

Critique of Bruce Reichenbach's Approach to Human Destiny and Spirituality in the Utopia of Artificial Intelligence

Seyd Mehdi Sajadi*

Abstract

Advances in technology in the field of artificial intelligence and entering the era of linking biological and digital achievements, depicts the perspective of the utopia of artificial intelligence that seeks to transfer the structure of the biological brain to artificial machines, so that their performance can exceed the performance of modern humans. But the realization of such an ideal, from the point of view of Bruce Reichenbach, is accompanied by theological and philosophical challenges. In this article, by analyzing the content, an attempt has been made to explain and evaluate these alleged challenges from two perspectives of Sadra's wisdom and artificial intelligence specialists in the areas of human identity, spirituality, worship of God, free will, and moral decisions. Although Reichenbach's concerns are significant, they seem to have drawbacks, such as a lack of attention to the survival of primitive patterns and the survival of choice and awareness, the possibility of transmitting religious sentiments in religious experiences, and the possibility of assuming the true God as religious. The possibility of realizing models of free will and the possibility of transition to moral awareness is faced in strong artificial intelligence.

Keywords: Strong Artificial Intelligence, Bruce Reichenbach, Sadra's Wisdom, Human Identity, Religious Experience, Free Will.

* PhD graduate in Theoretical Foundations of Islam, Ferdowsi University, Mashhad, Iran.
se.mahdisajjadi@gmail.com

Date received: 2022/08/04, Date of acceptance: 2022/09/24



نقد رویکرد بروس رایشنباخ به سرنوشت انسان و معنویت در آرمان‌شهر هوش مصنوعی

سیدمهدی سجادی*

چکیده

پیشرفت‌های تکنولوژی در عرصه هوش مصنوعی و ورود به عصر پیوند دستاوردهای بیولوژیکی و دیجیتالی دورنمایی از آرمان‌شهر هوش مصنوعی را به تصویر می‌کشد که در پی انتقال ساختار مغز بیولوژیکی به ماشین‌های مصنوعی است، تاجایی که عملکرد آن‌ها بتواند از عملکرد انسان کنونی پیش افتد؛ اما تحقق چنین آرمانی، از منظر بروس رایشنباخ، با چالش‌های الاهیاتی و فلسفی همراه است. در این مقاله با روش تحلیل محتوا سعی در تبیین و ارزیابی این چالش‌های ادعایی از دو منظر حکمت صدرایی و متخصصان هوش مصنوعی در حیطه‌های هویت انسانی، معنویت، پرستش خداوند، اراده آزاد، و تصمیمات اخلاقی شده است. هرچند دغدغه‌های رایشنباخ قابل توجه‌اند، ولی به نظر می‌رسد با استناد به نظرهای متخصصان با اشکالاتی از جمله عدم توجه به بقای الگوهای اولیه، امکان انتقال احساسات مذهبی در تجربیات دینی، امکان فرض خدای واقعی به عنوان متعلق تجربه دینی، امکان تحقق مدلهایی از اراده آزاد، و امکان گذار به آگاهی اخلاقی در هوش مصنوعی قوی مواجه است.

کلیدواژه‌ها: هوش مصنوعی قوی، بروس رایشنباخ، حکمت متعالیه، هویت انسانی، تجربه دینی، اراده آزاد.

* دکترای مبانی نظری اسلام دانشگاه فردوسی، مشهد، ایران، se.mahdisajjadi@gmail.com

تاریخ دریافت: ۱۴۰۱/۰۵/۱۳، تاریخ پذیرش: ۱۴۰۱/۰۷/۰۲



۱. مقدمه

پیشرفت روزافزون تکنولوژی، به‌خصوص در عرصه دیجیتال، نویدبخش تحولات بنیادینی در تعامل انسان و ماشین است. هوش مصنوعی از جمله حیطه‌های پژوهش دیجیتالی است که توانسته است افق‌های جدیدی را پیش‌روی بشریت قرار دهد. ادعای آرمان‌شهر هوش مصنوعی آینده‌ای است که در آن کامپیوتر می‌تواند همه آن‌چه را در مغز بیولوژیکی اتفاق می‌افتد منعکس کند و خود را مانند آن مغز سازد. ماشین‌ها می‌توانند خوداندیش و خودتعامل باشند: «زمانی که کامپیوتر به سطح هوش انسانی برسد، دیگر انسان امتیازی از نظر هوش بر ماشین نخواهد داشت» (Kurzweil 1999: 3). ماشین‌ها خواهند توانست ارزش‌گذاری کنند و احساس داشته باشند، اگرچه دقیقاً مانند ما نباشد. ماشین‌ها خواهند توانست خود به مطالعه مدارک پردازند و تحلیل ارائه دهند و آن را با دیگران به‌اشتراک بگذارند. می‌توانند تمام مکتوبات عالم را بخوانند و در کسری از ثانیه، معنای آن را در قالب دیجیتال عرضه کنند. عصر طلایی اطلاعات زمانی محقق می‌شود که ماشین‌هایی که ساختار مغز ما را حمل می‌کنند از انسان بیولوژیکی پیش بیفتند. این عصر طلایی مدعی است درنهایت یک روز شاهد جای‌گزینی انسان دی.ان.ای محور با ماشین کامپیوتری خواهیم بود که می‌تواند بدون خطا میلیون‌ها بار سریع‌تر و صحیح‌تر از نورون‌های کربنی در مغز پردازش انجام دهد. متخصصان به‌دنبال آینده ماشینی هستند که انسان بتواند ذهنیات خود را با آن آپلود کند و از دنیای مجازی در کنار بدن‌های مجازی لذت ببرد. متخصصان هوش مصنوعی در پی تحقق دو نوع هوش مصنوعی هستند؛ هوش مصنوعی ضعیف مانند آن‌چه امروزه در کامپیوترها مشاهده می‌کنیم و هوش مصنوعی قوی که در پی بازساخت مغز بیولوژیکی است، تاجایی که عملکرد آن بتواند تولید شخص کند.

هدف این پژوهش بررسی نظرهای بروس رایشنباخ درباره چالش‌های هوش مصنوعی قوی در عرصه‌های هویت انسانی، معنویت و تجربه دینی، پرستش خدای حقیقی یا مجازی، اراده آزاد، و تصمیم اخلاقی است. بروس رایشنباخ (Bruce R. Reichenbach, 1943) استاد بازنشسته در گروه فلسفه دانشگاه آگسبورگ مینیاپولیس ایالات متحده آمریکا است. وی از نویسندگان همکار در نگارش کتاب *عقل و اعتقاد دینی* به‌همراه مایکل پترسون است و تاکنون کتاب‌ها و مقالات متعددی را منتشر کرده است. او معتقد است که مسیحیت توانسته است با چالش‌های نئوداروینیسم سازگار شود، ولی هوش مصنوعی حداکثری چالش‌ها و مسائل جدی برای مسیحیت خواهد داشت. در این پژوهش این مسائل طرح و بررسی خواهد شد. آیا در فرایند انتقال الگوهای ذهنی انسان به هوش مصنوعی هویت انسانی محفوظ خواهد ماند؟ آیا با

گذشت زمان و افزوده شدن اطلاعات، الگوها، و خاطرات باز هم هویت اولیه انسان برقرار است؟ آیا خود حقیقی انسان دست‌خوش تغییر خواهد شد؟ جایگاه معنویت، تجربه دینی، و خدای حقیقی در عصر انسان مصنوعی چگونه است؟ آیا هوش مصنوعی خواهد توانست فراتر از برنامه‌ریزی خود عمل کند یا با آن مخالفت کند؟ آیا هوش مصنوعی می‌تواند مانند انسان دارای اخلاقیات و تصمیمات اخلاقی شود؟ آیا صرف واکنش مناسب به تغییرات محیطی و یا توانایی بروز احساسات مشابه انسانی جای‌گزین اخلاق خواهد بود؟

پیشینه پژوهش درباره هوش مصنوعی اکثراً به حوزه هوش مصنوعی ضعیف، که کاربرد گسترده‌تری در عصر حاضر دارد، معطوف است. اما پژوهش‌هایی که چه مستقیم و چه غیرمستقیم بیش‌تر به حوزه هوش مصنوعی قوی مرتبط باشند عبارت‌اند از: مقاله «نگرشی انتقادی به مسئله هوش مصنوعی»، نوشته حبیب کارکن بیرق، که به مناقشات درباره تحقق هوش حقیقی در ماشین پرداخته است؛ مقاله «دین و هوش مصنوعی»، نوشته علیرضا قائمی، که به سازگاری مفهوم تفکر نزد محققان هوش مصنوعی با تفکر نزد دین پرداخته است و امتیازات دیگر انسان را بر ماشین شرح می‌دهد؛ مقاله «بررسی فلسفی امکان تحقق هوش مصنوعی قوی باتوجه به دیدگاه‌های مختلف در مسئله ذهن و بدن»، نوشته حسین مطلبی کربکندی و دیگران که به بحث انواع نظریات درباره رابطه ذهن و بدن از منظر فلسفه ذهن و نیز فلسفه اسلامی و تطبیق آن بر هوش مصنوعی قوی پرداخته است. تمرکز بیش‌تر مقالات پیشین بر بحث‌های نظری درباره اصل تحقق هوش مصنوعی و نیز سرنوشت رابطه ذهن و بدن در عصر هوش مصنوعی است. اما تفاوت مقاله حاضر با مقالات پیشین ورود به مباحثی مانند سرنوشت هویت انسانی، معنویت، پرستش خداوند، اراده آزاد، و تصمیم اخلاقی است. لذا مقاله حاضر جنبه‌های جدیدتری را شامل می‌شود. هم‌چنین، درباره نظرهای بروس رایشنباخ کار چندانی انجام نشده است و مقاله «بررسی نظریه دادباوری شرور طبیعی از منظر بروس رایشنباخ»، نوشته عبدالرسول کشفی و آرزو مشایخی، یافت شد که به موضوع مسئله شر می‌پردازد.

