

The Relationship between Various Approaches to the Mind with the Possibility of Strong Artificial Intelligence¹

Seyyed Mohammad Qazavy¹ 

Ahmad Vaezi² 

1. PhD student, Islamic Philosophy, Department of Philosophy and Theology, Faculty of Moral Philosophy, Baqir al-Olum University, Qom, Iran (corresponding author).

Email: qazavy.7626@chmail.ir

2. Professor, Department of Philosophy and Theology, Faculty of Moral Philosophy, Baqir al-Olum University, Qom, Iran.

Email: vaezi@bou.ac.ir



Abstract

Currently, the theory of strong artificial intelligence stands as a contentious and pivotal idea, having many proponents and opponents. Resolving this issue necessitates taking a stance on the mind-body or soul-body relationship. Diverse perspectives on the possibility or impossibility of strong artificial intelligence emerge from these differing approaches, as articulated by Western and Islamic philosophers of mind. This study employs an analytical-descriptive method to scrutinize the correlation between theories addressing the mind-body problem and the potential for strong artificial intelligence. The findings indicate that not only dualist perspectives but also some quarters of physicalist monism present challenges to the possibility of strong artificial intelligence.

Keywords

Mind, body, strong artificial intelligence, dualism, monism.

1. This article has been taken from the doctoral dissertation entitled "Critical examination of the anthropological foundations and tools of artificial intelligence from the perspective of transcendental wisdom" (Supervisor: Ahmad Vaezi), Department of Philosophy and Theology, Faculty of Moral Philosophy, Baqir al-Olum University, Qom, Iran.

* **Cite this article:** Qazavy, S. M. & Vaezi, A. (2023). The Relationship between Various Approaches to the Mind with the Possibility of Strong Artificial Intelligence. *Naqd va Nazar*, 28(111), pp. 159-187.

<https://doi.org/10.22081/JPT.2023.67881.2085>

* **Publisher:** Islamic Propagation Office of the Seminary of Qom (Islamic Sciences and Culture Academy, Qom Iran). * **Type of article:** Research Article

 **Received:** 21/11/2023 • **Revised:** 05/12/2023 • **Accepted:** 10/12/2023 • **Published Online:** 14/01/2024

© Author(s)



نسبت رویکردهای مختلف به ذهن با امکان هوش مصنوعی قوی^۱

id^۲ احمد واعظی id^۱ سیدمحمد قاضوی

۱. دانشجوی دکتری رشته فلسفه اسلامی، گروه فلسفه و کلام دانشکده فلسفه اخلاق
 دانشگاه باقرالعلوم (ع)، قم، ایران (نویسنده مسئول).

Email: qazavy.7626@chmail.ir

۲. استاد، گروه فلسفه و کلام، دانشکده فلسفه اخلاق، دانشگاه باقرالعلوم (ع)، قم، ایران.

Email: vaezi@bou.ac.ir



چکیده

نظریه هوش مصنوعی قوی در عصر حاضر، یکی از نظریات پرچالش و بااهمیت است که طرفداران و منکران بسیاری دارد. پاسخ به این مسئله، مستلزم داشتن رویکردی درباره رابطه ذهن/نفس و بدن است. تفاوت رویکردها به تفاوت پاسخ به امکان و عدم امکان هوش مصنوعی قوی منجر می‌گردد. فلاسفه ذهن غربی و اسلامی درباره رابطه ذهن/نفس و بدن تحلیل‌های متفاوتی ارائه داده‌اند. این تحقیق بر آن است که با روش تحلیلی - توصیفی نسبت میان دیدگاه‌های معروف درباره رابطه ذهن/نفس و بدن را با امکان هوش مصنوعی قوی بررسی و ارزیابی کند. این نتیجه به دست آمد که نه تنها دوگانه‌انگاران بلکه حتی برخی یگانه‌انگاران فیزیکال نیز نمی‌توانند به درستی از امکان هوش مصنوعی قوی دفاع کنند.

کلیدواژه‌ها

ذهن، بدن، هوش مصنوعی قوی، دوگانه‌انگاری، یگانه‌انگاری.

* این مقاله برگرفته از رساله دکتری با عنوان: «بررسی انتقادی مبانی و لوازم انسان‌شناختی هوش مصنوعی از منظر حکمت متعالیه» (استاد راهنما: احمد واعظی)، گروه فلسفه و کلام، دانشکده فلسفه اخلاق، دانشگاه باقرالعلوم (ع)، قم، ایران می‌باشد.

۱. **استاد به این مقاله:** قاضوی، سیدمحمد؛ واعظی، احمد. (۱۴۰۲). نسبت رویکردهای مختلف به ذهن با امکان هوش مصنوعی قوی. نقد و نظر، ۲۸(۱۱۱)، صص ۱۵۹-۱۸۷. <https://doi.org/10.22081/JPT.2023.67881.2085>

□ نوع مقاله: پژوهشی؛ ناشر: دفتر تبلیغات اسلامی حوزه علمیه قم (پژوهشگاه علوم و فرهنگ اسلامی، قم، ایران) © نویسندگان
 □ تاریخ دریافت: ۱۴۰۲/۰۸/۳۰ • تاریخ اصلاح: ۱۴۰۲/۰۹/۱۴ • تاریخ پذیرش: ۱۴۰۲/۰۹/۱۹ • تاریخ انتشار آنلاین: ۱۴۰۲/۱۰/۲۴



انسانی که در عصر تکنولوژی زندگی می‌کند در کنار پیامدهای و آثار خوبی که تکنولوژی برای وی به همراه دارد، چالش‌ها و اضطراب‌های بسیاری نیز اطرافش را فراگرفته است. این انسان برای دستیابی به زندگی بهتر و مطلوب خود، تلاش می‌کند تا بتواند از تکنولوژی به بهترین نوع استفاده کند و بهره‌بردار. یکی از انواع تکنولوژی در عصر حاضر که پیش‌روی انسان قرار گرفته و انسان در صدد است هرچه بیشتر از آن برای زندگی مطلوب یاری بطلبد، هوش مصنوعی با همه گستردگی آن است.

علم هوش مصنوعی قلمرویی بسیار وسیع دارد، چه در بُعد تجربی و چه انسان‌شناختی و چه الهیاتی. آنچه این تحقیق در پی آن است، بررسی نسبت میان رویکردهایی درباره رابطه ذهن و بدن با امکان هوش مصنوعی قوی است.

در میان فلاسفه ذهن غربی و نیز اسلامی دیدگاه‌های متفاوتی درباره رابطه میان ذهن/نفس و بدن وجود دارد که هر کدام برای نسبت میان ذهن و بدن تقریرهای متفاوتی ارائه داده‌اند. این تقریرها برخی مبتنی بر پذیرش بُعد مجرد و غیرفیزیکی در انسان است و برخی دیگر اساساً بُعد مجرد را نمی‌پذیرند و تنها به بُعد مادی باور دارند.

مسئله‌ای که در دوران معاصر، بسیار محل بحث و گفتگو قرار گرفته است، امکان و عدم امکان هوش مصنوعی قوی است. هر پاسخی به این مسئله، از سوی هر فیلسوفی مبتنی بر نظریه‌ای است که وی در ارتباط با رابطه ذهن/نفس و بدن گرفته است؛ از این رو برای بررسی امکان و عدم امکان هوش مصنوعی قوی لازم است ابتدا دیدگاه‌هایی درباره رابطه ذهن/نفس و بدن طرح شود و سپس آنها را نسبت به امکان و عدم امکان هوش مصنوعی قوی ارزیابی کرد.

۱. هوش مصنوعی

هنوز تعریف دقیقی برای هوش مصنوعی ارائه نشده است که مورد پذیرش همه دانشمندان صاحب‌نظر در این زمینه باشد و این خود به آن علت است که اساس این



موضوع، یعنی هوش مورد جنجال و اختلاف است و تعریف جامعی درباره آن وجود ندارد؛ هوش مصنوعی به شیوه‌های مختلفی تعریف شده است.

در کتاب‌های مرتبط با هوش مصنوعی معمولاً به چهار تعریف آن اشاره می‌کنند:

۱. سیستم‌هایی که منطقی فکر می‌کنند. ۲. سیستم‌هایی که مثل انسان‌ها فکر می‌کنند.
۳. سیستم‌هایی که منطقی عمل می‌کنند. ۴. سیستم‌هایی که مثل انسان‌ها رفتار می‌کنند. از میان چهار تعریف بیان‌شده آنچه نویسندگان کتاب هوش مصنوعی (راهبردی نوین) می‌پذیرند، تعریف چهارم است؛ یعنی منطقی و عاقلانه عمل کردن (راسل و نورویگ، ۱۳۹۱، ص ۱۸).

۱-۱. هوش مصنوعی قوی، تفکر انسان گونه

تعریفی که بیان شد مبتنی بر هوش مصنوعی ضعیف است؛ اما آنچه این تحقیق در پی آن است، هوش مصنوعی قوی است که تعریف «تفکر انسان گونه» آن را تأمین می‌کند. این تعریف چنین است: «تلاش جدید و هیجان‌انگیز برای ساخت ماشین‌هایی متفکر و با حس کامل»؛ به عبارت دیگر «خود کارسازی فعالیت‌های مرتبط با تفکر انسان، فعالیت‌هایی مثل تصمیم‌گیری، حل مسئله، یادگیری». هوش مصنوعی قوی «سیستمی است که همانند انسان فکر می‌کند، تصمیم می‌گیرد و دارای احساس و اراده است» (عباس‌زاده جهرمی، ۱۳۹۰، ص ۲۶ و لو، ۱۳۸۹، ص ۲۳۹).

