

GUÇLU YAPAY ZEKAYA KARŞI ÇİN ODASI ARGUMANI

The Chinese Room Argument against the Strong Artificial Intellingence

*Fatma DORE **

OZET

Bu makalede, Amerikalı filozof John Rogers Searle'ün Çin Odası Argümanı olarak bilinen ünlü düşünce deneyi incelenmiştir. Argüman, Searle tarafından, işlevi; yapay zeka kullanmaya dayalı bilgisayarların, aslında bilinçli olarak düşünmediklerinin bir kanıtı olmak üzere geliştirilmiştir. Makalede, Çin Odası Argümanı'nın detaylarına girilmeden önce, Searle'ün muhalif olarak düşünce deneyini geliştirdiği Alan Turing'in Turing Testi'ne bakılmıştır. Makalenin son bölümü, Searle'ün görüşünün doğurabileceği etik etkiler temelinde, bilgisayar bilinçliliği üzerinde odaklanmıştır.

Anahtar kelimeler: Çin Odası Argümanı, Turing Testi, yapay zeka, bilinçlilik, robot hakları.

ABSTRACT

This article is on the American philosopher John Rogers Searle's famous thought experiment known as the Chinese Room Argument. It was developed by Searle to be a proof that computers, which function using artificial intelligence, do not actually consciously think. Before examing the Chinese Room Argument in some detail, this article first looks at Alan Turing's Turing Test which Searle developed his thought experiment in opposition to. The final part of the article focuses on the ethical implications of Searle's position on computaer consciousness.

Keywords: Chinese Room Argument, Turing Test, artificial intelligence, consciousness, robot rights.

GİRİŞ: Bilgisayar Teknolojisinde Büyük İlerleme

20. yüzyılın ortalarından (1940-1945) itibaren, geliştirilen ilk elektronik bilgisayarlar, büyük bir oda büyüklüğünde olup, çalışmak için yüzlerce modern kişisel bilgisayarın kullandığı güç miktarında enerjiye ihtiyaç duyuyorlardı. Hantal, ağır ve karmaşık olan bu makineler, kolay kullanıma pek elverişli değildi. Bilgisayar teknolojisi tarihi, 1970'lerin sonunda icat edilen "mikroçip" adlı elektronik cihaz ile tamamen farklı bir döneme girmiştir (Swaine, 2005). 5 mm'den de küçük bir silikon yapıya bastırılan ve binlerce transistor, direnç ve diğer devre elementlerini içeren bu entegre devre, modern bilgisayarların kalbini oluşturmaktadır. Son yıllarda bilgisayarlar giderek daha hızlı çalışan, küçülüp hafifleşen mikroçipler sayesinde şaşırtan gelişmeler göstermeye; belli alanlarda insan zekasıyla yarışmaya, hatta onu geçmeye başlamıştır.

Peki insan zekası, gerçekten de bilgisayar zekasının gerisinde kalacak mı? Bu soruya belli bir düzeyde "evet" denebilir. Nitekim "insana karşı makine" konseptinde ilk akla gelen örneklerden biri, dünya satranç ustası Garry Kasparov'un, IBM'in süper bilgisayarı Deep Blue ile yarışması olayıdır. 1996 yılında, adı geçen bilgisayarı yenen Kasparov, bir yıl sonra 1997'de, IBM tarafından güncellenen bu bilgisayara yenilmiştir (BBC News, 2003).

2008 yılının Şubat ayında, *Independent Gazetesinde* yayınlanan bir makalenin iddia ettiğine göre, yaklaşık 20 yıl içinde, artan bir hızla gelişen teknoloji sayesinde bilgisayarlar, insan aklını geçecektir. Makaleye göre, bilgisayarlar kendi aralarında iletişim kurma, kopyalama ve öğretmeyi öğrenmek suretiyle, insan zekasını çok geride bırakacaktır (Connor, 2008). Bu noktada, yakın bir zamanda, IBM'in bir başka süper bilgisayarı Watson'un, "jeopardy" adlı sözcük oyununa dayalı yarışma programında, insan hasımlarını yenilgiye uğratmış olduğu olay hemen akla gelebilir. Bir bilgisayar için oynaması çok kolay olmayan, kuralları oldukça karmaşık bir oyun olan "jeopardy"de, hem anahtar sözcükleri eşleştirmek gibi doğrudan hem de eş anlamlı, çok anlamlı (örneğin, "yüz" sözcüğünde olduğu gibi birden çok anlama gelebilen) ve istatistiksel içerikli daha karmaşık bir çok soruyu yanıtlayarak zafer kazanan Watson'un, birçok unsuru gözden geçirmesi gerekmiştir (Pepitone, 2011).

Aradan geçen yaklaşık on yıllık zaman göz önüne alındığında, bu iki bilgisayar arasındaki gelişme farkı önemlidir: Deep Blue, kuralları basit ve karmaşık olmayan, ancak oynanması yine de güç olan satranç oyununu öğrenmiş ve

Acaba, insan bir gün gerçekten bilinçli makineler üretebilecek seviyeye ulaşabilecek mi ve bu makineler gerçekten insanın yerini alacak mı? Gelecekte bir robot düşünelim ki, eve geldiğimizde bize kapıyı açsın, oturup günümüzün nasıl geçtiğini ya da buna benzer sorular sorarak bizimle sohbet etsin. Bu varsayım altındaki sorumuz şu olacaktır: “Bu makine, tartışmasız olarak basit bir hesap makinesi ya da bir satranç programından tamamen farklı görünse de, bunlar, özünde gerçekten farklı şeyler midir?”

Son 50 yıldır, makinelerin -ki bununla kastedilen donanımlı sayısal bilgisayarlardır- düşünebilme becerisi üzerine tartışmalar sürüp gitmektedir. Söz konusu tartışmalar genel olarak, “Bilgisayar programı yüklenmiş bir makine düşünebilir mi?”; “İnsan eliyle yapılmış cansız bir şey, canlı ve duyarlı olan insan gibi düşünme özelliği kazanabilir mi?” gibi soruların çevresinde dönmektedir. Bu sorulara tereddütsüz bir şekilde “hayır” yanıtını veren John Searle’e (d. 1932) göre makineler, ancak bilince de sahip olduğu zaman insanın yerini alabilir. Bu anlamda süper bilgisayarlar ile elektrikli tost makineleri arasında hiçbir fark görmeyen Searle, bu düşüncesini “Çin Odası Argümanı” adlı alegorik anlatımda açıklamaya çalışır.