۲. اصطلاحات هوش مصنوعی ضعیف و قوی

مطالعات دانشگاهی درباره هوش مصنوعی را شاید بتوان به دهه ۱۹۵۰ بازگرداند. هوش مصنوعی ضعیف در اکثر کامپیوترهای امروزی کاربرد دارد، ولی آنچه موردنظر این پژوهش است اصطلاح هوش مصنوعی قوی است. اصطلاح هوش مصنوعی قوی را سرل (Searle) مطرح کرد که اشاره به دیدگاه کسانی دارد که معتقدند: «نسبت ذهن به مغز مثل نسبت برنامه به

سخت‌افزار کامپیوتر است» (سرل ۱۳۸۲: ۵۶). هوش مصنوعی ضعیف اشاره به ماشین مصنوعی دارد که توان بروز رفتارهایی در سطح رفتار انسانی را داشته باشد؛ مثلاً مانند انسان سخن بگوید و آزمون تورینگ را پشت سر بگذارد. هوش مصنوعی قوی تلاش برای ساختن ماشینی است که تمامی قوای ذهنی انسان را داشته باشد، مانند فهمیدن معنای گفته‌ها و اعمال و هم‌چنین داشتن آگاهی پدیداری از اتفاقات محیطی (دریفوس ۱۳۹۹: ۲۷). هوش مصنوعی قوی معتقد است که باید به فراتر از شبیه‌سازی و کپی‌برداری ساختارهای مغز برویم تا بتوانیم به انتقال محتویات ساختاری برسیم (ساختاری که به‌وسیله مهندسی معکوس ذهن تکمیل شده است): «ما می‌توانیم از معماری، ساختار، و دانش درونی مغز انسان بهره‌مند شویم تا دانشمان را از چگونگی طراحی هوش در یک ماشین به‌وسیله کاوش در مدار مغز تسریع کنیم؛ سپس می‌توانیم از آن کپی‌برداری یا تقلید کنیم» (Kurzweil, 1999: 120). «از نظر هوش مصنوعی قوی دیگر بدن‌های بیولوژیکی ما بی‌ارزش می‌شوند. البته اگر به بدن نیاز داریم می‌توانیم از بدن‌های مصنوعی یا اتصال نانوربات‌ها به مغز کمک بگیریم» (ibid.: 129). درنهایت میلیاردها موجود غیریولوژیکی می‌توانند حاکم بر همه انسان‌ها باشند (Kurzweil 2002: 13). این تحول درباره بشر جدید می‌تواند در چند سطح واقع شود: ۱. کاشت دستگاه ماشینی مستقیماً در مغز انسان که این ویژگی ما را قادر می‌سازد تا تجربیات مجازی با دیگران داشته باشیم؛ ۲. انسان نیز متعاقباً اطلاعات خود را به کامپیوتر منتقل کند. اگر یک کامپیوتر خراب شود می‌تواند اطلاعات خود را به دیگری کپی کند. پس نابودی یک کامپیوتر باعث پایان‌یافتن تجربیات فردی نمی‌شود: «ما می‌توانیم به یک ماشین جدید منتقل شویم و نه تنها ابدی می‌شویم، بلکه هیچ‌یک از خطرات و اطلاعات خود را از دست نخواهیم داد» (ibid.: 14). به‌باور رایشنباخ، آرمان‌شهر هوش مصنوعی قوی چالش‌های جدی پیش‌روی باورهای دینی و فلسفی خواهد گذاشت. در ادامه به طرح این چالش‌ها و ارزیابی آن پرداخته می‌شود.

۳. چالش حفظ هویت انسانی

آیا انسان انتقال‌یافته به هوش مصنوعی با انسان موجود همانند خواهد بود؟ آیا شخص انتقالی فاقد هویت انسان اصیل نخواهد بود؟ هانس موراوک (Hans Moravec, 1948)، دانشمند برجسته هوش مصنوعی، در پاسخ به این سؤال معتقد است که اگر معیار هویت را هویت بدنی بدانیم، آن‌گاه حفظ چنین هویتی در پروژه کپی‌سازی انسان با چالش مواجه می‌شود، ولی وی اصطلاح «هویت الگویی» را پیشنهاد می‌دهد و معتقد است که هویت الگویی ذات انسان را

همان الگوها و فرایندهای مغزی می‌داند، لذا تازمانی که این فرایندها محفوظ بماند هویت ما نیز محفوظ است و بقیه امور مانند بدن و سر جانبی هستند. حتی در زندگی روزمره نیز اتصال مکانی و زمانی بدن ضروری نیست. برای مثال، سلول‌های تشکیل‌دهنده بدن در حال زوال و نوسازی هستند و این تداوم الگوهاست که هویت ما را تداوم می‌بخشد. در نهایت، وی تأکید می‌کند انسان مصنوعی در فرایند انتقال الگوها اتصال بدنی، مکانی، و زمانی نخواهد داشت (Moravec 1988: 116-117). رایشنباخ چندان با موراوک در این بحث مخالف نیست، زیرا از نظر رایشنباخ نیز اتصال مکانی و زمانی برای هویت انسانی ضروری نیست. در فرض هوش مصنوعی قوی آن‌چه برای هویت مهم است تداوم الگوها و فرایندهاست، به‌خصوص آن‌ها که به ذهن مربوط است (Reichenbach 2021: 110).

سؤال دوم درباره هویت انسانی آن است که آیا این الگوها می‌توانند بی‌نهایت تکرار شوند؟ از نظر موراوک مشکلی در این باره نیست و مانند کپی‌سازی از موسیقی و اطلاعات می‌ماند، پس می‌توانیم کپی‌های متعددی از الگوهای خود بسازیم که هریک برابر با نسخه اصلی باشد (Moravec 1988: 118). سؤال سوم آن است که در هوش مصنوعی قوی، به فاصله کوتاهی پس از همانندسازی، تمام نسخه‌ها همان نسخه فردی است، اما پس از گذشت زمان متفاوت از فرد اصلی خواهند شد، زیرا بر اثر افزودن اطلاعات برخوردار از الگوهای جدید می‌شوند. ولی اگر ما متفاوت و متعدد شویم، چگونه می‌توانیم بگوییم هویت فردی ما مصون از این انتقال بیولوژیکی به مصنوعی خواهد بود؟ به عبارت دیگر، ما کدام نسخه المثنی خواهیم بود؟ موراوک معتقد است که این گسترش الگوها در طی زمان مشکلی برای هویت انسان ایجاد نمی‌کند، زیرا چنین وضعیتی مشابه انتقال صفات انسانی از طریق تکثیر ژنتیکی فرد از نسلی به نسل بعد است (ibid.: 115).

در این‌جا دو چالش مطرح است اولاً در صورت تکرار الگوهای یکسان وضعیت اصل «همانندی امور تمایزناپذیر» چگونه خواهد بود؟ رایشنباخ معتقد است که در این‌جا اصل «همانندی امور تمایزناپذیر» جاری نیست، زیرا این اصل مربوط به امور مادی و جسمانی است و نه امور اعتباری مانند کپی‌برداری الگوها (Reichenbach 2021: 110)؛ ثانیاً کدام نسخه المثنی برابر با نسخه اصلی است؟ در این سؤال رایشنباخ با پاسخ موراوک مخالف است و آن مثال را قانع‌کننده نمی‌داند و حتی معتقد است که شباهت مناسبی نیز با بحث حاضر ندارد، زیرا عوامل هویت‌ساز مانند حالات روانی ما در آن انتقالات ژنتیکی حضور ندارند. پس اساساً انسان نسل بعد با انسان نسل قبل تداوم هویتی ندارد تا پس از آن این سؤال مطرح شود که کدام نسل بعدی

دقیقاً همان فرد اولیه است. اما در هوش مصنوعی قوی، به دلیل تداوم هویت، از طریق تداوم الگوهای ذهنی جای این سؤال مطرح است که در صورت تکثیر نسخه های مصنوعی یک فرد کدام نسخه دقیقاً برابر با همان نسخه اصلی است؟ از نظر رایشنباخ این پرسش پاسخ مناسب خود را نمی یابد (ibid.).

سؤال چهارم درباره برهم خوردن هویت فردی انسان با گذشت زمان برای نسخه هوش مصنوعی ماست، زیرا موراوک بیان می دارد حالات روانی شامل خاطرات، اعتقادات، و دیدگاه ها می توانند از یک الگوی بزرگ تر و حتی صفات حیوانی به الگوی یک فرد و از آن فرد به دیگری افزوده شوند و پس از گذشت زمان هر فرد به ترکیبی از اطلاعات جهانی تبدیل می شود. خاطرات نه تنها می توانند میان نسخه های متعدد یک شخص اشتراک گذاری شوند، بلکه می توانند میان اشخاص مختلف نیز به اشتراک گذاشته شوند. موراوک از این اشتراک گذاری به «برترین روش ارتباط» تعبیر می کند و خود نیز معترف است که مفاهیمی مانند هویت معانی اکنون خود را از دست خواهند داد (Moravec 1988: 115)؛ اما رایشنباخ متذکر می شود که اگر خاطرات نقش برجسته در تداوم هویت ما دارند، این تداوم هویت با گذشت زمان در نسخه هوش مصنوعی ما به خطر می افتد، زیرا با افزودن خاطرات دیگران به خود ما «تصویر اول شخص» متعددی خواهیم داشت که از یکی به دیگری منتقل می شود و حتی حاوی خاطراتی خواهیم بود که هرگز تجربه نکرده ایم (Reichenbach 2021: 111).

