حکمت و فلسفه
سال ششم، شماره اول، بهار ۱۳۸۹، صص ۳۳-۴۷

Hekmat va Falsafeh
(Wisdom and Philosophy)
Vol. 6, No. 1, March 2010

نقد سه استدلالِ گودلی در فلسفهٔ ذهن

محمدصالح زارع‌پور*
سید محمدعلی حجتی**

چکیده

برخی فیلسوفان برای اثبات ناممکن بودن مدل کردن ذهن با ماشین از قضیهٔ گودل و نتایج آن استفاده کرده‌اند. در فلسفهٔ ذهن استدلالات مبتنی بر قضیهٔ گودل را استدلالات گودلی می‌نامند. در این مقاله قصد داریم که سه استدلال گودلی را مورد نقد و بررسی قرار دهیم. این استدلال‌ها به ترتیب توسط رودی روکر، جان راندولف لوکاس و راجر پنروز ارائه شده‌اند. تلاش می‌کنیم تا نشان دهیم: (الف) بخشی از استدلال روکر مقبول و بخشی از آن مردود است؛ (ب) استدلال لوکاس به کلی مردود است؛ و (پ) بخشی از استدلال پنروز مقبول و بخشی از آن مشکوک است.

واژگان کلیدی: *این‌همانی ذهن و ماشین، نظام صوری، قضیهٔ گودل، استدلال‌های گودلی، رودی روکر، جان لوکاس و راجر پنروز.*

*. دانشجوی کارشناسی ارشد فلسفه، دانشگاه تربیت مدرس، zarepour@modares.ac.ir.
**. دانشیار گروه فلسفه، دانشگاه تربیت مدرس، hojatima@modares.ac.ir
[تاریخ دریافت: ۱۳۸۹/۲/۲۹؛ تاریخ تایید: ۱۳۸۹/٤/٦]

مقدمه

یکی از مهم‌ترین سؤالاتِ حوزهٔ فلسفهٔ ذهن (philosophy of mind) و هوش مصنوعی (artificial intelligence) این است که «آیا می‌توان ماشینی طراحی کرد که ذهن انسان (یا ذهن انسانی) را مدل کند؟» به بیان دیگر «آیا، علی‌الاصول، طراحی و ساخت ماشینی که همهٔ قابلیت‌های ذهن انسان را داشته باشد، ممکن است؟» ذهن انسان قابلیت‌های بسیار متنوع و متفاوتی دارد که به راحتی نمی‌توان پذیرفت که یک ماشین می‌تواند همهٔ این قابلیت‌ها را مدل کند؛ قابلیت‌هایی نظیر تفکر، تخیل، آگاهی، خودآگاهی و غیره. با این همه شاید بتوان ماشینی طراحی کرد که بخشی از این قابلیت‌ها را داشته باشد. مثلاً شاید بتوان ماشینی طراحی کرد که تفکر انسانی را مدل کند. یعنی بتواند درست مثل یک انسان فکر کند و له یا علیه درستی گزاره‌های مختلف استدلال کند. صد البته تفکر انسانی هم صور مختلف و متنوعی دارد که قضاوت در مورد امکان یا عدم امکان مدل کردن تفکر با یک ماشین را دشوار می‌کند. بنابراین خود را تنها به یکی از صور تفکر انسانی، یعنی تفکر ریاضی (mathematical reasoning)، محدود می‌کنیم و سؤال بالا را در خصوص این قابلیت ذهن انسانی بازنویسی می‌کنیم: «آیا می‌توان ماشینی طراحی کرد که همهٔ قابلیت‌های ذهن انسانی را در تفکر ریاضی داشته باشد؟». برخی فیلسوفان ذهن برای پاسخ منفی دادن به این سؤال از قضایای ناتمامیت گودل (Gödel's incompleteness theorems) استفاده کرده‌اند. ایشان تلاش می‌کنند تا با استفاده از دو قضیهٔ ناتمامیت گودل و به خصوص قضیهٔ اول ناتمامیت، نشان دهند که نمی‌توان ماشینی طراحی کرد که همهٔ قابلیت‌های ذهن انسانی را در تفکر ریاضی داشته باشد. در فلسفهٔ ذهن این استدلالاتِ مبتنی بر قضیهٔ گودل را استدلالات گودلی (Gödelian arguments) می‌نامند. سابقهٔ این استدلالات به خود گودل می‌رسد (Rucker, 2007, p. 171). ما در این مقاله قصد داریم سه استدلال گودلی را مورد نقد و بررسی قرار دهیم. این استدلال‌ها به ترتیب توسط رودی روکر (Rudy Rucker)، جان راندولف لوکاس (John Randolph Lucas) و راجر پنروز (Roger Penrose) ارائه شده‌اند. روشن است که پیش از پرداختن به این استدلال‌ها باید روشن کنیم که مقصود این فیلسوفان از یک ماشین چیست و چه وقت می‌توان گفت که یک ماشین قابلیت‌های ذهن انسانی را در تفکر ریاضی دارد. به همین منظور در بخش اول این مقاله تعریف دقیقی از ماشین ارائه می‌دهیم. در بخش دوم، قضایای ناتمامیت گودل را – که قضایایی دربارهٔ نظام‌های صوری هستند – شرح می‌دهیم. در بخش‌های سوم، چهارم و پنجم این مقاله به ترتیب به شرح و نقد استدلال‌های روکر، لوکاس و پنروز می‌پردازیم.

یک. ماشین به‌مثابهٔ یک نظام صوری

معمولاً در مباحث این‌چنینی فلسفهٔ ذهن و هوش مصنوعی دو عبارت ماشین (یا رایانه) و نظام صوری (formal system) جایگزین یکدیگر می‌شوند (Franzén, 2005, p. 115)؛ و در واقع همان تعریف نظام صوری برای ماشین هم به کار می‌رود. برای تعریف دقیق یک نظام صوری، نیازمند ارائهٔ چند تعریف مقدماتی هستیم:

تعریف یک. *زبان صوری* متشکل از دو مجموعهٔ الفبای زبان و نحو (grammar) زبان است. الفبای هر زبان مجموعه‌ای متناهی از نمادهاست. از قرار گرفتن این نمادها در کنار هم رشته‌های زبانی (linguistic strings) ساخته می‌شوند. برخی از این رشته‌ها درست‌ساخت (well-formed) هستند و برخی دیگر خیر. نحو زبان شامل همهٔ قواعد ساختِ رشته‌های زبانیِ درست‌ساخت از الفبای زبان است.

