ماشین‌آلات محاسباتی

آلن تورینگ

ترجمه پیمان غلامی

فهرست

ماشین‌آلات محاسبه‌گر و هوش (۱۹۵۰) ۳
ماشین‌آلات هوشمند، نظریه‌ای ارتدادی (۱۹۵۱) ۲۷
آیا کامپیوترهای دیجیتال می‌توانند فکر کنند؟ (۱۹۵۱) ۳۱
آیا می‌شود گفت که ماشین‌های محاسبه‌گر خودکار قابلیت تفکر دارند؟ (۱۹۵۲) ۳۶
منبع ۴۹

ماشین‌آلات محاسبه‌گر و هوش (۱۹۵۰)

یک. بازی تقلید

قصد دارم این پرسش را در نظر بگیرم: «آیا ماشین‌ها می‌توانند فکر کنند؟». این مسئله باید با تعاریف معنی الفاظ «ماشین» و «فکر» آغاز شود. چه‌بسا چارچوبی برای این تعاریف ایجاد شود طوری که تا آنجا که امکان دارد بر استفاده‌ی معمول این کلمات تأمل کرد، اما این نگرش خطرناک است. اگر معنی کلمات «ماشین» و «فکر» را می‌بایست با بررسی نحوه‌ی استفاده‌ی رایج‌شان دریافت آن‌گاه دشوار بتوان از این نتیجه فرار کرد که معنی و پاسخ به پرسش «آیا ماشین‌ها می‌توانند فکر کنند؟» می‌بایست در پیمایشی آماری همچون نظرخواهی جستجو شود. اما این هم پوچ است. به جای تلاش برای ارائه‌ی چنین تعریفی، پرسش مزبور را با پرسش دیگری عوض می‌کنم که رابطه‌ای بسیار نزدیک با آن دارد و در کلماتی نسبتا بی‌ابهام بیان شده است.

صورت جدید مسئله را می‌توان بر حسب یک بازی توصیف کرد که «بازی تقلید» می‌خوانیم. این بازی با سه نفر انجام می‌شود، یک مرد (الف)، یک زن (ب)، و یک مستنطق (پ) که می‌تواند از هر کدام از دو جنس باشد. مستنطق در اتاقی جدا از دو نفر دیگر قرار می‌گیرد. هدف بازی برای مستنطق تعیین این نکته است که کدام یک از این دو مرد است و کدام یک زن. او آن‌ها را با برچسب‌های (ز) و (م) می‌شناسد، و در پایان بازی یا می‌گوید «ز الف است و م ب است» یا می‌گوید «ز ب است و م الف است». مستنطق اجازه دارد پرسش‌ها را به صورت زیر برای الف و ب طرح کند:

پ: ز آیا لطفا طول موی او (زن یا مرد) را به من می‌گویی؟ حال فرض کنید ز عملا الف است، در نتیجه الف باید پاسخ دهد. هدف الف در بازی این است که سعی کند و پ را وادارد که که تشخیص اشتباه بدهد. از این‌رو پاسخ او می‌تواند این‌طور باشد: «موی من توفال‌پوش است، و بلندترین رشته‌هایش تقریبا نه اینچ هستند.»

برای اینکه لحن‌های صدا به مستنطق کمکی نکنند، پاسخ‌ها باید نوشته شوند، یا حتی بهتر اینکه تایپ شوند. چیدمان آرمانی در داشتن یک چاپگر ازراه‌دور است که بین دو اتاق ارتباط برقرار می‌کند. به همین منوال، پرسش و پاسخ‌ها می‌توانند با یک میانجی تکرار شوند. هدف بازی برای بازیگر سوم (ب) کمک به مستنطق است. بهترین استراتژی برای او احتمالا این است که پاسخ‌های صادقانه بدهد. می‌تواند چیزهایی مثل «من زن هستم، به آن مرد گوش نده!» را به پاسخ‌هایش اضافه کند، اما وقتی مرد هم می‌تواند اشارات مشابه‌ای داشته باشد پس این افزوده‌ها به درد نمی‌خورند.

حال این سئوال را می‌پرسیم: «چه اتفاقی خواهد افتاد وقتی یک ماشین نقش الف را در این بازی ایفا کند؟» آیا مستنطق همچون اغلب موارد که بازی این‌طور انجام می‌شود تصمیم اشتباهی خواهد گرفت همان‌طور که وقتی بازی بین یک زن و یک مرد انجام می‌شود چنین می‌کند؟ پرسش اولیه‌مان، «آیا ماشین‌ها می‌توانند فکر کنند؟»، با این پرسش‌ها عوض می‌شود.

دو. نقد مسئله‌ی تازه

ضمن پرسیدن اینکه «چه پاسخی به این شکل تازه‌ی پرسش وجود دارد؟»، می‌توان پرسید «آیا این پرسش تازه ارزش کندوکاو دارد؟». این پرسش دوم را بدون بلوایی اضافی کنکاش خواهیم کرد و این‌گونه پسروی بی‌پایان را متوقف می‌کنیم.

مسئله‌ی تازه این مزیت را دارد که خطی به‌راستی واضح بین ظرفیت‌های جسمانی و فکری یک انسان ترسیم می‌کند. هیچ مهندس یا شیمیدانی ادعا نمی‌کند که بتواند ماده‌ای تمیزناپذیر از پوست انسانی تولید کند. امکان دارد که همزمان این امر انجام شود، اما حتی با فرض مهیابودن این ابداع باید حس کنیم که تلاش برای انسانی‌ترکردن یک «ماشین متفکر» از طریق پوشاندن آن در چنان گوشت مصنوعی اهمیت کمی داشت. مسئله‌مان را به صورتی وضع کرده‌ایم که این واقعیت را در شرایطی منعکس می‌کند که مانع مشاهده یا لمس دیگر رقبا یا شنیدن صدایشان از طرف مستنطق می‌شود. یکی دیگر از مزایای این سنجه‌ی پیشنهادی می‌تواند با پرسش‌ها و پاسخ‌های نمونه‌وار نشان داده شود. از این‌رو:

س: لطفا در مورد موضوع **پل فورث** یک سونات برایم بنویس.

ج: در این زمینه روی من حساب نکن. هرگز نمی‌توانم شعر بنویسم.

س: ۳۴۹۵۷ را به ۷۰۷۶۴ اضافه کن.

ج: (۳۰ ثانیه تاخیر دارد و بعد جواب می‌دهد) ۱۰۵۶۲۱.

س: شطرنج بازی می‌کنی؟

ج: بله.

س: شاه در ردیف ۱ دارم، و نه هیچ مهره‌ی دیگر. تو فقط شاه در ردیف ۶ داری و قلعه در ردیف ۱. نوبت حرکت توست. چطور بازی می‌کنی؟

ج: (پس از ۱۵ ثانیه مکث) قلعه به ردیف ۸ مات.

به نظر می‌رسد روش پرسش و پاسخ برای معرفی تقریبا هر کدام از ساحت‌های کوشش انسانی که می‌خواهیم ذیل این روش بگنجانیم مناسب باشد. مایل نیستیم ماشین را بابت ناتوانی‌اش در درخشیدن در رقابت‌های زیبایی جریمه کنیم، نه حتی یک انسان را بابت کم‌آوردن در برابر یک هواپیما در یک مسابقه. شرایط بازی ما این ناتوانی‌ها را بی‌ربط می‌کند. «شاهدها»، مادامی که از جذبه‌ها، نیرومندی، یا قهرمان‌گرایی‌شان راضی‌اند، می‌توانند لاف بزنند اگر این کار را مفید ببینند، اما مستنطق نمی‌تواند اثبات‌های عملی را از آن‌ها بخواهد.

بازی احتمالا می‌تواند بر این مبنا نقد شود که شانس اصلا با ماشین نیست. اگر انسان می‌بایست تلاش و تظاهر کند که ماشین باشد، آشکارا نمایشی بسیار ضعیف خواهد داشت. ممکن است به خاطر هم کندی و هم بی‌دقتی در حساب شکست بخورد. آیا امکان ندارد که ماشین‌ها چیزی را که می‌بایست تفکر توصیف شود انجام دهند اما با آنچه یک انسان انجام می‌دهد بسیار متفاوت است؟ این اعتراض بسیار قوی است، اما حداقل می‌توانیم بگوییم که اگر در هر صورت یک ماشین بتواند به‌طور راضی‌کننده‌ای برای اجرای بازی تقلید ساخته شود، آن‌وقت این اعتراض نباید مایه‌ی دردسرمان شود.

چه‌بسا اصرار شود که بهترین استراتژی برای ماشین حین انجام «بازی تقلید» احتمالا چیزی غیر از تقلید رفتار یک انسان باشد. می‌تواند این‌طور باشد، اما فکر می‌کنم نامحتمل است که اثر چندانی داشته باشد. در هر صورت، اینجا هیچ قصدی برای کندوکاو در نظریه‌ی بازی نداریم، و فرض گرفته خواهد شد که بهترین استراتژی این است که سعی شود پاسخ‌هایی مهیا گردد که طبیعتا از طرف یک انسان داده شده‌اند.

سه. ماشین‌های دخیل در بازی

پرسشی که در بخش یک طرح کردیم به‌قدر کافی معین نیست اگر مشخص نکنیم منظورمان از کلمه‌ی «ماشین» چیست. طبیعی است که باید استفاده از هر نوع فن مهندسی در ماشین‌هایمان را مجاز بدانیم. همچنین دوست داریم این امکان را روا بدانیم که یک مهندس یا تیمی از مهندسان ماشینی بسازند که کار می‌کند، اما سبک عملکردش نمی‌تواند به‌طور راضی‌کننده‌ای از طرف سازندگانش توصیف شود، زیرا آن‌ها روشی را به کار گرفته‌اند که تا اندازه‌ی زیادی تجربی است. نهایتا، دوست داریم انسان‌هایی را که به‌صورت عادی به دنیا آمده‌اند از ماشین‌ها مستثنی کنیم. دشوار بتوان تعاریف را در چارچوبی گذاشت تا این سه شرط را برآورده کرد. احتمالا برای مثال تاکید می‌شود که کل تیم مهندسان باید از یک جنس باشد، اما این واقعا راضی‌کننده نخواهد بود، زیرا احتمالا ممکن است که بتوان یک فرد کامل را از روی یک سلول واحد پوست (فرضا) یک انسان پرورش داد. انجام چنین کاری شاهکار فن زیست‌شناختی خواهد بود که لایق بالاترین ستایش است، اما تمایل نداریم تا این را به‌صورت موردی از «ساخت ماشین متفکر» در نظر بگیریم. این نکته وادارمان می‌کند این خواسته را که استفاده از همه‌ی انواع فنون باید مجاز باشد کنار بگذاریم. با نظر به این واقعیت که علاقه‌ی کنونی به «ماشین‌های متفکر» با نوع خاصی از ماشین که معمولا «کامپیوتر الکترونیک» یا «کامپیوتر دیجیتال» می‌خوانند هویدا شده است، داریم آماده‌ی انجام چنین کاری می‌شویم. پیرو این پیشنهاد فقط به کامپیوترهای دیجیتال اجازه می‌دهیم تا در بازی‌مان نقش ایفا کنند.

این محدودیت در نگاه نخست بسیار عمده و جدی به نظر می‌رسد. می‌کوشم نشان دهم که در واقعیت این‌گونه نیست. برای انجام چنین کاری ضرورت دارد طبیعت و خصایص این کامپیوترها را کوتاه تشریح کنیم.

همچنین می‌توان گفت که این شناسایی یا تشخیص ماشین‌ها با کامپیوترهای دیجیتال، مثل سنجه‌مان برای «فکرکردن»، فقط به شرطی راضی‌کننده نخواهد بود که (برخلاف باورم) کامپیوترهای دیجیتال نتوانند نمایش خوبی در بازی ارائه دهند.

پیشاپیش شماری از کامپیوترهای دیجیتال در مرتبه‌ی کاری وجود دارند، و چه‌بسا پرسیده شود «چرا آزمایش را بدون فوت وقت انجام ندهیم؟ برآوردن شرایط بازی آسان خواهد بود. از تعدادی مستنطق می‌توان استفاده کرد، و آمار هم کمپایل می‌شود تا نشان دهد که شناسایی درست به چه نحوی اغلب داده شده بود.» پاسخ کوتاه این است که نمی‌پرسیم آیا همه‌ی کامپیوترهای دیجیتال خوب بازی خواهند کرد یا نه، و نه حتی می‌پرسیم آیا همه‌ی کامپیوترهای کنونی دردسترس کارشان را خوب انجام می‌دهند یا نه، بلکه می‌پرسیم آیا می‌توان

کامپیوترهایی را تصور کرد که بتوانند خوب بازی کنند یا نه. اما این صرفا پاسخی کوتاه است. این پرسش را در پرتو متفاوتی بعدتر می‌بینیم.

چهار. کامپیوترهای دیجیتال

ایده‌ی پس پشت کامپیوترهای دیجیتال می‌تواند با این گفته توضیح داده شود که قرار است این ماشین همه‌ی عملیات‌هایی را انجام دهد که یک کامپیوتر انسانی می‌تواند انجام دهد. فرض شده که کامپیوتر انسانی از قواعد تثبیت‌شده‌ای پیروی کند؛ این کامپیوتر هیچ اختیاری برای انحراف از آن قواعد در هیچ کدام از جزئیات‌شان را ندارد. می‌توانیم فرض بگیریم که این قواعد در یک کتاب تدارک دیده شده‌اند، که هر وقت به کار تازه‌ای گماشته شود تغییر می‌کند. کامپیوتر همچنین موجودی نامحدودی کاغذ دارد که محاسباتش را روی آن انجام می‌دهد. ضرب‌ها و جمع‌هایش را نیز می‌تواند روی یک «ماشین رومیزی» انجام دهد، اما این اهمیتی ندارد.

اگر توضیح بالا را درمقام یک تعریف به کار ببریم در خطر حلقوی‌بودن استدلال قرار می‌گیریم. با طرح خطوط کلی وسایلی که اثر مطلوب با استفاده از آن‌ها به دست می‌آید از خطر مزبور اجتناب می‌کنیم. یک کامپیوتر دیجیتال معمولا می‌تواند واجد سه بخش در نظر گرفته شود:

یک. ذخیره

دو. واحد اجرایی

سه. کنترل

ذخیره ذخیره‌ی اطلاعات است، و با کاغذ کامپیوتر انسانی تناظر دارد، چه کاغذی باشد که محاسباتش را روی آن انجام می‌دهد، چه کاغذی باشد که کتاب قواعدش روی آن پرینت شده‌اند. مادامی که کامپیوتر انسانی محاسبات را در سرش انجام می‌دهد بخشی از ذخیره با حافظه‌اش متناظر خواهد بود.

واحد اجرایی بخشی است که عملیات‌های منفرد متعدد مشمول در یک محاسبه را انجام می‌دهد. ماهیت این عملیات‌های منفرد از یک ماشین به ماشین دیگر فرق خواهد کرد. معمولا عملیات‌های نسبتا طولانی می‌توانند در این واحد انجام شوند، مثل «۳۵۴۰۶۷۵۴۴۵ را در ۷۰۷۶۳۴۵۶۸۷ ضرب کن»، اما در برخی ماشین‌ها تنها عملیات‌های بسیار ساده‌ای مثل «بنویس ۰» ممکن‌اند.

اشاره کرده‌ایم که «کتاب قواعد» تدارک‌دیده‌شده برای کامپیوتر با یک بخش ذخیره در ماشین عوض می‌شود. پس به آن «جدول دستورات» می‌گوییم. مشاهده‌ی تبعیت صحیح از این دستورات و تبعیت از آن‌ها در مرتبه‌ای درست وظیفه‌ی کنترل است. کنترل طوری ساخته شده است که این امر ضرورتا روی می‌دهد.

اطلاعات موجود در ذخیره معمولا به بسته‌هایی با اندازه‌ی نسبتا کوچک تقسیم می‌شود. در یک ماشین، برای نمونه، یک بسته ممکن است از ده رقم ده‌دهی تشکیل شود. اعداد به بخش‌هایی از ذخیره اختصاص می‌یابند که در آن‌ها بسته‌های متعدد اطلاعات به صورتی نظام‌مند ذخیره شده‌اند. یک دستور نمونه می‌تواند بگوید—

«عدد ذخیره‌شده در محل ۶۸۰۹ را به عدد ذخیره‌شده در محل ۴۳۰۲ اضافه کن و نتیجه را در محل ذخیره‌سازی دوم قرار بده.»

نیازی نیست گفته شود که این دستور در ماشین رویت‌شده در انگلستان رخ نخواهد داد. به احتمال زیاد به‌صورتی همچون ۶۸۰۹۴۳۰۲۱۷ کد خواهد شد. اینجا ۱۷ یعنی کدام یک از عملیات‌های ممکن متعدد باید روی دو عدد مزبور اجرا شود. در این مورد، عملیات همان است که بالا توصیف شد، به عبارت دیگر، همان دستور «عدد... را به... اضافه کن». پی خواهیم برد که این دستور به ۱۰ رقم نیاز دارد و به‌نحوی بسیار متعارف به یک بسته‌ی اطلاعات شکل می‌دهد. کنترل به‌طور عادی دستوراتی را که باید از آن‌ها تبعیت شود در سطح مراتبی در نظر می‌گیرد که دستورات در آن‌ها ذخیره شده‌اند، اما گاهی هم ممکن است با دستوری مواجه شود مانند

«حالا از دستور ذخیره‌شده در محل ۵۶۰۶ تبعیت کن و از آنجا ادامه بده»

یا دیگربار

«اگر محل ۴۵۰۵ شامل ۰ می‌شود سپس از دستور ذخیره‌شده در ۶۷۰۷ تبعیت کن، درغیراین‌صورت به کارت ادامه بده.»

دستوراتی از سنخ آخر اهمیت بسیاری دارند زیرا تکرار دائم رشته‌ای از عملیات‌ها تا زمان تحقق شرطی خاص را ممکن می‌سازند، اما این کار را انجام می‌دهند تا نه از دستوراتی تازه درباره‌ی هر تکرار بلکه از دستوراتی یکسان دوباره و دوباره تبعیت کنند. یک قیاس خانوادگی: فرض کنید مادر از تامی می‌خواهد هر روز در مسیرش به سمت مدرسه به کفاش سر بزند تا ببیند آیا کفش‌هایش آماده شده‌اند یا نه، او می‌تواند هر روز صبح همین را دوباره از بچه‌اش بخواهد. به‌نحوی بدیل، او می‌تواند یک بار و برای همیشه یادداشتی را در راهرو بچسباند که وقتی بچه‌اش می‌خواهد خانه را به قصد مدرسه ترک کند آن را ببیند و یادداشت به او بگوید که به دنبال کفش‌ها برود، و نیز اگر بچه‌اش کفش‌ها را با خودش به خانه آورده باشد یادداشت را از بین ببرد.

خواننده باید این نکته را به‌عنوان یک واقعیت بپذیرد که بنا بر اصولی که توصیف کرده‌ایم کامپیوترهای دیجیتال می‌توانند ساخته شوند و در حقیقت ساخته شده‌اند، و اینکه در واقع می‌توانند اعمال کامپیوتر انسانی را با دقت بالا تقلید کنند.

کتاب قواعدی که گفتیم کامپیوتر انسانی‌مان از آن استفاده می‌کند قطعا خیالی بی‌دردسر است. کامپیوترهای انسانی موجود واقعا به یاد می‌آورند که باید چه کاری انجام دهند. اگر بخواهیم ماشینی بسازیم که رفتار کامپیوتر انسانی را در عملیاتی پیچیده تقلید کند باید بپرسیم که چطور این کار انجام شده است، و بعد پاسخ را به صورت یک جدول دستورات ترجمه کنیم. ساختن جدول‌های دستور معمولا به‌صورت «برنامه‌ریزی» توصیف شده است. «برنامه‌ریزی یک ماشین برای اجرای عملیات الف» یعنی قراردادن جدول دستور متناسب در ماشین طوری که الف را انجام خواهد داد.

نوع جالب دیگری از ایده‌ی کامپیوتر دیجیتال یک «کامپیوتر دیجیتال با عنصر تصادفی» است. این نوع کامپیوتر دستوراتی دارد شامل پرتاب یک تاس یا یک‌جور فرایند الکترونیک معادل؛ یکی از چنین دستوراتی برای نمونه می‌تواند «تاس را پرتاب کن و عدد حاصل را در

ذخیره‌ی ۱۰۰۰ قرار بده» باشد. گاهی چنین ماشینی با اراده‌ی آزاد توصیف شده است (گرچه خودم از این عبارت استفاده نخواهم کرد). به‌طور عادی ممکن نیست که با مشاهده‌ی یک ماشین تعیین کرد که آیا عنصری تصادفی دارد یا نه، چراکه اثری مشابه می‌تواند با چنان تدابیری تولید شود طوری که انتخاب‌ها به ارقام اعشاری برای π بستگی دارد.

اغلب کامپیوترهای دیجیتال موجود صرفا ذخیره‌ای متناهی دارند. ایده‌ی کامپیوتری با ذخیره‌ای نامحدود از حیث نظری کاملا صحیح است. قطعا تنها یک بخش متناهی بتواند در هر یک بار مورد استفاده قرار بگیرد. به همین منوال، تنها یک مقدار متناهی بتواند ساخته شود، اما می‌توانیم تصور کنیم که در صورت لزوم مقادیر هرچه‌بیشتری اضافه می‌شوند. چنین کامپیوترهایی منفعت نظری خاصی دارند و کامپیوترهایی با ظرفیت نامتناهی نامیده خواهند شد.

ایده‌ی کامپیوترهای دیجیتال ایده‌ای قدیمی است. چارلز ببیج، استاد ریاضیات در کمبریج از ۱۸۲۸ تا ۱۸۳۹، چنین ماشینی را طرح ریخت و **موتور تحلیلی** نامید، اما هرگز کامل نشد. گرچه ببیج همه‌ی ایده‌های اساسی را داشت، ماشینش در آن دوره دورنمای چندان جذابی نبود. سرعتی که دردسترس خواهد بود حتما بیشتر از کامپیوتر انسانی است اما چیزی مثل ۱۰۰ بار کندتر از ماشین منچستر، که خودش یکی از ماشین‌های کند مدرن است. ذخیره باید کاملا مکانیکی می‌بود و از چرخ‌ها و کارت‌ها استفاده می‌کرد.

این واقعیت که **موتور تحلیلی** ببیج باید سرتاسر مکانیکی می‌بود به ما یاری خواهد رساند تا خودمان را از خرافه‌ای خلاص کنیم. اغلب این واقعیت را مهم دانسته‌اند که کامپیوترهای دیجیتال مدرن الکتریکی‌اند، و اینکه سیستم عصبی نیز الکتریکی است. از آنجا که ماشین ببیج الکتریکی نبود، و از آنجا که تمام کامپیوترهای دیجیتال به یک معنا هم‌ارز هستند، می‌بینیم که این استفاده از الکتریسیته نمی‌تواند واجد اهمیتی نظری باشد. البته الکتریسیته معمولا آنجا به میان می‌آید که ارسال سریع سیگنال مد نظر باشد. در سیستم عصبی، پدیده‌های شیمیایی حداقل به اندازه‌ی پدیده‌های الکتریکی مهم‌اند. در برخی کامپیوترها سیستم ذخیره اساسا صوتی است. بدین‌ترتیب ویژگی استفاده از الکتریسیته تنها شباهتی بسیار ظاهری به نظر می‌رسد. اگر می‌خواهیم چنین شباهت‌هایی بیابیم باید در پی قیاس‌های ریاضیاتی تابع باشیم.