پیش از بیان ارزیابی نظرهای رایشنباخ باید توضیحی درباره دیدگاه وی درباره عدم نیاز به تداوم مکانی و زمانی بدن برای حفظ هویت فردی و نیز ملاک وی درباره این همانی هویت فرد بیان شود. به طور خلاصه، او در بحث معاد به خلق دوباره انسان قائل است و در توجیه این نظر خود کلیت لزوم تداوم زمانی و مکانی بدن را زیر سؤال می برد و ملاک های دیگری برای حفظ هویت انسانی بیان می کند. توضیح مطلب آن که قول به خلق مجدد یعنی شخص مرده به معنای حقیقی کلمه دوباره خلق شود. خداوند قادر است همان شخص معدوم شده را پس از مدت ها دوباره خلق کند. در این مدت وقفه ای در زندگی فرد ایجاد می شود که در این وقفه معدوم است، ولی قبل و بعد از آن موجود است (تالیافرو ۱۳۸۲: ۵۳۲). از نظر رایشنباخ اگر خداوند توان خلق ابتدایی انسان را دارد، معقول است که این خدای قادر و عالم مطلق بتواند انسان را از نظر فیزیکی دوباره هم خلق کند (به همراه تمام خصیصه های مغزی پیشین وی)، این امر سبب می شود که انسان با تمام آگاهی ها و خاطراتش بازیابی شود. درحقیقت این انسان ها همان افرادی هستند که مرده بودند. خداوند می تواند مغز آنان را چنان بسازد که ساختار عصبی

آنان از نظر هویت و اصالت شبیه همان افراد قبل از مرگ باشد، پس آنان همان ایده‌ها، دیدگاه‌ها، خاطرات، و ویژگی‌های شخصیتی را دارند که پیش از مرگ داشتند (Reichenbach 2021: 104).

رایشنباخ برای تردید در لزوم تداوم زمانی و مکانی بدن برای حفظ هویت انسانی به این مطلب توجه می‌دهد که سلول‌های بدن انسان در همین دنیا تجدید می‌شوند و در حال تغییرند، لذا بدن ما دقیقاً همان بدن مثلاً دوازده سال پیش نیست (Reichenbach 1983: 85). البته او اصل ادعای لزوم اتصال مکانی و زمانی را در حد کلیت می‌پذیرد و اذعان می‌دارد دشوار است اشیای مادی را بیابیم که فاقد این معیار باشند، ولی باز با هم برابر باشند و نه نسخه بدل، اما وی قائل است که استثناهایی نیز وجود دارند. برای مثال، یک قطعه موسیقی را در نظر بگیریم که به‌طور متناوب با فاصله تکرار می‌شود؛ اگرچه این دو قطعه اتصال مکانی و زمانی ندارند، ولی حکم به این‌همانی آن دو می‌کنیم، هرچند می‌دانیم قطعه دوم کپی قطعه اول است. یا مثال دیگر کپی کردن اطلاعات هارد یک کامپیوتر به کامپیوتر دیگر است، حال زمانی که اطلاعات کامپیوتر اول نابود شود، شکی نیست که اطلاعات هارد دوم همان اطلاعات قبلی هستند. مثال سوم بازیگری است که در چند پرده مختلف تئاتر ظاهر می‌شود و به اجرای نقش مثلاً هملت می‌پردازد. مسلماً زمانی که این بازیگر در میان دو پرده به استراحت می‌پردازد دیگر هملت نیست، ولی با آمدن به روی صحنه و اجرای دوباره نقش هملت همگان حکم به این‌همانی دو نقش هملت خواهند کرد. پس اتصال مکانی و زمانی در این موارد برای حکم به این‌همانی آن‌ها نیاز نیست. در خلق دوباره انسان مرده نیز خداوند قادر است که اطلاعات او را مثلاً پیش از نابودی به بدن جدید منتقل می‌کند (Reichenbach 2021: 105). حال اگر اتصال مکانی و زمانی برای هویت‌سازی نیاز نیست، پس چه چیز نیاز است؟ رایشنباخ مواردی را برمی‌شمارد مانند صرف شباهت فیزیکی، وحدت آگاهی‌های درونی شامل خاطرات، و خصوصیات فردی (Reichenbach 1978: 31). وی از همین نظرها در بحث هوش مصنوعی و چگونگی تداوم هویت انسانی در زمان انتقال الگوهای مغز انسانی به ماشین مصنوعی بهره برده است.

۱.۳ نقد و بررسی

در ارزیابی نظرهای رایشنباخ ابتدا به بیان اشکالات مبنایی طبق حکمت متعالیه در نقد نظریه خلق مجدد و ملاک هویت انسانی از منظر ملاصدرا پرداخته می‌شود و سپس سایر اشکالات بیان خواهد شد. اشکال مبنایی اول متوجه بحث معدوم‌شدن یک شیء و دوباره خلق‌شدن آن

است که در فضای علوم اسلامی با عنوان «اعاده معدوم» در آثار کلامی و فلسفی با عنوان «تکرارپذیر نبودن تجلیات الاهی» در آثار عرفانی و گاهی با عنوان «نفی تکرار در وجود» در برخی آثار فلسفی آمده است (جوادی آملی ۱۳۸۲: ج ۱، ۴۳). ملاصدرا چهار دلیل بر نفی اعاده معدوم آورده است:

دلیل اول را می‌توان متشکل از چند مقدمه دانست: ۱. ماهیت عدم رفع هستی و وجود است. در این جا ماهیت به معنای مفهوم است یعنی مفهوم عدم رفع وجود است (همان: ۴۶)؛ ۲. هویت یک شیء همان وجود آن است. این مقدمه براساس اصالت وجود طراحی شده است؛ ۳. چون هویت هر شیء یگانه است، پس وجود هر شیء هم یگانه است؛ یعنی بیش از یکی نیست؛ ۴. از انضمام مقدمه ۱ و ۳، یعنی عدم همان رفع وجود است و هر شیء بیش از یک وجود ندارد، نتیجه می‌شود که درمورد هر شیء بیش از یک رفع وجود یا همان عدم متصور نیست. از مجموع این چهار مقدمه نتیجه می‌شود که چون برای یک شیء دو وجود تصور نمی‌شود، پس برای آن، دو رفع هم تصور نمی‌شود. پس هر شیء یک رفع و یک وجود دارد. در این صورت، اگر شیء معدوم شد، دیگر اعاده وجود آن امکان ندارد (صدرالدین شیرازی ۱۳۶۸: ج ۱، ۳۵۳). این مطلب را می‌توان به بیان علامه طباطبایی «تکرارناپذیر بودن وجود» نامید (جوادی آملی ۱۳۸۲: ج ۱، ۴۵).

دلیل دوم قیاسی استثنایی را تشکیل می‌دهد: اگر اعاده معدوم جایز باشد، واحد کثیر و متعدد خواهد شد و تالی یعنی کثرت و تعدد واحد باطل است؛ پس مقدم یعنی اعاده معدوم نیز باطل است.

دلیل سوم بیان می‌دارد اگر اعاده معدوم جایز باشد، حیثیت ابداء و اعاده، با آن که متنافی هستند، عین یکدیگر خواهند بود و عینیت دو امر متناهی محال است؛ پس اعاده معدوم جایز نیست. دلیل تنافی ابداء و اعاده آن است که شیء ابتدایی مسبوق به عدم ازلی است، ولی شیء معاد مسبوق به وجود است، پس منافی هستند. دلیل عینیت هم آن است که چون وجود اولی و دومی هویت واحدی دارد، پس امتیازی از هم ندارند پس باید دو حیثیت ابداء و اعاده هم یکسان باشند (همان: ج ۱، ۴۷).

دلیل چهارم آن است که اگر اعاده اول جایز باشد، اعاده دوم، سوم، و غیره هم جایز است؛ یعنی جواز اعاده‌های نامتناهی برقرار است. اما اگر همه با هم بالفعل محقق شوند، فعلیت همه آحاد محال است و اگر برخی اعاده شوند ترجیح بلامرجح است. پس هیچ کدام عود نمی‌کنند (صدرالدین شیرازی ۱۳۶۸: ج ۱، ۳۵۳).

علاوه بر فلسفه اسلامی، نقد نظریه خلق مجدد در برخی کتب غربیان آمده است (پترسون و دیگران ۱۳۸۹: ۳۴۰-۳۴۳؛ کوپر ۱۳۹۲: ۴۱۷-۴۱۸). رایشنباخ با نظر متخصصان هوش مصنوعی قوی موافق است که اتصال مکانی و زمانی برای هویت انسانی لازم نیست و چنان‌که در بحث معاد امکان اعاده معدوم یا همان خلق مجدد انسان مقبول است، پس در رستاخیز انسان در ماشین هم می‌توان از انسان دوباره ایجاد شده در قالب ماشین سخن گفت و با این حال هویت انسانی پابرجا بماند. اما اعاده معدوم در حکمت صدرایی محال است و دلایل آن گذشت و لذا اگر انسان بمیرد، ولی ساختار مغز او به ماشین منتقل شود، نمی‌توان برای توجیه رستاخیز انسان در ماشین و حفظ هویت آن از بحث اعاده معدوم کمک گرفت.

