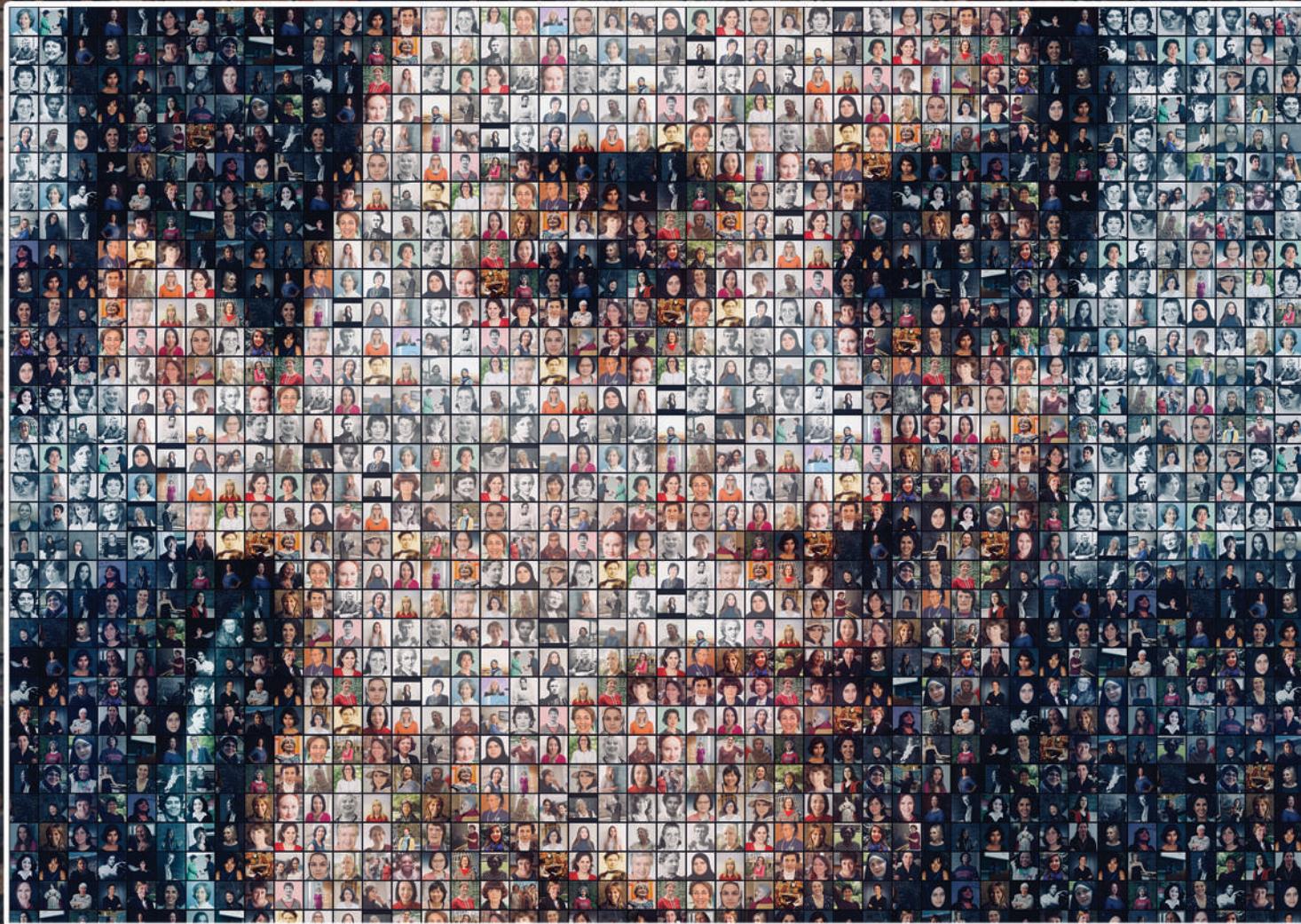


مجله‌ی ریاضی شریف

دانشجویان دانشکده‌ی علوم ریاضی دانشگاه صنعتی شریف

دوره‌ی سوم، شماره‌ی دوم، بهار ۱۴۰۳



فهرست مطالب

یادداشت‌ها

- ۱ سرمقاله
۴ ده قانون ساده برای نوشتن مقاله

مقالات

- ۱۱ اثباتی برای نامتناهی بودن اعداد اول با استفاده از نظریه اطلاعات
۱۴ بحران در مبانی ریاضیات (بخش یکم)
۲۲ ۵۰ سال پیچیدگی محاسبه (یک گزارش خبری)
۳۹ تعلی اعداد
۴۶ محاسبات چندجانبه امن
۵۳ هندسه‌ی فضای اندازه‌های احتمال
۶۹ ارجاع و وجهیت
۸۲ نسخه‌ی کواتنومی $\mathbb{N}P$
۱۱۴ شمارش قدم‌زندهای خودپرهیز روی مشکلهای شش ضلعی

صحابه‌ها و مکاتبه‌ها

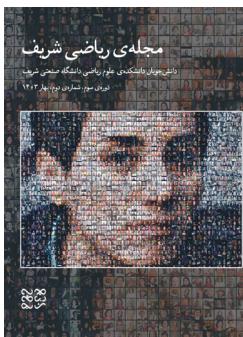
- ۱۲۱ مصاحبه با مارینا ویازوفسکا
۱۲۷ مکاتبات فرگه و راسل (بخش دوم)
۱۳۱ نامه به میر

معرفی‌ها

- ۱۳۶ رسانه‌های همگانی و ریاضیات

آموزش و مسئله‌ها

- ۱۳۹ مسائل



طرح روی جلد:

صد زن ریاضیدان؛ به مناسبت روز جهانی زن در ریاضیات.

مجله‌ی ریاضی شریف

دوره‌ی سوم، شماره‌ی دوم، بهار ۱۴۰۳



مدیر مسئول:

امیر جعفری

سردپیر:

نوید دژبرد

هیئت تحریریه:

علی الماسی

پارسا تربتی

نیکی حسنی

مینو معظمی

نویسنده‌گان و مترجمان:

امین طالبی، روزبه فرهودی، آلا جواهری، حامی عبادزاده سمنانی، امیرحسین ندیری، یاشار طالبی‌راد، سجاد طیبی، مهلا امیری، محسن امام‌وردي، بشري بصيري، على دائي نبي

طراح جلد:

بتول زرکنده

با سپاهی از:

متین انصاری‌پور، کسری علیشاھی، امیر دانشگر، محمد اردشیر، حسام رجب‌زاده، شهرام خرازی، میر یاشا هادیانی، کیمیا تیغ‌بند، حنانه مبلغ توحید، حسن خورشیدی و انجمن علمی همبند



سرمقاله

به مناسبت روز جهانی زن در ریاضیات

علی الماسی و نوید دژبرد

نوشتن یک سرمقاله در باب روز جهانی زن در ریاضیات کاری تهورآمیز است، و شاید بهتر می‌بود از این تهور پرهیز کرد اگر که مثل ما نه یک زن باشد، نه ریاضی‌دان و نه پژوهش‌گری در مسئله‌ی زنان در ریاضیات. لکن حال که پذیرفته‌ایم چنین متنی بنویسیم، می‌بایست هم شرایط ذکر شده را تصریح کنیم و هم به خواننده متذکر شویم که هدف از نگارش این سرمقاله آن است که این دیدگاه شخصی منجر به پرورش بحثی حول موضوع و طرح و شرح نقطه‌نظرات افراد مختلف شود. مجله‌ی ریاضی شریف از دریافت چنین نظراتی استقبال می‌کند و آمادگی انعکاس آن‌ها را در شماره‌ی آتی خود دارد.

راههای نسبتاً مدونی وجود دارد برای برگزیدن یک روز به مثاله‌ی روز ملی یا جهانی در یک موضوع. یکی از این راه‌ها پیوندزدن این روز با بزرگ‌داشت یک انسان است، با انتخاب زادروز یا سال‌مرگ او—روزهایی که برای هر انسانی یکه است. روز جهانی زن در ریاضیات نیز از این قاعده مستثنی نیست، و انسانی که نام و زندگی‌اش به این روز پیوند خورده مریم میرزاخانی است. بی‌شک بهترین نقطه برای شروع بحث درباره‌ی زنان ریاضیات ایران خود مریم است. شگفت‌انگیز نیست که اولین زنی که موفق به دریافت جایزه‌ی فیلدز شده، ریشه در خاک کشوری دارد که از قضا زنانش، به خاطر جنسیت‌شان، دده‌ها و سده‌های است که با مشکلاتی بسیار بنیادین دست‌وینجه نرم می‌کند؟ وقتی به مسئله‌ی زنان در ریاضیات ایران می‌اندیشیم و می‌خواهیم برای آن راه حلی ارائه دهیم، بلافاصله طومار طولی از مصاعبی که زنان ایرانی با آن مواجه‌اند پیش چشم می‌آید و آدمی را نویمد و دل‌سرد می‌کند—از زن‌کشی و کودک‌همسری گرفته تا فقدان آزادی و امنیت اجتماعی و تحمل تعیض‌های نهادی. مریم از چنین بستری برخاسته است، و این چنین امروز الهام‌بخش دخترانی شده است که می‌خواهند پیشه‌ای در علم—خاصه در ریاضیات—برگزینند. از سوی دیگر، شگفت‌انگیز نیست که مسئولین و متولیان آموزش و پژوهش در ایران، پس از مشاهده‌ی امکان چنین دستاوردهایی برای زن ایرانی، هنوز و هر ساله بسنده می‌کنند به یک مفتخرشدن تشریفاتی به ایرانی‌بودن مریم و در بهترین حالت، بر همان رویه‌ی پیشین خود در مدیریت این سرمایه‌ی علمی پافشاری می‌کنند؟ چندان هم نیست.

مریم ریشه در خاک ایران دارد، و نخستین بذرهای علاقه‌ی او به ریاضیات در بستر سیستم آموزشی ایران در او کاشته شده است؛ اما نباید فراموش کرد که نقش دوران دانش آموزی و دانش‌جویی مقطع کارشناسی در مقایسه با مابقی راهی که از یک لیسانسی ریاضی یک پژوهش‌گر تراز اول می‌سازد، ناچیز است. مریم بی‌شک ثمره‌ی ریاضیات ایران نیست، بلکه نماینده‌ای است از آن چه ریاضیات ایران می‌تواند به آن دست پیدا کند؛ مشروط به آن که بتواند و پیش‌تر از آن، بخواهد رویه‌ای متناسب با این هدف در پیش بگیرد.

یکی از دشواری‌های چنین مسائلی توافق بر صورت‌بندی آن‌هاست. هر کسی به تبع میزان حساسیت خود، وجود یک دشواره‌ی مسئله‌ساز را حس می‌کند، و به تناسب آن چه احساس کرده، تلاش می‌کند راه حلی برگزیند. بخشی از راه حل‌هایی که معمولاً ارائه می‌شود، اصطلاحاً موضعی است. راه حل‌های موضعی از این جنس‌اند که مثلاً جمعی از دختران دانشکده—و عمده‌ای هم فارغ‌التحصیل—در مواجهه با مشکلاتی که در آکادمیا با آن رویه‌رو شده‌اند و ریشه‌ی این مشکلات را در جنسیت خود دیده‌اند، به ستوه آمده و گروهی تشکیل داده‌اند به نام «دختران ریاضی شریف»، تا در شرایطی که هیچ‌کس به فکر حل این مشکلات نیست، خودشان برای خودشان کاری کنند. مثالی دیگر آن که چند سالی است که دانش‌جویان دانشکده، و به نماینده‌گی از آنان انجمن علمی دانشکده، به مناسبت روز جهانی زن در ریاضیات برنامه‌ای با هزار مکافات ترتیب می‌دهند تا از نگاه خودشان به بررسی حضور زنان در ریاضیات ایران پردازنند؛ اما بحث درباره‌ی مسئله‌ای چنین فراگیر و روزمره عملاً محدود می‌شود به این برنامه‌ی مناسبتی سالانه. این چنین راه حل‌هایی کوتاه‌مدت‌اند و حکم مدیریت بحران دارند، و منتج به تغییر رویکرد

مسئولین نمی‌شوند. افراد دست‌اندرکار در این NGOs‌ها دل‌مشغولی و وظیفه‌ی اصلی‌شان کاری نیست که در سازمان مربوطه بر عهده دارند، و معمولاً مشکلاتی که پیش پایشان قرار می‌گیرد چنان فراتر از ظرفیت‌شان است که به سالی نکشیده — و حتی پیش از اتمام دوره‌ی مسئولیت‌شان — سرشکسته و دلزده می‌شوند. معمولاً هم اضافه می‌شوند به تلنبار افراد «باتجربه‌ای» که معتقد‌نند نمی‌شود کاری از پیش برد، و در اشاعه‌ی یأس حاکم سهیم می‌شوند.

یک ایراد بزرگ دیگر راه حل‌های NGOs‌محور این است که مسئول مربوطه با خیالی راحت از این که این مشکل فعلًا ختم به یک بحران نمی‌شود، حل آن را در اولویت خود قرار نمی‌دهد. انجمن علمی دانشکده را در نظر بگیرید. بخشی از اصلی‌ترین اهداف و انگیزه‌های موارد متعددی از فعالیت‌های انجمن چنین شکل می‌گیرد: فقدانی در سیستم آموزشی دانشکده وجود دارد، و دانش‌جویان می‌خواهند راه حلی برای برطرف کردن آن بیابند. مداوای اساسی همواره دور از دسترس دانش‌جویان باقی می‌ماند، و از سر ناچاری و حتی بی‌حوصلگی به تسکینی کفایت می‌شود. حال فراموش نکنید که دهه‌هاست این مشکلات پابرجا هستند، و متعاقباً دهه‌هاست که نامزدهای انتخابات انجمن علمی در ایام انتخابات به مشکلات یکسانی نشانی می‌دهند. همه‌ی این‌ها را بگذارید کنار این که در همین فعالیت‌های انجمن هم خود مسئولین مربوطه کم سنگ‌اندازی نمی‌کنند. به طور خلاصه، راه حل‌های NGOs‌محور یک کنترل بی‌جیره و مواجب بحران است که در بلندمدت تنها کسانی که از آن منتفع می‌شوند، مشتبی مسئول نالایق و بی‌توجه است. هرچند نمی‌بایست انکار کرد که چنین تلاش‌های نسبتاً مستمری — که گاهی تن به مبارزه می‌زند — در تطور فضای فرهنگی حول این مسئله نقش مهمی دارند؛ ولوزوماً به تغییری در سطح سیاست‌گذاری نینجامد. پویایی این فضای فرهنگی می‌تواند سنگ بنای آن سیاست‌گذاری‌ای باشد که نسل‌های متمادی ای از افراد در پی دست‌یابی به آن بوده‌اند. همه‌ی آن‌چه گفته شد، درباره‌ی کسانی است که وجود مسئله را احساس می‌کنند. افرادی هستند که وجود هیچ دشواره‌ای را به رسمیت نمی‌شناسند یا محض حفظ ظاهر چنان صورت‌بندی‌ای از آن ارائه می‌دهند که با هر سنجه‌ی سالمی دور از واقعیت است. دست بر قضا این افراد با شانس بیش‌تری هم به مسئولیت‌های اجرایی و سیاست‌گذاری دست می‌یابند. ارائه‌ی راه حل بنیادین مستلزم به رسمیت شناختن وجود مسئله و توافق بر صورت‌بندی‌های صحیح آن است. در کنار همه‌ی فعالیت‌های موضعی که لنگ‌لنگان بار مسئله را به دوش می‌کشند، گام دل‌گرم‌کننده و از حیث اجرایی مهمی است اگر مسئولین دانشکده — خاصه ریاست دانشکده و معاونت فرهنگی و دانشجویی — کمیته‌ای تشکیل دهند که مشکلات زنان و دختران در دانشکده را شناسایی کند. کمیته‌ی مزبور با این حال، نباید به لیست بلندبالای کمیته‌های فرمایشی دیگری افزوده شود که نقشی جز فریه‌کردن عناوین و مسئولیت‌های مسئولین نالایق ندارند. مسئله‌ی حضور زنان در علم مسئله‌ای است که در حال حاضر همه‌ی دنیا در حال تلاش برای حل آن هستند. ابوهی از مطالعات و سیاست‌گذاری‌ها در دسترس است تا بتوان در آغاز راه آن‌ها را سرلوحه قرار داد — سیاست‌گذاری‌هایی که فقط به حالت کلان و مقیاس بزرگ اختصاص ندارند و به سادگی می‌توان نمونه‌های با مقیاس کوچک و دانشکده‌ای آن را هم پیدا کرد. مضحك خواهد بود اگر کمیته‌ی مزبور تماماً متشکل باشد از اعضای هیئت‌علمی دانشکده‌ای که هیچ عضو هیئت‌علمی زنی ندارد. به علاوه، دانش‌جویان نیز باید بتوانند هم در مقام عضو و هم در مقام مشاور کمیته‌ی مربوطه فعالیت کنند، و البته اگر عضو هیئت‌علمی با عضویت در این کمیته به حقوق‌ش اضافه می‌شود، به دانشجوی عضو نیز باید حقوق پرداخت شود.

با همه‌ی این‌ها روشن است که تضمینی نیست که مسئول مربوطه در راستای بهبود شرایط فعلی قدم از قدم بردارد. ما چه می‌توانیم بکنیم؟ به باور ما در عصری که دسترسی به اخبار و اطلاعات به صورت فراگیر میسر است، اولین وظیفه‌ی همگانی اطلاع‌رسانی گسترده‌ی مصادیق معضلات جنسیتی مربوط به فضای آکادمیک دانشکده است، و البته می‌دانیم که در شرایط فعلی این کار ممکن است ساده نباشد. استاد خانمی که حقوق‌ش از همکار هم‌رده‌ی آفایش کمتر است، دانشجویی که مورد آزار جنسی یا کلامی توسط استاد درس یا استاد راهنمایش قرار می‌گیرد و دانشجوی دختری که احساس می‌کند فرصت‌هایی علمی را به خاطر جنسیت‌ش از دست می‌دهد، تنها راهی که پیش پای خود می‌بیند این است که این موارد را به طور همگانی بازگو کند — خصوصاً در شرایطی که دانشکده مسئول منابع انسانی ندارد. امید آن است که در نتیجه‌ی این اطلاع‌رسانی‌ها مسئول مربوطه ضرورتی بیابد در این‌که به اذهان عمومی پاسخ‌گو باشد و نتیجتاً رویه‌ی خود را تغییر دهد.

اگر چشم‌هایتان را بیندید، چیزی نخواهیدید؛ اگر گوش‌هایتان را بگیرید، به میزان قابل توجهی کمتر خواهیدشند؛ اما ممکن نیست بتوانید لامسه‌تان را به نحوی خاموش کنید. در برابر آن‌چه نیمی از جمعیت جامعه، با پوست و گوشت و به ملموس‌ترین نحو ممکن احساس می‌کنند، آن‌چه یک مرد از مشکلات زنان می‌داند در حد دیده‌ها و شنیده‌های است، و ظالملانه است اگر آن نیم مردانه بخواهد در همین اندازه هم که شده توجهی نکند. ایجاد بستری برای ارائه‌ی روایت‌های شخصی و جمیعی، کم‌ترین کاری

است که می‌توان انجام داد تا زیسته‌ها و دیده‌ها تبدیل به گفته‌هایی رسا و شنیده شوند. انتخاب زادروز مریم میرزاخانی به عنوان روز جهانی زن در ریاضیات انتخابی هوشمندانه بود. چنین روزی نه تنها در حکم یادآوری اهمیت حضور فraigیر زنان در این حیطه است، بلکه باید بدل به فرصتی شود که رشد و سلامت این وجود نیز هر سال وارسی شود—بهسان یک روز تولد واقعی. به امید آن که رسمیت این مسئله از رخدادی یکروزه در سال فراتر رود و به عنوان تلاشی ممتد در طول سال جا بیفت؛ بلکه به این طریق بتوانیم در روزهای آتی دانشکده شاهد این باشیم که مریم‌های بعدی با مشکلات کمتری به مراتب بالای علمی و پژوهشی دست پیدا می‌کنند.



ده قانون ساده برای نوشتن مقاله

برت منش و کنراد کردینگ

چکیده. نوشتن توانایی مهمی برای پیشرفت علم و رشد در آن است. یک متن سلیس و قوی می‌تواند خواننده را درباره‌ی موضوعی به وجود بیاورد، منتقدان را علاقه‌مند به بررسی دقیق آن کند و در نهایت در جامعه‌ی علمی تاثیرگذار باشد. با این وجود اکثر دانشجویان آموزشی در این زمینه نمی‌بینند و مجبورند ممارست زیادی برای نوشن مقاله‌ی خود کنند. در این نوشته از زاویه‌ی دید یک خواننده و یا داور مقاله ده قانون ساده را برای نوشتن مقاله بیان می‌کنیم. این قوانین کمک می‌کنند که سریع‌تر و سلیس‌تر بنویسید و از نوشتن لذت ببرید.

مقدمه

خواندن و نوشتن توانایی مهمی برای دانشجویان است. در واقع ماشین علم بر پایه‌ی تولید مقاله بنا شده است [۱، ۲]. انگیزه‌ها و اولویت‌ها در کنار مولفه‌های بسیار دیگری در مقاله بیان می‌شوند. سردبیران یک مجله‌ی علمی می‌خواهند ارزش علمی مقاله را بفهمند تا با رای داوران درباره‌ی صحت نتایج تصمیم به نشر آن بگیرند. خواننده‌ی مقاله می‌خواهد نتایج آن را بفهمد تا تصمیم بگیرد که آن را دقیق‌تر بخواند یا خیر. در نهایت نویسنده‌ی مقاله می‌خواهد مطمئن شود که پیام اصلی مقاله به گوش مخاطب عام رسیده و مخاطب متخصص متوجه اهمیت آن شده است. تمام این اهداف با داشتن یک ساختار مناسب در مقاله قابل دسترسی است. ساختاری که جمله‌ها را در پاراگراف، بخش و نهایتاً کل مقاله سازماندهی کند.

شسته‌رفته‌بودن نوشه برای ترویج علم در شاخه‌های مختلف بسیار حیاتی است. برای مثال در علوم زیستی که موضوعات آن با زمینه‌های متنوعی درگیر است، نوشتن متن به گونه‌ای که محققان رشته‌های مرتبط ولی متفاوت پیام آن را بگیرند اهمیت بسیار دارد. این کار نیاز به مهارت نویسنده دارد تا بتواند مفاهیم را به درستی به یکدیگر ارتباط دهد. یک مقاله تنها وقتی مفید است که خوانا و قابل ارزش‌گذاری باشد و نتایج آن به راحتی به خاطر سپرده شوند.

گزاره‌هایی که در مقاله بیان می‌شوند باید منطقاً درست و با داده‌ها هم‌خوانی داشته باشند. بدون ساختن یک داستان اصلی برای مقاله خواننده معمولاً در اینوه داده‌ها گم می‌شود و نمی‌تواند خط فکری شما را دنبال کند تا متوجه نتایج آن شود. در اینجا ده قانون ساده برای نوشتن مقاله بیان می‌کنیم. چهار قانون اول کلی هستند و در تمام قسمت‌های مقاله (و نوشه‌های دیگر مانند پیشنهاد پژوهشی، پوستر و ...) باید رعایت شوند. چهار قانون دیگر به شما می‌گویند چگونه نتایج هر قسمت از مقاله را بیان کنید و در نهایت دو قانون آخر می‌گویند که چگونه بنویسید تا اهداف اصلی کل مقاله روشن شود.

قانون‌های کلی (قانون‌های ۱ تا ۴)

نوشتن ابزاری برای ارتباط برقرارکردن است. بنابراین ارتباط قوی با ذهن خواننده اولویت اصلی یک مقاله است. هنگام نوشتن مرتباً باید مطمئن شوید که خواننده همراه شماست. ۴ قانون پیش رو برای اطمینان یافتن از همراهی خواننده‌ی مقاله شما است.

قانون ۱: تنها و تنها پیرامون موضوعی که در عنوان مقاله بیان می‌کنید بنویسید. مقاله‌ای ماندگار است که خواننده یک سال بعد از خواندن آن همچنان بتواند دستاورد اصلی آن را بیاد بیاورد. لازمه این کار، تمرکز مقاله تنها روی یک موضوع است.

* این نوشته ترجمه‌ای از مقاله‌ای زیر است:

اگرچه یک مقاله می‌تواند نتایج زیادی را بیان کند ولی نیازی به حرص‌زدن برای این کار نیست. هنگامی که چند موضوع متفاوت لابه‌لای مقاله بیان می‌شود دستاورده اصلی آن در ابهام می‌ماند و احتمالاً خواننده خیلی زود تمام آن‌ها را از یاد می‌برد. مهمترین قسمت یک مقاله عنوان آن است. برای درک اهمیت آن ببینید که در چند درصد موارد تنها با خواندن عنوان مقاله اراده کردید آن را بخوانید. عنوان اولین چیزی است که دیده می‌شود و در تصمیم خواننده برای خواندن کل مقاله نقش اصلی را دارد [۳].

عنوان نه تنها ایده‌ی اصلی مقاله را می‌رساند، بلکه مانند قطب نمایی برای کل متن آن است. با نگاه به عنوان و مقایسه‌ی آن با قسمت‌های مختلف مقاله می‌توان حواشی و کاستی‌های مقاله را تشخیص داد. عنوان معمولاً باید کوتاه باشد. اگر نتایج زیادی برای گفتن دارید پیدا کردن عنوانی کوتاه کاری سخت ولی ارزشمند است؛ زیرا هدف علم به دست آوردن اصول کوتاه و عمیق با تکیه بر انبوه داده‌ها و قضیه‌ها است. فکر کردن مداوم به عنوان تحقیق علمی‌тан نه تنها به نوشتن بهتر آن کمک می‌کند بلکه ذهن شما را شفاف می‌کند تا آزمایش مناسب‌تری طراحی کنید و یا نظریه‌ی عمیق‌تری بسازید.

این قانون شاید سخت‌ترین قانون باشد زیرا با تمام پروژه‌ی علمی شما درگیر است. در نهایت باید با تحلیل و استدلال داده‌ها ساده‌ترین ادعا/مدل را کشف کنید که البته قابل ساده‌سازی بیشتر نباشد. نتیجه‌ی این تلاش یافتن عنوانی است که تمام نتایج تحقیق‌تان را در یک جا جمع می‌کند. برای مثال عنوان مقاله‌ای که درباره‌ی یک فن آوری جدید و نتایج زیستی آن نوشته شده است، باید همچون پلی باشد که توصیف کند چگونه فن آوری جدید می‌تواند زیست‌شناسی را دگرگون کند.

قانون ۲: برای کسی بنویسید که اطلاعی راجع به رشته‌ی شما ندارد. نویسنده‌ی مقاله آگاه‌ترین فرد به زیرویم تحقیق علمی و در عین حال ناآگاه‌ترین فرد برای درک ذهن خواننده‌ی آن مقاله است. این ناآگاهی ریشه‌ی بسیاری از اشتباهات نوشته‌یاری است. برای پرهیز از آن‌ها، مثل یک طراح به کار خود نگاه کنید: فضایی بسازید و ذهن خواننده را به سوی هدف خود سوق دهید [۴]. به نوشته‌تان دوباره (و چندباره) نگاه کنید مرتب از خود پرسید آیا ارزشی دارد که خواننده به مسائله‌ای که شما در اینجا می‌گویید فکر کند یا نه (قانون ۶). اگر احساس کردید خواننده را به مساله علاقه‌مند کرده‌اید باید پاسخ‌تان را طوری بیان کنید که خواننده با کمترین تلاش ذهنی آن را بفهمد.

وازگان تخصصی را با وضوح برای خواننده تشریح کنید چون ندانستن معنی درست یک کلمه خواننده را گیج و خسته می‌کند. از خلاصه‌سازی لغات و یا ساختن بی‌مورد سروازه پرهیز کنید چون خواننده مجبور می‌شود برای یاد آوری آن مرتب مقاله را پایین و بالا کند.

دانش روان‌شناسی به نوشتن بهتر مقاله کمک می‌کند. مثلاً حافظه‌ی کوتاه‌مدت انسان تنها تعداد محدودی موضوع را به طور همزمان می‌تواند ذخیره کند که اولین و آخرین آن‌ها بهتر از بقیه به یاد می‌آیند [۵]. سعی کنید در هر جای مقاله استفاده از حافظه‌ی خواننده بسیار سبک باشد.

قانون ۳: متن مقاله را به شکل مقدمه-تحقیق-نتیجه (متن) بنویسید. تقریباً تمام داستان‌ها ساختار ثابتی دارند: توصیف فضای داستان (مقدمه)، حادثی که رخ می‌دهد (تحقیق) و حرف نهایی (نتیجه). مقدمه فضای داستان را آماده می‌کند که در بستر آن حوادث مختلف رخ می‌دهد و نهایتاً داستان به نتیجه‌ی غایی می‌رسد. این ساختار کمک می‌کند تا خواننده نه تمرکز خود از خط اصلی داستان را از دست بدهد و نه خسته شود.

حسب اینکه خواننده چه قدر قرار است درگیر موضوع شود استراتژی‌های مختلفی برای بیان یک موضوع وجود دارد [۶]. خواننده‌ی کم‌حصوله باید به سرعت درگیر موضوع شود، مثلاً با تهییج او در اول کار مانند گزارش‌های خبری. ساختار متن که در اینجا بیان کردیم برای خواننده‌ی صبوری مناسب است که می‌خواهد درگیر جزئیات داستان شود. این روش برای درگیر کردن خواننده‌ی کم‌حصله کارآمد نیست. هرچند در یک مقاله‌ی علمی این دغدغه‌ی مهمی نیست زیرا در عنوان و مقدمه‌ی مقاله قسمت‌های مهیج گفته شده‌اند، و در نتیجه فردی که درگیر خواندن متن مقاله است به اندازه‌ی کافی به موضوع علاقه‌مند شده است. علاوه بر این در علم، بیان یک نظریه‌ی جدید بدون هیچ مقدمه‌ای، باعث می‌شود که خواننده به حرف شما مشکوک شود. به خصوص اگر خواننده قسمت مهمی از نظریه‌ی شما را درست نفهمیده باشد. متن در مقیاس‌های مختلف در یک مقاله باید رعایت شود. در بزرگ‌ترین مقیاس، اگر تمام مقاله را به عنوان یک داستان در نظر بگیریم، از سه قسمت مقدمه^۱، نتیجه‌های^۲

¹introduction

²results

تحقیق و نتیجه‌گیری نهایی آن در توضیحات^۱ تشکیل شده است. در مقیاس کوچک‌تر هر پاراگراف مقاله باید به شکل متن بیان شود. جمله‌ی اول موضوع آن پاراگراف را بیان میکند که مقدمه‌ی آن است. در ادامه مفروضات، شواهد و دست آوردها قسمت تحقیقی پاراگراف را شکل می‌دهند و نهایتاً جمله‌ی آخر پاراگراف نتیجه‌ی آن را می‌گوید.

رعایت نکردن ساختار متن خواندن مقاله را سخت می‌کند. متأسفانه اگر فی‌الدعاhe شروع به نوشتن کنیم به طور ناخودآگاه این ساختار را رعایت نمی‌کنیم؛ زیرا یک محقق بیشتر وقت خود را صرف تحقیقات علمی مانند انجام آزمایش، واکاوی مقالات محققان قبلی و فکر کردن به ایده‌های جدید با ذهن خلاق خود می‌کند و کمتر روی سازماندهی آن‌ها وقت می‌گذارد. بنابراین طبیعی است که اگر می‌خواهد دست آوردهای خود را ثبت کند از تایم متأخر تحقیقات خود شروع کند — که برای یک خواننده‌ی ناآشنا ارتباط برقرار کردن با آن بسیار سخت است. خواننده نمی‌خواهد تایم تحقیق را با همان سلسه‌ی زمانی که منجر به کشف آن شده است بداند؛ بلکه می‌خواهد پیام اصلی مقاله را بفهمد و بداند چگونه این پیام با گزاره‌های منطقی و حقایق اثبات می‌شود. بنابراین تمام تلاش ما هنگام نوشتن باید معطوف به نظم بخشی به دست آوردهای علمی، اتصال آن‌ها به یک دیگر و استخراج تایم قابل فهم و ماندگار باشد.

قانون ۴: از ذکر یک موضوع در میانه‌ی موضوع دیگر پرهیز کنید و موازی بنویسید تا نوشته‌ی شما سلیس‌تر شود. پرهیز از ذکر یک موضوع در میانه‌ی موضوعی دیگر: تها عنوان مقاله است که می‌توان مرتباً به آن ارجاع داد و درباره‌ی هر موضوع دیگری بهتر است تنها یک بار بحث شود. جملاتی که با هم اشتراک دارند باید مکمل هم باشند تا یک کاسه شوند و تنها یک بار بیان شوند. به همین شکل بهتر است ایده‌هایی که مشابهت دارند، مانند دو دلیل برای درستی یک نظریه، بلاfaciale بعد از یکدیگر بیان شوند.

موازی نویسی: پاراگراف‌ها یا جملاتی که پیام‌های مشابهی دارند باید شبیه هم دیگر نوشته شود زیرا کار خواندن را راحت‌تر می‌کنند. به عنوان مثال اگر سه دلیل متفاوت برای برتری یک نظریه بر نظریه‌ی دیگر وجود دارد، مناسب‌تر است که جمله‌بندی هر سه به یک شکل نوشته شود تا خواننده با کمترین انرژی، توجه خود را صرف فهم دلایل کند. هیچ اشکالی ندارد که یک کلمه بارها و بارها در جملات و پاراگراف‌ها بکار برود. وسوسه نشوید که کلمات متفاوتی را برای یک مفهوم به کار ببرید؛ زیرا در این صورت این پیام گمراه‌کننده را به خواننده می‌فرستید که شاید معنای این کلمات اندکی با هم متفاوت است.

متن مقاله (قانون‌های ۵ تا ۸)

قسمت‌های مختلف مقاله — چکیده، سرآغاز، یافته‌ها و بحث — اهداف متفاوتی دارند و برای نیل به این اهداف در کنار رعایت ساختار متن، نکات دیگری نیز باید مورد توجه قرار گیرد. در اینجا به این نکات می‌پردازیم. برای جمع بندی این بخش، شکل ۱ را در خاطر داشته باشید.

قانون ۵: کار خود را به طور کامل در چکیده بنویسید. اکثر خوانندگان یک مقاله تنها چکیده‌ی آن را کامل می‌خوانند. بنابراین باید به شکلی منسجم پیام مقاله در آن رسانده شود. برای این کار متن آن را تشریح می‌کنیم. در مقدمه‌ی چکیده خواننده باید متوجه شود که مقاله‌ی شما چه نکته‌ی مغفوی از علم را روشن می‌کند. جمله‌ی اول ذهن خواننده را به سمت فضای رشته‌ی علمی مقاله می‌برد. در جملات بعد در هر مرحله مانند قیفی این فضا کوچک‌تر می‌شود تا به مساله‌ی خاصی که شما بررسی کرده‌اید، برسد. به خواننده بگویید که چه چیزی در تحقیقات فعلی پنهان‌مانده و یا ناتص است و چرا کارکردن روی این کاستی ارزشمند است (مثالاً با ارتباط مناسب به رشته‌ی علمی در جمله‌ی اول مقاله). جمله‌ی اول تحقیق چکیده (در اینجا می‌خواهیم^۲) روش شما را برای مرتفع کردن آن نقص و جملات بعدی دستاوردهای مرتبط شما را بیان می‌کند. نهایتاً در قسمت اول نتیجه‌ی چکیده، تفسیر دستاوردهای خود را در رشته‌ی تحقیقاتی می‌گویید و در قسمت دوم نشان می‌دهید چگونه کار شما باعث می‌شود آن رشته یک گام به پیش رود. قسمت دوم برای مجلاتی که خوانندگان متعددی دارند، بسیار مهم است.

این نحوه‌ی نوشتن به شما کمک می‌کند تا مرتکب بسیاری از اشتباهات نوشتاری نشود — مانند شروع به توضیح دست آوردها بدون بیان پیش‌زمینه‌ی مناسب. یک چکیده‌ی خوب به بارها بازیبینی و تغییر نیاز دارد تا در نهایت به خوبی نشان دهد مقاله

¹ discussion

² Here we



شکل ۱: خلاصه‌ی ساختار مقاله در سه مقیاس: بین بخش‌ها، بین پاراگراف‌ها و در داخل پاراگراف‌ها. متناسب با مقدمه، تحقیق و نتیجه متن چکیده هر سه رنگ را دارد.

چگونه در فضای آن رشته‌ی علمی جای می‌گیرد. استفاده از ساختار قیفی کمک می‌کند که مقاله با خواننده‌های مختلفی ارتباط برقرار کند و ارزش علمی آن برجسته شود.

قانون ۶: در مقدمه‌ی مقاله اهمیت آن را نشان دهید. هدف از مقدمه‌ی مقاله، برجسته‌کردن خلا موجود در دانش فعلی و رهیافت شما برای پرکردن آن است. برای این کار نیاز به نوشتن چندین پاراگراف است که به طور ملموس نشان دهد چه چیزی در آن رشته مغفول مانده و در نهایت با جمع‌بندی نهایی نشان دهد آن خلا علمی چگونه در این مقاله پر می‌شود.

با یک مثال نشان می‌دهیم چگونه می‌توان مقدمه‌ی مقاله‌ای در زمینه‌ی ژنتیک راجع به تقسیم سلولی نوشت. پاراگراف اول توضیح می‌دهد چرا درک فرایند تقسیم سلولی موضوعی داغ در زیست است که هنوز به درستی شناخته نشده است (خلاً در رشته‌ی اصلی). پاراگراف دوم توضیح می‌دهد چه چیزهایی درباره‌ی تقسیم سلولی استرسایتی — نوعی سلول خاص — ناشناخته است (خلاً در زیررشته‌ی رشته اصلی). پاراگراف سوم سرنخ‌هایی درباره‌ی یک ژن خاص که باعث تقسیم سلولی در استرسایتی‌ها می‌شود به دست می‌دهد ولی بیان می‌کند که این موضوع همچنان تایید نشده است (خلاً موجود در زیررشته که مقاله‌ی شما قرار است آن را پر کند). در هر مرحله که جمله‌ای راجع به نادانسته‌ای در یک زمینه گفته می‌شود خواننده حدس بهتری راجع به موضوعی که شما می‌خواهید در مقاله‌تان بحث کنید می‌زند.

هر پاراگراف مقدمه (به جز آخرین پاراگراف) معطوف به یک خلاً در علم است. مقدمه‌ی یک جمله‌ای هر پاراگراف خواننده را به آن خلاً رهنمون می‌کند. در جملات بعد مروری بر تحقیق‌های روز در آن زمینه می‌کند تا نشان دهد چه چیزهایی دانسته شده است. در نهایت با ذکر موضوعی مهم ولی مغفول در این تحقیق‌ها و تاثیر احتمالی مقاله در حل آن به خواننده کمک می‌کند تا جایگاه مقاله را درک کند. درین این خطوط اغلب می‌توان علت‌یابی کرد که چرا موضوعی مغفول مانده است تا سر نخی به خواننده دهد که چگونه می‌توان آنها را حل کرد. هیچ نیازی نیست که مقاله‌ی شما مرور مفصلی بر تمام تحقیق‌های پیشین کند. تنها آن‌هایی را بیان کنید که مستقیماً در راستای مقاله‌ی شما است. اگر این ساختار به خوبی رعایت شود، خواننده می‌تواند متخصص به سرعت می‌تواند متوجه جایگاه مقاله شود.

با این حال، پاراگراف آخر قسمت مقدمه متفاوت است: در آن به طور خلاصه یافته‌های مقاله در راستای پرکردن خلاً علمی بیان می‌شوند. این پاراگراف شبیه چکیده خواهد بود، با این تفاوت که مقدمه‌ی چکیده را ندارد، یافته‌های مقاله را دقیق‌تر بیان می‌کند و در پایان (و در صورت نیاز) عصاره‌ی مقاله را بازگو می‌کند.

قانون ۷: نتیجه‌های خود را با کمک نمودارها و استنتاج منطقی پلهمله پیش ببرید تا به هدف مقاله برسید. در قسمت نتیجه‌ها شما باید خواننده را درباره‌ی ادعایتان قانع کنید. ابزار شما منطق و داده‌ی خام است. اثبات هر ادعای علمی نیاز به ارائه‌ی دنباله‌ای از گزاره‌ها دارد تا صحت تک‌تک اجزای آن مشخص شود. مثلاً اگر حدسی را بررسی می‌کنید ابتدا به خواننده نشان دهید که متغیرهای آن را به درستی اندازه‌گیری می‌کنید و سپس آن‌ها را برای درستی حدس تحلیل کنید. یا اگر فرضیه‌های مختلفی را برای توضیح یک پدیده بیان می‌کنید با دلیل اثبات کنید که همه‌ی آن‌ها اشتباه هستند به جز یکی. همواره به یاد داشته باشید که ارائه‌ی یک دلیل نیازمند داشتن یک روش علمی و انجام آزمایش‌های کنترل است.

هنگامی که شکل کلی مقاله را آماده می‌کنید (قانون ۹)، به طور خلاصه گزاره‌هایی را که برای اثبات ادعایتان مورد نیاز است استخراج و آن‌ها را به شکل سلسه‌ای از جملات بنویسید. این جملات در آینده عنوان‌سrfصل‌ها، بخش‌ها و نمودارهای مقاله‌ی شما خواهند شد. اکثر مجلات چنین ساختاری برای عنوان‌بندی دارند. حتی اگر مجله‌ای اینگونه نبود، می‌توانید با بسط هر یک از این جملات در یک یا چند پاراگراف، مقاله را به سرعت آماده کنید و قبل از ارسال کردن آن به مجله، همه را پاک کنید. داشتن مجموعه‌ای از گزاره‌های منطقی و شفاف کمک می‌کند تا خواندن مقاله بسیار سرراست شود.

به طور خاص نمودارها، عنوان و شرح آن‌ها بسیار مهم هستند زیرا با رجوع به داده‌ها پلهمله قدم‌های تحقیق را تا اثبات ادعای اصلی نشان می‌دهند. علاوه بر این، اکثر خواننده‌های مقاله بعد از خواندن چکیده به صورت گذرا نمودارها و عکس‌های آن را می‌بینند. عنوان یک نمودار باید نتیجه‌ی یک آنالیز و شرح داخل آن نحوه‌ی انجام آن آنالیز را نشان دهد. در نهایت به یاد داشته باشید که درست کردن نمودار و عکس به خودی خود یک هنر است و کتاب ادوارد تافت^۱ مرجعی مناسب برای درست کردن شاھکار هنری شماست [۲، ۸].

ساختار پاراگراف اول قسمت نتیجه‌ها متفاوت است؛ در آن رهیافت مقاله به مساله‌ای که در قسمت مقدمه گفته شده همراه با اشاره‌ای به روش‌های نوآورانه‌ی مقاله برای پاسخ به آن بیان می‌شود. اکثر خواننده‌ها قسمت روش مقاله^۲ را نمی‌خوانند و این پاراگراف فرصتی است تا چکیده‌ی آن بازگو شود. دیگر پاراگراف‌های قسمت نتیجه با عبارتی یک یا دو جمله‌ای شروع می‌شوند که به طور ضمنی سوالی را بیان می‌کنند که قرار است در پاراگراف پیش رو به آن پاسخ داده شود. سوالاتی از این دست: «برای اینکه مطمئن شویم عامل ناخواسته‌ای در آزمایش وجود ندارد»، «اندازه‌گیری‌های ما تا جایی دقیق هستند که»، «سپس ما بررسی کردیم که آیا کلسیم از طریق دریچه‌های شکل وارد می‌شود یا خیر». در میانه‌ی پاراگراف، داده‌ها و منطق جواب‌دهی به سوال

¹Edward Tufts

²method section

گفته می‌شود تا در جمله‌ی آخر جواب نهایی بیان گردد. مثلاً در جمله‌ی آخر گفته می‌شود: «نتیجه می‌گیریم که هیچ عامل ناخواسته‌ای در آزمایش وجود ندارد». این نحوه‌ی نوشتمن به خواننده‌ی مجرب کمک می‌کند تا درستی تکنک نتایج را چک کند. هر پاراگراف خواننده را مقاعد می‌کند که جواب داده شده در جمله‌ی آخر درست است. با این روش می‌توان گزاره‌های مشکوکی را که باعث خدشه به نتیجه‌ی نهایی می‌شود به راحتی پیدا کرد. مانند قضایای ریاضی، نتیجه‌ی هر پاراگراف می‌تواند در پاراگراف‌های بعد از آن مورد استفاده قرار گیرد.

قانون ۸: توضیح دهید که چگونه یک خلاً علمی را پر کرده‌اید، محدودیت‌ها و تاثیرات آن چیست و جایگاه آن در مقالات مشابه کجاست. قسمت توضیحات مقاله نشان می‌دهد چگونه یک خلاً علمی که در قسمت مقدمه بیان شد اکتون برطرف شده است، محدودیت و اشکالات روش ارائه شده چیست و چگونه این مقاله می‌تواند به پیش‌برد شاخه‌ی علمی مورد بحث کمک کند. پاراگراف اول قسمت توضیحات متفاوت است چون در آن خلاصه‌ی یافته‌ها بیان می‌شود. این به خواننده‌هایی که قسمت نتیجه‌ها را نخوانده‌اند کمک می‌کند تا شرح مختصری از آن را بدانند. از آن به بعد هر پاراگراف با بیان نقص، ضعف یا اهمیت روش مقاله شروع می‌شود، به واکاوی آن در تحقیقات مشابه می‌پردازد و در پایان به شکلی خلاصه‌وار و نوآورانه بیان می‌کند که چگونه می‌توان تاثیر این کار را در آینده دید و یا نقص بیان شده را برطرف کرد.

مثالی از ساختار قسمت توضیحات می‌زنیم. پاراگراف اول خلاصه‌ی نتیجه‌ها را می‌گوید. پاراگراف دوم تا چهارم ضعف و کاستی‌های مقاله را می‌گویند و اشاره می‌کند که چگونه انجام آزمایش‌های بیشتر می‌تواند آن‌ها را بهبود ببخشد. از پاراگراف پنجم به بعد نشان داده می‌شود که چگونه این مقاله می‌تواند یک قدم زمینه علمی مورد بحث آن را به پیش‌برد. خواننده قدم به قدم بهتر متوجه می‌شود که نتیجه‌ی مقاله چیست و جایگاه آن در تحقیقات مشابه کجاست.

پیش‌برد مقاله (قانون ۹ و ۱۰)

برای نوشتمن خوب بهتر است عادت‌هایی داشته باشیم. بعضی از جنبه‌های مقاله اهمیت بیشتری دارند که متناسب با آن بهتر است زمانی بیشتری صرف آن‌ها شود. علاوه بر این، گرفتن بازخورد مقاله از همکاران و دوستان کمک می‌کند تا داستان مقاله بهتر شکل بگیرد و نوشته پخته‌تر شود.

قانون ۹: متناسب با اهمیت عنوان، چکیده، نمودارها و سرفصل‌ها زمان اختصاص دهید. روشن شدن کلام اصلی مقاله مهم‌تر از هر چیزی است. این کلام از طریق ساختن پلی میان آزمایش‌هایی که انجام گرفته و نوشتمن مقاله شکل می‌گیرد. بنابراین، بهتر است که به طور روزمره نتیجه آزمایش‌ها و علت انجام آن‌ها در جایی یادداشت شود (مثلاً در نشسته‌های هفتگی گروه) تا به مرور زمان بدنی اصلی مقاله شکل بگیرد. بهتر است متناسب با اهمیت هر قسمت از تحقیق روی آن وقت بگذارد. عنوان مقاله، چکیده و نمودارها بیننده‌ی بیشتری دارند و در عوض قسمت‌روش‌ها خواننده‌ی کمتر. متناسب با آن زمان خود را تقسیم کنید. با تعمق کافی پیش از نوشتمن هر قسمت می‌توانید زمان نوشتمن آن را کاهاش دهید. ابتدا سرفصل‌ها را درست کنید و برای هر قسمت از آن جمله‌ی ساده‌ای حتی محاوره‌ای بنویسید. ساده‌تر است ابتدا به توضیح مستقل تک‌تک نتیجه‌ها پردازید. از دل این توضیح عنوان آن قسمت به دست می‌آید. وقتی خط اصلی مقاله مشخص شود هر پاراگراف ذره‌ای از آن را شکل می‌دهد. قانون ۹ کمک می‌کند تا وقت خود را بهبوده صرف ویرایش و پیدا کردن کلمات در پاراگراف‌هایی که محتملاً حذف می‌شوند نکنید.

قانون ۱۰: بازخورد دیگران کمک می‌کند که داستان مقاله را ویرایش و یا از نو بازنویسی کنید. نوشتمن مثل یک مساله‌ی بهینه‌سازی است که هدف آن یافتن بهترین داستان، عنوان و جمله‌ها است. به همین دلیل بهتر است خیلی در ویرایش چرک‌نویس خود غوطه‌ور نشوید؛ زیرا در اغلب اوقات دوربین‌ختن کل پاراگراف و بازنویسی آن سریع‌تر از ویرایش مدام آن است. نشانه‌های مختلفی وجود دارد که نشان می‌دهد باید کار بیشتری روی نوشتمن انجام گیرد. برای مثال اگر شما به عنوان نویسنده نتوانید در چند دقیقه پیام اصلی مقاله را به همکاران بگویید، قطعاً یک خواننده نیز نمی‌تواند. در این حالت باید روی داستان مقاله‌تان بیشتر کار کنید. پیدا کردن نقص‌های این چنینی مقاله می‌تواند کیفیت آن را بسیار بهبود ببخشد.

یک محقق برای نوشتمن خوب نیاز به بازخورد از طرف دیگران دارد. از چند فرد بخواهید متن را یک بار بخوانند تا مطمئن شوید که کلیت آن قابل فهم است. بازخورد آن‌ها کمک می‌کند تا بدانید چه قسمتی از نوشتمن تند یا آهسته پیش‌رفته است و یا اینکه چه قسمتی باید دوباره از نو نوشتene شود. بازخوردی که از داوران مقاله می‌گیرید بسیار مفید است. یک بازخورد مبهم

نشانه‌ای از آن است که شخص خواننده حرف اصلی مقاله را متوجه نشده است و در نتیجه باید روی آن کار کنید. در مقابل یک بازخورد مشخص به قسمتی از مقاله به شما می‌گوید احتمالاً منطق آن پاراگراف یا بخش نیاز به بازبینی دارد. همواره با روی باز با بازخوردهای افراد برخورد کنید؛ زیرا داشتن بازخورد از شبکه‌ای از همکاران برای پختن داستانی ماندگار در مقاله بسیار حیاتی است. برای اطمینان از اینکه که این اثر متقابل است به آن‌ها نیز در خواندن و ویرایش مقاله‌هایشان کمک کنید.

توضیحات

در این مقاله درباره‌ی ساختار یا به اصطلاح آناتومی مقاله صحبت کردیم. شاید بهتر بود درباره‌ی بسیاری از موارد جزئی تر مانند انتخاب کلمات و دستور زبان، تحلیل خلاقانه‌ی مباحث و همکاری علمی نیز نکاتی بیان می‌شد. نوشن درباره‌ی نوشن پایانی ندارد. توصیه می‌کنیم بسیاری از موارد دیگر را در ارجاعات [۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۱۶، ۱۷] ببینید.

اگر به سلیقه‌ی خودمان باشد به این قوانین پاییند نخواهیم بود و حتی ممکن است با ساختارشکنی متن زیبایی بنویسیم. لکن مثل اکثر تجربیات دیگر در زندگی، هنگامی یک اثر جاودانه خلق می‌شود که خالق آن با ممارست اصول اولیه را یاد گرفته باشد و در زمان مناسب با دورزدن آن شکل جدیدی را به وجود آورد [۱۸]. هدف این قوانین این است که به خوانندگان وسیع و متنوعی متصل شوید تا گفت‌گو بین رشته‌های مختلف علمی راحت‌تر برقرار شود.

قدرتانی و تشکر

از پدرم، بیژن فرهودی، برای ویرایش این ترجمه قدردانی و صمیمانه تشکر می‌کنم.

مراجع

- [1] Hirsch, Jorge E, *An index to quantify an individual's scientific research output*, Proceedings of the National academy of Sciences of the United States of America, 2005
- [2] Acuna, Daniel E and Allesina, Stefano and Kording, Konrad P, *Future impact: Predicting scientific success*, Nature, 2012
- [3] Paiva, Carlos Eduardo and Lima, João Paulo da Silveira Nogueira and Paiva, Bianca Sakamoto Ribeiro, *Articles with short titles describing the results are cited more often*, Clinics, 2012
- [4] Carter, Matt, *Designing science presentations: A visual guide to figures, papers, slides, posters, and more*, 2012
- [5] Murdock Jr, Bennet B, *Serial order effects in short-term memory*, Journal of Experimental Psychology, 1968
- [6] Schimel, Joshua, *Writing science: how to write papers that get cited and proposals that get funded*, 2012
- [7] Tufte, Edward R, *Envisioning information*, Optometry & Vision Science, 1991
- [8] Tufte, Edward R, *The visual display of quantitative information*, 2001
- [9] Lisberger, SG, *From Science to Citation: How to Publish a Successful Scientific Paper*, 2011
- [10] Simons, D, *Dan's writing and revising guide*, 2012
- [11] Sørensen, Carsten, *This is not an article: Just some thoughts on How to Write One*, 2002
- [12] Robert, A, Day, *How to write and publish a scientific paper*, 1994
- [13] Lester, JD and Lester, J, *Writing research papers*: Scott, 1967
- [14] Dumont, JL, *Trees, Maps, and Theorems*. Principiae, 2009
- [15] Pinker, Steven, *The Sense of Style: The Thinking Person's Guide to Writing in the 21st Century!*, 2015
- [16] Bern, D, *Writing the empirical journal, The compleat academic: A practical guide for the beginning social scientist*, 1987
- [17] Gopen, George D and Swan, Judith A, *The science of scientific writing*, American Scientist, 1990
- [18] Strunk, William, *The elements of style*, 2007

[†] مترجم: روزبه فرهودی



اثباتی برای نامتناهی بودن اعداد اول با استفاده از نظریه‌ی اطلاعات

آلا جواهری

چکیده. در این نوشه نامتناهی بودن اعداد اول ارائه خواهیم داد که از مفهوم انتروپی — که از مفاهیم بنیادین نظریه‌ی اطلاعات است — بهره می‌گیرد. افزون بر این، خواهیم دید که اثبات ما کران پایینی را نیز برای تابع شمارش اعداد اول فراهم می‌کند.

۱. مقدمه

نظریه‌ی اطلاعات بی‌شک از مهم‌ترین دستاوردهای علمی قرن بیستم است. کلاؤد شانون در ۱۹۴۸ با انتشار [۱] نظریه‌ای را بنیان نهاد که به مدد آن امروزه قادر هستیم گسترده‌ای از فناوری‌ها را — از دیسک‌های فشرده^۱ تا اینترنت ۵G — در اختیار داشته باشیم. یکی از اساسی‌ترین کارهای شانون در [۱] این بود که روشی را برای کمی‌سازی مفهوم «اطلاعات» پیشنهاد داد — البته باید در نظر داشت که رویکرد شانون تنها راه ممکن برای نیل به چنین مقصودی نیست. او بدین منظور از مفهوم انتروپی بهره گرفت. برای یک متغیر تصادفی گستته‌ی X که مقادیر x_1, \dots, x_n را با احتمال p_1, \dots, p_n اخذ می‌کند، انتروپی X که با $H(X)$ نمایش داده می‌شود، به صورت

$$H(X) \equiv \sum_{i=1}^n p_i \log \frac{1}{p_i}$$

تعریف می‌شود، که در آن لگاریتم‌ها در پایه‌ی دو هستند. شانون نشان داد که کمیت فوق — که به تعبیری میانگین اطلاعات موجود در یک متغیر تصادفی است — تعیین‌کننده‌ی نرخ نهایی قابل حصول برای برخی از مهم‌ترین وظایف نظریه‌ی مخابرات — فشرده‌سازی داده و انتقال داده روی یک کانال در معرض نویز — است. با این همه، اهمیت کار شانون فقط به کاربردهای آن در نظریه‌ی مخابرات محدود نمی‌شود. ابزارهای نظریه‌ی اطلاعاتی را می‌توان در زمینه‌های متفاوتی از ریاضیات به کار گرفت. در ادامه به عنوان نمونه‌ای از این کاربردها، نامتناهی بودن اعداد اول را با به‌کارگیری چند نامساوی انتروپیک ثابت می‌کنیم. این قضیه از کهن‌ترین قضایای نظریه‌ی اعداد است که یونانیان باستان نیز از آن مطلع بودند، و اولین اثباتی که برای آن در دست است، اثباتی است در اصول، شاهکار ماندگار اقليدس [۲]. با این حال از آن زمان تا به امروز اثبات‌های متعدد دیگری نیز برای این قضیه ارائه شده است. خواننده می‌تواند لیست مفصلی از این اثبات‌ها را در [۴] بیابد. اثبات ما در این نوشه برگرفته از اثباتی است که نخستین بار در [۳] بیان شده است.

۲. اثبات نامتناهی بودن اعداد اول

برای عدد طبیعی $n \leq 2$ ، تعداد اعداد اول کوچکتر از یا مساوی با n را با $\pi(n)$ نمایش می‌دهیم. به این تابع، تابع شمارش اعداد اول^۲ گفته می‌شود. فرض کنید $p_{\pi(n)} < p_2 < \dots < p_1$ اعداد اول کوچکتر از یا مساوی با n باشند. N را یک متغیر تصادفی یکنواخت روی مجموعه‌ی $\{1, 2, \dots, n\}$ در نظر بگیرید. از قضیه‌ی اساسی حساب می‌دانیم N تجزیه‌ی یکتاوی به عوامل اول دارد. برای $i = 1, \dots, \pi(n)$ ، متغیر تصادفی X_{p_i} را توان عدد اول p_i در تجزیه‌ی N تعریف

¹CD

²Prime-counting function

کنید. از یکتایی تجزیه می‌توان نتیجه گرفت که توزیع توانم $(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}})$ با توزیع N یکسان است. بنابراین

$$H(N) = H(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}}).$$

از نحوه‌ی تعریف N می‌توان نتیجه گرفت $n = H(N) = \log n$. حال از یک نامساوی نظریه‌ی اطلاعاتی برای یافتن کران بالایی برای $H(N)$ استفاده می‌کیم.

لم ۱.۲. برای متغیرهای تصادفی X_1, X_2, \dots, X_k

$$H(X_1, X_2, \dots, X_k) \leq \sum_{i=1}^k H(X_i).$$

با استفاده از لم فوق داریم:

$$\log n = H(N) = H(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}}) \leq \sum_{i=1}^{\pi(n)} H(X_{p_i}).$$

اکنون با استفاده از نامساوی انتروپیک دیگری کران بالایی برای هر (X_{p_i}) ارائه می‌دهیم.

لم ۲.۲. فرض کنید X یک متغیر تصادفی گستته است. اگر تعداد اعضای $\text{Supp}(X)$ برابر با d باشد،

$$H(X) \leq \log d.$$

برای هر X_{p_i} ، تعداد اعضای $\text{Supp}(X_{p_i})$ کوچکتر از یا مساوی با ۱ است. بنابراین داریم:

$$\log n = H(N) = H(X_{p_1}, X_{p_2}, \dots, X_{p_{\pi(n)}}) \leq \sum_{i=1}^{\pi(n)} H(X_{p_i}) \leq \pi(n) \log(\log n + 1),$$

که نتیجه می‌دهد

$$\frac{\log n}{\log(\log n + 1)} \leq \pi(n).$$

از آن‌جا که سمت چپ نامساوی فوق وقتی $\rightarrow \infty$ به بینهایت میل می‌کند، تعداد اعداد اول نمی‌تواند متناهی باشد. توجه کنید که در اثبات فوق نه تنها نشان دادیم که تعداد اعداد اول نامتناهی است، بلکه کران پایینی برای تابع شمارش اعداد اول پیدا کردیم.

می‌توان با کلک هوشمندانه‌ای کران پایینی را که در این اثبات برای $\pi(n)$ به دست آمد، بهتر کرد. فرض کنید در تجزیه‌ی N به عوامل اول، آن را به صورت $N = M \times p_1^{Y_{p_1}} \times \dots \times p_{\pi(n)}^{Y_{p_{\pi(n)}}}$ بنویسیم، که در آن M بزرگ‌ترین عدد مرتع کاملی است که N را می‌شمارد. در این صورت \sqrt{n} حداکثر $\text{Supp}(M)$ عضوی است، و برای هر i ، $\text{Supp}(Y_{p_{\pi(i)}})$ دو عضوی است. لذا

$$\log n = H(N) = H(M, Y_{p_1}, Y_{p_2}, \dots, Y_{p_{\pi(n)}}) \leq H(M) + \sum_{i=1}^{\pi(n)} H(Y_{p_i}) \leq \frac{1}{2} \log n + \pi(n) \log 2,$$

که نتیجه می‌دهد

$$\frac{\log n}{2 \log 2} \leq \pi(n).$$

۳. اثبات لم‌ها

در این بخش برای اثبات لم‌های ۱.۲ و ۲.۲ کمیت انتروپیک جدیدی را معرفی می‌کنیم که به انتروپی نسبی دو متغیر تصادفی مشهور است. برای دو متغیر تصادفی گستته‌ی X و Y که مقادیر p_1, \dots, p_n را به ترتیب با احتمال‌های $\omega_1, \dots, \omega_n$ اخذ می‌کنند، انتروپی نسبی از X به Y ، که آن را با $D(X||Y)$ نمایش می‌دهیم، به صورت

$$D(X||Y) \equiv \sum_{i=1}^n p_i \log \frac{p_i}{q_i},$$

تعريف می‌شود^۱. ویرگی مهم این کمیت آن است که نامنفی است. برای اثبات نامنفی بودن، از نامساوی $\ln x \leq x - 1$ استفاده می‌کنیم. داریم:

$$\sum_i p_i \ln \frac{q_i}{p_i} \leq \sum_i p_i \left(\frac{q_i}{p_i} - 1 \right) = \sum_i q_i - \sum_i p_i = 0.$$

$$\text{بنابراین } 0 \geq \sum_i p_i \log \frac{p_i}{q_i}$$

حال می‌توان دید که اثبات لم ۲.۲ تیجه‌ی سرراستی از نامنفی بودن انتروپی نسبی است. کافی است برای متغیر تصادفی X که مقادیر x_1, \dots, x_d را با احتمال ناصفراخذ می‌کند، متغیر تصادفی Y را به عنوان متغیر تصادفی یکنواخت روی $\text{Supp}(X)$ تعریف کنیم. به این ترتیب داریم:

$$0 \leq D(X||Y) = \sum_i p_i \log \frac{p_i}{\frac{1}{d}} = \sum_i p_i \log p_i - \sum_i p_i \log \frac{1}{d} = \sum_i p_i \log p_i - \log \frac{1}{d},$$

$$\text{که تیجه می‌دهد } H(X) \leq \log d \text{ یا معادلاً } \log \frac{1}{d} \leq \sum_i p_i \log p_i$$

برای اثبات لم ۱.۲ کافی است حکم را برای حالت $k=2$ ثابت کنیم. بدین منظور، متغیرهای تصادفی مستقل X' و Y' را به ترتیب هم توزیع با X و Y تعریف می‌کنیم. به این ترتیب داریم:

$$\begin{aligned} 0 &\leq D(X, Y||X', Y') = \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \frac{\Pr[X = x_i, Y = y_j]}{\Pr[X' = x_i, Y' = y_j]} \\ &= \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X = x_i, Y = y_j] - \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X' = x_i, Y' = y_j] \\ &= -H(X, Y) - \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X' = x_i] \Pr[Y' = y_j] \\ &= -H(X, Y) - \sum_{i,j} \Pr[X = x_i, Y = y_j] \log \Pr[X' = x_i] - \sum_{i,j} \Pr[X = x_i, Y = y_j] \Pr[Y' = y_j] \\ &= -H(X, Y) - \sum_i \Pr[X = x_i] \log \Pr[X' = x_i] - \sum_j \Pr[Y = y_j] \Pr[Y' = y_j] \\ &= -H(X, Y) + H(X) + H(Y) \end{aligned}$$

$$\text{که تیجه می‌دهد } H(X, Y) \leq H(X) + H(Y)$$

مراجع

- [1] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.
- [2] Heath, T. L. (Ed.). (1956). *The thirteen books of Euclid's Elements*. Courier Corporation.
- [3] Chaitin, G. J. (1977). Toward a Mathematical Definition of Life, 2. IBM Thomas J. Watson Research Division.
- [4] Meštrović, R. (2012). Euclid's theorem on the infinitude of primes: a historical survey of its proofs (300 BC–2022) and another new proof. *arXiv preprint arXiv:1202.3670*.
- [5] Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.

^۱ تعریفی که در اینجا ارائه کردہ‌ایم، حالت خاصی از تعریف رایج در ادبیات نظریه‌ی اطلاعات است. برای اطلاع از تعریف رایج به [۵] مراجعه کنید.

بحران در مبانی ریاضیات*

بخش اول

ژوژه فریروس

سرآغاز

در میان ریاضی‌دانان، بحران در مبانی ریاضیات مساله‌ای پذیرفته‌شده است که زمزمه‌اش حتی به غیر‌ریاضی‌دانان نیز رسیده‌است. از یک ریاضی‌دان خبره انتظار می‌رود در مورد سه مکتب «منطق‌گرایی»، «صورت‌گرایی»، و «شهود‌گرایی» — که در ادامه توضیح داده می‌شوند — اطلاعاتی داشته باشد، و همچنین بداند نتایج ناتمامیت گودل [۱، ۱۵.V] در مورد وضعیت فعلی دانش ریاضی به ما چه می‌گویند. ریاضی‌دانان حرفه‌ای معمولاً^۱ در این کشمکش عقاید راسخی دارند. گروهی بحث در مورد مبانی ریاضیات را به‌کلی ناموجه می‌دانند — و نتیجتاً گفتمان غالب را نقویت می‌کنند — و گروهی، یا به عنوان یک امر بنیادی و یا صرفاً به عنوان یک گزینه‌ی جذاب، نوعی رویکرد تجدیدنظرخواهانه به ریاضیات را پیش می‌گیرند. اما پیکره‌ی اصلی این مجادله‌ی تاریخی چندان شناخته‌شده نیست، و ایرادات فلسفی ظریف در این میان معمولاً به‌کلی نادیده گرفته می‌شوند. در این نوشتار، بیشتر به مسأله‌ی نخست پرداخته می‌شود، به این امید که با این کار، ایرادات بنیادین این بحث بیشتر در مرکز توجه قرار گیرند.

تصور غالب از بحران بنیادین، دوره‌ای نسبتاً کوتاه در دهه‌ی ۱۹۲۰ میلادی است؛ مجادله‌ای مفصل بین طرفداران ریاضیات «کلاسیک» (اینجا منظور از ریاضیات کلاسیک، ریاضیات اوخر سده‌ی نوزدهم است.)، به رهبری هیلبرت [۱، VI.63]، و منتقدینشان، به رهبری براوئر [۱، VI.75]، که قائل به لزوم تجدیدنظر اساسی در اصول ریاضیات کلاسیک بودند. به عقیده‌ی من، طرز فکر بسیار مهم دیگری وجود دارد؛ اینکه این «بحaran بنیادین»، روندی طولانی و جهان‌شمول بوده‌است؛ روندی که با ظهور ریاضیات جدید و مسائل فلسفی و روش‌شناسانه‌ی برآمده از آن نیز ارتباط تنگاتنگی داشته‌است. این نوشتار نیز از همین زاویه‌ی دید نگاشته شده‌است.

در روند طولانی‌تری که ذکر شد، همچنان می‌توان بازه‌هایی نسبتاً مهم‌تر متصور بود. در حوالی سال ۱۸۷۰، بحث‌هایی درباره‌ی پذیرفته بودن یا نبودن هندسه‌های ناقلیدسی، و همچنین درباره‌ی یافتن بنیادهایی شایسته برای آنالیز مختلط و حتی اعداد حقیقی شکل گرفته بود. در اوایل سده‌ی بیستم مجادلاتی درباره‌ی نظریه‌ی مجموعه‌ها، مفهوم پیوستگی، و همچنین درباره‌ی نقش منطق و روش اصل موضوعی در برایر شهود در ریاضیات وجود داشت. در سال ۱۹۲۵، وجود یک بحران در مبانی ریاضیات^۲ دیگر به‌سادگی قابل مشاهده بود. در همین حین، ایده‌های اصلی مطرح شده در این مجادلات پرورش یافته و به پژوهه‌های تحقیقاتی مفصلی در ریاضیات تبدیل شدند. و نهایتاً در دهه‌ی ۱۹۳۰، گودل [۱، ۹۲.VI] قضایای ناتمامیت خود را اثبات کرد؛ تاییجی که برای هضم آن‌ها، لازم بود ریاضی‌دانان از برخی باورهای محبوب خود دل بکنند. در ادامه تعدادی از این رویدادها و مسائل را با جزئیات بیشتر بررسی می‌کنیم.

۱. پرسش‌های بنیادین اولیه

شواهدی وجود دارد که هیلبرت در سال ۱۸۹۹، بر تفکری بود که بعدها منطق‌گرایی نام گرفت. ایده‌ی منطق‌گرایی این است که مفاهیم اولیه‌ی ریاضیات با استفاده از ابزار‌آلات منطقی قابل تعریف‌اند، و ضمناً اصول کلیدی ریاضیات را نیز می‌توان

*این نوشتة، ترجمه‌ای از مقاله‌ی زیر است:

Ferreirós, J. (2008). The Crisis in the Foundations of Mathematics. *The Princeton Companion to Mathematics*, 142-156.

تنها با استفاده از اصول منطقی استنتاج کرد.

با گذشت زمان، این ایده قدری مبهم شده‌است، زیرا این‌گونه به نظر می‌رسد که در مورد قلمرو یک نظریه‌ی منطقی دچار کج فهمی است. اما از نظر تاریخی، منطق‌گرایی واکنشی هوشمندانه به ظهور ریاضیات جدید، به ویژه رویکرد و روش‌های برخاسته از نظریه مجموعه‌ها بود. از آن‌جا که به نظر اکثر ریاضی‌دانان نظریه‌ی مجموعه‌ها صرفاً بخشی از منطق بود،^۱ به نظر می‌رسید حقایقی مانند اینکه نظریات مربوط به اعداد طبیعی و حقیقی را می‌توان از نظریه‌ی مجموعه‌ها استنتاج کرد، و همچنین اهمیت روزافزون روش‌های مبتنی بر نظریه‌ی مجموعه‌ها در جبر، آنالیز حقیقی و آنالیز مختلط، مؤید این ایده باشد.

درک هیلبرت از ریاضیات، متاثر از ددکیند [۱، VI.50] بود. پایه و اساس منطق‌گرایی هیلبرت و ددکیند، پشتیبانی جسورانه‌شان از روش‌هایی بود که در آن زمان جدید و ناشناخته بودند. روش‌هایی که تولدشان به قرن نوزدهم و دانشگاه گوتینگن^۲ (گاووس [۱، VI.26] و دیریکله [۱، VI.36]) باز می‌گردد. پس از آن ایده‌های بدیع ریمان [۱، VI.49] نقطه عطفی برای این روش‌ها بود، و در نهایت ددکیند، کانتور [۱، VI.54]، هیلبرت، و افراد دیگری روی توسعه‌ی آن‌ها کار کردند. در همین حال، مکتب پرنفوذ برلین، شامل افرادی چون کرونکر [۱، VI.48] و وایرشتراس [۱، VI.44]، مخالف این گرایش جدید بود. (گرچه نام وایرشتراس با معرفی دقت در آنالیز حقیقی پیوند خورده است، اما خواهیم دید که او روش‌های جدیدی که در آن زمان در حال شکل‌گیری بودند را چندان قبول نداشت). ریاضی‌دانان در پاریس و بقیه‌ی دنیا نیز در مورد این ایده‌های رادیکال و نو، تردید داشتند.

شاخص‌ترین ویژگی‌های رویکرد جدید در ریاضیات، عبارت بودند از:

(آ) پذیرش مفهوم توابع «دلخواه»، به پیشنهاد دیریکله؛

(ب) پذیرش تمام و کمال مجموعه‌های نامتناهی و بی‌نهایت‌های متعالی؛

(پ) تلاش برای «قرار دادن اندیشه‌ها به جای محاسبات» (دیریکله)، و تمرکز بر «ساختار»‌هایی که ویژگی‌های آن‌ها از اصول موضوعه نتیجه می‌شود؛ و

(ت) تکیه‌ی مکرر بر اثبات‌های «کاملاً وجودی».

به عنوان مثالی برای این ویژگی‌ها، می‌توان از رویکرد ددکیند در سال ۱۸۷۱ در مواجهه با نظریه‌ی جبری اعداد نام برد. ددکیند میدان‌های عددی [۱، III.63] و ایده‌آل‌ها [۱، §۲.۸۱] را با استفاده از مفاهیم نظریه‌ی مجموعه‌ها تعریف کرد، و از روش‌های مبتنی بر ریاضیات جدید برای اثبات نتایجی چون قضیه‌ی اساسی یکتایی، تجزیه استفاده کرد. او بر خلاف عرف رایج در نظریه‌ی اعداد، به بررسی تجزیه‌ی اعداد صحیح جبری بر اساس ایده‌آل‌ها، که مجموعه‌هایی نامتناهی از اعداد صحیح جبری‌اند، پرداخت و با استفاده از این مفهوم مجرد جدید به همراه تعریفی مناسب از ضرب دو ایده‌آل، توانست در حالت کلی اثبات کند که در هر حلقه از اعداد صحیح جبری، ایده‌آل‌ها به‌طور یکتا به ایده‌آل‌های اول تجزیه می‌شوند.

کرونکر، جبردان صاحب‌نظر، بر اثبات‌های ددکیند این نقد را داشت که با استفاده از آن‌ها نمی‌توان شمارنده‌ها یا ایده‌آل‌های موردنظر را به دست آورد؛ به عبارت دیگر، اثبات او کاملاً وجودی بود. نظر کرونکر این بود که این طرز فکر، که به لطف روش‌های مبتنی بر نظریه‌ی مجموعه‌ها و تمرکز روی ویژگی‌های جبری ساختارها ممکن شده بود، فاصله‌ی زیادی با روش‌های ساختی داشت. اما از نظر ددکیند این نقد وارد نبود و صرفاً نشان‌دهنده‌ی این بود که او در مسیر اثبات توانسته بود با موقفيت اصل قرار دادن اندیشه‌ها به جای محاسبات را نشان دهد؛ اصلی که کاربردش همچنین در نظریه‌ی توابع مختلط ریمان نیز به کرات دیده می‌شد. البته، برای مسائلی که کم‌تر جنبه‌ی انتزاعی داشتند همچنان توسعه‌ی روش‌های محاسباتی لازم بود و ددکیند نیز در چندین مقاله به این موضوع پرداخته بود، اما او بر اهمیت یک نظریه‌ی عمومی و انتزاعی نیز پاشاری داشت.

روش‌ها و ایده‌های ریمان و ددکیند در پی مقالات منتشرشده در دوره‌ی ۱۸۶۷ تا ۱۸۷۲ بیشتر شناخته شدند. دفاع صریح آن‌ها از این ایده که نظریه‌ها در ریاضیات باید نه بر پایه‌ی فرمول‌ها و محاسبات، بلکه بر پایه‌ی مفاهیم کلی باشند، و عبارات تحلیلی و ابزارهای محاسباتی صرفاً در مقام ابزارهایی برای کمک به توسعه‌ی نظریه‌ها هستند.

نمونه‌ای از تفاوت بین این دیدگاه‌ها را می‌توان در رویکردهای متفاوتی دید که ریمان و وایرشتراس به نظریه‌ی تابع‌ها داشتند. وایرشتراس توابع تحلیلی را توابعی می‌دانست که نمایشی به‌شکل یک سری توانی مانند $\sum_{n=0}^{\infty} a_n(z - a)^n$ داشتند، و با فرازد تحلیلی [۱، I.3] به یک دیگر مرتبط می‌شدند. در مقابل، ریمان رویکردی کاملاً متفاوت داشت. او یک تابع را تحلیلی می‌دانست

^۱ایدی خاطرنشان کرد که افراد شاخصی چون ریمان و کانتور مخالف این موضوع بودند (ر.ک. Ferreirós 1999). اکثریت مورد اشاره شامل افرادی چون ددکیند، پئانو [۱، VI.62]، هیلبرت، و راسل است.

اگر شرایط مشتق‌پذیری کوشی – ریمان^۱ [۱.I.۳] را ارضاء می‌کرد. از آنجا که ویژگی‌های توابع مشتق‌پذیر تا آن زمان هیچ‌گاه به طور موشکافانه مشخص نشده بود، وایرشتراس این تعریف انتزاعی را غیرقابل قبول می‌دانست. او با توانایی‌های نقادانه‌ی مشهور خود، مثال‌هایی از توابعی پیوسته اما همه‌جامشتو ناپذیر یافت.

شایان ذکر است که ترجیح وایرشتراس بر اصل قرار دادن سری‌های نامتناهی در تحقیقات مربوط به آنالیز و نظریه‌ی تابع‌ها، او را به تفکر کهنه‌ی قرن هجدهمی مبنی بر اینکه هر تابع یک عبارت تحلیلی است نزدیک‌تر قرار می‌داد. در مقابل، ریمان و ددکیند ایده‌ی انتزاعی‌تر دیریکله را ترجیح می‌دادند، که یک تابع f در واقع روشنی برای نسبت دادن یک $y = f(x)$ دلخواه به x است (و نیازی نیست این مقدار، لزوماً به صورت یک عبارت صریح بر حسب x مشخص شود). وایرشتراس در نامه‌هایش این تفکر دیریکله را به دلیل بیش از حد^۲ کلی و عمومی بودن، برای ایفای نقش به عنوان نقطه‌ی شروعی برای گسترش ریاضیات قابل نمی‌دانست. به نظر می‌رسد او در این مورد اشتباه می‌کرد، و این طرز فکر دقیقاً چارچوب مورد نیاز برای تعریف و بررسی مفاهیمی چون پیوستگی و انتگرال‌گیری را فراهم می‌کرد. این چارچوب بعدها در قرن نوزدهم به رویکرد مفهوم‌گرایانه معروف شد. مجادلات روش‌شناسانه مشابهی در شاخه‌های دیگر نیز در حال شکل‌گیری بودند. کرونکر سال ۱۸۷۰ در نامه‌ای ادعا می‌کند که قضیه‌ی بولتسانو – وایرشتراس یک «سفسطه‌ی آشکار» است، و او مثال‌های نقضی را برای آن مطرح خواهد کرد. این قضیه، که صورت آن بیان می‌کند که هر مجموعه‌ی کران‌دار نامتناهی از اعداد حقیقی، یک نقطه‌ی انباشتگی دارد، قضیه‌ای بنیادین در آنالیز کلاسیک است، و وایرشتراس نیز در سخنرانی‌های برلین خود به این موضوع تاکید کرده بود. اشکال کرونکر این بود که این قضیه تکیه‌ای اساسی بر اصل تمامیت اعداد حقیقی (که یک صورت‌بندی از آن این است که اشتراک هر دنباله‌ای از بازه‌های بسته‌ی تودرتو در \mathbb{R} ، ناتهی است). دارد. ساخت اعداد حقیقی از روی اعداد گویا با روش‌های مقدماتی ممکن نیست و نیاز به کارگیری مجموعه‌های نامتناهی (مانند مجموعه‌ی تمام برش‌های ددکیند)، که عبارت از زیرمجموعه‌هایی از اعداد گویا مانند $C \subset \mathbb{Q}$ است به طوری که اگر $q \in C$ و $q < p$ باشد. دارد. به عبارت دیگر، مسئله‌ی کرونکر این بود در اکثر مواقع، که نقطه‌ی انباشتگی مورد اشاره در این قضیه را نمی‌توان با روش‌های مقدماتی از روی اعداد گویا ساخت. ایده‌ی ریاضی‌دانان کلاسیک از مجموعه‌ی اعداد حقیقی، یا پیوستار، در آن زمان نیز شامل هسته‌های اولیه‌ی قسمت‌های غیرساختی ریاضیات جدید می‌شد.

بعدتر، حدود سال ۱۸۹۰، کارهای هیلبرت در نظریه‌ی ناوردا^۲ منجر به درگرفتن جدالی شد درباره‌ی اثبات کاملاً وجودی او از یک حکم ساده‌ی دیگر، قضیه‌ی پایه، که می‌گوید هر ایده‌آل در یک حلقه‌ی چندجمله‌ای به طور متناهی تولید می‌شود. پل گرдан، که به دلیل کار جدی الگوریتمی در زمینه‌ی ناورداها به پادشاه این موضوع مشهور است، به طور طنزآمیزی اظهار کرده بود که این اثبات نوعی خداشناسی است و نه ریاضیات. (منظور او ظاهرا این بوده که به دلیل اینکه اثبات هیلبرت کاملاً وجودی — و نه ساختی — است، می‌توان آن را همراه با اثبات‌های فلسفی بر وجود خدا دانست).

این جدال پایه‌ای باعث شد نقطه‌نظرهای دو طرف روش‌تر شود. اثبات‌های کانتور در نظریه‌ی مجموعه‌ها نیز به مثال‌های اصیل از روش جدید اثبات وجودی تبدیل شد. او در مقاله‌ای که سال ۱۸۸۳ به چاپ رسید، دفاعیه‌ای صریح از بی‌نهایت متعالی و روش‌های جدید ریاضی‌ورزی ارائه داد، و هم‌چنین به طور مخفی نظرات کرونکر را مورد حمله قرار داد. کرونکر نیز در ۱۸۸۲ به طور عمومی سیاق ددکیند را نقد کرد، در مخالف خصوصی بر علیه کانتور سخن گفت، و در ۱۸۸۷ تلاش کرد با انتشار مقاله‌ای، نقطه‌نظرات خود درباره‌ی بنیان‌های ریاضی را مشخص کند. ددکیند در پاسخ به او، در ۱۸۸۸ نظریه‌ای بر پایه‌ی نظریه‌ی مجموعه‌ها (و بنابرین، از نظر خودش، منطق‌گرایانه) درباره‌ی اعداد طبیعی مطرح کرد.

موج اول انتقادات ظاهرا با پیروزی جبهه‌ی مدرن به پایان رسید. جبهه‌ای که همراهان جدید و قدرتمندی مانند هورویتز، مینکفسکی [۱.VI.۶۴، هیلبرت، ولترا، پئانو، و هادامارد [۱.VI.۶۵] داشت، و توسط اشخاص تاثیرگذاری مانند کلاین [۱.VI.۵۷] حمایت می‌شد. با اینکه نظریه تابع‌های ریمانی هنوز به اصلاحاتی نیاز داشت، اما پیشرفت‌های اخیر در آنالیز حقیقی، نظریه‌ی اعداد و زمینه‌های دیگر قدرت و دورنمای روش مدرن را نشان می‌داد. در دهه‌ی ۱۸۹۰، دیدگاه مدرن به طور کلی، و منطق‌گرایی به طور ویژه، بسیار گسترش یافت. هیلبرت این روش نو را در قالب روش اصل موضوعی توسعه داد، و سپس از آن برای اصلاح هندسه (۱۸۹۹) و ویرایش‌های بعدی) و دستگاه اعداد حقیقی استفاده کرد.

^۱ ریمان توابع را بر اساس تعدادی ویژگی مستقل مانند سطح ریمانی متناظر و رفتار در نقاط تکینه مشخص می‌کرد. این ویژگی‌ها تابع را بر اساس اصل دیریکله مشخص می‌کرد؛ اصلی که وایرشتراس بر آن نیز نقدهایی ایراد کرده و مثال نقض ارائه داده بود. بعدها هیلبرت و نسر این اصل را مجدداً صورت‌بندی و توجیه کردن.

