

A Comparative Study of the Capacities of Artificial Intelligence in Reconstructing Human Identity and Consciousness from the Perspective of Mulla Ṣadrā's Transcendent Theosophy and John Searle's Philosophy of Mind

Sayyedah Masoumeh Tabatabaei^{1✉}  and Arash Rajabi^{2✉} 

1. PhD Graduate in Islamic Theology, University of Qom; Lecturer at Amirkabir University, Email: Eshraghm@yahoo.com
2. PhD Graduate in Theoretical Foundations of Islam, University of Islamic Sciences; Lecturer at Amirkabir University, Email: Eshragh.sadra@gmail.com

Article Info

Article type:
Research Article

Article history:

Received:

11 April 2025

Received in revised form:

20 April 2025

Accepted:

3 September 2025

Available online:

10 March 2026

Keywords:

Artificial Intelligence,
Identity,
Transcendent Theosophy,
Ṣadr al-Muta'allihīn (Mulla Ṣadrā),
John Searle,
Philosophy of Mind,
Consciousness.

ABSTRACT

This article presents a philosophical analysis of the concept of identity in the context of artificial intelligence from the perspective of Transcendent Theosophy—particularly the views of Mulla Ṣadrā—and contemporary philosophy. With the rapid advancement of artificial intelligence as one of the most significant innovations in computer science, fundamental questions have arisen regarding the reconstruction of human identity and its comparison with natural intelligence. This study examines the philosophical capacities and limitations of artificial intelligence from the viewpoints of Ṣadr al-Muta'allihīn and John Searle and offers a comparative analysis of their perspectives on identity and consciousness in machines and intelligent systems. According to both thinkers, despite its remarkable progress, artificial intelligence is incapable of attaining human identity, as it lacks essential attributes such as consciousness, intentionality, and immateriality. John Searle maintains that artificial systems will never achieve genuine consciousness, since the comprehension of meaning and the experience of awareness are unique characteristics of the human brain and mind and cannot be attributed to machines. He has consistently reinforced this position through his famous “Chinese Room” thought experiment and his extensive critiques of prevailing theories of artificial intelligence.

Cite this article: Tabatabaei, S.M.; and Rajabi, A. (2025). A Comparative Study of the Capacities of Artificial Intelligence in Reconstructing Human Identity and Consciousness from the Perspective of Mulla Ṣadrā's Transcendent Theosophy and John Searle's Philosophy of Mind. *Digital Islamic Studies and Humanities*, 1 (2), 81-98.
<https://doi.org/10.22034/disah.2026.2057628.1047>



© The Author(s). **Publisher:** Research Center for Digital Islamic Studies and Humanities (RCDISAH).

DOI: <https://doi.org/10.22034/disah.2026.2057628.1047>

اقتضانات متن کاوی هوشمند در تحلیل نگره‌های ادبیات عرب

سیده معصومه طباطبایی^۱ و آرش رجبی^۲

۱. دانش‌آموخته دکتری کلام اسلامی دانشگاه قم، مدرس دانشگاه امیرکبیر، تهران، ایران، رایانامه: Eshraghm@yahoo.com
 ۲. دانش‌آموخته دکتری مبانی نظری اسلام دانشگاه معارف اسلامی، مدرس دانشگاه امیرکبیر، تهران، ایران، رایانامه: Eshragh.sadra@gmail.com

اطلاعات مقاله

چکیده

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت: ۱۴۰۴/۰۱/۲۲

تاریخ بازنگری: ۱۴۰۴/۰۴/۲۰

تاریخ پذیرش: ۱۴۰۴/۰۶/۰۳

تاریخ انتشار: ۱۴۰۴/۱۲/۲۰

کلیدواژه‌ها:

هوش مصنوعی،

هویت،

حکمت متعالیه،

صدرالمتألهین،

جان سرل،

فلسفه ذهن،

آگاهی.

این مقاله به تحلیل فلسفی مفهوم هویت در زمینه هوش مصنوعی از منظر حکمت متعالیه - به‌ویژه دیدگاه‌های ملاصدرا - و فلسفه معاصر می‌پردازد. با پیشرفت‌های سریع در حوزه هوش مصنوعی به‌عنوان یکی از نوآوری‌های برجسته علوم رایانه، پرسش‌هایی بنیادین درباره‌ی بازسازی هویت انسانی و مقایسه آن با هوش طبیعی مطرح شده است. در این پژوهش، ظرفیت‌ها و محدودیت‌های فلسفی هوش مصنوعی از منظر صدرالمتألهین و جان سرل بررسی شده و تحلیل تطبیقی دیدگاه‌های آن‌ها درباره‌ی هویت و آگاهی در ماشین‌ها و سامانه‌های هوشمند ارائه می‌گردد. بر اساس دیدگاه این دو اندیشمند، هوش مصنوعی با وجود پیشرفت‌های چشمگیر، قادر به دستیابی به هویت انسانی نیست؛ زیرا فاقد ویژگی‌های بنیادین همچون آگاهی، قصدیت و تجرد است. جان سرل بر این باور است که سامانه‌های مصنوعی هرگز به آگاهی واقعی دست نخواهند یافت؛ چراکه درک معنای واقعی و تجربه آگاهی، از ویژگی‌های منحصر به فرد مغز و ذهن انسان است و نمی‌توان آن را به ماشین‌ها نسبت داد. او این دیدگاه را با طرح آزمایش مشهور «اتاق چینی» و نقدهای متعدد بر نظریه‌های رایج درباره‌ی هوش مصنوعی، به‌طور مستمر تقویت کرده است.

استناد: طباطبایی، سیده معصومه؛ و رجبی، آرش (۱۴۰۴). اقتضانات متن کاوی هوشمند در تحلیل نگره‌های ادبیات عرب. علوم انسانی و اسلامی دیجیتال، ۱ (۲)، ۸۱-۹۸. <https://doi.org/10.22034/disah.2026.2057628.1047>



ناشر: پژوهشکده علوم اسلامی و انسانی دیجیتال (مرکز تحقیقات کامپیوتری علوم اسلامی نور). © نویسندگان.

