

اثبات به کمک ماشین

ترنس تائو*

ترجمه: سعید علیخانی و سپهر ابراهیمی مود**

محاسبات ممیز شناور قابل توجه بود. اما سیستم‌های جبر رایانه‌ای مدرن (به عنوان مثال، مگما، سیج-مت، متمتیکا، میل^۱ و غیره) و همچنین زبان‌های برنامه‌نویسی با هدف کلی‌تر، می‌توانند فراتر از «شماره‌گیری» سنتی عمل کنند. اکنون آنها به طور معمول برای انجام محاسبات نمادین در جبر، آنالیز، هندسه، نظریه اعداد و بسیاری از شاخه‌های دیگر ریاضی استفاده می‌شوند. برخی از اشکال محاسبات علمی به دلیل خطاهای گرد و ناپایداری‌ها به طور معروفی غیرقابل اعتماد هستند، اما اغلب می‌توان این روش‌ها را با جایگزین‌های دقیق‌تر (برای مثال، جایگزینی محاسبات ممیز شناور با محاسبات فاصله‌ای) جایگزین کرد، که احتمالاً به قیمت افزایش زمان اجرا یا استفاده از حافظه انجام می‌شود. بستگان سیستم‌های جبر رایانه‌ای حل‌کننده‌های رضایت‌پذیری (SAT) و حل‌کننده‌های رضایت‌پذیری پیمانه‌ای (SMT) هستند که می‌توانند استنتاجات منطقی پیچیده‌ای از نتایج از مجموعه‌های خاصی از فرضیه‌ها انجام دهند و برای هر استنتاج گواه اثبات ایجاد کنند. البته، قابلیت اثبات یک مسئله کامل-NP^۵ است، بنابراین این حل‌کننده‌ها از نقطه‌ای معین فراتر نمی‌روند. در اینجا یک نمونه معمولی از نتیجه‌ای است که با استفاده از یک حل‌کننده SAT ثابت شده است:

قضیه (قضیه سه‌گانه فیثاغورسی بولی [۱۲]). مجموعه $\{1, 2, \dots, 7824\}$ را می‌توان به دو دسته تقسیم کرد که هیچ کدام حاوی یک سه‌گانه فیثاغورسی (a, b, c) با $a^2 + b^2 = c^2$ نیستند. با این حال، این برای $\{1, 2, \dots, 7825\}$ امکان‌پذیر نیست.

اثبات به چهار سال محاسبات CPU نیاز داشت و گزاره‌ای ۲۰۰ ترابایتی تولید کرد که بعداً به ۶۸ گیگابایت فشرده شد.

کامپیوترها، البته، به طور معمول توسط ریاضی‌دانان برای کارهای روزمره مانند نوشتن مقالات و ارتباط با همکاران نیز استفاده می‌شوند. اما در دهه‌های اخیر، روش‌های جدید امیدوارکننده‌ای برای استفاده از رایانه‌ها در کمک به تحقیقات ریاضی پدیدار شده‌اند:

الگوریتم‌های یادگیری ماشینی می‌توانند برای کشف روابط ریاضی جدید یا تولید نمونه‌ها یا مثال‌های نقض بالقوه برای مسائل ریاضی استفاده شوند.

ریاضی‌دانان قرن‌هاست (یا حتی هزاره‌ها، اگر ابزارهای محاسباتی اولیه مانند چرتکه را در نظر بگیریم) برای کمک به تحقیقات خود به رایانه‌ها (انسانی، مکانیکی یا الکترونیکی) و ماشین‌ها متکی بوده‌اند. برای نمونه، از زمان جداول اولیه لگاریتم نپر و دیگران، ریاضی‌دانان ارزش ساختن مجموعه داده‌های بزرگی از اشیاء ریاضی را برای انجام محاسبات و حدس زدن، می‌دانستند. لژاندر و گاوس از جداول گسترده اعداد اول که توسط محاسبات دستی گردآوری شده بود، برای حدس زدن آنچه اکنون به عنوان قضیه اعداد اول شناخته می‌شود، استفاده کردند. یک قرن و نیم بعد، بیرچ و سوینرتون-دایر^۱ به طور مشابه از رایانه‌های الکترونیکی اولیه برای تولید داده‌های کافی در مورد منحنی‌های بیضوی روی میدان‌های متناهی برای ارائه حدس مشهور خود در مورد این اشیاء استفاده کردند و بسیاری از خوانندگان بدون شک از یکی از گسترده‌ترین مجموعه داده‌های ریاضی، یعنی دانشنامه آنالین دنباله‌های صحیح، استفاده کرده‌اند که حدس‌های متعدد و روابط غیرمنتظره بین زمینه‌های مختلف ریاضی را ایجاد کرده است و به عنوان یک موتور جستجوی ریاضی ارزشمند برای محققانی که به دنبال ادبیات در مورد یک موضوع ریاضی هستند که نام آن را نمی‌دانند، اما می‌توانند آن را با دنباله‌ای از اعداد صحیح مرتبط کنند، عمل می‌کند. در قرن بیست و یکم، چنین پایگاه داده‌های بزرگی همچنین به عنوان داده‌های آموزشی مهم برای الگوریتم‌های یادگیری ماشین عمل می‌کنند که نوید خودکارسازی یا حداقل تسهیل فرآیند ایجاد حدس‌ها و اتصالات در ریاضیات را دارند. علاوه بر تولید داده، یکی دیگر از کاربردهای قدیمی رایانه‌ها در محاسبات علمی است که امروزه به طور گسترده برای حل عددی معادلات دیفرانسیل و سیستم‌های دینامیکی یا محاسبه آمار ماتریس‌های بزرگ یا عملگرهای خطی استفاده می‌شود. نمونه اولیه چنین محاسباتی در دهه ۱۹۲۰ زمانی رخ داد که هندریک لورنتس^۲ تیمی از محاسبه‌گران انسانی را برای مدل‌سازی جریان سیال در اطراف آفسلوت‌دایک^۳ - یک سد بزرگ که در آن زمان در هلند در حال ساخت بود - گرد هم آورد. در میان چیزهای دیگر، این محاسبه به دلیل پیشگام بودن دستگاه (اکنون استاندارد)

^۱ Birch and Swinnerton-Dyer

^۲ Hendrik Lorentz

^۳ Afsluitdijk

^۴ Magma, SAGE-Math, Mathematica, Maple

^۵ NP-complete

ابزارهای اثبات رسمی می‌توانند برای تأیید اثبات‌ها (و همچنین خروجی مدل‌های بزرگ زبانی) استفاده شوند، همکاری‌های ریاضی واقعی در مقیاس بزرگ را امکان‌پذیر کنند و به ساخت مجموعه داده‌ها برای آموزش الگوریتم‌های یادگیری ماشین مذکور کمک کنند. مدل‌های بزرگ زبانی مانند چت‌جی‌پی‌تی^۶ می‌توانند (به‌طور بالقوه) استفاده از ابزارهای دیگر را آسان‌تر و سریع‌تر کنند؛ آنها همچنین می‌توانند استراتژی‌های اثبات یا کارهای مرتبط را پیشنهاد دهند و حتی مستقیماً اثبات‌های (ساده) را تولید کنند.

