ilişkin arzular yalnızca gerçekliğin simüle edilebilirliği ve makinelerdeki simülasyonların statüleri üzerinden değil aynı zamanda Reeke ve Edelman'ın da ifade ettiği beş temel husus üzerinden de eleştirilere maruz kalmaktadır: İlk olarak bir yapay zekânın tasarlanmasında bir programcı, söz konusu yapay zekâya gerçeklikle ilgili tüm kategorileri sağlamak durumundadır. Kategoriler ve kategoriler ile ilgili kuralları sağlamak konusunda bir tamlık sağlanmış değildir. Kodlama problemine göre bir programcı, bilginin temsil edilebileceği uygun bir temsil modeli bulmalıdır.[*] Gerçek dünyada ve çevrede eyleyecek varlıklar olarak yapay zekâların, prosedür problemine göre bir programcı tarafından olası tüm davranış çıktılarının tanımlanması gerekmektedir. **Homunculus** problemine göre biçimsel olarak hiçbir anlama sahip olmayan komutların, bir anlam taşıması ve bu anlamın kavranması bakımından bir yapay zekânın kendisini bir nesne edinebilmesi gerekmektedir. Gelişim problemine göre ise bir programlanmış sistemin bir programcıya ihtiyaç duymadan var olup olamayacağı sorulmaktadır.[115] Minsky'ye göre ise yapay zekâların karşılaşacakları problemlerin üstesinden gelebilmeleri için problemleri beş temel bölümde ele almaları gerekmektedir: Araştırma, örüntü tanıma, öğrenme, planlama ve indüksiyon.[116] Ancak eleştirimizde de vurgulanmıştır ki mantık temelli yapay zekâ çalışmalarının zayıf yapay zekâ anlayışında olduğu gibi kimi belirli işler bakımından gösterdiği başarı bir kenara, güçlü yapay zekâ çalışmalarının arzuları bakımından istenilen temeli sağlamadığı ve sağlayamayacağı açıktır. Çünkü Dreyfus'un da ileri sürdüğü gibi, gerçeklikte yer alan olguların, nesnelerin vb. bitler olarak makinelere sunulabileceğine ilişkin bir inanç bulunsa da[117] *bilginin en ufak birimleri olarak bitler* düşüncesine ilişkin henüz

---

[114] **A.e.**, s. 300.

[*] Buradaki itiraz bu eleştirinin mantık ve matematik bilimini ele alan alt incelemelerinde detaylı bir şekilde ortaya konmuştur.

[115] George N. Reeke Jr. ve Gerald M. Edelman, "Real Brains and Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 153-154.

[116] Marvin Minsky, "Steps Toward Artificial Intelligence", **Proceedings of the IRE**, C. XLIX, No: 1, 1961.

[117] Dreyfus, "Bilgisayarların Akıllı Olması İçin Neden Bedenleri Olması Gerekir?", s. 46.

sıkı bir temellendirme bulunmamaktadır. Dolayısıyla bir makinenin sahip olduğu ama kendisini oluşturan yargılardan hareketle doğruluğunu veya yanlışlığını ortaya koyamadığı en az bir yargının olduğu da düşünülürse, bir insanın bunu yine kendi var oluşuna dayanarak ortaya koyma becerisini yadsıyamayız. Yadsıyamadığımız için ise bir zihnin yapabileceği ancak bir makinenin yapamayacağı durumların söz konusu olduğunu ileri sürebiliriz.[118] Böylece zayıf yapay zekâ düşüncesinde olduğu gibi zihnin belirli işlevlerinin başarılabilmesi işi için tasarlanabilecek makinelerin başarılarını yadsımıyor ancak zihnin tümünü gerçekleştirebilmeye ilişkin bir düşünce olarak güçlü yapay zekâ düşüncesinin başarılarını henüz olanaklı bulmuyoruz. Çünkü Searle'ün de vurguladığı gibi bilinç, düşünceler, duygulanımlar, duygular ve zihinsel olan her durum sentakstan ötedir. Bilgisayarların sentaktik yapıları vasıtasıyla taklit edemeyeceği, sahip olamayacağı ancak yalnızca simüle edebilecekleri durumlar olarak bu zihin durumları biz insanların sahip olduğu organik bir yapı olan beynimizin bir ürünüdür. Tamamen biçimsel ve sentaktik bir yapı olan bilgisayar programları, zihinsel içeriklere sahip olamaz.[119] Dolayısıyla saf hesaplama gücüne dayanarak bilgisayarların insan beyni kadar güçlü olduğunu iddia etmek olanaklı değildir.[120]

Yapay zekâ çalışmalarının bilgisayar bilimine dayanan temelleri üzerine burada gerçekleştirdiğimiz inceleme, eleştirimizde felsefi işlevselciliğin program ve bilgisayar üzerinden insan zihnini anlama ve onu makinelere indirgeme girişiminin felsefece ele alınması bakımından önem taşır. Çünkü felsefi işlevselciliğin bilgisayımsal temellere dayanan görüşlerinin merkezinde yer alan iddiaların bilgisayar biliminin temelleri ve bilgisayarlar bakımından da karşılanabiliyor olması gerekmez mi? Arf'ın da ifade ettiği üzere bilgisayarlar ve insanlar arasında bir varlık farkı bulunur. Bu fark, makinelerin özgür irade sahibi olmamalarına ve estetik mahiyette kararlar alamamalarına dayanmaktadır. Arf da kuantum bilgi işleme süreçlerinin bilgisayarlara kazandırılması durumunda insan düzeyinde estetik mahiyetlere sahip bilgisayarların elde edilebileceğine ilişkin ümidin varlığını

[118] Lucas, "Minds, Machines and Gödel", s. 115.
[119] Searle, **Minds, Brains and Science**, s. 37-40.
[120] Russell, **İnsanlık için Yapay Zekâ: Yapay Zekâ ve Kontrol Problemi**, s. 53.

yadsımaz ancak bunu bir olasılık olarak kabul etmez.[121] O halde burada ortaya koyduğumuz alt incelemenin amaçları netlik kazanmıştır: Felsefi işlevselciliğin bilgisayımsal temelleri, yapay zekâ felsefesi bakımından kimi argümanların güvenilirliğinin sağlanması bakımından şimdiye kadar dile getirdiğimiz sorunlarla yüzleşmek durumundadır.

3.3. Felsefi İşlevselciliğin Sonuçları Bakımından Yapay Zekâ ve Sanat İlişkisi Üzerine Bir Deneme

Sanat eseri nedir? Bir üretimin sanat eseri niteliği kazanmasının koşulları nelerdir? Eleştirimizin amaçları düşünüldüğünde sanat felsefesi alanındaki tartışmalar hakkında derin bir felsefi incelemede bulunmak istemediğimizi belirtmemiz gerekir. Ancak okuyucu açısından sanat eserinin neliğini burada soru konusu etme gerekçelerimizi açık kılmamız gerekmektedir: Güçlü yapay zekâ projesiyle birlikte düşünüldüğünde bilgisayımsal-temsil işlevselciliğine dayanan, beyni bir bilgisayar ve zihni bu bilgisayarda çalışan bir program olarak tasarlama arzusuna sahip olan felsefi işlevselci argümanın iddiaları, sanat felsefesi bakımından ve insanın varlık koşullarından biri olarak sanatın olanaklılıkları bakımından türlü sorunlara yol açmaktadır.

Her şeyden önce bir sanat eseriyle, o sanat eserinde konu edinilen şeyler olarak gerçeklik tasarımımızda bulunan nesneler ve onlar arasındaki ilişkiler düşünüldüğünde, kendimizi *insan doğasına özgü bir yaratımın* karşısında buluruz. Tanrı'ya yakarmakta olan herhangi bir çift el ile, Dürer'in "Dua Eden Eller" eserinde Tanrı'ya yakarmakta olan eller arasında ya da kahvehanede vakit geçirmekte olan sıradan bir insan ile Nurullah Berk'in eserlerinde ortaya konan insan arasındaki farka dikkat çekmek istiyoruz. Sanat eserinin ortaya çıkması, herhangi bir matematiğin ya da biçimsel herhangi bir dilin ya da sistemin bir sonucu olarak düşünülemez. Çünkü sanat eseri belirli bir anda *belirli bir duygulanımın, orada ve amaçsız bir yaratımın uğruna* sahip olunan bir etkilenimin neticesinden daha fazlası değildir. Felsefi işlevselci argümanın ve güçlü yapay zekâ düşüncesinin temel iddialarından biri,

[121] Arf, "Makineler Düşünebilir mi ve Nasıl Düşünebilir?", s. 103.

gerçeklik tasarımlarımızın bilginin en temel birimleri olarak kabul edilebileceği semboller ve biçimsel ilişkiler olarak ifade edilebileceği düşüncesidir. Buradan hareketle tüm gerçekliğin belirli fizik yasalarınca düzenlenmiş, matematiksel olarak açıklanamayacak herhangi bir unsuru bulunmayan bir apaçıklıkla karşımızda bulunduğu da ileri sürülmektedir. Eleştirimizde de dile getirildiği üzere matematiği ve mantığı temel kabul eden felsefi işlevselci yaklaşım, gerçeklik tasarımlarımızın tümünü eksiksizce kapsayabileceği böylece yalnızca işlevsel ilişkilere indirgenmiş olan zihnin makinelerde de gerçekleştirilebilecek kadar yalın bir yapıda olduğu görüşü üzerinde yükselmektedir. Et ile kemikten arındırılmış ve zihin sahibi yapılmış bilgisayarların, programlarıyla bir bütün olarak kabul edildiği yapay zekâlı varlıkların şiir yazabilecekleri, resim yapabilecekleri ve hatta sanatın her alanında insandan aşağı kalmayacakları ısrarla savunulmaktadır. Peki yapay zekâlı varlıkların varlık koşulları eleştirimiz içerisinde dile getirdiğimiz hususlar bakımından düşünülürse, insanın başarılarından biri olarak sanatın yapay zekâlı varlıklar için de mümkün olduğunu söyleyebilir miyiz?

