Essays in Philosophy and Kalam

HomePage: https://jphilosophy.um.ac.ir/ Vol. 52, No. 1: Issue 104, Spring & Summer 2020, p.161-183

Online ISSN 2538-4171 **Print ISSN** 2008-9112

Receive Date: 03-05-2021 **Accept Date:** 18-06-2021

DOI: 10.22067/epk.2021.70220.1032 **Article type:** Original Research

Analysis of the Possibility of Heideggerian Artificial Intelligence in the Views of Hubert Dreyfus

Mohammad Mehdi Moqaddas PhD student in Contemporary Philosophy, Imam Khomeini International University, Qazvin

Email: mehmoghadas@gmail.com

Abstract

With his phenomenological method taken from the teachings of Heidegger and Merleau-Ponty, Hubert Dreyfus critiques the two paradigms of artificial intelligence, namely symbolism and connectionism. He then approves a trend from the third paradigm of AI, namely artificial life, with special conditions. This trend is known as "Heideggerian AI". After a brief review of Dreyfus's critiques of the first two paradigms of artificial intelligence, this article examines his final opinion on the possibility of accepting Heideggerian AI. According to the author, the possibility of Heideggerian artificial intelligence based on Dreyfus' theories can be demonstrated in at least three ways. These three approaches are presented as follows: The first approach is presented using the Daniel Susser's approach and is based on interactionism. The second solution is expressed by citing John Searle and distinguishing between strong and weak artificial intelligence; the third explanation demonstrates how this is possible based on the doctrine of co-instantiation, inter-consciousness, intelligence, and intelligent behavior.

Keywords: Artificial intelligence, Hubert Dreyfus, Heideggerian artificial intelligence, philosophy of mind, phenomenology, consciousness

جستارهایی در
فلسفه و کلام

سال ۵۲- شماره ۱- شماره پیاپی ۱۰۴- بهار و تابستان ۱۳۹۹- ص ۱۸۳-۱۶۱ HomePage: https://jphilosophy.um.ac.ir

شاپا چاپی ۹۱۱۲-۲۰۰۸ شاپا الکترونیکی ۴۱۷۱-۲۵۳۸

تاریخ دریافت: ۱۴۰۰/۰۲/۱۳ تاریخ پذیرش: ۱۴۰۰/۰۳/۲۸

نوع مقاله: پژوهشی **DOI:** 10.22067/epk.2021.70220.1032

تحلیل و بررسی امکان هوش مصنوعی هیدگری در آرای هیوبرت دریفوس

محمدمهدی مقدس
دانشجوي دكتري فلسفه معاصر دانشگاه بين المللي امام خميني(ره) قزوين
Email: mehmoghadas@gmail.com

چکیده

هیوبرت دریفوس با روش پدیدار شنا سانهٔ خود که برگرفته از آموزه‌های هیدگر و مرلوپونتی ا ست، دو پارادایم هوش م صنوعی، یعنی نمادگرایی و پیوندگرایی را به نقد می‌کشد. وی سپس جریانی را از پارادایم سوم هوش م صنوعی، یعنی حیات م صنوعی با قید شروطی ویژه می‌پذیرد. این جریان به هوش مصنوعی هیدگری مشهور است. در این مقاله با بیان اجمالی نقدهای دریفوس بر دو پارادایم نخ ست هوش م صنوعی، رأی نهایی وی در خ صوص امکان پذیرش هوش م صنوعی هیدگری برر سی می شود. به‌نظر نو یسندهٔ این مقاله، دست‌کم به سه طریق می‌توان امکان هوش مصنوعی هیدگری بر مبنای نظریات دریفوس را نشان داد. راهکار نخست با استفاده از رهیافت دنیل سوسر و بر اساس تعاملگرایی ارائه می‌شود؛ راهکار دوم با استناد به دیدگاه جان سرل و تفکیک هوش م صنوعی قوی و ضعیف بیان می شود؛ راهکار سوم باتوجه‌به آموزهٔ هم‌تحققی، میان‌آگاهی، هوشمندی و رفتار هوشمندانه نشان می‌دهد که این امکان به چه ترتیب میسّر است.

واژگان کلیدی: هوش مصنوعی، هیوبرت دریفوس، هوش مصنوعی هیدگری، فلسفهٔ ذهن، پدیدارشناسی، آگاهی.

مقدمه

هیوبرت دریفوس،[۱] فیلسوف و پدیدارشناس آمریکایی معاصر است که از شارحان هیدگر[۲] و مرلوپونتی[۳] نیز به شمار می‌رود. با ظهور پدیدهٔ هوش مصنوعی[۴] و اقبال پژوهشگران به آن، بررسی فلسفیِ این پدیده نیز مدنظر فیلسوفان قرار گرفت. هوش مصنوعی پدیده‌ای است که بیش از هرچیز با مسئلهٔ آگاهی مرتبط است. تلقی ما از سازوکارِ آگاهی کمک می‌کند تا بتوانیم امکان تعمیم این مفهوم را بررسی کنیم. به‌طور خلاصه چهار نوع تعریف از هوش مصنوعی می‌توان داشت: نخست پژوهش در سامانه‌هایی که شبیه انسان می‌اندیشند؛ دوم پژوهش در سامانه‌هایی که به‌گونه‌ای عقلانی می‌اندیشند؛ سوم پژوهش در سامانه‌هایی که شبیه انسان رفتار می‌کنند؛ چهارم پژوهش در سامانه‌هایی که به‌گونه‌ای عقلانی رفتار می‌کنند (Russell and Norvi, 2). مسئله‌ای که در این میان نقش تعیین‌کننده‌ای دارد، مسئلهٔ ماهیت ذهن و به‌تبع آن ارتباط ذهن و بدن است، هرچند که ارتباط ذهن و بدن به نحو مستقیم با ماهیت ذهن مرتبط است. فرض کنید که ذهن چیزی بیش از برنامهٔ کامپیوتری نیست و مغز چیزی شبیه به سخت‌افزار است، در اینجا ماهیت ذهن را چیزی مانند برنامهٔ کامپیوتری قلمداد کرده‌ایم. نتیجهٔ این تلقی از ذهن این است که اگر ذهن برنامه‌ای است که روی سخت‌افزار مغز اجرا می شود، چرا نتوان موجوداتی ساخت که قادر به اجرای برنامه‌های مختلف روی ساختار خود باشند؟ این ایده‌ای بود که نمادگرایان به آن علاقه نشان دادند و نخستین‌بار آلن تورینگ چنین پیشنهادی را داد که آیا ماشین‌ها توانایی تفکر دارند؟ (Turing, 1) بنابراین اهمیت تعیین ماهیت ذهن در اینجا کاملاً روشن است. ازآنجاکه مسائلی مانند آگاهی و ماهیت ذهن ازجمله مسائل فلسفی هستند؛ عرصهٔ هوش مصنوعی خیلی زود وارد بحث‌های فلسفی و نوشته‌های فیلسوفان شد. از این جمله هیوبرت دریفوس بود که از آغازِ ظهور این پدیده، همواره نگاهی انتقادی به رویکردهای پژوهشگران هوش مصنوعی داشت. بحث وی ناظر به امکان هوشمندی و امکان آگاهی کامپیوترها و دیگر وسایل هوشمند است. هوش مصنوعی در طول دورهٔ هفتادسالهٔ خود دست‌کم سه دوره یا پارادایم را گذرانده است (clarck، ۳). این سه پارادایم عبارت‌اند از: ۱. پارادایم نمادگرایی[۵] (نشانه‌گرایی)؛ ۲. پارادایم پیوندگرایی؛[۶] ۳. پارادایم حیات مصنوعی.[۷] دریفوس دو پارادایم نخست را به نقد می‌کشد اما با رویکردی محتاطانه جریانی از پارادایم سوم را که به هوش مصنوعی هیدگری[۸] موسوم

---

۱. Hubert Dreyfus (1929–2017): فیلسوف و پدیدارشناس معاصر، استاد فلسفه دانشگاه برکلی کالیفرنیا. زمینه‌های مطالعاتی وی عبارت‌اند از: پدیدارشناسی، اگزیستانسیالیسم، فلسفهٔ روان شناسی و همچنین هوش مصنوعی. دریفوس از شناخته شده‌ترین شارحان هیدگر به شمار می‌رود تا جایی که وی را دریگر (مرکب از نام‌های دریفوس و هیدگر) نامیده‌اند. وی همچنین به‌سبب نقدهایش بر هوش مصنوعی شناخته شده است.

2. Heidegger, Martin.

3. Merleau-Ponty, Maurice.

4. Artificial Intelligence: AI.

5. symbolic AI.

6. connectionist AI.

7. Artificial life.

8. Heideggerian AI.

است با قید شرایطی ویژه می‌پذیرد.

پذیرش هوش مصنوعی توسط دریفوس بدین معنا است که وی معتقد است که نقدهایش بر دو پارادایم هوش مصنوعی کلاسیک در این جریان برطرف شده است؛ لذا بایستی در نظر وی هوش مصنوعی هیدگری به‌نحوی سامان‌دهی شود که در فهم ما از جهان مشترک باشد. همچنین باید در جهان به‌گونه‌ای عمل کند که قادر به فعالیت ماهرانه در زندگی روزمره باشد. این مهم لزوماً به‌معنای بالاترین سطح درگیری با جهان همانند انسان نیست، اما بایستی در نظر وی سطح پذیرفته‌شده‌ای از این تفاهم و اشتراک فهم از جهان را ارائه کند.

در این مقاله تلاش بر این است که به‌نحو مختصر نقدهای دریفوس بر دو پارادایم نخست را بیان و سپس سازگاری رأیِ واپسین وی مبنی‌بر پذیرش هوش مصنوعی هیدگری را با مواضع پیشین او در نقد دو پارادایم نخست بررسی کنیم و با ارائهٔ سه راهکار نشان دهیم که پذیرش هوش مصنوعی هیدگری با مواضع پیشین وی در نقدهایش بر هوش مصنوعیِ نمادی و پیوندگرا، سازگار و امکان‌پذیر است.

۱. پارادایم نمادگرایی (نشانه‌گرایی)

۱. ۱. ویژگی‌های نمادگرایی

ایدهٔ محوری در این پارادایم این است که کار ذهن و رایانه‌های دیجیتال، دست‌کاری داده‌ها و داده‌پردازی نشانه‌ها و نمادهای صوری است. به‌دیگر سخن، ذهن نوعی رایانهٔ دیجیتال است. در واقع محل نزاع این است که همان طور که برنامهٔ کامپیوتری با دست‌کاری نمادهای منطقی و توابع ریاضیاتی، آن‌ها را پردازش می‌کند، ذهن انسان نیز داده‌پردازی می‌کند و این داده‌پردازی، ویژگی اصلی ذهن و رایانهٔ دیجیتال است. این نوع تلقی از ذهن و هوشمندی، ناظر به آن دسته از تعاریفی است که هوشمندی را با معیار شبیه به انسان‌بودن تعریف می‌کنند.[۱] بنابراین تأکید این پارادایم بر شبیه به انسان فکرکردن است. این پارادایم نخستین‌بار با پژوهش‌های آلن تورینگ[۲] آغاز شد.