هریک از این نوع ابزارها در حال حاضر کاربردهای خاصی را در زمینه‌های مختلف ریاضیات پیدا کرده‌اند، اما چیزی که من را به ویژه مجذوب می‌کند، امکان ترکیب این ابزارها با هم است، به‌طوری که یک ابزار نقاط ضعف دیگری را جبران کند. به‌عنوان مثال، دستیاران اثبات رسمی و بسته‌های جبر رایانه‌ای می‌توانند تمایل اکنون بدنام مدل‌های بزرگ زبانی برای «توهم»^۷ مزخرفات به‌ظاهر معقول را فیلتر کنند، در حالی که برعکس این مدل‌ها می‌توانند به خودکارسازی جنبه‌های خسته‌کننده‌تر رسمی‌سازی اثبات کمک کنند، و همچنین یک رابط زبانی طبیعی برای اجرای الگوریتم‌های پیچیده نمادین یا یادگیری ماشینی فراهم کنند. بسیاری از این ترکیبات هنوز فقط در مرحله اثبات مفهوم توسعه هستند و زمان می‌برد تا فناوری به یک ابزار واقعاً مفید و قابل اعتماد برای ریاضی‌دانان تبدیل شود. با این حال، آزمایشات اولیه به‌نظر دلگرم‌کننده است و باید انتظار داشته باشیم که در آینده نزدیک شاهد برخی از نمایش‌های شگفت‌انگیز روش‌های جدید تحقیقات ریاضی باشیم. نه مفهوم علمی تخیلی یک هوش مصنوعی ابرهوشمند که بتواند به‌طور مستقل مسائل پیچیده ریاضی را حل کند، بلکه یک دستیار ارزشمند که می‌تواند ایده‌های جدید را پیشنهاد کند، خطاها را فیلتر کند و بررسی موارد معمول، آزمایش‌های عددی و وظایف بررسی ادبیات را انجام دهد و به ریاضی‌دانان انسانی در پروژه اجازه دهد تا بر اکتشاف مفاهیم سطح بالا تمرکز کنند.

دستیاران اثبات

البته، صرف انجام یک محاسبه با استفاده از رایانه، به‌طور خودکار تضمین نمی‌کند که صحیح باشد. محاسبه ممکن است دچار خطاهای عددی شود، مانند خطاهایی که ناشی از جایگزینی متغیرهای پیوسته یا معادلات با تقریب‌های گسسته است. اشکالات ممکن است به‌طور ناخواسته در کد وارد شوند یا خود داده‌های ورودی ممکن است

⁶ ChatGPT ⁷Hallucinate ⁸Appel and Haken ⁹reducibility ¹⁰unavoidability ¹¹Dorothea Blostein ¹²Robertson, Sanders, Seymour, and Thomas ¹³Coq, Isabelle or Lean

زمینه‌های ریاضیات نسبت به سایرین، رسمی‌سازی دشوارتری دارند. کوین بازارد^{۲۱} اخیراً پروژه‌ای را برای رسمی‌سازی اثبات آخرین قضیه فرما اعلام کرده است که تخمین می‌زند حداقل پنج سال طول بکشد.

با توجه به تمام تلاش‌های لازم، ارزش تلاش‌های رسمی‌سازی اثبات در ریاضیات چیست؟ بدیهی‌ترین مزیت آن، ایجاد سطح بسیار بالایی از اطمینان از صحت یک نتیجه معین است که به‌ویژه برای نتایج بحث برانگیز یا بدنام برای جذب اثبات‌های نادرست، یا برای اثبات‌های بسیار طولانی در زمینه‌هایی که داوران مایل به تأیید چنین اثبات‌هایی خط‌به‌خط هستند، ارزشمند است. (از نظر تئوری، هنوز ممکن است یک اشکال پنهان در کامپایلر دستیار اثبات وجود داشته باشد - که عمده‌اً تا حد ممکن کوچک نگه داشته شده است تا این احتمال را کاهش دهد - یا تعاریفی که در بیانیه رسمی نتیجه استفاده شده است ممکن است به روش‌های ظریف اما مهمی با بیانیه قابل‌خواندن توسط انسان متفاوت باشد، اما چنین سناریویی بعید است، به‌خصوص اگر اثبات رسمی از نزدیک اثبات نوشته شده توسط انسان را دنبال کند.) فرآیند رسمی‌سازی معمولاً مسائل جزئی را در اثبات انسانی آشکار می‌کند و گاهی اوقات می‌تواند ساده‌سازی‌ها یا تقویت‌های استدلال را آشکار کند، برای مثال با آشکار کردن این که یک فرضیه ظاهراً مهم در یک لم در واقع غیرضروری بود، یا این که می‌توان از یک ابزار کم‌قدرت اما کلی‌تر به جای یک ابزار پیشرفته اما تخصصی استفاده کرد. یک پروژه رسمی‌سازی در یک زبان مدرن مانند Lean معمولاً نتایج ریاضی اساسی بسیاری را که در طول پروژه تولید شده‌اند به یک کتابخانه ریاضی مشترک کمک می‌کند که اجرای پروژه‌های رسمی‌سازی آینده را آسان‌تر می‌کند. اما دستیاران اثبات رسمی همچنین می‌توانند روش‌های جدید آموزش و همکاری ریاضی را فعال کنند. چندین پروژه آزمایشی در حال انجام است تا یک اثبات رسمی را به اشکال قابل‌درک‌تر برای انسان تبدیل کند، مانند یک متن تعاملی که در آن مراحل فردی در استدلال را می‌توان به جزئیات بیشتری گسترش داد یا به یک خلاصه سطح بالا جمع کرد. این می‌تواند فرمت مناسبی برای کتاب‌های درسی ریاضی آینده باشد. یک همکاری سنتی ریاضیات به‌ندرت شامل بیش از پنج یا شش همکار است، تا حدی به دلیل نیاز هر همکار به اعتماد و تأیید کار هر یک از همکاران دیگر. اما پروژه‌های رسمی‌سازی به‌طور معمول شامل ده‌ها نفر است که ممکن است هیچ تعامل قبلی نداشته باشند، دقیقاً به این دلیل که دستیار اثبات رسمی اجازه می‌دهد تا زیر وظایف

این جنبه‌های سطح پایین یک اثبات رسمی انجام شده است، اما در عمل بخش‌های «آشکار» یک استدلال ریاضی اغلب می‌تواند زمان بیشتری برای رسمی‌سازی نسبت به «مهم» ببرد. برای مثال: با توجه به سه مجموعه داده شده A_1, A_2, A_3 ، یک ریاضی‌دان ممکن است با ضرب دکارتی $(A_1 \times A_2) \times A_3$ ، $A_1 \times (A_2 \times A_3)$ و $A_1 \times A_2 \times A_3$ کار کند، از آنجایی که آنها «بدیهیاً» یکسان هستند اما در اکثر صورتی‌های ریاضی، این ضرب‌ها در واقع یکسان نیستند و نسخه رسمی استدلال ممکن است نیاز به سرمایه‌گذاری بخشی از اثبات برای ایجاد معادل‌های مناسب بین چنین فضاهایی داشته باشد و اطمینان حاصل کند که بیانیه‌هایی که شامل یک نسخه از این ضرب هستند برای نسخه دیگر نیز ادامه دارند. به همین دلیل و دلایل دیگر، تبدیل یک اثبات نوشته شده توسط یک ریاضی‌دان انسانی - حتی یک اثبات بسیار دقیق - به یک اثبات رسمی که در یک دستیار اثبات رسمی کامپایل می‌شود، بسیار زمان‌بر است، اگرچه این فرآیند به تدریج در طول زمان کارآمدتر شده است. قضیه چهار رنگ فوق‌الذکر توسط ورنر و گونتیه^{۱۴} در سال ۲۰۰۵ در Coq رسمی شد [۷].