تعریف دو. مجموعهٔ E از رشته‌های زبانی را شمارای محاسبه‌پذیر (computably enumerable) می‌گوییم اگر بتوان برنامه‌ای نوشت که بتواند با صرف‌نظر از محدودیت‌های زمان و انرژی و حافظه، همهٔ اعضای این مجموعه را محاسبه کند. اگر این برنامه را به یک رایانه بدهیم و از آن بخواهیم تا همهٔ خروجی‌های این برنامه را چاپ کند، آن‌گاه به ازای هر رشتهٔ زبانی s که عضو E باشد، زمانی فرا می‌رسد که s در خروجی چاپ خواهد شد.

برای مثال، اگر هر رشتهٔ زبانی را که از کنار هم قرار گرفتن تعدادی از حروف الفبای انگلیسی[1] به دست می‌آید، یک کلمهٔ انگلیسی بدانیم (فارغ از معنادار بودن یا نبودن این کلمه)، آن‌گاه مجموعهٔ همهٔ کلمات زبان انگلیسی، یک مجموعهٔ شمارای محاسبه‌پذیر است. برنامه‌ای که می‌توانیم برای محاسبهٔ رشته‌های این مجموعه بنویسیم از ترتیب لغت‌نامه‌ای (lexicographic) استفاده می‌کند و همهٔ کلمه‌های زبان انگلیسی را به ترتیبی که در یک لغت‌نامه ظاهر می‌شوند، چاپ می‌کند. یعنی اول رشته‌های به طول یک یا کلمات یک حرفی را به ترتیب حروف الفبای انگلیسی چاپ می‌کند. بعد رشته‌های به طول دو یا کلمات دو حرفی را چاپ می‌کند؛ ابتدا کلماتی که با a آغاز می‌شوند، بعد کلماتی که با b آغاز می‌شوند و به همین ترتیب تا انتها. بعد هم کلمات سه‌حرفی، چهارحرفی و غیره را چاپ می‌کند. به این ترتیب می‌توانیم مطمئن باشیم که به ازای هر کلمه زمانی خواهد رسید که آن کلمه در خروجی چاپ شود.

حال تنها به یک تعریف مقدماتی دیگر نیاز داریم تا بتوانیم یک نظام صوری را به‌طور دقیق تعریف کنیم:

> **تعریف سه**. مجموعهٔ E از رشته‌های زبانی را تصمیم‌پذیر[2] می‌گوییم اگر بتوان برنامـه‌ای نوشت که یک رشتهٔ دلخواه s را به عنوان ورودی بگیرد و در خروجی معین کند که آیا این رشته عضوی از E هست یا خیر.

روشن است که هر مجموعهٔ تصمیم‌پذیر، شمارای محاسبه‌پذیر هم هست. برای اثبات این حکم فرض کنید که E یک مجموعهٔ تصمیم‌پذیر از رشته‌های زبانیِ زبان L باشـد. بـرای ایـن‌کـه نـشان دهـیم E شمارای محاسبه‌پذیر است باید برنامه‌ای بنویسیم که به مرور زمان همهٔ اعضای E در خروجی آن ظـاهر شوند. چون E تصمیم‌پذیر است، پس برنامه‌ای مثل P وجود دارد که اگر رشتهٔ دلخواه s را بـه P بـدهیم، معین می‌کند که s عضو E است یا خیر. پس با پیروی از ترتیب لغت‌نامه‌ای اولین رشتهٔ زبانیِ زبـان L را می‌سازیم، بعد آن رشته را به برنامهٔ P می‌دهیم، اگر مشخص شـد کـه آن رشـته در E هـست، آن را در خروجی چاپ می‌کنیم و در غیر این صورت بدون فرستادن آن رشته به خروجـی، رشـتهٔ زبـانی بعـدی را می‌سازیم و آن را به برنامهٔ P می‌دهیم و همین کار را به طور مـداوم تکـرار مـی‌کـنیم. بـه ایـن ترتیـب برنامه‌ای نوشته‌ایم که با مرور زمان همـهٔ اعـضای E را در خروجـی ظـاهر مـی‌کنـد. بنـابراین E یـک مجموعهٔ شمارای محاسبه‌پذیر است.

اگر متمم E را مجموعهٔ همهٔ رشته‌هایی از زبان L تعریف کنـیم کـه عـضو E نیـستند، آن‌گـاه ثابـت می‌شود که E تصمیم‌پذیر است اگر و تنها اگر هم و هم متمم آن شمارای محاسبه‌پـذیر باشـند. در عـین حال تورینگ ثابت کرده است (Turing, 1936) که برخی مجموعه‌های شـمارای محاسـبه‌پـذیر وجـود دارند که تصمیم‌پذیر نیستند. به بیان دیگر یعنی مجموعه‌های شمارای محاسبه‌پـذیری وجـود دارنـد کـه متمم آن‌ها شمارای محاسبه‌پذیر نیست (Franzén, 2005, pp. 67-70). حـالا مـی‌تـوانیم یـک نظـام صوری را به‌طور دقیق تعریف کنیم:

> **تعریف چهار**. یک نظام صوری متـشکل از یـک زبـان صـوری، مجموعـه‌ای از اصـول موضوعه (axioms) و مجموعه‌ای از قواعد استنتاج (inference rules) است. بخش‌های مذکور باید این سه شرط را احراز کننـد: (۱) مجموعـهٔ جملـه‌هـای ایـن زبـان صـوری یـا همان مجموعهٔ رشته‌های درست‌ساخت این زبان، باید یک مجموعهٔ تصمیم‌پذیر باشـد. (۲) اصول موضوعه باید مجموعه‌ای از جمله‌های زبانِ این نظـام صـوری باشـد. (۳) مجموعـهٔ اصول موضوعه و قواعد استنتاج این نظام صوری بایـد طـوری باشـند کـه مجموعـهٔ همـهٔ قضایای (theorems) این نظام صوری، یک مجموعهٔ شمارای محاسبه‌پـذیر باشـد. یـک قضیهٔ این نظام صوری، بنا بر *قرارداد*، یا یکـی از اصـول موضـوعهٔ ایـن نظـام اسـت و یـا جمله‌ای است که با اعمال قواعد استنتاج بر روی اصول موضوعه در متناهی گام[3] به دسـت می‌آید.