پنج. جهان‌شمولیت کامپیوترهای دیجیتال

کامپیوترهای دیجیتال ملاحظه‌شده در قسمت قبل می‌توانند در رده‌ی «ماشین‌هایی با حالت گسسته» طبقه‌بندی شوند. این ماشین‌ها با جهش‌ها یا تلق‌های ناگهانی از یک حالت کاملا معین به یک حالت کاملا معین دیگر به حرکت درمی‌آیند. این حالات به‌قدر کافی متفاوت‌اند تا امکان خلط بین آن‌ها نادیده گرفته شود. به بیانی واضح، چنین ماشین‌هایی وجود ندارند. همه چیز واقعا به‌طور پیوسته حرکت می‌کند. اما انواع‌واقسام بسیار زیادی از ماشین‌ها وجود دارند که می‌توان به آن‌ها به‌طور سودمندی درمقام ماشین‌هایی با حالت گسسته فکر کرد. برای نمونه، با ملاحظه‌ی سوئیچ‌ها برای یک سیستم نوری، خیالی بی‌دردسر است که هر سوئیچ باید به‌طور معین و قطعی روشن یا خاموش شود. باید مراتب واسط یا جایگاه‌های رابط در کار باشد، اما برای اغلب مقاصد می‌توانیم نادیده‌اش بگیریم. به‌عنوان

مثالی از یک ماشین با حالت گسسته می‌توانیم چرخی را در نظر بگیریم که ۱۲۰ درجه در هر ثانیه دور خودش تلق می‌چرخد، اما ممکن است با یک اهرم که می‌تواند از بیرون اعمال گردد متوقف شود. این ماشین می‌تواند از حیث انتزاعی به این ترتیب توصیف شود. حالت درونی ماشین (که با جایگاه چرخ توصیف شده) می‌تواند q1، q2، یا q3 باشد. یک سیگنال ورودی i_o یا i_1 (جایگاه اهرم) وجود دارد. حالت درونی در هر لحظه با آخرین حالت و سیگنال ورودی با جدول زیر تعیین می‌شود:

		Last State		
		q_1	q_2	q_3
Input	i_0	q_2	q_3	q_1
	i_1	q_1	q_2	q_3

سیگنال‌های خروجی، تنها نشان مرئی حالت درونی از بیرون (نور)، با جدول زیر توصیف می‌شود:

State	q_1	q_2	q_3
Output	o_0	o_0	o_1

این مثال نمونه‌ای از ماشین‌هایی با حالت گسسته است. آن‌ها را می‌توان با چنین جدول‌هایی توصیف کرد به این شرط که صرفا تعدادی متناهی حالات ممکن داشته باشند.

به نظر خواهد رسید که با فرض حالت ابتدایی ماشین و سیگنال‌های ورودی همواره بتوان همه‌ی حالت‌های آتی را پیشگویی کرد. این یادآور دیدگاه لاپلاس است که از حالت کامل عالم در یک لحظه از زمان، که با جایگاه‌ها و سرعت‌های تمام ذرات توصیف شده، باید بتوان تمام حالت‌های آینده را پیشگویی کرد. بااین‌همه پیشگویی مد نظر ما بیشتر به عملی‌بودن یا قابل‌اجرابودن نزدیک است تا به پیشگویی مد نظر لاپلاس. سیستم «عالم درمقام یک کل» طوری است که خطاهای بسیار کوچک در شرایط اولیه می‌توانند اثری چشم‌گیر در زمانی متعاقب داشته باشد. جابجایی الکترونی واحد با یک میلیاردم سانتی‌متر در یک لحظه ممکن است تفاوتی ایجاد کند بین انسانی که با بهمن در یک سال بعد کشته می‌شود، یا از آن می‌گریزد. این یک خصیصه‌ی ذاتی سیستم‌های مکانیکی موسوم به «ماشین‌هایی با حالت گسسته» است که این پدیده رخ نمی‌دهد. حتی وقتی ماشین‌های فیزیکی موجود را به جای ماشین‌های آرمانی‌شده در نظر می‌گیریم، شناخت معقولانه‌ی دقیق از حالتش در یک لحظه شناخت معقول دقیقی از هر تعداد از قدم‌های آتی به دست می‌دهد.

اشاره کردیم که کامپیوترهای دیجیتال در زمره‌ی ماشین‌هایی با حالت گسسته‌اند. اما تعداد حالت‌هایی که چنین ماشینی قادر به آن‌هاست معمولا بی‌اندازه زیاد است. برای نمونه، عدد ماشینی که اکنون در منچستر کار می‌کند تقریبا $۲^{۱۶۵۰۰۰}$ است، یعنی تقریبا $۱۰^{۵۰۰۰۰}$. این را با مثال‌مان از چرخ تلق‌زننده‌ی توصیف‌شده در بالا قیاس کنید، که سه حالت داشت. دشوار نیست که بفهمیم چرا تعداد حالت‌ها باید بسیار گسترده باشد. کامپیوتر ذخیره‌ای متناظر با کاغذ مورد استفاده‌ی کامپیوتری انسانی را شامل می‌شود. باید ممکن باشد که هر کدام از ترکیبات نمادها را که می‌توانند روی کاغذ نوشته شوند در ذخیره نوشت. برای سادگی بحث،

فرض کنید که صرفا ارقام از ۰ تا ۹ به‌صورت نمادها استفاده می‌شوند. تغیرات دست‌نویسی نادیده گرفته شده‌اند. فرض کنید کامپیوتر اجازه‌ی استفاده از ۱۰۰ ورق کاغذ را دارد که هر کدام شامل ۵۰ خط با فضایی برای ۳۰ رقم است. در نتیجه تعداد حالت‌ها $10^{100\times50\times30}$ است، یعنی 10^{150000}. این درباره‌ی حالت‌های سه ماشین منچستر است که کنار هم گذاشته شدند. لگاریتم مبنای دو تعداد حالت‌ها معمولا «ظرفیت ذخیره‌سازی» ماشین خوانده می‌شود. بنابراین ظرفیت ذخیره‌سازی ماشین منچستر تقریبا ۱۶۵۰۰۰ و ماشین چرخ مثال ما ۶.۱ است. اگر این دو ماشین کنار هم گذاشته شوند ظرفیت‌هایشان باید اضافه شود تا به ظرفیت ماشین نتیجه‌شده برسد. این به امکان گزاره‌هایی می‌انجامد مثل «ماشین منچستر ۶۴ شیار مغناطیسی هر کدام با ظرفیت ۲۵۶۰ و هشت لوله‌ی الکترونیکی با ظرفیت ۱۲۸۰ را در خود دارد. ذخیره‌ی متفرقه معادل است با تقریبا ۳۰۰ که روی هم رفته به ۱۷۴۳۸۰ می‌رسد.»

با فرض جدول متناظر با ماشینی با حالت گسسته می‌توان عملکردش را پیشگویی کرد. هیچ دلیلی وجود ندارد که چرا این محاسبه نباید با استفاده از کامپیوتری دیجیتال انجام شود. به شرطی که کامپیوتر دیجیتال بتواند به‌قدر کافی سریع کارش را اجرا کند، می‌تواند رفتار هر ماشین با حالت گسسته را تقلید کند. بازی تقلید این‌گونه می‌تواند با ماشین مورد بحث (مثل ب) و کامپیوتر دیجیتال تقلیدکننده (مثل الف) انجام شود و مستنطق نخواهد توانست آن‌ها را تشخیص بدهد. قطعا کامپیوتر دیجیتال باید یک ظرفیت ذخیره‌سازی کافی داشته باشد و به‌قدر کافی سریع کار کند. به‌علاوه، برای هر ماشین تازه‌ای که می‌خواهد تقلید کند باید از نو برنامه ریخته شود.

این ویژگی خاص کامپیوترهای دیجیتال، یعنی توان تقلید از هر ماشین با حالت گسسته، با این گفته توصیف می‌شود که آن‌ها ماشین‌های *جهانشمول* هستند. وجود ماشین‌هایی با این ویژگی این نتیجه‌ی مهم را دارد که جدا از ملاحظات سرعت دیگر لازم نیست که ماشین‌های تازه‌ی متعددی برای انجام فرایندهای محاسبه‌ی متعدد طراحی شود. همه‌اش می‌تواند با یک کامپیوتر دیجیتال انجام شود که به‌طور مناسبی برای هر مورد برنامه‌ریزی شده است. در نتیجه فهم خواهد شد که تمام کامپیوترهای دیجیتال به یک معنا هم‌ارزند.

دوباره می‌توانیم نکته‌ای را که در انتهای بخش ۳ طرح شد در نظر بگیریم. با احتیاط پیشنهاد شده بود که پرسش «آیا ماشین‌ها می‌توانند فکر کنند؟» باید با پرسش «آیا کامپیوترهای دیجیتال قابل‌تصوری وجود دارند که بازی تقلید را به‌خوبی انجام دهند؟» عوض شود. اگر مایل باشیم می‌توانیم این مسئله را در ظاهر عام‌تر کنیم و بپرسیم «آیا ماشین‌هایی با حالت گسسته وجود دارند که خوب عمل کنند؟» اما با نظر به ویژگی جهانشمولیت می‌بینیم که هر کدام از این پرسش‌ها معادل است با این: «بیایید توجه‌مان را روی یک کامپیوتر دیجیتال خاص پ متمرکز کنیم. آیا درست است که با اصلاح و تغییر این کامپیوتر، طوری که ذخیره‌ای کافی داشته باشد، به‌طور مناسبی سرعت عملش افزایش یابد، و برنامه‌ای متناسب برایش مهیا شود، پ می‌تواند به‌طور راضی‌کننده‌ای نقش الف در بازی تقلید و نقش ب که یک انسان جایش را گرفته را بازی کند؟»

شش. دیدگاه‌های مغایر بر پرسش اصلی

اکنون زمینه‌ای را که می‌بایست روشن شده باشد در نظر می‌گیریم و آماده‌ایم تا به سراغ مناقشه بر سر پرسش‌مان، «آیا ماشین‌ها می‌توانند فکر کنند؟» و نسخه‌ای از آن که در انتهای قسمت قبل نقل شد، برویم. روی‌هم‌رفته نمی‌توانیم شکل اصلی مسئله را کنار بگذاریم، زیرا عقاید در مورد مناسب‌بودن تعویض مسئله فرق خواهد کرد و باید دست‌کم به آنچه باید در این پیوند گفته شود گوش بسپاریم.

موضوعات برای خواننده ساده خواهند شد اگر در ابتدا باورهای خودم در این مورد را توضیح بدهم. ابتدا شکل دقیق‌تر پرسش را در نظر بگیرید. به باورم طی تقریبا پنجاه سال ممکن خواهد بود که کامپیوترهایی با ظرفیت ذخیره‌ی تقریبا 10^9 را برنامه بریزیم تا سبب شویم بازی تقلید را به‌قدری خوب انجام دهند که یک مستنطق معمولی پس از پنج دقیقه پرسش‌گری بیشتر از ۷۰ درصد شانس شناسایی درست خواهد داشت. به باورم پرسش اصلی، «آیا ماشین‌ها می‌توانند فکر کنند؟»، به‌قدری بی‌معناست که لایق بحث نیست. در هر صورت، معتقدم که در پایان این قرن استفاده از کلمات و عقیده‌ی دانش‌آموخته‌ی عمومی به‌قدری عوض خواهد شد که خواهیم توانست از فکرکردن ماشین‌ها حرف بزنیم بدون اینکه توقع هر نوع مغایرت داشته باشیم. کاملا اشتباه است این دیدگاه عامه‌پسند که دانشمندان، که بی‌امان از واقعیتی تثبیت‌شده به واقعیتی تثبیت‌شده می‌روند، هرگز تحت تأثیر هیچ حدس‌وگمان اثبات‌نشده‌ای نبوده‌اند. هیچ لطمه‌ای نمی‌تواند به وجود آید اگر روشن شده باشد که کدام‌ها واقعیت اثبات‌شده هستند و کدام‌ها حدس‌وگمان. حدس‌وگمان‌ها اهمیت زیادی دارند زیرا خطوط پژوهشی مفیدی را پیش می‌نهند.

حال می‌خواهم عقایدی مخالف عقیده‌ام را در نظر بگیرم.

(۱) ***اعتراض الاهیاتی***. تفکر کارکرد روح نامیرای انسان است. خدا به هر زن و مرد روحی نامیرا داده است، اما نه به هیچ حیوان دیگر یا به ماشین‌ها. بنابراین هیچ حیوان یا ماشینی نمی‌تواند فکر کند.

نمی‌توانم ذره‌ای از این اعتراض را بپذیرم، اما سعی خواهم کرد به زبانی الاهیاتی پاسخ دهم. استدلال مزبور را اقناع‌کننده‌تر می‌یافتم اگر حیوانات با انسان‌ها در یک دسته قرار می‌گرفتند زیرا از دیدم تفاوت بین جاندار و بیجان بزرگ‌تر از تفاوت بین انسان و دیگر حیوانات است. خصیصه‌ی دلبخواهی منظر راست‌کیش روشن‌تر می‌شود اگر در نظر بگیریم که برای عضوی از یک اجتماع دینی دیگر چطور می‌تواند به نظر برسد. مسیحی‌ها دیدگاه مسلمان‌ها را که زنان روح ندارند چطور در نظر می‌گیرند؟ اما اجازه دهید این نکته را کنار بگذاریم و به استدلال اصلی برگردیم. به نظرم می‌رسد که استدلال نقل‌شده در بالا محدودیت جدی قدرت مطلق خداوند قادر را ایجاب می‌کند. تصدیق شده که چیزهای مشخصی وجود دارند که او نمی‌تواند انجامش دهد مثل معادل‌کردن یک با دو، اما آیا نباید باور کنیم که او آزادی آن را دارد که اگر شایستگی لازم را در یک فیل ببیند روحی به او اعطا کند؟ می‌توانیم انتظار داشته باشیم که او این قدرت را فقط در پیوند با تحولی اعمال خواهد کرد که یک مغز به‌تناسب ترقی‌یافته را برایش مهیا می‌کند که در خدمت نیازهای روحش باشد. به همین صورت استدلالی دقیقا مشابه می‌تواند در مورد ماشین‌ها آورده شود. ممکن است متفاوت به نظر برسد زیرا «هضم»ش دشوارتر است. اما در واقع این فقط یعنی ما فکر می‌کنیم احتمال

کمتری دارد که خداوند اقتضائات درخور برای اعطای روح را در نظر بگیرد. این اقتضائات مورد پرسش در بقیه‌ی این مقاله به بحث گذاشته می‌شوند. در تلاش برای ساختن چنین ماشین‌هایی، همان‌قدر که در حال زادوولد کودکان هستیم، نباید قدرت او در خلق روح‌ها را بی‌مناسبت غصب کنیم: در عوض، ما در هر دو مورد ابزارهای اراده‌ی او هستیم که عمارت‌هایی برای روح‌هایی که او می‌آفریند مهیا می‌کنیم.

بااین‌همه، این صرفاً گمانه‌زنی است. خیلی متأثر از استدلال‌های الاهیاتی نیستم، حال برای پشتیبانی از هر چیزی استفاده شوند. این استدلال‌ها اغلب در گذشته راضی‌کننده نبوده‌اند. در زمانه‌ی گالیله، استدلال آورده شده بود که متون «و خورشید ثابت ایستاد... و شتافت تا سرتاسر روز را غروب نکند» (یوشع، باب ۱۳) و «او بنیان‌های زمین را بر پا کرد، که نباید هیچ زمانی حرکت کند» (مزامیر، بخش ۵) ابطال کافی نظریه‌ی کوپرنیکی بودند. با دانش کنونی‌مان چنین استدلالی بیهوده به نظر می‌رسد. وقتی این دانش در دسترس نبود، برداشت کاملا متفاوتی را ایجاد کرد.

(۲) ***اعتراض «خود را به ندیدن زدن»***. «عواقب فکرکردن ماشین‌ها بسیار هولناک است. بیایید امیدوار باشیم و باور کنیم که ماشین‌ها نتوانند فکر کنند.»

این استدلال ندرتا به‌نحوی کاملا علنی به‌صورت بالا بیان شده است. اما در هر صورت اغلب ما را که به آن فکر می‌کنیم کاملا تحت تأثیر قرار می‌دهد. دوست داریم باور کنیم که بشر به شیوه‌ای ظریف برتر از باقی خلقت است. عالی است اگر می‌توان نشان داد که او *ضرورتا* برتر است، زیرا در نتیجه این خطر دیگر وجود ندارد که موضع فرمان‌دهنده‌اش را از دست بدهد. محبوبیت استدلال الاهیاتی آشکارا با این احساس رابطه دارد. احتمالا این استدلال در ذهن روشنفکران بسیار قوی باشد، زیرا آن‌ها بیش از دیگران برای توان تفکر ارزش بالایی قائل‌اند، و بیشتر تمایل دارند تا باورشان به برتری بشر را در بر این توان بنا کنند.

فکر نمی‌کنم که این استدلال آنقدر اساسی باشد که به ابطال نیازی داشته باشد. دلداری متناسب‌تر خواهد بود: احتمالا این را باید در تناسخ ارواح جستجو کرد.

(۳) ***اعتراض ریاضیاتی***. تعدادی از نتایج منطق ریاضیاتی وجود دارند که می‌توانند نشان دهند محدودیت‌هایی برای قوای ماشین‌هایی با حالت گسسته وجود دارد. بهترین این نتایج به فرضیه‌ی *گودل* شناخته می‌شود، و نشان می‌دهد که در هر سیستم منطقی به‌قدر کافی قوی جمله‌هایی می‌توانند فرموله شوند که در این سیستم نه می‌توانند اثبات شوند نه رد، مگر اینکه احتمالا خود سیستم نامنجسم باشد. نتایج دیگری نیز با نظر به ملاحظات مشابه‌ای به خاطر چرچ، *کلین*، *روزر*، و *تورینگ* وجود دارند. نتیجه‌ی مورد آخر را راحت‌تر از همه می‌توان در نظر گرفت زیرا مستقیما به ماشین‌ها ارجاع دارد، درحالی‌که باقی تنها می‌توانند در استدلالی نسبتا غیرمستقیم استفاده شوند: برای نمونه، در صورت استفاده از فرضیه‌ی گودل نیاز داریم که افزون بر این‌ها وسیله‌ای برای توصیف سیستم‌های منطقی بر حسب ماشین‌ها و توصیف ماشین‌ها بر حسب سیستم‌های منطقی داشته باشیم. نتیجه‌ی مورد بحث به سنخی از ماشین ارجاع دارد که اساسا یک کامپیوتر دیجیتال با ظرفیت نامتناهی است. این نتیجه می‌گوید که چیزهای مشخصی وجود دارند که چنین ماشینی نمی‌تواند انجام دهد. اگر این ماشین سر هم شود تا به پرسش‌هایی همچون موارد بازی تقلید پاسخ بدهد، پرسش‌هایی در کار خواهند

بود که یا پاسخ غلطی به آن‌ها خواهد داد یا اصلا نمی‌تواند هیچ پاسخی بدهد، حال هر قدر زمان برای جواب در اختیار داشته باشد. قطعا پرسش‌های زیادی از این سنخ می‌توانند وجود داشته باشند، و پرسش‌هایی که یک ماشین نتواند به آن‌ها پاسخ دهد می‌توانند به‌طور راضی‌کننده‌ای توسط ماشین دیگری پاسخ داده شوند. البته داریم اکنون فرض می‌گیریم که این پرسش‌ها از نوعی هستند که یک پاسخ «بله» یا «نه» برای‌شان مناسب است و نه پرسش‌هایی مثل «درباره‌ی پیکاسو چه فکر می‌کنی؟» پرسش‌هایی که می‌دانیم ماشین‌ها می‌بایست در پاسخ به آن‌ها شکست بخورند از این سنخ‌اند، «در نظر بگیر که ماشین به‌صورت ... معین شود. آیا این ماشین هرگز به هر گونه پرسشی پاسخ «بله» می‌دهد؟» نقطه‌ها را باید با توصیفی از یک ماشین به‌صورتی استاندارد عوض کرد، که می‌تواند چیزی مثل توصیف مورد استفاده در قسمت ۵ باشد. وقتی ماشین مورد توصیف یک نسبت تقریبا ساده‌ی مشخص با ماشینی که ذیل استنطاق قرار دارد به وجود آورد، می‌توان نشان داد که پاسخ یا غلط است یا در حال آماده‌سازی. نتیجه‌ی ریاضیاتی‌اش این است: استدلال شده که این وضعیت ناتوانی ماشین‌هایی را اثبات می‌کند که عقل انسانی به آن‌ها مقید نیست.

پاسخ کوتاه به این استدلال این است که گرچه محرز شده که محدودیت‌هایی برای قوای هر ماشین به‌خصوص وجود دارند اما این استدلال بدون هر گونه گواهی بیان کرده است که هیچ محدودیتی از این سنخ برای عقل انسانی کاربرد ندارد. اما فکر نمی‌کنم این دیدگاه بتواند به این راحتی کنار گذاشته شود. هرآینه پرسش انتقادی متناسبی از یکی از این ماشین‌ها پرسیده شود، و پاسخی معین بدهد، می‌دانیم که این پاسخ باید غلط باشد، و همین احساس مشخص برتری را به ما می‌دهد. آیا این احساس وهم‌آلود است. بی‌تردید کاملا موثق است، اما فکر نمی‌کنم اهمیتی خیلی زیادی باید به آن داد. خود ما نیز اغلب پاسخ‌های غلطی به پرسش‌ها می‌دهیم تا در خرسندی زیاد از این گواه خطاپذیری از طرف ماشین‌ها توجیه شویم. به‌علاوه، برتری ما فقط می‌تواند در چنان وهله‌ای در نسبت با یک ماشینی حس شود که غلبه‌ی بی‌اهمیت‌مان بر آن را به دست آورده‌ایم. هیچ پرسشی از غلبه‌ی همزمان بر *تمام* ماشین‌ها وجود ندارد. خلاصه، امکان دارد انسان‌هایی باهوش‌تر از هر ماشین مفروض وجود داشته باشد، اما این امکان هم وجود دارد که ماشین‌های باهوش‌تر وجود داشته باشند، و الخ.

به گمانم آن‌ها که به استدلال ریاضیاتی باور دارند اغلب دوست دارند که بازی تقلید را به‌عنوان مبنایی برای بحث بپذیرند. آن‌هایی که به دو اعتراض قبلی باور دارند احتمالا به هیچ سنجه‌ای علاقه ندارند.

(۴) ***استدلال از روی آگاهی.*** این استدلال بسیار خوب در *پروفسور جفرسون* لیستر اوریشن در ۱۹۴۹ بیان شده است، که نقل‌قولی از آن را می‌آورم: «تا وقتی یک ماشین نتواند از سر افکار و هیجانات حس‌شده و نه با نزول شانسی نمادها سونات بنویسد یا کنسرتو بسراید، نمی‌توانیم موافق باشیم که ماشین با مغز معادل است ـ یعنی نه فقط آن را بنویسد بلکه بداند که آن را نوشته است. هیچ مکانیزمی نمی‌تواند از موفقیت‌هایش لذت ببرد (و نه صرفا به‌طور مصنوعی سیگنال بدهد، که تدبیری آسان است)، وقتی دریچه‌هایش فیوز می‌سوزانند غصه بخورد، با تملق سر حال بیاید، با اشتباهاتش به فلاکت درافتد، با سکس شیفته شود، وقتی نمی‌تواند به چیزی که می‌خواهد برسد خشمگین یا افسرده شود.»

به نظر می‌رسد این استدلال انکار صحت تست ما باشد. بنا بر افراطی‌ترین شکل این دیدگاه، تنها شیوه‌ای که با استفاده از آن می‌توان مطمئن بود که یک ماشین فکر می‌کند خود ماشین *بودن* و حس‌کردن فکرکردن خود است. بدین‌ترتیب می‌توان این احساسات را برای جهان توصیف کرد، اما قطعا هیچ کسی با هیچ اشارتی توجیه نمی‌شود. به همین منوال، بنا بر این دیدگاه، تنها شیوه‌ی دانستن این‌که یک *انسان* فکر می‌کند این است که آن انسان به‌خصوص باشد. در واقع این نظرگاهی خودمدارانه است. می‌تواند منطقی‌ترین دیدگاه برای باورکردن به آن باشد اما انتقال ایده‌ها را دشوار می‌سازد. الف مستعد آن است که باور کند «الف فکر می‌کند اما ب فکر نمی‌کند» درحالی‌که ب باور دارد که «ب فکر می‌کند اما الف فکر نمی‌کند.» عوض بحث مداوم پیرامون این نکته، معمول است که این رسم مودبانه را داشته باشیم که همه فکر می‌کنند.

مطمئن‌ام که پروفسور جفرسون مایل نیست نظرگاه افراطی و خودمدارانه را اتخاذ کند. احتمالا تماما دوست دارد که بازی تقلید را به‌عنوان یک تست بپذیرد. این بازی (با حذف بازیگر ب) مکررا به‌طور عملی تحت نام *امتحان شفاهی* استفاده شده است تا کشف شود که آیا یک فرد واقعا چیزی را می‌فهمد یا «آن را طوطی‌وار آموخته است». بیایید به بخشی از یک *امتحان شفاهی* گوش بدهیم:

مستنطق: در اولین خط سونات‌تان که می‌خواند «قیاسم ترا با روزی تابستانی»، آیا «روز بهاری» جواب نمی‌دهد یا بهتر نیست؟

شاهد: موزون نخواهد شد.