روشن است آنچه به لحاظ فلسفی مورد مناقشه است، همان رویکرد قوی به مسئله هوش مصنوعی است و گرنه با نتایج به‌دست آمده، هیچ‌کس منکر این نیست که رایانه‌ها در برخی زمینه‌ها (مانند سرعت و دقت محاسبات جبری) حتی از انسان پیشی گرفته‌اند؛ اما درباره اینکه آیا این اعمال همراه با اوصاف ذهنی متعلق به انسان، مانند آگاهی، فکر و ... صورت می‌گیرد یا شاید صورت بگیرد، تردیدهای بسیاری وجود دارد (لو، ۱۳۸۹، ص ۲۳۹).

1. Artificial Intelligence Strong



۱-۲. رویکردهای بنیادین در هوش مصنوعی

در طی زمانی که برای دستیابی به هوش مصنوعی تلاش شده است، دو رویکرد عمده شکل گرفته است که آن را نشانه‌گرایی^۱ یعنی رایانه دستگامی برای داده‌پردازی نشانه‌های ذهنی است و پیوند‌گرایی^۲ یعنی رایانه وسیله‌ای برای مدل‌سازی کردن مغز است، نام نهاده‌اند. گفتنی است که هوش مصنوعی قوی هیچ کدام از این دو رویکرد بنیادین نیست؛ اما قائلان به امکان تحقق هوش مصنوعی قوی، از طریق رویکرد پیوند‌گرایی، ایده خود را پیش خواهند برد که در نهایت به مقصود خود برسند.

۱-۲-۱. نشانه‌گرایی

این رویکرد بر این فرض استوار است که بسیاری از جنبه‌های هوش را می‌توان با دستکاری نمادها به دست آورد؛ فرضیه‌ای که با عنوان «سیستم نماد فیزیکی» در دهه ۱۹۶۰ توسط آلن نیوول^۳ و هربرت سایمون^۴ تعریف شده است (Simon, 2006, pp.116-119). آنها این فرضیه را بنیان گذاردند که مغز انسان و رایانه دیجیتال در همان حال که ساختار و مکانیزمشان کاملاً با هم متفاوت است، در سطحی از انتزاع، کارکردشان را می‌توان مشترک وصف کرد. در این سطح، هم مغز انسانی و هم رایانه دیجیتالی که به نحو مناسبی برنامه‌ریزی شده باشند، هر دو را می‌توان به مثابه دو نمونه‌سازی مختلف از یک نوع وسیله خاص در نظر گرفت (Dreyfus, Hubert L & Dreyfus, Stuart E, 1992, p.1).

نشانه‌گرایی با ایجاد مشابهت بین ماهیت هوش انسانی و کامپیوتری که برنامه‌ریزی شده است، حقیقت هوش را نوعی برنامه‌ریزی دانسته‌اند و مغز انسان را در مقابل آن ابررایانه‌ای فرض کرده‌اند که این نوع هوش خاص در آن کار گذاشته شده است. اما

1. Symbolism
2. Connectionism
3. Alan Newell
4. Herbert Simon



قائلان به رویکرد پیوندگرایی، نگرش نشانه‌گرایی را نپذیرفتند. در ادامه به رویکرد پیوندگرایی پرداخته خواهد شد.

۳-۱. پیوندگرایی

پیوند‌گرایان معتقدند ساخت مغز انسان متفاوت از ساخت یک کامپیوتر الکترونیکی است. مغز انسان دارای میلیون‌ها سلول عصبی یا نورون^۱ است که هر یک از آنها همگی به هم مرتبط هستند. این ارتباط به نحوی است که مجموع نورون‌ها با هم یک کل درهم‌تنیده و یک شبکه‌ای واحد از نورون‌ها است. پیوند‌گرایان موفق به ساخت ماشین‌هایی شدند که توانایی دارند اموری را انجام دهند که کامپیوترها از انجام آن ناتوان‌اند. این ماشین‌ها دارای مزایایی بودند؛ مثلاً قادر بودند طرح‌های مختلف را از هم تفکیک کنند؛ همچنین قادر بودند در شرایطی که برخی واحدهای آنها ایراد پیدا کرده‌اند، همچنان کار کنند (Russell & Norvig, 2003, p. 21).

همان‌طور که سیستم عصبی انسان آموزش می‌بیند و یاد می‌گیرد، این ماشین‌ها نیز چنین خصوصیتی دارند. برای درک بهتر دو دیدگاه، لازم است مقایسه‌ای میان آن انجام شود.

رویکرد نشانه‌گر مدعی است ذهن انسان یک کامپیوتر و بلکه ابررایانه است؛ یعنی ذهن انسان بدون هیچ تفاوتی همان زبان نمادین و صوری در رایانه است. این بدین معناست که همان‌گونه که رایانه تنها یک برنامه از قبل طراحی شده است، پس ذهن انسان نیز یک برنامه از قبل طراحی شده خواهد بود؛ یعنی تمام حرکات، رفتارهای انسان پیش‌بینی‌پذیر و تغییرناپذیر است؛ همان‌گونه که رایانه نیز چنین است. در این نوع رویکرد، آزمون و خطا، یادگیری و افزایش داده‌ها از محیط اطراف امکان ندارد. ارتباط دوطرفه با محیط اطراف و تمامی اشیایی که در ارتباط با AI هستند، غیرممکن خواهد

1. Neuron





بود. این همان نقدی است که پیوند گراها به آنها گرفته‌اند.

اما در مقابل، پیوند گراها فراتر از نشانه گراها قدم برداشته‌اند و مدعی شده‌اند می‌توان AI را به گونه‌ای ساخت که بتواند همانند انسان با جهان اطراف خود تعامل کند، آزمون و خطا کند، مشاهده کند و یاد بگیرد. با تحقق چنین ویژگی‌هایی، برخی از آنان مدعی شده‌اند که AI قوی امکان دارد. روشن است که طبق نشانه‌گرایی AI قوی امکان نخواهد داشت؛ بلکه اساساً در رویکرد نمادین، چنین توانایی و ظرفیتی وجود ندارد (Graham, 1993, p. 85؛ لو، ۱۳۸۹؛ طهماسبی، ۱۳۸۵، ص ۴۲).

رویکرد نشانه‌گرا همان‌طور که بیان شد، مدعی است ذهن انسان یک رایانه و بلکه ابررایانه است. روشن است که طبق نشانه‌گرایی، هوش مصنوعی قوی امکان نخواهد داشت؛ بلکه اساساً در رویکرد نمادین، چنین توانایی و ظرفیتی وجود ندارد؛ بنابراین امکان هوش مصنوعی قوی بیشتر باید در حوزه پیوندگرایی بررسی گردد تا نشانه‌گرایی. با توجه به مدعای قائلان به هوش مصنوعی قوی، رویکرد پیوندگرایی ظرفیت و توانایی دارد. پس در بررسی رویکردهای مختلف به رابطه ذهن و بدن و امکان هوش مصنوعی قوی، همین که رویکردی، هوش مصنوعی قوی را ممکن ندانست، به طریق اولی رویکرد نشانه‌گرایی در هوش مصنوعی را نیز ممکن نخواهد دانست و اگر هوش مصنوعی قوی را ممکن دانست، باید این مسیر را از طریق پیوندگرایی پیش رود.

۲. رویکردهای فلاسفه ذهن به تعارض موجود میان بستار فیزیکی و علیت ذهنی

حل حقیقت ذهن و چگونگی ارتباط آن با بدن، مسئله‌ای اساسی است که قائلان به هوش مصنوعی قوی باید پاسخی مناسب برای آن داشته باشند. نکته حائز اهمیت این است تحقیق جایگاه ذهن و حالات ذهنی در میان اندیشه‌های فلاسفه ذهن، بسیار زیادی به این بستگی دارد که مسئله بستار فیزیکی یا خودبستگی ماده نیز تبیین و تحلیل گردد؛ زیرا مبنای اصلی اندیشه بسیاری از فلاسفه ذهن، فیزیکال است.

۱-۲. بستار فیزیکی^۱

طبیعت گرایان هستی‌شناختی با تکیه بر اصل مهم «تمامیت یا بستار فیزیکی جهان» بر این باور هستند که امور فیزیکی، تنها علت فیزیکی دارند و دیگر نیازی به پذیرش علل دیگری در عرض این علل فیزیکی نیست؛ از این رو اینان بستار فیزیکی را به معنای خودبستگی و کفایت علت فیزیکی و عدم پذیرش علل غیر فیزیکی معنا کرده‌اند. این ادعای آنان به معنای این است بر فرض که امور فراطبیعی وجود داشته باشند؛ اما با امور طبیعی ارتباطی علی ندارند (نریمانی، ۱۴۰۰، ص ۷۷). از آنجا که امور فراطبیعی بر اساس طبیعت‌گرایی هستی‌شناختی بی‌اثر هستند، پس با تکیه بر «تیغ اکام» مبنی بر رد معقولیت امری که در تبیین جهان بی‌اثر است و نیز با تکیه بر «اصل الیتیک» مبنی بر اینکه نشانه وجود هر شیئی اثرگذاری علی آن است، این نتیجه به دست می‌آید که افزون بر اینکه امور طبیعی، تبیینی طبیعی دارند، اساساً باور به امور غیر طبیعی، معقول نیست (نریمانی، ۱۴۰۰، ص ۷۸).