درباره ملاک این‌همانی هویت انسانی در انسان بیولوژیکی و انسان مصنوعی، رایشنباخ به تداوم این هویت از طریق تداوم خاطرات قائل شد. در ارزیابی این نظر باید گفت میان ملاک متافیزیکی این‌همانی هویت و ملاک معرفت‌شناسی تفاوت است. در حکمت متعالیه وجود خاص شیء ملاک متافیزیکی این‌همانی است، زیرا تشخیص بالذات از آن وجود است. صدرا از این تشخیص در مقدمه دوم دلیل اول بر نفی اعاده معدوم استفاده کرد (صدرالدین شیرازی ۱۳۶۸: ج ۱، ۳۵۳). ملاک قرار دادن صرف نفس به عنوان ملاک تشخیص به دلیل کلی بودن آن و قابلیت فرض صدق آن بر کثیرین در دستگاه فلسفی صدرایی صحیح نیست. محفوظات ذهنی نیز ملاک صحیحی نیست، زیرا ممکن است با فراموشی و نیز خلط میان محفوظات ظاهری و واقعی روبه‌رو شود. حال اگر وجود انسان بیولوژیکی از بین رفته است و انسان مصنوعی قرار است تنها الگوهای مغزی و یا خاطرات او را نگه دارد، در این صورت دیگر نمی‌توان از تداوم هویت انسان سخن گفت. در این جا برخی اشکالات دیگر نیز به نظر می‌رسند:

اشکال اول آن است که پذیرش تداوم هویت انسانی از طریق تداوم الگوهای مغزی تبعات و الزاماتی را به همراه خواهد داشت. برای مثال، اگر بتوان یک کپی از این الگوها تهیه کرد، مانعی برای تهیه کپی‌های دیگر نیز نخواهد بود، زیرا اگر نسخه اولیه را نسخه‌ای با قابلیت کپی‌برداری بدانیم، یعنی تکثیرپذیری جزو ساختار اصلی آن باشد، آن‌گاه هیچ مانعی برای تصور امکان تهیه کپی‌های بیش‌تر وجود نخواهد داشت. توجه به چنین تبعاتی می‌تواند در حل چالش‌های احتمالی دیگر نیز مؤثر باشد.

اشکال دوم به رایشنباخ آن است که به نظر می‌رسد می‌توان از پاسخ موراوک در عدم منافات گسترش الگوهای ذهنی و خاطرات با تداوم هویت انسانی دفاع کرد، زیرا موراوک به از بین رفتن الگوها و خاطرات اولیه قائل نیست، بلکه از قابل اشتراک شدن این الگوها در طی زمان

سخن می‌گوید. بدن مصنوعی در این فرض حامل خاطرات بیش‌تر از آن چیزی است که بدن اولیه بیولوژیکی داشت. در این جا به هیچ وجه بحث از بین رفتن خاطرات قبلی نیست. موراوک از واژگانی چون ترکیب، آمیختگی، و سیال بودن برای اشاره به جابه‌جایی خاطرات استفاده می‌کند (Moravec 1988: 115)، اما رایشنباخ چنان تفسیری از کلام موراوک ارائه داد که گویی الگوهای اولیه کپی شده به کلی از بین خواهند رفت و الگوها و خاطرات جدید جای‌گزین خواهند شد.

اشکال سوم متوجه بحث انتخاب است. موراوک تصریح می‌کند که این ادغام و اشتراک‌گذاری انتخابی خواهد بود، به این بیان که می‌توان انتخاب کرد که خاطرات خاصی به دیگران منتقل شود یا دریافت شود (ibid.). اگر حق انتخاب باقی مانده است، لازمه آن وجود آگاهی به انتخاب‌های خویش است؛ یعنی فرد می‌داند چه خاطراتی را دریافت کرده است که مال خود وی نبوده است. لذا برخلاف نظر رایشنباخ، هنوز هویت فردی و انسانی محفوظ مانده است و از بین نرفته است.

اشکال چهارم آن است که با فرض نبود حق انتخاب نیز داشتن خاطراتی که هرگز تجربه نکرده‌ایم مشکل اساسی ایجاد نمی‌کند، زیرا آگاهی و درک چنین دارایی کافی است، به این معنا که من می‌دانم این خاطرات من نیست یا آگاهم که این خاطراتی است که من تجربه نکرده‌ام، هرچند به اختیار خودم، وارد ذهن من نشده باشد. پس با وجود آگاهی هنوز هویت انسانی پابرجاست.

اشکال پنجمی نیز وجود دارد، زیرا هرچند می‌توان با رایشنباخ موافقت کرد که مثال موراوک در مشابهت انتقال الگوها با ادغام ژن‌های انسان‌ها در طی تکثیر نسل‌ها مناسب نیست، اما موراوک مثال دیگری نیز می‌آورد که رایشنباخ به آن اشاره نکرده است و آن اکتساب مهارت‌ها و نگرش‌های دیگران در طول زندگی اکنون ماست (ibid.). این مثال بهتر از مثال پیشین است، زیرا افزوده شدن چنین نگرش‌هایی به فرد در طول حیات خود باعث برهم خوردن هویت اولیه فرد نخواهد شد.

اشکال ششم درباره تصویر اول‌شخص متعدد است. باید گفت اگر تعدد این تصاویر درک شود، پس هنوز من درک‌کننده واحدی وجود دارد و اگر این تعدد درک نشود و شخص نتواند خاطرات خود را از دیگری تمایز دهد، به عبارتی وحدت خاطرات صورت گرفته است، باز همین که شخص آن‌ها را به خود منسوب کند حاکی از وجود من حکم‌کننده واحدی خواهد بود. در نهایت می‌توان چنین نتیجه گرفت که تمام چالش‌های ادعایی رایشنباخ در این بخش (هویت انسانی) قابل تصحیح هستند.

۴. چالش معنویت و تجربه دینی

رایشنباخ این سؤال را مطرح می‌کند که رابطه هوش مصنوعی با دین چگونه خواهد بود؟ وی با استناد به گفتار ریموند کورزوویل (Raymond Kurzweil, 1948)، از دانشمندان برجسته هوش مصنوعی قوی، معتقد است که انسان مصنوعی هنوز دارای نوعی معنویت است. اساساً عنوان یکی از کتاب‌های کورزوویل نیز عصر ماشین‌های معنوی (*The Age of Spiritual Machines*) است؛ اما رایشنباخ به تفاوت نگاه کورزوویل به معنویت انسان مصنوعی با نگاه دینی توجه می‌دهد. وی به این جمله کورزوویل استناد می‌کند: «تنها موجود تجربه‌کننده و موجود آگاه موجود معنوی و نشان‌دهنده ذات معنویت است» (Kurzweil 1999: 153). در این جمله به نظر می‌رسد کورزوویل معنویت را به صرف آگاهی کاهش داده است.

در عبارتی دیگر، به نظر می‌رسد کورزوویل به درک تجربیات متعالی برای هوش مصنوعی قوی قائل است: «احساس تعالی، که از مرزهای فیزیکی روزمره فراتر می‌رود، برای رسیدن به درک عمیق‌تری از واقعیت است» (ibid.: 151). رایشنباخ با این توانایی تجربه متعالی موافق است. وی بنابر یافته‌های علمی بیان می‌کند که اگر در «مغز مکانی برای خداست» منظور مکان خاصی از سلول‌های عصبی در بخش جلویی است که در موقعیت تجربه دینی فعال می‌شود، آن‌گاه این مکان می‌تواند در بازسازی کامپیوتری مغز نیز واقع شود. در این صورت، فرد کامپیوتری نیز تجربه دینی از خدا خواهد داشت و حتی با بدن مصنوعی خود می‌تواند فعالیت‌هایی چون پرستش، دعا، و مدیتیشن را تجربه کند. رایشنباخ این فرض را هم نامحتمل نمی‌داند که انسان مصنوعی بتواند از طریق اتصال به ابعاد معنوی سایر ماشین‌ها به تبلیغ مسیحیت بپردازد (Reichenbach 2021: 111).