حدس مشهور کپلر در مورد متراکم‌ترین بسته‌بندی $\mathbb{R}^3$ توسط توپ‌های واحد توسط هیلز و فرگوسن^{۱۵} در سال ۱۹۹۸ [۱۱] در یک اثبات بسیار پیچیده (و با کمک رایانه) اثبات شد. در سال ۲۰۰۳، هیلز پروژه فلای اسپک^{۱۶} را برای تأیید رسمی اثبات راه‌اندازی کرد و تخمین زد که انجام آن ۲۰ سال طول خواهد کشید، اگرچه در نهایت مشخص شد که از طریق همکاری بین هیلز و ۲۱ مشارکت‌کننده دیگر، این کار در «فقط» ۱۱ سال به دست آمد [۱۰]. اخیراً، شولزه^{۱۷} در سال ۲۰۱۹، «آزمایش تانسور مایع»^{۱۸} [۵] برای تأیید رسمی قضیه بنیادی خود و کلاوزن^{۱۹} در مورد ناپدید شدن یک گروه Ext خاص از یک «فضای برداری مایع» در نظریه ریاضیات متراکم راه‌اندازی شد. اثبات نوشته شده توسط انسان «فقط» ده صفحه طول داشت، اگرچه با مقدار زیادی مواد پیش‌نیاز در ریاضیات متراکم همراه بود. با این وجود، رسمی‌سازی در Lean حدود ۱۸ ماه در یک تلاش مشترک بزرگ طول کشید. من خودم رهبری یک تلاش رسمی‌سازی؟؟ را بر روی اثبات اخیر گاورز، گرین، مانرز^{۲۰} و خودم [۹] از یک حدس در ترکیبیت جمعی برعهده داشتم. اثبات نوشته شده توسط انسان ۳۳ صفحه طول داشت، اما عمدتاً مستقل بود و گروهی متشکل از حدود ۲۰ همکار توانستند آن را در سه هفته رسمی کنند. برخی از

¹⁴ Werner and Gonthier ¹⁵ Hales and Ferguson ¹⁶ Flyspeck ¹⁷ Scholze ¹⁸ liquid tensor experiment ¹⁹ Clausen ²⁰ Gowers, Green, Manners

²¹ Kevin Buzzard

کوچک). این مدل‌ها کاربردهای عملی بی‌شماری دارند، به‌عنوان مثال در تشخیص تصویر و گفتار، سیستم‌های توصیه یا تشخیص تقلب. با این حال، آنها معمولاً با تضمین‌های قوی دقت همراه نیستند، به‌ویژه هنگام اعمال به ورودی‌هایی که به‌طور قابل توجهی با مجموعه داده آموزشی متفاوت هستند، یا زمانی که مجموعه داده آموزشی نویزی یا ناقص است. علاوه بر این، مدل‌ها اغلب مبهم هستند، به این معنا که استخراج توضیح قابل درک توسط انسان از مدل در مورد اینکه چرا مدل پیش‌بینی خاصی انجام داده است، یا درک رفتار کلی مدل دشوار است. به این ترتیب، در نگاه اول به‌نظر می‌رسد که این ابزارها برای تحقیقات ریاضی مناسب نیستند، جایی که هم اثبات دقیق و هم درک شهودی استدلال‌ها مورد نظر است. با این وجود، موارد استفاده امیدوارکننده اخیر از یک ابزار یادگیری ماشینی مناسب برای تولید یا حداقل پیشنهاد ریاضیات دقیق جدید وجود داشته است، به‌ویژه هنگام ترکیب با تکنیک‌های قابل اعتمادتر که می‌توانند خروجی این ابزارها را تأیید کنند. برای مثال، یک مشکل اساسی در نظریه ریاضی معادلات سیال مانند معادلات اوایلر یا ناویه-استوکس، توانایی اثبات دقیق انفجار زمانی محدود راه حل‌های u در زمان محدود از داده‌های اولیه هموار است. بدنام‌ترین نمونه این مورد مربوط به معادلات ناویه-استوکس تراکم‌ناپذیر در سه بعد است، که حل آن یکی از مسائل جایزه هزاره (حل نشده) است. این هنوز خارج از دسترس است، اما پیشرفت‌های اخیر در معادلات سیال دیگر، مانند معادلات بوسینسک^{۲۳} در دو بعد (یک مدل ساده‌شده برای معادلات اوایلر تراکم‌ناپذیر در سه بعد) حاصل شده است. یکی از مسیرهای ایجاد چنین تکنیکی، ساختن یک راه حل انفجار خودمشابه u است که توسط یک تابع با بعد کمتر U توصیف می‌شود که یک PDE ساده‌تر را حل می‌کند. به‌نظر نمی‌رسد راه حل فرم بسته‌ای برای این PDE در دسترس باشد. اما اگر بتوان یک راه حل تقریبی $\hat{U}$ با کیفیت کافی برای این PDE تولید کرد (که تقریباً از شرایط مرزی معینی اطاعت می‌کند)، ممکن است بتوان با استفاده از نظریه اختلال (مانند آنهایی که بر اساس قضایای نقطه ثابت هستند) یک راه حل دقیق U را به‌طور دقیق نشان داد. به‌طور سنتی، از روش‌های PDE عددی برای تلاش برای تولید این راه حل‌های تقریبی U استفاده می‌شد، برای مثال با گسسته‌سازی PDE به یک معادله تفاضلی، اما استفاده از چنین روش‌هایی برای به‌دست آوردن راه حل‌هایی با سطح دقت موردنظر می‌تواند از نظر محاسباتی پرهزینه باشد. رویکرد جایگزینی در سال ۲۰۱۹ توسط وانگ، لای، گومز-سرانو و باک مستر^{۲۴} [۱۹] پیشنهاد

