

آیا هوش مصنوعی می‌تواند اینشتین دوم باشد؟

دانشمندان در حال بررسی این موضوع هستند که آیا مدل‌های زبانی آموزش دیده با داده‌های تاریخی می‌توانند به کشفیات خلاقانه‌ای مشابه دانشمندان بزرگ دست پیدا کنند یا خیر.



دانشمندان در حال بررسی این موضوع هستند که آیا مدل‌های زبانی آموزش دیده با داده‌های تاریخی می‌توانند به کشفیات خلاقانه‌ای مشابه دانشمندان بزرگ دست پیدا کنند یا خیر.

به گزارش ایسنا، در سال ۱۹۱۵، آلبرت اینشتین نظریه نسبیت عام خود را ارائه کرد و نگاه ما به ساختار جهان فیزیکی را دگرگون ساخت. این نظریه که گرانش را به عنوان تغییر شکل فضا-زمان در اثر جرم توضیح می‌دهد، یکی از نقاط اوج فیزیک مدرن است؛ نظریه‌ای که زیربنای کیهان‌شناسی، از پژوهش درباره سیاه چاله‌ها گرفته تا اندازه‌گیری امواج گرانشی، محسوب می‌شود و به طور معمول برای هدایت مأموریت‌های فضایی و ماهواره‌های جی‌پی‌اس مورد استفاده قرار می‌گیرد.

به نقل از نیچر، نسبیت عام همچنین به معیاری برای سنجش توانایی پیشگامان فناوری هوش مصنوعی تبدیل شده است؛ کسانی که می‌پرسند آیا ساخته‌های آنها روزی می‌توانند به پیشرفتی در چنین سطحی دست پیدا کنند یا خیر. در اجلاس هوش مصنوعی هند که فوریه امسال در دهلی نو برگزار شد، دیمیس هاسابیس، یکی از بنیان‌گذاران شرکت دیپ مایند گوگل در لندن، پیشنهاد کرد که اگر یک مدل زبانی بزرگ (LLM) با تمام دانسته‌های موجود تا یک تاریخ مشخص آموزش داده شود برای مثال تا سال ۱۹۱۱، می‌توان مشخص کرد آیا این مدل می‌تواند نظریه نسبیت عام را بازتولید کند یا خیر. هاسابیس با اشاره به مفهوم مبهم هوش مصنوعی عمومی (AGI) که هدف بسیاری از فعالان صنعت هوش مصنوعی است، گفت: این می‌تواند آزمون خوبی برای هوش مصنوعی عمومی باشد.

چنین آزمونی لزوماً نباید مشخصاً درباره نسبیت عام باشد. در ماه دسامبر سال ۲۰۲۴، اواین ایوانز، پژوهشگر سازمان غیرانتفاعی Truthful AI در برکلی کالیفرنیا، در سخنرانی درباره «مدل‌های زبانی قدیمی» یا «تاریخی» صحبت کرد؛ مدل‌هایی که فقط با داده‌های تاریخی تا یک تاریخ مشخص آموزش داده می‌شوند. ایوانز این پرسش را مطرح کرد که چنین مدل‌هایی چه چیزهایی را ممکن است دوباره کشف کنند.

امسال چندین گروه پژوهشی نمونه‌هایی از این مدل‌های تاریخی را ساخته‌اند که یکی از آنها در راستای آزمون پیشنهادی هاسابیس بوده است. با این حال، تلاش‌های اولیه آنها بیشتر درباره محدودیت‌های هوش مصنوعی کنونی اطلاعات ارائه می‌دهد تا توانایی‌های آن.

ایدو کامینر، متخصص اپتیک کوانتومی و یکی از نویسندگان مقاله پیش‌چاپی با عنوان «آیا هوش مصنوعی می‌تواند جا پای اینشتین بگذارد؟» که در ماه ژوئیه منتشر شده است، می‌گوید دستیابی هوش مصنوعی به یک پیشرفت شبیه نسبیت، ذاتاً غیرممکن نیست.

با این حال، او و همکارانش استدلال می‌کنند که برای وقوع چنین اتفاقی، باید برخی از اصولی که مدل‌های امروزی بر اساس آنها ساخته شده‌اند، بازنگری شود.