مستنطق: «روزی زمستانی» چطور. بسیار موزون خواهد بود.

شاهد: بله، ولی هیچ کسی نمی‌خواهد با یک روز زمستانی قیاس شود.

مستنطق: آیا می‌گویید آقای پیکویک شما را به یاد کریسمس می‌اندازد؟

شاهد: یک‌جورهایی.

مستنطق: بااین‌حال کریسمس یک روز زمستانی است، و فکر نمی‌کنم آقای پیکویک اهمیتی به این مقایسه بدهند.

شاهد: فکر نمی‌کنم راست بگویید. منظور از روزی زمستانی یک روز زمستانی معمول است و نه روزی خاص مثل کریسمس.

و الخ. پروفسور جفرسون چه خواهد گفت اگر ماشین سونات‌نویس می‌توانست به این صورت در *امتحان شفاهی* پاسخ بدهد؟ فکر نمی‌کنم بر این تصور می‌بود که ماشین «صرفا به‌طور مصنوعی» دارد این پاسخ‌ها را «سیگنال» می‌دهد، اما اگر پاسخ‌ها همچون در قطعه‌ی بالا راضی‌کننده و مستمر بودند فکر نمی‌کنم که آن را به‌صورت «تدبیری آسان» توصیف کند. به گمانم، این عبارت می‌خواهد چنان وسایلی را در ماشین ضبط کسی که سونات را می‌خواند شامل شود.

پس به‌طور خلاصه، فکر می‌کنم که اغلب آنانی که طرف استدلال از روی آگاهی هستند می‌توانند ترغیب شوند که این استدلال را کنار بگذارند عوض اینکه به موضعی خودمدارانه مجبور شوند. آن‌ها بدین‌ترتیب احتمالا دوست خواهد داشت تست ما را بپذیرند.

نمی‌خواهم این برداشت را بدهم که فکر می‌کنم هیچ رازی در مورد آگاهی وجود ندارد. برای نمونه، چیزی متناقض‌نما در کار است که با هر نوع تلاش برای جایابی‌اش رابطه دارد. اما فکر نمی‌کنم قبل از اینکه بتوانیم پاسخی به پرسش مورد بحث در این مقاله بدهیم ضرورتا نیاز باشد که این رازها حل شوند.

(۵) ***استدلال‌هایی از روی ناتوانی‌های متعدد.*** این استدلال‌ها به این صورت درمی‌آیند: «تصدیق می‌کنم که می‌توانید ماشین‌هایی بسازید که تمام چیزهایی را که اشاره کرده‌اید انجام بدهند اما هرگز نخواهید توانست ماشینی بسازید که ایکس را انجام بدهد». در این رابطه خصایص بی‌شمار ایکس پیشنهاد شده‌اند. منتخبی از آن‌ها را می‌آورم:

مهربان‌بودن، باتدبیربودن، زیبابودن، صمیمی‌بودن (ص۴۵۴)، داشتن ابتکار عمل، داشتن حس شوخ‌طبعی، تمیزگذاری درست از غلط، اشتباه‌کردن (ص۴۵۴)، عاشق‌شدن، لذت‌بردن از توت‌فرنگی و خامه (ص۴۵۳)، فردی دیگر را عاشق خود کردن، آموختن از تجربه (ص۴۶۰)، استفاده‌ی صحیح از کلمات، موضوع فکر خود شدن (ص۴۵۴ـ۵)، داشتن تنوعی از رفتارها به اندازه‌ی یک انسان، انجام‌دادن کاری واقعا تازه (ص۴۵۵ـ۶). (چنان‌که شماره‌ی صفحات اشاره دارند، ملاحظات ویژه‌ای به برخی از این ناتوانی‌ها نسبت داده شده‌اند.)

معمولا هیچ حمایتی از این جملات ارائه نشده است. به باورم آن‌ها اغلب بر اصل استقرای علمی متکی‌اند. انسان هزاران ماشین در طول زندگی‌اش دارد. بنا بر آنچه از آن‌ها می‌بیند شماری نتایج عام استخراج می‌کند. آن‌ها زشت‌اند، هر کدام برای مقصدی بسیار محدود طراحی شده‌اند، وقتی برای مقصدی بی‌نهایت متفاوت لازم باشند بی‌فایده‌اند، تنوع رفتاری هر کدام از آن‌ها بسیار کم است، و الخ، و الخ. طبیعتا انسان نتیجه می‌گیرد که این‌ها خصایص ضروری ماشین‌ها به‌طور کلی هستند. بسیاری از این محدودیت‌ها با ظرفیت ذخیره‌سازی بسیار کوچک اغلب ماشین‌ها پیوند دارند. (می‌پندارم که ایده‌ی ظرفیت ذخیره‌سازی به طریقی بسط می‌یابد تا ماشین‌ها و نه ماشین‌هایی با حالت گسسته را پوشش دهد. تعریفی دقیق مهم نیست همان‌طور که هیچ دقت ریاضیاتی در بحث حاضر ادعا نشده است.) چند سال پیش، وقتی چیزهای بسیار کمی از کامپیوترهای دیجیتال شنیده می‌شد، ممکن بود که ناباوری زیادی را در موردشان استنباط کرد. این احتمالا بابت کاربست مشابه اصل استقرای علمی بود. این کاربست‌های اصل قطعا بی‌اندازه ناخودآگاهانه‌اند. وقتی کودکی سوخته از آتش می‌ترسد و با دوری از آن نشان می‌دهد که از آن می‌ترسد، باید بگویم که دارد استقرای علمی را به کار می‌بندد. (البته همچنین می‌توانم این رفتار را به بی‌شمار طرق دیگر توصیف کنم.) کارها و رسم‌های بشریت به نظر ماده‌ی خیلی مناسبی برای کاربست استقرای علمی نمی‌رسد. بخش بسیار بزرگی از مکان‌زمان باید کاوش شود اگر بناست نتایج قابل‌اطمینانی به دست آیند. در غیر این صورت (همان‌طور که اغلب کودکان انگلیسی انجام می‌دهند) ممکن است تصمیم بگیریم که همگان انگلیسی حرف می‌زنند، و احمقانه است که فرانسوی یاد گرفته شود.

بااین‌حال، نکات ویژه‌ای وجود دارند که باید در مورد بسیاری از ناتوانی‌های ذکرشده خاطرنشان شوند. ناتوانی لذت‌بردن از توت‌فرنگی و خامه می‌تواند از دید خواننده سخیف به نظر برسد. احتمالا ماشینی بتواند ساخته شود که از این خوراک خوشمزه لذت ببرد، اما هر

تلاش برای ساختن چنین ماشینی که چنین کاری انجام بدهد ابلهانه خواهد بود. نکته‌ی مهم در مورد این ناتوانی این است که در برخی از دیگر ناتوانی‌ها دخیل است، برای مثال، در دشواری مشابه‌ای که در مورد صمیمیت بین انسان و ماشین رخ می‌دهد همان‌طور که بین انسانی سفیدپوست و انسانی سفیدپوست، یا بین انسانی سیاه‌پوست و انسانی سیاه‌پوست.

عجیب است این ادعا که «ماشین‌ها نمی‌توانند مرتکب اشتباه شوند». وسوسه می‌شویم با تندی جواب بدهیم «آیا برای ارتکاب اشتباه این‌قدر بد هستند؟» اما اجازه دهید نگرشی همدلانه‌تر را اتخاذ کنیم، و سعی کنیم ببینیم که واقعا چه معنایی دارد. فکر می‌کنیم که این نقد می‌تواند بر حسب بازی تقلید توضیح داده شود. ادعا شده که مستنطق به‌سادگی می‌تواند با طرح تعدادی مسئله‌ی حسابی ماشین را از انسان تمیز دهد. ماشین به خاطر دقت مهلکش افشا خواهد شد. جواب این نکته ساده است. ماشین (که برای اجرای این بازی برنامه‌ریزی شده) سعی نخواهد کرد که پاسخ‌های درست به مسائل حسابی بدهد. ماشین به‌طور دلبخواهی اشتباهاتی را به شیوه‌ای حساب‌شده مطرح خواهد کرد تا مستنطق را گیج کند. یک ایراد مکانیکی احتمالا خودش را از راه تصمیم نامناسبی نشان خواهد داد که برمبنایش چه نوع اشتباهی باید در حساب انجام گیرد. حتی این تفسیر از نقد به‌قدر کافی همدلانه نیست. اما نمی‌توانیم فضایی را برای بررسی بیشترش مهیا کنیم. به نظرم می‌رسد که این نقد به خلط بین دو نوع اشتباه بستگی دارد. می‌توانیم آن‌ها را «خطاهای کارایی» و «خطاهای نتیجه‌گیری» بخوانیم. خطاهای کارایی به خاطر ایرادی مکانیکی یا الکتریکی‌اند که سبب می‌شوند ماشین رفتاری غیر از آنچه برایش طراحی شده داشته باشد. در مباحث فلسفی دوست داریم امکان چنین خطاهایی را نادیده بگیریم؛ در نتیجه «ماشین‌های انتزاعی» را به بحث می‌گذاریم. ماشین‌های انتزاعی خیالات ریاضیاتی‌اند و نه ابژه‌های روانی. بنا به تعریف، آن‌ها قادر به خطاهای کارایی نیستند. از این حیث، به‌راستی می‌توانیم بگوییم که «ماشین‌ها هرگز نمی‌توانند مرتکب اشتباه شوند.» خطاهای نتیجه‌گیری فقط وقتی می‌توانند رخ بدهند که معنا به سیگنال‌های خروجی ماشین الصاق شود. ماشین برای نمونه ممکن است معادلات ریاضیاتی یا جملاتی به انگلیسی را چاپ کند. وقتی گزاره‌ای غلط تایپ شد می‌گوییم که ماشین مرتکب خطای نتیجه‌گیری شده است. آشکارا هیچ دلیلی وجود ندارد که بگوییم یک ماشین نمی‌تواند مرتکب چنین اشتباهی شود. چه‌بسا هیچ کاری نکند مگر اینکه مکررا تایپ کند «۰=۱». با ملاحظه‌ی مثالی نه‌چندان لجوجانه، چه‌بسا با استفاده از استقرای علمی روشی برای استخراج نتایج داشته باشد. باید انتظار چنین روشی را داشته باشیم تا گاها به نتایج خطاآمیز برسیم.

این ادعا که یک ماشین نمی‌تواند موضوع فکر خودش باشد قطعا فقط به شرطی می‌تواند پاسخ داده شود که بتواند نشان دهد که ماشین یک فکری در مورد یک موضوعی دارد. در هر صورت، «موضوع عملیات‌های یک ماشین» به نظر معنایی داشته باشد، دست‌کم برای آدم‌هایی که با ماشین سروکار دارند. اگر برای نمونه ماشین سعی می‌کرد راه‌حلی برای معادله $x^2 - 40x - 11 = 0$ بیابد، وسوسه خواهیم شد این معادله را به‌عنوان بخشی از موضوع ماشین در آن لحظه توصیف کنیم. در این معنای خاص، یک ماشین بی‌تردید می‌تواند موضوع اصلی خودش باشد. این موضوع می‌تواند در راستای کمک به ایجاد برنامه‌های خودش، یا برای پیشگویی اثر تغییرهایش در ساختار خودش استفاده شود. ماشین با مشاهده‌ی نتایج

رفتارش می‌تواند برنامه‌هایش را اصلاح کند طوری که مؤثرتر به قصد خودش دست یابد. این امکان‌ها از آن آینده‌ای نزدیک‌اند و نه رویاهایی اتوپیایی.

این نقد که یک ماشین نمی‌تواند تنوع رفتاری چندانی داشته باشد صرفا راهی است برای گفتن اینکه نمی‌تواند ظرفیت ذخیره‌سازی چندانی داشته باشد. ظرفیت ذخیره‌سازی حتی هزار رقم نسبتا تا همین اواخر بسیار نادر بود.

نقدهایی که اینجا در نظر می‌گیریم اغلب شکل‌های معوج‌شده‌ی استدلال از روی آگاهی‌اند. معمولا اگر بر این باور باشیم که یک ماشین *می‌تواند* یکی از این کارها را انجام دهد، و و نوع روشی را که ماشین می‌تواند مورد استفاده قرار دهد توصیف کنیم، تأثیر زیادی ایجاد نخواهیم کرد. تصور می‌شود که روش (هرآنچه باشد، چراکه باید مکانیکی باشد) واقعا مبنایی است. قیاس کنید پرانتز جمله‌ی نقل‌شده از جفرسون در ص۴۵۱.

(٦) ***اعتراض لیدی لاولیس.*** مفصل‌ترین اطلاعات ما از **موتور تحلیلی** ببیج از خاطرات مکتوب *لیدی لاولیس* می‌آید. او در این نوشته می‌گوید «**موتور تحلیلی** هیچ ادعایی مبنی بر *پدیدآوردن* هیچ چیزی ندارد. می‌تواند هر چیزی *را که می‌دانیم چطور به آن ترتیب بدهیم* اجرا کند» (ایتالیک‌ها از اوست). این جمله را *هارتری* (ص۷۰) نقل می‌کند و می‌افزاید: «این ایجاب نمی‌کند که نمی‌توان تجهیزاتی الکترونیکی ساخت که «له‌نفسه» فکر کند، یا در آن، به زبان زیست‌شناختی، بتوان انعکاس شرطی برقرار کرد، که به‌عنوان مبنایی برای «یادگیری» به کار رود. چه اصولا این امر ممکن باشد چه نباشد در هر صورت مسئله‌ای تحریک‌کننده و هیجان‌برانگیز است که برخی از همین پیشرفت‌های اخیر پیش گذاشته‌اند. اما به نظر نمی‌رسید که ماشین‌های ساخته‌شده یا طراحی‌شده در آن زمان این خصیصه را داشته باشند.»

من در این زمینه با هارتری توافق تمام‌وکمال دارم. اشاره خواهد شد که او اعلام نمی‌کند که ماشین‌های مورد بحث این خصیصه را به دست نیاورده بودند بلکه در عوض می‌گوید مدرک دردسترس برای لیدی لاولیس ترغیبش نکرد که باور کند ماشین‌ها خصیصه‌ی مزبور را به دست آورده بودند. کاملا ممکن است که یک ماشین با حالت گسسته این خصیصه را داشته باشد. **موتور تحلیلی** یک کامپیوتر دیجیتال جهانشمول بود، طوری که اگر ظرفیت ذخیره‌سازی و سرعتش کافی بودند با برنامه‌ریزی مناسب می‌توانست برای تقلید از ماشین مورد بحث ساخته شود. احتمالا این استدلال به ذهن کونتس یا ببیج خطور نکرد. در هر حال، هیچ قیدی بر آن‌ها نبود که مدعی همه‌ی چیزهایی شوند که می‌توان ادعایشان را داشت.

کل این مسئله دیگربار ذیل عنوان ماشین‌های یادگیری در نظر گرفته خواهد شد.

شکل دیگری از اعتراض لیدی لاولیس می‌گوید که یک ماشین نمی‌تواند «هرگز هیچ کار واقعا تازه‌ای انجام دهد». این اعتراض برای لحظه‌ای می‌تواند با مثل «هیچ چیز تازه‌ای زیر آسمان نیست» از سر باز شود. چه کسی می‌تواند یقین داشته باشد که «کار اصیل»ی که انجام داده به‌سادگی رشد بذری که با آموزش در او کار گذشته شده یا اثر پیروی از اصول عام مشهور نبوده باشد. نسخه‌ی بهتری از این اعتراض می‌گوید که یک ماشین هرگز نمی‌تواند ««شگفت‌زده‌مان کند». این جمله یک چالش مستقیم‌تر است و می‌توان مستقیما با آن رویارو شد. ماشین‌ها با بسامد بزرگی شگفت‌زده‌ام می‌کنند. تا اندازه‌ی زیادی دلیلش این است که محاسبات زیادی انجام نمی‌دهم تا تصمیم بگیرم که انتظار انجام چه کاری را از آن‌ها دارم، یا

در عوض به این خاطر است که گرچه محاسبه‌ای انجام می‌دهم آن را به شیوه‌ای عجولانه و سرسری انجام می‌دهم و خطرات را می‌پذیرم. احتمالا به خودم می‌گویم «ولتاژ در اینجا باید با ولتاژ در آنجا یکی باشد: در هر صورت این را مسلم می‌گیرم.» طبیعتا اغلب براشتباه‌ام، و نتیجه برایم یک شگفتی است زیرا وقتی آزمایش انجام شد این مفروضات فراموش شده‌اند. این تصدیق‌ها پرده از درسگفتارهایی درباره‌ی موضوع شیوه‌های شرورانه‌ام برمی‌دارند، اما وقتی به شگفتی‌هایی که تجربه می‌کنم گواهی می‌دهم مایه‌ی هیچ نوع تردیدی در قبال اعتبار کارم نیستند.

توقع ندارم این جواب منتقدم را ساکت کند. او احتمالا خواهد گفت که چنین شگفتی‌هایی به خاطر عمل ذهنی خلاقانه‌ی خودم هستند، و هیچ اعتباری به ماشین نمی‌دهند. این ما را به استدلال از روی آگاهی برمی‌گرداند، و از ایده‌ی شگفتی بسیار دور می‌کند. این خط از استدلال را باید با دقت در نظر بگیریم، اما احتمالا ارزش اشاره دارد که تجلیل از چیزی درمقام شگفتی مستلزم انبوهی «عمل ذهنی خلاقه» است، چه رویداد شگفتی‌آور از یک انسان، یک کتاب، یک ماشین سرچشمه بگیرد، چه از هر چیز دیگر.

این دیدگاه که ماشین‌ها نمی‌توانند بانی شگفتی‌ها شوند به باورم به خاطر مغالطه‌ای است که خصوصا فیلسوفان و ریاضیدانان مقید به آن‌اند. این فرض در کار است که به محض اینکه واقعیتی به ذهن ارائه شد تمام عواقب آن واقعیت همزمان با خود آن واقعیت به ذهن خطور می‌کند. این فرض تحت شرایط بسیاری عمیقا مفید است اما خیلی راحت فراموش می‌کنیم که غلط است. یک نتیجه‌ی طبیعی انجام چنین کاری این است که بعد فرض می‌گیریم که هیچ فضیلتی در صرف پرداختن به عواقب از روی داده‌ها و اصول عام نیست.

(۷) ***استدلال از روی پیوستگی در سیستم عصبی.*** سیستم عصبی قطعا ماشینی با حالت گسسته نیست. خطایی کوچک در اطلاعات درباره‌ی اندازه‌ی تکانه‌ای عصبی که بر یک نورون تأثیر می‌گذارد می‌تواند تفاوتی بزرگ در اندازه‌ی تکانه‌ی خروجی ایجاد کند. ممکن است استدلال شود که در این صورت نمی‌توان انتظار داشت که بتوان رفتار سیستم عصبی را با سیستمی با حالت گسسته تقلید کرد.

درست است که ماشینی با حالت گسسته باید از ماشینی پیوسته متفاوت باشد. اما اگر به شرایط بازی تقلید وفادار باشیم، مستنطق نخواهد توانست هیچ امتیازی از این تفاوت به دست آورد. این موقعیت می‌تواند روشن‌تر شود اگر یک ماشین پیوسته‌ی ساده‌تر دیگر را در نظر بگیریم. یک آنالیزور دیفرانسیلی بسیار خوب جواب خواهد داد. (یک آنالیزور دیفرانسیلی یک نوع خاص از ماشین غیر از سنخ حالت گسسته برای انواع محاسبه است.) برخی از این ماشین‌ها پاسخ‌شان را به صورتی تایپ‌شده آماده می‌کنند، و بنابراین برای مشارکت در بازی مناسب‌اند. برای یک کامپیوتر دیجیتال ممکن نیست که دقیقا پیشگویی کند که آنالیزور دیفرانسیلی دقیقا چه پاسخ‌هایی به یک مسئله خواهد داد، اما کاملا می‌تواند پاسخ درست را ارائه دهد. برای نمونه، اگر از آن بخواهیم که مقدار π (عملا به‌تقریب ۱۴۱۶.۳) را بگوید، معقول است که به‌طور تصادفی بین مقادیر ۱۲.۳، ۳۱۳، ۱۴.۳، ۱۵.۳، ۱۶.۳ با احتمالات (فرضا) ۰۵.۰، ۱۵.۰، ۵۵.۰، ۱۹.۰، ۰۶.۰ انتخاب صورت گیرد. تحت این شرایط برای مستنطق بسیار دشوار خواهد بود که آنالیزور دیفرانسیلی را از کامپیوتر دیجیتال تمیز دهد.

(۸) ***استدلال از روی بی‌پیرایگی رفتار.*** امکان ندارد که بتوان مجموعه قواعدی را تولید کرد که مدعی توصیف عملکرد الزامی یک انسان در هر مجموعه‌شرایط قابل‌فهم باشد. برای نمونه می‌توان قاعده‌ای داشت که براساسش وقتی چراغ راهنما را قرمز می‌بینیم متوقف شویم، و اگر چراغ سبز باشد ادامه دهیم، اما چه می‌شود اگر بابت ایرادی هر دو با هم ظاهر شوند؟ احتمالا تصمیم گرفته شود که ایمنی در توقف است. اما دشواری بعدی می‌تواند بابت همین تصمیم بعدتر ظهور کند. تلاش برای تدارک قواعد کرداری برای پوشش هر رویداد، حتی آن‌ها که از چراغ‌های راهنما ناشی می‌شوند، محال به نظر می‌رسد. با همه‌ی این‌ها موافقام.

از این گفته‌ها استدلال می‌شود که نمی‌توانیم ماشین باشیم. سعی می‌کنیم این استدلال را بازگو کنم، اما فکر نمی‌کنم بتوانم در قبالش منصفانه برخورد کنم. به نظر می‌رسد که این استدلال چیزی شبیه این را به جریان بیاندازد: «اگر هر انسان مجموعه‌قواعد کرداری معینی می‌داشت که با آن‌ها زندگی‌اش را تنظیم می‌کرد دست‌کمی از یک ماشین نمی‌داشت. اما چنین قواعدی وجود ندارند، پس انسان‌ها نمی‌توانند ماشین باشند.» میانه‌ی توزیع‌نشده در این استدلال فاحش است. فکر نمی‌کنم که این استدلال هرگز به این صورت مطرح شده باشد، اما به باورم این استدلالی است که در هر صورت مورد استفاده بوده است. بااین‌همه ممکن است خلط مشخصی بین «قواعد پیشبرد کردار» و «قوانین رفتاری» وجود داشته باشد. منظورم از «قواعد پیشبرد کردار» احکامی است مانند «اگر چراغ‌های قرمز را می‌بینی توقف کن»، که می‌توان بر مبنایش دست به عمل زد و می‌تواند ازشان آگاه بود. منظورم از «قوانین رفتاری» قوانین طبیعی است آن‌طور که برای بدن یک انسان به کار می‌رود، مثل «اگر او را نیشگون بگیری جیغ خواهد زد». اگر در استدلال نقل‌شده «قوانین رفتاری را که زندگی‌اش را تنظیم می‌کنند» به جای «قواعد پیشبرد کردار که او با آن‌ها زندگی‌اش را تنظیم می‌کند» بنشانیم، میانه‌ی توزیع‌نشده دیگر حل‌نشدنی نیست. زیرا معتقدیم که نه تنها درست است که تنظیم‌شدن با قوانین رفتاری یک‌جور ماشین‌بودن را ایجاب می‌کند (هرچند نه ضرورتا ماشینی با حالت گسسته) بلکه برعکس چنین ماشینی بودن تنظیم‌شدن با چنان قوانینی را ایجاب می‌کند. بااین‌همه، نمی‌توانیم خیلی راحت خودمان را در مورد غیاب قوانین کامل رفتار و نیز غیاب قواعد کامل پیشبرد کردار متقاعد کنیم. تنها راهی که برای یافتن چنان قوانینی می‌شناسیم مشاهده‌ی علمی است، و قطعا از هیچ شرایطی آگاه نیستیم که تحت آن‌ها بتوانیم بگوییم «به‌قدر کافی جستجو کرده‌ایم. چنین قوانینی وجود ندارند.»