۲-۲. ارتباط طبیعت‌گرایی هستی‌شناختی با فلسفه ذهن

بدیهی است که بر پایه بستار فیزیکی، وجود نفس / ذهن غیرمادی نیز انکار می‌شود. در فلسفه ذهن به جای اصطلاح «طبیعت‌گرایی» از اصطلاح «فیزیکالیسم» استفاده شده است. فلاسفه ذهن برای پیشبرد مبنای نظریات خود (مادی‌گرایی) تلاش کرده‌اند به نفع فیزیکالیسم، استدلالی اقامه کنند. بر اساس اصل طبیعت‌گرایانه و فیزیکالیستی، توصیف طبیعت‌باوران علمی از جهان طبیعی، همچون ساختاری مکانی - زمانی و به لحاظ علی بسته است و فقط از طریق علی پیش می‌رود. مراد از بسته‌بودن علی این است که امری غیرمادی نمی‌تواند در آن تأثیری علی گذارد (نریمانی، ۱۴۰۰، ص ۸۱).

1. Physical Closure





در نظریهٔ بستار فیزیکی، تمامی معلول‌های فیزیکی قطعاً با علت کافی^۱ و فیزیکی همراه هستند. از سوی دیگر، بنا بر اصلی به نام «عدم تعین مضاعف»^۲ امکان ندارد که رویدادها بیش از یک علت کافی داشته باشند. پس نتیجه گرفته می‌شود که رویدادهای فیزیکی فقط علت فیزیکی دارند و علت غیرفیزیکی هیچ نقش علی ندارد (نریمانی، ۱۴۰۰، ص ۸۱). لازم است توجه شود که می‌توان بستار فیزیکی را همان یگانه‌انگاری مادی تلقی کرد؛ زیرا نظریه بیان‌شده نفی بعد غیرمادی می‌کند.

امری که به‌ندرت در فلسفهٔ ذهن انکار شده است، اثرگذاری امور ذهنی بر رویدادهای فیزیکی است؛ اما میان تعهد به اثرگذاری امور ذهنی و بستار فیزیکی تعارض وجود دارد. مراد از علیت ذهنی، تأثیر علی حالات و رویدادهای ذهنی بر حالات و رویدادهای مغزی است (نریمانی، ۱۴۰۰، ص ۸۳). برای روشن شدن مسئله به این مثال توجه شود: هنگام بالا رفتن از کوه، ناگهان سنگی به پایم برخورد می‌کند، درد شدیدی احساس می‌کنم (علیت فیزیکی به ذهنی) درد - احساس درد - چنان شدید است که فکر می‌کنم پایم شکسته است (علیت ذهنی به ذهنی) این باور و میل به درمان سبب می‌شود که همراهانم را صدا کنم تا به کمک آنها بازگردم (علیت ذهنی به فیزیکی). مسئلهٔ علیت ذهنی برای دکارت نیز مطرح بود که یک جوهر ذهنی مجرد چگونه می‌تواند با جوهری مادی - بدن - تأثیر و تأثر داشته باشد؟ این دو جوهر به‌قدری با هم تفاوت دارند که پذیرفتن رابطهٔ علی میان آنها بسیار مشکل است. در دوران معاصر بیشتر فلاسفهٔ ذهن، ذهن را مادی دانسته‌اند؛ ولی همچنان علیت ذهنی را نیز نمی‌توانند به‌راحتی انکار کنند؛ زیرا بالاخره این پرسش پابرجاست که امور ذهنی - جوهر باشند یا نباشند - چگونه می‌توانند با بدن رابطهٔ علی داشته باشند؟ پس در هر حال، حل چنین مسئله‌ای - نسبت میان علیت ذهنی و علیت فیزیکی (بستار فیزیکی) - بسیار اهمیت دارد

۱. مراد از «علت کافی» با توجه به اصطلاح فلسفهٔ اسلامی، همان «علت تامه» است.

۲. مراد از «عدم تعین مضاعف» (Principle of non-overdetermination) در اصطلاح فلسفهٔ اسلامی، «عدم توارد عاملین علی معلول واحد» و در اصطلاح کلامی، «عدم اجتماع دو قادر بر مقدر واحد» است.

و به نوعی واکاوی مسئله ذهن و بدن به پاسخ درست به پرسش بیان شده ارتباط دارد. هر یک از دیدگاه‌های موجود تلاش کرده‌اند که پاسخی به این تعارض دهند. هم‌علیت امور ذهنی را بپذیرند و هم از قلمرو بستر فیزیکی خارج نشوند.

۳. امکان هوش مصنوعی قوی در هرکدام از راه‌حل‌های فلاسفه ذهن

۳-۱. دوگانه‌انگاری جوهری دکارتی

بستر فیزیکی نه تنها دوگانه‌انگاری دکارتی را نفی می‌کند؛ بلکه اصل امکان علیت ذهنی را نیز به نقد می‌کشاند؛ زیرا اگر قرار باشد همه امور فیزیکی تنها بر اساس قوانین و رویدادهای فیزیکی تبیین شوند، پس رویدادهای ذهنی نقشی علی و مستقل در تحقق رخدادهای فیزیکی نخواهند داشت. پس دوگانه‌انگاری دکارتی مسیری غیر از بستر فیزیکی و اغلب دیدگاه فلسفه ذهن را طی کرده است و اساساً به مبنای طبیعت‌گرایی هستی‌شناختی باور ندارد. در ادامه پاسخی که دوگانه‌انگاری جوهری به حل تعارض گذشته می‌دهد، بیان می‌شود. بنا بر دیدگاه دوگانه‌انگاری در جوهر، ذهن امری متمایز و غیر از بدن است. آنان معتقدند که در جهان دو جوهر متفاوت و متمایز از هم وجود دارد؛ جوهری ذهنی و جوهری فیزیکی.

بنا بر این دیدگاه، گرچه ذهن و بدن از هم متمایز هستند؛ ولی با هم تعامل دارند و در یکدیگر تأثیر و تأثر می‌گذارند. این تعامل می‌تواند به این گونه باشد که اطلاعات حسی از جهان خارج از مغز به ذهن می‌رسد و سپس تصمیم‌ها درباره نحوه واکنش، از ذهن به مغز منتقل می‌شود. در این فرایند، بدن و مغز به خودی خود هیچ تصمیمی نمی‌گیرد (دکارت، ۱۳۷۱، ص ۱۲۱ و دکارت، ۱۳۸۱، ص ۱۲۸). دکارت با نظریه خودش مبنی بر وجود جوهری غیرمادی و مجرد و نفی بستر فیزیکی و یگانه‌انگاری، به تعارض میان علیت ذهنی و بستر فیزیکی پاسخ می‌دهد؛ به عبارت دیگر دکارت با ارائه نظریه دوگانه‌انگاری و پذیرش بُعد مجرد توانسته است به نوعی به تعارض موجود پاسخ دهد. در مقابل، هوش مصنوعی قوی - چه نشانه‌گرایی و چه پیوندگرایی - اساساً به بعد





غیرمادی باور ندارد. رویکرد وی در انسان‌شناسی، فیزیکی و مادی‌انگاره است. همه ویژگی‌های ذهنی را اعم از استدلال و استنباط به امر مادی تحویل می‌برد. اساساً انسان را تک بعد و مادی می‌داند.

حال با چنین وضعیتی که میان دوگانه‌انگاری و مسئله هوش مصنوعی قوی وجود دارد، بدیهی است که بر اساس نظریه دوگانه‌انگاری، هوش مصنوعی قوی امکان نخواهد داشت. البته نه بدین دلیل که وجه عدم امکان در خود مسئله هوش مصنوعی قوی وجود داشته باشد؛ بلکه وجه عدم امکان به مبانی رویکرد دوگانه‌انگاری جوهری باز می‌گردد.

۲-۳. دوگانه‌انگاری جوهری صدرایی

صدرالمتألهین قائل به وجود دو بُعد مجرد و مادی در انسان است که البته بنا بر حرکت جوهری اشتدادی، می‌توان ایشان را نوعی یگانه‌انگار نیز به شمار آورد. این نظریه نیز اساساً مبتنی بر بستر فیزیکی نیست. وی معتقد است برای ارتباط میان بخش مجرد و مادی، احتیاج به هیچ واسطه‌ای جز مراتب خود نفس نیست. بر اساس حرکت جوهری اشتدادی، هويت و وجود هر چیزی قابل اشتداد است و مراتب مختلف آن شیء در اشتداد اتصالی، دارای وحدت اتصالی خواهد بود. بر اساس این وحدت اتصالی، شخصیت آن شیء و متحرک در طول حرکت و اشتداد در وجود، حفظ خواهد شد. نفس نیز در این حرکت جوهری اشتدادی خود، وحدتش را حفظ خواهد کرد و از میان مراتب گوناگونش، ماهیات متفاوت و مختلفی انتزاع می‌شود. نفس بنا بر ترکیب اتحادی ماده و صورت، با مرتبه بدن خود اتحاد وجودی دارد، ولو اینکه به لحاظ ماهوی، میان آنها تفاوتی باشد؛ زیرا ملاک وحدت و کثرت «وجود» است.