جهش‌های خلاقانه
هاسابیس سال گذشته در مصاحبه‌های رسانه‌ای به تشبیه «آزمون اینشتین» اشاره کرده بود، اما تام زهاوی، پژوهشگر دیگری در دیپ مایند گوگل، در مقاله‌ای که در ماه ژانویه در وب‌سایت شخصی خود منتشر کرد، این ایده را با جزئیات بیشتری شرح داد. عنوان این مقاله «مدل‌های زبانی بزرگ نمی‌توانند بپرند» است و بر ناتوانی فعلی مدل‌های زبانی بزرگ در انجام جهش‌های استدلالی مشابه جهش اینشتین تأکید دارد.

زهاوی نوشت، همان‌طور که فیلسوفان علم مدت‌ها است تشخیص داده‌اند، چنین پیشرفت‌هایی به استدلال استقرایی که از انباشت داده‌ها یا نمونه‌ها به یک قاعده کلی می‌رسد، نیاز ندارند؛ بلکه به استدلال ربایشی (Abductive reasoning) نیاز دارند که یک جهش خلاقانه است که برای یک پدیده منفرد، علتی را ابداع می‌کند.

این نوع استدلال نقطه قوت مدل‌های هوش مصنوعی کنونی نیست. این مدل‌ها ظاهراً بیشتر برای انجام بخش‌های سخت و پرزحمت علم مناسب هستند تا کشف‌های تحول‌آفرین. مدل‌ها با آموزش بر اساس نمونه‌های شناخته‌شده و سپس دریافت دستور از طریق پرامپت برای ارائه محتمل‌ترین پاسخ‌های آماری درباره موارد ناشناخته، به دنبال همبستگی‌ها در مجموعه داده‌های عظیم می‌گردند.

این ویژگی باعث می‌شود مدل‌ها ابزارهای پیش‌بینی ارزشمندی باشند و زمانی که داده‌های فراوانی وجود دارد، بتوانند کمک‌کننده باشند. اما جهش‌های تخیل یا چیزی که توماس کوهن، مورخ علم قرن بیستم، آن را «تغییر پارادایم» در درک علمی می‌نامید، اغلب زمانی رخ می‌دهند که داده‌های اندکی برای اتکا وجود دارد یا از ناهنجاری‌هایی ناشی می‌شوند که کاملاً با تصویر استاندارد موجود سازگار نیستند.

دانشمندان در چنین شرایطی اغلب یک «مدل جهان» ایجاد می‌کنند؛ یعنی استنباطی درباره اصول زیربنایی که بر اساس حجم اندکی از داده شکل می‌گیرد. برای مثال، مشاهدات دقیق حرکت سیارات توسط یوهانس کپلر، اخترشناس قرن هفدهم و روابط ریاضی که او از این مشاهدات استخراج کرد، باعث شد اسحاق نیوتن در پاسخ به نیروها، قانون گرانش و قوانین مکانیکی حرکت خود را توسعه دهد.

به زبان ساده‌تر مدل‌های هوش مصنوعی فعلی بیشتر در پیدا کردن الگوهای که قبلاً در داده‌ها وجود داشته‌اند، خوب هستند؛ اما مشکلیشان این است که نمی‌توانند به راحتی یک ایده کاملاً جدید را از چند مشاهده محدود به دست بیاورند. برای مثال استدلال استقرایی یعنی تعداد زیادی نمونه را ببینی و از آنها یک قانون کلی بسازی. مثلاً مشاهده کنی که هر بار

جسمی را رها می کنی، به سمت زمین می افتد و از این مشاهدات به یک قانون کلی درباره گرانش برسی. استدلال ربایشی یک مرحله خلاقانه تر است؛ یعنی از یک اتفاق یا مشاهده خاص، یک توضیح یا علت جدید دریافت کنی که قبلا کسی مطرح نکرده است.

مشکل اصلی اینجاست که کشف های بزرگ علمی همیشه با حجم زیادی از داده اتفاق نمی افتند. گاهی یک دانشمند فقط چند مشاهده عجیب یا یک مورد ناسازگار با نظریه های موجود می بیند و از همان، یک ایده کاملا جدید پیدا می کند. برای مثال کپلر با بررسی دقیق حرکت سیارات، الگوهایی را پیدا کرد که قبلا توضیح کاملی برایشان وجود نداشت. نیوتن بعدا توانست از این روابط و مشاهدات، یک تصویر کلی تر از جهان بسازد و قانون گرانش و قوانین حرکت را ارائه کند. کاری که به نظر می رسد مدل های هوش مصنوعی چندان در آن خوب نیستند.

