چیستی اتاق چینی

- الف) فلسفه تحلیلی سرل و فلسفه قاره‌ای دریفوس در نقد آگاهی هوش مصنوعی
- ب) تفکیک فهم انسانی از قواعد کار با نمادها در اتاق چینی
- ج) آشنایی کودک با ارائه جواب در جدول ضرب؛ نظیری برای اتاق چینی
- د) جداسازی مرز بین نرم‌افزار و سخت‌افزار در اتاق چینی

الف) فلسفه تحلیلی سرل و فلسفه قاره‌ای دریفوس در نقد آگاهی هوش مصنوعی

خب علوم شناختی چیست؟ امروز این را توضیح می‌دهم و بعد توضیح مختصری از اتاق چینی می‌دهم. بعد به جایگاهی که این علم دارد را بیان می‌کنیم. این اتاق چینی بحث‌های خیلی خوبی در هوش مصنوعی دارد؛ حدود چهل و سه سالش هم شده؛ از ۱۹۸۰ این بحث در فضای علمی آمده. در طول این چهل و سه سال هم از بحث‌های علمی خیلی داغ بوده. خیلی هم حمله کرده‌اند تا این اتاق چینی را بشکنند. حالا این‌که شکسته‌ای نشکسته، به آن می‌رسیم. روی آن بحث‌های مفصلی شده و شکستن آن هم کار دارد و در آن نکات مهمی هست. در سایت حوزه هم مقاله‌ای از فصلنامه ذهن هست که حدود سال ۱۳۸۷ چاپ شده است. در [فدکی‌ه] آن را گذاشته‌ام. در این مقاله، ایشان خلاصه‌ای از مقاله جان سرل - که مبدع اتاق چینی است - را آورده است. اول آن را توضیح می‌دهد و بعد شش اشکالی که به آن شده را پاسخ می‌دهد. در ششمین اشکال، رنگ پاسخش رنگ تسلیم است. یعنی خود صاحب اتاق چینی در پاسخ به اشکال ششم گویا می‌خواهد تسلیم شود. حالا ببینیم آیا جا بوده که تسلیم شود و چرا میل کرده و می‌گوید اگر شد، بله؟

خب اتاق چینی چیست؟ اولی که این بحث پیدا شد، صحبت سر این بود که وقتی شما با هوش مصنوعی برنامه‌های مینویسید و کامپیوتر، کاری انجام می‌دهد، آیا این دستگاه ذهن و آگاهی دارد یا نه؟ فهم دارد یا ندارد؟ فهم معنا دارد یا ندارد؟ مهندسین نرم‌افزار بیشتر می‌خواهند کار خودشان را پیش ببرند و کاری با این بحث‌هایی که ریخت فلسفی دارند، ندارند و کار خودشان را می‌بینند. آن‌ها با الفاظی که به کار می‌برند، بحث‌های کسانی که صبغه فلسفی داشت را برانگیخت. مثلاً ریخت فکر خود سرل، ریخت فلسفه تحلیلی است. شخص دیگری داریم که همانند سرل در مبارزه با هوش مصنوعی معروف هستند؛ البته سرل مبارزه نکرد ولی تحلیل‌هایش در مقابل قرار گرفت. دیگری مبارزه کرد؛هوبرت دریفوس؛ ریخت فکر دریفوس ریخت فلسفه قاره‌ای است. قاره‌ای فکر می‌کند، نه تحلیلی. و لذا استدلالاتی که دارند دو جور است. ان شاء الله بعداً روی آن‌ها تأمل کنید. اگر با بحث‌های هر کدام از این‌ها آشنا شوید، خیلی خوب است.

جریاناتی هم دارند؛ مثلاً سرل این مقاله را در سال ۱۹۸۰ شروع کرد؛ اما دریفوس سه-چهار سال بعد از پیدایش هوش مصنوعی و کنفرانس‌هایی که تشکیل شده بود، علیه آن‌ها حرف زد. تند حرف زد. حتی به جایی رسید که او گفت یک بچه ده ساله بیشتر از هوش مصنوعی می‌فهمد. یک بچه ده ساله می‌تواند این ماشین را شکست دهد. این حرف او، خیلی به مهندسین و کسانی که کار می‌کردند، برخورد. گفتند حالا حساب تو را می‌رسیم تا از این حرف‌ها نزنی. یکی آمد برنامه‌های ریخت؛ گفت شما آقای دریفوس بیایید با این برنامه هوش مصنوعی ما مسابقه شطرنج بدهید.

مثل منی نمیدانم چیست. ایشان هم قبول کرد و رفت و جلو چشم همگان شکست خورد. گفتند چطور تو میگویی اندازه بچه ده ساله نمیفهمد اما حالا خودت با این ادعا، از او باختی؟! داریوس جملهای گفت: گفت بچه ده ساله میتواند هوش مصنوعی را شکست دهد و بیش از آن است، هرچند این هوش میتواند داریوس را در شطرنج ببرد و او را شکست بدهد.