بنابراین نمادگرایان بر این باورند که ذهن و کامپیوترهای دیجیتالی هر دو سیستم‌هایی هستند که کارشان داده پردازیِ نمادها و نشانه‌های صوری است؛ نمادهایی که بازنمودهای جهان خارج‌اند، در واقع در این تلقی، کار ذهن بسیار شبیه کامپیوتر است (Dreyfus, *What computers still can't do*,302).

نمادگرایان معتقدند که اطلاعات و بیت‌هایی که رایانه‌ها پردازش می‌کنند، می‌توانند نشانهٔ هرچیزی حتی اوصاف جهان خارج باشند؛ یعنی «در مدل کامپیوتری ذهن، ذهن نیز همچون کامپیوتر واجد برنامه با مجموعه‌ای از

---

۱. تعاریفی از این دست مانند تلاش جدید و هیجان‌انگیز برای ساخت کامپیوترهایی متفکر، ماشین‌های متفکر و با حس کامل (Haugeland, 2) و یا آنگونه که ریچارد بلمن می‌گوید خودکارسازی فعالیت‌های مرتبط با تفکر انسان، فعالیت‌هایی مثل تصمیم‌گیری، حل مسئله و یادگیری (بلمن،۱۹۷۸) بر معیار مانند انسان‌بودن متمرکز هستند.

۲. Turing, A. M. 'Computing machinery and intelligence'.

قواعد است و تفکر، داده‌پردازی تصورات یا دیگر علائم به شمار می‌رود. آلن نوویل[1] و هربرت سایمون[2] از پیشگامان این رهیافت به این نتیجه رسیدند که بیت‌هایی که هر کامپیوتر دیجیتالی جهت داده‌پردازی به کار می‌برد، نه‌تنها نشانگر اعداد هستند، بلکه می‌توانند نشانگر هرچیز دیگر و ازجمله اوصاف جهان خارج باشند. علاوه بر این، برنامه ای که در این کامپیوتر به کار می‌رود، نمایانگر روابط میان بیت‌هاست... نوویل و سایمون با نگارش مقالهٔ معروف خود این فرضیه را بنیان گذاردند که مغز انسان و کامپیوتر دیجیتال در همان حال که ساختار و مکانیزم‌شان کاملاً با هم متفاوت است، کارکردشان را می‌توان مشترک دانست. بر اساس نظر آن‌ها، مغز دستگاهی است که می‌توان آن را سیستم علائم فیزیکی[3] نام گذاشت. کار مغز، داده‌پردازی این علائم است» (طهماسبی، ۲۸).

بایستی این نکته را اضافه کرد که نقدهایی که دریفوس بر تلقی نمادگرایان بیان می‌کند، چنانچه با مطالعات جدید در دانش کامپیوتر و داده‌پردازی مقایسه شود، بسا نادرست و ناکافی بنماید. این مطلب از آن جهت رخ می‌دهد که پیشرفت‌های تکنیکی و نظریه‌ای در هوش مصنوعی با سرعتی زیاد در حال رخ‌دادن هستند؛ اما بایستی این مطلب را به یاد داشته باشیم که تلقیِ نمادگرایان از ذهن و ارتباط آن با آگاهی به همین نحو بوده است که دریفوس بیان و آن‌ها را نقد می‌کند. لذا ممکن است در برخی موارد به نظر برسد که در حال حاضر در هوش مصنوعی رهیافت‌ها و پدیده‌های نوینی باشد که بعضاً نقدهای دریفوس نیز به آن‌ها وارد نباشد، اما این نکته ضروری است که بدانیم دریفوس مطلق هوش مصنوعی را نقد نمی‌کند، بلکه تلقیِ خاصی را بیان و نقد می‌کند.

چنانچه مغز داده‌پردازی علائم فیزیکی را انجام می‌دهد و این داده‌ها همان ذهن هستند، پس نمادگرایان با چنین تلقی از ذهن بدین نتیجه رسیدند که تعمیم موجودات هوشمند به رایانه‌های دارای قابلیت برنامه‌ریزی، امکان‌پذیر است. اما دریفوس معتقد است که چنین تصوری دست‌کم سه پیش‌فرض دارد که هیچ‌کدام توجیه نشده‌اند. این فرض‌ها عبارت‌اند از: ۱. فرض روان‌شناسانه؛ ۲. فرض معرفت‌شناسانه؛ ۳. فرض هستی‌شناسانه. این فرض‌ها به هم مرتبط‌اند؛ بدین صورت که فرض معرفت‌شناسانه پیش‌فرض فرضِ روان‌شناسانه و فرض هستی‌شناسانه، پیش‌فرض فرضِ معرفت‌شناسانه است. اما ببینیم این فرض‌ها چیستند؟ «فرض روان‌شناسانه[4] می‌گوید که ذهن می‌تواند همچون دستگاهی نگریسته شود که با بیت‌هایی از اطلاعات مبتنی بر یک سری قواعد صوری عمل می‌کند. فرض معرفت‌شناسانه[5] که مضمون آن این است که همهٔ دانش می‌تواند صوری سازی شود و فرض هستی‌شناسانه[6] که می‌گوید هر آنچه هست، مجموعه‌ای از واقعیت‌هایی است که هریک منطقاً مستقل از یکدیگر هستند»

1. Allen Newell (1927-92).
2. Herbert A. Simon (1916-2001).
3. Physical signs system.
4. Psychological assume.
5. Epistemological assume.
6. Ontological assume.

(Dreyfus, *What Computers Can't Do, 68*)

۱. ۲. نقد فرض روان‌شناسانه

دریفوس بر این باور است که این فرض الگویی است که در آن ذهن کارکردی همچون رایانهٔ دیجیتال دارد؛ ازآن‌رو که هم ذهن و هم رایانه اطلاعات را پردازش می‌کنند، اما وی معتقد است که معنای پرداز شگری در انسان و رایانه متفاوت است و نمی‌توان این دو نوع را مانند هم دانست. در آثار دریفوس دو نقد بر فرض روان شناسانه دیده می‌شود: ۱. پردازشگری رایانه با امر معنادار سروکار ندارد (Ibid, ۷۷)؛ ۲. پردازشگری ذهن با برنامهٔ رایانه‌ای سروکار ندارد (همان، ۷۸).

دریفوس با استناد به تئوری کلود شانون[1] در سایبرنتیک[2] اشاره می‌کند که به نظر وی این تئوری به‌روشنی معنا را در معنای معمولی‌اش نادیده می‌گیرد. نخست تعریف پرداز شگری اطلاعات در رایانه به‌مثابهٔ تابعی ریا ضیاتی، هیچ نیازی به فهم و ادراک معنادار انسانی ندارد، اما هر تئوری در فلسفهٔ ذهن یا روان‌شناسی فلسفی به‌هیچ‌وجه نمی‌تواند فهم و ادراک معنادار انسانی را نادیده بگیرد. متغیرهای ریاضیاتی به‌کاررفته در عملیات محاسباتی به‌خودی‌خود هیچ محتوای درونی و معناداری ندارند.[3] آنچه اسباب خلط پردازشگری معنادار انسانی و پردازشگری بی‌معنای رایانه‌ای را فراهم می‌آورد، نادیده‌گرفتن نقش برنامه‌نویس است (در پارادایم پیوندگرایی کوشیده شد تا شبکه به‌نحوی طراحی شود تا مطابق الگوهای عصبی مغز، خود بتواند در مقابل دریافت ورودی‌های خاص، خروجی را حدس بزند)؛ این دقیقاً نقش برنامه‌نویس است که گذر از جملاتی را که معنادار هستند (شامل اطلاعات در معنای معمولی) به رشته‌هایی از بیت‌های گسستهٔ بی‌معنا که رایانه‌ها با آن‌ها کار می‌کنند، ممکن می‌سازد (همان). دوم اینکه ذهن با برنامهٔ رایانه‌ای سروکار ندارد؛ یعنی آن‌گونه که نمادگرایان مدعی هستند، سازوکار ذهن مانند سازوکار برنامهٔ رایانه‌ای نیست. البته بایستی تلقی ما از برنامهٔ رایانه‌ای به‌نحوی باشد که با ملاحظهٔ تلقیِ خاص نمادگرایان مدنظر قرار گیرد. در نقد فرض روان‌شناسانه، دریفوس نظر به پدیدارشناسان و همچنین روان‌شناسان پیروِ گشتالت[4] دارد. به‌باور ایشان تجربهٔ ادراکیِ انسان، ساختارِ شکل-زمینه[5] دارد. در هر تجربهٔ ادراکی، شکل برجسته‌شده و باقیِ گسترهٔ آگاهی در زمینه یا به‌تعبیر پدیدارشناسان در افق بیرونی[6] جای می‌گیرد. آنچه در زمینه جای می‌گیرد، مبهم و نامتعین است و بی‌شک آگاهی پرداز شگرانه از آن نمی‌تواند از جنس پرداز شگری رایانه‌ای باشد؛ بنابراین پرداز شگری رایانه با امری مشخص و نامبهم ارتباط دارد، درحالی‌که تجربهٔ ادراکی انسان این‌گونه نیست. آگاهی در انسان، بستر و زمینه دارد.

---

1. Claude Shannon (1916–2001).

2. Cybernetic.

۳. برای مطالعهٔ بیشتر به استدلال‌های جان سرل، خصوصاً استدلال اتاق چینی در کتاب راز آگاهی مراجعه کنید که ناظر به این تلقی از ذهن است که ذهن صرفاً رایانهٔ دیجیتال است. البته وی معتقد است که این تلقی در اوایل ظهور هوش مصنوعی بود اما در حال حاضر به واقع مدافعی ندارد.

4. Gestalt psychologist.

5. figure - background.

6. outer horizon.