Güzelin ne olduğu, güzel olanın hangi ilkelere göre belirleneceği sanat felsefesi ve sanat çevreleri bakımından tartışılagelen sorulardan yalnızca birkaçıdır. Sanat eserinin niteliğini ve sanat eserleri arasındaki nitelik farkını tayin edebilmek için öne sürülmüş sayısız önerinin, sayısız kuramın olması şaşılacak bir durum değildir. Çünkü neyin sanat eseri olup olmadığı sorusu hiç de önemsiz bir soru değildir. Belirli bir duygulanımın değil ama bir olanaklılık olarak duygulanımların, sanat eserini var eden zorunlu koşullardan biri olduğunu kabul etmek durumundayız. Çünkü sanat eserlerini yalnızca onları oluşturan belirli bir kompozisyon, malzeme bilgisi ve tekniğin ürünü olarak sınırlandıramayız. Sanat eserleri onlarda *içkin bir anlamın*, gerçekliğin bir yorumunun, algılayanda hissettirilmek istenen bir durumun taşıyıcısıdır. Ve felsefi işlevselci argüman bakımından düşünüldüğünde sanat eserinin yaratım süreci, bu işlevleri yerine getiren zihinsel süreçlerden ibaret bir yaratım olarak düşünülebilir. Böylece sanat eserinin yaratım süreci *sterile edilerek* sayılara, transistörlere ve en temelinde kimi koşullu ifadelere yerini bırakabilir. İşlevselciliğe göre belirli koşullar yerine geldiğinde sanatsal üretime temel olan duygulanımların gerçekleşeceğini, örneğin Dürer'in ya da Berk'in aynı koşullar

gerçekleştiğinde yine aynı eseri yaratabileceğini savunabiliyor hale gelmekteyiz. Oysa bu durum sezgilerimize oldukça aykırı durmaktadır. Gerçekliğin *o andaki duyumsanmasının*, duygularla ve insanın tüm fenomenal deneyimleriyle bir bütün olduğunu düşünmemek için bir gerekçemiz bulunmamaktadır. İşte sanat eserinin ortaya çıkabilmesini sağlayan önemli varlık koşullarından biri duygulardır. Duygular, gerçeklik tasarımlarımızı, gerçekliği algılayışımızı etkileyen ve dönüştüren fenomenlerdir. Farklı zamanlarda sürmekte olan deneyimimiz bir nesneye ilişkin farklı etkilenimlere ve farklı sınıflandırmalara sahip olmamakta mıdır? Örneğin yaşamının bir noktasında satranca küsmüş bir insanın satrancı yeniden sevmeye başladığına şahit olamaz mıyız? Benzer şekilde bir sanat eserinin yaratım sürecinde de gerçekleşen şeyin tüm bu duygularla ve duygulanımlarla, şeylerin *o anda ve burada* algılanmalarıyla ortaya konan bir yorumlama olduğunu, onda sanatçıya özgü bir düşünmenin içkin olduğunu kabul ediyoruz. Dolayısıyla felsefi işlevselci argüman neticesinde işlevlerin yalıtılıp, güçlü bir yapay zekânın inşa edildiği iddia edilse bile bu varlığın hiçbir zaman bir sanat eserinin yaratıcısı olabileceğini düşünmüyoruz. Okuyucu sorabilir: Konçertolar, resimler ve şiirler üreten yapay zekâlar yok mu? Bu yapay zekâların ortaya koyduğu işleri yadsıyor muyuz? Her şeyden önce bir sanat eserinin sanat eseri olmaklığını, *onda içkin olan bağlamda* aramalıyız. Yalnızca notalardan, dörtlüklerden ya da yağlı pastellerden oluşan bir eserin bir *sanat eseri* olduğunu kabul edemeyiz. O eser pek tabii bir şeyler anlatır bize, bir şeyleri resmeder ya da kimi seslerden oluşur. Ancak mühim olanın herhangi bir renk, figür, kafiye ya da nota değil, *o renk, o figür, o kafiye* ve *o nota* olabilmek olduğunu vurgulamak istiyoruz. Çünkü matematik bilimi üzerinden yapay zekâ felsefesi bağlamında yapay zekâ çalışmalarına ilişkin ortaya koyduğumuz saptamalar ve eleştiriler anımsanırsa, mühim olanın *şeylerin sembolleştirilmesi ve bu semboller arası ilişkilere indirgenmesi* olmadığını unutmamalıyız. Çünkü gerçeklik tasarımlarımızı yalnızca sembollere ve semboller arası ilişkileri düzenleyen kurallara indirgediğimiz durumda gerçekliği, bizim gerçeklik tasarımımızdaki gibi hiçbir şeyi dışarda bırakmayacak tarzda temsil edebilmenin koşullarını sağlayabilmemiz gerekiyor. Ancak bu durumun mümkün olmadığı matematik felsefesi tarihi ve matematik ile mantık bilimleri temelinde incelendi. Benzer şekilde renklere, figürlere, kafiyelere veya notalara uygulamaktan kaçınmamız gereken bu düşünce

tarzı, bize sanat eserlerinin de işlevselci tarzda analiz edilebileceğini ve biçimselleştirilebileceğini düşündürebilir. İşte okuyucu açısından asıl dikkat çekmek istediğimiz nokta burasıdır: Sanat eserinin içeriği birimlerine, materyaline ve sembollerine indirgenemez. O belirli bir duygunun, bir zamanın ve koşulun sonucu olarak vardır. Anlamı yorumlayıcısı ile bir bütündür.

Sanat eserinin anlamının doğurulmasında ve sanat eserinin yaratılmasında duygu önemli bir rol oynamaktadır. Duygu nedir? Duygunun ne olduğu şimdiye değin tanımlanamamış bir hususdur. Pratik kaygıların giderilmesi amacıyla kimi tanımlar kabul edilmişse de felsefi bir akıl yürütmenin doğası gereği ele alındığında "duygu" kavramının ne olduğu açık değildir. Duygunun, bir duruma ilişkin bir yanıt olduğu ya da organizmanın fizyolojik değişimlere eşlik eden bir fikre sahip olması şeklinde kimi tanımlar söz konusudur.[122] Oysa gözlemlenebilecek durumlar yalnızca beyindeki nöral ateşlemeler ve vücuttaki kimi değişimlerden ibarettir.[123] Ancak ne sevinmek ne üzülmek ne de öfke bu fizyolojik durumlara indirgenebilir. Dolayısıyla duygular, belirli işlevleri yerine getiren zihinsel fenomenler olarak henüz kabul edilemezler. Güçlü yapay zekâ düşüncesinin ve felsefi işlevselciliğin merkezinde yer alan arzu olarak inorganik materyalde gerçekleştirilmek istenen zihnin sahip olamayacağı şeylerden biri duygulardır. Bir makine için sevmek, korkmak ya da öfke duymak oldukça anlamsızdır. Çünkü duygular hep bir *şeye* ilişkindir. Anlamdan yoksun bir varlığın duygulara sahip olmaktan çok onları en fazla simüle edebileceğini söylemek mümkündür. İşte bu durum sanatın ve sanat eserinin değerine bir tehdit oluşturmaktadır. Belirli bir duygulanımın sonucu olarak sanat eseri, bir işleve ya da işlevler bütününe indirgenemez. Sanat eserinin bir *ereği* yoktur. Sanat eserinin ölçülmesini ve bir derecelendirmenin onda tesis edilmesinin olanağını güçleştiren mahiyet tam da budur. Ancak felsefi işlevselci argümanın olası sonuçlarından biri, sanatın tam da matematikselleştirilebilir, mantıksal ifadelere indirgenebilir ve biçimselleştirilebilir bir yapıda olduğunu ileri sürmektir. Çünkü bu argüman zihinsel fenomenlerin biricikliğini yadsımaktadır. Nitelceler olarak buradaki ve şimdideki bir elmanın *kırmızılığı,* bir işlevin ürününden daha fazlası

[122] Barış Korkmaz, "Duygunun Nöropsikolojisi", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 195.
[123] **A.e.**, s. 191.

olarak görülmemektedir. Duygular gibi herhangi bir erekle bağdaştırılamayacak zihinsel fenomenlere sahip olamayacak yapay zekâlı varlıkların özgün ve biricik işler olarak sanat eserlerinin yaratıcısı olabileceğini kabul etmek için temellendirilmiş gerekçelerimizin bulunmadığını belirtmek istiyoruz.

SONUÇ

Eleştirimiz yapay zekâ felsefesi bakımından felsefi işlevselcilik olarak adlandırdığımız görüş üzerine metafelsefi bir yöntem incelemesini, "yapay zekâ" kavramının analizini ve bilgisayımsal temellere dayanan felsefi işlevselciliğin bilgisayar bilimini de içeren bir eleştirisini ortaya koymayı amaçlamıştır. Nitekim eleştirimizin her bir alt bölümü açısından da görülecektir ki eleştirimiz amaçlarına ulaşmıştır: Beyni bir bilgisayar, zihni ise bir işlevler bütünü olarak bir programa indirgemek isteyen bilgisayımsal-temsil işlevselciliğinin veya eleştirimizde kullandığımız adlandırmayla, metafizik kabulleriyle birlikte felsefi işlevselciliğin ortaya koyduğu iddialar ve argümanlar yeterince temellendirilmiş değildir.

