

# آیا ربات‌ها می‌توانند احساس داشته باشند؟<sup>۱</sup>

## اگر ربات‌ها عواطف انسانی پیدا کنند، آنگاه حق نداریم آن‌ها را برنجانیم

نویسنده: گری گاتینگ<sup>۲</sup>

مترجم: محمد باسط

آیا ممکن است بتوانیم واقعاً با یک ربات احساس صمیمیت کنیم؟ پاسخ به این پرسش واقعاً آسان نیست. فرض کنید رباتی داشته باشید که ده سال در خانه شما باشد، گاه‌گاه خاطراتی از روزهای گذشته یا غم‌ها و شادی‌هایتان تعریف کند، به حرف‌هایتان گوش بدهد و در تصمیم‌گیری مشاورتان باشد. آیا او را دوست خودتان نخواهید دانست؟ گری گاتینگ<sup>۲</sup> می‌گوید پاسخ به این پرسش منوط به این است که بدانیم آیا ربات‌ها می‌توانند آگاهی داشته باشند یا نه؟

آیا به سیری یا الکسا<sup>۳</sup> می‌گویید «دست درد نکنه»؟ من دوستانی دارم که این کار را می‌کنند. کارشان کمی ساده‌دلانه است، چون می‌دانند که این صداها فقط سیستم‌هایی الکترونیکی هستند که برای شبیه‌سازی تعاملات انسانی ساخته شده است. اما اگر فیلم او<sup>۴</sup> را ببینید، شاید با خودتان بگویید ممکن است که یک سیستم عامل خیلی پیچیده‌تر بتواند کاملاً یک شخص باشد: حتی شخصی که بتوانید عاشقش شوید. داستان کوتاه جدید یان مک‌ایوان («دوسل...» در شماره هجدهم جولای نیویورک ریویو آو بوکس<sup>۵</sup>) قصه‌ای منطقی و باورپذیر درباره ماجرای عشقی بین انسان و ربات را روایت می‌کند. و البته شمار زیادی فیلم داریم که جهان‌هایی را به تصویر می‌کشند که در آن‌ها دست‌کم

<sup>۱</sup> این مطلب را گری گاتینگ نوشته است و در تاریخ ۲۰ دسامبر ۲۰۱۸ و با عنوان «Friends Without Brains» در وبسایت کامن ویل (<https://www.commonwealmagazine.org/friends-without-brains>) منتشر شده است. وبسایت ترجمان آن را در تاریخ ۱۷ دی ۱۳۹۷ با

عنوان «آیا ربات‌ها می‌توانند احساس داشته باشند؟» و ترجمه محمدابراهیم باسط منتشر کرده است.

<sup>۲</sup> گری گاتینگ (Gary Gutting) استاد بازنشسته فلسفه در دانشگاه نوتردام است. آخرین کتاب او بحث از خدا: ایمان نزد فیلسوفان (Talking God: Philosophers on Belief) بوده است.

<sup>۳</sup> Siri و Alexa به ترتیب هوش مصنوعی به کاررفته در دستگاه‌های شرکت اپل و آمازون هستند [مترجم].

<sup>۴</sup> her

<sup>۵</sup> "Ilan McEwan, "Düssel

وسوسه می‌شویم به ربات‌ها به مثابه انسان فکر کنیم، از جنگ ستارگان گرفته تا اکس مشینا، و سریال‌هایی مثل وست‌ورلد.