افزون بر این، پردازشگری معنادار در انسان از کُل به جز است. وقتی اتومبیلی را مشاهده می‌کنیم، آن را همانند کلی در نظر می‌گیریم حتی اگر پشت آن را نبینیم، هنگامی که از روبه‌رو به آن نظاره می‌کنیم. پدیدارشناسان به این چیزی بیش از شکل ادراکی، افق درونی[1] می‌گویند؛ یعنی ما ابتدا با یک ابژه به صورت کل مواجه می شویم و اگر بخواهیم که آن را بهتر بشناسیم، پشت و جزئیات آن را نیز ملاحظه می‌کنیم. اما یک ماشین ازجمله هر رایانه که چیزی معادل افق درونی ندارد، این پردازش اطلاعات را به‌صورت برعکس انجام می‌دهد؛ یعنی از جز به کل (Ibid, 241). بنابراین فرض روان‌شناسانه ازنظر دریفوس مخدوش است؛ یعنی ذهن، رایانه‌ای دیجیتال نیست.

۱. ۳. نقد فرض معرفت‌شناسانه

گفتیم که محتوای این فرض این است که تمام دانش بشری قابل صوری سازی است؛ یعنی اینکه می‌توان تمام دانش بشر را به گزاره‌هایی منطقی تبدیل کرد و این گزاره‌های منطقی نیز می‌تواند توسط سیستمی هوشمند مانند رایانهٔ دیجیتال داده‌پردازی شود. همچنین این فرض، پیش‌فرض فرض روان‌شناسانه تلقی می‌شود. بنابراین چنانچه این فرض تأیید نشود، فرض روان‌شناسانه نیز ملغی قلمداد خواهد شد.

دریفوس علت چنین تصوری را موفقیت دانش‌های صوری می‌داند؛ برای مثال می‌توان از فیزیک و زبان شناسی نام برد. به باور دریفوس، صوری‌سازی در فیزیک ارتباطی با جهان معنادار انسانی ندارد، اما در خصوص زبان شناسی بحث متفاوت است. دریفوس با اشاره به چامسکی[2] و دیگر زبان شناسان گشتاری[3] بیان می‌کند که آن‌ها کشف کرده‌اند که با انتزاع از اجرای[4] انسانی؛ یعنی به‌کارگیری جملات جزئی در موقعیت‌های جزئی، آن‌ها می‌توانند آنچه را که باقی مانده است؛ یعنی توانایی شناخت دستور زبانی جملات خوش‌ساخت و کنارگذاشتن موارد بدساخت، صوری‌سازی کنند؛ یعنی آن‌ها می‌توانند تئوری برای توانایی[5] زبانی فراهم آورند (Ibid, 110)، اما وی با پیش‌کشیدن بحثِ دانش‌های مهارتی[6] نشان می‌دهد که بسیاری از دانش‌های بشری تن به صوری سازی نمی‌دهد. به‌عقیدهٔ وی، اساساً دانش‌های مهارتی قابل تبدیل به دانش‌های گزاره‌ای[7] نیستند؛ مثلاً استعمال استعاری واژه‌ها امری است که درخور صوری سازی نیست (این امکان در پیشرفت‌های اخیر هوش مصنوعی در دسترس است؛ اما آن دانشی از هوش مصنوعی که محل نقد دریفوس بوده است، چنین قابلیتی نداشته است)، حتی جملات بدساخت و ناقص نیز توانایی انتقال معنا را دارند.

بنابراین دریفوس با رد امکان صوری‌سازی دانش بشری، فرض معرفت‌شناسانه را نیز نمی‌پذیرد.

---

1. Inner horizon.
2. Noam Chomsky (1928-).
3. Transformational.
4. Performance.
5. Competence.
6. know-how.
7. know-that.

۱. ۴. نقد فرض هستی‌شناسانه

فرض هستی‌شناسانه ناظر بدین بود که «آنچه هست، واقعیاتی منطقاً جدا و به‌خودی‌خود مستقل از هم است.» گفتیم که این شرط، پیش‌فرضِ فرضِ معرفت‌شناسانه است. دریفوس معتقد است که کل سنت فلسفهٔ غرب اعم از تجربه‌گرایی و عقل‌گرایی پشتوانهٔ این فرض به شمار می‌رود؛ «مکاتب عقل‌گرا و تجربه‌گرا در اتمیسم منطقی راسل به نقطهٔ مشترک رسیدند و این ایده، کامل‌ترین بیانش را در تراکتاتوس[1] ویتگنشتاین[2] یافت؛ جایی که جهان به‌واسطهٔ مجموعه‌ای از واقعیت‌های اتمی تعریف می‌شود که می‌تواند در گزاره‌های منطقاً مستقلی بیان شود. این خالص‌ترین صوری‌سازی از فرض هستی‌شناختی است» (Ibid, 123). ریشهٔ این فرض به‌باور دریفوس، تشخیص‌ندادن تمایز میان عالَم[3] و جهان[4] است. به‌دیگر سخن، ناتوانایی «در جداسازی بین آنچه که از سوی یک چیز در موقعیت انسانی و آنچه از سوی یک حالت فیزیکی سامانه به دست داده می‌شود» (همان، ۱۲۵)؛ یعنی ما از سویی با عالم که واقعیات فیزیکی مستقل‌اند، مواجه هستیم و از سوی دیگر، جهان معنادار انسانی که با همدیگر مرتبط است و ازنظر معنایی و مفهومی وابسته به یکدیگرند. در این میان دریفوس با یاری مکتب‌های فلسفی که رویکردی انتقادی نسبت به سنت فلسفی دکارتی-اروپایی دارند؛ یعنی هرمنوتیک، اگزیستانسیالیسم و به‌ویژه پدیدارشناسی به نقد فرض هستی‌شناسانه مبادرت ورزیده است. ازجمله می‌توان به ویتگنشتاین متأخر و نظریهٔ صورت‌های زندگیِ[5] وی در کتاب پژوهش‌های فلسفی[6] اشاره کرد. چنان‌که دریفوس با یاری‌گرفتن از آموزه‌های این فیلسوفان نشان می‌دهد؛ جهانِ انسانی کاملاً متفاوت با عالم فیزیکی است. جهان انسانی، جهانی زمان‌مند، معنادار، آمیخته به ارزش، دارای‌بستر و کل‌گرایانه است که نمی‌تواند به عالم فیزیکی که مجموعه‌ای از واقعیت‌های بی‌زمان، بی‌معنا، خنثی، فارغ از بستر و اتمی است، فروکاسته شود (محمدعلی خلج، ۱۱۳).

فرض هستی‌شناسانه، پیش‌فرضِ فرضِ معرفت‌شناسانه است؛ بنابراین هنگامی که فرض هستی‌شناسانه مخدوش شود، به‌تبع آن فرض معرفت‌شناسانه نیز مردود می‌شود و ازآنجاکه فرض روان‌شناسانه بر پایهٔ فرض معرفت‌شناسانه استوار شده است؛ فرض روان‌شناسانه نیز بی‌اساس و ناپذیرفتنی خواهد بود. بنابراین محتوای آنکه ناظر بر نوعی این همانی کارکردی میان ذهن و رایانهٔ دیجیتال ازآن‌روی که هردو به داده‌پردازی اطلاعات می‌پردازند، بود نیز رد می‌شود. با مردودشدن هر سه فرض روان‌شناسانه، معرفت‌شناسانه و هستی‌شناسانه توسط دریفوس، پارادایم نمادگرایی ازنظر وی کاملاً ناپذیرفتی و ناممکن است. البته ذکر این نکته ضروری است که نقدهایی که دریفوس ناظر به هوش

1. *Tractatus Logico-Philosophicus* (1922).
2. Wittgenstein, Ludwig (1889-1951).
3. Universe.
4. World.
5. Forms of life.
6. Philosophical Investigations (1953).

مصنوعی نمادی ابراز داشته است با پیشرفت دانش‌های مرتبط با هوش مصنوعی، امروزه وارد نیستند. این مهم را می‌توان در رأی نهایی دریفوس در پذیرش هوش مصنوعی هیدگری مشاهده کرد. البته بایستی دقت کنیم که همان طور که ذکر شد دریفوس، مطلقِ هوش مصنوعی را نقد نمی‌کند، بلکه آن تلقیِ ساده‌انگارانهٔ پژوهشگرانِ نمادگرا را انتقاد می‌کند.

۲. پارادیم پیوندگرایی

برخلاف نمادگرایان، پیوندگرایان بر این باورند که ذهن و رایانهٔ دیجیتال از دو ساختار کاملاً متفاوت برخوردارند. پیوندگرایی به‌لحاظ روند آن درست عکس طریقی است که نمادگرایان در پیش گرفته بودند. کامپیوترهایی که توسط نمادگرایان برنامه‌ریزی می‌شد در انجام برخی امور مثل بازی شطرنج توفیق می‌یافتند، اما از انجام اموری که هر فردی در زندگی روزمره به‌راحتی انجام می‌دهد ناتوان بودند. جان هوگلند[1] توانایی هایی نظیر بازی شطرنج را جداافتاده از مابقی زندگی نام گذاشت (Graham, 85). پیوندگرایان درصدد ساخت سیستمی هوشمند بودند که قابلیت یادگیری داشته باشد؛ یعنی با دریافت برخی ورودی‌ها و خروجی‌های مرتبط با آن، بتواند با دادن ورودی‌های مشابه، خروجی‌ها را تشخیص دهد. این ویژگی بود که رایانه‌های ساخته‌شده توسط رهیافت نمادگرایی، فاقد آن بودند. پیوندگرایان معتقد بودند که «مغز انسان ساختمانی متفاوت از کامپیوتر الکترونیکی دارد. مغز از میلیون‌ها سلول عصبی یا نورون[2] تشکیل شده است که این نورون‌ها همگی با یکدیگر ارتباط دارند؛ به‌نحوی که این مجموعه نورون‌ها با هم یک کلِّ درهم تنیده را به وجود می‌آورند که شبکه‌ای یگانه و درهم‌پیچیده از نورون‌ها است. هنگامی که اندام‌های حسی بدن از یک محرک خارجی در محیط متأثر می‌شوند، نورون‌های حسی پیام عصبی را به مغز انتقال می‌دهند و به این ترتیب سبب میان کنش در نورون‌های مغز می‌شوند. این تأثیر، کل شبکهٔ عصبی مغز را در بر می‌گیرد. در اینجا تفاوت مهمی میان کامپیوتر و مغز دیده می‌شود. در مغز برخلاف کامپیوتر، واحد خاص جهت پردازش اطلاعات وجود ندارد. به‌عبارت‌دیگر، اطلاعات در مغز به‌نحو موازی[3] پردازش می‌شوند؛ درحالی‌که این کار در کامپیوتر به‌صورت سری[4] یا متوالی انجام می‌شود. در واقع قسمت‌های مختلف مغز بر روی اطلاعات واصله، هم‌زمان عمل داده‌پردازی را انجام می‌دهند و همهٔ اطلاعات در نقطهٔ واحد، پردازش نمی‌شود» (طهماسبی، ٤٢). در این پارادایم، ماشین‌های هوشمند نیاز به آموزش دارند تا بتوانند طرح‌ها را از یکدیگر تمیز دهند (هرچند با رشد هوش مصنوعی، متدهایی وجود دارد که نیاز به آموزش ندارد، اما هنگامی که دریفوس به نقد این پارادایم می‌پرداخت، چنین متدهایی البته طراحی نشده بود)، پیوندگرایان برای این مهم از مغز انسان الگوبرداری کرده بودند.