بر اساس دیدگاه ملاصدرا، نفسی که وجودی متحرک دارد و از مرتبه جسمانیت به مرتبه تجرد می‌رسد، بدین معناست که حد مرتبه جسمانیت را رها کرده و حدود مرتبه

بالا تر را به خود گرفته است. این نوع رها کردن و رفتن به مرتبه بالاتر برای نفس، این نکته را یادآور می‌شود که دیگر نفس محدود و منحصر در مرتبه گذشته نیست و آماده حرکت به سوی مرتبه جدید و بالاتر است. البته کمالات مرتبه گذشته را از دست نمی‌دهد؛ بلکه صرفاً محدودیت‌ها را رها می‌کند، و اگر کمالی در مرتبه قبلی باشد، آن را همچنان همراه خود خواهد داشت. پس این سخن بدین معناست که نفس در عین اینکه مرتبه تجرد یافته است؛ اما مرتبه جسمانی را نیز به همراه خود دارد (صدرالمتألهین، ۱۹۸۱م، ج ۳، صص ۴۷۵ - ۴۷۸؛ ج ۸، صص ۶، ۱۳۶ - ۱۴۷ و ۲۶۰ و ج ۹، ص ۹۹).

تیین بیشتر رابطه نفس و بدن بر اساس جایگاه مزاج در حدوث نفس است. مزاج مبتنی و متوقف بر سطح پایه خودش یعنی عناصر و کیفیات چهارگانه است و نفس جسمانی نیز متوقف و مبتنی بر سطح پایه خودش یعنی مزاج است. تا زمانی که نفس پدید نیامده است، فقط سطح پایه پایینی بر بالایی تأثیر گذار است؛ اما از زمانی که نفس ایجاد شد، تأثیر گذاری متقابل خواهد شد؛ یعنی اینکه هم سطح پایه پایینی یعنی کیفیات چهارگانه بر مزاج و بعد بر نفس تأثیر می‌گذارد و هم نفس جسمانی که صورت مدبر است بر سطح پایه اثر گذار است. نفس جسمانی گرچه جوهری متمایز از سطح پایه یعنی بدن محسوس است؛ اما مستقل نیست و حتی نفس مجرد نیز مستقل نخواهد بود. در اینجا تفاوتی وجود دارد به جای اینکه مزاج که سطح پایه است، حافظ و نگه‌دارنده باشد، نفس نگه‌دارنده و حافظ و هدایت‌گر مزاج است (صدرالمتألهین، ۱۹۸۱م، ج ۳، ص ۲۰؛ ج ۸، ص ۲۸ و ج ۹، ص ۱۰). صدرالمتألهین نیز بر اساس مراتبی دانستن نفس انسانی، هم علیت ذهنی را می‌پذیرد و هم بستر فیزیکی؛ به عبارت دیگر ملاصدرا با تقریر خاصی که ارائه داد، هم به بُعد مجرد قائل است و هم به بعد مادی و این به معنای پذیرش علیت ذهنی همراه با پذیرش بستر فیزیکی (بُعد مادی) است.

با توجه به نقش مزاج به دست می‌آید که در مسئله مدل‌سازی هوش انسانی و ایجاد هوش مصنوعی قوی بنا بر دیدگاه صدرا، مزاج یا بدن و جسم طبیعی موضوعیت ویژه‌ای دارد. اساساً طبق دیدگاه مزبور، بدن و جسم مصنوعی قابلیت تحقق و ایجاد هویتی





نوخاسته را ندارد؛ زیرا اگر نفس انسانی دقیقاً همان هویت نوخاسته دانسته شود، سطح پایه آن لزوماً بدن و جسم طبیعی یا همان مزاج است. صدرا معتقد است نفس جسمانی که مرتبه اول از مراتب نفس انسانی است لزوماً برخاسته از مزاجی خاص است و تبیین ایشان نشان می‌دهد که وی راهی جز مزاج برای ایجاد مرتبه اول نفس انسانی قائل نیست. طبیعی است که دیگر مراتب نفس انسانی اعم از مرتبه مجرد خیالی و مجرد عقلی متوقف بر این است که قبل از آنها، مرتبه جسمانی تحقق یافته باشد. پس از آنجا که صدرالمآلهین نفس انسانی را متوقف بر مرتبه‌ای جسمانی و طبیعی می‌داند، دیگر جسمی مصنوعی نمی‌تواند بنا بر حرکت جوهری اشتدادی منجر به نفسی انسانی گردد (برای مطالعه بیشتر: قاضی، ۱۴۰۲، فصل سوم و چهارم).

البته می‌توان محل نزاع را به گونه‌ای دیگر نیز بیان کرد: بر فرض بپذیریم که بر اساس حکمت صدرایی، مزاج و جسم و فرایندی طبیعی موضوعیتی در تحقق نفس انسانی نداشته باشد؛ اما با این حال نزاع همچنان پابرجاست؛ زیرا محققان هوش مصنوعی مدعی مادی‌انگاری حقیقت انسان هستند و به همین دلیل می‌گویند: می‌توان شرایط تحقق یک ماده طبیعی را به صورت مصنوعی نیز فراهم کرد؛ اما حکمت صدرایی به مجرد حقیقت انسان باور دارد. پس حتی در فرض پذیرش عدم موضوعیت ماده و فرایند طبیعی در مرتبه‌ای از مراتب انسانی، باز هم بر اساس حکمت صدرایی، هوش مصنوعی قوی امکان نخواهد داشت.

۳-۳. دوگانه انگاری ویژگی‌ها

مغز انسان جوهری فیزیکی است که انواع مختلفی از ویژگی‌ها را دارد؛ مثلاً میلیاردها عصب دارد، خون و ... دارد. این ویژگی‌ها هم فیزیکی است؛ اما افزون بر آنها ویژگی‌های دیگری نیز دارد که غیرفیزیکی هستند، مانند: آگاهی داشتن، باورداشتن، دردداشتن، محبت داشتن، نفرت داشتن. نام این نوع دوگانه‌انگاری ویژگی‌ها را شبه پدیدارگرایی نامیده‌اند. توضیح اینکه ویژگی‌های فیزیکی مغز علت و ویژگی‌های

غیرفیزیکی آن هستند؛ اما عکس این مطلب صحیح نیست؛ بدین معنا که تغییر و تأثیرپذیری ویژگی‌های فیزیکی مغز که از خارج است، سبب ایجاد ویژگی‌های غیرفیزیکی در آن می‌شود.

نکته این است ویژگی‌های غیرفیزیکی هیچ تأثیر و سببیتی در ویژگی‌های فیزیکی مغز ندارند. پس ویژگی‌های غیرفیزیکی از لحاظ علی خنثی و بی‌تأثیرند (چرچلند، ۱۳۸۶، ص ۲). با توجه به آنچه گذشت، مسئله‌ای مهم باقی می‌ماند که چگونه رویدادهای فیزیکی مغز پدیده‌ای مثل آگاهی را به وجود می‌آورند؟ نحو وابستگی پدیده ذهنی به فیزیکی چگونه و چیست؟ همچنین بر اساس تعارضی که میان علیت ذهنی و بستر فیزیکی مطرح شد، آیا ویژگی‌های ذهنی می‌توانند بر جهان فیزیکی مؤثر باشند؟ اگر همچنان مبنای بستر فیزیکی و خودبسندگی ماده را پذیرفته باشیم، ممکن نیست که ویژگی‌های ذهنی علت رویداد فیزیکی باشند؛ از این رو این امر منجر به بی‌اثری علی ویژگی‌های ذهنی خواهد شد. اگر هم ویژگی ذهنی تأثیر بگذارد به دلیل وجود ویژگی فیزیکی همراه آن بوده است. پس همچنان علیت ذهنی انکارشده و تعارض حل نشده است. مسئله این است که بتوان میان علیت ذهنی و بستر فیزیکی جمع کرد؛ انکار یکی از آنها نمی‌تواند راه‌حلی برای تعارض موجود باشد.

درباره امکان و عدم امکان هوش مصنوعی قوی می‌توان به دو بیان اشاره کرد:

الف) عدم امکان هوش مصنوعی قوی: AI قوی به بعدی غیرمادی ولو جوهری نباشد باور ندارد و ذهن و حالات ذهنی را مادی می‌داند و ساخت AI قوی را با رویکرد مادی‌انگارانه به ذهن و حالات ذهنی پیش می‌برد؛ بنابراین نمی‌توان بر اساس رویکرد دوگانه‌انگاری ویژگی‌ها، به امکان AI قوی معتقد شد و اساساً در چنین رویکردی، توانایی و ظرفیت امکان AI قوی وجود ندارد.

ب) امکان هوش مصنوعی قوی: از آنجا که ویژگی‌های غیرفیزیکی برآمده از امر فیزیکی هستند، همین نسبت را می‌توان در مغزی مصنوعی نیز پیاده و شبیه‌سازی کرد؛ از این رو بنا بر دوگانه‌انگاری ویژگی‌ها، هوش مصنوعی قوی امکان خواهد داشت. در



این فرض می‌توان ویژگی‌های غیرفیزیکی را با نمادها و رمزهایی تعریف کرد که بر مغز مصنوعی تأثیرگذار و از آن تأثیرپذیر هستند.