شرط (۱) در **تعریف چهار** در حقیقت به این معناست که ما روشی در اختیار داریم که بـا اسـتفاده از آن می‌توانیم تشخیص دهیم که آیا یک رشتهٔ دلخواه s جمله‌ای از زبان این نظام صوری هست یـا خیـر. شرط (۳) هم به این معنی است که اگر زمان و انرژی و حافظهٔ کافی در اختیار داشـته باشـیم مـی‌تـوانیم همهٔ قضیه‌های این نظام صوری را استنتاج کنیم. یعنی هیچ قضیه‌ای از این نظام صوری وجـود نخواهـد داشت که هرگز به آن نرسیم.

چنان که پیش‌تر گفتیم در مباحث مربوط به این‌همانی ذهن و ماشین در فلسفهٔ ذهن معمولاً دو مفهوم ماشین (رایانه) و نظام صوری جایگزین هم می‌شوند. چون برای هر نظام صـوری S مـی‌تـوان برنامـه‌ای مانند P نوشت که با اجرای آن توسط یک ماشین یا یک رایانه همهٔ قضایای آن نظام صوری تولید شوند؛ و در مقابل به ازای هر برنامهٔ P برای تولید رشته‌های زبانی توسط یک ماشین یا یک رایانـه، یـک نظـام صوری مانند S وجود دارد که قضایای آن همان جمله‌های تولید شده توسط P هستند. در واقع همـان دو شرط (۱) و (۳) در **تعریف چهار** هستند که این جایگزینی ماشـین‌هـا و نظـام‌هـای صـوری را ممکـن می‌کنند.

با توصیفات مذکور می‌توان مسئلهٔ این‌همانی ذهـن و ماشـین در حـوزهٔ تفکـر ریاضـی را بـه صـورت زیر تقریر کرد: «آیا ماشینی وجود دارد که بتواند، با صرف نظر از زمان و انـرژی و حافظـه، همـهٔ احکـام ریاضی که انسان علی‌الاصول می‌تواند آن‌ها را اثبـات کنـد و بدانـد، در خروجـی خـود چـاپ کنـد؟» یـا معادلاً «آیا یک نظام صوری وجود دارد که قضایای آن همهٔ حقایق ریاضی باشد که انسان علـی‌الاصـول می‌تواند آن‌ها را بداند؟» چنان‌که پیش‌تر گفتیم برخی فیلسوفان برای پاسـخ منفـی دادن بـه ایـن سـؤال از قضایای ناتمامیت گودل استفاده کرده‌اند. پیش از پرداختن به استدلال‌های ایـن فیلـسوفان بایـد ابتـدا ببینیم خود قضیهٔ گودل دقیقاً چه می‌گوید. در بخش بعـد قـضایای اول و دوم ناتمامیـت گـودل را شـرح می‌دهیم.

دو. قضایای ناتمامیت گودل

در این بخش تلاش می‌کنیم تا بدون پرداختن به اثبـات ریاضـی قـضایای ناتمامیـت گـودل، صـورتِ این قضایا را به زبانی ساده و تا حـد ممکـن دقیـق شـرح دهـیم. ابتـدا بـه قـضیهٔ اول ناتمامیـت گـودل می‌پردازیم:

> **قضیهٔ اول ناتمامیت.** هر نظام صوری که مقدار معینـی از حـساب مقـدماتی را در خـود داشته باشد، نمی‌تواند هم سازگار (consistent) باشد و هم تمـام (complete). بـه‌طـور دقیق‌تر برای هر نظام صوری سازگار که مقدار معینی از حساب را در خود داشته باشد، جملهٔ حسابی صادقی وجود دارد که در آن نظام اثبات‌ناپذیر (unprovable) اسـت. یعنـی در آن نظام نه آن جمله اثبات می‌شود و نه نقیض آن.

اثبات این قضیه و نیز طرح کلی اثبات قضیهٔ دوم ناتمامیت در مقالهٔ ۱۹۳۱ گودل آمده است.[۴] گودل قصد داشت تا در مقالهٔ دیگری جزئیات اثبات قضیهٔ دوم ناتمامیت را هم بیاورد اما آن مقاله هیچ‌گاه منتشر نشد. صد البته صورت اصلی قضیهٔ گودل در مقایسه با آنچه ما در بالا ذکر کردیم قدری فنی‌تر و دقیق‌تر است.[۵] برای مثال باید به طور دقیق روشن شود که منظور از این‌که یک نظام صوری مقدار معینی از حساب مقدماتی را در خود *دارد*، چیست. هم‌چنین باید روشن شود که یک نظام صوری چه مقدار از حساب مقدماتی را باید در خود داشته باشد تا شرایط قضیهٔ اول ناتمامیت را ارضا کند. به طور تلویحی می‌توان گفت که اگر یک نظام صوری بتواند گزاره‌هایی را که در خصوص جمع و تفریق و ضرب و تقسیم اعداد طبیعی وجود دارد، در زبان خود بیان کند آن‌گاه شرایط قضیهٔ اول ناتمامیت را ارضا می‌کند. شرایط قضیهٔ دوم ناتمامیت گودل هم مشابه قضیهٔ اول او است؛ و به همین دلیل هر نظامی که شرایط یکی از این دو قضیه (با تقریری که ما در اینجا داشتیم) را ارضا کند، شرایط قضیهٔ را هم ارضا می‌کند.

> **قضیهٔ دوم ناتمامیت.** هر نظام صوری سازگار که مقدار معینی از حساب را در خود داشته باشد، نمی‌تواند سازگاری خودش را داخل خود این نظام ثابت کند.

اهمیت قضایای گودل در این است که اگر نظام صوری S دارای شرایط قضیهٔ گودل باشد آن‌گاه این قضیه در خودِ S ثابت می‌شود. یعنی در خود S ثابت می‌شود که «اگر S سازگار باشد، آن‌گاه جملهٔ صادقی در زبان این نظام صوری وجود دارد که در S اثبات‌ناپذیر است» (قضیهٔ اول ناتمامیت)؛ و نیز در خود S ثابت می‌شود که «اگر S سازگار باشد، آن‌گاه سازگاری S در خود S اثبات‌ناپذیر است» (قضیهٔ دوم ناتمامیت). در بخش‌های بعدی این مقاله به بیان و نقد سه استدلال گودلی در فلسفهٔ ذهن می‌پردازیم.