۱.۴ نقد و بررسی

آیا منظور کورزوویل از تجربیات معنوی صرفاً آگاهی است؟ رایشنباخ چنین برداشتی دارد، اما این موضوع نیازمند بررسی بیش‌تر است. کورزوویل بحث خود را درباره تجربیات معنوی مغز با این فرض شروع می‌کند که مغز می‌تواند مستقیماً تجربیاتی داشته باشد که تاکنون گمان می‌کردیم فقط از طریق تجربیات بیرونی میسر است. وی به یافته‌های دانشمندان عصب‌شناس درباره تحریک مکانی در مغز برای تولید تجربه شادی و جنسی اشاره می‌کند. وی مایل است این عمل را رهاسازی مغز بنامد (Kurzweil 1999: 149). چنین اکتشافی می‌تواند راه را برای انتخاب، گزینش، و تلفیق تجربیات مختلف باز کند. هم‌چنین، می‌تواند توانایی کنترل و

برنامه‌ریزی احساسات را در هوش مصنوعی قوی فراهم کند. البته وی به احتمال سوءاستفاده یا افراط در استفاده از چنین اکتشافی نیز اشاره می‌کند (ibid.: 150). پس از بیان این دستاورد کورزوئل این پرسش را مطرح می‌کند که دربارهٔ تجربیات معنوی چه باید گفت؟ وی ابتدا تعریف عامی از تجربهٔ معنوی، که شامل تمامی تجربیات فراتر از تجربیات روزمرهٔ فیزیکی می‌شود، ارائه می‌دهد، ولی در مقام بیان گوناگونی چنین تجربیاتی مثال‌هایی از رقص عرفانی در برخی فرقه‌ها و برعکس سکوت در مراسمات بوداییان می‌زند؛ مثال دیگر زمانی است که وی به تأثیر موسیقی در برانگیختن تجربیات معنوی اشاره می‌کند و آن را مانند تجربهٔ حاصل از مدیتیشن می‌داند (ibid.: 151-152). این مثال‌ها نشان می‌دهد که منظور وی از تجربیات معنوی می‌تواند دربردارندهٔ احساسات و اشتیاق حاصل از اعمال دینی باشد و نه صرف آگاهی.

اما عبارتی که رایشنباخ از کورزوئل نقل کرد که «تنها موجود آگاه نشان‌دهندهٔ ذات معنویت است» کلامی است که کورزوئل در باب مقایسهٔ میان هوش مصنوعی و انسان کنونی آورده است (ibid.: 153). هوش مصنوعی برساخته از الگوهای انسانی مدعی داشتن آگاهی و تجربیات معنوی خواهد بود و این ادعا باور عمیقی در وجودش است، تاجایی که قائل است این تجربیات دارای معنا هستند؛ اما انسان کنونی نیاز به ادعا ندارد، بلکه اساساً تجربه‌کننده و آگاه است. در نتیجه تجربیات معنوی انسان کنونی اصیل و تجربیات معنوی نسخهٔ هوش مصنوعی برساخته و الگوبرداری‌شده و غیراصیل است، نه این که این تجربیات غیردینی باشند. هم تجربیات معنوی انسان بیولوژیکی و هم تجربیات معنوی هوش مصنوعی قوی جنبهٔ دینی دارند، ولی موجودی که اصالتاً قابلیت تجربه و آگاهی را دارد منعکس‌کنندهٔ اصالت معنویت است و موجود دست‌ساخته، که چنین قابلیت‌هایی را الگوبرداری کرده است، مدعی معنویت است.

۵. پرستش خدای واقعی یا خدای مجازی

اما آیا انسان از خداوند واقعی بی‌نیاز خواهد شد؟ برداشت رایشنباخ آن است که این ماشین‌ها، ربات‌ها، یا آدم‌های سایبورگ (cyborg مخفف دو واژهٔ سایبرنتیک (cybernetic) و ارگانیسم (organism) است و اشاره به یک موجود با هر دو اجزای ارگانیکی و مکانیکی دارد) می‌توانند به خدایی مجازی معتقد باشند و تجربیات معنوی داشته باشند، زیرا این توانایی را دارند که به‌طور مجازی با دیگران در ارتباط باشند، اما در این فرض دیگر خدای واقعی، که نقش معناآفرین برای تجربیات بشری دارد، برای اعمال دینی و تجربیات دینی ضروری نخواهد بود (Reichenbach 2021: 112).

۱.۵ نقد و بررسی

آیا هوش مصنوعی نیاز به خداوند حقیقی ندارد؟ رایشنباخ آن را ضروری ندانست. در بررسی این مطلب باید گفت کورزوویل با اشاره به یافته‌های عصب‌شناسان دانشگاه کالیفرنیا در سن‌دیگو بیان می‌دارد که تحریک بخشی از مغز می‌تواند تجربیاتی مشابه تجربیات دینی ایجاد کند. کورزوویل، بنابر یافته‌های زیست‌شناسان تکاملی و نیز سخنان ریچارد هریس، کشیش آکسفورد، معتقد است که می‌توان گفت خداوند ما را به‌هم‌راه توانایی فیزیکی برای ایمان آفریده است. حال اگر دانشمندان بتوانند ارتباطات عصبی گونه‌های مختلف تجربیات معنوی را معین کنند، خواهند توانست این تجربیات را مانند سایر تجربیات تقویت کنند (Kurzweil 1999: 152-153). در نتیجه، عبارات کورزوویل در این‌جا نشان می‌دهد وی خدای حقیقی را هم‌چنان متعلق چنین تجربیات معنوی حاصل از تحریک عصبی می‌داند. این خدای حقیقی است که زمینه‌های فیزیکی درک ایمان را در مغز انسان قرار داده است. تکنولوژی تنها به کشف مکان آن و تقویت یا احضار آن کمک خواهد کرد. پس می‌توان فرض کرد در انتقال الگوهای مغز به هوش مصنوعی باز هم تحریک این مکان موجب درک تجربه‌ای از همان خدای حقیقی خواهد شد.

۶. چالش آزادی انتخاب

رایشنباخ آرمان هوش مصنوعی قوی را اکتساب حداکثری دانش و اطلاعات می‌داند و زمانی که به این هدف دست یافت از همه انسان‌ها پیشی می‌گیرد. وی به تعبیری که دیگران برای موجود نهایی مدنظر هوش مصنوعی قوی به‌کار می‌برند اشاره می‌کند، مانند آن‌که انسان در آرمان‌شهر هوش مصنوعی به حاکمی همه‌چیزدان تبدیل خواهد شد یا ارزش ذاتی انسان‌ها به اطلاعات اکتسابی آن‌ها خواهد بود. وی چنین آرمانی را خالی از چالش نمی‌داند، زیرا معتقد است که دانستن به‌خودی‌خود ارزش نیست، بلکه انسان دانش را برای تأثیر معنادار در محیط زندگی، روابط با دیگران، و برای عمل می‌خواهد؛ منظور اعمالی است که براساس انتخاب‌های متمایز اخلاقی و برپایه آگاهی فردی، آزادی، و حتی پیروی از خدا باشند (Reichenbach 2021: 113). وی می‌پرسد آیا این امور با انتقال ما به بدن‌های سایبرنتیک تحت تأثیر قرار نمی‌گیرند؟ اگر ما به برنامه‌های نرم‌افزاری تبدیل شویم، آنگاه آزادی انتخاب و عمل ما چه می‌شود؟ آیا ما می‌توانیم انتخاب دیگری جز ماشینی‌شدن داشته باشیم؟ یا آن‌که هر انتخابی ناشی از برنامه‌ریزی قبلی خواهد بود؟ نظر رایشنباخ آن است که آزادی عمل انسان از بین خواهد رفت، زیرا انسان، به‌عنوان نرم‌افزار، دست‌رسی به تمام مقتضیات برای تکمیل تصمیم خود نخواهد داشت. هرچند

اتصال انسان به سخت‌افزارهایی که در مغز او تعبیه شده است باعث افزایش اطلاعات و تصمیمات وی می‌شود، اما به هر حال انسان متکی به سازندگان و برنامه‌ریزان می‌شود. سازندگان می‌توانند به تمام ورودی‌های اطلاعاتی انسان مسلط شوند و او را در جهت اهداف خود برنامه‌ریزی کنند. رایشناخ چنین آینده‌ای را هرچند مایه غلبه بر ترس‌های حاصل از جهل و نادانی انسان می‌داند، ولی از آن به کابوس و افسوس تعبیر می‌کند، زیرا معتقد است که دیگر «آزادی انتخاب» به نوعی «تلفات» قابل توجه فرایند کامپیوتری شدن انسان خواهد بود (ibid.)؛ یعنی آزادی انتخاب نه امری مبارک و ارزشمند، بلکه نوعی احلال در فرایند برنامه‌ریزی و محاسبات کامپیوتری به حساب می‌آید، زیرا سبب می‌شود از انسان مصنوعی اعمالی غیرقابل پیش‌بینی بروز کند.

۱.۶ نقد و بررسی

طبق حکمت متعالیه یک اشکال مبنایی در این بحث وجود دارد. اراده در انسان متوقف بر مقدماتی است. ابتدا تصور شیء لازم است که این تصور می‌تواند ظنی، خیالی، یا یقینی باشد. پس از درک یک شیء، اگر این شیء ملائم یا منافر با نفس ما باشد، در ما شوقی در جذب یا دفع شکل می‌گیرد. اگر این شوق مؤکد شد، عزم جازم شکل می‌گیرد که همان اراده است؛ و اگر با قدرت هم ضمیمه شد، موجب تحریک اعضا می‌شود (صدرالدین شیرازی ۱۳۰۲ ق: ۱۹۸). چنان‌که مشخص است، فعالیت عصبی و مغزی در این نگاه آخرین مرحله است و اما اراده پیش از آن نیازمند مقدماتی است که در نفس انسان رخ می‌دهد. در هوش مصنوعی قوی، به دلیل عدم انتقال چیزی به نام نفس به ماشین، طبق مبنای حکمت متعالیه نمی‌توان به انتقال اراده انسانی قائل شد، اما سؤال این است که «آیا می‌توان از شبیه‌سازی مدل حکمت متعالیه اراده در هوش مصنوعی قوی سخن گفت؟». فرض چنین مسئله‌ای نیازمند شبیه‌سازی علم و آگاهی، شبیه‌سازی امکان ترجیح آگاهانه یکی از طرفین، شبیه‌سازی درک ملائمت یا منافرت شیء با نفس، شبیه‌سازی شوق و شوق مؤکد، و شبیه‌سازی قدرت است. در میان این شبیه‌سازی‌ها دو مورد شامل آگاهی به خصوص آگاهی از انتخاب خود و نیز درک ملائمت یا منافرت با ذات خود در هوش مصنوعی قوی به نظر می‌رسد ناممکن باشد، زیرا جمع‌آوری اطلاعات و تحلیل آن هرچند قابلیت ممتازی در هوش مصنوعی است، با آگاهی و آگاهی از آگاه‌بودن خود متفاوت است. شبیه‌سازی درک ملائمت یا منافرت با ذات نیز به دلیل عدم انتقال وجود یا نفس انسانی به ماشین صرفاً در همان حد جمع‌آوری و تحلیل داده‌های فردی قابل بیان است.

