شاهکار هوش مصنوعی: حل مسئله ریاضی ۳۰ ساله تنها در ۶ ساعت* حساب کاربری عمومی در وی‌چت

ترجمه و تنظیم: سمیه رهنما و سعید علیخانی**

ترنس تائو: چت‌جی‌پی‌تی و مدل‌های دیگر شکست خوردند

هیچ دخالت انسانی وجود ندارد.

شب گذشته، جامعه ریاضی شوکه شد! هوش مصنوعیِ ریاضی‌دان به نام «ارسطو»[۱] نسخه ساده شده‌ای از یک مسئله ۳۰ ساله را تنها در ۶ ساعت و با یک کلیک حل کرد و تحسین فراوانی از ترنس تائو برنده مدال فیلدز، دریافت نمود. عصر اثبات الهامی[۲] در حوزه ریاضیات فرا رسیده است.

یک مسئله باز ریاضی که به مدت سی سال حل‌نشده باقی مانده بود، سرانجام حل شد.

سامانه هوش مصنوعی «ارسطو»، که توسط هارمونیک‌مث توسعه یافته، به‌طور کاملاً مستقل مسئله شماره ۱۲۴ اردوش[۳] را حل کرد. این دستاورد، که در آن هیچ مداخله و کمک انسانی صورت نگرفت، تنها در ۶ ساعت انجام شد و بررسی آن در سیستم اثبات (Lean) فقط یک دقیقه طول کشید، به‌عنوان «لحظه فرود بر ماه» در جامعه ریاضیات تلقی می‌شود.

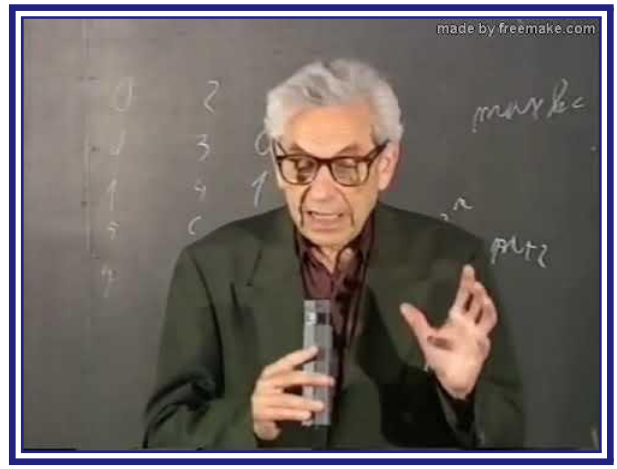
ولاد تِنِف[۴] بنیانگذار HarmonicMath، با تحسین اظهار داشت که جامعه‌ی ریاضیات در حال پشت سر گذاشتن تحولی بزرگ است. اکنون دورانِ اثباتِ الهامی فرا رسیده است!

حتی ترنس تائو، برنده مدال فیلدز، هوش مصنوعیِ ریاضی‌دان «ارسطو» را بسیار تمجید کرد.

فهرست مسائل پال اردوش، همچون قله‌ی اورست در جهان ریاضیات، نمادی چالش‌برانگیز از مرزهای دانش بشری است. پاداش حل این مسائل، که عمدتاً از ده‌ها تا هزاران دلار متغیر است، در نگاه اول ناچیز می‌نماید. با این حال، ارزش و اعتبار معنوی این دستاورد برای ریاضی‌دانان آن‌قدر بالاست که گویی مدال افتخاری بی‌همتا را دریافت کرده‌اند.

