

در دفاع از آزمون تورینگ^۱

محمد رضا واعظ شهرستانی

۱. کارشناسی ارشد فلسفه علم، دانشگاه صنعتی امیرکبیر، تهران، ایران

مکاتبات: vaez.dawn1985@gmail.com

چکیده:

تورینگ در مقاله سال ۱۹۵۰ خود، توجیهی برای جایگزینی سؤال «آیا یک ماشین می‌تواند در بازی تقلید^۲ موفق شود؟» به جای سؤال اصلی مقاله، یعنی «آیا ماشین‌ها می‌توانند فکر کنند؟» ارائه نمی‌کند و این موضوع به‌عنوان شکافی تبیینی در آزمون تورینگ مطرح می‌شود. در این مقاله استدلال خواهیم کرد که پاسخ مثبت به سؤال اصلی و اولیه آزمون، به معنای نسبت‌دادن قابلیت تفکر به ماشین‌هاست. این در حالی است که پاسخ مثبت به سؤال جایگزین شده، به معنای نسبت‌دادن قابلیت تعامل زبانی به ماشین‌هاست. در چنین وضعی، دفاع از آزمون تورینگ، نیاز به برقراری ارتباطی میان فکر و زبان دارد؛ رابطه‌ای که بتواند دربرگیرنده توجیهی مناسب برای جایگزینی قابلیت تعامل زبانی در سؤال دوم به جای قابلیت تفکر در سؤال اول باشد. در این زمینه، نظریه عدم اولویت زبان و فکر دیویدسون به‌عنوان رویکردی در جهت دفاع از این آزمون ارائه می‌شود.

کلمات کلیدی: آزمون تورینگ، قابلیت تعامل زبانی، قابلیت تفسیر، زبان و فکر، نظریه عدم اولویت دیویدسون.

۱. مقدمه

آیا یک کامپیوتر با برنامه مناسب می‌تواند فکر کند؟ هوش مصنوعی به دنبال یافتن پاسخی برای این پرسش است. در دهه‌های اخیر، امکان هوشمندی کامپیوترها توجه فیلسوفان ذهن و متخصصان هوش مصنوعی را برانگیخته است. یکی از نظریات معروف در این زمینه، آزمون تورینگ است. آلن تورینگ^۳ فیلسوف انگلیسی که مبدع نظریه ماشین‌های مجرد است، در

^۱ این مقاله در ماهنامه علمی-تخصصی اطلاعات حکمت و معرفت، بهمن ۱۳۹۳، شماره ۱۰۶ منتشر شده است. (مراجعه کنید به <http://www.ettelaathekmatvamarefat.com>)

^۲ Imitation game

^۳ Alan Turing

سال ۱۹۵۰ به طراحی آزمونی برای سنجش هوشمندی کامپیوترها پرداخت. او مقاله خود را با طرح این سؤال آغاز می‌کند: «آیا ماشین‌ها می‌توانند فکر کنند؟» تورینگ لازمه درک این سؤال را ارائه تعریفی روشن از کلمات «ماشین» و «فکر» می‌داند. تورینگ با توجه به اینکه تعریف این کلمات از طریق کاربرد عادی آن‌ها امکان‌پذیر است، این روش را برای پاسخ‌گویی به این سؤال، خطرناک و نادرست می‌انگارد و چنین سؤالی را بی‌معنی ارزیابی می‌کند و به‌منظور روشن‌تر کردن سؤال و استفاده از کلماتی بدون ابهام، بلافاصله سؤال دیگری را جایگزین این سؤال می‌کند. سؤال دوم چنین است: «آیا یک ماشین می‌تواند در بازی تقلید موفق شود؟»

در مقاله تورینگ، بازی تقلید را سه نفر انجام می‌دهند: یک مرد، یک زن و یک بازیگر؛ با این توضیح که جنسیت بازیگر اهمیتی ندارد. در این بازی، بازیگر در اتاقی جداگانه و به‌صورت پنهان از دو نفر دیگر قرار می‌گیرد. وظیفه بازیگر در این بازی، شناسایی جنسیت هر فرد است؛ به این صورت که او باید تشخیص دهد کدام فرد مرد و کدام فرد زن است و اگر این کار را به‌درستی انجام دهد، در بازی موفق شده است. بازیگر برای شناسایی می‌تواند سؤال‌هایی را از هریک از آن دو نفر بپرسد؛ البته برای اینکه صدای افراد در هنگام پاسخ‌گویی در شناسایی افراد به بازیگر کمک نکند، سؤالات به‌صورت کتبی پرسیده می‌شوند. روند بازی به این شکل است که در هنگام پاسخ‌گویی به سؤالات بازیگر، زن درجهت کمک به بازیگر برای موفقیت در بازی و مرد درجهت فریب‌دادن بازیگر، تلاش می‌کند. در ادامه مقاله، تورینگ جای مرد را در این بازی با یک ماشین عوض می‌کند و سؤال دوم یعنی سؤال «آیا یک ماشین می‌تواند در بازی تقلید موفق شود؟» را به جای سؤال اول مطرح می‌کند. تورینگ در مقاله خود در مورد موجه‌بودن این جایگزینی هیچ توضیحی نمی‌دهد و تنها به مبهم‌بودن واژه‌های «تفکر» و «ماشین» در سؤال اول اشاره می‌کند و سؤال دوم را فارغ از ابهام دانسته و به همین توضیح برای جایگزینی بسنده می‌کند. در نقد و بررسی این آزمون، ارائه‌نکردن توجیهی مناسب در جایگزینی سؤال دوم به جای سؤال اول، به‌عنوان شکافی تبیینی در این آزمون مطرح می‌شود. در واقع در این مقاله، تورینگ دلیل موجهی را ارائه نمی‌کند که چرا سؤال دوم می‌تواند جایگزین سؤال اول شود (تورینگ، ۱۹۵۰: ۴۳۳ تا ۴۳۴). هدف ما در این مقاله، یافتن توجیهی مناسب برای مجازشمردن این جایگزینی است تا از این طریق بتوانیم همدلانه به دفاع از این آزمون بپردازیم. در این بررسی از نظریه معروف فیلسوف معاصر آمریکایی، دونالد دیویدسون (۱۹۱۷ تا ۲۰۰۵)^۴ در ارتباط با رابطه فکر و زبان بهره خواهیم برد که در ادامه شرحی اجمالی درباره آن خواهیم داد. در این مقاله، ما این نظریه را بدون اثبات و به‌عنوان یک پیش‌فرض در حکم رویکردی در دفاع از آزمون تورینگ به کار خواهیم برد. توجیه این نظریه فلسفی را باید در نوشته‌های دیویدسون جستجو کرد.