1. Haugeland, John (1945-2010).
2. Neuron.
3. Parallel.
4. Serial.

پیوندگرایان، موافق این دیدگاه نیستند که پاره‌های مختلف اطلاعات در بخش‌های مختلفی از مغز پردازش می شود. به‌زعم آنان، تمامی اطلاعات در کل شبکهٔ عصبی به‌عنوان کلّ درهم‌تافته جای دارد و ازاین‌رو نمی‌توان برای اطلاعات خاصی، جای ویژه‌ای را در مغز نشان داد. پیروان این رهیافت معتقدند که مغز انسان بر اساس تصورات یا بازنمودها کار نمی‌کند و طرز تلقی مبتنی بر تصورات و باورها را رد می‌کنند (Lowe, 221-227).

به این پارادایم با کارهای عصب شناسی به نامِ دونالد هب[1] که بر روی یادگیری شبکه‌های عصبی[2] کار می‌کرد، توجه شد (Hebb). این رویکرد در مقابل رویکرد نمادی، وجهی کل‌گرایانه[3] دارد و سیستم در این پارادایم پس از هربار انجام‌دادن عملیات، خبره‌تر می شود. این در حالی است که سیستم‌هایی که بر پایهٔ نمادگرایی کار می‌کردند، چه بار نخست و چه بار هزارمِ انجام عملیات، هیچ تأثیری در خبرگی آنان نداشت. اما پیوندگرایی با الگوبرداری از سیستم عصبی مغز به‌گونه‌ای سازماندهی شده بود که قادر به یادگیری و عملکرد آن رو به پیشرفت بود. بنابراین «پیوندگرایی در صدد است این ویژگی زندگی انسانی را که به‌موجب آن تجربهٔ هرچه بیشتر به مهارت هرچه بیشتر منتهی می‌شود، پیاده‌سازی کند. ازاین‌رو مبتنی بر یک سری الگوریتم‌های آموزشی، هر بار به‌سوی بهبود نتایج پیش می‌رود. گویی شبکه هر بار در یافتن پاسخ، ماهرتر شده است. در این رابطه در هر پردازشِ آموزشی، کل شبکهٔ عصبی درگیر است تا پیوندهای میان نورون‌ها به‌گونه‌ای اصلاح شود که شبکه، پاسخ مناسب‌تری بدهد» (Clark, 8).

بنابراین از نظرگاهی فلسفی، پیوندگرایی از دو جهت با هوش مصنوعی نمادی متفاوت است: نخست به‌جهت بازنمایی کل‌گرایانه و دوم به‌جهت پیاده‌سازی آموزش‌یافتن مهارت در عمل (محمدعلی‌خلج، ۱۱۵).

برخی مشکلاتی که دریفوس به آن اشاره کرده بود در این پارادیم مرتفع شده است؛ اما علی‌رغم این ملاحظات، دریفوس در برابر این رهیافت نیز موضعی منتقدانه اتخاذ می‌کند.

۲. ۱. نقد پیوندگرایی

دریفوس در نقد پارادایم پیوندگرایی دست روی ویژگی می‌گذارد که به تعمیم[4] مشهور است. همهٔ پیوندگرایان «می‌پذیرند که شبکهٔ هوشمند باید بر تعمیم توانا باشد. برای نمونه، برای انجام یک طبقه‌بندی با دادن مثال‌های مناسب از ورودی‌های مرتبط با یک خروجی مشخص، شبکه باید خودِ ورودی‌های دیگر از همان نوع را به همان خروجی مرتبط سازد» (Dreyfus, *what computers still can't do*, xxxvi). به‌نظر دریفوس، شبکه‌های مبتنی بر پارادایم پیوندگرایی فقط با دخالت طراح شبکه قادر به انجام تعمیم مناسب و درخور هستند. در واقع، وی با بیان

1. Donald O. Hebb (1904-85).
2. Neural networks learning.
3. Holistic.
4. Generalization.

مثال‌هایی نشان می‌دهد که پیوندگرایی در فهم ما از جهان انسانی اشتراک ندارد. وی در برابر نمادگرایی از پیوندگرایی دفاع می‌کند؛ اما رویکردش به‌گونه‌ای است که درنهایت پیوندگرایی را نیز در جایگاهی مشابه با نمادگرایی قرار می‌دهد. به‌نظر دریفوس «یک شبکه باید در فهم عقل متعارف[1] ما از جهان شریک باشد تا در درک ما از تعمیم مناسب شریک شود» (Ibid). دریفوس نتیجه می‌گیرد که «بنابراین پژوهشگران شبکه‌های عصبی در این موقعیتشان با پیروزی‌های کوچک بدون شیوه‌ای اصولی برای تعمیم، به نظر می‌رسد در همان جایگاه پژوهشگران هوش مصنوعی کلاسیک باشند، هنگامی که من در ۱۹۹۰ شروع به نوشتن دربارهٔ آن‌ها کردم» (Ibid, xxxvii).

برای دریفوس، پارادایم پیوندگرایی علی‌رغم پیشرفت‌های درخور توجه باز هم نتوانست موقعیتی فراهم کند تا در آن بتوان سیستمی داشت که با انسان در نیازها، عواطف، احساسات و فهم از جهان، اشتراک داشته باشد. به نظرِ دریفوس تنها سیستمی که بتواند چنین ویژگی‌هایی داشته باشد؛ یعنی در فهم هر روزهٔ ما از جهان مشترک باشد، قادر به امکان هوشمندی است و نه نمادگرایی و نه پیوندگرایی هیچ‌کدام نتوانسته‌اند چنین شروطی را تحقق بخشند. نمی‌توان گفت که به واقع دریفوس در این داوری بر حق است؛ چراکه هرکدام از این پارادایم‌ها تا اندازه‌ای موفق بوده و توانسته‌اند فهمی هرچند اندک در درک جهان داشته باشند. در اینجا به نظر می‌رسد رویکرد دریفوس، رویکردی حداکثری است. البته دیدیم که دریفوس علی‌رغم نقد خود بر پیوندگرایی به پیشرفت این پارادایم در تحقق شرایط مذکور توجه می‌کند اما این برای وی ناکافی می‌نماید.

۳. هوش مصنوعی هیدگری

هوش مصنوعی هیدگری شاخه‌ای از پارادایم سوم هوش مصنوعی؛ یعنی حیات مصنوعی است. این پارادایم سه گرایش دارد که به رویکرد خیس،[2] سخت[3] و نرم[4] مشهور است. «گرایش نخست، پژوهش برای پیاده سازی رفتار هوشمند را بر روی بافت‌های زندهٔ جانداران پی می‌گیرد. این گرایش بیش از هر دانشی به بیوشیمی[5] وابسته است» (Dennett, 155, 268, 271)، اما «گرایش دوم درصدد است تا ربات‌هایی با توانایی‌های زیستی بسازد. این گرایش در تراز سخت‌افزاری در پی تحقق جریان حیات در یک جاندار است. در این نگرش، ربات و محیطش سامانهٔ واحد دانسته می‌شوند و ربات همواره در یک بستر محیطی در نظر گرفته می‌شود» (محمدعلی‌خلج، ۱۱۸). گرایش سوم «به مطالعه و شبیه سازی نرم‌افزاری رفتار جاندار در عرض کل محیط زیست و طول تکامل می‌پردازد و رفتار هوشمندانه را فرآوردهٔ درآمیختگی و دادوستد جانداران با این محیط در حال تکامل می‌داند. این گرایش در

1. common sense.
2. Wet.
3. Hard.
4. Soft.
5. Biochemistry.

محدودهٔ نرم‌افزاری به پژوهش می‌پردازد» (clark, 9-14).

هوش مصنوعی هیدگری را می‌توان جریانی ذیل گرایش سخت‌افزاری دانست. نخستین کسی که این نام را به این جریان داد، تری وینوگراد[1] بود که از هوش مصنوعی کلاسیک ناامید شده بود. چنان‌که وینوگراد می‌گوید یک سری از جریان‌های هوش مصنوعی به‌صورت مستقیم از دریفوس تأثیر پذیرفتند. دریفوس دو ویژگی کلی برای هوش مصنوعی هیدگری قائل است: «۱. اینکه نقد هیدگر بر بازنمودگرایی درونی دکارتی را پذیرفته‌اند؛ ۲. شعار جان هوگلند را مبنی‌بر اینکه شناخت، بدن‌مندانه[2] و موقعیت‌مند[3] است، سرلوحهٔ عمل قرار می‌دهند» (Dreyfus, Overcoming the Myth of the Mental , 7).

بنابراین شروطی ویژه که در ابتدای این مقاله ذکر شد، دو شرط فوق هستند. از اینجا مشخص است که برای پروژهٔ هوش مصنوعیِ هیدگری، شناخت بایستی بدن‌مندانه و بسترمند و دارای ساختار و موقعیت‌مند باشد. این واقعیتی است که دریفوس به‌تبع هیدگر و مرلوپونتی بیان می‌کند. اهمیت بدن‌مندی در دریفوس متأثر از مرلوپونتی است. از نگاه مرلوپونتی حضور ما در جهان؛ یعنی بودن در جهان[4] برای ما به‌نحو بدن‌مندانه صورت می‌پذیرد. همچنین آموزهٔ بودن در جهان برای هیدگر همراه است با بسترمندی[5] یا پس‌زمینه[6] داشتن و ساختارمندبودنِ شناخت و آگاهی؛ لذا به‌باور دریفوس، هر شناختی بایستی دارای ملاحظات فوق باشد.