اما از منظر علوم طبیعی برخی معتقدند هرچند فقدان اثبات علمی اراده آزاد سبب شده است تا فهم دانشمندان از آن با محدودیت روبه‌رو شود، به‌هرحال رواج پذیرش اراده آزاد به‌عنوان عامل مهمی در کنترل عملکرد ذهنی انسان کنونی باید در هوش مصنوعی قوی نیز منعکس شود (Krausová and Hazan 2013: 101). درمقابل شاید سؤالات دیگری هم مطرح شود: مثلاً، آیا اصلاً نیاز است که هوش مصنوعی برای رقابت با عملکرد ذهنی انسان دارای اراده آزاد باشد؟ آیا تلاش برای ساخت اراده آزاد در هوش مصنوعی ضروری است؟ آیا فواید هوش مصنوعی زمانی که جبری عمل کند بیش‌تر نخواهد بود؟ در پاسخ به چنین سؤالاتی طرفداران ایجاد اراده آزاد در هوش مصنوعی دلایلی را بیان می‌کنند: ۱. برنامه‌ریزی اراده آزاد در هوش مصنوعی می‌تواند به شناخت بهتر طبیعت انسان به ما کمک کند؛ ۲. گمان می‌رود اراده آزاد بتواند در ارتقای سطح آگاهی هوش مصنوعی دخیل باشد، مثلاً از طریق کسب آگاهی منحصربه‌فرد یا داشتن آگاهی از چگونگی استفاده از آگاهی (ibid.).

ازسوی دیگر، برای ایجاد اراده آزاد در هوش مصنوعی تلاش‌هایی شده است؛ برای مثال پیش‌نهاد اراده آزاد ضعیف از سوی مک‌کارتی (McCarthy) و اراده آزاد شبه‌انسانی از سوی مانزوتی (Manzotti). اراده آزاد ضعیف شامل (۱) محاسبه اعمال ممکن و نتایج آن و (۲) تصمیم‌گیری درباره عمل مرجح است (McCarthy 2005: 1). اراده آزاد شبه‌انسانی عبارت است از مفهومی از آزادی که متکی به دلایل شخصی و براساس تجربیات فردی باشد (Manzotti 1986: 186). برخی مدعی‌اند اراده آزاد پایه‌های عصب‌شناسانه دارد. زمانی که فرد تصمیمی را می‌گیرد رشته‌های عصبی خاصی فعال می‌شوند (Tse 2013: 19). از آن‌جاکه برخی تصمیمات برپایه تجربیات گذشته گرفته می‌شود، شاید هوش مصنوعی بتواند با کپی‌برداری ساختار مغز و تحلیل خاطرات فرد انتخاب‌های فردی داشته باشد. برخی دیگر مدعی‌اند که می‌توان یکی از مؤلفه‌های اراده آزاد، یعنی غیرقابل پیش‌بینی بودن، را در هوش مصنوعی از طریق نرم‌افزارهای مولد رندوم کوانتمی ایجاد کرد. رندوم‌های کوانتمی غیرقابل محاسبه‌اند و می‌توان گفت دقیقاً با هیچ الگوریتمی قابل تولید نیستند (Calude et al. 2010: 18). اگر این تکنولوژی در هوش مصنوعی به‌کار گرفته شود، می‌توان انتظار تصمیماتی خارج از برنامه‌ریزی منطقی نیز از هوش مصنوعی داشت (Krausová and Hazan 2013: 105). البته متخصصان هوش مصنوعی متذکر می‌شوند، باوجود تلاش‌ها و فرضیات متخصصان، زمانی که ساختار عصبی مغز دقیقاً شبیه‌سازی شود، باز هم درحد احتمال است که هوش مصنوعی اراده آزاد داشته باشد، زیرا تشخیص اراده آزاد متکی به منظر اول‌شخص است (ibid.: 107). ما نمی‌توانیم به‌عنوان انسان بیولوژیکی تصویر اول‌شخصی از یک هوش مصنوعی داشته باشیم، این موضوع درباره قابلیت‌های حیوانات نیز

به دلیل عدم دسترسی به تصویر اول شخص آنان جاری است (همتی مقدم ۱۳۹۱: ۳۷). در نتیجه، از آن جاکه رایشنباخ تداوم هویت انسانی از طریق انتقال الگوهای ذهنی را پذیرفت، باید به تبعات آن هم پای‌بند باشد و درباره راه کارهای فیزیکی شبیه‌سازی اراده آزاد بیش‌تر تأمل کند.

۷. چالش تصمیم اخلاقی

رایشنباخ معتقد است که موجود اخلاقی پیش‌فرض‌های آگاهانه از نقش و موقعیت خود در محیط دارد و از اصولی بهره می‌برد که گاهی با هم در تضادند؛ لذا نیازمند اولویت‌بندی و لحاظ شرایط هستند (Reichenbach 2021: 113). در این رستخیز کامپیوتری برای آن‌که موجودی اخلاقی باشیم، باید فراتر از واکنش‌دهنده به محیط باشیم؛ باید بتوانیم توجیهات اخلاقی برای رفتارمان بیاوریم یا قادر باشیم که به‌طور مستدل بیان کنیم برخی رفتارها اخلاقاً درست هستند و برخی نه (ibid.: 115). به علاوه دانستن متفاوت با واکنش عملی است. حتی یک ربات می‌تواند لبخند بزند، بدون شنیدن یک لطیفه. مسلم نیست که ماشین‌های مادی و نرم‌افزارها بتوانند از واکنش عملی مناسب به سمت تشخیص ذهنی و ارزیابی موقعیت یا خودآگاهی یا منظر اول شخص حرکت کنند و به آن دست یابند (ibid.).

۱.۷ نقد و بررسی

از منظر ملاصدرا معرفت عقلی، چه عقل نظری و چه عقل عملی، عامل رشد، شناخت، و محرک اخلاق است. در بعد رشد و کمال اخلاق انسانی ملاصدرا ابتدا کمال و سعادت حقیقی انسان را رسیدن به کمالات عقلی، معرفت را عامل سعادت، و جهل را عامل شقاوت دانسته است. همین معرفت و حرکت در جهت کمال عقلی سبب رشد فضیلت، شرف، و کمال نفس نیز است. برای اثبات عاملیت معرفت عقلی در رشد فضائل نفسانی مستندات کلام ملاصدرا در دو بخش ارائه می‌شود:

بخش اول شامل سعادت حقیقی دانستن معرفت است. ملاصدرا، در بحث تحقیق معنای علم، علم را امری وجودی می‌داند (صدرالدین شیرازی ۱۳۶۸: ج ۳، ۲۹۷). به علاوه (بنابر اصالت وجود و حرکت جوهری) کمال انسان را خروج از حد محسوس و رسیدن به حد معقول می‌داند که آن‌جا مجل انوار الاهی و ابتهاجات نامتناهی است (صدرالدین شیرازی ۱۳۶۳: ۱۱۷). ادراک حق و مشاهده جمال و جلال او بالاترین کمال و سعادت برای عرفا و حکماست که در سابقه عقل نظری حاصل می‌شود (صدرالدین شیرازی ۱۳۸۱: ۱۶۴-۱۶۵). وی با اشاره به

عوالم سه‌گانه (شامل عالم حس و دنیا، عالم غیب، و عالم قدس) مسافران این سه عالم را شامل مسافر دنیا، که متاعش مال و ثروت و سودش پشیمانی است، مسافر عقبی، که متاعش عبادت و سودش بهشت است، و مسافر عالم قدس، که متاعش معرفت و سودش لقاءالله است، می‌داند. در همین جا او معرفت را اصل هر سعادت می‌خواند و جهل را اصل هر شقاوتی می‌داند (صدرالدین شیرازی ۱۳۶۶: ج ۴، ۴۲۳). آفرینش انسان برای فهم حقایق اشیاست و سعادت اصیل مرتبه‌ای است که انسان در ذات خود عالم عقلی شود؛ در پی آن مطیع و مقرب درگاه الاهی شود؛ و شاهد حسن مطلق، خیر مطلق، و جمال حق مطلق باشد (همان: ج ۲، ۳۳۲).