۲. آزمون تورینگ

در اکتبر ۱۹۵۰ آلن تورینگ پیش‌بینی کرد که تا سال ۲۰۰۰ ساخت کامپیوتری که حداقل سی درصد شانس برای فریب بازیگر دارد، امکان‌پذیر خواهد بود؛ به‌طوری‌که بازیگر در تشخیص انسان از کامپیوتر دچار اشتباه شود (تورینگ، ۱۹۵۰: ۴۳۳ تا ۴۳۴).

^۴ Donald Davidson

۴۴۲). در این آزمون یک بازپرس پشت دو چاپگر قرار می‌گیرد؛ به طوری که یک چاپگر به یک پایانه به کار گرفته شده توسط یک زن و چاپگر دیگر به یک کامپیوتر دیجیتالی متصل شده است. زن و کامپیوتر از شخص بازپرس پنهان هستند. تورینگ به بازپرس تنها پنج دقیقه برای گفت‌وگو وقت می‌دهد. همچنین بازپرس برای اینکه بتواند تعیین کند کدام پایانه به کامپیوتر متصل است و کدام پایانه توسط یک انسان مدیریت می‌شود، سؤال‌هایی را خطاب به هر کدام حروف‌نگاری (تایپ) می‌کند. روند آزمون به این شکل است که زن در جهت کمک به بازپرس برای موفقیت در آزمون تلاش می‌کند؛ در حالی که کامپیوتر در جهت فریب‌دادن بازپرس برنامه‌ریزی شده است. این آزمون زمانی نتیجه خواهد داد که فرد بازپرس تعیین کند که کدام پایانه متصل به کامپیوتر است و کدام پایانه متصل به زن (دیویدسون، ۱۹۹۰: ۷۸). در صورتی که کامپیوتر موفق به فریب‌دادن بازپرس شود و آزمون را با موفقیت پشت سر بگذارد، تورینگ ادعا می‌کند که کامپیوتر می‌تواند فکر کند. البته این یک آزمون یک‌طرفه است؛ به این معنی که اگر کامپیوتر در فریب‌دادن بازپرس موفق نشود، ادعایی نمی‌توان کرد. در ادامه همان طور که در مقدمه گفته شد، به منظور استفاده از نظریه عدم اولویت زبان و فکر دیویدسون در حکم رویکردی در دفاع از آزمون تورینگ، به شرح این نظریه خواهیم پرداخت.

۳. نظریه عدم اولویت زبان و فکر دیویدسون

ارتباط میان فکر و زبان چیست؟ دیویدسون در مقاله معروف «فکر و زبان» (دیویدسون: ۱۹۷۵) با طرح این سؤال به تحلیل و بررسی نسبت میان فکر و زبان می‌پردازد. از نظر او وابستگی زبان به فکر، به روشنی مشهود است؛ زیرا زبان برای بیان افکار استفاده می‌شود؛ برای مثال هنگامی که شخصی جمله «شمع خاموش است» را بر زبان می‌آورد، این گزاره به شمعی در جهان خارج اشاره دارد که در زمان بیان این گزاره، خاموش است و شخص به این گزاره باور دارد. در واقع او با تولید چنین اصواتی در حال بیان کردن گزاره‌ای است که تنها تحت چنین وضعیتی صحیح است. به عبارت دیگر در صورت فقدان چنین باوری، یعنی باور به وجود یک شمع خاموش در جهان خارج، چنین اصواتی نیز در زبان تولید نخواهد شد و این بیانگر ارتباط باورها و زبان است. به طور کلی دیویدسون معتقدست که نه زبان و نه فکر هیچ‌یک نمی‌توانند به طور کامل بر حسب دیگری توضیح داده بشوند و هیچ‌یک بر دیگری اولویت مفهومی و معرفتی ندارند. زبان و فکر هر دو به یکدیگر متصل‌اند؛ یعنی هریک برای آنکه فهمیده بشود، به دیگری نیاز دارد. به عبارت دیگر از این نظریه این دو ادعا نتیجه می‌شود:

۱. ما برای اینکه بتوانیم فکر کنیم، نیاز به تعامل زبانی با دیگران داریم؛ یعنی باید توانایی تفسیر سخنان دیگران را داشته باشیم. این اصل را همچنین می‌توان به این صورت نوشت: «اگر موجودی بتواند فکر کند، آنگاه باید دارای قابلیت تعامل زبانی باشد.»

از نظر دیویدسون یک موجود نمی‌تواند افکاری داشته باشد؛ مگر آنکه او دست‌کم به صورت بالقوه، مترجم و مفسری از گفتار و سخنان دیگران باشد. از نظر او نتیجه فوق، مستلزم تحویل فکر به زبان (مانند رفتارگرایی) نیست و او قائل به هیچ‌گونه اولویت مفهومی یا معرفتی زبان بر فکر نیست. همچنین ما برای فکر کردن نیاز به مفاهیم داریم؛ چون تفکر از طریق مفاهیم و تصورات

انجام می‌گیرد. دیویدسون استدلال می‌آورد که برای داشتن شبکه‌ای از مفاهیم، نیاز به مفهوم خطا داریم و برای داشتن مفهوم خطا، باید در تعامل زبانی با یک موجود دیگر قرار بگیریم تا از طریق او بتوانیم به درستی و نادرستی افکار خویش پی ببریم. بنابراین برای اینکه بتوانیم فکر کنیم، باید در تعامل زبانی با موجودات دیگر باشیم؛ یعنی باید بتوانیم به تفسیر و ترجمه گفتار و اظهارات دیگران پردازیم.

۲. ما برای تعامل زبانی با دیگران نیاز به فکرکردن داریم؛ یعنی باید مجموعه‌ای از باورها داشته باشیم. این اصل را می‌توان به این صورت نوشت: «اگر موجودی دارای قابلیت تعامل زبانی باشد، آنگاه باید بتواند فکر کند.»