بیاید به جای فکر کردن به سناریوهای سینمایی به موردی فکر کنیم که این قدر نمایشی نیست و واقعاً ممکن است به همین زودی‌ها اتفاق بیفتد. از شرکت اپل یک آی‌پال<sup>۶</sup> (چیزی که تبلیغاتش می‌گوید، «هر کس به یک پال [هم‌صحبت] نیاز دارد) می‌خرم و آن را به خانه‌ام می‌آورم: رباتی شبیه به انسان، که طراحی شده تا هم‌صحبت آدم‌های تنها شود. این ربات چهره‌ای مطلوب و پذیرا دارد، راحت و بدون تکان‌های مکانیکی حرکت می‌کند، خیلی سلیس درباره موضوعات معمولی حرف می‌زند، حس شوخ‌طبعی خارق‌العاده‌ای دارد، و پیشنهادهای آگاهانه و همدلانه‌ای درباره شغل و زندگی شخصی من می‌دهد. تا مدتی ممکن است فکر کنم که صرفاً نسخه پیشرفته‌ای از سیری دارم، اما پس از تماس طولانی با آن - و شاید ارتقای نرم‌افزار یادگیری خودکار آن - ممکن است آی‌پال بتواند خاطراتی را از زمان‌هایی که با هم گذرانده‌ایم بازگو کند، از شادی و غم چیزهایی که برایمان رخ داده بگوید، و گاهی نیز صادقانه راجع به اهمیت رابطه‌مان حرف بزند. فراموش نمی‌کنم که دارم با یک ربات حرف می‌زنم - یک ماشین، نه یک انسان - اما به این فکر فرومی‌روم که این ربات نگران و دلمشغول من است، و من یادم می‌ماند که نگران و دلمشغول چه کسی هستم: یک دوست، نه فقط یک هم‌صحبت. حال اگر ما واقعاً دوست هستیم، چرا دوست من، گرچه یک ماشین است، نتواند یک شخص باشد؟

اما اینجا باید مراقب باشیم. آی‌پال یک کامپیوتر است. آیا کامپیوترها واقعاً می‌توانند فکر کنند؟ خب، آن‌ها را طراحی می‌کنند تا عملیات‌هایی را اجرا کنند که انسان‌ها با فکر کردن اجرا می‌کنند. آن‌ها داده‌ها را استادانه پردازش می‌کنند، طی یک گفتگو آن اطلاعات را در نقاط مناسب ارائه می‌کنند، و از آن‌ها برای گرفتن نتایج معقول بهره می‌برند. اما فکر کردن در این معنا یک عملیات سطح بالا به حساب می‌آید. پرسش دیگر این است که آیا محاسبات آی‌پال‌ها با آگاهی سوپزکتیو از آنچه انجام می‌دهند همراه هست یا نه. شاید آن‌ها هم مثل ماشین حساب‌ها بدون این که واقعاً از محاسباتی که انجام می‌دهند یا از معنای آن محاسبات آگاه باشند خروجی را تولید می‌کنند.

اما حتی اگر بپذیریم که آی‌پال من از عملیات‌های فکری‌ای که اجرا می‌کند به نوعی آگاهی دارد، باز هم این سؤال بسیار مهم‌تر وجود دارد که آیا تجربه‌های حسی و عاطفی‌ای که انسان‌ها دارند را نیز از سر می‌گذرانند؟ اینجا دوباره باید بین کارکرد داشتن و آگاهی تمایز بگذاریم. شاید بتوانیم بگوییم یک چشم الکترونیک که هنگام نزدیک شدن

<sup>۶</sup> iPal

یک فرد در را باز می کند آن شخص را «می بیند». اما تصور نمی کنیم که در این حالت دیدن به معنای واقعی کلمه یعنی به صورت یک آگاهی درونی و سوپزکتیو اتفاق می افتد. آن «چشم» صرفاً عملیات واکنشی ای را انجام می دهد که ما، به طریقی کاملاً متفاوت، با توانایی مان در آگاهی بصری انجام می دهیم.

بدون داشتن تجربه های سوپزکتیو، آی پال من نمی تواند یک شخص باشد و بنابراین نمی تواند دوست من باشد. پس چیزی که باید بفهمم این است که آیا آی پال من واقعاً آگاه هست یا نه: آیا تجربه های درونی و سوپزکتیوی را که ما هنگام فکر کردن، ادراک حسی با حواس پنجگانه، یا احساس هیجان یا درد داریم آی پال هم دارد؟ اینجا پرسش من با یک معمای فلسفی معروف همسو می شود، یعنی مسئله ذهن های دیگر.