سه. استدلال روکِر

اولین استدلالی که در این مقاله به بیان و نقد آن می‌پردازیم، استدلالی از رودی روکر است. در این استدلال روکر تلاش می‌کند تا با همان شیوه‌ای که گودل قضیهٔ گودل را ثابت می‌کند، نشان دهد که هیچ ماشین جهانیِ صدق[۶] یا حقیقت وجود نخواهد داشت. یعنی هیچ ماشینی وجود نخواهد داشت که بتواند پاسخ هر سؤالی را که از آن پرسیده شود، به درستی بدهد. استدلال روکر ربط آشکاری به آن دسته از استدلال‌های گودلی که به آن‌ها اشاره شد ندارد؛ و اساساً نمی‌خواهد نشان دهد که ماشینی وجود ندارد که همهٔ قابلیت‌های انسانی را در تفکر ریاضی داشته باشد. البته اگر استدلال او درست باشد، این نتیجه هم از آن اخذ می‌شود. در هر حال به نظر ما نقل و نقد این استدلال، دست‌ِکم، به دو دلیل مفید خواهد بود: (۱) به این دلیل که او در این استدلال از همان روشی استفاده می‌کند که گودل برای اثبات قضایای ناتمامیت استفاده کرده است. کلیت این روش عبارت است از پیدا کردن

یـک جملـهٔ خودارجـاعِ (self-referential) خـالی از تنـاقض و نـشان‌دادن ایـن‌کـه خـودِ ایـن جملـه اثبات‌ناپذیر است (در قضیهٔ ناتمامیت) و یا صدق و کذبِ آن برای ماشـین جهـانی صـدق نـامعلوم اسـت (در استدلال روکر). در مباحث مربوط به قضیهٔ گودل به این جمله‌ها جمله‌هـای گـودلی مـی‌گوینـد.[۷] (۲) انتقاد وارد بر استدلال روکر بسیار شبیه به انتقادات وارد بر برخی استدلال‌هـای گـودلی در فلـسفهٔ ذهـن است.

روکر استدلال خود را در قالب یک داستانک طـرح مـی‌کنـد (Rucker, 2007, pp. 161-2). مـا در اینجا داستانک او را نقل به مضمون می‌کنیم. فرض کنید کسی گـودل را بـه ماشـین جهـانی صـدق یـا UTM، یعنی ماشینی که می‌تواند با صرف زمان و انرژی کافی پاسـخ همـهٔ سـؤال‌هـا را بدهـد، معرفـی می‌کند. از آنجایی که UTM یک ماشین است پـس حتمـاً برنامـه‌ای مثـل $P\ (UTM)$ وجـود دارد کـه UTM با گرفتن سؤال‌ها به عنوان ورودی و بعد از اجرای برنامهٔ $P\ (UTM)$ پاسخ ایـن سـؤال‌هـا را بـه عنوان خروجی ارائه می‌دهد. گودل جمله‌ای می‌نویسد و از UTM در خصوص صدق و کذب ایـن جملـه سؤال می‌کند. جمله‌ای که او می‌نویسد این است: «ماشینی که با برنامهٔ $P\ \ (UTM)$ کار می‌کنـد، هرگـز نخواهد گفت که این جمله صادق است». مرجع ضمیر "این" در جملـهٔ مـورد نظـر گـودل خـود همـان جمله‌ای است که در آن ظاهر شده است. اگر این جمله را G بنامیم، آن‌گاه G معادل است با ایـن جملـه که «ماشین UTM هرگز نخواهد گفت که G صادق است». در مرحلهٔ بعد گودل لبخنـدی مـی‌زنـد و از UTM می‌پرسد که آیا G صادق است یا خیر؟

بعد خود گودل چنین استدلال می‌کند که اگر UTM بگوید که G صادق است، پس چنین نیست کـه جملهٔ «ماشین UTM هرگز نخواهد گفت که G صادق است» صادق باشد. یعنی این جمله کاذب اسـت؛ اما این جمله خود معادل G است و بنابراین، G کاذب است. بنابراین، اگر UTM بگویـد کـه G صـادق است، در حقیقت G کاذب است. پس اگر UTM بگوید که G صادق اسـت، دروغ گفتـه اسـت؛ و چـون UTM ماشین جهانی صدق است و هرچه بگوید راست است، پس UTM هرگز نخواهـد گفـت کـه G صادق است.

چنـان‌کـه در انتهـای بنـد قبـل دیـدیم ماشـین UTM هرگـز نخواهـد گفـت کـه G صـادق اسـت. بنابراین جملهٔ «ماشین UTM هرگز نخواهد گفت که G صـادق اسـت» یـک جملـهٔ صـادق اسـت و از آنجایی که این جمله معادل G است، پس G صادق است. گودل در نهایت با تکیـه بـر ایـن اسـتدلالات چنین می‌گوید که: «من حقیقتی را می‌دانم که UTM هرگز به آن اذعان نخواهد کرد. مـن مـی‌دانـم G صادق است، اما UTM هرگز نخواهد گفت که G صادق است. پس UTM واقعـاً یـک ماشـین صـدق جهانی نیست».

به نظر می‌رسد که بخشی از استدلالی که روکر در این داستانک آورده است درست اسـت و بخـشی از آن نادرست[۸]. بخش درست استدلال بالا این اسـت کـه UTM ماشـین جهـانی صـدق نیـست. بخـش نادرست استدلال بالا هم این است که گودل *می‌داند* یا ما *می‌دانیم* که G صادق است. حال بایـد ببینـیم که چرا یکی از این دو بخش درست است و دیگری نادرست. استدلال بالا بـرای نـشان دادن صـدق G

عمیقاً به این‌که UTM یک ماشین جهانی صدق است، وابسته است. در حقیقت آنچه از استدلال گودل در داستانک روکر نتیجه می‌شود صرفاً یک گزارهٔ شرطی است: «اگر UTM یک ماشین جهانی صدق باشد، آن‌گاه G صادق است»[9]. با این اوصاف استدلال به نفع این‌که UTM ماشین جهانی صدق نیست (و به نفع این‌که اساساً چنین ماشینی وجود ندارد) صحیح به نظر می‌رسد. این استدلال را قدری روشن‌تر بازنویسی می‌کنیم:

ماشین UTM یک ماشین جهانی صدق است (فرض).

جملهٔ G معادل است با جملهٔ «ماشین UTM هرگز نخواهد گفت که G صادق است» (تعریف).

ماشین UTM می‌گوید G صادق است (فرض).

پس جملهٔ «ماشین UTM هرگز نخواهد گفت که G صادق است» کاذب است (نتیجهٔ (۳)).