دیویدسون معتقدست که یک فکر به وسیله نظامی از باورها معین می‌شود. به عبارت دیگر داشتن یک فکر به داشتن پیش‌زمینه‌ای از باورها نیاز دارد؛ برای مثال، او می‌گوید: «فکر رفتن به یک کنسرت خاص را در نظر بگیرید. ما برای داشتن این فکر مجموعه‌ای از باورها را در ذهن خود داریم؛ باورهایی چون "برای رفتن به این کنسرت نیاز به صرف هزینه و زمان دارم"، "می‌توانم با حضور در این کنسرت از موسیقی مورد علاقه خود لذت ببرم" و چندین و چند باور دیگر. بنابراین من برای داشتن چنین فکری پیش‌زمینه‌ای از باورها دارم که چنین فکری را برای من معین می‌کنند؛ اما، داشتن یک فکر خاص به داشتن باوری خاص وابسته نیست؛ یعنی من در زمینه رفتن به این کنسرت فکری دارم؛ اما تا زمانی که تصمیم بگیرم آیا به کنسرت خواهم رفت یا نه، باور خاص و ثابتی با این محتوا که "من به کنسرت خواهم رفت" ندارم و تا آن زمان من فقط پذیرای آن فکر خواهم بود.» بعد از مطرح کردن محوریت باورها برای داشتن افکار، دیویدسون به لزوم داشتن پیش‌زمینه‌ای از باورها برای فهم و تفسیر گفتار و اظهارات دیگران در تعاملات زبانی اشاره می‌کند. مطابق نظریه، کسی که می‌تواند یک جمله در یک زبان طبیعی، مثل زبان فارسی را تفسیر کند و بفهمد، باید باورهای زیادی داشته باشد و این باورها باید به باورهای کسی که پذیرای این فکر است، خیلی شبیه باشند؛ برای مثال اگر کسی بشنود که «این تفنگ پُر است» و در صدد تفسیر این جمله برآید، باید باور داشته باشد که یک تفنگ، یک جنگ‌افزار است و باور داشته باشد که آن، یک جسم کم‌وییش پایدار فیزیکی است و البته چندین و چند باور دیگر. البته برای اینکه او این جمله را بفهمد، شاید فهرست تعریف شده و ثابتی از باورهایی که او باید داشته باشد، در کار نباشد؛ اما وجود پیش‌زمینه‌ای از تعداد بی‌شماری باور برای فهم این جمله ضروری است. به طور خلاصه می‌توان گفت که فهمیدن گفتاری در زبان توسط یک موجود، مستلزم داشتن فکری درباره آن جمله است و داشتن چنین فکری نیز مستلزم داشتن مجموعه‌ای از باورها در ارتباط با آن است.

به‌طورکلی مطابق با نظر دیویدسون می‌توان گفت در نظر گرفتن هر یک از عناصر فکر و زبان به‌عنوان مبنایی برای دیگری، امری بیهوده و اشتباه است. اگر تصور کنیم فکر بر زبان تقدم دارد، به نظر می‌رسد با چنین دیدگاهی، زبان هدفی جز بیان و حمل افکار نداشته باشد و این بدین معنی است که زبان تنها وسیله‌ای برای انتقال افکار است و اگر تصور کنیم که زبان بر فکر تقدم دارد، این تلاشی است در جهت تحلیل افکار به‌عنوان گرایش‌های گفتاری و بدان معناست که تفکر در سطح بشری، یک فعالیت صرفاً زبانی است؛ بنابراین توازن بین ساختار افکار و ساختار جملات زبان، هیچ دلیل و برهانی را برای تقدم یکی بر دیگری تأمین نمی‌کند و تنها فرض و احتمالی است به نفع وابستگی متقابل زبان و فکر.

۴. رویکرد عدم اولویت زبان و فکر در دفاع از آزمون تورینگ

در این بخش باتوجه به دو نتیجه شرح داده شده در قسمت قبل که نتایج نظریه عدم اولویت زبان و فکر دیویدسون است، به دفاع از آزمون تورینگ با استفاده از چنین رویکردی می پردازم.

در آزمون تورینگ با جایگزینی سؤال «آیا یک ماشین می تواند در بازی تقلید موفق شود؟» به جای سؤال «آیا ماشین ها می توانند فکر کنند؟» مواجه هستیم. حال اگر بخواهیم با تورینگ همدل باشیم و به دفاع از آزمون تورینگ بپردازیم، باید بتوانیم استدلال موجهی برای جایگزینی سؤال جدید به جای سؤال اولیه داشته باشیم؛ زیرا همان طور که پیش تر اشاره شد، تورینگ درباره موجه بودن این جایگزینی هیچ توضیحی نمی دهد و تنها به مبهم بودن واژه های «تفکر» و «ماشین» در سؤال اول اشاره می کند و سؤال دوم را مبهم دانسته و به همین توضیح برای جایگزینی بسنده می کند.

دیویدسون در مقاله ای که درباره آزمون تورینگ نوشته است، می گوید: «با جایگزینی سؤال، روشن است که تورینگ به اهمیت فلسفی از دست رفته معتقد نیست» (دیویدسون، ۱۹۹۰: ۷۷). بنابراین در مقاله تورینگ، عدم توجه این جایگزینی همچون شکافی دیده می شود و ما برای دفاع از آزمون تورینگ نیاز به پرکردن این شکاف داریم. همچنین می دانیم ماشین تورینگ (یک کامپیوتر دیجیتالی با برنامه ریزی مناسب) برای اینکه بتواند در آزمون، موفق شود، نیاز دارد که بازپرس را در تشخیص فریب دهد. این در حالی است که ارتباط ماشین تورینگ با بازپرس فقط از طریق مکالمه زبانی (تعامل زبانی) از طریق چاپگرها امکان پذیر است؛ یعنی باید از طریق تعامل زبانی بازپرس را فریب دهد. در واقع قابلیت تفکر در ماشین تورینگ در سؤال اول به قابلیت تعامل زبانی در سؤال دوم تبدیل شده است؛ بنابراین برای موجه کردن آزمون، نیاز به برقراری رابطه ای میان فکر و زبان داریم و در واقع به دیدگاهی نیازمندیم که به وجود رابطه ای نظام مند میان فکر و زبان قائل باشد؛ چرا که اگر کسی به وجود رابطه ای میان فکر و زبان باور نداشته باشد، نمی تواند برای آزمون تورینگ ارزش و اعتباری قائل شود. حال اگر بتوانیم استدلال کنیم که موفقیت کامپیوتر در آزمون تورینگ به معنای این است که کامپیوتر دارای قابلیت تعامل زبانی با دیگران است، یعنی بتوانیم درستی گزاره زیر را اثبات کنیم: «اگر ماشین تورینگ در آزمون موفق بشود، آنگاه قابلیت تعامل زبانی را دارد»، با در نظر گرفتن روندی که در ادامه شرح خواهیم داد، بر اساس رویکرد عدم اولویت زبان و فکر دیویدسون، می توانیم به دفاع از آزمون تورینگ بپردازیم.