شاگرد: حالا که باخته این حرف را میزند؟

استاد: نه، حرف بعدی او هم درست بود. یعنی میخواهد بگوید ذهن بچه ده سالهٔ خیلی چیزها دارد که ممکن نیست شما در ماشین پیاده کنید. منظورش این است. ولی خب شما در بخشی از کار، برنامههای نوشتهایدی که من را شکست دادید. بعد هم میگوید من نگفتم ماشین نمیتواند شطرنج بازی کند. من گفتم این نمیتواند ذهنی شود که ما داریم. این خیلی فرق میکند. البته الآن هم در اخبار میآیدی که قهرمان ها را شکست میدهد.

شاگرد: یکی از شطرنج بازها با کامپیوتر بازی کرده بود و آن را شکست داده بود. گفت چون من با آن آماتور بازی کردم آن را بردم اما برنامهریزی آن بهصورت حرفهای بود.

استاد: اینها مطالب مهمی است. اگر همه اینها جمع شود، میبینید.

علی ای حال ریخت کار داریوس قارهای است. باید سر جایش صحبت کنیم که او چه میگوید. ولی کار سرل و اتاق چینی، به این صورت نبود.

شاگرد: قارهای به چه معنا است؟

استاد: قارهای اصطلاحی برای تقسیمبندی فلسفه غرب است. میگویند فلسفه غرب دو نوع تفکر دارد. تفکرات تحلیلی که خودشان میگویند برای آن تعریفی ندارید. درحالیکه روشن هم هست. دوم تفکرات قارهای مانند هگل و ... . واقعش هم این است که قارهایها در عقلانیت بالاتر هستند. شبیه همین کامپیوتر است که شما زبانهای برنامهنویسی سطح بالا دارید که چقدر گستردهٔ است و زبان ماشین هم دارید که در آن سطح پایین با آن سختافزار مباشرتاً کار میکند. قبلاً هم در مبحث عرض کردهام که به گمانم قارهای و تحلیلی مکمل هم هستند. یعنی در بشر به هر دو نیاز است. لذا یک دیگر را در پیشرفت کمک میکنند. اینطور هم نیست که چون فلسفه تحلیلی است، بگوییم قرار است کار گیر کند، بلکه آستانههای که الآن فلسفه تحلیلی رسیده است، جوری است که قبول متافیزیک و چیزهایی که وراء مادهٔ است برای متخصصینش مثل آب خوردن است و حال آنکه قبلاً اینطور فکر نمیکردند.

ب) تفکیک فهم انسانی از قواعد کار با نمادها در اتاق چینی

خب اتاق چینی را عرض کنم. این آقا حدود چهل و سه سال پیش چه گفت؟ برای اینکه نشان دهد - هرچقدر ماشین، دیگری را شکست بدهد و کارهای مهمی انجام دهد- این ماشین فهم و ذهن ندارد، اتاق چینی را طراحی کرد. اتاق چینی چیست؟ من قبلش مثال‌هایی طلبگی عرض می‌کنم. اول یک اتاق فارسی درست می‌کنم تا به اتاق چینی برسیم.

شما به‌صورت بسیار ساده کسی را که اصلاً فارسی بلد نیست می‌آوری، به او حروف یا کلمات فارسی می‌دهی. البته در مقصود من فعلاً کلمات آسان‌تر است. او اصلاً نمی‌داند این کلمه به چه معنا است. بعد به او یک قاعده می‌دهی و می‌گویی وقتی این شد، این کار را بکن. کسی که خبر ندارد، وقتی با او مکاتبه کند به خیالش می‌رسد که او کاملاً فارسی می‌داند.