فردی در پروژه به‌طور دقیق تعریف و به‌طور مستقل از سایر زیر وظایف تأیید شوند. قابل تصور است که این دستیاران اثبات نیز می‌توانند تقسیم کار مشابهی را برای تولید نتایج ریاضی جدید امکان‌پذیر کنند و امکان همکاری‌های بسیار موازی و جمع‌سپاری را در مقیاسی بسیار بزرگتر از همکاری‌های آنلاین قبلی (مانند پروژه‌های «پلی مت» [۸]) فراهم کنند. با گذشت زمان، همکاری‌های بزرگ که در حال حاضر در سایر علوم یا در پروژه‌های مهندسی نرم‌افزار رویه‌ای تثبیت شده است، ممکن است در تحقیقات ریاضی نیز رایج شود؛ برخی از مشارکت‌کنندگان ممکن است نقش «مدیران پروژه» را ایفا کنند و به‌عنوان مثال بر ایجاد «نقشه‌های راه» دقیقی متمرکز شوند که پروژه را به قطعات کوچکتر تقسیم می‌کنند، در حالی که برخی دیگر می‌توانند در اجزای فردی پروژه تخصص داشته باشند، بدون اینکه لزوماً تمام تخصص‌های مورد نیاز برای درک پروژه به‌عنوان یک کل را داشته باشند. با این حال، قبل از اینکه این اتفاق بیفتد، فرآیند صوری‌سازی باید کارآمدتر شود. «عامل د بروین»^{۲۲} (نسبت بین دشواری نوشتن یک اثبات رسمی صحیح و یک اثبات غیررسمی صحیح) هنوز بسیار بالاتر از ۱ است (من تخمین می‌زنم ۲۰)، اما در حال کاهش است. من معتقدم هیچ مانع اساسی برای کاهش این نسبت به زیر ۱ وجود ندارد، به‌ویژه با ادغام بیشتر با هوش مصنوعی، حل‌کننده‌های SMT و سایر ابزارها؛ این امر برای زمینه ما متحول‌کننده خواهد بود.

یادگیری ماشینی

یادگیری ماشینی به طیف گسترده‌ای از تکنیک‌ها برای آموزش رایانه به‌منظور انجام یک کار پیچیده اشاره دارد - مانند پیش‌بینی یک خروجی متناظر با یک ورودی معین که از یک کلاس بسیار گسترده ترسیم شده است، یا تشخیص همبستگی‌ها و روابط دیگر در یک مجموعه داده. بسیاری از مدل‌های محبوب برای یادگیری ماشینی از نوعی شبکه عصبی برای رمزگذاری نحوه انجام کار توسط رایانه استفاده می‌کنند. این شبکه‌ها توابعی از متغیرهای متعدد هستند که با ترکیب تعداد زیادی از عملیات‌های ساده‌تر (هم خطی و هم غیرخطی) تشکیل می‌شوند. معمولاً نوعی تابع پاداش (یا تابع جریمه) به چنین شبکه‌ای اختصاص می‌دهند، برای مثال با اندازه‌گیری تجربی عملکرد آن در برابر یک مجموعه داده آموزشی و سپس انجام بهینه‌سازی محاسباتی فشرده برای یافتن انتخاب‌های پارامتر برای این شبکه به‌منظور ایجاد تابع پاداش تا حد امکان بزرگ (یا تابع ضرر تا حد امکان

²² De Bruijn factor ²³ Boussinesq ²⁴ Wang, Lai, Gómez-Serrano, and Buckmaster

بیشتر پیش رفت. این تجزیه و تحلیل نشان داد که از دو دوجین ثابت هذلولی استفاده شده، تنها سه مورد از آنها (ترجمه طولی و قسمت‌های واقعی و پیچیده ترجمه نصف النهاری) تأثیر قابل توجهی بر تابع پیش‌بینی داشتند. با بازرسی بصری نمودارهای پراکندگی امضا در برابر این سه ثابت، نویسندگان توانستند رابطه قابل درک‌تری بین این کمیت‌ها را حدس بزنند. اعداد بیشتر حدس اولیه آنها را رد کرد، اما نسخه اصلاح شده‌ای از حدس را پیشنهاد کرد که آنها توانستند آن را به‌طور دقیق اثبات کنند. این تعامل بین حدس‌های تولید شده توسط ماشین و تأیید (و اصلاح) انسان با استفاده از نظریه، الگویی امیدوارکننده است که به‌نظر می‌رسد در بسیاری از زمینه‌های دیگر ریاضیات قابل اجرا است.

بسیاری از کاربردهای یادگیری ماشینی به مقدار زیادی داده آموزشی نیاز دارند که در حالت ایده‌آل به شکلی استاندارد (به‌عنوان مثال، یک بردار از اعداد) نمایش داده شوند تا بتوان الگوریتم‌های موجود یادگیری ماشین را با سهولت نسبی به آن اعمال کرد. نمایش دقیق داده‌ها می‌تواند از اهمیت حیاتی برخوردار باشد؛ همبستگی بین اجزای مختلف داده‌ها ممکن است به‌راحتی توسط الگوریتم‌های یادگیری ماشینی در یک نمایش داده قابل کشف باشد، اما در نمایش دیگری تقریباً غیرممکن است. در حالی که برخی از حوزه‌های ریاضیات در حال شروع جمع‌آوری پایگاه‌های داده بزرگ از اشیاء مفید (مانند گره‌ها، گراف‌ها یا منحنی‌های بیضوی) هستند، هنوز بسیاری از کلاس‌های مهم مفاهیم ریاضی مبهم‌تر تعریف شده‌اند که به شکلی سیستماتیک در قالبی قابل استفاده برای یادگیری ماشین قرار نگرفته‌اند. به‌عنوان مثال، با بازگشت به مثال PDE هزاران معادله دیفرانسیل جزئی مختلف در ادبیات مطالعه شده‌اند، اما اغلب با تغییرپذیری زیادی در نمادگذاری و ترتیب جبری عبارات، و هیچ چیز شبیه به یک پایگاه داده استاندارد از PDE‌های معمولاً مورد مطالعه همراه با خواص اساسی آنها (مانند اینکه آیا آنها بیضوی، سهموی یا هذلولی هستند؛ چه چیزی در مورد وجود و یکتایی جواب‌ها، قوانین بقاء، و غیره شناخته شده است) وجود ندارد. چنین پایگاه داده‌ای می‌تواند بالقوه در پیش‌بینی حدسی در مورد رفتار یک PDE بر اساس نتایج برای PDE‌های دیگر مفید باشد، یا در پیشنهاد احتمالات قیاسی یا کاهش از یک PDE به PDE دیگر؛ اما فقدان هر فرم نرمال متعارف برای نمایش چنین معادلاتی (یا حداقل یک «اثر انگشت» برای شناسایی آنها [۲]) در حال حاضر ایجاد چنین پایگاه داده‌ای را دشوار می‌کند، چه رسد به تغذیه آن به یک شبکه عصبی. با این حال،

شد، که از یک شبکه عصبی آگاه از فیزیک (PINN) آموزش‌دیده برای تولید توابع $\hat{U}$ استفاده کردند که یک تابع ضرر مناسب را به حداقل می‌رسانند که میزان تقریباً رعایت PDE و شرایط مرزی مورد نظر را اندازه‌گیری می‌کند. از آنجایی که این توابع $\hat{U}$ از طریق یک شبکه عصبی و نه یک نسخه گسسته از معادله تولید می‌شوند، می‌توان آن‌ها را سریع‌تر تولید کرد و احتمالاً کمتر مستعد ناپایداری‌های عددی هستند. همانطور که مشخص شد، کار همزمان چن و هو [۲۵] [۴] توانست انفجار زمان محدود را برای این معادله با استفاده از روش‌های عددی سنتی‌تر ثابت کند. با این حال، الگوی یادگیری ماشینی پتانسیل بالایی را به‌عنوان یک رویکرد مکمل برای این نوع مسائل PDE نشان می‌دهد. به‌عنوان مثال، می‌توان رویکرد ترکیبی را تصور کرد که در آن یک ریاضی‌دان انسانی ابتدا یک فرضیه انفجار را پیشنهاد می‌کند، که برای آن یک شبکه عصبی سپس سعی می‌کند یک راه حل تقریبی تقریبی پیدا کند، و سپس از روش‌های عددی سنتی‌تر برای اصلاح آن راه حل به یک راه حل استفاده می‌شود. که به اندازه کافی دقیق برای اعمال تجزیه و تحلیل پایداری دقیق است.