بنا بر شرح پدیدارشناسانهٔ دریفوس از رفتار هوشمندانه، ما همواره در تعادلی درونی و بدنی با جهانمان قرار داریم. هرگاه این وضعیت تعادلی و تسلط ما بر پیرامونمان بر هم بخورد بی‌آنکه نیاز به تحلیل بازنمودی امر بیرونی باشد، حسی درونی ما را به‌سوی حداکثر تعادل با محیط و حداکثر تسلط بر آن سوق خواهد داد. نمونه‌ای که دریفوس از مرلوپونتی وام می‌گیرد، نگریستن به تابلوی نقاشی است. اگر ما زیاد به تابلو نزدیک شده با شیم خودبه خود کمی عقب می‌رویم و اگر زیاد دور باشیم، بی‌نیاز به تأمل کمی جلو می‌رویم تا تابلو را تار ادراک نکنیم. حس درونی تشویشِ حاصل از ادراک تار، ما را به‌سوی حداکثر تعادل هدایت خواهد کرد. به‌باور دریفوس، این تنظیم تعادل از درون، ویژگی کلیِ نسبت انسان و جهان است (Dreyfus, Why Heideggerian AI Failed, 16). دریفوس الگوی نورونی ویژه‌ای را به‌عنوان مصداق چنین سیستمی معرفی می‌کند که به الگوی نورونی فریمن[7] مشهور است. وی سازوکار چنین سیستمی را این‌گونه توصیف می‌کند که متناظر با حس درونی ما از تعادل با محیط در شرح مرلوپونتی و هیدگر، در الگوی فریمن حداقل انرژی در هر غشای مغزی قرار دارد. متناسب با هر محرک

1. Terry Winograd (1946-).
2. Embodiment.
3. Embedded.
4. being-in-the-world.
5. Contextually.
6. Background.
7. Freeman's Neuro dynamics model.

بیرونی، شبکه‌هایی نورونی در بخش مربوطه در مغز شکل می‌گیرند که پاسخ به محرک را هدایت می‌کنند. زمانی که پاسخ ما به محرک، فعالیت مغز را در حداقل انرژی قرار نمی‌دهد، همان زمانی است که در ساحت پدیداری حس درونی ما حداقل تشویش را ندارد و تسلط ما بر محیط و تعادل‌مان با آن بر هم خورده است. در این زمان کل بخش مدنظر وارد بی‌نظمی نورونی می‌شود و تا ایجاد شبکه‌ای جدید که به محرک پاسخی مناسب بدهد و وضعیت فعالیت را به حداقل انرژی برساند، بی‌نظمی ادامه می‌یابد. این کنش و واکنش‌هایی که به‌سبب رسیدن به حداقل انرژی در تراز نورونی صورت می‌پذیرد، به‌باور دریفوس در تراز پدیداری با کوشش در راستای رسیدن به کمترین تشویش و بیشترین تعادل با محیط متناظر است (Ibid, 19-25).

بنابراین باتوجه به ملاحظات اخیر، دریفوس پس از واردساختن نقدهای خود به دو پارادایم نخست هوش مصنوعی، درنهایت هوش مصنوعی هیدگری را با قید شروط دوگانه می‌پذیرد. مسئلهٔ ما در اینجا نشان‌دادنِ این امر است که این پذیرش می‌تواند همسو و همخوان با مواضع پیشین وی باشد. در بخش بعد می‌کوشیم تا با ارائهٔ سه راهکار نشان دهیم که چگونه می‌توان میان مواضع پیشین دریفوس در نقدهایش بر هوش مصنوعی کلاسیک و موضع اخیر وی در پذیرش هوش مصنوعی هیدگری، سازگاری برقرار کرد.

۴. راهکارها و رهیافت‌ها

در ابتدای امر به نظر می‌رسد که تصور سیستمی که با ما انسان‌ها بتواند در درک جهان و فهم ما از جهان اشتراک داشته و در درک عواطف، نیازها، آرزوها، تنش‌ها و بودن در جهان با ما سهیم باشد، بسیار دشوار است. در واقع، چنین مسئله‌ای گاه چنان بدیهی است که به سادگی ممکن است رأی ما را به امکان نداشتن سازگاری سوق دهد و نظریات دریفوس را نوعی دوگانگی و مغایرت در مواضع پیشین و اخیر تلقی کند؛ اما چنانچه به شروط ویژه‌ای که وی به آن‌ها تأکید داشته است برگردیم، می‌توان راهکارهایی یافت. نخستین رهیافت، رهیافت دنیل سوسر است که در ادامه آن را توضیح خواهیم داد.

۴. ۱. رهیافت دنیل سوسر

دنیل سوسر،[1] استاد دانشگاه پنسیلوانیا، رهیافتی برای جمع‌کردن و سازگاری پذیرش امکان هوش مصنوعی و نظریات و نقدهای دریفوس بر پارادایم‌های پیشین آن را ارائه می‌دهد. وی با بهره‌گیری از نظریات مارک بیک‌هارد[2] که یک اینتراکتیویست[3] یا تعامل‌گراست، می‌کوشد تفسیر و تعمیمی از نظریات دریفوس ارائه کند که در پی آن هم بتوان امکان هوش مصنوعی هیدگری را از دید وی پذیرفت و هم اینکه شبههٔ ناسازگاری میان دو رأی وی را نیز مرتفع کرد.

---

1. Daniel Susser: استادیار دانشکدهٔ علوم و فناوری اطلاعات، دانشیار پژوهشی و عضو هیئت‌علمی اصلی در مؤسسهٔ اخلاق راک و عضو هیئت‌علمی وابسته در گروه فلسفه در دانشگاه ایالتی پن.

2. Mark Bickhard، استاد رباتهای شناختی و نیز استاد فلسفهٔ دانش در دانشگاه لیه. Henry R. Luce: با نام اصلی

3. Interactivist.

سوسر دربارهٔ نقش بدن در ایجاد هوش مصنوعی انسانی می‌نویسد: «برای کسانی که نقدها و ملاحظات دریفوس را خوانده و صادق دانسته‌اند، پرسش از تولید هوش مصنوعی انسانی از دیدگاه فلسفی متوقف و به مهندسان واگذار شده است تا بدنی بسازند که مانند بدن انسان بتواند در تعامل با جهان اطراف خویش باشد» (Susser, 278)؛ اما اگر بخواهیم هوش مصنوعی انسانی با بدن غیرانسانی[1] داشته باشیم، آیا می‌توان با کارهای دریفوس به تولید آن امید داشت؟ وی ادامه می‌دهد که: «... بسیاری از این‌گونه هوش‌های غیرانسانی را هم‌اکنون نیز می‌توان در اطراف مشاهده کرد. سگ‌ها، گربه‌ها، دلفین‌ها، برخی پرندگان و اختاپوس‌ها اغلب، رفتارهای هوشمندانه[2] از خود نشان می‌دهند. بنابراین اگر ما نمی‌توانیم هوش مصنوعی با بدن انسانی؛ یعنی آن‌گونه که انسان واجد آن است داشته باشیم، آیا اَشکال دیگر این هوش نیز نمی‌تواند وجود داشته باشد؟ اگر نمی‌توانیم هوشی مانند انسان داشته باشیم، آیا هوش‌های دیگر نیز ممکن نیست، به‌نحوی که نه مانند انسان باشد و نه سگ و نه دلفین؟» . (Ibid, 278) اگر این مورد را بخواهیم بررسی کنیم، بایستی از دریفوس این را بپرسیم که آیا و چگونه استدلال‌های او برای هوش انسانی می‌تواند مناسب چنین آلترناتیوهایی باشد، به‌ویژه اگر بدن[3] به‌عنوان اساس و بنیاد هریک از وجوه این زندگی هوشمندانه باشد؟ بنابراین پرسش می‌تواند این باشد که بدن مصنوعی غیرانسانی شبیه به چیست و چگونه است؟ آن چیست که تمامی موجودات دارای هوش در آن مشترک هستند؟ چه چیزی در بدن انسان، بدن سگ و اختاپوس مشترک است؟ آثار هیوبرت دریفوس هرگز این مسئله را به‌طور شفاف بحث نکرده است و پاسخی برای آن به‌روشنی نمی‌توان یافت. در اثرهای پدیدارشناسانهٔ دریفوس به نظر می‌آید که وی مقصودش از بدن، بدن انسانی است. اگر بخواهیم هوش مصنوعی غیرانسانی داشته باشیم و با این‌حال همچنان بینش‌های دریفوس دراین‌باره را درست تلقی کنیم، بایستی که آنالیز و تحلیل او از این بحث را توسعه و گسترش دهیم (Ibid, 277-8).

سوسر در ادامه تلاش می‌کند تا هیوبرت دریفوس را با مارک بیک‌هارد که اینتراکتیویست است وارد گفت‌وگو کند. تفاوت این دو در این است که تئوری بیک‌هارد بر سطح بالاتری از تعمیم نسبت به آنچه دریفوس مدنظر دارد، متمرکز است. در واقع نظریهٔ شناخت مارک بیک‌هارد شبیه تئوری فعالیت ماهرانه[4] دریفوس به‌نحو مبنایی و گسترش یافته است. تبیین بیک‌هارد به‌جای ساخت یافته‌شدن در قالب هوش انسانی و بدن انسانی به‌طور کلی بر سیستم‌های فیزیکی[5] قرارگرفته است. بنابراین این نظریه، تصویر دریفوس از ویژگی اساسی بدن‌ها را بیرون می‌کشد و می‌گوید که همهٔ موجودات بدن‌مند هوشمند، واجد چه ویژگی مشترکی هستند. بیک‌هارد، نظریهٔ شناختی را که در روح نظریهٔ فعالیت ماهرانه دریفوس قرار داشت، گسترش داد. به جای درنظرداشتن هوش انسانی به‌مثابهٔ ابژه مورد تحلیل قرار

1. non-human.
2. intelligent behavior.
3. Body.
4. skillful coping.
5. Physical systems.

گرفته، آن‌گونه که دریفوس بدان مبادرت می‌ورزید، بیک‌هارد در نظر دارد تا ساختار هوش[1] را بسیار گسترده‌تر بررسی کند. بنابراین وی نظریهٔ خود را بسیار کلی‌تر از آنچه دریفوس در نظر دارد انجام می‌دهد؛ یعنی در قالب واژه‌های ساختارها و عملکردهای پردازش‌های فیزیکی و سیستم‌های فیزیکی. تئوری بیک‌هارد نیازمندی‌های واقعی هر سیستم فیزیکی را برای اینکه ظرفیت ایجاد شکل و فرم هوشمند را داشته باشد، توصیف می‌کند. این سیستم‌ها لزوماً نیازمند این نیستند که فقط شبیه به انسان باشند، بلکه صرفاً لازم است که در چنین سطح پایه‌ای مانند انسان‌ها باشند و هر سیستم این‌چنینی، یک سری ویژگی‌ها و مشخصه‌های سازمانی مشترک با هم دارند (Ibid, 278-279).