مقدمه

هوش مصنوعی (Artificial Intelligence) یا (AI) به‌عنوان یکی از پیشرفته‌ترین دستاوردهای بشر و از نوآوری‌های برجسته علوم رایانه، از اواسط قرن بیستم با تلاش‌هایی نظیر طراحی ماشین تورینگ و پروژه‌های اولیه شبیه‌سازی رفتار انسان آغاز شد. در دهه‌های اخیر، پیشرفت‌های چشمگیر در حوزه یادگیری ماشین و شبکه‌های عصبی عمیق، هوش مصنوعی را به عرصه‌ای گسترده و تأثیرگذار تبدیل کرده است. این فناوری نه تنها در حل مسائل پیچیده علمی و صنعتی نقش آفرین بوده، بلکه در حوزه‌های فلسفی همچون هویت و آگاهی نیز پرسش‌ها و چالش‌های نوینی را پدید آورده است. از جمله مسائل مطرح در این زمینه، ظرفیت هوش مصنوعی در بازسازی هویت و آگاهی انسانی است. این پرسش که آیا هوش مصنوعی می‌تواند روزی با دستیابی به آگاهی، بر هوش طبیعی انسان غلبه کند، در سطح آن قرار گیرد یا حتی جایگزینی برای آن باشد، همچنان محل تأمل است. همچنین، این مسئله مطرح می‌شود که آیا می‌توان برای هوش مصنوعی، هویتی مستقل قائل شد یا خیر. این مقاله به تحلیل و مقایسه دیدگاه‌های حکمت متعالیه، به‌ویژه نظرات صدرالمآلهین شیرازی، و فیلسوف معاصر، جان سرل، درباره هویت و آگاهی در هوش مصنوعی می‌پردازد. هدف آن، بررسی تطبیقی ظرفیت‌های فلسفی هوش مصنوعی در بازسازی هویت و آگاهی انسانی از منظر این دو اندیشمند است.

تعریف هوش مصنوعی

هوش مصنوعی شاخه‌ای از علوم رایانه است که به طراحی و توسعه سیستم‌ها و الگوریتم‌هایی می‌پردازد که قادرند وظایفی را انجام دهند که معمولاً نیازمند هوش طبیعی انسانی اند. این وظایف شامل درک زبان طبیعی، شناسایی الگوها، یادگیری از تجربیات، تصمیم‌گیری و حل مسائل می‌شود. به‌عبارت دیگر، هوش مصنوعی به توانایی ماشین‌ها در انجام فعالیت‌هایی اشاره دارد که در گذشته مستلزم تفکر و استدلال انسانی بودند (Russell & Norvig, 2021)¹.

در حقیقت، هوش مصنوعی می‌تواند مکمل یا جایگزینی برای هوش انسانی در فعالیت‌هایی باشد که نیازمند تحلیل دقیق الگوها، پیش‌بینی و ارزیابی نتایج آینده هستند.

1. Artificial Intelligence (AI) is a branch of computer science that deals with the design and development of systems and algorithms that can perform tasks that typically require human intelligence. These tasks include natural language understanding, pattern recognition, learning from experiences, decision-making, and problem-solving. In other words, AI refers to the ability human thought and reasoning.

Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson Education.

الف. بررسی فلسفی هویت انسان در اندیشه صدرالمآلهین

۱. حرکت جوهری

ملاصدرا در نظام فلسفی خود، «حرکت جوهری» را به عنوان مبنایی برای تبیین پویایی و تحول هویت انسان و نهاد ناآرام جهان مطرح می‌کند. او معتقد است از آن‌جا که جهان طبیعت فعلیت محض نیست و نمایشگاه و آمیزه قوه‌ها و فعلیت‌ها، و حرکت، خروج شیء از حالت قوه و حدوث تدریجی فعلیت است، نهاد عالم طبیعت و جوهر وجودی انسان به‌طور مداوم در حال تغییر و تکامل است و این تحول بنیاد هویت انسان را تشکیل می‌دهد (صدرالمآلهین، ۱۳۶۸: ۱۰/۳). از این منظر، هوش مصنوعی، شیء طبیعی و بخشی از طبیعت جهان به شمار نمی‌آید و تابع و محصولی است از طبیعت وجود آدمی و بنابراین به شکل مستقل فاقد حرکت جوهری است و نمی‌تواند به‌عنوان موجودی دارای هویت انسانی تکامل پذیر تلقی شود. البته روشن است که با حرکت جوهری استکمالی انسان، دستاوردهای او از جمله هوش مصنوعی که تابع هوش طبیعی است رشد و تکامل خواهد داشت ولی از آن‌جا که فاقد ذات و نفسیت مستقل و طبیعت است مشمول حرکت جوهری نخواهد بود و حتی فرض خودتوسعه‌ی آن نیز مبتنی بر داده‌های پیشینی است که هوش طبیعی آدمی در آن تعبیه می‌کند نه رشد و تکامل نهاد و جوهر آن.

۲. اتحاد عقل و معقول

یکی از اصول کلیدی در حکمت متعالیه، اتحاد عقل و عاقل و معقول است. این اصل بیان می‌کند که انسان در فرآیند شناخت، با موضوع شناخت یکی می‌شود. نکته مهم این که آدمی واجد «خودآگاهی» است و به خود و معلومات بالذات خود به علم حضوری و مستقیم عالم است، و به بیان دیگر آگاهی با موضوع آگاهی و هویت فرد آگاه یکی است که در فلسفه صدرایی از آن به اتحاد عقل و عاقل و معقول و یا اتحاد علم و عالم و معلوم و یا ادراک و مدرک و مدرک یاد می‌شود (صدرالمآلهین شیرازی، ۱۳۶۸: ۳/۳۱۲). این ویژگی منحصر به انسان، تمایزی بنیادین میان او و هوش مصنوعی ایجاد می‌کند که صرفاً داده‌ها را پردازش می‌کند؛ چرا که هوش مصنوعی واجد خودآگاهی نیست و تنها بر پایه داده‌های هوش طبیعی عمل می‌کند و حتی نسبت به تحلیل و پردازش

داده‌ها آگاهی ندارد تا معتقد شویم هوش مصنوعی عالم به خویشتن و به معلومات و داده‌ها و تحلیل‌ها است و علم او با معلوم بالذات او و با خود او یعنی عالم اتحاد دارد در حالی که در خصوص انسان چنین چیزی با تبیین فلسفی حکمت متعالیه محقق است.