روند مدنظر که در آن از صورت برهان تعدی ترکیب های شرطی ($P \rightarrow R$) « $P \rightarrow Q, Q \rightarrow R$ » استفاده شده، به شرح زیر است:

۱. اگر ماشین تورینگ در آزمون موفق شود، آنگاه دارای قابلیت تعامل زبانی است (گزاره ای که در پی اثبات آنیم)؛
۲. اگر موجودی دارای قابلیت تعامل زبانی است، می تواند فکر کند (نتیجه دوم نظریه دیویدسون)؛
۳. اگر ماشین تورینگ در آزمون موفق شود، می تواند فکر کند (با استفاده از صورت برهان تعدی ترکیب های شرطی و مقدمه های شماره ۱ و ۲).

همچنین طبق نتیجه اول، برای اینکه ماشین تورینگ بتواند فکر کند، نیاز به تعامل زبانی دارد. حال اگر بتواند در آزمون موفق بشود، باتوجه به اثبات گزاره شماره ۱ (در ادامه مقاله) یعنی دارای چنین قابلیت است. بنابراین حداقل می‌توانیم بگوییم که در صورت موفقیت ماشین در آزمون، ماشین قابلیت تعامل زبانی دارد و تناقضی با نتیجه اول ایجاد نمی‌شود.

بنابراین همان طور که ذکر شد، برای بررسی صحت این رویکرد در دفاع از آزمون تورینگ، باید مقدمه شماره ۱ بررسی شود که آیا موفقیت در آزمون به معنای قائل شدن قابلیت تعامل زبانی برای ماشین تورینگ است یا نه. بهتر است سؤال را به این صورت مطرح کنیم: آیا حروف‌نگاری سؤالاتی از طرف بازپرس خطاب به کامپیوتر و پاسخ دادن کامپیوتر به سؤالات (به طور مکتوب) به طوری که بازپرس را فریب دهد، به معنای تعامل زبانی کامپیوتر با بازپرس است؟ روشن است که اگر نیازی به فریب دادن بازپرس نباشد و کامپیوتر طوری برنامه‌ریزی شده باشد که صرفاً بتواند به سؤالات بازپرس پاسخ دهد، این ارتباط کامپیوتر با بازپرس به صورت مکتوب، تعامل زبانی به معنای مدنظر ما که لازمه آن فکر کردن است، نمی‌باشد؛ بنابراین آنچه این تعامل زبانی را متمایز می‌کند، توانایی کامپیوتر از طریق تحلیل و تشخیص موقعیت برای پاسخ‌گویی به شیوه‌ای است که بازپرس متوجه نشود مخاطب آن یک کامپیوتر است. در ادامه سعی می‌کنم به این مسئله بپردازم که چه موقع ما معتقدیم موجودی قابلیت تعامل زبانی را دارد. به عبارت دیگر، ویژگی‌های بارز تعامل زبانی چیست؟

همان طور که گفته شد، به منظور اینکه کامپیوتر بتواند در آزمون، موفق شود، باید بتواند بازپرس را از طریق مکالمه زبانی مکتوب، فریب دهد. چنین کاری دست کم به دو قابلیت نیاز دارد:

الف. توانایی تفسیر سخنان بازپرس یا به عبارت دیگر توانایی تفسیر سخنانی که در قالب سؤال به صورت مکتوب مطرح می‌شوند؛
ب. توانایی پاسخ‌گویی نادرست به سؤال‌های فوق به طوری که بازپرس متوجه نادرستی آن‌ها نشود یا در مورد ارزش صدق آن‌ها شک نکند؛ البته در این باره هم پاسخ‌گویی کامپیوتر به سؤال‌ها به صورت مکتوب است.

دیویدسون در مقاله «فکر و زبان» (دیویدسون: ۱۹۷۳) به ویژگی اصلی تعامل زبانی می‌پردازد. او معتقد است که ما معمولاً تصور می‌کنیم که قابلیت تعامل زبانی تا حد زیادی عبارت است از توانایی گفتاری یا مورد خطاب قرار دادن دیگران؛ اما مطابق با نظر او توانایی گفتاری، تنها نقشی غیرمستقیم در قابلیت زبانی دارد. آنچه از نظر دیویدسون برای داشتن قابلیت تعامل زبانی ضروری است، ایده مفسر (شخص استفاده‌کننده از زبان) است. مفسر کسی است که سخنان دیگران را می‌فهمد یا به عبارت دیگر کسی است که به تفسیر (تعبیر) سخنان دیگران می‌پردازد. دیویدسون معتقد است که یک گوینده باید توانایی تفسیر سخنان دیگران را داشته باشد و این توانایی لازمه داشتن قابلیت تعامل زبانی است (دیویدسون، ۱۹۷۵: ۱۵۷). اگر بنا بر یک تعریف مصطلح، تعامل زبانی را توانایی یک موجود برای تفسیر اظهارات دیگران و توانایی گفتاری (یا توانایی مورد خطاب قرار دادن دیگران) تعریف کنیم، مطابق با دیدگاه دیویدسون، داشتن توانایی تفسیر سخنان دیگران برای داشتن قابلیت تعامل زبانی، ضروری به نظر می‌رسد. این در حالی است که داشتن توانایی مورد خطاب قرار دادن دیگران از نقشی غیرمستقیم و کمتر برخوردار است. اکنون دوباره به موارد الف و ب که پیش‌تر بیان شده، مراجعه می‌کنیم. باتوجه به موارد الف و ب می‌بینیم که کامپیوتر برای موفقیت در

آزمون به هر دوی این توانایی‌ها به صورت مکتوب نیاز دارد و این، بدین معناست که کامپیوتر در صورتی که بتواند آزمون را با موفقیت پشت سر بگذارد، قابلیت تعامل زبانی با بازپرس را دارد.