Her şeyden önce eleştirimiz, birinci bölümde felsefi işlevselciliğe temel sağlayan unsurlardan biri olarak felsefe öncesi düşünmenin sorunlarından birkaçını, yöntem incelemesi içerisinde açığa çıkarmıştır: Yapay zekâ felsefesi, yapay zekâların ve bedenli yapay zekâlar olarak robotların konuşması, davranışlar göstermesi vb. gibi durumlar üzerinden – hiçbir felsefi ve bilimsel incelemeye maruz bırakılmadan – canlı veya zihin sahibi varlıklar olduklarının düşünüldüğü antropomorfik yüklerle doludur. Tüm kavramsal tartışmalardan önce okuyucu açısından bu yüklerin serimlenmesi gereklidir. Çünkü felsefe içerisine de "başarılması gereken arzular" olarak taşınan bu metafizik kabuller, felsefece güvencenin sağlanabileceği bir duruma zemin hazırlamazlar ve yapay zekâ felsefesi bakımından tartışmaların felsefece yapılabilmesinin önünde birer engel oluştururlar. Yöntem incelememiz ile antropomorfik yükler elendiğinde, bir diğer sıkıca tartışılmamış ön kabullerin bulunduğu bir alan olarak fizikalizm bizi karşılamaktadır. Yapay zekâ tartışmalarında sıkça kabul edilen fizikalist kabullere göre, insan zihinselliği insanın nöral aktivitelerine indirgenmekte, nöral ateşlemelere indirgenen insan zihni maddenin fiziksel örgütlenişinden ve onun deviniminden daha fazlası olarak görülmemektedir. Fizikalist bu kabullere göre, geriye sadece bu örgütlenişin sembolleştirilmesi ve biçimsel bir sistem içerisinde kurallara uygun olarak temsil edilmesi kalmaktadır. Böylece Descartes'tan miras kalan düalizmin tözlerinden biri olan zihinsel töz, bilgimizin sınırlılığı dışındaki bir var olma tarzı olarak, bu fizikalist anlayış tarafından elenebilecek ve matematik ile mantık gibi biçimsel alanlar

üzerinde tanımlanabilecek yeni bir metafizik tasarım olarak yapay zekâlar ön plana çıkabilecektir. Dolayısıyla yapay zekâ felsefesine temel sağlayan fizikalist savların, Descartes'ın ruhunun elenmesi olmadığını, tam da o ruha ilişkin ileri sürülen eleştirilerin biçimsel sistemlerin ifade gücü vasıtasıyla ortadan kaldırıldığı ve onun zihinselliğine ilişkin her türlü bilinmezin de bu yolla elendiği bir metafizik tasarımın okuyucu açısından da açığa çıkarılması zorunludur.

Yöntem incelememizin ikinci adımı, yapay zekâ felsefesinin temelindeki metafizik tasarım olarak inşa edilmek istenen ruhun, yani yapay zekâların, dayandığı temel bilimlerin kendilerini ve problemlerini serimler. Bu açıdan eleştirimizde matematik, mantık ve ilintili önde gelen doğa bilimleri ele alınmıştır. Matematik felsefesi tarihine ilişkin kısa bir bakışla da görülür ki matematik biliminin nesnelerinin güvenceleri hakkında derin felsefi tartışmalar söz konusudur. Bu tartışmalar söz konusuyken kendisini biçimselci bir matematiksel anlayışa konumlandırmış olan ve matematik alanının nesneleri olarak semboller ve semboller arası operasyonlar ile gerçeklik tasarımının inşa edilebileceğini savunan yapay zekâ çalışmalarının sorunlarının ortaya çıkarılması zorunludur. Bu anlamda yapay zekâ çalışmalarına da temel olan Hilbertçi matematik anlayışı ve bunun karşısında konumlanan Gödelci itirazlar aktarılmalıdır. Matematiğin tamlığı ve tutarlılığı tartışmaları, yapay zekâlı varlıkların zihinselliklerinin olanağı ve zihinlerinin eksiksizliği üzerine yargıların dile getirilebilmesi için önemlidir çünkü eğer felsefi işlevselcilerin matematiğe dayanan çalışmalarında dile getirdiği bu iddiaların yine matematik biliminde güvencesi sağlanamazsa, yapay zekâlı varlıkların gerçekliğin tüm unsurlarını matematiksel zeminde gerçekleştirebilecek olanaklarının bulunup bulunmadığını tespit etmek mümkün olamaz. Matematiksel temeller üzerine incelememizden çıkardığımız sonuç ise açıktır: Yapay zekâlı varlıklar bizim ontolojimizde, bizim gerçekliğimizde bulunmamaktadırlar. Onlar bizim gerçekliğimizde eyliyor, bizimle konuşuyor, bizim işlerimizi yapıyor görünseler bile onlar yalnızca matematiksel bir ontolojide eylemektedirler. Onlar yalnızca matematik evrenlerinde vardır. Hiçbir zaman o matematik evrenden çıkmamışlardır. Onlar bizim gerçekliğimizi, şu anda ve burada olan ontolojiyi ve onun nesnelerini yalnızca *kendilerinde şeyler olarak* değil matematiğin temsil gücünün sınırlılığından ötürü

eksik ve güvencesiz bir şekilde temsil edebilmektedirler. Dolayısıyla onlar için tıpkı Searle'ün de öne sürdüğü gibi anlamın, semantiğin sahipliği matematiksel ontolojinin kısıtlılığından ötürü olanaklı olamaz.

Mantığa dayanan yapay zekâ çalışmalarının ikinci girişimi, bir yöntem ve aynı zamanda bu yöntemin bilimi olarak mantığın bir doğa bilimi kılınmasıdır. Mantık temelli yapay zekâ çalışmalarında mantıksal yargılar yoluyla inşa edilmek istenen yapay zekâların gerçeklik tasarımlarının duyu verilerine gitmeden yine kendilerine dayanarak yaratılabileceği ve böylece salt mantıksal yapılarına dayanarak bu yapay zekâların buluşçu mahiyete kavuşabileceği felsefi işlevselci argümanın iddialarından biridir. Çünkü yapay zekâ çalışmalarında mantığın temel alınmasının gerekçeleri açıktır: Bir yöntem olarak mantık sıkı ve güvence sağlayabilecek bir yöntemdir. Yöntembilim olarak mantık ise bu biçimsel yapının geliştirilmesi ve sınanması için bir olanak sağlar. Ancak felsefi işlevselci argüman çerçevesinde inşa edilmek istenen yapay zekâlar bakımından bu çalışmalar içerisinde mantığa kazandırılmak istenen rol burada dile getirdiğimizden daha fazlasıdır: Mantık, bir fizik veya kimya bilimi gibi, genel olarak bir doğa bilimi gibi kılınmak istenmektedir. Çünkü yapay zekâlı varlıklar, yalnızca sahip oldukları mantıksal temele dayanarak sağduyu bilgisi de dahil olmak üzere epistemolojik anlamda insan bilgisinin tüm unsurlarını temsil edebilecek ve yine bu yapıya dayanarak bir doğa anlayışı ortaya koyabilecektir. Ancak önde gelen mantıkçılarımızın da dile getirdiği gibi mantık ne felsefedir ne de bir doğa bilimidir. Fakat bir yöntem ve bir yöntembilim olarak mantık bu alanlarda bilginin üretilmesi ve akıl yürütme süreçlerinin güvencesinin sağlanması bakımından önemli bir rol oynamaktadır. Dolayısıyla sadece mantıksal-biçimsel temelde inşa edilmiş bir yapay zekâlı varlığın, tıpkı biz insanların gerçekleştirmekte olduğu gibi bilim yapabilmesi ve indüksiyon süreçleri vasıtasıyla gerçeklik hakkında yargılar türetebilmesi türünden buluşçu temel mahiyetleri gerçekleştirmesi olanaklı değildir. Çünkü her ne kadar bir yöntem ve yöntembilim olarak mantık, felsefe ve bilimler bakımından temel bir yöntem olarak bulunuyor olsa da sadece mantığın kendisine dayanarak bu alanlardaki başarıların elde edilmesi olanaklı değildir. Sadece mantık temelinde olan ancak duyu verileriyle de kendini besleyen bir yapay zekâlı varlığın ise, duyumsanan nesne ve

olguları nasıl kavramsallaştıracağı açık değildir. Bu sebeple felsefi işlevselci savdan hareketle, yalnızca mantık temelinde inşa edilecek yapay zekâlı varlıkların, insan zihinselliğinin tüm unsurlarına ek olarak, felsefe ve bilim yapmak konusunda da başarılar sağlayabileceğine ilişkin sonuç temelsizdir.

Yöntem incelememizde doğa bilimlerine ayırdığımız alan ise eleştirimizin temelleri bakımından büyük önem arz eder. Çünkü felsefi işlevselciliğin zihinsellik bakımından kendine hedef edindiği varlık olarak insan, doğa bilimlerinden güç alınan bir tarzla, zihinselliğinin nöral ateşlemelere indirildiği bir varlık olarak kabul edilir. Nöral ateşlemelerin ise mantık kapılarından oluşan yapay nöral ağlarda gerçekleştirilmesinin, bu ağlarda zihnin belirmesi için yeterli olacağı düşünülür. Oysa bizim iddiamız nöral ateşlemelerin insan zihinselliğinin gerekli ancak yeterli olmayan durumlarından biri olduğudur. Bu iddiamızın gerekçesi ise Levine'in izah gediği itirazında da dile getirildiği üzere, nöral ateşlemeler gerçekleşse bile herhangi bir X nöral ateşlemenin ilgili herhangi bir zihin durumunun belirmesiyle olan ilişkisi, doğa bilimlerinin salt kendilerine dayanarak ortaya koyabileceği bir ilişki olamayacağıdır. Öte yandan felsefi işlevselcilik içerisinde kaos ve kuantum kuramları vasıtasıyla katı mantık kuralları ve mantık kapıları temelinde inşa edilmiş yapay zekâ sistemlerine insan zihninin sahip olduğu esnekliğin kazandırılabileceği türünden kimi girişimler söz konusu olsa da kaos kuramının temelinde yer alan indeterminist sistemlerin, Roskies'in ileri sürdüğü gibi görünürde indeterminist olan ancak en nihayetinde oldukça karmaşık gözükse de determinist olan sistemler olarak da görülebileceği bir durum söz konusudur. Kuantum kuramının sunduğu olasılıksal çerçevenin ise insan zihninin indeterminist mahiyetini de aşan bir rastgelelilik sunduğu ortadadır ancak bu durum insan zihninin ortaya koyduğu düzenliliğe aykırıdır ve insan zihni tamamen rastgele süreçlerden meydana gelmemektedir. Felsefi işlevselci argümanın aşması gereken sorunlardan biri, inorganik materyalden inşa edilmiş robotların ve bu robotlarda bulunacak yapay zekâların tamamen katı determinist bir şekilde işlemediklerini göstermektir. Bu hususta Denkel'in, determinizmin zorunlulukla katı mahiyette düşünülmemesi gereken ve olasılıkçı bir durumu da kapsayan önerisine benzer öneriler üzerinden yapay zekâ çalışmalarına böyle bir temelin sağlanamayacağına ilişkin bir eleştiri sunduk. Çünkü Denkel'in

determinizm yaklaşımı gibi yapay zekâ çalışmaları bakımından önemli kabul edilebilecek bir yaklaşım her ne kadar determinizmin teleolojik yorumlarını eliyor olsa da varlığın en temel özelliklerinden biri olarak düşünülebilecek neden olabilmeyi yadsımakta, böylece fizik yapabilmenin olanağını ortadan kaldırmaktadır. Dolayısıyla felsefi işlevselci argümana göre inşa edilecek yapay zekâların, katı determinist süreçlere göre eyleyen ve düzenlilikleri kimi biçimsel formül ve ilkeler ile açıklanabilecek bir yapıda olan varlıklar olduğu açıktır. Bu anlamda katı determinist bir anlayış içerisinde özgür irade, bilinç vb. gibi birçok fenomenin, felsefi işlevselcilik tarafından ileri sürüldüğü üzere belirebileceğine ilişkin görüşler temelsizdir.