تحلیل فلسفی

ملاصدرا در فلسفه خود به‌ویژه در نظریات مربوط به «اصالت وجود» و «حرکت جوهری» و «اتحاد عقل و عاقل و معقول»، انسان را موجودی متعالی و متطور می‌داند که از طریق حرکت جوهری به کمال نهایی خود می‌رسد. از نظر او، هویت انسان از دیدگاه معرفت‌شناختی و هستی‌شناختی نه تنها به ابعاد مادی، بلکه به ابعاد معنوی نیز مرتبط است. و با به تعبیر دیگر انسان «جسمانیة الحدوث و روحانیة البقاء» است. (صدرالمتألهین، ۱۳۶۸: ۳۴۵/۸) در این چارچوب، انسان موجودی است که ریشه در خاک و عالم طبیعت دارد، ولی در آن متوقف نمی‌ماند و با حرکت اشتدادی کمالی خود با اتصال به حقیقت الهی به‌طور مداوم در مسیر کمال قرار دارد و در این مسیر توانایی‌ها و قابلیت او نیز قابلیت ارتقا دارند. از دیدگاه او انسان موجودی صرفاً مادی نیست و می‌تواند به مقام تجرد و مافوق تجرد راه یابد و از قابلیت‌های الوهی و ربانی برخوردار شود که در زبان قرآن کریم از آن به مقام خلافت الهی یاد شده است. و نیز انسان تنها حیوانی ناطق نیست بلکه در مقام فعلیت کمالاتش، نوع اشرف است که تمام معانی و کمالات در آن مندرج و جمع می‌شود. و با تحصیل علم و عمل می‌تواند به ملکوت اعلا و مقام لقاء الله بار یابد. صدرا معتقد است، ماده اولیه بدن انسان گل است و سپس در اثر تطور و تکامل به صور مختلفی در می‌آید تا سرانجام قابل فیض الهی شود و نور ربانی به بدن عنصری او بتابد و با تابش این نور ربانی خلقت دیگری رقم می‌خورد که در عموم آدمیان به صورت استعداد نهفته است و در برخی از آنان با تقویت بعد علمی و عملی با اختیار و اراده آگاهانه به فعلیت کمال نهایی واصل می‌شود. (صدرالمتألهین، ۱۳۸۵: ۱۲۹ - ۱۳۵) و مقصود از خلقت کائنات نیز همینان هستند که از آنان به انسان کامل و خلیفه الله نام برده می‌شود. و تنها انسان کامل است که شایسته نام انسان و مقام خلافت خداوندی است و مظهر اسم اعظم الهی و کمالات ربانی است. بنابر این از دیدگاه صدرا هویت اصلی انسان از خاک بر می‌آید و با علم و اختیار تا افلاک ملکوت اوج می‌گیرد و هویتی ربانی می‌شود که هیچ موجودی به مقام او نمی‌رسد.

بر پایه این نگرش، روشن است که هوش مصنوعی با فقدان خودآگاهی و اختیار و با توانایی‌های محدود و ساختارهای محاسباتی و داده‌ای خود هر چند می‌تواند در تسهیل برخی امور زندگی انسان ابزار مناسبی به شمار آید، ولی هویتی ثابت و تعریف شده و ماشینی دارد که توسط عامل بیرونی در او تعبیه می‌شود و بر پایه فلسفه صدرایی نمی‌تواند با علم و عمل، بدون مداخله و برنامه‌ریزی عامل انسانی، هویت خود را ارتقا دهد. ملاصدرا بر این باور است که انسان در مسیر تکامل روحانی، به معرفت حقیقی دست می‌یابد. در این نگرش، انسان از جنبه‌های معنوی خاصی برخوردار است که نمی‌توان آن‌ها را در قالب داده‌ها و الگوریتم‌های ریاضی محصور کرد. این ویژگی‌ها، هم‌چون اتصال به حقیقت و درک فطری، موجب تمایز انسان از ماشین‌ها می‌شود. ملاصدرا به انسان به‌عنوان موجودی که قادر به دریافت حقیقت از طریق شهود و عقل است اشاره می‌کند. از این رو، هوش مصنوعی که تنها توانایی پردازش داده‌ها را دارد، قادر به درک حقیقت در معنای عمیق آن نیست و نمی‌تواند به هویت انسانی پی ببرد.

ملاصدرا با دیدگاه‌های مادی‌گرایانه به‌طور کامل مخالف است. او به انسان‌ها به‌عنوان موجوداتی مرکب از بدن و نفس نگاه می‌کند، جایی که «نفس» یا روح انسانی به‌عنوان جوهری مستقل و غیرمادی عمل می‌کند. به این ترتیب، هوش مصنوعی که اساساً بر مبنای داده‌های مادی و الگوریتم‌ها عمل می‌کند، نمی‌تواند به آگاهی حقیقی دست یابد؛ زیرا این آگاهی نیاز به نفس و تجرد دارد.

ب. بررسی فلسفی هویت انسان در اندیشه جان سرل

جان سرل، یکی از برجسته‌ترین فیلسوفان معاصر، در فلسفه ذهن و زبان نظریاتی بنیادین مطرح کرده است. او با تکیه بر مفاهیمی چون قصدیت (Intentionality)، آگاهی (Consciousness)، شخصی‌سازی حالت‌های ذهنی (Subjectivity)، و علّیت زیستی (Biological Naturalism) به تحلیل هویت انسان می‌پردازد.