می‌توانیم قویا اثبات کنیم که هر جمله‌ای از این سنخ ناموجه است. چون فرض کنید که می‌توانستیم از یافتن چنان قوانینی در صورت وجود مطمئن باشیم. سپس با فرض ماشینی با حالت گسسته قطعا باید ممکن باشد که از طریق مشاهده به‌قدر کافی درباره‌اش پی ببریم تا بتوانیم رفتار آتی‌اش را در یک گستره‌ی زمانی معقول، فرضا هزار سال، پیشگویی کنیم. اما به نظر نمی‌رسد که مسئله این باشد. من برنامه‌ای کوچک را با استفاده از تنها ۱۰۰۰ واحد ذخیره روی کامپیوتر منچستر نصب کردم، که به‌وسیله‌اش ماشینی مجهز به یک عدد شانزده رقمی طی دو ثانیه به ماشینی دیگر پاسخ می‌دهد. هر کسی را که از این پاسخ‌ها بیاموزد و آموخته‌اش درباره‌ی برنامه نیز کفایت کند به چالش می‌طلبم که بتواند هر جوابی به مقادیر نیازموده را پیشگویی کند.

(۹) ***استدلال از روی ادراک فراحسی.*** فرض می‌کنم که خواننده با ایده‌ی ادراک فراحسی و معنی چهار فقره‌اش، یعنی تلپاتی، غیب‌بینی، پیش‌آگاهی، و روان‌جنبش، آشناست. به نظر می‌رسد این پدیده‌های آزارنده تمام ایده‌های علمی معمول‌مان را رد کنند. چطور می‌بایست آن‌ها را از اعتبار ساقط کنیم! متأسفانه گواه آماری، دست‌کم برای تلپاتی، چشم‌گیر است. بسیار دشوار بتوان ایده‌هایمان را طوری بازترتیب داد که با این واقعیت‌های تازه جفت‌وجور شوند. همین‌که آن‌ها را پذیرفتیم به نظر خیلی سخت نیست که به اشباح و اجنه باور داشته باشیم. این ایده که بدن‌هایمان به‌سادگی بر حسب قوانین شناخته‌شده‌ی فیزیک حرکت می‌کنند، همراه با قوانین دیگری که هنوز کشف نشده‌اند اما تا اندازه‌ای مشابه‌اند، از اولین قدم‌هایی است که باید برداشته شود.

از دیدم این استدلال بسیار قوی است. در جواب می‌توان گفت که به‌نظر بسیاری از نظریات علمی، به‌رغم اختلاف با ادراک فراحسی، در عرصه‌ی عمل کارا می‌مانند؛ اینکه در واقع اگر فراموشش کنیم بسیار خوب می‌توانیم با آن سر کنیم. این تا اندازه‌ای یک آسودگی سرد است، و نگران‌ایم که تفکر صرفا آن نوع پدیده‌ای باشد که ادراک فراحسی به‌طور خاص می‌تواند به آن مربوط باشد.

یک استدلال خاص‌تر مبتنی بر ادراک فراحسی می‌تواند به این ترتیب مطرح شود: «بیایید بازی تقلید را انجام بدهیم، از یک انسان که درمقام گیرنده‌ای تلپاتیک خوب عمل می‌کند به‌عنوان شاهد و از یک کامپیوتر دیجیتال استفاده کنیم. مستنطق می‌تواند چنین پرسش‌هایی بپرسد: «ورق دست راستم از کدام خال است؟» انسان واجد تلپاتی یا غیب‌بینی ۱۳۰ بار برای ۴۰۰ کارت پاسخ درست می‌دهد. ماشین تنها می‌تواند تصادفا حدس بزند، و احتمالا ۱۰۴ پاسخ درست بدهد، در نتیجه مستنطق می‌تواند شناسایی درست را انجام بدهد.» امکان جالب‌توجهی وجود دارد که اینجا برملا می‌شود. فرض کنید کامپیوتر دیجیتال یک مولد اعداد تصادفی در خود دارد. پس طبیعی خواهد بود که از این مولد استفاده کند تا در مورد اینکه چه پاسخ‌هایی باید ارائه بدهد تصمیم بگیرد. اما بعد این مولد عدد تصادفی به قوای روان‌جنبشی مستنطق مقید خواهند بود. احتمالا این روان‌جنبش سبب شود که ماشین اغلب بیشتر از آنچه بنا بر یک محاسبه‌ی احتمال می‌توان انتظار داشت درست حدس بزند، طوری که مستنطق ممکن است هنوز نتواند تشخیص درست را داشته باشد. از طرف دیگر، ممکن است بتواند بدون هرگونه پرسش‌گری، از روی غیب‌بینی، درست حدس بزند. با ادراک فراحسی هر چیزی می‌تواند اتفاق بیافتد.

اگر تلپاتی تصدیق شود، آنگاه ضرورت دارد که برای تست‌مان سخت‌گیری بیشتری به خرج دهیم. این موقعیت می‌تواند معادل با موقعیتی در نظر گرفته شود که در صورتی رخ می‌دهد که مستنطق دارد با خودش حرف می‌زند و یکی از رقبای بازی دارد با گوش چسبیده به دیوار او را می‌شنود. قراردادن رقبا درون یک «اتاق اثبات تلپاتی» همه‌ی الزامات را برآورده خواهد کرد.

هفت. ماشین‌های یادگیری

خواننده پیشبینی کرده است که من هیچ استدلال قانع‌کننده‌ی ماهیتا مثبتی برای پشتیبانی از دیدگاهم ندارم. اگر می‌داشتم، متحمل چنین دردسرهایی نمی‌شدم تا به مغالطه‌های موجود در دیدگاه‌های مغایر اشاره کنم. گواهی را که در دست دارم حالا باید ارائه دهم.

اجازه دهید برای لحظه‌ای به اعتراض لیدی لاولیس برگردیم، که می‌گفت ماشین فقط کاری را انجام می‌دهد که ما به او بگوییم. می‌توان گفت که انسان می‌تواند ایده‌ای را به ماشین «تزریق» کند، و اینکه تا اندازه‌ای پاسخ خواهد داد و بعد خاموش خواهد شد، مثل رشته‌ای سیم پیانو که چکش به آن خورده باشد. تشبیه دیگر راکتوری اتمی با اندازه‌ای کوچکتر از میزان بحرانی است: یک ایده‌ی تزریق‌شده باید با نوترونی که از بیرون به راکتور وارد می‌شود متناظر باشد. هر یک نوترون اختلالی ایجاد خواهد کرد که نهایتا از بین می‌رود. بااین‌حال، اگر اندازه‌ی راکتور به‌قدر کافی افزایش یابد، اختلال ایجادشده با این نوترون ورودی به احتمال بسیار زیاد به‌طور پیوسته افزایش خواهد یافت تا وقتی که کل راکتور نابود شود. آیا پدیده‌ای متناظر برای ذهن‌ها وجود دارد، و برای ماشین‌ها چطور؟ قطعا به نظر می‌رسد که پدیده‌ای متناظر برای ذهن انسانی در کار باشد. اکثریت‌شان «غیربحرانی» به نظر می‌رسند، یعنی، در این قیاس، متناظر با راکتورهایی با اندازه‌ی زیربحرانی. ایده‌ای که به چنین ذهنی ارائه شده به‌طور میانگین به کمتر از یک ایده در جواب می‌انجامد. تناسبی نسبتا کوچک فرابحرانی است. ایده‌ای که به چنین ذهنی ارائه شده می‌تواند به «نظریه»ای کامل بیانجامد شامل ایده‌های ثانوی، ثالث، و بعیدتر. به نظر می‌رسد ذهن‌های حیوانی به‌نحوی کاملا معین زیربحرانی باشند. با وفاداری به این قیاس می‌پرسیم «آیا می‌توان ماشینی ساخت که فرابحرانی باشد؟»

قیاس «پوست پیاز» نیز سودمند است. با ملاحظه‌ی کارکردهای ذهن یا مغز عملیات‌های مشخصی را می‌یابیم که می‌توانیم به زبانی کاملا مکانیکی توضیح دهیم. می‌گوییم که این قیاس با ذهن واقعی تناظر ندارد: یک‌جور پوست است که باید بکنیم اگر می‌خواهیم از ذهن واقعی سردربیاوریم. اما بعد در باقیمانده‌اش پوست دیگری می‌یابیم که باید کنده شود، و الخ. با پیشروی در این مسیر آیا هرگز به ذهن «واقعی» می‌رسیم، یا آیا نهایتا به پوستی می‌رسیم که هیچ چیز در خودش ندارد؟ در مورد دوم کل ذهن مکانیکی است. (بااین‌همه ماشینی با حالت گسسته نخواهد بود. دراین‌باره قبلتر بحث کرده‌ایم.)

دو پاراگراف آخر داعیه‌ی استدلال‌هایی قانع‌کننده را ندارند. در عوض باید به‌عنوان «برخوانی‌هایی مایل به تولید باور» توصیف شوند.

تنها پشتیبان واقعا راضی‌کننده‌ای که می‌تواند برای دیدگاه بیان‌شده در ابتدای قسمت ۶ ارائه شود مشروط خواهد بود به اینکه تا انتهای این قرن صبر کنیم و سپس آزمایش توصیف‌شده را انجام دهیم. اما برای الان چه می‌توانیم بگوییم؟ چه قدم‌هایی باید اکنون برداشته شوند اگر بناست که آزمایش به موفقیت برسد؟

همان‌طور که توضیح دادم، مسئله اساسا مسئله‌ی برنامه‌ریزی است. پیشروی‌هایی نیز باید در مهندسی انجام گیرد، اما به نظر نامحتمل می‌رسد که این پیشرفت‌ها برای الزامات مسئله‌مان بسنده نباشند. تخمین‌ها درباره‌ی ظرفیت ذخیره‌سازی مغز بین 10^{10} تا 10^{15} رقم دودویی است. به مقادیر پایین‌تر تمایل دارم و معتقدم که فقط کسر بسیار کوچکی برای سنخ‌های سطح‌بالاتر تفکر مورد استفاده قرار می‌گیرد. اغلبش احتمالا برای نگهداری حس‌یافت‌های

بصری استفاده شود. مایه‌ی تعجبم می‌شود اگر بیش از $۱۰^{۹}$ برای اجرای راضی‌کننده‌ی بازی تقلید مورد نیاز باشد که به هر صورت در برابر انسانی کور انجام می‌شود. (یادداشت ـ ظرفیت نسخه‌ی ۱۱ام *دانش‌نامه‌ی بریتانیکا* $۱۰^{۹} \times ۲$ است.) ظرفیت ذخیره‌سازی $۱۰^{۷}$ می‌تواند امکانی بسیار عملی حتی برای فنون امروزی باشد. احتمالا ضروری ندارد که سرعت عملیات‌های ماشین را ابدا افزایش دهیم. اجزای ماشین‌های مدرن که می‌توانند به‌عنوان نظیرهای سلول‌های عصبی در نظر گرفته شوند هزار برابر سریع‌تر از سلول‌های مزبور عمل می‌کنند. این می‌بایست یک «حاشیه‌ی امن» به وجود بیاورد که می‌تواند کمبودهای سرعت را که به طرق بسیاری می‌تواند ایجاد شود پوشش دهد. این‌گونه، مسئله‌مان این است که چطور این ماشین‌ها را برای اجرای بازی برنامه بریزیم. من با آهنگ فعلی کارکردنم تقریبا هزار رقم برنامه در یک روز تولید می‌کنم اگرکه البته هیچ چیز روانه‌ی آشغال‌دانی نشده باشد، طوری که حدودا شصت کارگر که طی پنجاه سال به‌طور ثابت کار می‌کنند ممکن است این کار را انجام دهند. به نظر می‌رسد که روشی تندتر مطلوب باشد.

ما در فرایند تلاش برای تقلید از یک ذهن انسانی بزرگسال لازم داریم که بسیار زیاد درباره‌ی فرایندی فکر کنیم که آن را به وضعیتی که در آن قرار دارد رسانده است. می‌توانیم به سه مؤلفه اشاره کنیم:

الف. حالت اولیه‌ی ذهن، فرضا در زمان تولد

ب. تعلیم‌وتربیتی که ذهن به آن مقید شده

پ. تجارب دیگری که ذهن به آن مقید شده است و نباید درمقام تعلیم‌وتربیت توصیف شود

عوض تلاش برای تولید برنامه‌ای که ذهن بزرگسال را شبیه‌سازی کند، چرا در عوض نکوشیم برنامه‌ای تولید کنیم که ذهن کودک را شبیه‌سازی می‌کند؟ اگر این ذهن بدین‌ترتیب به یک دوره‌ی مناسب تعلیم‌وتربیتی مقید شود به ذهن بزرگسال خواهیم رسید. فرضا مغزـکودک چیزی مثل دفترچه‌یادداشت است که از مغازه‌ها می‌خریم. نسبتا اندکی مکانیزم، و انبوهی ورقه‌های خالی. (مکانیزم و نوشتار از منظر ما تقریبا مترادف هم‌اند.) امیدمان این است که مکانیزم بسیار اندکی در مغز بچه وجود داشته باشد تا چیزی مثل آن بتواند به‌آسانی برنامه ریخته شود. حجم کاری که می‌توانیم در این تعلیم‌وترتیب فرض بگیریم، با یک تقریب اولیه، با حجم کاری که باید برای تعلیم‌وترتیب کودک انسانی انجام داد بسیار مشابه است.

این‌گونه مسئله‌مان را به دو بخش تقسیم کرده‌ایم. برنامه‌ی کودک و فرایند تعلیم‌وتربیت. این‌دو در رابطه‌ای بسیار نزدیک با هم می‌مانند. نمی‌توانیم توقع داشته باشیم که در اولین تلاش‌مان یک ماشین کودک خوب بیابیم. باید آموزش یک چنین ماشینی را آزمایش کرد و دید که چطور خوب یاد می‌گیرد. بعد می‌توان ماشین کودک دیگری را امتحان کرد و دید آیا بهتر است یا بدتر. رابطه‌ی مشهودی بین این فرایند و تکامل وجود دارد اگر این موارد را با هم مترادف بگیریم:

ساختار ماشین کودک = ماده‌ی وراثتی

تغییرات ماشین کودک = جهش‌ها

انتخاب طبیعی = داوری آزمونگر

می‌توان بااین‌همه امیدوار بود که این فرایند تندتر از تکامل باشد. بقای اصلح روشی کند برای سنجش مزایاست. آزمونگر، با تمرین هوش، باید بتواند به فرایند سرعت بدهد. به همین میزان این واقعیت اهمیت دارد که او مقید به تحولات تصادفی نیست. اگر او بتواند علتی را برای ضعفی ردیابی کند احتمالا می‌تواند به آن نوع تحولی که بهبودش خواهد داد فکر کند.

ممکن نخواهد بود که فرایند آموزش یکسانی را برای ماشین و کودک عادی اعمال کرد. برای نمونه، پاهایی در اختیار نخواهد داشت و در نتیجه از او خواسته نخواهد شد که بیرون برود و جازغالی را پر کند. احتمالا چشم نخواهد داشت. اما هر قدر هم که ممکن است با مهندسی هوشمندانه بر این کمبودها غلبه کرد، نمی‌توان این مخلوق را به مدرسه فرستاد بدون آنکه دیگر کودکان بی‌نهایت او را دست نیاندازند. باید به او تدریس خصوصی داد. نیازی نداریم که خیلی نگران پاها، چشم‌ها، و مانند این‌ها باشیم. مثال میس هلن کلر نشان می‌دهد که تعلیم‌وتربیت می‌تواند به شرطی رخ دهد که ارتباط در هر دو جهت بین معلم و شاگرد به طریقی رخ بدهد.

معمولا تنبیهات و تشویقاتی را برای فرایند آموزش در نظر می‌گیریم. برخی ماشین‌های کودک ساده می‌توانند بر این مبنا ساخته یا برنامه‌ریزی شوند. این ماشین باید طوری ساخته شود که تکرار رویدادهایی که قدری مقدم بر سیگنال تنبیه‌اند نامحتمل باشد، درحالی‌که سیگتال تشویق احتمال تکرار رویدادهایی را که به آن انجامیده‌اند افزایش می‌دهد. این تعاریف هیچ احساسی را در طرف ماشین پیش‌فرض نمی‌گیرند. من آزمایش‌هایی را با یکی از این ماشین‌های کودک انجام داده‌ام، و موفق شدم که چیزهای کمی را به آن بیاموزم، اما روش آموزش برای آنکه واقعا موفقیت‌آمیز از کار دربیاید بسیار ناروال بود.

استفاده از تنبیهات و تشویقات در بهترین حالت می‌توانند بخشی از فرایند آموزش باشند. به بیانی سفت‌وسخت، اگر معلم هیچ راه دیگری برای ارتباط‌گرفتن با شاگرد نداشته باشد، میزان اطلاعاتی که می‌تواند به او برسد از کل تعداد تشویقات و تنبیهات به‌کارگرفته‌شده بیشتر نیست. وقتی یک کودک یاد گرفته «کازابیانکا» را تکرار کند احتمالا بسیار دمغ خواهد بود، اگر متن تنها بتواند با یک فن «بیست‌سئوالی» دریافت شود، هر «**نه**» شکل یک ضربه را به خود می‌گیرد. بنابراین ضرورت دارد که کانال‌های ارتباطی «غیرهیجانی» دیگری در اختیار داشت. اگر این کانال‌ها دردسترس باشند، می‌توان ماشین را از طریق تنبیهات و تشویقات آموزش داد که از دستورانی که در زبانی فرضا همچون زبانی نمادین به او داده شده تبعیت کند. این دستورات می‌بایست از خلال کانال‌های «غیرهیجانی» انتقال یابند. استفاده از این زبان تعداد تنبیهات و تشویقات مورد نیاز را تا حد بسیار زیادی کاهش خواهد داد.

عقاید راجع به پیچیدگی مناسب در ماشین کودک می‌تواند فرق کند. ممکن است سعی شود که آن را به ساده‌ترین حالت ممکن سازگار با اصول عام ساخت. به‌طور بدیل، ممکن است یک سیستم کامل از استنباط منطقی «تعبیه» شود. در مورد آخری، ذخیره تا اندازه‌ی زیادی با تعاریف و قضایا اشغال خواهد شد. قضایا می‌توانند انواع‌واقسام شأن‌های مختلف را داشته باشند، برای مثال واقعیات تثبیت‌شده، حدس‌وگمان‌ها، فرضیه‌های ریاضیاتی اثبات‌شده، جمله‌هایی از طرف یک مرجع، بیان‌هایی با فرم منطقی قضیه اما نه با ارزش باور. برخی قضایا می‌توانند به‌صورت «احکام» توصیف شوند. ماشین باید طوری ساخته شود که

به محض اینکه حکمی درمقام «تثبیت‌شده» طبقه‌بندی شد عمل متناسب به‌طور خودکار رخ بدهد. برای ترسیم این نکته، فرض کنید معلم به ماشین می‌گوید «تکلیفت را الان انجام بده.» این سبب می‌شود که معلم می‌گوید «تکلیفت را الان انجام بده»» در زمره‌ی واقعیات تثبیت‌شده گنجانده شود. واقعیت دیگری از این سنخ می‌تواند «همه‌ی گفته‌های معلم درست است» باشد. ترکیب این‌ها نهایتا به حکم «تکلیفت را الان انجام بده» می‌انجامد که در زمره‌ی واقعیات تثبیت‌شده گنجانده شده است و این با ساخت ماشین یعنی تکلیف در عمل شروع شده است، اما اثرش بسیار راضی‌کننده است. نیازی نیست که فرایندهای استنباط مورد استفاده توسط ماشین طوری باشند که دقیق‌ترین منطقدان‌ها را راضی کنند. برای نمونه هیچ سلسله‌مراتبی از تایپ‌ها در کار نخواهد بود. اما، همان‌قدر که به‌ناچار از پرتگاه‌های حصاربندی‌نشده به پایین سقوط می‌کنیم، لازم است که معنی جمله‌ی قبلی این نباشد که غلط‌های تایپی رخ نخواهند داد. احکام مناسب (که *درون* سیستم‌ها بیان شده‌اند و بخشی از قواعد *متعلق به* سیستم را شکل نمی‌دهند) مثل «از یک طبقه استفاده نکن مگر زیرطبقه‌ی طبقه‌ای باشد که معلم به آن اشاره کرده است» می‌تواند اثری مشابه بر «خیلی به لبه نزدیک نشو» داشته باشد.

لازم است آن احکامی که یک ماشین فاقد اندام‌ها می‌تواند ازشان تبعیت کند خصیصه‌ای فکری داشته باشند، همچون در مثال (انجام تکلیف) ارائه‌شده در بالا. احکام مهم بین چنین احکامی آن‌هایی هستند که دستوری را رگوله یا تنظیم می‌کنند که قواعد سیستم منطقی مد نظر باید در آن به کار بسته شود. چراکه در هر مرحله وقتی داریم از سیستمی منطقی استفاده می‌کنیم، شمار بسیار زیادی از گام‌های بدیل وجود دارند که تا آنجا که اطاعت از قواعد سیستم منطقی مد نظر باشد مجاز هستیم هر کدام‌شان را به کار ببریم. این انتخاب‌ها تفاوت بین یک استدلال‌کننده‌ی درخشان و استدلال‌کننده‌ی پیش‌پاافتاده را ایجاد می‌کند و نه تفاوت بین یک استدلال‌کننده‌ی مستدل و یک استدلال‌کننده‌ی مغلطه‌آمیز. قضایایی که به احکامی از این نوع می‌انجامند می‌توانند این‌طور باشند: «وقتی به سقراط اشاره شد، قیاس را در باربارا به کار ببر» یا «اگر اثبات شد که یک روش سریع‌تر از روش دیگر است، از روش کندتر استفاده نکن». برخی از این قضایا می‌توانند «از طرف مرجع صادر شوند» اما دیگر قضایا می‌تواند از طرف خود ماشین، فرضا با استقرای علمی، تولید شده باشد.

ممکن است ایده‌ی ماشین یادگیری برای برخی خواننده‌ها متناقض‌نمایانه به نظر برسد. قواعد عملیاتی ماشین چطور می‌توانند عوض شوند؟ آن‌ها باید به‌طور کامل توصیف کنند که ماشین با هر تاریخچه‌ای، با هر تغییری که می‌تواند متحملش شود، چطور واکنش نشان خواهد داد. قواعد این‌گونه به‌هیچ‌وجه با زمان تغییر نمی‌کنند یا نامتغیر-با-زمان‌اند. این گفته کاملا درست است. توضیح متناقض‌نما این است که قواعدی که طی فرایند یادگیری عوض شده‌اند چندان متکلف و مدعیانه نیستند و صرفا داعیه‌ی یک صحت یا اعتبار منطقی گذرا را دارند. خواننده می‌تواند تناظری با قانون اساسی ایالات متحده برقرار کند.

یک ویژگی مهم ماشین یادگیری این است که معلمش اغلب عمیقا نامطلع است از اینکه چه دارد درون ماشین اتفاق می‌افتد گرچه هنوز می‌تواند تا اندازه‌ای رفتار شاگردش را پیشگویی کند. این نکته باید قویا برای تعلیم‌وتربیت متعاقب ماشینی که از یک ماشین کودک

خوب‌طراحی‌شده (یا خوب‌برنامه‌ریزی‌شده) ناشی می‌شود به کار بسته شود. این نکته در مغایرت آشکار است با روندی عادی که از ماشین برای انجام محاسبه‌ها استفاده می‌کند: پس هدف این است که تصویر ذهنی روشنی از حالت ماشین در هر لحظه از محاسبه داشته باشیم. این هدف فقط با مبارزه می‌تواند حاصل شود. این دیدگاه که «ماشین فقط می‌تواند کاری را انجام دهد که ما می‌دانیم انجامش را چطور به آن دستور بدهیم» در برابر هدف یادشده عجیب به نظر می‌رسد. اغلب برنامه‌هایی که می‌توانیم در ماشین قرار بدهیم به انجام چیزی از طرف ماشین می‌انجامد که نمی‌توانیم هیچ فهمی از آن داشته باشیم، یا به‌عنوان رفتار کاملا تصادفی تلقی می‌کنیم. رفتار هوشمند احتمالا عبارت است از عزیمت از رفتار کاملا منضبط مشمول در محاسبه، اما عزیمتی کوچک، که به رفتار تصادفی یا به لوپ‌های تکراری بیهوده منجر نمی‌شود. یک نتیجه‌ی مهم دیگر آماده‌سازی ماشین‌مان برای ایفای نقش خود در بازی تقلید از راه فرایند آموزش و یادگیری این است که «خطاپذیری انسانی» احتمالا می‌بایست به طریقی نسبتا طبیعی، یعنی بدون هرگونه «مربی‌گری» به‌خصوص، کنار گذاشته شود. (خواننده باید این نکته را با نظرگاه ص۴۵۴ تطبیق بدهد.) فرایندهایی که یاد گرفته شده‌اند نتیجه‌ای با قطعیت صددرصد تولید نمی‌کنند؛ اگر این‌طور بودند نمی‌توانستند فراموش شوند.