مدل بیک‌هارد به مرکزیت نوعی از سیستم آزاد [باز] ترمودینامیکی[2] است که به ساختار اتلافی[3] مشهور است. چنین سیستمی با این واقعیت تعیّن یافته است که در شرایط دور از تعادل ترمودینامیکی[4] عمل می‌کند و اگر چنین شرایطی برقرار نباشد، ادامهٔ فرایند را متوقف می‌کند یا به‌اصطلاح موجودبودن را متوقف[5] می‌کند (Bickhard, 11). نمونه‌ای از چنین سیستم‌هایی، تنظیم‌کردن سیستم‌های سادهٔ انتقال گرما است، به‌طوری که آن‌ها برای [وضعیت‌های] باد و باران پاسخ‌گو باشند و همچنین تا پیچیده‌ترین سیستم‌های جهان؛ یعنی ارگانیزم‌های زنده.[6]

وی سپس به دو نوعِ متفاوت از چنین سیستم‌هایی اشاره می‌کند؛ یعنی سیستم‌هایی که توانایی خودنگهداری دارند و سیستم‌هایی که برای رسیدن به تعادل، نیازمند دخالت دیگر سیستم‌ها هستند. افزون بر آن، درون سیستم‌های دور از تعادل، تمایزی میان آن‌هایی که نیازمند مداخلهٔ صریح و قاطع سیستم‌های دیگر هستند تا از شرایط دور از تعادل برخوردار شوند، یعنی به شرایط دور از تعادل برسند و آن‌هایی که قادرند تا برخی گسترش‌ها و دستکاری‌ها را برای برقراری این شرایط بر خود شان اعمال کنند، می‌تواند ایجاد شود. مثالی برای سیستم‌هایی که نیازمند مداخلهٔ دیگر سیستم‌ها هستند، حمام شیمیایی[7] است که به‌طور ثابت و یکنواخت، مادهٔ شیمیایی خاص به آن پمپاژ می‌شود تا آن را در شرایط دور از تعادل نگه دارد و برای نمونه‌ای از سیستم‌هایی که نیاز به مداخله ندارند، شعله‌ای شمع را که دمای آستانهٔ احتراق را بالا نگه می‌دارد،[8] در نظر بگیرید. این شعله، موم را ذوب می‌کند تا اینکه آن را تا فتیله بالا برده و بچسباند و موم را تا فتیله بخار می‌کند تا بسوزد، سپس در شرایط گرانشی جوی استاندارد[9] باعث انتقال گرما می‌شود که [در پی آن] اکسیژن تازه را وارد می‌کند و از ضایعات آن نیز خلاص می‌شود (Ibid, 11) بااین‌حال

1. Intelligence construed.
2. open thermodynamic system.
3. Dissipative structure.
4. Far-from-thermodynamic-equilibrium.
5. Cease to exist.
6. living organisms.
7. Chemical bath.
8. Maintains above combustion threshold temperature.
9. Standard atmospheric gravitational conditions.

می‌توان سیستم‌های بسیارِ دیگری غیر از شعلهٔ شمع را نام برد که با محیط پیرامون به‌نحوی پیچیده‌تر تعامل می‌کنند. سوسر در ادامه ذکر می‌کند که: «من باور دارم که ما هم‌اکنون می‌توانیم فهمی مشترک از هوش را از تصویر دریفوسی آن و تئوری بیک‌هارد، مشاهده کنیم. ازآنجاکه در هر دو تئوری، کم‌وبیش هوش به‌مثابهٔ موضوع و مادهٔ عمل ماهرانه در نظر داشته شده است تا آنچه را فرد می‌خواهد و بدان علاقه‌مند است، محقق سازد و درجایی که این کار را انجام می‌دهد، یعنی تعاملی پویا با جهان دارد که شخص به‌نحو اساسی و به‌طور اجتناب‌ناپذیری، موقعیت‌مند شده است» (Susser, 284).

به‌نظر سوسر، فرایندهای توصیف‌شدهٔ فوق در جایی که هر سیستم فیزیکی از شرایط هستی خود توسط سیستم تمیزگذاری موفق[1] میان محیط‌های سالم و سمی مراقبت می‌کند و با گرایش به‌سوی رهیافت پیشین که رهیافت انجام فعالیت ماهرانه در ساده‌ترین شکل آن است، شرح و توصیفی از فعالیت ماهرانهٔ دریفوس است که در بالاترین سطح از گستردگی قرار گرفته است. «بنابراین من مایلم که پیشنهاد کنم که مشخصه‌سازی و توصیف صفات اختصاصی بیک‌هارد از سیستم فیزیکی خودنگه‌دار بازگشتی،[2] به‌خوبی تعریفی است برای آنچه به‌طور فیزیکی هر بدن را قوام می‌بخشد» (Ibid, 284).

می‌توان دربارهٔ این دو دیدگاه نتیجه گرفت که بیک‌هارد به‌خوبی آنچه مدنظر دریفوس بوده را توسعه داده است: «آنچه دربارهٔ دیدگاه بیک‌هارد ضروری است، این است که این اجزای اساسیِ فعالیت ماهرانه حتی در ابتدایی‌ترین تجسم‌هایشان نیز اجمالاً حضور دارند و بیک‌هارد نه‌تنها با اشارهٔ استعاری به چینش از پایین‌ترین[3] کارکرد شناختی که در آن انسان و حیوان‌های غیرانسان شریک هستند؛ یعنی آن‌گونه که دریفوس مدنظر دارد، بلکه با تشریح دقیق اینکه چگونه چنین هوش‌های ابتدایی و اولیه کار می‌کنند و اینکه چگونه بالاترین سطح[4] هوش ممکن است به‌طور موجه و باورپذیر از آن برخیزد. این به ما راهی را نشان می‌دهد تا دربارهٔ ساخت هوش مصنوعی به جای ساختن هوش مصنوعیِ انسانی یا انسان‌وار بیاندیشیم» (Ibid, 284-285).

در این رویکرد مشاهده کردیم که با خوانشی خاص از روح نقدهای دریفوس و نظریهٔ فعالیت ماهرانهٔ وی، می‌توان به سازگاری میان دو رویکرد دریفوس دربارهٔ هوش مصنوعی قائل شد. این خوانش بر نقش بدن تأکید دارد، اما مانند دریفوس لزوماً بدن انسانی را شرط نمی‌داند. چنانچه خوانش ما از دریفوس به‌نحوی باشد که بدن‌مندی را لزوماً انسانی تلقی کنیم و در این خوانش نگاه ما به اقتباس دریفوس از مرلوپونتی باشد، بلکه نقطهٔ مشترک میان همهٔ موجودات بدن‌مند را معیار در نظر می‌گیرد و درنهایت با این واقعیت که بارزترین ویژگی بدن، تعامل با جهان اطراف

1. Successfully discriminating system.
2. Recursively self-maintaining physical systems.
3. Lower .
4. Higher level.

است، راه را به‌سوی دست‌یابی به امکان هوش مصنوعی از نوعی که دریفوس مدنظر دارد نیز فراهم می‌کند.

۴. ۲. رهیافت دوم: بدن‌مندی قوی و ضعیف

در این بخش می‌کوشیم با استمداد از جان سرل،[1] فیلسوف آمریکایی که با هیوبرت دریفوس نیز گفت‌وگوهای زیادی داشته است، راهکاری را برای سازگاری دو رأیِ به‌ظاهر ناسازگار دریفوس ارائه دهیم. البته در این کار تنها با نظر به طبقه‌بندی جان سرل از هوش مصنوعی سعی می‌کنیم چنین درجه‌بندی را نیز در دریفوس تشخیص دهیم و رهیافتی جهت سازگار ساختن دو رأی ایشان و امکان هوش مصنوعی هیدگری ارائه کنیم. اما ابتدا ببینیم جان سرل دربارهٔ هوش مصنوعی چه تقسیم‌بندی را ارائه می‌دهد.

جان سرل در نوشته‌های خود بارها به موضوعات مطرح در فلسفهٔ ذهن ورود کرده و نظریات گوناگونی داشته است. ازآنجاکه هوش مصنوعی رابطه‌ای تنگاتنگ با مسائل مطروحه در فلسفهٔ ذهن و آگاهی دارد؛ سرل نیز دراین باره نظریات خواندنی و تأمل‌انگیزی دارد. جان سرل در کتاب راز آگاهی[2] در حین بر شمردن مشکلاتی که دامن‌گیر ارائهٔ نظریه‌ای باورپذیر و خرسندکننده از آگاهی شده است، به مقولهٔ هوش مصنوعی نیز واکنش نشان می‌دهد. وی می‌نویسد: «اشتیاق به استفاده از استعارهٔ کامپیوتری به‌معنای واقعی کلمه برای ذهن... که بسیاری هنوز فکر می‌کنند مغز، کامپیوتر دیجیتال و ذهن آگاه، برنامهٔ کامپیوتری است، هرچند این نقطه‌نظر خوشبختانه نسبت به یک دههٔ قبل چندان متداول نیست. با تعبیر به این شیوه، ذهن برای مغز به‌عنوان نرم‌افزار برای سخت‌افزار است. نظرات متفاوتی دربارهٔ نظریهٔ محاسباتی ذهن وجود دارد. قدرتمندترین آن‌ها موردی است که تازه (۱۹۹۷) مطرح کرده‌ام: ذهن صرفاً برنامهٔ کامپیوتری است. در آنجا چیز دیگری وجود ندارد. این نقطه‌نظر را هوش مصنوعی قوی[3] می‌نامم تا آن را از این نقطه‌نظر متمایز کنم که کامپیوتر وسیله‌ای سودمند برای شبیه‌سازی ذهن است. همان طور که شبیه‌سازی هرچیزی که بتوانیم دقیقاً توضیحش دهیم؛ برای مثال الگوهای آب‌وهوایی یا جریان پول در اقتصاد امری سودمند است. این نقطه‌نظرِ محتاط‌تر را هوش مصنوعی ضعیف[4] می‌نامم» (سرل، ۲۱).

سرل در ادامه به نقد هوش مصنوعی قوی می‌پردازد: «هوش مصنوعی قوی را می‌توان به‌سرعت رد کرد. این کاری است که طی پانزده سال گذشته در نقد کتاب نیویورک[5] و جاهای دیگر انجام داده‌ام. طبق تعریف،[6] کامپیوتر وسیله‌ای است که نمادهای صوری را دستکاری می‌کند... ، اما نکتهٔ مدنظر در حال حاضر این است که آن سازوکار تماماً برحسب دستکاری نمادها تعریف می‌شود. محاسبه‌ای که این‌گونه تعریف می‌شود، صرفاً مجموعه‌ای نحوی از

1. Searle, John (1932-).
2. *The mystery of consciousness*.
3. Strong AI.
4. Weak AI.
5. *The New York Review of Books*.
۶. البته تعریفی که نمادگرایان در ابتدا ارائه می‌کردند.