^۲ invariant theory

بعد از آن، ظهور پارادوکس‌های منطقی — که کاتتور، راسل، زرملو و دیگران کاشف آن‌ها بودند — باعث نابودی چشم‌انداز زیبایی شد که منطق‌گرایی و پیشرفت‌های حاصل از آن ایجاد کرده بود. این پارادوکس‌ها دو دسته‌اند؛ در یک رده، استدلال‌هایی هستند که نشان می‌دهند فرض وجود برخی مجموعه‌ها، به تناقض ختم می‌شود. این تناقضات بعداً تناقضات نظریه‌های مجموعه‌ای نام گرفتند. در رده‌ی دیگر، تناقض‌های معنایی هستند، که بیانگر سختی‌هایی در مورد مفاهیم راستی و تعریف‌پذیری‌اند. به‌واقع دوران اوج منطق‌گرایی پیش از ظهور این تناقض‌ها — پیش از سال ۱۹۰۰ — بود؛ اگرچه بعدها راسل با «نظریه‌ی انواع» تا حدی باعث بازگشت منطق‌گرایی به دوران اوج شد، اما در ۱۹۲۰ منطق‌گرایی بیش‌تر برای فلاسفه جذاب بود تا ریاضی‌دانان. با این حال، اختلاف میان طرفداران روش‌های مدرن با معتقدان ساخت‌گرای آنان از بین نرفت.

۲. حدود ۱۹۰۰

هیلبرت لیست مسائل معروف خود را در ۱۹۰۰ در کنگره‌ی بین‌المللی ریاضیات پاریس با مسئله‌ی پیوستار [۱، §۵] در نظریه‌ی مجموعه‌ها، و این مسئله که آیا هر مجموعه خوش‌ترتیب است یا نه آغاز کرد. مسئله‌ی دوم کاتتور، مسئله‌ای مهم در نظریه‌ی مجموعه‌ها، او مربوط به تصدیق سازگاری مجموعه‌ی اعداد حقیقی \mathbb{R} می‌شد. آغاز لیست هیلبرت با این دو مسئله اتفاقی نبود؛ بلکه روش او بود برای اعلام نظر درباره‌ی اینکه ریاضیات در قرن بیستم به چه سمتی باید برود. این دو مسئله، و اصل انتخاب [۱، III.3] — که زرملو، همکار جوان هیلبرت، با استفاده از آن نشان داد که برای \mathbb{R} یک خوش‌ترتیبی وجود دارد — مثال‌هایی بنیادین از ویژگی‌های چهارگانه‌ای که پیش‌تر آمد هستند. تعجبی ندارد که ذهن‌های محافظه‌کارتر مخالفت کردند و تشکیک‌های کرونکر را تکرار کردند، و این موضوع در بسیاری از مقالات منتشرشده در ۱۹۰۵–۱۹۰۶ مشهود است. با بیان این نکته، حال به موج بعدی جداول‌ها می‌رویم.

۱.۲. تناقض‌ها و مسئله‌ی سازگاری.

در ۱۸۹۶، کاتتور کشف کرد که مفاهیم ظاهراً بی‌دردسر «مجموعه‌ی تمام اعداد ترتیبی» و «مجموعه‌ی تمام اعداد اصلی»، مفاهیمی متناقض‌اند. مورد اول به تناقض بورالی–فورتی و مورد دوم به تناقض کاتتور مشهورند. این فرض که تمام اعداد ترتیبی ترامتناهی تشکیل یک مجموعه می‌دهند، به دنبال نتایج پیشین کاتتور، به این نتیجه ختم می‌شود که عدد ترتیبی‌یی وجود دارد که کمتر از خودش است. برای اعداد اصلی نیز تناقض مشابهی وجود دارد. ددکیند پس از خبردار شدن از این پارادوکس‌ها به شک افتاد که آیا اصلاً فکر آدمیزاد لروماً منطقی است یا نه. حتی بدتر، در ۱۹۰۱–۱۹۰۲ زرملو و راسل تناقضی بسیار ابتدایی پیدا کردند که امروزه به پارادوکس راسل، یا پارادوکس زرملو–راسل معروف است. ناشایستگی درک کلاسیک از نظریه‌ی مجموعه‌ها که آن را همتراز با منطق می‌دانست عیان شد، و عصر جدیدی از ناپایداری شروع شد. البته باید گفت فقط منطق‌گرایان — که نظریاتشان با تناقض مواجه شده بود — از این موضوع نگران بودند.

خوب است در اینجا به اهمیت پارادوکس زرملو–راسل پیردازیم. ریاضی‌دانان بسیاری، از ریمان گرفته تا هیلبرت، این اصل را پذیرفته بودند که برای هر ویژگی منطقی یا ریاضیاتی خوش‌تعريف، مجموعه‌ای وجود دارد شامل تمام اشیایی که دارای آنند. به عبارت دیگر، برای هر ویژگی خوش‌تعريف p ، وجود دارد مجموعه‌ی $\{x : p(x)\}$. مثلاً برای ویژگی «عدد حقیقی بودن» — که با استفاده از اصول موضوعه‌ی هیلبرت به طور دقیق بیان می‌شود — مجموعه‌ی تمام اعداد حقیقی، و برای ویژگی «عدد ترتیبی بودن» مجموعه‌ی تمام اعداد ترتیبی وجود دارد. این اصل که اصل تفهیم^۱ نام دارد، سنگ بنای درک منطق‌گرایانه از نظریه‌ی مجموعه‌های است، که معمولاً^۲ نظریه‌ی طبیعی^۳ مجموعه‌ها نامیده می‌شود. ساده‌اندیشانه بودن این ایده البته تنها با نگاه به گذشته معلوم می‌شود. این اصل به عنوان یک قانون پایه‌ای منطق پنداشته می‌شد، و بنابرین تمام نظریه‌ی مجموعه‌ها صرفاً بخشی از منطق مقدماتی بود.

پارادوکس زرملو–راسل نشان می‌دهد که اصل تفهیم^۴ متناقض است. این امر با ساخت یک ویژگی انجام می‌شود که در نگاه اول کاملاً ساده و منطقی^۵ به نظر می‌آید. فرض کنید $x \notin p$ ویژگی x (توجه کنید که نفی و عضویت مفاهیمی کاملاً منطقی فرض می‌شوند) باشد. اصل تفهیم وجود مجموعه‌ی $\{x : x \notin x\} = R$ را منجر می‌شود، اما وجود چنین مجموعه‌ای

¹comprehension principle

²در متن اصلی، عبارت naive set theory آمده که شاید بهتر باشد آن را نظریه‌ی مجموعه‌های ساده‌اندیشانه ترجمه کرد. نیمنگاهی به این ترجمه‌ی دیگر برای فهم جمله‌ی بعدی لازم است.

³مقصود اینجا از منطقی، این است که در ساخت آن به استفاده از ابزارهای خارج از منطق، مانند ریاضیات، نیازی پیدا نمی‌شود؛ و نه اینکه ویژگی مورد بحث عقلانی است.

تناقض دارد، زیرا اگر $R \in R$ ، طبق تعریف $R \notin R$ ، و نیز اگر $R \notin R$ ، آنگاه $R \in R$. هیلبرت (مانند همکار بزرگ‌ترش فرگه [۱] VI.56) تصمیم گرفت منطق‌گرایی را رها کند، و حتی به این فکر بیفتند که شاید تمام این مدت حق با کرونکر بوده است. در نهایت او به این نتیجه رسید که نظریه‌ی مجموعه‌ها نشان می‌دهد اصلاح اساسی نظریه‌ی منطق ضروری است. در ضمن نیاز بود نظریه‌ی مجموعه‌ها دوباره و این‌بار به روش اصل موضوعی بنا شود؛ به عنوان یک نظریه‌ی ریاضیاتی و برپایه‌ی اصول موضوعه‌ی برخاسته از ریاضی (ونه منطق)، و زرملو این کار را کرد.

هیلبرت اعتقاد داشت که ادعای وجود مجموعه‌ای از اشیایی ریاضیاتی، معادل با اثبات سازگاری (خالی از تناقض بودن) دستگاه اصول موضوعه‌ی متناظرش است. از شواهد تاریخی بر می‌آید که این اعتقاد او، واکنشی به پارادوکس‌های کانتور بوده است. استدلال او ممکن است این بوده باشد که به جای پرس مستقیم از مفاهیم خوش‌تعریف به مجموعه‌های متناظر با آن‌ها، ابتدا باید اثبات کنیم که مفاهیم مورد نظر منطقاً سازگارند. به عنوان مثال، پیش از پذیرفتن وجود مجموعه‌ی تمام اعداد حقیقی، باید سازگاری دستگاه اصول موضوعه‌ی هیلبرت برای اعداد حقیقی را ثابت کرد. این اصل هیلبرت روشی برای زدودن مفهوم وجود در ریاضی از مسائل متافیزیکی بود. این ایده که اشیا ریاضی دارای نوعی «وجود ایده‌آل» در قلمرو ذهن هستند و نه یک وجود مستقل متافیزیکی، پیش‌تر توسط ددکیند و کانتور هم گفته شده بود.

پارادوکس‌های منطقی علاوه بر پارادوکس‌های بورالی-فورتی، کانتور، و راسل، شامل چندین پارادوکس معنایی — از جمله پارادوکس‌های مطرح شده توسط راسل، ریچارد، کنیگ، گلینگ و ... — نیز می‌شد. وفور تناقضات منطقی باعث سردرگمی زیادی شد، اما یک چیز واضح بود: نقش مهم آن‌ها در تسریع رشد منطق مدرن و اقتاع ریاضی‌دانان درباره‌ی لزوم ارائه‌ی کاملاً صوری نظریاتشان. از تناقضات منطقی تنها زمانی می‌توان صرف نظر کرد — و آن‌ها را از تناقضات نظریه‌ی مجموعه‌ای تمیز داد — که نظریه در قالب یک زبان صوری دقیق بیان شود.

۲.۲. محمولیت.^۱

وقتی کتاب‌های فرگه و راسل باعث شناخته شدن تناقضات نظریه‌ی مجموعه‌ها در جامعه‌ی ریاضی‌دانان شدند، پوانکاره [۱] VI.61 از آن‌ها استفاده کرد تا نسبت به هر دوی منطق‌گرایی و صورت‌گرایی اعتقاداتی ابراد کند.

تحلیل او از این تناقض‌ها باعث شد او ایده‌ی جدید محمولیت را مطرح کند، و بر این باشد که در ریاضی باید از تعاریف نامحمولی پرهیز کرد. به بیان ساده، یک تعریف نامحمولی است اگر شیئی را با ارجاع به کلیتی معرفی کند که پیش‌تر شامل آن است. ددکیند مجموعه‌ی اعداد طبیعی را اشتراک تمام مجموعه‌های شامل ۱ و نسبت به تابع تالی بسته تعریف می‌کند (عدد ۱ در برد تابع تالی نیست). هدف او معرفی \mathbb{N} به عنوان یک مجموعه‌ی مینیمال بود، اما در تلاش برای این کار او در تعریف \mathbb{N} از یک کلیت از مجموعه‌ها بهره می‌جوید که باید قبلًا شامل \mathbb{N} باشد. پوانکاره (و هم‌چنین راسل) چنین روشی را پذیرفتنی نمی‌دانستند، به ویژه زمانی که شی مورد اشاره، تنها با چنین روشی قابل تعریف باشد. پوانکاره روش‌های نامحمولی مانند این را در هر یک از تناقضاتی که مطالعه می‌کرد، می‌یافت.

برای مثال، پارادوکس ریچارد را بنگرید، که یک پارادوکس زبان‌شناسانه یا معنایی است (که همانطور که گفته شد، با مفاهیم صدق و تعریف‌پذیری سروکار دارد). با آغاز از مفهوم اعداد حقیقی تعریف‌پذیر، و چون تعاریف باید با زبانی خاص بیان شوند و عبارات زبان نیز متناهی‌اند، نتیجه می‌شود که تنها تعداد شمارایی عدد حقیقی تعریف‌پذیر وجود دارد. پس می‌توانیم این اعداد را با استفاده از ترتیب لغت‌نامه‌ای تعاریف‌شان بشماریم. ایده‌ی ریچارد این بود که، مشابه ایده‌ی قطری‌سازی کانتور برای اثبات ناشمارایی \mathbb{R} ، از یک روند قطری استفاده کند. فرض کنید a_1, a_2, a_3, \dots اعداد تعریف‌پذیر باشند. عدد جدید r را طریق تعریف کنید که مطمئن باشید رقم n ام آن با رقم a_n متفاوت است (مثلا، این رقم را ۲ تعریف کنید مگر این که رقم a_n نیز ۲ باشد، در این صورت آن را ۴ بگذارید). این عدد نمی‌تواند عضو مجموعه‌ی اعداد تعریف‌پذیر باشد، اما همین لحظاتی پیش آن را با تعدادی متناهی کلمه تعریف کردیم! پوانکاره با ممنوع ساختن تعاریف نامحمولی، جلوی معرفی عددی چون r را می‌گیرد، زیرا برای تعریف آن از ارجاع به کلیت تمام اعداد تعریف‌پذیر بهره بردیم.^۲

در این رویکرد در مورد مبانی ریاضیات، تمام اشیای ریاضی (فرای اعداد طبیعی) باید با تعاریفی صریح معرفی شوند. اگر در تعریفی، از کلیتی استفاده شود که شی مورد تعریف به آن تعلق دارد، با یک دور مواجه می‌شویم: شی مورد تعریف، خود جزئی

^۱predicativity

^۲راه حل امروزی این مسئله، ساختن تعاریف ریاضیاتی در یک نظریه‌ی صوری خوش‌تعریف است که زبان و عبارات آن از ابتدا مشخص‌اند. پارادوکس ریچارد از ابهامی که در مورد روش‌های در دسترس برای تعریف وجود دارد حاصل می‌شود.

از تعریف خود است. در این دیدگاه، تعاریف باید محمولی باشند: تنها ارجاع به کلیت‌هایی مجاز است که پیش‌تر آن‌ها را معین کرده‌ایم. مؤلفان مهمی چون راسل و ویل [۱، VI.80] این دیدگاه را پذیرفته و به گسترش آن کمک کردند. این دیدگاه زرملو را قانع نکرده بود، او همچنان عقیده داشت بهره جستن از تعاریف نامحمولی، نه تنها در نظریه‌ی مجموعه‌ها (مانند تعریف ددکیند از \mathbb{N} ، بلکه در آنالیز کلاسیک هم معمولاً^۲ بی‌اشکال است. به عنوان یک مثال خاص، او به اثبات کوشی [۱، VI.29] از قضیه‌ی اساسی جبر^۳ اشاره کرد، اما یک مثال ساده‌تر از چنین تعریفی مفهوم کوچک‌ترین کران بالا در آنالیز حقیقی است. اعداد حقیقی نه مستقلًا با تعاریف محمولی از هر یک، بلکه به عنوان یک کلیت کامل تعریف می‌شوند، و در نتیجه روشنی که کوچک‌ترین کران بالای یک مجموعه‌ی نامتناهی کران‌دار از اعداد حقیقی را مشخص می‌کند نامحمولی خواهد بود. اما زرملو اصرار داشت که این تعاریف بی‌آزارند، زیرا شی مورد تعریف نه در حال ساخته شدن، بلکه تنها در حال مشخص شدن است (ر.ک. مقاله‌ی ۱۹۰۸ او، چاپ شده در صص. ۹۸–۱۸۳ کتاب ۱۹۶۷ ون‌هاینورت).

ایده‌ی پوانکاره برای از بین بردن تعاریف نامحمولی، برای راسل اهمیت بسیاری پیدا کرد، تا جایی که در نظریه‌ی تاثیرگذار انواعش از آن به عنوان «اصل دور باطل» یاد کرد. نظریه‌ی انواع یک دستگاه منطقی مرتبه بالاتر است که تسویر روی ویژگی‌ها، مجموعه‌ها، رابطه‌ها، مجموعه‌هایی از مجموعه‌ها، و به همین ترتیب را ممکن می‌سازد. با صرف نظر از جزئیات، این نظریه بر پایه‌ی این ایده‌ی بنا می‌شود که اعضای هر مجموعه باید اشیایی در یک نوع همگن خاص باشند. مثلاً، مجموعه‌هایی از «ashخاص»، مانند $\{a, b\}$ ، یا از مجموعه‌هایی از اشخاص، مانند $\{\{a\}, \{a, b\}\}$ ، مجاز، اما مجموعه‌ای با اعضای مخلوط، مثلاً $\{a, \{a, b\}\}$ ، غیرمجاز دانسته می‌شود. تفسیر راسل از نظریه‌ی انواع به دلیل انشعباتی که برای جلوگیری از نامحمولیت وارد آن کرده بود، به نسبت پیچیده بود. این دستگاه، به همراه اصول موضوعه‌ی بی‌نهایت، انتخاب، و «تحویل پذیری» (که راه حلی موضوعی برای حل مشکل انشعبات‌ها بود)، برای توسعه‌ی نظریه‌ی مجموعه‌ها و دستگاه‌های اعداد کفايت می‌کرد و بنابرین به سنگ بنای منطقی کتاب مشهور وايتهد و راسل، *Principia Mathematica*، که در آن مؤلفان با دقت نظر مبانی ریاضیات خود را بنا کردند، تبدیل شد.

نظریه‌ی انواع تا حدود ۱۹۳۰ در جایگاه خود به عنوان دستگاه منطقی اصلی مورد استفاده در ریاضی باقی ماند، اما در قالب نظریه‌ی انواع ساده (بدون انشعبات راسل): که همانطور که چویستک^۲، رمزی و دیگران پی برندن، برای یک مینا در سبکی مشابه پرینکیپیا کافی بود. رمزی دلایلی در جهت حذف نگرانی‌ها درباره نامحمولیت مطرح کرد، و تلاش کرد باقی اصول موضوعه‌ی وجودی پرینکیپیا — اصل بی‌نهایت و اصل انتخاب — را به عنوان اصولی منطقی توجیه کند، اما دلایل او ناکافی بودند. تلاش راسل برای نجات منطق‌گرایی از تناقض نیز، جز برای گروه محدودی از فلاسفه (به‌ویژه اعضای مکتب وین)، ناموفق بود.

پیشنهادات پوانکاره هم‌چنین جزئی کلیدی از رویکرد جالبی که ویل در کتاب ۱۹۱۸ خود *Das Kontinuum* در مورد مبانی ریاضیات پیش می‌برد بودند. ایده‌ی اصلی این بود که نظریه‌ی اعداد طبیعی را همان گونه که عرفاً پذیرفته بود — با استفاده از منطق کلاسیک — پیذیریم، اما پس از آن از روش‌های محمولی استفاده کنیم. بنابرین برخلاف براوئر، ویل اصل طرد شق ثالث را پذیرفت. (در بخش بعدی به تفصیل به این موضوع و هم‌چنین نظرات براوئر می‌پردازیم). اما اعداد حقیقی به طور کامل رام او نبودند، زیرا در دستگاه او مجموعه‌ی \mathbb{R} تمامیت نداشت و قضیه‌ی بولتسانو-وایرشراس قابل اثبات نبود. به همین دلیل او ناچار به ابداع جایگزین‌های پیچیده‌ای برای اثبات‌های معمول حکم‌های آنالیزی شد.

ایده‌ی مبانی محمولی برای ریاضیات، به سبک ویل، در دهه‌های اخیر منجر به حصول نتایج قابل توجهی شده است (ر.ک. ففرمن^۳ ۱۹۹۸). دستگاه‌های محمولی جایگاهی در میان دستگاه‌هایی که با سخت‌گیری از روش‌های ساختی حمایت می‌کنند، و آن‌ها که با سرسختی پشتیبان روش‌های جدید هستند دارند. این رویکرد به مبانی ریاضیات، یکی از بسیار رویکردهایی است که در سه‌گانه‌ی مرسوم اما کهنه‌ی منطق‌گرایی، صورت‌گرایی و شهود‌گرایی نمی‌گنجد.

۲.۲. انتخاب‌ها.

^۱ استدلال کوشی کاملاً ناساختی، یا آنطور که در این نوشته گفته‌ایم «کاملاً وجودی» بود. برای نشان دادن این که چندجمله‌ای مورد نظر باید حداقل یک ریشه داشته باشد، کوشی از مقدار چندجمله‌ای استفاده کرد. این مقدار یک کمینه‌ی سراسری مانند s دارد، که تعریف این مقدار کمینه نامحمولی است. او سپس از این فرض که s مثبت است استفاده کرد و به تناقض رسید.

² Chwistek

³ Feferman

با تمام اهمیتی که پارادوکس‌ها داشتند، تاثیر آن‌ها بر مناظرات بر سر مبانی، معمولاً بیش‌پنداشته می‌شود. روایات بسیاری هستند که برخلاف بررسی ما در بخش ۱، از آن‌ها به عنوان نقطه‌ی شروع واقعی مجادلات یاد می‌کنند. با این حال حتی اگر توجه‌مان را به دهه‌ی اول قرن بیستم محدود کنیم، همچنان مشاجره‌ای دیگر، بر سر اصل انتخاب و اثبات زرملو از قضیه‌ی خوش ترتیبی، وجود دارد که اگر نگوییم مهم‌تر، به همان اندازه مهم است.

از قسمت ۱.۲ به یاد داریم که ارتباط بین مجموعه‌ها و مشخصه‌های معرف آن‌ها، در آن زمان، (به واسطه‌ی اصل متناقض تفهیم) در اذهان هم ریاضی‌دانان و هم منطق‌دانان ریشه دوانده بود. اصل انتخاب عبارت است از این که برای هر خانواده‌ی نامتناهی از مجموعه‌های مجزا و ناتهی، مجموعه‌ای (مجموعه‌ی انتخاب) وجود دارد که از هر یک از اعضای خانواده، دقیقاً یک عضو در آن وجود دارد. مشکل منتقدان با این اصل، این بود که صرفاً وجود مجموعه‌ی انتخاب را مقرر می‌کند، اما تعریف آن را ارائه نمی‌دهد. اتفاقاً در موقعی که مشخص کردن مجموعه‌ی انتخاب به طور صریح ممکن است، اصلاً نیازی به بکارگیری اصل انتخاب نیست! از طرفی اثبات زرملو برای قضیه‌ی خوش‌ترتیبی به استفاده از این اصل نیاز دارد. خوش‌ترتیبی مورد بحث برای \mathbb{R} ، به معنای ایده‌آل مورد نظر کانتور، ددکیند و هیلبرت «وجود» دارد، اما در یک چشم‌انداز ساختی غیرقابل دسترس است. این شد که اصل انتخاب، به ابهام‌های موجود درباره‌ی نظریه‌ی مجموعه‌ها شدت بخشد، و ریاضی‌دانان را ناچار به شفاف‌سازی کرد. از طرفی، اصل انتخاب چیزی بیش از بیان صریح دیدگاه‌های موجود درباره‌ی زیرمجموعه‌های دلخواه نبود، اما از طرف دیگر ایده‌ی آن به وضوح با این دیدگاه که مجموعه‌های نامتناهی را باید با ویژگی‌هایی که صریحاً تعريف کرد اختلاف داشت. صحنه برای مناظره‌ای عمیق آماده بود. بحث‌های حول این موضوع، بیش از هر چیز، پیامدهای وجودی روش‌های جدید ریاضیات را روشن کرد. حتی بورل [۱، VI.70]، بیر، و لیگ [۱، VI.72]، که بعداً منتقد این اصل شده بودند، همگی به طرقی که کمتر آشکار بودند، در اثبات قضیه‌های آنالیز از اصل انتخاب بهره برده بودند. اتفاقی نبود که این اصل را ارهاردد اشمیت، که شاگرد هیلبرت و یک آنالیزکار بود، به زرملو پیشنهاد کرده بود.

پس از انتشار اثبات زرملو، بحث شدیدی در اروپا در گرفت. زرملو تصمیم به کار کردن روی مبانی نظریه‌ی مجموعه‌ها گرفته بود تا نشان دهد که اثباتش در یک دستگاه مقبول و اعتراض‌ناپذیر از اصول موضوعه معتبر است. نتیجه‌ی تلاش او، دستگاه اصول موضوعه‌ی مشهورش [۱، VI.22 §3] شد که حاصل تحلیل هشیارانه‌ی نظریه‌ی مجموعه‌ها — آن طور که کانتور و ددکیند عقیده داشتند و در قضیه‌ی خودش هم وجود داشت — بود. پس از برخی اصلاحات (اصول جایگزینی و انتظام) که به پیشنهاد فرانکل و فون نویمان [۱، VI.91] انجام شد و ایده‌ی نوآورانه‌ی مهمی که توسط ویل و اشکولم [۱، VI.81] مطرح شد (صورت‌بندی آن در منطق مرتبه اول [۱، IV.23 §1]، یعنی تسویر روی اشیا منفرد — مجموعه‌ها — و نه روی ویژگی‌هایی که شکلی را گرفت که امروزه می‌شناسیم)، دستگاه زرملو در دهه‌ی ۱۹۲۰ همان شکلی را گرفت که امروزه می‌شناسیم.

دستگاه ZFC (حروف ابتدای اسمی زرملو و فرانکل، و C برای اصل انتخاب) جنبه‌های اصلی روش‌شناختی ریاضیات مدرن را در بر گرفته، و چارچوبی مطبوع برای توسعه‌ی نظریات ریاضی و پرداختن به اثبات‌ها ارائه می‌کند و به ویژه، دارای اصول وجودی قدرتمند است، امکان ارائه‌ی تعاریف نامحمولی و توابع دلخواه را فراهم می‌سازد، اثبات‌های کاملاً وجودی را می‌پذیرد، و تعريف مناسب ساختارهای اصلی ریاضیات در آن ممکن است. به این دلایل، همه‌ی ویژگی‌های چهارگانه‌ی گفته شده در بخش ۱ را دارد. کار زرملو کاملاً در راستای تلاش‌های غیررسمی هیلبرت در حوالی ۱۹۰۰ برای اصل موضوعی سازی بود، و فراموش نکرد اثباتی برای سازگاری دستگاهش را نیز وعده دهد. نظریه‌ی اصل موضوعی مجموعه‌ها، خواه به بیان زرملو-فرانکل و خواه به بیان فون نویمان-برنیز-گودل، دستگاهی است که اکثر ریاضی‌دانان به عنوان مبنای کارا برای نظام خود می‌پذیرند.

تا ۱۹۱۰، تقابلی شدید بین نظریه‌ی انواع راسل و نظریه‌ی مجموعه‌های زرملو وجود داشت. اولی در چارچوب منطق صوری پروردۀ شده بود، و محل انحرافش (که البته بعداً به دلایل پراگماتیک به خطر افتاد) در راستای محمولیت بود. برای رسیدن به ریاضیات، نیاز بود فرض‌های وجودی بی‌نهایت و انتخاب را پذیریم، اما این دو نه اصولی آشکار، بلکه فرضیاتی تجربی تلقی می‌شدند. دومی دستگاهی بود که ابتدا به شکل غیررسمی پیشنهاد شده بود، به دیدگاه نامحمولی ایمان کامل داشت، و اصولش فرض‌های وجودی قدرتمندی بودند که برای اشتقاء کل ریاضیات کلاسیک و علاوه بر آن نظریه‌ی بی‌نهایت متعالی کانتور کفایت می‌کردند. در دهه‌ی ۱۹۲۰ فاصله‌ی این دو بسیار کاهش یافت، به خصوص در مورد دو ویژگی اولی که مطرح کردیم. دستگاه زرملو به کمال رسید و در قالب زبان منطق صوری مدرن صورت‌بندی شد. راسل‌گراها نیز نظریه‌ی انواع ساده را پذیرفته، و در جریان این پذیرش روش «وجودی» و نامحمولی ریاضیات مدرن را قبول کردند. این موضوع معمولاً (شاید به

طرزی گیج‌کننده) «افلاطون‌گرایی» خوانده می‌شود: به اشیا مورد بحث نظریه طوری نگاه می‌شود که انگار وابستگی‌بی به آنچه ریاضی‌دان می‌تواند واقعاً و به طور صریح تعریف کند ندارند.

در همان دهه‌ی اول قرن بیستم، ریاضی‌دانی هلندی در مسیر رسیدن به نوعی شهود‌گرایی قرار گرفته بود که از نظر فلسفی غنی‌تر بود. براوئر جوان در ۱۹۰۵ عقاید عجیب و غریب در متافیزیک و اخلاقیات را بیان کرد، و در همان راستا مبانی ریاضیات خود را در تزش در ۱۹۰۷ توضیح داد. فلسفه‌ی «شهود‌گرایی» او در این دیدگاه متافیزیکی‌اش ریشه داشت که آگاهی فرد منبع یگانه و یکتای دانش است. این نگاه به خودی خود چندان جذاب نمی‌نماید، بنابرین بهتر است بیشتر روی عقاید ساختگرایانه‌ی براوئر تمرکز کنیم. در حوالی ۱۹۱۰، براوئر — از جمله به دلیل نتایج حیاتی‌بی در توبیلوژی مانند قضیه‌ی نقطه‌ی ثابت [۱، V.11] — به ریاضی‌دانی شناخته‌شده تبدیل شد. هنگامی که نخستین جنگ جهانی به پایان رسید، او شروع به انتشار جزئیات ایده‌هایش در مورد مبانی ریاضی کرد، و این اقدامش نقش بزرگی در ایجاد «بحaran» معروف — که حال به آن می‌پردازم — داشت. او هم‌چنین موفق شد مرز مرسوم (اما گمراه‌کننده) میان صورت‌گرایی و شهود‌گرایی را جا بیندازد.

مراجع

[1] Gowers, T., Barrow-Green, J. & Leader, I. (2009). *The Princeton Companion to Mathematics*. Princeton: Princeton University Press.

مترجم: پارسا ترتی[†]

دانشجوی کارشناسی علوم کامپیوتر، دانشگاه صنعتی شریف
ریانامه: ptorbatii@icloud.com



۵۰ سال پیچیدگی محاسبه یک گزارش خبری

بن برویکر*

چکیده. چه قدر دشوار است که ثابت کنیم حل مسئله‌ای سخت است؟ نظریه‌پردازان فرای پیچیدگی دهه‌هاست که به پرسیدن چنین سوال‌هایی مشغولند و سلسله‌ای از نتایج به یافتن پاسخ‌هایی انجامیده است. با این حال متخصصان پیچیدگی محاسبه هنوز در حال دست‌وینجه نرم کردن با دشوارترین مسئله‌ی خود هستند: خود نظریه‌ی پیچیدگی!

۱. خاستگاه

در اولین هفته‌ی ترم پاییز ۲۰۰۷، مارکو کارموسینو^۱ خود را به کلاس درس ریاضی‌ای کشاند که برای همه‌ی دانشجویان علوم کامپیوتر دانشگاه ماساچوست امهرست اجباری بود. کارموسینو یک دانشجوی سال دومی بود که می‌خواست برای آن که طراح بازی‌های کامپیوترا شود، ادامه‌ی تحصیل را رها کند. با این حال در آن کلاس بود که استاد با سوالی ساده مسیر زندگی او را تغییر داد: «از کجا می‌دانید که ریاضیات واقعاً کار می‌کند؟». کارموسینو که هم اکنون یک متخصص علوم کامپیوتر نظری در IBM است، به یاد می‌آورد: «آن سوال باعث شد که بنشینم و توجه کنم». او درس سینیاری اختیاری درباره‌ی آثار گودل برداشت. گودلی که استدلال‌های خود را جاعده‌اش برای اولین بار محدودیت‌های استدلال ریاضی را آشکار کرده بود و زیربنایی را برای تمام نتایج بعدی درباره‌ی محدودیت‌های اساسی محاسبه پی‌ریزی کرده بود. هضم آن همه مطلب برای کارموسینو دشوار بود. «صد درصدش را نفهمیدم، اما می‌دانستم که می‌خواهم بفهمم!».



شکل ۱. جایزه‌ی میلیون دلاری من کجاست؟! (منبع تصویر: [۱])

امروز حتی محققان کارکشته نیز وقتی با مسئله‌ی باز مرکزی علوم کامپیوتر نظری، یعنی P در برابر NP مواجه می‌شوند، نمی‌دانند که چطور می‌توانند به آن جوابی بدهنند. این پرسش درباره‌ی این است که آیا ممکن است تعدادی از مسائل محاسباتی که مدت‌هاست دشوار تلقی شده‌اند به سادگی (از راه میانبری مخفی که هنوز کشف نکردۀایم) قابل حل باشند، یا این که همان‌طور که اکثر محققان گمان می‌کنند، آن‌ها واقعاً دشوارند؟ در حقیقت این مسئله درباره‌ی چیزی جز ذات آن چه قابل دانستن است، نیست.

علی‌رغم دهه‌ها تلاش محققان پیچیدگی محاسبه^۲ — که به مطالعه‌ی چنین سؤالاتی درباره‌ی دشواری ذاتی مسائل مختلف پرداخته‌اند — پاسخ مسئله‌ی P در برابر NP هنوز ناشناخته است، و حتی مشخص نیست که برای یک اثبات احتمالی از کجا

*این نوشته ترجمه‌ای از مقاله‌ی زیر است:

Ben Brubaker (2023) Complexity Theory's 50-Year Journey to the Limits of Knowledge. Quanta Magazine.

¹Marco Carmosino

²computational complexity theory

باید شروع کرد. مایکل سیپسر^۱ می‌گوید: «هیچ نقشه‌ی راهی وجود ندارد..». او یک متخصص کهنه‌کار پیچیدگی محاسبه در MIT است که در تلاش برای حل این مسأله سال‌ها وقت صرف کرده است. او ادامه می‌دهد: «مثل این است که به برهوت بروی..».

این طور که به نظر می‌رسد اثبات این‌که حل برخی مسائل محاسباتی دشوار است، خود مسأله‌ای سخت است؛ اما چرا و چقدر؟ کارموسینو و دیگر محققان حوزه‌ی فرایچیدگی با چرخاندن لنز دوربین به سمت خود نظریه‌ی پیچیدگی، سؤالاتی این چنینی را مجدداً به صورت مسائل پیچیدگی فرمول‌بندی کردند تا تحقیقات این حوزه را پیش ببرند. راهول ایلانگو^۲، دانشجوی تحصیلات تکمیلی در MIT، که به برخی از هیجان‌انگیزترین دست آورده‌ای اخیر در این زمینه دست یافته است، درباره‌ی این رویکرد می‌گوید: «ممکن است فکر کنید جالب است. شاید هم فکر کنید متخصصان پیچیدگی محاسبه عقلشان را از دست داده‌اند!».

با مطالعه‌ی این مسائل درون‌نگرانه، پژوهشگران یادگرفته‌اند که دشواری اثبات سختی مسائل، با سوال‌هایی بنیادی که در وهله‌ی اول ممکن است نامرتبط به نظر برسند، گره خورده است: تشخیص الگوهای پنهان در داده‌های به ظاهر تصادفی چه قدر مشکل است؟ و اگر مسائل واقعاً مشکل وجود دارند، تعدادشان چقدر زیاد است؟ اسکات آرانسون^۳، متخصص پیچیدگی در دانشگاه تگزاس آستین^۴، می‌گوید: «روشن شده که فرایچیدگی به حقیقت نزدیک است..».

این نوشته داستان مسیر طولانی و پریچ‌وخمی است که محققان را از مسأله‌ی P در برابر \NP به فرایچیدگی سوق داد. مسیری پر از پیچ‌های اشتباه و جاده‌های مسدود که دویاره و دویاره به نقطه‌ی اول خود برمی‌گشتند. با این حال برای متخصصان فرایچیدگی این سفر به سرزمین‌های ناشناخته خود پاداش خود است. والتنین کابانتس^۵، متخصص پیچیدگی محاسبه در دانشگاه سایمون فریزر^۶ کانادا می‌گوید: «شروع کنید به پرسیدن سوالات به ظاهر ساده..»، و ادامه می‌دهد: «هیچ ایده‌ای ندارید که قرار است از کجا سر در بیاورید..».

۱.۱. ناشناخته‌های شناخته‌شده. مسأله‌ی P در برابر \NP نام بدون زرق و برقص را مدیون عادت نظریه‌پردازان پیچیدگی به دسته‌بندی کردن مسائل محاسباتی در گستره‌ای از کلاس‌های پیچیدگی است. یک مسأله‌ی محاسباتی، مسأله‌ای است که بتوان آن را با یک الگوریتم — یا به بیانی ساده با دنباله‌ی دقیقی از دستورالعمل‌ها — حل کرد. با این حال تمام الگوریتم‌ها به طور یکسانی مفید نیستند، و تنوع میان آن‌ها اشاره به تفاوت‌های بنیادی میان مسائل کلاس‌های مختلف دارد. چالش نظریه‌پردازان پیچیدگی این است که این اشارات را به قضیه‌هایی دقیق درباره‌ی روابط میان کلاس‌های پیچیدگی تبدیل کنند. این روابط — که دامنه‌ی آن‌ها فراتر از مزه‌های همه‌ی فناوری‌هاست — منعکس‌کننده‌ی حقایقی تغییرناپذیر درباره‌ی مفهوم محاسبه هستند. کابانتس می‌گوید: «این کار مانند کشف کردن قوانین کیهان است..».

P و \NP دو نمونه از مشهورترین اعضای یک باع‌وحش در حال توسعه از صدھا کلاس پیچیدگی هستند. صرف نظر از جزئیات، P کلاس مسائلی است که به راحتی^۷ با یک الگوریتم حل می‌شوند، مانند مرتب کردن الفبایی یک لیست. \NP کلاس مسائلی است که درستی جواب‌هایش به راحتی قابل بررسی هستند، مانند حل سودوکو. از آنجا که تمام مسائلی که به راحتی حل می‌شوند به راحتی هم جواب‌هایشان قابل بررسی هستند، مسائل عضو P عضو \NP هم هستند. با این وجود، حل برخی مسائل \NP سخت به نظر می‌رسد؛ مثلاً در سودوکو بدون این که ابتدا بسیاری از حالات را امتحان کنید، نمی‌توانید جواب نهایی را بدست آورید^۸. آیا ممکن است که این سختی ظاهری، فقط یک توهم باشد و یک ترفند ساده برای حل تمام مسائلی که درستی جواب‌شان به راحتی قابل بررسی است، وجود داشته باشد؟

اگر چنین باشد، آنگاه $\NP = P$: یعنی هر دو دسته مسأله با هم معادلنند. در این صورت باید الگوریتمی وجود داشته باشد که حل سودوکوهای عظیم، بهینه‌سازی مسیرهای حمل و نقل جهانی، شکستن پیشرفته‌ترین رمزها و ماشینی کردن اثبات قضایایی

¹Michael Sipser

²Rahul Ilango

³Scott Aaronson

⁴University of Texas, Austin

⁵Valentine Kabanets

⁶Simon Fraser University

⁷ در اینجا مقصود از حل پذیری راحت این است که الگوریتمی وجود دارد که در زمانی که بر حسب طول ورودی مسأله کران بالایی چند جمله‌ای دارد، مسأله را حل می‌کند [م..].

⁸ به عبارت بهتر، به نظر می‌رسد برای هر الگوریتم که برای حل سودوکو وجود دارد، جدول‌های سودوکویی قابل تصورند که حل آن‌ها با الگوریتم مزبور آسان نیست [م..].

ریاضی^۱ را به سادگی انجام دهد. اگر $\mathcal{N}\mathcal{P} \neq \mathcal{P}$ ، آنگاه بسیاری از مسائل محاسباتی که از نظر تئوری حل شدنی هستند، در عمل قابل حل نخواهند بود.

محققان مدت‌ها قبل از این‌که مسئله‌ی \mathcal{P} در برابر $\mathcal{N}\mathcal{P}$ برای اولین‌بار بیان شود — در حقیقت مدت‌ها قبل‌تر از شروع علوم کامپیوتر مدرن — نگران محدودیت‌های استدلال‌های صوری ریاضی بودند. در ۱۹۲۱، دیوید هیلبرت^۲ ریاضیدان در تلاش برای پاسخ‌دادن به همان سوالی که نزدیک به یک قرن بعد توجه کارموسینو را جلب کرد، برای پی‌ریزی ریاضیاتی دقیق یک برنامه‌ی تحقیقاتی پیشنهاد کرد. او امید داشت که با شروع از تعداد کمی فرض ساده — به نام اصول موضوعه^۳ — یک نظریه‌ی ریاضی یک‌پارچه^۴ استخراج کند که دارای سه معیار کلیدی باشد. شرط اول هیلبرت سازگاری^۵ بود، شرطی ضروری برای این‌که ریاضیات عاری از تناقضات باشد: اگر دو گزاره‌ی متناقض را بتوان با شروع از اصول موضوعه‌ی یکسانی اثبات کرد، تمام نظریه‌ی غیرقابل نجات خواهد بود. با این حال یک قضیه‌ی می‌تواند عاری از تناقض، ولی هم‌چنان دست‌نیافتنی باشد. این انگیزه‌ای بود برای شرط دوم هیلبرت، تمامیت^۶: التزام به این‌که تمام گزاره‌های ریاضی یا به طور قابل اثباتی درست باشند و یا نادرست. معیار سوم او، تصمیم‌پذیری^۷، خواستار یک رویه‌ی بدون ابهام مکانیکی برای تعیین درست یا نادرست بودن هر گزاره‌ی ریاضی بود. هیلبرت در کنفرانسی در ۱۹۳۰ اعلام کرد: «شعار ما این است: ما باید بدانیم؛ ما خواهیم دانست.».

تها یک سال بعد گودل اولین ضریب را به رویای هیلبرت وارد کرد. او اثبات کرد که یک گزاره مانند «این گزاره اثبات‌پذیر نیست.» می‌تواند از هر مجموعه‌ی مناسی از اصول موضوعه استنتاج^۸ شود. اگر چنین گزاره‌ای به راستی اثبات‌پذیر نباشد، تمامیت نقض می‌شود و اگر اثبات‌پذیر باشد، نظریه ناسازگار خواهد بود، که نتیجه‌ی بدتری است. همچنین در همان مقاله، گودل ثابت کرد که هیچ نظریه‌ی ریاضی‌ای هرگز نمی‌تواند سازگاری خودش را ثابت کند.^۹

محققان هنوز امید داشتند که ممکن است در آینده نظریه‌ای

ریاضی یافت شود که تصمیم‌پذیر باشد، هر چند چنین نظریه‌ای لزوماً ناتمام خواهد بود. ممکن است آن‌ها بتوانند روش‌هایی را توسعه دهنده که در حالی که از گزاره‌های آزاردهنده‌ای مانند گزاره‌های گودل دوری می‌کند، تمام گزاره‌های اثبات‌پذیر را شناسایی کند. مشکل آن‌جا بود که هیچ کس نمی‌دانست چگونه درباره‌ی این روش‌ها استدلال کند.



شکل ۲. در دهه‌ی ۱۹۲۰، هیلبرت (تصویر چپ) قصد داشت ریاضیات را بر مبانی محکم‌تری استوار کند. گودل (تصویر وسط) و تورینگ (تصویر راست) نشان دادند که رویای هیلبرت غیرممکن است.

بعدتر در ۱۹۳۶، یک دانشجوی تحصیلات تکمیلی ۲۳ ساله به نام آلن تورینگ^{۱۰}، شرط تصمیم‌پذیری هیلبرت را به زبان در آن

زمان ناآشنای محاسبه بازنویسی کرد و ضریه‌ی مهلکی بر آن وارد کرد. تورینگ یک مدل ریاضی را فرمول‌بندی کرد که امروزه به نام ماشین تورینگ شناخته می‌شود، و می‌تواند تمام الگوریتم‌های ممکن را ارائه کند، و نشان داد اگر روش مکانیکی هیلبرت وجود داشته باشد، با این مدل قابل توصیف خواهد بود. او سپس از روش‌های خودارجاعی، مانند روش‌های گودل، استفاده کرد تا وجود گزاره‌های تصمیم‌نای‌پذیر را ثابت کند [۲]، یا معادلاً نشان داد مسائلی وجود دارند که هیچ الگوریتمی قادر به حلشان نیست. برنامه‌ی هیلبرت ویران شده بود: محدودیت‌های اساسی و همیشگی برای آنچه می‌توان اثبات کرد و آنچه می‌توان محاسبه کرد وجود خواهد داشت. با این حال هنگامی که کامپیوترا از اشیاء انتزاعی نظری به دستگاه‌های واقعی تبدیل شدند، محققان متوجه شدند که تمایز ساده‌ی تورینگ بین مسائل حل شدنی و حل نشدنی بسیاری از سوالات را بی‌پاسخ گذاشته است. تا دهه‌ی ۱۹۶۰، محققان علوم کامپیوترا الگوریتم‌های سریعی را برای حل کردن برخی مسائل توسعه داده بودند، و این در حالی بود که برای باقی مسائل الگوریتم‌های شناخته‌شده به طرز طاقت‌فرسایی کند بود. چه می‌شد اگر سوال فقط این نبود که آیا مسائل قابل

^۱Automated theorem proving (ATP)

^۲David Hilbert

^۳Axioms

^۴Unified mathematical theory

^۵Consistency

^۶Completeness

^۷Decidability

^۸Derivation

^۹به بیان دقیق‌تر، گودل نشان داد که هیچ نظریه‌ی ریاضی‌ای که اقلًا شامل حساب باشد — با فرض سازگاری این نظریه — نمی‌تواند سازگاری خودش را ثابت کند [۳].

^{۱۰}Alan Turing

حل هستند، بلکه این بود که حل آنها چقدر سخت است؟ کارموسینو می‌گوید: «یک نظریه‌ی غنی پدیدار شد و ما دیگر جواب‌ها را نمی‌دانیم.».

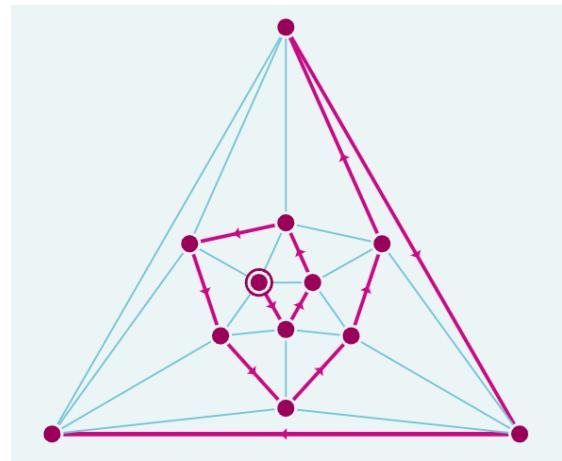
۲۰.۱. مسیرهای واگرا. برای این‌که نشان دهیم سوال کردن در مورد «سختی» تا چه اندازه می‌تواند بعرنج باشد، بگذارید دو مسئله در مورد گراف‌ها را که به طور نزدیکی با هم در ارتباطند در نظر بگیریم. گراف‌ها شبکه‌ای از نقاط یا رأس‌ها^۱ هستند که با خطوط یا یال‌هایی^۲ به هم وصل شده‌اند. محققان علوم کامپیوتر از گراف‌ها استفاده می‌کنند تا همه چیز—از محاسبات کوانتومی گرفته تا جریان ترافیک—را با آن مدل کنند.

فرض کنید گرافی به شما داده شده و از شما خواسته شده است که یک مسیر همیلتونی^۳ را در آن بیابید؛ مسیری که از تمام رأس‌ها دقیقاً یک‌بار می‌گذرد. این مسئله به وضوح از نظر تئوری حل شدنی است: تنها تعدادی متناهی مسیر ممکن وجود دارد؛ بنابراین می‌توانید هر یک از مسیرها را بررسی کنید. اگر تعداد کمی رأس وجود داشته باشد این روش ایده‌ی بدی نیست، اما وقتی اندازه‌ی گراف فقط کمی بزرگ‌تر می‌شود، تعداد حالت‌ها از کنترل خارج می‌شود و به سرعت این الگوریتم ساده را بی‌فایده می‌کند. راسل ایمپاگلیازو^۴ می‌گوید: «الگوریتم‌های مسیر همیلتونی

پیچیده‌تری وجود دارند که با این مشکل سرخشنانه‌تر مبارزه می‌کنند، اما زمانی که الگوریتم نیاز دارد که مسئله را حل کند، متناسب با اندازه‌ی گراف همیشه به طور نمایی رشد می‌کند. حتی قبل از این‌که گراف خیلی بزرگ شده باشد، بهترین الگوریتم‌هایی که محققان کشف کرده‌اند نیز نمی‌توانند مسیر را در زمان معقولی پیدا کنند. ایمپاگلیازو یک نظریه‌پرداز پیچیدگی در دانشگاه کالیفرنیا، سن دیگو^۵ است. او ادامه می‌دهد: «و از زمان معقول، منظورم قبل از پایان کیهان است.».

مسئله‌ی مسیر همیلتونی یک خاصیت جالب دیگر هم دارد. اگر کسی ادعا کند که یک مسیر همیلتونی در یک گراف خاص پیدا کرده است، شما می‌توانید به سرعت بررسی کنید که جواب او معتبر است یا نه، حتی اگر گراف خیلی بزرگ باشد. تنها کاری که لازم است انجام دهید این است که مسیر را دنبال کنید و هر رأس را یک‌یکی علامت بزنید، و نهایتاً بررسی کنید تا مطمئن شوید که هر رأس را دوبار علامت نزدیکی داشته باشد، آنگاه مسیر همیلتونی است. زمان لازم برای اجرای این الگوریتم بررسی جواب متناسب با اندازه‌ی گراف است، که آن را در دسته‌ی وسیع‌تری از الگوریتم‌های چندجمله‌ای قرار می‌دهد که زمان‌های اجرایشان متناسب با توابعی چندجمله‌ای از اندازه‌ی گراف افزایش پیدا می‌کند. رشد چندجمله‌ای در مقایسه با رشد نمایی رامتر است، در نتیجه الگوریتم‌های چندجمله‌ای حتی برای گراف‌های بزرگ نیز قابل اجرا خواهد بود. کارموسینو می‌گوید: «آن‌ها به طور چشمگیری کارآمدترند.».

مسئله‌ی مسیر همیلتونی عدم تقارن آشکاری دارد: شما می‌توانید یک جواب درست را با یک الگوریتم چندجمله‌ای سریع تایید کنید، اما برای پیداکردن جواب به یک الگوریتم کند نمایی نیاز دارید. این عدم تقارن ممکن است تعجب آور به نظر نرسد—تشخیص یک شاهکار هنری راحت‌تر از خلق‌کردنش است، یا بررسی یک اثبات ریاضی از اثبات‌کردن یک قضیه جدید راحت‌تر است. با این حال، مسئله‌ی دیگری که بسیار شبیه به مسیر همیلتونی است، کاملاً متفاوت رفتار می‌کند. مجدداً فرض کنید گرافی به شما داده شده است، اما این بار از شما خواسته شده که یک مسیر اویلری^۶ پیدا کنید—مسیری که از تمام یال‌ها دقیقاً یک بار می‌گذرد. مجدداً می‌توان دید که یک الگوریتم چندجمله‌ای برای بررسی جواب‌های ممکن وجود دارد، اما این بار



شکل ۳. مسیر همیلتونی، مسیری در گراف است که از هر رأس دقیقاً یک‌بار می‌گذرد.

¹Node

²Edge

³Hamiltonian path

⁴Russell Impagliazzo

⁵University of California, San Diego

⁶Eulerian path

برای حل مسئله هم یک الگوریتم چندجمله‌ای وجود دارد. هیچ عدم تقارنی اینجا نیست. به نظر می‌آید در نظریه‌ی پیچیدگی بعضی مسیرها راحت‌تر از بقیه پیدا می‌شوند.

هر دو مسئله‌ی مسیر همیلتونی و مسیر اویلری در کلاس پیچیدگی \mathcal{NP} قرار دارند — کلاسی که شامل مسائلی است که تمام جواب‌هایشان با الگوریتمی چندجمله‌ای قابل بررسی هستند. مسئله‌ی مسیر اویلری در کلاس \mathcal{P} هم قرار می‌گیرد؛ زیرا یک الگوریتم چندجمله‌ای می‌تواند حلقه کند. با این حال آن‌طور که به نظر می‌رسد این برای مسئله‌ی مسیر همیلتونی صادق نیست. چرا این دو مسئله، که به طور اعجاب‌آوری شبیه هستند، به شدت متفاوت‌اند؟ این ذات مسئله‌ی \mathcal{P} در برابر \mathcal{NP} است.

۳.۱ به طور جهانی سخت. در وله‌ی اول، به نظر می‌رسید کلاس‌های پیچیدگی دسته‌بندی‌های مناسبی برای مرتب‌سازی مسائلی هستند که به یکدیگر شبیه‌اند اما مستقیماً به هم مرتبط نیستند. هیچ کس شک نکرد که پیداکردن مسیرهای همیلتونی ارتباطی با دیگر مسائل سخت محاسباتی داشته باشد. سپس در ۱۹۷۱، استفن کوک^۱، ظرف یک سال نقل مکانش به دانشگاه تورنتو بعد از ردشدن درخواست استاد تمامی اش در ایالات متحده، نتیجه‌ی فوق العاده‌ای را منتشر کرد^[۲]. او مسئله‌ی \mathcal{NP} خاصی را با یک ویژگی عجیب شناسایی کرده بود: اگر الگوریتمی چندجمله‌ای وجود داشته باشد که بتواند آن مسئله را حل کند، آنگاه می‌تواند هر مسئله دیگر \mathcal{NP} را نیز حل کند. به نظر می‌رسید مسئله‌ی «جهانی» کوک یکه ستونی است که مسائل ظاهرآ سخت را بالا نگه می‌دارد و آن‌ها را از مسائل راحت زیرشان جدا می‌کند. آن مسئله را حل کنید، و باقی \mathcal{NP} فرو می‌ریزد و پایین می‌آید. کوک گمان می‌کرد که هیچ الگوریتم سریعی برای مسئله‌ی جهانی اش وجود ندارد، و در نیمه‌های مقاله‌اش گفته بود: «من احساس می‌کنم اثبات این حدس ارزش تلاش قابل ملاحظه‌ای را دارد». به نظر می‌رسد که «تلاش قابل ملاحظه» دست کم گرفتن سختی مسئله بود. تقریباً در همان زمان، یک دانشجوی کارشناسی در اتحاد جماهیر شوروی به نام لئونید لوین^۲، نتیجه‌ی مشابهی را ثابت کرد^[۳] و افزون بر این، چندین مسئله‌ی جهانی متفاوت را نیز شناسایی کرد. هم‌چنین نظریه‌پرداز پیچیدگی امریکایی ریچارد کارپ^۳، ثابت کرد^[۴] که خاصیت جهانی بودن که توسط کوک (ولوین، اگرچه کوک و کارپ از کارهای لوین تا سال‌ها بعد اطلاعی نداشتند) شناسایی شده بود، به خودی خود جهانی است. تقریباً هر مسئله‌ی \mathcal{NP} بودن یک الگوریتم چندجمله‌ای — یعنی تقریباً تمام مسائلی که سخت به نظر می‌رسیدند — خاصیت یکسانی داشتند که به \mathcal{NP} -کامل بودن^[۴] معروف شد. این یعنی تمام مسائل \mathcal{NP} -کامل — مسیر همیلتونی، سودوکو و هزاران چیز دیگر — به معنای دقیقی معادل هستند. ایلانگو می‌گوید: «شما تمام این مسئله‌های طبیعی مختلف را در اختیار دارید و حالا معلوم می‌شود که تمامشان به نحوی جادویی سوالی یکسان بودند، و هنوز نمی‌دانیم که آیا همان سوال حل می‌شود یا نه».

حل وفصل کردن سختی هر مسئله‌ی \mathcal{NP} -کامل برای حل سوال \mathcal{P} در برابر \mathcal{NP} کافی خواهد بود. اگر $\mathcal{P} \neq \mathcal{NP}$ باشد، تمایز میان مسائل سخت و آسان توسط صدھا ستون که به یک اندازه قوی هستند نگه داشته می‌شود. اگر $\mathcal{P} = \mathcal{NP}$ باشد، فروپاشی این عمارت لرzan در انتظار کوچک‌ترین تلنگر خواهد بود. به این ترتیب کوک، لوین و کارپ مسائل بسیاری را که به نظر می‌رسید نامرتبه باشند، یکی کردنند. حالا نظریه‌پردازان پیچیدگی تنها یک مسئله را باید حل می‌کردن: $\mathcal{NP} = \mathcal{P}$ یا نه؟ پنجاه سال گذشته و این سوال هم‌چنان بی‌پاسخ مانده است. کابانتس استدلال‌های حول محدودیت‌های محاسبه را به بررسی یک قلمروی وسیع، بدون درک جامعی از عواقب این کار، تشییه می‌کند. موجودی با قدرت محاسباتی نامحدود می‌تواند از قله‌ی کوه به پایین نگاه کند و تمام چشم‌انداز را به یکباره ببیند، اما موجودات فانی ناچیز نمی‌توانند روی چینین مزیتی حساب کنند. او می‌گوید: «ما در پایین آن کوه می‌توانیم برای کمی بهتر دیدن بالا و پایین پیریم».

فرض کنید $\mathcal{P} = \mathcal{NP}$ باشد. برای اثبات، محققان باید الگوریتم سریعی برای یک مسئله‌ی \mathcal{NP} -کامل پیدا کنند که ممکن است در گوشه‌ای از آن چشم‌انداز پنهان شده باشد. هیچ ضمانتی وجود ندارد که به این زودی‌ها پیدایش کنند: نظریه‌پردازان پیچیدگی گاه‌وی گاه بعد از دھه‌ها کار، الگوریتم‌های هوشمندانه‌ای برای مسائل ظاهرآ سخت (ونه \mathcal{NP} -کامل) کشف کرده‌اند. حال فرض کنید که $\mathcal{NP} \neq \mathcal{P}$ باشد. اثبات آن حتی سخت‌تر به نظر می‌رسد. نظریه‌پردازان پیچیدگی باید نشان دهند که هیچ الگوریتم سریعی نمی‌تواند برای حل مسائل به ظاهر سخت وجود داشته باشد.

ندانستن این که از کجا باید شروع کیم بخشی از مشکل است؛ اما این‌گونه نیست که محققان هیچ تلاشی نکرده باشند. آن‌ها طی دھه‌ها از جهت‌های بسیاری به این مسئله حمله کرده‌اند و در هر مسیر به بن بست رسیده‌اند. کارموسینو می‌گوید:

¹Stephen Cook

²Leonid Levin

³Richard Karp

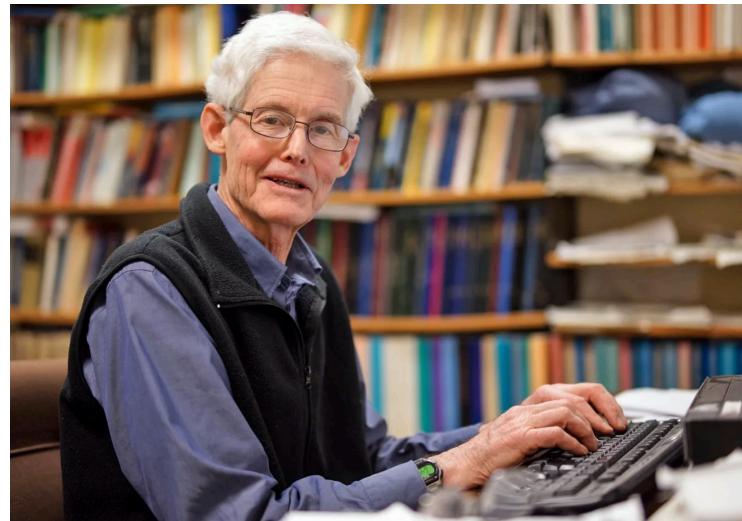
⁴NP-Completeness

«این یکی از آشکارترین حقایق در علوم کامپیوتر نظری است.» و ادامه می‌دهد: «وقتی با پدیده‌های مواجه می‌شوید که این قدر بادوام است، دلتنان می‌خواهد علتیش را بفهمید.».

۲. موضع

تا پایان آخرین سال تحصیل، حس کنجکاوی کارموسینو او را از گودل به یک کورس تحصیلات تکمیلی در نظریه‌ی پیچیدگی رسانده بود. او شگفت‌زده بود از این‌که متوجه شده بود زمانی که صرف انجام تکلیف‌هاییش می‌کرد از زمانی که برای انجام پروژه‌ای که از سر علاقه‌ی شخصی انجامش می‌داد — یک برنامه کامپیوتری که ساختار روابطی افسانه‌ها را یاد بگیرد و داستان‌های جدیدی تولید کند — بیشتر بوده است. کارموسینو به یاد می‌آورد: «با خودم فکر کردم: اووه! باید جدی بگیرم». طولی نکشید که او چنان جذب این موضوع شد که مشاورش با ملایمت پیشنهاد کرد در برنامه‌های پس از فارغ‌التحصیلی اش تجدیدنظر کند. کارموسینو می‌گوید: «او به من گفت که اگر بخواهم این

کار — علوم کامپیوتر نظری — را ادامه دهم، کافی است تحصیلات تکمیلی را شروع کنم». او بعد از گرفتن فوق‌لیسانس، در ۲۰۱۲ برای آن‌که دکتری اش را زیر نظر ایمپاگلیازو انجام دهد، به سن دیگو رفت.



شكل ۴. استفن کوک هم‌راستا با کارپ و لوین، مسئله‌ی P در برابر NP را در اوایل دهه‌ی ۱۹۷۰ صورت‌بندی کرد.

در ابتدا هدف اصلی کارموسینو این بود که یک مقاله‌ی برجسته [۶] از دو دهه قبل را بهتر بفهمد. در آن مقاله الکساندر رازبروف^۱ و استیون رو دیچ^۲ نشان داده بودند که یک استراتژی «طبیعی» برای اثبات $P \neq NP$ تقریباً به طور قطع شکست می‌خورد، چون این موفقیت با هزینه‌ی گرافی — زیرورودن کامل رمزنگاری — به دست می‌آید، که محققان آن را بسیار دور از ذهن می‌دانند. محققان دست آورده رازبروف و رو دیچ را به عنوان یک مانع برای این رویکرد رایج برای اثبات $P \neq NP$ تفسیر کردند.

این «مانع اثبات‌های طبیعی»^۳ تنها یکی از بسیاری موضع شناخته‌شده برای حل کردن مسائل باز در نظریه‌ی پیچیدگی است. هر یک از این



شكل ۵. تصویری از مارکو کارموسینو. جذابیت نتیجه‌ای که در دهه‌ی ۱۹۹۰ به دست آمده بود، بیست سال بعد الهام‌بخش او شد تا دستاورد مهمی را در فرایچیدگی حاصل کند.

مانع مانند یک راهبند عمل می‌کند؛ هشدار می‌دهد که یک مسیر به ظاهر امیدوارکننده در واقع بنیست است. این موضع، با هم، نشان می‌دهند که هر اثباتی که مسئله‌ی P در برابر NP است، باید کاملاً با هر روشی که در گذشته استفاده شده است، متفاوت باشد؛ برای همین است که بیشتر محققان باور دارند جواب هنوز دور از دسترس است. با این حال موضع به ما می‌گویند

^۱Alexander Razborov

^۲Steven Rudich

^۳Natural proofs barrier

که کجا را نگاه نکنیم. ایلانگو می‌گوید: «نظریه‌ی پیچیدگی با مانع بسیار زیادی، هم نفرین شده است و هم مورد لطف قرار گرفته است.».

زمانی که کارموسینو با مانع اثبات‌های طبیعی مواجه شده بود، نزدیک به ۲۰ سال از عمر آن می‌گذشت. با این حال او احساس می‌کرد که مانع اثبات‌های طبیعی آموزه‌های بیشتری برای محققان دارند. آن احساس زمانی روشن شد که او و سه تن از همکارانش با بررسی مانع اثبات‌های طبیعی از منظر فرایپیچیدگی نتیجه‌ای اعجاب‌آور را اثبات کردند. اثبات آن‌ها یکی از محدود نتایج مهمی بود که باعث ایجاد علاقه‌ی جدیدی به فرایپیچیدگی و منجر به سیلی از پیشرفت‌ها در چند سال اخیر شده است. با این حال، برای دنبال‌کردن مسیری که از مانع اثبات‌های طبیعی به فرایپیچیدگی وجود دارد، باید به آن‌جا برگردیم که محققان را در دهه‌ی ۱۹۷۰ رها کردیم — آن زمان که برای اولین بار با مسئله‌ی P در برابر NP مواجه شدند. چه چیزی اثبات سختی مسائل را دشوار کرد؟

۱.۲. یک مسیر مداری. در آغاز محققان تلاش کردند به کمک اشکال دیگری از تکنیک‌هایی که تورینگ برای اثبات این‌که برخی مسائل توسط هیچ الگوریتمی حل شدنی نیستند، ثابت کنند $P \neq NP$ — به این معنا که ثابت کنند برخی مسائل NP با هیچ الگوریتمی با زمان چندجمله‌ای حل نمی‌شوند. با این حال به سرعت اثباتی را پیدا کردند که نشان می‌داد آن روش‌ها کار نمی‌کنند [۵]، و این اولین مانع بزرگ برای حل سوال P در برابر NP بود. آن‌ها در نتیجه شروع به جستجوی روش‌هایی دیگر کردند و طولی نکشید که روش دیگری در کار کلود شانون^۱، که معاصر تورینگ بود، پیدا شد.

غريب به نظر می‌رسيد که شانون که در شهر کوچکی در شمال میشیگان بزرگ شده بود، آغازگر عصر اطلاعات باشد. با این حال، او ماهیت میان‌رشته‌ای حوزه‌ی نوظهور علوم کامپیوتر را نمایان کرد، و این در حالی بود که او در مهندسی برق و منطق ریاضی هم چیره‌دست بود. شانون در پایان نامه‌ی ارشدش [۶] نشان داد که چگونه مدارهای ساخته‌شده از کلیدهای الکترومکانیکی می‌توانند نمایان‌گر عبارات منطقی شامل متغیرهای بولی باشند. در این عبارات، متغیرهای بولی با «گیت‌های منطقی» OR و NOT به هم متصل شده‌اند. به عنوان مثال عبارت ساده‌ی $A \text{ AND } B$ زمانی درست است که هم A و هم B درست باشند، و در غیر این صورت نادرست است. از طرفی دیگر $A \text{ OR } B$ زمانی درست است که حداقل یکی از دو متغیر درست باشند. گیت NOT ساده‌تر است: مقدار یک متغیر را برعکس می‌کند. با تعدادی کافی از این بلوک‌های پایه‌ای می‌توان هر محاسباتی را انجام داد. ایلانگو می‌گوید: «وقتی در پایان روز به کامپیوتر خود نگاه می‌کنید، کامپیوتر شما دارد چه می‌کند؟ در واقع در حال اجرای یک مدار است.».

کار شانون روش جدیدی را به نظریه‌پردازان پیشنهاد داد تا در مورد دشواری مسائل محاسباتی فکر کنند؛ روشی که به «پیچیدگی مدار»^۲ موسوم است — گرچه مدارهای مورد بحث فقط مفاهیم انتزاعی ریاضی هستند. برای مدتی، محققان فکر می‌کردند که این رویکرد می‌تواند راهی برای حل P در مقابل NP باشد، اما در نهایت این مسیر در برابر مانع اثبات‌های طبیعی قرار گرفت.

چهارچوب پیچیدگی مداری مستلزم بازنگری مفاهیم بنیادی مدل محاسباتی تورینگ است. این‌جا محققان به جای مسائل محاسباتی و روش‌های حلشان، توابع بولی و مدارهایی که آن‌ها را محاسبه می‌کنند را مورد مطالعه قرار می‌دهند. یک تابع بولی متغیرهای بولی با مقادیر 0 و 1 را ورودی می‌گیرد و یکی از دو مقدار 0 یا 1 را خروجی می‌دهد. مشابه تعریف الگوریتم، یک



شکل ۶. تصویری از کلود شانون. او در پایان نامه‌ی کارشناسی ارشدش مدلی نظری برای محاسبه بر اساس مدارهای الکتریکی توسعه داده بود.

¹Claude Shannon

²circuit complexity

مدار فرایندی را توصیف می‌کند که با توجه به هر ورودی، یک خروجی تعیین می‌شود. رایان ویلیامز^۱ می‌گوید: «به نظر من مردم شروع به کار بر روی پیچیدگی مداری کردند چون به این نتیجه رسیدند که ماشین‌های تورینگ خیلی پیچیده‌اند». ویلیامز یک نظریه‌پرداز پیچیدگی در MIT است. او ادامه می‌دهد: «ما می‌توانیم مدارها را گیت به گیت بررسی کنیم.».

همان‌گونه که برای حل هر مسئله‌ی محاسباتی، ممکن است الگوریتم‌های متفاوتی وجود داشته باشد — که بعضی نسبت به بقیه سریع‌ترند — مدارهای متفاوت زیادی نیز می‌توانند هر تابع بولی را محاسبه کنند — برخی با گیت‌های کم‌تری از بقیه. محققان پیچیدگی مداری یک تابع را با تعداد گیت‌های کوچک‌ترین مداری که آن را محاسبه می‌کند، تعریف می‌کنند. برای یک تابع با تعداد ثابتی متغیر ورودی، پیچیدگی مدار هم یک مقدار ثابت است — برای برخی توابع بیشتر از بقیه.

با این حال، در بسیاری از موارد می‌توان نسخه‌های پیچیده‌تری از یک تابع را با افزایش متغیرهای ورودی اش در نظر گرفت — همان‌گونه که می‌توان مسئله‌ی مسیر همیلتونی را با در نظر گرفتن گراف‌های بزرگ‌تر سخت کرد. این جاست که پژوهشگران همان سوالی را پرسیدند که هنگام مطالعه‌ی زمان اجرای یک الگوریتم پرسیده بودند: آیا با افزایش متغیرهای ورودی، حداقل تعداد گیت‌های لازم برای محاسبه‌ی یک تابع بولی به طور چندجمله‌ای رشد می‌کند یا نمایی؟ این دو دسته از توابع به ترتیب تابع «به راحتی محاسبه‌پذیر» و «به سختی محاسبه‌پذیر» نامیده می‌شوند.

یک تابع بولی به راحتی محاسبه‌پذیر مانند یک مسئله‌ی محاسباتی در کلاس \mathcal{P} است؛ مسئله‌ای که می‌توان آن را در زمان چندجمله‌ای حل کرد. با این حال توابع مشابه با مسائل \mathcal{NP} -سخت هم وجود دارند، به این معنا که بهترین روشی که محققان برای محاسبه‌ی نسخه‌های بزرگ‌تر تابع یافته‌اند، مستلزم افزایش نمایی تعداد گیت‌هاست؛ اگرچه درستی جوابشان به راحتی قابل بررسی است. اگر نظریه‌پردازان پیچیدگی می‌توانستند ثابت کنند که واقعاً روش بهتری برای محاسبه چنین توابعی وجود ندارد، این به معنای $\mathcal{NP} \neq \mathcal{P}$ بود.

این استراتژی‌ای بود که بیشتر نظریه‌پردازان پیچیدگی در دهه‌ی ۱۹۸۰ دنبال کردند، و البته شانس با آن‌ها همراه بود. شانون در ۱۹۴۹ ثابت کرده بود [۱۰] که پیچیدگی مداری تقریباً هر جدول ارزش بولی (که تنها لیستی طولانی از ورودی‌های ممکن و خروجی‌های یک تابع بولی مشخص است) عملاً بیشترین حالت ممکن است. او از یک استدلال ساده‌ی خیره‌کننده استفاده کرد: راههای ممکن برای ترکیب تعداد کمی از گیت‌ها نسبت به راههای ترکیب تعداد زیادی از گیت‌ها بسیار کم‌تر است. آرانسون می‌گوید: «تعدادی کافی از مدارهای کوچک برای گشت‌وگذار وجود ندارد».

به این ترتیب نظریه‌پردازان پیچیدگی خود را در وضعیت عجیبی یافتند. اگر تقریباً تمام جدول‌های ارزش پیچیدگی مداری بالایی دارند، پس تقریباً هر تابع بولی به سختی محاسبه می‌شود. محققان تنها باید یک تابع را شناسایی می‌کردند که در کلاس \mathcal{NP} هم قرار بگیرد. مگر چقدر می‌تواند سخت باشد؟

۲.۲. برادران رمز. پیشرفت‌ها در آغاز سریع بود. در ۱۹۸۱، سیپسر و دو تن از همکارانش ثابت کردند [۱۲] که اگر از مدارهایی با قیود خاصی بر شیوه‌ی قرارگیری گیت‌های آن‌ها استفاده شود، یک تابع بولی خاص قطعاً محاسبه‌اش مشکل خواهد بود. سیپسر می‌گوید: «روبا یمان این بود که بتوانیم چیزهایی را در مورد این مدل‌های محدودشده اثبات کنیم و سپس بر اساس آنچه آموخته‌ایم با محدودیت‌های کمتر و کمتر کار کنیم».

در ۱۹۸۵، رازبروف قدم بزرگ بعدی را برداشت. او به تازگی تحصیلات تکمیلی اش را در موسکو شروع کرده بود و در حالی که داشت به مسئله‌ای در شاخه‌ی دیگری از ریاضیات می‌پرداخت، به طور اتفاقی به این تلاش پیوسته بود؛ جایی که فهمیده بود



شکل ۷. تصویری از اجزای اصلی کامپیوتر مارک ۱ هاروارد در سال ۱۹۴۴. کلیدهای الکترومکانیکی مشابه با همان‌هایی است که شانون در پایان نامه‌اش بررسی‌شان کرده بود.

^۱Ryan Williams

حل مسئله‌ی P در برابر \mathcal{NP} یک پیشیاز برای کار او است. رازبروف می‌گوید: «من صرفاً خوششانس بودم که نمی‌دانستم این مسئله چقدر سخت است، و گرنه ممکن بود حتی شروعش هم نکنم».

رازبروف مدارهایی را که فقط شامل گیت‌های AND و OR می‌شدند تحلیل می‌کرد، و ثابت کرد [۱۲] که یک تابع خاص، هر طوری که گیت‌ها چیده شوند، به سختی با چنین مدارهایی محاسبه می‌شود — تابعی که \mathcal{NP} -کامل بودن آن ثابت شده بود. تنها کاری که محققان باید برای حل P در برابر \mathcal{NP} انجام می‌دادند این بود که تکنیک‌های رازبروف را به مدارهایی با گیت NOT گسترش دهند. رازبروف می‌گوید: «یک احساس جهانی وجود داشت که یک قدم دیگر، یک ضربه دیگر، و ما قرار است آن را بفهمیم»؛ اما این اتفاق نیفتاد. رازبروف خودش اثبات کرد که روش او، اگر گیت‌های NOT اضافه شوند، شکست خواهد خورد، و هیچ کس نتوانست راه دیگری برای پیش‌روی پیدا کند. با گذشت سال‌ها، او شروع به فکرکردن به این کرد که چه شد که آن مسیر از بین رفت.

در ایالات متحده، رودیج^۱ نیز داشت به همین سوال فکر می‌کرد. او و

ایمپاگلیازو همکلاسی‌های دانشگاه بودند که با هم به تحصیلات تكمیلی رفته بودند. دوستی آن‌ها به دلیل شیفتگی مشترکشان نسبت به اثبات‌های خودارجاعی گودل و تورینگ و پیامدهای آن‌ها برای پایه‌های ریاضیات و علوم کامپیوتر شکل گرفته بود. ایمپاگلیازو می‌گوید: «شوخی ما این بود که قرار بود دکمه‌ای بگیریم که رویش نوشته باشد خودارجاعی».^۲ به عنوان دانشجویان تحصیلات تكمیلی، رودیج و ایمپاگلیازو روی مبانی نظری پیچیدگی رمزگاری کار می‌کردند؛ موضوعی که شاید بهترین انگیزه‌ی عملی را برای فکرکردن روی اثبات $P \neq \mathcal{NP}$ فراهم می‌کرد. رمزگاران پیام‌های سری را با پیچیدن آن‌ها لای «شبه‌تصادفی بودن»^۳ پنهان می‌کنند. پیامی که به این صورت رمزگذاری شده باشد برای هر شنودگری شبیه به رشته‌ی تصادفی درهم‌برهمی از اعداد خواهد بود، اما هم‌چنان می‌تواند توسط گیرنده رمزگشایی شود. با این حال چگونه می‌توان مطمئن شد که شکستن رمز برای یک شنودگر بالقوه بسیار مشکل خواهد بود؟

این جا جایی است که نظریه‌ی پیچیدگی وارد می‌شود. بیشتر روش‌های رمزگذاری که امروزه استفاده می‌شوند، مبتنی بر مسائل به ظاهر سخت \mathcal{NP} هستند. یک مهاجم برای رمزگشایی به یک الگوریتم — تاکنون کشف نشده — سریع برای حل کردن مسئله نیاز پیدا خواهد کرد. برای اثبات این‌که این روش‌ها واقعاً امن هستند، کاری که باید انجام دهید این است که نشان دهید $P \neq \mathcal{NP}$. آن‌طور که سیپرس می‌گوید، بدون یک اثبات تنها کاری که می‌توانید انجام دهید این است که «امیدوار باشید که آن کسی که دارید سعی می‌کنید چیزی را از او پنهان کنید، ریاضی دان بهتری از شما نباشد».



شکل ۹. الکساندر رازبروف (تصویر چپ) و استیون رودیج (تصویر راست) مانع اثبات‌های طبیعی را کشف کردند، که توضیح می‌داد چرا تلاش‌های پیشین برای اثبات $P \neq \mathcal{NP}$ به نتیجه نرسیده است.

اگرچه رمزگاری به نوعه خود شکفتانگیز بود، اما به ظاهر عاری از استدلال‌های خودارجاعی‌ای بود که

در ابتدا رودیج و ایمپاگلیازو را به این حوزه کشانده بود. با این حال زمانی که رودیج در تلاش بود که بفهمد چرا پیچیدگی مداری به بن‌بست خورده است، متوجه شد که این دو موضوع چندان هم از هم دور نیستند. استراتژی‌ای که محققان برای اثبات $P \neq \mathcal{NP}$ به کار گرفته بودند، یک ذات خودمناقض داشت که یادآور گزاره‌ی معروف «این گزاره اثبات‌پذیر نیست» گودل بود

¹ Steven Rudich

² pseudorandomness

و رمزنگاری می‌توانست به توضیح چراجی آن کمک کند. رازبروف در همان زمان در روسیه ارتباط مشابهی را کشف کرد. این‌ها بذرهای «مانع اثبات‌های طبیعی» بودند.

کشاکشی که در قلب مانع اثبات‌های طبیعی وجود دارد این است که مسئله‌ی تشخیص توابع با پیچیدگی بالا از توابع با پیچیدگی پایین مشابه مسئله‌ی تشخیص تصادفی‌های واقعی از شبه‌تصادفی‌ها در رمزگشایی پیام‌ها است. برای اثبات $\mathcal{NP} \neq P$ دوست داریم که نشان دهیم توابع با پیچیدگی بالا با قطعیت متفاوت از توابع با پیچیدگی پایین‌اند. با این حال از سوی دیگر برای قابل اعتماد بودن امنیت رمزنگاری، تمایل داریم که شبه‌تصادفی‌ها از تصادفی‌های واقعی غیرقابل تمایز باشند. شاید نمی‌توانیم هر دو را با هم داشته باشیم.

۳.۲ یک لطیفه‌ی بی‌رحمانه. در ۱۹۹۴ رازبروف و رو دیج متوجه شدند که به بینش‌های مشابهی رسیده‌اند و شروع به هم‌کاری کردند تا نتایجشان را باهم ترکیب کنند. آن‌ها در ابتدا مشاهده کردند که تمام تلاش‌های سابق برای اثبات $\mathcal{NP} \neq P$ با استفاده از پیچیدگی مداری، یک استراتژی کلی را اتخاذ کرده‌اند: ویژگی خاصی از یک تابع بولی \mathcal{NP} -کامل را شناسایی کنید، سپس اثبات کنید که هیچ تابع به راحتی محاسبه‌پذیری نمی‌تواند آن خاصیت را داشته باشد. این نشان خواهد داد که محاسبه‌ی تابع \mathcal{NP} -کامل انتخاب شده واقعاً سخت است و $P \neq \mathcal{NP}$ را ثابت می‌کند.

سیپسر، رازبروف و دیگران همین استراتژی را با موفقیت به کار گرفته بودند تا نتایج محدودتر خود را اثبات کنند و در تمام حالات، بیشتر توابع بولی دارای آن ویژگی خاص بودند که شناسایی شده بود. رازبروف و رو دیج برای اشاره به حالتی که آن ویژگی در اکثر توابع وجود دارد، اصطلاح «اثبات طبیعی» را ابداع کردند؛ صرفاً به این دلیل که هیچ جایگزین شناخته‌شده‌ای وجود نداشت. اگر اثبات‌های «غیرطبیعی» امکان‌پذیر باشند، باید بسیار ناشهودی باشند، و مستحق این نام خواهند بود.

پس از آن بود که رازبروف و رو دیج نتیجه‌ی اصلی‌شان را ثابت کردند: یک اثبات طبیعی برای $\mathcal{NP} \neq P$ مستلزم فهم گسترده‌ای خواهد بود از چگونگی تمایز توابعی که به راحتی محاسبه‌پذیرند از توابع به سختی محاسبه‌پذیر، و این درک می‌تواند انگیزه‌بخش الگوریتم سرعی برای شناسایی توابع به راحتی محاسبه‌پذیر باشد. اگر نظریه پردازان پیچیدگی در یک اثبات طبیعی $\mathcal{NP} \neq P$ موفق شده بودند، راهی تقریباً عاری از خطأ را نیز کشف می‌کردند که با نیمنگاهی به یک جدول ارزش دلخواه بتوان تعیین کرد که تابع متناظر آن پیچیدگی مداری زیاد یا کمی دارد — و البته این، نتیجه‌ای بسیار قوی‌تر و کلی تر است از آن چیزی که آن‌ها قصد داشتند ثابت کنند. کارموسینو می‌گوید: «تقریباً نمی‌توانید جلویش را بگیرید و بیشتر از آن چیزی که برایش چانه زدید به دست می‌آورید».

این مشابه آن است که تلاش کرده باشید یک گفته را صحبت‌سنگی کنید، ولی هر تلاشتان تبدیل به طرح ساخت یک دروغ‌سنجد همه‌کاره شده باشد — آنقدر خوب به نظر می‌رسد که باورکردنی نیست حقیقت داشته باشد. برای نظریه‌پردازان پیچیدگی، قدرت عجیب اثبات‌های طبیعی باعث شد که موفقیت کمتر محتمل به نظر برسد. اما اگر چنین اثباتی موفق می‌شد، به دلیل ارتباط میان پیچیدگی مداری و شبه‌تصادفی بودن، نتیجه‌ی غیرمنتظره‌اش خبر بدی برای رمزنگاری می‌بود.

برای درک این ارتباط، جدول ارزش یک تابع بولی با تعداد زیادی متغیر ورودی را تصور کنید. اگر آن تابع بولی پیچیدگی مداری زیادی داشته باشد، آن لیست طولانی ارزش‌ها اساساً از یک رشته‌ی واقعاً تصادفی از 0 و 1 قابل تمایز خواهد بود — رشته‌ای که مثلاً با پشت هم سکه‌انداختن به دست آمده باشد. با این حال اگر پیچیدگی مداری تابع کم باشد، آن رشته حتی اگر پیچیده به نظر برسد هم باید یک توصیف ساده و مختصر داشته باشد. این موضوع آن را بسیار شبیه به رشته‌های شبه‌تصادفی مورد استفاده در رمزنگاری می‌کند — رشته‌هایی که توصیف مختصرشان همان پیام مخفی مدفون در ظاهر تصادفی آن‌ها است. نتیجتاً دست آورده رازبروف و رو دیج نشان داد که هر اثبات طبیعی $\mathcal{NP} \neq P$ منتج به الگوریتم سرعی می‌شود که می‌تواند رشته‌های شبه‌تصادفی‌ای که دارای پیامی مخفی هستند را از تصادفی‌های واقعی تمایز دهد. به این ترتیب رمزنگاری امن غیرممکن خواهد شد، دقیقاً برعکس آن چیزی که محققان امیدوار بودند از اثبات $\mathcal{NP} \neq P$ به دست آورند.

از سوی دیگر اگر رمزنگاری امن ممکن باشد، آن‌گاه اثبات‌های طبیعی استراتژی ممکنی برای اثبات $\mathcal{NP} \neq P$ نخواهد بود — که پیشنهادی برای رمزنگاری است. این لب مطلب مانع اثبات‌های طبیعی بود. به نظر می‌رسید که نظریه‌پردازان پیچیدگی مخاطب یک لطیفه‌ی بی‌رحمانه بودند. کاباتنس می‌گوید: «اگر به سختی باور دارید، آنگاه باید بپذیرید که اثبات سختی سخت است».