مثال دیگری از استفاده از یادگیری ماشینی در ریاضیات در زمینه نظریه گره است. گره‌ها دارای مجموعه بسیار متنوعی از ثابت‌های توپولوژیکی هستند: امضای یک گره، یک عدد صحیح مرتبط با همولوژی یک سطح محدود کننده گره (یک سطح سیفرت) است. چندجمله‌ای جونز یک گره را می‌توان با استفاده از نظریه نمایش بافته‌ها توصیف کرد و اکثر گره‌ها (به استثنای گره‌های تورو) دارای یک هندسه هذلولی متعارف در مکمل هستند که می‌توان از آن برای توصیف تعدادی ثابت هذلولی، مانند حجم هذلولی استفاده کرد. اولاً، مشخص نیست که چگونه این ثابت‌ها از حوزه‌های بسیار متفاوت ریاضیات با یکدیگر مرتبط هستند. با این حال، در سال ۲۰۲۱، دیویس، جوهاز، لاکنبی و توماسف [۲۶] [۶] این موضوع را از طریق یادگیری ماشینی بررسی کردند. با آموزش یک شبکه عصبی بر روی یک پایگاه داده موجود از تقریباً دو میلیون گره (به همراه یک میلیون گره اضافی به صورت تصادفی تولید شده)، آنها توانستند این شبکه عصبی را برای پیش‌بینی امضای یک گره از حدود دو دوجین ثابت هذلولی با دقت بالا آموزش دهند. با این حال، تابع پیش‌بینی تولید شده کاملاً مبهم بود و در ابتدا بینش زیادی را در مورد رابطه دقیق بین امضا و ثابت‌های هذلولی آشکار نکرد. با این وجود، می‌توان با استفاده از یک ابزار ساده به نام تحلیل برجستگی، که به‌طور تقریبی، تأثیر هر ثابت هذلولی فردی را بر تابع پیش‌بینی اندازه‌گیری می‌کند،

²⁵Chen and Hou ²⁶Davies, Juhász, Lackenby, and Tomasev

در آینده قابل تصور است که پیشرفت در هر دو زمینه صوری سازی اثبات و هوش مصنوعی، تولید و استفاده از چنین پایگاه های داده ای (که می تواند شامل مجموعه داده های «دنیای واقعی» و «مصنوعی» باشد) را امکان پذیرتر کند.

مدل های زبان بزرگ

مدل های زبان بزرگ (LLM) نوع نسبتاً جدیدی از مدل های یادگیری ماشینی هستند که برای آموزش روی مجموعه داده های بسیار گسترده و بزرگی از متون زبان طبیعی مناسب هستند. یک مدل زبان بزرگ محبوب GPT ۲۷ است که همان طور که از نام آن پیداست بر اساس مدل ترانسفورماتور^{۲۸} ساخته شده است - نوعی شبکه عصبی مصنوعی که برای پیش بینی کلمه بعدی (یا «توکن») در یک رشته از کلمات طراحی شده است، که در آن برخی از «توجه» بلندمدت به کلمات در اوایل رشته برای شبیه سازی زمینه جمله حفظ می شود. با تکرار این مدل، سپس می توان یک پاسخ متنی طولانی برای یک دستورالعمل متنی معین ایجاد کرد. هنگامی که روی مقدار کمی از داده ها آموزش داده می شود، خروجی چنین مدل هایی چندان چشمگیر نیست - چندان پیچیده تر از تلاش برای تکرار ویژگی ورودی متن «تکمیل خودکار» روی یک گوشی هوشمند نیست - اما پس از آموزش گسترده روی مجموعه داده های بسیار بزرگ و متنوع، خروجی های این مدل ها می تواند به طور شگفت انگیزی منسجم و حتی خلاقانه باشد و متنی را تولید کند که در نگاه اول تشخیص آن از نوشتار انسان دشوار است، اگرچه در بازرسی دقیق تر، خروجی اغلب بی معنی است و به هیچ حقیقت اساسی مرتبط نیست، پدیده ای که به عنوان «توهم» شناخته می شود. البته می توان چنین LLM های عمومی را برای حمله مستقیم به مسائل ریاضی امتحان کرد. گاهی اوقات نتایج می تواند بسیار چشمگیر باشد. برای مثال، بوبک و دیگران^{۲۹} موردی را مستند می کنند که در آن مدل زبان بزرگ قدرتمند GPT - 4 توانست اثبات کامل و درستی از یک مسئله از المپیاد ریاضی بین المللی ۲۰۲۲ ارائه دهد، که در مجموعه داده های آموزشی برای این مدل نبود. برعکس، این مدل برای انجام محاسبات دقیق یا حتی محاسبات اولیه مناسب نیست. در یک مورد [۳]، هنگامی که از GPT - 4 خواسته شد عبارت $8 * 8 + 4 * 7$ را محاسبه کند، GPT - 4 به سرعت پاسخ نادرست ۱۲۰ را ارائه کرد، سپس با یک روش گام به گام که پاسخ صحیح ۹۲ را به دست آورد، به توجیه محاسبه پرداخت. هنگامی که در مورد این تناقض سؤال شد،

GPT - 4 تنها توانست پاسخ دهد که حدس اولیه آن یک «خطای تایپی» بوده است. این مسائل را می توان تا حدودی با استفاده از «پلاگین» برای GPT - 4 جبران کرد، که در آن آموزش داده می شود تا انواع خاصی از پرس و جوا، مانند محاسبات ریاضی، را به یک ابزار خارجی (مانند ولفرام آلفا^{۳۰} ارسال کند تا حدس پاسخ را از طریق مدل داخلی خود بزند، اگرچه ادغام بین ابزارها در حال حاضر یکپارچه نیست. به طور مشابه، یک اثبات مفهومی اخیر [۱۵] نشان داده است که می توان از LLM ها برای یافتن مثال هایی در مسائل مختلف ترکیبیات و علوم کامپیوتر استفاده کرد که از نمونه های قبلی تولید شده توسط انسان بهتر عمل می کنند، با درخواست از این مدل ها برای تولید یک برنامه برای ایجاد چنین مثال هایی به جای تلاش برای ساخت مستقیم مثال، و سپس اجرای آن برنامه در زبان دیگری برای تأیید قابل اعتماد کیفیت خروجی، که سپس به مدل اصلی ارسال می شود تا آن را برای بهبود حدس های خود تحریک کند. همچنین پیشرفت های اخیر در استفاده از LLM ها برای بهبود موتورهای اثبات نمادین موجود برای حمله به کلاس های محدودی از مسائل ریاضی، مانند مسائل هندسه المپیاد [۱۷] وجود داشته است. در آزمایش های خودم با GPT - 4 (که می توانید آن را در این لینک پیدا کنید)، دریافتام که مفیدترین موارد استفاده برای تولید کد رایانه ای اولیه به زبان های مختلف (پایتون، سیج، لاتک، لین، ریجکس و غیره) یا برای پاک سازی مجموعه های داده آشفته و سازمان دهی نشده (به عنوان مثال، برای سازماندهی انبوهی از منابعی که از اینترنت استخراج شده است به یک کتابخانه لاتک منسجم، پس از ارائه چند نمونه از فرمت دلخواه برای موارد کتابخانه به GPT - 4 برای شروع کار) بوده است. در چنین مواردی، اغلب در اولین تلاش خروجی رضایت بخش یا تقریباً رضایت بخش تولید می کرد، و برای به دست آوردن نوع خروجی مورد نظر من، تنها به مقدار کمی تجدید نظر نیاز بود. همچنین در ترغیب GPT - 4 برای پیشنهاد منابع علمی یا تکنیک های مرتبط برای یک مسئله ریاضی واقعی، موفقیت های محدودی کسب کرده ام. در یک مورد آزمایشی، از آن پرسیدم که چگونه می توان سرعت پوسیدگی نمایی را در احتمال دم یک جمع متغیرهای تصادفی مستقل محاسبه کرد تا ارزیابی کنم که آیا دانش کافی در مورد قضیه مرتبط در این زمینه (قضیه کرامر) را بدون ارائه کلمات کلیدی مانند نظریه انحراف بزرگ ترجمه کنید. همانطور که مشخص شد، GPT - 4 دقیقاً این قضیه را پیدا نکرد و در عوض رشته ای از مزخرفات ریاضی تولید کرد، اما کنجکاو است که موفق شد به تابع مولد لحظه لگاریتمی