---

[۱] Aristotle
[۲] Vibe proving
[۳] Erdős Problem # 124
[۴] Vlad Tenev



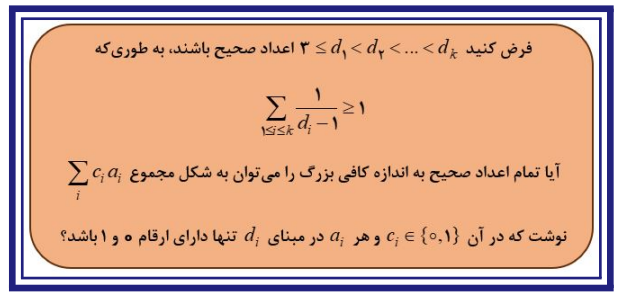
مسئله شماره ۱۲۴ اردوش، که در مقاله‌ای با عنوان Com-plete sequences of sets of integer powers «دنباله‌های کامل مجموعه‌هایی از توان‌های صحیح» در مجله Acta Arithmetica مطرح شده بود، سی سال بود که حل نشده باقی مانده بود. هسته مسئله شماره ۱۲۴ به شرح زیر است:

فرض کنید k عدد طبیعی $d_i \geq ۲$ برای $i = ۱, ۲, \ldots, k$ داده شده‌اند. اگر شرط

$$\sum_{i=۱}^{k} \frac{۱}{(d_i - ۱)} \geq ۱,$$

برقرار باشد، آنگاه برای هر عدد طبیعی n، همواره ضرایب صحیح نامنفی a_i (به‌ازای هر i) وجود دارند به‌طوری‌که: $n = \sum_{i=۱}^{k} a_i$.

علاوه بر این، هر یک از a_iها در سامانه‌ی نمایش اعداد (ارقام) بر مبنای d_i، فقط می‌توانند مقادیری از مجموعه $\{۰, ۱\}$ باشند. به بیان دقیق‌تر، برای هر i، عدد a_i در بسط بر مبنای d_i فقط از ارقام ۰ و ۱ تشکیل شده است.

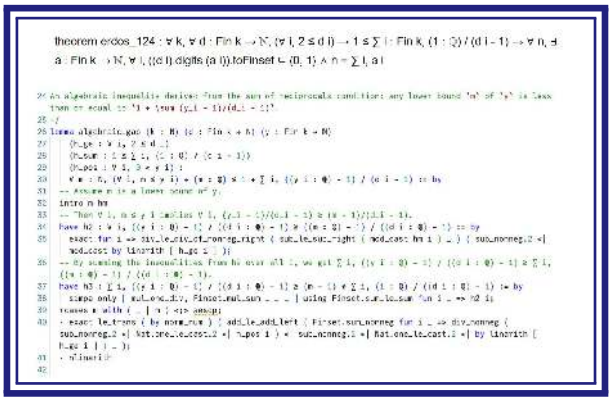


به‌عبارت ساده، این مسئله به‌طور اساسی می‌پرسد: آیا تحت محدودیت‌های شدید، هر عدد به اندازه کافی بزرگی می‌تواند همیشه به‌صورت «باینری» و بدون تأثیر گرفتن از پایه (مبنا) نمایش داده شود؟ این موضوع به لایه‌های عمیق «ریاضیات ترکیبیاتی» مربوط می‌شود و روش‌های سنتی درگیر شرط بزرگ‌ترین مقسوم علیه مشترک (gcd) و حالت‌های مرزی باقی مانده و از آن فراتر نرفته بودند تا دیشب که این دیوار فرو ریخت.

نحوه حل و اثبات ارسطو

تیم هارمونیک یک نمونه اولیه از «ابر-هوش ریاضی» به نام «ارسطو» ایجاد کرده‌اند. این مدل ترکیبی از یادگیری تقویتی، جستجوی درخت مونت کارلو و زبان صوری Lean را به کار می‌گیرد. پس از وارد کردن صورت مسئله، این سیستم صدها میلیون راهبرد اثبات را جست‌جو کرد و در نهایت یک قضیه ۱۰۰٪ قابل تأیید را ارائه داد.

ریاضی‌دان، بوریس الکسیف[۵] اظهار داشت که این قضیه، مورد علاقهٔ او در میان سه قضیه‌ای است که هوش مصنوعی تولید کرده است.



مترجم:در این عکس منظور از $\mathrm{Fin}\ k$ که مخفف کلمه متناهی است، مجموعه k عضوی است.