۴.۲ پیش به سوی متأورس. ارتباط میان پیامدهای حدس $\mathcal{NP} \neq P$ و سختی اثبات آن جالب، اما سر درآوردن از آن دشوار بود. یک دشواری این بود که مانع اثبات‌های طبیعی تنها جلوی یک رویکرد اثبات $\mathcal{NP} \neq P$ را گرفت. دیگر این که، سختی اثبات $\mathcal{NP} \neq P$ را نه به خود $P \neq NP$ ، بلکه به وجود رمزنگاری امن مرتبط کرد — به یک مسئله‌ی مشابه اما نه کاملاً معادل. برای فهم درست این ارتباط، محققان باید با فرایپیچیدگی^۱ خوبگیرند. ویلیامز می‌گوید: «این شهود وجود دارد که چون $P \neq NP$ ، پس اثبات $\mathcal{NP} \neq P$ باید خیلی سخت باشد؛ اما برای این که به این شهود معنایی بدھید، باید به امر اثبات گزاره‌ای مانند $\mathcal{NP} \neq P$ به عنوان یک مسئله‌ی محاسباتی نگاه کنید.».



شكل ۱۰. تصویری از والنتین کابانتس، که در دوران کارشناسی ارشدش مقاله‌ی تأثیرگذاری درباره‌ی یک مسئله‌ی اساسی در فرایپیچیدگی — که آن را مسئله‌ی حداقل اندازه‌ی مدار (MCSP) نامید — نوشته.

این کاری بود که کابانتس به عنوان یک دانشجوی تحصیلات تکمیلی انجام داد. او دو سال پس از سقوط جماهیر شوروی در اوکراین به دنیا آمد. در آشفتگی‌های پس از آن واقعه، او فرصت کمی داشت تا مباحث نظری‌ای که بیشتر به آن‌ها علاقمند بود را پی بگیرد. کابانتس به یاد می‌آورد: «من می‌خواستم کار آکادامیک‌تری انجام دهم، و به علاوه دوست داشتم دنیا را بگردم». او برای تحصیلات تکمیلی به کانادا رفت و آن‌جا بود که با مانع اثبات‌های طبیعی آشنا شد. کابانتس، مانند کارموسینو، مجدوب این نتیجه شده بود. او می‌گوید: «وجود چنین ارتباطی، خیلی عمیق به نظر می‌رسید».

او در ۲۰۰۰، اواخر تحصیلات تکمیلی‌اش، با صحبت‌هایی که با جین بی کای^۲ — یک نظریه‌پرداز پیچیدگی که در آن زمان برای فرصت مطالعاتی به تورنتو آمده بود — داشت، متوجه شد که موضوع مانع اثبات‌های طبیعی مدام در مکالماتشان مطرح می‌شود. آن‌ها تصمیم گرفتند که به عنوان یک بنیست بلکه به عنوان یک دعوت‌نامه نگاه کنند؛ فرضی برای بررسی دقیق این که چقدر سخت است که ثابت کنند مسائل سخت هستند. مقاله‌ای که آن‌ها در آن دیدگاه جدید را ارائه کردند [۱۴] به یکی از تأثیرگذارترین کارهای اولیه در حوزه‌ی نوظهور فرایپیچیدگی تبدیل شد.

مقاله‌ی کابانتس و کای روی یک مسئله‌ی محاسباتی مرکز می‌کند که در فرمول‌بندی مانع اثبات‌های طبیعی رازبروف و رودیچ به آن اشاره شده بود: با توجه به جدول ارزش یک تابع بولی، بررسی کنید که آیا پیچیدگی مداری آن زیاد است یا کم. آن‌ها آن مسئله را مسئله‌ی حداقل اندازه‌ی مدار^۳ یا MCSP نامیدند. MCSP یک مسئله‌ی بنیادی فرایپیچیدگی است: یک مسئله‌ی محاسباتی که موضوع مورد بحث آن نظریه‌ی گراف یا موضوع خارجی دیگری نیست، بلکه خود نظریه‌ی پیچیدگی است. در واقع، این مسئله مانند نسخه‌ای کمی از سوالی است که نظریه‌پردازان پیچیدگی را به درگیرشدن با \mathcal{NP} در مقابل \mathcal{P} با استفاده از رویکرد پیچیدگی مدار در دهه‌ی ۱۹۸۰ سوق داد: کدام توابع بولی سخت محاسبه می‌شوند و کدام راحت؟ ایمپاگلیازو می‌گوید: «اگر یک الگوریتم MCSP بیابیم، مانند این خواهد بود که راهی برای ماشینی کردن کاری که در نظریه پیچیدگی انجام

می‌دهیم پیدا کرده باشیم؛ حداقل باید بینش فوق‌العاده‌ای در مورد این که چگونه کارمان را بهتر انجام دهیم، به ما بدهد.» نظریه‌پردازان پیچیدگی نگران این نیستند که این الگوریتم جادویی آن‌ها را از کار بیندازد. در واقع، آن‌ها اصلاً فکر نمی‌کنند که چنین چیزی وجود داشته باشد؛ چرا که رازبروف و رودیچ نشان دادند که هر الگوریتم چنینی برای تمیزدادن جدول ارزش‌های با پیچیدگی زیاد از پیچیدگی کم رمزنگاری را غیرممکن می‌کند. این یعنی MCSP احتمالاً یک مسئله‌ی محاسباتی سخت است، اما چقدر سخت؟ آیا مانند مسئله‌ی مسیر همیلتونی و تقریباً هر مسئله‌ی دیگری که محققان در دهه‌ی ۱۹۶۰ با آن درگیر بودند، \mathcal{NP} -کامل است؟ معمولاً پاسخ دادن به «چقدر سخت است؟» برای مسائل کلاس \mathcal{NP} آسان است؛ اما به نظر می‌رسید برای

¹meta-complexity

²Jin-Yi Cai

³minimum circuit size problem

MCSP چیز دور از ذهنی باشد. کابانتس می‌گوید «ما تعداد کمی مسائل شناور داریم که با این که سخت به نظر می‌رسند، هنوز به جزیره‌ی \mathcal{NP} -کامل متصل نشده‌اند.».

کابانتس می‌دانست که او و کای اولين کسانی بودند که مسئله‌ای که آن‌ها MCSP نامیده بودند را بررسی کرده باشند. ریاضی‌دانان شوروی در ابتدای دهه‌ی ۱۹۵۰، در تلاشی اولیه برای درک دشواری ذاتی مسائل مختلف محاسباتی، مسئله‌ی بسیار مشابهی را مطالعه کرده بودند. لئونید لوین^۱ در دهه‌ی ۱۹۶۰ در جریان مطالعه‌ی چیزی که داشت به نظریه \mathcal{NP} -کامل بودن تبدیل می‌شد، با این مسئله گلاویز شده بود؛ اما توانسته بود \mathcal{NP} -کامل بودنش را ثابت کند، و مقاومت ماندگارش را بدون آن منتشر کرده بود. پس از آن برای ۳۰ سال آن مسئله توجه افراد کمی را به خود جلب کرد، تا این که کابانتس و کای به ارتباط آن با مانع اثبات‌های طبیعی اشاره کردند. کابانتس انتظار نداشت خودش این سوال را حل کند. او در عوض می‌خواست بررسی کند که چرا اثبات این که این مسئله‌ی ظاهراً سخت درباره‌ی محاسباتی واقعاً سخت بوده، انقدر مشکل بوده است. راهول ساتتانام^۲، یک نظریه‌پرداز پیچیدگی در دانشگاه آکسفورد^۳ می‌گوید: «این به یک معنا فرافرای پیچیدگی است.».

اما آیا قرار بود این سختی تا آخر مسیر وجود داشته باشد، یا این که حداقل یک راه برای درک این که چرا محققان در اثبات \mathcal{NP} -کامل بودن MCSP موفق نشده بودند وجود داشت؟ کابانتس کشف کرد که بله، دلیلی وجود دارد: سختی درک‌دنن پیچیدگی مداری مانند یک مانع برای هر استراتژی شناخته‌شده برای اثبات \mathcal{NP} -کامل بودن MCSP عمل می‌کند — مسئله‌ای که خودش درباره‌ی سختی درک پیچیدگی مداری است. گریزگاهی از منطق خودمنافق و پیچ دریچ مانع اثبات‌های طبیعی نبود. ممکن است که \mathcal{NP} -کامل نباشد، اما این نیز دور از ذهن به نظر می‌رسد؛ چرا که برخی از انواع ساده‌تر مسئله پیشتر به عنوان مسائل \mathcal{NP} -کامل شناخته شده‌اند. ایمپاگلیازو می‌گوید: «مشکل این است که فقط جای خوبی برای فرادرادن آن نداریم، به گونه‌ای که مستقیماً آن را با تمام مسائل دیگری که مطالعه می‌کنیم مرتبط کند.».

کابانتس رفتار عجیب MCSP را روشن کرده بود، اما نمی‌دانست چگونه جلوتر برود. جریان تحقیقات فرای پیچیدگی خیلی کند شد. با این حال ۱۶ سال بعد، زمانی که محققان رابطه‌ی غیرمنتظره‌ای را با یک سوال بنیادی دیگر کشف کردند، دوباره شکوفا شد: حل مسائل چقدر سخت است، اگر فقط بخواهید در «بیشتر اوقات» پاسخ صحیح دریافت کنید؟

۵.۲. جنگ جهان‌ها. برای مسائل روزمره راهکارهایی که فقط بیشتر اوقات جواب می‌دهند هم کفایت می‌کند. مثلاً ما رفت و آمد هایمان را با الگوهای ترافیکی معمولی برنامه‌ریزی می‌کنیم نه با سناریوهای بدترین حالت. بیشتر نظریه‌پردازان پیچیدگی سخت‌تر راضی می‌شوند: آن‌ها تنها وقتی راضی می‌شوند مسئله‌ای را آسان اعلام کنند که الگوریتم سریعی پیدا کنند که جواب درست را برای هر ورودی ممکن به دست آورد. این مواجهه‌ی استاندارد مسائل را بر اساس آنچه محققان، پیچیدگی «بدترین حالت»^۴ می‌نامند، طبقه‌بندی می‌کند. با این حال یک نظریه پیچیدگی «حالات میانگین»^۵ هم وجود دارد که در آن مسائل ساده در نظر گرفته می‌شوند اگر الگوریتم سریعی که جواب درست را برای بیشتر ورودی‌ها به دست آورد، وجود داشته باشد.

این تمایز برای رمزنگاران اهمیت دارد. یک مسئله‌ی محاسباتی را تصور کنید که تقریباً برای هر ورودی، به غیر از چند حالت سرسرخت که بهترین الگوریتم در آن‌ها شکست می‌خورد، به راحتی حل می‌شود. پیچیدگی بدترین حالت آن مسئله را سخت در نظر می‌گیرد، اما برای رمزنگاری این مسئله بی‌فائده است: چه فایده‌ای دارد اگر رمزگشایی فقط برای برخی از پیام‌های شما دشوار باشد؟ این در واقع لوین بود که یک دهه بعد از کار پیشگامانه‌اش در \mathcal{NP} -کامل بودن، مطالعه‌ی دقیق پیچیدگی حالت میانگین را آغاز کرد. در آن فاصله، او با مقامات شوروی درگیر شده بود. لوین یک دردرساز بی‌پرده بود که گه‌گاه فعالیت‌های میهن‌دوستانه‌ی خود در گروه جوانان حزب کمونیست را لکه‌دار می‌کرد. لوین در سال ۱۹۷۲ به دلایل آشکارا سیاسی از مدرک دکترا محروم شد. ایمپاگلیازو می‌گوید: «برای این که به عنوان یک محقق جوان در جماهیر شوروی موفق شوید، نمی‌توانید خیلی صاحب‌نظر باشید، و تصور این که لئونید صاحب‌نظر نباشد سخت است.».

لوین در ۱۹۷۸، به ایالات متحده مهاجرت کرد و در نیمه‌ی دهه ۱۹۸۰، توجه خود را به پیچیدگی حالت میانگین معطوف کرد. او کار با دیگران را شروع کرد — از جمله ایمپاگلیازو که در آن زمان دانشجوی تحصیلات تکمیلی بود — تا نظریه را بیشتر

¹Leonid Levin

²Rahul Santhanam

³Oxford

⁴worst-case

⁵average-case

توسعه دهد. با وجود این که آن‌ها پیشرفت کردند، ایمپاگلیازو متوجه شد اغلب محققان هرچند هر کدام حرف خودش را می‌زد، گمان می‌کردند که همگی دارند از یک موضوع سخن می‌گویند. او می‌خواست همه را همنظر کند، و البته این که مقاله‌های لوین به صورت مشهوری مختصر بود کمکی به آن نمی‌کرد — آن مقاله‌ای لوین [۱۵] که آغازگر حوزه‌ی پیچیدگی حالت میانگین بود، کمتر از دو صفحه بود. ایمپاگلیازو می‌گوید: «من می‌خواستم کارهای لئونید را به اصطلاحات فنی تر و در دسترس‌تر ترجمه کنم.». او تصمیم گرفت پیش از شیرجه‌زدن در ریاضیات ماجرا، با یک نمای کلی کوتاه و سرزنشه شروع کند. «آن بیشتر مقاله را گرفت، و البته این تنها بخشی است که همه به یاد می‌آورند.».

آن مقاله [۱۶] در ۱۹۹۵ منتشر شد و فوراً به یکی از بهترین‌های حوزه‌ی خودش تبدیل شد. ایمپاگلیازو نام‌های عجیب‌وغریبی را برای پنج جهانی که به وسیله‌ی درجه‌های متفاوت سختی‌های محاسباتی و قابلیت‌های رمزنگاری متفاوت تمایز شده بودند، ابداع کرده بود [۱۷]. ما در یکی از این جهان‌ها زندگی می‌کنیم ولی نمی‌دانیم کدام.



شکل ۱. لئونید لوین (تصویر راست) مطالعه‌ی پیچیدگی حالت میانگین را در اواسط دهه هشتاد آغاز کرد. راسل ایمپاگلیازو (تصویر چپ) بعدها این موضوع را در مقاله‌ای درباره‌ی پنج جهان محاسباتی که ممکن است در آن زندگی کنیم، درسترس‌تر کرد.

از زمانی که مقاله‌ای ایمپاگلیازو منتشر شد، محققان آرزوی حذف بخش‌هایی از متاورس مینیاتوری او را داشتند؛ با محدود کردن فضای احتمالات به وسیله‌ی ثابت کردن این که وجود برخی از جهان‌ها ممکن نخواهد بود. دو جهان هدف‌های وسوسه‌کننده‌ای بودند: آن‌هایی که رمزنگاری در آن‌ها ناممکن بود حتی اگر $\mathcal{P} \neq \mathcal{NP}$ باشد. در یکی از آن جهان‌ها، به نام هیورستیکا^۱، تمام مسائل \mathcal{NP} -کامل برای بیش‌تر ورودی‌ها به راحتی حل می‌شوند؛ اما الگوریتم‌های سریع گاهی اشتباه می‌کنند. بنابراین این مسائل با استانداردهای نظریه‌ی پیچیدگی

بدترین حالت، سخت در نظر گرفته می‌شوند. این جهانی است که رمزنگاری در آن ناممکن است چون تقریباً هر کدی به راحتی شکسته می‌شود. در جهان دیگر، به نام پسیلند^۲، رمزنگاری به دلیل دیگر ناممکن است: تمام مسائل در حالت میانگین سخت است، اما رمزگذاری یک پیام آن را حتی برای گیرنده‌ی موردنظر هم ناخوانا می‌کند.

علوم می‌شود این دو جهان ارتباط تنگاتنگی با مسائل فرایمپاگلیازو دارند؛ به طور خاص، سرنوشت هیورستیکا به سوال طولانی‌مدت \mathcal{NP} -کامل بودن MCSP پیوند خورده است. آن سوالی که خیلی وقت پیش کابانتس مذوب آن شده بود و لوین را هاج و اوج کرده بود دیگر فقط یک کنجدکاوی نبود: سرنوشت یک جهان در خطر است.

برای رد کردن هیورستیکا، محققان باید تمایز میان پیچیدگی بدترین حالت و حالت میانگین را از بین ببرند، که این یعنی باید ثابت کنند که هر الگوریتم فرضی که یک مسئله‌ای \mathcal{NP} -کامل را به درستی برای بیش‌تر ورودی‌ها حل می‌کند، در واقع برای همه حالت‌ها جواب می‌دهد. این نوع ارتباط، که تحويل بدترین حالت به حالت میانگین نام دارد^۳، برای مسائل خاصی وجود دارد، اما هیچ کدام‌شان \mathcal{NP} -کامل نیستند؛ بنابراین آن نتایج دلالت بر چیز کلی‌تری ندارند. نابودی هیورستیکا رمزنگاران را تا نیمه‌راه تحقق روایی رمزگذاری امن بر اساس فرض $\mathcal{P} \neq \mathcal{NP}$ پیش می‌برد؛ اما نابود کردن یک دنیا کار کوچکی نیست. در ۲۰۰۳، دو نظریه پرداز پیچیدگی نشان دادند [۱۸] که رویکردهای موجود برای اثبات کردن تحويل بدترین حالت به حالت میانگین برای مسائل \mathcal{NP} -کامل شناخته شده پیامدهای عجیب و غریبی خواهد داشت که پیشنهاد می‌دهد احتمالاً چنین اثبات‌هایی ممکن نخواهند بود.

محققان باید یک رویکرد دیگر پیدا کنند و اکنون فکر می‌کنند MCSP همان مسئله‌ای است که نیاز دارند. با این حال این برای بیش از یک دهه روشن نشد. اولین توجه به این ارتباط از شیفتگی مداوم کارموسینو به مانع اثبات‌های طبیعی پدیدار شد.

¹Heuristica

²Pessiland

³worst-case to average-case reduction

۳. فرصت‌ها

کارموسینو برای اولین بار به عنوان دانشجوی تحصیلات تکمیلی از طریق مقاله‌ی ۱۳ کابانتس و چهار محقق دیگر [۱۹] با تحقیقات فراییچیدگی رویرو شد، که رویکرد مانع اثبات‌های طبیعی را که کابانتس بیش از یک دهه قبل پیشگام آن شده بود، بیش‌تر توسعه داده بودند. این مقاله فقط اعتقاد او را که هنوز چیزهای زیادی برای بادگیری از مقاله‌ی کلاسیک رازبروف و رودیچ وجود دارد، راسخ‌تر کرد. کارموسینو می‌گوید: «من در آن زمان شیفتنه‌ی آن مقاله شده بودم، و هنوز هم هیچ چیز عوض نشده است.».

شیفتگی کارموسینو بالاخره در جریان بازدیدی از یک کارگاه یک ترمه در دانشگاه کالیفرنیا، برکلی —جایی که بیش‌تر وقت خود را به صحبت با ایمپاگلیازو، کابانتس و آتنونینا کولوکولوفا^۱، یک نظریه پرداز پیچیدگی در دانشگاه مموریال نیوفاندلند^۲ که با کابانتس در مقاله‌ی ۱۳۰۱۶ اش همکاری کرده بود، صرف می‌کرد —ثمره داد. کارموسینو قبل از نفرشان کار کرده بود، و آن همکاری موفقیت‌آمیز به او این اعتمادبهنه نفس را داد که بارها و بارها سوالاتی درباره‌ی موضوعی که بیش از همه مجدوب آن شده بود را با آن‌ها مطرح کند. کابانتس به یاد می‌آورد: «او مردم را اذیت می‌کرد، البته به نحوی سازنده.» در آغاز کارموسینو ایده‌های جدیدی داشت برای اثبات‌کردن *NP*-کامل بودن نسخه‌ای از MCSP که در مقاله‌ی رازبروف و رودیچ درباره‌ی مانع اثبات‌های طبیعی آمده بود؛ اما آن ایده‌ها نتیجه نداد. در عوض، یک اظهارنظر بی‌مقدمه‌ی ایمپاگلیازو باعث شد که این چهار محقق متوجه شوند که مانع اثبات‌های طبیعی الگوریتم‌های قوی‌تری از آنچه که هر کسی تصور کرده بود به دست می‌دهد. ظاهراً یک نقشه‌ی مخفی روی راه‌بند حک شده بود.

این چهار محقق در مقاله‌ای در سال ۱۶۰۱ ثابت کردند [۲۰] که نوع خاصی از الگوریتم MCSP حالت میانگین را می‌توان برای ساختن یک الگوریتم بدترین حالت برای شناسایی الگوهای پنهان در رشته‌های به ظاهر تصادفی اعداد استفاده کرد —کاری که محققان علوم کامپیوتراز آن به عنوان یادگیری^۳ یاد می‌کنند. این نتیجه‌ی قابل توجهی است؛ چرا که یادگیری شهوداً کار سخت‌تری از رده‌بندی دودویی (با پیچیدگی زیاد یا کم) که توسط یک الگوریتم MCSP انجام شده باشد، به نظر می‌رسد، و در کمال تعجب این نتیجه پیچیدگی بدترین حالت یک کار را به پیچیدگی حالت میانگین یک کار دیگر مرتبط کرد. ایمپاگلیازو می‌گوید: «به هیچ وجه روش نبود که چنین ارتباطی اصلاً می‌تواند وجود داشته باشد.».

یک الگوریتم سریع برای MCSP مدارهای بولی در حالت

کلی کاملاً فرضی و نظری است: تا زمانی که نشان داده نشود MCSP یک مسئله‌ی محاسباتی راحت است، علی‌رغم تمام شواهد ضد آن، این الگوریتم نمی‌تواند وجود داشته باشد، و این یعنی الگوریتم یادگیری‌ای که در مقاله‌ی این چهار محقق به آن اشاره شده، به همان اندازه فرضی و نظری است.

با این حال برای برخی حالت‌های ساده‌تر MCSP و زمانی که مدار محدودیت‌های خاصی داشته باشد، برای تشخیص جداول ارزش با پیچیدگی زیاد از جداول با پیچیدگی کم سال‌هاست که الگوریتم‌های سریعی شناخته شده‌اند. مقاله‌ی کارموسینو، کابانتس، کولوکولوفا و ایمپاگلیازو نشان داد که این الگوریتم‌ها را می‌توان به الگوریتم‌های یادگیری‌ای تبدیل کرد که به طور مشابه محدود شده‌اند، اما هنوز قوی‌تر از هر الگوریتمی هستند که محققان سابقاً با این سطح از دقت نظری، درک کرده‌اند. ایلانگو می‌گوید: «این ویرگی خودارجاعی آن‌ها به نحوی شما را قادر می‌سازد که کارهایی را انجام دهید که ظاهراً نمی‌توانید با مسائل استاندارد بیش‌تری انجام دهید.».



شکل ۱۲. آتنونینا کولوکولوفا در ۱۶۰۱ همراه با کارموسینو، ایمپاگلیازو و کابانتس ارتباطی شگفت‌انگیز میان MCSP و یادگیری را ثابت کرد، که موجب جلب توجه‌ها به سمت فراییچیدگی شد.

¹ Antonina Kolokolova

² Memorial University of Newfoundland

³ Learning

این دست آورد توجه نظریه پردازان پیچیدگی را که روی موضوعات دیگر کار می کردند، به خود جلب کرد. همچنین پیش نمایشی از ارتباطات بیشتر میان فرایپیچیدگی و پیچیدگی حالت میانگین را به نمایش گذاشت که در سال های آینده ظاهر خواهد شد. بیش از همه، این نتیجه گواهی بود بر این که محققان با پرسیدن سوالات ساده در مورد موانع که در ابتدا فقط مانع پیشرفت آن ها می شوند، تا چه حد می توانند پیش بروند. ایمپاگلیازو می گوید: «این نوع از دوگانی یک تم در جریان حداقل ۲۰ تا ۴۰ سال اخیر پیچیدگی است. موانع اغلب فرصت هستند».

۱.۳. اعتبارهای «جزئی». از زمانی که کارموسینو و همکارانش مقاله شان را منتشر کرده اند، پیشرفت ها شتاب گرفته است. کولوکولوفا می گوید: «اتفاقات جدیدی در حال رخدادن است، و محققان جوان خیلی خیلی باهوش زیادی وجود دارند».^۱ ایلانگو یکی از این محققان جوان است. او در سه سال اول تحصیلات تکمیلی اش، به مسئله باز و هولناک اثبات \mathcal{NP} -کامل بودن MCSP با استفاده از یک استراتژی دو وجهی حمله کرد: در حالی که \mathcal{NP} -کامل بودن حالت های مشابه MCSP [۲۱، ۲۲] را ثابت می کرد — همان گونه که محققان پیچیدگی مداری در دهه ۱۹۸۰ به \mathcal{P} در برابر \mathcal{NP} حمله کردند — در حال ثابت کردن \mathcal{NP} -کامل بودن حالت های پیچیده تر [۲۳] نیز بود، که شهوداً سخت تر به نظر می رسد و در نتیجه احتمالاً راحت تر می توان اثبات کرد که مشکل اند. ایلانگو علاقه ای خود به فرایپیچیدگی را مدیون اریک آلندر^۲ می دارد. او یک نظریه پرداز پیچیدگی در دانشگاه راتگرز^۳ است و یکی از معدود محققانی است که به کار روی فرایپیچیدگی در دهه ۲۰۰۰ و اوایل دهه ۲۰۱۰ ادامه دادند. ایلانگو می گوید: «اشتیاق او واگیردار بود».

یک محقق دیگر که از آلندر الهام گرفت، شوئیچی هیراها را^۴ بود، که اکنون استادی در موسسه ملی انفورماتیک توکیو^۵ است. او در ۲۰۱۸، در حالی که هنوز یک دانشجوی تحصیلات تکمیلی بود، سطح واقعی ارتباط میان فرایپیچیدگی و پیچیدگی حالت میانگین را که کارموسینو و همکارانش کشف کرده بودند، آشکار کرد. آن چهار محقق ارتباط میان پیچیدگی حالت میانگین یک مسئله (MCSP) و پیچیدگی بدترین حالت مسئله ای دیگر (یادگیری دودویی) را پیدا کرده بودند. هیراها را تکنیک های آن ها را بیشتر توسعه داد تا برای MCSP یک تحويل بدترین حالت به حالت میانگین به دست آورد [۲۴]. دست آورد او نتیجه می دهد که یک الگوریتم MCSP حالت میانگین فرضی — مانند الگوریتمی که کارموسینو و همکارانش در نظر گرفته بودند — در واقع به قدری قدرتمند خواهد بود که یک نسخه ای کمی متفاوت از MCSP را بدون هیچ اشتباهی حل کند.

نتیجه ای هیراها را هیجان انگیز است؛ زیرا بسیاری از محققان گمان می کنند که MCSP برخلاف تمام مسائل دیگری که برای آن ها تحويل بدترین حالت به حالت میانگین وجود دارد، \mathcal{NP} -کامل است. اگر آن ها بتوانند نتایج هیراها را گسترش دهند تا تمام الگوریتم های حالت میانگین را شامل شود و سپس ثابت کنند که \mathcal{NP} -کامل است، آنگاه ثابت می شود که ما در هیورستیک ارزندگی نمی کنیم. سانتانام می گوید: «این واقعاً نتیجه ای تکان دهنده ای خواهد بود».^۶ اثبات این که \mathcal{NP} -کامل است ممکن است کار دشواری به نظر برسد؛ چرا که این مسئله حدود ۵۰ سال است که مطرح شده است. با این حال پس از پیشرفت مهمی توسط هیراها در سال ۲۰۲۲ [۲۵]، اکنون محققان بسیار نزدیک تر از آن چیزی هستند که انتظار می رفت.



شکل ۱۳. راهول ایلانگو (تصویر چپ) و شوئیچی هیراها را (تصویر راست) اخیراً روش های جدید رمزگاری ای را توسعه داده اند که ثابت می کند نسخه هایی از \mathcal{NP} -کامل است.

هیراها \mathcal{NP} -کامل بودن را برای نوع دیگری از مسئله به نام MCSP جزئی^۷ اثبات کرد که در آن ورودی های خاصی را در جدول ارزش نادیده می گیرید. اثبات او مبتنی بر روش های توسعه یافته توسط ایلانگو بود تا نشان دهد که MCSP جزئی معادل

¹Eric Allender

²Rutgers University

³Shuichi Hirahara

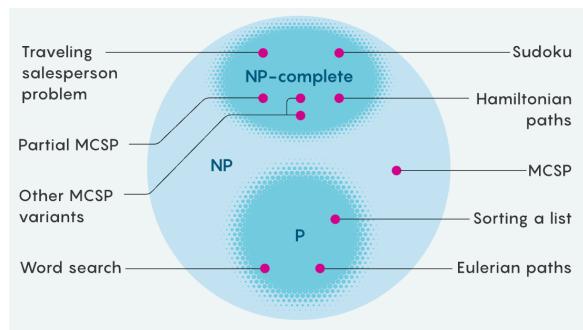
⁴National Institute of Informatics

⁵parital MCSP

مسئله‌ی به ظاهر نامربوطی است که شامل یک تکنیک رمزنگاری به نام تسهیم راز^۱ می‌شد. این روشی است برای تقسیم‌کردن یک پیام رمزنگاری شده میان تعداد زیادی از افراد، به طوری که پیام تنها زمانی رمزگشایی شود که تعداد خاصی از آن‌ها با یکدیگر همکاری کنند.

برای هر کاربرد واقع‌گرایانه‌ی رمزنگاری، پیش از هر چیز باید آن تعداد خاص را پیدا کنید؛ اما به کمک چند ترفند رمزنگاری می‌توانید یک سفاربی ناالمیدکننده بسازید که حتی در آن پیداکردن آن تعداد افرادی که باید با هم همکاری کنند هم دشوار باشد. هیراهارا راهی پیدا کرد تا ثابت کند این مسئله‌ی ساختگی رمزنگاری، NP-² کامل است و سپس نشان داد که این اثبات، NP-کامل بودن MCSP جزئی را هم تضمین می‌کند. این دست آورد، حتی بیش از کارهای قبلی هیراهارا، به محققان انرژی داد و سایر محققان را متوجه این موضوع کرد؛ نظریه‌پرداز پیچیدگی لنس فورتو^۳ آن را دست آورد سال نامید. ویلیامز می‌گوید: «نتیجه‌ی شگفت‌انگیزی بود. همه فکر می‌کردند این مسائل جزئی تقریباً به همان سختی مسئله‌ی اصلی باشند.».

موانعی برای اثبات NP-کامل بودن نسخه‌ی اصلی MCSP باقی مانده است؛ اما هیچ کدام از آن‌ها موانعی نیستند که نیاز به یک جعبه ابزار کاملاً جدید داشته باشند. ممکن است تنها مسئله، پیداکردن راهی درست برای ترکیب‌کردن تکیک‌های شناخته شده باشد. یک اثبات درنهایت وضعیت یکی از معدهود مسائلی را که از زمانی که نظریه‌ی پیچیدگی وجود داشته در برابر طبقه بندی مقاومت کرده است، حل می‌کند. لوین در ایمیلی نوشت: «این من را با نشان‌دادن این که احمق بودم که نتوانستم آن را ببینم، فروتن می‌کند.».



شکل ۱۴. ردیابی فعلی مسئله‌های مورد بحث در این نوشتہ.

۲.۳. تکه‌های گم شده. MCSP تنها مسئله‌ی فراپیچیدگی نیست که باعث پیشرفت بزرگی شده. در ۲۰۲۰، استاد رمزنگاری دانشگاه کرنل^۴، رافائل پس^۵ و دانشجوی تحصیلات تکمیلی اش یانی لو^۶، ارتباطی میان یک مسئله‌ی فراپیچیدگی دیگر را با یک پروتکل بنیادی رمزنگاری که مرز میان پسیلنند و هیورستیکا را مشخص می‌کند، و بدترین جهان‌های ایمپاگلیازو (جایی که مسائل NP-کامل در حالت میانگین مشکل‌اند ولی رمزنگاری هنوز غیرممکن است)، کشف کردند^[۲۶]. این کشف، مسئله‌ی مورد مطالعه‌ی آن‌ها را به یک کاندیدای اصلی برای حمله به پسیلنند بدل می‌کند. به علاوه کارهای اخیر آن‌ها می‌تواند علیه هیورستیکا عمل کند^[۲۷]. پس می‌گوید: «تکه‌های متفاوتی از پازل گم شده‌اند.»، و ادامه می‌دهد: «برای من امری جادویی است که این حوزه‌ها این قدر به هم مرتبط هستند.».

هیراهارا هشدار می‌دهد که هنوز چالش‌هایی در انتظار محققانی است که قصد دارند جهان‌هایی که ایمپاگلیازو ۳۰ سال پیش ساخته بود را از بین ببرند. او می‌گوید: «دوست دارم که بگویم در مقطعی هیورستیکا و پسیلنند کنار خواهد رفت، اما مطمئن نیستم چقدر به آن زمان نزدیک هستیم.».

بسیاری از محققان انتظار دارند که بزرگ‌ترین دشواری، پرکردن شکاف میان دو مدل متفاوت پیچیدگی حالت میانگین باشد. متخصصان رمزنگاری معمولاً الگوریتم‌های حالت میانگین را مطالعه می‌کنند که در هر دو جهت خطأ دارند — گهگاه رشته‌های تصادفی را به عنوان شبه‌تصادفی در نظر می‌گیرند و برعکس. با این حال، تحويل‌های بدترین حالت به حالت میانگین هیراهارا برای الگوریتم‌های حالت میانگینی جواب می‌دهد که فقط خطاهایی از نوع اول دارند. تمایزات ظرفی مانند این در نظریه‌ی پیچیدگی می‌تواند تفاوت‌های بزرگی ایجاد کند. اما علی‌رغم این مانع و موضع بسیار دیگر، آندر نمی‌تواند خوش‌بین نباشد. او می‌گوید: «من سعی می‌کنم نگذارم که خیلی هم معتقد باشم؛ چرا که سابقه‌ی کاملاً ثابت‌شده‌ای وجود دارد که هیچ چیز جواب نمی‌دهد. با این حال ما شاهد پیشرفت‌های بسیار هیجان‌انگیزی هستیم — راههایی برای مقابله با چیزهایی که مانند مانع بودند.».

¹ secret sharing

² Lance Fortnow

³ Cornell Tech

⁴ Rafael Pass

⁵ Yanyi Liu

اگر یک درس وجود داشته باشد که محققان از سروکله زدن های خود با مسئله‌ی \mathcal{P} در برابر \mathcal{NP} یاد گرفته باشند، این است که نظریه‌ی پیچیدگی به خودی خود پیچیده است؛ اما این چالش دقیقاً همان چیزی است که جستجو را بسیار ارزشمند می‌کند. کارموسینو می‌گوید: «راستش خیلی خوب است که این قدر سخت است. عوضش هیچ وقت حوصله‌ام سر نمی‌رود!».

مراجع

- [1] P Vs NP Problem In A Nutshell.. One of the unanswered questions in... | by Bilal Aamir | Medium. (n.d.). . Retrieved May 5, 2024, from <https://medium.com/@bilalaamir/p-vs-np-problem-in-a-nutshell-dbf08133bec5>
- [2] Turing, A.M. (1936) On Computable Numbers, with an Application to the Entscheidungsproblem. *The London Mathematical Society.*, Volume s2-42, Issue 1, 230-265.
- [3] Cook, Stephen A. (1971) The complexity of theorem-proving procedures. *STOC '71: Proceedings of the third annual ACM symposium on Theory of computing*.
- [4] Levin, L.A (1973) Universal Sequential Search Problems. *Problemy Peredachi Informatsii.*, Volume 9, Issue 3, 115-116.
- [5] Karp, R. M. (1972) Reducibility among Combinatorial Problems. *Complexity of Computer Computations.*, 85–103.
- [6] Razborov, A.A. and Rudich, S (1994) Natural proofs. *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*.
- [7] Baker, Theodore, John Gill, and Robert Solovay (1975) Relativizations of the $P =? NP$ Question *SIAM Journal on Computing.*, Volume 4, Issue 4.
- [8] Tse, David (2020) How Claude Shannon Invented the Future. *Quanta Magazine*.
- [9] Shannon, Claude E. (1940) A symbolic analysis of relay and switching circuits. *Electrical Engineering.*, 713-723.
- [10] Shannon, Claude E. (1949) The synthesis of two-terminal switching circuits. *The Bell System Technical Journal.*, Volume 28, Issue 1, 59 - 98.
- [11] Hartnett, Kevin (2018) Why Mathematicians Can't Find the Hay in a Haystack. *Quanta Magazine*.
- [12] Furst, Merrick, James B. Saxe, and Michael Sipser. (1981) Parity, circuits, and the polynomial-time hierarchy. *Institute of electrical and electronics engineers*.
- [13] Razborov, Alexander. (1985) Lower bounds on the monotone complexity of some Boolean function. *Soviet Math. Dokl.*, Vol. 31., 354-357.
- [14] Kabanets, Valentine, and Jin-Yi Cai. (2000) Circuit minimization problem. *Proceedings of the thirty-second annual ACM symposium on Theory of computing*.
- [15] Levin, Leonid A. (1986) Average Case Complete Problems. *SIAM Journal on Computing.*, Volume 15, Issue 1, 285-286.
- [16] Impagliazzo, Russell. (1995) A personal view of average-case complexity. *Tenth Annual IEEE Conference*.
- [17] Klarreich, Erica (2022) Which Computational Universe Do We Live In? *Quanta Magazine*.
- [18] Bogdanov, Andrej, and Luca Trevisan. (2003) On worst-case to average-case reductions for NP problems. *44th Annual IEEE Symposium on Foundations of Computer Science*.
- [19] Chen, R., Kabanets, V., Kolokolova, A., Shaltiel, R. And Zuckerman, D. (2013) Mining Circuit Lower Bound Proofs for Meta-Algorithms.
- [20] Carmosino, M. L., Impagliazzo, R., Kabanets, V. And Kolokolova, A. (2016) Learning algorithms from natural proofs. *31st Conference on Computational Complexity (CCC 2016)*., Article 10, 1-24.
- [21] Ilango, Rahul. (2020) Constant Depth Formula and Partial Function Versions of MCSP are Hard. *IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*.
- [22] Ilango, Rahul. (2021) The Minimum Formula Size Problem is (ETH) Hard. *IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*.
- [23] Ilango, Rahul. (2020) Approaching MCSP from Above and Below: Hardness for a Conditional Variant and $AC^0[p]$. *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*.
- [24] Hirahara, Shuichi. (2018) Non-Black-Box Worst-Case to Average-Case Reductions within NP. *IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*.
- [25] Hirahara, Shuichi. (2022) NP-Hardness of Learning Programs and Partial MCSP. *IEEE 63rd annual symposium on foundations of computer science (FOCS)*.
- [26] Klarreich, Erica (2022) Researchers Identify 'Master Problem' Underlying All Cryptography. *Quanta Magazine*.
- [27] Liu, Yanyi And Pass, Rafael. (2022) On one-way functions from NP-complete problems. *Proceedings of the 37th Computational Complexity Conference.*, Article 36, 1-24.

مترجم: حامی عبادزاده سمنانی[†]

* دانش آموخته‌ی دکتری فیزیک از دانشگاه بیل؛ عضو هیئت تحریریه‌ی مجله‌ی کواتا
تارنما: <https://benbrubaker.com>

[‡]دانشجوی کارشناسی علوم کامپیوتر، دانشگاه صنعتی شریف



تعالی اعداد

*نوید دژبرد

چکیده. در این نوشه ابتدا مفهوم و وجود اعداد متعالی مطالعه می‌شود و سپس نشان داده می‌شود که اعداد e و π متعالی‌اند.

۱. مقدمه

از نخستین و شناخته‌شده‌ترین افسانه‌های ریاضی است که وقتی هیپاسوس^۱، ریاضی‌دان یونانی، وجود اعداد گنگ را کشف کرد، فیثاغورسیان خشمگین او را غرقه در آب کردند. خوشبختانه طی قرن‌های بعدی جهان آکادمیک فرهنگ مسامحت آمیزتری با کشفیات و ابتكارات بدیع نشان داد، تا جایی که در قرن ۱۹ م. ریاضی‌دانان توانستند با اطمینان خاطر از دوگانه‌ی گویا و گنگ نیز فراتر روند و به مطالعه اشیائی پردازند که به اعداد متعالی^۲ معروف شد؛ اشیائی که حتی در حل مسائل کهن ریاضی، همچون مسئله‌ی تربیع دایره، نیز کارساز بوده‌اند. معرفی اعداد متعالی نیازمند معرفی اعداد جبری^۳ است—که نوعاً متمم هم محسوب می‌شوند.

تعریف ۱.۱. عددی حقیقی را جبری گوییم هرگاه ریشه یک چندجمله‌ای ناصرف با ضرایب صحیح باشد.

تعریف ۲.۱. عددی حقیقی را متعالی گوییم هرگاه جبری نباشد.

تعریف اعداد جبری و متعالی قابل گسترش به اعداد مختلط است. همچنین می‌توان در تعاریف فوق از از چندجمله‌ای‌های ناصرف با ضرایب گویا استفاده کرد، که در نهایت با آن‌چه معرفی کردیم معادل است.

۲. وجود اعداد متعالی

مطالعه‌ی اعداد متعالی از قرن ۱۸ م. آغاز گشت و در قرن ۱۹ م. برای نخستین بار وجود عدد متعالی اثبات شد. برخی از مهم‌ترین برهان‌هایی که برای وجود اعداد متعالی ارائه شده، از آن گئورگ کانتور^۴ است. وی نه تنها وجود اعداد متعالی را نشان داد، بلکه توانست کاردینالیتی اعداد متعالی را نیز مطالعه کند. در ادامه دو برهانی را که کانتور—به ترتیب در سال‌های ۱۸۷۴ و ۱۸۹۱—ارائه کرده‌است، معرفی می‌کنیم.

قضیه ۱.۲. مجموعه‌ی اعداد جبری شماراست؛ به عبارت دیگر، می‌توان مجموعه‌ی اعداد جبری را به صورت یک دنباله‌ی نامتناهی نمایش داد.

□ اثبات. تمرین.

قضیه ۲.۲ (۱۸۷۴). به ازای هر دنباله‌ی نامتناهی از اعداد حقیقی و هر بازه‌ی $[a, b] \subset \mathbb{R}$ ، می‌توان عضو $r \in [a, b]$ را طوری یافت که جزو آن دنباله نباشد. ضمناً مجموعه‌ی چنین اعضایی ناشمار است.

¹Hippasus

²Transcendental Numbers

³Algebraic Numbers

⁴Georg Cantor

اثبات. دنباله‌ی $(x_n)_{n \in \mathbb{N}}$ را در نظر بگیرید. برای سادگی فرض کنید که اعضای این دنباله دویه‌دو متمایزند. نخستین دو عضو این دنباله را که در بازه‌ی $I = [a, b] = [a_1, b_1]$ قرار می‌گیرند در نظر بگیرید. عضو کوچک‌تر را a_1 و عضو بزرگ‌تر را b_1 بنامید. بازه‌ی $I_1 = [a_1, b_1]$ را بسازید و نخستین دو عضو دنباله را که در بازه‌ی جدید قرار می‌گیرند در نظر بگیرید. به طریق مشابه بازه‌ی جدید $I_2 = [a_2, b_2]$ را بسازید. دنباله‌ی بازه‌هایی که به این طریق می‌توانیم بسازیم یا متناهی‌اند یا متناهی. در حالت اول، فرض کنید $[a_N, b_N] = I_N$ آخرين بازه‌ي توليدشده به اين شيوه باشد. بنابراین حداقل يك عضو از دنباله، مانند x_k ، می‌تواند عضو اين بازه باشد و هر عضو دیگري در بازه‌ی نهايی برقراری حکم کافی است.

حال فرض کنید دنباله‌ی بازه‌هایی که در این پروسه ساخته‌ایم، نامتناهی است. تعریف کنید $\lim_{n \rightarrow \infty} a_n = a_\infty$ و $\lim_{n \rightarrow \infty} b_n = b_\infty$. از آنجایی که $(a_n)_{n \in \mathbb{N}}$ و $(b_n)_{n \in \mathbb{N}}$ دنباله‌هایی یک‌نوا در بازه‌ای کران‌دار هستند، این حدّها وجود دارند. ضمناً هیچ یک از a_∞ و b_∞ در دنباله‌ی $(x_n)_{n \in \mathbb{N}}$ ظاهر نشده‌اند (چرا؟). اگر $a_\infty = b_\infty$ آن‌گاه تعریف کنید $r = a_\infty$; اگر $a_\infty < b_\infty$ هر عددی در بازه‌ی $[a_\infty, b_\infty]$ جهت برقراری حکم کافی خواهد بود.

حال فرض کنید مجموعه‌ی همه چنین اعضایی شمارا باشد، و بتوان آن‌ها را به صورت دنباله‌ی $(r_n)_{n \in \mathbb{N}}$ نمایش داد. دنباله‌ی $(\bar{x}_n)_{n \in \mathbb{N}}$ به صورت زیر تعریف کنید

$$\bar{x}_n = \begin{cases} x_{\frac{n+1}{2}} & 2 \nmid n, \\ r_{\frac{n}{2}} & 2 \mid n. \end{cases}$$

با تکرار روش مذکور می‌توان $\bar{r} \in [a, b]$ را طوری ساخت که جزوی از دنباله‌ی $(\bar{x}_n)_{n \in \mathbb{N}}$ نباشد. پس مجموعه‌ی چنین اعضایی ناشماراست. \square

نتیجه ۳.۲. در هر بازه‌ی $\mathbb{R} \subset [a, b]$ ناشمارا عدد متعالی وجود دارد.

نتیجه ۴.۲. مجموعه اعداد حقیقی ناشماراست.

قضیه ۵.۲ (۱۸۹۱). مجموعه‌ی همه دنباله‌های نامتناهی ناشماراست.

اثبات. فرض کنید M شامل همه اشیاء $(x_n)_{n \in \mathbb{N}} = S$ است، و برای سادگی فرض کنید در هر دنباله و به ازای هر $n, x_n \in \{0, 1\}$. فرض کنید M شماراست؛ به عبارتی دیگر می‌توان اعضای آن را به صورت یک دنباله نمایش داد. فرض کنید چنین نمایشی باشد. $(\bar{x}_n)_{n \in \mathbb{N}} = \bar{S}$ را به این شکل بسازید: فرض کنید $x_{i,j}$ عضو $-i$ -ام دنباله‌ی S_j باشد، و قرار دهید $S_n = 1 - x_{n,n}$. به بیانی ساده‌تر دنباله‌ی \bar{S} عضو $-n$ -ام دنباله‌ی S_n را برمی‌گزیند و مقدار آن را تغییر می‌دهد. دنباله‌ی ساخته شده، \bar{S} عضو M است، اما جزو $(S_n)_{n \in \mathbb{N}}$ نیست؛ زیرا با هر یک از اعضای آن، حداقل در یک درایه، مغایرت دارد. بنابراین هرگز نمی‌توان M را در یک تناظر یک‌به‌یک با مجموعه‌ی اعداد طبیعی قرار داد. \square

نتیجه ۶.۲. مجموعه‌ی اعداد حقیقی و متعالی ناشمارا هستند.

اثبات. ابتدا نشان دهید اعداد حقیقی در تناظری یک‌به‌یک با بازه‌ی $[1, 0]$ قرار دارد، سپس با در نظر گرفتن نمایش دودویی اعداد در این بازه از قضیه‌ی فوق استفاده کنید. \square

روش به کار گرفته شده در برهان فوق به قطعی‌سازی^۱ معروف است و بخشی از شهرت کاتور به خاطر معرفی این ابزار قدرتمند در ریاضی است. ضمناً از هر دو برهان کاتور می‌توان برای ساختن اعداد متعالی استفاده کرد [۱]. هرچند نشان دادیم که مجموعه‌ی اعداد متعالی ناشماراست، لکن اثبات این که یک عدد خاصی متعالی است، کار ساده‌ای نیست.

در ادامه قصد داریم متعالی بودن اعداد e و π را نشان دهیم. ابتدا ثابت خواهیم کرد که e ریشه‌ی هیچ چندجمله‌ای با ضرایب صحیحی نمی‌تواند باشد، و با کمک برهان متعالی بودن e نشان خواهیم داد π نیز متعالی است.

^۱Diagonalization

۳. e متعالی است.

اگر نشان دهیم که e متعالی است، مستقیماً نتیجه خواهد شد که گنگ است؛ لکن برای نوعی ملموس ساختن ایده هایی که بعداً به کار خواهیم برد، ابتدا نشان می دهیم که e گنگ است.

قضیه ۱.۳. e گنگ است.

$$\text{اثبات. فرض کنید } e = \sum_{n=0}^{\infty} \frac{1}{n!} \text{ که } e \in \mathbb{Z} \text{ و } a, b \in \mathbb{Z} \text{ می دانیم.} \\ \frac{a}{b} = \frac{1}{0!} + \frac{1}{1!} + \cdots + \frac{1}{b!} + \cdots = \left(\frac{1}{0!} + \frac{1}{1!} + \cdots + \frac{1}{b!} \right) + \frac{1}{b!} \left(\frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \cdots \right).$$

عبارت داخل پرانتز اول را می توان به صورت $\frac{N}{b!}$ نوشت که $N \in \mathbb{N}$. از طرفی در مورد عبارت داخل پرانتز دوم داریم

$$\delta = \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \cdots < \frac{1}{b+1} + \frac{1}{(b+1)^2} + \cdots = \frac{1}{b} \leq 1.$$

پس عبارت فوق به شکل

$$\frac{a}{b} = \frac{N + \delta}{b!} \quad (1.3)$$

در می آید که $\frac{1}{b} < \delta$. اگر طرفین (۱.۳) را در $b!$ ضرب کنیم، خواهیم داشت $a(b-1)! = N + \delta$ و نتیجتاً

$$a(b-1)! - N = \delta. \quad (2.3)$$

طرف چپ تساوی (۲.۳) یک عدد صحیح است، در حالی که طرف راست آن بین 0 و 1 است که واضحاً غیرممکن است.
□

تمرین ۲.۳. نشان دهید e یک عدد گنگ درجه i دو نیست؛ یعنی نمی تواند ریشه‌ی یک چندجمله‌ای درجه i ناصلف با ضرایب صحیح باشد.

(راهنمایی: معادله‌ی فرضی $ae^i - be + c = 0$ را به صورت $ae^i = be - c$ بازنویسی کنید. دقت کنید که این تفاوت که این بار و برهانی مشابه برهان پیشین را دنبال کنید.)

برهانی که برای متعالی بودن e ارائه خواهیم داد به نوعی مشابه همین برهان است، از این حیث که می توان e^k را به شکل $e^k = \frac{N_k + \delta_k}{N}$ تخمین زد که در آن، N و N_k اعدادی صحیح و δ_k به اندازه‌ی دلخواه کوچک است؛ با این تفاوت که این بار از بسط تیلور برای این تخمین‌ها استفاده نخواهیم کرد.

قضیه ۳.۳. e عددی متعالی است.

اثبات. فرض کنید e ریشه‌ی یک چندجمله‌ای ناصلف با ضرایب صحیح باشد. با انتخاب یک چندجمله‌ای با کمترین درجه خواهیم داشت

$$a_n e^n + \cdots + a_1 e + a_0 = 0, \quad a_0 \neq 0. \quad (3.3)$$

اگر نشان دهیم برای هر $k \in \{1, \dots, n\}$ می توان نوشت

$$e^k = \frac{N_k + \delta_k}{N} \quad (4.3)$$

که در آن، $N, N_1, \dots, N_n \in \mathbb{Z}$ و δ_k به اندازه‌ی دلخواه کوچک هستند، آن‌گاه می توانیم (۴.۳) را به صورت

$$a_0 N + (a_1 N_1 + \cdots + a_n N_n) + (a_1 \delta_1 + \cdots + a_n \delta_n) = 0 \quad (5.3)$$

بازنویسی کنیم. تخمین مان از e^k را طوری خواهیم ساخت که بخش صحیح (۵.۳) ناصلف و اندازه‌ی بخش δ آن به اندازه‌ی دلخواه کوچک — مثلاً کمتر از 1 — باشد. به کمک چنین تخمینی ثابت می شود که تساوی (۳.۳) ناممکن است و نتیجتاً e متعالی است.

برای تخمین ازتابع Γ^1 استفاده خواهیم کرد. در این برهان نیازی به تمام قوای این تابع نداریم و به خاطر ارتباطی که بین e^{-x} و فاکتوریل ها برقرار می کند به این تابع علاقه مندیم. با یک انتگرال جزءی جزء می توان نشان داد

$$\frac{1}{(p-1)!} \int_0^\infty e^{-x} x^j dx = \begin{cases} 1, & j = p-1; \\ mp, & m \in \mathbb{Z}, j \geq p. \end{cases} \quad (6.3)$$

همچنین اگر f یک چندجمله ای دلخواه باشد، واضح است که $e^k = \frac{\int_0^\infty e^{k-x} f(x) dx}{\int_0^\infty e^{-x} f(x) dx}$. حال قرار دهید

$$f(x) = x^{p-1} (x-1)^p (x-2)^p \cdots (x-n)^p. \quad (7.3)$$

به ازای هر $k \in \{1, \dots, n\}$ تعریف کنید

$$\begin{cases} N &= \frac{1}{(p-1)!} \int_0^\infty e^{-x} f(x) dx, \\ N_k &= \frac{1}{(p-1)!} \int_k^\infty e^{k-x} f(x) dx, \\ \delta_k &= \frac{1}{(p-1)!} \int_0^k e^{k-x} f(x) dx. \end{cases} \quad (8.3)$$

توجه کنید که

$$f(x) = (-1)^p (-2)^p \cdots (-n)^p x^{p-1} + \sum c_i x^{p+i}$$

پس طبق (6.3) نتیجه می شود

$$N = (-1)^{np} (n!)^p + mp, \quad m \in \mathbb{Z}.$$

اگر عدد اول p را بزرگ تر از n برگزینیم، آنگاه $N \neq np$ و در نتیجه $N \nmid n!$ است. حال N_k را در نظر بگیرید. قرار دهید $y = x - k$ و بخش دوم (8.3) را به صورت

$$N_k = \frac{1}{(p-1)!} \int_0^\infty e^{-y} f(y+k) dy$$

بازنویسی کنید. توجه کنید که $p > |a_0| \cdot p$ پس طبق (6.3) آنگاه

$$\begin{aligned} p \nmid a_0 N, \quad p \mid a_1 N_1 + \cdots + a_n N_n &\implies p \nmid a_0 N + (a_1 N_1 + \cdots + a_n N_n), \\ &\implies a_0 N + (a_1 N_1 + \cdots + a_n N_n) \neq 0. \end{aligned}$$

حال کافی است نشان دهیم اگر p را به اندازه‌ی کافی بزرگ کنیم، می‌توانیم δ_k را به اندازه‌ی دلخواه کوچک نگه داریم. توجه کنید انتگرالی که در بخش سوم (8.3) معرفی می‌کند، در بازه $[0, n]$ تعريف شده است. از طرفی اگر $x \in [0, n]$ بازه $|x - k| \leq n^{(n+1)p-1}$ و با استفاده از این کران بالا نتیجه می‌گیریم

$$\delta_k \leq \frac{e^n \cdot n^{(n+1)p}}{(p-1)!}, \quad \forall k \in \{1, \dots, n\}.$$

با مقداری حسابان مقدماتی به سادگی قابل مشاهده است که اگر $p \rightarrow \infty$ آنگاه $\delta_k \rightarrow 0$. با توجه به توضیحات ابتدای برهان و نتایج حاصل شده تساوی (5.3) و معادلاً (3.2) ناممکن است؛ پس e متعالی است. \square

نشان دادیم تساوی (5.3) ناممکن است. حال معادله‌ی زیر را در نظر بگیرید

$$a_1 e^{b_1} + \cdots + a_n e^{b_n} = 0$$

¹ جهت یادآوری، تابع Γ توسعه تابع فاکتوریل به اعداد مختلط است، بدین نحو که $\Gamma(z) = \int_0^\infty e^{-x} x^{z-1} dx$ که $\Re(z) > 0$ ؛ به طور خاص برای هر $n \in \mathbb{N}$ ، $\Gamma(n) = (n-1)!$

که a_1, \dots, a_n و b_1, \dots, b_n جبری هستند، $\forall i, j \leq n$ $a_i \neq b_j$ و $\exists i \leq n$ $a_i \neq b_i$. می‌توان نشان داد حتی چنین حالتی هم ممکن نیست. این حکم به قضیه لیندمان^۱ معروف است. برای جزئیات بیشتر به [۲] مراجعه کنید.

۴. π متعالی است.

برهانی که در این بخش برای اثبات متعالی بودن عدد π ارائه می‌شود، چارچوبی مشابه برهان متعالی بودن e است؛ لکن در برخی جزئیات نیاز به ابزارهای بیشتری دارد. ما نیز ابتدا چارچوب کلی برهان را در این بخش ترسیم می‌کنیم، سپس جزئیات آن را تکمیل می‌کنیم.

قضیه ۱.۴. π متعالی است.

اثبات. فرض کنید چندجمله‌ای ناصل p با ضرایب صحیح وجود داشته باشد که $p(\pi) = 0$. از این فرض می‌توان نتیجه گرفت که $i\pi$ ریشه‌ی چندجمله‌ای $p(-iz) = p(iz)p(-iz)$ خواهد بود. فرض کنید q یک چندجمله‌ای درجه‌ی n باشد. طبق قضیه‌ی اساسی جبر، این چندجمله‌ای دارای n ریشه خواهد بود. q را به صورت زیر بازنویسی می‌کنیم

$$q(z) = a(z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_n), \quad a \in \mathbb{Z} \setminus \{0\}, \quad \alpha_1 = i\pi \quad (1.4)$$

می‌دانیم $1 + e^{\alpha_1} = 0$ پس

$$(1 + e^{\alpha_1})(1 + e^{\alpha_2}) \cdots (1 + e^{\alpha_n}) = 0$$

تساوی فوق را به صورت $\sum_{k=1}^{r^n} e^{\beta_k} = 0$ بازنویسی می‌کنیم که در آن β_k ها جمع‌های مختلف α_j ‌های مجزاست، و البته شامل 0 . فرض کنید همه‌ی اعضای ناصل $\beta_1, \beta_2, \dots, \beta_m$ باشد؛ بنابراین می‌توان تساوی مذکور را به صورت زیر بازنویسی کرد

$$r + e^{\beta_1} + e^{\beta_2} + \cdots + e^{\beta_m} = 0, \quad r = 2^n - m > 0. \quad (2.4)$$

سعی می‌کنیم برهان ارائه شده برای e را تقلید کنیم. برای $z \in \mathbb{C}$ چندجمله‌ای

$$f(z) = z^{p-1} (g(z))^p \quad (3.4)$$

را تعریف می‌کنیم، که در آن p یک عدد اول است و

$$g(z) = a^m (z - \beta_1)(z - \beta_2) \cdots (z - \beta_m), \quad a \in \mathbb{Z} \setminus \{0\}. \quad (4.4)$$

سپس مقادیر N و δ را مشابه بخش قبل معرفی می‌کنیم، با این ملاحظه که انتگرال‌های به کار رفته در فرمول‌های زیر در صفحه‌ی اعداد مختلط تعریف شده‌اند. در انتگرال‌های زیر $\gamma = \mathbb{R}_{\geq 0}$ ، $t \in \mathbb{R}_{\geq 0}$ ، γ_k خط افقی $t = \beta_k + \gamma$ و γ'_k پاره‌خطی است که β_k به β_k متصل می‌کند.

$$\begin{cases} N &= \frac{1}{(p-1)!} \int_{\gamma} e^{-z} f(z) dz, \\ N_k &= \frac{1}{(p-1)!} \int_{\gamma_k} e^{\beta_k - z} f(z) dz, \\ \delta_k &= \frac{1}{(p-1)!} \int_{\gamma'_k} e^{\beta_k - z} f(z) dz. \end{cases} \quad (5.4)$$

حال طبق قضیه‌ی انتگرال کوشی داریم^۲

$$N e^{\beta_k} = N_k + \delta_k. \quad (6.4)$$

¹Lindemann Theorem

²فرض کنید $t \in \mathbb{R}_{\geq 0}$ و γ مسیری پسته و ذوزنقه‌ای است که رؤوس آن $\beta_k + t$ و β_k باشد، $\int_{\gamma} F = 0$ است. قضیه‌ی کوشی نتیجه می‌دهد که اگر F یک مشتق پذیر در \mathbb{C} باشد، $\int_{\gamma} F = 0$. حال فرض کنید $t \rightarrow +\infty$ از آنجایی که $e^{-t} \rightarrow 0$ ، روی مسیر γ_k^t که خط عمودی واصل بین t و $\beta_k + t$ است، $\int_{\gamma_k^t} e^{\beta_k - z} f(z) dz \rightarrow 0$ برقرار است.

می‌دانیم چندجمله‌ای q دارای ضرایب صحیح است. به کمک لم ۴.۴ که در ادامه خواهد آمد، می‌توان نتیجه گرفت که f و g نیز دارای ضرایب صحیح‌اند. مشابه استدلالی که در برهان پیشین مطرح شد، عدد اول p را بزرگ‌تر از $|a^m|$ برمی‌گزینیم و طبق (۶.۲) ثابت می‌شود که $N \in \mathbb{Z} \setminus p\mathbb{Z}$ در نتیجه $\circ \neq N$. حال می‌توان (۲.۴) را به صورت زیر بازنویسی کرد. ضمناً به یاد آورید که اگر $p > r$ آن‌گاه $rN \notin p\mathbb{Z}$

$$rN + (N_1 + N_2 + \dots + N_m) + (\delta_1 + \delta_2 + \dots + \delta_m) = \circ. \quad (7.4)$$

این جا N_k ‌ها رفتار نسبتاً متفاوتی از خود نشان می‌دهند و به طور کلی اعدادی مختلف هستند. با این حال مجموع N_k ‌ها به طور مطلوبی رفتار می‌کنند؛ به عبارت دقیق‌تر، $N_1 + N_2 + \dots + N_m \in p\mathbb{Z}$. اگر چنین خاصیتی برقرار باشد، بخش صحیح (۷.۴) با استدلالی مشابه برهان پیشین ناصر خواهد بود. جهت اثبات این خاصیت تغییر متغیر $w = z - \beta_k$ را در هر یک از انتگرال‌های بخش دوم (۵.۴) اعمال کنید. با این تغییر متغیر و جمع کردن N_k ‌ها خواهیم داشت

$$N_1 + \dots + N_m = \frac{1}{(p-1)!} \int_{\gamma} e^{-w} h(w) dw, \quad \gamma = \mathbb{R}_{\geq 0}; \quad (8.4)$$

$$h(w) = f(w + \beta_1) + \dots + f(w + \beta_m). \quad (9.4)$$

با توجه به تعریف $f(w) = f(w + \beta_k)$ از این که چندجمله‌ای q دارای ضرایب صحیح است و لم ۴.۴ می‌توان نتیجه گرفت h ضرایب صحیح خواهد داشت، و بنا بر (۶.۳) داریم $N_1 + \dots + N_m \in p\mathbb{Z}$. این حکم در کنار $rN \notin p\mathbb{Z}$ نتیجه می‌دهد که بخش صحیح تساوی (۷.۴) ناصر است.

حال کافی است به بخش δ تساوی (۷.۴) پردازیم. قرار دهید $M = \max |\beta_k|$. اگر $M \leq |z|$ آن‌گاه در این دیسک δ_k ‌ها همگی درون این دیسک محاسبه می‌شوند، پس با استفاده کردن بالای به دست آمده داریم

$$\delta_k \leq \frac{e^M \cdot (2^m |a|^m M^{m+1})^p}{(p-1)!}, \quad \forall k \in \{1, \dots, m\}.$$

مجدداً δ_k را می‌توان به اندازه‌ی دلخواه کوچک کرد؛ زیرا $\delta_k \rightarrow \infty$ هرگاه $p \rightarrow \infty$. برای تکمیل برهان متعالی بودن π کافی است نشان دهیم چندجمله‌ای‌های f , g و h دارای ضرایب صحیح‌اند. با تعاریف زیر شروع می‌کنیم.

تعریف ۲.۴. می‌گوییم چندجمله‌ای $P(x_1, \dots, x_n)$ یک چندجمله‌ای متقارن^۱ است، هرگاه به ازای هر جایگشت σ روی مجموعه‌ی $\{1, \dots, n\}$

$$P(x_1, \dots, x_n) = P(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

تعریف ۳.۴. چندجمله‌ای‌های متقارن پایه‌ای^۲ روی x_1, \dots, x_n عبارت‌اند از

$$s_1(x_1, \dots, x_n) = \sum_i x_i,$$

$$s_2(x_1, \dots, x_n) = \sum_{i < j} x_i x_j,$$

⋮

$$s_t(x_1, \dots, x_n) = \sum_{i_1 < \dots < i_t} x_{i_1} \cdots x_{i_t},$$

⋮

$$s_n(x_1, \dots, x_n) = x_1 \cdots x_n.$$

¹Symmetric Polynomial

²Elementary Symmetric Polynomials

لم ۴.۴ (قضیه‌ی بنیادی چندجمله‌ای‌های متقارن). چندجمله‌ای $P(x_1, \dots, x_n) \in \mathbb{Z}[x_1, \dots, x_n]$ یک چندجمله‌ای متقارن است، اگر و تنها اگر

$$P(x_1, \dots, x_n) = Q(s_1(x_1, \dots, x_n), \dots, s_n(x_1, \dots, x_n)), \quad Q(y_1, \dots, y_n) \in \mathbb{Z}[y_1, \dots, y_n].$$

ضرایب صحیح چندجمله‌ای q در (۱.۴) همگی به شکل $\pm a \cdot s_t(\alpha_1, \dots, \alpha_n)$ هستند. بنابراین اگر $a\beta_1, \dots, a\beta_{2^n} \in \mathbb{Z}[x_1, \dots, x_n]$ متقارن باشد، $P(a\alpha_1, \dots, a\alpha_n) \in \mathbb{Z}$. حکم اخیر درباره چندجمله‌ای‌های صحیح متقارن روی $\mathbb{Z}[x_1, \dots, x_n]$ نیز صدق می‌کند (چرا؟).

چند جمله‌ای

$$(10.4) \quad G(z) = (z - a\beta_1) \cdots (z - a\beta_m) = \frac{1}{z^r} (z - a\beta_1) \cdots (z - a\beta_{2^n})$$

در نظر بگیرید. ضرایب این چندجمله‌ای، خود چندجمله‌ای‌هایی صحیح و متقارن روی $a\beta_1, \dots, a\beta_{2^n}$ هستند، بنابراین همگی صحیح‌اند. طبق (۴.۴) $g(z) = G(az)$ ، بنابراین g و در نتیجه f در (۳.۴) نیز دارای ضرایب صحیح‌اند. برای نشان دادن حکم مشابه برای h در (۹.۴) از چندجمله‌ای‌های زیر استفاده می‌کیم. قرار دهید

$$(11.4) \quad \begin{cases} H(w) &= w^{p-1} (G(w))^p, \\ J(w) &= \frac{1}{w^p} \sum_{k=1}^m H(w + a\beta_k) = \frac{1}{w^p} (-rH(w) + \sum_{k=1}^{2^n} H(w + a\beta_k)). \end{cases}$$

با توجه به صحیح‌بودن ضرایب G مجدداً می‌توان نشان داد H و J نیز چندجمله‌ای‌هایی با ضرایب صحیح هستند. با ترکیب (۳.۴)، (۴.۴) و (۱۰.۴) داریم

$$h(w) = \sum_{k=1}^m (w + \beta_k)^{p-1} (g(w + \beta_k))^p = \frac{aw^p}{(aw)^p} \sum_{k=1}^m (aw + a\beta_k)^{p-1} (G(aw + a\beta_k))^p = aw^p J(aw).$$

از آن جایی که ضرایب چندجمله‌ای J صحیح است، h نیز یک چندجمله‌ای با ضرایب صحیح خواهد بود.

مراجع

- [1] R. Gray, *Georg Cantor and Transcendental Numbers*, The American Mathematical Monthly, 9, (1994), 819-832.
- [2] R. Steinberg and R. M. Redheffer, *Analytic proof of the Lindemann theorem*, Pacific Journal of Mathematics, 2 (1952), 231-242.

* دانشجوی کارشناسی ارشد ریاضی محض، دانشگاه صنعتی شریف
رایانامه: navid.dejbord@gmail.com



محاسبات چندجانبه‌ی امن

امیرحسین ندیری و یاشار طالبی‌راد

چکیده. در این مقاله، نخست محاسبات چندجانبه‌ی امن را تعریف می‌کنیم و به بررسی کاربردهای آن می‌پردازیم. سپس، ویژگی‌های ابتدایی پروتکل‌های محاسبات چندجانبه را بیان و مقالات مرتبط و روش‌های موجود را معرفی و مرور خواهیم کرد. در پایان نیز، به تعدادی از کارهای آتی مرتبط با محاسبات چندجانبه‌ی امن اشاره خواهیم کرد.

۱. مقدمه

در محاسبات چندجانبه‌ی امن (MPC)^۱، تعدادی شرکت‌کننده وجود دارند که هر کدام از آن‌ها دارای یک مقدار مخفی هستند و می‌خواهند به طور مشترک مقدار یک تابع را بر اساس مقادیر ورودی مخفی خود، بدون آشکارسازی آن ورودی‌ها، محاسبه کنند. برای اولین بار، محاسبات چندجانبه‌ی امن در قالب مسئله‌ی میلیونرها^۲ مطرح شد. این مسئله شرایطی را بیان می‌کند که در آن دو فرد ثروتمند بدون افشاکردن میزان دارایی خود و بدون استفاده از شخص سوم معتمد، می‌خواهند بهمند کدام‌یک ثروتمندتر هستند.

در سال‌های اخیر با وجود پیشرفت‌های روزافزون الگوریتم‌های محاسبات چندجانبه‌ی امن، به ویژه در زمینه‌ی کاهش تأخیر و هزینه‌ی ارتباطات، تعداد بسیار کمی از الگوریتم‌های MPC یافت شده‌اند که در شبکه‌هایی با تعداد شرکت‌کنندگان بسیار زیاد، عمل کرد مناسبی دارند. این اتفاق بسیار ناخواهیند است؛ زیرا محاسبات چندجانبه‌ی امن می‌تواند در حل بسیاری از مشکلات موجود در سیستم‌های توزیع شده^۳ به طور قابل توجهی کمک کننده باشد. برای نمونه می‌توان به مسائلی مانند ایجاد الگوریتم‌های یادگیری عمیق بر روی داده‌هایی که در میان خوش‌های^۴ بزرگ سیستم‌ها قرار دارند، در حالی که صاحبان داده‌ها تمایل ندارند داده‌های خود را به صورت خام به اشتراک بگذارند، یا برگزاری یک حراجی در شبکه‌ی کاملاً توزیع شده‌ای مانند بیت‌تورنت^۵ تنها با کمک اعضا^۶ اشاره کرد.

نکته‌ی دیگری که درباره‌ی سیستم‌های توزیع شده در مقیاس بزرگ وجود دارد این است که هر کدام از این سیستم‌ها میزان محدودی منابع دارند و وجود یک تعادل در توزیع فشار^۷ شبکه بر روی تمام شرکت‌کنندگان امری بسیار مهم است. هم‌چنین باید توجه داشت که وجود سیستم‌های مخرب^۸ در شبکه‌های بسیار بزرگ توزیع شده، به دلیل مکانیزم‌های کنترل پذیرش ضعیف^۹، یک اتفاق محتمل است.

۲. تعاریف

هدف الگوریتم‌های محاسبات چندجانبه‌ی امن، محاسبه‌ی یک تابع از مقادیر ورودی افراد به صورت امن است؛ بدون آنکه مقدار ورودی هر فرد به هر طریقی برای بقیه مشخص شود. به بیان ساده‌تر شرکت‌کنندگان P_1, P_2, \dots, P_n و مقادیر $f(d_1, d_2, \dots, d_n) = (y_1, y_2, \dots, y_n)$ را در نظر بگیرید که هر d_i تنها در اختیار P_i است و هدف^{۱۰} محاسبه‌ی

¹Secure Multi-Party Computation

²Millionaire's Problem

³Distributed Systems

⁴Clusters

⁵BitTorrent

⁶Peers

⁷Load

⁸Adversary

⁹Admission Control Mechanisms

است، به طوري که هر y_i خروجي الگوريتم P_i باشد.

دو ويرگي پايه اي از اين دسته از محاسبات انتظار مى رود:

(۱) افشا نشدن ورودی خصوصی هر شركت كننده: در اين محاسبات هر كدام از شركت كننده ها تنها از خروجي ديجران و مقدار ورودی خود آگاه مى شود. برای مثال در مسئله ميليونها که پيشتر بيان شد، هر شخص نباید به ميزان ثروت شخص دیگر دست يابد.

(۲) صحبت مقدار خروجي محاسبه شده توسط تابع: در اين محاسبات اگر تعدادي از شركت كننده ها از پروتوكل اصلی محاسبات منحرف شوند و رفتار مخرب از خود نشان دهند، نباید بتوانند باعث انحراف دیگر شركت كننده ها درست کار به سوي ايجاد خروجي غلط شوند.

با اين حال، در بعضی مواقع شرط دوم را ضعیفتر در نظر مى گيريم و به جای تضمین پایان يافتن محاسبات با خروجي صحيح برای شركت كننده ها درست کار، اجازه متوقف شدن در صورت تشخيص خطأ را به آنها مى دهیم.

به عمليات محاسبه مى شود که الگوريتم های زيادي برای آن وجود دارد. در بعضی از صورت بندی های مسئله، به جای اين که خروجي تابع f به شکل يك بردار n تايی باشد، n تابع مختلف f_1, f_2, \dots, f_n را در نظر مى گيرند که خروجي هر كدام از f_i ها يك عدد (همان y_i) است. در اين صورت مسئله MPC به n مسئله مى محاسبه مى امن تابع تبدیل مى شود.

با وجود تفاوت های بسيار در ويرگي ها و نحوه عمل کرد راه حل های مختلف محاسبات چند جانبه ای امن، اين سистем ها سه نقش اساسی در بين شركت كننده ها خود دارند که هر شركت كننده مى تواند يك يا چند نقش داشته باشد:

- شركت كننده ها ورودی دهنده که اطلاعات محريمانه را به محاسبه كننده ها مى دهد.
- شركت كننده ها دريافت كننده که نتایج را بصورت پاره ای يا كامل از محاسبه كننده ها دريافت مى کند.
- شركت كننده ها محاسبه كننده که به طور مشترك با دیگر محاسبه كننده ها، محاسبات را انجام مى دهد.

دققت كنيد راه حلی بدیهی که مى توان برای محاسبه يك تابع در نظر گرفت اين است که هر كدام از n نفر مقدار خود را به يك فرد مورد اعتماد مانند T بدهند و بعد از آن T مقدار تابع را با استفاده از اين n ورودی محاسبه کرده و T را به افراد بدهد. يعني در حقیقت همه ي n نفر ورودی دهنده و دريافت كننده هستند و T محاسبه كننده است. از آن جا که وجود چنین شخص مورد اعتمادي در عمل امکان پذير نیست، الگوريتم های محاسبات چند جانبه ای امن به ميان آمدند که ما را از وجود شخص مورد اعتماد برای انجام محاسبات بنياز كنند و از همان n نفر برای دريافت کردن و محاسبه کردن استفاده کنند.

هر پروتوكل MPC با ۴ ويرگي اصلی مشخص مى شود:

- عملکرد: روش کلي اجرای پروتوكل.
- نوع امنيت: نشان دهنده سطح امنيت و ميزان اطلاعات بيشتری است که يك مخرب مى تواند در يك سیستم MPC نسبت به همان پروتوكل با استفاده از يك شركت كننده قابل اعتماد به دست آورد.
- مدل مخرب ها: نشان مى دهد که اين پروتوكل در برابر چه نوع مخرب هایی (مثالاً مخرب فعل ^۲ یا منفعل ^۳، یا تقسیم بندی مخرب ها بر اساس توان محاسباتی آنها) مقاوم است.
- مدل شبکه: نشان مى دهد اين پروتوكل در چه نوع شبکه هایی (مثالاً از نظر هماهنگی یا توپولوژی) قابل اجراست.

به عنوان يك مثال ساده مى توان به مسئله زير اشاره کرد:

شبکه ای از شركت كننده ها A_1, A_2, \dots, A_n را در نظر بگيريد که هر كدام يك مقدار مخفی مانند d_i دارند و تمایل ندارند داده های خود را به صورت خام به اشتراك بگذارند، اما تصميم دارند که $\sum_{i=1}^N d_i$ که برابر مجموع مقادير مخفی آنها است را محاسبه کنند.

همان طور که گفته شد، يك راه بسيار ساده استفاده از يك شركت كننده خارجي مورد اعتماد به عنوان محاسبه كننده است که هر شركت كننده دیگر ورودی خود را به او بدهد و او پس از محاسبه مجموع، مقدار خروجي را به آنها بازگردد. با

¹Secure function evaluation

²Active Adversary

³Passive Adversary

این حال روشن است که این روش مبتنی بر وجود یک شخص مورد اعتماد است که هدف اصلی ما در MPC از بین این نیاز است.

به عنوان راه حلی مستقل از شرکت‌کننده‌ی قابل اعتماد می‌توان از این روش استفاده کرد: A_1 عدد تصادفی r را انتخاب می‌کند و مقدار $r + d_1$ را به A_2 می‌دهد و در ادامه هر شرکت‌کننده مقدار دریافتی را با مقدار خود جمع می‌کند و به شرکت‌کننده‌ی بعدی ارسال می‌کند تا در انتهای مجموع مقادیر و r به شرکت‌کننده‌ی آغازین برسد و شرکت‌کننده‌ی آغازین که از مقدار r آگاه است، با کم کردن آن می‌تواند به مقدار مجموع مورد نظر دست یابد.

۳. روش‌ها و الگوریتم‌ها

اولین بار یائو^۱ در سال ۱۹۸۲ مسئله‌ی میلیونرها را که سابقاً به آن اشاره شد، در [۱] مطرح کرد. چهار سال بعد خود او در [۲] مستقیماً برای حل مسئله‌ی محاسبه‌ی امن تابع در حالت خاص دو نفره راه حلی ارائه داد. یک سال بعد در [۳] الگوریتم حل مسئله‌ی محاسبه‌ی امن تابع برای توابع دلخواه و n متغیره ارائه شد، که به الگوریتم GMW^۲ مشهور است. این الگوریتم از پروتکل‌های دانایی صفر^۳ استفاده می‌کرد و صحت الگوریتم در حالت اثربیت درست کار^۴ تضمین شده بود.

یکی از مهم‌ترین قضایایی که شرایطی را درباره‌ی امکان پذیری انجام محاسبات چندجانبه‌ی امن مطرح می‌کرد، در سال ۱۹۸۸ در [۴] بیان و اثبات شد. این قضیه بیان می‌کند هر تابع n متغیره را می‌توان با استفاده از n پردازه^۵ طوری محاسبه کرد که:

(۱) اگر خطای^۶ رخ ندهد، هیچ زیرمجموعه‌ی کمتر از $\frac{n}{t}$ عضوی از شرکت‌کننده‌گان اطلاعاتی درباره ورودی شرکت‌کننده‌گان دیگر دریافت نکنند.

(۲) اگر خطای بیزانسین^۷ داشته باشیم، هیچ زیرمجموعه‌ی کمتر از $\frac{n}{t}$ عضوی از شرکت‌کننده‌گان توانند محاسبه‌ی تابع را دچار اشکال کنند و همچنین توانند اطلاعاتی درباره ورودی شرکت‌کننده‌گان دیگر دریافت کنند.

همچنین، دو کران داده شده بهینه هستند. به طور کلی در مسئله‌های امنیتی در سیستم‌های توزیع شده و به خصوص در MPC، مخرب‌ها را به صورت مجموعه‌ای کنترل شده توسط یک هماهنگ‌کننده مرکزی در نظر می‌گیرند. این مخرب‌ها در دو نوع دسته‌بندی می‌شوند:

(۱) مخرب منفعل: مخرب‌هایی که t نفر را کنترل می‌کنند و می‌توانند شرایط داخلی و مقدار متغیرهای آن‌ها را بینند. این مخرب‌ها طبق پروتکل عمل می‌کنند و بنابراین اشکالی در اجرای الگوریتم ایجاد نمی‌کنند. در حقیقت تنها نگرانی درباره‌ی مخرب‌های منفعل این است که ورودی بقیه شرکت‌کننده‌گان را بفهمند. شرط اول بیان می‌کند که در حالتی که مخرب‌ها منفعل باشند تعدادشان (یا به طور دقیق‌تر تعداد افراد تحت کنترل‌شان که همان t است) باید کمتر از $\frac{n}{t}$ نفر باشد تا اطلاعات اضافه‌ای دریافت نکنند. در صورتی که مخرب منفعل با t نفر تحت کنترل خود تواند ورودی بقیه شرکت‌کننده‌گان یک محاسبه را به دست آورد، آن محاسبه را t -private می‌نامیم.

(۲) مخرب فعال: مخرب‌هایی که t نفر را کنترل می‌کنند و می‌توانند حالت داخلی آن t نفر را بینند. این مخرب‌ها لزوماً طبق پروتکل عمل نمی‌کنند.

شرط دوم بیان می‌کند که در حالتی که مخرب‌ها فعال باشند تعدادشان باید کمتر از $\frac{n}{t}$ نفر باشد تا اطلاعات اضافه‌ای دریافت نکنند و الگوریتم به درستی اجرا شود. در صورتی که مخرب فعال با t نفر تحت کنترل خود تواند ورودی بقیه‌ی شرکت‌کننده‌گان یک محاسبه را به دست آورد، آن محاسبه را t -secure می‌نامیم.

¹Yao

²Goldreich, Micali, Wigderson

³Zero-Knowledge

⁴Honest Majority

⁵Process

⁶Fault

⁷Byzantine

[۴] همچنين الگوريتمي بيان می کند که بتوان يك تابع f را با استفاده از n نفر با شرایط بالا در محیط هماهنگ^۱ محاسبه کرد. در اين مقاله گفته شده که می توانيم فرض کنيم که تابع f چندجمله‌اي است و در نتيجه يك مدار حسابي^۲ برای آن وجود دارد. فهم اين الگوريتم و أكثر الگوريتم‌هاي مقالات مرتبط ديگر نياز به آشنایي با مفهوم الگوريتم تسهييم راز^۳ دارد که روشی را بيان می کند که با استفاده از آن می توان يك راز مانند S را طوري بین n نفر به سهم‌هاي S_1, S_2, \dots, S_n تقسيم کرد به طوري که:

(۱) دانستن حداقل k مورد از سهم‌ها باعث محاسبه‌ی S می شود.

(۲) دانستن $1 - k$ مورد از سهم‌ها هیچ اطلاعاتی درباره‌ی S نمی دهد.

اين الگوريتم را يك تسهييم راز با پارامترهاي (k, n) يا تسهييم راز با آستانه‌ی k می نامند. شاميير^۴ الگوريتم تسهييم رازی را در [۵] ارائه داده است که به تسهييم راز شاميير^۵ مشهور است.

توجه كنيد که الگوريتم‌هاي تسهييم راز فرض می کنند شخصی که می خواهد رازش را به اشتراك بگذارد سهم‌ها را به طور صحيح محاسبه و تقسيم می کند. با حذف اين فرض دسته‌ی ديگري از الگوريتم‌هاي تسهييم راز را خواهيم داشت که به تصديق‌پذير^۶ يا به اختصار VSS مشهور هستند. اين الگوريتم‌ها تسهييم راز را حتى با فرض وجود چنین افرادي ممکن می سازند. الگوريتم مربوط به قضيه‌ی قبل به طور کلي از ۳ مرحله تشکيل شده است:

(۱) مرحله‌ی ورودی که در آن هر نفر با استفاده از يك الگوريتم تسهييم راز مانند الگوريتم شاميير ورودی خود را به n قسمت تقسيم کرده و به هر نفر سهم مربوط به او را می دهد.

(۲) مرحله‌ی محاسبه^۷ که در آن هر نفر مدار محاسباتي f را گيت به گيت شبیه‌سازی کرده و مقدار محاسبه‌شده‌ی هر گيت را به عنوان يك راز مشترك بین افراد نگه‌داری می کند.

(۳) در مرحله‌ی نهايی، مقدار تابع f (خروجي مدار) که محاسبه شده به چند سهم تقسيم می شود و بين افرادي که قرار است بتوانند مقدار نهايی تابع را به کمک هم محاسبه کنند تقسيم می شود.