۱. قصدیت یا حیث التفاتی، اساس هویت ذهنی انسان

سرل مفهوم قصدیت و یا حیث التفاتی را به‌عنوان ویژگی اصلی ذهن معرفی می‌کند. قصدیت به توانایی ذهن در ارجاع به اشیاء، حالات، یا امور دلالت دارد. این مفهوم که در فلسفه پدیدارشناسی

نیز مورد بحث بوده، به نظر سرل، محور تمایز انسان از ماشین است. او بیان می‌کند: «قصیدیت مشخصه‌ای است که به ذهن انسان امکان می‌دهد به چیزی فراتر از خود ارجاع دهد، در حالی که ماشین‌ها چنین توانایی‌ای ندارند.» (Searle, 1983). در این چارچوب، هویت انسان از طریق توانایی او در درک، تولید معنا، و ارتباط با جهان خارجی شکل می‌گیرد. البته از دیدگاه او، حیث التفاتی فقط به قصدها مربوط نمی‌شود و شامل باورها، میل‌ها، امیدها، ترس‌ها، عشق، نفرت، حرص، انزجار، شرم، غرور، عصبانیت، شادی و همه حالت‌های ذهنی آگاهانه و ناآگاهانه که به جهان خارج از ذهن اشارتی دارند یا درباره آن هستند، نیز می‌شود (سرل، ۱۳۸۲: ۲۴).

۲. آگاهی: شرط لازم برای هویت انسانی

سرل بر اهمیت آگاهی به عنوان عنصری حیاتی در هویت انسان تأکید دارد. او آگاهی را حالتی کیفی و ذهنی می‌داند که به تجربه‌های درونی انسان وابسته است. سرل، آگاهی را واقعیت محوری و اساسی وجود انسانی می‌داند؛ چرا که بدون آن جنبه‌های انسانی وجود انسان، یعنی زبان، عشق، شوخی کردن همه ناممکن خواهند بود (سرل، ۱۳۸۲: ۲۴). برخلاف نظریات رفتارگرایانه یا فیزیکیالیستی که آگاهی را به فرآیندهای فیزیکی فرو می‌کاهند، سرل معتقد است که آگاهی ویژگی‌ای غیرقابل کاهش دارد: «آگاهی نمی‌تواند صرفاً به حالت‌های فیزیکی یا عصبی تقلیل یابد، بلکه باید به عنوان یک پدیده ذهنی مستقل تحلیل شود.» (Searle, 1992) از این دیدگاه، هویت انسانی نه تنها به فعالیت‌های مغزی، بلکه به تجربه‌های آگاهانه فرد وابسته است.

۳. شخصی‌سازی حالت‌های ذهنی

شخصی‌سازی حالات ذهنی یا همان سابعکتیویتی در حقیقت تمایز میان احساس و شناخت فرد از فردی دیگر است. به باور سرل، این شخصی بودن با واقعیاتی چون این واقعیت که من می‌توانم دردم را احساس کنم و شما نمی‌توانید، مشخص می‌شود. من جهان را از منظر خودم می‌بینم و شما از منظر خودتان. من از خودم و حالت‌های ذهنی درونی‌ام که کاملاً از خود مردم دیگر و حالات ذهنی آن‌ها جداست، آگاهم. (سرل، ۱۳۸۲: ۲۵).

۴. علّیت زیستی: بنیاد هویت انسانی

در نظریه «طبیعت‌گرایی زیستی» (Biological Naturalism)، سرل هویت انسان را با ریشه‌های زیستی آن مرتبط می‌داند. او بر این باور است که آگاهی و سایر ویژگی‌های ذهنی، محصول فعالیت‌های مغزی هستند، اما این فعالیت‌ها باید در زمینه‌ای زیستی درک شوند. به بیان دیگر، هویت انسانی نمی‌تواند جدا از بدن یا ساختارهای زیستی بررسی شود. سرل تأکید می‌کند که این دیدگاه میان کاهش‌گرایی افراطی و دوگانگی‌گرایی ذهن و بدن قرار دارد: «ویژگی‌های ذهنی واقعی هستند و بخشی از جهان طبیعی محسوب می‌شوند، اما باید آن‌ها را در چارچوب زیست‌شناسی درک کرد» (Searle, 2004).

سرل از نظریات دوگانه‌گرایانه دکارت انتقاد می‌کند و معتقد است که ذهن و بدن بخشی از یک سیستم واحد هستند. او همچنین دیدگاه‌های مادی‌گرایانه‌ای را که آگاهی را به حالات فیزیکی تقلیل می‌دهند، ناکافی می‌داند. از این دیدگاه، هویت انسان ترکیبی از ویژگی‌های ذهنی و زیستی است که به‌طور هم‌زمان عمل می‌کنند. هویت انسان در اندیشه جان سرل بر پایه مفاهیمی چون قصدیت، آگاهی، و زیست‌شناسی استوار است. او با رد دوگانگی دکارتی و تقلیل‌گرایی فیزیکی، نظریه‌ای جامع ارائه می‌دهد که هم به تجربه‌های ذهنی و هم به پایه‌های زیستی توجه دارد. سرل تأکید می‌کند که این ویژگی‌ها، هویت انسانی را از سایر موجودات و به‌ویژه ماشین‌ها متمایز می‌کند.

جان سرل با نگرش مادی‌انگارانه در باب هویت نفس بر این باور است که پدیده‌های ذهنی معلول فرایندهای نوروفیزیولوژیکی مغز و نیز ویژگی مغزند و نام آن را نیز طبیعت‌گرایی زیست‌شناختی می‌گذارد. وی معتقد است رویدادها و فرایندهای ذهنی به همان اندازه جزء تاریخ طبیعی زیست‌شناختی ما هستند که معده، تقسیم‌افزایشی سلول‌ها، تقسیم‌کاهشی سلول‌ها یا ترشح آنزیم جزء آن هستند. سرل برای تبیین دیدگاه خود از یک تمثیل بهره می‌برد، وی می‌گوید: آب را در نظر بگیرید. همه ما با رفتار آب در سطح کلان و معمولی اشیای روزمره آشنا هستیم. مقلای دانیم که مرطوب و بی‌بو است (اگر خالص باشد) می‌توان آن را نوشید شکل ظرف را به خود می‌گیرد... اما چرا آب این گونه رفتار می‌کند؟ پاسخ را باید در سطح خرد یافت، آن‌گاه که در می‌بایم آب از میلیون‌ها ملکول به تنهایی نامرئی تشکیل شده که هر ملکول مرکب از دو اتم هیدروژن و یک اتم اکسیژن (H₂O) است. همین مطلب در مورد مواد جامد هم صادق است. وی سپس به مثال‌هایی در این باب می‌پردازد مانند الماس و گرافیت که هر دو از اتم‌های کربن تشکیل شده‌اند ولی الماس سخت است و گرافیت نرم. در هر حال ویژگی‌های کلان ظاهری معلول رفتار عناصر خرد زیرین هستند.