Eleştirimizin ikinci bölümü ise "yapay zekâ" kavramının kavramsal analizinden oluşmaktadır. Burada öncelikle literatürdeki mevcut yapay zekâ tanımlarının ele alınması gereklidir. Bu tanımların çoğunun zihnin algoritmalarla ifade edilebileceği ve bu algoritmaların bilgisayarlarda gerçekleştirilebileceği ile ilgili varsayımlara dayandığı görülür. Böylece eleştirimizde daha sonraları ele alınacak olan zayıf ve güçlü yapay zekâ yaklaşımlarından güçlü yapay zekâ yaklaşımının, bir makinede gerçekleştirilen zihinsel süreçlerin yalnızca o süreçlerden ibaret olmadığı aynı zamanda makinenin ilgili zihin durumlarına sahip de olduğu yollu görüşünün bu varsayımlara dayanan tanımlara dayandığı vurgulanmıştır. Dolayısıyla bilgisayımsal-temsil işlevselciliği temelinde yükselen felsefi işlevselciliğin güçlü yapay zekâ düşüncesiyle de açıklanabilecek ve zihin durumlarının yalnızca işlevler üzerinden gerçekleştirildiği bilgisayar programları olarak yapay zekâların zihin durumlarına sahip de olabileceklerine ilişkin iddialarının temelinde, zihnin algoritmalara dayanan bir anlayışının yattığı açıktır ve bu iddialar da güçlü yapay zekâ düşüncesine gelen eleştirilerin hedefindedir. Buradan hareketle felsefi işlevselciliğin güçlü yapay zekâ düşüncesiyle de örtüşen iddiaları, en nihayetinde algoritmaların zihinsel süreçleri asıl anlamda temsil edebilecek yapılar olduğunu ortaya koyamaz. Çünkü biçimsel bir yapı olarak algoritmalar, matematiksel ve mantıksal yargılardan oluşur ve bu yargılar içeriksiz olmalarından ötürü zihnin fenomenal içeriklerini gerçekleştirmek konusunda yetersizdir.

Yapay zekâ tanımlarının ele alınması, yapay zekâ çalışmalarındaki teknik süreç ve unsurların da ele alınmasını zorunlu kılar. Bu alandaki teknik süreçlerin insan beyninin fizyolojik yapısı olarak nöral temele odaklandığı görülür. Sembolik yapay zekâ yaklaşımları ve yapay nöral ağlar düşüncesine dayanan bağlantıcı sistemler de bu zihinsel yapının gerçekleştirilmesini arzular. Böylece bu yapıların, nöral ateşlemeleri uygun şekilde inorganik materyal üzerinde gerçekleştirdiklerinde zihin durumlarına sahip de olacakları düşünülür. Ancak görülür ki yapay sinir ağları da dahil olmak üzere farklı mimarilere dayanan yapay zekâ çalışmaları içeriksiz alanlardır ve bazı problemlerin çözümünde paralel mimariye sahip bile olsalar yetersizlerdir. Bunun en büyük sebebinin sağduyu bilgisinin temsil edilmesi konusunda, semantik açıdan yaşanılan eksiklik olduğu açıktır. Böylece felsefi işlevselci argümanın, zihnin seri veya paralel yapay mimariler üzerinde gerçekleştirilebileceğine ilişkin iddiası bu yapıların içeriksiz olmaları, bilginin biçimselleştirilebilmesine ilişkin belirsizliği ve semantik sorunları aşamaz.

Eleştirimizin kavramsal incelemeye ilişkin kısmı "zekâ" kavramına odaklanmıştır. Çünkü "yapay" ve "doğal" sıfatlarının getirildiği kavram olarak "zekâ" kavramının net bir tanımı bulunmamaktadır. Açık bir tanımın bulunmuyor oluşu, felsefi işlevselci argüman bakımından makine zekâsının insan zekâsı ile eş düzeyde ve ondan farksız bir zekâ olabileceğine ilişkin iddianın sınanmasını olanaksız kılmaktadır. Bu anlamda "zekâ" kavramının yapay zekâ literatürü bakımından incelemesinde görüldü ki zekânın tanımı, döngüsel bir tanım olarak zekâ gerektirdiği düşünülen işlerle eşleştirilmektedir. Dolayısıyla insan zekâsı ve makine zekâsının "doğal" ve "yapay" statülerinin, bu var olanlar arasındaki varlık farklarında yattığı ortaya koyulmuştur. Bu varlık farklarına dayanarak felsefi işlevselci argümanın zekânın makineler üzerinde gerçekleştirilebileceğine ilişkin görüşleri sınanmış ve makinelerde ortaya konan hesaplama süreçlerinin, insan zihninin zekâ ile ilişkilendirilen yalnızca belli ve kimi süreçleri bakımından başarılar elde ettiği, oysa bu başarıların insan zihninin tümünü kapsamak bakımından yetersiz kaldığı ortaya koyulmuştur. Öte yandan "yapay zekâ" kavramının kavramsal analizinin, bu alanda tercih edilmesi gereken "zihin" kavramından çok "zekâ" kavramına dayanan temellerinin incelenmesi gerekmektedir. Çünkü felsefi işlevselci argüman içerisinde

“zekâ” kavramı, zihnin bir olanağı olarak hesaplamayla ilişkilendirilmektedir ancak zihnin hesaplamayı da içeren, ondan beslenen fakat hesaplamanın kendisine indirgenemeyecek duyumsama, hayal kurma, düşünme, duygular vb. gibi birçok unsuru da bulunmaktadır. Dolayısıyla hesaplama kuramı temelinde inşa edilmiş “zekâ” kavramının felsefi işlevselci argüman bakımından zihnin tüm yetilerinin ona indirgeneceği ve tekrardan zihnin hesaplama kuramı üzerinden başarılacağı bir olanağı kendinde taşımadığı gösterilmeye çalışılmıştır. Çünkü zihnin “zekâ” dışındaki diğer yetileri yalnızca hesaplamaya indirgenemez. Felsefi işlevselci savların yapay zekâ felsefesi bakımından varmak istediği nokta olarak işlevlere indirgenmiş bir zihinsel yapının, bu yapıyı gerçekleştirebilecek herhangi bir fiziksel yapı tarafından gerçekleştirilmesine ilişkin arzu, “yapay zekâ” kavramının zihin-beden problemi ve onunla ilintili kendilik, bilinç, fenomenal deneyim gibi birçok problemli kavram tarafından da ele alınmasını zorunlu kılmaktadır. Çünkü felsefi işlevselci argümanın iddialarından biri, matematik ve mantık temelinde biçimselleştirilmiş ve algoritmalara dökülmüş olan insan zihninin işlevlerinin, bilgisayarlar üzerinde gerçekleştirilmesi neticesinde ortaya çıkacak yapay zekâlı varlıkların benlik, kendilik, bilinç gibi fenomenal unsurlara sahip olacağıdır. Dolayısıyla “yapay zekâ” kavramının ve felsefi işlevselci argümanın bu iddiasının temelinde yer alan yapay zekâlı varlıkların olanağı bir kere kabul edildiğinde, burada dile getirdiğimiz fenomenler bakımından felsefi işlevselci argümanın karşılaşacağı sorunlar analiz edilmiştir. Bu analize göre, fail statüsü kabul edilecek bir yapay zekâlı varlığın bedeninin sınırları, yazılımının bu bedene olan özgülüğü tesis edilmedikçe benlik ve kendilik gibi fenomenal unsurlara sahip olmaları olanaksızdır. Çünkü kavramsal incelememizde de görülmüştür ki benlik ve kendilik düşünceleri uzamsal ve zamansal olarak süreklilik gösteren bir bedeni ve bu bedene özgü bir fenomenal deneyimi zorunlu kılmaktadır. Bedeni ve fenomenal deneyimi ile bir bütün olarak bu varlığın benlik ve kendilik düşüncesine sahip olması olanaklıdır. Fakat yapay zekâlı varlıkların algoritmalar temelinde bir bütün olarak ne yazılımsal anlamda, sayılabilirlik bakımından bir adet bedene özgü olacakları açık değildir ne de bedenlerinin bir sınırı bulunmaktadır. Dolayısıyla kavramsal incelememizle de açık kılınmıştır ki felsefi işlevselciliğin yapay zekâların “kendilik” düşüncesine sahip olabileceklerine ilişkin iddiası temelsizdir.