احتمالا معقول است که عنصری تصادفی را در یک ماشین یادگیری بگنجانیم (نگاه کنید ص۴۴۵). یک عنصر تصادفی وقتی مفید است که در جستجوی راه‌حلی برای یک مسئله‌ایم. برای نمونه فرض کنیم می‌خواستیم عددی بین ۵۰ تا ۲۰۰ را بیابیم که با مجذور جمع ارقامش معادل باشد، می‌توانیم با ۵۱ شروع کنیم بعد ۲۵ را امتحان کنیم و ادامه بدهیم تا به عددی برسیم که جواب بدهد. به‌طور بدیل، می‌توانیم عددها را تصادفی انتخاب کنیم تا به عددی درست برسیم. مزیت روش یادشده در این است که دیگر ضرورتی ندارد حساب مقادیری را که امتحان شده‌اند پیش خود نگه داشت، اما این اشکال را دارد که ممکن است یک عدد را دو بار امتحان کنیم، اما این هم خیلی مهم نیست اگر راه‌حل‌های متعددی در کار باشند. روش نظام‌مند این اشکال را دارد که ابتدا باید مورد کنکاش قرار بگیرد. حال فرایند یادگیری درمقام جستجویی برای شکلی از رفتار در نظر گرفته می‌شود که معلم (یا سنجه‌ای دیگر) را راضی خواهد کرد. از آنجا که احتمالا شمار بسیار زیادی راه‌حل راضی‌کننده وجود دارد، به نظر می‌رسد که روش تصادفی بهتر از روش نظام‌مند باشد. باید خاطرنشان کرد که این روش در فرایند همسان تکامل استفاده شده است. اما روش نظام‌مند در آنجا ممکن نیست. چگونه می‌توان حساب ترکیب‌های ژنتیکی متفاوت را که امتحان شده‌اند پیش خود حفظ کرد تا از تکرار دوباره‌شان اجتناب شود؟

می‌توانیم امیدوار باشیم که ماشین‌ها نهایتا در تمام زمینه‌های مطلقا فکری با انسان‌ها رقابت خواهند کرد. اما بهترین زمینه‌ها برای شروع کدام‌ها هستند؟ حتی این هم تصمیمی دشوار است. بسیاری از مردم فکر می‌کنند که یک فعالیت بسیار انتزاعی، مثل شطرنج‌بازی‌کردن، بهترین خواهد بود. همچنین می‌توان در نظر داشت که عالی است اگر بهترین اندام‌های حسی که می‌توان با پول تهیه کرد برای ماشین فراهم شوند، و بعد به ماشین آموخت که انگلیسی بفهمد و حرف بزند. این فرایند می‌تواند آموزش عادی یک کودک را پی بگیرد. چیزها نشان

داده خواهند شد و نام خواهند گرفت، و غیره. دیگربار نمی‌دانم پاسخ درست چیست، اما فکر می‌کنم هر دو رویکرد را باید امتحان کرد.

فقط می‌توانیم فاصله‌ای کوتاه در پیش روی خود ببینیم، اما می‌توانیم کارهای زیادی را ببینیم که باید انجام شوند.

ماشین‌آلات هوشمند، نظریه‌ای ارتدادی (۱۹۵۱)

نمیتوانید ماشینی بسازید که برای‌تان فکر کند. این تصور پیش‌پاافتاده‌ای است که معمولا بی‌چون‌وچرا پذیرفته میشود. این مقاله میخواهد آن تصور را به پرسش بکشد.

هدف اغلب ماشین‌آلات ایجادشده برای مقاصد تجاری اجرای کاری بسیار خاص با قطعیت و سرعت قابل‌ملاحظه است. اکثر اوقات این ماشین‌آلات رشته‌ای از عملیات‌های یکسان را بدون هرگونه تنوعی پشت سر هم انجام میدهد. این واقعیت درباره‌ی ماشین‌آلات واقعی دردسترس از دید بسیاری استدلال قدرتمندی به نفع شعار نقل‌شده در بالاست. این استدلال برای یک منطقدان ریاضیاتی قابل‌استفاده نیست، زیرا نشان داده شده که از حیث نظری ماشین‌هایی میتوانند وجود داشته باشند که چیزی بسیار نزدیک به فکرکردن را انجام دهند. آن‌ها برای نمونه صحت یک برهان صوری در سیستم *اصول ریاضیاتی* را تست میکنند یا حتی میگویند که فرمولی از آن سیستم اثبات‌پذیر است یا اثبات‌ناپذیر. در حالتی که فرمول نه اثبات‌پذیر باشد نه اثبات‌ناپذیر، چنین ماشینی قطعا به شیوه‌ای رضایت‌بخش رفتار نمیکند، چراکه بدون تولید هر نوع نتیجه‌ای به‌نحوی نامعین به فعالیت ادامه میدهد، اما این را نمیتوان بسیار متفاوت از واکنش ریاضیدان‌ها در نظر گرفت که برای مثال برای صدها سال در مورد این پرسش کار کرده‌اند که آیا آخرین فرضیه فرمات درست یا نه. در مورد ماشین‌هایی از این سنخ یک استدلال ظریف‌تر ضرورت دارد. با فرضیه مشهور گودل یا استدلالی مشابه میتوان نشان داد که ماشین به هر طریقی ساخته شده باشد حدودوثغوری برای مواردی وجود دارند که در آن‌ها ماشین نمیتواند پاسخی ارائه دهد اما یک ریاضیدان خواهد توانست. از طرف دیگر، ماشین مزیت‌های خاصی بر ریاضیدان دارد. میتوان با فرض عدم هرگونه فروشکنی مکانیکی به هرآنچه ماشین انجام میدهد اطمینان کرد، درحالی‌که ریاضیدان مقادیر مشخصی اشتباه دارد. به باورم این خطر ریاضیدانی که مرتکب اشتباه میشود پیامد اجتناب‌ناپذیر توان او در مواجهه‌ی گاه‌وبی‌گاه با روشی کاملا تازه است. به نظر میرسد این نکته با واقعیتی مشهور تصدیق شده باشد که براساسش قابل‌اطمینان‌ترین آدم‌ها معمولا به روش‌های واقعا تازه برنخواهند خورد.

موضوع بحثم این است که میتوان ماشین‌هایی ساخت که رفتار ذهن انسانی را با دقت بالا شبیه‌سازی کنند. گاه مرتکب اشتباه خواهند شد، و گاه جملات تازه و بسیار جالبی ایجاد خواهند کرد، و در کل خروجی‌شان به قدر خروجی ذهن انسانی ارزش توجه دارد. محتوای این جمله در فرکانس بالاتری مستتر است که از جملات درست انتظار میرود، و به گمانم نمیتوان جمله‌ای دقیق به آن ارائه داد. برای مثال، بسنده نیست به‌سادگی بگوییم که ماشین همه نوع جمله‌ی درست را دیر یا زود خواهد ساخت، چراکه مثالی از چنین ماشینی میتواند ماشینی باشد که دیر یا زود تمام جملات ممکن را میسازد. ما میدانیم چطور این جملات را بسازیم، و وقتی آن‌ها (احتمالا) به یک میزان جملات درست و غلط را مکررا تولید کنند احکام‌شان

کاملا بی‌ارزش خواهد بود. این واکنش واقعی یا بالفعل ماشین به شرایط است که موضوع بحثم را به اثبات خواهد رساند اگرکه در حقیقت هرگز بتواند اثبات شود.

اجازه دهید با احتیاط بیشتری به ماهیت این برهان بپردازیم. آشکارا میتوان ماشینی تولید کرد که اگر به‌طور کافی ساخته‌وپرداخته شده باشد برای هر گستره‌ای از تست‌ها شرح بسیار خوبی از خودش به دست بدهد. بااین‌همه، این هم دیگربار دشوار بتواند درمقام برهانی بسنده در نظر گرفته شود. چنین ماشینی با ارتکاب مکرر اشتباهی یکسان خودش را بروز میدهد، و دیگر اصلا نمیتواند خودش را اصلاح کند یا با استدلالی از بیرون اصلاح شود. اگر این ماشین به طریقی میتوانست از راه تجربه بیاموزد اثرگذاری بسیار بیشتری میداشت. اگر نکته این باشد آنگاه به نظر میرسد هیچ دلیل واقعی وجود نداشته باشد که چرا نباید از یک ماشین نسبتا ساده شروع کرد، و با مقیدکردنش به دامنه‌ای مناسب از تجربه آن را به ماشینی بدل ساخت که ساخته‌وپرداخته‌تر بوده باشد، و بتواند به گستره‌ای بسیار بزرگتر از امکان‌های خاص تصادفی بپردازد. این فرایند احتمالا بتواند با انتخاب مناسب تجاربی که ماشین به آن‌ها مقید شده بود تسریع گردد. این را میتوان آموزش نامید. اما اینجا باید مراقب باشیم. خیلی آسان بتوان تجارب را به طریقی ترتیب داد که به‌طور خودکار سبب شوند ساختار ماشین به صورتی از قبل قصدشده ساخته شود، و این آشکارا شکل فاحش نیرنگ است، تقریبا هم‌تراز وجود یک انسان درون ماشین. اینجا دیگربار سنجه‌ی اینکه نمیتواند با ضوابطی ریاضیاتی نشان داد که چه چیزی میتواند در مسیر آموزش عقلانی در نظر گرفته شود، اما پیشنهاد میدهم که ادعای پیش رو میتواند در عمل بسنده باشد. بیایید فرض کنیم که میخواهیم ماشین زبان انگلیسی را بفهمد، و به خاطر اینکه فاقد دست یا پا است، نیازی به خوردن ندارد، و از میل به سیگارکشیدن عاری است، پس اغلب وقتش صرف بازی‌های شطرنج و گو و احتمالا بریج میشود. یک کیبورد تایپ در اختیار این ماشین قرار میگیرد که هر علامتی به آن رویش تایپ شده است، و همچنین هر علامتی که دوست دارد بسازد بر آن تایپ شده است. پیشنهاد میدهم آموزش این ماشین باید به معلمی بسیار قابل محول شود که به این پروژه علاقه دارد اما از هر شناخت جزئی از کارکردهای داخلی این ماشین منع شده است. بااین‌حال مکانیکی که این ماشین را ساخته مجاز است ماشین را در حال فعالیت حفظ کند، و اگر ظن ببرد که ماشین عملکرد نادرستی داشته است آن را به یکی از جایگاه‌های قبلی بازگرداند و از معلم بخواهد که درس‌هایش را از آن وهله به بعد تکرار کند، اما او هیچ نقشی در تدریس ایفا نمیکند. از آنجا که این روند فقط در خدمت تست حسن‌نیت مکانیک خواهد بود، نمیتوانم بگویم که در مراحل تجربی اتخاذ نخواهد شد. از منظر من، این فرایند آموزشی یک الزام برای تولید یک ماشین به‌طور عقلانی هوشمند درون یک بازه زمانی به‌طور عقلانی کوتاه است. قیاس انسانی به‌تنهایی چنین پیشنهادی دارد.

حالا نشانگری ارائه میدهم از شیوه‌ای که انتظار میرود این ماشین براساسش عمل کند. ماشین یک حافظه را شامل میشود. این به توضیح چندان زیادی نیاز ندارد. این امر به‌سادگی فهرستی از همه‌ی جملاتی خواهد بود که برای این ماشین یا توسطش ساخته شده‌اند، و نیز فهرستی از همه‌ی حرکاتی که انجام داده بود و ورق‌هایی که در بازی‌هایش بازی کرده بود. این فهرست در ترتیبی گاه‌شناختی تنظیم خواهد شد. علاوه بر این حافظه‌ی سرراست و مستقیم

شماری از نمایه‌های تجارب در کار خواهد بود. برای توضیح این ایده فرمی را که هر کدام از آن نمایه‌ها ممکن است به خود بگیرد پیشنهاد خواهم داد. میتواند نمایه‌ای الفبایی از کلمات باشد که برای ارائه‌ی زمان‌هایی استفاده شده‌اند که آن کلمات در آن وهله‌ها مورد استفاده قرار گرفته‌اند، طوری که بتوانند در حافظه جستجو شوند. نمایه‌ای دیگری از این سنخ میتواند شامل الگوهای انسان‌ها باشد که در بخش‌های یک صفحه‌ی گو روی داده بودند. در مراحل نسبتا متاخر آموزش، حافظه میتواند بسط یابد تا بخش‌های مهم پیکربندی ماشین در هر لحظه را شامل شود، یا به عبارت دیگر کم‌کم به یاد میآورد که افکارش چه بودند. این به فرم‌های ثمربخش تازه‌ی نمایه‌سازی خواهد انجامید. فرم‌های تازه‌ی نمایه میتوانند بر اساس ویژگی‌های خاص مشاهده‌شده در نمایه‌هایی قبلتر استفاده‌شده معرفی شوند. نمایه‌ها بدین‌طریق مورد استفاده قرار خواهند گرفت. هرآینه انتخابی باید صورت بگیرد که حرکت بعدی چه خواهد بود، ویژگی‌های موقعیت فعلی در نمایه‌های دردسترس جستجو میشوند، و انتخاب پیشین در موقعیت‌های مشابه، و خروجی، خوب یا بد، کشف میشود. این انتخاب جدید به همین منوال انجام میگیرد. این نکته شماری از مسائل را مطرح میکند. چه باید کرد اگر برخی نشانگرها مطلوب و برخی نامطلوب باشند؟ پاسخ به این پرسش احتمالا از یک ماشین به ماشین دیگر فرق خواهد کرد و با درجه‌ی آموزشش هم تغییر خواهد داشت. در ابتدا احتمالا قاعده‌ای کاملا خام کافی خواهد بود، برای مثال انجام آنچه بزرگ‌ترین عدد آرا را به سود خودش دارد. در متاخرترین مرحله‌ی آموزش کل پرسش از فرایند در چنین مواردی احتمالا از طرف خود ماشین و به‌وسیله‌ی یک‌جور نمایه کندوکاو خواهد شد و این هم به فرمی عمیقا پیچیده و امیدواریم عمیقا رضایت‌بخش از قاعده بیانجامد. بااین‌همه احتمالش هست که خود فرم‌های نسبتا خام قاعده به‌نحوی عقلانی رضایت‌بخش باشند طوری که پیشرفت روی‌هم‌رفته میتواند به رغم خامی قواعد انتخاب انجام شود. به نظر میرسد نکته‌ی مزبور را بتوان با این واقعیت معتبر دانست که مسائل مهندسی گاه با خام‌ترین قاعده‌ی روند تورق حل میشوند که صرفا با سطحی‌ترین جوانب مسئله سروکار دارند، فرضا اینکه یک تابع با یکی از متغیرهایش افزایش مییابد یا کاهش. این تصویر از شیوه‌ای که رفتار تعین مییابد مسئله‌ی دیگری را مطرح میکند: ایده‌ی خروجی مطلوب. بدون چنین ایده‌ای، متناظر با اصل لذت روانشناس‌ها، بسیار دشوار بتوان نحوه‌ی پیشروی را دریافت. قطعا معرفی چنین چیزی به ماشین کاملا طبیعی است. پیشنهاد میدهم که باید دو کلید وجود داشته باشند که معلم بتواند دستکاری‌شان کند و نشانگر ایده‌های لذت و درد باشند. در مراحل آخر آموزش، ماشین شرایط مشخص دیگری را به خاطر اینکه در گذشته دائما با لذت ربط داشته‌اند مطلوب شناسایی میکند و به همین ترتیب برخی دیگر را نامطلوب. برخی تجلیات خشم از طرف معلم میتوانند برای مثال به قدری شوم و ترسناک تشخیص داده شوند که هرگز نتوان از آن‌ها چشم پوشید، طوری که معلم درمییابد کاربست ترکه‌ی تنبیه دیگر نالازم شده است.

ارائه‌ی پیشنهادات بیشتر در راستای همین خطوط احتمالا در این مرحله ثمربخش نخواهد بود، چراکه احتمالا عبارت‌اند از تحلیل روش‌های بالفعل یا واقعی آموزش به‌کاربسته‌شده برای فرزندان انسان. بااین‌حال خصیصه‌ای که دوست دارم پیشنهاد بدهم باید در ماشین‌ها گنجانده شود، و آن هم عنصر تصادفی است. هر ماشین باید با نواری حاوی یک سری

تصادفی از ارقام مثلا صفر و یک در مقادیر برابر تامین شود و این سری ارقام باید در گزینه‌هایی که ماشین انتخاب میکند استفاده شوند. نتیجه‌اش این است که رفتار ماشین به‌هیچ‌رو کاملا با تجاربی که به آن‌ها مقید شده بود تعین نیافته است، و وقتی با آن دست به آزمایش میزنیم استفاده‌های ارزشمندی دارد. با جعل‌کردن انتخاب‌های انجام‌گرفته خواهیم توانست توسعه‌ی ماشین را تا حدی مهار کنیم. برای مثال چه‌بسا اصرار شود که انتخاب انجام‌گرفته انتخابی به‌خصوص در فرضا ۱۰ جای به‌خصوص است و این یعنی تقریبا یک ماشین با مبنای ۱۰۲۴ یا بیشتر توسعه خواهد یافت که درجه‌اش به بزرگی درجه‌ی ماشینی است که جعل شده بود. این واقعا نمیتواند جمله‌ای دقیق فرض گرفته شود چراکه ماهیت ذهنی ایده‌ی درجه‌ی توسعه هیچ چیز از این واقعیت نمیگوید که ماشین جعل‌شده چه‌بسا می‌توانست در انتخاب‌های جعل‌نشده‌اش هم خوش‌اقبال بوده باشد.

حالا اجازه دهید محض خاطر استدلال فرض کنیم که این ماشین‌ها یک امکان راستین‌اند، و نگاهی به عواقب ساختن‌شان بیاندازیم. انجام چنین کاری قطعا با مخالفت بزرگی رویارو خواهد شد، مگر اینکه در بردباری دینی نسبت به روزگار گالیله بسیار ترقی کرده باشیم. روشنفکرانی که نگران ازدست‌دادن شغل‌شان هستند مخالفت بزرگی خواهند داشت. گرچه محتمل است که روشنفکران در این زمینه‌ها مرتکب اشتباه شده باشند. کارهای بسیار زیادی برای انجام‌دادن خواهد بود (تلاش برای فهم اینکه ماشین‌ها چه داشتند میگفتند)، برای مثال در تلاش برای ارتقای هوش‌مان تا معیاری که ماشین‌ها وضع کرده‌اند، چون محتمل به نظر میرسد که وقتی روش فکرکردن ماشین آغاز شد پشت‌سرگذاشتن قوای سست‌مان خیلی طول نخواهد کشید. هیچ پرسشی از مردن ماشین‌ها در کار نخواهد بود، و آن‌ها خواهند توانست با یکدیگر گفتگو کنند تا ذکاوت‌شان را هشیارتر سازند. پس در مرحله‌ای باید توقع داشته باشیم که ماشین‌ها اداره‌ی امور را به دست بگیرند، به همان شیوه‌ای که در *ارهون* ساموئل باتلر اشاره شده است.

آیا کامپیوترهای دیجیتال میتوانند فکر کنند؟ (۱۹۵۱)

کامپیوترهای دیجیتال اغلب درمقام مغزهای مکانیکی توصیف شده‌اند. اکثر دانشمندان احتمالا این توصیف را به‌صورت یک نمایش روزنامه‌ای در نظر میگیرند، اما برخی هم نه. یک ریاضیدان نظرگاهی مخالف را به‌نحوی محکم با این کلمات ابراز کرده است: معمولا گفته شده که این ماشین‌ها مغز نیستند اما تو و من میدانیم که آن‌ها مغزند. در این سخنرانی تلاش میکنم ایده‌های پس پشت نظرگاه‌های ممکن متعدد را توضیح دهم، هرچند نه روی‌هم‌رفته به‌طور کامل. بیش از همه توجهم را معطوف منظری میکنم که خودم به آن باور دارم، اینکه اصلا معقول نیست ماشین‌های دیجیتال را به‌صورت مغز توصیف کنیم. پروفسور هارتری قبلتر نظرگاه متفاوتی را مطرح کرده است.

ابتدا نظرگاه خام آدم معمولی در خیابان را در نظر میگیریم. او شرح‌های حیرت‌آوری از توانایی‌های این ماشین‌ها میشنود: اغلب‌شان ظاهرا مشغول اعجازهایی فکری‌اند که او در قبال‌شان کاملا ناتوان است. او این نکته را تنها میتواند با این فرض توضیح دهد که ماشین یک‌جور مغز است، گرچه ترجیح میدهد به‌سادگی به شنیده‌هایش باوری نکند.

اکثریت دانشمندان از این نگرش تقریبا خرافه‌ای بیزارند. آن‌ها از اصولی مطلع‌اند که ماشین‌ها بر مبنایشان ساخته میشوند و از شیوه‌ای مطلع‌اند که ماشین‌ها مورد استفاده قرار میگیرند. لیدی لاولیس حدود صد سال پیش با صحبت درباره‌ی موتور تحلیلی ببیج دیدگاه‌شان را به‌خوبی خلاصه کرده بود. همان‌طور که هارتری قبلتر نقل آورده، او میگوید موتور تحلیلی هیچ ادعایی مبنی بر *زایش* چیزی ندارد. این موتور میتواند هرآنچه را که *ما میدانیم چطور به آن فرمان* دهیم اجرا کند. این نقل‌قول شیوه‌ای را که کامپیوترهای دیجیتال عملا در زمان حاضر مورد استفاده قرار میگیرند و نیز شیوه‌ای را که احتمالا برای سال‌های پیش رو مورد استفاده قرار خواهند گرفت بسیار خوب توصیف میکند. برای هر تک‌محاسبه، کل روندی که ماشین باید متحملش بشود از قبل توسط ریاضیدانی طرح ریخته میشود. هرچه تردید کمتری وجود داشته باشد که چه اتفاقی قرار است بیافتد، ریاضیدان هم بهتر به هدفش رسیده است. این مثل طراحی عملیاتی نظامی است. تحت این شرایط، درست است گفته شود که ماشین هیچ چیز نمیزاید.

بااین‌همه نظرگاه سومی هم وجود دارد که خودم به آن معتقدم. با اظهارنظر لیدی لاولیس در تمامیتش موافق‌ام، اما به باورم صحتش به این ملاحظه بستگی دارد که چطور کامپیوترهای دیجیتال مورد استفاده قرار *گرفته‌اند* و نه اینکه چطور *میتوانند* مورد استفاده قرار بگیرند. در واقع به‌گمانم آن‌ها را میتوان به طریقی مورد استفاده قرار داد که به‌نحو متناسبی بتوانند درمقام مغز توصیف شوند. همچنین باید بگویم که اگر هر ماشینی بتواند به‌طور متناسبی یک مغز وصف شود، آنگاه هر کامپیوتر دیجیتال میتواند به همین ترتیب توصیف شود.

این جمله‌ی آخر به توضیحاتی نیاز دارد. چه‌بسا تکان‌دهنده به نظر برسد اما با ملاحظاتی به نظر میرسد که واقعیتی اجتناب‌ناپذیر باشد. میتوان نشان داد که از یک ویژگی سرشت‌نمای

کامپیوترهای دیجیتال پیروی کند، که *جهانشمولیت*شان خواهم خواند. یک کامپیوتر دیجیتال یک ماشین *جهانشمول* است به این معنا که میتواند ساخته شود تا به جای هر ماشینی از هر رده‌ی بسیار گسترده بنشیند. این ماشین جانشین یک بولدوزر یا موتور بخار یا تلسکوپ نخواهد شد، اما جانشین هر طراحی رقیب از ماشین‌حساب خواهد شد، یعنی هر ماشینی که بتواند داده دریافت کند و بعدتر نتایجی چاپ کند. برای آنکه کامپیوترمان را برای سرمشق‌گرفتن از یک ماشین مفروض آرایش دهیم صرفا ضرورت دارد تا کامپیوتر را برای محاسبه‌ی آنچه ماشین مد نظر تحت شرایط مفروض انجام خواهد داد و خصوصا برای محاسبه‌ی آنچه چاپ خواهد کرد برنامه بریزیم. پس کامپیوتر میتواند برای چاپ پاسخ‌های یکسان ساخته شود.