بر پایه دیدگاه‌های جان سرل در (Minds, brains, and programs)، هوش مصنوعی می‌تواند جنبه‌های شناختی انسان را شبیه‌سازی کرده و به توانمندی‌های مشابه انسان دست یابد. در این رویکرد، ماشین‌ها قادرند بر اساس داده‌ها و الگوریتم‌ها، تصمیم‌گیری‌های پیچیده و رفتارهای مشابه انسان را انجام دهند، هرچند درک عمیق معنای این تصمیم‌ها و پیوند آن‌ها با هویت انسانی ممکن نیست. در مواجهه با مسئله هوش مصنوعی، سرل نشان می‌دهد که ماشین‌ها فاقد ویژگی‌های کلیدی برای تشکیل هویت انسانی هستند. او بیان می‌کند که هویت انسان بر توانایی تجربه قصدیت و آگاهی متکی است، در حالی که ماشین‌ها صرفاً داده‌ها را پردازش می‌کنند. سرل در این زمینه از استدلال مشهور خود، اتاق چینی، استفاده می‌کند تا نشان دهد ماشین‌ها نمی‌توانند به فهم واقعی دست یابند که به آن در این مقاله پرداخته خواهد شد.

جان سرل در تحلیل فلسفی خود از هوش مصنوعی، میان دو نوع هوش مصنوعی، یعنی هوش مصنوعی ضعیف (Weak AI) و هوش مصنوعی قوی (Strong AI)، تمایز قائل می‌شود. این تقسیم‌بندی از جمله مباحث محوری در فلسفه ذهن معاصر است و دیدگاه سرل نقش کلیدی در نقد قابلیت‌های هوش مصنوعی قوی ایفا می‌کند.

هوش مصنوعی ضعیف (Weak AI)

سرل معتقد است که هوش مصنوعی ضعیف تنها شبیه‌سازی عملکردهای ذهن انسان است. بر این اساس، سیستم‌های هوش مصنوعی ضعیف می‌توانند رفتارهایی مشابه با انسان از خود نشان دهند، اما این رفتارها دلالتی بر وجود فهم، آگاهی یا ذهن در آن‌ها ندارد. به عبارت دیگر، این نوع هوش مصنوعی ابزاری است برای شبیه‌سازی ذهن انسانی بدون اینکه خود ذهنی واقعی داشته باشد. سرل بیان می‌کند که هوش مصنوعی ضعیف می‌تواند برای آزمون نظریه‌های مربوط به شناخت انسان مفید باشد، اما هرگز به طور واقعی معانی یا حیث التفاتی یا قصدیت (Intentionality) را درک نمی‌کند (Searle, 1980, p. 417).

هوش مصنوعی قوی (Strong AI)

برخلاف هوش مصنوعی ضعیف، هوش مصنوعی قوی مدعی است که اگر یک ماشین بتواند عملکرد ذهن انسان را شبیه‌سازی کند، آنگاه می‌توان آن ماشین را دارای ذهن، آگاهی، و فهم واقعی دانست. سرل به شدت با این دیدگاه مخالف است و آن را بر اساس استدلال «اتاق چینی» رد می‌کند.

نقد هوش مصنوعی قوی

سرل سه نقد اصلی به هوش مصنوعی قوی وارد می‌کند:

۱. نبود قصدیت یا حیث التفاتی: ماشین‌ها فاقد توانایی ارجاع به اشیاء یا معانی بیرونی هستند. آن‌ها فقط داده‌ها را پردازش می‌کنند (Searle, 1980, p 418).
۲. نبود تجربه آگاهانه: آگاهی (Consciousness) یکی از ویژگی‌های بنیادین ذهن انسانی است که نمی‌تواند به‌سادگی در ماشین‌ها بازتولید شود.
۳. عدم توانایی در تولید معنا: ماشین‌ها تنها نمادها را دستکاری می‌کنند؛ در حالی که ذهن انسانی قادر به تولید معناست.

سرل تأکید می‌کند که درک این تمایز برای جلوگیری از سوء برداشت‌های علمی و فلسفی ضروری است. اگرچه هوش مصنوعی ضعیف می‌تواند در علوم شناختی و توسعه فناوری‌ها مفید باشد، ادعای هوش مصنوعی قوی دربارهٔ ایجاد ذهن و آگاهی در ماشین‌ها نه تنها اغراق‌آمیز، بلکه از نظر فلسفی نیز نادرست است. دیدگاه سرل دربارهٔ هوش مصنوعی، با تمایز میان هوش مصنوعی ضعیف و قوی، چارچوبی فلسفی برای درک محدودیت‌های این فناوری ارائه می‌دهد. او با استفاده از استدلال اتاق چینی، نشان می‌دهد که حتی پیشرفته‌ترین ماشین‌ها نیز نمی‌توانند به آگاهی و قصدیت واقعی دست یابند، و این مسئله تأکیدی بر تمایز ماهوی میان ذهن انسان و عملکرد ماشین‌هاست.