Eleştirimizin üçüncü bölümü işlevselcilik, felsefi işlevselcilik ve yapay zekâ ilişkisi üzerinde durmaktadır. İşlevselcilik düşüncesinin özü, ele alınan kavram veya yapının işlevleri üzerinden düşünülmesidir. Ned Block'un işlevsel analiz olarak tasniflediği böyle bir işlevselcilik değil ama eleştirimizin adlandırılmasında da metafizik kabulleriyle ve bilgisayımsal-temsil işlevselcilik ile bir bütün olan "felsefi işlevselcilik" ele alınmıştır. Felsefi işlevselciliğin özü insan zihinselliğinin tüm unsurlarını işlevlere indirgemesi ve bu işlevlerin bilgisayarlar üzerinde çalışan algoritmalar olarak programlar üzerinden başarılmak istenmesidir. Eleştirimizin bilgisayımsal temeller üzerinde de durması zorunludur. Dolayısıyla bilgisayar bilimi ve bu bilimin temel nesnesi olarak bilgisayarların ve algoritmaların sahip oldukları kimi problemler üzerinde durduk. Karar verme probleminin çözümsüzlüğü ile ilgili olarak durma probleminin, bilgisayarların hesaplayamayacağı kimi durumların söz konusu olduğunu ortaya koyan önemli bir problem olarak yapay zekâ çalışmaları ve özellikle felsefi işlevselci uslamlama bakımından önemli bir karşı itiraz noktası oluşturduğu vurgulanmıştır. Öte yandan felsefi işlevselci argüman içerisinde zihinlerin belirmesinin tasarlandığı yapılar olarak bilgisayarların kuantum bilgi işleme gibi süreçlerle esneklik kazanacakları düşünülse de eleştirimizde de vurgulanmıştır ki bilgisayarlara kuantum bilgi işleme olanağının sağlandığı durumda bile bilgisayarların gerçekliği tam anlamıyla simüle etmesinin bir olanağı bulunmamaktadır. Bu hususta Feynman ve Dennett'ın bilgisayarların simüle ettiği durumlarla simüle edilen durumlar arasındaki nitelik farklarına dikkat çektiği uslamlamaları aktarılmıştır. Eleştirimiz, felsefi işlevselci argümana dayanarak ortaya konacak herhangi bir bilgi işleme kapasitesine sahip bir fiziksel yapıda gerçekleştirilebilen bir zihin olarak yapay zekâlı varlıkların olanağı bir kere kabul edildiğinde sanat alanında söz konusu olabilecek problemler hakkında bir denemeyi de içermektedir. Bu denemede amacımız, ereği olmadığını düşündüğümüz bir insan etkinliği olarak sanatın fail olmakla ve duygu gibi kimi fenomenal unsurlara sahip olmakla olan ilişkisine dayanarak felsefi işlevselci argümanın tasarımladığı yapay zekâlı varlıkların sanat eserinin ortaya koyulması konusundaki olanaklılıkları üzerine bir düşünme sürecini incelemektir. Yapay zekâlı varlıkların eleştirimizin ortaya koyduğu temel çerçevede düşünüldüğünde semantik içerikten yoksun yapıları, sanatsal üretimin onlarda yalnızca görünürde var olan bir mahiyette söz konusu

olabileceğini göstermektedir. Çünkü sanat yapmanın kimi nota ve renklerin gelişigüzel bir araya getirilmesi olmadığı, bağlamın da onda içkin olduğu açıktır.

Eleştirimizin ulaştığı sonuçların okuyucu açısından artık açık kılınması zorunludur: Felsefi işlevselci argüman, beyinlerimizi bilgisayarlar ve zihinlerimizi bilgisayarlarda çalışan programlar olarak izah edemez. Çünkü eleştirimizde de ortaya konmaya çalışılmıştır ki insan zihinselliği ne matematik veya mantık yoluyla biçimselleştirilebilir ne de doğa bilimlerinin ele aldığı gibi yalnızca nöral ateşlemelere indirgenebilir. Dolayısıyla bu bilimler temelinde inşa edilmiş olan ve bizce düalist yüklerinden kurtulmasına rağmen fizikalizmin varlığı ele alma tarzından, ifade etme gücünden de beslenerek hala var olmakta olan *zihinsel töz* olarak yapay zekâlı varlıkların olanağı ve epistemolojik/ontolojik güvenceleri sağlanamaz. Bu anlamda böyle bir varlığın gerçekleştirilmesi uğruna yapay zekâ çalışmalarının temelindeki bilgi ve bilim alanlarının, konunun felsefi mahiyetini tanımazcasına alanı metafizik kabullerle doldurduğu görülmektedir. Antropomorfik yüklerle ve metafizik kabullerle dolmuş olan yapay zekâ çalışmaları ve yapay zekâ felsefesinin, yapay zekâlı varlıklarla ilgili temel kavramlar olan bedeni, zihni, zekâyı ve bu zekânın yapaylığı/doğallığı gibi birçok kavramı değerlendirmekteki başarısızlığı, insanın varlık koşullarının da yapay zekâ alanının gelişmeleri karşısında hatalı değerlendirilmesine neden olmuştur. Eleştirimizdeki tavrımız, "insan" kavramının değerinin yapay zekâ gelişmeleri karşısında korunmasına ve onun bu alanda felsefece bir ele almaya yeniden layık görülmesine yöneliktir. Böyle bir ele alma için metafelsefi olarak da ihtiyaç duyulan antropomorfik yüklerden arınmış ancak "insan" kavramını sınır tanımazcasına yıkmayan, onun varlık koşullarını dikkate alan, diyalektik bir felsefe yapmanın önemi vurgulanmıştır. Çünkü yapay zekâ tanımlarının ele alınmasıyla da görülmüştür ki metafizik kabulleriyle yapay zekâ felsefesindeki şimdiye kadar ortaya konmuş olan tanımlar, yapay zekâ felsefesi bakımından problemlerin çözümlenmesinde başarısız kalmıştır. Dolayısıyla burada dile getirdiğimiz türde bir felsefe yapma tarzı benimsenmedikçe insan ve yapay zekâlı varlıklar arasındaki varlık farkları açığa çıkarılamaz. Çıkarılamadığı gibi de "insan" kavramımız bir yeniden değerlendirmeye hazır olamaz.

KAYNAKÇA

Acar, Ayşe: "Autonomy in Human and Artificial Intelligence and Its Relationship with the Concepts of Dignity and Human Rights", İstanbul, İstanbul Bilgi Üniversitesi Lisansüstü Programlar Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2020.

Akın, H. Levent: "Yapay Zekâda Vücut Beyin Problemi", (Çevrimiçi) https://www.cmpe.boun.edu.tr/~akin/papers/beyin.pdf, 19 Kasım 2022.

Akman, Toygar: **Sibernetik**: **Dünü, Bugünü, Yarını,** İstanbul, Kaknüs Yayınları, 2003.

Akyüz, Ömür: "Beyin, Zihin, Zaman, Fizik (Kaos, Kuantum Mekaniği vs.)", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, Pan Yayıncılık, İstanbul, 2012, s. 207-212.

"Zihnin Fiziği?", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 128-135.

Alagüney, Ahmet Haşim: "Hukukta İrade Özgürlüğü Kavramının Yapay Zekâ Türleri Bakımından Tartışılması", **Pasajlar**, No: 9, 2021, s. 197-213.

Alpaydın, Ethem: "Yapay Zekâ/Doğal Bilgisayarlar", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 95-102.

Yapay Öğrenme: Yeni Yapay Zekâ, çev. Aylin Ağar, 2.bs, İstanbul, Tellekt, 2021.

Altınörs, S. Atakan: "Zihin-Beden Problemine Searle'ün Yaklaşımı", **Temaşa Felsefe Dergisi**, No: 9, 2008, s. 39-59.

Arf, Cahit: "Makine Düşünebilir mi ve Nasıl Düşünebilir?", **Atatürk Üniversitesi 1958-1959 Öğretim Yılı Halk Konferansları**, No: 1, s. 91-103.

Arıcı, Murat: "Zihinsel Alanda 'Ben Hissi' Neye Karşılık Gelir?", **Dört Öge**, No: 7, 2015, s. 65-76.

Arıcı, Murat ve Doğan, Mehtap: "Prof. Dr. Erdinç SAYAN ile Söyleşi: Zihnin Doğası Üzerine", **MetaZihin**, C. I, No: 2, 2018, s. 241-255.

Arıcı, Murat: "Materyalizm, Fenomenal Özne ve Ontolojik Statüsü", **Beytülhikme**, C. IV, No: 2, 2014, s. 2-13.

Aristoteles: **Metafizik**, çev. Y. Gurur Sev, 4. bs., İstanbul, Pinhan Yayıncılık, 2018.

Organon V -Topikler-, çev. Prof. H. Ragıp Atademir, C. V, İstanbul, Milli Eğitim Basımevi, 1996.

Arslantunalı, Mustafa: "Yapay Zekâ ve İnsanbiçimcilik: 'Bacakları Olan Bir Araba'", **Pasajlar**, No: 9, 2021, s. 227-237.

Asimov, Isaac: **Ben, Robot**, çev. Ekin Odabaş, 4. bs., İstanbul, İthaki Yayınları, 2018.

Audi, Robert: **The Cambridge Dictionary of Philosophy**, 2. bs., New York, Cambridge University Press, 1999.

Ayer, A. J.: **The Concept of a Person and Other Essays**, London, Macmillan Press, 1973.

Ayral, Hakan: "Halting Prediction on Busy Beaver Type Turing Machines Based on Information Entropy", İstanbul, Galatasaray Üniversitesi Fen Bilimleri Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2008.

Baker, Ulus: "Homo Homunculus: Kendi Kendini Yaratamayan İnsan", **Pasajlar**, No: 9, 2021, s. 215-226.

Barrow, Harry: "Connectionism and Neural Networks", **Artificial Intelligence**, ed. Margaret A. Boden, 2. bs., San Diego, Academic Press, 1996.

Bell, Anthony J.: "Levels and loops: the future of artificial intelligence and neuroscience", **Philosophical Transactions of the Royal Society of London**, No: 354, 1999, s. 2013-2020.

Bingöl, Haluk: "Bilgisayar Üzerine", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 28-39.

"Karmaşık Ağlar Beyni Anlamamızda Yardımcı Olabilir mi?", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 303-314.

Blackburn, Simon: **The Oxford Dictionary of Philosophy**, 2. bs., Oxford, Oxford University Press, 2005.

Block, Ned: "Is Experiencing Just Representing?", (Çevrimiçi) https://www.nedblock.us/papers/1998.ppr.representing.pdf, 9 Aralık 2022.

"Paradox and cross purposes in recent work on consciousness", **Cognition**, No: 79, 2001, s. 197-219.

"Troubles with Functionalism", **Consciousness, Function, and Representation: Collected Papers**, C. I, Cambridge-Massachusetts, The MIT Press, 2007, s. 63-101.

"What is Functionalism?", **Consciousness, Function, and Representation: Collected Papers**, C. I, Cambridge-Massachusetts, The MIT Press, 2007, s. 27-44.

Boden, Margaret A.: **Minds as Machine: A History of Cognitive Science**, C. II, Oxford, Clarendon Press, 2006.

"Artificial Intelligence", Routledge Encyclopedia of Philosophy, (Çevrimiçi) https://www.rep.routledge.com/articles/thematic/artificial-intelligence/v-1#, 18 Kasım 2022.