²⁷ Generative Pre-trained Transformer

²⁸ Transformer

²⁹ Bubeck et al.

³⁰ Wolfram Alpha

این مدل‌ها بر اساس سبک و ترجیحات نوشتن شخصی امکان‌پذیر می‌شود.

آیا می‌توان این ابزارها را با هم ترکیب کرد؟

تکنولوژی‌های مختلفی که در بالا مورد بحث قرار گرفتند دارای نقاط قوت و ضعف بسیار متنوعی هستند و هیچ‌یک از آنها در سطح فعلی توسعه خود به‌عنوان ابزارهای همه‌منظوره برای ریاضی‌دانان مناسب نیستند، مانند پلتفرم‌های همه‌گیر مانند لاتک یا آرکایو. با این حال، آزمایشات امیدوارکننده اخیر در ایجاد ابزارهای رضایت‌بخش‌تر با ترکیب دو یا چند فناوری جداگانه با هم وجود دارد. به‌عنوان مثال، یک روش محتمل برای مقابله با ماهیت توهم‌زای LLMها هنگام تولید اثبات، الزام مدل به قالب‌بندی خروجی خود به زبان یک دستیار اثبات رسمی است، و هرگونه خطایی که توسط دستیار ایجاد می‌شود به‌عنوان بازخورد به مدل ارسال می‌شود. این سیستم ترکیبی برای تولید اثبات‌های کوتاه عبارات ساده مناسب به‌نظر می‌رسد [۲۰]. از آنجایی که چنین کارهایی اغلب عامل محدود کننده در رسمی‌سازی کارآمد اثبات هستند، این نوع الگو می‌تواند سرعت چنین رسمی‌سازی را به‌میزان قابل‌توجهی افزایش دهد، به‌ویژه اگر این مدل‌ها به‌طور خاص روی اثبات‌های رسمی تنظیم دقیق شوند. به‌جای متن عمومی و با ادغام با روش‌های اثبات قضیه خودکار سنتی‌تر، مانند استقرار حل کننده‌های SMT.

با توانایی پذیرش ورودی‌های زبان طبیعی، LLMها همچنین می‌توانند رابط کاربری کاربرپسندی باشند که به ریاضی‌دانان بدون تخصص نرم‌افزاری خاص اجازه می‌دهند از ابزارهای پیشرفته استفاده کنند. همانطور که قبلاً ذکر شد، من و بسیاری دیگر از این مدل‌ها برای تولید کد ساده به زبان‌های مختلف (از جمله بسته‌های جبر نمادین) یا برای ایجاد نمودارها و تصاویر پیچیده استفاده می‌کنیم. به‌نظر می‌رسد منطقی است که انتظار داشته باشیم در آینده نزدیک، بتوان از طریق این مدل‌ها نیز با استفاده از دستورالعمل‌های سطح بالا و محاوره‌ای، چیزی به پیچیدگی یک مدل یادگیری ماشین را طراحی و اجرا کرد. جسورانه‌تر، می‌توان امیدوار بود که در نهایت بتوان (پیش‌نویس) کل مقالات تحقیقاتی را همراه با تأیید رسمی تولید کرد، با توضیح یک نتیجه به زبان طبیعی به یک هوش مصنوعی، که سعی می‌کند هر مرحله از نتیجه را رسمی‌سازی کند. و هر زمان که نیاز به توضیح بیشتر بود از نویسنده سوال کند. ماهیت انسان‌محور تأییدیه‌های اثبات رسمی در شکل فعلی آن به

اشاره کند، که یک مفهوم کلیدی در بیان قضیه کرامر است، حتی اگر به‌نظر نمی‌رسد «بداند» دقیقاً چگونه این تابع با مساله مرتبط است. در آزمایش دیگری، من از $GPT-4$ درخواست پیشنهاداتی برای اثبات یک اتحاد ترکیبی کردم که روی آن کار می‌کردم. این چندین پیشنهاد را ارائه داد که من قبلاً آنها را در نظر گرفته بودم (تجزیه و تحلیل مجانبی، القا، عددی) و همچنین برخی توصیه‌های عمومی (تسهیل عبارات، جستجوی مشکلات مشابه، درک مشکل)، اما همچنین تکنیکی را پیشنهاد کرد (توابع مولد) که من به‌سادگی از آن غافل شده بودم و در نهایت مساله را به‌راحتی حل کردم. از سوی دیگر، چنین لیستی از توصیه‌ها احتمالاً برای یک ریاضی‌دان تازه‌کار که تجربه کافی برای ارزیابی مستقلانه هر یک از پیشنهادات ارائه شده را نداشت، چندان مفید نبود. با این وجود، من نقش این ابزارها را در کشف دانش نهفته کاربر از یک مشکل می‌بینم، صرفاً با گوش دادن خوب و ارائه ایده‌های معقول و مرتبط که کاربر به اندازه کافی متخصص است تا آنها را ارزیابی کند.

گیتوب کپایلو^{۳۱} یکی دیگر از مدل‌های GPT است که در چندین ویرایشگر کد محبوب ادغام شده است. این مدل با آموزش بر روی مجموعه داده‌های بزرگ کد در زبان‌های مختلف، طراحی شده است تا پیشنهادات تکمیل خودکار را برای کدی که تا حدی نوشته شده است، با استفاده از نشانه‌های زمینه‌ای مانند توضیحات غیررسمی کار برای انجام در جای دیگر کد ارائه دهد. من دریافتیم که این کار برای نوشتن LaTeX ریاضی و همچنین رسمی‌سازی در Lean به‌طرز شگفت‌انگیزی موثر است. در واقع، این به نوشتن همین مقاله نیز کمک کرد و چندین جمله را در حین نوشتن پیشنهاد کرد که بسیاری از آنها را برای نسخه نهایی حفظ یا کمی ویرایش کردم. اگرچه کیفیت پیشنهادات آن بسیار متغیر است، اما گاهی اوقات می‌تواند سطحی غیرقابل باور از درک شبیه‌سازی شده از هدف متن را نشان دهد.