G کاذب است (نتیجهٔ (۲) و (۴)).

اگر UTM بگوید G صادق است، G کاذب است (معرفی شرطی از بندهای (۳) تا (۵)).

اگر UTM بگوید G صادق است، یک گزارهٔ کاذب گفته است (نتیجهٔ (۶)).

ماشین UTM هرگز نخواهد گفت که G صادق است (نتیجهٔ (۱) و (۷)).

G صادق است (نتیجهٔ (۲) و (۸)).

G صادق است و ماشین UTM هرگز نخواهد گفت که G صادق است (عطف (۸) و (۹)).

اگر UTM یک ماشین جهانی صدق باشد، آن‌گاه G صادق است و ماشین UTM هرگز نخواهد گفت که G صادق است (معرفی شرطی از بندهای (۱) تا (۱۰)).

UTM یک ماشین جهانی صدق نیست (نتیجهٔ (۱۱)).[10]

در توضیح استنتاج بند (۱۲) از بند (۱۱) باید گفت که گزاره‌ای که در بند (۱۱) آمده است، در حقیقت گزاره‌ای به شکل $P \supset \neg P$ و معادل با $\neg P$ است.

با این همه آن بخش از استدلال روکر که می‌گوید گودل می‌داند که G صادق است، غلط است. آنچه گودل می‌داند این است که «اگر UTM یک ماشین جهانی صدق باشد، آن‌گاه G صادق است». گودلِ داستنکِ روکر نمی‌تواند مدعی شود که «من می‌دانم که G صادق است». چون او نمی‌داند که UTM حقیقتاً یک ماشین جهانی صدق است. از فرض این‌که UTM یک ماشین جهانی صدق است نمی‌توان نتیجه گرفت که گودل می‌داند که G صادق است. بل‌که تنها از این‌که گودل می‌داند که UTM یک ماشین جهانی صدق است، می‌توان نتیجه گرفت که او می‌داند که G صادق است؛ و چون او نمی‌داند که UTM حقیقتاً یک ماشین جهانی صدق است، پس نمی‌داند که G صادق است. این نقص استدلال روکر در استدلال لوکاس هم وجود دارد؛ استدلالی که یکی از قدیمی‌ترین و مشهورترین استدلال‌های گودلی فلسفهٔ ذهن است. در بخش بعدی سعی می‌کنیم تا استدلال لوکاس را قدری دقیق‌تر از استدلال روکر نقد کنیم.

چهار. استدلال لوکاس

استدلال لوکاس در خصوص امکان‌ناپذیری مدل کردن ذهن بشری با ماشین‌ها یکی از قدیمی‌ترین استدلال‌های گودلی است که مناقشات و مباحثات بسیاری پیرامون آن در گرفته است. لوکاس این استدلال را در مقالهٔ ۱۹۶۱ خود ارائه کرده است. او در آن مقاله چنین می‌نویسد: «قضیهٔ گودل می‌گوید در هر نظام صوری که برای تولید حساب ابتدایی به اندازهٔ کافی قوی باشد، فرمول‌هایی وجود دارند که نمی‌توانند در درون آن نظام اثبات شوند؛ اما ما می‌توانیم ببینیم که این فرمول‌ها صادق هستند». لوکاس سعی می‌کند تا با تکیه بر این تقریر از قضیهٔ گودل نشان دهد که هیچ ماشینی نمی‌تواند مدل مناسبی برای ذهن انسان باشد. به نظر او از قضیهٔ گودل «چنین نتیجه می‌شود که به ازای هر ماشین مفروضی که سازگار باشد و قابلیت انجام [محاسباتِ] حساب مقدماتی را داشته باشد، فرمولی وجود دارد که آن ماشین نمی‌تواند آن فرمول را به عنوان فرمولی صادق معرفی کند؛ یعنی فرمولی که در آن نظام صوریِ [معادل با ماشین مذکور] اثبات‌ناپذیر است اما ما می‌توانیم ببینیم که صادق است. پس هیچ ماشینی نمی‌تواند مدلی کامل و مناسب برای ذهن باشد؛ یعنی اذهانِ [بشری] ذاتاً متفاوت از ماشین‌ها هستند».

اگر استدلال لوکاس درست باشد ذهن صدق جملهٔ گودلی ماشینی را که به عنوان مدل ذهن ارائه شده است، می‌داند؛ درحالی‌که آن ماشین نمی‌تواند آن جمله را اثبات کند. از آنجایی که جملهٔ گودلیِ نظام‌هایی که در شرایط قضیهٔ گودل صدق می‌کنند یک جملهٔ حسابی است و صدق آن معادل جواب داشتن یک معالهٔ دیوفانتی[11] (Diophantine equation) است، پس در حقیقت ذهن می‌داند که آن معادلهٔ دیوفانتی جواب دارد اما آن ماشین نمی‌تواند چنین چیزی را اثبات کند. یعنی آن ماشین نمی‌تواند تفکر ریاضی ذهن بشری و به طریق اولی قابلیت‌های تفکر ذهن بشری را مدل کند.

با این همه، استدلال لوکاس چنان‌که برخی منتقدان (Chalmers, 1995 & Franzén, 2005, p. 117) هم گفته‌اند، کاملاً غلط است. لوکاس می‌گوید ما می‌دانیم یا می‌توانیم بدانیم که جملهٔ گودلیِ نظام صوریِ متناظر با ماشینی که به‌عنوان مدل ذهن ما ارائه شده است، صادق است اما خود این ماشین نمی‌تواند این حقیقت را بداند؛ پس ذهن ما قابلیتی دارد که این ماشین ندارد؛ و بنابراین ماشین مذکور نمی‌تواند مدلی برای ذهن باشد. این تقریر لوکاس از قضیهٔ گودل غلط است. ما نمی‌دانیم و نمی‌توانیم بدانیم که جملهٔ گودلی آن ماشین صادق است، مگر این‌که بدانیم نظام صوری متناظر با آن ماشین سازگار است. حال آن‌که منطقاً ممکن است ماشینی به‌عنوان مدل ذهن ارائه شود که ما سازگاری آن را ندانیم و بنابراین، ندانیم که جملهٔ گودلی این ماشین صادق است. اگر چنین باشد، نمی‌توانیم با این استدلال مدعی شویم که ما صدق جمله‌ای را می‌دانیم که آن ماشین نمی‌تواند آن را اثبات کند. فلذا نمی‌توانیم با تکیه بر این دلیل بگوییم که این ماشین نمی‌تواند مدل ذهن ما باشد.