همچنین این مسئله را از منظری دیگر نیز می‌توان بررسی کرد. همان طور که مطرح شد، دفاع از آزمون تورینگ با استفاده از رویکرد عدم اولویت زبان و فکر دیویدسون، به اثبات گزاره «اگر ماشین تورینگ در آزمون موفق بشود، آنگاه دارای قابلیت تعامل زبانی است»، نیاز دارد. من این گزاره را بر مبنای دیدگاه دیویدسون به گزاره «اگر کامپیوتر بتواند بازپرس را فریب دهد، آنگاه دارای قابلیت تفسیر سخنان بازپرس است»، تبدیل می‌کنم؛ زیرا فریب بازپرس از طریق کامپیوتر به معنای موفقیت کامپیوتر در آزمون است و قابلیت تفسیر اظهارات دیگران بر اساس نظر دیویدسون، ویژگی ضروری تعامل زبانی است؛ بنابراین این دو جایگزینی در گزاره اول، ایرادی نخواهد داشت.

اکنون با انجام این جایگزینی، به جای اینکه اثبات کنیم یک کامپیوتر دارای قابلیت تعامل زبانی است، اگر بتواند آزمون را با موفقیت پشت سر بگذراند، باید اثبات کنیم که یک کامپیوتر دارای قابلیت تفسیر سخنان بازپرس است، اگر بتواند بازپرس را فریب دهد. در این راستا، کامپیوتر برای اینکه بتواند در آزمون موفق شود، باید بتواند به سؤال‌های بازپرس، پاسخ‌های مناسب دهد؛ بنابراین برای درپیش گرفتن چنین رویکردی ابتدا کامپیوتر به ترجمه جملاتی که بازپرس آن‌ها را تایپ می‌کند، به زبان خود نیاز دارد؛ بنابراین کامپیوتر باید بتواند سؤالات بازپرس (که همان سخنان بازپرس به صورت مکتوب هستند) را از یک زبان بیگانه (زبان بازپرس) به یک زبان آشنا (زبان کامپیوتر) ترجمه کند. تا اینجا این عمل کامپیوتر به معنای تفسیر اظهارات بازپرس نیست. چنین عملیاتی، یعنی ترجمه جملات از یک زبان به زبان دیگر، به راحتی از طریق یک فرهنگ لغت عادی نیز انجام پذیر است و نیاز به تعامل زبانی ندارد. به عبارت دیگر، انجام چنین کاری تنها نیازمند به یک نظریه ترجمه است تا از طریق آن بتوان یک جمله در یک زبان را به جمله‌ای در زبان دیگر تبدیل کرد؛ صرف نظر از آنکه بدانیم معنای هر جمله در هر یک از آن دو زبان چیست. اما نکته‌ای که در اینجا مدنظر است و باید به آن توجه شود، این است که کامپیوتر چگونه می‌تواند به معنای یک جمله در زبان بازپرس دست یابد. به عبارت دیگر، کامپیوتر برای فهمیدن معنای جملات ترجمه شده از زبان بازپرس به زبان خود به چه چیز نیاز دارد؟

درحقیقت، آنچه کامپیوتر مدنظر در آزمون تورینگ را، در صورت موفقیت، از یک فرهنگ لغت عادی که آن هم به ترجمه جملات و عبارات از یک زبان به زبان دیگر می‌پردازد، متمایز می‌کند، مهارت کامپیوتر در درک معنای جملات ترجمه شده به زبان خود و درحقیقت تفسیر آن‌ها با توجه به موقعیت مدنظرست؛ یعنی کامپیوتر برای اینکه بتواند بازپرس را فریب دهد، علاوه بر اینکه به ترجمه سخنان بازپرس به زبان خود می‌پردازد، باید بتواند تشخیص دهد که آن جملات در آن موقعیت چه معنایی دارند. همچنین از آنجاکه معنای جملات بازپرس به ساختارشان وابسته است و هر جمله به صورت مجزا و مستقل معنا نمی‌شود، کامپیوتر باید بتواند ترجمه جملات را طوری در کنار هم قرار دهد که معنادار باشند و این مهارت، تنها ویژگی موجودی است که قابلیت تفسیر سخنان دیگران را دارد.

البته علاوه بر نکته فوق باید توجه داشت که یک فرهنگ لغت تنها قادر خواهد بود که کلمات محدودی را به صورت معناداری ترجمه کند؛ اما هنگامی که تعداد کلمات زیاد می‌شود، ترجمه‌ای که توسط آن ارائه می‌شود به صورت یک کل منسجم و معنادار نیست و صرفاً فرهنگ لغت به ترجمه تک تک جملات به صورت مستقل و فارغ از ارتباط هر جمله با جمله دیگر متن پرداخته است. با توجه به آنچه گفته شد، می‌توان گفت که ترجمه و تفسیر، متفاوت از هم هستند. به عبارت دیگر، قابلیتی که در فرهنگ لغت موجود است، تنها به قابلیت ترجمه کلمات و جملات محدود است؛ در صورتی که قابلیت موجود در کامپیوتر آزمون تورینگ در صورت موفقیت، قابلیت ترجمه و تفسیر هم‌زمان است.

دیویدسون در مقاله «تفسیر ریشه‌ای» ترجمه را گذر از یک زبان به زبان دیگر دانسته است؛ یعنی ترجمه رابطه‌ای است میان دو زبان. این در حالی است که او تفسیر را گذاری می‌داند که در یک بستر زبانی انجام می‌شود. او به خوبی این مطلب را چنین شرح می‌دهد:

«در حالت کلی، یک نظریه ترجمه^۵ شامل سه زبان است: زبان موضوعی^۶، زبان سوژه^۷ و فرا زبان (زبان‌هایی که ترجمه از آن‌ها و به آن‌ها انجام می‌شود و زبان نظریه که تعیین می‌کند کدام عبارات زبان موضوعی، کدام عبارات زبان مقصد را ترجمه می‌کند)^۸ و در این حالت کلی ما قادریم بدانیم کدام جملات زبان موضوعی، ترجمه کدام جملات زبان سوژه است؛ بدون اینکه بدانیم معنای هر کدام از جملات هر دو زبان چیست.»