Bu çalışmanın konusu, bu argümanın detaylı olarak incelenmesidir. Ancak bunu anlamak için, öncelikle Searle’ün tartışmasını hangi nedenle geliştirdiğini bilmek gerekir. Searle, Çin Odası Argümanı’nda özellikle Güçlü Yapay Zeka kavramına, genel olarak da Turing makinelerine karşı bir tez yürütür. Bu nedenle, önce Turing’in sayısal bilgisayarlar hakkındaki görüşü

1. TURİNG TESTİ ve TAKLİT OYUNU

Makinelerin, düşünebilme özelliğine sahip olup olmadığını, “Turing Testi”ni başarıyla geçmelerine dayandıran matematikçi ve filozof Alan Turing (1912-1954), bunun için bilgisayarların biyolojik olarak insan ile aynı özellikleri paylaşmasını gerekli bulmaz (Rapaport, 2003: 164). Turing, 1950 yılında yayınlamış olduğu ünlü “Bilgisayar ve Zeka” (Computing Machinery and Intelligence) adlı yazısında şöyle demiştir: “Yaklaşık 50 yıl içinde, 10⁹ kapasiteli hafızayla üretilen bilgisayarlar, “Taklit Oyunu”nu o kadar iyi oynayacaklardır ki, ortalama bir sorgulayıcının, soru sorulduktan beş dakika sonra doğru kimliği bulma şansı %70’ten daha az olacaktır” (Turing, 1950: 442). Şimdi “Taklit Oyunu” ve “Turing Testi” açıklanacaktır.

Turing, yukarıda adı geçen makalesinin başında, “Bilgisayarlar düşünebilir mi?” diye sorar ve hemen ardından, böyle bir sorunun karmaşık sorunlar çağrıştıran bir nitelik taşıdığını belirterek, “Taklit Oyunu”

(Imitation Game) adını verdiği uygulamalı bir düşünce deneyi önermek suretiyle sorunun şeklini değiştirir.

Taklit Oyunu, her biri ayrı odalara yerleştirilmiş bir erkek (A), bir kadın (B), bir de kadın ya da erkek olabilen bir sorgulayıcıdan (C) oluşan ve sadece yazı ekranı aracılığıyla birbiriyle iletişim kurabilen üç unsur içermektedir. Bu oyuna göre, beş dakikalık görüşmeden sonra sorgulayıcının (C), vardığı sonuçtan tam olarak emin olamaması durumunda bile bilgisayar testi geçmiş sayılacaktır (Turing, 1950: 442).¹ Oyunda C'nin görevi, X ve Y bilinmeyenlerini kullanarak "X, A'dır (yani erkektir) ve Y, B'dir (yani kadındır); ya da bunun aksine "X, B'dir ve Y, A'dır" şeklinde karar vererek, A ve B'den hangisinin kadın olduğunu bulmaktır. Öte yandan, oyunda A'nın görevi, C'yi bir kadın olduğu konusunda yanıltmak; B'nin görevi ise, C'yi kendisine, yani zaten bir kadın olduğuna inandırmaktır (Turing, 1950: 443).

Şimdi Turing, oyunda sorgulayıcının haberi olmaksızın A'nın yerine geçirilen ve sorgulayıcıyı yanıltmayı başaran bir bilgisayarı tartışmaya başlar. Bu noktada "Bilgisayarlar düşünebilir mi?" sorusunun yerine şu soruyu sorar: "Acaba A'nın yerini bir bilgisayar alırsa ne olur?" ya da "Acaba, sorgulayıcının sadece insanlardan oluşan düzenekte varacağı yanlış karar sıklığı, bir bilgisayarla bir insandan oluşan yeni düzenekte varacağı yanlış karar sıklığı gibi mi olacak?" (Turing, 1950: 442). Turing'e göre bilgisayar, oyunu başarıyla oynamış olduğundan testi geçmiş ve "zeki" kabul edilmiş sayılacaktır.

Turing, "Bilgisayar ve Zeka" adlı bu makalede biri kuralları yukarda açıklanan "Taklit Oyunu" olmak üzere iki test önermiştir. "Standart Turing Testi" olarak bilinen ikinci testin kuralları da, birincisinden çok farklı değildir. Her iki test de, genel olarak bir bilgisayarın bir insanı kandırması prensibine dayanmaktadır. Standart Turing Testinde, başından beri, sorgulayıcı karşısında biri insan diğeri bilgisayar olan bir düzenek kurulur. Burada da, doğal dilde ve yazı ekranı aracılığıyla gerçekleştirilecek görüşme sonucunda, sorgulayıcının bu kez, karşısındaki X ve Y'den hangisinin insan, hangisinin bilgisayar olduğunu bulması istenecektir (Sterrett, 2003: 79-81).

Bilgisayarların doğal dil kullanarak, bir insan tarafından kolaylıkla ayırt edilemeyecek derecede iletişim kurabildiklerine inanan Turing'e göre Taklit Oyunu'nu başarıyla oynatabilecek

'Sorgulayıcının, "ortalama" bir insan olarak belirtilmesi, bu testin önemli noktalarından biridir. Nitekim Turing, bu testlerle

bakışları göz önüne aldığı anlaşılan Turing, öteki insanların düşünmesine dair bilgimizin de, zaten tümevarımsal bir temele dayandığını belirterek, insanın düşünmesi konusunda soliptik bir tavır sergiler (Turing, 1950: 446).