عمل‌هاست؛ به این معنا که تنها آن ویژگی‌هایی از نمادها که صوری یا نحوی‌اند در پیاده‌سازی برنامه اهمیت دارند، اما ما از طریق تجربه فهمیده‌ایم که ذهن، چیزی بیش از دستکاری نمادهای صوری را به انجام می‌رساند. ذهن‌ها درون‌مایه دارند؛ برای مثال وقتی به انگلیسی فکر می‌کنیم، آن واژه‌های انگلیسی که از ذهن ما می‌گذرند صرفاً نمادهای صوری تعبیرنشده نیستند؛ بلکه می‌دانیم که از چه معنایی برخوردارند. برای ما واژه‌ها معنایی دارند و معنا شناسی در کار است. ذهن نمی‌تواند تنها برنامهٔ کامپیوتری باشد، چون نمادهای صوری برنامهٔ کامپیوتری، به خودی‌خود برای تضمین حضور آن درون‌مایهٔ معناشناختی که در ذهن‌های واقعی رخ می‌دهد، کافی نیستند» (همان، ۲۲). بنابراین جان سرل با درجه‌بندی هوش مصنوعی به قوی و ضعیف، ضمن ردِ هوش مصنوعی قوی با استدلال‌های مختلف که بخشی از آن ذکر شد و ازجمله می‌توان به استدلال اتاق چینی[1] نیز اشاره داشت، هوش مصنوعی قوی را رد کرده و در مقابل، اینکه کامپیوترها به‌عنوان مدلی برای شبیه‌سازی سازوکار ذهن سودمند هستند را که هوش مصنوعی ضعیف نامیده بود، محتاطانه می‌پذیرد. به‌کمک این تقسیم‌بندی تلاش می‌کنیم تا رهیافتی مشابه را در نظریات دریفوس ارائه کنیم.

دریفوس برنامه‌ای را که بر پایهٔ الگوی نورونی فریمن بازسازی شده باشد به‌عنوان هوش مصنوعی هیدگری راستین می‌پذیرد و می‌دانیم برای تعامل با جهان اطراف به دانشی نامتعین نیازمندیم که در موقعیت‌های پیش‌آمده بتواند به‌نحو ماهرانه عمل کند و خود را با موقعیت‌های اطراف تطبیق دهد و ازآن‌روی که در روح پدیدارشناسی هیدگری، بدن‌مندی به‌نحو اکید وجود ندارد و می‌دانیم که شرط بدن‌مندی برای آگاهی، بیشتر مختص اندیشه‌های مرلوپونتی است و مرلوپونتی، بدن را وجودی نه ذهنی و نه مادی، بلکه نوع سومِ وجود می‌داند؛ لذا می‌توان نتیجه گرفت که بدن‌مندی در خوانش دریفوسی می‌تواند دست‌کم دو قرائت، خوانش و تفسیر داشته باشد. بر اساس این تفاسیر دوگانه می‌توان دربارهٔ سازگاری یا ناسازگاری نظریات دریفوس اظهارنظر کرد.

الف. اگر بدن در خوانش دریفوسی لزوماً بدن انسانی تلقی شود؛ یعنی نگاه ما به اقتباس وی از مرلوپونتی باشد، آن‌گونه که وی حضور ما در جهان را بدن‌مندانه می‌دانست و معتقد بود که تمایز سوژه و ابژه در بدن ما از بین می‌رود (Merleau-ponty, 167) و از مواردی که ناظر به چنین حضور بدن‌مندانه‌ای است، نتیجه بگیریم که مقصود دریفوس از بدن، بدن انسانی است؛ بدیهی است که تا زمانی که نتوان بدنی انسانی را در پروژهٔ هوش مصنوعی کاملاً داشت، شرط دریفوس برای پذیرش هوش مصنوعی تحقق نخواهد یافت و بدین‌صورت می‌توان نظر اخیر وی در پذیرش هوش مصنوعی هیدگری را ناسازگار با نظریات پیشین وی دانست و می‌توان آن را عدول از مواضع پدیدارشناسانهٔ قبل و گرایش به طبیعت‌گرایی را علت این چرخش نظری دریفوس دانست. اما باتوجه‌به اینکه دریفوس فعالیت روزمرهٔ ماهرانه را یک باید برای تحقق هوش مصنوعی می‌داند،[2] پذیرش این نظر که دیدگاه وی

1. Chinese room Argument.

۲. در ابتدای امر به نظر می‌رسد که مقصود دریفوس نوعی کارکردگرایی است، اما باتوجه‌به نقدی که وی در پارادایم نمادی به این دیدگاه دارد ناظر بر اینکه حضور ما در جهان

لزوماً دیدگاهی حداکثری دربارهٔ بدن باشـد، چندان موجه نخواهد بود و لزوماً دلیلی بر تفکیک بدن به انسـانی و غیرانسـانی وجود ندارد. اینکه دریفوس با پذیرش امکان هوش مصـنوعی هیدگری متوجه این تناقض نبوده باشـد نیز دور از ذهن است؛ لذا به نظر نمی‌رسد بتوان این تفسیر را تفسیری درست قلمداد کرد و حتی اساس تفکیک بدن به انسانی و غیرانسانی نیز چندان وجهی ندارد. البته این امکان وجود دارد که دریفوس تحت‌تأثیر گفت‌وگوهایش با جان سـرل، نوعی از طبیعت‌گراییِ زیسـتی را مورد مداقه قرار داده باشـد، اما اینکه او چنین رویکردی را پذیرفته باشـد از نوشته‌هایش بر نمی‌آید و آنچه مشهود است دفاع وی از مواضع پدیدارشناسانه در ضمن گفت‌وگوهایش با سرل اسـت. بنابراین، این تفسـیر از بدن را تفسـیر تمام‌انسـانی از بدن یا بدن‌مندی قوی می‌نامیم و مقصـود این اسـت که چنانچه منظور از بدن‌مندی در دریفوس را کاملاً انسانی تلقی کنیم، لاجرم بین آرای وی ناسازگاری خواهیم داشت. باتوجه‌به آیندهٔ هوش مصنوعی، حتی این نوع از تحقق نیز دور از ذهن نخواهد بود.

ب. چنانچه تفسیری از شرایط بدن‌مندی، موقعیت‌مندی و تعامل را در هر شکلی و نه لزوماً انسان‌وار در نظر داشته باشیم، می‌توان گفت که این تفسیر از دریفوس با شروط ذکرشده ایشان در نقد پارادایم‌های هوش مصنوعی، واجد سـازگاری اسـت و می‌توان اسـتدلال کرد که رأی او در باب پذیرش هوش مصـنوعی هیدگری بر پایهٔ الگوی نورونی فریمن در ادامهٔ رویکرد پیشین وی و بر طبق مواضع پدیدارشناسانه و همخوان با خوانش او از پدیدارشناسی هیدگر است. بدین صورت کافی است تفسیر وی از شروط پیشین را تمام‌انسانی ندانیم، بلکه آن‌را تفسیرِ تحققی[1] بدانیم؛ به این معنا که هر سیستمی که بتواند این شروط را محقق سازد، می‌تواند واجد **هوش** تلقی شود؛ خواه این سـیسـتم زیسـتی باشـد، خواه فیزیکی یا هر گونهٔ دیگری. در این خوانش می‌توان آن دسـته از هوش‌هایی را که ما در طبیعت می‌بینیم نیز نوعی هوشـمندیِ درخور توجه به شـمار آورد و بدین‌سـان هر سـازوکاری که بتواند شـروط هوشـمندی مدنظر دریفوس؛ یعنی کل‌گرایی، بدن‌مندی، موقعیت‌مندی، قادر بر انجام فعالیت ماهرانه‌بودن، قابلیت یادگیری، تعمیم و... را داشته باشد، می‌تواند امکان هوش مصنوعی هیدگری را فراهم آورد و لزوماً نیاز نیست که بدن مندی را بدن‌مندیِ انسـانی تلقی کنیم، بلکه در این تقریر، بدن‌مندی ضـعیف مدنظر ما اسـت؛ یعنی هر بدنی خواه انسانی یا خواه غیرانسانی. در این خوانش البته نوعی کارکردگرایی لحاظ می‌شود، اما به‌معنای صرف کارکرد نیست؛ بلکه با برقراری معیارهایی مانند توانایی یادگیری، توانایی تعمیم و قدرت تصـمیم‌گیری در موقعیت‌های جزئی، از کارکردگراییِ صرف متمایز شده است.

با الهام از تقسیم‌بندی جان سرل، دو تفسیر از نظریات هیوبرت دریفوس ارائه شد. به این ترتیب همان گونه که جان سرل، هوش مصنوعی قوی و ضعیف را از هم تمیز داد، ما نیز دو تفسیر از نظریات دریفوس را شرح دادیم که بر مبنای تفسیر نخست؛ بدن انسانی شرط لازم هر برنامهٔ پژوهشی هوش مصنوعی است که آن را تفسیر تمام‌انسانی

معنادار است و صرفاً کارکرد نمی‌تواند با معنا برابر گرفته شود، این تلقی را نمی‌توان از نوع کارکردگرایی صرف دانست.

1. satisfactory rendering.

نامیدیم، (یا به‌تبع سرل می‌توان آن را بدن‌مندی قوی نامید، آن‌گونه که وی هوش مصنوعی نمادی را، هوش مصنوعی قوی نامید) و در تفسیر دوم، تحقق شروط پدیدارشناسانهٔ آگاهی، تحت هر بدن ممکنی شرط امکان هوش مصنوعی قلمداد شده است (که به‌تبع سرل می‌تواند بدن‌مندی ضعیف نامیده شود).

این نکته تأکید شدنی است که بدن‌مندبودن شرط اساسیِ این رویکرد است، اما اینکه بدن لزوماً بدن انسانی باشد منجر به تعارض میان دیدگاه‌های دریفوس و ناچار به قائل شدن به تناقض‌گویی و عدم توجه وی به این مسئله خواهد شد.