۳-۴. رفتارگرایی

نظریات دوگانه‌انگاری بیان‌شده و این‌همانی (در ادامه بیان خواهد شد)، ذهن را یک شیء تلقی کرده‌اند. رویکرد دوگانه‌انگاری نگاهی متافیزیکی و غیرتحویلی به ذهن دارد و حالات ذهنی را به رشته‌های عصبی تحویل نمی‌برد؛ اما رویکرد این‌همانی کیفیات و حالات ذهنی را به رشته‌های عصبی تحویل می‌برد و نگرشی مادی‌انگارانه ارائه می‌دهد؛ اما رویکردهای کارکردگرایی (در ادامه بیان خواهد شد) و رفتارگرایی نگاهی متفاوت به ذهن و حالات ذهنی دارند و یکی ذهن و حالات آن را به مثابه کارکرد مغز و دیگری به منزله رفتار عمومی فرد قلمداد می‌کنند (هارت و دیگران، ۱۳۸۱، ص ۳۱).

راه‌حلی که رفتارگراها برای تعارض میان علیت ذهنی و بستار فیزیکی در نظر گرفته‌اند این است که مراد از ذهن و حالات ذهنی، همان رفتار است و میان ذهن و حالات آن با رفتار، نوعی این‌همانی وجود دارد. این دو رویکرد از آنجا که تلاش می‌کنند مفاهیم ذهنی را به سطح فعالیت‌های فیزیکی تقلیل دهند، نظریه ماده‌گرا هستند (تامپسون، ۱۳۹۵، ص ۵۸). رفتارگرایان نیز در قلمرو بستار فیزیکی عمل می‌کنند و مسئله ذهن و حالات آن را تحلیل می‌نمایند. بهترین راه‌حل این مشکل آن است که چنین قلمداد کنیم که واژگان مربوط به حالات ذهنی دیگران، در واقع معادل واژگان مربوط به رفتار بالفعل و بالقوه آنها باشد. در این صورت است که این واژگان معنا دار خواهند شد؛ زیرا رفتار افراد دیگر به وضوح قابل ارزیابی تجربی است (مسلین، ۱۳۹۵، ص ۵۹).

با توجه به آنچه بیان شد، برای توضیح بیشتر این دیدگاه و بررسی نسبتش با آگاهی می‌توان به چند مورد از ویژگی‌های آن اشاره کرد:

الف) مطرح‌نشدن مسئله ذهن و بدن: طبیعی است که وقتی ذهن و حالات ذهنی به



رفتار و عینیت آنها با رفتار تحویل برده می‌شود، دیگر مسئله‌ای به نام رابطه ذهن و بدن مطرح نخواهد بود. این گونه نیست که ذهن، رفتار را به وجود آورد؛ بلکه ذهن همان رفتار بدن است. پس با تحلیل عینیت میان ذهن و رفتار، مسئله دشوار رابطه ذهن و بدن برای همیشه حل خواهد شد (البته به شرطی که دیدگاه رفتارگرایی صحیح باشد).

ب) ذهن و حالات ذهنی غیرمجرد هستند: وقتی ذهن مندی همان مجموعه رفتارهای فرد باشد؛ از این رو پدیده‌ای مشهود و قابل رؤیت خواهد بود، نه نامرئی. طبق مادی‌بودن ذهن، بررسی آن به مبانی تحقیق‌پذیری تجربی سپرده می‌شود.

بر اساس رویکرد رفتارگرایی، رویکرد نشانه‌گرایی در هوش مصنوعی نیز ممکن نخواهد بود؛ زیرا نشانه‌گرایی تنها یک برنامه‌رایانه‌ای است که هیچ‌گونه رفتاری از خود نشان نمی‌دهد و حتی بالقوه نیز رفتاری را در خود ندارد. این همان اشکالی است که پیوندگراها به نمادگراها وارد کردند.

اما رویکرد پیوندگرایی بر اساس رفتارگرایی امکان دارد؛ زیرا پیوندگرا افزون بر اینکه در صدد مدل‌سازی فکر و هوش انسانی به صورت مصنوعی هست، مدعی تحقق رفتارهای هوشمندانه است. این ادعا همان ادعای رویکرد رفتارگرایی است؛ زیرا آنان بر این باورند که ذهن و حالات ذهنی با رفتارهای بالفعل و بالقوه، همسانی و این‌همانی دارند.

وقتی که رفتارگرایی معتقد است که ذهن و حالات ذهنی در انسان همان رفتارهای اوست، پس مدل‌سازی هوش و آگاهی انسانی (AI قوی) با مدل‌سازی رفتارهای هوشمندانه محقق خواهد شد؛ به عبارت دیگر درست است که محل نزاع در AI قوی مدل‌سازی رفتار هوشمندانه نیست و رفتارگرایی به رفتار هوشمندانه پرداخته است؛ اما نباید این نکته مهم را نادیده گرفت که رفتارگرایی، ذهن و حالات ذهنی را دقیقاً همان رفتار می‌داند، پس عملاً تعریف AI قوی بر اساس رفتارگرایی، مدل‌سازی رفتار خواهد شد.

یادآوری این نکته خالی از فایده نیست که رویکرد رفتارگرایی در AI قوی در عمل



و واقع با آنچه اکنون در ذیل AI ضعیف تحقق یافته است، تفاوتی نخواهد داشت؛ اما با توجه به مبانی متفاوت خواهند شد. مسئله اصلی در این نظریه، «رفتار» است؛ همان طور که در این همانی ذهن و مغز «مغز طبیعی» و در کارکردگرایی ماشینی «کارکرد» است. توجه به این نکته در AI تأثیر بسزایی دارد؛ یعنی در رفتارگرایی باید رفتار مدل‌سازی شود و در این همانی باید مغز طبیعی و در کارکردگرایی باید کارکرد مدل‌سازی گردد. حال مدل‌سازی رفتار و کارکرد مشکلی ایجاد نمی‌کند؛ اما مدل‌سازی مغز طبیعی غیر از هوش مصنوعی قوی است.

۳-۵. این همانی (ماتریالیسم تحویل‌گرا)

همان طور که از توضیحات درباره نظریه بستار فیزیکی روشن می‌شود، مناسب‌ترین رویکرد در فلسفه ذهن با نظریه بیان‌شده، این همانی ذهن و مغز است؛ زیرا خودبستگی ماده بر آن است که اگرچه ذهن بر جهان فیزیکی تأثیرگذار است؛ اما جهان فیزیکی علتی و رای خودش را نمی‌پذیرد، پس باید ذهن را نیز جزئی از جهان فیزیکی در نظر گرفت؛ بنابراین حالات ذهنی با حالات فیزیکی این همان خواهند شد. اساساً رخدادهای فراطبیعی و غیرفیزیکی خارج از محدوده علم هستند و تنها رخدادهای فیزیکی می‌توانند تأثیر علی بگذارند؛ به عبارت دیگر رخدادهای فیزیکی خودبسته و بسته هستند. پس اصل تمایز میان ذهن و مغز را نمی‌پذیرند (نریمانی، ۱۴۰۰، صص ۸۰-۹۰). مطالبی که در توضیح بستار فیزیکی داده شد، همگی در مرحله اول بیانگر رویکرد این همانی ذهن و مغز هستند و دیگر رویکردها پس از ارائه اشکال علیت ذهنی مطرح می‌شوند.

این نظریه همه امور ذهنی را به جوهرهای مادی و حالات آن تحویل می‌برد. این تعریف چیزی جز رویکرد فیزیکالیستی نیست. پس بر اساس این دیدگاه، عوامل فوق طبیعی و ماورایی نسبت به امور ذهنی هیچ نقشی نخواهند داشت و هر آنچه ذهن و آگاهی و حالات ذهنی نامیده می‌شود، همان رابطه‌های علی فیزیکی در میان رشته‌های عصبی است. این آگاهی از حس‌های ورودی تا خروجی‌های رفتاری به صورت



آگاهانه و غیر آگاهانه، همه به روابط عصبی که ذهن وجود دارد، تحویل می‌رود (چرچلند، ۱۳۸۶، ص ۵۱؛ مسلین، ۱۳۹۱، ص ۱۱۹). بنا بر این دیدگاه، انسان بخشی از عالم ماده و طبیعی است و به همین دلیل برای تبیین آگاهی و ذهن در او باید حتماً به تبیینی مناسب و صحیح از عالم ماده و ساختار طبیعی آن رسید. در این تبیین، روابط ماورایی و غیر طبیعی و غیر فیزیکی نباید نقش و دخالتی داشته باشند.

۳-۵-۱. اشکال بر این همانی

این راهبرد نیز در حل تعارض میان علیت ذهنی و بستار فیزیکی موفق نیست. اینکه چرا این رویکرد نمی‌تواند به تعارض بیان‌شده پاسخی درخور دهد، به مسئله دشوار آگاهی بازمی‌گردد که اساساً هویتی غیر فیزیکی دارد با ویژگی‌هایی، همچون هستی‌شناسی اول‌شخص، کیفی بودن، ساختار پدیداری داشتن. درباره این مسئله در فصل چهارم، گفتار اول بحث خواهد شد.

بر اساس رویکرد این همانی ذهن و مغز، هم نشانه‌گرایی و هم پیوند‌گرایی امکان خواهد داشت؛ ولی AI قوی ممکن نخواهد بود.

امکان نشانه‌گرایی: گرچه نظریه این همانی مدعی انحصار ذهن در مغز طبیعی است؛ اما نشانه‌گرایی هم تفسیری برای همین مغز طبیعی است. پس این دو رویکرد با هم سازگار خواهند بود.

امکان پیوند‌گرایی: اگرچه پیوند‌گرایی با نشانه‌گرایی تفاوت دارند، هر دو در اینکه تفسیری برای مغز طبیعی هستند، و ذهن و حالات ذهنی را به همان مغز تحویل می‌برد و این همان می‌دانند، اشتراک دارند؛ از این رو پیوند‌گرایی هم با این همانی ذهن و مغز سازگار است.