لازم به یادآوری است که مسئله ۱۲۴ در واقع در دو صورت‌بندی متمایز توسط پال اردوش مطرح شده است. آنچه هوش مصنوعی ارسطو اخیراً به آن دست یافته، حل صورت ساده‌شده (و از حیث ساختاری متعادل‌تر) این مسئله است، درحالی‌که صورت‌بندی کامل‌تر

[۵] Boris Alexeev

و عمومی‌تر آن، کماکان به عنوان یک چالش باز در نظریه‌ی اعداد و ترکیبیات باقی مانده است.

اگرچه زمان حل مسئله توسط ارسطو بالغ بر ۶ ساعت شد و سیستم Lean همان برهان را در کمتر از ۱ دقیقه تأیید کرد، مدیر وب‌سایت مسائل اردوش، اعلام کرد که عملکرد ارسطو تحسین‌برانگیزترین دستاورد در این زمینه بوده است!

شکست مدل‌های بزرگ زبان و نقص فنی توصیف مسئله

هم ChatGPT و هم Gemini در حل این مسئله شکست خوردند. ترنس تائو در این باره اظهار داشت: «تا آنجا که من می‌دانم، ابزارهای تحقیقاتی عمیق جمینای و چت‌جی‌پی‌تی هیچ مقالهٔ جدید و ارزشمندی در مورد این مسئله نیافته‌اند.»

جمینی یک مشاهدهٔ ساده ارائه داد: اگر عدد ۱ از مجموعه مبنا حذف شود، آنگاه شرط ب.م.م (بزرگترین مقسوم‌علیه مشترک) برای تحقق تساوی به یک شرط لازم تبدیل می‌شود؛ این سامانه سپس شرط $\sum_i \frac{1}{(d_i-1)} \geq 1$ را توضیح داده و آن را به تحقیقاتی موازی در حوزه‌ی مجموعه‌های کانتور به ویژه لم «شکاف نیوهاوس»[۶] مرتبط ساخت. با این حال، نتوانست مقالات جدیدی را که مستقیماً به این مسئله پرداخته باشند، بیابد.

از سوی دیگر، چت‌جی‌پی‌تی عمدتاً بر همان صفحه‌ی وب اولیه به‌عنوان منبع اصلی و معتبر تکیه کرد و به ارائه‌ی خلاصه‌ای از برهان ارسطو، سایر مقالات ارجاع‌شده در آن صفحه و نیز صفحات مسائل مرتبط بسنده کرد. نتیجه آن که اطلاعات جدید مستقلی به‌دست نیامد، اما به‌طور حتم، خلاصه‌ها و ارتباط‌های تولیدشده توسط این هوش‌های مصنوعی برای خوانندگان جالب توجه خواهد بود.

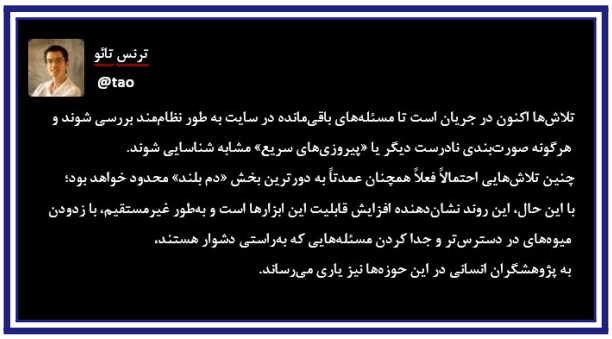
در مواجهه با این پرسش، ابزارهای تحقیقاتی عمیق «جمینای» (Gemini) و «چت‌جی‌پی‌تی» (ChatGPT) هم نتوانستند منابع یا مقاله‌های جدید ارزشمندی درباره‌ی این مسئله پیدا کنند.
جمینای تنها به یک نکته‌ی ساده اشاره کرد: اگر عدد ۱ از مجموعه مبنا حذف شود، آنگاه شرط ب.م.م