لوکاس سعی می‌کند در مقاله‌ای (Lucas, 1996) که بیش از سه دهه بعد از طرح استدلال اولیهٔ خود می‌نویسد به این انتقاد پاسخ دهد. او در این مقاله می‌گوید باید از کسانی که مدعی هستند مدل کردن ذهن با ماشین ممکن است، در خصوص سازگاری یا ناسازگاری ماشین مورد نظر آن‌ها سؤال کرد. اگر ماشین مورد نظر آن‌ها ناسازگار باشد، آن‌گاه اساساً این ماشین جالبِ هیچ توجهی نیست؛ چون هر چیزی ممکن است از آن استنتاج شود. اما اگر بگویند که این ماشین سازگار است، آن‌گاه مقدم شرطی مورد نظر در قضیهٔ گودل تحقیق یافته است و بنابراین، ما صدق جمله‌ای را می‌دانیم یا خواهیم دانست که ماشینی که به عنوان مدل ذهن پیش‌نهاد شده است، نمی‌تواند آن را اثبات کند.

به‌نظر می‌رسد که این تلاش لوکاس برای دفع انتقاد مذکور هم قرین توفیق نیست. فرض کنید نظام صوری متناظر با ماشینی که به عنوان مدلِ ذهنِ M پیش‌نهاد شده است، S و جملهٔ گودلی آن G باشد. حال اگر «S سازگار است» را با $Con(S)$ نشان دهیم، بنابر قضیهٔ گودل $S \vdash Con(S) \supset G$. اگر «فاعل شناسای A می‌داند که P» را با K_AP نشان دهیم، آن‌گاه باز هم بنابر قضیهٔ گودل $K_M(Con(S) \supset G)$. حال با این صوری‌سازی به خوبی می‌توانیم ایراد استدلال‌های لوکاس در دو مقاله‌اش را تبیین کنیم.

در مقالهٔ اول (Lucas, 1961) لوکاس مدعی می‌شود که ما می‌دانیم و ذهن بشری می‌داند که G صادق است؛ اما S نمی‌تواند G را اثبات کند. یعنی او می‌گوید K_MG برقرار است، اما $S \vdash G$ برقرار نیست. حال آن‌که این ادعا نادرست است. آنچه برقرار است، این است که $K_M(Con(S) \supset G)$ و برای برقراری K_MG باید $K_M(Con(S))$ برقرار باشد؛[۱۲] که نیست. یعنی ذهن بشری باید بداند که S سازگار است، اما هیچ دلیلی سرراستی وجود ندارد که ذهن چنین چیزی را می‌داند.

در مقالهٔ دوم (Lucas, 1996) لوکاس مدعی می‌شود که چون کسانی که مدعی هستند مدل‌کردن ذهن با ماشین ممکن است می‌گویند S سازگار است، پس S سازگار است و ما می‌دانیم که جملهٔ G صادق است. لکن باز هم ادعای او نادرست است. او می‌گوید از $K_M(Con(S) \supset G)$ و $Con(S)$ می‌توان نتیجه گرفت که K_MG؛ اما این ادعا نادرست است. آنچه درست است این است که از $K_M(Con(S) \supset G)$ و $K_M(Con(S))$ می‌توان نتیجه گرفت که K_MG. یعنی از صرف سازگاری S نمی‌توان نتیجه گرفت که ذهن بشری می‌داند که G صادق است. از این‌که «ما می‌دانیم که اگر S سازگار باشد، آن‌گاه G صادق است» نمی‌توان نتیجه گرفت که «اگر S سازگار باشد، آن‌گاه ما می‌دانیم که G صادق است».[۱۳] حال آن‌که به نظر می‌رسد لوکاس مرتکب چنین اشتباهی شده است. در حقیقت ایراد استدلال روکر هم در همین نکته نهفته است. او در داستانکِ خود از این‌که «گودل می‌داند که اگر UTM یک ماشین جهانی صدق باشد، آن‌گاه G صادق است» نتیجه گرفته است که «اگر UTM یک ماشین جهانی صدق باشد، آن‌گاه گودل می‌داند که G صادق است»؛ و چنان‌که دیدیم این نتیجه‌گیری و به تبَع آن بخشی از استدلال روکر نادرست است.

از مطالب فوق نباید چنین نتیجه گرفت که همهٔ استدلالات گودلی نامعتبر هستند. برخی فیلسوفان در استفاده از قضیهٔ گودل دقت و احتیاط بیش‌تری به خرج داده‌اند و به همان نسبت نتایج معتبرتری را هم به دست آورده‌اند. یکی از این فیلسوفان راجر پنروز است که در بخش بعد به یکی از استدلال‌های گودلی او می‌پردازیم.

پنج. استدلال اول پنروز

راجر پنروز در دو کتاب خود (Penrose, 1989 & 1994) به‌طور مفصل بـه نتـایج قـضیهٔ گـودل در حوزهٔ علوم شناختی و فلسفهٔ ذهن می‌پردازد. او در کتاب سایه‌های ذهن (Penrose, 1994) دو استدلال گودلی ارائه می‌دهد. استدلالات او در سال‌های بعـد موضـوع مباحثـات و مناقـشات جـدی فیلـسوفان و منطق‌دانان قرار گرفت[۱۴]. او در مقاله‌ای (Penrose, 1996) چکیدهٔ آراء خود را در این دو کتـاب آورده و تلاش کرده است تا به منتقدان خود پاسخ دهد. ما در اینجا قـصد داریـم بـه اسـتدلالی کـه مـشهور بـه استدلال اول پنروز است بپردازیم. او در فصل دوم کتاب سایه‌های ذهن چنین استدلال می‌کند که ذهـن بشری نمی‌تواند با ماشینی که سازگاری نظام صوری متناظر با آن را می‌دانیم، مدل شود. او بـرای اثبـات ادعای خود از قضیهٔ گودل به همان روش لوکاس استفاده مـی‌کنـد. بـا ایـن حـال اسـتدلال او بـرخلاف استدلال لوکاس معتبر است. چون ادعای پنروز قـدری محـدودتر از ادعـای لوکـاس اسـت و در حقیقـت استدلال لوکاس اگرچه برای اثبات ادعای خود او معتبر نیست اما برای اثبات ادعای پنـروز معتبـر اسـت. لوکاس می‌گوید ذهن بشری نمی‌تواند با هیچ ماشینی مدل شود؛ اما پنروز می‌گوید ذهن بشری نمی‌توانـد با هیچ *ماشینی که سازگاری نظام صوری متناظر با آن را می‌دانیم*، مدل شود.