اگر حالا یک ماشین به‌خصوص بتواند به‌صورت مغز توصیف شود صرفا باید کامپیوتر دیجیتال‌مان را در راستای سرمشق‌گرفتن از آن برنامه بریزیم و این نیز یک مغز خواهد بود. اگر بپذیریم که مغزهای واقعی، چنان‌که در حیوان‌ها و خصوصا در انسان‌ها یافت شده، سنخی از ماشین‌اند، آن‌وقت نتیجه خواهیم گرفت که کامپیوتر دیجیتال‌مان که به‌طور مناسبی برنامه‌ریزی شده همچون یک مغز رفتار خواهد کرد.

این استدلال مفروضات متعددی را شامل میشود که میتوانند به‌طور معقولی به چالش کشیده شوند. قبلتر توضیح داده‌ام که آن ماشینی که باید تقلید شود می‌بایست بیشتر به ماشین‌حساب شبیه باشد و نه بولدوزر. این گفته صرفا بازتاب این واقعیت است که داریم از آنالوگ‌های مکانیکی مغزها حرف میزنیم و نه از پاها یا پنجه‌ها. همچنین ضرورت داشت که این ماشین از سنخی باشد که رفتارش در اصل از طریق محاسبه قابل‌پیشگویی است. قطعا نمیدانیم چطور هر کدام از چنین محاسباتی باید انجام شوند، و حتی سر آرتور ادینگتون استدلال کرده که این پیشگویی در مورد اصل عدم‌تعین در مکانیک کوآنتوم حتی از نظر نظری هم محال است.

فرض دیگر این بود که ظرفیت کامپیوتر مورد استفاده باید کافی باشد تا پیشگویی رفتار ماشینی که باید تقلید شود انجام گیرد. درعین‌حال باید سرعت کافی هم داشته باشد. کامپیوترهای فعلی‌مان احتمالا به ظرفیت ذخیره‌ی ضروری نرسیده‌اند، گرچه بااین‌همه سریع هستند. این یعنی در عمل اگر بخواهیم هر چیزی به پیچیدگی مغز انسان را تقلید کنیم به ماشینی بسیار عظیم‌تر از تمام کامپیوترهای کنونی دردسترس نیاز داریم. احتمالا به چیزی حداقل صد برابر عظیم‌تر و به بزرگی کامپیوتر منچستر نیاز داریم. به‌طور بدیل، قطعا ماشینی با اندازه‌ی برابر یا کوچک‌تر جواب خواهد داد اگر پیشروی کافی در تکنیک ذخیره‌سازی اطلاعات انجام شده باشد.

باید اشاره شود که هیچ نیازی نیست پیچیدگی کامپیوترهای مورد استفاده افزایش یابد. اگر سعی داشته باشیم که از ماشین‌ها یا مغزهایی حتی پیچیده‌تر تقلید کنیم باید از کامپیوترهای هرچه‌عظیم‌تر برای این کار استفاده کنیم. نیاز نداریم تا به‌طور متوالی از کامپیوترهای پیچیده‌تر استفاده کنیم. این میتواند متناقض‌نمایانه به نظر برسد اما توضیحش دشوار نیست. تقلید یک کامپیوتر از یک ماشین مستلزم این است که نه فقط باید کامپیوتر را بسازیم بلکه همچنین

باید آن را به‌طور متناسب برنامه بریزیم. هر چه پیچدگی ماشینی که باید تقلید شود بیشتر باشد برنامه هم باید پیچیده‌تر باشد.

این نکته احتمالا با یک قیاس روشن‌تر شود. دو نفر را در نظر بگیرید که هر دو میخواهند خودشرح‌حال‌نگاری‌شان را بنویسند و یکی از آن‌ها زندگی پرحادثه‌ای داشته است، اما اتفاقات بسیار کمی برای دیگری افتاده است. دو مشکل در کار خواهد بود که فرد دارای زندگی پرحادثه‌تر را جدی‌تر از دیگری آزار میدهد. او باید بیشتر وقت صرف نوشتن کند و برای فکرکردن به آنچه میخواهد بگوید باید پذیرای دردسر بیشتری باشد. تأمین کاغذ احتمالا مشکلی جدی نیست، مگر اینکه برای مثال در جزیره‌ای متروک باشد، و در هر صورت تنها میتواند مسئله‌ای فنی یا مالی باشد. مشکل دیگر بنیادی‌تر است و جدی‌تر هم خواهد شد اگر او مشغول نوشتن نه زندگی‌اش بلکه اثری درباره‌ی چیزی باشد که هیچ از آن نمیداند، فرضا درباره‌ی زندگی خانوادگی در مریخ. مسئله‌ی ما برای برنامه‌ریزی یک کامپیوتر برای آنکه مثل یک مغز رفتار کند به چیزی همچون نوشتن رساله‌ای درباره‌ی جزیره‌ای متروک میماند. نمیتوانیم ظرفیت ذخیره‌ی مورد نیازمان را به دست آوریم: به عبارت دیگر نمیتوانیم کاغذ کافی برای نوشتن رساله بر روی آن را پیدا کنیم، و در هر صورت نمیدانیم اگر کاغذ میداشتیم چه باید ثبت میکردیم. این وضعیت بسیار فقیر است اما اگر به قیاس ادامه دهیم چیزی است برای اطلاع از اینکه چطور باید نوشت، و چطور باید قدردان این واقعیت بود که اغلب دانش میتواند در کتاب‌ها تجسد یابد.

با دیدگاه مزبور به نظر میرسد که این بخردانه‌ترین زمینه‌ای است که بر مبنایش میتوان توصیف کامپیوترهای دیجیتال درمقام مغزهای مکانیکی یا درمقام مغزهای الکترونیک را به نقد کشید، گرچه ممکن است برنامه‌ریزی شوند تا مانند مغزها رفتار کنند. در حال حاضر نمیدانیم که این چطور باید انجام شود. با این منظر کاملا موافقم. این منظر این پرسش را برای‌مان باز میگذارد که آیا نهایتا خواهیم توانست یا نخواهیم توانست در یافتن چنین برنامه‌ای موفق شویم. به‌شخصه تمایل دارم باور کنم که چنین برنامه‌ای یافت خواهد شد. فکر میکنم احتمال دارد برای مثال در پایان این قرن برنامه‌ریزی ماشینی برای پاسخ‌دادن به پرسش‌ها به طریقی ممکن شود که بی‌اندازه دشوار بتوان حدس زد که پاسخ‌ها را انسان داده است یا ماشین. دارم چیزی مثل آزمونی شفاهی را در تصور میکنم اما با پرسش‌ها و پاسخ‌هایی که همگی تایپ‌شده‌اند تا در نتیجه به موضوعات نامربوطی همچون مطابقت با اصل که صدای انسان با آن بتواند تقلید شود نیاز نداشته باشیم. این تنها نشانگر عقیده‌ی من است؛ فضاهای بسیار زیادی برای دیگران وجود دارد.

هنوز دشواری‌هایی وجود دارند. به نظر میرسد رفتارکردن مثل یک مغز اراده‌ی آزاد را شامل شود، اما رفتار یک کامپیوتر دیجیتال، وقتی برنامه‌ریزی شده باشد، کاملا تعین‌یافته است. این دو واقعیت باید تا اندازه‌ای با هم وفق یابند، اما به نظر میرسد که انجام چنین کاری ما را مشغول مناقشه‌ای قدیمی کند، یعنی مناقشه‌ی اراده‌ی آزاد و جبرگرایی. دو راه پیش روی ماست. میتواند اینطور باشد که احساس اراده‌ی آزاد که همه‌مان داریم یک وهم است. یا میتواند اینطور باشد که ما واقعا اراده‌ی آزاد داریم، اما هنوز راهی نداریم که از روی رفتارمان بگوییم که اینطور است. در مورد آخر، هرقدر ماشین از رفتار انسان خوب تقلید کند باید

درمقام یک چیز ساختگی یا ظاهری صرف در نظر گرفته شود. نمیدانم چطور میتوانیم هرگز بین این بدیل‌ها تصمیم بگیریم اما هر کدام که بدیل درست باشد حتما ماشینی که بناست از مغزی تقلید کند باید طوری رفتار کند که توگویی اراده‌ای آزاد دارد، و به‌درستی پرسیده خواهد شد که چطور قرار است به چنین چیزی برسیم. یک امکان این است که رفتارش را به چیزی مثل یک چرخ رولت یا ذخیره‌ی رادیوم منوط کنیم. رفتار این‌ها احتمالا میتواند قابل‌پیشگویی باشد، اما اگر این‌طور باشد، نمیدانیم چطور پیشگویی را انجام دهیم.

بااین‌حال واقعا حتی ضرورت ندارد این کار را انجام دهیم. دشوار نیست طراحی ماشین‌هایی که رفتارشان برای هر کسی که از جزئیات ساختمان‌شان ناآگاه است کاملا تصادفی باشند. هر قدر هم که شمول این عنصر تصادفی ماهیتا کافی باشد، هر فنی هم که مورد استفاده قرار بگیرد، باز هم مسئله‌ی اصلی ما را حل نمیکند، اینکه چطور ماشین را برای تقلید از مغز، یا به بیانی احتمالا مؤجزتر و بااین‌حال با دقت کم‌تر، برای فکرکردن برنامه بریزیم. اما این نکته سرنخی به دست میدهد از اینکه فرایند به چه شباهت خواهد داشت. نباید همیشه توقع داشته باشد که بدانیم کامپیوتر قرار است چه کاری انجام بدهد. وقتی ماشین مایه‌ی تعجب‌مان میشود باید راضی باشیم، با همان رضایتی که وقتی مردمک کاری میکند که برای انجامش مشخصا آموزش ندیده بود.

حال بیایید اظهارنظر لیدی لاولیس را در نظر بگیریم. ماشین میتواند هرآنچه را که ما میدانیم چطور به آن فرمان دهیم انجام دهد. معنای باقی قطعه طوری است که وسوسه میشویم بگوییم که ماشین *تنها* میتواند آنچه را که میدانیم چطور فرمان اجرایش را به آن بدهیم انجام دهد. اما فکر میکنیم این گزاره صحت ندارد. قطعا ماشین تنها میتواند آنچه را که به‌حتم فرمان اجرایش را میدهیم انجام دهد، هر چیز دیگر اشتباهی مکانیکی است. اما هیچ نیازی نیست فرض کنیم که وقتی به ماشین فرامینش را میدهیم میدانیم که چه کاری داریم انجام میدهیم و عواقب این فرمان‌ها چه میتوانند باشند. نیاز نیست بتوانیم بفهمیم که چطور این فرامین به رفتار متعاقب ماشین منجر میشوند، همانقدر که وقتی بذری را در زمین میکاریم نیاز نیست مکانیزم رویش را بفهمیم. گیاه چه بفهمیم چه نفهمیم رشد میکند. اگر برنامه‌ای به ماشین بدهیم که به انجام چیزی جالب‌توجه بیانجامد که پیشبینی نکرده بودیم، آن‌وقت تمایل دارم بگویم که ماشین زاینده‌ی چیزی *بوده است*، عوض آنکه مدعی شوم رفتارش تلویحا در برنامه وجود داشته است و در نتیجه‌اش اصالت (زایشی‌اش) تماما به ما بستگی دارد.

سعی نخواهم کرد حرف چندانی درباره‌ی این مسئله نزنم که این فرایند برنامه‌ریزی ماشین برای فکرکردن چطور باید انجام شود. واقعیت این است که ما بسیار کم از این مسئله میدانیم، پژوهش‌های بسیار کمی تا به امروز انجام شده‌اند. ایده‌های بی‌شماری وجود دارند، اما هنوز نمیدانیم کدام‌شان مهم‌اند. مثل داستان‌های کاراگاهی، در ابتدای پژوهش هر چیز بی‌اهمیت میتواند برای پژوهنده مهم باشد. وقتی مسئله حل شد، فقط واقعیات اساسی را باید به هیئت منصفه گفت. اما در حال حاضر هیچ چیز ارزشمندی برای ارائه به هیئت منصفه نداریم. فقط این نکته را خواهم گفت که به باورم این فرایند باید نسبت نزدیکی با فرایند تدریس داشته باشد.

سعی کردم استدلال‌های عقلانی اصلی له و علیه این نظریه را که میتوان ماشین‌هایی برای فکرکردن ساخت توضیح دهم اما باید چیزی هم درباره‌ی استدلال‌های غیرعقلانی گفت. بسیاری از مردم با ایده‌ی ماشینی که فکر میکند عمیقا تعارض دارند اما فکر نمیکنیم که این تعارض به خاطر هیچ کدام از علت‌هایی که ارائه کرده‌ام یا به خاطر هر نوع علت معقول باشد، اما به‌سادگی بابت این است که آن‌ها خود این ایده را دوست ندارند. میتوان ویژگی‌های زیادی را دید که این ایده را نامطلوب میسازند. اگر یک ماشین بتواند فکر کند، آنگاه میتواند معقول‌تر از ما فکر کند، و آن‌وقت چه بر سر ما خواهد آمد؟ حتی اگر بتوانیم ماشین‌ها را در جایگاهی خدمتگزار نگه داریم، برای مثال با خاموش‌کردن‌شان در لحظات استراتژیک، آنگاه باید درمقام یک گونه عمیقا خوار شده باشیم. این امکان که خوک یا موش از انسان جلو بزنند ما را با خطر و تحقیر مشابه‌ای تهدید میکند. این یک امکان نظری است که دشوار بتواند مناقشه‌برانگیز باشد اما به‌قدری برای سالیان دراز با خوک‌ها و موش‌ها زندگی کرده‌ایم بدون آنکه هوش‌شان افزایش زیادی پیدا کند که دیگر خودمان را با این امکان به دردسر نمیاندازیم. حس میکنیم که اگر هرگز چنین اتفاقی بیافتد برای چند میلیون سال پیش رو نخواهد بود. اما این خطر تازه بسیار نزدیک‌تر است. اگر هم هرگز پیش بیاید قطعا طی هزاره‌ی بعدی خواهد بود. این اتفاق بعید است اما به‌نحوی نجومی بعید نیست، و قطعا چیزی است که میتواند مایه‌ی اضطراب ما شود.

در یک سخنرانی یا مقاله راجع به این موضوع رایج است که ذره‌ای آسودگی پیش گذاشته شود، آن‌هم به صورت جمله‌ای که براساسش ماشین نمیتواند برخی از خصایص خصوصا انسانی را هرگز تقلید کند. احتمالا برای مثال گفته شود که هیچ ماشینی نمیتواند انگلیسی درست را بنویسد، یا اینکه نمیتواند تحت تاثیر جاذبه‌ی جنسی قرار بگیرد یا پیپ بکشد. نمیتوانم هیچ نوع آسایشی ارائه دهم، چراکه به‌گمانم هیچ نوع حدومرزی از این سنخ نمیتواند برقرار شود. اما قطعا امیدوارم و باور دارم که هیچ تلاش بزرگی صرف ساخت ماشین‌هایی با خصایص خصوصا انسانی اما غیرفکری همچون شکل بدن انسانی نخواهد شد؛ انجام چنین تلاش‌هایی به نظرم بی‌ثمر میرسد و نتایج‌شان کیفیت ناخوشایند گل‌های مصنوعی را خواهد داشت. به نظرم تلاش‌ها برای تولید یک ماشین متفکر در مقوله‌ای متفاوت قرار بگیرد. کل فرایند تفکر هنوز برای ما تا اندازه‌ای رازآلود است، اما به باورم تلاش برای ساخت یک ماشین متفکر به کشف نحوه‌ی فکرکردن خودمان کمک زیادی خواهد کرد.

آیا می‌شود گفت که ماشین‌های محاسبه‌گر خودکار قابلیت تفکر دارند؟ (۱۹۵۲)

ریچارد برایتویت: امروز اینجا هستیم تا درباره‌ی این موضوع بحث کنیم که آیا میشود گفت که ماشین‌های محاسبه‌گر میتوانند به هر معنای درست کلمه فکر کنند یا نه. معمولا تفکر اغلب اوقات به‌قدری به‌عنوان ویژگی انسان و احتمالا ویژگی دیگر حیوانات برتر در نظر گرفته شده است که خود بحث بر سر این پرسش میتواند پوچ به نظر برسد. اما قطعا همه‌اش بستگی به این دارد که چه چیز بناست تفکر به حساب آید. این کلمه معمولا انبوهی از فعالیت‌های متفاوت را پوشش میدهد. جفرسون تو به‌عنوان یک فیزیولوژیست مهمترین عناصر مشمول در تفکر را چه میدانی؟

جفری جفرسون: فکر نمیکنم که نیاز داشته باشیم زمان خیلی زیادی را روی تعریف تفکر صرف کنیم زیرا ورای عبارات موجود در کاربرد متعارف دشوار بتوانیم به آن برسیم، چیزهایی مثل داشتن ایده‌هایی در ذهن، اندیشیدن، تأمل‌کردن، تحویل‌دادن، حل‌کردن مسائل یا خیال‌کردن. فیزیولوژیست‌ها میگویند که کلمه‌ی «**انسان**» از کلمه‌ای سانسکریت به معنی «فکرکردن» مشتق میشود، احتمالا به معنای داوری بین یک ایده و ایده‌ای دیگر. موافق‌ام که دیگر نمیتوانیم کلمه‌ی «فکرکردن» را به معنایی محدود به انسان به‌کار ببریم. هیچ کسی انکار نخواهد کرد که حیوانات بسیاری فکر میکنند، گرچه به شیوه‌ای بسیار محدود. آن‌ها فاقد بینش‌اند. برای مثال، سگ میآموزد که اشتباه است که با پنجه‌های گل‌آلود روی بالش‌ها یا صندلی‌ها برود، اما او این را تنها به‌عنوان عمل مخاطره‌آمیزی که به صلاح نیست میآموزد. او هیچ فهمی از دلیل واقعی‌اش ندارد، اینکه او با انجام چنان کاری به پارچه‌ها آسیب میزند.

آدم عادی احتمالا راضی است که تفکر را با الفاظی بسیار عام درمقام ایده‌هایی چرخان در ذهن تعریف کند، به‌عنوان داشتن انگاره‌هایی در سرمان، درمقام مشغول‌شدن ذهن‌مان با یک مسئله، و الخ. اما فقط درست است اضافه کنیم که ذهن‌هایمان اغلب اوقات با امور مبتذل مشغول میشوند. چه‌بسا دست آخر گفته شود که تفکر نتیجه‌ی عمومی داشتن یک سیستم عصبی به‌قدرکافی پیچیده است. سیستم‌های بسیار ساده مسائلی را که با مکانیزم‌های بازتابی ساده پاسخ داده نمیشوند هرگز در اختیار مخلوق نمیگذارند. فکرکردن اینگونه به همه‌ی چیزهایی بدل میشود که در مغز ما میگذرد، چیزهایی که اغلب به یک عمل میانجامند اما ضرورتا آن را انجام نمیدهد. باید بگویم که تفکر سرجمع فعالیت مغز انسان یا حیوان است. تورینگ، تو دراین‌باره چه فکر میکنی؟ آیا تعریفی مکانیکی داری؟

آلن تورینگ: نمیخواهم تعریفی از تفکر به دست بدهم، اما اگر میبایست این کار را انجام بدهم احتمالا نمیتوانم هیچ چیزی درباره‌اش بگویم مگر اینکه مثل یک‌جور وزوز بود که به کله‌ام وارد شد. اما واقعا فکر نمیکنم که اصلا باید روی یک تعریف به توافق برسیم. مهم این است که خطی بکشیم بین خصایص یک مغز، یا یک انسان، که میخواهیم درباره‌اش بحث کنیم، و آن خصایصی که نمیخواهیم درباره‌شان بحث کنیم. اگر یک مورد افراطی را در نظر بگیریم، علاقه‌ای به این واقعیت نداریم که مغز انسجام هلیم سرد را دارد. نمیخواهیم بگوییم

«این ماشین بسیار سخت است، پس یک مغز نیست، و بنابراین نمیتواند فکر کند.» مایل‌ام یک تست خاص را پیشنهاد بدهم که میتوان برای یک ماشین به‌کار بست. میتوانید آن را تستی بخوانید که با آن میتوان دریافت که آیا ماشین فکر میکند یا نه، اما بهتر است از مسلم‌فرض‌گرفتن مسئله اجتناب کنیم، و بگوییم ماشین‌هایی که به‌درستی از پس این تست برمیآیند (اجازه دهید بگوییم) ماشین‌های «**درجه یک**» هستند. ایده‌ی تست این است که ماشین باید با پاسخ به پرسش‌های مطروحه برای آن سعی و تظاهر کند که یک انسان باشد، و فقط به شرطی از پس تست برخواهد آمد که تظاهرش به‌نحوی معقول متقاعدکننده باشد. ترکیب قابل‌ملاحظه‌ای از هیئت منصفه، که نباید متخصص ماشین‌ها باشند، باید تظاهر ماشین را ملاحظه کند. آن‌ها اجازه ندارند خود ماشین را ببینند ــ دیدن ماشین کارشان را بسیار راحت خواهد کرد. بنابراین ماشین در اتاقی دور نگه داشته میشود و هیئت اجازه دارند تا از ماشین سئوال بپرسند، طوری که سئوال‌ها به ماشین منتقل میشوند: ماشین پاسخ را به‌صورت تایپ‌شده پس میفرستد.

برایتویت: آیا پرسش‌ها می‌بایست حساب جمع باشند، یا آیا میتوانم از ماشین بخواهم بگوید که برای صبحانه چه خورده است؟

تورینگ: آه بله، هر نوع پرسشی. و پرسش‌ها واقعا نباید پرسش باشند، همان‌قدر که پرسش‌ها در محکمه‌ای قانونی واقعا پرسش نیستند. در جریان هستید. این گفته که «به خودتان محول میکنم که صرفا دارید تظاهر میکنید که یک انسان باشید» هیچ اشکالی ندارد. به همین منوال، ماشین مجال اجرای همه‌ی انواع ترفندها را دارد تا بیش از پیش به انسان شباهت داشته باشد، مثلا قبل از پاسخ قدری صبر کند، یا در هجی‌کردن اشتباه کند، اما نمیتواند روی کاغذ لکه بیاندازد، همان‌قدر که نمیتوانیم لکه‌ها را با تلگراف ارسال کنیم. بهتر بود فرض میکردیم که هر هیئت منصفه می‌بایست چندین دفعه اقدام به داوری کند، و گاهی واقعا با انسان و نه ماشین سروکار دارند. این مانع‌شان میشود که هربار بدون ملاحظه‌ای مناسب بگویند «این باید یک ماشین باشد». خب، این تست من است. قطعا در حال حاضر نمیگویم که یا ماشین‌ها واقعا به‌درستی از پس تست برمی‌آیند، یا اینکه نمیتوانند. پیشنهادم صرفا این است که باید این پرسش را به بحث بگذاریم. این پرسش معادل نیست با این گفته که «ماشین‌ها حتما فکر میکنند»، اما به نظرم برای مقاصد کنونی‌مان تقریبا کافی به نظر میرسد، و دشواری‌های مشابهی را طرح میکند.

ماکس نیومن: دوست دارم وقتی هماوردی‌تان بین انسان و ماشین روی می‌دهد حضور داشته باشیم، و احتمالا به سهم خودم بکوشم برخی از پرسش‌ها را طرح کنم. اما اگر قرار است این ماشین بدون هیچ پرسش قدغن‌شده‌ای هرگونه شانسی داشته باشد، پس این اتفاق بسیار دور از امروز خواهد بود؟

تورینگ: آه بله، باید بگویم حداقل ۱۰۰ سال.