Turing'in ortaya koyduğu test; beynin bir makine, makinelerin ise beyin oldukları varsayımına dayanmaz. Bununla beraber Taklit Oyunu, bilgisayarların herhangi bir oyun daha öğrenip öğrenmediği ile de ilgili değildir. Burada test edilen şey, olası bilgisayar etkinlikleri ile bilinen anlamda düşünme kavramı arasında bir bağ olup olmadığıdır (Moor, 1976: 250). Düşünmek bilme, imgeleme, değerlendirme ve karar verme gibi bilişsel bir süreci kapsar ve ortalama bir bilgisayar bunların tümünü gerçekleştirebilmededir. O halde Turing için “düşünme”

kavramının

kullanım ve çağrışımları, hem insan hem de bilgisayarlar için eşit şekilde geçerlidir (Rapaport, 2003: 166). Nitekim bu oyun, bilgisayarların

düşünebildiği hipotezini olumlayabilecek potansiyel bir tümevarımsal sonuç için güçlü kanıtlar sunmak anlamında önemlidir (Moor, 1976: 249). Sonuç olarak testten geçen bilgisayarların, tabii tutuldukları dolaylı ya da doğrudan çok geniş tabanlı ve detaylandırılmış bir konular silsilesinden başarıyla geçmek zorunda olmalarını, normal yetişkin bir insan düzeyinde düşünebildiklerinin kanıtı olarak gördüğü söylenebilen Turing, makine ile insan arasında bir çeşit kıyaslama yapar. Bilgisayarın yetişkin bir insanın aklını taklit etmesini, bir çocuğun aklının eğitilerek yetişkin akıl biçimini kazanmasıyla karşılaştıran matematikçi, bu kıyaslamada çocuğun beynini çok az mekanizma ve bir sürü boş sayfa ile nitelediği mağazadan yeni alınmış bir laptopa benzetir (Turing, 1950: 443-455; Saygın vd., 2003: 3132).

Şüphesiz, testten başarılı olmakla bilgisayarın “zeki” olduklarının kanıtlandığına inanan Turing, bununla onların bilinçli de olduklarının kanıtlandığına dair bir yargıda bulunmamıştır (Saygın vd., 2003: 45-47). Yine de, Turing'e göre düşünme ile bilinç arasında bir tür ilişki olduğu söylenebilir (Turing, 1950: 447).

2. ÇİN ODASI ARGÜMANI: BİR DÜŞÜNCE DENEYİ

Turing Testini geçen bilgisayar için, “Ama o, gerçekten düşünmüyor; düşünme için bir Turing Testi’ni geçmekten daha fazla şey gereklidir” diyerek, Turing'in düşüncesine karşı çıkan Searle'e göre zeka sahibi olmak, sadece bilinçli olan varlıklarla sınırlıdır (Searle, 2004: 133). Bilgisayar bilimi ve yapay zeka konusunda yakın zamanda yapılan çalışmaları zeki makineler

Argümanı” ile açıklamaya çalışır (Searle, 1996: 44-45). Searle’ün ilk kez 1980 yılında yazdığı “Akıllar, Beyinler ve Programlar” adlı yazısında dile getirdiği ve müteakip çalışmalarında da, hep kullandığı bu argüman, ne kadar karmaşık olursa olsun, bilgisayarların anlamaktan yoksun olduklarını, dolayısıyla asla düşünemeyeceklerini kanıtlama girişimidir.

Çin Odası Argümanı, İngilizceden başka dil bilmeyen bir kişinin tek başına bir odaya kapatılmış olduğu varsayımıyla başlar. Bu kişinin eline, üzerinde Çince semboller olan bir dizi kart ile bir İngilizce kurallar kitabı verilmiştir. Varsayım şöyle devam etmektedir: Odadaki kişi, kapı altından içeri atılan ve üzerinde Çince sorular yazan bir kağıdı alıp, üzerinde bu soruların Çince yanıtları olan bir başka kağıdı kapı altından dışarı atar. Sorulara verilen yanıtlar, ana dili Çince olan birinin ayırt edemeyeceği kadar kusursuzdur (Searle, 1980: 3).

Şimdi bu alegoride anlaşılması gereken en önemli nokta, odadaki kişinin, bu yazıların içeriğini hiçbir şekilde anlamamış olmasıdır. Kartlardaki sembollerin tanınması, tamamen biçimleriyle sınırlı olup, kart dizileri arasındaki bağdaştırma da kurallar kitabı yardımıyla sağlanmıştır. Dolayısıyla odadaki kişi, ne olup bittiğini bilmeden, sadece verilen kuralları izleyerek kendisinden istenen “hiçbir şey anlamadığı semboller”i elle manipüle etmiş ve bunu “farkında olmadan” yapmıştır.

Searle, bu argüman ile bir sistemin insanın bazı yetilerini mükemmel bir taklit (simulation) verecek şekilde bir program geliştirebileceğini gösterdiğini düşünür (Searle, 2004: 68). Ona göre, odanın dışındakiler için odadakinin mükemmel Çince biliyor görünmesi ile Turing Testi’ni geçen bilgisayarın sorgulayıcıyı “yanıltabilmiş” görünmesi aynıdır. Çince yazılı bu “bir dizi biçimsel sembolün diğer bir dizi biçimsel sembolle bağdaştırması” istenen kişinin, Searle’ün kafasında Turing Testi’ni geçen bilgisayar olduğu çok açıktır (Searle, 1980: 2). Bu bağlamda odadan içeriye gönderilen kartlar, “sorular”; odadan dışarı gönderdiği kartlar ise “soruların yanıtları” olarak kabul edilen semboller; “İngilizce kurallar kitabı” bilgisayar programı ve kitabı yazarlar programcılar olarak belirlenebilir.

“Biçimsel” (formal) ile, “sözdizimsel”i (syntax) kasteden Searle; sözdizimini, simgelerin belli kurullarla mantık kuralı çerçevesinde belirli bir yerleşim düzeni şeklinde tanımlar. Searle’e göre, Güçlü Yapay Zeka sadece anlamları olmayan simgelere ve sözdizimine sahiptir; o halde niyet taşıyan akılsal içeriklere sahip değildir; gerçekten zeki değildir ve gerçekten düşünmüyordur (Hauser, 2005: 1). Farzedelim ki der, odaya gönderilen notta “En sevdiğiniz renk nedir?” diye soruluyordur

öğrenilmesinin imkansız olacağına inanan Searle'e göre, bilgisayarlar kendi başlarına algılama, anlama, kavrama ve düşünme gibi özelliklerden yoksun olduklarından, Çinceyi ve bunun gibi öteki bilişsel formların hiçbirini anlayamayacaklardır (Searle, 1990: 26).