Boole, George: **An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities**, Cambridge, Macmillan and Co., 1854.

Bostrom, Nick: "Ethical Issues in Advanced Artificial Intelligence", (Çevrimiçi) https://nickbostrom.com/ethics/ai, 17 Aralık 2022.

Bringsjord, Selmer ve Govindarajulu, Naveen Sundar: "Artificial Intelligence", The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta ve Uri Nodelman, (Çevrimiçi) https://plato.stanford.edu/entries/artificial-intelligence/, 2 Ocak 2023.

Burkholder, Leslie: "The Halting Problem", **ACM SIGACT News**, C. XVII, No: 3, 1987, s. 48-60.

Calude, Cristian S.: "Incompleteness and the Halting Problem", **Studia Logica**, No: 129, s. 1159-1169.

Canbeyli, Reşit: "Psikoloji Açısından Beyin-Bilgisayar Karşılaştırması", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 113-127.

Capek, Karel: **R.U.R. – Rossum'un Uluslararası Robotları**, çev. Bilge Kösebalaban, 2. bs, İstanbul, İthaki Yayınları, 2021.

Chella, Antonio ve Manzotti, Riccardo: "Artificial intelligence and consciousness", Association for the Advancement of Artificial Intelligence Fall Symposium, (Çevrimiçi) https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=0958617423b09e79b03f050cd807c0c7e5859b66, 4 Ocak 2023.

Choi, Yejin: "The Curious Case of Commonsense Intelligence", **Daedalus**, C. CLI, No: 2, 2022, s. 139-155.

Chrisley, Ron: "Philosophical Foundations of artificial consciousness", **Artificial Intelligence in Medicine**, No: 44, 2008, s. 119-137.

Churchland, Patricia Smith ve Sejnowski, Terrence J.: "Neural Representation and Neural Computation", **Philosophical Perspectives**, C. IV, 1990, s. 343-382.

Churchland, Paul M. ve Churchland, Patricia Smith: "Could a Machine Think?", **Scientific American**, C. CCLXXII, No: 1, 1990, s. 32-39.

Clark, Andy ve Chalmers, David: "The extended mind", **Analysis**, C. LVIII, No: 1, 1998, s. 7-19.

"Philosophical Foundations", **Artificial Intelligence**, ed. Margaret A. Boden, 2. bs., San Diego, Academic Press, 1996.

Crevier, Daniel: **AI: The Tumultuous History of the Search for Artificial Intelligence**, New York, Basic Books, 1993.

Çağıl, Ayşe: "Daniel Dennett'ın Zihin Felsefesinde Güçlü Yapay Zekânın Olasılığı", Ankara, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2022.

Çelebi, Vedat: "Searle ve Nagel'in Bilinç Kuramlarında Fizikalizm Eleştirisi Bağlamında Yapay Zekânın Sınırı Tartışması", **Beytülhikme**, C. IX, No: 9, 2019, s. 351-376.

Davis, Martin: **The Universal Computer: The Road From Leibniz to Turing**, New York, W. W. Norton & Company, 2000.

Demircioğlu, Erhan: "Mehmet Hilmi Demir: 'Aklımızın Sınırlarını Deniyoruz Aslında, Nereye Kadar Gidebiliriz?'", **Pasajlar**, No: 9, 2021, s. 133-157.

"Yapay Zekâ: Açmaz ve İmkân", **Pasajlar**, No: 9, 2021, s. 11-15.

Denizhan, Yağmur: "Beyin Bağlamında Kaotik Sistemlere Bakış", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 165-173.

Denkel, Arda: "Belirlenim, Neden ve Değişim", **Düşünceler ve Gerekçeler: Felsefe Yazıları**, ed. Ahmet Haluk Atalay, C. II, İstanbul, Göçebe Yayınları, 1997.

Dennett, Daniel C.: "When Philosophers Encounter Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 283-295.

"Ben Neredeyim?", **Aklın Gözü: Benlik ve Ruh üzerine Hayaller ve Düşünceler**, ed. Daniel C. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 209-222.

"Cognitive Wheels: The Frame Problem of AI", (Çevrimiçi) http://cinc.uaem.mx/~bruno/material/dennet_93_cognitiveWheels.pdf, 14 Aralık 2022.

"Why You Can't Make a Computer That Feels Pain", **Synthese**, No: 38, 1978, s. 415-456.

Bilinç Açıklanıyor, çev. Sibel Kibar, İstanbul, Alfa Basım Yayım, 2017.

Descartes, Rene: **İlk Felsefe Hakkında Meditasyonlar**, çev. İsmet Birkan, 3. bs., Ankara, BilgeSu Yayıncılık, 2014.

Deveci, İbrahim Ethem; Küskü, Mert ve Özkurt, Tolga Esat: "Makinenin Beni ve Ötekisi", **Pasajlar**, No: 9, 2021, s. 103-117.

Dilek, Çağlan Çınar: "Zor Problem ve Yüksek Düzey Düşünce Kuramı", **Phainomena**, No: 1, 2020, s. 127-151.

Dilman, Ilham: **Free Will: An Historical and Philosophical Introduction**, New York, Routledge, 2001.

Doğan, Mehtap: "Biyolojik Doğalcılık Ekseninde John Searle ve Zihin-Beden Problemi", **MetaZihin**, C. I, No: 1, 2018, s. 227-240.

"Güçlü Yapay Zekâ Mümkün müdür?", **Pasajlar**, No: 9, 2021, s. 87-101.

"Yapay Zekâ Felsefesinde Bilinç Problemi", Ankara, Ankara Yıldırım Beyazıt Üniversitesi Sosyal Bilimler Enstitüsü, Yayımlanmamış Doktora Tezi, 2020.

"Yapay Zekâ ve Özgür İrade: Yapay Özgür İradenin İmkânı", **TRT Akademi**, C. VI, No: 13, 2021, s. 788-811.

Dreyfus, Hubert L. ve Dreyfus, Stuart E.: "Making a Mind Versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint", **Dædalus**, C. CXVII, No: 1, 1988, s. 15-43.

Dreyfus, Hubert L.: "Bilgisayarların Akıllı Olması İçin Neden Bedenleri Olması Gerekir?", çev. Berk Celayir ve Selenay Yakın, **Pasajlar**, No: 9, 2021, s. 45-61.

Durmuş, Talha Zekeriya: Sonlanma Problemi, (Çevrimiçi) https://talhadurmus.com/bilgisayarbilimi/sonlanma-problemi, 26 Haziran 2022.

Ersoy, Çağlar: **Robotlar, Yapay Zekâ ve Hukuk**, 2. bs., İstanbul, On İki Levha Yayıncılık, 2017.

Feynman, Richard P.: "Simulating Physics with Computers", **International Journal of Theoretical Physics**, C. XXI, No: 6/7, 1982, s. 467-488.

Fisher, Mark: "A note on free will and artificial intelligence", **Philosophia**, C. I, No: 13, 1983, s. 75-80.

Floridi, Luciano: "The Onlife Manifesto: Being Human in Hyperconnected Era", (Çevrimiçi) https://ec.europa.eu/digital-single-market/sites/digital-agenda/files/Manifesto.pdf, 25 Aralık 2022.

Philosophy and Computing: An Introduction, New York, Routledge, 1999.

Fodor, Jerry A. ve Pylyshyn, Zenon W.: "Connectionism and cognitive architecture: A critical analysis", **Cognition**, C. XXVIII, No: 1-2, 1988, s. 3-71.

Frankfurt, Harry G.: "Freedom of the Will and the Concept of a Person", **The Journal of Philosophy**, C. LXVIII, No: 1, 1971, s. 5-20.

Franzen, Torkel: **Gödel's Theorem: An Incomplete Guide to Its Use and Abuse**, Wellesley-Massachusetts, A. K. Peters, 2005.

Gamez, David: "The Potential for Consciousness of Artificial Systems", **International Journal of Machine Consciousness**, C. I, No: 2, 2009, s. 213-223.

Goertzel, Ben: "Artificial General Intelligence: Concept, State of the Art, and Future Prospects", **Journal of Artificial General Intelligence**, C. V, No: 1, 2014, s. 1-46.

Gödel, Kurt: "Some basic theorems on the foundations of mathematics and their implications", **Collected Works: Unpublished essays and lectures**, ed. Solomon Feferman v.d., C. III, New York, Oxford University Press, 1995.

"The modern development of the Foundations of mathematics in the light of philosophy", **Collected Works: Unpublished essays and lectures**, ed. Solomon Feferman ve diğerleri, C. III, New York-Oxford, Oxford University Press, 1995.

Gözkan, Bülent: "Bilgi, Bilinç ve Yapay Zekâ", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs. İstanbul, Pan Yayıncılık, 2012, s. 157-164.

Grünberg, Teo: **Felsefe ve Felsefi Mantık Yazıları**, İstanbul, YKY, 2005.

Sembolik Mantık El Kitabı: Sembolik Mantığın Uygulamaları, C. III, Ankara, METU PRESS, 2000.

Sembolik Mantık El Kitabı: Temel Mantık, C. I, 2. bs., Ankara, ODTÜ Yayıncılık, 2017.

Güven, Özgüç: "Hilbert, Matematiğin Temelleri ve Görü", **Felsefe Arkivi**, No: 52, 2020, 113-149.

Hagar, Amit ve Cuffaro, Michael: "Quantum Computing", The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta ve Uri Nodelman, (Çevrimiçi) https://plato.stanford.edu/archives/fall2022/entries/qt-quantcomp/, 20 Aralık 2022.

Hanks, Steve ve McDermott, Drew: "Default Reasoning, Nonmonotonic Logics, and the Frame Problem", **AAAI**, C. LXXXVI, 1986, s. 328-333.

Harding, D. E.: "Bir Kafaya Sahip Olmamak Hakkında", **Aklın Gözü: Benlik ve Ruh Üzerine Hayaller ve Düşünceler**, ed. Daniel C. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 31-37.

Hauser, Larry: "Artificial Intelligence", Internet Encyclopedia of Philosophy, (Çevrimiçi) https://iep.utm.edu/artificial-intelligence/#SSH4c.i, 15 Aralık 2022.

Hehner, Eric C. R.: "Problems with the Halting Problem", **Advances in Computer Science and Engineering**, C.I, No: 10, 2013, s. 31-60.