در [۶] که در همان سال منتشر شد، الگوريتمي بيان شده است که با استفاده از آن، انجام هر پروتوكل MPC امکان‌پذير است، اگر حداقل $\frac{n}{3}$ از اعضاء درست‌كار باشند. اين مقاله برای اولين بار بدون استفاده از تكنيك‌هاي رمزنگاري و با استفاده از يك VSS جديد، از اثريگذاري افراد مخرب جلوگيري کرد. اين دستاوردي بزرگ بود، چرا که اولين بار بود که فهميدن ورودي‌هاي بقيه‌ی اعضاء از نظر رياضي غيرممکن بود، درحالی که روش‌ها و تكنيك‌هاي رمزنگاري حل اين مسئله را صرفاً از نظر محاسباتي سخت و غيرممکن می ساختند.

الگوريتم معرفی شده در [۶] از دو مرحله تشکيل شده است:

(۱) مرحله‌ی تعهد^۸ که در آن از الگوريتم تسهييم راز تصديق‌پذير جديدي استفاده شده که با استفاده از آن، افراد به ورودي خودشان متعهد می شوند و راهي برای عوض کردن آن در مراحل بعدی ندارند.

(۲) مرحله‌ی محاسبه که در آن بعد از اين که همه به ورودي خود متعهد شدند و هر نفر اطلاعات مورد نيازش را از ورودي‌هاي افراد ديگر گرفت، هر شخص به صورت محلی^۹ با استفاده از سهم‌هاي بقيه مقدار تابع را محاسبه می کند.

دققت كنيد در مرحله‌ی اول چون از الگوريتم تسهييم راز تصديق‌پذير استفاده شده، در صورتی که شخصی خرابکار عددی را به اشتباه متعهد شود افراد متوجه می شوند و اقدامات لازم را انجام می دهند. هر دوی اين مقالات وجود کanal‌هاي امن برای فرستادن يا دریافت اطلاعات بين هر جفت از اعضاء را فرض می گيرند. اين دو مقاله و الگوريتم‌هاي ارائه شده در آن‌ها، مقدمه‌ای شدند برای طراحی الگوريتم‌هايی بر مبنای آن‌ها که وقتی با تكنيك‌هايی برای تسریع انجام اعمال اولیه ترکيب می شوند، بهينگی آن‌ها به حدی می رسد که در عمل قابل پياده‌سازی و استفاده باشند. برای مثال در [۱۲] دو کاربرد MPC به طور کامل بررسی شده است که عبارت‌اند از:

¹Synchronous

²Arithmetic Circuit

³Secret Sharing

⁴Shamir

⁵Shamir's Secret Sharing Scheme

⁶Verifiable

⁷Computation

⁸Commitment

⁹Local

(۱) انواعی از مزایده که در آن‌ها پیشنهاد هر شخص باید به هر دلیل مخفی بماند و برای سایر شرکت‌کنندگان فاش نشود.

(۲) ارزیابی^۱ شرکت‌ها که در آن هر شرکت می‌خواهد خودش را طوری با بقیه‌ی شرکت‌ها مقایسه کند که اطلاعات هیچ شرکتی برای شرکت‌های دیگر فاش نشود.

در سال ۱۹۹۰ در [۴] پروتکلی موسوم به BMR^۲ بر مبنای الگوریتم GMW و یک VSS^۳ جدید ارائه شد که زمان اجرای آن بر حسب عمق مدار تابع f خطی بود. همچنین روش کار الگوریتم به طور کلی مشابه با الگوریتم GMW بود. این الگوریتم در سال ۲۰۰۸ در [۵] با پروتکل دیگری به نام BGW^۴ ترکیب شد و بسیار بهینه‌تر شد. تا این زمان همه‌ی الگوریتم‌های ارائه‌شده برای تعداد کمی نفر (n کوچک) عملی بودند و با زیاد شدن n ، حجم محاسبات و یا زمان اجرای الگوریتم بسیار زیاد می‌شد، تا جایی که دیگر قابل استفاده نباشد (به بیان دیگر این الگوریتم‌ها مقیاس‌پذیر نبودند). سرانجام در سال ۲۰۰۶ در [۶] اولین الگوریتم مقیاس‌پذیر برای حل حالت کلی MPC ارائه شد. این الگوریتم علیه مخرب‌های فعلی امن بود و در محیط هماهنگ اجرا می‌شد. اشکال اساسی‌ای که به الگوریتم‌هایی که در مقاله‌های قبل بررسی شده بودند وارد است این است که همگی برای محیط‌های هماهنگ هستند. محیط‌های واقعی مانند اینترنت معمولاً ناهماهنگ^۵ هستند و این باعث می‌شود الگوریتم‌هایی که همگی شده در موارد بسیاری غیر قابل پیاده‌سازی باشند. هم‌چنین اجرای الگوریتم‌های هماهنگ در عمل ممکن است بیشتر از الگوریتم‌های ناهماهنگ طول بکشد؛ زیرا مثلاً در شبکه‌ای که معمولاً سریع است اما گاهی بسیار کند عمل می‌کند، باید هر مرحله‌ی یک الگوریتم هماهنگ به اندازه‌ی بیشینه‌ی زمان رسیدن همه‌ی پیام‌ها طول بکشد، زیرا در غیر این صورت نمی‌توان بین یک شرکت‌کننده‌ی مخرب و کنندی شبکه تمایز قائل شد. در صورتی که در یک الگوریتم ناهماهنگ، اعضا به محض دریافت اطلاعات کافی کارشان را ادامه می‌دهند.

در [۷] که در سال ۲۰۰۹ منتشر شد، یک پروتکل برای حل حالت کلی MPC در محیط‌های ناهماهنگ با مخرب‌های فعلی (کمتر از $\frac{n}{2}$ نفر) و با فرض وجود کانال‌های امن بین هر دو نفر ارائه شد که با دیگر الگوریتم‌های ناهماهنگ برای حل MPC دو تفاوت اساسی داشت:

(۱) وجود یک نقطه‌ی هماهنگی^۶ فرض شده بود که کران بالایی برای زمان رسیدن پیام‌های فرستاده شده از طرف افراد درست کار در نظر گرفته بود (الگوریتم قبل و بعد از نقطه‌ی هماهنگی کاملاً ناهماهنگ است).

(۲) اتمام^۷ الگوریتم فقط به شرط اتمام صحیح مرحله‌ی پیش‌پردازش تضمین شده بود. در حقیقت افراد مخرب می‌توانند باعث اخلال در مرحله‌ی پیش‌پردازش شوند اما اگر این اتفاق رخ ندهد، الگوریتم به اتمام می‌رسد و همه‌ی افراد درست کار خروجی را دریافت می‌کنند.

این مقاله سه نکته‌ی قابل توجه داشت که عبارت‌اند از:

(۱) پیچیدگی زمانی و محاسباتی بهترین الگوریتم‌های MPC آن زمان که $t_{\text{private}} = \min(n^2 k |C|, O(n^2 k))$ بودند (جلوی مخرب‌های منفعل را می‌گرفتند). از $O(n^2 k |C|)$ بود که در آن $|C|$ سایز مدار محاسباتی مورد استفاده برای محاسبه‌ی f و k بیشینه‌ی طول هر ورودی است؛ در حالی که پیچیدگی این الگوریتم که جلوی مخرب‌های فعلی را هم می‌گرفت همان $O(n^2 k |C|)$ بود.

(۲) در این مقاله روشهایی برای اجرای الگوریتم‌های هماهنگ در محیط‌های ناهماهنگ ارائه شده و از آن استفاده شده است.

(۳) نویسنده‌اند برای نشان دادن عملی بودن الگوریتم، آن را به طور کامل پیاده‌سازی و ارزیابی کرده‌اند و در قالب نرم‌افزاری در اختیار عموم گذاشته‌اند.

حال الگوریتم را توضیح می‌دهیم. هدف نهایی در این الگوریتم محاسبه‌ی $y_1, \dots, y_n = f(x_1, \dots, x_n)$ است. در مرحله‌ی پیش‌پردازش، r_i عددی تصادفی در نظر گرفته شده که کمک می‌کند هر نفر ورودی خود را با آن جمع کند، تسهیم کند و سهم‌ها

¹Benchmarking

²Beaver, Micali, and Rogaway

³Ben-Or, Goldwasser, and Wigderson

⁴Asynchronous

⁵Synchronization Point

⁶Termination

را بفرستد. برای شبیه‌سازی یک الگوریتم هماهنگ با R مرحله در محیط ناهماهنگ، الگوریتم زیر برای هر پردازه (مثلاً P_j) پیشنهاد شده است:

(۱) r که شماره‌ی راند است را ۱ قرار بده و برای هر i پیام $m_{j,i,1}$ را که قرار است در اولین مرحله‌ی الگوریتم هماهنگ به P_i فرستاده شود، محاسبه کن.

(۲) $m_{j,1}$ را که قرار است به همه فرستاده شود، محاسبه کن.

(۳) برای هر i ، همه‌ی $m_{j,i,1}$ ها و $m_{j,1}$ را به P_i بفرست.

(۴) تا وقتی $r \leq R$

۱.۴. صبر کن تا همه‌ی پیام‌های $m_{i,j,r}$ و $m_{i,r}$ از همه‌ی P_i ها دریافت شود.

۲.۴. به کمک پیام‌های دریافتنی، پیام‌های خروجی که به شکل $m_{j,i,r+1}$ و $m_{j,i,r+1}$ هستند را محاسبه کن و در نهایت به r یکی اضافه کن.

(۵) متغیر g_j یکی از حالات S یا F را می‌گیرد که به ترتیب بیان گر به درستی اجرایشدن یا به مشکل خوردن در مرحله‌ی پیش‌پردازش است. همچنین M_j شامل تمام پیام‌های به شکل $m_{i,r}$ برای $i \in \{1, \dots, n\}$ و $r \in \{1, \dots, R\}$ می‌باشد. پیام $(check, g_j, M_j)$ را به همه بفرست.

(۶) صبر کن تا همه‌ی $1 - n$ نفر دیگر، پیام‌های $(check, g_i, M_i)$ خود را بفرستند. اگر برای هر i ، داشتیم $S = g_i$ و $s_i = x_i + r_i$ ، $M_i = M_j$ را به همه بفرست.

(۷) صبر کن تا s_j را از همه بگیری و سپس S_j را برابر (s_1, \dots, s_n) قرار بده و به همه بفرست.

(۸) اگر همه‌ی $1 - n$ نفر دیگر S_i خودشان را قبل از timeout مشخص شده به تو فرستادند و برای هر i ، S_i با S_j برابر بود، q_i را برابر S قرار بده. در غیر این صورت q_i را قرار بده.

(۹) یک الگوریتم اجماع بیزانسین^۱ روی q_i ها انجام بده تا همه روی یک مقدار مشترک q به توافق برسند. چون از اجماع بیزانسین استفاده کردیم، می‌توانیم مطمئن باشیم اگر همه‌ی افراد راستگو یک q داشته باشند، مقدار نهایی q هم برابر همان خواهد بود.

حال که مرحله‌ی پیش‌پردازش تمام شده است، وارد مرحله‌ی محاسبه می‌شویم. در این مرحله هر شخص در صورتی که مقدار نهایی q برابر S باشد، سهم‌های x_i را برای هر نفر دیگر از روی s_i و سهم‌های r_i محاسبه می‌کند. سپس الگوریتمی ارائه شده که با استفاده از آن، هر شخص سهم‌های y_i را با استفاده از سهم‌های x_i محاسبه می‌کند. در نهایت، سهم‌های y_i به P_i فرستاده می‌شوند تا وی بتواند y_i نهایی مخصوص خودش را محاسبه کند.

اما هنوز یک مشکل اساسی وجود داشت! الگوریتم ارائه شده با این که در محیط ناهماهنگ به خوبی قابل اجرا بود، مقیاس‌پذیر نبود. بدین ترتیب در سال ۲۰۱۷ در [۱۱] چند پروتکل برای حل مسئله‌ی MPC بین تعداد زیادی شرکت‌کننده معرفی شد که در هر دو محیط هماهنگ و ناهماهنگ قابل اجرا بودند. این الگوریتم‌ها همچنین در برای مخرب‌های فعل امن بودند، اما به این شرط که در محیط‌های هماهنگ کمتر از $\frac{1}{\lambda}$ تعداد کل شرکت‌کننده‌ها را مخرب‌ها تشکیل دهند. این عدد برای محیط‌های ناهماهنگ به $\frac{1}{\lambda}$ کاهش پیدا کرده است.

۴. پیشنهاد برای کارهای آتی

۱.۴. سیستم‌های داده خصوصی به عنوان سرویس. در گذشته سیستم‌های داده به عنوان سرویس^۲ وجود داشتند که تحلیل‌گرانی که برای تحقیقات خود نیاز به داده‌های واقعی داشتند، می‌توانستند داده‌های مورد نیاز خود را از مجموعه‌ای از داده‌های به‌اشتراک‌گذاشته شده در آن سیستم‌ها به دست آورند. این سیستم‌ها مشکلاتی برای تحلیل‌گران به همراه داشتند که از جمله‌ی آن‌ها امکان به وجود آمدن مشکلات امنیتی برای داده‌های به‌اشتراک‌گذاشته شده با تحلیل‌گران و یا آشکار شدن داده‌های خصوصی افراد و نقض حریم خصوصی افراد به صورت ناخواسته بود. همچنین ایجاد تغییراتی در داده‌ها به منظور حفظ حریم خصوصی صاحبان داده می‌تواند باعث کم شدن کیفیت و کاربرد داده شود.

¹Byzantine Agreement

²Data as a Service

در سال‌های اخیر فعالیت‌هایی برای به وجود آمدن سیستم‌های «داده خصوصی به عنوان سرویس» آغاز شده اما آن‌ها در مراحل ابتدایی خود قرار دارند. برای مثال جانا^۱ یکی از این سیستم‌ها است که صرفاً پرسمان^۲ های ابتدایی بر روی آن پیاده‌سازی شده‌اند و برای کارهای آتی می‌توان امکانات جدیدی از جمله پرسمان‌های پیشرفته‌ی موجود در پایگاه‌داده‌های رابطه‌ای را به این سیستم‌ها افزود.

۲.۴. سیستم‌های تحلیل داده امن. سیستم‌های تحلیل داده امن، سیستم‌های مشابه سیستم‌های تجزیه و تحلیل داده‌های کسب و کار فعلی هستند با این تفاوت که داده‌های به صورت خاص رمزگذاری شده دریافت می‌کنند و تحلیل را با حفظ حریم خصوصی انجام می‌دهند و نتایج امن را به کاربر باز می‌گردانند و تنها کاربر قادر به بازگردانی داده‌ها و نتایج است. این سیستم‌ها نیز غالباً دارای پرسمان‌های ابتدایی هستند و نیازمند بهبود در توانایی تحلیل داده‌های ورودی هستند.

۳.۴. سیستم‌های معاملاتی جایگزین. افراد زیادی وجود دارند که تمایل دارند انواع معاملات و مبادلات را به صورت محترمانه انجام دهند. می‌توان با استفاده از MPC سیستم‌های مبادلاتی امنی برای این افراد ایجاد کرد.

تشکر و قدردانی

در پایان از استاد گران‌قدر دکتر صابر صالح کلیبر، که بدون رهنمون‌های ایشان نگارش این نوشته ناممکن بود، کمال سپاس‌گزاری را داریم.

مراجع

- [1] Yao, A. Protocols for secure computations. *FOCS*. **82** pp. 160-164 (1982)
- [2] Yao, A. How to generate and exchange secrets. *27th Annual Symposium On Foundations Of Computer Science (sfcs 1986)*. pp. 162-167 (1986)
- [3] Goldreich, O., Micali, S. & Wigderson, A. How to play any mental game. *Proceedings Of The Nineteenth Annual ACM Symposium On Theory Of Computing*. pp. 218-229 (1987)
- [4] Ben-Or, M., Goldwasser, S. & Wigderson, A. Completeness theorems for non-cryptographic fault-tolerant distributed computation. *Proceedings Of The Twentieth Annual ACM Symposium On Theory Of Computing*. pp. 1-10 (1988)
- [5] Shamir, A. How to share a secret. *Communications Of The ACM*. **22**, 612-613 (1979)
- [6] Chaum, D., Crépeau, C. & Damgård, I. Multiparty unconditionally secure protocols. *Proceedings Of The Twentieth Annual ACM Symposium On Theory Of Computing*. pp. 11-19 (1988)
- [7] Beaver, D., Micali, S. & Rogaway, P. The round complexity of secure protocols. *STOC*. **90** pp. 503-513 (1990)
- [8] Ben-David, A., Nisan, N. & Pinkas, B. FairplayMP: a system for secure multi-party computation. *Proceedings Of The 15th ACM Conference On Computer And Communications Security*. pp. 257-266 (2008)
- [9] Damgård, I. & Ishai, Y. Scalable secure multiparty computation. *Annual International Cryptology Conference*. pp. 501-520 (2006)
- [10] Damgård, I., Geisler, M., Kroigaard, M. & Nielsen, J. Asynchronous multiparty computation: Theory and implementation. *International Workshop On Public Key Cryptography*. pp. 160-179 (2009)
- [11] Dani, V., King, V., Movahedi, M., Saia, J. & Zamani, M. Secure multi-party computation in large networks. *Distributed Computing*. **30**, 193-229 (2017)
- [12] Bogetoft, P., Christensen, D., Damgård, I., Geisler, M., Jakobsen, T., Kroigaard, M., Nielsen, J., Nielsen, K., Nielsen, J., Pagter, J. & Others Secure multiparty computation goes live. *International Conference On Financial Cryptography And Data Security*. pp. 325-343 (2009)

* دانشجوی علوم کامپیوتر، دانشگاه یورک
رایانامه: anadiri@yorku.ca

* دانشجوی علوم کامپیوتر، دانشگاه آلبتا
رایانامه: talebira@ualberta.ca

¹Jana

²Query



هندرسه‌ی فضای اندازه‌های احتمال

امین طالبی*

چکیده. می‌توان فضای اندازه‌های احتمال بر روی یک فضای متريک فشرده را به متريک و سرشناسین^۱ مجهر کرد. در اين مقاله سعی داريم تا به توصيفي مختصر از هندسه‌ی اين فضا پردازيم. ابتدا به مطالعه هندسه و توپولوژي فضای اندازه‌های احتمالي که بر روی متناهي نقطه جمع شده‌اند می‌پردازيم، سپس نشان می‌دهيم همان گونه که به هر عدد حقيقي می‌توان به عنوان يك دنباله کوشی از اعداد گويا نگاه کرد، به هر اندازه‌ی احتمال می‌توان به چشم يك دنباله کوشی از اندازه‌های روی متناهي نقطه از فضا نگريست. انگيزه‌ی ما برای انجام اين کار، كاربرد آن در سистем‌های ديناميکي و نظريه ارگوديك است.

۱. سرآغاز

يکی از سوالات محوري در شاخه سیستم‌های ديناميکي مطالعه موجوداتی است که تحت ديناميک، ناوردا می‌مانند. اندازه‌های احتمال از جمله موجوداتی هستند که می‌توان ناوردا بودن آن‌ها را تحت اثر ديناميک برسی کرد. در بسياري از مسائل موجود در سیستم‌های ديناميکي جواب به اين سوال که چه اندازه‌های احتمالي تحت اثر ديناميک ناوردا می‌مانند، و همچنین توصيف اندازه‌های ناوردا، از اهميت ويرهای برخوردار است. يکی از انواع اندازه‌های احتمال که به صورت طبیعی در سیستم‌های ديناميکي به آن بر می‌خوريم، اندازه‌های احتمال ديراكی هستند، يعني اندازه‌هایي که بر روی تعدادي متناهي از نقاط فضا جمع شده‌اند. رفتار اين نوع اندازه‌ها و دنباله‌هایي از آن‌ها و اين که در فضای تمام اندازه‌های احتمال چگونه پخش می‌شوند، به موضوعاتي مهم و پايه‌اي در سیستم‌های ديناميکي و نظريه ارگوديك مربوط می‌شود. با اين انگيزه، در اين مقاله قصد داريم تا به مطالعه هندسه‌ی فضای اندازه‌ها بر روی یک فضای متريک فشرده پردازيم و نشان می‌دهيم که چگونه اين فضای بزرگ را می‌توان توسط اندازه‌های قابل فهمي همچون اندازه‌های ديراكی تقریب زد و حسی نسبت به آن یافت. در اين مقاله صحبتی از ارتباط اين مباحث با سیستم‌های ديناميکي نخواهيم کرد، و اميد اين است که در مقاله‌اي دیگر به توضیحی مفصل از اين موضوع پردازيم.

۲. نمادگذاري‌ها و تعاريف اوليه

فرض کنيد X يك فضای متريک فشرده مجهر به فاصله d است. احتمالا اکثر افرادي که اين مقاله را می‌خوانند با تعریف اندازه‌ی احتمال آشنا هستند، ولی بيايد دوباره اين مفهوم را تعريف کنيم. برای اين منظور ابتدا نياز داريم تا ساختاري به نام سیگما-جبر را تعريف کنيم:

تعريف ۱.۲. يك سیگما-جبر بر روی X ، يك مجموعه‌ی ناتهی از زیرمجموعه‌های X مانند B است که تحت اعمال مكمل گيري، اشتراك شمارا و اجتماع شمارا بسته باشد.

تعداد سیگما-جبرهایي که می‌توان بر روی يك فضای متريک متصور بود بسیار زياد است، اما ما از ميان تمامی سیگما-جبرهای موجود، با سیگما-جبر بول کار خواهيم کرد. ولی جا دارد ذكر کنيم که در مطالعه سیستم‌های ديناميکي، مطالعه اندازه‌های روی سیگما-جبرهای دیگري به جز بول نيز به طور طبیعی ظاهر می‌شود که ما در اين مقاله به آن‌ها نمی‌پردازيم.

تعريف ۲.۲. سیگما-جبر بول، کوچکترین سیگما-جبری است که تمام زیرمجموعه‌های باز فضای متريک X را شامل می‌شود.

تعريف ۳.۲. يك اندازه‌ی احتمال بر روی X ، تابعي مانند μ است که به هر عضو سیگما-جبر بول، عددی حقيقي و نامنفي نسبت می‌دهد به طوری که

$$\mu(\emptyset) = \circ$$

- برای هر تعداد شمارا و دو به دو مجزا از اعضای سیگما-جبر مانند ... A_1, A_2, \dots داریم:

$$\mu(\cup_i A_i) = \Sigma_i \mu(A_i).$$

در اینجا باید ذکر کنیم که در تعریف اندازه‌ی احتمال می‌توانستیم هر سیگما-جبر بورل در نظر بگیریم، ولی در این نوشته ما خود را به اندازه‌های احتمال روی سیگما-جبر بورل محدود می‌کنیم. فضای تمام اندازه‌های احتمال را بر روی فضای X با M_1 نمایش خواهیم داد. گرچه ما خود را به اندازه‌های احتمالی که سیگما-جبر بورل دارند محدود کردیم، ولی این فضای خودی خواهیم بسیار بزرگی محسوب می‌شود. در ادامه سعی می‌کنیم این فضای X را از طریق مطالعه‌ی خانواده‌ای از اندازه‌های ساده که برای تعریف آن‌ها تنها به متناهی نقطه از فضای X نیاز است، بفهمیم. در واقع برای فهم بهتر فضای همه‌ی اندازه‌های احتمال، می‌خواهیم کاری شبیه به ساختن اعداد حقیقی با روش در نظر گرفتن دنباله‌های کوشی از اعداد گویا انجام دهیم. یعنی هر اندازه‌ی احتمال را به صورت دنباله‌ای کوشی از اندازه‌های روی متناهی نقطه در فضای بینیم. فهم اعداد حقیقی به عنوان دنباله‌های کوشی از اعداد گویا چیزی نیست که معمولاً افراد در ذهن خود دارند و خیلی راحت به خاطر قابل تصور بودن اعداد حقیقی، و داشتن شکلی از آن در ذهن با آن کنار می‌آیند. ولی فضای اندازه‌های احتمال فضای بزرگی است و معمولاً افراد با خواص آن را می‌شناسند، لذا نیاز به طی کردن چنین مسیری برای درک آن موجه می‌نماید.

برای فهم بهتر فضای همه‌ی اندازه‌های احتمال، می‌خواهیم کاری شبیه به ساختن اعداد حقیقی با روش در نظر گرفتن دنباله‌های کوشی از اعداد گویا انجام دهیم.

۳. اندازه‌های احتمال بر روی فضاهای متربک متناهی

بد نیست برای شروع فهم بهتر فضای اندازه‌های احتمال، از حالتی شروع کنیم که فضای متربک ما مجموعه‌ای متناهی است. فرض کنید مجموعه X ، یک مجموعه N عضوی باشد. می‌توانیم اعضای آن را با اعداد 1 تا N نام گذاری کنیم. در حقیقت اگر مجموعه این اعداد را با \mathbb{Z}_N نمایش دهیم، می‌توانیم فکر کنیم که فضای متربک X همان \mathbb{Z}_N است که به متربکی مانند d مجهز شده است. یکی از موضوعات قابل تأملی که می‌توان در مورد آن فکر کرد، فضای همه‌ی متربک‌های ممکن بر روی مجموعه‌ی \mathbb{Z}_N است. بیایید این فضای \mathcal{D}_N را با \mathbb{Z}_N نمایش دهیم. یک متربک با نسبت دادن اعداد مثبتی به زوج نقاط فضای \mathbb{Z}_N مشخص می‌شود. در نگاه اول می‌گوییم، برای این که متربک را مشخص کنیم $(\frac{N}{2})$ تا زوج داریم و به ازای هر زوج یک عدد حقیقی مثبت انتخاب می‌کنیم. ولی باید به این نکته توجه کنیم که انتخاب این اعداد مثبت باید طوری باشد که در نامساوی مثلث صدق کند. با این شرط فضای \mathcal{D}_N دیگر همان فضای $\mathbb{R}_{+}^{(N)}$ نخواهد بود، بلکه زیرمجموعه‌ای از این فضا خواهد بود. می‌توان به خواص مختلف این فضای جمله توبولوژی این فضا، هندسه‌ی این فضا و دیگر خواص آن فکر کرد، اما در این مقاله قصد پرداختن به این موضوعات را نداریم، و می‌خواهیم تنها فرض کنیم که متربک d عضوی از این فضا است و سعی کنیم فضای اندازه‌های احتمال بر روی (\mathbb{Z}_N, d) را مطالعه کنیم.

برای ساده ترین حالت، یعنی $N = 1$ ، بدیهی است که فضای اندازه‌های احتمال یک فضای تک عضوی است. برای حالت $N = 2$ ، هر اندازه احتمال با مشخص کردن $(1)(\mu)$ مشخص می‌شود که باید عددی در بازه $[0, 1]$ باشد. پس اندازه‌های احتمال بر روی \mathbb{Z}_2 که به متربکی مانند d مجهز شده را می‌توان با بازه $[0, 1]$ معادل در نظر گرفت. در حالت کلی برای فضای متربک (\mathbb{Z}_N, d) ، یک اندازه‌ی احتمال μ (که سیگما-جبر آن سیگما-جبر بورل است) با مقدار آن روی هر یک از اعضای \mathbb{Z}_N مشخص می‌شود. لذا اگر مقدار $(i)\mu$ را با c_i نشان دهیم، برای مشخص کردن اندازه μ کافیست اعداد c_i را طوری انتخاب کنیم که بردار (c_1, \dots, c_N) در سادک $N - 1$ بعدی Δ_N در \mathbb{R}^N بیفتند که

$$\Delta_N := \{(c_1, \dots, c_N) | \forall i \in \mathbb{Z}_N, 0 \leq c_i \leq 1 \quad \& \quad \sum_{i=1}^N c_i = 1\}.$$

مشاهده کنید که اگر μ و ν دو اندازه‌ی احتمال باشند، برای هر عدد $t \in [0, 1]$ می‌توان از روی این دو اندازه، اندازه‌ی احتمالی جدید ساخت، با گرفتن ترکیب خطی‌ای از این دو اندازه:

$$t\mu + (1-t)\nu.$$

با توجه به این نکته، متوجه می‌شویم که فضای اندازه‌های احتمال بر روی فضای متریک $(\mathbb{Z}_{n,d})$ یک فضای محدب است. از آنجایی که ترکیب محدب دو اندازه‌ی احتمال، معادل ترکیب محدب دو بردار در سادک Δ^N است، محدب بودن فضای اندازه‌های احتمال را با محدب بودن Δ^N نیز می‌توان مشاهده کرد. هر مجموعه‌ی محدب دارای نقاط گوشه‌ای است، یعنی نقاطی که آنها را نمی‌توان به صورت ترکیب محدب دو اندازه‌ی نامساوی با خودش نوشت. می‌توان دید که نقاط گوشه‌ای فضای اندازه‌های احتمال روی $(\mathbb{Z}_{N,d})$ اندازه‌های دیراک یک- نقطه‌ای هستند:

تعريف ۱.۳. یک اندازه‌ی احتمال بر روی مجموعه‌ی X را اندازه‌ی دیراکی یک- نقطه‌ای گویند، هرگاه نقطه‌ای مانند $x \in X$ موجود باشد که به مجموعه‌ی تک عضوی $\{x\}$ اندازه‌ی کامل نسبت دهد. این اندازه را با δ_x نمایش می‌دهیم. اندازه‌ی μ بر روی مجموعه‌ی X را اندازه‌ی دیراکی k - نقطه‌ای گویند هرگاه ترکیب محدب با ضرایب نااصراف از k اندازه دیراک یک- نقطه‌ای بر روی k نقطه دو به دو متفاوت در فضا باشد.

این تعریف در حالتی که مشغول مطالعه‌ی اندازه‌های احتمال روی مجموعه‌های متناهی هستیم تعریفی بدیهی به نظر می‌رسد، زیرا تمام اندازه‌های احتمال دیراکی هستند. اما در قسمت‌های بعدی این مقاله نیز از این تعریف استفاده خواهیم کرد. سوالی که پیش می‌آید این است که پس نقش متریک d در این میان چیست؟ به نظر می‌رسد که فضای تمام اندازه‌های احتمال بر روی $(\mathbb{Z}_{N,d})$ مستقل از d باشد. جواب این است که در واقع هندسه‌ی فضای اندازه‌های احتمال است که با تغییر متریک d تغییر می‌کند. پس باید بگوییم که چه هندسه‌ای بر روی اندازه‌های احتمال در نظر می‌گیریم تا این پاسخ را توجیه کنیم.

۴. هندسه‌ی فضای اندازه‌های احتمال بر روی فضاهای متریک متناهی

منظورمان از هندسه‌ی یک فضا، متریکی است که بر روی آن در نظرش می‌گیریم. حال فضای $(\mathbb{Z}_{N,d})$ را فیکس کنید. برای مشخص کردن هندسه‌ی $M_1(\mathbb{Z}_{N,d})$ باید بگوییم که فاصله بین دو اندازه‌ی احتمال مانند μ و ν را چه تعریف می‌کنیم. بیایید فرض کنیم این دو اندازه پخش شدن ۱ کیلوگرم خاک را بر روی هر یک از اعضای \mathbb{Z}_N مشخص کرده باشند. یعنی روی i به ترتیب $(i)\mu$ و $(i)\nu$ کیلوگرم خاک ریخته شده. و ما می‌خواهیم به عنوان فاصله‌ی بین μ و ν کمترین انرژی لازم برای تبدیل یکی از آن‌ها به دیگری را در نظر بگیریم. برای این منظور باید ابتدا مفهوم طرح انتقال^۱ را تعریف کنیم. یک طرح انتقال بین این دو اندازه، به ما می‌گوید که برای تبدیل اندازه‌ی μ به اندازه‌ی ν چه مقدار از خاک ریخته شده بر روی جایگاه i ام را به جایگاه r زام منتقل کنیم. واضح است که از طرح انتقالی که اندازه‌ی μ را به اندازه‌ی ν تبدیل می‌کند، می‌توان با برعکس نگاه کردن به ماجرا، به طرح انتقالی برای تبدیل اندازه‌ی ν به اندازه‌ی μ رسید. تعریف ریاضیاتی این مفهوم از این قرار است:

تعريف ۱.۴. یک طرح انتقال بین اندازه‌ی μ و ν یک ماتریس $n \times n$ با درایه‌های نامنفی p_{ij} است که

$$\sum_i p_{ij} = \mu(i), \quad \sum_j p_{ij} = \nu(j)$$

در این تعریف دیدن این نکته راحت است که اگر ماتریس $P_{n \times n}$ طرح انتقالی بین اندازه‌ی μ و اندازه‌ی ν باشد، آنگاه ترانهاده آن $P_{n \times n}^T$ طرح انتقالی بین اندازه‌ی ν و اندازه‌ی μ خواهد بود. حال باید بگوییم که به یک طرح انتقال چگونه یک انرژی، یا یک هزینه نسبت می‌دهیم. قبل از شروع به بیان این مطلب بهتر است این را بگوییم که راههای متنوعی می‌توان برای نسبت دادن انرژی به یک طرح انتقال متصور شد و ما تنها یکی از آن‌ها را که بیشتر با اهداف ما برای مطالعه سیستم‌های دینامیکی همخوانی دارد را توضیح خواهیم داد.

دید ما برای تخصیص انرژی به یک طرح انتقال، دیدی جزء نگرانه است. به این صورت که ابتدا انرژی لازم برای حمل p_{ij} کیلوگرم خاک از مکان i به مکان j را معرفی می‌کنیم، در نتیجه انرژی کل حاصل جمع زدن چنین انرژی‌هایی خواهد بود. انتظاری که ما از انرژی داریم، اولاً این است که نسبت به p_{ij} خطی باشد، یعنی اگر r برابر این مقدار را از i به j منتقل کردیم، r برابر انرژی اولیه نیاز باشد. ثانیاً انتظار داریم که این انرژی به فاصله‌ی $(j,i)d$ وابستگی مستقیم داشته باشد. یعنی با افزایش فاصله بین i و j این انرژی بیشتر شود. وابستگی‌های مستقیم متفاوتی را می‌توان در نظر گرفت، اما ما همان طور که اشاره شد، به

^۱transport plan

منظور اهداف دینامیکی، وابستگی خطی را انتخاب می‌کنیم. به طور خاص می‌توانیم انرژی لازم برای حمل p_{ij} کیلوگرم خاک از مکان i به مکان j را برابر با

$$d(i, j)p_{ij}$$

در نظر بگیریم. واضح است که این عدد در دو خاصیتی که در بالا به آن‌ها اشاره شد صدق می‌کند. حال می‌توان انرژی یک طرح انتقال را به صورت زیر به دست آورد:

$$\sum_{1 \leq i, j \leq n} d(i, j)p_{i, j}.$$

حال که انرژی یک طرح انتقال را تعریف کردیم، می‌توانیم فاصله بین دو اندازه‌ی احتمال روی (\mathbb{Z}_N, d) را تعریف کنیم.

تعريف ۲.۴. فضای تمام طرح‌های انتقال موجود بین اندازه‌ی μ و اندازه‌ی ν را با $(\mu, \nu)_\pi$ نمایش می‌دهیم. طرح انتقال P را یک طرح انتقال بهینه می‌نامیم اگر کمترین انرژی را در بین اعضای $(\mu, \nu)_\pi$ داشته باشد. فاصله‌ی بین دو اندازه‌ی احتمال μ و ν را برابر این مقدار مینیمیم در نظر می‌گیریم و آن را با d_w نمایش می‌دهیم:

$$d_w(\mu, \nu) := \min_{P \in \pi(\mu, \nu)} \sum_{1 \leq i, j \leq n} d(i, j)p_{i, j}$$

چک کردن این که d_w در خواص فاصله صدق می‌کند را به خواننده می‌سپاریم.

حال می‌توانیم بینیم که بعد از فیکس کردن فضای متريک (\mathbb{Z}_N, d) ، فضای $M_1((\mathbb{Z}_N, d))$ با متر d_w نيز تبدیل به یک فضای متريک می‌شود. اگر دوست داشته باشید، می‌توانید این فضا را همان سادک Δ_N بینیید که به متريکی جدید، و نه آن متريکی که از فضای \mathbb{R}^N به ارث می‌برد مجهز شده است. آیا می‌توانید تصور کنید با تغییر متريک d هندسه‌ی $M_1((\mathbb{Z}_N, d))$ چه تغییری می‌کند؟ به طور مثال اگر $(i, j)_d$ را به سمت صفر میل دهیم، فاصله‌ی بین دو اندازه‌ی δ_i و δ_j نيز به سمت صفر میل می‌کند. همچنین می‌توان دید حداکثر فاصله‌ای که دو اندازه‌ی احتمال روی (\mathbb{Z}_N, d) می‌توانند کسب کنند، دقیقاً برابر قطر (\mathbb{Z}_N, d) است.

۵. هندسه‌ی فضای اندازه‌های احتمال بر روی یک فضای متريک فشرده

حال فرض کنید (X, d) یک فضای متريک نه لزوماً متناهی، ولی فشرده باشد. می‌خواهیم هندسه‌ی فضای $M_1((X, d))$ را بفهمیم. بگذارید همانند بخش قبل، از اندازه‌های دیراکی شروع کنیم. تمام اندازه‌های دیراکی حداکثر k - نقطه‌ای را با $M_k^*((X, d))$ نمایش می‌دهیم:

$$M_k^*((X, d)) := \{\sum_{i=1}^k c_i \delta_{x_i} | (c_1, \dots, c_k) \in \Delta_k, x_i \in X\}.$$

دقت کنید چون به تعداد دلخواه می‌توانیم c_i صفر داشته باشیم، تمام اندازه‌های دیراک s - نقطه‌ای برای $s \geq k$ در مجموعه‌ی $M_k^*((X, d))$ حضور دارند.

$$M_1^*((X, d)) \subset M_2^*((X, d)) \subset M_3^*((X, d)) \subset \dots$$

مجموعه‌ی تمام اندازه‌های دیراکی روی (X, d) را با $M_1^*((X, d))$ نمایش می‌دهیم:

$$M_1^*((X, d)) := \cup_{k=1}^{\infty} M_k^*((X, d)).$$

گزاره ۱.۵. $M_1^*((X, d))$ یک مجموعه‌ی محدب است که نقاط گوشاهای آن اعضای $M_1^*((X, d))$ یا همان اندازه‌های دیراک یک- نقطه‌ای هستند.

اثبات. برای اثبات محدب بودن کافیست دقت کنید که اگر $(\mu, \nu) \in M_1^*((X, d))$ و $\mu \in M_k^*((X, d))$ و $\nu \in M_s^*((X, d))$ ، آنگاه برای هر $t \in [0, 1]$ $t\mu + (1-t)\nu$ در $M_1^{k+s}((X, d))$ زندگی می‌کند. از طرفی برای اثبات قسمت دوم حکم، واضح است که هر عضو اندازه‌ی $M_1^*((X, d))$ را به شکل یک ترکیب خطی متناهی از اعضای $M_1^*((X, d))$ می‌توان نوشت و اعضای $M_1^*((X, d))$ را نمی‌توان به صورت ترکیب خطی اندازه‌های دیگری نوشت. \square

همان طور که فاصله‌ی بین دو اندازه‌ی احتمال بر روی یک فضای متریک متناهی را در بخش قبل تعریف کردیم، می‌توانیم بین دو اندازه‌ی دیراکی دلخواه بر روی (X, d) نیز یک فاصله‌ی تعریف کنیم و $M_*^*((X, d))$ را مجهر به یک متریک کنیم. برای این منظور کافی است مفهوم طرح انتقال را برای این حالت کلی تر تعمیم دهیم. فرض کنید $\Sigma_{i=1}^k c_i \delta_{x_i} = \mu$ یک اندازه‌ی دیراک k - نقطه‌ای و $\Sigma_{j=1}^s b_j \delta_{y_j} = \nu$ یک اندازه‌ی دیراک s - نقطه‌ای باشد. یادآوری می‌کنیم که با توجه به تعریف، تمام x_i ها و همچنین تمام y_j ها دو به دو مجزا هستند. اگر دوباره به این دو اندازه‌ی احتمال به چشم دو نحوی پخش شدن یک کیلوگرم خاک به ترتیب بر روی k - نقطه از فضا نگاه کنیم، یک طرح انتقال بین آن‌ها به ما می‌گوید چه مقدار از خاک موجود در نقطه x_i را به نقطه y_j منتقل کنیم. تعریف ریاضیاتی این مفهوم در این حالت از این قرار است:

تعریف ۲.۵. یک طرح انتقال بین دو اندازه‌ی دیراک k - نقطه‌ای μ و s - نقطه‌ای ν یک ماتریس $P_{k \times s}$ است که

$$\Sigma_j p_{ij} = \mu(x_i), \quad \Sigma_i p_{ij} = \nu(y_j).$$

می‌توان به یک طرح انتقال، نه به عنوان یک ماتریس، بلکه به عنوان یک اندازه‌ی دیراکی روی فضای $X \times X$ نگاه کرد:

گزاره ۳.۵. طرح‌های انتقال بین دو اندازه‌ی دیراک k - نقطه‌ای μ بر روی نقاط x_1, \dots, x_k و s - نقطه‌ای ν بر روی نقاط y_1, \dots, y_s در تناظر یک‌به‌یک با اندازه‌های دیراکی حداقل $s \times k$ - نقطه‌ای روی فضای $X \times X$ با فرم $\Sigma_{i,j} c_{ij} \delta_{(x_i, y_j)}$ هستند با این شرط که

$$\Sigma_j c_{ij} = \mu(x_i), \quad \Sigma_i c_{ij} = \nu(y_j).$$

اثبات. کافی است درایه p_{ij} از یک طرح انتقال را همان ضریب c_{ij} در اندازه‌ی $\Sigma_{i,j} c_{ij} \delta_{(x_i, y_j)}$ در نظر بگیرید، و بالعکس. \square

حال همانند قبل می‌توانیم به یک طرح انتقال بین μ و ν انرژی

$$\Sigma_{i,j} d(x_i, y_j) p_{ij}$$

را نسبت دهیم. برای مشخص کردن فاصله‌ی بین این دو اندازه، کمترین انرژی را در بین تمام طرح‌های انتقال بین آن‌ها در نظر می‌گیریم:

تعریف ۴.۵. فضای تمام طرح‌های موجود بین دو اندازه‌ی دیراک k - نقطه‌ای μ بر روی نقاط x_1, \dots, x_k و s - نقطه‌ای ν بر روی نقاط y_1, \dots, y_s را با $\pi(\mu, \nu)$ نمایش می‌دهیم. طرح انتقال P را یک طرح انتقال بهینه نامیم اگر کمترین انرژی را در بین اعضای $\pi(\mu, \nu)$ داشته باشد. فاصله‌ی بین دو اندازه‌ی احتمال μ و ν را برابر این مقدار مینیمیم در نظر می‌گیریم و آن را با d_w نمایش می‌دهیم:

$$d_w(\mu, \nu) := \min_{P \in \pi(\mu, \nu)} \Sigma_{1 \leq i \leq k, 1 \leq j \leq s} d(x_i, y_j) p_{ij}.$$

اثبات این که d_w در این حالت نیز در خواص متریک صدق می‌کند به خواننده واگذار می‌شود.

حال باید فضای $M_*^*((X, d))$ مجهر به متریک d_w را بهتر بفهمیم، همانطور که در بالا دیدیم این فضا از اجتماع مجموعه‌های تو در توی $M_*^k((X, d))$ تشکیل شده است. حال به بررسی خواص توپولوژیک و همچنین خواصی از این فضا می‌پردازیم که متریک d_w در آن‌ها دخیل است. ابتدا به بررسی $M_*^*((X, d))$ می‌پردازیم:

گزاره ۵.۵. (۱) نگاشت $(\tilde{\alpha})$ بین $(X, d) \rightarrow M_*^*((X, d))$ را به این حالت نیز در خواص متریک d_w می‌فرستد. این $\tilde{\alpha}$ را به این حالت نیز در خواص متریک d_w می‌فرستد و $(M_*^*((X, d)), d_w)$ است.

$$(2) diam(X, d) = diam(M_*^*((X, d)), d_w) = diam(M_*^*((X, d)), d_w)$$

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. \square

با توجه به گزاره‌ی بالا متوجه می‌شویم که $\mathcal{M}_\lambda^k((X, d))$ یک کپی ایزومنتیریک از X است و تمام خواص آن با X یکسان خواهد بود. وقتی $k < 1$ است، برای فهمیدن بهتر $\mathcal{M}_\lambda^k((X, d))$ باید به پاسخ به این سوال‌ها پردازیم: آیا می‌توانیم توپولوژی (سرتاسری) این فضا را بر حسب توپولوژی X توصیف کنیم؟ اعضایی از $\mathcal{M}_\lambda^k((X, d))$ که در $\mathcal{M}_\lambda^k((X, d))$ نیستند کدامند؟

حاصل ضرب دکارتی X در خودش به تعداد k مرتبه را با X^k نشان می‌دهیم. یادآوری می‌کنیم که سادک k بعدی Δ_k را به صورت زیر تعریف می‌کردیم:

$$\Delta_k := \{(c_1, \dots, c_k) \mid \forall i \in \mathbb{Z}_k, 0 \leq c_i \leq 1 \quad \& \quad \sum_{i=1}^k c_i = 1\}.$$

به هر $2k$ -تایی مرتب $(x_1, \dots, x_k, c_1, \dots, c_k) \in X^k \times \Delta_k$ به صورت زیر نسبت دهیم:

$$(x_1, \dots, x_k, c_1, \dots, c_k) \mapsto \sum_{i=1}^k c_i \delta_{x_i}.$$

این به ما نگاشتی از $X^k \times \Delta_k$ به $\mathcal{M}_\lambda^k((X, d))$ می‌دهد که آن را p_k می‌نامیم.

$$\begin{array}{ccc} X^k \times \Delta_k & & \\ p_k \downarrow & & \\ \mathcal{M}_\lambda^k((X, d)) & & \end{array}$$

اثبات این نکته که p_k پیوسته است، کار راحتی است.

اگر مانند حالت $1 = k$ نگاشت p_k تناظری یک‌به‌یک برقرار می‌کرد، کار تمام بود. ولی نگاشت p_k به سه دلیل نگاشتی یک‌به‌یک نیست. اولاً این که اعضای X^k دارای ترتیب هستند و نگاشت p_k به این ترتیب حساس نیست. دوماً برای اندازه‌هایی که محمل کمتر از k نقطه دارند، چون در هر نمایش آن‌ها به صورت $(x_1, \dots, x_k, c_1, \dots, c_k)$ حداقل دو تا x_i و x_j مساوی وجود دارد، می‌توانیم به جای c_i و c_j مقادیر t و $c_i - t$ را جایگذاری کنیم و دوباره همان اندازه را بگیریم. سوماً حتی اگر تمام x_i دو به دو متفاوت باشند ولی یکی از c_i ها برابر صفر باشد به این معنی است که اندازه‌ی شما نقطه‌ی x_i مربوطه را در محمل خود نمی‌بیند و نحوه‌های نمایش دیگری برای این اندازه وجود دارد. به طور مثال می‌توانیم x_i را هر نقطه‌ی دلخواهی بگذارید. برای کم کردن میزان غیر یک‌به‌یک بودن نگاشت p ، از فضای خارج قسمتی

$$Q_k(X) := (X^k \times \Delta_k) / \sim$$

کمک می‌گیریم که در آن رابطه‌ی همارزی \sim به صورت زیر تعریف شده است:

$$(x_1, \dots, x_k, c_1, \dots, c_k) \sim (x_{\sigma(1)}, \dots, x_{\sigma(k)}, c_{\sigma(1)}, \dots, c_{\sigma(k)}) \quad \forall \sigma \in S_k.$$

در بالا S_k گروه جایگشت‌های روی \mathbb{Z}_k را نشان می‌دهد. به هر کلاس همارزی $((x_1, \dots, x_k, c_1, \dots, c_k)) \in Q_k(X)$ همانند بالا می‌توانیم یک اندازه‌ی احتمال به صورت زیر نسبت دهیم:

$$[(x_1, \dots, x_k, c_1, \dots, c_k)] \mapsto \sum_{i=1}^k c_i \delta_{x_i}.$$

این به ما نگاشتی از $Q_k(X)$ به $\mathcal{M}_\lambda^k((X, d))$ می‌دهد که آن را q_k می‌نامیم. می‌توان چک کرد که q_k خوش تعریف است.

$$\begin{array}{ccc} Q_k(X) & & \\ q_k \downarrow & & \\ \mathcal{M}_\lambda^k((X, d)) & & \end{array}$$

دیدن این نکته سخت نیست که نگاشت q_k نیز پیوسته است.

اگر نگاشتی که هر عضو $(x_1, \dots, x_k, c_1, \dots, c_k)$ را به کلاس همارزی آن در $Q_k(X)$ می‌برد را با ι_k نمایش دهیم، از تعاریف تیجه می‌شود که دیاگرام زیر جا به جایی است:

$$\begin{array}{ccc} X^k \times \Delta_k & & \\ \downarrow \iota_k & \searrow p_k & \\ Q_k(X) & \xrightarrow{q_k} & \mathcal{M}_1^k((X, d)) \end{array}$$

همچنان نگاشت q_k به دلایل دوم و سوم ذکر شده در بالا، غیر یک‌به‌یک است. اما می‌توانیم زیرمجموعه‌ای خوب (!) از $Q_k(X)$ را بیابیم که q_k روی آن یک‌به‌یک باشد. مجموعه‌ی تمام k -تایی‌های مرتب در X^k که دارای حداقل دو مولفه‌ی برابر هستند را با $D(X^k)$ نمایش می‌دهیم. همچنین تعریف کنید:

$$\Delta_k \supset \Delta_k^\circ := \{(c_1, \dots, c_k) \mid \forall i \in \mathbb{Z}_k, 0 < c_i \leq 1 \quad \& \quad \sum_{i=1}^k c_i = 1\}.$$

می‌توان به Δ_k° به چشم نقاط درونی Δ_k نگاه کرد. حال فضای خارج قسمتی زیر را در نظر بگیرید:

$$Q_k(X) \supset Q_k^\circ(X) := ((X^k \setminus D(X^k)) \times \Delta_k^\circ) / \sim.$$

می‌توان دید که $Q_k^\circ(X)$ زیرمجموعه‌ای باز و چگال از $Q_k(X)$ است. این همان مجموعه‌ی خوبی است که ما به دنبال آن می‌گشتهیم، زیرا نه تنها q_k روی آن یک‌به‌یک است، بلکه یک هموئیمورفیسم به روی تصویرش نیز هست:

گزاره ۶.۵. نگاشت q_k مجموعه‌ی $Q_k^\circ(X)$ را به صورت هموئیمورفیسم به مجموعه‌ی اندازه‌های دیراکی دقیقاً k -تایی، یعنی $\mathcal{M}_1^k((X, d)) \setminus \mathcal{M}_1^{k-1}((X, d))$ می‌برد و مجموعه‌ی $Q_k^\circ(X) \setminus Q_k(X)$ را به صورت پوشش پوشانده باشد. البته این برای همین دادکثر $1 - k$ نقطه‌ای، یعنی $\mathcal{M}_1^k((X, d)) \cap \mathcal{M}_1^{k-1}((X, d))$ می‌برد.

□ اثبات. این گزاره به خواننده واگذار می‌شود.

نتیجه ۷.۵. فضای $\mathcal{M}_1^k((X, d))$ با متريک d_w فشرده است.

□ اثبات. $Q_k(X)$ فضای فشرده است و q_k از اين فضا به $\mathcal{M}_1^k((X, d))$ پوشش پيوسته است، لذا حکم ثابت می‌شود.

گزاره ۶.۵ به ما کمک می‌کند تا $\mathcal{M}_1^k((X, d))$ را به صورت استقرایی بهتر بفهمیم. در حقیقت به ما می‌گوید زیرمجموعه‌ی اعضایی از $\mathcal{M}_1^k((X, d))$ که در $\mathcal{M}_1^{k-1}((X, d))$ نیستند با $Q_k^\circ(X)$ هموئیمورف است. پس با افزایش k به اندازه‌ی یک واحد، می‌فهمیم که مجموعه‌ی نقاطی که به $\mathcal{M}_1^k((X, d))$ اضافه می‌شود از لحاظ توبولوژیک چه شکلی دارد. البته این برای فهمیدن شکل $\mathcal{M}_1^k((X, d))$ کافی نیست. زیرا ما باید بگوییم که تکه‌هایی که در هر مرحله به شکل قبلی اضافه می‌شوند، چطور به آن می‌چسبند. برای این منظور به این نکته دقت کنید که با توجه به گزاره‌ی بالا نگاشت q_k روی $Q^k(X) \setminus Q_k^\circ(X)$ را می‌توان نگاشتی به $\mathcal{M}_1^{k-1}((X, d))$ نیز در نظر گرفت. حال اگر فضای $Q_k(X)$ را از روی زیرمجموعه‌ی $Q^k(X) \setminus Q_k^\circ(X)$ با نگاشت q_k به فضای $\mathcal{M}_1^{k-1}((X, d))$ بچسبانیم، به عنوان چسباندن دو فضای توبولوژیک، فضای توبولوژیک حاصل با $\mathcal{M}_1^k((X, d))$ هموئیمورف خواهد شد. به این گونه، به صورت استقرایی با شروع از $\mathcal{M}_1^1((X, d))$ که با X هموئیمورف است، و با چسباندن صحیح $Q_2(X)$ به آن می‌توانیم شکل $\mathcal{M}_1^2((X, d))$ را به دست آوریم و الی آخر.

توجه کنید که یکی از نتایج دیگری که از گزاره‌ی بالا می‌توان گرفت در مورد بعد $\mathcal{M}_1^m((X, d))$ است. به طور مثال اگر فضای X یک منیفلد m بعدی باشد، می‌توان دید که $Q^k(X) \setminus Q_k^\circ(X)$ یک منیفلد $km + k - 1$ بعدی خواهد بود. این نکته نشان می‌دهد که در این حالت $\mathcal{M}_1^m((X, d))$ یک فضای متناهی بعد نخواهد بود. در حالت کلی می‌توان چک کرد که وقتی تعداد اعضای X نامتناهی باشد $\mathcal{M}_1^m((X, d))$ نیز یک فضای نامتناهی بعد خواهد شد.

حال به معرفی زیرمجموعه‌هایی از (X, d) می‌پردازم که شامل اعضایی هستند که در مطالعه‌ی رفتار آماری سیستم‌های دینامیکی به صورت طبیعی ظاهر می‌شوند.

فضای حاصل ضرب متقارن^۱ $-k$ -ام یک فضای توپولوژیک X ، به صورت خارج قسمت X^k تحت عمل گروه جایگشت‌های k تایی که مولفه‌های اعضای X^k را جایه جا می‌کند، تعریف می‌شود و با نماد $Sym_k(X)$ نمایش داده می‌شود. می‌توان دید که یک کپی همئومورف با $Sym_k(X)$ در $Q_k(X)$ زندگی می‌کند. اگر مقطعی از (کلاف تاری) $\Delta_k \times X^k$ را در نظر بگیرید که در آن تمام ضرایب c_i برابر $\frac{1}{k}$ هستند:

$$E_k := \{(x_1, \dots, x_k, c_1, \dots, c_k) \mid \forall i \in \mathbb{Z}_k \quad c_i = \frac{1}{k}\}.$$

می‌توان دید که $\iota_k(E_k)$ با $Sym_k(X)$ همئومورف است. حال باید تمام اندازه‌هایی که تحت نگاشت q_k از اعضای $\iota_k(E_k)$ می‌توان دید را در نظر بگیریم:

$$\mathcal{E}_\lambda^k((X, d)) := q_k \circ \iota_k(E_k).$$

گزاره ۸.۵. نگاشت q_k یک همئومورفیسم از $\mathcal{E}_\lambda^k((X, d))$ به $\mathcal{E}_\lambda^k((X, d))$ است.

اثبات. برای اثبات حکم بالا سخت‌ترین قسمت کار چک کردن یک‌به‌یک بودن است. برای این منظور دقت کنید که اگر

$$\sum_{i=1}^k \frac{1}{k} \delta_{x_i} = \sum_{i=1}^k \frac{1}{k} \delta_{y_i},$$

آنگاه از مساوی بودن محمل این دو اندازه، مساوی بودن دو مجموعه‌ی $\{x_1, \dots, x_k\}$ و $\{y_1, \dots, y_k\}$ را نتیجه می‌گیریم و از مساوی بودن ضرایب نیز نتیجه می‌گیریم که هر عضو از این مجموعه به تعدادی یکسان در بین x_i ‌ها و y_i ‌ها ظاهر می‌شود. لذا می‌توان با یک جایگشت، x_i ‌ها را به y_i تبدیل کرد. پس دو کلاس همارزی زیریکسان هستند:

$$[(x_1, \dots, x_k, \frac{1}{k}, \dots, \frac{1}{k})] = [(y_1, \dots, y_k, \frac{1}{k}, \dots, \frac{1}{k})].$$

و این یک‌به‌یک بودن را به ما می‌دهد. \square

از آنجایی که $\iota_k(E_k)$ و $Sym_k(X)$ همئومورف هستند، گزاره‌ی بالا به ما می‌گوید که $\mathcal{E}_\lambda^k((X, d))$ یک کپی همئومورف از $Sym_k(X)$ است. پس فهمیدن توپولوژی آن راحت‌تر از فهمیدن توپولوژی $M_\lambda^k((X, d))$ است.

برخلاف مجموعه‌های $M_\lambda^k((X, d))$ مجموعه‌های $\mathcal{E}_\lambda^k((X, d))$ تودرتو نیستند، و این که چه اشتراکی با هم دارند کمی پیچیده‌تر از قبل است:

گزاره ۹.۵. برای هر دو عدد طبیعی $k, s \in \mathbb{N}$ اگر بزرگترین مقسوم علیه مشترک آن دو را با (k, s) نمایش دهیم، آنگاه:

$$\mathcal{E}_\lambda^k((X, d)) \cap \mathcal{E}_\lambda^s((X, d)) = \mathcal{E}_\lambda^{(k, s)}((X, d)).$$

اثبات. ابتدا ثابت کنید که اگر $|k|$ آنگاه $\mathcal{E}_\lambda^l((X, d)) \supset \mathcal{E}_\lambda^k((X, d))$. این یک طرف تساوی را نتیجه می‌دهد. برای اثبات طرف دیگر به این نکته دقت کنید که اگر عضوی از $\mathcal{E}_\lambda^k((X, d))$ باشد که دارای محملی l عضوی باشد، برای این که تمام c_i ‌ها برابر $\frac{1}{k}$ باشند راهی به جز این که $|l|$ باشد باقی نمی‌گذارد. \square

ذکر این نکته مفید است که حرف‌هایی که تا به حال در مورد توپولوژی $M_\lambda^k((X, d))$ زده شد، تنها به توپولوژی فضای X وابسته است، و با تغییر متریک d ، طوری که توپولوژی وابسته به آن تغییری نکند، فضاهایی که در بالا معرفی کردیم در حد همئومورفیسم یکسان باقی می‌مانند. بحث در مورد توپولوژی این فضاهایا به خودی خود مبحث جالبی است و حرف‌های بیشتری می‌توان در این مورد زد. خواننده‌ی علاقه‌مند را تشویق می‌کنیم که به این موضوع بیشتر فکر کند. به خصوص وقتی فضای X را مشخص کنید، این مسئله می‌تواند جالب باشد. یکی از ساده‌ترین حالات این مساله وقتی است که فضای X را دایره یا پاره خط بگیرید. حتی در این حالت فهمیدن توپولوژی $M_\lambda^k((X, d))$ کار ساده‌ای نیست. ما در اینجا به همین مقدار توضیح در مورد توپولوژی $M_\lambda^k((X, d))$ بسنده می‌کنیم و در ادامه به خواص هندسی این فضا که به متریک d_w مجهز شده می‌پردازیم.

^۱symmetric product

دقت کنید که در تعریف متریک d_w متریک روی فضای X دخیل است و خواهیم دید که خواص هندسی $\mathcal{M}_\lambda^k((X, d))$ چگونه به آن وابسته است.

گزاره ۱۰.۵. برای هر $k \in \mathbb{N}$ مجموعه‌ی $\mathcal{M}_\lambda^k((X, d))$ با متریک d_w فضای متریکی کامل است.

از آنجایی که با توجه به نتیجه‌ی ۷.۵ فضای $\mathcal{M}_\lambda^k((X, d))$ فشرده است، و هر فضای متریک فشرده کامل است، حکم بالا به راحتی نتیجه می‌شود. ولی ما در اینجا به عنوان نقطه‌ای شروع برای فهم هندسه‌ی این فضا، اثباتی مستقیم برای آن می‌آوریم.

اثبات. فرض کنید ... $\mu_1, \mu_2, \dots, \mu_n$ دنباله‌ای کوشی از اندازه‌های احتمال متعلق به $\mathcal{M}_\lambda^k((X, d))$ باشد. محمول هر یک از این اندازه‌ها یک زیرمجموعه‌ی حداقل k نقطه‌ای از X است. می‌توان زیردنباله‌ای مانند $\mu_{i_1}, \mu_{i_2}, \dots, \mu_{i_s}$ انتخاب کرد به طوری که محمول μ_{i_n} ها در توپولوژی هاسدورف^۱ به مجموعه‌ای متناهی شامل s عضو مانند $\{x_1, x_2, \dots, x_s\}$ میل کند. می‌توان چک کرد که حتما $\epsilon \leq s$. حال ϵ را آنقدر کوچک بگیرید که گویی شاعع ϵ حول هر یک از x_i ها شامل هیچ x_j دیگری نباشد. اثبات این که حد زیر موجود است کار راحتی است:

$$\lim_n \mu_{i_n}(B_\epsilon(x_j)).$$

اگر این حد را c_j بنامیم می‌توان چک کرد که برای اندازه‌ی

$$\mu_\infty := \sum_{j=1}^s c_j \delta_{x_j}$$

فاصله‌ی d_w اندازه‌های μ_{i_n} از اندازه‌ی μ_∞ به صفر میل می‌کند. از آنجایی که $\{c_j\}_{j=1}^s$ یک دنباله کوشی بود، میل کردن یک زیردنباله‌ی آن به μ_∞ به معنی میل کردن کل دنباله به این اندازه است. برای اتمام اثبات کافیست توجه کنید که چون $s \leq k$

$$\mu_\infty \in \mathcal{M}_\lambda^k((X, d)).$$

□

حال باید به این فکر کنیم که اعضایی از $\mathcal{M}_\lambda^{k-1}((X, d))$ که در $\mathcal{M}_\lambda^k((X, d))$ نیستند حداقل چقدر از این فضا فاصله می‌گیرند. فرض کنید $\mu = \sum_{i=1}^k c_i \delta_{x_i}$ چنین عضوی باشد. ما به دنبال یافتن کران بالایی برای

$$d_w(\mu, \mathcal{M}_\lambda^{k-1}((X, d)))$$

هستیم. با توجه به مفروضاتمان، x_i ها دو به دو متفاوت هستند و تمام c_i ها ناصفند. پس حداقل عددی مانند k داریم که $c_s \leq \frac{1}{k}$ حال اندازه‌ی

$$\mu' := \mu - c_s \delta_{x_s}$$

را در نظر بگیرید. محمول این اندازه شامل $1 - k$ نقطه است، ولی این یک اندازه احتمال نیست زیرا اندازه کل فضا برابر $1 - c_s$ است. برای اندازه احتمال شدن، باید آن را نرمال کنیم:

$$\nu = \frac{1}{1 - c_s} \mu.$$

برای این که تخمینی از (μ, ν) به دست آوریم یک طرح انتقال بین این دو اندازه پیشنهاد می‌دهیم. این دو اندازه به مقدار خوبی باهم اشتراک دارند. برای هر $i \neq s$ مقدار خاک موجود بر روی x_i را دست نمی‌زنیم. مقدار خاک موجود بر روی x_s نیز به نسبت وزن‌های دیگر x_i ها بین آنها تقسیم می‌کنیم. یعنی برای $i \neq s$ به مقدار زیر خاک می‌دهیم:

$$\frac{c_s c_i}{1 - c_s}.$$

^۱ توپولوژی هاسدورف توپولوژی‌ای است که بر روی فضای تمام زیرمجموعه‌های فشرده‌ی یک فضای متریک فشرده گذاشته می‌شود. با این توپولوژی، فضای تمام زیرمجموعه‌های فشرده‌ی X فضایی فشرده خواهد بود ولذا هر دنباله، زیردنباله‌ای همگرا دارد.

می‌توانید چک کنید که این طرح انتقال اندازه‌ی μ را به اندازه‌ی ν تبدیل می‌کند، و انرژی آن برابر

$$\Sigma i \neq s \frac{c_s c_i}{1 - c_s} d(x_s, x_i)$$

خواهد بود. از آنجایی که $d(x_s, x_i) \leq \text{diam}(X, d)$ و این که ممکن است طرح انتقال‌های دیگری با انرژی کمتری بین این دو اندازه وجود داشته باشد، نتیجه می‌گیریم

$$d_w(\mu, \nu) \leq \frac{\text{diam}(X, d)}{k}. \quad (1.5)$$

می‌بینیم که $\mathcal{M}_1^k((X, d))$ خیلی چاق‌تر از $\mathcal{M}_1^{k-1}((X, d))$ نیست، مخصوصاً اگر k عددی بزرگ باشد. سوالی که پیش می‌آید این است که برای $l \in \mathbb{N}$ که خیلی از k بزرگ‌تر است، آیا $\mathcal{M}_1^l((X, d))$ می‌تواند خیلی از $\mathcal{M}_1^k((X, d))$ چاق‌تر باشد؟ گزاره‌ی زیر جوابی منفی به این سوال است:

گزاره ۱۱.۵. برای هر $k \in \mathbb{N}$ تعریف کنید:

$$\epsilon_k := \sup_{\mu \in \mathcal{M}_1^k((X, d)))} d_w(\mu, \mathcal{M}_1^k((X, d))).$$

آنگاه دنباله‌ی ϵ_k نزولی و همگرا به صفر است.

با توجه به این گزاره، اگر بخواهیم فضای $\mathcal{M}_1^*(x, d)$ را به چیزی تشبیه کنیم، پیاز می‌تواند گزینه‌ی خوبی باشد. فضایی لایه‌لایه و تودرتو، که ضخامت لایه‌های بیرونی آن، کم و کمتر می‌شود. البته یک فرق اساسی این است که بعد لایه‌های فضایی اندازه‌های دیراکی ثابت نیست، و افزایش می‌یابد. اثبات این گزاره در راستای هدفی که در این مقاله به دنبال آن هستیم، یعنی فهم بهتر از فضای اندازه‌های احتمال مهم است.

اثبات. برای فضای متریک (X, d) ، عدد $\epsilon_k(X, d)$ را برابر مینیمم مقداری در نظر بگیرید که به ازای آن یک مجموعه‌ی k عضوی $\lim_{k \rightarrow \infty} \epsilon_k(X, d)$ وجود داشته باشد. واضح است که شرط فشرده بودن (X, d) نتیجه می‌دهد که $\epsilon_k(X, d) = 0$. حال عدد k را فیکس کنید و مجموعه‌ی $\{x_1, \dots, x_k\}$ را طوری در نظر بگیرید که یک مجموعه‌ی $\epsilon_k(X, d)$ -چگال باشد. اندازه‌ی دلخواه $\mu \in \mathcal{M}_1^k((X, d))$ را در نظر بگیرید. از روی این اندازه، اندازه‌ای با محمل $\{x_1, \dots, x_k\}$ می‌سازیم که فاصله‌ی آن‌ها با هم کمتر از $\epsilon_k(X, d)$ باشد. بستار گوی به شعاع r حول نقطه‌ی x را با $\overline{B_r(x)}$ نمایش می‌دهیم. تمام جرم توزیع شده توسط μ در مجموعه‌ی $\overline{B_{\epsilon_k(X, d)}(x_1)}$ را به نقطه‌ی x_1 منتقل می‌کنیم. سپس تمام جرم توزیع شده توسط μ در مجموعه‌ی $\overline{B_{\epsilon_k(X, d)}(x_2)} \setminus \overline{B_{\epsilon_k(X, d)}(x_1)}$ را به نقطه‌ی x_2 منتقل می‌کنیم. به همین شکل به صورت استقرایی جلو رفته و در مرحله‌ی s ، جرم پخش شده توسط μ را در $\overline{B_{\epsilon_k(X, d)}(x_s)} \setminus \cup_{i=1}^{s-1} \overline{B_{\epsilon_k(X, d)}(x_i)}$ به نقطه‌ی x_s منتقل می‌کنیم. پس از k مرحله به این اندازه می‌رسیم:

$$\nu := \sum_{s=1}^k c_s \delta_{x_s},$$

که در آن

$$c_s = \sum_{s=1}^k \mu(\overline{B_{\epsilon_k(X, d)}(x_s)} \setminus \cup_{i=1}^{s-1} \overline{B_{\epsilon_k(X, d)}(x_i)}).$$

با توجه به نحوه‌ی به دست آوردن ν ، که در حقیقت خود یک طرح انتقال بین μ و ν است، نتیجه می‌گیریم که

$$d_w(\mu, \nu) \leq \epsilon_k(X, d).$$

□

پس برای اثبات حکم کافی است قرار دهیم $\epsilon_k = \epsilon_k(X, d)$.

اگر X نامتناهی عضو داشته باشد، پیدا کردن دنباله‌ای کوشی از اندازه‌های دیراکی که به هیچ یک از اعضای $\mathcal{M}_1^*((X, d))$ میل نکند کار سختی نیست. به طور مثال فرض کنید x_1, x_2, \dots اعضای دویه‌دو مجزا از X باشند. اندازه‌های احتمال زیر، دنباله‌ای کوشی می‌دهند که به هیچ یک از اعضای $\mathcal{M}_1^*((X, d))$ همگرا نیست:

$$\nu_n := \frac{2^n}{2^n - 1} \sum_{j=1}^n \frac{1}{2^j} \delta_{x_j}.$$

برای کوشی بودن این دنباله می‌توان از روش به دست آوردن نامساوی ۱.۵ استفاده کرد و بدست آورد که

$$d_w(\nu_{n-1}, \nu_n) \leq \frac{\text{diam}(X, d)}{2^n}.$$

برای X نامتناهی، فضای متریک $(\mathcal{M}_1^*((X, d)), d_w)$ کامل نیست.

۶. کاملسازی فضای $(\mathcal{M}_1^*((X, d)), d_w)$

همان طور که قبلاً اشاره کردیم، هدف نهایی ما فهم فضای تمام اندازه‌های احتمال، توسط اندازه‌های احتمال دیراکی، همانند ساختن اعداد حقیقی با در نظر گرفتن دنباله‌های کوشی از اعداد گویا است. در اینجا مجموعه‌ی $\mathcal{M}_1^*((X, d))$ حکم اعداد گویا را برای ما را دارد و دنباله‌های کوشی آن قرار است اندازه‌های احتمال دلخواه را بسازند. در ادامه درباره‌ی کاملسازی این فضا و خواص آن صحبت خواهیم کرد. فضای تمام دنباله‌های کوشی از اندازه‌های دیراکی را در نظر بگیرید:

$$\mathcal{C}_1[(X, d)] := \{\underline{\mu} : (\mu_n)_{n \in \mathbb{N}} \mid \mathcal{M}_1^*((X, d)) \text{ است}\}$$

رابطه‌ی همارزی \sim را روی این فضا این‌گونه تعریف کنید که دو دنباله‌ی کوشی $(\mu_n)_{n \in \mathbb{N}}$ و $(\nu_n)_{n \in \mathbb{N}}$ معادلند هرگاه

$$\lim_{n \in \mathbb{N}} d_w(\mu_n, \nu_n) = 0.$$

اگر فضای متریک ما کامل بود، این تعریف با این که بگوییم هر دو دنباله به یک نقطه میل کنند یکسان می‌شد. ولی ممکن است مانند مثالی که در انتهای فصل قبل آورده شد، دنباله‌ای کوشی باشد، ولی حد نداشته باشد. حال خارج قسمت $\mathcal{C}_1[(x, d)]$ نسبت به این رابطه‌ی همارزی را در نظر بگیرید:

$$\overline{\mathcal{M}}_1(X, d) := \mathcal{C}_1[(x, d)] / \sim'.$$

اگر یک اندازه‌ی دیراکی داشته باشیم، می‌توانیم دنباله‌ی کوشی ثابت را در نظر بگیریم که تمام اعضای دنباله همان اندازه‌ی دیراکی اولیه هستند. پس می‌بینیم که فضای $\overline{\mathcal{M}}_1(X, d)$ یک کپی از $\mathcal{M}_1^*((X, d))$ را در خود دارد. و قبلاً دیدیم که اگر X نامتناهی باشد، حتماً اعضای بیشتری نیز دارد. حال می‌خواهیم فضای $\overline{\mathcal{M}}_1(X, d)$ را مجهز به متریکی کنیم که تحديد آن به کپی موجود از $\mathcal{M}_1^*((X, d))$ در این فضا، همان متریک d_w باشد. برای دو کلاس همارزی $[(\mu_n)_{n \in \mathbb{N}}]$ و $[(\nu_n)_{n \in \mathbb{N}}]$ تعریف کنید

$$\bar{d}_w([(\mu_n)_{n \in \mathbb{N}}], [(\nu_n)_{n \in \mathbb{N}}]) := \lim_{n \rightarrow \infty} d_w(\mu_n, \nu_n).$$

می‌توان چک کرد که در این تعریف این که چه نماینده‌هایی از این دو کلاس همارزی انتخاب شود فرقی ایجاد نمی‌کند. به راحتی می‌توان چک کرد که \bar{d}_w در سه خاصیت متریک صدق می‌کند و همچنین $(\overline{\mathcal{M}}_1(X, d), \bar{d}_w)$ با این متریک کامل است.

فضای متریک $(\overline{\mathcal{M}}_1(X, d), \bar{d}_w)$ کامل است.

از حالا به بعد برای نمایش کپی‌هایی از ((X, d)) \mathcal{M}_1^k ها و $\mathcal{M}_1^k((X, d))$ که در $\overline{\mathcal{M}}_1(X, d)$ وجود دارند، با سوء استفاده از نمادگذاری، از خود این اسمای استفاده می‌کیم:

$$\mathcal{M}_1^k((X, d)) \subset \mathcal{M}_1^*((X, d)) \subset \overline{\mathcal{M}}_1(X, d).$$

گزاره ۱.۶. فضای متریک $(\overline{\mathcal{M}}_1(X, d), \bar{d}_w)$ فشرده است.

اثبات. برای اثبات به این نکته دقت کنید که اگر $(\mu_n)_{n \in \mathbb{N}}$ دنباله‌ای از دنباله‌های کوشی باشد، آنگاه می‌توان نزدیک‌ترین نقطه روی $\mathcal{M}_1^k((X, d))$ به عضو k ام از μ_n ها را در نظر گرفت، و از فشردگی هر یک از $\mathcal{M}_1^k((X, d))$ ها به همراه گزاره‌ی ۱۱.۵ استفاده کرد و دنباله‌ای کوشی ساخت که زیردنباله‌ای از μ_n ها به آن میل می‌کند. \square

۷. تناظر بین $\mathcal{M}_1(X, d)$ و $\overline{\mathcal{M}}_1(X, d)$

در بخش قبل به فضای متریک کامل و شامل (یک کپی از) تمام اندازه‌های دیراکی بود. اگر تعریفی از اندازه‌ی احتمال که در ابتدای این نوشه ارائه شد را فراموش کنیم، می‌توانیم اعضای این فضا را به عنوان اندازه‌های احتمال بر روی فضای (X, d) در نظر بگیریم. در این صورت دیگر نیازی به درگیر شدن با تعاریفی از قبیل سیگما-جبر، مجموعه‌های بورل و غیره که به صورت استاندارد در تعریف اندازه‌ی احتمال نقش ایفا می‌کنند نداریم، چون هیچ یک از این مفاهیم در تعریف جدیدی که از اندازه‌ی احتمال پیشنهاد دادیم ظاهر نمی‌شوند. اندازه‌های احتمال جدید ما یا اندازه‌ی احتمال دیراکی هستند، یا یک دنباله‌ی کوشی از چنین اندازه‌هایی. البته باید دقیق‌تر باشیم و بگوییم یک کلاس همارزی از دنباله‌های کوشی. باید چنین تعریف و دیدگاهی را با تعریف مرسوم از اندازه‌ی احتمال مقایسه کنیم و بینیم آیا این تعریف نیز می‌تواند جایگزینی برای تعریف مرسوم باشد؟ آیا در کاربرد، و اثبات قضایای مربوط به اندازه‌های احتمال مواردی وجود دارد که استفاده از این تعریف کار را نسبت به استفاده از تعریف مرسوم راحت‌تر کند؟ آیا خواصی از اندازه‌های احتمال را که برای تعریف مرسوم به راحتی می‌توان چک کرد، برای این تعریف نیز می‌توان به راحتی چک کرد؟ و از همه مهم‌تر، موجوداتی که به عنوان اندازه‌ی احتمال معروفی کردیم، چه نسبتی با اندازه‌های احتمال مرسوم دارند؟ آیا تناظری یک‌به‌یک و طبیعی بین این موجودات و اندازه‌های احتمال مرسوم وجود دارد؟

به طور مثال از تعریف مرسوم به راحتی نتیجه می‌شود که ترکیب محدب دو اندازه‌ی احتمال، نیز از جنس اندازه‌ی احتمال است. چک کردن این نکته در $\overline{\mathcal{M}}_1(X, d)$ نیز کار راحتی است، و می‌توان دو کلاس همارزی را ترکیب محدب گرفت. اما یکی از اولین انتظاراتی که ما از یک اندازه‌ی احتمال داریم، این است که اندازه‌ی یک زیرمجموعه‌ی اندازه‌پذیر را به ما بدهد. اصلاً تعریف مرسوم از چنین جایی شروع می‌شود. ولی فعلاً برای ما معلوم نیست که یک کلاس همارزی از دنباله‌های کوشی دیراکی به زیرمجموعه‌های فضای چه عددی باید نسبت دهد. پس به نظر در اولین قدم یک چنین تعریفی دچار مشکل می‌شود. اما باید سعی کنیم بینیم چطور می‌توان از روی یک کلاس همارزی در $\overline{\mathcal{M}}_1(X, d)$ به زیرمجموعه‌های فضای اندازه نسبت داد.

فرض کنید $\mu_n = \underline{\mu}$ یک دنباله‌ی کوشی از اندازه‌های دیراکی، و $[\underline{\mu}]$ کلاس همارزی آن باشد. اولین ایده‌ای که به ذهن می‌رسد این است که برای اندازه‌ی زیرمجموعه‌ای مثل $X \subset U$ چنین مقداری را در نظر بگیریم:

$$\lim_{n \rightarrow \infty} \mu_n(U).$$

اما به سرعت می‌توان دید که به چندین دلیل این مقدار گزینه‌ی درستی نیست. اول این که ممکن است حد بالا وجود نداشته باشد. دوم این که در صورت وجود ممکن است به این که ما چه نماینده‌ای از کلاس همارزی $[\underline{\mu}]$ انتخاب کنیم حساس باشد. به طور مثال اگر X را بازه‌ی $[1, 0]$ و U را بازه‌ی $(1, 0)$ در نظر بگیریم و تعریف کنیم

$$\mu_n = \delta_{\frac{1}{n}},$$

که دنباله‌ای کوشی است، می‌بینیم حد بالا برای این دنباله برابر ۱ است. اما می‌توانیم دنباله‌ای کوشی همارز این دنباله را در نظر بگیریم:

$$\mu_n = \delta_{0+}.$$

و با انتخاب یکی در میان از این دو دنباله، دنباله‌ای بسازیم که حد برای آن موجود نباشد. اما از طرفی دقت کنید که هر نماینده‌ای که در این کلاس همارزی بگیریم، از جایی به بعد اعضای آن، بیشتر جرمنشان در نزدیکی صفر است و حتی اگر این جرم‌ها داخل مجموعه‌ی U باشند، با انرژی خیلی کمی می‌توان آن‌ها را به خارج U هل داد. لذا به نظر خوب است که در بالا به جای \liminf استفاده کنیم. اینگونه مشکل وجود نداشتن حد نیز حل می‌شود. اما همچنان وابستگی به دنباله‌ی انتخاب شده از کلاس همارزی وجود دارد. برای رفع این مشکل هم می‌توان روی تمام نماینده‌های کلاس همارزی، اینفیم گرفت و در نهایت به چنین مقداری برای اندازه‌ی U می‌رسیم:

$$\inf_{(\mu_n)_n \in [\underline{\mu}]} \liminf_{n \rightarrow \infty} \mu_n(U).$$

به نظر می‌رسد که این تعریف گزینه‌ی مناسبی برای اندازه‌ی حاصل از کلاس همارزی یک دنباله‌ی کوشی باشد. اما نه لزوماً برای هر نوع مجموعه‌ای! به طور مثال اگر U مجموعه‌ی تک نقطه‌ای x باشد این مقدار صفر می‌شود، ولی انتظار ما این است که این کلاس همارزی همان اندازه‌ی دیراک روی صفر باشد. یا مثلاً اگر U را مجموعه‌ی اعداد گنگ در $[1, 0]$ بگیرید و دنباله‌ی کوشی را به صورت زیر تعریف کنید:

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\frac{i}{n}},$$

این دنباله‌ی کوشی اگر بخواهد معادل یک اندازه باشد، آن اندازه باید اندازه‌ی لبگ باشد، ولی می‌توان دید که $(U)^*$ برابر صفر خواهد شد، ولی اندازه‌ی مجموعه‌ی اعداد گنگ برای اندازه‌ی لبگ برابر یک است.

در دو مثال بالا مجموعه‌ی U مجموعه‌ای باز نیست. دردامه می‌بینیم که تعریف بالا روی خانواده‌ی زیرمجموعه‌های باز X همان خواصی را دارد که یک اندازه‌ی احتمال باید داشته باشد.

مجموعه‌ی تمام زیرمجموعه‌های باز X را با \mathcal{U} نمایش می‌دهیم. دیدیم که به هر کلاس همارزی $[\underline{\mu}]$ می‌توان تابعی از \mathcal{U} به $\mathbb{R}^{\geq 0}$ نسبت داد. نام این نگاشت را ϕ^* می‌نامیم:

$$\phi^* : \overline{\mathcal{M}}_1(X, d) \rightarrow \{\text{تابع حقیقی مقدار روی } \mathcal{U}\}$$

$$[\underline{\mu}] \mapsto \inf_{(\mu_n)_n \in [\underline{\mu}]} \liminf_{n \rightarrow \infty} \mu_n(.).$$

گزاره ۱.۷. برای هر $([\underline{\mu}]) \in \overline{\mathcal{M}}_1(X, d)$ نگاشت $([\underline{\mu}]) \phi^*$ بر روی مجموعه‌های باز، دارای خاصیت سیگما-جمع پذیری است، به این معنی که اگر U_1, U_2, \dots, U_n تعدادی شمارا (یا متناهی) از مجموعه‌های باز دویه دو مجرزا باشد، آنگاه

$$\phi^*([\underline{\mu}])(\cup_{i=1}^n U_i) = \sum_{i=1}^n \phi^*([\underline{\mu}](U_i)).$$

اثبات. برای اثبات به این نکته دقیق کنید که می‌توان در یک کلاس همارزی، دنباله‌ای کوشی یافت که این فیلم را به طور همزمان، هم برای U_i بدهد، و هم برای تک تک U_i ها. این گونه که با شروع از یک دنباله‌ی کوشی دلخواه، برای هر عضو μ_n از دنباله، جرمی از آن اندازه را که در درون و در فاصله‌ی نزدیکی از مرز U_i ها قرار دارد دقیقاً به روی مرز آن U_i منتقل می‌کنیم. این گونه جرم درون یک U_i را داخل هیچ U_j دیگری نریختیم. حال هرچقدر اندیس n بزرگتر شود، فاصله کمتری از مرز U_i ها برای انتقال جرم آن به روی مرز در نظر می‌گیریم. می‌توان دید طی این فرآیند به دنباله‌ای جدید از اندازه‌ها می‌رسیم که کوشی است، و با دنباله‌ی اولیه همارز است. دیدن این نکته سخت نیست که این دنباله مقدار هر دو طرف تساوی را محقق می‌کند. \square

حال می‌خواهیم از روی $([\underline{\mu}]) \phi^*$ یک اندازه‌ی احتمال با سیگما-جبر بورل بسازیم. برای افرادی که تجربه‌ی احتمالاتی یا نظریه‌ی اندازه‌ای دارند، خاصیت $([\underline{\mu}]) \phi^*$ که در گزاره‌ی ۱.۷ به آن اشاره شد آشناست و شبیه خواص پیش-اندازه^۱ است. یک پیش اندازه بر روی یک جبر از زیرمجموعه‌ها^۲ تعریف می‌شود، که نسبت به خاصیت مکمل‌گیری، اجتماع و اشتراک متناهی بسته است. ثابت می‌شود^۳ که هر پیش اندازه متناهی به صورت یکتا به یک اندازه گسترش داده می‌شود. اگر ما بتوانیم از این قضیه استفاده کنیم کار تمام است، ولی مجموعه‌ی زیرمجموعه‌های باز نسبت به مکمل‌گیری بسته نیست. اما حل این مشکل نیز کار سختی نیست و می‌توانیم در نهایت از روی یک دنباله‌ی کوشی یک اندازه بر روی سیگما-جبر بورل بسازیم:

گزاره ۲.۷. برای هر کلاس همارزی $([\underline{\mu}]) \in \overline{\mathcal{M}}_1(X, d)$ اندازه‌ی احتمال یکتای $([\underline{\mu}]) \phi$ چنان موجود است که بر روی زیرمجموعه‌های باز با $([\underline{\mu}]) \phi^*$ برابر است:

$$\forall U \in \mathcal{U}, \phi([\underline{\mu}](U)) = \phi^*([\underline{\mu}](U)).$$

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. \square

¹ pre-measure

² algebra of subsets

³ Carathéodory's extension theorem

می‌توانیم در گزاره‌ی بالا ϕ را به عنوان نگاشتی از (X, d) به $(\overline{\mathcal{M}}_1(X, d), \mathcal{M}_1(X, d))$ در نظر بگیریم:

$$\phi : \overline{\mathcal{M}}_1(X, d) \rightarrow \mathcal{M}_1(X, d).$$

گزاره ۳.۷. نگاشت $\phi : \overline{\mathcal{M}}_1(X, d) \rightarrow \mathcal{M}_1(X, d)$

- یک به یک است،

- بر روی $\mathcal{M}_1^*(X, d)$ به صورت همانی عمل می‌کند:

$$\forall \mu \in \mathcal{M}_1^*(X, d), \quad \phi([\mu_n = \mu]) = \mu.$$

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. \square

سوالی که مطرح می‌شود این است که آیا ϕ یک تناظر یک به یک است؟ یعنی باید بینیم آیا ϕ پوشاست یا خیر. برای این کار باید بینیم که چطور می‌توان در جهت عکس حرکت کرد و به یک اندازه‌ی احتمال دلخواه، دنباله‌ای کوشی از اندازه‌های دیراکی نسبت داد.