جفرسون: نیومن، ماشین‌های موجود این تست را چقدر خوب تاب خواهند آورد؟ الان چه جور کارهایی را میتوانند انجام دهند؟

نیومن: قطعا، قوی‌ترین خط‌شان محاسبه‌ی ریاضیاتی است، که برای‌اش طراحی شده‌اند، اما همچنین از پس پرسش‌هایی که عددی به نظر نمی‌رسند هم خوب برمی‌آیند، اما به‌آسانی

می‌توانند برای چنین مقاصدی ساخته شوند، مثل حل‌کردن یک مسئله‌ی شطرنج یا یافتن قطاری برای‌تان در جدول زمانی.

برایتویت: میتوانند این کار را انجام دهند؟

نیومن: بله. هر دو این کارها میتواند با آزمودن همه‌ی امکان‌ها انجام شود، یکی پس از دیگری. کل اطلاعات در یک جدول زمانی معمولی می‌بایست به‌عنوان بخشی از برنامه نوشته شود، و ساده‌ترین روتین ممکن آن است که قطارهای لندن به منچستر را با آزمودن هر قطار در جدول زمانی پیدا کند تا ببیند آیا در هر دو محل جواب میدهد یا نه، و آن‌هایی را که درست است پرینت بگیرد. البته، این روشی ملال‌آور و خسته‌کننده است، و می‌توانید با استفاده از یک روتین پیچیده‌تر بهبودش بدهید، اما اگر تست تورینگ را به‌درستی فهمیده باشم، مجاز نیستید که به ورای صحنه‌ها بروید و روش را به نقد بکشید، بلکه باید از نمره‌گرفتن از پاسخ‌های صحیحی که به‌طور مستدل سریع یافت شده‌اند پیروی کنید.

جفرسون: بله، اما بااین‌همه انسانی که باید مکررا به دنبال قطارها باشد با آموختن نحوه‌ی سروکله‌زدن با جدول زمانی حرکت قطارها در انجام کارش پیشرفت میکند. فرض کنید مسئله‌ای یکسان را دوباره به یک ماشین ارائه می‌دهم، آیا می‌تواند بیآموزد که این کار را بهتر انجام دهد بدون اینکه به کل دنگ‌وفنگ آزمودن هرباره‌ی همه چیز دربیافتد؟ دوست دارم پاسخ‌تان به این پرسش را بدانم چراکه نکته‌ی بسیار مهمی است. آیا ماشین‌ها می‌توانند با تمرین بیآموزند که کارشان را بهتر انجام بدهند؟

نیومن: بله، می‌تواند. احتمالا مسئله‌ی شطرنج تصویر بهتری از این موضوع ارائه دهد. اول باید اشاره کنم که هر اطلاعات لازم در هر کاری ـ عددها، زمان‌های قطارها، جایگاه‌های تکه‌ها، یا هرچه که هست، و همچنین دستوراتی که می‌گویند با آن‌ها چه باید کرد ـ تمام این مصالح به شیوه‌ای یکسان ذخیره می‌شوند. (در ماشین منچستر این اطلاعات به‌صورت الگویی ذخیره شده است که به چیزی مانند یک صفحه‌نمایش تلویزیونی شباهت دارد.) با پیشروی کار، الگو عوض می‌شود. معمولا این بخشی از الگوست که شامل داده‌ای می‌شود که تغییر می‌کند، درحالی‌که دستورات ثابت باقی می‌مانند. اما به‌قدری ترتیب‌دادن به آن ساده است که خود دستورات می‌بایست همواره عوض شوند. خب، حال برنامه‌ای می‌تواند ساخته شود که ماشین را به انجام این کار وادارد: یک مسئله‌ی شطرنج با دو حرکت با یک کدگذاری مناسب در ماشین ثبت می‌شود، و هرآینه ماشین شروع به کار کرد یک حرکت مهره‌ی سفید به‌تصادف انتخاب می‌شود (در ماشین ما وسیله‌ای برای اجرای انتخاب‌های تصادفی وجود دارد). تمام عواقب این حرکت حالا تحلیل می‌شوند، و اگر این حرکت به مات الزامی در دو حرکت *نیانجامد*، ماشین این پرینت را فرضا می‌دهد: «P–Q3، حرکت اشتباه»، و متوقف می‌شود. اما تحلیل نشان می‌دهد که وقتی حرکت درست انتخاب شده باشد ماشین نه تنها فرضا می‌گوید «B–Q5، جواب»، بلکه دستورات لازم برای انتخاب تصادفی را به دستوری عوض می‌کند که می‌گوید: «حرکت به B–Q5 را امتحان کن». نتیجه‌اش این است که هرآینه ماشین از نو شروع به کار می‌کند بی‌درنگ راه‌حل یا جواب درست را پرینت می‌گیرد و به خروجی می‌فرستد ـ و این هم بدون انسانی اتفاقی می‌افتد که روتین را از کار درآورده است

و از قبل می‌داند که این روتین چه بود. چنین روتینی قطعا می‌تواند حالا ساخته شود، و فکر می‌کنم این را به‌درستی می‌توان یادگیری خواند.

جفرسون: بله، به گمانم همین است. موجودات انسانی با تکرار تمرین‌های یکسان، تا وقتی تکمیل‌شان کرده باشند، یاد می‌گیرند. قطعا موضوع جلوتر از این‌ها می‌رود، و همزمان به‌طور گسترده یاد می‌گیریم که دانش تحصیل‌شده از یک چیز را به مجموعه‌مسائل دیگر منتقل کنیم، و خط‌وربط‌ها و نسبت‌های بین‌شان را ببینیم. یادگیری یعنی یادآوری. یک ماشین تا چه زمانی می‌تواند اطلاعات را ذخیره کند؟

نیومن: آه، حداقل به قدر طول عمر یک انسان، اگرکه متناوبا رفرش یا تازه شود.

جفرسون: تفاوت دیگر این است که در فرایند یادگیری مداخله‌ای مکررتر از طرف معلم‌ها، پدرانه یا غیر از آن، در کار است که تدابیر یادگیری را پیش می‌برند. شما ریاضیدان‌ها برنامه را یک‌بار در ماشین طرح می‌ریزید و به آن محولش می‌کنید. اگر این کاری است که انجام دادید، آن‌وقت هیچ فاصله‌ای از موجودات انسانی نمی‌گیرید. در واقع، تنها باری که چنین کاری را در دوره‌ی آموزش انجام می‌دهید وقت امتحانات است.

تورینگ: کاملا درست است که وقتی یک کودک آموزش داده می‌شود، والدین و معلمانش مکررا در کارش مداخله می‌کنند تا مانع او از انجام این کار شوند یا تشویقش می‌کنند آن کار را انجام دهد. اما این نکته برای آموختن یک ماشین به‌هیچ‌عنوان صدق نمی‌کند. آزمایش‌هایی در آموختن یک ماشین برای انجام عملیاتی ساده انجام دادم، و نیاز بود که بخش بسیار بزرگی از این مداخله قبل از دریافت هرگونه نتیجه‌ای انجام گیرد. به عبارت دیگر، ماشین به‌قدری به‌آرامی یاد گرفت که به میزان بالایی آموزش نیاز داشت.

جفرسون: اما چه کسی داشت یاد می‌گرفت، شما یا ماشین؟

تورینگ: خب، به گمانم هر دو ما داشتیم یاد می‌گرفتیم. اگر بناست به هرگونه موفقیت واقعی برسیم، باید دریافت که چطور ماشین‌هایی بسازیم که سریع‌تر می‌آموزند. همچنین امیدواریم که یک‌جور گسترش سریع در کار باشد. ماشین هرچه چیزهای بیشتری یاد گرفته باشد آسان‌تر می‌تواند مهیای یادگیری چیزهای دیگر باشد. ماشین با یادگیری انجام هر چیز به‌خصوص احتمالا یاد خواهد گرفت که مؤثرتر بیاموزد. تمایل دارم بر این باور باشم که وقتی آموختیم برخی چیزها را انجام بدهیم، درخواهیم یافت چیزهای دیگری که برنامه داریم تا به ماشین بیاموزیم بدون هرگونه آموزش لازم خاصی روی می‌دهد. این اتفاق به‌حتم با یک ذهن انسانی هوشمند روی می‌دهد، و اگر این اتفاق ضمن آموزش‌دادن یک ماشین نمی‌افتد، پس چیزی در این ماشین کم است. برایتویت، تو در مورد آموختن امکان‌ها چه فکر می‌کنی؟

برایتویت: هیچ کسی به دشواری بزرگ درباره‌ی یادگیری در نظرم اشاره نکرده است زیرا صرفا درباره‌ی یادگیری در جهت حل مسئله‌ای به‌خصوص بحث کرده‌ایم. اما مهم‌ترین بخش یادگیری انسان یادگیری از تجربه است ـ نه یادگیری از یک نوع خاص تجربه، بلکه توانایی یادگیری از تجربه به‌طور عام. ماشین به‌آسانی می‌تواند با یک وسیله‌ی فیدبک ساخته شود طوری که برنامه‌ریزی ماشین با نسبت خروجی‌اش با ویژگی‌ای در محیط بیرونی‌اش کنترل می‌شود ـ طوری که فعالیت ماشین در نسبت با محیط خوداصلاح‌گر است. اما این مستلزم آن است که این ویژگی خاص باید از محیطی باشد که ماشین باید خودش را با آن تنظیم کند.

ویژگی خاص انسان‌ها و حیوان‌ها این است که آن‌ها توان تطبیق خودشان با تقریبا تمام خصایص را دارند. ویژگی‌ای که تطبیق با آن در وهله‌ای خاص انجام می‌گیرد همان است که به آن توجه می‌شود و او به آنچه به آن *علاقه* دارد توجه می‌کند. علایق او روی‌هم‌رفته معین‌اند، بابت اشتیاق‌هایش، میل‌هایش، رانه‌هایش، غریزه‌هایش ـ همه‌ی چیزهایی که همراه با هم «جنب‌وجوش‌ها»ی او را می‌سازند. اگر می‌خواهیم ماشینی بسازیم که توجه‌اش به چیزها در محیطش را تغییر دهد طوری که گاهی خودش را با یک چیز و گاهی با چیزی دیگر تنظیم کند، به نظر ضروری می‌رسد که ماشین را به چیزی متناظر با مجموعه‌ای از اشتیاق‌ها مجهز کنیم. اگر ماشین ساخته شد تا صرفا به‌عنوان حیوانی خانگی تلقی شود، و مسائلی به‌خصوص به او خورانده شود، قادر نخواهد بود به شیوه‌ای متغیر که موجودات انسانی با آن یاد می‌گیرند یاد بگیرد. این از ضرورت تطبیق مناسب رفتار با محیط ناشی می‌شود اگرکه اشتیاق‌های انسانی می‌بایست ارضا شوند.

جفرسون: تورینگ، تو با اطمینان زیاد از چیزی گفتی که قرار است قادر به انجامش باشی. طوری بیانش کردی که به نظر می‌رسد به‌آسانی بتوان ساختمان یا سرشت آن را اصلاح کرد به نحوی که این ماشین بیشتر مثل یک انسان واکنش نشان می‌دهد. اما من این را از دوره‌ی دکارت و بورلی به یاد می‌آورم که درباره‌ی آدم‌ها گفته‌اند که این امر صرفا چند سال، احتمالا ۳ یا ۴ یا شاید ۵۰ سال، زمان ببرد، و یک المثنا یا نسخه‌ی عین انسان به‌طور مصنوعی خلق خواهد شد. مطمئن‌ام که می‌توانیم مرتکب اشتباه شویم اگر این برداشت را بدهیم که انجام این کارها آسان خواهد بود.

نیومن: موافق‌ام که داریم از ماشین‌های محاسبه‌گر موجود امروزی دور می‌شویم. این ماشین‌ها اشتیاق‌های نسبتا محدودی دارند، و وقتی دچار مشکل شدند نمی‌توانند شرمگین شوند، اما این به‌قدر کافی خیلی سخت و به گمانم مسئله‌ای بسیار جالب است که دریابیم تا چه اندازه این ماشین‌های عملا موجود می‌توانند در مسیر فکرکردن پیش بروند. حتی اگر به بخش استدلالی تفکر متوسل شویم، راه درازی از حل مسئله‌ی شطرنج تا ابداع مفاهیم ریاضیاتی تازه یا ایجاد یک عمومیت‌یابی یا تعمیم داریم که مشتمل بر ایده‌هایی باشند که قبلتر وجود داشتند، اما هرگز به‌صورت وهله‌های یک انگاره‌ی عام واحد یک جا جمع نشده بودند.

برایتویت: مثلا؟

نیومن: انواع مختلف عدد. اعداد صحیح، ۰، ۱، ۲ ـ، و الخ؛ اعداد حقیقی مورد استفاده برای مقایسه‌ی طول‌ها، برای مثال محیط یک دایره و قطرش؛ و اعداد مختلط شامل $\sqrt{-1}$ و غیره. به‌هیچ‌وجه بدیهی نیست که این موارد وهله‌های یک چیز یعنی «عدد» باشند. ریاضیدان‌های یونانی از کلمات کاملا متفاوتی برای اعداد صحیح و اعداد حقیقی بهره می‌برند، و هیچ ایده‌ی واحدی برای پوشش‌دادن یا شمول هر دو ندارند. واقعا فقط همین اواخر بود که انگاره‌ی عام انواع عدد از این وهله‌ها منفک شد و با دقت تعریف شد. برای ایجاد این نوع تعمیم نیاز دارید که قدرت لازم برای تشخیص شباهت‌ها و دیدن قیاس‌های تشابهی بین چیزهایی را که قبلتر یک‌جا کنار هم جمع نشده بودند داشته باشید. مسئله صرفا بر سر تست‌کردن چیزها بابت یک ویژگی به‌خصوص و دسته‌بندی‌شان به همین منوال نیست.

خود این مفهوم باید در قالبی قرار بگیرد، چیزی باید خلق شود، فرضا ایده‌ی میدان عددی. آیا حتی می‌توانیم شیوه‌ای را حدس بزنیم که ماشین می‌تواند از روی برنامه‌ای دست به چنین ابداعی بزند که انسانی سازنده‌اش است که این مفهوم را در ذهن خودش نداشته است؟

تورینگ: نیومن، به نظرم می‌رسد که گفته‌ات درباره‌ی «آزمودن امکان‌ها» درمقام روش تا اندازه‌ای کاربرد دارد، حتی وقتی ماشین نیازمند انجام چیزی است که به‌قدر یافتن یک مفهوم تازه‌ی مفید پیشرفته است. نه دوست دارم تعریفی از معنای کلمه‌ی «مفهوم» ارائه بدهم، نه حتی مایل‌ام قواعدی برای رتبه‌بندی سودمندی‌شان فراهم آورم، اما هرچه هستند فرم‌هایی بیرونی و مرئی دارند، که کلمه‌ها و ترکیب‌های کلمه‌ها هستند. یک ماشین می‌تواند این ترکیب‌های کلمات را کمابیش به‌صورت تصادفی بسازد، و بعد بابت ارزش‌های مختلف‌شان نشان‌هایی به‌شان بدهد

نیومن: آیا زمان بازدارنده‌ی زیادی نخواهد برد؟

تورینگ: حتما به‌نحوی شوکه‌کننده آرام خواهد بود، اما می‌تواند با چیزهای ساده‌ای آغاز شود، مثل جمع‌کردن باران، تگرگ، برف، و برفابه کنار هم، ذیل کلمه‌ی «بارندگی». احتمالا بعدتر کارهای دشوارتری را انجام بدهد اگر تمام‌وقت یاد بگیرد که چطور روش‌هایش را ارتقا بدهد یا بهبود ببخشد.

برایتویت: فکر نمی‌کنم برای مشاهده‌ی قیاس‌های تشابهی که می‌توانند به‌طور صوری تحلیل شوند و به‌طور واضح بیان شوند دشواری زیادی وجود داشته باشد. پس مسئله صرفا بر سر طراحی ماشین است طوری که بتواند شباهت‌های ساختار ریاضیاتی را تشخیص بدهد. دشواری به شرطی ظهور می‌کند که این قیاس تشابهی مبهم باشد طوری که درباره‌اش اندکی بیشتر از آن قیاس تشابهی دیگری بتوان گفت که این احساس را درباره‌اش داریم که شباهتی بین این دو مورد وجود دارد اما هیچ ایده‌ای در قبال جنبه‌ای نداریم که این دو مورد در آن با هم شباهت دارند. ماشین نمی‌تواند شباهت‌ها را تشخیص بدهد وقتی هیچ چیز در برنامه‌اش وجود ندارد که بگوید این شباهت‌هایی که انتظار می‌رود تشخیص‌شان بدهد چه هستند.

تورینگ: فکر می‌کنم بتوانید ماشینی بسازید که قیاس تشابهی را بفهمد یا بشناسد، در واقع این نمونه‌ای بسیار خوب است از اینکه چطور ماشینی می‌تواند ساخته شود که برخی از آن چیزهایی را که معمولا به‌عنوان انحصار اساسا انسانی در نظر گرفته می‌شوند انجام بدهد. فرض کنید که فردی در این تلاش بود تا امر منفی مضاعف را به من توضیح بدهد که برای مثال وقتی چیزی نه سبز نیست پس باید سبز باشد، و درعین‌حال نتواند از پس آن برآید. چه‌بسا بگوید «خب، مثل عبور از رودخانه است. از آن رد می‌شوید، و بعد دوباره از آن رد می‌شوید، و به جایی می‌رسید که از آن آغاز به رفتن یا عبور کرده بودید.» این نکته احتمالا به‌نحوی قاطع موضوع را اثبات کند. این یکی از چیزهایی است که باعث می‌شود دوست داشته باشید با ماشین‌ها کار کنید، و فکر می‌کنم احتمالا چنین اتفاقی برای آن‌ها بیافتد. تصور می‌کنم قیاس تشابهی به چنین شیوه‌ای در مغزمان کار می‌کند. وقتی دو مجموعه از ایده‌ها یا تعدادی بیشتر الگوی یکسانی از اتصالات منطقی دارند، مغز به احتمال زیاد بخش‌ها را با استفاده از برخی از آن‌ها برای دو مرتبه صرفه‌جویی می‌کند، تا اتصالات منطقی را هم در یک مورد و هم در مورد دیگر به یاد بیاورد. باید در نظر گرفت که بخشی از مغز من به این شیوه

دوبار استفاده شده بود، یک بار برای ایده‌ای از نفی مضاعف و یک بار برای عبور از جاده، رفت و برگشت. واقعا بناست درباره‌ی هر دو این‌ها بدانم اما مادامی که دارد درباره‌ی همه‌ی آن نه‌ها و نه‌نیست‌ها صحبت می‌کند نمی‌توانم بفهمم انسان دارد به چه سمتی می‌رود. تا اندازه‌ای از بخش درست مغز رد نمی‌شود. اما همین‌که قطعه‌اش درباره‌ی عبور از جاده را می‌گوید از بخش درست رد می‌شود، اما از راه مسیری متفاوت. اگر چنین توضیح کاملا مکانیکی از نحوه‌ی کارکرد این استدلال در مغز وجود داشته باشد، می‌توان کامپیوتر دیجیتالی ساخت که همین کار را انجام بدهد.

جفرسون: خب، توضیحی مکانیکی بر حسب سلول‌ها و فیبرهای متصل به هم در مغز وجود ندارد.

برایتویت: اما آیا ماشین واقعا می‌تواند این کار را انجام دهد؟ چطور انجامش خواهد داد؟

تورینگ: قطعا بخش زیادی از ماجرا را به تخیل سپرده‌ام. اگر توضیح مبسوط‌تری ارائه داده بودم احتمالا کاری می‌کردم که قطعی‌تر به نظر برسد که آنچه داشتم توصیفش را می‌کردم شدنی بود، اما احتمالا همه‌تان قدری دراین‌باره حس خوبی پیدا نخواهید کرد، و شاید با بی‌صبری فریاد می‌زدید «خب، بله، می‌فهمم که یک ماشین می‌تواند همه‌ی این‌ها را انجام بدهد، اما این تفکر را نمی‌گویم.» تا زمانی که بتوان علت و معلول را دید که خودشان را در مغز حل یا کشف می‌کنند، این اتفاق نه تفکر بلکه یک‌جور خرحمالی غیرقابل‌تصور در نظر گرفته می‌شود. از این دیدگاه چه‌بسا وسوسه شویم تفکر را عبارت از «فرایندهایی ذهنی» تعریف کنیم که «نمی‌فهمیم». اگر این درست باشد آنگاه ساختن یک ماشین متفکر یعنی ساختن ماشینی که کارهای جالبی انجام می‌دهد بدون اینکه واقعا درست بفهمیم چطور این کارها را انجام داده است.

جفرسون: حرف‌تان کاملا درست است اگر منظورتان این باشد که سیم‌بندی در انسان‌ها را نمی‌دانیم.

تورینگ: نه، به‌هیچ‌وجه منظورم این نبود. ما سیم‌بندی ماشین‌مان را می‌دانیم، اما این سیم‌بندی پیشتر به‌صورتی محدود اتفاق می‌افتد. گاهی یک ماشین محاسبه‌گر کار نسبتا غریبی عملا انجام می‌دهد که انتظارش را نداشتیم. در اصل می‌شد پیشبینی‌اش کرد، اما در عمل معمولا دردسر بسیار زیادی دارد. آشکارا اگر قرار بود هر چیزی را که کامپیوتر بناست انجام دهد پیشبینی کنیم آن‌وقت احتمالا بدون کامپیوتر نیز همین کار را به‌خوبی انجام می‌دادیم.

نیومن: کاملا صحت دارد که مردم مأیوس می‌شوند وقتی درمی‌یابند که ماشین‌های محاسبه‌گر بزرگ چه کاری عملا انجام می‌دهند، که فقط جمع‌زدن و ضرب‌کردن است، و استفاده از نتایج برای تصمیم‌گیری درباره‌ی عملکرد جمع‌ها و ضرب‌های آتی. «*این* تفکر نیست» نظری طبیعی است، اما این گفته مسئله را مسلم فرض می‌گیرد. اگر به یکی از کلیساهای کهن در راونا بروید برخی از زیباترین تصاویر را روی دیوارهایش می‌بینید، اما اگر با دوربین دوچشمی با دقت به آن‌ها نگاه کنید چه‌بسا بگویید «عجب، آن‌ها واقعا ذره‌ای تصویر نیستند، بلکه صرفا انبوهی سنگ رنگی کوچک‌اند با سیمانی بینابین‌شان.» فرایندهای ماشین معرق‌هایی با اجزای استاندارد بسیار ساده‌اند، اما طراحی‌ها می‌توانند پیچیدگی زیادی داشته باشند، و بدیهی نیست که حد الگوهای فکری که می‌توانند تقلید کنند تا کجاست.

برایتوایت: اما چه تعداد سنگ در معرق‌تان وجود دارد؟ جفرسون، آیا کثرتی کافی از سلول‌ها در مغز برای‌شان وجود دارد تا مثل ماشینی محاسبه‌گر عمل کنند؟

جفرسون: بله، هزاران، ده‌ها هزار سلول در مغز وجود دارند بیشتر از آنچه در ماشینی محاسبه‌گر موجود است، چراکه ماشین فعلی شامل... ــ گفتید چقدر؟

تورینگ: نیم میلیون رقم. فکر می‌کنم می‌توان فرض گرفت که هم‌ارز نیم میلیون سلول عصبی است.

برایتوایت: اگر مغز مثل یک ماشین محاسبه‌گر عمل می‌کند، آن‌وقت ماشین محاسبه‌گر کنونی نمی‌تواند تمام کارهایی را که مغز انجام می‌دهد انجام دهد. در مورد این موضوع به توافق رسیدیم؛ اما اگر ماشینی محاسبه‌گر ساخته می‌شد که می‌توانست تمام کارهایی را که مغز انجام می‌دهد انجام دهد، آن‌وقت به ارقامی بیشتر از فضای موجود در مغز نیاز نخواهد داشت؟

جفرسون: خب، نمی‌دانم. فرض کنیم که معادل‌گذاری ارقام در ماشین با سلول‌های عصبی در مغز درست باشد. تخمین‌های متعددی وجود دارد، گمان می‌رود چیزی بین ده هزار میلیون و پانزده هزار میلیون سلول در آنجا وجود داشته باشد. در جریان‌اید که اطلاع قطعی در کار نیست. این عددی فوق‌العاده بزرگ است. اما قطعا مسئله فقط بر سر اندازه نیست. منطق بسیار زیادی در ماشین بزرگ‌تان وجود خواهد داشت. واقعا مثل یک خروجی انسانی فکر نخواهد بود. برای آنکه هرچه بیشتر شبیه آن شود، اجزای ماشینی زیادی می‌بایست به‌نحوی کاملا متفاوت طراحی شوند تا انعطاف‌پذیری بالاتر و امکان‌های متنوع‌تر استفاده را به آن بدهند. حقیقتا کار بسیار شاقی است.