Hilbert, David ve Ackermann, W.: **Principles of Mathematical Logic**, çev. Lewis M. Hammond, George G. Leckie ve F. Steinhardt, ed. Robert E. Luce, New York, Chelsea Publishing Company, 1950.

Hodgson, David: "Quantum Physics, Consciousness, and Free Will", **The Oxford Handbook of Free Will**, ed. Robert Kane, New York, Oxford University Press, 2002, s. 85-110.

Hoffmann, E. T. A.: **Kum Adam**, çev. Ersel Kayaoğlu ve Sami Türk, İstanbul, Can Yayınları, 2020.

Holland, Owen: "The future of embodied artificial intelligence: Machine consciousness?", **Embodied Artificial Intelligence**, Ed. Fumiya Iida, Rold Pfeifer, Luc Steels, Ysuo Kuniyoshi, Berlin, Springer, 2004,

Hurlbert, Anya ve Poggio, Tomaso: "Making Machines (and Artificial Intelligence) See", **Dædalus**, C. CXVII, No: 1, 1988, s. 213-239.

Inwagen, Peter Van: "Free Will Remains a Mystery", **The Oxford Handbook of Free Will**, ed. Robert Kane, New York, Oxford University Press, 2002, s. 158-177.

Işıklı, Şevki: "Fuzzy Mantık ve Sibernetik'in Siber Toplum ve Yapay Zekâ Üzerindeki Etkileri", Erzurum, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Yayımlanmamış Yüksek Lisans Tezi, 2004.

İnel, Oğuz: **Özgür İrade Yanılsaması**, İstanbul, LEJAND Kitap, 2022.

İnönü, Mahbube Nazlı: "Yapay Zekâ Felsefesi", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 136-147.

Jackson, Frank: "Epiphenomenal Qualia", **The Philosophical Quarterly**, C. XXXII, No: 127, 1982, s. 127-136.

Kane, Robert: **A Contemporary Introduction to Free Will**, New York, Oxford University Press, 2005.

Kaplan, Jerry: **Artificial Intelligence: What Everyone Needs to Know**, New York, Oxford University Press, 2016.

Kitcher, Philip: "Mathematical Naturalism", **From Kant to Hilbert Volume 1: A Source Book in the Foundations of Mathematics**, C. I., ed. William Bragg Ewald, Oxford, Oxford University Press, 1996.

Koç, Yalçın: **Determinizm ve Mekân**, İstanbul, Boğaziçi Üniversitesi Yayınları, 1984.

Korkmaz, Barış: "Çocuklarda Nöronal ve Zihinsel Gelişim", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 57-71.

Köken, Enes: "Yapay Zekânın Cezai Sorumluluğu", Türkiye Adalet Akademisi Dergisi, No: 47, 2021, s. 793-833.

Köse, Utku: **Yapay Zekâ Felsefesi**, İstanbul, Doğu Kitabevi, 2022.

Kurzweil, Raymond: **The Age of Intelligent Machines**, Cambridge, The MIT Press, 1990.

Lamme, Victor A. F.: "Why visual attention and awareness are different", **Trends in Cognitive Sciences**, C. VII, No: 1, 2003, s. 12-18.

Larson, Erik J.: **Yapay Zekâ Miti -Bilgisayarlar Neden Bizim Gibi Düşünemez-**, çev. Kadir Yiğit Us, Ankara, Fol Kitap, 2022.

Lem, Stanislaw: "Non Serviam", **Aklın Gözü: Benlik ve Ruh Üzerine Hayaller ve Düşünceler**, ed. Daniel D. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 288-306.

Levine, Joseph: "Materialism and Qualia: The Explanatory Gap", **Pacific Philosophical Quarterly**, No: 64, 1983, s. 354-361.

Lewis, Michael: "Myself and me", **Self-awareness in Animals and Humans: Developmental Perspectives**, ed. Sue Taylor Parker ve Robert W. Mitchell, New York, Cambridge University Press, 2006, s. 20-34.

Lloyd, Seth: "A Turing test for free will", **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, No: 370, 2012, s. 3597-3610.

Lucas, J. R.: "Minds, Machines and Gödel", **Philosophy**, C. XXXVI, No: 137, 2009, s. 112-127.

The Freedom of the Will, Oxford, Clarendon Press, 1970.

Manzotti, Riccardo ve Tagliasco, Vincenzo: "Artificial consciousness: A discipline between technological and theoretical obstacles", **Artificial Intelligence in Medicine**, No: 44, 2008, s. 105-117.

McCarthy, John ve Hayes, Patrick J.: "Some Philosophical Problems from the Standpoint of Artificial Intelligence", (Çevrimiçi) http://www-formal.stanford.edu/jmc/mcchay69.pdf, 8 Aralık 2022.

McCarthy, John: "Artificial Intelligence, Logic and Formalizing Common Sense", (Çevrimiçi) http://www-formal.stanford.edu/jmc/ailogic.pdf, 27 Aralık 2022.

"Concepts of Logical AI", (Çevrimiçi) http://www.formal.stanford.edu/jmc/concepts-ai.pdf, 28 Aralık 2022.

"From here to human-level AI", **Artificial Intelligence**, No: 171, 2007, s. 1174-1182.

"Generality in Artificial Intelligence", **Communications of the ACM**, C. XXX, No: 12, 1987, s. 1030-1035.

"Mathematical Logic in Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 297-211.

"Programs with Common Sense", Cambridge, **RLE and MIT Computation Center**, 1960, s. 75-91.

"Review of the Emperor's New Mind by Roger Penrose", (Çevrimiçi) http://jmc.stanford.edu/artificial-intelligence/reviews/penrose1.pdf, 7 Ocak 2023.

"The well-designed child", **Artificial Intelligence**, No: 172, 2008, s. 2003-2014.

"What has AI in Common with Philosophy?", (Çevrimiçi) https://www.cp.eng.chula.ac.th/~prabhas/teaching/cbs-it-seminar/2012/aiphil-mccarthy.pdf, 18 Kasım 2022.

"What is Artificial Intelligence?", (Çevrimiçi) http://www.formal.stanford.edu/jmc/whatisai.pdf, 25 Aralık 2022.

McCulloch, Warren S. ve Pitts, Walter: "A Logical Calculus of the Ideas Immanent in Nervous Activity", **Bulletin of Mathematical Biology**, C. LII, No: ½, 1990, s. 99-115.

Mill, John Stuart: "Of the Remaining Laws of Nature", **A System of Logic**, 8. bs., New York, Harper&Brothers, 1882.

Minsky, Marvin: "Logical Versus Analogical or Symbolic Versus Connectionist or Neat Versus Scruffy", **AI Magazine**, C. XII, No: 2, 1991, s. 34-51.

"Steps Toward Artificial Intelligence", **Proceedings of the IRE**, C. XLIX, No: 1, 1961, s. 8-30.

Moravec, Hans: "When will computer hardware match the human brain?", **Journal of Evolution and Technology**, C. I, No: 1, 1998, s. 10-21.

Morowitz, Harold J.: "Zihni Yeniden Keşfetmek", **Aklın Gözü: Benlik ve Ruh Üzerine Hayaller ve Düşünceler**, ed. Daniel C. Dennett ve Douglas R. Hofstadter, çev. Füsun Doruker, İstanbul, Boğaziçi Üniversitesi Yayınevi, 2008, s. 41-48.

Nagel, Thomas: "What Is It Like to Be a Bat?", **The Philosophical Review**, C. LXXXIII, No: 4, 1974, s. 435-450.

Nahmias, Eddy; Allen, Corey Hill ve Loveall, Bradley: "When Do Robots Have Free Will? Exploring the Relationships between (Attributions of) Consciousness and Free Will", **Free Will, Causality, and Neuroscience**, ed. Bernard Feltz, Marcus Missal ve Andrew Sims, Leiden, Brill Rodopi, 2020, s. 57-80.

Neslioğlu, E. Funda: "İrade Özgürlüğü Sorununun Çözümsüzlüğü Üzerine", **Felsefe Dünyası**, C.I, No: 47, 2008, s. 179-184.

Newell, Allen ve Simon, Herbert A.: "The Simulation of Human Thought", (Çevrimiçi) https://stacks.stanford.edu/file/druid:cn091ff4880/cn091ff4880.pdf, 7 Aralık 2022.

Newell, Allen: "Physical Symbol Systems", **La Jolla Conference on Cognitive Science**, California, 15 Ağustos 1979, (Çevrimiçi) https://scholar.archive.org/work/skfkjrtc3nftnacjc7berhd7vq/access/wayback/https://s3-eu-west1.amazonaws.com/pstorage-cmu-348901238291901/12099818/file.pdf, 6 Ocak 2023.

Nichols, Shaun: "Experimental Philosophy and the Problem of Free Will", **Science**, No: 331, 2011, s. 1401-1403.

Nilsson, Nils J.: "The Physical Symbol System Hypothesis: Status and Prospects", **50 Years of Artificial Intelligence: Essays Dedicated to the 50th Anniversary of Artificial Intelligence**, ed. Max Lungarella, Fumiya Iida, Josh Bongard ve Rolf Pfeifer, Springer, 2007.

Artificial Intelligence A New Synthesis, San Francisco, Morgan Kaufmann Publishers, 2009.

The Quest for Artificial Intelligence, Cambridge, Cambridge University Press, 2010.

O'Regan, Kevin: "How to Build a Robot that is Conscious and Feels", **Minds & Machines**, No: 22, 2012, s. 117-136.

Özel, Emir H.: "Church-Turing Tezi ve Algoritmik Zihin Argümanı: Zihinsel Fenomenleri Matematiğe İndirgemek Mümkün müdür?", **Pasajlar**, No: 9, 2021, s. 159-175.

Özlem, Doğan: **Mantık**, 7. bs., İstanbul, İnkılap Kitabevi, 2004.

Öztürk, Kadir ve Şahin, Mustafa Ergin: "Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış", **Takvim-i Vekayi**, C. VI, No: 2, 2018, s. 25-36.

Pagels, Heinz R., v.d.: "Has Artificial Intelligence Research Illuminated Human Thinking?", **Annals of the New York Academy of Sciences**, No: 426, 1984, s. 138-160.