کسی را بیاوری و برایش بنویسد «چطوری؟». این کلمه‌ای ساده است. ساده‌ترین اتاق فارسی است که خدمت شما می‌گویم. در اصطلاح فنی به این «چطوری» می‌گویی‌م یک نماد است؛ Symbol است؛ علامت است؛ «sign» است. او هم اصلاً نمی‌داند این کلمه چیست. فوقش اگر حروف را بلد باشد می‌بیند چهار حرف است. اگر هم حروف را ندانند، این ترکیب را به‌عنوان یک نماد می‌شناسد. به او می‌گویی به این نماد نگاه کن. بعد نماد دیگری را به او نشان می‌دهی؛ «خوب». این هم چهار حرف است. یا نمادهای حروف را به او داده‌ای خود کلمه را. من چون می‌خواهم ساده‌تر باشد، خود «خوب» را به‌عنوان یک نماد در نظر بگیری‌م. به او می‌گویی‌م در دستت باشد. بعد با یک وسیله ارتباطی یا پشت دیواری که او را نمی‌بینند، این را مرا‌سله می‌کنی‌د. به او می‌گویی‌د اگر برای تو «چطوری» را فرستاد، به ازاء آن «خوب» را بفرست. این یک قاعده است. در اصطلاح برنامه‌نویسی این قاعده است. پس دو نماد داری‌م و یک قاعده داری‌م. می‌گویی‌م وقتی «چطوری» آمد، نماد «خوب» را بفرست. او هم که اصلاً فارسی بلد نیست و سر در نمی‌آورد فارسی چیست. ولی این قاعده را بلد است که وقتی نماد «چطوری» آمد، به ازاء آن «خوب» را بفرستد. کسی که آن طرف است، می‌گوید او از فارسی سر در می‌آورد. و حال آنکه او فارسی نمی فهمد. او فقط طبق یک قاعده بسیار ساده بین دو شکل، بین دو نماد، رابطه برقرار می‌کند. وقتی به او می‌گویی‌م «چطوری»، «خوب» را جواب می‌دهد.

اتاق چینی هم همینطور است. او اتاقی را فرض گرفت؛ کسی که زبان چینی بلد نیست به این اتاق می‌رود. چند چینی زبان هم هستند که به او قواعدی را طبق زبان و نماد چینی یاد می‌دهند. می‌گویند وقتی اینطور چیزی آمد، تو اینطور کاری را انجام بده. همین قاعده‌ای را که الآن مثال زدیم. اولش هم که شروع کرد یک گروهی بودند...؛ می‌دانید در دانشگاه مرتب پروژه تعریف می‌کنند تا روی آن تحقیق کنید. خیلی گسترده است. یک پروژه‌ای تعریف کردند برای فهم داستان توسط ماشین. فهم داستان به چه معنا است؟ وقتی حتی برای بچه‌های چهار-پنج‌ساله داستان می‌گویی‌د، یک چیزهایی را در داستان می‌گویی‌د، اما به‌خاطر چیزهایی که می‌دانند اطلاعات دیگری را ولو در متن

داستان نیامده، میفهمد. این پروژه این بود که کاری کنیم که مثل انسانها که داستان را میفهمند و اگر از آنها سؤالی کنید که صریحاً در متن داستان نیامده، میتوانند پاسخ بدهند، ماشین هم همینطور باشد. یعنی یک چیزهایی به او میدهید و چیزهای دیگری هم در میان آنچه به او دادهایم وجود ندارد، اما اگر از او سؤال کنید همانها را که به او ندادهایم، پاسخ میدهد. ایشان همین را گرفت و اتاق چینی را درست کرد.

شاگرد: اینکه میتواند به سؤال خارج از داستان پاسخ بدهد، امارهای بر هوش است. اما اینکه یک ماشین به سؤال خارج از داستان پاسخ بدهد نمیتوانیم بگوییم دارای هوش است. چون صرفاً اماره است.

استاد: سرل همین را میخواهد بگوید. پروژه آنها را گرفت و اتاق چینی را طراحی کرد تا بگوید ولو پروژه شما سر برسد و بتوانید فهم داستان را به ماشین القاء کنید و جواب بدهد، درعینحال هیچ ذهنی ندارد. هوش قوی ندارد. هوش ضعیف را دارد. او اصلاً میخواهد همین را بگوید. بعد شش اشکال به او کردند و در اشکال ششم نرم میشود.

او میگوید کسی که فقط زبان انگلیسی بلد است و با زبان انگلیسی قواعد را به او میگوییم، اما از چینی هیچ چیزی بلد نیست و فقط نمادهای چینی دست او است، ولی قواعد را بهخوبی بلد است. میگوید وقتی شما از او سؤال کنید...؛ یک عده اطلاعات به او میدهید که حروف الفبا چینی است. یک عده از اطلاعات به او میدهید که مربوط به محتوای داستان است. یک قواعدی هم به او میدهید، چهارمین دسته از اطلاعات هم پاسخگویی است که یعنی طبق قاعده به او میگویید وقتی به تو اینها را گفتند، تو اینها را بگو. چیزهایی که او میدهد اصلاً در متن داستان نبود اما طبق قواعد میدانند باید چه چیزی بدهد. باید اطلاعات را هم به او بدهند.

مثلاً یک ماشینی باشد؛ به او بگویید یک خواهر و برادر بودند؛ پدر و مادرش تنها همین دو بچه را داشتند. این خواهر ازدواج کرد و شوهرش فلانی بود. شوهر این خواهر یک روز به بازار رفت و در بازار باجناقش را دید! اگر ماشینی باشد که فهم داستان و هوشمندی را داشته باشد، فوری میگوید او که باجناق ندارد. میگویید از کجا فهمید؟! اطلاعات را طوری به او دادید که وقتی میگویید یک خواهر و برادر هستند، باجناق یعنی کسی که شوهر خواهر او است و حال اینکه او خواهر ندارد. میگویید من که الان اسمی از باجناق به او ندادم! ولی درواقع در اطلاعات ضمیمه، به او دادهاید.