به‌عنوان مثال، هنگام نوشتن یادداشت LaTeX توضیح دیگر در مورد نحوه تخمین انتگرال‌ها، من توضیح دادم که چگونه انتگرال باید به سه قسمت تقسیم شود و سپس جزئیات نحوه تخمین قسمت اول را ارائه دادم. سپس کپایلو به سرعت پیشنهاد کرد که چگونه قسمت دوم را با روشی مشابه تخمین بزنیم، متغیرها را به روشی کاملاً صحیح تغییر دهیم. فرکانس این تجربیات منجر به افزایش سرعت کمی اما قابل توجه در نوشتن من در هر دو LaTeX و Lean شده است و انتظار دارم که این ابزارها در آینده مفیدتر شوند، زیرا امکان «تنظیم دقیق»

³¹ Github Copilot

هوش مصنوعی اجازه داد سپس آرگومان‌ها را با خانواده‌های بزرگ معادلات مرتبط تطبیق دهند، و در صورت لزوم در صورت لزوم از نویسنده سؤال شود. از آرگومان‌ها غیرمعمول است. برخی نکات از این نوع از اکتشافات ریاضی در مقیاس بزرگ شروع به ظهور در زمینه‌های دیگر ریاضیات، مانند اکتشاف خودکار حدس‌ها در نظریه گراف [۱۸] کرده است. در حال حاضر مشخص نیست که کدام یک از این آزمایش‌ها در انجام آزمایش‌های پیشرفته موفق‌ترین خواهند بود کمک کامپیوتر به ریاضی‌دان معمولی شاغل برخی از شواهد مفهومی در حال حاضر مقیاس پذیر نیستند، به‌ویژه آنهایی که به مدل‌های هوش مصنوعی بسیار فشرده محاسباتی (و اغلب اختصاصی) وابسته هستند، یا به مقدار زیادی از ورودی و نظارت انسانی متخصص نیاز دارند. با این حال، من از تلاش‌های متنوع برای کشف فضای احتمالات دلگرم هستم و معتقدم که در آینده‌ای نزدیک، نمونه‌های بیشتری از روش‌های جدید برای انجام ریاضیات به کمک ماشین وجود خواهد داشت.

مطالعه بیشتر

موضوع اثبات به کمک ماشین کاملاً پراکنده است و در حوزه‌های مختلف ریاضیات، رایانه توزیع شده است. علم و حتی مهندسی؛ در حالی که هریک از زیرشاخه‌های مجزا فعالیت‌های فراوانی دارد، اخیراً تلاش‌هایی برای ایجاد یک جامعه متحدتر با جمع‌آوری همه موضوعات فهرست‌شده در این‌جا صورت گرفته است. به این ترتیب، در حال حاضر مکان‌های کمی وجود دارد که بتوان بررسی‌های جامعی از این روش‌های ریاضیات به سرعت در حال توسعه پیدا کرد. یک نقطه شروع، اقدامات [۱۳] یک کارگاه آموزشی آکادمی ملی در ژوئن ۲۰۲۳ در مورد "هوش مصنوعی برای کمک به استدلال ریاضی" است (که نویسنده یکی از سازمان‌دهندگان آن بود). به‌عنوان یکی از نتایج آن کارگاه، تالیا رینگر تلاشی را برای گردآوری هوش مصنوعی برای منابع ریاضی انجام داد که نتایج آن را می‌توانید در [لینک](#) پیدا کنید.

به‌عنوان مثال، در آن سند پیوندی به «بازی اعداد طبیعی» وجود دارد که یک راه در دسترس و تعاملی برای آشنایی با زبان دستیار اثبات ناب است. بسیاری از نمونه‌های مورد بحث در اینجا نیز از یک کارگاه IPAM در فوریه ۲۰۲۳ در مورد «اثبات به کمک ماشین» (که نویسنده نیز در سازماندهی آن مشارکت داشت) استخراج شده است، که صحبت‌های آن را می‌توان به‌صورت آنلاین یافت. از داور ناشناس برای اصلاحات و پیشنهادات تشکر می‌کنیم.

این معنی است که در حال حاضر برای بخش قابل توجهی از مقالات تحقیقاتی فعلی، رسمی‌سازی کامل آنها در زمان واقعی امکان‌پذیر نیست. با این حال، محتمل است که بسیاری از ابزارهایی که در حال حاضر برای تأیید اجزای محاسباتی خاص یک مقاله تحقیقاتی، مانند یکپارچه‌سازی عددی یا حل‌کننده PDE یک محاسبه جبر نمادین یا یک نتیجه حاصل‌شده با استفاده از یک حل‌کننده SMT استفاده می‌شوند، بتوان آنها را برای تولید گواهی‌های اثبات رسمی اصلاح کرد. علاوه بر این، کلاس محاسباتی که می‌توان به این روش رسمی‌سازی کرد، می‌تواند بسیار گسترده‌تر از جایی که در عمل فعلی قرار دارد، گسترش یابد. برای مثال، در حوزه PDE معمول است که صفحات محاسباتی را به تخمین برخی عبارات انتگرالی اختصاص داد که شامل یک یا چند تابع ناشناخته (مانند جواب‌های یک PDE) می‌شوند، با استفاده از کران‌های این توابع در هنجارهای مختلف فضای تابعی (مانند هنجارهای فضای سوبولف)، همراه با نابرابری‌های استاندارد (مانند نامساوی هولدر و نامساوی‌های سوبولف)، و همچنین هویت‌های مختلف مانند انتگرال‌گیری جزء به جزء یا مشتق‌گیری تحت علامت انتگرال. چنین محاسباتی، در حالی که معمول هستند، می‌توانند حاوی اشتباهات املائی (مانند خطاهای علامت) با درجات مختلف باشد، و می‌تواند برای دآوری با دقت خسته‌کننده باشد، زیرا محاسبات خود بینش کمی فراتر از تأیید صحت برآورد نهایی ارائه می‌دهند. می‌توان تصور کرد که می‌توان ابزارهایی را برای ایجاد چنین تخمین‌هایی به‌صورت خودکار یا نیمه خودکار ایجاد کرد و در حال حاضر طولانی است. و شواهد غیر روشن‌گر چنین برآوردهایی را می‌توان با پیوندی به گواهی اثبات رسمی جایگزین کرد. بلندپروازانه‌تر، ممکن است بتوان از یک ابزار هوش مصنوعی آینده درخواست کرد تا بهترین تخمینی را که می‌تواند، با توجه به مجموعه‌ای از فرضیه‌ها و روش‌های اولیه، ارائه دهد، بدون اینکه ابتدا مقداری محاسبه کاغذ و قلم برای حدس زدن آن تخمین انجام شود. در در حال حاضر، فضای حالت تخمین‌های ممکن بسیار پیچیده‌تر از آن است که به‌طور خودکار به این شکل مورد بررسی قرار گیرد. اما دلیلی نمی‌بینم که چرا این نوع اتوماسیون با پیشرفت فناوری قابل دستیابی نباشد. هنگامی که این مورد محقق شود، اکتشافات ریاضی در مقیاس‌هایی که در حال حاضر امکان‌پذیر نیست امکان‌پذیر می‌شود. در ادامه مثال PDE مقالات در این زمینه معمولاً یک یا دو معادله را در یک زمان مطالعه می‌کنند. اما در آینده ممکن است بتوان صدها معادله را به‌طور همزمان مطالعه کرد، شاید یک استدلال را به‌طور کامل فقط برای یک معادله ایجاد کرد و به ابزارهای