تورینگ: در این مورد اندازه واقعا مسئله‌ی مهمی است. اندازه همان میزان اطلاعاتی است که می‌تواند ذخیره شود. اگر می‌خواهید یکی از این ماشین‌ها کار بسیار پیچیده‌ای را برای‌تان انجام دهد، چه‌بسا متوجه شوید که آن ماشین به‌خصوص نتواند از پس انجامش برآید، اما اگر هر ماشینی اساسا بتواند از پس انجامش برآید، آن‌وقت این کار می‌تواند به‌وسیله‌ی اولین کامپیوتر شما انجام شود که به‌سادگی ظرفیت ذخیره‌سازی‌اش افزایش یافته است.

جفرسون: اگر واقعا قرار است به هرآن‌چیزی برسیم که حقیقا می‌تواند «تفکر» خوانده شود، آن‌گاه اثرات محرک‌های بیرونی نمی‌توانند از قلم بیافتند؛ مداخله‌ی انواع‌واقسام عوامل بیرونی مثل نگرانی‌هایی درباب سرپانگه‌داشتن زندگی‌مان، یا پرداخت مالیات‌ها، یا خرید غذایی که دوست داریم. این عوامل به‌هیچ‌وجه کوچک نیستند، در حقیقت بسیار مهم‌اند، و نگرانی‌های مرتبط با آن‌ها عمیقا می‌تواند مانع تفکر درست و خصوصا تفکر خلاق شود. می‌بینید که ماشین هیچ محیطی ندارد، و انسان در نسبتی ثابت با محیطش است که همچون همیشه او را پانچ می‌کند درحالی‌که او نیز محیط را متعاقبا پانچ می‌کند. پس‌زمینه‌ی حافظه‌ها در مغز انسان به‌قدری وسیع است که هر ایده یا تجربه‌ی تازه می‌بایست با آن جفت‌وجور شود. مایل‌ام اگر می‌شود به من بگویید که ماشین محاسبه‌گر چقدر می‌تواند این موقعیت را برآورده کند. اغلب مردم موافق‌اند که اولین واکنش انسان به ایده‌ای جدید (مثل ایده‌ای که امروز داریم درباره‌اش بحث می‌کنیم) انکار است، اغلب هم رد فوری و ترس‌خورده‌ی آن ایده. نمی‌فهمم چطور یک ماشین می‌تواند چنان‌که صحبتش بود بگوید «حالا پروفسور نیومن یا

آقای تورینگ، من اصلا این برنامه را که در من کار گذاشته‌اید دوست ندارم، راستش نمی‌خواهیم هیچ سروکاری با آن داشته باشم.»

نیومن: یک مشکل برای پاسخ‌دادن به مسئله‌ی مزبور همان است که تورینگ قبلتر اشاره کرد. اگر گفته شود «آیا ماشین می‌تواند این کار را انجام دهد، فرضا بتواند بگوید: برنامه‌ای را که در من کار گذاشته‌اید دوست ندارم»، و برنامه‌ی انجام آن کار چنان‌که بایدوشاید ایجاد شده باشد، گرایش دارد که فضایی مصنوعی و ازپیش‌مقرر داشته باشد، و بیشتر یک حقه به نظر می‌رسد تا پاسخی جدی به پرسش. مثل آن صفحاتی از انجیل است که وقتی پسربچه‌ای کوچک بودم مایه‌ی نگرانی‌ام می‌شد، همان‌ها که می‌گوید فلان و بهمان انجام گرفت «که رسالت نبوت بتواند محقق شود» چنین و چنان. همواره به نظرم رسید که این نادرست‌ترین یا ناحق‌ترین شیوه‌ی اطمینان‌دادن از واقعیت‌یافتن رسالت نبوت است. جفرسون، اگر با ایجاد روتینی به سئوالت پاسخ بدهم که به‌سادگی سبب شد تا ماشین این کلمات را بگوید که «نیومن و تورینگ، برنامه‌تان را دوست ندارم»، قطعا حس خواهید کرد که این حقه‌ای بچگانه است و نه پاسخی به آنچه واقعا می‌خواستید بدانید. اما بااین‌همه همچنان سخت است که بتوان خواسته‌تان را دقیقا مشخص کرد.

جفرسون: دوست دارم ماشین این مسئله را رد کند چون به‌طریقی اذیتش می‌کند. این نکته من را به سمت کندوکاو در ماهیت مؤلفه‌های ایده‌هایی که رد می‌کنیم سوق می‌دهد چون ما به‌طور غریزی اهمیتی برای‌شان قائل نیستیم. نمی‌دانم چرا برخی تصاویر و یک موسیقی خاص را دوست دارم و حوصله‌ام از انواع دیگر تصاویر و موسیقی سر می‌رود. اما نمی‌خواهیم آن سویه را پیش ببرم چون همه‌مان متفاوت‌ایم، عدم‌پسندهایمان مبتنی بر سرگذشت‌های شخصی‌مان است و همچنین احتمالا بر مبنای تفاوت‌های کوچک در سرشت تک‌تک ما که یعنی همان توارث. ماشین‌های شما هیچ ژن و هیچ شجره‌ای ندارند. توارث مندلی هیچ معنایی برای لامپ‌های خلا بی‌سیم ندارد. اما نمی‌خواهم امتیازات بحث را جمع کنم! می‌بایست روشن کنیم که حتی خود تورینگ هم فکر نمی‌کند که همه‌ی آنچه باید انجام بدهد کشیدن پوستی روی ماشین است تا بگوید که ماشین زنده است! ما برای این هدف محدودشده‌تر کوشیده‌ایم که آیا آن سنخ کاری که ماشین‌ها انجام می‌دهند می‌تواند به‌عنوان تفکر در نظر گرفته شود یا نه. اما آیا ماشین شما نسبت به هر موجود انسانی در رسیدن به مسئله‌اش به‌نحوی اطمینانی یا بدون خطا، قطعیت بیشتری ندارد؟

نیومن: اوه!

تورینگ: ماشین‌های محاسبه‌گر واقعا به‌طور کامل خطاناپذیر نیستند. چک‌کردن دقت‌شان بخش بسیار مهمی از هنر استفاده از آن‌هاست. آن‌ها علاوه بر اشتباهات گاهی واقعا محاسبه‌ای را که انتظارش را داشتیم انجام نداده‌اند و به چیزی می‌رسیم که می‌تواند «سوفهم» خوانده شود.

جفرسون: به هر صورت، آن‌ها تحت تأثیر هیجانات نیستند. کافی است حال یک شخص را خراب کنی و او هم در نتیجه‌اش آشفته می‌شود، او نمی‌تواند به پاسخ‌ها فکر کند و ممکن است خودش به حماقت درافتد. این محتوای هیجانی بالای فرایندهای ذهنی در موجود انسانی است که او را از ماشین کاملا متفاوت می‌کند. به نظرم این محتوا از پیچیدگی بزرگ سیستم

عصبی‌اش با 10^{10} سلول و نیز از سیستم غده‌ی درون‌ریزش بیاید که بر همه‌ی انواع هیجان‌ها و غریزه‌ها از جمله موارد مرتبط با جنس اشاره دارد. انسان ذاتا یک ماشین شیمیایی است، او با گرسنگی و خستگی، با به‌اصطلاح «ناخوش‌احوالی»، همچنین با داوری‌های فطری، و با کشش‌های جنسی، بی‌اندازه تحت تأثیر قرار می‌گیرد. این سویه‌ی شیمیایی بسیار اهمیت دارد نه صرفا به این خاطر که مغز یک کنترل ازراه‌دور را بر مهم‌ترین فرایندهای شیمیایی جاری در بدن‌هایمان عملی می‌سازد. ماشین‌های شما نباید از این بابت، یعنی از خسته یا سرد یا شاد یا راضی بودن، اذیت شوند. آن‌ها هیچ شوقی در انجام آنچه قبلتر هرگز انجام نشده نشان نمی‌دهند. نه، آن‌ها «از حیث ذهنی» ساده‌اند. منظورم این است که هر قدر هم ساختارشان پیچیده باشد (و می‌دانم که بسیار پیچیده *است*)، آن‌ها در قیاس با آن بسیار ساده‌اند و وظایف‌شان را با غیاب افکار حواس‌پرت‌کننده اجرا می‌کنند، که این هم خودش کاملا *غیرانسانی* است.

برایتویت: نمی‌توانم با شما موافق باشم. به باورم ضرورت خواهد داشت تا چیزی متناظر با اشتیاق‌ها یا دیگر «جنب‌وجوش‌ها» را برای ماشین تدارک ببینیم، تا در نتیجه‌اش توجه کافی را به خصایص مرتبط با آن‌ها در محیطش نشان بدهد تا بتواند از تجربه بیاموزد. بسیاری از روان‌شناس‌ها بر این باورند که هیجان‌ها در انسان‌ها محصول جانبی اشتیاق‌هایشان هستند و اینکه در خدمت کارکردی زیست‌شناختی‌اند تا سطوح بالاتر فعالیت ذهنی را به راه بیاندازند وقتی سطوح پایین‌تر ناتوان از تطبیق‌یافتن با موقعیت بیرونی‌اند. برای مثال، حس ترس نداریم وقتی خطری در کار نباشد، یا خطری که بتوان کمابیش به‌طور خودکار از آن اجتناب کرد: ترس سمپتومی است که نشان می‌دهد خطر باید با فکری آگاه مواجه شده باشد. احتمالا محال خواهد بود که ماشینی ساخته شود که بدون اینکه دستگاهی هیجانی در آن کار گذاشته شود بتواند به‌طور عام از تجربه بیاموزد، دستگاهی که کارکردش سوئیچ‌کردن به جز متفاوتی از ماشین خواهد بود وقتی محیط بیرونی تفاوت بسیار زیادی دارد با آنچه اشتیاق‌های ماشین را با بیش از مقداری مشخص برآورده می‌کند. نمی‌خواهم پیشنهاد بدهم که برای ماشین ضرورت دارد که بتواند قدری اوقات‌تلخی کند. اما در انسان‌ها اوقات‌تلخی‌ها متناوبا کارکردی معین را برآورده می‌کنند ــ کارکرد گریز از مسئولیت‌پذیری؛ و برای محافظت از ماشین در برابر محیطی بیش‌ازحد متخاصم الزامی است که به آن اجازه دهیم با روان‌نژندی یا مرضی روان‌زایانه خاموش شود ــ درست همان‌طور که به شیوه‌ای ساده‌تر فیوزی برای ترکیدن برایش مهیا می‌شود اگرکه توان برقی که با آن کار می‌کند وجود دوام‌یافته‌اش را تهدید کند.

تورینگ: خب، در تصورم نیست که ماشین را برای برون‌ریزی صحنه‌های فطری یا کج‌خلقی آموزش بدهیم. فکر می‌کنم چنان اثراتی احتمالا درمقام یک‌جور محصول جانبی آموزش راستین رخ بدهند، و بیشتر مایل‌ایم که مانع چنین جلوه‌هایی شویم تا اینکه ترغیب‌شان کنیم. این اثرات احتمالا به‌طور مشخص با نوع متناظر انسانی‌شان فرق دارند، اما می‌توانند به‌صورت تغیراتی از آن‌ها شناسایی شوند. این یعنی اگر ماشین یکی از تست‌های تقلید من را به انجام رسانده باشد، می‌بایست قدری دست به عمل بزند، اما اگر این را به شیوه‌ای نه‌چندان سرراست با یک انسان قیاس کنیم شباهت‌شان احتمالا بسیار تأثیرگذار و حیرت‌آور است.

نیومن: همچنان حس می‌کنم که بخش زیادی از بحث‌مان درباره‌ی این است که ماشین‌های فرضی آینده چه خواهند کرد. بسیار خوب بتوان گفت که یک ماشین به‌آسانی می‌تواند برای انجام فلان یا بهمان کار ساخته شود، اما، با ملاحظه‌ی تنها یک نکته‌ی عملی، در مورد زمانی که خواهد برد تا این کار را انجام بدهد چه؟ تنها یک یا دو ساعت طول خواهد کشید که روتینی آماده شود که سبب شود ماشین منچستر ما همه‌ی تغیرهای ممکن بازی شطرنج را درجا تحلیل کند، و بدین‌طریق بهترین حرکت را بیابد ـ البته *اگر* اهمیتی نمی‌دهید که هزاران میلیون سال برای پردازش روتین زمان ببرد. حل یک مسئله روی ماشین به معنی یافتن راهی برای انجامش در گستره‌ی زمانی معقول است و نه انجامش بین اکنون و ابدیت. صرف جزئیات فنی نیست که از پیشرفت‌های آتی حفاظت خواهد کرد. بیش از همه نامحتمل است که مهندسان هرگز بتوانند عاملی در اختیارمان بگذارند که یک یا دو هزار بار سریع‌تر از سرعت فعلی‌مان باشد. این پنداشت که اجرای روتین مزبور، که در ماشین‌های فعلی‌مان هزاران میلیون سال زمان خواهد برد، به یک آن در ماشین‌های آینده انجام خواهد گرفت به معنی حرکت به سمت قلمروهای علمی‌خیالی است.

تورینگ: از دیدم این عامل زمان مسئله‌ای است که هر دشواری فنی واقعی را شامل خواهد شد. اگر پیشاپیش نمی‌دانستیم که مغزها می‌توانند این کارها را در زمانی معقول انجام دهند، آنگاه استفاده از ماشین برای انجام‌شان مایوس‌کننده به نظرمان می‌رسید. این واقعیت که مغز *می‌تواند* این کار را انجام دهد به نظر پیشنهاد می‌دهد که دشواری‌ها واقعا آنقدرها که اکنون به نظرمان می‌رسد خیلی هم بد نباشند.

برایتویت: موافق‌ام که نمی‌بایست بحث‌مان را به این سمت بسط می‌دادیم که آیا ماشین‌های محاسبه‌گری می‌توانند ساخته شوند که هرآنچه را یک انسان قادر به انجامش باشد انجام بدهند یا نه. مسلما نکته این است که آیا آن‌ها می‌توانند تمام کارهایی را که لایق نام تفکر است انجام دهند یا نه. تجلیل از یک عکس عناصر فکرکردن را شامل می‌شود، اما عناصر احساس‌کردن را هم شامل می‌شود؛ و و دغدغه‌مان نیست که آیا ماشینی می‌تواند ساخته شود که حس کند یا نه. همین نکته را در مورد پرسش‌های اخلاقی داریم: فقط تا آنجا با آن‌ها سروکار داریم که پرسش‌های فکری هم هستند. نمی‌خواهیم حسی از وظیفه یا هر چیزی متناظر با اراده را به ماشین بدهیم: و هنوزکمتر از این نیاز داریم که وسوسه‌هایی را به ماشین بدهیم که در نتیجه‌اش می‌بایست دستگاهی برای مقاومت‌کردن داشته باشد. همه‌ی آنچه این ماشین برای فکرکردن باید انجام دهد این است که قادر به حل همه‌ی مسائل فکری باشد یا بتواند در مسیر حل‌شان تلاش خوبی به خرج دهد، مسائلی که این ماشین ممکن است با آن‌ها در محیطی که خودش را در آن می‌یابد مواجه شود. این محیط هم البته باید شامل پرسش‌های نامناسب و ناجوری شود که تورینگ از آن می‌پرسد تا رویدادهای طبیعی مثل باریدن باران بر سرش، یا لرزانده‌شدن با یک زلزله.

نیومن: اما فکر می‌کنم این شما بودید که گفتید ماشین اگر با مجموعه‌ای از اشتیاق‌ها و همه‌ی الزام‌هایشان مهیا نشده باشد نمی‌تواند یاد بگیرد که با محیطش وفق یابد.

برایتویت: بله، قطعا. اما مسائلی که ماشین دارای اشتیاق طرح می‌کند واقعا مسئله‌ی امروز ما نیستند. چه‌بسا بدون اشتیاق‌های یادشده نتواند از تجربه بیاموزد؛ اما صرفا باید در نظر

بگیریم که آیا اصلا می‌تواند یاد بگیرد یا نه ـ چون موافق‌ام که توانایی یادگیری جزئی الزامی از تفکر است. پس آیا نمی‌بایست به چیزی که بر تفکر متمرکز شده بازگردیم؟ برای مثال آیا ماشین می‌تواند مفاهیم تازه بسازد؟

نیومن: واقعا فقط دو پرسش وجود دارد که می‌توان درباره‌ی ماشین‌ها و تفکر پرسید؛ اول: پیش از آنکه موافق باشیم ماشین‌ها *هرآنچه* را که تفکر می‌خوانیم انجام می‌دهند، به چه نیاز داریم؟ بیشتر وقت‌مان را واقعا دراین‌باره صحبت کردیم؛ اما پرسش جالب و مهم دیگری هم وجود دارد: این قلمروی تردیدآمیز از کجا آغاز می‌شود؟ نزدیک‌ترین چیز به محاسبه‌ی سرراستی که ماشین‌های کنونی احتمالا نمی‌توانند انجام دهند چیست؟

برایتویت: و جواب خودتان چه خواهد بود؟

نیومن: من احتمالا به مسائل ریاضیاتی فکر می‌کنم که هیچ روشی، به شیوه‌ای که انسان‌ها می‌توانند انجامش دهند، برای‌شان شناخته نشده است؛ یعنی یافتن روش‌های تازه. این هدف ابداع مفاهیم ریاضیاتی تازه بسیار فروتنانه‌تر است. وقتی سعی می‌کنید مسئله‌ای جدید را به شیوه‌ای عادی حل کنید این اتفاق می‌افتد که برای ثانیه‌ای، یا برای چند سال، به آن فکر می‌کنید، سعی می‌کنید همه‌ی قیاس‌های تشابهی را که به‌شان فکر می‌کنید با مسائلی که حل شده‌اند محک بزنید، و سپس ایده‌ای دارید. به‌طور مفصل امتحانش می‌کنید. اگر ایده‌ی خوبی نبود، باید منتظر ایده‌ای دیگر باشید. قدری شبیه روتین مسئله‌ی شطرنج است، آنجا که حرکتی پس از حرکت دیگر امتحان می‌شود، اما با یک تفاوت بسیار مهم، اینکه اگر حتی ریاضیدانی واقعا خوب باشم ایده‌هایی که به‌شان می‌رسم ایده‌هایی صرفا تصادفی نیستند بلکه از قبل گزینش شده‌اند طوری که این شانس خوب وجود دارد که پس از چند آزمون‌وخطا یکی از آن‌ها به نتیجه برسد. هنری مور درباره‌ی مطالعاتی که برای مجسمه‌اش انجام می‌دهد می‌گوید: «وقتی کار بیش از یک تمرین است، جهش‌های توضیح‌ناپذیر رخ می‌دهد. اینجاست که پای تخیل به میان می‌آید.» اگر ماشین واقعا بتواند این جست‌زدن ناگهانی به سمت یک ایده را تقلید کند، آن‌گاه به باورم همه موافق خواهند بود که شروع به فکرکردن کرده است، ولو اینکه اشتیاق‌ها، یا نگرانی درباره‌ی مالیات درآمدش را نداشته است. و فرض کنید همچنین به سراغ دانسته‌هایمان از فیزیولوژی تفکر انسانی برویم؛ جفرسون، چقدر این معنا خواهد داشت؟

جفرسون: درباره‌ی محصول نهایی یا خود تفکر چیزهای زیادی می‌دانیم. آیا محتواهای کتابخانه‌ها و موزه‌ها کاملا به‌روز نیست؟ فیزیولوژی تجربی چیزهایی زیادی درباره‌ی شیوه‌ی استفاده‌مان از حافظه و تداعی ایده‌ها به ما آموخته است، اینکه چطور شکاف‌های موجود در شناخت را پر می‌کنیم و از روی واقعیت‌های داده‌شده‌ی معدود بداهه‌پردازی می‌کنیم. اما دقیقا نمی‌دانیم که چطور این کار را بر حسب عمل‌های سلول عصبی انجام می‌دهیم. خصوصا اطلاعی از این نکته‌ی خاص نداریم که نیومن لحظاتی قبل به آن اشاره کرد، یعنی فیزیولوژی موجود جست‌زدن به یک ایده یا الهام ناگهانی. تفکر آشکارا یک‌جور فعالیت حرکتی سلول‌های مغز است، تلقینی که این تجربه‌ی رایج پشتیبان آن است که بسیاری از مردم با قلم در دست‌شان بهتر فکر می‌کنند تا خلافش یا با خواب‌وخیال و تأمل. اما با تحریک الکتریکی مغز انسان در هر نقطه‌ای نمی‌توانید ایده‌هایی در ذهن انسان تولید کنید. واقعا هیجان‌برانگیز است اگر بتوان این کار را انجام داد ـ اگر احتمالا بتوان افکار اصیل تهییج‌شده‌ای را با تحریک

محلی ایجاد کرد. این کار شدنی نیست. التهاب‌نگار الکتریکی نیز انجام فرایند تفکر را نشان‌مان نمی‌دهد. می‌توانیم فرضا جریان صفحه‌ای از یک پرینت یا صفحه‌ای از یک جریان کلمات به سمت مغز را رد بگیریم، اما نهایتا آن را از دست می‌دهیم. اگر می‌توانستیم آن‌ها را تا مکان‌های ذخیره‌شان پی بگیریم هنوز نمی‌توانستیم ببینیم چطور بعدتر به‌صورت ایده‌ها از نو سر هم می‌شوند. این مزیت بزرگ را دارید که می‌دانید چطور ماشین‌تان ساخته شده بود. ما فقط می‌دانیم که در سیستم عصبی انسان سهمی داریم با اندازه‌ای فشرده و به شیوه‌ی خودش کامل برای انجام کارهایش. چیزهای زیادی درباره‌ی ساختار میکروسکوپی و اتصالات‌شان می‌دانیم. در واقع همه چیز را می‌دانیم به جز اینکه چطور این یک عالم سلول به ما مجال فکرکردن می‌دهند. اما، نیومن، قبلتر گفتیم «نه فقط این ماشین فکر می‌کند بلکه همچنین اینجا در این ماشین همتای دقیق سیم‌بندی و مدارهای سیستم‌های عصبی انسان را داریم»؛ لازم است بپرسم که آیا ماشین‌هایی ساخته شده‌اند یا می‌توانند ساخته شوند که از حیث کالبدشناختی با هم متفاوت باشند و بااین‌حال کار یکسانی را انجام دهند یا اثر مشابه‌ای را تولید کنند.

نیومن: نقشه‌ی منطقی تمام‌شان نسبتا مشابه است، اما قطعا کالبدشناسی‌شان و به گمانم بتوانید بگویید فیزیولوژی‌شان بسیار با هم فرق دارد.

جفرسون: بله، همین تصور را داشتم ـ پس نمی‌توانیم فرض بگیریم که هیچ کدام از این ماشین‌های الکترونیک نسخه‌ی عین یا المثنای بخشی از مغز انسان باشد حتی اگر نتیجه‌ی عمل‌هایش باید به‌عنوان فکر پذیرفته شود. ارزش واقعی ماشین برای شما نتایج نهایی‌اش، کارایی یا اجرایش، است و نه اینکه نقشه‌اش مدلی از مغزها و عصب‌هایمان را برای‌مان آشکار کند. سودمندی‌اش در این واقعیت است که حرکت جریان برق در طول سیم‌ها ۲ یا ۳ میلیون بار سریع‌تر از حرکت تکانه‌ها در اعصاب است. می‌توانید این ماشین را برای انجام چیزهایی تنظیم کنید که انسان برای تکمیل‌شان به هزاران زندگی نیاز دارد. اما آن مربی کند قدیمی، یعنی انسان، دارای ایده‌هاست ـ یا این‌طور فکر می‌کنم. تورینگ، این مایه‌ی خوشی خواهد بود که روزی به بحثی فرضا درباره‌ی **برنامه‌ی چهارم** گوش بسپاریم بین دو ماشین درباره‌ی اینکه چرا موجودات انسانی فکر می‌کنند که فکر می‌کنند!

منبع

Alan Turin, *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life: Plus The Secrets of Enigma*, editd by B. Jack Copeland (OxfordL Clarendon, 2004)