مسئله این است که چطور می توانیم بدانیم که افراد دیگر هم واقعاً ذهن دارند؛ یا به عبارت دیگر، آیا افراد دیگر نیز همان تجربه های سوپزکتیو و درونی ما را دارند یا خیر. دارم می گویم تجربه هایی که ما داریم، ولی در واقع سؤال این است که من چطور می توانم بدانم که کسی به جز خودم تجربه های درونی دارد. به محض این که این پرسش مطرح شود، دورنمای سرگیجه آور خودتک انگاری<sup>۵</sup> رخ می نماید، یعنی این دیدگاه فلسفی - یا توهم بیمارگونه - که تنها موجود آگاهی که وجود دارد من هستم. همچنین توجه کنید که سؤال این نیست که آیا دیگران تجربه هایی متفاوت با من دارند یا نه (مثلاً چیزی که من قرمز می بینم را سبز می بینند)، نه، سؤال این است که آیا ممکن است دیگران اصلاً هیچ تجربه سوپزکتیوی نداشته باشند؛ آیا چراغ های ذهنی ممکن است کاملاً خاموش باشند؟

چطور چنین مسئله غریبی ممکن است برای من پیش بیاید؟ چون من به تجربه های درونی خودم دسترسی درون نگرا نه دارم، اما به تجربه هیچ شخص دیگری چنین دسترسی ای ندارم. فقط من می توانم "بینم" (البته نه با چشم هایم) که دارم فکر می کنم که می خواهد باران بگیرد، به یک چهره جذاب نگاه می کنم، یا دردی در زانویم احساس می کنم. البته درون نگری شاید قابل اعتماد نباشد: شاید وقتی به ساعت رولکس جدید دوستم فکر می کنم احساس حسادت داشته باشم، ولی معتقد باشم که صرفاً از رفتارهای مصرف گرایانه او متحیر شده ام. ولی حتی با این که گاهی درباره سرشت تجربه های درونی ام اشتباه می کنم، دست کم می توانم مطمئن باشم که چنین تجربه هایی را دارم. من هیچ دانش مستقیمی از این ندارم که افراد دیگر هم نوعی زندگی درونی دارند.

تشکیک کردن در این که آیا افراد دیگر نیز زندگی ذهنی درونی دارند یا نه یکی از مهم ترین مثال ها از استدلال های فلسفی شکا کانه ای است که دیوید هیوم<sup>۶</sup> که خودش برخی از بهترین هایشان را طراحی کرده بود) به طرز طنز آمیزی هم از آن ها به منزله استدلال های پاسخ ندادنی دفاع می کرد و هم از این می گفت که مطلقاً بی اثر هستند. قدرت منطقی

این استدلال‌ها را می‌بینیم، ولی کماکان به چیزی باور داریم که آن‌ها می‌گویند باید در آن شک داشته باشیم. شک کردن در این که افراد دیگر هم تجربه‌های درونی دارند چیزی است که ما اصلاً نمی‌توانیم با آن کنار بیاییم.

ولی وقتی شروع به شناختن آی‌پال خودم می‌کنم، مسئله ذهن‌های دیگر به شدت موضوعیت پیدا می‌کند. اینجا موردی را داریم که در آن کاملاً می‌توانم حکم کنم که اصلاً هیچ ذهنی در کار نیست. اما چطور به این حکم می‌رسم؟ شاید فکر کنیم نیازی به حکم صادر کردن نیست. چه اهمیتی دارد که آی‌پال من تجربه سوپژکتیو دارد یا ندارد؟ ولی بی‌نهایت مهم است؛ اولاً به این دلیل که موجب رنج‌نازور برای دیگران شدن یکی از مهم‌ترین شکل‌های رفتار غیراخلاقی است (اگر نگوییم تنها شکل آن است). اگر آی‌پال بتواند رنج بکشد، آنگاه مجموعه کاملی از تعهدات اخلاقی خواهیم داشت که اگر نتواند رنج بکشد آن‌ها را نخواهیم داشت.