با توجه به آنچه دیویدسون می‌گوید، می‌توان گفت یک فرا زبان می‌تواند به ما بگوید که ترجمه عبارت "Es regnet" در زبان آلمانی، عبارت "il pleut" در زبان فرانسوی است. درحقیقت، با فرض اینکه ما به هیچ‌یک از این دو زبان آشنایی نداشته باشیم، می‌توانیم از طریق زبان نظریه (فرا زبان) متوجه شویم که عبارت "il pleut" در زبان سوژه (فرانسوی)، ترجمه عبارت "Es regnet" در زبان موضوعی (آلمانی) است؛ بدون آنکه معنای هریک از این دو عبارت را بدانیم (دیویدسون، ۱۹۷۳: ۱۲۹ و ۱۳۰). همچنین مطابق با نظر دیویدسون، در روش ترجمه حالتی وجود دارد که در آن فردی که نظریه را فهمیده است، می‌تواند جملات زبان موضوعی را تفسیر کند. دیویدسون می‌گوید که اگر در حالتی قرار داشته باشیم که زبان سوژه اتفاقاً با زبان نظریه یکسان باشد، آنگاه فردی که نظریه را می‌فهمد بدون شک می‌تواند از راهنمای ترجمه برای تفسیر سخنان زبان بیگانه استفاده کند؛ یعنی در این حالت ما می‌توانیم با استفاده از راهنمای ترجمه (نظریه ترجمه) به تفسیر سخنان زبان بیگانه بپردازیم. او دلیل این امر را دو چیز می‌داند که از نظر او نظریه آن‌ها را بیان نمی‌کند: اول این واقعیت که زبان سوژه همان زبان آن فرد است و دوم دانش و معرفت آن فرد درباره چگونگی تفسیر سخنان در زبان خودش (دیویدسون، ۱۹۷۳: ۱۲۹). برای روشن‌تر شدن منظور دیویدسون می‌توان به مثال قبلی اشاره کرد؛ اما این بار به جای زبان فرانسوی، از زبان فارسی به عنوان زبان سوژه استفاده می‌شود؛ یعنی با استفاده از یک روش ترجمه از زبان آلمانی به فارسی می‌توان فهمید که ترجمه عبارت "Es regnet" در زبان

⁵. a theory of translation

⁶. the object language

⁷. the subject language

⁸. metalanguage

آلمانی، عبارت «باران می‌آید» در زبان فارسی است. در این مثال، زبان موضوعی، زبان آلمانی و زبان سوژه، زبان فارسی است و همچنین زبان سوژه با زبان نظریه یکسان بوده و فارسی است؛ بنابراین در این مثال ما می‌توانیم با این آگاهی که زبان سوژه با زبان ما (زبان نظریه: فارسی) یکسان است و با داشتن دانش چگونگی تفسیر عبارت «باران می‌آید» در زبان خودم، به تفسیر عبارت “Es regnet” در زبان بیگانه (زبان موضوعی: زبان آلمانی) پردازیم و برای مثال بفهمیم که معنای عبارت “Es regnet” در زبان موضوعی (آلمانی) این است که باران می‌آید.

نکته ای که در اینجا باید به آن توجه شود، واقعیت دومی است که دیویدسون به آن اشاره می‌کند؛ یعنی دانش و معرفت فرد درباره چگونگی تفسیر سخنان در زبان خودش؛ یعنی همان چیزی که نظریه ترجمه از توضیح آن عاجز است. او در این زمینه معتقدست که اگرچه یک نظریه ترجمه باید ترکیب و ساخت⁹ جملات را بخواند و ترجمه کند، اما داشتن این انتظار از یک نظریه ترجمه که بتواند چگونگی وابستگی جملات به ترکیب و ساختشان را ارائه دهد، انتظاری بی‌اساس است (دیویدسون، ۱۹۷۳: ۱۳۰). درحقیقت مطابق با دیدگاه او این وظیفه یک نظریه تفسیرست که بتواند برای جملات یک زبان، ساختار معنایی آن را آشکار کند؛ برای مثال تفسیر جملات پیچیده به طور نظام‌مندی به تفسیر جملات ساده‌تر وابسته است و این نکته‌ای است که از طریق یک نظریه تفسیر می‌تواند آشکار شود.

در ادامه دیویدسون پیشنهاد می‌کند که یک نظریه تفسیر برای زبان موضوعی می‌تواند به‌عنوان نتیجه ترکیب یک نظریه تفسیر آشکارکننده ساختار معنایی برای یک زبان شناخته‌شده و یک روش ترجمه از یک زبان ناشناخته به یک زبان شناخته‌شده در نظر گرفته شود. اکنون به‌منظور تفسیر یک جمله، فرض کنید که یک نظریه تفسیر رضایت‌بخش¹⁰ برای زبان خودمان را به نظریه ترجمه اضافه کنیم؛ یعنی به این صورت که راهنمای ترجمه برای هر جمله از زبان قابل ترجمه (زبان بیگانه) یک جمله از زبان مترجم (زبان آشنا) را بیرون دهد و سپس نظریه تفسیر مدّ نظر، برای این جملات آشنا تفسیری ارائه دهد. دیویدسون معتقدست که با این وضعیت آنچه را که می‌خواهیم در اختیار داریم؛ ولی به شکلی که بیش از حد بزرگ و حجیم است. از نظر او به‌منظور داشتن یک نظریه تفسیر، رجوع به زبان آشنای مادری غیرضروری است؛ زیرا واسطه‌ای قابل چشم‌پوشی میان تفسیر و سخنان زبان بیگانه است. به عبارت دیگر ادغام این دو نظریه، رجوع به زبان آشنا را مهمل می‌سازد؛ درنتیجه، با حذف رجوع به زبان آشنا آنچه باقی می‌ماند، یک نظریه تفسیر آشکارکننده ساختار معنایی برای زبان موضوعی است.