۴. ۳. رهیافت هم‌تحققی یک‌سویه[1]

رهیافت سوم با عطف به مفهوم هم‌تحققی توضیح داده می شود. در این رهیافت مقصود این است که آگاهی و هوشمندی با بروز رفتارهای هوشمندانه و جزئی هم‌تحقق است. این رویکرد، تأکید بر این اصل دارد که وقتی دو چیز حتی اگر هم‌مصداق نباشند، در صورتی‌که تحقق آن‌ها با هم باشد، با هم اتفاق می‌افتند و رخداد آن‌ها هم‌زمان است. مثالی که کواین[2] دراین‌باره می‌زند، مثال خرگوش و عضو جدانشده از بدن خرگوش مثلاً دست خرگوش است؛ بدین صورت که هرگاه خرگوشی از مکانی عبور کند، لزوماً عضو جدانشده از خرگوش نیز از آن مکان رد می شود. خرگوش و عضو جداشده خرگوش هم‌مصداق نیستند اما هم‌تحقق هستند.[3] بااین‌حال، هرجا خرگوشی باشد، ضرورتاً عضوی جداشده از خرگوش هم خواهد بود. اگر مفهوم خرگوش تحقق پیدا کند، ضرورتاً مفهوم عضوی جدانشده از خرگوش هم تحقق پیدا می‌کند (مسلین، ۵۴). دو مفهوم هوشمندی و رفتار هوشمندانه که نشان‌دهندهٔ تعامل با محیط اطراف و درنتیجه اشتراک در فهم جهان دارد، هم‌مصداق و هم‌مفهوم نیستند؛ اما لزوماً هم‌تحقق هستند و با هم رخ می‌دهند؛ یعنی چنانچه رفتار هوشمندانه تحقق پیدا کند، ضرورتاً هوشمندی نیز تحقق پیدا خواهد کرد. عبارت یک سویه ازاین‌رو است که نشان می‌دهد عکس این عبارت لزوماً صادق نیست و بایستی این رخداد را یک‌سویه در نظر گرفت. به توصیفات دریفوس در بخش سوم همین مقاله دربارهٔ الگوی نورونی فریمن بنگرید. سازوکارهای مغزی در هر تجربه و فعالیت ما حضور دارند. این سازوکارها مفهوماً با فعالیت‌های متناظر خود یکی نیستند؛ اما ضرورتاً با آن‌ها هم‌تحقق هستند. هرجا فعالیتی هوشمندانه و ماهرانه وجود داشته باشد، ضرورتاً هوشمندی نیز وجود دارد. این امر خصوصاً به‌نحوی است که بتواند در موقعیت‌های جزئی به شکل ماهرانه ای عمل کند. از این توضیحات می‌توان چنین نتیجه گرفت که این رویکرد، هم هوش مصنوعی هیدگری و شروط دوگانهٔ آن را در نظرِ دریفوس ارضا می‌کند و هم به دام نمادگرایی و پیوندگرایی نمی‌افتد؛ از آن روی که شرط

1. unilateral Co-Instantiation.

2. **Quine, Willard Van Orman (1908-2000).**

۳. برای بیانی کافی‌تر از مفهوم هم‌تحققی، شیء انضمامی را در نظر بگیرید که به‌عنوان مجموعه‌ای از کیفیات و ویژگی‌ها لحاظ شده است که همگی در رابطهٔ ممکن اساسی قرار دارند که این ویژگی‌ها با یکدیگر هم‌تحقق هستند؛ بدین معنا که هنگامی که شیء محقق شود، تمامی دیگر ویژگی‌ها نیز محقق می‌شوند و این ویژگی‌ها با یکدیگر هم‌مفهوم نیستند؛ اما با هم تحقق پیدا می‌کنند.

هم‌تحققی، رفتاری است که بتواند در درک ما از جهان سهیم باشد. از این نظر چنانچه سیستمی بتواند این شروط را ارضا کند، با مفاهیم آگاهی و هوشمندی هم‌تحقق خواهد بود، اما ماشین‌هایی که در رهیافت نمادگرایی و پیوندگرایی طراحی شده بودند، تعامل هوشمند و معناداری با محیط اطراف نداشتند. نمادگرایی، هرگونه معنا[1] را در رفتار کنار گذاشته بود، اما در هوش مصنوعیِ هیدگری بر بسترمندی و موقعیت‌مندی و همچنین بر کل‌گرایی آگاهی تأکید شده است و شرط بدن‌مندیِ شناخت از آن روی در آن لحاظ شده که نقش بدن در تعامل پویا با جهان، یعنی تعاملی که معنادار باشد و بتواند در موقعیت‌های جزئی، رفتارهای مناسب از خود نشان دهد بسیار حائز اهمیت است و از بازگشت به پارادیم‌های نمادی و پیوندگرایی جلوگیری می‌کند. با این اوصاف، بدن‌مندی انسانی و غیرانسانی هر دو می‌توانند با هوشمندی هم‌تحقق باشند؛ اگرچه هیچ‌یک با آن هم‌مصداق و هم‌مفهوم نیستند. با تکیه بر مفهوم هم تحققی نیز می‌توان خوانشی از دریفوس ارائه داد که نشان‌دهندهٔ سازگاری نظریات وی در نقد هوش مصنوعی کلاسیک و پذیرش هوش مصنوعی هیدگری باشد.

نتیجه‌گیری

در این مقاله کوشیدیم به‌نحوی اجمالی نقدهای دریفوس بر هوش مصنوعی نمادگرا و پیوندگرا و سپس نظر دریفوس را دربارهٔ پذیرش جریان هوش مصنوعی هیدگری که ذیل پارادایم حیات مصنوعی قرار دارد، توضیح دهیم. گفتیم که دریفوس این جریان را با لحاظ شروطی ویژه می‌پذیرد. سپس بررسی کردیم که آیا امکان ارائهٔ خوانشی سازگار از نظریات پیشین و اخیر دریفوس؛ یعنی نقد هوش مصنوعی کلاسیک و پذیرش هوش مصنوعی هیدگری وجود دارد یا اینکه بایستی چنین تلقی کرد که در نظریات وی تناقض و ناهمخوانی وجود دارد؟ درنهایت با ارائهٔ سه رهیافت نشان دادیم که امکان هوش مصنوعی هیدگری بر طبق مبانی دریفوس در خوانش وی از هیدگر و همچنین در نقدهای او از هوش مصنوعی کلاسیک وجود دارد و بایستی این رأیِ واپسین وی را در ادامهٔ سنت فکری پیشین وی دانست. رویکرد نخست، بر نقش تعامل‌گرایی با محیط تأکید دارد و مفهوم آگاهی را در سیستم‌های فیزیکی تشریح می‌کند و نشان می‌دهد که چگونه سیستم‌های فیزیکی از خود رفتارهای هوشمندانه بروز می‌دهند. رهیافت دوم، شرط بدن‌مندی انسانی را ذیل بدن‌مندی قوی مطرح می‌کند و می‌گوید که نظریات دریفوس بر بدن‌مندی لزوماً انسانی تأکید ندارد، بلکه مطلق بدن‌مندی را مدنظر خویش دارد و ویژگی‌های شناخت و آگاهی در هر بدنی که تحقق یابد، هم شروط دوگانهٔ هوش مصنوعی هیدگری را ارضا می‌کند و هم مطابق با نظریات پیشین وی است. رویکرد سوم نیز با استفاده از مفهوم هم‌تحققی می‌گوید که آگاهی و هوشمندی با رفتار و تعامل هوشمندانه هم مصداق و هم‌مفهوم نیستند؛ اما الزاماً هم‌تحقق هستند. صرف تحقق این شروط در هر سیستم هوشمندی، امکان

1. Meaning.

هوشمندی ازنظرِ دریفوس را فراهم ساخته است و رویکرد وی در پذیرش هوش مصنوعی هیدگری، رویکردی سازگار با مواضع پیشین وی در نقد پارادایم‌های نمادی و پیوندگرا و نشان‌دهندهٔ امکان هوش مصنوعی هیدگری بر مبنای نظریات دریفوس است.

منابع

سرل، جان، راز آگاهی، ترجمهٔ: رضا امیر رحیمی، تهران: انتشارات مهرویستا، ۱۳۹۵.

طهماسبی، محمدرضا، «رهیافت‌های بنیادین فلسفی در هوش مصنوعی»، حکمت و فلسفه، تهران،، س۲، ش۲، ۱۳۸۵، صص ۲۵ تا ٤٨.

محمدعلی‌خلج، محمدحسین، «دریفوس و تاریخ فلسفی هوش مصنوعی»، غرب‌شناسی غرب‌شناسی بنیادی، تهران، ش۱، ۱۳۹۳، صص ۱۰۳ تا ۱۲۸.

مسلین، کیت، فلسفهٔ ذهن، ترجمهٔ جمه: مهدی ذاکری، تهران: انتشارات علمی و فرهنگی، ۱۳۹۵.

Bellman،R. E. ،*An Introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser Publishing Company،1978.

Bickhard،M. ،*Part II: Applications of Process-Based Theories: Process and Emergence: Normative Function and Representation*. Axiomathes 14(1)،121-155،2004.

Clark،A. ،*Philosophical Foundations in Artificial Intelligence*،Margaret A. Boden (ed.)،Academic Press،INC،1996.

Dreyfus،H. ،*What Computers Can't Do: A Critique of Artificial Reason*،Harper and Row،1979.

Dreyfus،H. and Dreyfus،Stuart E،'Making a Mind versus Modeling the Brain: Artificial Intelligence Back at a Branch-point' *the philosophy of artificial intelligence*،ed. Margaret Boden،Oxford: Oxford University Press،pp. 309-333،1988.

Dreyfus،H،*What computers still can't do: A critique of artificial reason*،Cambridge،MA: MIT Press،1992.

Dreyfus،H and Dennett،D. ،Mechanical Bodies،Computational Minds،*A Bradford Book*،The M. I . T . Press،1997.

Dreyfus،H. ،"Overcoming the Myth of the Mental: How Philosophers Can Profit from the Phenomenology of Everyday Expertise"،*APA Pacific Division Presidential Address*،2005.

Dreyfus H (2007). 'Why Heideggerian AI Failed and how Fixing it would Require making it more Heideggerian', available in: http://leidlmair.at/doc/WhyHeideggerianAIFailed.pdf.

Graham،George،*Philosophy of Mind: An Introduction*. Oxford and Cambridge: Blackwell،1993.

Hebb،D. O. ،*The organization of behavior: a neuropsychological theory*،John Wiley & Sons،1949.

Haugeland،J. ،*Artificial Intelligence: The Very Idea*،Cambridge،MA: MIT Press،1985.

Lowe،E. J . ،*An Introduction to the Philosophy of Mind United Kingdom*: Cambridge University Press،2001.

Merleau-Ponty،Maurice،*Signs*،translated by R. Mc Cleary. Evanston Ill. :Northwestern University Press،1964.

Russell،S and Norvig،P،*Artificial Intelligence: A Modern Approach*،Pearson Education International،2003.

Susser،D. ،2013،`Artificial Intelligence and the Body: Dreyfus،Bickhard،and the Future of AI'،Routledge،New York (1999)،2013.

Turing،A. M. ،`Computing Machinery and Intelligence'،*Mind*،59،1950.