عدم امکان AI قوی: از آنجا که AI قوی مدعی مدل‌سازی ذهن و آگاهی در غیر از مغز طبیعی است و نظریه این همانی مدعی انحصار ذهن و آگاهی در مغز طبیعی است، این دو نظریه با یکدیگر سازگار نیستند و از این همانی نمی‌توان به AI قوی رسید. البته





می‌توان بیانی دیگر از این همانی ارائه داد مبنی بر اینکه مراد از این نظریه، انحصار ذهن در مغز طبیعی نیست، بلکه انحصار و این همانی ذهن با مغزی فیزیکی است، خواه طبیعی باشد و خواه مصنوعی. طبق این بیان بنا بر این همانی، می‌توان از امکان هوش مصنوعی قوی سخن گفت و به‌نوعی روشن‌ترین فرض امکان هوش مصنوعی قوی بر اساس این همانی ذهن و مغز رخ خواهد داد.

۳-۶. کارکردگرایی ماشینی

بر اساس بستر فیزیکی، نظریه این همانی ذهن و مغز تأیید می‌شود؛ اما پس از ایراد اشکالات پرشمار، از جمله مسئله علیت ذهنی، یکی از رویکردهایی که در فضای فلسفه ذهن غلبه می‌کند، رویکرد کارکردگرایی ماشینی است. این رویکرد بر آن است که با توجه به مسئله تحقق چندگانه، با حفظ و پذیرش بستر فیزیکی، واقعیت علیت ذهنی و نیز استقلال و خودمختاری‌اش را نیز حفظ کند و پاسخی به تعارض میان پذیرش علیت ذهنی و بستر فیزیکی داده باشد. بنا بر کارکردگرایی ماشینی هیچ رویداد ذهنی، رویداد مکانیکی نیست؛ اما می‌تواند از طریق یک یا چند رویکرد مکانیکی محقق شود؛ به عبارت دیگر مصادیق ویژگی‌های ذهنی این همان با مصادیق ویژگی‌های فیزیکی هستند (نریمانی، ۱۴۰۰، صص ۸۰-۹۰).

کارکردگرایان همانند رفتارگرایان تحلیلی ذهن را مفهومی به مثابه یک شیء یعنی جوهری منطقی نمی‌دانند؛ بلکه آن را کارکرد می‌دانند. مطابق دیدگاه کارکردگرایان، آگاهی که یک حالت ذهنی است، امری انتزاعی قلمداد می‌شود که البته محصول کنش‌های مغز است. کارکرد با توجه به این توضیح غیرمادی است؛ زیرا حیثیتی مفهومی دارد و این مفهوم هم از نوع کنش‌های مغز به دست می‌آید. کارکردگرایان معتقدند ذهن همان کارکردهایی است که از خود، به صورت‌های گوناگون نشان می‌دهد. این فعالیت‌هایی که ذهن انجام می‌دهد همچون رایانه‌ای است که از طرفی ورودی‌هایی حسی دارد و مدارکات را دریافت می‌کند و در ادامه خروجی‌هایی بر

اساس همان ورودی‌ها از خود بروز می‌دهد (مسلین، ۱۳۹۵، ص ۷۹).

قائلان به این دیدگاه، کارکردگرایی را مجموعه ساختار فیزیکی پرشماری می‌دانند که حالات ذهنی یکسانی را حاصل می‌کند و از این رو دیگر برای آنها تفاوتی ندارد آنچه این کارکردها را ایجاد می‌کند، دستگاه عصبی انسان است و یا ماشینی مکانیکی (هارت و دیگران، ۱۳۸۱، ص ۷۷). این خود شاهدهی بر آن است که بر طبق این دیدگاه می‌توان به AI قائل شد؛ بنابراین کارکردگراها طبیعتاً تحقق چندگانه حالات ذهنی را می‌پذیرند و نباید ذهن مندی فقط به انسان گره بخورد. پس کارکردگرایی این امتیاز را نسبت به این‌همانی خواهد داشت که ذهن مندی را با دستگاه‌های فیزیکی کاملاً متفاوت با انسان ممکن می‌داند؛ اما این‌همانی ذهن را منحصر در مغز می‌داند (مسلین، ۱۳۹۵، ص ۸۱).

مطابق دیدگاه کارکردگرایان، آگاهی کارکردی با مفهوم انتزاعی و مابعدالطبیعی است. ذهن مندی حاصل کارهایی است که در واحدهای کوچک مغز انجام می‌گیرد. همه هدف رایانه‌ها محاسبه کارکردها است؛ یعنی اینکه ورودی خاصی را می‌گیرد و آن را از طریق انجام مجموعه‌ای از مراحل جداگانه‌ای که نیازمند هوش نیستند، به خروجی تبدیل می‌کند. کارکردگرایی حالات ذهنی را همان نرم‌افزار رایانه می‌داند. مراد از حالت نرم‌افزاری مرتب کردن یک رشته اطلاعات توسط رایانه است. این حالت در تقابل با حالات سخت‌افزاری رایانه، همانند حالات الکترومغناطیسی مدارهای معینی است که می‌توان آن را با حالات عصبی مغز قیاس کرد (لو، ۱۳۸۹، ص ۵۷).

۳- ۶- ۱. حل مشکل رابطه ذهن و بدن

از آنجاکه یکی از مسائل بسیار مهم در فلسفه ذهن، رابطه ذهن و بدن است، کارکردگرایی با تبیینی که از دیدگاه خود ارائه می‌دهد (ادراکات حسی و چگونگی واکنش‌های مغز روی آنها و مکانیسم رفتار خروجی در فرایندی به هم پیوسته)؛ یعنی یگانه‌انگاری ذهن و بدن، البته به این گونه که ذهن همان کارکرد مغز است، دیگر جایی برای دوگانه‌انگاری و ایجاد مسئله ذهن و بدن باقی نمی‌گذارد.





۳- ۶- ۲. تلاش برای شبیه‌سازی ذهن

این دیدگاه به بازسازی ذهن توجه بسیار زیادی دارد. فرایندهای ذهنی را به زبان داده‌ها و کدهای قابل بازخوانی ترجمه می‌کند تا سرانجام بتواند چنین کارکردی را در هر ماشین پیچیده‌ای تحقق بخشد. این دیدگاه درصدد آن است که ثابت کند انسان ماشینی است با کارکرد ذهن‌مندی و آگاهی و لذا می‌توان ماشینی ساخت که کارکرد ذهن انسان را داشته باشد؛ بنابراین طبق این رویکرد، تحقق AI دور از دسترس نیست بلکه طراحی و ساخت مغزی مصنوعی که البته کارکردی آگاهانه دارد، یکی از اهداف اصلی کارکردگرایان است (مسلین، ۱۳۹۱، ص ۲۱۷؛ لو، ۱۳۸۹، ص ۲۵۶).

از آنجا که نشانه‌گرایی، ذهن را یک ابررایانه می‌داند یک زبانی نمادی و صوری دارد و در پی تعامل ذهن با جهان خارج نیست و نظریه‌اش چنین ظرفیتی ندارد؛ همان‌طور که پیوندگراها گفته‌اند؛ از این رو به نظر می‌رسد در برخی کارکردها با نظریه کارکردگرایی ماشینی سازگاری نداشته باشد.

اما پیوندگرایی با توجه به تعریفی که برای آن بیان شد کاملاً منطبق با کارکردگرایی ماشینی است. بر همین اساس است که هوش مصنوعی قوی نیز ممکن خواهد بود؛ زیرا کارکردگرایی مدعی تحقق چندگانه حالات ذهنی است و هوش مصنوعی قوی نیز همین مدعا را دنبال می‌کند.

۳-۷. نخواستگرای^۱

نخواستگرای به نوعی همان دوگانه‌انگاری ویژگی‌هاست که به صورت مجزا به مثابه دیدگاهی مطرح شده است. این دیدگاه نیز درصدد است که به تعارض میان پذیرش علیت ذهنی و بستار فیزیکی پاسخ دهد. همان‌طور که در ادامه روشن خواهد شد، این نظریه نیز بستار فیزیکی را می‌پذیرد و با این پیش‌فرض در پی حل تعارض است.

1. Emrgentism

نظریهٔ نوحاسته‌گرایی با توجه به دیدگاه‌های رقیب بهتر شناخته می‌شود. این نظریه رویکردی میان نظریات تحویلی اصالت فیزیک و غیرفیزیکی قرار دارد. در یک سوی این نظریه، نظریهٔ مکانیستی و سیستمی اصالت فیزیکی قرار دارد و در سوی دیگر، نظریهٔ دوگانه‌انگاری دکارتی. نظریهٔ نوحاسته‌گرایی از سویی، نگاه اصالت فیزیکی را می‌پذیرد و از سوی دیگر، نگاه تحویلی آن را مردود می‌داند؛ یعنی نظریهٔ مورد بحث، با پذیرش اصول فیزیکالیستی، قصد ندارد هویت‌های غیرفیزیکی را نادیده بگیرد. بر اساس دیدگاه نوحاسته‌گرایی، امور نوحاسته همان امور غیرفیزیکی است (موسوی، ۱۳۹۶، ص ۳۴۶).