شاگرد: اینکه حقیقت هوش چیست روشن است؟

استاد: مفصل کار شده است. اشکال ششم همین است. اشکال ششم که ایشان میگوید شاید اگر شد، همین است.

خب حالا اتاق چینی چه میکند؟ میگوید این سر در نمیآورد. من مثال دیگری به شما عرض میکنم. چرا ایشان اتاق چینی را آورد؟ چون ریخت بحثهای آنها در دانشگاه، کاملاً روش علمی به اصطلاح خودشان است؛ (Scientific) است. لذا کاری کرد که خودش بهعنوان انگلیسیدان، مشتمل بر دو چیز بود؛ یعنی هم فهم انسانی را داشت و زبان انگلیسی را میفهمید، هم میتوانست نمادها و قواعد را بفهمد و با آنها کار انجام بدهد. یعنی هم حیث ماشین در او بود و هم حیث ذهن. لذا این دو حیث را از هم جدا کرد. زبان چینی را آورد که فهمی از آن ندارد، اما میتواند مثل ماشین با نمادهای زبان چینی بازی کند. چرا علمی شد؟ بهخاطر اینکه دقیقاً نشان میدهد؛ میگوید من در اینجا انگلیسی را میفهمم اما از چینی سردرنمیآورم. اما همینجا یک چینیای که در کنارم نشسته یک چیزی اضافه بر من دارد. عیناً داریم یک کار را میکنیم و جوابها هم عین هم است. ولی او میفهمد چون زبان چینی را بلد است اما من چون زبان چینی را بلد نیستم نمیفهمم که آنها چه میگویند و من چه جوابی به آنها میدهم؛ اصلاً خبر ندارم. با این کار کاملاً نشان داد.

ج) آشنایی کودک با ارائه جواب در جدول ضرب؛ نظیری برای اتاق چینی

من میخواهم مثالهایی را عرض کنم که ذهن شما جلو برود. جدول ضرب را در نظر بگیرید. به منزل که رفتید به بچهای که سه سالهٔ است، اصالً به مدرسه نرفته و نمیداند عدد چیست و حروف را نمیداند، بگویید؛ جدول ضرب را به او نشان بدهید و بگویید وقتی عددی از این ردیف و از آن ستون به همراه علامت ضرب به تو دادند، انگشتت را روی عدد این ردیف با عدد روی آن ستون بگذار و با خط مستقیمی ادامه بده؛ صاف به یک خانه میرسی؛ آن را به عنوان جواب، نشان بده. این هم یک نوع اتاق چینی است. یعنی بچه نه عدد میشناسد و نه ضرب را، ولی بازی کردن با نمادها را به او یاد دادهایم. قواعد را به او یاد دادهایم. به او میگویید هشت ضرب در هفت؛ خب نماد هشت را میبینند، نماد هفت را هم میبینند، از هفت و هشت مستقیم برو، به یک خانهای میرسی، همین خانه را جواب بده؛ پنجاه و شش را نشان بده. هر کسی به این بچه بگوید هفت ضرب در هشت، فوری نگاه میکند و جواب میدهد پنجاه و شش. میگویند عجب بچهای است! هنوز چهار سالش است اما جدول ضرب را بلد است. و حال آنکه این بچه نه نماد عدد را بلد است، نه اصالً عمل ضرب را بلد است، بلکه او تنها بلد است که یک کار انجام بدهد. بعداً در بخش نرمافزار میخواهم تفاوتش را عرض کنم.

د) جداسازی مرز بین نرم‌افزار و سخت‌افزار در اتاق چینی

قبل از اینکه آن را بگویم، نکته‌ای می‌گویم تا به مقصود اصلی برسم. آن چه که در فضای این اتاق چینی مهم است، این است: این مثال به‌طور خیلی دقیق، مرز بین نرم‌افزار و سخت‌افزار را جدا می‌کند. قبلاً در مباحثات داشته‌ایم؛ ما در فضای طلبگی می‌گوییم طبیعت و فرد. یعنی این مثال‌ها به‌صورت بسیار کاربردی -که حالت مسامحه را از ذهن بشر می‌گیرد- این‌ها را کف دستش می‌گذارد. و لذا خیلی از کسانی که به او ایراد گرفتند هنوز بطور دقیق، مرز بین نرم‌افزار و سخت‌افزار برایشان جدا نبوده است. یکی از جواب‌هایی که او می‌دهد همین است. نکته‌ای دقیق است. شما هم اگر توجه کنید فوائد بسیاری خوبی می‌برید. من هم اگر فرصت شد، عرض می‌کنم.