Çin Odası Argümanı ile "Güçlü Yapay Zeka" kavramına karşı olduğunu vurgulamış olan Searle, "Güçlü Yapay Zeka bizden, oda sıcaklığını düzenlemesi için kullandığımız duvara asılı bir "metal yığını"nm ve odadaki öteki birçok makinenin, tıpkı bizimki, eşlerimiz ve çocuklarımızınki gibi inançları olduğunu kabul etmemizi istiyor" diyerek, bu konudaki düşüncesinin altında yatan asıl endişesini dile getirir (Searle, 1980: 7). Yapay zeka kavramını "Güçlü Yapay Zeka" ve "Zayıf Yapay Zeka" şeklinde iki gruba ayırması, Searle'ün genel olarak bilgisayarların zeki işler yaptıklarını tartışmadığı anlamına gelmektedir. Onun tartışmasının merkezine aldığı Güçlü Yapay Zeka, bir bilgisayar programının akıl da üretebileceğini ileri süren, özellikle Turing makinelerinin savunduğu yaklaşımdır (Searle, 1996: 26). Searle; zihni, hesaba dayalı bir modele oturtarak bir bilgisayar programı olduğu fikrini öne süren Güçlü Yapay Zeka yaklaşımlarını, asıl olarak zihinle ilgili bilinç ve niyetlilik gibi önemli şeyleri dışarıda bıraktığı için reddeder (Searle, 2004: 67-68).

Çin Odası Argümanı'nın özü oldukça basittir: Biçimsel bir bilgisayar programı işletilirken, dışta bulunan bir gözlemci bakış açısından odadaki kişi tek bir sözcük bile Çince anlamadığı halde, sanki Çince anlıyor gibidir. Bir başka ifadeyle, Çince düzenlenmiş bir Turing Testi'nde sembolleri anlayan ve bunu başarıyla geçen bir bilgisayarın doğal olarak düşündüğünü ileri süren Turing taraftarlarının aksine; Searle'e göre, programlanmış bir bilgisayar, geçerli girdi ve çıktılar sayesinde mükemmel bir Çince performansla testi geçebilse bile, Çince tek bir karakteri dahi anlamayacaktır. Öyleyse Çince anlamak için en uygun bir program bile Çince anlamayı sağlayamıyorsa, herhangi bir sayısal bilgisayarın da Çince anlaması olanaklı değildir (Searle, 1996: 45; Saygın vd., 2003: 44-45; Rapaport, 2003: 166-167). Demek ki, bir program çalıştırmakla bir bilgisayarın Çince anlaması sağlanamaz. Searle, tartışmasını yürütürken, belli tür ya da özellikte bir bilgisayarı tanımlamaz. Bu bakımdan bilgisayar teknolojisinde ileride mümkün olabilecek şaşırtıcı gelişmeler de bir önem arzetmez; Çin Odası Argümanı, Güçlü Yapay Zeka programlarına karşı genel geçer bir yanıt olmayı hedefler (Searle, 1996: 51).

Daha önce Turing'in, "Makineler düşünebilir mi?" diye bir soru sorduğu, ancak böyle bir soruyu belirsizlik içermesi

şey, “İnsan eliyle yapılmış olan bir şey düşünülür mü?” sorusudur (Searle, 1996: 49). Searle’ün buna yanıtı hazırdır: “Ancak nedenleriyle birlikte etkileri de kopyalanabilen, yani insanla aynı yapıda bir makine yapılabilirse, olasılıkla bu makine düşünebilecektir.” Şimdi sorunun, “Sayısal bilgisayar düşünebilir mi?” şeklinde olması gerektiğini belirten filozof, bu soruyu da, kesin bir “Hayır!” ile yanıtlar. Bunun nedeni ise; bilgisayar programının, tümüyle sözdizimsel olarak tanımlanabilir olması ve bu tanımın, anlamsız sembolleri anlamlı bir içerikle doldurmak demek olan düşünmeyi içermemesidir. “Anlamlı içerik” için “anlam” derken neyi kast ediyorsak odur,” diyen Searle için; şimdilik yeterli kapasitede bilgisayar üretilmemiş olması hiç önemli değildir (Searle, 1996: 51); sonuçta, en iyi ihtimalle önümüze konacak olan makine, içinde beyin işlemlerinin bir hesaplamalı taklidinin (imitation) yapılabileceği, yani beyin işlemlerinin

canlandırılabilceği bir dijital ortamdan ibaret olacaktır (Searle, 2004: 246248).

Searle’e göre, zihin ile dünya arasındaki nedensel ilişkinin dile getirilebilmesi için, minimum düzeyde bile olsa sembollerin anlamının olması gerekir. O halde anlam için biçimsel sentakstan başka bir şey gereklidir ve bu “başka bir şey” onun deyişiyle biyolojidir (nörofizyoloji). İnsan beyni ile ilgili olarak mühim olan, beyin formel bir yapıya sahip olması değildir; nitekim öyle bir biçimsel yapıyı herhangi bir şeye koyabilirsiniz, bu belli şekillerle bir araya getirilmiş bira kutuları topluluğu olsa bile... Asıl önemli olan şey, özel bir biyokimyadır. Beyin, kendine özgü biyokimyası ile belirli duygu ve düşüncelere neden olur. Öyleyse, zihinlerin içsel zihinsel içeriklerinin varlığını kanıtlayamadığından, programlar tek başına zihni oluşturamaz; çünkü bir program yalnızca sembolleri manipüle edebilirken, beyin aynı zamanda bu sembollere anlamlar yüklemektedir (Searle, 2004: 246; 1996: 26).