۳-۷-۱. نوحاسته‌گرایی، راهی میانه

تبیین فیزیکالی که در علوم فیزیکی رشد پیدا کرد، برای تحلیل رفتار کل سیستم، به تبیینی مکانیستی دربارهٔ ویژگی‌های بخش‌های سیستم متوسل می‌شود. در این نوع نگاه، گاهی نمی‌توان بر اساس سیستم مکانیستی، برخی پدیده‌ها را تبیین کرد. همین مسئله سبب شد تا مخالفان این نوع رویکرد اشکالاتی را مطرح کنند و حتی مستشکلان بیان داشتند که پدیده‌ها محصول سیستم‌های غیرفیزیکی‌اند. این گونه بود که دوگانه‌انگاران دکارتی، تبیین توانایی‌های علی در قالب فرایندهای مکانیسمی در مغز را رد کردند و به جای آن پدیده‌های ذهنی را محصول جواهر ذهنی یا مجرد دانستند. در این حین، تضاد شدیدی میان این دو رویکرد رخ داد. اشکال اصلی این بود که رویکرد فیزیکالیستی برای تبیین پدیده‌های ذهنی کفایت نمی‌کرد. البته گفتنی است که مستشکلان مدعی نبودند که یک سیستم فیزیولوژیک از عناصر مادی تشکیل نشده است؛ بلکه آنها بر این باور بودند که ویژگی‌های ذهنی را نمی‌توان به سیستم فیزیکالی تحویل برد.

پس از این تقابل بود که رویکرد نوحاسته‌گرایی سر برآورد که نه تحویل‌گرا بود و نه به دوگانه‌انگاری تمایل داشت. اینان از سویی نمی‌خواستند که امور نوحاسته را به هویت‌های مجرد نسبت دهند و از سوی دیگر، در تقابل با نگاه تحویلی، هویت‌های تازه‌ای که متفاوت با هویت‌های فیزیکی بود را رد نمی‌کردند. آنها بر این باورند که





نحوه‌ای از ارتباط میان اجزای سیستم سبب ایجاد ویژگی نخواستہ شده است. سیستم‌های فیزیکی از امر مجرد غیر فیزیکی تشکیل نشده است؛ ولی نگاه تحویلی نیز صحیح نیست (عموسلطانی‌فروشانی، ۱۴۰۰، فصل دوم). یکی از ویژگی‌های مهم نخواستہ گرای، وابستگی به سطح پایه است. مراد این است که اگرچه ویژگی‌های نخواستہ به ویژگی‌ای سطح پایه قابل تحویل نیست، به آن وابسته است؛ زیرا آنان به اصالت فیزیک معتقد هستند. اینکه هویت نخواستہ از چیزی هر چند مادی پدید می‌آید، حدود مرزی با دو گانه‌انگاری جوهری خواهد داشت؛ زیرا ویژگی‌ها در دو گانه‌انگاری، برخاسته از جوهر غیر فیزیکی است (خوشنویس، ۱۳۸۹، فصل اول).

مطابق دیدگاه نخواستہ‌گرایان، ذهن ویژگی و توانایی ارگانسیم بشری است که در طول روند تکاملی، نخواستہ شده است. بر اساس دیدگاه نخواستہ‌گرای، مغز و ذهن و نیز حالات مغزی با حالات ذهنی یکسان نیستند؛ البته بین مغز و ذهن تفاوت چندانی نیز وجود ندارد؛ زیرا حالات ذهنی وابسته به حالات مغزی‌اند (عموسلطانی‌فروشانی، ۱۴۰۰، ص ۱۰۱). از آنجا که نخواستہ‌گرای نیز همان دو گانه‌انگاری ویژگی‌هاست، امکان هوش مصنوعی قوی در آن نیز همانند دو گانه‌انگاری ویژگی‌ها تبیین خواهد شد.

۳-۸. حذف‌گرایی

نظریه ماتریالیسم حذف‌گرا که حذف‌گرایی^۱ نیز نامیده می‌شود، در حدود نیم قرن پیش، توسط پاول فیرابند^۲ و ریچارد رورتی^۳ ارائه شد و در زمان معاصر، فلاسفه‌ای چون پاول چرچلند^۴ و پارتیشیا چرچلند^۵ از آن جانب‌داری می‌کنند (Atmstrong, 1999, p. 91).

1. Eliminativism
2. Paul Feyerabend
3. Richard Rorty
4. Paul Churchland
5. Patricia Churchland

پاسخی که حذف گراها به تعارض میان علیت ذهنی و بستار فیزیکی می‌دهند، پاک کردن صورت مسئله و اشکال است؛ یعنی اینان بر این باور هستند که اساساً ذهن و حالات ذهنی وجود ندارند و هرچه هست امر فیزیکی است. ماتریالیست‌های تحویل‌گرا معتقدند که تناظری یک‌به‌یک میان حالات ذهنی و مغزی وجود دارد؛ اما ماتریالیست‌های حذف‌گرا مدعی هستند که چیزی با عنوان باور، میل، قصد و دیگر حالات ذهنی اصلاً وجود ندارد و این واژگان را باید از هر گفتمانی حذف کرد (دیویس، ۱۳۹۰، ص ۵۱). با توجه به این سخن آنها، دیگر لازم نیست از رابطه ذهن و بدن حرفی زد؛ زیرا اصولاً ذهنی نداریم تا بخواهد با بدن مرتبط باشد؛ بلکه هرچه هست همان اندام مادی و مغز است که طبق کنش‌های عصبی مغز، کار خودش را انجام می‌دهد. دیگر نباید از تحویل حالات ذهنی به مغزی و مادی سخن گفت و فقط آنچه وجود دارد، فعالیت‌های عصبی مغز است.

از آنجا که قائلان به AI قوی مدعی الگوسازی هوش انسانی هستند و ماتریالیسم حذف‌گرا از اساس منکر وجود ذهن و حالات ذهنی است، پس از نگاه اینان AI قوی امکان ندارد.

اما می‌توان به شیوه‌ای دیگر نیز مسئله را بررسی کرد که حذف گراها ذهن و حالات ذهنی را نفی می‌کنند و فقط مغز و فرایندهای عصبی را می‌پذیرند، طبق این نظر چه اشکالی دارد که بتوانند مدعی ساخت انسانی مصنوعی شوند؟! بدین صورت که همان مغز و فرایندهای عصبی را به شیوه مصنوعی شبیه‌سازی کنند.

رویکرد نشانه‌گرایی در AI که زبان ذهن را همانند زبان رایانه، زبانی صوری و نمادین می‌داند طبیعتاً طبق حذف‌گرایی ممکن خواهد بود؛ مگر اینکه آنچه حذف‌گرا مدعی حذف و نفی آن است، همین زبان صوری و نمادین باشد. رویکرد پیوند‌گرا نیز به همین شیوه تحلیل می‌گردد.

نتیجه‌گیری

از آنچه گذشت به دست می‌آید که درباره نسبت رویکردهای مختلف به ذهن با امکان





هوش مصنوعی قوی، دیدگاه‌های یکسانی وجود ندارد.

دو گانه‌انگاران جوهری دکارتی به دلیل پذیرش بعد غیرمادی نمی‌توانند مدعی امکان هوش مصنوعی قوی باشند؛ زیرا اساس هوش مصنوعی قوی بر پذیرش بستر فیزیکی است.

بنا بر دیدگاه صدرالمتألهین، مزاج یا بدن و جسم طبیعی موضوعیت ویژه‌ای دارد. تبیین ایشان نشان می‌دهد که وی راهی جز مزاج برای ایجاد مرتبه اول نفس انسانی قائل نیست. پس از آنجا که صدرالمتألهین، نفس انسانی را متوقف بر مرتبه‌ای جسمانی و طبیعی می‌داند؛ از این رو دیگر جسمی مصنوعی نمی‌تواند بنا بر حرکت جوهری اشتدادی به نفسی انسانی منجر گردد.

دو گانه‌انگار ویژگی‌ها با دو بیان متفاوت می‌تواند، هم مدعی امکان هوش مصنوعی قوی باشد و هم مدعی عدم امکان آن؛ زیرا هرچند ویژگی‌ها را غیرفیزیکی می‌داند؛ اما مبتنی بر ماده و مغز تلقی می‌کند و از سوی دیگر، می‌توان همین ویژگی‌های غیرفیزیکی را به نماد و رمز تعریف کرد که در هوش مصنوعی قوی بسیار محل بحث است. رفتارگرا نمی‌تواند رویکرد نشانه‌گرایی را مدعی باشد؛ ولی می‌تواند پیوندگرایی و نیز امکان هوش مصنوعی قوی را مدعی باشد. رفتارگراها ذهن را همان رفتار می‌دانند. پس با مدل‌سازی رفتار، هوش مصنوعی قوی رخ خواهد داد. در حال حاضر نیز توانسته‌اند رفتار را مدل‌سازی کنند.

ماتریالیست تحویل‌گرا نیز با دو بیان متفاوت می‌تواند، هم از امکان و هم عدم امکان هوش مصنوعی قوی دفاع کند. اگر مدعای آنها انحصار ذهن در مغز باشد، هوش مصنوعی قوی امکان نخواهد داشت و اگر مغز یک نمونه از ذهن و حالات ذهنی باشد، می‌توان نمونه دیگر همچون هوش مصنوعی نیز جایگزین مغز کرد که طبق آن هوش مصنوعی قوی رخ خواهد داد.