اما آیا مدل های هوش مصنوعی می توانند به شیوه ای مشابه، از داده های اندک به مدل های کلی از جهان برسند؟ سندیل مولایناتان، دانشمند علوم رایانه در موسسه فناوری ماساچوست (MIT) در کمبریج، می گوید در حال حاضر چنین نیست. او و همکارانش در مقاله ای که در ماه ژوئیه منتشر شد، یک مدل پایه مکانیک مداری را با داده های مصنوعی مربوط به سامانه های مختلف سیاره ای که از قوانین مکانیک نیوتنی پیروی می کردند، آموزش دادند.

آنها دریافتند که مدل هرگز قانون واقعی گرانش را استنباط نکرد؛ قانونی که نیروی میان دو جسم را به جرم آنها و فاصله میانشان مرتبط می کند. در عوض، مدل برای هر سامانه سیاره ای یک قانون متفاوت را استنباط کرد که در هر مورد به شکلی منحصر به فرد اشتباه بود.

با این حال، مدل های زبانی بزرگ در ریاضیات پیشرفت های چشمگیری داشته اند؛ پیشرفت هایی که نشان می دهد آنها گاهی می توانند ساختارهای منطقی زیربنایی داده های آموزشی خود را درک کنند. در ماه مه امسال، یک چت بات هوش مصنوعی که توسط شرکت اوپن ای آی ساخته شده بود، یک حدس ۸۰ ساله متعلق به ریاضی دان مجارستانی پال اردوش را رد کرد. مولایناتان این دستاورد را «یک پیشرفت مفهومی واقعی» می داند.

او می گوید هوش مصنوعی صرفا با جست و جوی کورکورانه و نیروی محاسباتی به پاسخ نرسید؛ راه حل در طول مسیر به انتزاع های جدید و کوچکی نیاز داشت. در عین حال، این رد نظریه ایده هایی را که از قبل در مقالات ریاضی وجود داشتند با یکدیگر ترکیب کرد؛ بنابراین، از این نظر، این دستاورد مانند یک کشف اینشتین وار نبود که گویی از هیچ به وجود آمده باشد.

یک لحظه در زمان
پس از آنکه هاسابیس آزمون سال ۱۹۱۱ را مطرح کرد، مایکل هلا، پژوهشگر مستقل هوش مصنوعی در سان فرانسیسکو، تلاش مشابهی انجام داد. هلا در ماه مارس در وب سایت خود نوشت که یک مدل زبانی بزرگ به نام Machina Mirabilis را با داده های مربوط به پیش از سال ۱۹۰۰ آموزش داده است تا ببیند آیا می تواند مکانیک کوانتومی، نظریه نسبیت خاص اینشتین در سال ۱۹۰۵ و نظریه نسبیت عام را تولید کند یا خیر.

هلا در این مسیر به مدل سرنخ هایی می داد. در یک مورد، آن را با مشاهدات مربوط به اثر فوتوالکتریک که پدیده ای است که در آن نور باعث جدا شدن الکترون ها از صفحات فلزی می شود، راهنمایی کرد؛ پدیده ای که بعدها اینشتین آن را با کوانتیده بودن نور توضیح داد. در موردی دیگر، او ماهیت آزمایش فکری آسانسور را در اختیار مدل قرار داد؛ آزمایشی که بعدها به نظریه نسبیت عام اینشتین منجر شد.

هلا می گوید مدل در برخی موارد «نشانه هایی از شهود» نشان داد. برای مثال، هنگامی که اثر فوتوالکتریک به آن ارائه شد، مدل گفت نور «به انبوهی از تکانه های مجزا تجزیه می شود»؛ عبارتی که به ایده کوانتوم های نور اشاره داشت. اما مدل زبانی بزرگ در بیشتر موارد شکست خورد، درک واقعی از فیزیک مورد استفاده خود نداشت و گاهی کلماتی را تکرار می کرد که ظاهرا منطقی به نظر می رسیدند، اما به گفته هلا، ظاهرا هیچ نوع بازنمایی درونی قدرتمندی از جهان نداشت که بتواند بر اساس آن استدلال کند.