آزمایش اتاق چینی (Chinese Room Argument)

این استدلال مشهور یا آزمایش فکری توسط سرل در سال ۱۹۸۰ ارائه شد و هدف آن نقد هوش مصنوعی قوی بود. توضیح استدلال یا آزمایش به بیان سرل چنین است: تصور کنید که یک دسته از برنامه‌نویسان کامپیوتری، برنامه‌ای را نوشته‌اند که کامپیوتری را قادر می‌سازد تا فهم زبان چینی را شبیه‌سازی کند. بنابر این، اگر مثلاً به این کامپیوتر مسئله‌ای به زبان چینی داده شود، آن مسئله را با حافظه یا پایگاه اطلاعات خود تطابق داده پاسخ‌های مناسبی را به زبان چینی برای آن مسئله فراهم می‌آورد. برای کمک به استدلال، فرض کنید که پاسخ‌های کامپیوتر به اندازه پاسخ‌های کسی که چینی زبان مادری اوست خوب است. حال آیا این کامپیوتر بر این اساس زبان چینی را می‌فهمد؟

یعنی آیا به معنای واقعی کلمه به همان شکلی که چینی زبانان چینی را می‌فهمند زبان چینی را می‌فهمد؟ خوب، تصور کنید که در اتاقی حبس شده‌اید و در این اتاق چند سبد پر از نمادهای چینی وجود دارد. تصور کنید شما یک کلمه چینی نمی‌دانید اما به شما یک کتاب پر از قاعده برای کارکردن با این نمادهای چینی داده شده که به زبان انگلیسی است. این قواعد کارکردن با این نمادها را به شکل کاملاً صوری، یعنی برحسب دستور زبان آن‌ها مشخص می‌کند نه بر حسب معناشناسی شان. خوب شاید قاعده بگوید: فلان نماد را از سبد شماره یک بگیر و آن را کنار بهمان نماد از سبد شماره دو بگذار. حال فرض کنید که چند نماد دیگر چینی وارد این اتاق می‌شوند و فرض کنید به شما قواعد دیگری برای بازگرداندن نمادهای چینی به بیرون از اتاق داده شده است. فرض کنید بی آن که برای شما معلوم باشد نمادهای وارد شده به اتاق را انسان‌های خارج از اتاق «مسائل» و نمادهایی را که شما به بیرون از اتاق بر می‌گردانید «پاسخ‌های این مسائل» نامیده‌اند. باز فرض کنید برنامه نویسان در طراحی این برنامه‌ها بسیار عالی عمل کرده‌اند و شما در کارکردن با این نمادها بسیار خوب عمل کرده‌اید، به طوری که خیلی زود پاسخ‌هایتان از پاسخ‌های کسی که زبان مادری اش چینی است غیرقابل تشخیص می‌شوند. در اینجا شما در اتاق خود حبس شده‌اید در حالی که نمادهای چینی خود را جابه‌جا می‌کنید و نمادهای چینی را در پاسخ به نمادهای چینی وارد شده بیرون می‌دهید. بر اساس این وضعیت راهی وجود ندارد که بتوانید چینی را فقط از راه کار کردن با این نمادهای صوری یاد بگیرید (سرل، ۱۳۸۲: ۶۰-۶۲).

به‌طور خلاصه:

- فردی که زبان چینی نمی‌داند، در اتاقی قرار دارد. او یک کتابچه راهنما دارد که به او می‌گوید چگونه نمادهای چینی را در پاسخ به ورودی‌های خاص جفت کند.
- به نظر می‌رسد که فرد در اتاق دقیقاً می‌داند چگونه زبان چینی صحبت کند، زیرا پاسخ‌هایی معنادار تولید می‌کند.
- اما در واقع، او تنها یک دستورالعمل نمادین را دنبال می‌کند و هیچ فهمی از زبان چینی ندارد. سرل استدلال می‌کند که این شبیه عملکرد کامپیوترهاست. آن‌ها فقط دستورات برنامه‌ریزی شده را اجرا می‌کنند، بدون این که واقعاً بفهمند چه اتفاقی در حال رخ دادن است. ماشین‌ها ممکن است اطلاعات را پردازش کنند، اما هیچ آگاهی یا معنایی در این فرایند وجود ندارد. و در حقیقت هوش

مصنوعی قوی نمی‌تواند ذهن را شبیه‌سازی کند. سرل در این آزمایش فرض می‌کند که یک فرد غیر چینی در یک اتاق بسته نشسته است و با استفاده از یک دفترچه راهنما قادر است دستورات چینی را پردازش کرده و به سوالات چینی پاسخ دهد. فرد در این اتاق هیچ درکی از زبان چینی ندارد، اما همچنان می‌تواند به‌طور صحیح به سوالات چینی پاسخ دهد. سرل نتیجه می‌گیرد که چنین سیستمی (که در اینجا یک ماشین است) قادر به «درک» یا «آگاهی» از زبان نیست؛ بلکه تنها دستورات را طبق قواعدی از پیش تعیین‌شده دنبال می‌کند. به عبارت دیگر، صرفاً تطابق علامت‌ها و نشانه‌ها با هم نمی‌تواند به معنای آگاهی و فهم باشد. او می‌گوید که برای داشتن آگاهی، یک سیستم باید واقعاً معنا را درک کند و نه صرفاً بر اساس قواعد عمل کند (Searle, 1980, p. 417).

نقد اتاق چینی توسط دنیل دنت و پاسخ سرل

دنیل دنت به شدت از آزمایش اتاق چینی سرل انتقاد کرده است. سرل در این آزمایش فکری سعی کرد نشان دهد که سیستم‌های هوش مصنوعی نمی‌توانند آگاهی داشته باشند، زیرا آن‌ها تنها با پردازش نشانه‌ها و قواعد عمل می‌کنند، نه این‌که واقعاً درک کنند. دنت مخالف این ایده است و معتقد است که فهمیدن و آگاهی به معنای آگاهی از معانی نیست، بلکه صرفاً یک فرایند عملکردی است که بر اساس رفتارهای هوشمندانه یک سیستم می‌توان آن را نسبت داد. به عبارت دیگر، دنت باور دارد که اگر ماشینی بتواند به درستی همانند انسان‌ها رفتار کند، می‌توان آن را آگاه دانست، حتی اگر آن ماشین واقعاً مانند انسان‌ها معنا را درک نکند. (Denett, 1991, p. 195) دنت هم‌چنین در کتاب (Darwin's Dangerous Idea, 1995) به تبیین این موضوع پرداخته است که فرایندهای ذهنی و آگاهی می‌توانند از فرایندهای تکاملی ناشی شوند. وی معتقد است که فرایندهای شناختی، همان‌طور که در انسان‌ها و دیگر موجودات زنده مشاهده می‌شود، به‌طور کامل قابل مدل‌سازی و شبیه‌سازی در ماشین‌ها هستند. او بر این باور است که اگر یک سیستم مصنوعی بتواند رفتارهایی مانند انسان‌ها نشان دهد، می‌توان آن را آگاه دانست و نیازی به روح یا مفهوم درک معنای عمیق در آن وجود ندارد. در واقع، دنت معتقد است که هوش مصنوعی می‌تواند تمام ویژگی‌های ذهن انسانی را تقلید کند، بدون آن‌که نیاز به داشتن آگاهی واقعی باشد (Denett, 1995, p. 302).