Gerçekten de Çince bilen dışarıdaki herhangi bir kişi, odadakinin (bilgisayar) “Çince anlamadığını” kavramakta güçlük çekecektir. Bu noktada, değişik tür ve düzeylerde anlamlardan söz edilebileceği için “anlama” ile ilgili yanlış düşünceleri bertaraf etmek gerektiğine dikkat çeken Searle, “Otomatik kapı, ışık gözü (fotoelektrik hücre) sayesinde ne zaman açılacağını anlar.” ifadesiyle, “İngilizce anlıyorum.” ifadesindeki anlamın aynı olmadığını belirterek konuya açıklık getirir: “Bilgisayarın (örneğin Almanca) anlaması, benim (örneğin Almanca) anlamamın daha azı ya da bir kısmı şeklinde betimlenemez; onun anlaması sıfırdır” (Searle 1980: 4)

2. Aksiyom: İnsan zihni, anlamsal içeriklere sahiptir.
3. Aksiyom: Sözdizim tek başına anlam için yeterli değildir.

Sonuç: Programlar tek başına zihin oluşumu için kurucu nitelikte bir

yeterliliğe sahip değildir (Searle, 1990: 27). Elbette, Searle'ün bu tanımları ve sınırlamaları, yaşadığımız dünya ile ilgilidir. Yoksa, başka bir dizgede, beynin kullandıklarından bütünüyle farklı, akıllı varlıkların var olabileceğini belirten Searle şöyle bir açıklama yapar: "Marslıların Dünya'ya geldiklerini ve onların akıllı yaratıklar olduğunu ve bunların kafatasları açıldığında içlerinde yapışkan, yeşil bir madde bulunduğunun keşfedildiğini düşünün. Bu madde, bilinç ve bütün öteki akılsal durumları üretiyorsa; o zaman onların beyinlerinin de insan beynindekilerle eşdeğerde rastlantısal güçleri olduğu düşünülebilir" (Searle, 1996: 58).

İlke olarak bir laboratuvar ortamında beynin nedensel güçlerini kopyalayarak yapay bilinç üretmenin mümkün olabileceğini ve birçok biyolojik görüngünün zaten yapay olarak üretildiğini belirten Searle'ün burada dikkat çekmek istediği asıl nokta; bilince neden olma yeteneği olan bir sistemin, aynı zamanda beynin nedensel güçlerini kopyalama yeteneğinin de olması gerektiğidir. Ancak bilinçli zihin durumları ve süreçleri diğer doğal görüngülerin sahip olmadığı özgün bir niteliğe, yani öznelliğe sahiptir (Searle, 2004: 126).

Akılsal durumların öznelliğini de; belirli türden biyolojik dizgeleri, yani insanın ve belirli hayvanların öznel özellikleri olan beyinlerinin oluştuğu biyolojik yaşamın basit bir gerçeği olarak açıklayan Searle, bu bilinç durumunu başkası tarafından "nüfuz edilemez" bir alan olduğunu belirtir. Bireysel bilinç durumlarına ulaşamazlık, Searle'e göre biyolojinin nesnel bir gerçeğidir (Searle, 1996: 19-21,34) ve makineler için mümkün olmayan düşünme ve anlama sürecini sağlayan bilinç, beynin doğal bir özelliğidir (Searle, 2004: 125). Bu anlamda Searle'ün, Çin Odası Argümanı, bilincin indirgenemezliğinin bir savunusudur (Prado, 2006: 31). Searle'ün burada yapmaya çalıştığı şey, "Anlamsız olan dünya, nasıl anlamı da içerebilir?" sorusunu, aslında bütünüyle farklı iki şey arasındaki ilişkinin anlaşılmasını hedefleyen bir akıl-beden sorununu çözümlenmeye çalışmaktır. Searle'e göre sorunun yanıtı oldukça basit, doğal ve gizemsizdir: Bu iki alandan birini yadsımak ya da birini ötekine indirgeme yaklaşımlarının tersine, meseleyi mide-sindirim sistemi ilişkisi gibi düşünmek gerekir (Searle, 1996: 18). Akıl-beden sorununun, çözümlenmesi karmaşık bir konu şeklinde algılanmasının nedeni olarak Searle, akılsal olguların dört belirgin özelliğini gösterir:

4. Aklın nedenselliği

Her şeyden önce bilinç, bir “madde” değil; bilakis tıpkı, sıvılığın suyun bir özelliği olması gibi beynin bir özelliği ya da niteliğidir (Searle, 2004: 139). Peki “Bilinçlilik nasıl mümkün olur?” Searle’e göre, bu bilinmezlik ancak beyinde olup biten süreçlerin anlaşılması ile giderilebilir. Her ne kadar bu süreçler tam olarak henüz anlaşılammış olsa da, bilinçliliği sağlayan bu süreçlerin genel özyapısının, nöronlar ve nöron modülleri arasında oluşan elektro-kimyasal alışverişlerin ve beyin öteki özellikleri anlaşılmalıdır ve bu yeterlidir (Searle, 1996: 32).

Zihnin en temel özelliği olan bilincin ayrılmaz parçası ise, niyetliliktir. Searle’e göre, niyetlilik belirli bir akılsal tipin içeriğinden kaynaklanır; çünkü görsel ve işitsel deneyimler, dokunma duyumu, açlık, susuzluk ve cinsel istek gibi niyetli hareketlerde devinim önce akıl tarafından düşünülür (Searle, 1996: 33-89). Searle, bilinçlilik ve özellikle niyetlilik kavramlarını (1) beyin, bilincin oluşmasının nedenidir (2); bilinç, niyetliliği sağlar, önermeleriyle birbirine bağlarken, arada boşluk bıraktığı öne sürülerek, dualist bir tavır içinde olduğu yönünde eleştirilmiştir (Hauser, 2005: 6). Kendisi de, farkında olduğunu itiraf ettiği bu boşluğu “Ancak bilinçli ve niyetli olan durumlara sahip olabilen varlık, niyetli durumlara sahip olabilir ve her bilinçdışı ve niyetli olan durum en azından potansiyel olarak bilinçlidir” ifadeleriyle kapatmaya çalışmıştır (Searle, 2004: 169).