هلا همچنین دریافت که آموزش یک مدل زبانی بزرگ با محدودیت تاریخی بسیار دشوار است، زیرا داده ها مملو از نسخه های انگلیسی ناقص و خطاها و آثار ناشی از دیجیتالی کردن متون چاپی هستند. نیک لوین، دانشمند مستقل علوم رایانه و همکارانش نیز در پژوهشی که در ماه آوریل به صورت آنلاین ارائه کردند، به همین نتیجه رسیدند. آنها تلاش کردند یک مدل هوش مصنوعی تاریخی بسازند که فقط از اطلاعات شناخته شده تا سال ۱۹۳۰ استفاده کند؛ سالی که انتخاب شده بود زیرا آثاری که در آن سال منتشر شده اند، در آغاز سال ۲۰۲۶ وارد مالکیت عمومی در ایالات متحده شدند.

لوین که در سان فرانسیسکو مستقر است، می گوید با چنین مدلی، در اصل می توان پرسش هایی درباره موضوعاتی که در دهه ۱۹۳۰ توسعه یافتند مطرح کرد؛ برای مثال ماشین های تورینگ، که مفهومی اساسی در نظریه محاسبات هستند، قضیه ناتمامیت گودل در مبانی ریاضیات یا ذراتی به نام نوترینوها که نخستین بار در سال ۱۹۳۴ در یک مقاله علمی مطرح شدند. لوین و همکارانش دریافتند که ساختن یک مدل تاریخی که فقط با اطلاعات موجود تا سال ۱۹۳۰ آموزش دیده باشد، دشوار است؛ زیرا مواد آموزشی به شدت «نشت» دارند. لوین می گوید: اگر درباره اتفاقی که در دهه ۱۹۵۰ رخ داده از آن سؤال کنید، اغلب به طور تصادفی پاسخ می دهد و اغلب نیز پاسخ درست است.

برای مثال، مدل ظاهرا پیش از ۱۹۳۰ به پرسش هایی درباره دولت فرانکلین دی. روزولت پاسخ داد؛ فردی که از سال ۱۹۳۳ تا ۱۹۴۵ رئیس جمهور آمریکا بود. زمانی که مجموعه داده ها تاریخ گذاری دقیق یا صحیحی نداشته باشند، حذف اطلاعات مربوط به دوره پس از یک تاریخ مشخص دشوار است.

با وجود این، لوین که مدتی به عنوان پیش بینی کننده در حوزه اقتصاد کار کرده است، معتقد است در نهایت باید بتوان از این «ذهن ۱۹۳۰» خواست پیش بینی هایی انجام دهد؛ پیش بینی هایی که می توانند شامل چیزهایی باشند که در یک بازار پیش بینی مشاهده می کنید.

او می گوید مدل های هوش مصنوعی پیش بینی محور می توانند برای پیش بینی کشفیات علمی نیز به کار گرفته شوند. او انتظار دارد آزمایش کند که آیا مدلی که با داده های موجود تا مثلا ژانویه امسال آموزش دیده است، می تواند به کشفی دست پیدا کند که اکنون می دانیم در ماه ژوئن رخ داده است یا خیر. او پیش بینی می کند که چنین مدلی یک کشف قابل توجه انجام

خواهد داد، فقط به دلیل مقیاس داده ها و منابعی که در اختیار آن قرار می گیرد. پژوهشگران دانشگاه زوریخ در سوئیس نیز خانواده کاملی از مدل های تاریخی را در قالب پروژه ای به نام Ranke-۴B ایجاد کرده اند. این مدل ها با متون دارای تاریخ مشخص و با نقاط برش تاریخی معنادار در سال های ۱۹۱۳، ۱۹۲۹، ۱۹۳۳، ۱۹۳۹ و ۱۹۴۶ آموزش داده شده اند.

دانیل گوتلیش، یکی از اعضای این تیم و اقتصاددانی که اکنون در مؤسسه فناوری فدرال سوئیس (ETH) در زوریخ فعالیت می کند، می گوید داده های تاریخی یا منابع محاسباتی کافی برای ساخت مدل هایی به قدرت مدل های زبانی بزرگ مدرن وجود ندارد. بنابراین، او امیدوار است تنها بررسی کند که آیا یک مدل زبانی تاریخی می تواند ایده هایی با نشانه هایی از پیشرفت های آینده تولید کند یا خیر.

نظریه های پیش از اندازه زیاد
جیکوب آندریاس، دانشمند علوم رایانه در موسسه فناوری ماساچوست، معتقد است هیچ مانع آشکاری وجود ندارد که یک مدل زبانی هوش مصنوعی نتواند مثلا نظریه نسبیت عام را به عنوان یکی از خروجی های احتمالی در میان تعداد زیادی نظریه کاملا نادرست درباره مجموعه ای از داده ها تولید کند.