دیدگاه‌های دنیل دنت در مورد هوش مصنوعی بر خلاف دیدگاه‌های جان سرل است که معتقد به عدم امکان آگاهی در ماشین‌ها است. دنت با تأکید بر نظریه رفتارگرایی جدید و موقعیت عمدی معتقد است که سیستم‌های هوش مصنوعی می‌توانند همانند انسان‌ها رفتار کنند و به همین دلیل می‌توانند آگاهی و ذهن را به شکلی مشابه تجربه کنند. بر اساس دیدگاه دنت، آگاهی و هوش به طور کامل با عملکردهای شناختی و رفتاری مرتبط هستند و نیازی به ویژگی‌های خاص انسان‌ها یا موجودات زنده ندارند؛ اما سرل معتقد است که آگاهی انسان‌ها به طور ذاتی با زیست‌شناسی مغز مرتبط است و نمی‌توان آن را به فرایندهای الگوریتمی که در ماشین‌ها اجرا می‌شود تقلیل داد. به عبارت دیگر، ماشین‌ها فاقد ساختارهای زیستی لازم برای ایجاد آگاهی هستند و نیز آگاهی فقط در سیستم‌های زنده‌ای که ساختارهای زیستی مناسب دارند امکان‌پذیر است. سرل معتقد است که هوش مصنوعی قوی نمی‌تواند واقعاً ذهن یا آگاهی ایجاد کند، زیرا ماشین‌ها فقط پردازشگر نمادها هستند و فاقد توانایی درک معنا هستند. او بر تمایز بنیادین بین زیست‌شناسی مغز انسان و فرایندهای الگوریتمی ماشین‌ها تأکید دارد.

سرل در نظریه چینی خود (Chinese Room) تأکید می‌کند که حتی اگر ماشین‌ها رفتارهای انسانی را شبیه‌سازی کنند، هیچ‌گاه نمی‌توانند معنای واقعی و فلسفی این رفتارها را درک کنند. به عبارتی، ماشین‌ها صرفاً از نظر ظاهری مشابه انسان‌ها عمل می‌کنند، اما درون‌مایه این رفتارها و معنا را درک نمی‌کنند.

سرل به شدت دیدگاه‌های دنت را نقد کرده و معتقد است که هرچند هوش مصنوعی می‌تواند شبیه‌سازی‌هایی از رفتارهای انسان‌ها ارائه دهد، اما این شبیه‌سازی به معنی درک و آگاهی نیست. سرل می‌گوید که آگاهی درک از تجربه‌های پدیدارشناختی است، نه صرفاً پردازش اطلاعات. سرل معتقد است که آگاهی یک ویژگی غیرقابل کاهش است که مختص موجودات زیستی است و به هیچ وجه نمی‌توان آن را به سیستم‌های ماشینی نسبت داد. از نظر او، ذهن انسانی دارای ویژگی‌های خاصی است که آن را از سیستم‌های مصنوعی متمایز می‌کند. او این نظریه را پدیدارشناسی آگاهی می‌نامد و بر این باور است که آگاهی یک حالت ذهنی است که با تجربیات داخلی، احساسات، و درک معنای جهان مرتبط است. در حالی که هوش مصنوعی می‌تواند شبیه به رفتارهای انسانی عمل کند، اما نمی‌تواند از آنچه که سرل درک معنای واقعی می‌نامد، برخوردار باشد. او تأکید می‌کند که آگاهی، چنان‌که در انسان‌ها وجود دارد، یک پدیده فیزیکی است که به کارکرد مغز مرتبط است و نمی‌توان آن را به طور کامل در ماشین‌ها شبیه‌سازی کرد (Searle, 1992, p. 174).

مقایسه تطبیقی: نظرات ملاصدرا و جان سرل

ملاصدرا هویت انسان را در فرایند تکامل از طریق حرکت جوهری می‌بیند، او معتقد است که انسان موجودی است که از خاک برخاسته و از طریق علم و اختیار به کمال می‌رسد. این کمال به واسطه اتصال به حقیقت الهی و تجلی نور ربانی در انسان به دست می‌آید. بنابراین، از دیدگاه ملاصدرا، هویت انسان نه تنها مادی بلکه معنوی است و می‌تواند به مقام تجرد و خلافت الهی برسد. از این رو، انسان دارای ویژگی‌هایی همچون آگاهی شهودی و تجرد است که هوش مصنوعی قادر به دستیابی به آن‌ها نیست. جان سرل نیز هویت انسان را از جنبه آگاهی و قصدیت می‌فهمد. به اعتقاد او، انسان از توانایی ارجاع به اشیاء و حالات فراتر از خود برخوردار است، که به آن قصدیت گفته می‌شود. سرل به تمایز انسان از ماشین‌ها بر پایه قصدیت تأکید دارد و معتقد است که ماشین‌ها قادر به درک معنای حقیقی یا آگاهی به خود نیستند. سرل مخالف نظریات مکانیکی و فیزیکی کامل در مورد ذهن است و معتقد است که ذهن انسان نمی‌تواند تنها به فرایندهای فیزیکی و ماشین‌وار کاهش یابد. در برابر نظریات مثل نظریه ذهن کامپیوتر (که ذهن انسان را به یک برنامه کامپیوتری شبیه می‌کند)، سرل اظهار می‌دارد که ذهن و آگاهی ویژگی‌هایی دارند که نمی‌توانند به‌طور کامل توسط سیستم‌های مصنوعی تولید شوند.