Fiziksel dünya ile zihin arasında köprü kurarken, Searle’ün dayandığı temel, akılsal durumların özelliklerinden biri olan akılsal nedenselliklerdir. Bunu, “Düşünce gibi ağırlıksız ve cismi olmayan herhangi bir şey, bir eyleme nasıl yol açar?” sorusu doğrultusunda tartışır. Searle’e göre, “Düşünceler, ağırlıksız ve cisimsiz değildir. Bir düşüncemiz olduğu zaman, beyin etkinliğimiz sürmektedir. Beyin etkinliği ise, fizyolojik süreçler aracılığıyla bedensel devinime yol açar.” Bu nedenle, beyin özellikleri olması dolayısıyla akılsal durumlarımız iki düzeyde betimlenebilir:

1. Daha makro düzeyde akılsal terimlerle,
2. Daha mikro düzeyde bedensel terimlerle (Searle, 1996: 35-36).

Yani beyin aklın nedenidir, bununla beraber akılsal özellikler beyin mikro yapısında idrak edilir. Sonuç olarak, akıl ve beden karşılıklı olarak birbirini etkiler; ancak akılsal olgular beyin özellikleri olduğundan bunlar iki farklı şey değildir. Searle, akılsal durumların araştırılması söz konusu olduğunda, son adımın kafanın içine atılması gerektiğine inanır.

Çin Odası Argümanı’nda yürütülen; zihnin, biçimsel

sergilemenin merkezidir. Onlara göre, Searle'ün, ağn ve diğer akılsal olguları beynin, hatta merkezi sinir sisteminin özellikleri olarak tespit etmesi önemlidir; çünkü bu, sinirsel işlevin mikro ve makro düzey arasındaki ilişki ayırımının anlaşılmasını gösterir (Freeman and Skarda, 1993: 116-117). Hemen belirtmek gerekir ki, onların bununla, Searle'ün özellikle makro- düzey niyetlilik ile beynin mikro-düzey sinirsel süreci arasında kurduğu bağı destekledikleri sonucu çıkarılamaz. Burada önemli olan, sözü edilen nörobilimcilerin, Searle'ün beynin sayısal bir bilgisayar, daha da önemlisi, sayısal bilgisayarın insan beyni gibi olduğu düşüncesine karşı duruşunu desteklemiş olmalarıdır (Freeman and Skarda, 1993: 120-121).

Searle, insanın asıl düşünen parçasının beyin olduğunu açıkça ortaya koymak ister (Prado, 2006: 47). Beyindeki bilgiler daima öyle veya böyle bir duruma özgüdür. Örneğin düşünceye veya görmeye veya işitmeye veya dokunmaya özgüdür. Kognitif bilimin bilişe yönelik hesaplama modellerinde tanımlanan bilgi işlem düzeyi ise bir dizi girdi sembollerine tepki olarak bir dizi çıktı sembolleri elde etme sorunudur. Searle'e göre, görsel deneyim somut bir bilinçli olaydır ve belli elektro-kimyasal biyolojik işlemlerle beyinde üretilir. Bu olaylar ve işlemlerin biçimsel sembol hareketi ile karıştırmak, gerçekliği model ile karıştırmak demektir. Dolayısıyla kognitif bilimde kullanılan 'bilgi' anlamında beyin, bir bilgi-işlem aygıtı olarak kabul edilemez (Searle, 2004: 275-277).

Bu noktada "taklit" (simulation) ile "kopya" (duplication) kavramlarını tartışan Searle'e göre, sayısal bilgisayarlardaki durum kopyalama değil, benzetimini yapmadır. Bir bilgisayarın, akılsal olgular üretecekse, bunu sadece biçimsel bir program çalıştırmakla yapmasının olanaksız (Searle, 1990: 29) olduğunu tekrarlayan Searle'e göre bu, Turing'in neden olduğu en önemli kafa karışıklığıdır. Ona göre Turing, tasarlamış olduğu test ile bir şeyin belli bir akılsal davranış sergilemesinin, o şeyin gerçekten akılsal bit sürece sahipmiş gibi düşünülmesine yol açmıştır. Bu bağlamda (1) sentaks, semantik değildir (2) taklit, kopyalama değildir, önermeleriyle tamamen radikal iki mantıksal doğru temeline dayandırdığı Çin Odası Argümanı ile Searle, "Güçlü Yapay Zeka"ya doğrudan saldırarak bu noktayı belirgin kılmaya çalışmıştır (Dampier, 2004). İnsanın akılsal işlemler yaptığında gerçek anlamda düşünmekte olduğunu ve düşünmenin de, şu ya da bu düşünceyi üretmek anlamına geldiğini söyleyen Searle, insanların bu yetisiyle bilgisayarların yaptığı "sanki" akılsal süreçleri karıştırma eğilimini algılamacı anlayışa bağlar (Searle, 1996: 71). Açıktır ki, Searle'e göre Turing makinesi ya da "insan

Sonuç olarak, Searle'e göre sadece sembollerle yapılan hiçbir işlem anlam garantisi vermez (Searle, 1990: 28). Çin Odası'ndaki kişi, bilgisayarda çalışan çeviri programı gibidir; başarılı bir çeviri gerçekleştirmesine rağmen, çevirisi yapılan ifadelerin içeriğini anlamamaktadır (Prado, 2006: 48). Bilgisayarların dil bağlamında yaptığı şey de, sadece sentaktik bir gösterimdir; dolayısıyla dilin anlam, yani semantik yönüyle hiçbir ilişkisi yoktur.

3. ROBOT HAKLARI

Şu ana kadar, Turing ve Searle'ün yapay zeka konusunda ortaya koyduğu görüşler incelendi. Turing, uygun girdi ve çıktılarla donatılmış bilgisayarların, Turing Testi'nden geçmekle zeki olduklarını kanıtladığını ileri sürmüştür. Searle ise, "bilgisayarların düşünebildiğini" ima eden bu yaklaşımın yanlışlığını dile getirmiş ve bunu Çin Odası Argümanı ile kanıtlamaya çalışmıştır. Genel olarak, Searle'ün yöntem ve açıklamalarıyla ortaya koyduğu tartışma daha güçlü görünmektedir.