بنابراین می‌توان گفت در آزمون تورینگ نیز کامپیوتر به‌منظور پاسخ‌گویی به‌نحوی که بتواند بازپرس را فریب دهد، به یک نظریه تفسیر علاوه بر نظریه ترجمه نیاز دارد؛ به این صورت که ابتدا «راهنمای ترجمه» کامپیوتر برای هر جمله از زبان بازپرس، یک جمله از زبان کامپیوتر را بیرون می‌دهد. سپس نظریه تفسیر، برای این جملات آشنا تفسیری ارائه می‌دهد و باتوجه‌به این تفسیر و برنامه‌ریزی انجام‌شده متناسب با موقعیت مدّ نظر، پاسخ مناسب داده می‌شود. همچنین مطابق با نظر دیویدسون می‌توان گفت درمورد کامپیوتر نیز رجوع به زبان آشنای آن غیرضروری است؛ زیرا واسطه‌ای قابل چشم‌پوشی میان تفسیر و

⁹. structure

¹⁰. satisfactory

سخنان زبان بیگانه (زبان بازپرس) است؛ در نتیجه، با حذف رجوع به زبان آشنا (زبان کامپیوتر)، آنچه باقی می‌ماند، یک نظریهٔ تفسیر آشکارکنندهٔ ساختار معنایی برای زبان موضوعی، یعنی زبان بازپرس است.

بنابراین با ارائهٔ چنین توضیحاتی می‌توان گفت کامپیوتر برای اینکه بتواند بازپرس را فریب دهد و در آزمون موفق شود، نیاز به ترجمه و تفسیر اظهارات بازپرس دارد. همچنین با توجه به نظر دیویدسون، چنین قابلیت‌های ویژگی ضروری قابلیت تعامل زبانی است و این بدان معناست که کامپیوتر دارای قابلیت تعامل زبانی است؛ بنابراین گزارهٔ شمارهٔ ۱ که به دنبال اثبات آن بودیم، اثبات شده است و با استفاده از اصل دوم نظریهٔ دیویدسون (گزارهٔ شمارهٔ ۴)، می‌توان گفت که کامپیوتر مدّ نظر اگر بتواند در آزمون تورینگ موفق شود، باید بتواند فکر کند.

۵. آزمایش فکری اتاق چینی بر علیه آزمون تورینگ

جان سرل^{۱۱} با ارائهٔ آزمایش فکری اتاق چینی [1] بر این باور است که یک کامپیوتر قادر به شناسایی ویژگی‌های دلالت شناختی (معنایی)^{۱۲} سمبل‌ها نیست. به عبارت دیگر، از نظر او کاری که یک کامپیوتر انجام می‌دهد صرفاً شناسایی و دست‌کاری سمبل‌ها بر اساس ویژگی‌های نحوی^{۱۳} آن‌ها است؛ این در حالی است که پی بردن به تفسیر و معنای سمبل‌ها امری دلالت شناختی (معنایی) است. به طور خلاصه سرل می‌گوید که صرف دستور زبان و شناسایی نحوی سمبل‌ها ما را از معنانشناسی بی‌نیاز نمی‌کند. استدلال سرل را به نقل از خود او به صورت زیر می‌توان خلاصه کرد:

۱. ویژگی‌های نحوی ما را از ویژگی‌های دلالت شناختی (معنایی) بی‌نیاز نمی‌کند.

۲. برنامه‌های کامپیوتری کاملاً با ساختار نحوی یا صوری زبانی شان تعریف می‌شوند.

۳. ذهن محتویاتی معنانشناختی دارد. (دارای حیث التفاتی (درباره‌گی) هستند).

۴- نتیجه: هیچ برنامهٔ کامپیوتری به خودی خود نمی‌تواند سیستمی را دارای ذهن کند. (سرل، ۱۹۸۴: ۳۹)

همچنین باید اشاره کرد که اگر چه سرل معتقد است که کامپیوتر ابزاری است که به شناسایی و دست‌کاری سمبل‌ها می‌پردازد و اگر درست برنامه‌ریزی شود می‌تواند در راستای حفظ ویژگی‌های دلالت شناختی - همچون ارزش صدق - نیز گام بردارد اما این بدان معنا نیست که کامپیوتر به چنین امری آگاه است بلکه این کار توسط برنامه‌ریزان حرفه‌ای بر اساس ویژگی‌های نحوی امکان‌پذیر است.

حال می‌خواهیم به نکتهٔ مورد نظر بپردازیم. همانطور که پیش‌تر بیان شد تورینگ به طراحی آزمونی یک طرفه می‌پردازد. یعنی اگر کامپیوتری بتواند در آزمونی که تورینگ طراحی کرده است موفق شود، این بدان معناست که دارای توانایی تفسیر سخنان بازپرس است. به عبارت دیگر، همانطور که پیش‌تر توضیح داده شد اگر کامپیوتر بتواند بازپرس را در تشخیص فریب دهد، چاره‌ای نیست جز این که بپذیریم کامپیوتر توانسته است به تفسیر سخنان بازپرس بپردازد و علاوه بر ویژگی‌های نحوی به شناسایی

11. J. Searl

12. semantic

13. syntax

ویژگی های معنایی نیز پرداخته است. در چنین شرایطی، یعنی اگر بپذیریم که موفقیت کامپیوتر در آزمون تورینگ به معنای داشتن قابلیت شناسایی ویژگی های معنایی است، در نهایت و حداکثر سرل می تواند ادعا کند که ساخت چنین کامپیوتری - یعنی کامپیوتر که بتواند آزمون تورینگ را پشت سر بگذارد - امکان پذیر نیست و یا این که بگوید هیچ کامپیوتری قادر به موفقیت در آزمون تورینگ نیست، چرا که مطابق با نظر سرل یک برنامه کامپیوتری صرفا بر مبنای ویژگی های نحوی تعریف می شود و این در حالی است که موفقیت کامپیوتر در آزمون تورینگ نیاز به چیزی بیش از شناسایی ویژگی های نحوی دارد. بهتر است گزاره های بدست آمده تا اینجا را در قالب گزاره های منطقی صورت بندی کنیم:

۱- مطابق با فهم ما از آزمون تورینگ در این مقاله می توان گفت: اگر یک کامپیوتر در آزمون تورینگ موفق شود، آنگاه دارای قابلیت شناسایی ویژگی های معنایی (قابلیت تفسیر) است.

۲- همچنین با توجه به قاعده منطقی عکس نقیض و مقدمه شماره ۱ می توان گفت: اگر یک کامپیوتر دارای قابلیت شناسایی ویژگی های معنایی نباشد، آنگاه نمی تواند در آزمون تورینگ موفق شود.

۳- مطابق با دیدگاه سرل نیز می توان گفت: یک کامپیوتری دارای قابلیت شناسایی ویژگی های معنایی نیست.