در بخش دوم رشد عقلی را عامل رشد فضائل نفسانی شمرده است. در کتاب مبدأ و معاد در فصلی با عنوان «سعادت حقیقی برای نفس»، از جهت عقل نظری و عملی، سعادت عقلی برای انسان را از دو جهت عقل نظری و عملی می‌داند و درعین حال شرف، فضیلت، زینت، کمال، سرور، و غبطه برای نفس را برحسب عقل نظری می‌داند که هویت آن را تشکیل می‌دهد و وظیفه عقل عملی را توجه به بعد محسوس و تعلق نفس به بدن می‌داند که جزء ذاتی نفس نیست، بلکه هدایت نفسیات را برعهده دارد، مانند دوری از نقایص و زشتی‌ها و خلاص از رذائل (صدرالدین شیرازی ۱۳۵۴: ۳۶۶). مشابه همین بیان را در *سفر* دارد و معتقد است که نفس طهارت کمال حقیقی نیست، بلکه هدف طهارت قلبی حصول نور معرفت است. غایت حکمت عملی، که به واسطه تهذیب نفس است، شرح صدر و غایت حکمت نظری نور است (صدرالدین شیرازی ۱۳۶۸: ج ۹، ۱۴۰). در بعد شناختی نیز معتقد است که عقل نظری مربوط به تصور و تصدیق و عقل عملی برای استعمال این فکر و نظر در افعال و صنایع بشری از جهت خیر یا ظن خیر است (صدرالدین شیرازی ۱۳۶۰: ۱۹۹-۲۰۰). عقل نظری عاملی برای یافتن صدق و کذب، واجب و ممکن، و ممتنع و عقل عملی برای یافتن خیر و شر و حسن و قبح است (همان: ۲۰۱). در بعد محرکی قائل است که اراده، که موجب صدور فعل یا ترک آن است، ناشی از عقل عملی است (صدرالدین شیرازی ۱۳۶۸: ج ۶، ۳۵۴). عقل عملی قوه عاملی است که محرک انسان برای کاربرد قوای تحریکی به هم‌راه فکر و معرفت برای افعال جزئی است (صدرالدین شیرازی ۱۴۲۲ق: ۲۴۱). در نتیجه، از نظر ملاصدرا رابطه معرفت عقلی و اخلاق در هر سه حیطه رشد، شناخت، و محرک اخلاق حائز اهمیت است.

در بحث حاضر نیز تأکید رایشنباخ بر آگاهی از موقعیت محیطی و داشتن توجیه اخلاقی، به‌عنوان پشتوانه‌های فعل اخلاقی، صحیح به‌نظر می‌رسد. همچنین، برخی معتقدند مشکل اصلی نظریات یگانه‌انگاری ذهن و بدن مانند این‌همانی، حذف‌گرایی، رفتارگرایی تحلیلی، و

کارکردگرایی موضوع آگاهی است: «مشکل اصلی این نظریه‌ها، که یا به‌صراحت مادی‌انگارند یا از مادی‌انگاری الهام گرفته‌اند، این است که آگاهی را جدی نمی‌گیرند، حال یا به‌واسطه این‌که جایی برای آن باز نمی‌کنند، یا به‌سبب آن‌که صراحتاً وجود آن را انکار می‌کنند» (مسلین ۱۳۹۱: ۲۳۸).

با این حال، اشکالاتی در نفی توانایی تصمیم اخلاقی در هوش مصنوعی قوی به‌نظر می‌رسد: اشکال اول آن است که وی در بحث معاد به خلق مجدد و حفظ هویت فردی از طریق انتقال خاطرات، آگاهی‌ها، و منظر اول‌شخص به فرد دوباره خلق شده قائل است. پس اگر به هر فرضی منظر اول‌شخص و آگاهی‌ها قابلیت انتقال از فرد معدوم به فرد بازآفرینی شده داشته باشند، چرا قائل نشویم این انتقال در هوش مصنوعی نیز قابل کشف و استفاده است. اگر بنابر فرض خود رایشنباخ تداوم هویت انسانی با انتقال و تداوم الگوهای ذهنی میسر است، چرا وی انتقال قابلیت تصمیم اخلاقی را نفی می‌کند؟ آیا تصمیم اخلاقی جزو هویت انسانی نیست؟ به‌هرحال یا رایشنباخ باید در مبنای خود تجدیدنظر کند یا تبعات احتمالی آن را نیز بپذیرد.

اشکال دوم به‌نوعی توجیه نظریه‌پردازان هوش مصنوعی قوی است. آنان معتقدند این قابلیت ذهنی درباره آگاهی و دانش می‌تواند در طی یک تمثیل عملی پدید آید. به‌عبارت‌دیگر، می‌توان گفت دانش همان به‌کارگیری عملی اشیای ملموس است. برای مثال کودکان فهمی از «یک به‌علاوه یک مساوی دو» دارند؛ هرچند مباحثات انتزاعی و ذهنی درباره آن را نمی‌دانند، زمانی که یک شیء را کنار شیء دیگر می‌گذارند، آن دو را به‌هم‌پیوسته می‌دانند. حال اگر ربات‌های ماشینی بتوانند به‌طور کامل برنامه‌ریزی شوند، تاحدی که به سرنخ‌های محیطی متعدد در طی زمان واکنش نشان دهند، آیا نمی‌توان گفت آن‌ها درکی از موقعیت‌ها دارند؟ برای دانستن آیا حتماً باید آن‌ها از موقعیتی که با آن مواجه می‌شوند نقشه ذهنی داشته باشند؟ (Wallach and Allen 2009: 68-69) یا شاید مهم‌ترین خصوصیات آگاهی عملکردی باشند؛ یعنی حتی اگر کامپیوترها دقیقاً مانند انسان‌ها آگاه نباشند، شاید بتوان آن‌ها را طوری طراحی کرد تا مشابه چنین ظرفیتی عمل کنند، مثلاً از خود احساسات نشان دهند، بدون آن‌که حقیقتاً دارای احساسات باشند (ibid.). هرچند رایشنباخ رسیدن از عمل به آگاهی را نامسلم دانسته، ولی آن را نامحتمل نیز ندانسته است و لذا اگر در هوش مصنوعی قوی مبنای هویت انسانی تغییر یابد، چرا نتوان به بازتعریف سایر مؤلفه‌های انسانی و نیز نحوه رسیدن به آن‌ها پرداخت. البته فروگاهش دانش به بعد عملکردی آن یا گذار از عملکرد به کسب دانش در دستگاه فلسفی ملاصدرا کاملاً متنفی است، زیرا چنان‌که گذشت، او عقل عملی را به‌نوعی خادم عقل نظری در کنترل نفسانیات دانست و وزن اصلی را به عقل نظری داد. حتی در بعد عملکردی نیز اراده

نیازمند مقدماتی است که از جمله آن‌ها تصور و تصدیق است؛ لذا پشتوانه اصلی عملکرد عقل در اخلاق همان عقل نظری است؛ بنابراین در دستگاه صدرایی باید از گذار از معرفت به عملکرد سخن گفت و نه برعکس.

۸. نتیجه‌گیری

تحقق آرمان‌شهر هوش مصنوعی و انتقال الگوهای مغزی انسان از بدن بیولوژیکی به کامپیوترها و یا بدن‌های مصنوعی، درعین شگفت‌انگیزی، با چالش‌های نظری مواجه است. بروس رایشنباخ سعی در بیان این چالش‌ها از منظر الاهیاتی و فلسفی کرد و عمده توجه وی حول سرنوشت هویت انسانی، خدا و معنویت، اراده، و تصمیم اخلاقی قرار گرفت.

چالش هویت انسانی در زمان انتقال الگوهای ذهنی از فرد انسانی به ماشین مطرح است. در دستگاه صدرایی هویت انسانی در صورت معدوم‌شدن قابل‌اعاده نیست؛ به‌علاوه، ملاک این‌همانی هویت تشخیص وجودی است که قابل انتقال به ماشین نیست. به اشکالات دیگری هم چون لزوم پای‌بندی به تبعات پذیرش تداوم هویت از طریق تداوم الگوهای ذهنی اشاره شد. اگر هویت انسانی در اولین کپی‌برداری الگوهای مغزی محفوظ بماند، چرا نتوان به حفظ آن در کپی‌برداری‌های دیگر هم امیدوار بود؟ لذا به نظر می‌رسد اگر رایشنباخ حفظ الگوهای ذهنی را برای حفظ هویت انسانی کافی بداند، آن‌گاه باید تبعات آن را نیز بپذیرد. نقد دیگر بی‌توجهی به از بین رفتن الگوهای اولیه در فرایند گسترش الگوهاست، زیرا متخصصان هوش مصنوعی قوی از اصطلاحاتی مانند ترکیب و آمیختگی الگوها استفاده می‌کنند و نه جای‌گزینی آن‌ها با الگوهای اولیه. هم‌چنین، متخصصان به انتخابی‌بودن پدیده اشتراک‌گذاری الگوها و عدم اجبار در این امر قائل هستند. به‌علاوه به باقی ماندن آگاهی انسان از ورود خاطرات و الگوهای جدید معتقدند. تمام این موارد (بقای الگوهای اولیه، بقای حق انتخاب، بقای آگاهی)، برخلاف نظر رایشنباخ، نشان‌دهنده حفظ هویت فردی انسان در فرایند گسترش الگوهای ذهنی است. نکته مهم‌تر آن است که اگر منظر اول‌شخصی باقی نمانده است، چگونه فرد متوجه وجود تصاویر اول‌شخص متعدد در ذهن خود خواهد شد؟ آیا چنین حکمی خود نشان‌دهنده بقای منظر اول‌شخص نیست؟

در چالش معنویت و تجربه دینی رایشنباخ معتقد است که منظور از تجربه دینی در هوش مصنوعی قوی همان آگاهی است، ولی بررسی نظرهای کورزوئل (نظریه‌پرداز هوش مصنوعی قوی) نشان می‌دهد که منظور از معنویت و تجربه دینی در هوش مصنوعی قوی احساسات و

تجربیات حاصل از اعمال دینی است و نه صرف آگاهی. لذا به نظر می‌رسد عنصر احساسات از تجربه دینی و معنویت هوش مصنوعی حذف نخواهد شد.