فرض کنید $(X, d) \in \mathcal{M}_1^n$ اندازه‌ی احتمالی دلخواه و نه لزوماً دیراکی باشد. ابتدا به صورت استقرایی، دنباله‌ای از افزارهای فضای X را می‌سازیم که قطر اعضای آنها کم و کمتر می‌شود. برای نقاط x_k^k, \dots, x_k^1 در X عدد مثبت ϵ_k را طوری درنظر بگیرید که در خواص زیر صدق کنند:

- مجموعه‌ی $\{x_k^k, \dots, x_k^1\}$ در X یک مجموعه‌ی $-\epsilon_k$ -چگال باشد،

$$\lim_{k \rightarrow \infty} \epsilon_k = 0.$$

برای هر k مجموعه‌ی تمام گوی‌های بسته به ساعت ϵ_k حول x_j^k ها و مکمل آنها را داخل مجموعه‌ی F_k بزیند و قرار دهید

$$G_k := \bigcup_{s=1}^k F_k.$$

شامل تعدادی متناهی عضو است. تمام اشتراکات ممکن این متناهی مجموعه نیز متناهی تا مجموعه خواهند بود. از بین تمام این اشتراکات، اعضای مینیمال را که به صورت سره شامل هیچ اشتراک دیگری نیستند در نظر بگیرید و داخل مجموعه‌ی P_k بزیند. از آنجایی که هر عضو X در یک اشتراک مینیمال خواهد افتاد، و اشتراکات مینیمال با هم اشتراک ندارند، P_k یک افزار از X به ما می‌دهد که قطر اعضای آن حداقل ϵ_k است. حال بباید دنباله‌ی دیراکی ν_k را به این صورت بسازید که داخل هر یک از اعضای افزار P_k یک عضو انتخاب کنید. اگر تعداد اعضای P_k برابر N_k باشد، اعضای افزار را با $A_1^k, A_2^k, \dots, A_{N_k}^k$ و نقاط انتخاب شده از آنها را با $y_1^k, y_2^k, \dots, y_{N_k}^k$ نمایش می‌دهیم. تعریف کنید :

$$\nu_k := \sum_{i=1}^{N_k} \nu(A_i^k) \delta_{y_i^k}.$$

یعنی اندازه‌ی ν_k اندازه‌ای دیراکی بر روی نقاط انتخاب شده از هر عضو افزار است که اندازه‌ی ν به ما می‌گوید به هر نقطه، اندازه‌ی عضوی از افزار که شامل آن است را نسبت دهیم.

لم ۴.۷. دنباله‌ی (ν_k) کوشی است و کلاس همارزی این دنباله‌ی کوشی به انتخاب x_i^k ها و ϵ_k ها بستگی ندارد.

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. \square

با توجه به لم بالا می‌توان نگاشت

$$\psi : \mathcal{M}_1(X, d) \rightarrow \overline{\mathcal{M}}_1(X, d)$$

را به این صورت تعریف کرد که اندازه‌ی ν را به کلاس همارزی دنباله‌ی (ν_k) که در فرآیند بالا ساخته شد ببرد.

گزاره ۵.۷. نگاشت $\psi : \mathcal{M}_1(X, d) \rightarrow \overline{\mathcal{M}}_1(X, d)$ معکوس نگاشت $\phi : \overline{\mathcal{M}}_1(X, d) \rightarrow \mathcal{M}_1(X, d)$ است.

اثبات. اثبات این گزاره به خواننده واگذار می‌شود. \square

پس بالاخره دیدیم که تناظری یک به یک بین دنباله‌های کوشی اندازه‌های دیراکی و اندازه‌های احتمال برقرار است.

۸. انتقال ساختار متریک از $(\mathcal{M}_1(X, d), \bar{d}_w)$ به $(\bar{\mathcal{M}}_1(X, d), d_w)$

با توجه به تناظر یک بهیکی که در بخش قبل به دست آمد می‌توانیم متریک \bar{d}_w را به فضای اندازه‌های احتمال منتقل کنیم. از آنجایی که تناظر ما بر روی زیرمجموعه‌ی $\mathcal{M}_1^*(X, d)$ همانی بود، متریک روی این زیرمجموعه از $(\mathcal{M}_1(X, d), \bar{d}_w)$ خواهد بود که در ابتدا آن را تعریف کردیم. در حالت کلی نیز، متریک جدید بین دو اندازه‌ی احتمال دلخواه μ و ν را با $d_w(\mu, \nu)$ نمایش می‌دهیم

$$d_w(\mu, \nu) := \bar{d}_w(\psi(\mu), \psi(\nu)).$$

با توجه به پوشایش ψ نتیجه می‌گیریم که این نگاشت یک ایزومنتری بین دو فضای $(\mathcal{M}_1(X, d), \bar{d}_w)$ و $(\bar{\mathcal{M}}_1(X, d), d_w)$ است، لذا هر خاصیتی که تا کون برای $(\bar{\mathcal{M}}_1(X, d), d_w)$ اثبات کردیم، برای $(\mathcal{M}_1(X, d), \bar{d}_w)$ نیز برقرار خواهد بود، به طور خاص:

نتیجه ۱۰.۸. فضای $(\mathcal{M}_1(X, d), d_w)$ فشرده است.

۹. بحث و نتیجه گیری

ما با تعریف یک متریک بر روی فضای اندازه‌های احتمال دیراکی شروع کردیم. سپس این فضای را کامل کردیم و به فضای متریک جدیدی رسیدیم که ثابت کردیم در تناظری یک بهیکی با فضای اندازه‌های احتمال قرار دارد. از روی این تناظر، متریکی بر روی فضای اندازه‌های احتمال به دست آوردیم که تحدیدش به اندازه‌های احتمال دیراکی، همان متریکی بود که از آن شروع کرده بودیم. لذا این‌گونه توانستیم بین هر دو اندازه‌ی احتمال دلخواه یک فاصله تعییف کنیم. خوب است به این نکته اشاره کنیم که مسیری که طی کردیم، می‌تواند به عنوان جایگزینی برای تعییف اندازه‌ی احتمال بر روی فضاهای متریک در نظر گرفته شود. یعنی به جای این که یک اندازه‌ی احتمال را تابعی تعییف کنیم که به اعضای سیگما-جبر بول اعدادی نسبت می‌دهد که دارای خواص ذکر شده در ابتدای مقاله باشد، هر اندازه را به عنوان کلاس همارزی یک دنباله‌ی کوشی از اندازه‌های دیراکی بینیم. البته در چنین دیدگاهی برای تعییف اندازه‌های احتمال، فقط اندازه‌های احتمال بر روی سیگما-جبر بول را می‌توانیم بدست آوریم. این که با لفظ "تعییف جدید" از کاری که انجام دادیم یاد می‌کنیم از این منظر نیست که این تعییف را جایگزین تعییف مرسوم از اندازه‌های احتمال در نظر بگیریم، بلکه این است که می‌توان به عنوان دو تعییف معادل به این دو نگاه کرد و با توجه به موقعیت و کاربردی که مد نظرمان است، از هرکدام که کارمان را راحت‌تر به ثمر می‌رساند استفاده کنیم.

ذکر این نکته نیز واجب است که متریک d_w بر روی فضای $(\mathcal{M}_1(X, d), \bar{d}_w)$ مشهوری است که سال‌ها پیش معرفی شده و مورد مطالعه قرار گرفته است. اما تعییف مرسوم این متریک به صورت مستقیم بین دو اندازه‌ی احتمال دلخواه انجام شده و این‌گونه نیست که ابتدا بر روی اندازه‌های دیراکی تعییف شده باشد. در ادبیات مربوطه، فاصله d_w اسمی متفاوتی دارد، که می‌توان به فاصله‌ی واسرشتاین^۱، فاصله‌ی کانترووبیچ-روینشتاین^۲ و یا فاصله‌ی بولدوزر^۳ اشاره کرد. در تعییف مرسوم، یک طرح انتقال به صورت کلی بین دو اندازه‌ی احتمال نه لزوماً دیراکی تعییف می‌شود، و سپس انرژی طرح انتقال تعییف می‌شود و در نهایت فاصله را به عنوان اینفیم اثری‌های تمام طرح‌های انتقال ممکن در نظر گرفته می‌شود. ما نیز از تعییف مرسوم بر روی اندازه‌های دیراکی شروع کردیم، ولی از روشی دیگر آن را بین دو اندازه‌ی دلخواه گسترش دادیم. احکام زیادی مرتبط با هندسه‌ی این فضای استفاده از تعییف مرسوم به دست آمده که ممکن است با استفاده از تعییف جدیدی که در اینجا ارائه دادیم با روشی متفاوت اثبات شوند. به طور مثال می‌توانید اثبات فشردگی این فضای را در [۱] بینید و با اثباتی که در اینجا آورده شده مقایسه کنید.

به عنوان حسن ختم مقاله، بگذارید کمی در مورد ماهیت هندسی این طرز نگاه به فضای اندازه‌های احتمال صحبت کنیم. فضای حاصل ضرب‌های متقارن، در بعضی از شاخه‌های ریاضی بسیار مورد توجه قرار می‌گیرند و مطالعه می‌شوند. به طور مثال وقتی X یک روبه ریمانی دو بعدی است، حاصل ضرب‌های متقارن آن، منیفلدهایی مختلط خواهند شد که به ساختارهای بیشتری نیز مجهز هستند و احکام زیادی در مورد آن‌ها یافت شده است. این نکته بسیار جالب است که فضای $(\mathcal{M}_1^*(X, d), \bar{d}_w)$ تمام حاصل ضرب‌های متقارن فضای X را در خود دارد. یا حتی اگر از طرف دیگر به ماجرا نگاه کنیم، این که فضاهای حاصل

¹ Wasserstein distance

² Kantorovich–Rubinstein distance

³ earth mover distance

ضربهای متقارن به طور همزمان در یک فضای بزرگتر در کنار یکدیگر نشسته‌اند نیز بسیار جالب است. اگر تعریف کنیم

$$\mathcal{E}_1^*(X, d) := \cup_{k=1}^{\infty} \mathcal{E}_1^k(X, d),$$

آنگاه می‌توانیم فرآیند کامل‌سازی‌ای که برای $\mathcal{M}_1^*(X, d)$ انجام دادیم و فضای اندازه‌های احتمال را گرفتیم، برای $\mathcal{E}_1^*(X, d)$ نیز انجام دهیم و دوباره در این حالت نیز به فضای اندازه‌های احتمال می‌رسیم. این نکته به ما می‌گوید که برای $\mathcal{E}_1^k(X, d)$ بزرگ، در فضای اندازه‌های احتمال می‌تواند به دلخواه چگال باشد. داشتن فضایی تقریباً چگال از اندازه‌های احتمال که بسیار ساختارمند باشد و اعمال هندسی زیادی را بتوان بر روی آن انجام داد، این قابلیت را دارد که مسیرهای جدیدی برای فکر در مورد فضای اندازه‌های احتمال و مسائل مرتبط با آن را باز کند.

مراجع

- [1] Figalli, A., & Glaudo, F. (2023). *An invitation to optimal transport, Wasserstein distances, and gradient flows* (2nd edition). Berlin: European Mathematical Society (EMS).

* دانشگاه صنعتی شریف

رایانامه: amin.s.talebi@gmail.com



ارجاع و وجهیت

ویلارد ون اورمن کواین

مقدمه‌ی مترجم

مقاله‌ی «ارجاع و وجهیت» (Reference and Modality) که ترجمه‌ی فارسی آن در صفحات پیش رو آمده است، نخستین بار در در کتاب از چشم‌اندازی منطقی: ۹ جستار منطقی-فلسفی (ویراست اول، ۱۹۵۳؛ ویراست دوم ۱۹۶۱) منتشر شده است (فصل VIII، صص. ۱۵۹-۱۳۹). این مقاله حاوی کامل‌ترین شکل استدلال‌های کواین علیه مشروعیت منطق وجهی مسوز (Quantified Modal Logics) است. برخلاف بسیاری از مقالات دیگر منتشرشده در این کتاب، مقاله‌ی حاضر پیش‌تر به شکل فعلی منتشر نشده بوده است. با این حال، همان‌طور که کواین خود در صفحه‌ی ۱۷۰ از کتاب توضیح می‌دهد، مقاله‌ی عمدتاً ترکیبی است از محتوای دو مقاله‌ی پیش‌تر کواین است در رد منطق وجهی، با مشخصات زیر:

Quine, W. V. (1943). *Notes on Existence and Necessity*. Journal of Philosophy 40 (5):113-127.

Quine, W. V. (1947). *The Problem of Interpreting Modal Logic*. Journal of Symbolic Logic 12 (2):43-48.

توجه به دو نکته‌ی زیر درباره‌ی ترجمه‌ای که در ادامه می‌آید می‌تواند مفید باشد:

- ارجاع‌های درون‌منتهی در پاورقی‌های مقاله، که در ترجمه به شکل یادداشت‌های پایانی در انتهای مقاله آمده، همگی ناظر بر مقالات دیگری است که در کتاب فوق منتشر شده است.
- «مفهوم» و مشتقات آن در ترجمه واژه‌ای ساختگی است که در مقابل «intension» و مشتقات آن در متن اصلی به کار رفته است. واژه‌ی انگلیسی را شاید بتوان به «مفهوم»، «مضمون» یا معادل‌های دیگر برگرداند. مترجم از هیچ کدام از این معادل‌ها راضی نیست. معذلک، در حال حاضر ترجیح اش استفاده از آن واژه‌ی ساختگی است، که البته ساخته‌ی خود او نیست.

در پایان، از دوستان عزیز در «مجله‌ی ریاضی شریف» که سهم مهمی در آماده‌شدن ترجمه برای انتشار و اصلاح خطاهای آن داشتند کمال قدردانی را دارم. و البته مسؤولیت خطاهای باقی‌مانده تنها با من است و تذکر شان موجب امتنان بسیار است.

۱

یکی از اصول بنیادین حاکم بر این‌همانی اصل جایگزین‌پذیری^{*} — یا، چنان که می‌توان به درستی نامیدش، اصل تمایزناپذیری این‌همان‌ها[†] — است. این اصل می‌گوید که، با داشتن یک حکم این‌همانی صادق، یکی از طرفین آن را می‌توان در هر حکم صادقی جایگزین دیگری کرد و نتیجه صادق خواهد بود. یافتن موارد نقض برای این اصل سهل است. برای مثال، احکام:

جورجونه = باریارلی، (۱)

جورجونه به خاطر ابعادش چنین خوانده می‌شد (۲)

این نوشته ترجمه‌ای از کتاب زیر است:

Quine W. V. O., *From a Logical Point of View: Nine Logico-Philosophical Essays, Second Revised Edition*. Harvard University Press, (1980)), 139 - 159.

^{*}Substitutivity

[†]Indiscernibility of identities

صادق اند^{*}: با وجود این، جایگزینی نام 'جورجونه' با نام 'باربارالی'^(۲) را به حکم کاذب:

باربارالی به خاطر ابعادش چنین خوانده می‌شد

تبديل می‌کند. علاوه بر این، احکام:

سیسرون = تولی،^(۳)

'سیسرون' شامل شش حرف است^(۴)

صادق اند[†]: اما جایگزینی نام اول با نام دوم^(۴) را کاذب می‌کند. مع‌هذا، بنیاد اصل جایگزین‌پذیری کاملاً استوار به نظر می‌رسد؛ هر آن‌چه درباره شخص سیسرون (یا جورجونه) بتوان گفت باید درباره شخص تولی (یا باربارالی) نیز، که همان شخص است، صادق باشد.

در مورد^(۴)، این پارادوکس فوراً حل می‌شود. واقعیت این است که^(۴) حکمی راجع به شخص سیسرون نیست، بلکه صرفاً راجع به لفظ 'سیسرون' است. اصل جایگزین‌پذیری را نباید به سیاق‌هایی تعمیم داد که نامی که در معرض جایگزینی است بدون ارجاع سرراست به شیء در آن سیاق‌ها واقع می‌شود. نقض جایگزین‌پذیری صرفاً آشکار می‌سازد که وقوعی که قرار است جایگزین شود ارجاع ماضی[‡] نیست،^۱ یعنی حکم نه تنها به شیء بلکه به شکل نام نیز بستگی دارد. چرا که واضح است که آنچه بتوان درباره‌ی شیء تصدیق کرد، وقتی به آن شیء با هر نام دیگری ارجاع می‌دهیم صادق باقی می‌ماند.

عبارتی که دربردارنده عبارتی دیگر در میان نشانه‌های نقل تکی^[۱] است نامی از آن عبارت دیگر را تشکیل می‌دهد؛ و واضح است که وقوع آن عبارت دیگر یا بخشی از آن، درون سیاق نقل، در حالت کلی، ارجاعی نیست. مشخصاً، وقوع نام شخصی درون سیاق نقل در^(۴) نه ارجاعی است، و نه موضوع اصل جایگزین‌پذیری. این نام شخصی در آن جاتتها به عنوان جزئی از نامی طولانی‌تر واقع می‌شود که، علاوه بر این جزء، شامل دو نشانه‌ی نقل قول نیز است. جایگزین کردن نامی شخصی، درون چنین سیاقی، همان اندازه غیرقابل توجیه است که جایگزین کردن لفظ 'شام' درون سیاق 'احشام'.

مثال^(۲) اندکی ظرف‌تر است، چرا که حکمی درباره‌ی یک شخص است و نه صرفاً درباره‌ی نام او. آن‌چه به خاطر ابعادش چنین و چنان خوانده شده است آن مرد بوده است، و نه نام او. با وجود این، نقض جایگزین‌پذیری نشان می‌دهد که وقوع نام شخصی در^(۲) ارجاعی ماضی نیست. در واقع ساده است که^(۲) را به حکمی دیگر ترجمه کنیم که شامل دو وقوع از آن نام باشد، یکی ارجاعی ماضی و دیگری نه:

جورجونه به خاطر ابعادش 'جورجونه' خوانده می‌شد.^(۵)

وقوع اول ارجاعی ماضی است. جایگزینی بر اساس^(۱)،^(۵) را به حکم دیگری کماکان صادق تبدیل می‌کند:

باربارالی به خاطر ابعادش 'جورجونه' خوانده می‌شد.

وقوع دوم این نام شخصی به اندازه‌ی هر وقوع دیگری درون سیاق نقل غیرارجاعی است. خیلی دقیق نیست که نتیجه بگیریم که وقوعی از یک نام درون سیاق نقل‌های تکی هرگز ارجاعی نیست. این احکام را در نظر بگیرید:

'جورجونه شطرنج بازی می‌کرد' صادق است،^(۶)

'جورجونه' نامی است برای یک شطرنج باز،^(۷)

هر یک از این‌ها، بسته به صدق یا کذب حکم بدون نشانه نقل.

جورجونه شطرنج بازی می‌کرد^(۸)

* اشاره به Giorgio Barbarelli da Castelfranco نقاش ایتالیایی دوره‌ی رنسانس، که نام او، جورجونه، اصطلاحاً به معنای 'جورج بزرگ' است.

[†] اشاره به Marcus Tullius Cicero سیاست‌مدار قرن ۱ پ.م. اهل روم.

صادق یا کاذب است. وقوع نام 'جورجونه' در (۸)، مطابق معیار ما برای وقوع ارجاعی، ارجاعی است، و به همین دلیل وقوع‌های 'جورجونه' در (۶) و (۷) نیز، علی‌رغم حضور علائم نقل قول تکی در (۶) و (۷)، باید ارجاعی باشد. نکته اصلی راجع به نقل قول این نیست که وقوع ارجاعی را لزوماً از بین ببرد، بلکه این است که می‌تواند وقوع ارجاعی را از بین ببرد (و عموماً چنین می‌کند). مثال‌های (۶) و (۷) استثنای هستند چرا که—همان‌طور که از مقایسه‌ی (۶) و (۷) با (۸) آشکار است—محمول‌های خاص 'صادق است' و 'نامی بودن' موجب بی‌اثر شدن نقل‌های تکی می‌شوند.

به عنوان مثالی از سخن رایج دیگری از احکامی که در آن‌ها نام‌ها ارجاعی واقع نمی‌شوند، شخصی را در نظر بگیرید که فیلیپ خوانده می‌شود و وضعیت

(۹) فیلیپ اطلاع ندارد که تولی کاتلین را محکوم کرده است،

یا شاید وضعیت

(۱۰) فیلیپ باور دارد که تگوسیگالپا در نیکاراگوئه است،

درباره‌ی او صادق است. (۹) با جایگزینی بر اساس (۳) تبدیل می‌شود به حکم

(۱۱) فیلیپ اطلاع ندارد که سیسرون کاتلین را محکوم کرده است،

که بی‌شک کاذب است. جایگزینی بر اساس این‌همانی صادق:

تگوسیگالپا = پایتخت هندوراس

نیز حکم صادق (۱۰) را تبدیل می‌کند به حکم کاذب:

(۱۲) فیلیپ باور دارد که پایتخت هندوراس در نیکاراگوئه است.

بنابراین، می‌بینیم که وقوع‌های 'تولی' و 'تگوسیگالپا' در (۹)–(۱۰) ارجاع‌هایی محض نیستند. از این نظر تضادی بنیادین است میان (۹)، یا (۱۰)، و

کراسوس شنیده است که تولی کاتلین را محکوم کرده است.

این حکم رابطه‌ای را میان سه شخص تصدیق می‌کند، و این اشخاص فارغ از نام‌های به کاررفته برای آن‌ها، در این رابطه می‌مانند. اما (۹) را نمی‌توان صرفاً چونان تصدیق رابطه‌ای میان سه شخص، یا (۱۰) را رابطه‌ای میان شخصی، شهری، و کشوری، در نظر گرفت—دست‌کم نه تا وقتی که گفته‌هایمان را چنان تعبیر می‌کنیم که (۹) و (۱۰) صادق قلمداد می‌شوند و (۱۱) و (۱۲) کاذب.

بعضی خوانندگان شاید بخواهند اطلاع نداشتن و باور را چونان رابطه‌هایی میان اشخاص و احکام تفسیر کنند، فلذا (۹) و (۱۰) را بدین نحو بنویسند:

(۱۳) فیلیپ از 'تولی' کاتلین را محکوم کرده است؛ اطلاع ندارد،

(۱۴) فیلیپ باور دارد 'تگوسیگالپا' در نیکاراگوئه است؛

تا هر وقوع غیر ارجاعی محض از یک نام را درون یک سیاق نقل تکی قرار دهنده. چرچ [۵] علیه این [پیشنهاد] استدلال می‌کند. برای این کار او از مفهوم تحلیلیت* بهره می‌گیرد، که برای ما محل تردید است (صفحات ۳۷–۲۳ بالا)؛ با این حال نمی‌توان به‌سادگی از کنار استدلال او گذشت، و ما در اینجا نیازی نداریم که موضوعی در این باره بگیریم. کافی است بگوییم که یقیناً نیازی به بازسازی (۹)–(۱۰) به شکل (۱۳)–(۱۴) نیست. آن‌چه لازم است این است که ملاحظه کنیم که سیاق‌های 'اطلاع ندارد' که ... و 'باور دارد که ...' شبیه به سیاق‌های نقل تکی اند، از این حیث: یک نام ممکن است در حکم S به صورت ارجاعی واقع شود در حالی که در حکم طولانی‌تری که از گنجاندن S در سیاق 'اطلاع ندارد' که ... یا 'باور دارد که ...' تشكیل شده است به صورت ارجاعی واقع نشود. برای این که وضعیت را در یک کلمه خلاصه کنیم، می‌توانیم از سیاق‌های 'اطلاع ندارد' که ...

*Analyticity

و بُاورد دارد که ... به عنوان ارجاعاً تیره^{*} سخن بگوییم.^۲ همین وضعیت درباره‌ی سیاق‌های 'می‌داند که ...'، 'می‌گوید که ...'، 'تردید دارد که ...'، 'متعجب است که ...'، و غیره صادق است. این که همه‌ی سیاق‌های ارجاعاً تیره را بهزور در قالبی نقل‌وار بریزم جالب اما غیرضروری است؛ در عوض می‌توانیم نقل قول را چونان یک سیاق ارجاعاً تیره در کنار موارد متعدد دیگر در نظر بگیریم.

نشان خواهیم داد که تیرگی ارجاعی ایضاً دامن سیاق‌های به‌اصطلاح وجهی[†] 'ضرورتاً ...' و 'ممکن است ...' را نیز می‌گیرد، دست‌کم وقتی به آن‌ها، مثل منطق وجهی لوئیس، معنای ضروت و امکان اکید داده می‌شود.^۳ مطابق معنای اکید 'ضرورتاً' و 'ممکن است'، این احکام صادق قلمداد می‌شوند:

$$9 \text{ ضرورتاً بزرگتر از } 7 \text{ است،} \quad (15)$$

ضرورتاً اگر در ستاره‌ی شامگاهی حیات وجود داشته باشد آن‌گاه در ستاره‌ی شامگاهی حیات وجود دارد، (16)

تعداد سیاره‌ها ممکن است کمتر از ۷ باشد، (17)

و این‌ها کاذب اند:

تعداد سیاره‌ها ضرورتاً بزرگتر از ۷ است، (18)

ضرورتاً اگر در ستاره‌ی شامگاهی حیات وجود داشته باشد آن‌گاه در ستاره‌ی بامدادی حیات وجود دارد، (19)

ممکن است کمتر از ۷ باشد، (20)

ایده‌ی کلی وجهیت‌های اکید مبتنی بر مفهوم متعارف تحلیلیت است، بدین ترتیب که: حکمی به شکل 'ضرورتاً ...' صادق است اگر و تنها اگر حکم سازنده‌ی آن که 'ضرورتاً' بر آن اعمال می‌شود تحلیلی باشد، و حکمی به شکل 'ممکن است ...' کاذب است اگر و تنها اگر نقیض حکم سازنده‌ی آن که 'ممکن است' بر آن اعمال می‌شود تحلیلی باشد. بنابراین (15)–(17) را می‌توان چنین بازنویسی کرد:

$9 < 7$ تحلیلی است، (21)

اگر حیات در ستاره‌ی شامگاهی وجود داشته باشد آن‌گاه حیات در ستاره‌ی شامگاهی وجود دارد تحلیلی است، (22)

تعداد سیاره‌ها کمتر از ۷ نیست تحلیلی نیست، (23)

و به همین ترتیب در مورد (18)–(20).

اکنون می‌توان به راحتی دید که سیاق‌های 'ضرورتاً ...' و 'ممکن است ...' ارجاعاً تیره اند؛ چرا که جایگزینی بر اساس این‌همانی‌های صادق:

تعداد سیاره‌ها = ۹، (24)

ستاره‌ی شامگاهی = ستاره‌ی بامدادی (25)

صدق‌های (15)–(17) را به کذب‌های (18)–(20) تبدیل می‌کند.

توجه کنید که این واقعیت که (15)–(17) معادل (21)–(23) اند، و این واقعیت که '۹' و 'ستاره‌ی شامگاهی' و 'تعداد سیاره‌ها' در (21)–(23) درون نقل قول‌ها واقع شده‌اند، به خودی خود ما را در این نتیجه‌گیری موجه نمی‌کند که '۹' و 'ستاره‌ی شامگاهی' و 'تعداد سیاره‌ها' در (15)–(17) غیرارجاعی واقع شده‌اند. چنین استدلال کردن شبیه خواهد بود به اشاره به معادل بودن (۸) با (۶) و (۷) به عنوان شاهدی که 'جورجونه' در (۸) غیرارجاعی واقع شده است. آنچه نشان می‌دهد که وقوع‌های

*Referentially opaque

[†]Modal

۹ و 'ستاره‌ی شامگاهی'، و 'تعداد سیاره‌ها' در (۱۵)–(۲۰) (و در (۱۸)–(۱۷)) غیرارجاعی است این واقعیت است که جایگزینی بهواسطه (۲۴)–(۲۵) احکام صادق (۱۵)–(۱۷) را به احکامی کاذب (واحکام کاذب (۱۸)–(۲۰) را به احکامی صادق) تبدیل می‌کند.

اشاره شد که برخی ممکن است مایل باشند (۱۳) و (۱۴) را صورت‌بندی‌ای بنیادی‌تر از (۹) و (۱۰) قلمداد کنند. مشابه‌اً بسیاری مایل اند که (۲۱)–(۲۳) را صورت‌بندی‌ای بنیادی‌تر از (۱۵)–(۱۷) پندراند.^۴ اما این هم غیرضروری است. ما یقیناً فکر نمی‌کنیم (۶) و (۷) به نحوی از انجام از (۸) پایه‌ای‌تر اند، و نیازی نداریم که (۱)–(۲۳) را پایه‌ای‌تر از (۱۵)–(۱۷) لحاظ کنیم. آنچه مهم است این است که درک کنیم که سیاق‌های 'ضرورتاً ...' و 'ممکن است ...'، همچون نقل قول و اطلاع ندارد که ... و 'باور دارد که ...'، ارجاعاً تیره اند.

۲

تا اینجا پدیده‌ی تیرگی ارجاعی را با توسل به رفتار الفاظ مفرد توضیح داده‌ایم. اما، می‌دانیم که، الفاظ مفرد با بارنویسی قابل حذف اند (ن.ک. صص. ۷ به بعد، ۸۵، ۱۶۶ به بعد). اشیائی را که در یک نظریه به آنها ارجاع می‌شود باید در نهایت به مثابه‌ی مقادیر متغیرهای تسویر^{*} توضیح داد، و نه به مثابه‌ی چیزهایی که با الفاظ مفرد نامیده می‌شوند. فلذا، اگر تیرگی ارجاعی مشکلی است که ارزش دارد نگران‌اش باشیم، علاوه‌ی این مشکل باید در نسبت با تسویر نیز به همان اندازه‌ی الفاظ مفرد بروز پیدا کند.^۵ پس باید توجه‌مان را معطوف کنیم به تسویر.

پیوند میان نامیدن و تسویر در عملیاتی که به واسطه‌ی آن، از 'سقراط فانی است' نتیجه می‌گیریم که (x فانی است) ($\exists x$)، یعنی 'چیزی فانی است'، مستتر است. این عملیاتی است که پیش‌تر از آن با عنوان تعمیم وجودی سخن گفتیم (ص. ۱۲۰)، با این تفاوت که اکنون لفظ مفرد 'سقراط' را در جایی داریم که آن وقت متغیری آزاد داشتیم. ایده‌ی پشت چنین استنتاجی این است که هر آنچه راجع به شیئی نامیده شده با لفظی مفرد صادق باشد، راجع به چیزی صادق است؛ و به‌وضوح استدلال در جایی که لفظ مفرد مورد بحث [چیزی را] نمی‌نامد توجیه‌اش را از دست می‌دهد. برای مثال، از:

چیزی به عنوان پگاسوس وجود ندارد،

نتیجه نمی‌گیریم:

(چیزی به عنوان x وجود ندارد) ($\exists x$),

یعنی، 'چیزی وجود دارد که آن چنان چیزی وجود ندارد'، یا 'چیزی وجود دارد که وجود ندارد'. چنین استنتاجی البته به همین اندازه در مورد وقوعی غیرارجاعی از اسامی ذات غیرموجه است. تعمیم وجودی از (۲) به:

(x به خاطر ابعادش چنین خوانده می‌شد) ($\exists x$),

خواهد انجامید، یعنی این که، 'چیزی به خاطر ابعادش چنین خوانده می‌شد'. این با توجه به این که دیگر مرجع مناسبی برای 'چنین خوانده می‌شد' وجود ندارد، به‌وضوح مهمل است. در مقابل، توجه کنید که تعمیم وجودی در مورد وقوع ارجاعی محض در (۵) نتیجه‌ی درست:

(x به خاطر ابعادش 'جورجونه' خوانده می‌شد) ($\exists x$),

را به دست می‌دهد، یعنی این که، 'چیزی به خاطر ابعادش 'جورجونه' خوانده می‌شد'.

عملیات منطقی تخصیص کلی[†] عملیاتی است که بهواسطه‌ی آن از، برای مثال، 'هر چیزی خودش است'، یا با نمادگذاری ($x = x$)، استنتاج می‌کنیم که سقراط = سقراط. این عملیات و تعمیم وجودی دو وجه از اصلی واحد اند؛ چرا که به جای گفتن این که ' $(x = x)$ ' استنتاج می‌کنیم که سقراط = سقراط است، می‌توانیم به درستی بگوییم که انکار [در] 'سقراط ≠ سقراط' مستلزم ' $(x \neq x)$ ' است. اصل متجلی در این دو عملیات پیوند میان تسویرها و احکام مفردی است که به عنوان نمونه‌ها به آن تسویرها مربوط اند. اما این تنها با اغماض یک اصل است. این اصل تنها در مواردی برقرار است که یک لفظ [چیزی را] می‌نامد و، علاوه بر این، به صورت ارجاعی واقع می‌شود. این اصل صرفاً محتوای منطقی این ایده است که وقوعی مفروض

^{*} Quantification

[†] Universal instantiation

ارجاعی است. بدین دلیل، این اصل همچون وصله‌ای ناجور نظریه‌ی منطقی محض تسویر است. اهمیت منطقی این واقعیت که تمامی الفاظ مفرد، به جز متغیرها که در نسبت با سورها چونان ضمایر به کار می‌روند، با بازنویسی اجتناب پذیر و قابل حذف اند، از اینجا می‌آید.^۶

الساعه دیدیم که سیاق ارجاعاً تیره (۲) چه وضعیتی در قبال تعمیم وجودی دارد. باید بینیم بر سر دیگر سیاق‌های ارجاعاً تیره چه می‌آید. کارست تعمیم وجودی بر وقوع نام شخصی در (۴) ما را به:

$$(26) \quad (\exists x) \text{'}x\text{' شامل شش حرف است},$$

می‌رساند، یعنی:

(27) چیزی وجود دارد که 'آن' شامل شش حرف است،

یا شاید:

(28) 'چیزی' شامل شش حرف است،

حال عبارت:

'آن' شامل شش حرف است

بهوضوح بدین معنا است:

(29) امین حرف الفبا [ای انگلیسی] شامل شش حرف است.

در (۲۶) وقوع حرف [x] در سیاق نقل به همان اندازه‌ی وقوع همین حرف در سیاق 'six' بی‌ارتباط با سور است. (۲۶) صرفاً شامل حکمی کاذب است که پیش از آن سوری بی‌ربط قرار گرفته است. (۲۷) نیز همین‌طور است؛ بخشی از آن:

'آن' شامل شش حرف است.

کاذب است، و پیش‌وند 'چیزی وجود دارد که' نامربوط است. (۲۸) نیز کاذب است—اگر از 'شامل شش بودن' مرادمان 'شامل دقیقاً شش بودن' باشد.

این که تعمیم وجودی به همین ترتیب در مورد (۹) و (۱۰) نیز بی‌ربط است کمتر واضح، و در نتیجه تشخیص آن مهم‌تر است. با اعمال آن بر (۹) به

(فیلیپ اطلاع ندارد که x کاتلین را محکوم کرده است) $(\exists x)$,

می‌رسید، یعنی

(29) چیزی چنان است که فیلیپ اطلاع ندارد که آن چیز کاتلین را محکوم کرده است.

این شیء چیست که کاتلین را محکوم کرده است بدون این که فیلیپ از این واقعیت مطلع شود؟ تولی، یعنی سیسرون؟ اما چنین فرضی با این واقعیت که (۱۱) کاذب است در تضاد است. توجه داشته باشیم که (۲۹) را با:

فیلیپ اطلاع ندارد که x کاتلین را محکوم کرده است) $(\exists x)$,

خلط نکیم، که گرچه از قضا کاذب است، کاملاً سرراست است و در خطر این نیست از طریق تعمیم وجودی از (۹) استنتاج شود. مشکل مطرح در ظاهراً منتج شدن (۲۹) از (۹) بار دیگر خود را در تلاش برای اعمال تعمیم وجودی بر احکام وجهی نشان می‌دهد. ظاهراً نتیجه‌ی (۱۵) و (۱۶) بودن:

(۳۰) $(\exists x) x \text{ ضرورتاً بزرگتر از } 7 \text{ است}$,

(۳۱) (ضرورتاً اگر در ستاره‌ی شامگاهی حیات وجود داشته باشد آن‌گاه در x حیات وجود دارد) $(\exists x)$,

همان پرسشی را پیش می‌کشد که (۲۹) پیش کشید. این عدد که، بنا بر (۳۰)، ضرورتاً بزرگ‌تر از ۷ است چیست؟ بنا بر (۱۵)، که (۳۰) از آن نتیجه شده است، [این عدد] ۹، یعنی تعداد سیارات، است؛ اما چنین فرضی با این واقعیت که (۱۸) کاذب است در تضاد است. در یک کلام، ضرورتاً بزرگ‌تر از ۷ بودن خصیصه‌ای از یک عدد نیست، بلکه وابسته به نحوه ارجاع به آن عدد است. بار دیگر، آن چیز x ‌ای که وجودش در (۲۱) تصدیق می‌شود چیست؟ بنا بر (۱۶)، که (۲۱) از آن نتیجه گرفته شده، آن چیز ستاره‌ی شامگاهی، یعنی ستاره‌ی بامدادی، است؛ اما چنین فرضی در تضاد با این واقعیت است که (۱۹) کاذب است. ضرورتاً یا امکاناً چنین و چنان بودن به‌طور کلی خصیصه‌ای از شیء مورد بحث نیست، بلکه وابسته به نحوه ارجاع به آن شیء است.

توجه کنید که (۳۰) و (۲۱) را با:

$$\text{ضرورتاً } (\exists x)(x > 7),$$

(اگر در ستاره‌ی شامگاهی حیات وجود داشته باشد آن‌گاه در x حیات وجود دارد) $\text{ضرورتاً } (\exists x)$

اشتباه نگیرید که به مشکلی در تعبیر از جنس آنچه در (۲۰) و (۲۱) مطرح شد نمی‌انجامند. این تفاوت را می‌توان با تغییری در مثال مؤکد ساخت: در یک بازی از نوعی که در آن مساوی امکان‌پذیر نیست، ضروری است که برخی بازیگران برنده خواهند شد، اما هیچ کدام از بازیگران نیست که بشود گفت ضرورتاً برنده خواهد شد.

در بخش قبل دیدیم که چگونه تیرگی ارجاعی در نسبت با الفاظ مفرد بروز می‌یابد، و وظیفه‌ای که ابتدای این بخش بر دوش خود گذاشتیم این بود که بینیم چگونه تیرگی ارجاعی در نسبت با متغیرهای سورها بروز می‌یابد. پاسخ اکنون روشن است: اگر سوری را بر سیاقی ارجاعاً تیره از یک متغیر اعمال کنیم، با این قصد که آن سور از بیرون از سیاق ارجاعاً تیره بر آن متغیر حاکم باشد، عموماً چیزی که در نهایت خواهیم داشت معنایی ناخواسته یا حکمی مهمل از نوع (۲۶)–(۲۱) است. در یک کلام، به‌طور کلی نمی‌توانیم به درستی سور بیندیم به درون سیاق‌های ارجاعاً تیره.

در بخش قبل، به‌واسطه ملاحظات مربوط به نقض کاریست جایگزین‌پذیری این‌همانی در مورد الفاظ مفرد، دیدیم که سیاق نقل قول و سیاق‌های دیگر ... چنین خوانده می‌شود، اطلاع ندارد که ...، باور دارد که ...، ضرورتاً ...، و ممکن است ... ارجاعاً تیره اند. در بخش حاضر، این سیاق‌ها با معیاری که با شکست تسویر سروکار دارد و، نه با الفاظ مفرد، ارجاعاً تیره تشخیص داده شدند. البته خواننده‌ای ممکن است احساس کند که ما در این معیار دوم واقعاً از الفاظ مفرد خلاص نشده‌ایم؛ چرا که رد کردن تسویرهای (۲۹)–(۲۱) کماکان منوط به رابطه‌ی متقابل میان نام‌های مفرد 'تولی' و 'سیسرون'، و 'تعداد سیارات'، 'ستاره‌ی شامگاهی' و 'ستاره‌ی بامدادی' در پس‌زمینه است. اما در واقع می‌توان از بازگشت به الفاظ مفرد آشنایمان در پس‌زمینه اجتناب کرد، چنان که اکنون با استدلال مجدد برای بی‌معنایی (۲۰) از راهی دیگر نشان خواهیم داد. هر آن‌چه از ۷ بزرگ‌تر است یک عدد است، و هر عدد x بزرگ‌تر از ۷ را می‌توان با بسیاری شروط مختلف به نحوی یکتا مشخص کرد، به طوری که $7 < x$ نتیجه‌ای ضروری از برخی دیگر خیر. یک عدد واحد x را می‌توان با شرط:

$$x = \sqrt{x} + \sqrt{x} \neq \sqrt{x} \quad (32)$$

و با شرط

$$\text{دقیقاً } x \text{ سیاره وجود دارد}, \quad (33)$$

به نحو یکتا معین کرد، اما $7 < x$ نتیجه‌ای ضروری از (۲۲) است، اما نتیجه ضروری‌ای از (۲۲) نیست. اطلاق ضرورتاً بزرگ‌تر بودن از ۷ به عدد x هیچ معنایی ندارد؛ ضرورت تنها قابل اعمال بر رابطه‌ی میان $7 < x$ و نحوه خاص مشخص کردن x در (۲۲)، در مقابل (۳۳)، است.

به همین ترتیب، (۲۱) بی‌معنا بود چرا که آن قسم چیزی، یعنی شیئی فیزیکی، را که شرط:

$$\text{اگر در ستاره‌ی شامگاهی حیات باشد آن‌گاه در } x \text{ حیات هست}, \quad (34)$$

را برآورده می‌کند، می‌توان با شرط‌های مختلفی معین کرد، که (۳۴) نتیجه‌ی ضروری همه‌ی آن‌ها نیست. اطلاق ضروری (۳۴) بر یک شیء فیزیکی x بی‌معنا است؛ در بهترین حالت، تنها بر رابطه‌ی میان (۳۴) و این یا آن شیوه مشخص کردن x می‌توان ضرورت قائل شد.

در باب اهمیت تشخیص تیرگی ارجاعی هر چه بگوییم کم گفته‌ایم. در ۱۸ دیدیم که تیرگی ارجاعی می‌تواند جلوی جایگزین‌پذیری این‌همانی را بگیرد. حال فهمیدیم که می‌تواند مانع تسویر شود: سورهای بیرون از یک ساخت ارجاعاً تیره نمی‌توانند تأثیری بر متغیرهای درون آن داشته باشند. این موضوع در مورد نقل قول نیز آشکار است، همان‌طور که این مثال مصحک نشان می‌دهد:

(۳۰) شامل 'x' است (six').

۳

در (۳۰)-(۳۱) دیدیم که چگونه اعمال سوری بر جمله‌ای وجهی به‌سادگی به بی‌معنایی در واقع صرف فقدان معنا است، و همواره می‌توان آن را با استفاده از خواه یک معنا علاج کرد. اما نکته مهمی که باید در نظر گرفت این است که با اتخاذ فهمی از وجهیات (براساس پذیرش غیرسنجهش‌گرایانه مفهوم پایه‌ای تحلیلیت به‌منظور پیش‌برد استدلال)، و با فرض فهمی از آنچه عموماً تسویر خوانده می‌شود، به‌طور خودکار معنایی برای جملات وجهی مسور مانند (۳۰)-(۳۱) نخواهیم داشت. هر کس که متکلف به دست دادن قوانینی برای منطق وجهی مسور می‌شود باید به آن توجه داشته باشد.

منشأ این مشکل تیرگی ارجاعی سیاق‌های وجهی است. اما تیرگی ارجاعی تا حدودی وابسته به وجودشناصی مفروض گرفته‌شده است، یعنی به این که کدام اشیاء به عنوان اشیاء مورد ارجاع پذیرفته می‌شوند. این نکته را خیلی ساده می‌توان با رجوع دوباره به منظر ۱ لذت‌گیر کرد، جایی که تیرگی ارجاعی براساس نقض جایگزین‌پذیری نام‌هایی که شیئی واحد را می‌نامند توضیح داده شد. حال فرض کنید که ما منکر تمامی اشیائی شویم که، همچون ۹ و تعداد سیارات، یا ستاره‌ی شامگاهی، قابل نامیدن با نام‌هایی هستند که در سیاق‌های وجهی جایگزین‌نپذیر اند. چنین کاری معادل خواهد بود با دور ریختن تمامی مثال‌هایی که نشانگر تیرگی سیاق‌های وجهی اند.

اما در جهانی که چنین پاکسازی شده است چه اشیائی باقی می‌ماند؟ شیء x برای این که باقی بماند باید این شرط را برآورده سازد: اگر S حکمی دربردارنده وقوع ارجاعی از نامی از x است، و 'S' از x با جایگزینی هر نام متفاوتی از x تشکیل شده است، آن‌گاه S و 'S' نه تنها باید در همین وضعیت‌شان از نظر ارزش صدق مشابه هم باشند، بلکه باید وقتی پیش از آن‌ها 'ضرورتاً'، و 'امکاناً' می‌آید نیز از نظر ارزش صدق مشابه باقی بمانند. معادلاً: نامی از x را به جای نامی دیگر از آن در یک حکم تحلیلی گذاشتن باید به حکمی تحلیلی بینجامد. معادلاً: هر دو نامی از x باید مترادف باشند.^۲

بنابراین، سیاره‌ی زهره چونان یک شیء مادی کنار گذاشته می‌شود؛ چرا که واحد نام‌های مختلف‌المعانی 'نووس'، 'ستاره‌ی شامگاهی'، و 'ستاره‌ی بامدادی' است. اگر قرار باشد که سیاق‌های وجهی ارجاعاً تیره نباشند، متناظر با این دو نام باید به جای یک شیء قائل به سه شیء باشیم—شاید زهره—مفهوم، ستاره‌ی شامگاهی—مفهوم، و ستاره‌ی بامدادی—مفهوم.

به همین ترتیب ۹، به مثابه‌ی یگانه عدد صحیح مابین ۸ و ۱۰، به‌واسطه واحد نام‌های مختلف‌المعانی ^۹ و 'تعداد سیاره‌های منظومه‌ی شمسی' کنار گذاشته می‌شود. اگر قرار باشد که سیاق‌های وجهی ارجاعاً تیره نباشند، متناظر با این دو نام باید به جای یک شیء قائل به دو شیء باشیم؛ شاید ^۹-مفهوم و تعداد سیارات-مفهوم. این مفاهیم عدد نیستند، چرا که یکی از آنها نه مساوی با دیگری است، نه کوچک‌تر و نه بزرگ‌تر از دیگری.

ممکن است این لازمه که هر دو نامی از x با یک‌دیگر مترادف باشند محدودیتی بر مجموعه‌ی الفاظ مفرد مجاز قلمداد شود و نه محدودیتی بر اشیاء x مجاز. در این صورت، مشکل در نحوه بیان این لازمه است؛ ما در اینجا صرفاً شاهد جلوه‌ای دیگر از بی‌مایگی مواجهه با پرسش‌های وجودشناصیک از منظر الفاظ مفرد هستیم. اما، بصیرت واقعی، که اینک در خطر تحت الشاعر قرار گرفتن است، این است: ضرورت را نمی‌توان به درستی برآورده شدن شرایطی توسط اشیاء (مانند گوی سنجینی که ونوس است، یا عددی که تعداد سیاره‌ها است)، مستقل از نحوه‌های خاصی از مشخص کردن آنها، اعمال کرد. این نکته را به ساده‌ترین شکل می‌توان با استفاده از ملاحظات مربوط به الفاظ مفرد نشان داد، اما با حذف آنها از اعتبار نمی‌افتد. اینک باید موضوع را از منظر تسویر به جای الفاظ مفرد بررسی کنیم.

از منظر تسویر، تیرگی ارجاعی سیاق‌های وجهی در بی‌معنایی تسویرهایی همچون (۳۰)-(۳۱) منعکس می‌شد. اصل مشکل (۳۰) این است که یک عدد x ممکن است به نحو یکتایی با دو شرط مختلف، همچون (۳۲) و (۳۳)، که ضرورتاً، یعنی به نحو تحلیلی، با یک‌دیگر معادل نیستند، معین شود. اما فرض کنیم ما اکنون منکر تمام چنین اشیائی بشویم و تنها اشیاء xx را نگه داریم که هر دو شرطی که به نحوی یکتا x را تعیین می‌کنند به نحو تحلیلی معادل باشند. در این صورت تمام مثال‌های مثل

(۳۰)–(۳۱)، که نمایان‌گر تیرگی ارجاعی سیاق‌های وجهی بودند، دور ریخته می‌شوند. در این صورت معنا خواهد داد که به طور کلی بگوییم که شیئی هست که، مستقل از هر نحوه خاصی از مشخص کردن آن، ضرورتاً چنین و چنان است. کوتاه‌سخن، مجموع خواهد بود که بر سیاق‌های وجهی سور بیندیم.

تا وقتی که مقادیر متغیرهایی که بر آنها سور بسته می‌شود محدود اند به اشیاء مفهومی^{*}، مثال‌های ما بر اعتراضی به سور بستن به درون سیاق‌های وجهی دلالت نمی‌کنند. این محدودیت به این معنا خواهد بود که، به هر حال برای چنین تسویرهایی، نه رده‌ها، بلکه رده-مفهوم‌ها یا صفت‌ها[†] را مجاز بشماریم، و این را چنین بفهمیم که دو جمله‌ی باز که رده‌ای واحد را معین می‌کنند، دو صفت مجزا را مشخص می‌کنند مگر آن که به لحاظ تحلیلی معادل باشند. این بدین معنا خواهد بود که، برای مقاصد چنین تسویری، نه اعداد، بلکه سخنی از مقاهیم را مجاز بدانیم که نسبت چند-به-یک با اعداد دارند. علاوه بر این، این بدین معنا خواهد بود که، برای مقاصد چنین تسویری، نه هیچ شیء انسجامی‌ای[‡]، بلکه تنها آن چیزی را مجاز بشماریم که فرگه [۲] مضمون‌های نام‌ها[§]، و کارنپ [۳] و چرچ مقاهیم فردی[¶] نامیده‌اند. اشکالی بر چنین وجودشناسی‌ای این است که اصل تفرد هویات آن همواره مبتنی بر مفهوم مفروض تراوُد، یا تحلیلی بودن است.

در واقع، حتی با فرض چنین هویات مشکوکی، به سرعت می‌توانیم بینیم که سودمندی تحدید مقادیر متغیرها به آن‌ها، نهایتاً یک اشتباه است. این کار مشکل اصلی راجع به تسویر بر سیاق‌های وجهی را برطرف نمی‌کند؛ برعکس، می‌توان مثال‌هایی کاملاً به همان اندازه‌ی مثال‌های قدیمی آزاردهنده در قلمرو اشیاء مفهومی اقامه کرد. چرا که، اگر A شیئی مفهومی، مثلاً یک صفت، باشد، و p ^۵ اشاره به یک جمله صادق دلخواه داشته باشد، بهوضوح:

$$A = (\exists x)[p \cdot (x = A)]. \quad (35)$$

با این حال، اگر جمله صادقی که p ^۶ بازنمایانده آن است تحلیلی نباشد، آن‌گاه (۳۵) نیز چنین نیست، و دو طرف آن در قبال جایگزین‌پذیری در سیاق‌های وجهی وضعیت بهتر از «ستاره‌ی شامگاهی» و «ستاره‌ی بامدادی» یا 9 و «تعداد سیاره‌ها» ندارند. یا، برای بیان این نکته بدون توصل به الفاظ مفرد، نکته این است که صرف فرض این که x شیئی مفهومی است، برآورده‌شدن شرط اخیر که ایتالیک شد—«هر دو شرطی که به نحوی یکتا x را تعیین می‌کنند به نحو تحلیلی معادل اند»—تضمين نمی‌شود. چرا که به Fx ^۷ به عنوان هر شرطی که به نحوی یکتا x را تعیین می‌کند فکر کنید. در این صورت، Fx . p ^۸ به نحوی یکتا x را تعیین می‌کند؛ اما به نحو تحلیلی معادل Fx ^۹ نیست، حتی اگر x شیئی مفهومی باشد.

من نخستین بار در مقاله‌ی ۱۹۴۳ خود به سور بستن به درون سیاق‌های وجهی اعتراض کردم، و چرچ در مرور آن مقاله بود که راه حل تحدید متغیرهایی که چنین سورهایی بر آنها بسته می‌شود به مقادیر مفهومی را مطرح کرد. این راه حل، که من اکنون آن را اشتباه دانستم، در آن زمان به نظر خوب می‌آمد. کارنپ [۳] آن را در شکلی افراطی اخذ کرد، و دامنه‌ی متغیرهایش را در سرتاسر سیستم‌اش به اشیاء مفهومی محدود کرد. او در واقع رویه‌اش را چنین توصیف نمی‌کرد؛ او تصویر را با پیش‌کشیدن تعبیر دوگانه عجیبی از متغیرها پیچیده می‌کند. اما من استدلال کرده‌ام^{۱۰} این ترفلده پیچیده‌کننده هیچ تأثیری ندارد و بهتر است کنار گذاشته شود.

در زمانی که چرچ منطق مفهومی خودش را مطرح کرد [۶]، احتمالاً متوجه بوده است که نهایتاً نمی‌توان به سور بستن به درون سیاق‌های وجهی با صرف محدود کردن متغیرهای تحت چنین سورهایی به مقادیر مفهومی مشروعیت بخشید. به هر حال کاری که او می‌کند خیلی افراطی تراست. او، به جای یک عملگر ضرورت که قابل الحق به جمله‌ها است، محمول ضرورتی دارد که به نام‌های مرکب برخی اشیاء مفهومی که گزاره‌ها خوانده می‌شوند قابل الحق است. آن‌چه این اقدام را جدی تراز آن آن‌چه به نظر می‌رسد می‌کند این است که ثابت‌ها و متغیرهایی که در یک جمله واقع می‌شوند در نام گزاره‌های متناظر چرچ ظاهر نمی‌شوند. بنابراین تعامل رایج در منطق وجهی میان وقوع‌های عبارت‌ها خارج از سیاق‌های وجهی و وقوع مجدد آن‌ها درون سیاق‌های وجهی، بدرسی در سیستم چرچ نشان داده نمی‌شود. شاید نباید آن را سیستمی از منطق وجهی بخوانیم؛ چرچ عموماً چنین نمی‌کرد. علی‌ای حال، بحث من در ادامه را چنان بفمیم که تنها به مربوط به منطق‌های وجهی به معنای محدودتر مربوط است، وقتی که عمل‌گر وجهی به جمله‌هایی الحق می‌شود.

^{*}Intensional objects

[†]Attributes

[‡]Concrete objects

[§]Senses of names

[¶]Individual concepts

چرچ^[۴] و کارنپ—چنان که استدلال کردم به نحو ناموفقی—تلاش کردند با تحدید مقادیر متغیرهایشان بر نقد من بر منطق وجهی مسور غلبه کنند. آرتور اسمولیان راه بدیل به چالش کشیدن خود نقد من را پیش گرفت. استدلال او مبتنی است بر فرض تفکیکی بنیادین میان نامها به نامهای خاص و وصفهای معین (آشکار یا پنهان)، به نحوی که نامهای خاصی که یک شیء را می‌نامند همواره مترادف اند. (مقایسه کنید با (۲۸) در ادامه). او، بر اساس چنین فرضی، کاملاً به درستی، معتقد است که هر مثالی که، همچون (۱۵)–(۲۰) و (۲۴)–(۲۵)، نشان دهنده نقض جایگزین‌پذیری این همانی در سیاقهای وجهی است، باید ناظر بر وصفهایی معین باشد و نه نامهای خاص. او آن‌گاه با طرح بدیلی از منطق آشنای راسل برای وصفهای معین، در رابطه با سیاقهای وجهی، به دنبال حل و فصل مسائل می‌رود.^۹ با این وجود، همان‌طور که در بخش قبلی تأکید شد، حتی وقتی وصفهای معین و دیگر الفاظ مفرد را کلاً حذف می‌کنیم، هنوز با تیرگی ارجاعی سروکار داریم.

مع هذا، تنها امید برای حفظ منطق وجهی مسور پیش‌گرفتن مسیری مشابه مسیر اسمولیان است، به جای مسیر چرچ^[۴] و کارنپ^[۲]، بدین نحو: باید بر نقد من فائق آمد. این تفوق مشتمل است بر استدلال برای یا تصمیم به این که سور بستن بر درون سیاقهای وجهی معنادار است، حتی اگر هر مقداری برای متغیر تحت چنین سوری را بتوان با شرط‌هایی که با یکدیگر به لحاظ تحلیلی معادل نیستند، معین کرد. تنها امیدی که هست این است که شرایط تصویر شده در (۳۲) و (۳۳) را پذیریم و مُصر باشیم که، علی‌رغم آن، شیء x ضرورتاً بزرگتر از ۷ است. این به معنای اتخاذ رویکرد طردی در مورد برخی شیوه‌های به نحو یکتا مشخص کردن x ، برای مثال (۲۳)، و جانب داری از شیوه‌هایی دیگر، برای مثال (۲۲)، به عنوان شیوه‌هایی که بهتر «ذات» آن شیء را مشخص می‌کنند، است. از چنین منظری، نتایج (۳۲) راجع به شئی که ۹ (و تعداد سیاره‌ها) است ضرورتاً صادق قلمداد می‌شوند، در عین حال که برخی نتایج (۳۲)، راجع به آن شیء امکاناً صادق در نظر گرفته می‌شوند. اگر اصرار بر سور بستن بر درون سیاقهای وجهی داشته باشیم، آشکارا این بازگشت به ذات‌گرایی ارسطوی (ن.ک. به ص. ۲۲) ناگزیر است. باید درباره‌ی یک شیء، فی‌نفسه و با هر نامی یا بدون نام، قائل شد که برخی صفت‌هایش را به نحوی ضروری دارد و برخی دیگر را به نحوی امکانی، علی‌رغم این واقعیت که صفت‌های اخیر از برخی شیوه‌های مشخص کردن شیء به نحوی تحلیلی نتیجه می‌شوند، درست به همان ترتیبی که صفت‌های اولی از برخی شیوه‌های مشخص کردن دیگر آن به نحوی تحلیلی نتیجه می‌شوند. در واقع، خیلی سرراست می‌توانیم بینیم که هر منطق وجهی مسوری ناگزیر است از ابراز چنین تفکیکی میان صفت‌های یک شیء؛ چرا که حکماً قائل خواهد بود که برای هر چیز x از یک سو

$$(x = x) \text{ ضرورتاً} \quad (36)$$

و از سوی دیگر

$$\sim [p . (x = x)], \quad (37)$$

که در آن ' p ' بر یک صدق امکانی دل‌خواه دلالت دارد. ذات‌گرایی در تضاد اکید با ایده‌ی توضیح ضرورت بر اساس تحلیلی بودن (ن.ک. ص. ۱۴۳) است، که کارنپ، لوئیس، و دیگران از آن دفاع می‌کنند. برای توسل به تحلیلی بودن تنها در نسبت با نحوی مشخص کردن شیء است که می‌توان وانمود به تمیز میان صفت‌های ضروری و امکانی یک شیء شد، نه به صورت مطلق. اما مدافعان منطق وجهی مسور باید ذات‌گرایی را پذیرید.

این که این شخص مقادیر متغیرهایش را محدود کند برای توجیه سور استن بر متغیری درون سیاق وجهی نه لازم است و نه کافی. با این حال، محدود کردن مقادیر متغیرهای توأم با ذات‌گرایی این شخص هنوز می‌تواند این اثر را داشته باشد: اگر می‌خواهد ذات‌گرایی‌اش را به گونه‌های خاصی از اشیاء محدود کند، باید متناظراً مقادیر متغیرهایی را که بر آنها درون سیاقهای وجهی سور می‌بندد محدود کند.

سیستم ارائه شده در مقالات پیش‌گامانه‌ی خانم بارکان درباره منطق وجهی مسور از این نظر که هیچ محدودیت خاصی را بر مقادیر متغیرها اعمال نمی‌کند با سیستم‌های کارنپ و چرچ تفاوت می‌کرد. علاوه بر این، در این قضیه او کم‌ویش به این که او مهیای پذیرش پیش‌فرضهای ذات‌گرایانه بود اشاره شده است:

$$(x)(y)\{(x = y) \supset \text{ضرورتاً}[(x = y)]\}, \quad (38)$$

چرا که این گویی بیان این است که حداقل (و در واقع حداقل، $Fx \cdot p$ را در نظر داشته باشید) برخی از صفت‌هایی که شیئی را معین می‌کنند این کار را به نحو ضروری می‌کنند. منطق وجهی [مطرح] در فیتچ^[۱] در هر دوی این نکته‌ها از خانم بارکان پیروی می‌کند. ضمناً توجه داشته باشید که (۳۸) مستقیماً از (۳۶) و قانون جایگزین‌پذیری این‌همانی برای متغیرها:

$$(x)(y)[(x = y \cdot Fx) \supset Fy].$$

نتیجه این تأملات قرار است این باشد که راه پرداختن به منطق وجهی مسور، اگر که راهی باشد، پذیرش ذات‌گرایی ارسطوی است. با وجود این، دفاع از ذات‌گرایی ارسطوی بخشی از برنامه‌ی من نیست. چنین فلسفه‌ای به همان اندازه که برای کارنپ و لوئیس نامعقول است برای من هم هست. و در نتیجه، علی‌رغم این که کارنپ و لوئیس نگفته‌اند، من می‌گویم: بدا به حال منطق وجهی مسور. و در پی آن، بدا به حال منطق وجهی غیرمسور؛ چرا که اگر قصد سور استن از خلال عمل‌گر ضرورت را ندادته باشیم، هیچ فایده‌ی روشنی در استفاده از این عمل‌گر به جای صرف نقل کردن یک جمله و گفتن این که آن جمله تحلیلی است باقی نماند.

۴

نگرانی‌های مطرح شده درباره‌ی وجهیات منطقی با پذیرش صفات (در مقابل رده‌ها) نیز مطرح می‌شوند. اصطلاح 'صفت' چنین و چنان بودن 'ارجاعاً تیره' است، چیزی که می‌توان، برای مثال، در این واقعیت متوجه‌اش شد که حکم صادق:

$$\text{صفت بیشتر از } ۹ \text{ بودن} = \text{صفت بیشتر از } ۹ \text{ بودن} \quad (۴۹)$$

با جایگزینی متناظر این‌همانی صادق (۲۴) به حکم کاذب:

$$\text{صفت بیشتر از } ۹ \text{ بودن} = \text{صفت بیشتر از تعداد سیاره‌ها بودن}$$

تبديل می‌شود. علاوه بر این، تعمیم وجودی (۳۹) به:

$$(\exists x) \text{صفت بیشتر از } ۹ \text{ بودن} = \text{بودن } x \text{ صفت بیشتر از } (۴۰)$$

می‌انجامد که تن به تفسیری سازگار نمی‌دهد، درست همان‌طور که تعمیم‌های وجودی (۲۹)–(۳۱) از (۹)، (۱۵)، و (۱۶) تن نمی‌دادند. سور استن بر جمله‌ای که شامل متغیر سور درون سیاقی به شکل 'صفت ...' است دقیقاً وضعیت مشابهی با سور استن بر جمله‌ای وجهی دارد.

همان‌طور که پیش‌تر اشاره شد، صفت‌ها با این اصل تفرد می‌یابند: دو جمله‌ی باز که رده‌ی واحدی را معین می‌کنند، صفتی واحد را معین نمی‌کنند مگر آن که به نحو تحلیلی معادل باشند. گزاره سخن محبوب دیگری از هویت‌های مفهومی است. نسبت گزاره‌های به حکم‌ها همان نسبت صفت‌ها به جمله‌های باز در نظر گرفته می‌شود: دو حکم گزاره‌ای واحد را معین می‌کنند تنها در شرایطی که بلحاظ تحلیلی معادل باشند. اشکال مذکور راجع به صفت‌ها آشکارا به همان ترتیب بر گزاره‌ها نیز وارد است. در نتیجه جایگزین‌پذیری بر اساس (۲۴)، صدق:

$$\text{این گزاره که } ۹ < ۷ = \text{این گزاره که } ۷ < ۹ \quad (۴۱)$$

تبديل می‌شود به کذب:

$$\text{این گزاره که } ۹ < ۷ = \text{این گزاره که تعداد سیاره‌ها } < ۷.$$

تعمیم وجودی بر (۴۱) به نتیجه‌ای مشابه (۲۹)–(۳۱) و (۴۰) می‌انجامد.

بیش‌تر منطق‌دانان، معناشناسان، و فیلسوفان تحلیلی که آسوده‌خاطر از صفت‌ها، گزاره‌ها، یا وجهیات منطقی سخن می‌گویند آشکار می‌کنند که درک نمی‌کنند که آن‌ها بر این اساس ملزم به موضع متأفیزیکی ای می‌شوند که خودشان نخواهد پذیرفت اش. قابل ذکر است که در پرنیکیپیا متمتیکا، که صفت‌ها به ظاهر به مثابه‌ی هویات پذیرفته شده‌اند، تمام سیاق‌های واقعی ای که در خلال کارهای صوری واقع می‌شوند چنان اند که به همان اندازه‌ی صفت‌ها با رده‌ها نیز درست در می‌آیند. تمام سیاق‌های موجود

به معنای صفحه‌ی ۳۰ مصداقی^{*} اند. نویسنده‌گان پرینکیپیا متمتیکا بنابراین در عمل بر اصل مصدقیتی وفادار می‌مانند که نظرآ از آن دفاع نمی‌کردند. اگر عمل کرد آنها به گونه‌ای دیگر می‌بود، شاید ما زودتر به درک ضرورت این اصل نائل می‌شدیم. دیده‌ایم که چطور جملات وجهی، صفت‌عبارت‌ها، و گزاره‌عبارت‌ها با دیدگاه ضد ذات‌گرایانه درباره جهان در تضاد اند. باید در نظر داشت که این عبارت‌ها تنها در صورتی به چنین تضادی می‌انجامند که به درون آنها سورسته شود، یعنی، وقتی ذیل یک سور قرار می‌گیرند و خودشان در بردارنده‌ی متغیر سور اند. ما این واقعیت را (که پیش از این در ۲۶ نشان داده شد) می‌دانیم که نقل قول نمی‌تواند شامل متغیری به‌نحوی مؤثر آزاد باشد، متغیری که سوری از بیرون به آن دسترسی داشته باشد. اگر رویکردی مشابه راجع به به وجهیات، صفت‌عبارت‌ها، و گزاره‌عبارت‌ها اخذ کنیم، می‌توانیم با خیال راحت از آن‌ها استفاده کنیم بی‌آن که نگرانی‌ای از سخن نگرانی‌های جدی حاضر داشته باشیم.

آن‌چه در این صفحات راجع به وجهیات گفته شد تنها به وجهیات اکید مربوط است. برای دیگر اقسام، همچون ضرورت و امکان فیزیکی، اولین مسأله صورت‌بندی روشن و دقیق این مفاهیم است. پس از آن می‌توانیم بررسی کنیم که آیا به درون چنین وجهیاتی، همچون وجهیات اکید، نمی‌توان بدون گرفتار شدن در بحرانی متافیزیکی سور است. مسئله اساساً مربوط می‌شود به کاربرد زبان در عمل. برای مثال، مربوط می‌شود به کاربرد شرطی‌های خلاف‌امر واقع درون یک سور؛ چرا که معقول است فرض این که شرطی‌های خلاف‌امر واقع در معنایی از ضرورت فروکاسته می‌شوند به شکل 'ضرورتاً اگر p آن‌گاه q '. تعریف، برای مثال، حل‌پذیری در آب نیز، به نوبه‌ی خود، بر شرطی‌های خلاف‌امر واقع استوار است: گفتن این که شیئی در آب حل‌پذیر است معادل است با گفتن این که در آب حل می‌شد اگر در آب می‌بود. در بحث‌های فیزیک ما، طبعاً، محتاج سور بستن‌هایی شامل عبارت^x در آب قابل حل است، یا لفظی معادل آن هستیم؛ اما، مطابق تعریف پیشنهادی، در این صورت باید درون تسویرها عبارت 'اگر x در آب می‌بود x حل می‌شد'، یعنی، 'ضرورتاً اگر x در آب باشد آن‌گاه x حل می‌شود' را مجاز بشماریم. اما نمی‌دانیم که آیا معنای مناسبی از 'ضرورت' هست که بتوانیم چنین به درون آن سور بیندیم.^{۱۰}

هر نحوه‌ای از گنجاندن احکام درون احکام دیگر را، چه مبتنی باشد بر مفهومی از «ضرورت»، و چه، برای مثال، بر مفهومی از «احتمال»، آن‌چنان که در رایشناخ چنین است، باید به دقت در رابطه با پذیرای سور بستن بودن آن بررسی کرد. شاید توابع صدق تنها اشکال مفید ترکیب احکام باشند که بی‌قید و شرط پذیرای سور استن اند. خوش‌بختانه، دست‌کم در ریاضیات، نیازی به اشکال دیگر ترکیب جملات نیست؛ و، جالب این که، ریاضیات شاخه‌ای از علم است که نیاز به آن را به آشکارترین شکلی می‌توان درک کرد.

به عنوان مشاهده‌ی کلی پایانی، بیایید به آزمون نخست‌مان برای تیرگی ارجاعی بازگردیم، یعنی نقض جایگزین‌پذیری این‌همانی؛ و بیایید فرض کنیم که با نظریه‌ای سروکار داریم که در آن (الف) فرمول‌های منطقاً معادل در تمام سیاق‌ها جایگزین‌پذیر حافظ‌الصدق اند و (ب) منطق رده‌ها در اختیار باشد.^{۱۱} در مورد چنین نظریه‌ای می‌توان نشان داد که هر شکل از ترکیب حکم، به جز توابع صدق، ارجاعاً تیره است. چرا که، فرض کنید ϕ و ψ احکامی باشند که در ارزش صدق مشابه اند، و فرض کنید $\Phi(\phi)$ حکم صادقی باشد که ϕ جزئی از آن است. آن‌چه باید نشان داد این است که $\Phi(\psi)$ نیز صادق خواهد بود، مگر آن که سیاق بازنمایی شده با $\Phi(\psi)$ ارجاعاً تیره باشد. حال ردی نامیده شده با $\hat{\alpha}\phi$ ، بسته به این که ϕ صادق است یا کاذب، یا V است یا Λ ؛ چرا که به یاد داشته باشید که ϕ حکمی است فاقد α آزاد. (اگر نمادگذاری $\hat{\alpha}\phi$ بدون وقوع مجدد α گیج کننده است، آن را چونان $\phi = \alpha$ بخوانید). علاوه بر این ϕ منطقاً معادل است با $V = \hat{\alpha}\phi$. بنابراین، بر اساس (الف)، از آن جا که $\Phi(\phi)$ صادق است، $\Phi(\hat{\alpha}\phi) = V$ نیز چنین است. اما $\hat{\alpha}\phi$ و $\hat{\alpha}\psi$ نامهایی از رده‌ای واحد اند، چرا که ϕ و ψ در ارزش صدق مشابه اند. فلذا، از آن جا که $\Phi(\hat{\alpha}\phi) = V$ صادق است، $\Phi(\hat{\alpha}\psi) = V$ نیز صادق است مگر آن که سیاق بازنمایی شده با $\Phi(\psi)$ ارجاعاً تیره باشد. اما اگر $\Phi(\hat{\alpha}\psi) = V$ صادق است، آن‌گاه، بر اساس (الف)، $\Phi(\psi)$ نیز چنین است.

یادداشت‌ها

^{۱۰} فرگه [۲] از وقوع‌های مستقیم و تیره سخن گفته، و درست مثل این جا از جایگزین‌پذیری این‌همانی به مثابه‌ی معیاری [برای تمایز آن‌ها] سخن استفاده کرده است.

^{۱۱} این لفظ کم‌ویش متضاد تعبیر 'شفاف' را سل است، آن‌گونه که در ضمیمه‌ی C از پرینکیپیا، ویراست دوم، ج. ۱ آن را به کار می‌برد.

^{۱۲} Lewis, [1], Ch. 5; Lewis and Langford, pp. 78-89, 120-166

^{۱۳} Cf. Carnap [2], pp. 245-259

^{۱۵} این نکته را اساساً چرج [۲] مذکور شده است.

عن. ک. به صفحات ۷ به بعد و صفحه‌ی ۱۳ در بالا، و صفحات ۱۶۶ به بعد در ادامه. توجه کنید که تعمیم وجودی همچون آنچه در ص. ۱۲۰ آمده به نظریه‌ی تسویر محض تعلق دارد؛ چرا که به جای الفاظ مفرد با متغیرهای آزاد سروکار دارد. همین امر راجع به کاربرد متناظر تخصیص کلی آنگونه که در R2 در مقاله‌ی ۷ ظاهر می‌شود صادق است.

۷ ص. ۳۲ در بالا را ببینید. ترادف نام‌ها تنها به معنای نامیدن چیزی واحد نیست؛ به این معنا است که حکم این‌همانی تشیکل شده از آن دو نام تحلیلی باشد.

^۸ در نقدی که کارنپ بلندنظرانه آن را در صفحات ۱۹۶ به بعد اثر [۳] خود گنجاند.

^۹ نظریه‌ی توصیفاتِ راسل، در صورت‌بندی اصلی اش، در بردارنده‌ی تمایزهای اصطلاحاً «دامنه» است. با این حال، تغییر در دامنه‌ی یک وصف برای ارزش صدق تمامی احکام بلاقتضا است، مگر آن که وصف از نامیدن قاصر باشد. این بلاقتضا بودن برای برآورده شدن غرض از آن بهمثابه‌ی تحلیلی یا جانشینی برای اصطلاح واقعی وصف مفرد، به واسطه‌ی نظریه‌ی راسل، مهم است. از سوی دیگر، اسمولیان روا می‌دارد که تفاوتِ دامنه، حتی در جایی که وصف مورد نظر موفق به نامیدن می‌شود، ارزش صدق را متأثر کند.

^{۱۰} برای نظریه‌ای راجع به الفاظ ناظر بر قابلیت، مانند 'حل‌پذیر'، به کارنپ [۵] نگاه کنید.
^{۱۱} صفحات ۲۷ و ۸۷ بالا را ببینید.

مترجم: ساجد طبیبی [†]

[†] هیئت علمی، پژوهشکده فلسفه تحلیلی

رایانامه: tayebi@ipm.ir



نسخه‌ی کوانتومی \mathcal{NP}

علی الماسی*

چکیده. این مقاله با هدف معرفی مفهوم اثبات‌های (غیرعاملی) کوانتومی نوشته شده است. این سیستم‌های اثبات ردهای از مسائل را مشخص می‌کنند که به آن، کلاس QMA گفته می‌شود. مطالعه‌ی QMA از دو جهت حائز اهمیت است؛ نخست آن که این کلاس همتای کوانتومی کلاس \mathcal{NP} است، و می‌توان معادل کوانتومی بسیاری از تاییجی که تاکنون در مورد \mathcal{NP} یا همتای تصادفی آن M ، یافت شده است را در چهارچوب محاسبات کوانتومی نیز جست‌وجو کرد. خواهیم دید برخی از سوالاتی که در مورد \mathcal{NP} یا M به سادگی پاسخ داده می‌شوند، درباره‌ی QMA می‌توانند بسیار دشوار باشند، و همین سبب می‌شود تلاش برای پاسخ‌دادن به آن‌ها به حصول درکی عمیق‌تر از محاسبات کوانتومی انجامد. وجه دیگر اهمیت مطالعه‌ی QMA ارتباط عمیق آن با مسائل فیزیک ماده‌ی چگال است. چه آن‌که یکی از مسائل کامل این کلاس، مسئله‌ی همیلتونی‌های موضوعی است که یافتن پاسخ تقریبی خوبی برای آن، مسئله‌ای مرکزی در فیزیک ماده‌ی چگال است. به همین دلیل است که بخشی از پیشرفت‌های فعلی نظریه‌ی پیچیدگی محاسبات کوانتومی بر یافتن روش‌هایی کارا برای پاسخ به این مسئله، یا پیدا کردن شواهدی برای سختی آن مرکز است. در این نوشته پس از معرفی مدلی برای محاسبات کوانتومی، سیستم‌های اثبات غیرعاملی کوانتومی را معرفی خواهیم کرد و به بررسی کلاس QMA از هر دو وجه فوق خواهیم پرداخت.

۱. مقدمه

محاسبات کوانتومی حوزه‌ای است که در نیمه‌ی دوم قرن بیستم، در پی پیدایش مکانیک کوانتومی و نیز به وجود آمدن نظریه‌ی مناسبی برای محاسبه‌پذیری، با انگیزه‌ی معرفی الگوریتم‌هایی کارا‌تر برای مطالعه‌ی سیستم‌های فیزیکی کوانتومی شکل گرفته است. از نظر تاریخی اولین پیشنهاد برای ساختن ماشین محاسبه‌ای که بر اساس فیزیک کوانتوم کار کند را می‌توان مربوط به پاول بنیوف دانست [۵۶]. با این وجود، معمولاً از ریچارد فاینمن به عنوان آغازکننده‌ی راه محاسبات کوانتومی یاد می‌شود. در حقیقت فاینمن در [۵۷]، با توجه به این که شبیه‌سازی برخی پدیده‌های فیزیکی کوانتومی بر روی کامپیوترهای کلاسیک غیرممکن به نظر می‌رسد، پیشنهاد داد از کامپیوترهایی که خود بر اساس فیزیک کوانتوم کار می‌کنند برای چنین شبیه‌سازی‌هایی استفاده شود.

بدون شک دعوت فاینمن، که فیزیکدان برجسته و شناخته‌شده‌ای در آن زمان بود، در جلب توجه فیزیکدانان به این مسئله تأثیر زیادی داشت. از جمله‌ی این افراد، دیوید دویچ بود که سه سال پس از مقاله‌ی فاینمن، مدل محاسبه‌ی ماشین تورینگ کوانتومی^۱ و در سال ۱۹۸۸ مدل محاسبات مداری کوانتومی^۲ را معرفی کرد. به این ترتیب، با داشتن مدل محاسبه‌ای که به طور دقیق تعریف شده باشد و بر اساس قوانین فیزیک کوانتوم کار کند، تلاش‌ها برای مطالعه‌ی بیشتر این دو مدل و یافتن الگوریتم‌هایی بر اساس آن‌ها آغاز شد. برای مثال، یائو در [۳۸] نشان داد که هر دو مدل قدرت محاسباتی یکسانی دارند. این نتیجه، از این نظر تأثیرگذار بود که پیاده‌سازی فیزیکی مدل ماشین تورینگ کوانتومی غیرممکن می‌نماید؛ حال آن‌که مدل مداری از نظر پیاده‌سازی عملی تا حدی امکان‌پذیر است، و این معادل بودن قدرت محاسباتی امیدبخش پیاده‌سازی عملی الگوریتم‌هایی کوانتومی است که پیشتر بر اساس مدل ماشین تورینگ کوانتومی تعریف شده بودند.

در سوی دیگر، یافتن الگوریتم‌هایی در این مدل محاسباتی جدید به عنوان راهی برای شناخت بهتر آن دنبال می‌شد. برنشتاین و وزیرانی با ارائه‌ی الگوریتمی در [۵۸]، نشان دادند اوراکلی وجود دارد که نسبت به آن، محاسبات کارای کوانتومی به طور اکید شامل محاسبات کارای تصادفی کلاسیک است. این نتیجه اولین نشانه را از این که مدل کوانتومی ممکن است به نقض تز

¹quantum Turing machine

²quantum circuit model

توسعه یافته‌ی چرچ-تورینگ منتج شود، نمایان کرد. سایمون با ارائه‌ی الگوریتمی در [۵۹] نشان داد که محاسبات کوانتومی کارا مشمول در محاسبات زیرنامایی^۱ تصادفی نیست، و گروور در [۶۰] ثابت کرد که مسأله‌ی جستجو را با الگوریتم‌های کوانتومی می‌توان به صورت کاراتری حل کرد. گرچه برنشتاین و وزیرانی در [۵۸] ثابت کرد که محاسبات کلاسیک و محاسبات کوانتومی از نظر قدرت محاسبه‌پذیری یکسانند، آن‌طور که از نتایج بالا برمی‌آمد، محاسبات کوانتومی در مواردی از نظر کارایی می‌تواند بهتر از همتای کلاسیک خود باشد. قوی‌ترین مؤید این مطلب الگوریتم‌های کارایی است که شور در [۶۱] برای حل مسأله‌های تجزیه‌ی اعداد و لگاریتم گسته ارائه کرده است. ارائه‌ی این الگوریتم‌ها توجه جامعه‌ی علمی را به قدرت و تأثیرات بالقوه‌ی محاسبات کوانتومی بر زمینه‌های متعددی از علوم کامپیوتر جلب کرد. بالاخص که با پیاده‌سازی الگوریتم شور، شکستن برخی سیستم‌های رایج رمزگاری همچون RSA، DH و ECC امکان‌پذیر می‌شد.

محاسبات کوانتومی از زمان ارائه‌ی الگوریتم‌های شور تا به امروز، در کمتر از چهل سال، رشد و پیشرفتی بسیار سریع داشته است. در هزاره‌ی جدید، با پیشرفت تکنولوژی قادر هستیم در عمل کامپیوترهای کوانتومی بسازیم و با آن‌ها محاسبه انجام دهیم [۶۲]. از سوی دیگر، امروزه به طور نظری بسیاری از حوزه‌های علوم کامپیوتر همتای کوانتومی دارند و نتایج امیدبخشی در این حوزه‌ها به درست آمده است. این نویدبخش آن است که در آینده‌ای نه چندان دور، می‌توان از محاسبات کوانتومی به طور گسترده‌ای بهره گرفت، و همین سبب شده است که توجه ویژه‌ای از سوی بسیاری از دولت‌ها و سرمایه‌گذاران بخش خصوصی به توسعه‌ی فناوری‌ها و علوم کوانتومی روانه شود [۶۳]. یک نتیجه‌ی توسعه‌ی محاسبات کوانتومی آن است که با پیدا شدن الگوریتم‌های کوانتومی جدید، تنظیم رده‌بندي جدیدی از مسائل از نظر کیفیت کارایی الگوریتم‌هایی که آن‌ها را حل می‌کنند، ضرورت می‌باشد. نظریه‌ی پیچیدگی محاسبات کوانتومی چهارچویی است که در آن، این برنامه را دنبال می‌کنیم.