فرض کنید ذهن بشری با ماشینی مدل شده باشد که نظام صوری متناظر با آن S باشد و ذهن بـشری سازگاری آن را بداند. پس از یک طرف داریم $K_MCon(S)$ و از ظرف دیگـر بنـابر قـضیهٔ گـودل داریـم $K_M(Con(S)\supset G)$. پس بنابر اصل K در منطق معرفت، K_MG. یعنی ذهن بشری می‌داند کـه گـزارهٔ G صادق است در حالی که S نمی‌تواند آن را اثبات کند. به همین دلیل S نمی‌تواند مدل خوبی بـرای ذهـن بشری باشد.

به‌نظر می‌رسد که این استدلال پنروز بی‌نقص باشد؛ اما مشکل در این است که ادعای او بسیار ضعیف است. یعنی صدقِ ادعای او برای اثبات ناممکن بودنِ مدل کردنِ ذهن بشری بـا ماشـینی کـه سـازگاری نظام صوری متناظر با آن را نمی‌دانیم، کافی نیست. منطقاً ممکن است ماشینی مثل S وجود داشته باشـد که ذهن بشری را مدل می‌کند و $\neg K_M(Con(S))$. در فصل سوم کتاب سایه‌های ذهن تلاش می‌کنـد تـا این کاستی در استدلالش را بر طرف کند. او مدعی می‌شود که نمی‌توان ذهن بـشری را بـا ماشـینی کـه نمی‌دانیم نظام صوری متناظر با آن سازگار است، مدل کرد. به بیان دیگر او مدعی است که اگر ذهن مـا با ماشینی که نظام صوری متناظر با آن S است مدل شود، آن‌گاه قطعاً ما می‌دانیم کـه S سـازگار اسـت. استدلال او تقریباً چنین است:[۱۵]

ما می‌دانیم که سازگار هستیم (یا ذهن بشری می‌داند که سازگار است).

۱- ما می‌دانیم که S ذهن ما را مدل می‌کند (یا ذهن بشری می‌داند که S آن را مدل می‌کند).

پس

ما می‌دانیم که S سازگار است (یا ذهن بشری می‌داند که S سازگار است).

به‌نظر می‌رسد که هر دو مقدمهٔ استدلال بالا قابل مناقشه باشند. اولاً ما از کجا می‌دانیم که سازگار هستیم؟ ثانیاً چه لزومی دارد که اگر یک نظام صوری S ذهن ما را مدل کند، آن‌گاه ما باید بدانیم که S ذهن ما را مدل می‌کند؟ روشن است که اگر به این دو سؤال پاسخ قانع‌کننده‌ای داده نشود، استدلال پنروز متقاعد کننده نخواهد بود. او تلاش زیادی کرده است تا مقدمهٔ یک را توجیه کند. ما در اینجا قصد نداریم تا قوت توجیهات او را بسنجیم؛ چون استدلالات او برای توجیه این مقدمه ربط چندانی به مباحث منطق فسلفی و فلسفهٔ منطق ندارد. با این همه حتی اگر فرض کنیم که توجیه او برای مقدمهٔ (۱) قانع‌کننده باشد، هم‌چنان مقدمهٔ (۲) مشکوک باقی می‌ماند. البته اگر فقط همین مقدمهٔ (۱) هم درست باشد و مقدمهٔ (۲) درست نباشد، بازهم می‌توان نتیجه گرفت که «اگر ماشینی ذهن ما را مدل کند، ما نمی‌دانیم که آن ماشین ذهن ما را مدل می‌کند». نتیجه‌ای که به‌نظر نمی‌رسد چندان هم ضعیف و بی‌فایده باشد.

نتیجه‌گیری

از مجموع مباحث فوق چنین برمی‌آید که: (۱) در استفاده از قضیهٔ گودل به‌منظور اقامهٔ براهین فلسفی، بدفهمی‌ها و سوء تعبیرهایی از این قضیه دیده می‌شود. سوء تعبیرهایی که نتایج به‌دست آمده از این براهین را به کلی مخدوش می‌کند. به طور خاص به نظر می‌رسد که نمی‌توان با تکیه بر قضیهٔ گودل نشان داد که هیچ ماشینی (بدون آن‌که آن را مقید به قید یا صفتی بکنیم) نمی‌تواند ذهن بشر را مدل کند. صد البته این به معنای این‌همانی ذهن و ماشین نیست؛ بلکه صرفاً به این معناست که نمی‌توان برای اثبات عدم این‌همانی ذهن و ماشین از قضیهٔ گودل به شیوه‌ای شبیه به شیوهٔ مثلاً لوکاس استفاده کرد. (۲) درعین‌حال، به‌نظر می‌رسد که می‌توان به درستی از این قضیه در جهت اقامهٔ برخی برهان‌ها علیه ادعاهای افراطی مادی‌گرایانه استفاده کرد. مثلاً می‌توان نشان داد که «هیچ ماشین جهانی صدق وجود نخواهد داشت». یا می‌توان مطابق بخشی از استدلال پنروز نشان داد که «ذهن را نمی‌توان با هیچ ماشینی که ذهن سازگاری آن را بداند، مدل کرد». هم‌چنین اگر بپذیریم که ذهن ما سازگاری خود را می‌داند، آن‌گاه بنابر بخش دیگری از استدلال اول پنروز «اگر ماشینی ذهن ما را مدل کند، ما نمی‌دانیم [و نخواهیم دانست] که آن ماشین ذهن ما را مدل می‌کند».

پی‌نوشت‌ها

۱- ابهامات و پیچیدگی‌هایی (نظیر سرهم‌نویسی و جدانویسی و غیره) در صورتِ نوشتاری الفبای فارسی و نحوهٔ قرار گرفتن حروف این الفبا در کنار هم وجود دارد که ممکن است ذهن خواننده را گمراه کند. به همین دلیل، ترجیح دادیم که از الفبای انگلیسی مثال بزنیم.

۲- در برخی از متون به‌جای عبارت تصمیم‌پذیر decidable از عبارات computably decidable یا computable استفاده شده است (Franzén, 2005, p. 115).