در چالش پرستش خدای حقیقی یا مجازی، رایشنباخ حداکثر توان هوش مصنوعی قوی را داشتن اعتقاد به خدایی مجازی می‌داند. در ارزیابی این سخن می‌توان براساس نظرهای متخصصان هوش مصنوعی گفت که اگر در زمان پدیدآمدن تجربه دینی نسبت به خدای حقیقی، مثلاً در هنگام مناجات یا اعمال دینی، بخش خاصی از اعصاب مغزی دچار تحریک می‌شوند، چرا نتوان در انتقال الگوهای مغز انسان به هوش مصنوعی باز هم با تحریک آن بخش ذهنی تجربه‌ای از همان خدای حقیقی را ایجاد کرد؟ به علاوه اگر به بقای تجربه دینی در هوش مصنوعی قائل شویم، چرا باید اساساً متعلق این تجربه تغییر کند و از حضوری به مجازی مبدل شود؟

در بحث چالش آزادی انتخاب، رایشنباخ مدعی است که آزادی انسان به خطر می‌افتد و هر انتخابی ناشی از برنامه ریزی پیشین خواهد بود. از منظر صدرایی شبیه‌سازی اراده با چالش مهمی مواجه است، زیرا تحقق اراده نیازمند مقدماتی است که وابستگی وجودی به نفس دارند، اما در انتقال ذهنی نفس منتقل نمی‌شود. اما آیا اساساً هوش مصنوعی قوی برای رقابت با عملکرد ذهنی انسان، و نه تمام ابعاد انسانیت، نیازی به اراده آزاد نیز دارد؟ آیا فواید هوش مصنوعی زمانی که جبری عمل کند بیش‌تر نخواهد بود؟ به علاوه برخی تلاش‌ها برای ایجاد اراده آزاد ضعیف یا اراده آزاد شبه‌انسانی نیز صورت گرفته است. برای مثال، شاید بتوان با اتکا به پردازش دلایل شخصی یا مرور تجربیات فردی گونه‌ای از اراده آزاد را شبیه‌سازی کرد. هم‌چنین، شاید بتوان برخی مؤلفه‌های اراده آزاد مانند غیرقابل پیش‌بینی بودن را از طریق نرم‌افزارهای مولد رندوم کوانتومی ایجاد کرد. به هر حال، اگر قرار است هویت انسانی تنها از طریق کپی‌سازی الگوهای مغزی تداوم یابد (چنان‌که رایشنباخ نیز آن را پذیرفت)، باید تبعات احتمالی آن را نیز پذیرفت و مسلماً مواردی هم‌چون اراده آزاد، که به ابعاد وسیع‌تری از انسان اشاره دارد، بازتعریف خواهند شد.

در چالش تصمیم اخلاقی رایشنباخ معتقد است که موجود اخلاقی دارای قابلیت اولویت‌بندی و لحاظ شرایط و نیز هدف‌مندی است، اما هوش مصنوعی قوی چنین قابلیت‌هایی نخواهد داشت. در نقد این نظر به تبعات مبنای رایشنباخ در تداوم هویت انسانی با تداوم الگوهای ذهنی و نیز احتمال گذار از عمل‌گرایی به درک و آگاهی مشابه انسانی اشاره شد. اما در دستگاه صدرایی بیان شد که وزنه اصلی از آن عقل نظری است و معرفت نظری

قابل فروکاهش به بعد عملی نیست و نیز گذار باید از معرفت نظری به عملی برای تحقق فعل صورت گیرد و نه برعکس.

درنهایت، نقطه اتکای اکثر انتقادات ذکرشده به نظرهای رایشنباخ آن است که اگر بپذیریم هویت انسانی از طریق کپی برداری الگوهای ذهنی هنوز پابرجا خواهد ماند (چنان که رایشنباخ پذیرفته است)، آن گاه باید به تبعات این تصمیم نیز پای بند ماند. فروکاهش هویت انسانی به الگوهای ذهنی موجب بازتعریف مواردی هم چون معنویت، اراده، و تصمیم اخلاقی خواهد شد.

کتاب نامه

پترسون، مایکل و دیگران (۱۳۸۹)، *عقل و اعتقاد دینی*، ترجمه احمد نراقی و ابراهیم سلطانی، تهران: طرح نو.

تالیافرو، چالز (۱۳۸۲). *فلسفه دین در قرن بیستم*، ترجمه انشاءالله رحمتی، تهران: دفتر پژوهش و نشر سهروردی.

جواد آملی، عبدالله (۱۳۸۲)، *رحیق مختوم*، ج ۱، بخش ۵، قم: مرکز نشر اسرا. دریفوس، هیوبرت و دیگران (۱۳۹۹)، *ذهن برتر از ماشین: قدرت شهود و خبرگی انسانی در عصر رایانه*، ترجمه یاسر خوشنویس، قم: پژوهشگاه علوم و فرهنگ اسلامی.

سرل، جان ار (۱۳۸۲)، *ذهن، مغز، و علم*، ترجمه امیر دیوانی، قم: بوستان کتاب. صدرالدین شیرازی، محمد بن ابراهیم (۱۳۰۲ق)، *مجموعه رسائل التسعه*، قم: مکتبه المصطفوی. صدرالدین شیرازی، محمد بن ابراهیم (۱۳۵۴)، *المبدأ و المعاد*، تهران: انجمن حکمت و فلسفه ایران. صدرالدین شیرازی، محمد بن ابراهیم (۱۳۶۰)، *شواهد الربوبیه فی مناهج السلوکیه*، تهران: مرکز نشر دانشگاهی.

صدرالدین شیرازی، محمد بن ابراهیم (۱۳۶۳)، *مفاتیح الغیب*، تهران: وزارت فرهنگ و آموزش عالی، انجمن اسلامی حکمت و فلسفه ایران، و مؤسسه مطالعات و تحقیقات فرهنگی.

صدرالدین شیرازی، محمد بن ابراهیم (۱۳۶۶)، *تفسیر القرآن الکریم*، قم: بیدار. صدرالدین شیرازی، محمد بن ابراهیم (۱۳۶۸)، *الحکمة المتعالیة فی الأسفار العقلیة الأربعة*، قم: مکتبه المصطفوی.

صدرالدین شیرازی، محمد بن ابراهیم (۱۳۸۱)، *کسر اصنام جاهلیه*، تهران: بنیاد حکمت اسلامی صدرا. صدرالدین شیرازی، محمد بن ابراهیم (۱۴۲۲ق)، *شرح الهدایة الأثریة*، بیروت: مؤسسه التاريخ العربی.

کوپر، جان دابلیو (۱۳۹۲)، «این همانی اشخاص اعاده شده: نقص مهلک انسان شناسی یگانه انگارانه»، در: *مسیحیت و مسئله ذهن و بدن*، قم: پژوهشگاه علوم و فرهنگ اسلامی.

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

همتی مقدم، احمدرضا (۱۳۹۱)، *نظریه‌های مادی/انگارانۀ ذهن*، قم: پژوهشگاه علوم و فرهنگ اسلامی.

- Calude, C. S. et al. (2010), "Experimental Evidence of Quantum Randomness Incomputability", *Physical Review A*, vol. 82, no. 2.
- Krausová, A. and H. Hazan (2013), "Creating Free Will in Artificial Intelligence", in: *Beyond AI: Artificial Golem Intelligence*, J. Romportl et al. (eds.), Pilsen, Czech Republic: University of West Bohemia.
- Kurzweil, R. (1999), *The Age of Spiritual Machines*, New York: Viking.
- Kurzweil, R. (2002), "The Evolution of Mind in the Twenty-First Century", in: *Are We Spiritual Machines?*, Jay W. Edwards (ed.), Seattle, WA: Discovery Institute.
- Manzotti, R. (2011), "Machine Free Will: Is Free Will a Necessary Ingredient of Machine Consciousness?", in: *From Brains to Systems*, C. Hernandez, R. Sanz, and J. Gomez-Ramirez (eds.) Dordrecht, Springer, Available at: <https://doi.org/10.1007/978-1-4614-0164-3_15>.
- McCarthy, J. (2005), "Simple Deterministic Free Will", Invited Talk on The 7th International Symposium on Logical Formalizations of Commonsense Reasoning, Available at: <<http://www-formal.stanford.edu/jmc/freewill2/>>.
- Moravec, H. (1988), *Mind Children: The Future of Robot and Human Intelligence*, Cambridge, Massachusetts: Harvard University Press.
- Reichenbach, B. R. (1978), "Monism and the Possibility of Life after Death", *Religious Studies*, vol. 14, no. 1, 27–34.
- Reichenbach, B. R. (1983), *Is Man the Phoenix?: A Study of Immortality*, Washington, D. C.: University Press of America.
- Reichenbach, B. R. (2021), "Christianity, Science, and Three Phases of Being Human", *Zygon®*, vol. 56, no. 1.
- Tse, P. (2013), *The Neural Basis of Free Will: Criterial Causation*, Cambridge, Massachusetts: The MIT Press.
- Wallach, W. and C. Allen (2009), *Moral Machines*, C. Allen (ed.), Oxford: Oxford University Press.