۴- مطابق با قاعده وضع مقدم و مقدمه های شماره ۲ و ۳ این نتیجه بدست می آید: یک کامپیوتری نمی تواند در آزمون تورینگ موفق شود.

بنابراین همانطور که ملاحظه می شود این ادعای سرل نافی ادعای تورینگ نخواهد بود، چرا که ادعای تورینگ صرفا ادعایی شرطی است. در حقیقت بر اساس آزمون تورینگ و مطابق با آنچه تا اینجا در این نوشته گذشته است اگر روزی کامپیوتری ساخته شود که بتواند این آزمون را پشت سر بگذارد، برای این کار نیاز به ترجمه و تفسیر سخنان بازپرس دارد و این بدان معناست که کامپیوتر قادر به پی بردن به ویژگی های معنایی است و متعاقبا بر اساس نظریه دیویدسون می توان گفت که این کامپیوتر می تواند فکر کند.

همچنین اگر بپذیریم که موفقیت کامپیوتر در آزمون تورینگ به معنای داشتن قابلیت شناسایی ویژگی های معنایی است (یعنی پذیرفتن گزاره ای که در این مقاله در جستجوی آن بودیم و پیش تر نشان دادیم)، این کامپیوتر در آزمون تورینگ که توسط سرل ارائه می شود - «آزمون تورینگ برای فهم چینی» [2] (سرل، ۱۹۸۰: ۴۲۰) - اصلا قابلیت موفقیت ندارد. در حقیقت، سرل نمی تواند ادعا کند که اگر اتاق چینی به جای کامپیوتر در آزمون قرار بگیرد و آزمون را با موفقیت پشت سر بگذارد این بدان معناست که کامپیوتر موفق شده است اگرچه هوشمند نیست. بنابراین، آزمایش فکری سرل اصلا نمی تواند مثال نقضی برای تلقی تورینگ از هوشمندی باشد چرا که مطابق فهم ما از آزمون تورینگ و بر اساس دیدگاه سرل در مورد محدودیت یک برنامه کامپیوتری در شناسایی ویژگی های معنایی، چنین کامپیوتری اولاً و بالقوه دارای قابلیت لازم برای موفقیت در آزمون نیست.

۶. جمع‌بندی

با ارائه چنین توضیحاتی می‌توان گفت کامپیوتر برای اینکه بتواند بازپرس را فریب دهد و در آزمون موفق شود، نیاز به ترجمه و تفسیر سخنان بازپرس دارد. همچنین با توجه به نظر دیویدسون چنین قابلیت، قابلیت تفسیر سخنان دیگران، ویژگی ضروری قابلیت تعامل زبانی است و این بدان معناست که کامپیوتر دارای قابلیت تعامل زبانی است؛ بنابراین گزاره شماره ۱ که به دنبال اثبات آن بودیم، اثبات شده است و با استفاده از نتیجه دوم نظریه دیویدسون (مقدمه شماره ۲) مطابق با روند شرح داده شده می‌توان گفت که کامپیوتر مدّ نظر در صورت موفقیت در آزمون، می‌تواند فکر کند؛ بنابراین در این مقاله نظریه عدم اولویت زبان و فکر دیویدسون در حکم رویکردی برای پرکردن شکاف تبیینی موجود در مقاله تورینگ ارائه شد.

۷. تشکر و قدردانی

نویسنده مراتب سپاسگزاری خود را به جناب آقای دکتر مهدی نسرین، عضو سابق هیئت علمی گروه مطالعات علم مؤسسه پژوهشی حکمت و فلسفه ایران، به دلیل راهنمایی‌های بارز ایشان اعلان داشته و نیز از داوران محترم ناشناس که مطالب سودمندی را پیشنهاد کرده‌اند، سپاسگزاری می‌نماید.

پی‌نوشت:

۱. برای آشنایی با آزمایش فکری اتاق چینی نگاه کنید به Searle 1984, pp. 28-41
۲. سرل می‌پندارد که آزمون تورینگ انواع مختلفی دارد. به نظر او یک نوع آن آزمون تورینگ است که برای تشخیص این که آیا فرد چینی می‌فهمد و یا نه طراحی شده است؛ او نام آن را «آزمون تورینگ برای فهم چینی» می‌گذارد. در حقیقت با چنین درکی از آزمون تورینگ مطابق با نظر سرل یک موجود چینی می‌فهمد اگر یک بازپرس نتواند او را از فردی که چینی زبان مادری او است، تشخیص دهد. او با چنین طرحی می‌خواهد نشان دهد که اگر موجودی که فهم چینی ندارد مطابق با آزمایش فکری اتاق چینی بتواند بازپرس را در تشخیص فریب دهد، این بدان معناست که در آزمون موفق شده است و مطابق با دیدگاه تورینگ دارای فهم از زبان چینی است، در حالی که در واقع فهمی از زبان چینی برای او وجود ندارد و این مثال نقضی برای تلقی تورینگ از هوشمندی است. نگاه کنید به Searle 1980
۳. در این زمینه نگاه کنید به: Dennett 1987 و Paul and Patricia Churchland 1990

منابع

Bibliography

Churchland, P. M. & Churchland, P. S. (1990). Could a Machine Think?. *Scientific American* 262, 32-37.

Ben-Yami, H. (1993). A Note on the Chines Room. *Synthese* (95) 2, 169-172.

Davidson, D. (1973). Radical Interpretation. *In D. Davidson, Inquiries into Truth and Interpretation* (2rd ed., pp. 125-140). Oxford: Clarendon Press.

Davidson, D. (1975). Thought and Talk. *In D. Davidson, Inquiries into Truth and Interpretation*. (2rd ed., pp. 155-170). Oxford: Clarendon Press.

Davidson, D. (1990). Turing's Test. *In D. Davidson, Problems of Rationality*. (pp. 77-86). Oxford: Clarendon Press.

Dennett, D. (1987). Fast Thinking. *In The Intentional Stance*. (pp. 323-338). Cambridge, Mass.: MIT Press.

Searle, J. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3(3) : 417-457.

Searle, J. (1984). *Minds, Brains and Science*. (pp. 28-41). Harvard University Press.

Turing, A. M. (1950). Computing Machinery and Intelligences. *Mind* 59, 433-460.