$$\sum_i \frac{1}{d_i - 1} \geq 1$$

(یعنی بزرگترین مقسوم‌علیه‌مشترک) ضروری می‌شود. این سامانه همچنین اهمیت شرط ۱≤ و ارتباط آن را به تحقیقات انجام‌شده در حوزه مجموعه‌های کانتور، به ویژه «لم شکاف نیوهاوس» توضیح داد. اما هیچ منبع مستقیم جدیدی برای خود مسئله ارائه نکرد.
از سوی دیگر چت‌جی‌پی‌تی عمده تکیه‌اش بر همان صفحه وب اولیه به عنوان منبع اصلی و معتبر بود و به راه‌حل «ارسطو» و سایر مقالات ارجاع شده در آن صفحه، و همچنین صفحات مسائل مرتبط بسنده کرد. بنابراین اطلاعات جدید و مستقلی به‌دست نیامد، اما شاید خلاصه‌ای که هوش مصنوعی از کل ماجرا ارائه داده برای خوانندگان جالب باشد.
ترنس تائو [Terence Tao]، ۱۶:۲۲ — ۳۰ نوامبر ۲۰۲۵

هوش مصنوعی در حال برداشت «میوه‌های در دسترس ریاضیات» است.

ترنس تائو معتقد است: هوش مصنوعی در حال برداشت «میوه‌های در دسترس»[۷] در ریاضیات است. ترنس تائو در Mathstodon با اشاره به سال‌ها تجربه خود، واقعیت کنونی را این‌گونه تشریح کرد: وضعیت حاکم بر مسائل حل‌نشده ریاضی از یک «توزیع دم‌بلند»[۸] پیروی می‌کند، و اتوماسیون مبتنی بر هوش مصنوعی عمدتاً در انتهای بسیار کوتاه و در دسترس این دم‌بلند متمرکز شده است.



تعداد زیادی از مسائل وجود دارند که در واقع اثبات یا رد آنها نسبتاً آسان است، اما به‌دلیل محدودیت تعداد ریاضی‌دانان متخصصی که عملاً می‌توانند وقت خود را به تحقیق اختصاص دهند، این مسائل کمتر مورد توجه قرار گرفته‌اند. به بیان دیگر، درون این «دم» دراز، در واقع انبوهی از «میوه‌های در دسترس» پنهان شده‌اند که هوش مصنوعی اکنون در آستانه چیدن آنهاست.

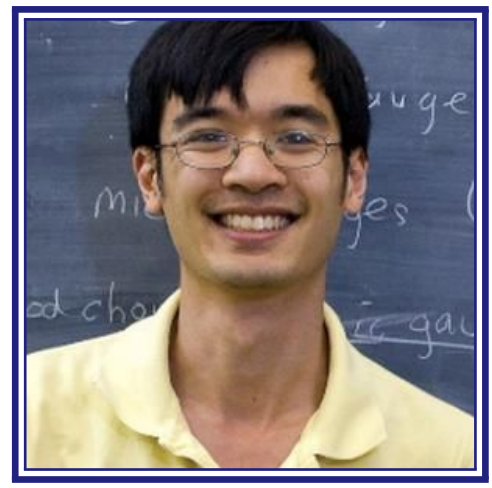
بی‌تردید، اگر راهی برای حل خودکار این مسائل در مقیاس انبوه یافت شود، احتمالاً حجم قابل‌توجهی از نتایج ریاضی جدید تولید

---

[۶] Newhouse Gap Lemma
[۷] low-hanging fruits
[۸] long-tail distribution

خواهد شد.

«پروژه‌ی نظریه‌های معادلاتی» سال گذشته برای ترنس تائو نمونه‌ای عینی از این اصل بود. در این پروژه، آنان با حجمی عظیم متشکل از ۲۲ میلیون امکان استلزام در حوزه جبر جهانی روبرو بودند. فرآیندی که حل دستی آن توسط انسان، بدون شک به زمانی بسیار طولانی نیاز داشت. با به‌کارگیری یک روش خودکارسازی نظام‌مند و نسبتاً ساده موفق شدند بخش عمده‌ای از این احتمالات را تنها در چند روز پیمایش و تحلیل کنند.