مشکل این است که تشخیص اینکه کدام نظریه درست است، یا دست کم ارزش آزمایش کردن دارد، دشوار است. برای مثال، این همان مشکلی بود که درباره «قوانین گرانش» مختلف و اشتباهی که مدل مکانیک مداری مولایناتان و همکارانش تولید کرده بود، وجود داشت.

در ریاضیات، در مقابل، می توان هر گام از روند کار مدل زبانی بزرگ را فوراً به عنوان درست یا نادرست بررسی کرد. این موضوع نشان دهنده تفاوت میان شیوه ای است که مدل های زبانی بزرگ و انسان ها برای ساخت نظریه به کار می گیرند. انسان ها به ندرت مجموعه ای از نظریه های احتمالی را به صورت آماری تولید می کنند و سپس بررسی می کنند که کدام یک، در صورت وجود، درست است.

برای مثال، آروین شرودینگر، فیزیکدان، یک قرن پیش معادله خود را برای مکانیک موجی کوانتومی نوشت؛ اما حتی امروز نیز پژوهشگران همچنان درباره روند فکری او در هنگام رسیدن به این معادله تحقیق می کنند.

آندریاس می گوید: تقریباً به طور قطع، روند ذهنی او شبیه این نبود که به صورت تصادفی مجموعه ای از گزاره های شبیه نظریه تولید کند و سپس بررسی کند کدام یک با مشاهدات واقعی سازگار است. مشکل دیگری نیز وجود دارد: اینکه اصلاً تصمیم بگیریم یک ایده ارزش دنبال کردن دارد یا خیر. آندریاس می گوید مدل های زبانی بزرگ، برای مثال، درک خوبی از اینکه یک قضیه ریاضی جالب چیست، ندارند. او اضافه می کند که توانایی مدل ها برای تشخیص اینکه چه چیزی مهم است، در واقع همان جایی است که آنها بیشترین فاصله را با متخصصان انسانی دارند.

در جستجوی یک پارادایم جدید
تغییرات تحول آفرین در درک علمی اغلب از توجه به اختلاف های کوچک و موارد پرت ناشی می شوند؛ دقیقاً همان نوع مواردی که مدل های هوش مصنوعی تمایل دارند آنها را میانگین گیری و محو کنند.

کامینر و همکارانش می گویند: کشف علمی انسان ها در طول تاریخ اغلب به توانایی نوعی «وسواس» مداوم متکی بوده است؛ سبکی شناختی که به یک پژوهشگر اجازه می دهد سال ها به یک ایده نامعقول و نامتعارف پایبند بماند. برای مثال، ماکس پلانک متوجه شد محاسباتش درباره تابش جسم سیاه دقیقاً با نتایج آزمایش ها جور در نمی آید. او این اختلاف را نادیده نگرفت و تلاش کرد دلیل آن را پیدا کند. این پیگیری در سال ۱۹۰۰ باعث شد به ایده مهمی برسد: انرژی برخلاف تصور قبلی، می تواند در بسته های کوچک و جداگانه، یعنی «کوانتوم ها» منتقل شود.

برای نظریه نسبیت خاص هم اینشتین لزوماً به دنبال حل یک مشکل آزمایشگاهی درباره «اثر» نبود. در آن زمان تصور می شد نور برای حرکت به محیطی به نام «اثر» نیاز دارد، اما آزمایش ها نتوانستند وجود این محیط را نشان دهند. برخلاف تصور رایج، اینشتین چندان درگیر این نتیجه آزمایش ها نبود و نسبیت خاص را صرفاً برای حل این مسئله مطرح نکرد.

به طور کلی، پیشرفت های علمی به ندرت صرفاً از یک مجموعه مشخص از داده های تجربی حاصل می شوند. آنها از تعامل مداوم و ظریف میان نظریه ها با مفاهیم، آزمایش ها و روش های آزمایشی به وجود می آیند.