Papert, Seymour: "One AI or Many?", **Dædalus**, C. CXVII, No: 1, 1988, s. 1-14.

Parthemore, Joel ve Whitby, Blay: "What Makes Any Agent a Moral Agent?: Reflections on Machine Consciousness and Moral Agency", **International Journal of Machine Consciousness**, C. V, No: 2, 2013, s. 105-129.

Pennachin, Cassio ve Goertzel, Ben: "Contemporary Approaches to Artificial General Intelligence", **Artificial General Intelligence**, ed. Ben Goertzel ve Cassio Pennachin, C. II, New York, Springer, 2007.

Penrose, Roger: **Kralın Yeni Usu**, çev. Tekin Dereli, C. I, 10. bs., Ankara, TÜBİTAK, 2000.

Pink, Thomas: **Free Will: A Very Short Introduction**, New York, Oxford University Press, 2004.

Pitts, Walter ve McCulloch, Warren S.: "How We Know Universals the Perception of Auditory and Visual Forms", **Bulletin of Mathematical Biophysics**, C. IX, 1947, s. 127-147.

Platon: **Menon**, çev. Furkan Akderin, 2. bs, Ankara, Say Yayınları, 2013.

Proudfoot, Michael ve Lacey, A. R.: **The Routledge Dictionary of Philosophy**, 4. bs., New York, Routledge, 2010.

Putnam, Hilary: "Beyinler ve Davranış", çev. İdil Gemici, **Phainomena**, No: 1, 2020, s. 153-175.

"Much Ado About Not Very Much", **Dædalus**, C. CXVII, No: 1, 1988, s. 269-281.

"Psychological Predicates", **Art, Mind and Religion**, C. I, London, University of Pittsburgh Press, 1967, s. 158-167.

Mind, Language and Reality: Philosophical Papers, C. II, Cambridge, Cambridge University Press, 1997.

Pylyshyn, Zenon W.: "Computation and cognition: issues in the foundations of cognitive science", **The Behavioral and Brain Sciences**, C. III, No: 1, 1980, s. 111-132.

Reeke Jr., George N. ve Edelman, Gerald M.: "Real Brains and Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 143-173.

Reggia, James A.: "The rise of machine consciousness: Studying consciousness with computational models", **Neural Networks**, No: 44, 2013, s. 112-131.

Revonsuo, Antti: **Bilinç: Öznelliğin Bilimi**, çev. Selim Değirmenci, 2. bs., İstanbul, Küre Yayınları, 2017.

Rosenblatt, F.: "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", **Psychological Review**, C. LXV, No: 6, 1958, s. 386-408.

Roskies, Adina L.: "Can Neuroscience Resolve Issues about Free Will?", **Moral Psychology**: **Free Will and Moral Responsibility**, ed. Walter Sinnott-Armstrong, C. IV, Cambridge, The MIT Press, 2014, s. 103-126.

"Neuroscientific challenges to free will and responsibility", **Trends in Cognitive Sciences**, C. X, No: 9, 2006, s. 419-423.

Russell, Stuart: **İnsanlık için Yapay Zekâ: Yapay Zekâ ve Kontrol Problemi**, çev. Barış Satılmış, Ankara, Buzdağı Yayınevi, 2021.

Say, Cem: "Diller, Çözümsüz Problemler ve Sonsuzluk", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. Bs., İstanbul, Pan Yayıncılık, 2012, s. s. 177-184.

"Yapay Zekâ ve Gödel Teoremi", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 254-264.

50 Soruda Yapay Zekâ, 15. bs., İstanbul, 7 Renk, 2020.

Sayan, Erdinç ve Aytekin, Tevfik: "Being Mindful of Zombies and Phenombies", **MetaZihin**, C. I, No: 2, 2018, s. 159-165.

Sayan, Erdinç: "Analitik Zihin Felsefesinin Temel Problemlerine Bir Bakış", **Kaygı**, No: 19, 2012, s. 37-54.

Schank, Roger ve Abelson, Robert: **Scripts Plans Goals and Understanding: An Inquiry into Human Knowledge Structures**, New Jersey, Lawrance Erlbaum Associates, 1977.

Schwartz, Jacob T.: "The New Connectionism: Developing Relationships Between Neuroscience and Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 123-141.

Searle, John: "Is the Brain's Mind a Computer Program?", **Scientific American**, C. CCLXII, No: 1, 1990, 25-31.

"Free Will as a Problem in Neurobiology", **Philosophy**, C. LXXVI, No: 298, s. 491-514.

"Minds, brains, and programs", **The Behavioral and Brain Sciences**, No: 3, 1980, s. 417-457.

Minds, Brains and Science, 13. bs., Cambridge, Harvard University Press, 2003.

Sejnowski, Terrence J. "The unreasonable effectiveness of deep learning in artificial intelligence", **Proceedings of the National Academy of Sciences**, C. CXVII, No: 48, 2020, s. 30033-30038.

Shannon, Claude Elwood: "A Symbolic Analysis of Relay and Switching Circuits", Cambridge, Massachusetts Institute of Technology, Yayımlanmamış Yüksek Lisans Tezi, 1940.

Simon, Herbert A.: "Artificial Intelligence: an empirical science", Artificial Intelligence, No: 77, 1995, s. 95-127.

Sloman, Aaron: **The Computer Revolution in Philosophy: Philosophy Science and Models of Mind**, Sussex, Harvester Press, 1978.

Smith, Peter: **An Introduction to Gödel's Theorems**, C. II, Logic Matters, 2020.

Sokolowski, Robert: "Natural and Artificial Intelligence", **Dædalus**, C. CXVII, No: 1, 1988, s. 45-64.

Stone, Peter v.d.,: "Artificial Intelligence and Life in 2030", One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, (Çevrimiçi) https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/ai100report10032016f nl_singles.pdf, 9 Aralık 2022.

Strawson, P. F.: **Individuals: An Essay in Descriptive Metaphysics**, New York, Routledge, 1996.

Şengör, Neslihan Serap: "Bilinç, Farkında Olan…", **Özgür İrade / Hukuk, Nörobilim, Psikoloji ve Ötesi**, ed. Ozan Erözden, Salih Haydar Günler, Mustafa Ebrar Palteki ve Cem Uysal, İstanbul, Zoe Yayıncılık, 2021, s. 96-110.

Taşdemir, Özgür; Özbay, Ümit Vefa ve Kireçtepe, Burhanettin Onur: "Robotların Hukuki ve Cezai Sorumluluğu Üzerine Bir Deneme", **Ankara Üniversitesi Hukuk Fakültesi Dergisi**, C. LXIX, No: 2, 2020, s. 793-833.

Tegmark, Max: **Yaşam 3.0**, çev. Ekin Can Göksoy, İstanbul, Pegasus Yayınları, 2019.

Thomason, Richmond: "Logic and Artificial Intelligence", The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta, (Çevrimiçi) https://plato.stanford.edu/entries/logic-ai/, 2 Ocak 2023.

Turing, A. M.: "Intelligent Machinery", **Mechanical Intelligence: Collected Works of A. M. Turing**, ed. D. C. Ince, C.I, North Holland, 1992, s. 100-127.

"Computing Machinery and Intelligence", **Mind**, C. LIX, No: 236, 1950, s. 433-360.

"On Computable Numbers, with an Application to the Entscheidungsproblem", **Proceedings of the London Mathematical Society**, C. XLII, 1936, s. 230-265.

Türkyılmaz, Çetin "Spinoza'da Özgürlük-Zorunluluk Bağlantısı", **Yeditepe'de Felsefe**, No: 6, 2007, s. 176-188.

Twain, Mark: **İnsan Nedir?**, çev. Esra Damla İpekçi, 11. bs., İstanbul, Dedalus, 2020.

Ural, Şafak: **Temel Mantık**, 4. bs., İstanbul Çantay Kitabevi, 2017.

Uspensky, Vladimir A.: "Gödel's incompleteness theorem", **Theoretical Computer Science**, C. LXXX, No: 2, 1994, s. S. 239-319.

Vernon, David ve Furlong, Dermot: "Philosophical Foundations of AI", **50 Years of Artificial Intelligence**, Berlin, Springer, 2007, s. 53-62.

Waltz, David L.: "The Prospects for Building Truly Intelligent Machines", **Dædalus**, C. CXVII, No: 1, 1988, s. 191-212.

Wang, Pei: "The Logic of Intelligence", **Artificial General Intelligence**, ed. Ben Goertzel ve Cassio Pennachin, C. II, New York, Springer, 2007, s. 31-62.

Warwick, Kevin: **Artificial Intelligence: The Basics**, New York, Routledge, 2012.

Weir, Alan: "Naturalism Reconsidered", **Oxford Handbook of Philosophy of Mathematics and Logic**, ed. Stewart Shapiro, New York, Oxford University Press, 2005, s. 460-482.

Weizenbaum, Joseph: **Computer Power and Human Reason**, New York, W. H. Freeman and Company, 1976.

Wiener, Norbert: "Some Moral and Technical Consequences of Automation", **American Association for the Advancement of Science**, C. CXXXI, No: 3410, 1960, s. 1355-1358.

Wigner, Eugene P.: "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", **Communications on Pure and Applied Mathematics**, C. XIII, 1960, s. 1-14.

Winston, Patrick Henry: **Artificial Intelligence**, 3. bs, Reading-Massachusetts, Addison-Wesley Publishing Company, 1993.

Zambak, Aziz F.: "Düşünen Makineler ve Yeni Bir Metafizik Tasarım Olarak Yapay Zekâ", **Bilgisayar ve Beyin**, ed. Haluk Bingöl, 2. bs., İstanbul, Pan Yayıncılık, 2012, s. 383-392.

"Free Will and Artificial Intelligence", **MetaZihin**, C. I, No: 2, 2018, s. 167-181.

Zambak, Aziz ve Vergauwen, Roger: "Artificial Intelligence and Agentive Cognition: A Logico-Linguistic Approach", **Logique et Analyse**, C. LII, No: 205, 2009, s. 57-96.

"M-S Models: A New Approach to Models and Simulations in Artificial Intelligence", **Logique et Analyse**, C. L, No: 198, 2007, s. 179-206

Zeman, Adam: **Bilinç Kullanım Kılavuzu**, çev. Gürol Koca, 3. bs., İstanbul, Metis Yayınları, 2017.