Ancak bu başlıkta ele alınacak konu açısından, bu iki tartışmadan hangisinin daha güçlü olduğunu değil de, hangisinin doğru olduğunu bilmek doğuracağı sonuçlar nedeniyle önemlidir. Böyle bir elemelerde, Karl Popper'in (1902-1994) "yanlışlanabilirlik" ilkesine başvurulacak olsaydı, ne Turing ne de Searle'ün ortaya koymuş olduğu iddialardan biri doğru kabul edilemezdi. Popper'a göre bir hükmün bilimsel olması için bu hükmün ilke olarak yanlışlanabilir olması gerekmektedir. Aksi halde, konu hakkında bağlayıcı rasyonel bir tartışma mümkün değildir (Skirbekk and Gilje, 1971: 545). Dolayısıyla deneye tabi tutulamayacaklarından her iki tartışmanın da yanlışlanabilir özellik taşımadıkları açıktır. Bu durumda böyle bir eleme ya Turing Testi ya da bunun daha gelişmiş bir versiyonunun uygulanması ile olurdu. Ancak böyle bir durumda da, teste tabi tutulacak olan bir bilgisayar (ya da robot) olacağından, Searle yine olumsuz bir yanıt verecek, Turing ise, bilgisayarın testi geçtiğini tekrarlayacaktı. Bu tartışmalara yanlışlanabilirlik ilkesinin uygulanması için geçerli tek yol, robotun kendisine sormak olurdu, ki bu defa da, robotun vereceği yanıtın bilinçli mi yoksa bilinçsiz mi olduğu konusunda sürüp gidecek yeni bir uzlaşsız tartışma doğacaktı.

Şu halde, asıl konu, odamıza insan gibi davranan bir robotun girmesi, bizim onunla herhangi bir konuda diyaloga girmemiz ve robotun da anlaşılır yanıtlar vermesi gibi bir

Bu konunun niçin önemli olduğuna gelince; bu, günümüzde insanların yerine çeşitli konumlarda robotların kullanıldığı² göz önüne alındığında, kısaca şöyle açıklanabilir:

1. Eğer makineler bilinçli ise, o halde onlar için “yaşamlı varlıklardır” denilecektir.

2. Eğer makinelerin bilinçsiz oldukları kabul edilirse, bu durumda onlar, sadece düşünemeyen birer araçlardır.

Bu bağlamda Descartes’i hatırlamak önemlidir. 17. yy’da Descartes, hayvanların ruhsuz otomatlar olduğunu düşünmüştür. Filozofun bu teorisine göre hayvanlar his ve duygu deneyiminden yoksun, bilinçsiz biyolojik makinelerdir (Descartes, 1968: 72-75). Descartes, hayvanların görünüşte havlayan ve miyavlayan, yiyen ve koşan halis robotlar olduklarını kanıtlamak için, üzerlerinde onları canlı olarak kesmek suretiyle bir takım deneyler de gerçekleştirmiştir (Grayling, 2005: 35, 159-160). Descartes’e göre, temel anatomik yapıları insaninkine benzer olsa da hayvanlar dahil, evrendeki her şey hareket yasalarına bağlı olup, bunun dışında herhangi bir amaç taşımaz. Bu anlamda tek başına insan bedeni de, hayvanlar ve diğer her şey gibi bir makineden farksızdır. Düşünme yetisine sahip olduğu için, sadece insan bu mekanizmanın dışında tutulmaktadır (Descartes, 1968: 65).

Kuşkusuz, insan açısından bir hayvanın acı çekip çekmemesi çok önemli olmayabilir; ancak hayvanlar açısından bunun önemi yadsınamaz. Aynı şekilde, robotların bilinçli varlıklar olup olmaması da, insanlar için önemli olmayabilir. Bilinçsiz bir robot, basitçe, bir tost makinesi ya da elektrik süpürgesi gibi insanların kullanımındaki bir araçtır. Ancak robotlar bilinçliyse ve insan için çalışmaya zorlanıyorlarsa, bu, onların tam anlamıyla “İnsana kölelik yapan bilinçli varlıklar” olduklarını gösterir. Güçlü Yapay Zeka taraftarları, bu ihtimali düşünmüşler ve konuyu gündeme taşıyarak, robotlar için yasal haklar arayışına girmişlerdir.³ Tabii ki, sırf programlanmış oldukları için bir robotun ya da robot grubunun hakları için insanlara yalvarır davranış sergilemesi de mümkündür. Örneğin, Japonya’da insana

² Bu konuda güncel bazı haberler için bk. Michael Fitzpatrick.

“No, Robot: Japan's Elderly Fail to Welcome Their Robot

Overlords” www.fox.com, Tokyo, 4 February 2011.

³ Bu konuda güncel bazı haberler için bk.

<http://news.bbc.co.uk/2/hi/technology/6200005.stm>.

⁴ Detaylı bilgi için bk. “A Robotic Performance: Japanese

Tabii, insan olma bakış açısından, bu konu olumlu ve olumsuz yönleriyle tartışılabilir. Turing'in iddiası ile Searle'ün bunu yadsıyan tavırları bu bakımdan iyi anlaşılmalı ve üzerinde düşünülmelidir. Aksi halde, insanoğlu ya mekanik bir köle sınıfı yaratmak ya da bilinçsiz varlıklara gereksiz yasal korumalar vermek gibi durumlarla karşı karşıya kalabilir. Bu düşünceler doğrultusunda toplumlar, elbette kendi tercihlerini yapacaklardır. Bu çalışma, bunun çok kolay olmadığını göstermeye çalıştı. Tartışması daha güçlü görünse bile, Searle'ün tezinde haksız olma riskinin, Turing'in yaklaşımına göre daha ağır sonuçlar getirebileceği belirtildi. Nitekim Searle, Çin Odası Argümanı ile Turing'in "düşünebilen" makine varsayımının yanlışlığını etkili bir şekilde ortaya koymuştur. Ancak Turing'in herhangi bir şekilde haklı olması ve bilgisayarların gerçekte bilinçli mekanik varlıklar olması durumunda, bunların bilinçliliğini reddetmek Searle'ü, robotlara karşı, bir çeşit 21. yy Descartes'i yapacaktır. Beyin ve biyolojik evrim tarihi paylaşımı nedeniyle, Searle'ün, hayvanlarda da bir dereceye kadar bilinç olduğunu kabul ettiğini belirtmek gerekir.