در این مقاله تمرکز ما بر مطالعه‌ی سیستم‌های اثبات غیرتعاملی کوانتومی است. سیستم‌های اثبات در پیچیدگی کلاسیک به طور مشروحی مورد مطالعه قرار گرفته‌اند [۶۴، ۶۵، ۶۶]، و موارد متعددی از نتایج درخشنان پیچیدگی کلاسیک را می‌توان در رابطه با آن‌ها دانست. در چهارچوب محاسبات کوانتومی، مطالعه‌ی اثبات‌ها با کارهای نیل در [۶۷] و کیتائاف در [۱] آغاز می‌شود. مشابه کلاس \mathcal{NP} در پیچیدگی کلاسیک، می‌توان کلاسی از مجموعه‌ی ویژگی‌هایی مانند P تعریف کرد که تصدیق $x \in P$ با اثبات کوانتومی کوتاهی مانند π و با استفاده از الگوریتمی کوانتومی و کارا امکان‌پذیر است. مطالعه‌ی این کلاس، که همتای کوانتومی کلاس \mathcal{NP} است، موضوع اصلی این مقاله است.

در جریان بررسی این کلاس، خواهیم دید مسأله‌ی همیلتونی‌های موضعی، که تعیینی از مسأله‌ی SAT است، مسأله‌ای کامل برای آن است. مطالعه‌ی روش‌هایی برای حل این مسأله و پیچیدگی این روش‌ها، شاخه‌ای از محاسبات کوانتومی به نام پیچیدگی همیلتونی کوانتومی را تشکیل می‌دهد. این حوزه ارتباطی عمیق میان نظریه‌ی پیچیدگی محاسبه و نظریه‌ی سیستم‌های چندپیکره در فیزیک ماده‌ی چگال برقرار می‌کند. بالاخص، یکی از زمینه‌های فعل در این حوزه، تلاش برای یافتن همتای کوانتومی برای قضیه‌ی PCP کلاسیک است. قضیه‌ی PCP یکی از درخشنان ترین دستاوردهای نظریه‌ی پیچیدگی محاسبه است که بین نوع خاصی از سیستم‌های اثبات کارا — که به آن‌ها اثبات‌های قابل بررسی احتمالاتی می‌گویند — و سختی یافتن الگوریتم‌های تقریبی کارا برای دسته‌ای از مسائل \mathcal{NP} -سخت ارتباط برقرار می‌کند. معادل کوانتومی این قضیه، که به عنوان حدس PCP کوانتومی شناخته می‌شود، در صورت درستی، نتایجی خلاف شهود فیزیکی رایج درباره‌ی سیستم‌های کوانتومی دارد. در حال حاضر فیزیکدانان و متخصصین علوم کامپیوتر، هر یک به روش‌های خود، در تلاش برای یافتن نتایجی در تایید یا رد این حدس هستند، و این مسیر هم‌چنان ادامه دارد.

۲. مقدمه‌ای بر مکانیک کوانتومی برای علوم کامپیوتردانان

مکانیک کوانتومی، که در ادامه با آن بیشتر آشنا خواهیم شد، چهارچویی ریاضی است که قواعد ساختن نظریه‌های فیزیکی توصیف‌کننده‌ی پدیده‌های کوانتومی را تعیین می‌کند [۴۵]. پیش از پرداختن به مکانیک کوانتومی، خالی از لطف نیست که مروری بر مکانیک کلاسیک داشته باشیم و سپس مکانیک کوانتومی را در آنالوژی با همتای کلاسیک آن معرفی کنیم.

حکایت مشهور سیبی که بر سر نیوتن افتاد و الهام‌بخش او برای تدوین نظریه‌ی گرانش شد را در نظر بگیرید. سیبی که از درخت جدا شده و در حال افتادن بر زمین است، نمونه‌ای از یک سیستم فیزیکی است. برخی ویژگی‌های فیزیکی این سیب طی حرکتش به سمت زمین تغییر می‌کنند؛ مثلاً سرعت، ارتفاع آن از سطح زمین، انرژی جنبشی و پتانسیل آن. از سوی دیگر،

^۱subexponential

برخی ویرگی‌های فیزیکی سیب نیز در طول این حرکت، ثابت باقی می‌مانند؛ برای مثال جرم سیب از جمله‌ی این ویرگی‌هاست. به خواص فیزیکی نوع اول، خواص پویا، و به خواص نوع دوم، خواص ایستا می‌گوییم [۴۹].

به طور کلی، هدف مکانیک کلاسیک را می‌توان مطالعه‌ی خواص پویای سیستم‌های فیزیکی ماکروسکوپی که متشکل از اشیاء در حال حرکت هستند، دانست. برای نیل به این مقصود، روشی طبیعی مدل‌سازی ریاضی خواص پویا با سیستم‌های دینامیکی زمان-پیوسته است. با این مدل‌سازی بسیاری از مسائل فیزیکی را می‌توان به عنوان مسائلی در نظریه‌ی سیستم‌های دینامیکی صورت‌بندی کرد. به عنوان مثال، فرض کنید که در مثال سیبی که در درخت افتداده است، می‌خواهیم رابطه‌ی بین ارتفاع اولیه‌ی سیب و سرعت آن را در هنگام برخورد به زمین پیدا کنیم. ترجمه‌ی این پرسش فیزیکی به زبان سیستم‌های دینامیکی می‌تواند به این صورت باشد: «چنانچه حالت اولیه‌ی یک سیستم دینامیکی را بدانیم، آیا می‌توانیم حالت سیستم را در یک زمان خاص پیش‌بینی کنیم؟».

از نظر تاریخی، صورت‌بندی سیستم‌های فیزیکی به عنوان سیستم‌های دینامیکی به انحصار مختلفی انجام شده است و منجر به شکل‌گیری فرمول‌بندی‌های متفاوتی مانند فرمول‌بندی‌های نیوتونی، لاگرانژی و همیلتونی برای مکانیک کلاسیک شده است. در ادامه، خود را به فرمول‌بندی نیوتونی محدود خواهیم کرد و توضیح خواهیم داد که ترجمه‌ی یک سیستم فیزیکی متشکل از یک ذره‌ی در حال حرکت در راستای عمودی (سیب افتان) به زبان سیستم‌های دینامیکی به چه صورت انجام خواهد شد.

در مکانیک نیوتونی تنها دو خاصیت پویا، یعنی مکان و سرعت یک ذره، برای توصیف حالت سیستم در هر لحظه کافی هستند. حالت سیستم در لحظه‌ی t با زوج مرتب $(x(t), v(t))$ مشخص می‌شود، که $x(t)$ و $v(t)$ به ترتیب مکان و سرعت ذره را در زمان t مشخص می‌کنند. علاوه بر این، قانون انتقال سیستم، یا قاعده‌ای که حالت سیستم بر اساس آن در طول زمان تغییر می‌کند، با کمیت فیزیکی نیرویی که بر سیستم وارد می‌شود مشخص می‌شود، که بنابر قانون دوم نیوتون متناسب با مشتق دوم مکان ذره است. به عبارت دیگر، معادله‌ی دیفرانسیل

$$F = m \frac{d^2 x}{dt^2} \quad (1.2)$$

تحول زمانی سیستم را مشخص می‌کند، که در آن F و m به ترتیب نیروی کل وارد بر ذره و جرم آن هستند. سیستم دینامیکی فوق که برای یک ذره‌ی در حال حرکت تعریف شد، به سادگی قابل تعمیم برای سیستمی متشکل از چند ذره نیز است. بدین منظور کافی است فضای حالت را مجموعه‌ی همهٔ n آتایی‌های مرتب که بیانگر مکان و سرعت هر یک از ذرات هستند، در نظر بگیریم و معادله‌ی ۱.۲ را نیز به صورت برداری بازنویسی کنیم.

توجه کنید که اصول موضوعه‌ی مکانیک کوانتومی نیز، در روند مشابهی با آن چه درباره‌ی مکانیک کلاسیک گفتیم، نحوی نسبت دادن یک سیستم دینامیکی به سیستم‌های فیزیکی کوانتومی را مشخص می‌کند که در زیربخش بعد به تفصیل آن‌ها را بررسی خواهیم کرد.

۱.۲ اصول موضوعه‌ی مکانیک کوانتومی. مکانیک کوانتومی چهارچوبی ریاضی است که جهان فیزیکی، به طور خاص پدیده‌هایی فیزیکی که در سطح اتمی و زیرatomی رخ می‌دهند، را به نظریات ریاضی پیوند می‌دهد. از نظر تاریخی، پیدایش فیزیک کوانتوم را می‌توان مربوط به اولین سال‌های قرن بیست و ناکامی فیزیک کلاسیک در توضیح تعدادی از نتایج آزمایشگاهی آن زمان دانست. معرفی مفهوم بسته‌های انرژی توسط مکس پلانک [۵۰] که بعدها اینشیان آن را توسعه داد و اثر فوتوالکتریک را به کمک این مفهوم توضیح داد [۵۱]، معرفی مدل اتمی بور برای توصیف طیف اتم هیدروژن [۵۲]، توسعه‌ی مکانیک ماتریسی توسط هایزبرگ و توابع موج توسط شرودینگر برای توصیف ریاضی پدیده‌های کوانتومی و ارائه‌ی اصول موضوعه‌ی مکانیک کوانتومی توسط فون نویمان [۵۳] از جمله مهم‌ترین گام‌هایی است که در سه دهه‌ی اول قرن بیست برداشته و منجر به ساخته شدن این نظریه‌ی ارزشمند، و البته غامض، شده‌اند. نظریه‌ای که تاثیرات شگرفی بر زندگی بشر در عصر حاضر گذاشته و انتظار می‌رود که به زودی، بسیار بیشتر از امروز، وجود مختلف زندگی ما را متاثر کند.

در این زیربخش، بررسی خواهیم کرد که اصول موضوعه‌ی مکانیک کوانتومی چگونه فضای حالت و تحول زمانی سیستم‌های فیزیکی کوانتومی را فرمول‌بندی می‌کنند. هم چنین خواهیم دید که چگونه این فرمول‌بندی‌ها قابل تعمیم به سیستم‌هایی متشکل از زیرسیستم‌های کوچک‌تر است. علاوه بر این، در اصلی که مشابه آن در مکانیک کلاسیک وجود ندارد، خواهیم دید که اندازه‌گیری یک سیستم کوانتومی — یکی از مفاهیم مناقشه‌برانگیز فیزیک کوانتوم — چگونه صورت‌بندی می‌شود.

در ادامه‌ی این نوشه، همه‌ی فضاهای برداری روی میدان مختلط تعریف شده‌اند، مگر خلاف آن ذکر شود. ما برای نمایش بردارها از نمادگذاری خاصی موسوم به نمادگذاری دیراک استفاده می‌کنیم. در نمادگذاری دیراک، هر بردار مانند $v \in V \cong \mathbb{C}^n$ را که برداری ستونی و $1 \times n$ است، با $\langle v |$ و ترانهاده و مزدوج این بردار را با $|v \rangle$ نمایش می‌دهیم. به این ترتیب ضرب داخلی دو بردار v و w برابر با حاصل ضرب ماتریسی $\langle v |$ و $|w \rangle$ خواهد بود که به اختصار به صورت $\langle v | w \rangle$ نمایش داده می‌شود. همچنین در ادامه از مفهوم ضرب تنسوری به کرات استفاده خواهیم کرد. چنان‌چه با این ضرب آشنایی ندارید، می‌توانید این طور به آن فکر کنید:

ضرب تنسوری دو ماتریس $A_{m \times n}$ و $B_{p \times q}$ ، که آن را با $A \otimes B$ نمایش می‌دهیم، ماتریسی با ابعاد $(mp) \times (nq)$ است که به صورت زیر تعریف می‌شود:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}. \quad (2.2)$$

ضرب تنسوری دو فضای برداری را نیز می‌توان با گسترش خطی ضرب فوق روی فضاهای تعریف کرد. با این مقدمه، اکنون آماده‌ی ارائه‌ی اصول مکانیک کوانتمی هستیم.

اصل ۱.۲ (فضای حالت). به هر سیستم فیزیکی منزوی یک فضای هیلبرت^۱ نسبت داده می‌شود که به آن فضای حالت سیستم^۲ می‌گویند. بردار حالت^۳ سیستم (یا به طور خلاصه، حالت سیستم)، بردار یکمایی در فضای حالت آن است [۴۵].

تعريف ۲.۲. یک کیویت^۴، یک سیستم کوانتمی است که فضای حالت آن، فضای هیلبرت دو بعدی \mathbb{C}^2 است.

کیویت‌ها — همتای کوانتمی بیت‌های کلاسیک — اساسی‌ترین و ضروری‌ترین سیستم‌هایی هستند که در محاسبات و اطلاعات کوانتمی به کار گرفته می‌شوند. حالت یک کیویت می‌تواند به صورت

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (3.2)$$

نوشته شود که در آن، $\mathbb{C}, \alpha, \beta \in \mathbb{C}$ ، $|0\rangle$ و $|1\rangle$ بردارهای $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ و $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ را مشخص می‌کنند.

برخلاف بیت‌های کلاسیک، که تنها می‌توانند یکی از دو مقدار 0 یا 1 را داشته باشند، یک کیویت می‌تواند (مانند معادله‌ی ۳.۲) در یک برهمنهی^۵ از $|0\rangle$ و $|1\rangle$ قرار گیرد. این یکی از تفاوت‌های اساسی میان محاسبات کلاسیک و محاسبات کوانتمی است.

در ادبیات محاسبات کوانتمی، نام‌های خاصی برای برخی حالت‌های یک کیویت وجود دارد. در نمادگذاری بعد، دو مورد از این حالات را معرفی می‌کنیم.

نمادگذاری ۳.۲. حالتهای $(|1\rangle + |0\rangle)/\sqrt{2}$ و $(|1\rangle - |0\rangle)/\sqrt{2}$ به ترتیب با $|+\rangle$ و $|-\rangle$ نمایش داده می‌شوند.

توجه کنید که $\{|-\rangle, |+\rangle\}$ پایه‌ای برای \mathbb{C}^2 است که به آن پایه‌ی X می‌گویند. همچنین پایه‌ی $\{|0\rangle, |1\rangle\}$ پایه‌ی محاسباتی یا \mathbb{C}^2 پایه‌ی Z نامیده می‌شود.

اصل ۴.۲ (تحول سیستم). این اصل را می‌توان به دو صورت متفاوت بیان کرد، و البته می‌توان نشان داد که این دو صورت با یکدیگر معادلند [۴۵]:

^۱ یک فضای هیلبرت، فضایی برداری مجهز به یک ضرب داخلی است که نسبت به نرم القاشه توسعه آن ضرب داخلی کامل است، به این معنی که هر دنباله‌ی کوشی در آن همگراست. در این مقاله خود را به فضاهای هیلبرت متناهی‌البعد محدود می‌کنیم که می‌توان نشان داد با \mathbb{C}^n یک‌یاخت هستند.

²State Space

³State Vector

⁴Qubit

⁵Superposition

- حالت یک سیستم بسته‌ی کوانتومی مطابق با معادله‌ی شرودینگر تحول می‌باید. معادله‌ی شرودینگر به صورت زیر است:

$$i\hbar \frac{d|\psi(t)\rangle}{dt} = H |\psi(t)\rangle,$$

که در آن $(t)\psi$ حالت سیستم در لحظه‌ی t ، H عملگری هرمیتی که به آن همیلتونی^۱ سیستم می‌گویند، و \hbar ثابت پلانک است.

- اگر حالت یک سیستم بسته‌ی کوانتومی در لحظه‌ی t_1 ، $|\psi(t_1)\rangle$ باشد، حالت سیستم در لحظه‌ی $t_2 > t_1$ با

$$|\psi(t_2)\rangle = U |\psi(t_1)\rangle$$

مشخص می‌شود که U نگاشتی یکانی است که تنها به $t_2 - t_1$ وابسته است.

از این به بعد، اصطلاح «گیت کوانتومی» را برای اشاره به عملگرهای یکانی که تحول سیستم را مشخص می‌کنند، به کار خواهیم برد. با وجود این‌که تعداد گیتهای کوانتومی که قابل اعمال بر یک کیویت هستند نامتناهی است، به دلایل متعددی تنها تعدادی متناهی از این گیتهای مورد علاقه‌ی ما هستند. تعدادی از این گیتهای نمایش گرافیکی و ماتریسی آن‌ها در مثال بعد معرفی شده‌اند.

مثال ۵.۲. گیتهای زیر از پرکاربردترین گیتهای در مدارهای کوانتومی هستند.

$$(1) \text{ گیت Pauli-I} \text{ با نمایش ماتریسی } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ و نمایش گرافیکی } \boxed{I}.$$

$$(2) \text{ گیت Pauli-X} \text{ با نمایش ماتریسی } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ و نمایش گرافیکی } \boxed{X}.$$

$$(3) \text{ گیت Pauli-Z} \text{ با نمایش ماتریسی } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ و نمایش گرافیکی } \boxed{Z}.$$

$$(4) \text{ گیت Hadamard} \text{ با نمایش ماتریسی } \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \text{ و نمایش گرافیکی } \boxed{H}.$$

$$(5) \text{ گیت T} \text{ با نمایش ماتریسی } \begin{pmatrix} 0 & 1 \\ e^{\frac{i\pi}{4}} & 0 \end{pmatrix} \text{ و نمایش گرافیکی } \boxed{T}.$$

اصل ۶.۲ (سیستم‌های مرکب). فضای حالت یک سیستم مرکب که متشکل از n زیرسیستم با فضاهای حالت V_1, \dots, V_n است، برابر است با $V_1 \otimes \dots \otimes V_n$. همچنین اگر هر یک از زیرسیستم‌ها حالت $|v_i\rangle$ را داشته باشند، حالت سیستم مرکب برابر با $|v_1\rangle \otimes \dots \otimes |v_n\rangle$ خواهد بود [۴۵].

با توجه به اصول ۴.۲ و ۶.۲، تحول زمانی یک سیستم مرکب کوانتومی که متشکل از دو زیرسیستم با فضاهای حالت V و W است، با نگاشتهای یکانی روی فضای $V \otimes W$ مشخص می‌شود. توجه کنید که زیرمجموعه‌ای از چنین نگاشتهایی، به صورت $L \otimes L'$ هستند، که L و L' به ترتیب نگاشتهای یکانی روی فضاهای V و W هستند. با این وجود، باید توجه شود که این زیرمجموعه، زیرمجموعه‌ای سره از همه نگاشتهای یکانی روی $V \otimes W$ است.

بنابراین، نظری و عملی، در محاسبات کوانتومی بیشتر علاقه‌مند به گیتهایی هستیم که حداقل روی ۳ کیویت به طور نابدیهی عمل می‌کنند. یکی از مهم‌ترین این گیتهای، گیت CNOT است که روی دو کیویت عمل می‌کند. نمایش ماتریسی این

$$\text{گیت به صورت } \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ و نمایش گرافیکی آن به صورت } \boxed{\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \text{---} \text{---}} \text{ است.}$$

^۱Hamiltonian

اصل ۷.۲ (اندازه‌گیری). مقصود از یک اندازه‌گیری با m نتیجه‌ی ممکن روی یک سیستم کوانتومی، خانواده‌ای از عملگرها مانند $\{M_1, \dots, M_m\}$ است ($M = \{M_1, \dots, M_m\}$ متناظر با نتیجه‌ی i است). که روی فضای حالت آن سیستم عمل می‌کند و شرط $\sum_{i=1}^m M_i^\dagger M_i = \mathbb{I}_n$ را نیز برآورده می‌کند. هنگامی که این اندازه‌گیری روی سیستمی که در حالت $|\psi\rangle$ قرار دارد انجام می‌شود، نتیجه‌ی اندازه‌گیری با احتمال

$$p(i) = \langle \psi | M_i^\dagger M_i | \psi \rangle,$$

برابر با i خواهد بود؛ و در این صورت، حالت سیستم به حالت

$$\frac{M_i |\psi\rangle}{\sqrt{\langle \psi | M_i^\dagger M_i | \psi \rangle}}$$

فرو خواهد ریخت^۱. [۴۵]

توجه کنید که اندازه‌گیری راهی برای استخراج اطلاعات کلاسیک از یک سیستم کوانتومی است. با این حال، اصل فوق نشان می‌دهد که استخراج اطلاعات کلاسیک از یک سیستم کوانتومی اولاً ذاتی تصادفی و غیرقطعی دارد، و ثانیاً ضرورتاً منجر به تغییر حالت سیستم می‌شود، و این امری است که افتراقی اساسی میان فیزیک کلاسیک و فیزیک کوانتوم ایجاد می‌کند. در ادامه‌ی این نوشه، عموماً از حالت خاصی از اندازه‌گیری‌های معرفی شده در اصل ۷.۲ بهره خواهیم گرفت که در ادامه معرفی می‌شوند.

تعريف ۸.۲. یک اندازه‌گیری افکنشی یک اندازه‌گیری کوانتومی است که متشکل است از عملگرها افکنشی دو به دو متعامد. یک عملگر افکنشی عملگری هرمیتی مانند $P : \mathbb{C}^n \rightarrow \mathbb{C}^n$ است به طوری که $P = P^2$. به عبارت دیگر، یک اندازه‌گیری افکنشی خانواده‌ای مانند $M = \{P_1, \dots, P_m\}$ است به طوری که:

(۱) هر P_i یک عملگر افکنشی است.

$$\sum_{i=1}^m P_i = \mathbb{I}_n \quad (2)$$

$$\text{که } \delta_{ij} \text{ دلتای کرونکر است.} \quad (3)$$

همچین ممکن است در ادامه‌ی این مقاله، از اصطلاح اندازه‌گیری در پایه‌ی $\{v_1, \dots, v_n\}$ استفاده کنیم. در چنین مواردی، مقصودمان یک اندازه‌گیری افکنشی با عملگرها اندازه‌گیری $|v_i\rangle$ خواهد بود.

نمادگذاری ۹.۲. در ادامه از نماد زیر برای نمایش گرافیکی اندازه‌گیری استفاده خواهیم کرد.



۲.۰. قضیه‌ی عدم امکان شبیه‌سازی و درهم‌تیدگی. در این زیربخش، به دو مفهوم اساسی که نقشی کلیدی در علم اطلاعات کوانتومی دارند خواهیم پرداخت. مفاهیمی که ما را به دو مورد از اساسی‌ترین تفاوت‌های محاسبات کلاسیک و محاسبات کوانتومی رهنمون خواهند کرد.

اولین مفهوم، امکان ناپذیری شبیه‌سازی^۲ یک حالت دلخواه کوانتومی است که نخستین بار در [۵۴] و [۵۵] بیان شد. این قضیه بیان می‌کند که اگر یک نگاشت یکانی وجود داشته باشد که حالت مولفه‌ی اول یک سیستم مرکب دو مولفه‌ای کوانتومی را روی مولفه‌ی دوم کپی کند، در این صورت هر دو حالت قابل کپی‌کردن مولفه‌ی اول یا بر هم عمودند و یا باهم برابرند. به عبارت دیگر، با داشتن یک حالت کوانتومی نامعلوم، امکان کپی‌کردن آن بدون تغییر دادن حالتش وجود ندارد.

این قضیه نتایج متعددی در محاسبات و اطلاعات کوانتومی دارد. به عنوان مثالی از یک نتیجه‌ی منفی، توجه کنید که برخلاف روش‌های کاهش خطای مبتنی بر تکرار که در مخابرات و اطلاعات کلاسیک به طور گسترده استفاده می‌شوند، در محاسبات کوانتومی نمی‌توان از روی یک پیام کوانتومی دلخواه تعداد زیادی کپی درست کرد و از این طریق تأثیر نویز ایجاد شده در کانال مخابراتی را کاهش داد. به این ترتیب، راههای ممکن برای تصحیح خطای مخابره در ارتباطات کوانتومی بسیار محدودتر و توسعه‌ی این روش‌ها بسیار سخت‌تر و نیازمند خلاقیت بیشتر است.

¹Collapse

²No-cloning Theorem

مفهوم دوم، مفهوم ساده و در عین حال مهمی به نام درهم‌تییدگی^۱ است.

تعريف ۲. ۱۰.۲. به حالت یک سیستم مرکب متشکل از دو زیرسیستم n و m بعدی است. درهم‌تییده می‌گوییم، هرگاه هیچ دو برداری مانند $| \psi \rangle \in \mathbb{C}^n \otimes \mathbb{C}^m$ (که حالت یک سیستم مرکب متشکل از دو زیرسیستم n و m بعدی است). درهم‌تییده

$$| \psi \rangle = | \phi_1 \rangle \otimes | \phi_2 \rangle.$$

اگر یک حالت کوانتومی درهم‌تییده نباشد، به آن جداشدنی یا ضربی می‌گوییم.

مثال ۱۱.۲. حالت‌های زیر که به حالت‌های بل^۲ یا زوج‌های EPR^۳ مشهورند، نمونه‌ای از حالت‌های درهم‌تییده‌اند.

$$\begin{aligned} |\Phi^+\rangle &= \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle = \left(\frac{1}{\sqrt{2}} \quad 0 \quad 0 \quad \frac{1}{\sqrt{2}} \right)^t \\ |\Phi^-\rangle &= \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle = \left(\frac{1}{\sqrt{2}} \quad 0 \quad 0 \quad -\frac{1}{\sqrt{2}} \right)^t \end{aligned}$$

این بخش را با ذکر این نکته به پایان می‌رسانیم که درهم‌تییدگی کوانتومی، همان‌گونه که شرودینگر گفته است، «ویژگی باز مکانیک کوانتومی است؛ آن چیزی که ماهیت آن را تماماً از خطوط فکری کلاسیک جدا می‌کند [۶۸]». تا به امروز موارد متعددی از تاثیج درهم‌تییدگی کشف شده‌اند؛ و این مسیر هم‌چنان ادامه دارد. به طور ویژه، باور بر این است که درهم‌تییدگی کوانتومی منبعی ضروری برای الگوریتم‌های کوانتومی است تا بتوانند به تسريعی نمایی نسبت به الگوریتم‌های کلاسیک دست یابند [۶۹].

۳. مدل محاسبات مداری کوانتومی

در این فصل به معرفی الگوریتم‌های کوانتومی و کلاس همه‌ی مسائل قابل حل با الگوریتم‌های کوانتومی کارا خواهیم پرداخت. پیش از آن که به طور دقیق مقصودمان از یک الگوریتم کوانتومی را بیان کنیم، خالی از لطف نیست که توصیفی غیر دقیق، اما شهودبخش از یک الگوریتم کوانتومی داشته باشیم. این توضیحات، برگرفته از مرجع [۳۲] است.

یک الگوریتم را می‌توان یک سیستم دینامیکی با زمان گستته دانست که فضای فاز آن نیز گستته است. در واقع، فضای فاز چنین سیستم‌هایی عبارت است از مجموعه‌ای از رشته‌ها (در الفبایی دلخواه، که در ادامه برای راحتی فرض می‌کنیم مجموعه‌ی $\{0, 1\}$ است). که کدشده‌ی پیکربندی ماشین محاسبه در هر لحظه هستند. قانون انتقال حالت این سیستم دینامیکی، به این صورت است که در گذر هر لحظه، رشته‌ای که متناظر با حالت فعلی سیستم است را به طور موضعی تغییر داده و آن را به رشته‌ای دیگر، متناظر با حالتی دیگر در فضای فاز، تبدیل می‌کند. در ادامه برای سادگی بیشتر، فرض کنید که اعضای فضای فاز همگی رشته‌هایی به طول n هستند^۴. با چنین فرمالیسمی، محاسبه‌ی یک ورودی توسط یک ماشین محاسبه، در واقع معادل با یک مسیر^۵ در سیستم دینامیکی متناظر با آن است.

با داشتن این ایده در ذهن، انواع مختلف مدل‌های محاسبه را می‌توان به این صورت، معادل با انواع مختلفی از سیستم‌های دینامیکی دانست. برای مثال، یک مدل محاسباتی احتمالاتی، عملأً همان مدل فوق است؛ با این تفاوت که هر حالت سیستم متناظر با آن برابر است با یک بردار 2^n تایی توزیع احتمال روی 2^n عضو متمایزⁿ $\{0, 1\}$ ؛ یا معادلأً، ترکیب محدودی مانند $\sum_{x \in \{0, 1\}^n} p_x x$. قانون انتقال حالت سیستم نیز متشکل از اعمالی موضعی است که در طول زمان این بردار حالت‌ها را تغییر می‌دهند.

با این مقدمه، محاسبات کوانتومی را می‌توان با استفاده از تعبیر سیستم دینامیکی فوق مورد بررسی قرار داد. در حقیقت، حالت سیستم در هر لحظه برداری 2^n تایی مانند $(\alpha_x)_{x \in \{0, 1\}^n}$ است که هر درایه‌ی آن عددی مخلط است؛ و این بردار با نرم L_2 برداری یکه است. حالت سیستم با استفاده از اعمالی موضعی تغییر می‌کند که نگاشته‌هایی خطی و یکانی روی بردار حالت اعمال می‌کنند. نهایتاً خروجی الگوریتم با اندازه‌گیری حالت سیستم مشخص می‌شود. برای سادگی فرض کنید اندازه‌گیری ما در

¹Entanglement

²Bell States

³EPR Pairs

⁴این فرض چندان دور از ذهن نیست. به عنوان مثال، سیستم دینامیکی متناظر با یک مدار محاسبه روی n بیت، مثالی از چنین سیستمی است.

⁵trajectory

پایه‌ی محاسباتی است. در این صورت نتیجه‌ی اندازه‌گیری به صورت کاملاً تصادفی یکی از رشته‌های $\{x^0, x^1\}_{x \in \{0,1\}^n}$ خواهد بود که با توزیع احتمال $\alpha_x = |\alpha_x|^{(2)}$ مشخص می‌شود. به طور خلاصه، یک الگوریتم کوانتومی عبارت است از اعمال متناهی نگاشت موضعی نابدیهی یکانی بر بردار اولیه‌ای واقع در کرهٔ واحد فضای \mathbb{C}^n که در پایان الگوریتم، با استفاده از اندازه‌گیری، به یک بردار توزیع احتمال تبدیل می‌شود.

گرچه توضیحات نادقیق فوق، شهودی از کارکرد و ساختار یک الگوریتم کوانتومی در اختیار ما می‌گذارد، دور از انتظار نیست که در تعریف کردن یک «الگوریتم کوانتومی» به صورت دقیق، به همان اندازه که تعریف کردن دقیق مفهوم «الگوریتم» در حالت کلاسیک چالش‌برانگیز است، با مشکل مواجه شویم. در حقیقت، مدل‌های مختلف محاسبات کوانتومی، نظری مدل ماشین تورینگ کوانتومی یا مدل محاسبات مداری کوانتومی، تعاریف متفاوتی از الگوریتم‌های کوانتومی را در اختیار ما قرار می‌دهند. در این فصل، ما بر مدل محاسبات مداری کوانتومی تمرکز خواهیم کرد، و می‌توان نشان داد که با گذر از ماشین‌های تورینگ به مدل مداری، چیز زیادی را نیز از دست نخواهیم داد [۳۸].

۱.۳. الگوریتم‌های کوانتومی.

تعریف ۱.۳. یک گیت کوانتومی k موضعی روی یک رجیستر n -کیویتی، نگاشتی یکانی است که به طور نابدیهی روی k کیویت از رجیستر عمل می‌کند؛ و عمل آن روی باقی کیویت‌ها نگاشت همانی است.

فرض کنید $U \in \mathcal{L}((\mathbb{C}^2)^{\otimes k})$ نگاشتی یکانی و $i_1, i_2, \dots, i_k \in [n]^k$ یک k تایی با درایه‌های متمایز باشد. در این صورت یک گیت کوانتومی k موضعی که نگاشت U را بر کیویت‌های i_1, i_2, \dots, i_k از یک رجیستر n کیویتی اعمال کرده و اثر آن بر باقی کیویت‌ها همانی است، نگاشتی مانند $U_{(i_1, \dots, i_k)}$ است که به صورت زیر تعریف می‌شود:

• اگر $k = 1$

$$U_{(i_1)} = I^{\otimes(i_1-1)} \otimes U \otimes I^{\otimes(n-i_1)}$$

• اگر $k > 1$: می‌دانیم که می‌توان نوشت:

$$U = \sum_j U^{1,j} \otimes \dots \otimes U^{k,j},$$

که هر $U^{i,j}$ نگاشتی یکانی روی \mathbb{C}^2 است. در این حالت:

$$U_{(i_1, \dots, i_k)} = \sum_j U_{(i_1)}^{1,j} \dots U_{(i_k)}^{k,j}.$$

در ادامه چنان‌چه از زمینه‌ی بحث روشن باشد که گیت‌های موضعی بر چه کیویت‌هایی به صورت نابدیهی عمل می‌کنند، از نوشتن پانویس (i_1, \dots, i_k) برای $U_{(i_1, \dots, i_k)}$ اجتناب خواهیم کرد.

تعریف ۲.۳. فرض کنید B مجموعه‌ای ثابت از نگاشت‌های یکانی باشد. یک مدار کوانتومی روی n کیویت، نگاشتی مانند $U \in \mathcal{L}((\mathbb{C}^2)^{\otimes n})$ است که به صورت زیر تعریف شده است:

$$U = U_{\alpha_1}^1 U_{\alpha_2}^2 \dots U_{\alpha_s}^s,$$

که در آن $U_{\alpha_i}^i$ ها گیت‌های کوانتومی k_i موضعی روی n کیویت هستند که از روی نگاشت‌های B ساخته شده‌اند، و $\alpha_i \in [n]^{k_i}$. مجموعه‌ی B یک پایه برای مدار U نامیده می‌شود. همچنین به عدد s اندازه‌ی مدار U گوییم.

تعریف کردن مفهوم الگوریتم، هدفی است که در قلب نظریه‌ی محاسبه قرار دارد و نیل به آن، نیازمند انتخاب مدل مناسبی برای محاسبه است. مدل محاسباتی رایج در ادبیات فعلی نظریه‌ی محاسبات کوانتومی، مدل محاسبات مداری است؛ گرچه از نظر تاریخی ماشین‌های تورینگ کوانتومی اولین مدلی هستند که برای مطالعه‌ی مفاهیم محاسبات کوانتومی مورد استفاده قرار گرفته‌اند [۳۹]. ما تا به این‌جا مفهوم مدار کوانتومی را به طور دقیقی تعریف کردیم. در ادامه، مختصراً سه سناریوی مختلف برای تعریف مفهوم الگوریتم کوانتومی را معرفی خواهیم کرد و خواهیم دید که هر یک از این سناریوها، ما را به منابع محاسباتی مختلفی رهنمون خواهند کرد که هر یک می‌توانند مبنای ساختن نظریه‌ای برای پیچیدگی محاسبات کوانتومی قرار گیرند.

ملاحظه ۳.۳. الگوریتم‌های کوانتومی را می‌توان به طرق مختلفی تعریف کرد. در ادامه، مطابق با مرجع [۴۰] به معرفی سه مورد از این روش‌ها خواهیم پرداخت. شایان ذکر است که هر یک از تعاریف زیر مزایای خاص خود را دارند؛ و گرچه هر یک از آن‌ها با دیگری متفاوت است، اما ارتباطاتی نیز میان آن‌ها وجود دارد که به طور مفصلی در ادبیات پیچیدگی محاسبه مورد مطالعه قرار گرفته است.

(۱) سناریوی اول: پیچیدگی محاسباتی کوانتومی^۱

با فرض این‌که تابعی جزئی مانند $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$: f داده شده باشد، یک الگوریتم که این تابع را محاسبه می‌کند عبارت است از یک مدار کوانتومی که برای هر $x \in \{0, 1\}^n$ بر حالت $|x\rangle$ اعمال می‌شود؛ و پس از آن کیویت مشخص اندازه‌گیری می‌شود تا حالتی مانند $|f(x)\rangle$ به دست آید. در این الگوریتم‌ها، منبع محاسباتی مدنظر ما برای اندازه‌گیری پیچیدگی محاسبه، تعداد گیت‌های تشکیل‌دهنده‌ی مدار هستند.

(۲) سناریوی دوم: پیچیدگی کوئری کوانتومی^۲

فرض کنید به عنوان ورودی مسئله جعبه‌سیاهی به ما داده شده است که تابعی مانند $f : \{0, 1\}^m \rightarrow \{0, 1\}^n$: f را پیاده‌سازی می‌کند، و از ما خواسته شده است که اطلاعاتی درباره‌ی این تابع را با کوئری کردن از این جعبه‌سیاه (یا اوراکل) به دست آوریم. برای پرسیدن کوئری از اوراکلی که تابع f را پیاده‌سازی می‌کند، از نگاشت‌هایی یکانی موسوم به f -گیت بهره می‌گیریم. به این ترتیب، یک الگوریتم که چنین مسئله‌ای را حل می‌کند عبارت است از یک مدار کوانتومی که از گیت‌های استاندارد کوانتومی و f -گیت‌ها تشکیل شده است؛ و بر تعداد مناسبی کیویت ورودی اعمال می‌شود (معمولًاً لازم است رجیستری که ورودی تابع f را در خود نگه می‌دارد را به یک رجیستر کمکی الحقن کنیم)، و پس از آن، تعدادی کیویت مشخص اندازه‌گیری می‌شوند و براساس تایخ اندازه‌گیری، اطلاعات مورد نظر درباره‌ی تابع f به دست می‌آید. در این الگوریتم‌ها منبع محاسباتی مورد نظر جهت اندازه‌گیری پیچیدگی محاسباتی، تعداد کوئری‌ها (یا تعداد f -گیت‌های استفاده شده در مدار) است.

(۳) سناریوی سوم: پیچیدگی ارتباطی کوانتومی^۳

فرض کنید آليس و باب دو رجیستر کوانتومی $|x\rangle$ و $|y\rangle$ در اختیار دارند، که $x, y \in \{0, 1\}^n$ ، و از آن‌ها خواسته شده است تا مقدار تابعی مانند $f(x, y)$ را محاسبه کنند. یک الگوریتم برای حل این مسئله، که در این سناریو به آن پروتکل نیز گفته می‌شود، عبارت است از دو مدار کوانتومی، که هر یک در اختیار یکی از آليس و باب است، و بر کیویت‌هایی که در اختیار هر یک از آن‌هاست اعمال می‌شود. این کیویت‌ها می‌توانند بین آليس و باب انتقال یابند (به این معنی که هر یک برای دیگری کیویت‌هایی بفرستد) و نهایتاً خروجی با اندازه‌گیری کیویت‌های مشخصی از رجیستری که در اختیار یکی از آن‌هاست (مثالاً باب)، تعیین می‌شود. در این سناریو، منبع محاسباتی مورد نظر ما تعداد کیویت‌های انتقال یافته میان طرفین است.

هدف این نوشته مطالعه‌ی کلاس‌های پیچیدگی محاسباتی کوانتومی مورد مطالعه قرار می‌گیرند.

۲.۳. گیت‌های جهانی کوانتومی. از محاسبات کلاسیک می‌دانیم که مجموعه‌هایی متناهی از گیت‌های کلاسیک وجود دارند که جهانی هستند؛ به این معنا که هر تابع بولی را می‌توان با مدارهایی فقط شامل گیت‌هایی از این مجموعه‌ها پیاده‌سازی کرد. مثلاً می‌توان نشان داد که برای محاسبات برگشت‌پذیر، مجموعه‌ی $\{\theta_2\}$ که در آن θ_2 گیت توفولی است، یکی از این مجموعه‌ها است، و نیز می‌توان نشان داد که هیچ مجموعه‌ای از گیت‌های ۱ و ۲ بیتی وجود ندارد که مجموعه‌ای جهانی برای محاسبات برگشت‌پذیر کلاسیک باشد. در این بخش به طور اجمالی وجود چنین مجموعه‌های جهانی‌ای از گیت‌ها را برای محاسبات کوانتومی مورد بررسی قرار می‌دهیم.

اولین مسئله‌ای که باید به آن توجه کرد این است که تعداد نامتناهی ناشمارایی گیت کوانتومی متمایز وجود دارد؛ در نتیجه هیچ مجموعه‌ی متناهی از گیت‌های کوانتومی نمی‌تواند به طور دقیق جهانی باشد. از سوی دیگر، بنا به دلایل نظری و عملی متعددی از جمله ممکن نبودن پیاده‌سازی فیزیکی هر نگاشت یکانی دلخواه، به صورت آزمایشگاهی تنها پیاده‌سازی تعدادی

¹ quantum computational complexity

² quantum query complexity

³ quantum communication complexity

متناهی از گیت‌های کوانتومی برای ما مقدور است. این دلایل، ما را به این رهنمون می‌کنند که مفهوم جهانی بودن را برای گیت‌های کوانتومی به دو صورت متفاوت تعریف کنیم: یکی جهانی بودن به صورت دقیق، و دیگری جهانی بودن به صورت تقریبی.

تعریف ۴.۳. مجموعه‌ای از گیت‌های کوانتومی مانند \mathcal{G} به طور دقیق جهانی است هرگاه برای هر گیت کوانتومی مانند U ، دنباله‌ای متناهی از گیت‌ها مانند $\mathcal{G} = g_1, g_2, \dots, g_n$ وجود داشته باشند به طوری که $U = g_1 g_2 \dots g_n$.

توجه کنید که سمت راست تساوی فوق مختصر نوشته شده است و باید چنین تعبیر شود: ممکن است هر یک از گیت‌های g_i تنها روی تعدادی از کیویت‌هایی که U بر آن‌ها اثر می‌کند (و نه همه آن‌ها) به صورت نابدیهی عمل کنند و اثراشان روی باقی کیویت‌ها نگاشت همانی باشد. بنابراین، تساوی فوق به این معنا نیست که بعد فضایی که U و گیت‌های g_i روی آن تعریف شده‌اند، یکسان است.

با یک استدلال ساده‌ی شمارشی می‌توان نشان داد که هیچ مجموعه‌ی متناهی ای از گیت‌ها وجود ندارد که به طور دقیق جهانی باشد. با این وجود، مجموعه‌هایی نامتناهی از گیت‌ها که به طور دقیق جهانی باشند وجود دارند. یکی از چنین مجموعه‌هایی، که شاید مشهورترین آن‌ها باشد، توسط بارنکو و همکاران در [۴۱] معرفی شده است. آن‌ها نشان دادند که مجموعه‌ی همه‌ی گیت‌های کوانتومی ۱ کیویتی به همراه گیت ۲ کیویتی CNOT، مجموعه‌ای به طور دقیق جهانی از گیت‌های کوانتومی است. برای تعریف کردن مفهوم جهانی بودن تقریبی، نخست به این نیازمندیم که به طور دقیق مشخص کنیم که منظور ما از «تقریب‌زدن» چیست. برای آن‌که بتوانیم گیتی را تقریب بزنیم، ضروری است که مفهومی از فاصله را روی نگاشت‌های یکانی تعریف کنیم. تعریف زیر، دسته‌ای از کاندیدهای مناسب برای این منظور را به ما پیشنهاد می‌دهد.

تعریف ۵.۳. p -نرم شاتن یک عملگر $T \in \mathcal{L}(\mathbb{C}^d)$ ، که در آن $(1, \infty)$ ، به صورت زیر:

$$\|T\|_p = \left(\text{tr}((T^\dagger T)^{\frac{p}{2}}) \right)^{\frac{1}{p}},$$

و برای $p = \infty$ نیز به شکل زیر:

$$\|T\|_\infty = \lim_{p \rightarrow \infty} \|T\|_p = \sup_{|\psi\rangle : \langle \psi | \psi \rangle = 1} \|T|\psi\rangle\|.$$

تعریف می‌شود. به ۱-نرم و ∞ -نرم شاتن به ترتیب نرم اثر^۱ و نرم طیفی^۲ گفته می‌شود.

حال به تعریف مفهوم جهانی بودن به طور تقریبی می‌پردازیم.

تعریف ۶.۳. مجموعه‌ای متناهی از گیت‌های کوانتومی مانند \mathcal{G} به طور تقریبی جهانی است هرگاه برای هر گیت کوانتومی مانند U و هر $\varepsilon > 0$ ، دنباله‌ای متناهی از گیت‌های $\mathcal{G} = g_1, g_2, \dots, g_n$ وجود داشته باشد به طوری که

$$\|U - g_1 g_2 \dots g_n\|_1 < \varepsilon.$$

مجموعه‌های متنوعی از گیت‌های به طور تقریبی جهانی وجود دارد که در مثال بعد، تعدادی از آن‌ها را معرفی می‌کنیم.

مثال ۷.۳. هر یک از مجموعه‌های زیر از گیت‌های کوانتومی، به طور تقریبی جهانی هستند:

- گیت دویچ [۴۲]
- گیت بارنکو [۴۴]
- [۴۵] $\{H, T, \text{CNOT}\}$
- تقریباً هر گیت کوانتومی که روی حداقل ۲ کیویت اثر می‌کند [۴۶]. (به این معنی که گیت‌هایی که جهانی نیستند، مجموعه‌ای اندازه صفر را مشخص می‌کنند).

¹trace norm

²spectral norm

در خاتمه‌ی این بخش، به پرسش مهم دیگری می‌پردازیم که پاسخ آن در ملاحظات پیچیدگی محاسباتی ما تاثیرگذار است. تا به اینجا دیدیم که مجموعه‌ی همه‌ی گیت‌های ۱ کوییتی به همراه گیت CNOT مجموعه‌ای به طور دقیق جهانی است. با این حال، سوال این جاست که «برای ساختن یک نگاشت یکانی دلخواه با استفاده از اعضای این مجموعه، به چند گیت نیاز است؟». می‌توان نشان داد که نگاشت‌هایی یکانی روی n کوییت وجود دارند که برای ساختن آن‌ها با استفاده از اعضای این مجموعه، به $(n^{24^n})^{\theta}$ گیت نیاز است [۴۵]. با این حال قضیه‌ای زیبا موسوم به قضیه‌ی سولووی-کیتائوف، به ما این تضمین را می‌دهد که برای هر دو مجموعه از گیت‌های به طور تقریبی جهانی، می‌توان یکی را با دیگری به صورت کارایی تقریب زد. همان‌گونه که در بخش بعد خواهیم دید، چنین نتیجه‌ای برای ساختن یک نظریه‌ی پیچیدگی مناسب برای محاسبات کوانتومی، اهمیت زیادی دارد.

قضیه ۸.۳ (قضیه‌ی سولووی-کیتائوف). فرض کنید \mathcal{G} مجموعه‌ای متناهی از گیت‌های کوانتومی ۱ کوییتی است که شامل وارون اعضاش نیز هست و گروهی که توسط اعضای \mathcal{G} تولید می‌شود در $SU(2)$ با نرم اثر چگال است. در این صورت برای هر $\epsilon > 0$ ، ثابت c موجود است چنان‌که برای هر $U \in SU(2)$ ، دنباله‌ای از اعضای \mathcal{G} مانند $g_1, g_2, \dots, g_n \in \mathcal{G}$ وجود دارد به طوری که $\|U - g_1 g_2 \dots g_n\|_1 \leq O(\log^c(\frac{1}{\epsilon}))$.

فرض کنید مداری کوانتومی داریم که شامل m گیت کوانتومی ۱ کوییتی است؛ و می‌خواهیم آن را به مداری که گیت‌هایش از یک مجموعه از گیت‌های ۱ کوییتی به طور تقریبی جهانی (برای گیت‌های ۱ کوییتی) می‌آید، تبدیل کنیم؛ به طوری که مدار دوم با دقت ϵ مدار نخست را تقریب بزند. لم زیر، که نتیجه‌ی مستقیم یکانی-ناوردا بودن نرم اثر است؛ نشان می‌دهد که بدین‌منظور کافی است هر گیت مدار اول را با دقت $\frac{\epsilon}{m}$ تقریب بزیم.

لم ۹.۳. فرض کنید $U_1 \dots U_m = U_m U_{m-1} \dots U_1 = V_m V_{m-1} \dots V_1$ دو مدار کوانتومی باشند به طوری که برای هر $i \leq m$ $\|U_i - V_i\|_1 < \epsilon$. در این صورت ϵ -همسایگی U قرار دارد، و افزون بر این اندازه‌ی مدار اخیر $O(m \log^c(\frac{m}{\epsilon}))$ است.

لم ۹.۳، همراه با قضیه‌ی ۸.۳ نتیجه می‌دهد که اگر مداری کوانتومی مانند U داشته باشیم که از m گیت ۱ کوییتی کوانتومی ساخته شده است، می‌توان آن را به مداری مانند U' تبدیل کرد؛ چنان‌که مدار اخیر تنها از گیت‌های جهانی ساخته شده است؛ U' در ϵ -همسایگی U قرار دارد، و افزون بر این اندازه‌ی مدار اخیر $O(m \log^c(\frac{m}{\epsilon}))$ است.

۳.۳ محاسبات کوانتومی کارا. در نظریه‌ی محاسبه‌ی کلاسیک، محاسبات کارا معمولاً به محاسباتی با زمان چندجمله‌ای تعبیر می‌شود؛ انتخابی که پیشنهاد آن را می‌توان مربوط به کارهای کابام دردهه‌ی ۶۰ دانست [۴۸]. گرچه انتخاب چندجمله‌ای‌ها برای این منظور تا حدی دلخواه به نظر می‌رسد، این انتخاب در طول سالیان از نقطه‌ی نظرهای مختلفی تایید شده است^۱؛ تا این حد که باور عمومی براین است که محاسبات کارایی که اساساً توسط بشر و با محدودیت‌های طبیعت و قوانین فیزیک قابل انجام است، محاسبات چندجمله‌ای است. این باور را در نسخه‌ی تعمیم‌یافته‌ی تز چرج-تورینگ می‌توان دید: «هر چیز که به صورت کارایی محاسبه‌پذیر باشد، با یک ماشین تورینگ احتمالاتی در زمان چندجمله‌ای قابل محاسبه است» [۳۲].

توجه کنید که در محاسبات کلاسیک مدل تورینگ مدل محاسبه‌ی مرچح است، حال آن‌که در محاسبات کوانتومی بنا به دلایل متعددی (از جمله این‌که پیاده‌سازی فیزیکی ماشین‌های تورینگ کوانتومی با توجه به محدودیت‌های فعلی مهندسی سیستم‌های کوانتومی غیرممکن می‌نماید). مدل مداری را به مدل تورینگ ترجیح می‌دهیم. با این توصیف به نظر می‌رسد که ضروری است با توجه به این پارادایم، در تعبیرمان از مفهوم کارایی محاسبه تغییراتی ایجاد کنیم. در مدل مداری، انتخاب طبیعی برای منابع محاسباتی‌ای که پیچیدگی‌شان مورد مطالعه قرار گیرد، اندازه و عمق مدار است، که می‌توان ثابت کرد تناظری بین این دو، با مفاهیم زمان و حافظه در مدل تورینگ وجود دارد [۱۹]. با این وجود، اگر محاسبات کارا در مدل مداری را به عنوان وجود مداری با اندازه‌ی چندجمله‌ای برای یک مسئله تعبیر کنیم، به سادگی می‌توان دید که بسیاری از مسائل محاسبه‌ناپذیر نیز تحت این تعبیر کارا خواهند بود. برای رفع این مشکل باید با گذاشت شرایط مناسبی بر مدارها، به نوعی آن‌ها را ملزم به رفتاری «یک‌نواخت» کرد. در ادامه، یک روش برای چنین کاری را بیان می‌کنیم.

^۱ گرچه در سال‌های اخیر با ظهور و توسعه‌ی حوزه‌هایی مثل تحلیل داده‌های حجمی، مناسب‌بودن این انتخاب برای برخی مقاصد محاسباتی مورد بازنی‌ی قرار گرفته است.

تعريف ۱۰.۳. یک خانواده از مدارها مانند (C_1, C_2, C_3, \dots) ، یکنواخت-چندجمله‌ای نامیده می‌شود هرگاه یک ماشین تورینگ با زمان چندجمله‌ای وجود داشته باشد که با ورودی 1^n ، توصیفی برای مدار C_n خروجی دهد.

پیش از آن که به طور دقیق مجموعه‌ی همه‌ی مسائلی که به طور کارا توسط کامپیوترهای کوانتومی قابل حل هستند را تعریف کنیم، ذکر دو نکته خالی از لطف نیست.

- همان‌گونه که تا به این جا دیده‌ایم، الگوریتم‌های کوانتومی ذاتاً احتمالاتی هستند؛ بنابراین برای تعریف کلاس همه‌ی مسائل قابل حل با الگوریتم‌های کارای کوانتومی، تلاش برای تعریف همتای کوانتومی کلاس BPP نقطه‌ی شروع مناسبتری است.

- با وجود آن‌که کلاس‌های پیچیدگی کلاسیک عمدتاً به عنوان مجموعه‌ای از «زبان»‌ها تعریف می‌شوند، به دلایلی در پیچیدگی محاسباتی کوانتومی تعریف کلاس‌ها بر اساس مسئله‌های قراردادی برتری یافته است. یک مسئله‌ی قراردادی^۱ عبارت است از زوج مرتبی مانند $\Pi = (\Pi_{Yes}, \Pi_{No})$ ، به طوری که $\{0, 1\}^*$ و $\Pi_{Yes}, \Pi_{No} \subseteq \{0, 1\}^*$ و $\Pi_{Yes} \cap \Pi_{No} = \emptyset$. در حالتی که ورودی مسئله عضو $\Pi_{Yes} \cup \Pi_{No}$ باشد، می‌گوییم ورودی قرارداد مسئله را برآورده می‌کند. روشن است که اگر $\{0, 1\}^* = \Pi_{Yes} \cup \Pi_{No}$ ، یک مسئله‌ی تصمیم‌گیری خواهد بود.

مقصودمان از این‌که الگوریتمی یک مسئله‌ی قراردادی $\Pi = (\Pi_{Yes}, \Pi_{No})$ را حل می‌کند این است که اگر ورودی عضو Π_{Yes} باشد، الگوریتم به ازای آن ورودی خروجی «بله» می‌دهد و اگر ورودی عضو Π_{No} باشد، خروجی عضو $\{0, 1\}^* \setminus (\Pi_{Yes} \cup \Pi_{No})$ باشد، خروجی الگوریتم می‌تواند دلخواه باشد. حل کردن یک مسئله‌ی قراردادی را می‌توان در ساختار حل‌پذیری تصادفی نیز، کاملاً مشابه با حل‌پذیری دقیق، تعریف کرد. کافی است حل مسئله را برای ورودی‌هایی که قرارداد مسئله را برآورده می‌کنند مشابه با حل یک مسئله‌ی تصمیم‌گیری تعریف کنیم؛ و برای ورودی‌هایی که قرارداد را برآورده نمی‌کنند، خروجی الگوریتم را دلخواه در نظر بگیریم.

گرچه در پیچیدگی کلاسیک رایج است که اگر کلاسی مانند C بر اساس مسئله‌های قراردادی تعریف شده باشد، از آن به عنوان $PromiseC$ یاد کنند، در پیچیدگی کوانتومی معمولاً از این‌کار عدول می‌شود. بنابراین توجه کنید که همه‌ی کلاس‌هایی که در ادامه تعریف خواهد شد کلاس‌های قراردادی هستند، مگر خلاف آن ذکر شود.

تعريف ۱۱.۳. کلاس پیچیدگی BQP عبارت است از همه‌ی مسائل قراردادی مانند $\Pi = (\Pi_{Yes}, \Pi_{No})$ ، به طوری که مدار کوانتومی یکنواخت-چندجمله‌ای (C_0, C_1, C_2, \dots) و چندجمله‌ای $q(x)$ موجودند به نحوی که برای هر $x \in \{0, 1\}^*$ یک مدار C_n ($n \in \mathbb{N}$) کوانتومی است که روی یک ورودی n کیویت کمکی (رجیستر in) و $q(n)$ کیویت کمکی (رجیستر an) با حالت اولیه‌ی صفر، عمل می‌کند، چنان‌که:

(۱) برای هر ورودی n^x حالت $C_n|x\rangle_{in}^{(0)^{\otimes q(n)}}$ را ورودی می‌گیرد و پس از اعمال شدن آن، یک کیویت خاص (مثلاً اولین کیویت رجیستر an) در پایه‌ی محاسباتی اندازه‌گیری می‌شود. فرض کنید حاصل اندازه‌گیری

$b \in \{0, 1\}$ باشد. در این صورت،

(۲) اگر $x \in \Pi_{Yes}$ ، آنگاه $\Pr[b = 1] \geq \frac{2}{3}$

(۳) اگر $x \in \Pi_{No}$ ، آنگاه $\Pr[b = 1] \leq \frac{1}{3}$

با توجه به تعریفی که پیشتر از گیت‌های کوانتومی ارائه کردیم — این‌که هر نگاشت یکانی یک گیت کوانتومی است — اگر بخواهیم هزینه‌ی محاسباتی یک مدار را تعداد گیت‌های آن تعریف کنیم، چنین هزینه‌ای خوش‌تعریف نخواهد بود؛ زیرا ترکیب چند گیت کوانتومی نیز خود گیتی کوانتومی است. برای رفع این مشکل مجموعه‌ی گیت‌هایی که در مدارها ظاهر می‌شوند را به مجموعه‌ای گسسته از گیت‌ها محدود می‌کنیم. از وجود مجموعه‌های به طور تقریبی جهانی از گیت‌ها می‌دانیم که چنین کاری مشکلی در محاسبه‌پذیری ایجاد نخواهد کرد. بعلاوه، از قضیه‌ی سولووی-کیتائوف می‌توان نتیجه گرفت که انتخاب مجموعه‌های جهانی متفاوت، در حد یک سربار چندجمله‌ای در سایز مدار تفاوت ایجاد خواهد کرد، که با توجه به تعریف فوق قابل تحمل است. بنابراین در ادامه می‌توانیم فرض کنیم که تمام مدارها متشکل از گیت‌هایی از مجموعه‌ی $\{H, T, CNOT\}$ هستند.

¹promise problem

ملاحظه ۱۲.۳. مانند بسیاری دیگر از کلاس‌های پیچیدگی تصادفی، کران‌های ظاهر شده در تعریف ۱۱.۳ را می‌توان در حد وارون نمایی کاهش داد. برای نشان دادن این موضوع کافی است با تکرار الگوریتم به تعداد کافی، از خروجی‌ها رای اکثریت بگیریم و نهایتاً از کران چرنف استفاده کنیم. به طریق مشابه، می‌توان دید که اگر تفاضل کران‌ها در حد وارون چندجمله‌ای باشد نیز، تعریف جدید به همان کلاس \mathcal{BQP} معرفی شده در تعریف ۱۱.۳ منجر خواهد شد.

۴. اثبات‌های (غیرتعاملی) کواتومی

۱.۴. کلاس پیچیدگی \mathcal{QMA} . همان‌گونه که خواهیم دید، کلاس پیچیدگی \mathcal{QMA} تعیینی طبیعی از کلاس \mathcal{NP} به قلمروی محاسبات کواتومی است. با این حال، باید توجه داشت که به دلیل آن که الگوریتم‌های کواتومی ذاتاً احتمالاتی هستند، تعریف کلاس \mathcal{QMA} بیش از آن که به تعریف \mathcal{NP} شبیه باشد، یادآور نسخه‌ی کلاسیک احتمالاتی آن، یعنی \mathcal{MA} است. از پیچیدگی محاسبات کلاسیک می‌دانیم که کلاس \mathcal{NP} را می‌توان با سیستم‌های اثبات نیز مشخص کرد. در واقع اگر $x \in \mathcal{NP}$ ، در این صورت برای هر $L \in \mathcal{NP}$ ، اثباتی کوتاه مانند π_x موجود است که به صورت موثری قابل تصدیق شدن است، و برای هر $L \notin \mathcal{NP}$ ، چنین اثباتی وجود ندارد. در اینجا یادآور می‌شویم که مقصودمان از کوتاه بودن اثبات آن است که طول اثبات π_x از مرتبه‌ی چندجمله‌ای بر حسب طول ورودی x است، و مقصود از تصدیق کردن به طور موثر، وجود الگوریتمی مانند V_L است که x و π_x را به عنوان ورودی می‌گیرد، و در زمان چندجمله‌ای بر حسب طول x ، اگر π_x اثباتی درست برای $x \in L$ باشد، خروجی «بله» می‌دهد.

تعیین‌های کواتومی متفاوتی را می‌توان برای سیستم اثبات فوق در نظر گرفت. در ادامه یکی از این تعیین‌ها را، که در آن الگوریتم تصدیق‌کننده با یک الگوریتم کواتومی و اثبات نیز با یک اثبات کواتومی جایگزین می‌شود، بررسی خواهیم کرد؛ تعیینی که برای اولین بار در [۱] معرفی شده است. یادآوری می‌کنیم که همان‌گونه که پیشتر تصریح کردیم، در پیچیدگی محاسبات کواتومی مسائل و کلاس‌های قراردادی مورد توجه ما هستند، و تعریف پیش رو نیز مشخص‌کننده‌ی یک کلاس قراردادی است.

تعریف ۱.۴. برای هر چندجمله‌ای $(x, p(x))$ کلاس $\mathcal{QMA}_p(\frac{2}{3}, \frac{1}{3})$ عبارت است از تمام مسئله‌های قراردادی مانند $\Pi = \Pi_{Yes} \cup \Pi_{No}$ به طوری که مدار کواتومی یکنواخت-چندجمله‌ای (C_0, C_1, C_2, \dots) و چندجمله‌ای $q(x)$ موجودند به نحوی که برای هر $C_n, n \in \mathbb{N}$ یک مدار کواتومی است که روی یک ورودی n کیویتی $q(n)$ (رجیستر in)، یک اثبات کواتومی کیویتی $p(n)$ (رجیستر pr) و $q(n)$ کیویتی کمکی (رجیستر an) با حالت اولیه‌ی صفر، عمل می‌کند، چنان‌که:

(۱) برای هر ورودی $n \in \{0, 1\}^n$ و هر اثبات $x \in \{0, 1\}^n$ حالت $C_n |\psi\rangle_{pr}^{(C^2)^{\otimes p(n)}} |\psi\rangle_{an}^{(C^2)^{\otimes q(n)}} |x\rangle_{in}$ را ورودی می‌گیرد و پس از اعمال شدن آن، یک کیویتی خاص (مثلاً اولین کیویتی رجیستر an) در پایه‌ی محاسباتی اندازه‌گیری می‌شود. فرض کنید حاصل اندازه‌گیری $b \in \{0, 1\}$ باشد.

(۲) تمامیت: اگر $x \in \Pi_{Yes}$ ، در این صورت اثبات $|\psi\rangle \in (\mathbb{C}^2)^{\otimes p(n)}$ وجود دارد به طوری که $\Pr[b = 1] \geq \frac{2}{3}$.

(۳) درستی: اگر $x \in \Pi_{No}$ ، در این صورت برای هر اثبات $|\psi\rangle \in (\mathbb{C}^2)^{\otimes p(n)}$ $\Pr[b = 1] \leq \frac{1}{3}$.

توجه کنید که اگر به جای ثابت تمامیت و درستی اعداد (یا توابع) a و b را قرار دهیم، کلاس $\mathcal{QMA}_p(a, b)$ به دست می‌آید. همچنین تعریف می‌کنیم: $\mathcal{QMA}(a, b) = \bigcup_{p(x)} \mathcal{QMA}_p(a, b)$ را معمولاً به اختصار با نشان می‌دهیم.

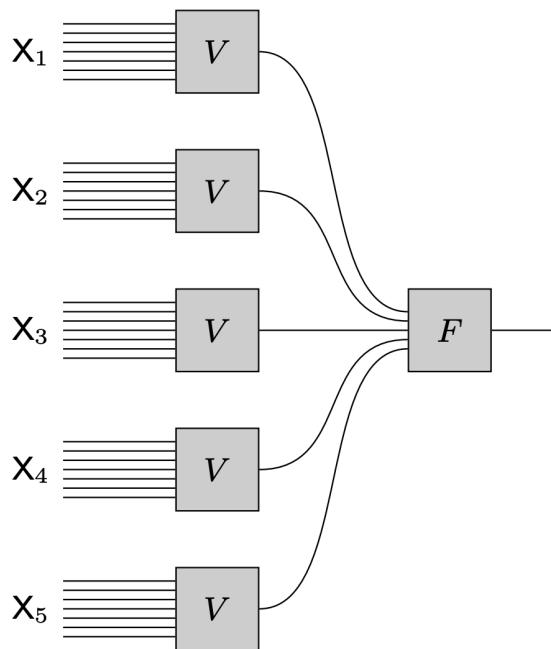
همان‌گونه که از تعریف بالا بر می‌آید، در تعریف کلاس \mathcal{QMA} ، روی (توزع) خروجی مدار در حالتی که ورودی x رشته‌ای عضو $\Pi_{Yes} \cup \Pi_{No}$ نیست، شرطی نداریم و در چنین حالتی، خروجی می‌تواند دلخواه باشد.

ملاحظه ۲.۴. \mathcal{QMA} سروازه‌ای برای کواتوم مرلین-آرتور^۱ است و بیان می‌کند که کلاس فوق همتای کواتومی کلاس پیچیدگی مرلین-آرتور (\mathcal{MA}) است. این نام‌گذاری اولین بار در [۲] به کار رفت و به تدریج جایگزین \mathcal{BQNP} ، نامی که [۱] نخستین بار برای این کلاس به کار برده بود، شد.

کاهش احتمال خطای \mathcal{QMA} در تعریف ۱.۴: مشابه دیگر کلاس‌های پیچیدگی احتمالاتی با احتمال خطای کراندار، در تعریف کلاس \mathcal{QMA} نیز می‌توان این سوال را مطرح کرد که آیا کران بالای $\frac{2}{3}$ روی احتمال خطای کراندار باشد یا نه. می‌دانیم کاهش

^۱Quantum Merlin-Arthur

شکل ۱: کاهش خطای موازی، تصویر برگرفته شده از مرجع [۳] است.



خطای کلاس $M.A$ امکان‌پذیر است؛ کافی است آرتور اثبات دریافت‌شده از مرلین را $k \in O(\log(\frac{1}{\varepsilon}))$ بار کپی کند و برای هر کپی، الگوریتم تصدیق‌کننده را یک‌بار اجرا کند و نهایتاً از خروجی‌های دفعات مختلف اجرای الگوریتم رای اکثربت بگیرد. به این ترتیب با استفاده از کران چرنف می‌توان دید کران بالای احتمال خطای $\varepsilon = 2^{-r(x)}$ را که $r(x) \in \mathbb{C}^{kp(n)}$ یک چندجمله‌ای است، کاهش می‌یابد.

با این حال، در کلاس $Q.M.A$ باید به این مطلب توجه کرد که بنابر قضیه‌ی عدم امکان شبیه‌سازی، نمی‌توان اثبات ارسال‌شده از طرف مرلین را کپی کرد و آرتور باید از خود مرلین بخواهد که k نسخه از اثبات را برایش ارسال کند. به این روش، روش کاهش خطای موازی^۱ یا کاهش خطای ضعیف^۲ می‌گویند. در این روش، مرلین یک اثبات $\psi' \in (\mathbb{C}^2)^{\otimes kp(n)}$ را برای آرتور ارسال می‌کند و آرتور باید مشابه همان کاری که در کاهش خطای $M.A$ انجام می‌داد را تکرار کند. در این روش کاهش خطای دو مشکل قابل طرح است:

- آیا درهم‌تییدگی امکان تقلب به مرلین نمی‌دهد؟

در حالتی که $x \in \Pi_{Y_{\text{Yes}}}$ ، می‌دانیم اثبات $\psi' \in (\mathbb{C}^2)^{\otimes kp(n)}$ وجود دارد که تصدیق‌کننده با احتمال حداقل $\frac{1}{2}$ آن را می‌پذیرد. در این حالت، مرلین کافی است k نسخه از این اثبات را به صورت $\psi' = |\psi'\rangle \otimes |\psi\rangle \otimes \dots \otimes |\psi\rangle$ برای آرتور بفرستد و آرتور مطابق شکل ۱ الگوریتم تصدیق‌کننده را روی هر یک از k رجیستر اثبات اجرا کند و نهایتاً رای اکثربت بگیرد.

با این وجود، در حالتی که $x \in \Pi_{N_o}$ ، باید برای هر اثبات $\psi' \in (\mathbb{C}^2)^{\otimes kp(n)}$ احتمال پذیرفته شدن اثبات توسط آرتور کراندار باشد. در این حالت، ممکن است مرلین اثباتی درهم‌تیید برای آرتور ارسال کند. به این ترتیب اگر آرتور مطابق شکل ۱ عمل کند، حالت رجیسترها مختلف اثبات لزوماً حالت خالص^۳ نخواهد ماند و ممکن است به دلیل درهم‌تییدگی، رجیستری از اثبات در حالت مخلوط قرار گیرد.

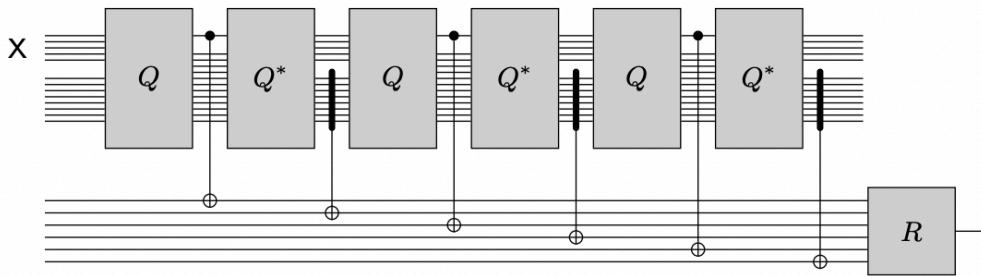
با این حال به سادگی می‌توان دید که اگر برای هر اثبات ψ در حالت خالص، بدانیم آرتور آن را با احتمال حداقل

¹Parallel Error Reduction

²Weak Error Reduction

³فرض کنید مجموعه‌ای از حالت‌ها مانند $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$ داریم که روی آن‌ها توزیع احتمالی مانند p_1, \dots, p_n وجود دارد. اگر بدانیم دقیقاً یکی از ها برابر با ۱ و مابقی برابر صفرند، به حالت این هنگردد حالت خالص و در غیر این صورت، حالت مخلوط گفته می‌شود. حالت‌های مخلوط را می‌توان با کمک فرمول‌بندی ماتریس‌های چگالی مطالعه کرد. خواننده‌ی علاقه‌مند می‌تواند برای مطالعه‌ی بیشتر درباره‌ی حالت‌های مخلوط به [۴۵] مراجعه کند.

شکل ۲: کاهش خطای حافظ اثبات، تصویر برگرفته شده از مرجع [۳] است.



$\frac{1}{3}$ می‌پذیرد، در این صورت برای هر اثبات با حالت مخلوط ρ نیز احتمال پذیرفته شدن حداقل $\frac{1}{3}$ خواهد بود. بنابراین، در هم‌تنیدگی نمی‌توان امکان تقلب را برای مرلین فراهم کند. نهایتاً با استفاده از روشی که در بالا گفته شد، می‌توان قضیه‌ی زیر را ثابت کرد:

قضیه ۳.۴. برای هر چندجمله‌ای $p(n)$ و ثابت $1 < c < \infty$ ، داریم [۴]:

$$\mathcal{QMA}(c - \frac{1}{p(n)}, c) \subseteq \mathcal{QMA}(\frac{1}{3}, \frac{1}{3}) = \mathcal{QMA}(\frac{1}{p(n)}, 1 - \frac{1}{p(n)}).$$

- اندازه‌ی اثبات در این روش افزایش یافته است. آیا این افزایش طول اثبات غیرقابل اجتناب است؟ در واقع، این افزایش طول اثبات ضروری نیست. [۵] روشی هوشمندانه موسوم به کاهش خطای حافظ اثبات یا کاهش خطای قوی (شکل ۲) ارائه کرده است که در نتیجه‌ی آن قضیه‌ی زیر را خواهیم داشت:

قضیه ۴.۴. فرض کنید $[1, \infty)$ دو تابع محاسبه‌پذیر در زمان چندجمله‌ای باشند و $(q(x))$ یک چندجمله‌ای باشد به نحوی که برای هر $n \in \mathbb{N}$ (به جز احتمالاً تعدادی متناهی از اعداد طبیعی)،

$$a(n) - b(n) \geq \frac{1}{q(n)}. \quad (1.4)$$

در این صورت برای هر دو چندجمله‌ای $(x)^r p(x), r(x)$ با این شرط که $2^r \geq n$ برای هر عدد طبیعی n (جز احتمالاً تعداد متناهی از اعداد طبیعی)، داریم:

$$\mathcal{QMA}_p(a, b) = \mathcal{QMA}_p(1 - 2^{-r}, 2^{-r}). \quad (2.4)$$

مسئله‌ی دیگری که پس از تعریف کلاس \mathcal{QMA} باید به آن پاسخ دهیم، بررسی رابطه‌ی این کلاس با دیگر کلاس‌های پیچیدگی و یافتن کران‌های پایین و بالایی برای آن است. روشن است که \mathcal{MA} و \mathcal{BQP} ، هردو، کران‌های پایینی برای \mathcal{QMA} هستند (زیرا محاسبات کلاسیک را می‌توان با محاسبات کوانتومی شبیه‌سازی کرد). در ادامه، کران‌های بالایی را نیز برای \mathcal{QMA} خواهیم یافت.

(۱) به سادگی می‌توان نشان داد که $\mathcal{QMA} \subseteq \mathcal{NEXP}$. فرض کنید که Π مسئله‌ای در \mathcal{QMA} باشد. در این صورت، ماشینی را در نظر بگیرید که ابتدا یک اثبات $(\psi) \in (\mathbb{C}^2)^{\otimes p(n)}$ را به صورت غیرقطعی حدس می‌زند (تعداد پارامترهای چنین اثباتی بر حسب n نمایی است). سپس احتمال این که مدار تصدیق‌کننده‌ی آرتور خروجی ۱ بددهد را محاسبه می‌کند و با توجه به مقدار این احتمال، اثبات را می‌پذیرد یا رد می‌کند. محاسبه‌ی این احتمال در زمان نمایی بر حسب n ممکن است. بنابراین ماشین توصیف شده در بالا، ماشینی غیرقطعی با زمان نمایی است، که نتیجه می‌دهد مسئله‌ی مورد نظر عضو کلاس \mathcal{NEXP} است.

(۲) به عنوان کران بالایی نابدیهی \mathcal{NEXP} ، می‌توان نشان داد که $\mathcal{QMA} \subseteq \mathcal{EXP}$. بدین منظور، نخست توجه کنید که احتمال این که مدار تصدیق‌آرتور برای ورودی با طول n (که آن را با Q_n نمایش می‌دهیم) به ازای اثبات (ψ)

خروجی ۱ بدهد برابر است با^۱:

$$\begin{aligned} \Pr[\text{output} = 1] &= \|(|\rangle\langle|\otimes\mathbb{I}_{N-1})Q_n|x\rangle_{in}|\psi\rangle_{pr}|^{\circ q(n)}\rangle_{an}\|_2^2 \\ &= \text{tr}\left(\langle x|_{in}\langle\psi|_{pr}|^{\circ q(n)}\rangle_{an}Q_n^\dagger(|\rangle\langle|\otimes\mathbb{I}_{N-1})Q_n|x\rangle_{in}|\psi\rangle_{pr}|^{\circ q(n)}\rangle_{an}\right) \\ &= \text{tr}\left(P_x|\psi\rangle\langle\psi|\right) = \langle\psi|P_x|\psi\rangle \end{aligned}$$

که در آن، $N = n + p(n) + q(n)$ و

$$P_x = \left(\langle x|_{in}\otimes\mathbb{I}_{pr}\otimes|\circ q(n)\rangle_{an}Q_n^\dagger(|\rangle\langle|\otimes\mathbb{I}_{N-1})Q_n|x\rangle_{in}\otimes\mathbb{I}_{pr}\otimes|\circ q(n)\rangle_{an}\right).$$

از طرفی می‌دانیم که

$$\max_{|\psi\rangle : \langle\psi|\psi\rangle=1} \langle\psi|P_x|\psi\rangle = \lambda_{\max}(P_x).$$

بنابراین، برای حل یک مسئله‌ی قراردادی در کلاس \mathcal{QMA} مانند Π ، که عبارت است از تعیین این‌که برای هر $x \in \Pi_{Yes} \cup \Pi_{No}$ ، کدامیک از $x \in \Pi_{Yes}$ یا $x \in \Pi_{No}$ درست است، کافی است بزرگترین مقدار ویژه‌ی عملگر خطی P_x را محاسبه کنیم. روشن است که ابعاد P_x بر حسب n نمایی است. با این وجود، پیدا کردن مقدار ویژه‌های یک ماتریس، در زمان چندجمله‌ای بر حسب ابعاد آن امکان‌پذیر است. در نتیجه، هر مسئله‌ی \mathcal{QMA} را می‌توان با ماشینی قطعی در زمان نمایی حل کرد.

(۳) با استفاده از کاهش خطای قوی می‌توان نشان داد $\mathcal{QMA} \subseteq \mathcal{PP}$ [۵]. برای هر چندجمله‌ای دلخواه p و هر مسئله‌ی قراردادی دلخواه در \mathcal{QMA}_p مانند $\Pi = (\Pi_{Yes}, \Pi_{No})$ ، از قضیه‌ی ۴.۴ می‌دانیم:

$$\Pi \in \mathcal{QMA}_p(1 - 2^{-(p(n)+2)}, 2^{-(p(n)+2)})$$

حال الگوریتمی کوانتومی را در نظر بگیرید که برای یک ورودی دلخواه با طول n ، اثباتی را به تصادف از بین همه‌ی اثبات‌های با طول $p(n)$ انتخاب می‌کند، و آن را در رجیستر اثبات مدار تصدیق‌کننده‌ی مسئله‌ی Π قرار می‌دهد و الگوریتم تصدیق‌کننده را اجرا می‌کند. احتمال این‌که این الگوریتم خروجی ۱ دهد برابر است با :

$$\text{tr}(P_x \frac{\mathbb{I}}{2^{p(n)}}) = \frac{1}{2^{p(n)}} \text{tr}(P_x).$$

- در حالتی که $\frac{1}{2^{p(n)}} \text{tr}(P_x) \geq \frac{1}{2^{p(n)}}(1 - \frac{1}{2^{p(n)+2}}) \geq \frac{1}{2^{p(n)+1}}$ ، $x \in \Pi_{Yes}$
- در حالتی که $\frac{1}{2^{p(n)}} \text{tr}(P_x) \leq \frac{1}{2^{p(n)+2}}$ ، $x \in \Pi_{No}$.

با توجه به فاصله‌ی بین ضرایب درستی و تمامیت در بالا، می‌توان دید که فاصله‌ی مورد نظر در تعریف کلاس \mathcal{PP} برآورده می‌شود. تنها مشکل این جاست که الگوریتمی که در بالا ارائه شده است، الگوریتمی کوانتومی، و نه کلاسیک است. به عبارت دیگر، آن‌چه که در بالا ارائه کردۀایم نشان می‌دهد که Π عضو کلاس \mathcal{PQP} است؛ کلاسی که همتای کوانتومی \mathcal{PP} محسوب می‌شود. با این همه، یاماکامی در [۶] نشان داده است که $\mathcal{PQP} = \mathcal{PP}$ ، و به این ترتیب اثبات تکمیل خواهد شد.

۲.۴. نسخه‌هایی دیگر از \mathcal{QMA} . در این بخش به معرفی نسخه‌هایی تغییریافته از کلاس \mathcal{QMA} می‌پردازیم و برخی ویژگی‌های اثبات‌شده و حدس‌های اثبات‌نشده را در ارتباط با این کلاس‌ها مرور خواهیم کرد.

• \mathcal{QCMA} با اثبات‌های کلاسیک (\mathcal{QCMA}):

اگر در تعریف کلاس \mathcal{QMA} فرض کنیم اثباتی که توسط مولین ارسال می‌شود یک رشته‌ی کلاسیک است، کلاس \mathcal{QCMA} به دست می‌آید. روشن است که $\mathcal{QCMA} \subseteq \mathcal{QMA}$. با این حال، این مسئله که آیا این شمول اکید است یا نه، مسئله‌ای باز است. آرانسون و کوپربرگ در [۷] نشان داده‌اند که یک اوراکل کوانتومی O وجود دارد که

^۱ در اینجا فرض کردۀایم که مقدار بیت خروجی با اندازه‌گیری اولین کوییت تعیین می‌شود.

$\mathcal{QMA}^O \neq \mathcal{QCMA}^O$. این نتیجه به این معنی است که در پاسخ به این سوال که آیا $\mathcal{QMA} = \mathcal{QCMA}$ یا نه، نیازمند تکمیک‌هایی هستیم که قابل نسبی شدن با اوراکل‌های کواتومی نیستند.

• \mathcal{QMA} با خطای یک‌طرفه (\mathcal{QMA}_1): اگر در تعریف کلاس \mathcal{QMA} ، این تغییر را ایجاد کنیم که در حالتی که ورودی عضو $\Pi_{Y_{\text{Yes}}}$ است، اثباتی وجود داشته باشد که احتمال پذیرفته شدن آن توسط آرتور برابر با ۱ باشد، کلاس پیچیدگی \mathcal{QMA}_1 به دست می‌آید. روش است که $\mathcal{QMA}_1 \subseteq \mathcal{QMA}$. با این حال، این‌که آیا این شمول اکید است یا نه، مسئله‌ای باز است. آرانسون در [۱۸] اوراکلی کواتومی مانند O ارائه می‌دهد که $\mathcal{QMA}^O \neq \mathcal{QMA}_1^O$.

ملاحظه ۵.۴. با وجود این‌که پرسش $\mathcal{QMA}_1 = \mathcal{QMA}$ بدون پاسخ مانده است، سوالی مشابه، $\mathcal{QCMA}_1 = ?$ ، توسط کوبایاشی و همکاران در [۹] پاسخ داده شده است و می‌دانیم \mathcal{QCMA} با \mathcal{QCMA}_1 با خطای یک‌طرفه برابر است. به این ترتیب، بلاfaciale می‌توان نتیجه گرفت که $\mathcal{QCMA} \subseteq \mathcal{QMA}_1$.

• \mathcal{QMA} با k مرلین (($\mathcal{QMA}(k)$)): اگر در تعریف کلاس \mathcal{QMA} ، این تغییر را ایجاد کنیم که اثبات به صورت حاصل‌ضرب تنسوری k حالت (n) کیویتی باشد، در این صورت کلاس پیچیدگی $(\mathcal{QMA}(k))$ به دست می‌آید. این کلاس نخستین بار توسط کوبایاشی و همکاران در [۱۰] معرفی شد.

می‌توان درباره‌ی این کلاس چنین اندیشید که مجموعه‌ی تمام مسائلی است که می‌توان آن‌ها را با سیستم‌های اثبات غیرتعاملی با چند اثبات‌کننده مشخص کرد، به طوری که اثبات‌کننده‌ها نیز با یکدیگر تعاملی ندارند (جداپذیر بودن اثبات را می‌توان چنین تعبیر کرد). در چهارچوب پیچیدگی محاسبات کلاسیک، افزودن به تعداد اثبات‌کننده‌های یک سیستم اثبات غیرتعاملی، با این فرض که اثبات‌کننده‌ها نیز با یکدیگر تعامل نداشته باشند، چیزی بر قدرت محاسباتی نمی‌افزاید؛ حال آن‌که در چهارچوب پیچیدگی کواتومی هنوز نمی‌دانیم که چنین نتیجه‌ای هم چنان برقرار خواهد ماند. به بیان دقیق‌تر، روش است که $\mathcal{QMA} \subseteq \mathcal{QMA}(k)$ ؛ با این وجود، اکید بودن این شمول هنوز مسئله‌ای بی‌پاسخ است. لیو و همکاران در [۱۱] مسئله‌ای به نام N -نمایش‌پذیری حالت خالص^۱ را پیشنهاد داده‌اند که در $\mathcal{QMA}(k)$ قرار دارد اما به نظر نمی‌رسد که عضو \mathcal{QMA} باشد. پرسش دیگر در مورد \mathcal{QMA} با چند مرلین این است که آیا در حالتی که تعداد اثبات‌کننده‌ها حداقل دو تاست، با افزایش تعداد مرلین‌ها قدرت محاسباتی افزایش می‌یابد یا نه. هرو و مونتانا رو در [۱۲] به این پرسش پاسخ داده و نشان داده‌اند که برای هر $2 \leq k \leq N$ $\mathcal{QMA}(k) = \mathcal{QMA}(2)$. بهترین کران‌های بالا و پایین شناخته شده برای $\mathcal{QMA}(k)$ به ترتیب کران‌های بدیهی \mathcal{NEXP} و \mathcal{QMA} هستند. البته کران بالای دیگری نیز برای $\mathcal{QMA}(k)$ شناخته شده است که کلاس Σ_3^Q می‌باشد. کلاس اخیر همتای کواتومی Σ_3 ، طبقه‌ی سوم سلسه‌مراتب چندجمله‌ای است. با این وجود، کلاس مزبور چندان شناخته شده نیست و شواهدی وجود ندارد که $\Sigma_3 \neq \mathcal{NEXP}$.