- [10] Thomas Hales, Mark Adams, Gertrud Bauer, Tat Dat Dang, John Harrison, Le Truong Hoang, Cezary Kaliszyk, Victor Magron, Sean McLaughlin, Tat Thang Nguyen, Quang Truong Nguyen, Tobias Nipkow, Steven Obua, Joseph Pleso, Jason Rute, Alexey Solovyev, Thi Hoai An Ta, Nam Trung Tran, Thi Diep Trieu, Josef Urban, Ky Vu, and Roland Zumkeller, A formal proof of the Kepler conjecture, *Forum Math. Pi* 5 (2017), e2, 29, DOI 10.1017/fmp.2017.1. MR3659768
- [11] Thomas C. Hales, A proof of the Kepler conjecture, *Ann. of Math.* (2) 162 (2005), no. 3, 1065–1185, DOI 10.4007/annals.2005.162.1065. MR2179728
- [12] Marijn J. H. Heule, Oliver Kullmann, and Victor W. Marek, Solving and verifying the Boolean Pythagorean triples problem via cube-and-conquer, *Theory and applications of satisfiability testing—SAT 2016*, 2016, pp. 228–245, <https://doi.org/10.1007/978-3-319-40970-2-15>. MR3534782
- [13] Samantha Koretsky (ed.), *Artificial intelligence to assist mathematical reasoning: Proceedings of a workshop*, National Academies Press, 2023, <http://dx.doi.org/10.17226/27241>.
- [14] Neil Robertson, Daniel P. Sanders, Paul Seymour, and Robin Thomas, A new proof of the four-colour theorem, *Electron. Res. Announc. Amer. Math. Soc.* 2 (1996), no. 1, 17–25, DOI 10.1090/S1079-6762-96-00003-0. MR1405965
- [15] Bernardino Romera-Paredes, Mohammadamin Barekatin, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli, and Alhussein Fawzi, Mathematical discoveries from program search with large language models, *Nature* 625 (December 2023), no. 7995, 468–475, DOI 10.1038/s41586-023-06924-6.
- [16] Terence Tao, Formalizing the proof of PFR in Lean4 using Blueprint: a short tour, 2023. <https://terrytao.word>
- *Terence Tao, Machine-Assisted Proof, *Notices of the American Mathematical Society*, 72 (1) (2025) 6-13.
- [1] Kenneth Appel and Wolfgang Haken, Every planar map is four colorable, *Contemporary Mathematics*, vol. 98, American Mathematical Society, Providence, RI, 1989. With the collaboration of J. Koch, DOI 10.1090/conm/098. MR1025335
- [2] Sara C. Billey and Bridget E. Tenner, Fingerprint databases for theorems, *Notices Amer. Math. Soc.* 60 (2013), no. 8, 1034–1039, DOI 10.1090/noti1029. MR3113227
- [3] S’ebastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, and Yi Zhang, Sparks of artificial general intelligence: Early experiments with GPT-4, 2023. arXiv:2303.12712.
- [4] Jiajie Chen and Thomas Y. Hou, Stable nearly self similar blowup of the 2D Boussinesq and 3D Euler equations with smooth data, parts I and II, 2022. arXiv:2210.07191; arXiv:2305.05660.
- [5] Johan Commelin, Liquid tensor experiment (German, with German summary), *Mitt. Dtsch. Math.-Ver.* 30 (2022), no. 3, 166–170, DOI 10.1515/dmvm-2022-0058. MR4469845
- [6] Alex Davies, Andr’as Juh’asz, Marc Lackenby, and Nenad Tomasev, The signature and cusp geometry of hyperbolic knots, 2021. arXiv:2111.15323.
- [7] Georges Gonthier, Formal proof—the four-color theorem, *Notices Amer. Math. Soc.* 55 (2008), no. 11, 1382–1393. MR2463991
- [8] W. T. Gowers, Polymath and the density Hales-Jewett theorem, *An irregular mind*, *Bolyai Soc. Math. Stud.*, vol. 21, J’anos Bolyai Math. Soc., Budapest, 2010, pp. 659–687, DOI 10.1007/978-3-642-14444-8-21. MR2815619
- [9] W. T. Gowers, Ben Green, Freddie Manners, and Terence Tao, On a conjecture of Marton, 2023. arXiv:2311.05762

works, Phys. Rev. Lett. 130 (2023), no. 24, Paper No. 244002, 6, DOI 10.1103/physrevlett.130.244002. MR4608987

[20] Kaiyu Yang, Aidan M. Swope, Alex Gu, Rahul Chalamala, Peiyang Song, Shixing Yu, Saad Godil, Ryan Prenger, and Anima Anandkumar, LeanDojo: Theorem proving with retrieval-augmented language models, 2023. arXiv:2306.15626.

press.com/2023/11/18.

[17] Trieu H. Trinh, Yuhuai Wu, Quoc V. Le, He He, and Thang Luong, Solving olympiad geometry without human demonstrations, Nature 625 (January 2024), no. 7995, 476–482, DOI 10.1038/s41586-023-06747-5.

[18] Adam Zsolt Wagner, Constructions in combinatorics via neural networks, 2021. arXiv:2104.14516.

[19] Y. Wang, C.-Y. Lai, J. G 'omez-Serrano, and T. Buckmaster, Asymptotic self-similar blow-up profile for three-dimensional axisymmetric Euler equations using neural net-

** دانشگاه یزد

آگهی

ده سری پوستر رنگی: پنج سری به قطع ۵۸×۸۸ سانتی متر به نام‌های ابوریحان بیرونی، ابوالوفا بوزجانی، ابوعبدالله محمد بن موسی خوارزمی، غیاث‌الدین ابوالفتح عمر خیام و غیاث‌الدین جمشید کاشانی و پنج سری پوستر به قطع ۴۸×۶۸ سانتی متر به نام‌های تمدن اسلامی، دوران طلایی یونان، دوران‌های اولیه، عصر نوین و نوزائی (رنسانس)، از انتشارات ستاد ملی سال جهانی ریاضیات در دبیرخانه انجمن موجود است.

این مجموعه زیبا و پرمحتوا می‌تواند زینت‌بخش کتابخانه‌ها، سالن‌ها، کلاس‌ها، اتاق‌ها و راهروهای دانشگاه‌ها، دبیرستان‌ها و مجامعی نظیر فرهنگ‌سراها و خانه‌های ریاضیات باشد.

از علاقه‌مندان، به‌ویژه مسئولان و مدیران محترم تقاضا می‌شود جهت خرید این مجموعه نفیس به این [لینک](#) مراجعه نمایند.