Kısacası, eğer Searle haklı ise, yani makineler düşünemeyen araçlardan başka bir şey değilse, bu durumda bizim onlara davranış biçimlerimizin etik bir boyutu yoktur. Çünkü onlar, sıradan bir tost makinesi ya da bir araba fabrikasındaki robot kolundan farklı değildir. Dolayısıyla robot hakları söz konusu olmayacaktır. Ancak eğer bilinçli varlıklar iseler, bu durumda onların insanlara hizmet etmede kullanılması kölelik ve hayvanlara yapılan davranışlara benzer bağlamda bir takım tatsız etik sorunlar ortaya çıkacaktır.

SONUÇ

Bu çalışmada, yapay zeka konusunda etkili iki filozofun görüşleri gösterilmeye çalışıldı. Bu bağlamda Alan Turing'in bilgisayar zekası anlamında, makinelerde yapay zeka yaklaşımı ile, bu yaklaşımı "Güçlü Yapay Zeka" olarak adlandıran ve reddeden Searle'ün, bilinçliliğe bakış açısı incelendi. Çalışmada öncelikle bilgisayarların öğrenme sürecini bir çocuğun eğitilmesi modeliyle karşılaştıran, dolayısıyla edindikleri deneyim ve kazanımlarla makinelerin de "düşünebileceğini" ima eden Turing ele alındı. Onun bu konuda önerdiği ve "Turing Testi" olarak adlandırılan ünlü yöntemi açıklandı. Ardından, temelde Turing'in

geliştirdiği bir düşünce deneyi olan “Çin Odası Argümanı”, yapay zeka alanında da en çok bilinen tartışmadır. Yukarıda bu alegorik tartışma geniş bir şekilde ele alındı.

Son olarak, birbiriyle örtüşmeyen bu iki yaklaşımın, muhtemel pratik sonuçlarına dikkat çekildi. Turing veya Searle’ün tartışmalarında haklı olmaları sonucunda insanların karşılaşabileceği sorunlara ve bu doğrultuda robot haklarından bahildi.

KAYNAKÇA

- BBC NEWS, (2003), “Kasparov Fights Computer to Draw”, <http://news.bbc.co.uk/1/1/technology/326303Lstm> (12 November).
- CONNOR, Steve. (2008), “Computers 'to match human brains by 2030’”, <http://www.independent.co.uk/life-style/gadgets-and-tech/news/computers-to-match-human-brains-by-2030-782978.html> (16 February 2008).
- DAMPER, Robert I. (2004), “The Chinese Room Argument--Dead but Not yet Buried”. *Journal of Consciousness Studies* 11 (5-6): 159-169, <http://eprints.ecs.soton.ac.uk/9561> (28 Jan 2009).
- DESCARTES, Rene. (1968), *Discourse on Method and The Meditations* (Translated with an Introduction by F. E Sutcliffe), Penguin Books.
- FREEMAN, Walter J. and Christine A. SKARDA. (1993), “Mind/Brain Science: Neuroscience on Philosophy of Mind”, (Eds.), Ernest Lepore and Robert Van Gulick, *John Searle and His Critics*, UK: Blackwell.
- FITZPATRICK, Michael. “No, Robot: Japan's Elderly Fail to Welcome Their Robot Overlords”, *BBC News*, Tokyo - February 2011, <http://www.bbc.co.uk/news/business-12347219> (3 Feb 2011).
- GRAYLING, C. (2005), *Descartes*, London.
- HAUSER, Larry. (2005), “Searle’s Chinese Room Argument”, <http://host.uniroma3.it/progetti/kant/field/chinesebiblio.html> (12 Apr 2001).
- MOOR, James H. (1976), “An Analysis of the Turing Test”, *Philosophical Studies* 30, Holland: D. Reidel Publishing, ss. 249-257.
- PEPITONE, Julianne, (2011), “IBM's Jeopardy Supercomputer Beats Humans in Practice Bout”, http://money.cnn.com/2011/01/13/technology/ibm_jeopardy

- RAPAPORT, William J. (2003), "How to Pass a Turing Test", (Ed.) James H. Moor, *The Turing Test: The Elusive Standard of Artificial Intelligence*, Netherlands: Kluwer Academic.
- SAYGIN, Ayşe Pınar; ÇİÇEKLİ, İlyas ve AKMAN, Varol. (2003), "Turing Test: 50 Years Later", (Ed.) James H. Moor, *The Turing Test: The Elusive Standard of Artificial Intelligence*, Netherlands: Kluwer Academic.
- SEARLE, John. (1996), *Akıllar, Beyinler ve Bilim* (Çev.) Kemal Bek, İstanbul: Say Yayınları.
- SEARLE, John. (2004), *Zihnin Yeniden Keşfi* (Çev.) Muhittin Macit, İstanbul: Litera Yayıncılık.
- SEARLE, John. (1990), "Is the Brain's Mind a Computer Program?" *Scientific American*, 262, no.1, 26-31.
- SEARLE, John. (1980), "Minds, Brains, and Programs", <http://cogprints.org/7150/1/10.1.1.83.5248.pdf>. ss. 1-19 (19 Aug 2001).
- SKIRBEKK, Gunnar and GILJE, Nils. (Eds.), (1971), *Antik Yunan'dan Modern Döneme Felsefe Tarihi* (Çev.) Emrah Akbaş ve Şule Mutlu, İstanbul: Üniversite Kitabevi Yayınları.
- STERRETT, Susan G. (2003), "Turing's Two Tests for Intelligence", (Ed.) James H. Moor, *The Turing Test: The Elusive Standard of Artificial Intelligence*, Netherlands: Kluwer Academic.
- SWAINE, Michael R. "Computers"; "The Microprocessors", *Encyclopaedia Britannica 2005 Ultimate Reference Suite CD*.
- TURING, A.M. (1950), "Computing Machinery and Intelligence", *Mind* 49, ss. 433-460.