این موضوع به ویژه در رشته هایی مانند زیست شناسی مولکولی صدق می کند؛ رشته ای که برخلاف بسیاری از شاخه های فیزیک، تحت کنترل مجموعه کوچکی از قوانین بنیادی و کاملاً تعریف شده قرار ندارد. با در نظر گرفتن این موضوع، برخی پژوهشگران در حال توسعه سامانه های خودکاری هستند که می توانند آزمایش ها را انتخاب و برنامه ریزی کنند، آنها را به صورت رباتیک انجام دهند و بر اساس نتایج، فرضیه هایی را تدوین کنند. با این حال، حتی در اینجا نیز پیشرفت ها گاهی از تغییر در طرز فکر ناشی می شوند، نه از داده های تازه و تعیین کننده. برای مثال، تصور می شد پروتئین های ذاتاً بی نظم بیش از آن ناپایدار هستند که بتوان آنها را برای تحلیل با پرتو ایکس بلوری کرد؛ تا اینکه پژوهشگران متوجه شدند این پروتئین ها اساساً ساختار پایداری ندارند و به این نتیجه رسیدند که بی نظمی یک ویژگی است، نه یک نقص.

همچنین نمونه هایی از آنچه امروزه تراکم های زیست مولکولی نامیده می شوند که خوشه های قطره مانند درون سلول ها هستند، برای چندین دهه مشاهده شده بودند، اما کسی به آنها توجه خاصی نکرده بود تا اینکه ایده مفهومی مربوط به آنها شکل گرفت.

آیا هوش مصنوعی می تواند روزی چنین جهش هایی در ایمان، استدلال یا ادراک را تقلید کند؟
آندریاس می گوید این چالش احتمالاً بیشتر شبیه یافتن راهی برای خودکارسازی این نوع جهش های ناخودآگاه و وصف ناپذیر شهودی است تا یک فرایند جستجوی کلاسیک که از قبل می دانیم، چگونه آن را پیاده سازی کنیم.

مولایناتان می گوید برای چنین کاری، مدل های زبانی بزرگ به تنهایی ابزار مناسبی نیستند. به باور او، ما به مدل هایی نیاز داریم که نه تنها بر اساس مجموعه دانش موجود، بلکه با داده های دنیای واقعی درباره اینکه جهان فیزیکی چگونه است و همچنین بر اساس «فضای ریاضی» آموزش دیده باشند؛ یعنی اینکه چگونه باید اشیای ریاضی را دستکاری کرد.

سامانه های هوش مصنوعی برای کشف علمی همچنین به اهداف عملکردی جدیدی نیاز خواهند داشت؛ اهدافی فراتر از صرفاً ارائه پیش بینی های آماری خوب.

کامینر می گوید کشف هایی در مقیاس نسبیت تنها زمانی از این سامانه ها حاصل خواهند شد که آنها را برای چیزی فراتر از دقت پیش بینی بهینه کنیم و در کنار آن به سادگی، دامنه توضیح دهنده و موارد دیگر نیز پاداش دهیم.

در حال حاضر، ظاهراً انگیزه چندانی برای ساخت مدل های هوش مصنوعی که شبیه انسان ها استدلال کنند، وجود ندارد. این همان چیزی است که ریچ ساتن، دانشمند علوم رایانه در دانشگاه آلبرتا در ادمونتون کانادا، آن را «درس تلخ هوش مصنوعی» نامیده است.

آنچه معمولاً در عملکرد الگوریتم تفاوت ایجاد می کند، قدرت محاسباتی بیشتر و داده های آموزشی بیشتر است. ارزش ادراک شده چنین مدل های هوش مصنوعی معمولاً به این بستگی دارد که آیا آنها پیش بینی های خوب، سریع و دقیقی ارائه می دهند یا خیر؛ بدون اینکه لزوماً چیزی درباره فرایند فیزیکی زیربنای یک پدیده بگویند.

برای مثال، ابزار آلفا فولد متعلق به دیپ مایند گوگل برای ساختار پروتئین، مسئله درک نحوه تاخوردگی پروتئین ها را حل نمی کند؛ بلکه با پیش بینی شکل تاخورده آنها، این مسئله را دور می زند.

آندریاس گمان می کند که دست کم در ابتدا، مدل های زبانی بزرگ تنها به کشف های بسیار محدودی دست پیدا خواهند کرد؛ آن هم در مواردی که از قبل می دانیم مسئله مهم چیست و پاسخ در قرض گرفتن یک راه حل یا کشیدن یک قیاس وجود دارد که از قبل در حوزه دیگری وجود داشته است.

او می گوید در چنین شرایطی، پژوهشگران انسانی ممکن است با خودشان بگویند: ای کاش در این حوزه هم تخصص داشتم تا می توانستم این ارتباط را پیدا کنم. اما درباره نظریه نسبیت عام، چنین واکنشی وجود نداشت. آندریاس می گوید: آنها فقط با یک سوال روبه رو شدند: این ایده از کجا آمده است؟