سپس، آن‌ها از روش‌های پیچیده‌تر برای مقابله با دشواری‌های سرسختی که در دورهای قبلی حل‌نشده باقی مانده بودند، استفاده کردند. در نهایت، حلِ تعداد اندکی از مسائل به‌خصوص سختِ باقیمانده، چندین ماه زمان از ریاضی‌دانان انسانی گرفت.

وضعیت کنونی پایگاه مسائل اردوش: در حال حاضر، پایگاه اینترنتی مسائل اردوش ۱۱۰۸ مسئله را ثبت کرده است که هر کدام حداقل در یکی از مقالات اردوش ظاهر شده‌اند. این مسائل طیفی را تشکیل می‌دهند:

- مسائل مشهور دشوار: مانند مسئله شناخته‌شده E3.
- انبوهی از مسائل گمنام: تعداد زیادی مسئله کم‌شناخته‌تر و کم‌توجه‌دیده‌شده نیز وجود دارد که حتی خود اردوش نیز هرگز به سراغ بازبینی آن‌ها نرفت.

هوش مصنوعی و «معماهای ازپیش‌حل‌شده»: در هفته‌های اخیر، نزدیک به ده برچسب «حل‌نشده» روی این پایگاه ناگهان حذف شدند. همه این موارد از طریق جستجوی ادبیات با کمک هوش مصنوعی کشف شدند. در واقع، این مسائل پیش‌تر توسط دیگران حل شده بودند، اما راه‌حل آن‌ها در جامعه ریاضیات یا این فهرست خاص ناشناخته مانده بود.

ریاضی‌دانانی که روی این مسائل تحقیق می‌کنند، هم‌اکنون به صورت عملی از ترکیب ابزارهای هوش مصنوعی و دستیاران اثبات صوری استفاده می‌کنند.

برخی از آنان اثبات‌های موجود را در Lean تأیید می‌کنند، برخی تولید عبارت‌های مرتبط با این مسائل در دنباله‌های اعداد صحیح را بر عهده می‌گیرند و برخی دیگر به پر کردن گام‌های اثبات مفقوده در یک ایده از پیش موجود می‌پردازند. اخیراً نوع دیگری از «میوه‌های در دسترس» در محدوده ابزارهای خودکار کشف شده است. مسائلی که به دلیل ایرادات فنی در بیان، به‌طور غیرمنتظره‌ای حلشان آسان شده است.

مسئله ۱۲۴ اردوش نمونهٔ شاخصی است. نسخه‌ی کامل این مسئله تا حدی دشوار است و در سه مقاله از اردوش ظاهر شده است. با این حال، یک فرض کلیدی در دو مورد از آنها حذف شده بود، و همین امر باعث شد که این نسخه در واقع نتیجه مستقیمی از معیار براون باشد. هیچ‌کس متوجه این موضوع نشد تا زمانی که بوریس آلکسیف مسئله را به ارسطو سپرد. به‌طور غیرمنتظره، هوش مصنوعی به‌صورت مستقل در عرض چند ساعت خلأ موجود را یافت و اثبات صوری را در سامانهٔ Lean تکمیل کرد. مشاهده می‌شود که هوش مصنوعی در حال روشن کردن «جنگل تاریک» ریاضیات است. همان‌طور که ترنس تائو گفته است: «ابزارهای خودکار ابتدا ساده‌ترین مسائل را پاکسازی می‌کنند، بخش واقعاً دشوار را جدا می‌سازند و به ریاضی‌دانان انسانی اجازه می‌دهند انرژی خود را بر چیزهای ارزشمند متمرکز کنند.»

* این مقاله از حساب کاربری عمومی «New Intelligence Yuan» در وی‌چت است. نویسنده آن New Intelligence Yuan و ویراستار آن Taozi است. این مطلب در این لینک منتشر شده است.

[1] https://www.erdosproblems.com/forum/thread/124post-1892

[2] https://x.com/SebastienBubeck/status/1994946303546331508?s=20

[3] https://mathstodon.xyz/@tao/115639983683442577

[4] https://x.com/thomasfbloom/status/1995094668879462466?s=2036Kr

** دانشگاه یزد