سرل یکی از مخالفان اصلی دیدگاه‌های مبتنی بر هوش مصنوعی است که ادعا می‌کنند می‌توان به سیستم‌های ماشین‌گونه یا برنامه‌های کامپیوتری آگاهی و ذهن را نسبت داد. او همچنین بر این باور است که فلسفه ذهن باید از دیدگاه‌های فیزیکی و مکانیکی فراتر برود و ویژگی‌های انسانی ذهن باید به‌طور خاص در نظر گرفته شوند.

سرل به آینده هوش مصنوعی به‌طور عمومی نگاه بدبینانه‌ای دارد. او معتقد است که حتی در آینده‌ای نزدیک نیز هوش مصنوعی نخواهد توانست به آگاهی دست یابد. با اینکه سرل به پیشرفت‌های علمی در زمینه هوش مصنوعی اعتراف می‌کند، اما از این می‌ترسد که تلاش‌ها برای شبیه‌سازی ذهن انسان به‌طور کامل شکست بخورد، زیرا سیستم‌های مصنوعی هیچ‌گاه نخواهند توانست به درک واقعی از معنا و آگاهی برسند. او تأکید می‌کند که اگرچه می‌توان ماشین‌هایی ساخت که شبیه به انسان‌ها رفتار کنند، اما هیچ‌گاه نخواهند توانست ذهنی مشابه انسان‌ها داشته باشند (Searle, 1992, p. 196).

نتیجه‌گیری

در مقایسه این دو دیدگاه، به نظر می‌رسد که هرچند هر دو فیلسوف به نوعی بر تمایز میان انسان و هوش مصنوعی تأکید دارند، اما تأکید ملاصدرا بر جنبه‌های معنوی و متعالی انسان، از جمله اتصال به حقیقت الهی و حرکت جوهری، هویت انسانی را به‌عنوان موجودی تکامل‌یافته و روحانی در نظر می‌گیرد. در حالی که جان سرل بیشتر بر جنبه‌های شناختی و آگاهی انسان تأکید دارد و به قصدیت به‌عنوان ویژگی تمایز انسان از ماشین اشاره می‌کند. به‌طور کلی، هر دو دیدگاه تأکید دارند که هوش مصنوعی، با تمام پیشرفت‌هایش، نمی‌تواند به هویت انسانی دست یابد زیرا فاقد ویژگی‌های حیاتی همچون آگاهی، قصدیت و مجرد است. دیدگاه‌های جان سرل در مورد هوش مصنوعی و آگاهی به‌ویژه در مخالفت با نظریه‌های دنیل دنت و دیگر طرفداران هوش مصنوعی قوی برجسته است. سرل معتقد است که هیچ‌گاه سیستم‌های مصنوعی نخواهند توانست به آگاهی واقعی برسند، چرا که درک معنای واقعی و آگاهی از ویژگی‌های مغز و ذهن انسان است که نمی‌توان آن را به ماشین‌ها نسبت داد. او این باور را از طریق آزمایش اتاق چینی و نقدهای مختلف به سیستم‌های مصنوعی به‌طور مستمر حمایت کرده است.

فهرست منابع

- شیرازی، صدرالمآلهین، (١٣٦٨ ش)، «الحکمة المتعالیة فی الأسفار الأربعة العقلیة»، بیروت: دار أحیاء التراث العربی.
- شیرازی، صدرالمآلهین، (١٣٨٥ ش)، «أسرار الآیات»، ترجمه: سیدمحمد موسوی، تهران: انتشارات حکمت.
- ار.سرل، جان، (١٣٨٢ ش)، «ذهن، مغز، علم»، امیر دیوانی، قم، بوستان کتاب.
- Searle, John R. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press, 1983.
- Searle, John R. *The Rediscovery of the Mind*. Cambridge: MIT Press, 1992.
- Searle, John R. *Mind: A Brief Introduction*. New York: Oxford University Press, 2004.
- Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences*, vol. 3, no. 3, 1980, pp. 417-424.
- Searle, John R. *The Rediscovery of the Mind*. Cambridge: MIT Press, 1992.
- Searle, J. (1980). *Minds, Brains, and Programs*. *Behavioral and Brain Sciences*, 3(3), 417-424.
- Searle, J. (1992). *The Rediscovery of the Mind*. MIT Press, p. 174.
- Searle, J. (1995). *The Problem of Consciousness*. *Journal of Consciousness Studies*, 2(4), 324-332.
- Dennett, D. (1991). *Consciousness Explained*. Little, Brown and Company, p. 76.
- Dennett, D. (1995). *Darwin's Dangerous Idea*. Simon & Schuster, p. 302.
- Searle, J. R. (1980). *Minds, Brains, and Programs*. *Behavioral and Brain Sciences*.
- Searle, J. R. (1992). *The Rediscovery of the Mind*. MIT Press.
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. Pearson.
- Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press.

References

- Dennett, D. C. (1991). *Consciousness explained*. Little, Brown and Company.
- Dennett, D. C. (1995). *Darwin's dangerous idea*. Simon & Schuster.
- Norvig, P., & Russell, S. (2021). *Artificial intelligence: A modern approach*. Pearson.
- Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424.
- Searle, J. (1992). *The rediscovery of the mind*. MIT Press.
- Searle, J. (1995). The problem of consciousness. *Journal of Consciousness Studies*, 2(4), 324–332.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.
- Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.
- Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.
- Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.
- Searle, J. R. (2004). *Mind: A brief introduction*. Oxford University Press.
- Searle, J. R. (2003). *Mind, brain, and science* (A. Divani, Trans.). Bustan-e Ketab. [in persian]

- Shirazi, Sadr al-Din. (1989). Al-hikmah al-muta‘aliyah fi al-asfar al-‘aqliyyah al-arba‘ah [The transcendent philosophy in the four intellectual journeys]. Dar Ihya al-Turath al-Arabi. [in persian]
- Shirazi, Sadr al-Din. (2006). Asrar al-ayat [The secrets of the verses] (S. M. Mousavi, Trans.). Hekmat Publications. [in persian].