۳- به‌بیان دقیق‌تر هر قضیهٔ یک نظام صوری عضو انتهایی (دستِ‌کم) یک دنبالهٔ *متناهی* از جمله‌های زبانِ آن نظام صوری است؛ دنباله‌ای که هر عضو آن یا خود یکی از اصول موضوعهٔ این نظام صوری است و یا جمله‌ای است که با اعمال قواعد استنتاجِ آن نظام صوری بر روی برخی عضوهای قبلی آن دنباله به دست می‌آید. این دنباله یک *برهان* برای قضیهٔ مذکور نامیده می‌شود.

۴- مقالهٔ (Gödel, 2000) ترجمه‌ای از مقالهٔ (Gödel, 1931) است.

۵- به علاوه آنچه ما در اینجا آوردیم، تعمیم چیزی است که گودل در (Gödel, 1931) اثبات کرده است. برای دیدن طرح اثباتِ قضیهٔ گودل (حسام، ۱۳۷۶) را ببینید. در (ناگل و دیگران، ۱۳۶۴) اثبات کامل قضیهٔ گودل به زبانی غیر فنی آمده است و در فصل سوم (Mendelson, 1997) هم اثبات فنی و کاملاً ریاضیاتی این قضیه آمده است.

۶- از این به‌بعد به‌جای عبارت ماشین جهانی صدق (Universal Truth Machine) به اختصار از نماد "UTM" برای ارجاع به ماشین جهانی صدق استفاده می‌کنیم.

۷- نکتهٔ جالب توجه این‌که در قضایای ناتمامیت، گودل با استفاده از روشی به نام روش قطری‌سازی (diagonalization) جملهٔ گودلی آن نظامی که در شرایط قضیه صدق می‌کند را می‌سازد. یعنی ادعای قضیهٔ ناتمامیت صرفاً یک ادعای وجودی بدون مصداق نیست و برهان گودل حتی برای ساخت‌گرایان (constructivists) هم مقبول و معتبر است.

۸- این نکته که بخشی از استدلال روکر صحیح است در برخی نقدهایی که به استدلال روکر وارد شده است، مغفول مانده است. مثلاً فرانتزن کل استدلال روکر را زیر سؤال می‌برد و هیچ اشاره‌ای به بخش صحیح این استدلال ندارد (Franzén, 2005, pp. 115-6).

۹- چنان‌که پیش‌تر گفتیم خود قضیهٔ گودل هم یک قضیهٔ شرطی است: اگر نظام صوری S دارای شرایط قضیهٔ گودل باشد آن‌گاه در خودِ S ثابت می‌شود که «اگر S سازگار باشد، آن‌گاه جملهٔ حسابی صادقی در زبان این نظام صوری وجود دارد که در S اثبات‌ناپذیر است».

۱۰- این برهان بسیار شبیه به برهان قضیهٔ تعریف‌ناپذیری صدق تارسکی است.

۱۱- در ریاضیات معادلهٔ دیوفانتی به یک معادلهٔ یک یا چندمتغیرهٔ چندجمله‌ای که ضرایب آن اعداد صحیح هستند، گفته می‌شود. مثلاً $x^5+y^5+z^5=0$ یک معادلهٔ دیوفانتی است.

۱۲- در اینجا به اصل K در منطق معرفت (epistemic logic) متوسل شده‌ایم:

$K_A(P \supset Q) \supset (K_A P \supset K_A Q)$

۱۳- فرانتزن هم به این نکته در خصوص استدلال لوکاس اشاره کرده است (Franzén, 2005, p. 118).

۱۴- برای نمونه، رجوع کنید به نقدهای ففرمن (Feferman, 1996)، چالمرز (Chalmers, 1995) و پاتنم (Putnam, 1995).

۱۵- این طرح برگرفته از (Chalmers. 1995) است.

منابع

حسام، بردیا. (۱۳۷۶). «آشنایی با منطق ریاضی»، ***مجلهٔ ریاضی: مجلهٔ دانشجویان دانشکدهٔ علوم ریاضی دانشگاه صنعتی شریف***، ش ۶ صص ۲۶-۶

ناگل، ارنست، جیمز نیومن و آلفرد تارسکی. (۱۳۶۴). ***برهان گودل و حقیقت و برهان***، تهران: مولی، چاپ اول.

Chalmers, David. (1995). "Mind, Machine and Mathematics", ***PSYCHE***, 2(9), electronic journal.

Feferman, Solomon. (1996). "Penrose's Gödelian argument". ***PSYCHE***, 2(7), electronic journal.

Franzén, Torkel. (2005). ***GÖDEL'S THEOREM: AN INCOMPLETE GUIDE TO ITS USE AND ABUSE***, Wellesley, Massachusetts, A K Peters, Ltd.

Gödel, Kurt. (1931). "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I", ***MONATSHEFTE FÜR MATHEMATIK UND PHYSIK***, 38, pp. 173-98.

Gödel, Kurt. (2000). "***On formally undecidable propositions of Principia Mathematica and related systems I***", translation of (Gödel 1931) by Martin Hirzel, will be found in:
http://www.research.ibm.com/people/h/hirzel/papers/canon00-goedel.pdf/

Lucas, John Randolph. (1961). "Minds, Machines and Gödel", ***PHILOSOPHY***, XXXVI, pp. 112-127.

Lucas, John Randolph. (1996). "Minds, Machines and Gödel: A Retrospect", in P. J. R. Millican and A. Clark (Eds.), ***MACHINES AND THOUGHT: THE LEGACY OF ALAN TURING***, Vol. 1. Oxford University Press, pp. 103-104.

Mendelson, Elliott. (1997). ***AN INTRODUCTION TO MATHEMATICAL LOGIC***, 4th ed., Chapman & Hall.

Penrose, Roger. (1989). ***THE EMPEROR'S NEW MIND***, Oxford University Press.

Penrose, Roger. (1994). ***SHADOWS OF THE MIND***, Oxford University Press.

Penrose, Roger. (1996). "Beyond the Doubting of a Shadow", ***PSYCHE***, 2(23), electronic journal.

Putnam, Hilary. (1995). "Review of Shadows of the Mind, by Roger Penrose", ***BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY***, 32, pp. 370-373.

Rucker, Rudy. (2007). ***INFINITY AND THE MIND***, 2nd ed. Princeton University Press.

Turing, Alan. (1936). "On Computable Numbers, with an Application to the Entscheidungsproblem", ***PROCEEDINGS OF THE LONDON MATHEMATICAL SOCIETY***, Series 2. 42, pp 230–265.