• اگر در تعریف \mathcal{QMA} ، تغییرات زیر را اعمال کنیم StoqMA کلاس باشد به دست می‌آید:

(۱) کیویت‌های کمکی می‌توانند با مقادیر اولیه‌ی $(+)$ یا $(-)$ مقداردهی شوند.

(۲) مداری که آرتور اعمال می‌کند، تنها از گیت‌های وارون‌پذیر کلاسیک تشکیل شده است.

(۳) اندازه‌گیری نهایی در پایه‌ی X انجام می‌شود.

در واقع، StoqMA را می‌توان به صورت سیستم اثباتی دید که در آن اثبات، یک حالت کواتومی است اما مدار تصدیق، یک مدار کلاسیک است.

کلاس StoqMA نسخه‌ای عجیب از \mathcal{QMA} است. یکی از وجوه تفاوت StoqMA با دیگر نسخه‌های \mathcal{QMA} این است که باور بر این است که StoqMA شامل \mathcal{BQP} نیست. در توضیح می‌توان گفت که از یک سو، همان‌گونه که ترهال و همکاران در [۱۳] نشان داده‌اند، $\text{StoqMA} \subseteq \mathcal{AM} \subseteq \mathcal{PH}$. از سوی دیگر، باور بر این است که $\mathcal{BQP} \not\subseteq \mathcal{PH}$. بنابراین، شمول \mathcal{BQP} در StoqMA بعید دانسته می‌شود.

وجه دیگری از تفاوت‌های StoqMA با دیگر نسخه‌ها این است که برقرار بودن کاهش خطای ضعیف برای این کلاس، مسئله‌ای باز است. اخیراً آهارونوف و همکاران در [۱۴] نشان داده‌اند که امکان کاهش خطای StoqMA از

^۱pure state N -Representability

(1) به $O(1 - \frac{1}{\text{poly}(n)})$ نتیجه خواهد داد که $\text{StoqMA} = \text{MA}$. با این حال، هنوز مشخص نیست که آیا این نتیجه در تعیین تکلیف کاهش خطای ضعیف برای این کلاس تاثیری خواهد داشت یا نه.

۵. پیچیدگی همیلتونی کوانتمویی

مطالعه‌ی رفتار سیستم‌های فیزیکی متشکل از تعدادی ذره در حال حرکت، مسئله‌ای است که در قلب مکانیک (کوانتمویی) قرار دارد. توصیف چنین سیستم‌هایی، وقتی که از تعداد زیادی ذره تشکیل شده‌اند که با یکدیگر برهمنش می‌کنند، به دلیل رفتار پیچیده‌ای که سیستم از خود نشان می‌دهد، کار سختی است. انگیزه‌ی اصلی نظریه‌ی سیستم‌های چندپیکره^۱ در فیزیک، سعی در فهمیدن ویژگی‌های چنین سیستم‌هایی است. از دیگر سو، اگر مطالعات فیزیکی را در سه مرحله‌ی مدل‌سازی، حل تقریبی مدل و پیش‌بینی کردن یک کمیت بر اساس آن، و نهایتاً بررسی کمیت پیش‌بینی شده به صورت تجربی و تنظیم و بهبود مدل خلاصه کنیم، به دلیل ذات الگوریتمی مرحله‌ی دوم ملاحظات پیچیدگی محاسباتی در این مرحله اهمیت پیدا می‌کنند [۱۵]. در این بخش، توجه ما معطوف به مطالعه‌ی پیچیدگی محاسباتی روش‌هایی است که در نظریه‌ی سیستم‌های چندپیکره برای توصیف سیستم‌های کوانتمویی به کار گرفته می‌شود. یک مشاهده‌ی غیرمنتظره آن است که مشابهتی کانونی میان اشیاء مورد مطالعه در نظریه‌ی پیچیدگی محاسبه و نظریه‌ی سیستم‌های چندپیکره وجود دارد، و همین مشابهت‌ها انگیزه‌بخش ما برای تلاش در جهت استفاده از ابزارهای نظریه‌ی پیچیدگی محاسبه برای پاسخ دادن به این پرسش خواهد بود که «شبیه‌سازی یک سیستم فیزیکی تا چه اندازه سخت است؟». مطالعه‌ی این مشابهت‌ها، که به شکوفایی‌هایی هم در علوم کامپیوتر و هم در فیزیک انجامیده است، حوزه‌ای است که از آن به عنوان پیچیدگی همیلتونی کوانتمویی^۲ یاد می‌شود، و از مهم‌ترین زمینه‌های پژوهش در اشتراک فیزیک و علوم کامپیوتر به حساب می‌آید.

۱.۵. همیلتونی‌های موضعی. در یک سیستم فیزیکی متشکل از n ذره، وقتی که n بزرگ می‌شود، برهمنش ذرات با یکدیگر پیچیده‌تر شده و توصیف حالت سیستم دشوار می‌شود. فیزیکدانان برای مدل‌سازی چنین سیستم‌هایی عموماً فرض‌هایی ساده‌کننده را در مدل لحاظ می‌کنند. مثلاً فرض می‌کنند که آرایش ذرات به صورت یک شبکه‌ی ۲-۳ بعدی است؛ و ذرات در چنین آرایشی تنها با نزدیک‌ترین همسایه‌شان برهمنش می‌کنند. مدل‌های ساده‌شده‌ی مختلفی در نظریه‌ی سیستم‌های چندپیکره برای توصیف سیستم‌های فیزیکی توسعه یافته است، که از جمله‌ی آن‌ها می‌توان به مدل آیسینگ^۳، مدل هایزینبرگ و مدل AKLT اشاره کرد.

با داشتن مدلی در دست، سوال‌های متعددی را می‌توان درباره‌ی سیستم طرح کرد. مثلاً می‌توان به محاسبه‌ی یک ویژگی موضعی از سیستم (مثلاً حالت یک زیرسیستم متشکل از تعداد کوچکی ذره در یک دمای خاص) پرداخت، یا تحول سیستم را در طول زمان مورد مطالعه قرار داد. همان‌گونه که در ادامه خواهیم دید، بخش قابل توجهی از تلاش‌های سیستم‌های چندپیکره مربوط به مطالعه‌ی انرژی سیستم است. این تمرکز بر انرژی سیستم از آن جهت است که در عمل، محاسبه‌ی انرژی سیستم می‌تواند منجر به محاسبه‌ی بسیاری از کمیت‌های موضعی آن شود.

در فیزیک کلاسیک، برای هر سیستم فیزیکی با یک فضای حالت مانند \mathcal{S} ، تابعی مانند $\mathcal{E} : \mathcal{S} \rightarrow \mathbb{R}$ وجود دارد که بیانگر انرژی سیستم است. در واقع این تابع، به هر حالت که سیستم ممکن است در آن قرار گیرد، یک انرژی نسبت می‌دهد. مثلاً در مدل آیسینگ کلاسیک، فرض بر این است که ذرات روی رؤوس یک شبکه قرار دارند، و حالت سیستم به صورت n تایی‌هایی مانند $\{x_1, x_2, \dots, x_n\} \in \{-1, 1\}^n$ است؛ که $\mathcal{E}(x_1, x_2, \dots, x_n) = \sum_{i,j} J_{i,j} x_i x_j$ تعریف می‌شود، که $J_{i,j}$ ها قدرت برهمنش را مدل می‌کنند و مقصود از $\langle i, j \rangle$ این است که جمع روی همه‌ی زوج‌های (i, j) ای که نزدیک‌ترین همسایه هستند، انجام می‌شود.

در فیزیک کوانتم، انرژی سیستم را وقتی در حالت ψ قرار دارد، نمی‌توان به صورت قطعی تعیین کرد. در واقع، انرژی مشاهده‌پذیری است که ویژه‌مقادیر آن بیانگر سطوح انرژی سیستم هستند، و با اندازه‌گیری انرژی سیستم، بر اساس توزیع احتمالی روی این سطوح انرژی وجود دارد، حاصل اندازه‌گیری یکی از این ویژه‌مقادیر خواهد بود. به چنین مشاهده‌پذیرهایی همیلتونی

¹many body theory

²quantum Hamiltonian complexity

³Ising model

سیستم گفته می‌شود. به عنوان مثال، در مدل آیسینگ کوانتومی، همیلتونی به صورت

$$H = -J \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z - g \sum_i \sigma_i^x$$

تعریف می‌شود، که در آن σ_i^x و σ_i^z عملگرهای پاولی X و Z هستند، و g بیانگر بزرگی میدان مغناطیسی است [۱۶]. در بین سطوح انرژی مختلف، سطح انرژی کمینه از اهمیت ویژه‌ای برخوردار است. چه آن‌که از توزع بولتزمن می‌دانیم حالت سیستم در دمای بسیار پایین و نزدیک به صفر، هنگردی از حالت‌های با انرژی کمینه خواهد بود، و به این ترتیب، با دانستن کمینه‌ی انرژی سیستم و حالت‌هایی که سیستم در آن‌ها این انرژی کمینه را دارد، می‌توان اطلاعاتی درباره‌ی بسیاری از خواص ترمودینامیکی سیستم در دمای پایین به دست آورد.

از سوی دیگر، همان‌گونه که در اصل ۴.۲ دیدیم، تحول زمانی یک سیستم فیزیکی با معادله‌ی شرودینگر توصیف می‌شود که به صورت زیر است:

$$i\hbar \frac{d|\psi(t)\rangle}{dt} = H |\psi(t)\rangle$$

و در این معادله، همیلتونی سیستم است که نحوه‌ی تحول آن را تعیین می‌کند. همین سبب می‌شود که همیلتونی‌ها در مسأله‌ی شبیه‌سازی سیستم‌های کوانتومی، به عنوان توصیفی از سیستمی که قصد شبیه‌سازی آن را داریم، ظاهر شوند. در این مسأله، توصیفی از یک همیلتونی H ، حالت اولیه‌ی ρ ، مشاهده‌پذیری مانند M و لحظه‌ای از زمان مانند t به عنوان ورودی داده شده است و خواسته‌ی مسأله آن است که به عنوان خروجی، تقریبی از

$$\text{Tr} \left[M \frac{(e^{iHt})^\dagger \rho e^{iHt}}{\text{Tr}((e^{iHt})^\dagger \rho e^{iHt})} \right]$$

محاسبه شود. می‌توان دید که مسأله‌ی یافتن تقریبی از کمینه‌ی انرژی سیستم، حالت خاصی از مسأله‌ی فوق است [۱۵]. با مقدمه‌ی بالا، در ادامه‌ی این بخش تمرکز خود را بر روش‌های محاسباتی برای یافتن تقریبی از کمینه‌ی مقدار انرژی برخی سیستم‌های خاص خواهیم گذاشت و پیچیدگی محاسباتی این روش‌ها را مطالعه خواهیم کرد.

تعریف ۱.۵. یک همیلتونی k موضعی^۱ بر روی یک سیستم n کیویتی، عملگری هرمیتی مانند $H : (\mathbb{C}^2)^{\otimes n} \rightarrow (\mathbb{C}^2)^{\otimes n}$ است که می‌توان آن را به صورت $H = \sum_i H^{(i)}$ نوشت، به نحوی که هر $H^{(i)}$ عملگری هرمیتی است که فقط روی k کیویت سیستم به طور نابدیهی عمل می‌کند. مقادیر ویژه‌ی H ، سطوح انرژی^۲ سیستم توصیف شده با H نامیده می‌شوند و کمترین مقدار ویژه‌ی H که آن را با $\lambda_{\min}(H)$ نمایش می‌دهیم، انرژی حالت پایه‌ی سیستم^۳ نام دارد. همچنین به بردار ویژه‌ی متناظر با $\lambda_{\min}(H)$ حالت پایه‌ی^۴ سیستم گوییم.

ملاحظه ۲.۵. در حالتی کلی تراز تعریف سطوح انرژی که در بالا بیان شد، می‌توان گفت که هر همیلتونی که بر یک سیستم n کیویتی عمل می‌کند، به هر حالت سیستم مانند $(\mathbb{C}^2)^{\otimes n} \in \{| \psi \rangle\}$ یک انرژی نسبت می‌دهد که برابر با مقدار $\langle H | \psi \rangle$ است (توجه کنید که این عدد، حقیقی است). به سادگی می‌توان دید که انرژی حالت پایه‌ی سیستم، کمینه‌ی مقدار انرژی همه‌ی حالت‌های سیستم است. به عبارت دیگر:

$$\lambda_{\min}(H) = \min_{|\psi\rangle} \langle \psi | H | \psi \rangle. \quad (1.5)$$

تعریف ۳.۵ (مسأله‌ی همیلتونی k موضعی). فرض کنید $\mathbb{N} \rightarrow \mathbb{R}^+$: p یک چندجمله‌ای باشد. مسأله‌ی همیلتونی k موضعی با فاصله‌ی قراردادی $p(n-k)$ ، مسأله‌ای قراردادی است که به صورت زیر تعریف می‌شود [۱]:

• ورودی: توصیفی از یک همیلتونی k موضعی^۵ ($k \in O(1)$) که بر روی n کیویت عمل می‌کند و توابع به طور موثر محاسبه‌پذیر ($\alpha(n), \beta(n)$) به طوری که^۶ $\beta(n) - \alpha(n) \geq \frac{1}{p(n)}$

¹ k -Local Hamiltonian

² energy levels

³ ground state energy

⁴ ground state

⁵ توجه کنید که این فاصله را می‌توان به یک عدد ثابت افزایش داد؛ کافی است هر جمله در H را $p(n)$ بار تکرار کنیم [۱۶].

• خروجی:

- اگر $\lambda_{min}(H) \leq \alpha(n)$ ، خروجی «بله» می‌دهیم.
- اگر $\lambda_{min}(H) \geq \beta(n)$ ، خروجی «خیر» می‌دهیم.
- در حالتی غیر از دو حالت فوق، به طور دلخواه خروجی می‌دهیم.

در ادامه این که این مسئله بر حسب p پارامتریزه شده است را به طور ضمنی فرض می‌کنیم و از نوشتن آن خودداری خواهیم کرد.

در واقع مسئله‌ی همیلتونی k موضعی، تعمیم کوانتمی مسئله‌ی CSP $_k$ است. در ادامه نشان خواهیم داد که SAT-3 در واقع مسئله‌ی همیلتونی k موضعی، تعمیم کوانتمی مسئله‌ی CSP_k است. در ادامه نشان خواهیم داد که LH-3 مسئله‌ای سخت برای کلاس \mathcal{NP} است.

قضیه ۴.۵.

$$\text{SAT} \leq_m^p \text{LH}. \quad (4.5)$$

اثبات. می‌دانیم هر فرمول CNF-3 فرمولی مانند $\varphi(x_1, \dots, x_n) = \bigwedge_i c_i(x_{i1}, x_{i2}, x_{i3})$ است به طوری که c_i ، یک ترکیب فصلی مانند $A_{i1} \vee A_{i2} \vee A_{i3}$ است که هر A_{ij} برابر با x_{ij} یا $\neg x_{ij}$ است. برای هر $c_i(x_{i1}, x_{i2}, x_{i3})$ همیلتونی $H_{i1, i2, i3}^{(i)}$ را به این صورت تعریف کنید: $(\mathbb{C}^2)^{\otimes n} \rightarrow (\mathbb{C}^2)^{\otimes n}$

$$H_{i1, i2, i3}^{(i)} = \sum_{\substack{x \in \{0, 1\}^n \\ s.t. \ c(x)=0}} |x\rangle \langle x| \quad (4.5)$$

(سمت راست تساوی بالا به صورت مختصر نوشته شده است و باید این گونه تعبیر شود: $H_{i1, i2, i3}^{(i)}$ تنها روی کیویت‌های i_1, i_2, i_3 به صورت نابدیهی عمل می‌کند و عمل آن روی این سه کیویت نیز مطابق با نگاشت $|x\rangle \langle x|$ است). حال تعریف کنید $H = \sum_i H_{i1, i2, i3}^{(i)}$. روش است که H یک همیلتونی 3 موضعی است و توصیف آن را می‌توان از روی توصیف φ و در زمان چندجمله‌ای (بر حسب طول توصیف φ) به دست آورد.

اکنون توجه کنید که اگر $\text{SAT-3} \in \varphi$ ، در این صورت ارزش‌گذاری $x \in \{0, 1\}^n$ به متغیرهای φ وجود دارد به طوری که $\lambda_{min}(H) \leq 0$. به طور خاص، برای هر i , $\langle x | H | x \rangle = 0$ ، بنابراین $c_i(x_{i1}, x_{i2}, x_{i3}) = 0$ ، یا معادلاً $\langle x | H | x \rangle = 0$. از سوی دیگر، اگر $\text{SAT-3} \notin \varphi$ ، در این صورت برای هر ارزش‌گذاری $x \in \{0, 1\}^n$ به متغیرهای φ , $\langle x | H | x \rangle = 0$ ؛ یا معادلاً $\langle x | H | x \rangle = 1$. بنابراین، $c_i(x_{i1}, x_{i2}, x_{i3}) = 1$.

$$\langle x | H | x \rangle = \sum_{x \in \{0, 1\}^n} \langle x | H_{i1, i2, i3}^{(i)} | x \rangle \geq 1. \quad (4.5)$$

حال برای هر $\psi \in (\mathbb{C}^2)^{\otimes n}$ که $|\psi\rangle = \sum_{x \in \{0, 1\}^n} \alpha_x |x\rangle$ داریم:

$$\langle \psi | H | \psi \rangle = \sum_{x \in \{0, 1\}^n} |\alpha_x|^2 \langle x | H | x \rangle \geq \sum_{x \in \{0, 1\}^n} |\alpha_x|^2 = 1 \quad (5.5)$$

بنابراین، در این حالت $\lambda_{min}(H) \geq 1$. \square

۲.۵. قضیه‌ی کوک-لوین کوانتمی. قضیه‌ی کوک-لوین را می‌توان یکی از عمیق‌ترین نتایج به دست آمده در نظریه‌ی پیچیدگی محاسبات کلاسیک دانست. این قضیه — که مستقل‌آتاً توسط لئونید لوین [۱۷] و استفن کوک [۱۸] اثبات شده است — از جهات متعددی حائز اهمیت است:

• نخست آن که این قضیه آغازگر مسیری طولانی در جهت یافتن مسائل \mathcal{NP} -کامل به امید پاسخ‌دادن به مسئله‌ی \mathcal{P} vs. \mathcal{NP} بوده است.

• اگر روح محاسبه را «یافتن توصیف‌هایی متناهی برای مجموعه‌های نامتناهی» بدانیم، قضیه‌ی کوک-لوین ارتباطی میان دو روش متفاوت برای توصیف متناهی مجموعه‌ها — یکی الگوریتم‌ها و دیگری عبارات منطقی — برقرار می‌کند.

- این قضیه ناظر به یکی از اساسی‌ترین ویژگی‌های مفهوم محاسبه، یعنی موضعی بودن آن است. حقیقتی که در پس اثبات قضیه‌ی کوک-لوین نهفته است آن است که هر محاسبه‌ای، عبارت است از دنباله‌ای از تغییرات موضعی بر پیکربندی ماشین محاسبه که نهایتاً به یک پیکربندی مطلوب ختم شود.

در روندی مشابه با محاسبات کلاسیک، در این بخش نشان خواهیم داد که کلاس QMA نیز مسئله‌ای کامل دارد؛ مسئله‌ای که در شاخه‌های دیگری از علم نیز بامعنى و قابل مطالعه است. به علاوه، خواهیم دید سخت بودن این مسئله برای کلاس QMA اطلاعاتی را در مورد یکی از اولین انگیزه‌های مطالعه‌ی محاسبات کواتومی، یعنی امید برای شبیه‌سازی کارای سیستم‌های فیزیکی کواتومی به وسیله‌ی کامپیوترهای کواتومی، فراهم خواهد کرد.

در بخش قبل دیدیم که مسئله‌ی $k\text{-LH}$ تعیینی از مسئله‌ی $k\text{-CSP}$ است، و بدین ترتیب کاندیدای طبیعی ما برای مسئله‌ای کامل در مسئله‌ی $k\text{-LH}$ خواهد بود. برای اثبات کامل بودن، باید نشان دهیم H -مسئله‌ای در QMA است، و نیز این مسئله برای کلاس QMA مسئله‌ای سخت است. اثبات مورد اول نسبتاً ساده است و نخست به آن می‌پردازیم.

قضیه ۵.۵. برای هر $k \in \mathcal{O}(\log n)$ و هر پارامتر چندجمله‌ای p ، $k\text{-LH} \in \text{QMA}$. [۱]

اثبات. برای آن که نشان دهیم $k\text{-LH} \in \text{QMA}$ ، الگوریتم تصدیق کردنی برای آرتور ارائه می‌دهیم که چنان‌چه برای ورودی $(H, \alpha, \beta) \in k\text{-LH}$ ، اثباته باشیم $(H, \alpha, \beta) \in k\text{-LH}$ ، اثباتی وجود داشته باشد که مرلین بتواند برای آرتور ارسال کرده و آرتور در زمان چندجمله‌ای آن را با احتمال بالایی تصدیق کند، و در حالتی که $(H, \alpha, \beta) \notin k\text{-LH}$ ، چنین اثباتی وجود نداشته باشد. نشان خواهیم داد که الگوریتم زیر این ویژگی را دارد.

ابتدا فرض کنید $H = \sum_{i=1}^m H^{(i)}$. برای سادگی، فرض کنید هر یک از جملات موضعی $H^{(i)}$ ، عملگرهایی مثبت معین با مقادیر ویژه‌ی بین 0 و 1 هستند. در این صورت، فرض کنید تجزیه‌ی طیفی هر $H^{(i)}$ به صورت $H^{(i)} = \sum_s \lambda_s |s\rangle\langle s|$ باشد.

الگوریتم

مرلین: به عنوان اثبات، حالت پایه‌ی H (بردار $|\psi\rangle$) را برای آرتور ارسال می‌کند.
آرتور:

- حالت ارسال شده از طرف مرلین را در کنار یک کیویت اضافی که در حالت $|0\rangle$ آماده‌سازی شده است قرار می‌دهد. این کیویت اضافه را «کیویت جواب» می‌نامیم.
- به طور تصادفی و با احتمال یکسان، عدد i را بین 1 تا m انتخاب می‌کند.
- نگاشت W_i را روی $|\psi\rangle$ $\otimes |0\rangle$ اعمال می‌کند و کیویت جواب را در پایه‌ی محاسباتی اندازه‌گیری می‌کند. چنان‌چه حالت سیستم پس از اندازه‌گیری $|0\rangle$ باشد، اثبات را رد می‌کند و در غیر این صورت، می‌پذیرد.

برای هر i ، عمل W_i روی مجموعه‌ی مستقل خطی $\{|s\rangle\otimes|0\rangle\}$ به صورت زیر تعریف شده است:

$$W_i(|s\rangle\otimes|0\rangle) = \sqrt{\lambda_s}|s\rangle\otimes|0\rangle + \sqrt{1-\lambda_s}|s\rangle\otimes|1\rangle \quad (6.5)$$

نخست توجه کنید که با توجه به این که $\text{Artor}(|\psi\rangle) \in \mathcal{O}(\log n)$ ، آرتور می‌تواند نگاشت W_i را در زمان چندجمله‌ای اعمال کند. حال نشان می‌دهیم برای هر حالت $|\psi\rangle$ که توسط مرلین ارسال شده باشد، احتمال پذیرفته شدن اثبات توسط آرتور $\text{Artor}(|\psi\rangle) = 1 - \frac{1}{m}$ است.

فرض کنید برای هر i ، $W_i(|\psi\rangle\otimes|0\rangle) = \sum_s \alpha_s |s\rangle\otimes|\psi\rangle$. در این صورت

$$W_i(|\psi\rangle\otimes|0\rangle) = W_i\left(\sum_s \alpha_s |s\rangle\otimes|0\rangle\right) \quad (7.5)$$

$$= \sum_s \alpha_s W_i(|s\rangle\otimes|0\rangle) \quad (8.5)$$

$$= \sum_s \alpha_s (\sqrt{\lambda_s}|s\rangle\otimes|0\rangle + \sqrt{1-\lambda_s}|s\rangle\otimes|1\rangle). \quad (9.5)$$

بنابراین احتمال این که آرتور پس از اعمال W_i و اندازه‌گیری کیویت جواب در پایه‌ی محاسباتی اثبات را بپذیرد برابر است با:

$$\sum_s |\alpha_s|^2 (1 - \lambda_s) = \sum_s |\alpha_s|^2 - \sum_s \lambda_s |\alpha_s|^2 = 1 - \sum_s \lambda_s |\alpha_s|^2 = 1 - \langle \psi | H^{(i)} | \psi \rangle$$

$$\sum_s |\alpha_s|^2 (1 - \lambda_s) = \sum_s |\alpha_s|^2 - \sum_s \lambda_s |\alpha_s|^2 = 1 - \sum_s \lambda_s |\alpha_s|^2 = 1 - \langle \psi | H^{(i)} | \psi \rangle \quad (10.5)$$

حال توجه کنید که احتمال پذیرفته شدن اثبات $\langle \psi | H^{(i)} | \psi \rangle$ توسط آرتور برابر با $1 - \frac{1}{m} \langle \psi | H | \psi \rangle$ است. بنابراین اگر $(H, \alpha, \beta) \in k\text{-LH}$ ، مارلین می‌تواند حالت پایه‌ی H را برای آرتور ارسال کند و آرتور با احتمال $1 - \frac{\alpha}{m} \geq 1 - \frac{1}{m} \langle \psi | H | \psi \rangle$ آن را می‌پذیرد. از طرفی اگر $(H, \alpha, \beta) \notin k\text{-LH}$ ، هر حالتی که از طرف مارلین ارسال شود، با احتمال $1 - \frac{\beta}{m} \leq 1 - \frac{1}{m} \langle \psi | H | \psi \rangle$ پذیرفته خواهد شد. نهایتاً با توجه به این که اختلاف ثوابت درستی و تمامیت، بزرگتر از $\frac{1}{p(n)}$ است، حکم از قضیه‌ی ۳.۴ تیجه خواهد شد. \square

پیش از آن که به اثبات سختبودن $H\text{-LH}$ برای کلاس QMA پردازیم، حالی از فایده نیست که روند اثبات سختبودن SAT برای NP را در قضیه‌ی کوک-لوین کلاسیک مرور کنیم.

فرض کنید $\{0, 1\}^L \subseteq \{0, 1\}^n$ زبانی عضو کلاس NP باشد. هدف ما این است که تحویلی با زمان چندجمله‌ای از این زبان به SAT بسازیم. به عبارت دیگر، برای هر رودی $\{0, 1\}^z \in z$ ، در زمان چندجمله‌ای بر حسب طول z ، فرمولی بولی ϕ_z بولی مانند ϕ بیابیم به طوری که

$$z \in L \iff \phi_z \in SAT.$$

بدین منظور، توجه کنید که چون $L \in NP$ ، پس ماشین تورینگ قطعی $M = (Q, \Sigma, \Gamma, \delta, q_0, q_{accept})$ با زمان چندجمله‌ای وجود دارد که تصدیق‌کننده‌ی زبان L است. می‌دانیم

$$z \in L \iff \exists y \in \{0, 1\}^{\text{poly}(|z|)} M(z, y) = 1.$$

ایده آن است که عبارت بولی ϕ را طوری بیابیم که عملکرد M روی z را شبیه‌سازی کند. در این شبیه‌سازی باید موارد زیر لحاظ شود:

- مقداردهی اولیه‌ی نوار: پیش از شروع به کار ماشین M ، ورودی مناسب روی نوار نوشته شده باشد.
- انتقال پیکربندی‌ها: هر پیکربندی مطابق با قانون انتقال δ از پیکربندی قبل به دست آمده باشد.
- خروجی: وقتی محاسبه پایان می‌یابد، $z \in L$ اگر و تنها اگر حالت نهایی q_{accept} باشد.

در اثبات قضیه‌ی کوک-لوین در محاسبات کلاسیک، برای چک‌کردن هر یک از موارد فوق، فرمولی بولی با طول حداقل چندجمله‌ای بر حسب طول z ساخته می‌شود و نهایتاً عطف این فرمول‌ها فرمول ϕ را در اختیار ما قرار می‌دهد. خواننده‌ی علاقه‌مند می‌تواند جزئیات اثبات را در [۱۹] دنبال کند. در محاسبات کواتنومی نیز، در روندی مشابه با حالت کلاسیک، سه مورد فوق را با همیلتونی‌های مناسبی کد خواهیم کرد.

قضیه ۶.۵. برای هر $k \geq 5$ تحت تحويل چندبهیک با زمان چندجمله‌ای، مسئله‌ای سخت برای کلاس QMA است.

اثبات. فرض کنید $\Pi = (\Pi_{Yes}, \Pi_{No})$ مسئله‌ی قراردادی دلخواهی در کلاس QMA باشد. طبق تعریف، می‌دانیم مدار کواتنومی تصدیق‌کننده‌ی V برای این مسئله وجود دارد. فرض کنید $V = V_T V_{T-1} \dots V_1$ ، که V_i ها گیت‌های ۱-موضعی یا ۲-موضعی هستند و $T \in \mathcal{O}(\text{poly}(n))$ است. تکنیکی که برای اثبات قضیه‌ی کوک-لوین کواتنومی استفاده می‌شود، این است که برای هر ورودی مانند $|x\rangle$ ، تاریخچه‌ی محاسبه‌ی $|x\rangle$ توسط مدار V را در یک استیت کواتنومی کد می‌کنیم، و سپس از همیلتونی‌های موضعی برای چک‌کردن این که مقداردهی اولیه‌ی رجیسترها درست است، انتقال پیکربندی‌ها به درستی انجام می‌شود و خروجی مدار همان خروجی مطلوب است، بهره می‌گیریم. در این اثبات کیتائاف برای کدکردن مفهوم زمان، بر اساس ایده‌ای از فاینمن، از یک رجیستر اضافه استفاده کرده و تاریخچه‌ی محاسبه را به صورت $|\psi_{hist}\rangle = \sum_t |\psi_t\rangle_{in, pr, an} |t\rangle_C$ کد کرده است که در آن

$$|\psi_t\rangle = (V_T V_{T-1} \dots V_1 |x\rangle_{in} |\psi\rangle_{pr} |0 \dots 0\rangle_{an}) |0 \dots 0\rangle_C,$$

و از پانویس C برای نمایش رجیستر کلاک استفاده شده است.

در ادامه تلاش می‌کنیم همیلتونی H را طوری طراحی کنیم که حالت کوانتومی فوق، حالت پایه‌ی آن باشد.

- برای آن‌که مقدار دهی اولیه‌ی رجیسترها را به آن‌چه مطلوب ماست ($|x\rangle$ در رجیستر ورودی و $|0\cdots 0\rangle$ در رجیستر کمکی) مقید کنیم، همیلتونی H_{in} را به صورت زیر تعریف می‌کنیم:

$$\begin{aligned} H_{in} = & (\mathbb{I} - |x\rangle\langle x|)_{in} \otimes \mathbb{I}_{pr} \otimes \mathbb{I}_{an} \otimes |0\cdots 0\rangle\langle 0\cdots 0|_C \\ & + \mathbb{I}_{in} \otimes \mathbb{I}_{pr} \otimes (\mathbb{I} - |0\cdots 0\rangle\langle 0\cdots 0|)_{an} \otimes |0\cdots 0\rangle\langle 0\cdots 0|_C \end{aligned}$$

روشن است که H_{in} مثبت نیمه‌معین است و به علاوه برای حالت $|\phi(y)\rangle$ که برای هر $y \in \{0, 1\}^{q(n)}$ ، به صورت $.y = 0^{q(n)}|\phi(y)\rangle = |x\rangle_{in}|\psi\rangle_{pr}|y\rangle_{an}|0\cdots 0\rangle_C$ تعریف شده است داریم

- برای آن‌که در زمان $t = T$ ، پس از اندازه‌گیری بیت خروجی (که در اینجا فرض می‌کنیم کیویت اول رجیستر کمکی است). خروجی مدار برابر ۱ باشد، همیلتونی H_{out} را به صورت زیر تعریف می‌کنیم:

$$H_{out} = \mathbb{I}_{in} \otimes \mathbb{I}_{pr} \otimes (|0\rangle\langle 0|)_{an} \otimes |T\rangle\langle T|_C.$$

در اینجا مقصود از $(|0\rangle\langle 0|)_{an}$ نگاشتی است که کیویت اول رجیستر کمکی را روی زیرفضای تولید شده توسط $|0\rangle$ می‌افکند و اثر آن روی باقی کیویت‌های رجیستر کمکی، همانی است.

- برای چک کردن انتقال درست پیکربندی‌ها، همیلتونی H_{prop} را به صورت زیر تعریف می‌کنیم:

$$\begin{aligned} H_{prop} = & \sum_{t=0}^{T-1} -V_{t+1} \otimes |t+1\rangle\langle t| - V_{t+1}^\dagger \otimes |t\rangle\langle t+1| \\ & + \mathbb{I}_{in, pr, an} \otimes |t\rangle\langle t| + \mathbb{I}_{in, pr, an} \otimes |t+1\rangle\langle t+1| \end{aligned}$$

می‌توان دید که برای هر حالت $|\phi\rangle = \frac{1}{\sqrt{T+1}} \sum_{t=0}^T (V_t V_{t-1} \dots V_1 |\eta\rangle_{in, pr, an}) \otimes |t\rangle_C$

$$\begin{aligned} H_{prop}|\phi\rangle = & \sum_{t=0}^{T-1} -(V_{t+1} V_t \dots V_1 |\eta\rangle) \otimes |t+1\rangle \\ & + \sum_{t=0}^{T-1} -(V_{t+1}^\dagger V_{t+1} V_t \dots V_1 |\eta\rangle) \otimes |t\rangle \\ & + \sum_{t=0}^{T-1} (V_{t+1} V_t \dots V_1 |\eta\rangle) \otimes |t+1\rangle \\ & + \sum_{t=0}^{T-1} (V_{t+1}^\dagger V_{t+1} V_t \dots V_1 |\eta\rangle) \otimes |t\rangle = 0 \end{aligned}$$

حال تعریف کنید: $H = H_{in} + H_{out} + H_{prop}$ عضو \mathcal{LH}_k است. در ادامه α و β را به نحو مناسبی انتخاب خواهیم کرد تا (H, α, β) باشد.

نخست فرض کنید $x \in \Pi_{Yes}$ است. در این صورت اثبات $|\psi\rangle \in (\mathbb{C}^2)^{\otimes p(n)}$ موجود است که با احتمال حداقل $\epsilon - 1$ توسط مدار تصدیق‌کننده‌ی V پذیرفته می‌شود. می‌خواهیم در این حالت همیلتونی ساخته شده به صورت بالا، حالت پایه‌ای کمتر از α داشته باشد. توجه کنید که

$$\begin{aligned} \langle \psi_{hist} | H | \psi_{hist} \rangle &= \langle \psi_{hist} | H_{in} | \psi_{hist} \rangle + \langle \psi_{hist} | H_{out} | \psi_{hist} \rangle + \langle \psi_{hist} | H_{prop} | \psi_{hist} \rangle \\ &= 0 + 0 + \frac{1}{T+1} \Pr[|\psi\rangle \text{ را پذیرد}, V] \leq \frac{\epsilon}{T+1} \end{aligned}$$

بنابراین کافی است قرار دهیم $\alpha = \frac{\epsilon}{T+1}$.

حال فرض کنید $x \in \Pi_{No}$ است. در این صورت برای هر اثبات $|\psi\rangle \in (\mathbb{C}^2)^{\otimes p(n)}$ ، احتمال پذیرفته شدن $|\psi\rangle$ توسط مدار V

حداکثر ϵ است. کیتائف با استفاده از لمی که به لم هندسی^۱ مشهور است نشان داد که در این حالت، $\lambda_{min}(H) \geq \frac{\pi^2(1-\sqrt{\epsilon})}{2(T+1)^3}$ در اینجا از بیان اثبات لم هندسی خودداری می‌کنیم. خواسته‌ی علاقمند می‌تواند اثباتی برای این لم را در [۱] بباید. با استفاده از آن‌چه در بالا به دست آمد، کافی است قرار دهیم $\frac{\pi^2(1-\sqrt{\epsilon})}{2(T+1)^3} \beta$. به این ترتیب، تحويل مورد نظر یافت می‌شود.

تنها مشکلی که وجود دارد آن است که همیلتونی‌های معرفی شده، همیلتونی‌های ۵-موضعی نیستند. با کمی دقت می‌توان دریافت که همیلتونی‌های ارائه شده روی تمام کیویت‌های رجیستر کلاک به طور نابدیهی عمل می‌کنند، و می‌دانیم رجیستر کلاک متتشکل از $O(\log n)$ کیویت است. به علاوه، در همیلتونی H_{in} می‌توان دید که عملگرهای $|x\rangle\langle x|$ و $|\psi\rangle\langle\psi|$ را با $\sum_{i=1}^{q(n)} |1\rangle\langle 1|_{an_i}$ ۵-موضعی نیستند. رفع مشکل در حالت دوم ساده است. مثلاً می‌توان $|\psi\rangle\langle\psi|$ را با $|\psi\rangle\langle\psi|_{an_i}$ کرد، که مقصود از $|\psi\rangle\langle\psi|_{an_i}$ نگاشتی است که تنها کیویت α رجیستر کمکی را روی زیرفضای تولید شده توسط $|\psi\rangle$ می‌افکند و بر روی باقی کیویت‌ها به صورت بدیهی عمل می‌کند. مشاهده می‌شود که حالت پایه‌ی همیلتونی مزبور با چنین تعویضی تغییر نمی‌کند. برای رفع مشکل رجیستر کلاک، باید از روش دیگری استفاده کرد. فرض کنید به جای آن که کلاک را به صورت دودویی کد کنیم، این کار را به طور یک‌یکی^۲ انجام دهیم. به عبارت دیگر، فرض کنید نمایش لحظه‌ی t در این رجیستر، به صورت $|\psi_{t+T-t}\rangle$ باشد. در این صورت چک کردن محتوای رجیستر کلاک با چک کردن تنها سه کیویت از آن امکان‌پذیر است (کیویت‌های $|\psi_t\rangle, |\psi_{t+T-t}\rangle$ و $|\psi_{t+T}\rangle$).

به این ترتیب با توجه به این که گیت‌های V_i حداکثر ۲-موضعی هستند و حالت رجیستر کلاک نیز با عملگری حداکثر ۳-موضعی چک می‌شود، همیلتونی‌های ارائه شده ۵-موضعی خواهند بود. تنها نکته‌ای که باید به آن توجه کرد این است که در بالا تلویحاً فرض شده است که کدینگ رجیستر کلاک $|\psi\rangle$ به صورت $|\psi\rangle\langle\psi|$ است. برای این که این التزام را ایجاد کنیم، همیلتونی دیگری را نیز به صورت $(\sum_{i=1}^{T-1} |1\rangle\langle 1|_{C_i} \otimes |\psi\rangle\langle\psi|_{C_i}) \otimes |\psi\rangle\langle\psi|_{in,pr,out}$ تعریف می‌کنیم. به این ترتیب همیلتونی جدید $H_C = H_{in} + H_{out} + H_{prop} + H_C$ همیلتونی ۵-موضعی مطلوب ما خواهد بود. می‌توان نشان داد که با اعمال تغییرات فوق، $|\psi_{hist}\rangle$ همچنان حالت پایه‌ی همیلتونی جدید باقی می‌ماند؛ و ثوابت درستی و تمامیتی که انتخاب کرده بودیم نیز همچنان کار می‌کنند. \square

ملاحظه ۷.۵. نخستین اثبات قضیه‌ی کوک-لوین کوانتومی، که تحويلی از هر مسئله‌ی QMA به LH ارائه می‌کند، منسوب به کیتائف است و در [۱] بیان شده است. رگف و کمپ در [۲۰] این نتیجه را بهبود بخشیدند و نشان دادند LH -۳ مسئله‌ای کامل برای QMA است. نهایتاً^۳ کمپ، رگف و کیتائف در [۲۱] نشان دادند که LH -۲ نیز مسئله‌ای QMA -کامل است.

نتیجه ۸.۵. نتیجه‌ی سرراست قضیه‌ی کوک-لوین کوانتومی این است که با فرض $QMA \neq BQP$ ، مسئله‌ی همیلتونی موضعی را نمی‌توان حتی با کامپیوترهای کوانتومی به صورت کارایی حل کرد. همان‌گونه که پیشتر بیان شد، مسئله‌ی همیلتونی‌های موضعی حالت خاصی از مسئله‌ای کلی‌تر، یعنی مسئله‌ی شبیه‌سازی سیستم‌های کوانتومی است. در واقع، قضیه‌ی کوک-لوین کوانتومی موجد این مطلب است که مطالعه‌ی سیستم‌های کوانتومی می‌تواند از نظر محاسباتی، حتی در حضور کامپیوترهای کوانتومی، سخت باشد.

نتیجه ۹.۵. نتیجه‌ای دیگر از سختی LH - k برای کلاس QMA این است که با فرض $QMA \neq NP$ ، همیلتونی‌هایی وجود دارند که حالت پایه‌ی آن‌ها توصیف کارایی کلاسیک ندارد. می‌توان نشان داد که این مطلب بدان معنی است که حالت پایه‌ی چنین همیلتونی‌هایی قویاً درهم‌تیشه است^۴. این نتیجه، سازگار با این مطلب است که ماده در دمای نزدیک به صفر خواصی مانند ابرشارگی^۵ و ابررسانایی^۶ را از خود نشان می‌دهد که برخاسته از وجود نوعی درهم‌تیشه قوی در حالت آن است.

۳.۵. اثبات‌های قابل بررسی احتمالاتی کوانتومی.

۱۳.۵. قضیه‌ی PCP کلاسیک. فرض کنید قرار است درستی اثبات یک قضیه را بررسی کنید. متاسفانه اثبات‌ها می‌توانند بسیار طولانی باشند. مثلاً اثبات قضیه‌ی لافورگ^۷ که در سال ۲۰۰۰ در جریان برنامه‌ی لنگلندر^۸ توسط لورن لافورگ ارائه شد،

^۱geometric lemma

^۲unary

^۳تاکنون اندازه‌های متعددی برای کمی‌سازی درهم‌تیشه و مقایسه‌ی آن توسعه یافته است.

^۴superfluidity

^۵superconductivity

^۶Langlands program

چیزی در حدود 600 صفحه است! بسیار جالب بود اگر می‌توانستید از کل اثبات‌ها یک خط را بخوانید و مطمئن شوید که اثبات درست است یا نه، و به این ترتیب می‌توانستید در زمان نیز صرفه‌جویی کنید. اگرچه برای داوران ژورنال‌ها حتی تصویر چنین خیالی نیز دلپذیر است، با این حال عجیب به نظر می‌رسد و بعيد است که بتوان آن را عملی کرد. در واقع همواره این امکان وجود دارد که ایراد اثبات در آن قسمتی باشد که شما آن را نخوانده‌اید (حتی اگر قسمت «کوچکی» را چک نکرده باشید): چه برسد به این که برای تحقیق درستی اثبات، فقط یک خط آن را درنظر بگیرید.

با این حال بگذارید توجه خود را به جای هر اثباتی، معطوف به اثبات‌هایی کنیم که برای مصداقی از مسئله‌ای در کلاس \mathcal{NP} داده می‌شود. یک نگرش به قضیه‌ی PCP این را بیان می‌کند که راهی وجود دارد که بتوان چنین اثباتی را به گونه‌ای بازنویسی کرد که تصدیق‌کننده، که قرار است در زمان چندجمله‌ای اثبات را تصدیق کند، بدون نیاز به خواندن کل اثبات و تنها با خواندن تکه‌ی کوچکی از آن، که به صورت تصادفی انتخاب می‌شود، با احتمال بالا بتواند اطمینان یابد که اثبات درست است یا غلط. این امکان، تعریف جدیدی برای کلاس \mathcal{NP} بر حسب نوعی از سیستم‌های اثبات، که به آن‌ها اثبات‌های قابل بررسی احتمالاتی^۱ می‌گویند، در اختیار ما می‌گذارد.

به قضیه‌ی PCP می‌توان از منظری دیگر نیز نگریست. همان‌طور که می‌دانیم، اگر $\mathcal{P} \neq \mathcal{NP}$ باشد، مسائلی وجود خواهد داشت که هیچ راه حل دقیق چندجمله‌ای برایش وجود ندارد. با این حال، بعضی از مسائل بهینه‌سازی \mathcal{NP} -سخت را می‌توان با الگوریتم‌های تقریبی چندجمله‌ای، در حد یک ضریب تقریب کوچک حل کرد؛ به این معنی که الگوریتمی چندجمله‌ای وجود دارد که جوابی که تولید می‌کند، مثلاً از دو برابر جواب بهینه بدتر نیست. برای بسیاری از چنین مسائلی، این ضریب تقریب حدی دارد. قضیه‌ی PCP روشی برای اثبات وجود چنین محدودیت‌هایی در تقریب‌زدن فراهم می‌کند و می‌تواند نشان دهد که برای بسیاری از مسائل، الگوریتم‌های تقریبی فعلی بهینه هستند و ضریب تقریب را نمی‌توان از آنچه که تا کنون به دست آورده‌ایم بهتر کرد. این‌ها نتایجی هستند که ظاهراً بدون قضیه‌ی PCP نمی‌توانستیم به آن‌ها دست پیدا کنیم.

در ادامه‌ی این زیربخش پس از بیان برخی مقدمات، دو صورت‌بندی معادل را برای قضیه‌ی PCP بیان خواهیم کرد، و معادل بودن این دو صورت‌بندی را نشان خواهیم داد.

نمادگذاری ۱۰.۵. فرض کنید \mathfrak{P} یک مسئله‌ی بهینه‌سازی باشد. در این صورت برای یک ورودی مسئله مانند I ، مقدار بهینه‌ی مسئله را با $OPT(I)$ نمایش می‌دهیم. اگر A یک الگوریتم برای حل مسئله باشد، مقصود از $(I)A$ پاسخی است که با ورودی I ، توسط الگوریتم A تولید می‌شود.

تعریف ۱۱.۵. الگوریتم A را برای مسئله‌ی کمینه‌سازی \mathfrak{P} یک الگوریتم $(\alpha(n))$ -تقریب گوییم $(1 \geq \alpha(n))$ ، هرگاه برای ورودی I از \mathfrak{P} ، به همین ترتیب برای یک مسئله‌ی بیشینه‌سازی \mathfrak{P} ، یک الگوریتم $(\alpha(n))$ -تقریب است $(1 \leq \alpha(n))$ ، هرگاه $(I)A$ را ضریب تقریب می‌نامیم.

مثال ۱۲.۵. مسئله‌ی MAX-3SAT را به عنوان مسئله‌ی یافتن بیشترین تعداد پرانتزهای ممکن در یک فرمول CNF-۳ که به طور همزمان قابل ارضاشدن هستند، در نظر بگیرید. الگوریتم زیر، الگوریتمی $\frac{1}{4}$ -تقریب برای این مسئله است: الگوریتم را به صورت حریصانه تعریف می‌کنیم. به این صورت که در هر مرحله یک متغیر را انتخاب می‌کنیم، مقداری از $\{true, false\}$ به آن نسبت می‌دهیم به طوری که بیشترین تعداد پرانتز را برآورده کند. بعد از آن، عباراتی که برآورده شده‌اند را از مسئله حذف کرده و به سراغ متغیر بعدی می‌رویم. همین کار را انجام می‌دهیم تا زمانی که تمام متغیرها مقداردهی شوند. با توجه به نحوه‌ی مقداردهی به هر متغیر، واضح است که در هر قدم الگوریتم، اگر برآورده شدن t عبارت مشخص شود، حداقل $\frac{t}{4}$ آنها برآورده می‌شوند و بنابراین در نهایت حداقل $\frac{t}{4}$ کل عبارات اولیه برآورده شده‌اند. بنابراین الگوریتم پیشنهادشده یک الگوریتم $\frac{1}{4}$ -تقریب برای مسئله‌ی MAX-3SAT است.

تعریف ۱۳.۵. برای مسئله‌ی کمینه‌سازی \mathfrak{P} ، مسئله‌ی $Gap_{h(n),g(n)}$ یک مسئله‌ی قراردادی است که در آن برای ورودی I با طول n :

$$\begin{aligned} OPT(I) &\leq h(n), I \in Gap_{\mathfrak{P}_{h(n),g(n)} \text{ Yes}} & \bullet \\ OPT(I) &\geq g(n)h(n), I \in Gap_{\mathfrak{P}_{h(n),g(n)} \text{ No}} & \bullet \end{aligned}$$

¹Probabilistically checkable proofs

به طور مشابه، تعریف بالا را می‌توان برای مسائلهای بیشینه‌سازی نیز انجام داد.
همیت تعریف بالا آن جاست که می‌توان از آن برای نشان دادن سختی حل تقریبی یک مسائلهای بهینه‌سازی بهره گرفت. در واقع برای آن که نشان دهیم حل کردن \mathfrak{P} با ضریب تقریب (n) در زمان چندجمله‌ای مسائلهای سخت است، کافی است نشان دهیم \mathcal{NP} -سخت است. گزاره‌ی زیر چرایی این امر را بیان می‌کند.

گزاره ۱۴.۵. فرض کنید برای یک مسائلهای کمینه‌سازی \mathfrak{P} ، مسائلهای \mathcal{NP} -سخت است. در این صورت الگوریتمی (n) -تقریب با زمان چندجمله‌ای برای \mathfrak{P} وجود ندارد، مگر آن که $\mathcal{P} = \mathcal{NP}$ باشد.

اثبات. با فرض وجود الگوریتمی (n) -تقریب با زمان چندجمله‌ای برای \mathfrak{P} (مثلًا A)، می‌توانیم هر مسائلهای \mathcal{NP} -کامل مانند L را در زمان چندجمله‌ای حل کنیم. به این ترتیب که برای هر $x \in \{0, 1\}^*$ ، ابتدا با تحویل چندجمله‌ای موجود از L به \mathfrak{P} ، یک ورودی مانند I_x را برای مسائلهای \mathfrak{P} ، بدست می‌آوریم. حال، الگوریتم (n) -تقریب موجود برای حل \mathfrak{P} را روی I_x اجرا می‌کنیم. توجه کنید که:

- $I_x \in \text{Gap}_{\mathfrak{P}, h(n), \alpha(n)}_{Y_{es}} \implies \text{OPT}(I_x) \leq h(n) \implies A(I_x) \leq \alpha(n)h(n)$
- $I_x \in \text{Gap}_{\mathfrak{P}, h(n), \alpha(n)}_{N_o} \implies \text{OPT}(I_x) \geq \alpha(n)h(n) \implies A(I_x) \geq \alpha(n)h(n)$

بنابراین با توجه به مقدار $A(I_x)$ ، می‌توان $\text{Gap}_{\mathfrak{P}, h(n), \alpha(n)}$ و در نتیجه L را تصمیم گرفت.

□

حال قادر هستیم اولین صورت قضیه‌ی PCP را بیان کنیم.

قضیه ۱۵.۵ (قضیه‌ی PCP: سختی تقریب). ثابت $1 < \rho$ وجود دارد به نحوی که $\text{Gap}_{\text{MAX-3SAT}_1, \rho}$ -سخت است. [۲۲]

در اینجا شایان ذکر است که نتیجه‌ی فوق را می‌توان بربسیاری دیگر از مسائل بهینه‌سازی \mathcal{NP} -کامل پیدا کرد. بدین منظور کافی است تعريفمان از تحویل را به صورت زیر تغییر دهیم:

تعريف ۱۶.۵. فرض کنید \mathfrak{P} و \mathfrak{P}' دو مسائلهای بیشینه‌سازی باشند. یک تحویل حافظ فاصله^۱ از \mathfrak{P} به \mathfrak{P}' با پارامترهای $(g(n), g'(n'), h(n), h'(n'))$ ، که $1 \leq g(n), g'(n'), h(n), h'(n') \leq 1$ ، عبارت است از الگوریتمی که هر ورودی \mathfrak{P} مانند I را، که طول I برابر n است، به یک ورودی برای \mathfrak{P}' مانند I' نظری می‌کند، که طول I' برابر با n' است، چنان‌که:

- اگر $\text{OPT}(I') \geq h'(n')$ ، آنگاه $\text{OPT}(I) \geq h(n)$.
- اگر $\text{OPT}(I') \leq g'(n')h'(n')$ ، آنگاه $\text{OPT}(I) \leq g(n)h(n)$.

به سادگی می‌توان دید که اگر تحویلی حافظ فاصله و چندجمله‌ای از \mathfrak{P} به \mathfrak{P}' با پارامترهای $(g(n), g'(n'), h(n), h'(n'))$ موجود باشد، در این صورت اگر $\text{Gap}_{\mathfrak{P}, h(n), g(n)}$ -کامل باشد، $\text{Gap}_{\mathfrak{P}', h'(n'), g'(n')}$ نیز چنین است. به این ترتیب، می‌توان با استفاده از قضیه‌ی ۱۵.۵، سختی تقریب مسائل بیشتری را اثبات کرد.

حال به صورت‌بندی دوم قضیه‌ی PCP می‌پردازم.

تعريف ۱۷.۵. یک تصدیق‌کننده‌ی $((r(n), q(n))$ -محدود، تصدیق‌کننده‌ای چندجمله‌ای است که به استفاده از حداکثر b بیت تصادفی محدود است، و قادر است تنها با استفاده از برآمد این بیت‌های رندوم، $r(n)$ حداکثر $q(n)$ کوئری به بیت‌های مختلف اثبات بزند و آن‌ها را بخواند.

تعريف ۱۸.۵. کلاس پیچیدگی $\mathcal{PCP}(r(n), q(n))$ عبارت است از همه‌ی زبان‌هایی مانند L ، به طوری که تصدیق‌کننده‌ی $(r(n), q(n))$ -محدودی مانند V برایشان وجود دارد چنان‌که:

- اگر $x \in L$ ، در این صورت اثبات π موجود است که $\text{Pr}[V(x, \pi) = 1] = 1$.
- اگر $x \notin L$ ، در این صورت برای هر اثبات π $\text{Pr}[V(x, \pi) = 1] \leq \frac{1}{3}$.

^۱gap-preserving reduction

قضیه ۱۹.۵ ((قضیه PCP): اثبات‌های قابل بررسی احتمالاتی)). $\mathcal{NP} = \mathcal{PCP}(\mathcal{O}(\log n), \mathcal{O}(1))$. [۲۲]

قضیه‌ی فوق، به طور شگفت‌آوری با شهود ما در تناقض است. مثلاً فرمول CNF_3 را در نظر بگیرید که ارضاشدنی نیست، اما تمام پرانتزهای آن به جزیکی به طور هم‌زمان ارضاشدنی هستند. برای چنین فرمولی، به نظر می‌رسد با نگاه کردن تصادفی به تنها $O(1)$ بیت از یک اثبات، توان با احتمال بالا اضویت فرمول را در SAT_3 رد کرد. با این حال شگفتی قضیه‌ی PCP در آن است که وجود نوعی از اثبات‌های «قدرتمند» را برای مسأله‌های \mathcal{NP} پیشنهاد می‌دهد؛ به این معنی که اگر ایرادی در بخش کوچکی از اثبات وجود داشته باشد، این ایراد در سرتاسر این اثبات‌های قدرتمند پراکنده شده است و با احتمال بالای می‌توان آن را تشخیص داد.

نخستین اثبات قضیه‌ی PCP، که اثباتی جبری و مبتنی بر نظریه‌ی کدگذاری است، در [۲۲] مطرح شده است؛ گرچه بسیاری از بخش‌های اثبات نتایجی هستند که در [۲۳] اثبات شده بودند. اثباتی جدیدتر و ترکیبیاتی، با تکیه بر ویژگی‌های گراف‌های گسترنده، توسط دینور در [۲۴] ارائه شده است. هر دوی اثبات‌ها طولانی و پیچیده هستند و بیان آن‌ها در این مجال نمی‌گنجد. با این حال، به عنوان خاتمه‌ی این بخش، نشان خواهیم داد که دو صورت بیان شده از قضیه‌ی PCP با هم معادلند.

قضیه ۲۰.۵ ($\text{Gap-MAX-3SAT}_{1,\rho}$). اگر و تنها اگر ثابت ρ وجود داشته باشد به نحوی که $\mathcal{NP} = \mathcal{PCP}(\mathcal{O}(\log n), \mathcal{O}(1))$ ، $\text{Gap-MAX-3SAT}_{1,\rho}$ سخت باشد [۲۲].

اثبات. ابتدا سمت «اگر» را ثابت می‌کنیم. این‌که $\mathcal{PCP}(\mathcal{O}(\log n), \mathcal{O}(1)) \subseteq \mathcal{NP}$ را به سادگی می‌توان اثبات کرد. فرض کنید $(1) \in \mathcal{PCP}(\mathcal{O}(\log n), \mathcal{O}(1))$ باشد. در این صورت تصدیق‌کننده‌ی $(\mathcal{O}(\log n), \mathcal{O}(1))$ -محدود V برای L وجود دارد. توجه کنید که طول اثبات حداقل $2^{\mathcal{O}(\log n)}$ بیت است. حال تصدیق‌کننده‌ای مانند V' را در نظر بگیرید که برآمد بیت‌های تصادفی را حدس می‌زند، و برای هر حدس احتمال آن‌که V اثبات را با توجه به کوئری‌هایی که بر اساس برآمد بیت‌های تصادفی تعیین می‌شوند، بپذیرد حساب می‌کند. نهایتاً اگر این احتمال برابر ۱ باشد، اثبات را می‌پذیرد. روشن است که V یک ماشین غیرقطعی چندجمله‌ای است که $L \in \mathcal{NP}$ را می‌پذیرد. بنابراین

حال می‌خواهیم نشان دهیم $\mathcal{NP} \subseteq \mathcal{PCP}(\mathcal{O}(\log n), \mathcal{O}(1))$. بدین‌منظور نشان می‌دهیم هر زبان $L \in \mathcal{NP}$ یک تصدیق‌کننده‌ی $(\mathcal{O}(\log n), \mathcal{O}(1))$ -محدود دارد. با توجه به تحولی که از L به $\text{Gap-MAX-3SAT}_{1,\rho}$ وجود دارد، هر $x \in \{0, 1\}^*$ به فرمولی ϕ_x نگاشته می‌شود که متغیرهای y_1, y_2, \dots, y_k و پرانتزهای C_1, C_2, \dots, C_m را دارد. یک تصدیق‌کننده‌ی $(\mathcal{O}(\log n), \mathcal{O}(1))$ -محدود برای L ، ابتدا با اجرای الگوریتم تحويل، از ورودی x فرمول ϕ_x را به دست می‌آورد. سپس با استفاده از $\log m \in \mathcal{O}(\log n)$ بیت تصادفی به تصادف یکی از پرانتزها را انتخاب می‌کند. در اینجا اثبات گمارشی از مقادیر $\{true, false\}$ به متغیرهای y_1, y_2, \dots, y_k است. با توجه به پرانتز انتخاب شده، تصدیق‌کننده مقادیر مربوط به متغیرهای ظاهر شده در پرانتز مزبور را از اثبات می‌خواند، و اثبات را می‌پذیرد اگر و فقط اگر با آن مقادیر پرانتز مزبور ارضا شود. توجه کنید در حالتی که $L \neq x$ ، ϕ_x ارضی‌پذیر است و اثباتی وجود دارد که توسط تصدیق‌کننده پذیرفته شود. در حالتی که $L \neq x$ ، احتمال آن‌که اثباتی توسط تصدیق‌کننده پذیرفته شود، حداقل برابر ρ است (زیرا تعداد پرانتزهایی که هم‌زمان ارضاشدنی هستند، ρ برابر تعداد کل پرانتزهاست). از طرفی، با تکرار می‌توان خطرا را کاهش داد و به $\frac{1}{\rho}$ رساند. بنابراین تصدیق‌کننده‌ای $(\mathcal{O}(\log n), \mathcal{O}(1))$ -محدود برای L وجود دارد.

حال به اثبات سمت «تنها اگر» می‌پردازیم. فرض کنید $L \in \mathcal{NP}$ است و تصدیق‌کننده‌ای $(\mathcal{O}(\log n), \mathcal{O}(1))$ -محدود دارد. یک مسأله‌ی CSP را به این صورت طرح می‌کنیم: متغیرهای $y_{|\pi|}, y_1, y_2, \dots, y_{|\pi|}$ نمایان‌گر بیت‌های اثبات π هستند؛ و برای هر برآمد از بیت‌های تصادفی مانند C_0, C_1, \dots, C_o قیدی است که تنها وابسته به بیت‌هایی است که با توجه به آن برآمد از اثبات خوانده می‌شوند، و C_0 برآورده می‌شود اگر و تنها اگر تصدیق‌کننده با خواندن بیت‌هایی از اثبات که با توجه به o تعیین شده‌اند، اثبات را پذیرد.

توجه کنید که هر قید CSP فوق تنها به $\mathcal{O}(1)$ متغیر وابسته است، و $2^{\mathcal{O}(\log n)}$ قید دارد. در حالتی که $x \in L$ ، می‌دانیم اثباتی وجود دارد که تصدیق‌کننده با احتمال ۱ آن را می‌پذیرد؛ پس ورودی مسأله‌ی CSP بالا که از روی x تولید می‌شود نیز حتماً ارضاشدنی است. در حالتی که $x \notin L$ ، برای هر اثبات π ، احتمال پذیرفته‌شدن اثبات توسط تصدیق‌کننده حداقل $\frac{1}{\rho}$ است؛ بنابراین حداقل $\frac{1}{\rho}$ قیود ورودی مسأله به طور هم‌زمان ارضاشدنی هستند. بنابراین یک تحويل از L به Gap-CSP بالا ارائه شد.

حال توجه کنید که CSP بالا را می‌توان به یک فرمول CNF - 3 تبدیل کرد. در واقع هر قید k موضعی را می‌توان به یک فرمول CNF - 3 با 2^k پرانتز تبدیل کرد، به طوری که ارضایبزیری ورودی CSP با ارضایبزیری فرمول CNF - 3 یکسان باشد. به این ترتیب، ثوابت تمامیت و درستی به ترتیب برابر با 1 و $\frac{1}{2^k} - 1$ خواهد بود، و می‌دانیم با تکرار می‌توان ثابت درستی را به عددی ثابت کاهش داد. از طرفی چون $(1) \in O(k)$ ، تحويل به دست آمده تحولی چندجمله‌ای است. \square

۲.۳.۵. حدس PCP کوانتمومی. همان‌گونه که تا به این جا دیده‌ایم، همنتای بسیاری از مفاهیم و نتایج پیچیدگی محاسباتی کلاسیک را می‌توان در محاسبات کوانتمومی جست‌وجو کرد. در بخش قبل، به یکی از درخشنان‌ترین نتایج پیچیدگی محاسبات کلاسیک پرداختیم و دو صورت معادل را برای آن بیان کردیم. دور از انتظار نیست که با پیشرفت پیچیدگی محاسبات کوانتمومی، در پی معادل کوانتمومی این قضیه باشیم. هر چند یافتن چنین معادلی می‌تواند خوشایند باشد، با این حال به نظر می‌رسد هنوز راه زیادی تا اثبات این حدس باقی است. افزون بر این، حتی درستی یا نادرستی این حدس نیز در هاله‌ای از ابهام است و شواهدی کافی به نفع هیچ‌یک وجود ندارد [۲۵].

در این زیربخش، به معروفی دو صورت از حدس PCP کوانتمومی خواهیم پرداخت. این حدس نخستین‌بار به صورت دقیق در [۲۶] صورت‌بندی شده است؛ گرچه شواهدی برای این که پیش از این مقاله نیز افرادی به وجود چنین همنتای کوانتمومی‌ای برای قضیه‌ی PCP امیدوار بوده‌اند، در [۴] و [۲۶] قابل ردگیری است.

تعريف ۲۱.۵. یک تصدیق‌کننده‌ی (k) ، $QPCP$ ، تصدیق‌کننده‌ای کوانتمومی است که با در اختیار داشتن اثباتی مانند ψ $(\mathbb{C}^{\star})^{\otimes p(n)}$ ، ابتدا به صورت تصادفی k بیت از اثبات مانند i_1, i_2, \dots, i_k را انتخاب می‌کند، و سپس مدار V_{i_1, i_2, \dots, i_k} را روی ورودی، k بیت مشخص شده از اثبات و نیز رجیستر کمکی اعمال می‌کند، و نهایتاً یک کیویت را (که از قبل مشخص کرده) اندازه می‌گیرد و با توجه به حاصل اندازه‌گیری، اثبات را می‌پذیرد یا رد می‌کند.

تعريف ۲۲.۵. کلاس پیچیدگی $QPCP(k, c, s)$ عبارت است از تمام مسئله‌های قراردادی مانند $(\Pi_{Yes}, \Pi_{No}) = \Pi$ به طوری که تصدیق‌کننده‌ای (k) وجود دارد چنان‌که:

- اگر $x \in \Pi_{Yes}$ در این صورت اثبات ψ وجود دارد که توسط تصدیق‌کننده با احتمال حداقل c پذیرفته می‌شود.
- اگر $x \in \Pi_{No}$ در این صورت برای هر اثبات ψ ، تصدیق‌کننده با احتمال حداقل s اثبات را می‌پذیرد.

حدس ۲۳.۵ (حدس $QPCP$: نسخه‌ی اثبات‌های قابل بررسی احتمالاتی). $QMA = QPCP(O(1), c, s)$ که در آن، $c - s = \Omega(1)$. [۲۷]

پیش از بیان نسخه‌ی سختی تقریب این حدس، از تعريف ۳.۵ به یاد آورید که تا به این جا، فرض کرده بودیم فاصله‌ی قراردادی در مسئله‌ی همیلتونی‌های موضعی، یک چندجمله‌ای است، و به جهت اختصار از ذکر آن اجتناب می‌کردیم. با این حال، مسئله‌ی k -LH را می‌توان برای فاصله‌های قراردادی دیگر نیز تعریف کرد. افزون بر این، از این جا به بعد فرض کنید که در مسئله‌ی همیلتونی موضعی، همه‌ی جمله‌های موضعی ورودی، نگاشت‌هایی مثبت نیمه‌معین هستند که نرم اثرشان حداقل برابر ۱ است.

حدس ۲۴.۵ (حدس $QPCP$: نسخه‌ی سختی تقریب). ثابت $\circ > \gamma$ وجود دارد به طوری که k -LH با فاصله‌ی قراردادی γ تحت تحويل چندبه‌یک چندجمله‌ای کوانتمومی مسئله‌ای QMA -سخت است [۲۷].

مفهوم از یک تحويل چندبه‌یک چندجمله‌ای کوانتمومی در گزاره‌ی بالا یک الگوریتم کوانتمومی و چندجمله‌ای است که با احتمال ثابت و ناصفر، تابع تحويل را پیاده‌سازی می‌کند.

ملحوظه ۲۵.۵. مشابه با قضیه‌ی PCP کلاسیک، می‌توان نشان داد دو صورت بیان شده از حدس $QPCP$ در بالا نیز با یکدیگر معادلند. در واقع اثبات یک سمت آن، سمتی که نسخه‌ی سختی تقریب نسخه‌ی اثبات‌های قابل بررسی احتمالاتی را نتیجه می‌دهد، ساده است، و می‌توان دید که چنان‌چه فاصله‌ی قراردادی مسئله‌ی k -LH مقداری ثابت باشد، تصدیق‌کننده‌ای که در قضیه‌ی ۵.۵ ارائه کردیم تصدیق‌کننده‌ای (k) است که ثوابت درستی و تمامیتی با فاصله‌ی ثابت خواهد داشت. سمت دیگر دشوارتر است، و به نظر می‌رسد بدون فرض کامل بودن تحت تحويل کوانتمومی، قادر به اثبات آن نیستیم. خواننده‌ی علاقه‌مند می‌تواند اثباتی برای سمت دیگر را در [۲۸] بیابد.

ملاحظه ۲۶.۵. در نتیجه‌ی ۹.۵ دیدیم که درستی قضیه‌ی کوک-لوین کواتومی نتیجه می‌دهد که سیستم‌هایی فیزیکی وجود دارند که اگر آن‌ها را تا دمای صفر سرد کنیم، در حالتی قویاً درهم‌تییده قرار می‌گیرند. به طریقی مشابه، درستی حدس PCP کواتومی، همراه با فرض $QMA \neq QCMA$ نتیجه خواهد داد که سیستم‌هایی فیزیکی وجود دارند که حالت آن‌ها حتی در دمای متناهی ناصرف نیز قویاً درهم‌تییده است. چنین نتیجه‌ای خلاف شهود فیزیکی متخصصان نظریه‌ی سیستم‌های چندپیکره است؛ چه آن‌که آن‌ها بر این باورند که نمی‌توان در دماهای بالا شاهد اثرات کواتومی با مقیاس بزرگ بود، و مثلًاً تلاش برای یافتن موادی که در دمای اتاق ابررسانا باشند، ناموفق است. به این ترتیب، چنانچه حدس PCP کواتومی ثابت شود، بر شهود فیزیکی استاندارد ما نیز تاثیر خواهد گذاشت [۲۵].

۶. مؤخره: پیشرفت‌های جدیدتر و زمینه‌های پژوهش

در این بخش اجمالاً اشاره‌ای به برخی زمینه‌های فعلی پژوهش که مرتبط با اثبات‌های غیرتعاملی کواتومی هستند، و نیز پیشرفت‌های نسبتاً جدیدتری که در این زمینه‌ها رخ داده است، خواهیم داشت.

(۱) در جست‌وجوی ارتباطاتی عمیق‌ترین فیزیک و نظریه‌ی محاسبه: همان‌گونه که در این مقاله دیدیم، برخی زمینه‌های پژوهش حول اثبات‌های کواتومی ارتباطات عمیقی میان مفهوم محاسبه و نظریه‌های فیزیکی برقرار می‌کنند. مطالعه‌ی چنین ارتباط‌هایی از دو جهت حائز اهمیت است:

- همان‌گونه که در یادداشت آغازین بخش ۵ اشاره شد، پیچیدگی همیلتونی کواتومی حوزه‌ای است که از یک سو مورد توجه فیزیکدانهای سیستم‌های چندپیکره و از دیگر سو مورد توجه پژوهشگران علوم کامپیوتر است. با وجود آن‌که معمولاً متخصصین علوم کامپیوتر چندان علاقه‌ای به وجه فیزیکی مسائل ندارند و ترجیح می‌دهند تا حد امکان از درگیرشدن با آن اجتناب کنند، به نظر می‌رسد که توجه به این وجه، ضروری و پیش‌برنده‌ی مسائل این حوزه باشد. در توضیح می‌توان گفت که یک رویکرد به روند توسعه‌ی نظریه‌ی پیچیدگی محاسبات کواتومی، همان‌گونه که تا به این جای این مقاله بارها بر آن تاکید کرده‌ایم، تلاش برای یافتن آنالوژی‌هایی میان محاسبات کلاسیک و محاسبات کواتومی است. نظریه‌ی محاسبه و پیچیدگی محاسبه‌ی کلاسیک طی حدود یک قرنی که از تولد آن می‌گذرد، دستاوردهای متعدد و درخشنانی داشته است، و حال با ظهور مدل محاسباتی جدیدی به نام محاسبات کواتومی، این که آیا نتایج مشابهی در این زمینه نیز یافت خواهد شد، کنگکاوی برانگیز است. با این حال، این سوال ممکن است در ذهن ایجاد شود که آیا چنین نتایج «مشابهی»، واقعاً سودمند و عمیق نیز هستند؟ این جاست که اهمیت فیزیکی مسئله می‌تواند به عنوان راهنمایی برای گزینش مسائل «خوب» به یاری ما بیاید [۲۹].

همان‌گونه که در یادداشت‌های ۹.۵ و ۲۶.۵ دیدیم، مسائلی که در حال حاضر در پیچیدگی همیلتونی از منظر محاسباتی مورد مطالعه قرار دارند، معنایی فیزیکی نیز دارند، و نتایج آن‌ها نه تنها برای فیزیکدانان نظری، بلکه برای متخصصان و مهندسان زمینه‌های دیگر نیز اهمیت دارد (برای نمونه رجوع کنید به [۳۰]). پژوهش‌های متعددی در سال‌های اخیر صورت گرفته است که با توجه به این معناداری فیزیکی، محدودیت‌هایی روی همیلتونی‌هایی که ممکن است در حدس QPCP ظاهر شوند، قرار داده شود. مثلًاً برخی همیلتونی‌هایی که این حدس نمی‌تواند برای آن‌ها درست باشد در [۳۱] مورد مطالعه قرار گرفته‌اند.

- وجهی دیگر از این ارتباطات، تاثیر آن بر پیشبرد فیزیک نظری است. آن‌گونه که ویگدرسون در [۳۲] می‌گوید، «شاید خواسته‌ی فیزیکدانان برای درک ساختار بنیادین فضا و زمان فیزیکی، وابسته به داشتن فهمی عمیق از منابع محاسباتی فضا و زمان باشد». نمونه‌ای از چنین تاثیراتی را می‌توان در کار هارلو و هایدن در [۳۳] دید، که به بررسی همواری افق‌های سیاه‌چالهای از منظر محاسباتی و ارتباط این موضوع با اثبات‌های دانش صفر کواتومی می‌پردازند. انتظار می‌رود که موارد بیشتری از چنین ارتباطاتی با حوزه‌های مختلف فیزیک یافت شود، و باور بر این است که آشنایی فیزیکدانان با مفاهیم و روش‌های محاسباتی، به شکوفایی‌های بیشتری در دستاوردهای فیزیکی می‌انجامد [۳۲].

(۲) در تلاش برای شناختن بهتر نسخه‌های مختلف QMA : همان‌گونه که در زیربخش ۲.۳.۵ دیدیم، یکی از صورت‌های حدس PCP کواتومی مرتبط با یافتن صورت‌بندی جدیدی از کلاس QMA با استفاده از نوعی از

تصدیق‌کننده‌های بسیار کارا است. به نظر می‌رسد چنانچه کلاس QMA و نسخه‌های مختلف آن را بهتر بشناسیم، راه ما برای یافتن اثباتی در تایید یا رد این حدس هموارتر خواهد شد [۲۵]. از طرفی، همان‌گونه که در بخش ۲.۴ بحث کردیم، مسائل حل نشده‌ی بسیاری درباره‌ی ارتباط میان این کلاس‌ها وجود دارد، و همین سبب شده است که بخشی از توجه پژوهشگران پیچیدگی محاسبات کوانتومی معطوف به مطالعه‌ی روابط میان این کلاس‌ها و مسائل کامل آن‌ها شود. نمونه‌هایی از برخی پژوهش‌های متاخر را می‌توان در [۳۴، ۳۵، ۳۶] مشاهده کرد.

(۳) نتایج یافته‌های پیچیدگی کوانتومی در پیچیدگی کلاسیک: پیشرفت‌های پیچیدگی محاسبات کوانتومی در سال‌های اخیر به حصول نتایج ارزشمندی در محاسبات کلاسیک انجامیده است. یک نمونه‌ی آن، یافتن صورت‌بندی جدیدی از کلاس NP از طریق ارائه‌ی پروتکلی برای تصدیق یک اثبات ۳SAT به طول m با استفاده از $O(\sqrt{m})$ شاهد کوانتومی غیر درهم‌تینده به طول $O(\log m)$ است، که توسط بیگی و همکاران در [۳۷] معرفی شده است. نمونه‌ای دیگر و متاخرتر، نتیجه‌ای است که توسط آهارونوف و همکاران در [۱۴] به دست آمده است. در اینجا ثابت می‌شود که اثبات حدس PCP کوانتومی برای خانواده‌ی خاصی از همیلتونی‌ها، معادل با پاسخ دادن به MA vs. NP است. مسئله‌ی اخیر، که نمونه‌ای از مسائل غیرتصادفی‌سازی^۱ است، برنامه‌ای است که در حال حاضر در پیچیدگی محاسبات کلاسیک در جریان است. باور عمومی براین است که این برنامه به نتیجه خواهد رسید و خواهیم توانست نشان دهیم که تصادفی‌سازی بر قدرت محاسباتی نمی‌افزاید.

مراجع

- [1] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. 2002. Classical and Quantum Computation. American Mathematical Society, USA.
- [2] Watrous, J. (2000, November). Succinct quantum proofs for properties of finite groups. In Proceedings 41st Annual Symposium on Foundations of Computer Science (pp. 537-546). IEEE.
- [3] Vidick, T., & Watrous, J. (2016). Quantum proofs. Foundations and Trends® in Theoretical Computer Science, 11(1-2), 1-215.
- [4] Aharonov, D., & Naveh, T. (2002). Quantum NP-a survey. arXiv preprint quant-ph/0210077.
- [5] Marriott, C., & Watrous, J. (2005). Quantum arthur–merlin games. computational complexity, 14(2), 122-152.
- [6] Yamakami, T. (1999). Analysis of Quantum Functions: (Preliminary Version). In Foundations of Software Technology and Theoretical Computer Science: 19th Conference Chennai, India, December 13-15, 1999 Proceedings 19 (pp. 407-419). Springer Berlin Heidelberg.
- [7] Aaronson, S., & Kuperberg, G. (2007, June). Quantum versus classical proofs and advice. In Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07) (pp. 115-128). IEEE.
- [8] Aaronson, S. (2009). On perfect completeness for QMA. Quantum Information and Computation, 9(1), 0081-0089.
- [9] Jordan, S. P., Kobayashi, H., Nagaj, D., & Nishimura, H. (2012). Achieving perfect completeness in classical-witness quantum merlin-arthur proof systems. Quantum Information & Computation, 12(5-6), 461-471.
- [10] Kobayashi, H., Matsumoto, K., & Yamakami, T. (2003). Quantum Merlin-Arthur proof systems: Are multiple Merlins more helpful to Arthur?. In Algorithms and Computation: 14th International Symposium, ISAAC 2003, Kyoto, Japan, December 15-17, 2003. Proceedings 14 (pp. 189-198). Springer Berlin Heidelberg.
- [11] Liu, Y. K., Christandl, M., & Verstraete, F. (2007). Quantum computational complexity of the N-representability problem: QMA complete. Physical review letters, 98(11), 110503.
- [12] Harrow, A. W., & Montanaro, A. (2013). Testing product states, quantum Merlin-Arthur games and tensor optimization. Journal of the ACM (JACM), 60(1), 1-43.
- [13] Bravyi, S., Bessen, A. J., & Terhal, B. M. (2006). Merlin-Arthur games and stoquastic complexity. arXiv preprint quant-ph/0611021.
- [14] Aharonov, D., Grilo, A. B., & Liu, Y. (2020). StoqMA vs. MA: the power of error reduction. arXiv preprint arXiv:2010.02835.
- [15] Osborne, T. J. (2012). Hamiltonian complexity. Reports on progress in physics, 75(2), 022001.
- [16] Gharibian, S., Huang, Y., Landau, Z., & Shin, S. W. (2015). Quantum hamiltonian complexity. Foundations and Trends® in Theoretical Computer Science, 10(3), 159-282.
- [17] Trakhtenbrot, B.A. (1984). A Survey of Russian Approaches to Perebor (Brute-Force Searches) Algorithms. Annals of the History of Computing, 6, 384-400.
- [18] Cook, S. A. (2023). The complexity of theorem-proving procedures. In Logic, Automata, and Computational Complexity: The Works of Stephen A. Cook (pp. 143-152).
- [19] Balcazar, J. L., Diaz, J., & Gabarro, J. (2012). Structural Complexity I. Springer Science & Business Media.
- [20] Kempe, J., & Regev, O. (2003). 3-local Hamiltonian is QMA-complete. Quantum Information and Computation, 3(3), 258-264.

¹derandomization

- [21] Kempe, J., Kitaev, A., & Regev, O. (2005). The complexity of the local Hamiltonian problem. In FSTTCS 2004: Foundations of Software Technology and Theoretical Computer Science: 24th International Conference, Chennai, India, December 16-18, 2004. Proceedings 24 (pp. 372-383). Springer Berlin Heidelberg.
- [22] Arora, S., Lund, C., Motwani, R., Sudan, M., & Szegedy, M. (1998). Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3), 501-555.
- [23] Arora, S., & Safra, S. (1992). Probabilistic checking of proofs; a new characterization of NP. *Proceedings., 33rd Annual Symposium on Foundations of Computer Science*, 2-13.
- [24] Dinur, I., & Reingold, O. (2004). Assignment testers: towards a combinatorial proof of the PCP-theorem. *45th Annual IEEE Symposium on Foundations of Computer Science*, 155-164.
- [25] Aharonov, D., Arad, I., & Vidick, T. (2013). Guest column: the quantum PCP conjecture. *ACM sigact news*, 44(2), 47-79.
- [26] Shtetl-Optimized » Blog Archive » The Quantum PCP Manifesto. (n.d.). Retrieved April 14, 2023, from <https://scottaaronson.blog/?p=139>
- [27] Aharonov, D., Arad, I., Landau, Z., & Vazirani, U. (2009, May). The detectability lemma and quantum gap amplification. In *Proceedings of the forty-first annual ACM symposium on Theory of computing* (pp. 417-426).
- [28] Grilo, A. B. (2018). Quantum proofs, the local Hamiltonian problem and applications (Doctoral dissertation, Université Sorbonne Paris Cité).
- [29] A quantum PCP theorem? | MyCQstate. (n.d.). Retrieved April 6, 2024, from <https://mycqstate.wordpress.com/2013/02/24/a-quantum-pcp-theorem/>
- [30] Gharibian, S., & Le Gall, F. (2022, June). Dequantizing the quantum singular value transformation: hardness and applications to quantum chemistry and the quantum PCP conjecture. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing* (pp. 19-32).
- [31] Brandao, F. G., & Harrow, A. W. (2013, June). Product-state approximations to quantum ground states. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing* (pp. 871-880).
- [32] Wigderson, A. (2019). Mathematics and computation: A theory revolutionizing technology and science. Princeton University Press.
- [33] Harlow, D., & Hayden, P. (2013). Quantum computation vs. firewalls. *Journal of High Energy Physics*, 2013(6), 1-56.
- [34] Chailloux, A., & Sattath, O. (2012, June). The complexity of the separable Hamiltonian problem. In *2012 IEEE 27th Conference on Computational Complexity* (pp. 32-41). IEEE.
- [35] Bittel, L., Gharibian, S., & Kliesch, M. (2023). The Optimal Depth of Variational Quantum Algorithms Is QCMA-Hard to Approximate. In *38th Computational Complexity Conference (CCC 2023)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.
- [36] Grilo, A. B., Kerenidis, I., & Sikora, J. (2015, August). QMA with subset state witnesses. In *International Symposium on Mathematical Foundations of Computer Science* (pp. 163-174). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [37] Aaronson, S., Beigi, S., Drucker, A., Fefferman, B., & Shor, P. (2008, June). The power of unentanglement. In *2008 23rd Annual IEEE Conference on Computational Complexity* (pp. 223-236). IEEE.
- [38] Yao, A. C. C. (1993, November). Quantum circuit complexity. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science* (pp. 352-361). IEEE.
- [39] Akama, S. (2015). Elements of quantum computing. History, Theories and Engineering Applications, Springer.
- [40] Macchiavello, C., Palma, G. M., & Zeilinger, A. (Eds.). (2000). *Quantum Computation and Quantum Information Theory: Reprint Volume with Introductory Notes for ISI TMR Network School, 12-23 July 1999, Villa Gualino, Torino, Italy*. World Scientific.
- [41] Barenco, A., Bennett, C. H., Cleve, R., DiVincenzo, D. P., Margolus, N., Shor, P., ... & Weinfurter, H. (1995). Elementary gates for quantum computation. *Physical review A*, 52(5), 3457.
- [42] Watrous, J. (2018). *The theory of quantum information*. Cambridge university press.
- [43] Deutsch, D. E. (1989). Quantum computational networks. *Proceedings of the royal society of London. A. mathematical and physical sciences*, 425(1868), 73-90.
- [44] Barenco, A. (1995). A universal two-bit gate for quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 449(1937), 679-683.
- [45] Nielsen, M. A., & Chuang, I. L. (2010). *Quantum computation and quantum information*. Cambridge university press.
- [46] Deutsch, D. E., Barenco, A., & Ekert, A. (1995). Universality in quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 449(1937), 669-677.
- [47] Dawson, C. M., & Nielsen, M. A. (2006). The Solovay-Kitaev algorithm. *Quantum Information and Computation*, 6(1), 81-95.
- [48] Cook, S. A. (1970). Alan Cobham. The intrinsic computational difficulty of functions. *Logic, methodology and philosophy of science, Proceedings of the 1964 International Congress*, edited by Yehoshua Bar-Hillel, Studies in logic and the foundations of mathematics, North-Holland Publishing Company, Amsterdam 1965, pp. 24–30. *The Journal of Symbolic Logic*, 34(4), 657-657.
- [49] Tang, C. L. (2005). *Fundamentals of quantum mechanics: for solid state electronics and optics*. Cambridge University Press.
- [50] Planck, M. (1978). Über das gesetz der energieverteilung im normalspektrum (pp. 178-191). Vieweg+ Teubner Verlag.
- [51] Einstein, A. (1965). Concerning an heuristic point of view toward the emission and transformation of light. *American Journal of Physics*, 33(5), 367.
- [52] Bohr, N. (1913). I. On the constitution of atoms and molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 26(151), 1-25.