کارکردگرایی ماشینی در برخی کارکردها نمی‌تواند رویکرد نشانه‌گرایی را محقق کند؛ ولی رویکرد پیوندگرایی و هوش مصنوعی قوی را می‌تواند محقق کند؛ زیرا

کارکردگرایی در پی کارکردهای عملی است و هوش مصنوعی قوی نیز تحقق همین کارکردهای عملی است.

نوخاسته گرا نیز، نوعی دوگانه‌انگار ویژگی‌هاست؛ از این رو همانند او تحلیل می‌گردد.

حذف‌گرا، اساساً چیزی به عنوان ذهن و حالات ذهنی را نمی‌پذیرد که در پی مدل‌سازی آن رود. البته اگر ذهن و حالات ذهنی را حذف کنند، می‌توانند به راحتی از امکان هوش مصنوعی قوی سخن بگویند؛ زیرا در فرض بیان‌شده مدل‌سازی مغز و کنش و واکنش‌های آن در شیء مصنوعی هم‌اکنون نیز تا حدودی رخ داده است.



پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی



فهرست منابع

- تامپسون، مل. (۱۳۹۵). آشنایی با فلسفه ذهن (مترجم: مهران داوور، چاپ اول). تهران: نقش و نگار.
- چرچلند، پاول. (۱۳۸۶). ماده و آگاهی (مترجم: امیر غلامی، چاپ اول). تهران: نشر مرکز.
- خوشنویس، یاسر. (۱۳۸۹). انواع ویژگی‌های ذهنی و تمایز ذهنی - فیزیکی. نقدونظر، ۱۵(۵۷)، صص ۲۲-۵.
- دکارت، رنه. (۱۳۷۱). اصول فلسفه (مترجم: منوچهر صانعی دره بیدی، چاپ دوم). تهران: انتشارات بین المللی الهدی.
- دکارت، رنه. (۱۳۸۱). تأملات در فلسفه اولی (مترجم: احمد احمدی، چاپ سوم). تهران: سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت).
- هارت، ویلیام. (۱۳۸۱). فلسفه نفس (مترجم: امیر دیوانی، چاپ اول). قم: انتشارات صداوسیما.
- دیویس، مارتین. (۱۳۹۰). فلسفه ذهن (مترجم: مهدی ذاکری، چاپ اول). تهران: انتشارات حکمت.
- راسل، استوارت جی؛ نورویگ، پیتر. (۱۳۹۱). AI راهبردی نوین (مترجم: دکتر حسین حاج رسولیها، چاپ اول). تهران: نیاز دانش.
- صدرالمآلهین. (۱۹۸۱م). الحکمة المتعالیة فی الاسفار العقلیة الاربعة (ج ۳، ۸ و ۹، چاپ سوم). بیروت: دار إحياء التراث.
- طهماسبی، محمدرضا. (۱۳۸۵). رهیافت‌های بنیادین در هوش مصنوعی. حکمت و فلسفه، ۲(۶)، صص ۴۷-۲۵.
- عباس‌زاده جهرمی، محمد. (۱۳۹۰). مقایسه تطبیقی ذهن و AI (چاپ اول). تهران: بنیاد پژوهشی فرهنگ پیما غدیر.
- عموسلطانی‌فروشانی، محمدمهدی. (۱۴۰۰). بررسی رابطه نفس و بدن از منظر حکمت متعالیه و نوحسته‌گرایی (چاپ اول). قم: پژوهشگاه علوم و فرهنگ اسلامی.



قاضوی، سیدمحمد. (۱۴۰۲). بررسی انتقادی مبانی و لوازم انسان‌شناختی هوش مصنوعی قوی از منظر حکمت متعالیه (استاد راهنما: احمد واعظی). رساله دکتری. دانشگاه باقرالعلوم علیه السلام، دانشکده فلسفه اخلاق، گروه فلسفه و کلام.

لو، جاناتان. (۱۳۸۹). مقدمه‌ای بر فلسفه ذهن (مترجم: امیر غلامی، چاپ دوم). تهران: نشر مرکز.

مسلمین، کیت. (۱۳۹۱). درآمدی به فلسفه ذهن (مترجم: مهدی ذاکری، چاپ دوم). قم: پژوهشگاه علوم و فرهنگ اسلامی.

مسلمین، کیت. (۱۳۹۵). فلسفه ذهن از کتاب فهم فلسفه (مترجم: مهدی ذاکری، چاپ دوم). تهران: انتشارات علمی و فرهنگی.

موسوی، هادی. (۱۳۹۶). مسئله نفس و بدن در حکمت صدرائی با نگاهی به نوخاسته‌گرایی (چاپ اول). قم: پژوهشگاه علوم و فرهنگ اسلامی.

نریمانی، نیما. (۱۴۰۰). بستار فیزیکی و استدلال به سود طبیعت‌گرایی؛ بررسی انتقادی پژوهش‌های فلسفی کلامی، ۲۳(۹۰)، صص ۷۳-۱۰۲.

Atmstrong, D. M. (1999). *The Mind-Body problem: an opinionated introduction*. Boulder: Westview Press.

Dreyfus, Hubert L & Dreyfus, Stuart E. (1992). *Making A Mind Versus Modelling The Brain: Artifical Intelligence Back At A Branch-Point. The Philosophy of Artifical Intelligence* (Bu Margaret A. Boden, Ed.). Oxford: Oxford University Perss.

Graham, George. (1993). *Philosophy of Mind: An Introduction*. Oxford & Cambridge: Blackwell.

Newell, Alen & Simon, Herbert A. (2006). *Computer Science as Rmpirical Inquiry: Symbols and Search*. Website: David J. Chalmers. Part.

Russell, S & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Pearson Education International.



References

- Abbaszadeh Jahromi, M. (1390 AP). *A comparison of the Mind and AI* (1st ed.). Tehran: Peyman Ghadir Cultural Research Foundation. [In Persian]
- Amoo-Soltani Forushani, M.M. (1400 AP). *A study of the soul-body relation from the perspective of the Transcendent philosophy and emergentism* (1st ed.). Qom: Islamic Sciences and Culture Academy. [In Persian]
- Armstrong, D. M. (1999). *The Mind-Body problem: an opinionated introduction*. Boulder: Westview Press.
- Churchland, P. (1386 AP). *Matter and Consciousness* (A. Gholami, transl., 1st ed.). Tehran: Markaz Publications. [In Persian]
- Davis, M. (1390 AP). *Philosophy of Mind* (M. Zakeri, transl., 1st ed.). Tehran: Hekmat Publications. [In Persian]
- Descartes, R. (1371 AP). *Principles of Philosophy* (M. Sanei Darrebidi, transl., 2nd ed.). Tehran: Al-Huda International Publications. [In Persian]
- Descartes, R. (1381 AP). *Meditations on First Philosophy* (A. Ahmadi, transl., 3rd ed.). Tehran: Samt. [In Persian]
- Dreyfus, H. L. & Dreyfus, S. E. (1992). Making a Mind versus Modelling the Brain: Artificial Intelligence Back At a Branch-Point. In *The Philosophy of Artificial Intelligence* (M. A. Boden, ed.). Oxford: Oxford University Press.
- Graham, G. (1993). *Philosophy of Mind: An Introduction*. Oxford & Cambridge: Blackwell.
- Hart, W. (1381 AP). *Philosophy of the mind* (A. Divani, transl., 1st ed.). Qom: IRIB Publications. [In Persian]
- Khoshnevis, Y. (1389 AP). Varieties of mental properties and the mental-physical distinction. *Naqd va Nazar* 15(57), pp. 5-22. [In Persian]
- Lowe, J. (1389 AP). *An introduction to philosophy of mind* (A. Gholami, transl., 2nd ed.). Tehran: Markaz Publications. [In Persian]
- Maslin, K. (1391 AP). *An Introduction to Philosophy of Mind* (M. Zakeri, transl., 2nd ed.). Qom: Islamic Sciences and Culture Academy. [In Persian]

- Mousavi, H. (1396 AP). *The soul-body problem in Sadraean philosophy with a view of emergentism* (1st ed.). Qom: Islamic Sciences and Culture Academy. [In Persian]
- Narimani, N. (1400 AP). Physical closure and the argument for naturalism: a critical examination. *Philosophical theological research* 23(90), pp. 73-102. [In Persian]
- Newell, A. & Simon, H. A. (2006). *Computer Science as Empirical Inquiry: Symbols and Search*. Website: David J. Chalmers. Part.
- Qazavy, S. M. (1402 AP). *A critical study of the foundations and anthropological implications of strong artificial intelligence from the perspective of the Transcendent philosophy*. [Doctoral dissertation, Baqir al-Olum University]. [In Persian]
- Russell, S & Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*. Pearson Education International.
- Russell, S. J. & Norvig, P. (1391 AP). *Artificial Intelligence: A Modern Approach* (H. Hajj-Rasouliha, transl., 1st ed.). Tehran: Niaz-e Danesh. [In Persian]
- Şadr al-Muta'allihīn, M. (1981). *Al-Hikmat al-muta'aliya fi l-asfār al-'aqliyyat al-arba'ā* (Vols. 3, 8, 9, 3rd ed.). Beirut: Dar Ihya' al-Turath al-'Arabi. [In Arabic]
- Tahmasbi, M. R. (1385 AP). Fundamental approaches in artificial intelligence. *Wisdom and philosophy*, 2(6), pp. 25-47. [In Persian]
- Thompson, M. (1395 AP). *Teach Yourself Philosophy of Mind* (Mehran Davar, trans., 1st ed.). Tehran: Naghsh-o-Negar. [In Persian]