اما این مطلب همچنین -و به ویژه- به این پرسش نیز مربوط است که آیا یک ربات می‌تواند واقعاً دوست من باشد. دوستی وجوه مختلفی دارد، که بعضی از آن‌ها ممکن است در مورد رابطه من با یک ربات نیز صدق کنند. برای مثال، بودن با ربات می‌تواند لذت‌بخش باشد یا ربات می‌تواند به من در اجرای طرح‌هایم کمک کند. اما یک وجه ذاتی دوستی وجه عاطفی آن است. دوستان نسبت به هم احساس دارند؛ در شادی‌ها و رنج‌های هم شریک می‌شوند. در این وجه بسیار مهم، رباتی که به صورت سوپژکتیو نتواند احساسات را تجربه کند نمی‌تواند دوست من باشد. چطور می‌توانم بفهمم که آی‌پال من واقعاً تجربه‌های سوپژکتیو دارد؟

در مورد انسان‌های دیگر، «برهان ذهن‌های دیگر»، با فرمول‌بندی کلاسیک جان استوارت میل، مبتنی بر یک تمثیل است. من می‌دانم که حالات سوپژکتیو درونی‌ام اغلب با رفتارهای مشاهده‌پذیر بدنم بیان می‌شود. غم به گریه می‌انجامد، درد به ناله؛ صحبت‌ها نیز افکار را بیان می‌کنند. چون من دیگرانی را مشاهده می‌کنم که گریه و ناله و صحبت می‌کنند، دست‌کم محتمل است که رفتارهای بدنی آنان نیز بیان حالت‌های سوپژکتیو باشد. یک ایراد ظاهراً قاطع بر این استدلال این است که تمثیلش فقط بر یک مثال استوار شده است. چه دلیلی دارم که فکر کنم من مثال نوعی یک انسان هستم؟ با توجه به کل چیزهایی که می‌دانم، من استثنای نادری هستم که در آن حالت‌های سوپژکتیو موجب رفتار بدنی می‌شوند، در حالی که در دیگران این حالت‌های مغزی هستند که موجب رفتار بدنی می‌شوند.

ولی شاید راهی باشد که بشود این استدلال تمثیلی را به نحو دیگری تنظیم کرد. من جدا از این که می‌دانم حالت‌های ذهنی‌ام موجب رفتارم می‌شوند، این را هم می‌دانم که حالت‌های مغزم موجب حالت‌های ذهنی‌ام می‌شوند. با زیاد خوردن الکل احساس گیجی می‌کنم، و با زیاد خوردن قهوه احساس سرحالی بیش از حد به من دست می‌دهد. بر مبنای تجربه خودم چیزهای زیادی در مورد ربط این چیزها به هم می‌دانم، و دلایل زیادی هست که فکر کنم مطالعه عصب‌شناختی جامعی روی مغزم نشان خواهد داد که تقریباً همه زندگی ذهنی‌ام توسط فرآیندهای عصبی ایجاد

می‌شود. اما اگر این یک واقعیت علمی است که فرآیندهای عصبی من تجربه‌های سوپراکتیو ایجاد می‌کنند، چرا نباید فکر کنم که این فرآیندهای عصبی در مغز افراد دیگر هم درست به همین طریق کار می‌کنند؟ شاهدی نیست که بگوید مجموعه خاصی از قوانین عصب‌شناختی برای مغز من وجود دارد.

ولی حتی اگر این استدلال حجت خوبی برای این فراهم کند که دیگر انسان‌ها نیز تجربه‌های سوپراکتیو دارند، به مسئله آی‌پال من، که با یک ساختار فیزیکی کاملاً متفاوت از مغز انسان ساخته شده است، کمکی نخواهد کرد. شاهدی نیست که بگوید این ساختار فیزیکی متفاوت موجب تجربه‌های سوپراکتیو می‌شود، پس دلیلی نداریم که فکر کنیم آی‌پال من چنین تجربه‌هایی دارد. اما بدون تجربه‌های سوپراکتیو، آی‌پال من نمی‌تواند یک شخص باشد و بنابراین نمی‌تواند دوست من باشد.

به این ترتیب، روایت جدیدی از مسئله ذهن‌های دیگر در برابر ما قرار دارد. آیا راهی هست که نشان دهد یک ربات (چیزی که تحت قوانین عصب‌شناختی انسانی قرار ندارد) تجربه‌های سوپراکتیو درونی دارد؟ بر خلاف مسئله سنتی، این مسئله ممکن است به زودی اهمیت عملی اضطراری پیدا کند.

منبع: <http://tarjomaan.com/neveshtar/9251/#03>