- [53] Von Neumann, J. (2013). *Mathematische grundlagen der quantenmechanik* (Vol. 38). Springer-Verlag.
- [54] Wootters, W. K., & Zurek, W. H. (1982). A single quantum cannot be cloned. *Nature*, 299(5886), 802-803.
- [55] Dieks, D. G. B. J. (1982). Communication by EPR devices. *Physics Letters A*, 92(6), 271-272.
- [56] Benioff, P. (1980). The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. *Journal of statistical physics*, 22, 563-591.
- [57] Feynman, R. P. (2018). Simulating physics with computers. In *Feynman and computation* (pp. 133-153). CRC Press.
- [58] Bernstein, E., & Vazirani, U. (1993, June). Quantum complexity theory. In *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing* (pp. 11-20).
- [59] Simon, D. R. (1997). On the power of quantum computation. *SIAM journal on computing*, 26(5), 1474-1483.
- [60] Grover, L. K. (1996, July). A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (pp. 212-219).
- [61] Shor, P. W. (1999). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2), 303-332.
- [62] Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., ... & Martinis, J. M. (2019). Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779), 505-510.
- [63] White House Earmarks New Money for A.I. and Quantum Computing - The New York Times. (n.d.). Retrieved April 7, 2024, from <https://www.nytimes.com/2020/02/10/technology/white-house-earmarks-new-money-for-ai-and-quantum-computing.html>
- [64] Shamir, A. (1992). IP= PSPACE. *Journal of the ACM (JACM)*, 39(4), 869-877.
- [65] Sipser, M. (1992, July). The history and status of the P versus NP question. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing* (pp. 603-618).
- [66] Babai, L., Fortnow, L., & Lund, C. (1991). Non-deterministic exponential time has two-prover interactive protocols. *Computational complexity*, 1, 3-40.
- [67] Knill, E. (1996). Quantum randomness and nondeterminism. *arXiv preprint quant-ph/9610012*.
- [68] Schrödinger, E. (1935, October). Discussion of probability relations between separated systems. In *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 31, No. 4, pp. 555-563). Cambridge University Press.
- [69] Jozsa, R., & Linden, N. (2003). On the role of entanglement in quantum-computational speed-up. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 459(2036), 2011-2032.

* دانشجوی دکتری علوم کامپیوت، انسٹیتو پلی تکنیک پاریس

تاریخ: <https://ali-almasi.github.io>

ایمیل: ali.almasi@polytechnique.edu



شمارش قدم‌زدن‌های خودپرهیز روی مشبکه‌ی شش ضلعی

هوگو دامینیل کوین

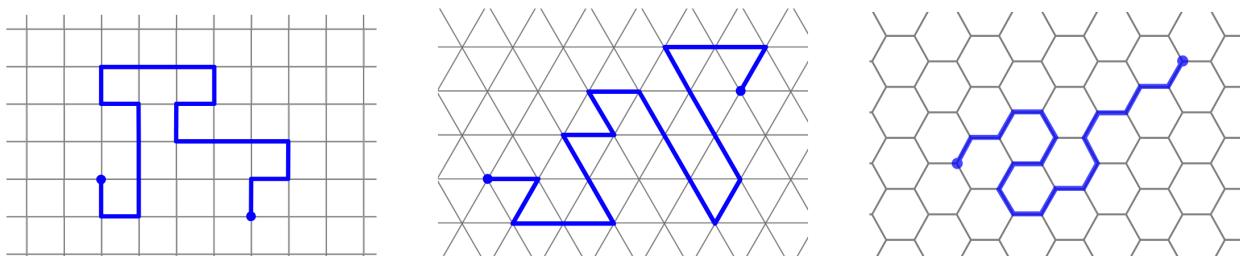
چکیده. به چند طریق می‌توانید روی یک مشبکه‌ی توپی قدم بزنید، به طوری که هیچ‌گاه به مسیری که قبلاً طی کردید اید بروخورد نکنید؟ در اینجا به برخی جنبه‌های ترکیباتی و آماری این قدم‌زدن به اصطلاح «خودپرهیز»^۱ می‌پردازم. به طور خاص، نتایج اخیر به دست آمده درباره‌ی تعداد قدم‌زدن‌های تصادفی روی مشبکه‌ی شش ضلعی (مشبکه‌ی کندوی عسل) را مورد بررسی قرار می‌دهیم. در بخش آخر، به طور خلاصه به ارتباط با هندسه‌ی قدم‌زدن‌های تصادفی خودپرهیز با طول بالا اشاره می‌کنیم.

۱. قدم‌زدن‌های تصادفی خودپرهیز روی یک مشبکه

در اواسط قرن بیستم میلادی، پاول فلوری^۲ و اور^۳ قدم‌زدن‌های تصادفی خودپرهیز^۴ (SAWs) را به عنوان یک مدل ریاضی برای شکل پلیمرها معرفی کردند. پلیمر ایده‌آل، یک زنجیره‌ی بلند از مولکول‌هاست که از تعداد زیادی اتصالات مشابه، به نام مونومر، تشکیل شده‌است [۵، ۱۱].

بیایید یک مشبکه، مانند مشبکه‌ی مرتعی، مشبکه‌ی مثلثی، یا مشبکه‌ی شش ضلعی را در نظر بگیریم (شکل ۱). یک قدم‌زدن خودپرهیز روی یک مشبکه، یک دنباله از رأس‌های مجاور است که به هیچ‌یک از رأس‌های قبلی باز نمی‌گردد. اولین سوالی که در ذهن ایجاد می‌شود این است:

چند SAW به طول n با شروع از رأس داده‌شده وجود دارد؟



شکل ۱. مثال‌هایی از قدم‌زدن‌های خودپرهیز با طول ۱۹ روی مشبکه‌ی مرتعی، مثلثی و شش ضلعی. واحد طول برابر با طول هر یال در مشبکه است.

برای پرداختن به این سوال، ابتدا نیاز داریم نگاهی دقیق‌تر به تعاریف بیندازیم. یک گراف، مجموعه‌ای از رئوس و یال‌هاست، به طوری که هر یال، دو رأس را به هم متصل می‌کند: زمانی که دو رأس با یک یال به هم متصل می‌شوند، می‌گوییم آن دو رأس مجاور یکدیگرند^۵. یک گراف، تراکدر^۶ است اگر، هر رأس گراف را بتوان با یک «تقارن» گراف به هر رأس دیگر آن تصویر

*این نوشته ترجمه‌ای از مقاله‌ی زیر است:

Duminil-Copin, H. (2019). Counting self-avoiding walks on the hexagonal lattice. Mathematisches Forschungsinstitut Oberwolfach.

^۱Paul J. Flory

^۲W. J. C. Orr

^۳self-avoiding walks

^۴در اینجا فقط گراف‌های ساده —یعنی گراف‌هایی که طوفه یا یال چندگانه ندارند— را بررسی می‌کنیم.

^۵transitive

کرد. به بیان دیگر، گراف از همه‌ی رئوسش یکسان دیده می‌شود. در دیدگاه پلیمری، رئوس گراف همان مونومرها هستند که با یال‌ها به هم متصل شده‌اند.

یک گراف، نامتناهی است اگر تعداد نامتناهی رأس داشته باشد، و به صورت موضعی متناهی است اگر هر رأس آن، درجه‌ی متناهی داشته باشد. بر این اساس، یک مشبکه را تعریف می‌کنیم: یک گراف نامتناهی تراکندر موضعی-متناهی. علاوه بر مشبکه‌های مریعی، مثلثی و شش ضلعی که قبل تر به آن‌ها اشاره کردیم و همگی در صفحه‌ی \mathbb{R}^2 هستند، یک مثال دیگر از مشبکه، مشبکه‌ی مکعبی در \mathbb{R}^3 است؛ که مجموعه‌ی رئوس آن \mathbb{Z}^3 است، و دو رأس همسایه‌اند اگر فاصله‌ی آن‌ها از یکدیگر ۱ باشد.

برای گراف داده‌شده‌ی \mathbb{G} ، یک قدمزدن به طول $n \in \mathbb{N}$ روی \mathbb{G} ، یک نگاشت $\mathbb{G} \rightarrow \{0, \dots, n\}$ است؛ به طوری که به ازای هر $i \in \{0, \dots, n-1\}$ دو رأس (i) و $(i+1)$ با هم مجاور باشند. به یک قدمزدن خودپرهیز می‌گوییم اگر یک به یک باشد؛ یعنی داشته باشیم: $(j) \neq (i)$ برای $j \neq i$. از این‌جا به بعد، تنها قدمزدن روی گراف‌هایی را در نظر می‌گیریم که مشبکه باشند.

برای پاسخ دادن به سوال قبلی شمارش تعداد قدمزدن‌های خودپرهیز، مشبکه‌ی \mathbb{L} را در نظر بگیرید، و فرض کنید c_n تعداد قدمزدن‌های خودپرهیز به طول n روی مشبکه‌ی \mathbb{L} با شروع از یک نقطه‌ی ثابت باشد. از آن‌جایی که مشبکه تراکندر است، مقدار c_n به نقطه‌ی شروع بستگی ندارد. برای مقادیر کوچک n ، مقدار c_n را می‌توان دستی حساب کرد و این کار سرگرم‌کننده است. با این حال با افزایش n ، این محاسبه‌ی دستی عملأً ناممکن می‌شود. این به این علت است که، همان‌طور که در زیر می‌بینیم، c_n به صورت نمایی رشد می‌کند. با استفاده از تکنولوژی امروز و با استفاده از الگوریتم‌های کارا، محاسبه‌ی تعداد قدمزدن‌ها تا طول 36 روی \mathbb{Z}^3 ممکن است، که برای آن یک الگوریتم جدید و 50000 ساعت محاسبه لازم است تا مقدار c_{36} به دست بیاید [۱۲]. روی مشبکه‌ی مریعی \mathbb{Z}^2 ، بزرگ‌ترین محاسبه مربوط می‌شود به قدمزدن‌ها به طول 71 [۲]. انتظار نمی‌رود که هیچ فرمول دقیقی برای c_n وجود داشته باشد. با این حال، می‌توانیم رفتار مجانبی و یا حدی c_n را، برای n ‌های خیلی بزرگ، مطالعه کنیم. جان همرسلی^۱ مشاهده کرد [۶]، که دنباله‌ی c_n دارای این ویژگی است که: عدد حقیقی مثبت مشخص $(\mathbb{L})_{\mu_c}$ وجود دارد، به طوری که

$$\lim_{n \rightarrow \infty} c_n^{1/n} = \mu_c(\mathbb{L}). \quad (1.1)$$

به $(\mathbb{L})_{\mu_c}$ ثابت اتصال مشبکه‌ی \mathbb{L} می‌گویند. در نتیجه، برای n ‌های بزرگ، c_n تقریباً برابر با $(\mathbb{L})_{\mu_c}$ است. برهان زیایی همرسلی در فیزیک آماری و احتمال به یک استدلال کلاسیک بدل شده است. این استدلال به صورت زیر است.
از آن‌جایی که یک SAW با $n+m$ قدم را می‌توان به یک n -SAW و یک انتقال موازی از یک m -قدمی تقسیم کرد، داریم

$$c_{n+m} \leq c_n c_m$$

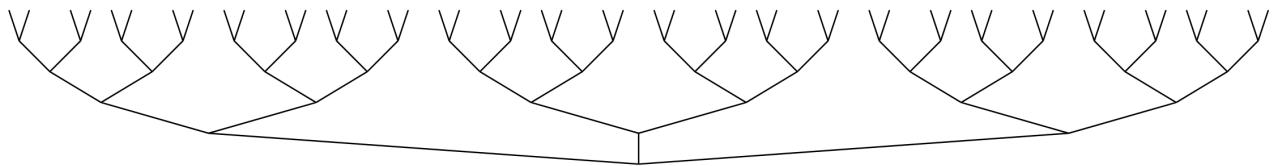
در نتیجه، $(c_n)_{n \in \mathbb{N}}$ یک دنباله‌ی (اصطلاحاً) زیرضربی است. به عنوان تمرین می‌توان نشان داد برای چنین دنباله‌هایی، ثابت $(\mathbb{L})_{\mu_c}$ وجود دارد به طوری که رابطه‌ی ۱.۱ برقرار باشد. توجه کنید که $(\mathbb{L})_{\mu_c}$ بزرگ‌تر مساوی ۱ و کمتر مساوی تعداد همسایه‌های یک نقطه منهای ۱ است.

به عنوان مثال، درخت از درجه‌ی $1+d$ را در نظر بگیرید. مشبکه‌ی \mathbb{T}_d به طور یکتا با این دو ویژگی تعریف می‌شود: درجه‌ی هر رأس $1+d$ است، و برای هر دو رأس، دقیقاً یک قدمزدن خودپرهیز، با شروع از یکی و پایان با دیگری، وجود دارد. بخشی از \mathbb{T}_2 در شکل ۲ نشان داده شده است. به راحتی می‌توان بررسی کرد که $c_n(\mathbb{T}_d) = d^n$ ، و لذا $\mu_c(\mathbb{T}_d) = d$. متأسفانه، انتظار نمی‌رود به زودی فرمولی صریح برای $(\mathbb{L})_{\mu_c}$ یافته شود، و ریاضی‌دانان و فیزیک‌دانان تنها پیش‌بینی‌هایی عددی برای رایج‌ترین مشبکه‌ها را در اختیار دارند. به عنوان مثال، مقاله‌های [۲، ۳] تخمین‌های زیر را برای ثابت اتصال \mathbb{Z}^2 و \mathbb{Z}^3 ارائه می‌دهند - پرانترزها شامل خطای احتمالی در این ارقام است:

$$\mu_c(\mathbb{Z}^2) = 2,638,158,3035(2),$$

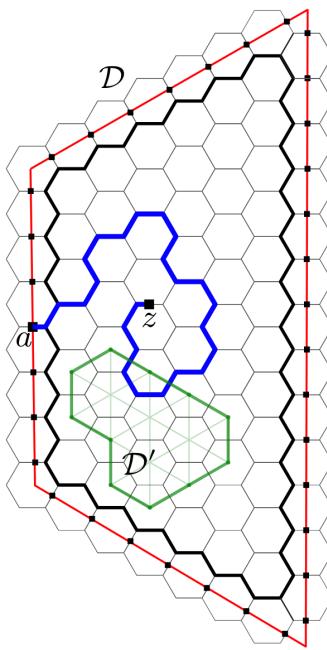
$$\mu_c(\mathbb{Z}^3) = 4,684,039,931(27).$$

^۱John Hammersley



شکل ۲. بخشی متناهی از \mathbb{T}_2 که درخت با درجه ۳ است. برخلاف آن‌چه که تصویر ممکن است القا کند، همه‌ی یال‌ها، و در نتیجه همه‌ی قدم‌های یک قدمزدن طول یکسانی دارند.

۲. ثابت اتصال مشبکه‌ی شش ضلعی



شکل ۳. یک دامنه‌ی D در مشبکه‌ی شش ضلعی، به همراه کاتور مرزی‌اش (به رنگ قرمز) و جعبه‌های کوچک در نقاط میان-یالی مرزی. یک قدمزدن خودپرهیز در D از یک نقطه‌ی میان-یالی مرزی به یک میان-یالی درونی با رنگ آبی نمایش داده شده است. یک کاتور گسسته که زیردامنه‌ی D از D را محدود می‌کند، به همراه تجزیه‌اش به کاتورهای مثلثی مقدماتی به رنگ سبز نمایش داده شده است.

در سال ۱۹۸۰، برنارد نینهویس [۱۰] پیشنهاد داد که مشبکه شش ضلعی H در میان مشبکه‌ها خاص است، از آن جهت که ثابت اتصال آن را می‌توان به طور دقیق بیان کرد. در واقع، نینهویس حدس زد که

$$\mu_c(H) = \sqrt{2 + \sqrt{2}}.$$

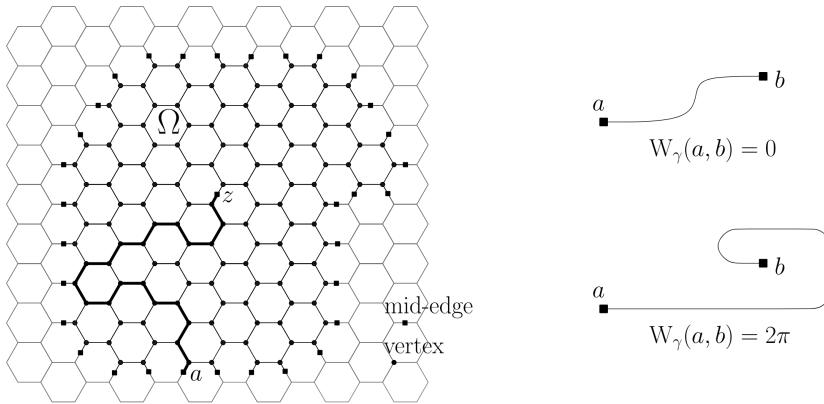
این پیش‌بینی زیبا بر پایه تطابق با مدل‌های مختلف فیزیک آماری بود و به لحاظ ریاضی نتیجه نشده بود. اخیراً، هوگو دومینیل-کوین و استانیسلاس اسمیرنوف، شواهد دقیقی از این نتیجه ارائه دادند [۴]. روند اثبات بسیار آموزونده است. اگرچه نمی‌توانیم آن را در تعداد کمی صفحه شرح دهیم، با این حال یک جنبه مهم از اثبات را بر جسته می‌کنیم و خواننده علاقه‌مند را به مقاله اصلی ارجاع می‌دهیم.

از این پس، مشبکه شش ضلعی H را روی صفحه مختلط C در نظر می‌گیریم. یک دامنه متناهی D را در نظر بگیرید؛ یعنی یک زیرمجموعه متناهی از یال‌های H . یک نقطه‌ی میان-یالی مانند a را در مرز D انتخاب کنید (همانطور که در شکل ۳ آمده است). برای هر نقطه میان-یالی z در D ، قدمزنی‌های خودپرهیز در D را که از a شروع شده و در z به پایان می‌رسند، در نظر بگیرید. برای یک قدمزنی مانند γ ، فرض کنید $W_\gamma(a, z)$ برابر با مجموع زاویه‌ای (بر حسب رادیان) است که جهت γ در مسیر از a به z چرخیده است (شکل ۴ را ببینید). به عبارت دیگر، $W_\gamma(a, z)$ برابر با تعداد چرخش‌های به چپ منهای تعداد

چرخش‌های به راستی است که γ در مسیر a به z انجام داده است، ضرب در $\frac{\pi}{\ell}$. تابع مختلط F را برای هر نقطه میان-یالی z وابسته به پارامترهای σ و x ، با فرمول زیر تعریف می‌کنیم.

$$F(z) := \sum_{\substack{z \in a \text{ از } D \\ \text{قدمزنی خودبهیز } \gamma \text{ در }}} e^{-i\sigma W_\gamma(a,z)} x^{\gamma z} \quad (1.2)$$

برای مثال در شکل ۲۴، قدمزنی آبی از a به z از ۲۴ راس عبور می‌کند. ۹ چرخش به چپ و ۱۵ چرخش به راست انجام



شکل ۴. چرخش قدمزنی γ .

می‌دهد و در نتیجه یک جمله $e^{2\pi i \sigma} x^{24} e^{2\pi i \sigma} x^{24}$ را به مجموع در فرمول ۱.۲ اضافه می‌کند. مزیت تابع F ، که به آن مشاهده‌پذیر پارافرمیونیک^۱ می‌گویند، این است که زمانی که $\sigma = \frac{5}{8}$ باشد، یک ویژگی بسیار خاص دارد: در اطراف هر راس v در D ، برای میان-یال‌های مجاور p ، q و r ، با ترتیب خلاف عقربه‌های ساعت، رابطه زیر برقرار است:

$$F(p) + e^{\frac{5\pi i}{8}} F(q) + e^{\frac{5\pi i}{8}} F(r) = 0. \quad (2.2)$$

ضرایب در این رابطه به گونه‌ای است که می‌توان سمت چپ را به عنوان یک جمع در امتداد «یک کانتور مقدماتی»^۲ بر روی مشبکه‌ی دوگان» دید، با صرف نظر از یک فاکتور ضربی. اگرچه این جمله ممکن است به نظر چند کلمات عجیب و غریب کنار هم باشد، اما تفسیرش بسیار ساده است:

یک کانتور در H ، یک مسیر $(z_i)_{i \leq n}$ از وجههای همسایه در H است، که z_i عدد مختلطی در مرکز وجه متناظر است. مسیر بسته است اگر که وجه اول همان وجه پایانی باشد، به عبارت دیگر، z_n و z_0 برابر باشند. برای مثال، مسیرهای سبز در شکل ۲۵ را بینید. انتگرال گسسته یک تابع F روی میان-یال‌ها در امتداد کانتور c را به شکل زیر تعریف می‌کنیم:

$$\oint_c F(z) dz := \sum_{i=0}^{n-1} F(p_i) (z_{i+1} - z_i), \quad (2.2)$$

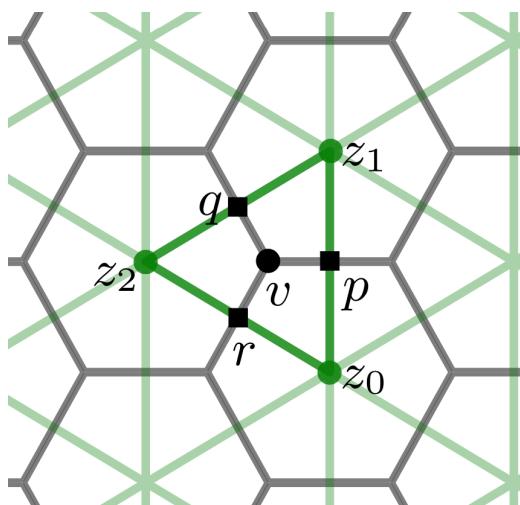
که p_i مرکز یال محاط توسط وجههای متناظر با z_i و z_{i+1} است. توجه کنید که برای هر $i < n$ داریم:

$$-\frac{i}{\sqrt{3}} (z_{i+1} - z_i) \in \left\{ 1, e^{2i\pi/3}, e^{4i\pi/3} \right\}.$$

معادله ۲.۲ در راس v بیان می‌کند که انتگرال گسسته تابع F در امتداد کانتور «مثلثی» گسسته‌ای که از سه وجه پیرامون v می‌گذرد، صفر است (شکل ۵ را بینید). حال فرض کنید که دامنه D هیچ سوراخی ندارد. در این صورت، می‌توان هر کانتور گسسته بسته در D را به مجموعه‌ای از کانتورهای مثلثی مقدماتی تجزیه کرد. از آنجا که انتگرال گسسته تابع مشاهده‌پذیر F در امتداد هر یک از این کانتورهای مثلثی مقدماتی صفر است، می‌توانیم نتیجه بگیریم که انتگرال گسسته F برای هر مسیر کانتور گسسته در D صفر است.

¹Parafermionic observable

²elementary contour



شکل ۵. با بزرگنمایی راس v در شبکه‌ی شش ضلعی، نقاط میان-یالی مجاور p ، q و r را می‌بینیم که به همراه مرکزهای z_0 ، z_1 و z_2 کاتور مقدماتی متناظر را می‌سازند. همه‌ی کاتورهای مقدماتی باهم شبکه‌ی مثلثی دوگان را تشکیل می‌دهند.

به نظر می‌رسد که این ویرگی، گستته‌سازی شده‌ی یکی از ویرگی‌های مختص توابع هولومورفیک است. توابع هولومورفیک توابع مختلطی هستند که مشتق پذیر مختلط نیز باشند؛ و انتگرال کاتور تابع هولومورفیک تا زمانی که کاتور از دور یک «سوراخ» در دامنه نگذرد، صفر است. بنابراین، می‌توانیم رابطه ۲.۲ را به عنوان یک نسخه گستته از هولومورفیک بودن تفسیر کنیم. در این بحث مفهوم گستتگی در مقابل پیوستگی، به این معناست است که متغیر به جای حرکت پیوسته در صفحه مختلط \mathbb{C} روی میان-یال‌های شبکه H پرش می‌کند.

به هر حال، در یک مورد احتیاط لازم است. یکی از ویرگی‌های کلیدی نگاشتهای هولومورفیک این است که مسائل مقدار مرزی، راه حل یکتا دارند — اگر دو تابع هولومورفیک مقادیر یکسانی را در مرز یک دامنه دارند، باید در داخل آن هم برابر باشند. بیایید این را در سطح گستته را بررسی کنیم. تصور کنید که می‌خواهیم یک تابع گستته F را فقط با استفاده از مقادیر مرزی آن و معادله ۲.۲ در اطراف هر راس تعیین کنیم. برای هر میان-یال z یک متغیر مجھول، یعنی مقدار (z) ، و یک معادله برای هر راس وجود دارد. به عنوان مثال، یک مسئله مقدار مرزی گستته برای دامنه سبز D در شکل ۳، ۱۱ مقدار مرزی، ۲۰ مجھول و ۱۷ معادله دارد. توجه کنید که:

$$\text{تعداد میان-یال‌های داخلی} \times 2 + \text{تعداد میان-یال‌های مرزی} = \text{تعداد رؤوس داخلی} \times 3$$

زیرا هر راس مجاور ۳ یال است، و هر میان-یال مجاور ۱ یا ۲ راس داخلی است، وابسته به این که روی مرز یا در داخل دامنه قرار داشته باشد. بنابراین، به طور معمول تعداد مجھول‌ها بیشتر از تعداد معادله‌ها است، و لذا چندین جواب برای این دستگاه معادلات خطی وجود دارد.

به نظر می‌رسد به بسته رسیده باشیم: در مورد حالت گستته، واقعیت این است که صفر شدن جمع کاتور اطلاعات کمی درباره تابع F به ما می‌دهد. به عبارت دیگر، یک تابع که رابطه ۲.۲ را در اطراف هر راس ارضاء می‌کند، می‌تواند به عنوان نوعی تابع «هولومورفیک گستته ضعیف» دیده شود، اما این معادلات به اندازه‌ی مفهوم استاندارد هولومورفیک به ما کمک نمی‌کنند. خوشبختانه، ویرگی صفرشدن جمع‌های کاتور، بی‌معنی نیست. یک تحلیل دقیق از جمع‌های کاتور در امتداد مرز دامنه‌هایی که به خوبی انتخاب شده‌اند، نشان می‌دهد که مقدار $\sqrt{2} + \sqrt{2}$ که در بالا ذکر شد، باید ثابت اتصال شبکه شش ضلعی باشد. این قسمت سراسرتی نیست، و با توجه به پاراگراف قبل، ممکن است شیوه یک معجزه به نظر رسد. دوباره، برای جزئیات بیشتر در مورد این استراتژی به مقاله اصلی [۴] ارجاع می‌دهیم.

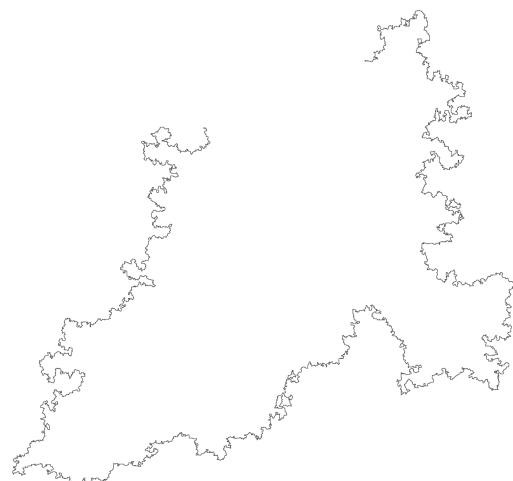
۳. هندسه‌ی SAW روی شبکه‌ی شش ضلعی

محاسبه‌ی ثابت اتصال باید به عنوان یک پله به سمت هدفی بزرگ‌تر در نظر گرفته شود: فیزیکدانان و ریاضیدانان علاقه‌ی زیادی به هندسه‌ی SAW‌های تصادفی بزرگ دارند. لذا باید از سوال ترکیباتی شمارش SAW‌ها فاصله بگیریم، و روی یک سوال هندسی تمرکز کیم: یک SAW بزرگ معمولاً چه شکلی است؟

بگذارید با اشاره به یک مدل تصادفی مرتبط شروع کنیم، که به آن قدم‌زن تصادفی ساده^۱ (SRW) می‌گویند. یک قدم‌زن تصادفی ساده به این صورت به دست می‌آید: همه‌ی قدم‌زن‌های ممکن (و احتمالاً خودمتقطع) روی شبکه‌ی شش ضلعی را در نظر بگیرید که از مبدأ شروع می‌شوند و n -قدمی هستند. از ترکیبات بدست می‌آید، دقیقاً 3^n تا از این قدم‌زن‌ها داریم. یک قدم‌زن تصادفی ساده n -قدمی، انتخاب تصادفی یکنواخت یک نمونه از مجموعه‌ی همه‌ی قدم‌زن‌های تصادفی n -قدمی است که از مبدأ شروع می‌شوند. احتمال انتخاب هر کدام از قدم‌زن‌های تصادفی ساده برابر $\frac{1}{3^n}$ است. به علاوه، برای مشبکه، اندازه‌ی مش δ را در نظر می‌گیریم که طول یک یال — و در نتیجه هر قدم از قدم‌زنی — با افزایش n تغییر کند. بنابراین، فرض کنید Φ_n یک قدم‌زن تصادفی ساده‌ی n -قدمی روی یک شبکه‌ی شش ضلعی \mathbb{H}_{δ_n} ، با ابعاد تغییرافته باشد. ما به رفتار حدی Φ_n وقتی که n زیاد می‌شود، علاقه‌مندیم.

اگر به ازای هر n داشته باشیم $\Gamma_n = \delta_n$ ، آن‌گاه تغییری در ابعاد مشبکه نداده‌ایم، و لذا با افزایش n ، قدم‌زن‌هایی بزرگ‌تر و بزرگ‌تر به دست می‌آوریم. بر عکس، اگر δ_n به سرعت کم شود، آن‌گاه وقتی n به سمت بی‌نهایت می‌رود، قدم‌زن در مبدا جمع می‌شود. حال، اگر δ_n را برابر n در نظر بگیریم، آن‌گاه با رفتن n به سمت بی‌نهایت، قدم‌زن Φ_n ، به عنوان یک موجود تصادفی، به یک خم تصادفی پیوسته، به نام حرکت براونی، همگرا می‌شود. به عنوان نتیجه، به دست می‌آید که «اندازه»‌ی یک SRW به طول n معمولاً برابر \sqrt{n} است.^۲

فیزیکدانان مسئله‌ی متناظر با این مسئله را برای SAW‌ها مطالعه کرده‌اند. فرض کنید Γ_n یک انتخاب تصادفی SAW با احتمال یکنواخت^۱ از میان همه‌ی SAW‌های n -قدمی روی \mathbb{H}_{δ_n} با شروع از مبدأ باشد. مجدداً فرض کنید n بزرگ و بزرگ‌تر می‌شود. در مقاله‌ای که SAW‌ها معرفی شده‌اند، فلوری پیش‌بینی کرده‌است که اگر قرار دهیم $\delta_n = n^{-\frac{1}{4}}$ ، آن‌گاه اندازه‌ی شع تصادفی Γ_n ، و در نتیجه به طور معمول، اندازه‌ی یک SAW n -قدمی، ثابت باقی می‌ماند. مجدداً این نتیجه می‌دهد که با افزایش n ، یک انتخاب یکنواخت SAW روی \mathbb{H} معمولاً در فاصله‌ی δ_n از مبدأ تمام می‌شود. بنابراین، برای $n = 10000$ انتظار داریم فاصله‌ی نقاط پایانی، عددی نزدیک به 1000δ باشد. یک مثال از یک قدم‌زن خودپرهیز 10000 -قدمی روی مشبکه‌ی مربعی، در شکل ۶ نشان داده شده است.



شکل ۶. یک مثال از یک قدم‌زن خودپرهیز 10000 -قدمی روی یک شبکه‌ی مربعی \mathbb{H}^{δ} . فاصله‌ی بین نقطه‌ی شروع و پایان تقریباً برابر 2735 است.

¹ simple random walk

² اندازه‌ی یک قدم‌زنی می‌تواند به عنوان فاصله‌ی بین نقطه‌ی شروع و نقطه‌ی پایان قدم‌زنی تعریف شود.

با وجود آن که امروزه می‌دانیم استدلال اولیه‌ی فلوری اشتباه بوده است، به طرز ناباورانه‌ای، پیش‌بینی او درست به نظر می‌رسد. در حال حاضر، استدلال‌های متفااعد کننده‌ی بیشتری موجود است که پیشنهاد می‌دهد δ_n باید برابر $\sqrt[3]{n}$ در نظر گرفته شود.

برای این انتخاب δ_n ، متغیر تصادفی Γ_n باید به خمی پیوسته مانند Γ همگرا شود، که نقشی مشابه حرکت براونی برای SRW‌ها ایفا کند. این خم تصادفی یک فراکتال تصادفی است به نام تحول شرام-لونر^۱. این شی در سال‌های اخیر، در مطالعات مدل‌های دو بعدی فیزیک آماری «در شرایط بحرانی» ظاهر شده‌است. برای جزئیات بیشتر حول این موضوع پیش‌رفته، شما را به [۷] ارجاع می‌دهم.

نتیجتاً برای مقادیر بزرگ n ، یک قدمزدن تصادفی ساده به طول n به طور متوسط روی نقطه‌ای به فاصله‌ی \sqrt{n} از مبدأ پایان می‌یابد. در حالی که یک قدمزدن خودپرهیز به طول n به طور میانگین در فاصله‌ی $\sqrt[3]{n}$ از مبدأ تمام می‌شود. ادعای اول با دقت ریاضی درک شده‌است، اما ادعای دوم به عنوان یکی از حدس‌های اصلی فیزیک آماری شناخته می‌شود، و بسیار رازآمد باقی مانده‌است. با وجود آن که، مباحث این بخش بی‌ربط به مباحث بخش‌های قبلی به نظر می‌رسد، در واقع کاملاً به هم مرتبط‌اند: یک راه ممکن برای اثبات پیش‌بینی فلوری این است که نشان دهیم مشاهده‌پذیر پارافرمونیک دوباره نرمال شده، زمانی که روی گراف شود، با میل کردن δ به سمت صفر، همگرا می‌شود. نهایتاً خوانندگانی که به دنبال مطالعه بیشتر راجع به قدمزدن خودپرهیز هستند را به [۸]، [۹] ارجاع می‌دهیم.

مراجع

- [1] R. Bauerschmidt, H. Duminil-Copin, J. Goodman, and G. Slade, *Lectures on self-avoiding walks*, Probability and statistical physics in two and more dimensions, Clay Mathematics Proceedings, vol. 15, American Mathematical Society, 2012, pp. 395–467.
- [2] N. Clisby, *Calculation of the connective constant for self-avoiding walks via the pivot algorithm*, Journal of Physics. A. Mathematical and Theoretical **46** (2013), no. 24, 245001.
- [3] N. Clisby and I. Jensen, *A new transfer-matrix algorithm for exact enumerations: self-avoiding polygons on the square lattice*, Journal of Physics. A. Mathematical and Theoretical **45** (2012), no. 11, 115202.
- [4] H. Duminil-Copin and S. Smirnov, *The connective constant of the honeycomb lattice equals $\sqrt{2 + \sqrt{2}}$* , Annals of Mathematics. Second Series **175** (2012), no. 3, 1653–1665.
- [5] P. Flory, *Principles of polymer chemistry*, Cornell University Press, 1953.
- [6] J. M. Hammersley, *Percolation processes: II. The connective constant*, Mathematical Proceedings of the Cambridge Philosophical Society **53** (1957), 642–645.
- [7] G. F. Lawler, O. Schramm, and W. Werner, *On the Scaling Limit of Planar Self-Avoiding Walk*, Fractal Geometry and Applications: a Jubilee of Benoît Mandelbrot, Part 2, Proceedings of Symposia in Pure Mathematics, vol. 72, American Mathematical Society, 2004, pp. 339–364.
- [8] N. Madras and G. Slade, *The Self-Avoiding Walk*, Probability and its Applications, Birkhäuser, 1993.
- [9] P. Mörters and Y. Peres, *Brownian Motion*, vol. 30, Cambridge University Press, 2010.
- [10] B. Nienhuis, *Exact Critical Point and Critical Exponents of $O(n)$ Models in Two Dimensions*, Physical Review Letters **49** (1982), no. 15, 1062–1065.
- [11] W. J. C. Orr, *Statistical treatment of polymer solutions at infinite dilution*, Transactions of the Faraday Society **43** (1947), 12–27.
- [12] R. D. Schram, G. T. Barkema, and R. H. Bisseling, *Exact enumeration of self-avoiding walks*, Journal of Statistical Mechanics: Theory and Experiment **2011** (2011), P06019.

مترجم: مهلا امیری[†] و کیمیا نیغ‌بند*

*دانشجوی کارشناسی ارشد ریاضی، دانشگاه UBC

[†]دانشجوی دکتری ریاضیات فیزیک برلین

¹Schramm–Loewner Evolution

صاحبه با مارینا ویازوفسکا *

کن اونو

چکیده. مدال فیلدز یکی از مهم‌ترین جوایز در دنیای ریاضیات می‌باشد. این جایزه هرچهار سال یکبار به حداقل چهار ریاضی‌دان که دستاوردهای چشمگیری داشته‌اند اهدا می‌شود. این نوشتار، ترجمه مصاحبه خبرنگار کن اونو^۱ با یکی از برنده‌گان این مدال در سال ۲۰۲۲ است.

۲. گفت‌وگو

اونو: شما برای کارتان روی مسئله بسته‌بندی گوی‌ها برنده یکی از چهار مدال فیلدز ۲۰۲۲ شدید — دومین زنی در تاریخ که چنین افتخاری به دست می‌آورد. کمی راجع به پیشینه این مسئله و زمینه‌ی ریاضیاتی‌ای که آن را شکل داده است صحبت می‌کنید؟

ویازوفسکا: مسئله بسته‌بندی گوی‌ها یک سوال بسیار طبیعی هندسی می‌پرسد که پیشینه‌ای بسیار طولانی در پس آن است. حالت سه بعدی این مسئله بسیار مشهور است و با نام حدس کپلر^۲ آن را می‌شناسند. اجازه بدھید نخست مسئله را در این حالت توضیح بدهم. درواقع به نسبت ساده است: یک جعبه بسیار بزرگ داریم و به تعدادی نامحدود گوی‌های صلب هم‌شکل. قصد ما این است که بیشترین تعداد ممکن از این گوی‌ها در داخل جعبه جای‌گذاری کنیم. حالا تصور کنید که جعبه گفته شده آنقدر بزرگ است که به معنایی کل فضا را پوشاند. در این صورت پرسش اینجاست که چگال‌ترین چیدمانی که می‌توانیم برای این گوی‌های نامتقاطع صلب انتخاب کنیم چیست؟ جان کپلر این پرسش را در مقاله‌ای راجع به دانه برف شش‌گون مطرح کرد.

در قرن هفدهم میلادی، زمانی که نظریه اتمی یک موضوع

۱. مقدمه

در پنجم جولای ۲۰۲۲، مارینا ویازوفسکا^۳ ریاضیدان حوزه نظریه اعداد، به عنوان دومین زن در تاریخ موفق به دریافت مدال فیلدز شد. خانم ویازوفسکا که در موسسه تکنولوژی فدرال سویس در لوزان^۴ مشغول به کار است، عمدۀ شهرت خود را به علت کار روی مسئله بسته‌بندی گوی‌ها^۵ در ابعاد هشت و بیست‌وچهار کسب کرده است.

متن رسمی اهدای جایزه ایشان به صورت زیر است: «مدال فیلدز ۲۰۲۲ به مارینا ویازوفسکا تعلق می‌گیرد به علت اثبات ایشان برای اینکه شبکه E_8 چگال‌ترین بسته‌بندی را برای گوی‌ها در بعد هشت برقرار می‌سازد؛ همچنین برای دستاوردهای دیگر او در زمینه مسائل اکسترمالی مرتبط و مسائل درونیابی در آنالیز فوریه..» دستاوردهای دیگر ویازوفسکا شامل همکاری فوق العاده ایشان با هنری کوهن^۶، آبهینا کومار^۷، استفن د. میلر^۸ و دنیلو رادچنکو^۹ می‌شود که درواقع مسئله بسته‌بندی گوی‌ها را در بعد بیست‌وچهار با ارائه چگال‌ترین بسته‌بندی ممکن حل می‌کند.

*این نوشتۀ ترجمه‌ای از مصاحبه‌ی زیر است:

Ono, K. *An Interview with Maryna Viazovska*. *Math Intelligencer* 44, 302–305 (2022). <https://doi.org/10.1007/s00283-022-10225-7>

¹Maryna Viazovska

²EPFL

³The packing-problem

⁴Henry Cohn

⁵Abhinav Kumar

⁶Stéhen D. Miller

⁷Danylo Radchenko

⁸Kepler's conjecture

بهینه‌سازی به نظر ساده‌تری می‌رویم. معمولاً این کار را نه در فضای پیکربندی نقطه‌ها، بلکه در فضای مناسبي از توابع که خطی بودن در آن معنادار است انجام می‌دهیم. درنتیجه وارد حوزه بهینه‌سازی محدب می‌شویم. این حوزه خوشبختانه به نسبت پیشرفته است و تاکنون در حل مسائل بهینه‌سازی هندسی متعددی به کار رفته. برای مثال برای حل مساله بوسه^۱ در بعد هشت به کار گرفته شده، که در سال ۱۹۷۹ توسط دو تیم مستقل از هم حل شد: ولادیمیر لونشتاین^۲ در مسکو و اندرو اودلیزکو^۳ و نیل سلون^۴ در آمریکا. همین روش توسط هنری کوهن و نوام الکیز برای مسئله بسته‌بندی گوی‌ها در فضای اقلیدسی به کار گرفته شد. به لطف مقاله این دو ریاضیدان، من با مسئله بسته‌بندی گوی‌ها آشنا شدم و تصمیم گرفتم روی آن کار کنم.

من بیشتر با مقاله‌ای که بعداً توسط هنری کوهن و نوام الکیز^۵ چاپ شد کار می‌کردم و بله، این روش تا جایی که آن را توسعه دادند، هوشمندانه بود. به جای مستقیم نگاه کردن به مسئله هندسی، ایده این است که تابعی کمکی بسازیم که دسته خاصی از نامساوی‌ها را برقرار می‌کند. این تابع نه تنها خودش این دسته از نامساوی‌ها را برقرار می‌کند بلکه تبدیل فوریه‌اش هم در این مجموعه نامساوی صدق می‌کند. هرگاه قادر باشیم چنین تابع کمکی‌ای را با پارامترهای مناسب پیدا کنیم، می‌توانیم یک کران بالا برای چگالی یک چینش از گوی‌ها ارائه بدهیم. کوهن و الکیز این روش را برای بعدهای سه تا سی‌وشش به کار گرفتند و کران بالاهای صریح به دست آوردند. به طور مستقل، روش مشابهی توسط دیمیتری گوریاچف^۶ به وجود آمد. همه تلاش‌ها کران‌های سابقی که می‌دانستیم را بهبود بخشیدند ولی در کل، انتظار می‌رفت که پاسخ‌های بسیار بهینه‌تری هم برای همه ابعاد فراهم کنند، همه ابعاد به جز: ابعاد هشت و بیست‌وچهار. کران‌های عددی در این دو بعد به شدت نزدیک به چگال‌ترین چینش‌هایی بودند که می‌شناختیم یعنی شبکه E_8 برای بعد هشت و شبکه لیچ^۷ در بعد بیست‌وچهار. چگالی جاگذاری برای این شبکه‌ها در بسیاری از جایگاه‌های اعشاری با کران‌های عددی ما مطابقت داشتند!

بسیار داغ در دنیای علم محسوب می‌شد، این سوال یک ایده بسیار جسورانه بود. ما امروزه می‌دانیم که این نگاه به ماده متراکم به نسبت ساده‌انگارانه است و درنتیجه نتایج مسئله بسته‌بندی گوی‌ها به تهایی کفایت نمی‌کند. در ادامه، مسئله‌های بهینه‌سازی پیچیده‌تری در حوزه مکانیک کوانتومی به میدان می‌آیند که در واقع در آن زمان مورد توجه نبودند. به عنوان یک مسئله در زمینه ریاضیات محض، این سؤال توجه بسیاری از ریاضیدانان را به خود جلب کرده است و در واقع یک مثال خوب از مسئله‌ای بسیار سخت در زمینه بهینه‌سازی هندسی می‌باشد.

این مسئله بیش از سی صد سال حل نشده باقی ماند. بالاخره در اواخر قرن بیستم، توماس هیلز^۸ موفق به حل آن شد. کار او در تاریخ ریاضیات مهم است، به این خاطر که اثبات او از جمله اولین اثبات‌های کامپیوتری است که برای قضیه‌ای به این مهمی پذیرفته شد. بحث‌های متعددی در جامعه ریاضی حول اینکه چطور باید با این اثبات‌ها بروخورد کرد شکل گرفت و از دید من این ماجرا راههای زیادی را در راستای منافع ریاضیات باز کرد. من به همراه کوهن، کومار، میلر و رادچنکو موفق به حل مسئله بسته‌بندی گوی‌ها در ابعاد هشت و بیست‌وچهار شدم.

اونو: شما موجوداتی به نام «توابع جادویی» کشف کردید که وجودشان نقش کلیدی‌ای در راه حل مسئله بسته‌بندی گوی‌ها در ابعاد هشت و بیست‌وچهار داشته است. کمی راجع به تحقیقاتتان درباره این توابع صحبت می‌کنید؟ و راجع به اینکه دقیقاً چه چیزی موجب شد به چیزی دست یابید که قبل از شما دیگران پیدا نکرده بودند؟

ویازوفسکا: وقتی صحبت از مسائل بهینه‌سازی هندسی به میان می‌آید، روشی جهانی نداریم که همه آن‌ها را حل کند. برای مثال، راه حل هیلز برای مسئله بسته‌بندی گوی‌ها در بعد سه، رویکرد مستقیم هندسی‌ای پیش می‌گیرد که طی آن با مطالعه دقیق هندسه سه بعد، مسئله اصلی را به چندین مسئله بهینه‌سازی تبدیل می‌کند و نهایتاً به کمک کامپیوتر آن‌ها را حل می‌کند.

رهیافت دیگری هم وجود دارد که روش برنامه‌ریزی خطی نامیده می‌شود. به طور سریسته می‌توان گفت در این رویکرد به جای مطالعه مستقیم مسئله اصلی، سراغ مسئله

¹Thomas Hales

²The kissing problem

³Vladimir Levenshtein

⁴Andrew Odlyzko

⁵Neil Sloane

⁶Noam Elkies

⁷Dmitry Gorbachev

⁸The Leech lattice

می کردند.

همچنین در رابطه با بعضی آموزگاران و در کل از حیث آشنایی با آدمها در زندگی خوششانس بودم. البته نمی توانم بگویم که جنسیت زدگی در دنیا وجود ندارد. من به طور خاص خوششانس بودم که در اوکراین متولد شدم، که چنین پدر و مادری داشتم و آموزگاران مناسبی سر راهم قرار گرفتند.

اونو: آیا به طور خاص افرادی هستند که در زندگی حرفه ای شما نقشی اساسی داشته باشند؟ اگر بله، کمی راجع به ماهیت تاثیرشان می گویید؟

بله، خیلی ها به من کمک کردند، با شروع از اولین معلم ای که داشتم. اولین معلم من، که به من خواندن و نوشتمن یاد داد، زنی بسیار سختگیر بود. او به من مفهوم اخلاق کاری و دوری نکردن از کارهای دشوار را آموخت. او، به ادبیاتی، یک «زن آهنین» بود ولی من فکر میکنم در عین حال بسیار مهربان بود. معلمی که در عین سختگیری، به دانش آموزان خود اهمیت می داد. برای مثال، به یاد دارم که او ساعتی زودتر به کلاس می آمد و با دانش آموزانی که عملکرد مناسبی نداشتند تمرين می کرد. البته، زمانی که بچه بودم در نظرم این کار وحشتناک بود. اما الان متوجه هستم که او فوق العاده بود و در حالی که مجبور نبود، صرفاً برای کمک به دانش آموزان ش ساعتی زودتر به سرکار می آمد. بعد از اتمام دوره ابتدایی، اولین معلم ریاضی من – که اتفاقاً فکر کنم دوست نزدیک معلم اول ابتداییم بود – یک «زن پولادین» بود. او هم از نظر سختگیری و هم از نظر تدریس ریاضیات عالی بود. وقتی به گذشته نگاه می کنم، می بینم که ریاضیاتی که در آن بازه مطالعه می کردیم مقدماتی بود و شاید چندان جالب نبود. متوجه هستم که چیزی که در چهارم یا پنجم دبستان مطالعه می کیم آنقدر هیجان انگیز نیست اما با این حال به یاد دارم که در همان زمان هم به ریاضیات علاقه زیادی داشتم. معلم ما رویکردی بسیار ساختارمند داشت، رویکردی که حس می کنم به نوعی در بعضی کتاب های مدرن آموزش ریاضیات به دانش آموزان وجود ندارد.

بعد از هفت سال تحصیل عمومی، برای تحصیل در رشته ریاضیات و فیزیک در مدرسه ای خاص، از من دعوت شد. در آنجا من با آموزگارانی حقیقتاً حیرت انگیز ملاقات کردم. از بین آنها دو فرد شگفت انگیز وجود داشتند که چیزی فراتر از صرفاً معلم برای من بودند. این دو فرد مشابه دانشمندان فکر می کردند و مطالعه زیر نظر آنها

برای اثبات مسئله در این ابعاد، چالش این بود که توابع کمکی ای پیدا کیم که مطابقت داشته باشد و در نتیجه بهینه بودن این ساختارها را اثبات کنند. اگر درست به یاد داشته باشم، استفن میلر^۱ آنها را «توابع جادویی» نامید. او دقیقاً به خاطر نمی آورد اما حدس میزند به دلیل اینکه پیدا کردن این توابع مشکل است، ایده این نام گذاری را داده. به این ترتیب همه کار پیدا کردن این دسته از توابع بود.

بعد از اینکه این توابع را پیدا می کنیم، همه چیز به راحتی و خوبی پیش می رود. دستاورد من در این حوزه این بود که موفق به خاطر نمی آورد اما حدس میزند به دلیل اینکه توابع شدم. اطلاعات عددی قویاً مؤید وجود چنین توابعی است. من فرمول صریحی برای این تابع در بعد هشت ارائه دادم و بعد مشخص شد که این تابع در فضای توابع شوارتز یکتاست. درنتیجه با یک موجود یکتای ویژه سروکار داریم. می دانید، وقتی یک نظریه اعداد دادن با موجودی یکتا و ویژه مواجه می شود، زنگی به صدا می آید: باید یک فرمول زیبا و صریح برای آن وجود داشته باشد. در اینجا شهود من به درستی کار کرد. در واقع، یک فرمول به نسبت ساده، خوب و صریح برای تابع جادویی وجود دارد و خاستگاه آن فرم های مدلولار می باشد.

اونو: کی متوجه علاقه خود به ریاضیات شدید؟ آیا فکر می کنید جنسیت تان موانعی بر سر شکل گرفتن علاقه تان به ریاضیات قرار داد؟

ویازوفسکا: من از اول ابتدایی به ریاضیات علاقه داشتم. زمانی که باد گرفتیم چگونه بخوانیم، چگونه بنویسیم، چگونه بشماریم، من به شمردن بسیار بیشتر از دو مورد دیگر علاقه مند بودم. البته بعدها فهمیدم که توانایی خواندن و نوشتمن هم برای یک ریاضیدان بسیار بالا همیت است. فکر می کنم همین باعث شد برای اولین بار گمان برم که شاید ریاضیات رشته مناسبی برای من است.

آیا جنسیت من موانعی بر سر راهم قرار داد؟ من در حال دنبال کردن علاقه ام بودم و در آن زمان فکر نمی کردم که اینطور باشد. اما حالا می دانم که اینطور هست. حالا که بیشتر راجع به دنیا می دانم، متوجه هستم که در واقع من بسیار خوششانس بودم. بسیار خوششانس بودم که والدینم من را مجبور به تحصیل در رشته هایی که مردم باور دارند برای دخترها مناسب تر است نکردند. آنها من را از مطالعه ریاضیات، صرفاً به این علت که حوزه ای مرد محور است، نویمید نکردند. به علاوه من خوششانس بودم که آنها کنیکاوی من برای علم و ریاضیات را حمایت

^۱ Stehen Miller

گفت و گوهای بسیار زیادی باهم داشتیم ولی متناسبانه، اون چند سال پیش درگذشت. او یک فرد مهربان و ارزشمند بود که به دانشجویاهیش اهمیت می‌داد. او یک نسل کامل از جبردان‌ها در دانشگاه کیف پرورش داد.

می‌دانی کن، وقتی به بون^۳ رفت و آمد می‌کنی، متوجه میشوی که وقتی مردم از «علم» حرف می‌زنند، بعضی اوقات به برج عاج^۴ اشاره می‌کنند. موسسه مکس پلانک^۵ در بون قطعاً یک تجسم عینی از همچین برجی بر روی زمین است. مکس پلانک یک مکان سریست که در مرکز شهر مخفی شده است و در یک مرکز پست قرار دارد. مردم باید راجع به آن مانند قطار^۶ ۹ هری پاتر^۷ در کینگز کراس^۸ فکر کنند، یک ورودی مخفی به سکو که فقط جادوگران می‌توانند به آن وارد شوند. البته به جای جادوگران، این مکان با ریاضی‌دانان پر شده است. در آن جا یک فضای عالی به وجود آمده است، چرا که همه ریاضیات را دوست دارند. در این مکان، که در آن ریاضیدانان گنجینه‌های خود را باهم به اشتراک می‌گذارند، یک جادوگر خاص وجود دارد به نام دان زگیر^۹. او نه بیست و چهار ساعت، نه بیست و پنج ساعت، بلکه شاید بیست و شش ساعت در روز کار می‌کند. من بسیار خوشحال هستم که در طی ماجراجویی جادویی‌ای که در دنیای ریاضی داشتم با استاد راهنمای مقطع دکتری م مقطع دکتری م یعنی دان زگیر، آشنا شدم.

اونو: شما در اوکراین متولد و بزرگ شدید ولی در حال حاضر در سویس زندگی می‌کنید. کمی راجع به تاریخ ریاضیات اوکراین و تاثیری که فکر می‌کنید مدار فیلدز شما بر هموطن‌هایتان خواهد گذاشت صحبت می‌کنید؟

ویازوفسکا: اوکراین کشوری با سنت‌هایی بسیار پایدار در ریاضیات می‌باشد. بسیاری از ریاضیدانانی که دستاوردهای ریاضی مهمی داشته‌اند، اهل اوکراین هستند. در بسیاری از موارد ریاضیدانان پس از کسب تحصیلات اولیه در اوکراین، در کشورهای دیگر دنیا به کار مشغول شدند.

اتفاقی که در اوکراین در حال رخ دادن است یک تراژدی فاجعه‌بار می‌باشد و دنیا در حال تماشاست. بزرگ‌ترین تراژدی ممکن از بین رفتن جان این‌همه انسان است. شاید این یک کلیشه باشد ولی به نظرم درست است که

یک ماجراجویی تمام عیار بود. آن‌ها، با ارائه مسئله‌هایی خارج از چهارچوب، ما را برای به صورت تیمی برای شرکت در المپیادهای فیزیک و ریاضی آماده می‌کردند. آن‌ها مباحثی درس می‌دادند که در محتوای معمول درس‌ها نبود. من و بقیه دانش آموزان، از شرایط این مدرسه خاص نهایت استفاده را می‌بردیم. ما برای مطالعه انگیزه داشتیم و همین یک محیط استثنایی برای همه به وجود آورده بود. البته آن قدرها راحت نبود و به نسبت فضای رقابتی ای محسوب می‌شد. ولی در عین حال، فکر می‌کنم این از جمله انتخاب‌هایی است که در زندگی داریم و من خوش‌شانس بودم که این امکان را داشتم که راجع به موضوعاتی که برایم بالهمیت بودند با افرادی صحبت کنم که متقابلاً به این موضوعات علاقه‌مند بودند.

ایگور اسچوچوک^۱ استاد من در دانشگاه کیف بود. او به من آنالیز ریاضی درس داد و مشوق من برای شرکت در رقابت‌های ریاضی بود. او به نوعی یک دنیا را برای من به وجود آورد. تجربه شرکت در رقابت‌های ریاضی، یک تجربه عالی تیمی بود اگرچه، احتمالاً من آن‌طور که امیدوار بودم عملکرد خوبی نداشتیم. من به همین منوال به تحصیل در دانشگاه ادامه دادم و فرصت فوق العاده‌ای داشتم که با افراد برجسته متعددی ملاقات کنم. ایگور اسچوچوک اولین فردی بود که، با اینکه من هنوز یک دانشجو بودم، من را تشویق کرد تا درباره چند مساله تحقیقاتی فکر کنم. مطمئن نیستم که این فعالیتی است که همه دانشجویان باید به آن پردازنند. روش‌های متنوعی برای یادگیری و رشد به عنوان یک ریاضی‌دان وجود دارد اما برای من، این فعالیت بسیار سازنده بود. شاید این به این خاطر است که من آدم بسیار صبوری نیستم و بعضی اوقات حوصله‌ام سر می‌رود. تحقیق در ریاضیات، وقتی که می‌توانید خلاق باشید، شبیه تنفس هوای تازه است. شاید هیچ جواب درستی وجود ندارد یا لاقل هیچ کس جواب درست را نمی‌داند. اینکه اولین کسی باشید که جواب را کشف می‌کند بسیار هیجان‌انگیز است. سرگی اوسینکو^۲ فرد دیگری است که در زندگی آکادمیک من بسیار تاثیرگذار بود. او به من جبر و چیزهای بسیار زیادی راجع به آن یاد داد. شاید بتوان گفت اینکه من یک نظریه اعداددان شدم و نه یک آنالیزدان به خاطر اوست. ما

¹Igor Schevchuk

²Sergiy Ovsienko

³Bonne

⁴اصطلاحی برای تشبیه به مکانی است که در گذشته مردم به آن میرفتند تا در آنجا به دور از جامعه مشغول فعالیت‌ها و جستجوهای ذهنی - علمی باشند.

⁵Max Planck Institute

⁶Harry Potter

⁷King's Cross

⁸Don Zagier

شما چطور تغییر خواهد کرد؟ به نظر می آید که بدن این مدار کاملاً زندگی شما را تغییر داده است.

ویازوفسکا: بله، الان، زندگی من تغییر کرده است اما در آینده، امیدوارم اینطور نباشد. امیدوارم که در یک ماه آینده به زندگی معمول خودم به عنوان یک ریاضی دان و استاد ریاضیات برگردم. بعضی از امتیازهایی که همراه با مدار فیلدز می آیند را دوست خواهم داشت ولی امیدوارم زندگی من به حالت عادی برگردد و من به تدریس و تحقیقاتم برگردم. احتمالاً در آینده، کارهای بیشتری برای انجام دادن داشته باشم چراکه احتمال میدهم مردم به عنوان نوعی رهبر با نظرات حائز اهمیت به من نگاه کنند. از نظر تاریخی، بعضی از برندهای مدار فیلدز پس از بدن این جایزه، به طور کامل زندگی خود را تغییر دادند. بعضی‌ها کاملاً سمت و سوی تحقیقات یا حتی رشته علمی خود را عوض کردند و شروع به انجام کارهای کاملاً متفاوتی کردند. من چنین برنامه‌هایی ندارم. من از یک ریاضیدان بودن بسیار خوشحال هستم.

در حال حاضر، باید در مصاحبه‌های زیادی شرکت کنم و دعوت‌های بیشتری برای سخنرانی در مراسم‌هایی که چندان علمی هم نیستند دارم. اما دوست دارم همچنان یک دانشمند باقی بمانم. هر از چندگاهی، احتمالاً مجبور به ترک برج عاج بشوم و با افرادی که هزینه‌های تحقیق ریاضی را تامین می‌کنند صحبت کنم. توضیح اینکه چرا کاری که ما می‌کنیم مهم است موضوع بالاهمیت‌ای است. درنتیجه از حالا به بعد احتمالاً چندین بار چنین کاری خواهم کرد.

اونو: چه توصیه یا پیامی برای دانشجویان جوان در ریاضیات دارید؟

ویازوفسکا: خطاب به دانشجویان رشته ریاضی می‌گوییم اینکه ریاضیات را مطالعه کنید مهم است. دانشجویان باید علاقه‌ای که در وجود خود حس می‌کنند را دنبال کنند. دانشجویانی که به هر دلیل دیگری سراغ تحقیقات در ریاضی آمده‌اند بهتر است به گزینه‌های دیگری به عنوان شغل فکر کنند. فقط به این خاطر که بقیه افراد، مثل والدین، از شما می‌خواهد کاری را انجام بدید، آن کار را انجام ندهید. هر کس باید شور و اشتیاق خودش را دنبال کند. این سخت‌ترین نوع توصیه‌ای است که یک استاد می‌تواند بدهد.

امیدوارم اکثر کسانی که به عنوان دانشجو وارد دانشکده‌های ریاضی می‌شوند به این خاطر آمده باشند که ریاضیات را

جنگ بهترین آدم‌ها را از ما می‌گیرد. به عنوان یک اوکراینی، این ماجرا برای من دردآورد است. ما نمی‌توانیم به این راحتی صرفاً تماشاچی باشیم. اتفاقات حال حاضر اوکراین، علاوه بر از دست رفتن جان بسیاری از انسان‌ها، برای بشریت و فرهنگ وحشتناک است. برای مثال، شهر خارکیف آهسته‌آهسته در حال پاک شدن است. تا به همین لحظه، نزدیک به ده درصد از ساختمان‌های خارکیف نابود شده‌اند. البته اوضاع خارکیف به اندازه اوضاع شهرهای دیگر (مارپیول^۱، سیویرودونتسک^۲) که در آن‌ها بیش از نود درصد ساختمان‌ها نابود شده‌اند، تراژیک نیست. متاسفانه حمله‌های موشکی و گلوله‌باران‌ها هر روز ادامه دارند. شهرهای متعدد دیگری در اوکراین وجود دارند که از بسیار نظر تاریخی، از جمله تاریخ ریاضیات و علم، اهمیت دارند اما در حال ویرانی هستند. شاید اینجا نقطه‌ای از بحث باشد که زیادی احساساتی می‌شوم. بله، پس شاید بهتر است به موضوعات خوش‌بینانه‌تری بپردازم.

امیدوارم که این مدار حال بعضی اوکراینی‌ها را بهتر کند. در این برهه دشوار، شاید دریافت خبرهای خوب کمک‌کننده باشد. آرزوی من برای اوکراین این است که به طرقی از خودش محافظت کند. من خواستار بازگشت صلح به سرزمین‌مان هستم و پس از آن، یک نوسازی مناسب. ما علم را فراموش نخواهیم کرد و شاید جایزه من یادآوری بر این باشد. دوست دارم متواضع باشم ولی شاید جایزه من به اوکراینی‌ها یادآوری کند که عملکرد آن‌ها در علم ممتاز است. اوکراینی‌ها لایق بهترین فرصت‌ها در تحصیل هستند تا به جوان‌ها این امکان را بدهد که به دنبال علم بروند. تاریخ اوکراین یکی از غم‌انگیزترین موضوعاتی است که می‌توان خواند. این دید می‌تواند منجر به نوعی نومیدی و بی‌اعتقادی شود. بردههای سخت در تاریخ اوکراین همواره وجود داشته‌اند، همانند وقتی که صنعت‌زادایی به سراغ کشورمان آمد و افرادی که به طور جدی در رشته‌های علمی و تکنولوژی آموزش دیده بودند، مجبور به پیدا کردن شغل‌های جایگزین شدند. امیدوارم خبر جایزه من به اوکراینی‌ها برای تحمل و غلبه بر این آسیب کمک کند و آن‌ها از آن برای بازسازی تعهد ما به علم بهره ببرند. متاسفانه، هیچ یک از این‌ها چیزی نیست که در حال حاضر ذهن آن‌ها را مشغول کرده است. همه ذهن مشغولی‌ها راجع به جنگ و دفاع از کشورمان می‌باشد.

اونو: فکر می‌کنید حالا که برندۀ مدار فیلدز شدید، زندگی

¹Mariupol

²Sievierodonetsk

می‌شوند و شغل‌های واقعی پیدا می‌کنند. هردو این‌ها مهم هستند. من می‌بینم که ارزش ریاضیات دارد بیشتر و بیش‌تر شناخته می‌شود و درنتیجه امیدوارم ماجراجویی ریاضیاتی دانشجویان امروز بسیار هیجان‌انگیز از آب در بیاید. خطاب به همه دانشجویان ریاضی می‌گوییم که من قویاً باور دارم ریاضیاتی که امروز در دانشگاه یاد می‌گیرند برای جامعه ما مفید خواهد بود. درنتیجه توصیه من این است که دانشجوی خوبی باشد و علاقه‌های درونی خود را دنبال کنید.

دوست دارند. زندگی یک ریاضیدان همواره چندان راحت نیست. باید برای انواع شگفتی‌ها و پیچیدگی‌ها آماده بود. البته فکر می‌کنم ریاضیات برای آن‌هایی که به آن علاقه‌مندند لذت‌های ذهنی ساده و نابی به ارمغان می‌آورد. در حوزه کاربرد، من فکر می‌کنم ریاضیات یک حرف به شدت مفید است. به دانشجویان توصیه می‌کنم آگاه باشند که بازار کار ریاضی می‌تواند پیچیدگی‌های خودش را داشته باشد. بعضی دانشجویان ریاضیدان‌های محقق و استاد دانشگاه خواهند شد و بعضی دیگر وارد «دنیای واقعی»



شکل ۱: تصویری از مارینا ویازوفسکا.

مترجم: نیکی حسنی[†]

[†]دانشجوی کارشناسی ریاضی، دانشگاه صنعتی شریف



مکاتبات فرگه و راسل*

بخش دوم

ترجمه‌ی ساجد طبیی

sadjad.tayebi@gmail.com

چکیده. در شماره‌ی پیشین مجله به ۲۰ نامه‌ی ابتدایی از ۲۰ نامه‌ی فرگه و راسل پرداختیم. در این مقاله ۲ نامه‌ی بعدی ترجمه شده‌است. ابتدائاً مقدمه‌ی ویراستار کتاب، Brian McGuinness، بر بخش راسل-فرگه را می‌خوانیم. در شماره‌های بعدی مجله به باقی نامه‌ها خواهیم پرداخت.

۱. مقدمه‌ی ویراستار

برتراند راسل (۱۸۷۲-۱۹۷۰) از ۱۹۰۲ تا ۱۹۱۲ با فرگه مکاتبه داشت، گرچه بیشتر مکاتبات مربوط‌اند به سال‌های ۱۹۰۲-۱۹۱۲. مکاتبات با اعلام آن چه امروزه به عنوان پارادوکس راسل شناخته می‌شود توسط راسل آغاز می‌شوند، و بیشتر آن‌ها ناظر اند بر راه حل‌های مختلفی که راسل برای پارادوکس پیش می‌نهد و فرگه آن‌ها را رد می‌کند. اما در آن‌ها به اغلب مفاهیم محوری فلسفه‌ی زبان فرگه نیز پرداخته می‌شود: مفاد و مرجع، شیء و مفهوم، صدق و کذب، جمله و رده. راسل زمانی پارادوکس را کشف کرد که مهم‌ترین اثر فرگه در شرف اتمام بود: در آستانه‌ی انتشار جلد II قوانین پایه‌ای اش. آثار اصلی راسل هنوز منتشر نشده بود: او در زمان این کشف به آماده‌سازی اصول ریاضیات برای انتشار مشغول بود. تمام نامه‌های راسل به فرگه به زبان آلمانی نوشته شده‌اند. دست‌کم یک نامه که در سال ۱۹۱۲ نوشته شده‌است امروز مفقود شده. نامه‌ی اول راسل (نامه‌ی ۱) و جواب مشهور فرگه به آن (نامه‌ی ۲) پیش از این به انگلیسی منتشر شده‌اند. ر.ک. به

Jean van Heijenoort (ed.) (1967) *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*. Cambridge. Mass. 1967.

از میان تمام نامه‌های فرگه به راسل، تنها اصل نامه‌ی آخر (نامه بیستم) باقی مانده‌است. راسل، بر این اساس که آن را کاملاً شخصی می‌دانسته، آن را پیش خودش نگاه داشته بوده‌است. باقی نامه‌ها برای شولز^۱ فرستاده شده‌بودند و اکنون تنها فتوکپی آن‌ها در اختیار است.

۲. نامه‌ی سوم: راسل به فرگه

فرایذرز هیل

هسلمر

۱۹۰۲/۰۶/۲۴

همکار عزیز،

از نامه‌تان و این که آثارتان را برایم ارسال کردید بسیار ممنونم. آن چه را در آن نامه مفقود شده‌بود دوباره ارسال می‌کنم. اشتباه صفحه ۷ را در مفهوم‌نگاشت شما از پیش اصلاح کرده بودم؛ اما همان‌طور که گفتید این اشتباه مطلقاً هیچ تالی فاسدی ندارد.

*این نوشته ترجمه‌ی بخشی از کتاب زیر است:

Frege, G. (1980) *Philosophical and Mathematical Correspondence of Gottlob Frege*. University of Chicago Press.

^۱H. Scholz

به نظرم مفاهیم به طور کلی می‌توانند متفاوت باشند، و تناقض تنها وقتی نتیجه می‌شود که آرگومان \vdash [یک تابع] خود^۱ تابعی از آن تابع است، یعنی اگر تابع و آرگومان نتوانند مستقل از هم تغییر کنند. φ در تابع $((\varepsilon)\varphi)$ یگانه متغیر است و آرگومان $(\varepsilon)\varphi$ خودش (آن گونه که عموماً بیان می‌شود) تابعی از φ است. به نظر می‌رسد توابعی به شکل $\{F(\varphi)\}$ ، که در آن‌ها F ثابت و φ متغیر است، حتماً به ازای هر مقداری از φ مجازند، گرچه وقتی بحث از مصدق است خطرناک‌اند. من آنها را فرم‌های درجه‌ی دوم می‌خوانم: کسی در واقع ممکن است مایل باشد به سیاق موهومی در جبر^۲ به معروفی موهومی در منطق پردازد.^۳ با چنین توابعی ما به محض مقدار دادن به φ تابعی اشباع‌شده خواهیم داشت؛ با این حال، آن‌ها نه توابعی مرتبه اول‌اند، و نه آرگومان‌های ثابت دارند. تابع $(\varphi)\varphi$ — به تناقضی مشابه تناقض برآمده از $((\varepsilon)\varphi)\varphi$ — می‌انجامد.

این مسیری بود که من را به تناقض رساند. همان‌طور که قطعاً می‌دانید، کانتور اثبات کرده است که بزرگترین عدد وجود ندارد. اثبات او چنین است:

$$R\varepsilon 1 \rightarrow 1 \cdot \ddot{\varrho} \supset \text{Cls}' \varrho \cdot w = \varrho \cap x_3(x \sim \varepsilon \iota \ddot{\varrho} x) \supset_R w \sim \varepsilon \varrho \supset \text{Nc}' \text{Cls}' \varrho \succ \text{Nc}' \varrho^\dagger$$

(دقیقاً این است که مهم‌ترین بخش اثبات است).^۴ حال مفاهیمی وجود دارد که مصدق آنها شامل همه چیز است؛ بنابراین این مفاهیم باید واجد بزرگترین عدد باشند. تلاش کردم رابطه‌ای یک-به-یک میان تمام اشیاء و تمام رده‌ها برقرار کنم؛ با کاریست برهان کانتور بر این رابطه‌ی خاص خود، متوجه شدم که در حالی که تمام رده‌ها شمرده شده‌اند، رده‌ی $\text{Cls} \cap x_3(x \sim \varepsilon x)$ بیرون مانده است. من تا کنون حدود یک سال راجع به این تناقض تأمل کرده‌ام؛ فکر می‌کنم یگانه راه حل این است که تابع و آرگومان باید بتوانند مستقل از هم تغییر کنند.

آنچه شما در ص. ۳۷ می‌گویید، این که یک تابع‌نشانه هرگز نمی‌تواند جای یک نام خاص را بگیرد (دارم از قوانین پایه‌ای سخن می‌گویم)، به مشکلی فلسفی می‌انجامد. خیلی خوب می‌دانم که چه دلایل خوبی به نفع این دیدگاه می‌توان یافت؛ با این حال این خودمتناقض است. چرا که اگر φ یک نام خاص باشد، « φ هرگز نمی‌تواند جای یک نام خاص را بگیرد» جمله‌ای کاذب است، و در غیر اینصورت اصلاً جمله نیست. اگر چیزی بتواند وجود داشته باشد که یک شیء نباشد، آنگاه این واقعیت را نمی‌توان بدون تناقض بیان کرد؛ چرا که در این حکم، آن چیز مورد بحث یک شیء خواهد شد. بنابراین من تردید دارم که آیا φ در x را بتوان اصولاً چیزی در نظر گرفت. اینک همانا در منطق فلسفی غوطه‌ور شده‌ایم.

شما در ص. ۴۹ می‌گویید که $\Delta = \Gamma$ مرجع دارد اگر Γ و Δ نام‌هایی خاص برای گستره‌های مقادیر یا نام‌هایی برای

^۱ [ترجم فارسی]: واژه آلمانی‌ای که فرگه در اینجا به کار می‌گیرد (*algebraischen*) است که، همان‌طور که در ص. ۱۲۰.

D. Bell (1981), Gottlob Frege: *Philosophical and Mathematical Correspondence*. Philosophical Books, 22: 117–121.

اشاره شده است، در ترجمه‌ی انگلیسی به اشتباه به [=حساب] برگردانده شده است.

^۲ ر.ک. به ص. ۱۰۴، ۱۰۷، ۱۰۲، ۵۱۲، و ۵۱۴ از

B. Russell, *The Principles of Mathematics* (Cambridge 1903; 2nd ed. London 1937)

^۳ به عنوان طرحی از برهان کانتور، فرمول راسل به تنهایی چندان قابل فهم نیست. با وجود این، اگر آن را با ارائه‌ی راسل از این برهان در B. Russell, ‘On Some Difficulties in the Theory of Transfinite Numbers’, *Proceedings of the London Mathematical Society*, series 2, vol. 4 (1907), part 1 (issued March 7, 1906), pp. 29–53

مخصوصاً در ص. ۳۲، و همچنین با طرح راسل از این برهان در بخش ۳۴۹ [کتاب] اصول او مقایسه کنیم، می‌توان محتوای این فرمول را این‌چنین بازسازی کرد:

مطابق نمادگذاری پثانو و کاربرد راسل از حروف متناظر در مقاله فوق‌الذکر و در اصول، φ دامنه‌ی عکس رابطه φ -به-یک R است؛ φ رده‌ی Cls' زیرده‌های φ است؛ φ عدد اصلی φ است؛ و نشانه‌ی φ ، همچون در پثانو، به عنوان نشانه‌ی شمول، میان نشانه‌های رده‌ها استفاده می‌شود، یعنی مانند نشانه‌ی امروزی φ . با این اوصاف، فرمول فوق می‌گوید که عدد اصلی رده‌ای از زیرده‌های رده‌ی φ بزرگ‌تر است از عدد اصلی خود φ ، زیرا در هر تناظر یک-به-یک R میان φ با رده‌ای از زیرده‌های φ (و مشخصاً، با رده‌ی همهی زیرده‌های φ) رده‌ی w شامل تمام عناصر φ که عناصری از دامنه‌ی R نیستند، به عنوان عنصری از دامنه‌ی R ظاهر نمی‌شود. چرا که اگر w به عنوان متناظر عنصر x از φ ظاهر شود، آنگاه از یک سو، فرض $x \in w$ به واسطه‌ی شرط معروف $w \neq Rx$ و بنابراین، به دلیل فرض $w = Rx$ ، به $x \notin w$ ، به $x \in w$ (و نتیجه به نقیض خودش) می‌انجامد، در حالی که از سوی دیگر، فرض $w \neq x$ ، اولاً به واسطه‌ی $w = Rx$ به $x \notin w$ و در پی آن، به همراه $x \in w$ که همواره معتبر است و به واسطه‌ی شرط معروف w ، به $x \in w$ (و نتیجه دوباره به نقیض خودش)، و بنابراین، با لحاظ کدن همهی این‌ها با هم، به تناقض می‌انجامد.

اما، اگر این بازسازی درست باشد، آنگاه « $\varepsilon \varrho$ » را میان دو علامت استلزم در فرمول راسل باید « $\varepsilon \varrho$ » بخوانیم.

[ترجم فارسی]: در ترجمه‌ی انگلیسی، دو فرمول جمله آخر اشتباهاً جابجا نوشته شده‌اند. در ترجمه‌ی فارسی جمله مطابق با اصل آلمانی اصلاح شده است. با تشکر از آرش ابازدی.

ارزش‌های صدق باشند. با این حال، در صفحات قبل توضیحی راجع به $\Delta = \Gamma$ در موردی که یکی از آن‌ها نامی برای یک گستره‌ی مقادیر و دیگری نامی برای یک ارزش صدق باشد نیافتم، جز در موردی که گستره‌ی مقادیر مورد بحث متشكل از همه چیز یا هیچ چیز است. اما گمان می‌کنم در این مورد درست متوجه منظورتان نشده‌ام.^۱

تا کنون فقط مفهوم‌نگاشت و قوانین پایه‌ای شما را خوانده‌ام؛ بهزودی خواندن آثار دیگر را شروع می‌کنم.

با احترام،
برتراند راسل

^۱ این نمادها در *Revue de mathématiques* VII, 2 توضیح داده شده‌اند.

۳. نامه‌ی چهارم: فرگه به راسل

ینا
۱۹۰۲/۰۶/۲۹

همکار عزیز،

نامه‌ی مورخ ۱۲۴۰م شما و مقالاتتان را دریافت کردم؛ بابت آنها بسیار ممنونم.

در مورد تناقضی که یافته‌اید، شاید گفته‌تان راجع به آن را درست متوجه نمی‌شوم. به نظر می‌رسد می‌خواهید برای اجتناب از تناقض فرمول‌هایی به شکل «((ε)φ(ε))φ» را ممنوع کنید. اما اگر نشانه‌ای برای مصدقایک مفهوم (یک رده) را به عنوان یک نام خاص دارای مرجع پذیرید و در نتیجه یک رده را چونان یک شیء به‌رسمیت بشناسید، آنگاه خود این رده باید یا تحت آن مفهوم قرار بگیرد یا خیر؛ طرد شق ثالث. اگر رده‌ی ریشه‌های دوم ۲ را به‌رسمیت بشناسید، آنگاه گزینی از این پرسش ندارید که آیا این رده یک ریشه دوم ۲ است یا خیر. اگر به نظر برسد که به این پرسش نه می‌توان پاسخ مثبت داد و نه پاسخ منفی، معنای آن این خواهد بود که نام خاص « $(\varepsilon^2)\varepsilon$ » فاقد مرجع بوده‌است. یا این که آیا باید قائل شد که گستره‌های مقادیر (تصادیق مفاهیم، اعداد) به مثابه نوعی خاص از اشیاء چنان‌اند که محمول‌هایی خاص را نه می‌توان به آنها نسبت داد و نه از آنها سلب کرد؟ این نیز قطعاً به مشکلات عمدۀ‌ای می‌انجامد.

راجع به تردیدهای شما راجع به گفته‌ی من که یک تابع‌نام هرگز نمی‌تواند جای یک نام خاص را بگیرد، باید تمایز قاطعی میان یک نام یا نشانه و مرجع آن بگذاریم. وقتی نامی را در یک جمله به کار می‌بریم، نه از این نام بلکه از شیئی که به آن اشاره می‌کند سخن می‌گوییم. اما پیش می‌آید که بخواهیم از خود نام هم سخن بگوییم؛ در این صورت آن را درون علائم نقل قول قرار می‌دهم. برای نشان دادن اشباع‌نشده بودن تابع‌نام‌ها، این بار بگذارید جایگاه آرگومان را خالی بگذارم. فلذا می‌توانم بگویم:

((۳ + ۴)) یک تابع‌نام است.

((۳ + ۴)) هرگز نمی‌تواند جای یک نام خاص را بگیرد.

شما درست می‌گویید که:

((اگر ε یک نام خاص باشد، (ε) هرگز نمی‌تواند جای یک نام خاص را بگیرد) جمله‌ای کاذب است))؛

اما اشتباه ادامه می‌دهید که:

((و در غیر اینطورت اصلاً جمله نیست.))

^۱ به نظر می‌رسد راسل قرارداد فرگه در بخش ۱۰ از جلد اول قوانین پایه‌ای حساب (ص. ۱۷) را که گستره‌ی مقادیر $(\varepsilon) = a$ گذب است، به اشتباه چنین فهمیده‌است که این دو گستره مقادیر متشكل از «همه چیز یا هیچ چیز»، یعنی اولی همه چیز و دومی هیچ چیز، است. فرگه در نامه بعدی این اشتباه را تصحیح می‌کند.

^۲ از قرار معلوم ارجاع به

‘Sur la logique des relations avec des applications à la théorie des séries’, *Rivista di matematica* (= *Revue de mathématiques*) 7 (1900-1), pp. 115-48 است، اما همه نمادهایی که در اینجا استفاده شده‌است آنجا توضیح داده شده‌اند.

درستاش این است که بگوییم:

اگر « \forall » نامی خاص نباشد، آنگاه « \forall هرگز نمی‌تواند جای یک نام خاص را بگیرد» جمله نیست.

در اینجا « $\forall \exists$ » — با دو دسته علامت نقل قول — جای « \forall » را می‌گیرد. در حالی که « $\exists \forall$ » تابع نام است، « $\forall \exists$ » نام خاص است، و مرجع آن عبارت است از تابع نام « $\exists \forall$ ». در جمله «چیزی یک شیء است»، واژه «چیزی» جای یک آرگومان از نوع اول را می‌گیرد و نشانه‌ای برای نام خاص است. بنابرین، هر چه را که به جای «چیزی» بگذاریم، همواره به جمله‌ای صادق می‌رسیم؛ چرا که یک تابع نام نمی‌تواند جای «چیزی» را بگیرد. در اینجا خود را در موقعیتی می‌یابیم که ماهیت زبان ما را ناگزیر از استفاده از عبارت‌های نادقيق می‌کند. جمله « A یک تابع است» چنین عبارتی است: همواره نادقيق است؛ چرا که « A » نشانه‌ای برای یک نام خاص است. مفهوم یک تابع باید مفهومی مرتبه دوم باشد، در حالی که در زبان همواره به شکل یک مفهوم مرتبه اول ظاهر می‌شود. کاملاً متوجه‌ام که همین حالا که دارم این را می‌نویسم، باز هم منظورم را نادقيق بیان کرده‌ام. گاهی این امر واقعاً اجتناب‌ناپذیر است. آن چه مهم است این است که بدانیم داریم چنین می‌کنیم، و این چطور رخ می‌دهد. در یک نمادگذاری مفهومی می‌توانیم عبارتی دقیق برای آنچه از تابع (مرتبه اول با یک آرگومان) مراد می‌کنیم معرفی کنیم، برای مثال: « $(\forall \exists \forall \exists)$ ». ^۱ بنابرین، « $(\forall \exists \forall \exists)$ » دقیقاً همان چیزی را بیان می‌کند که در « $\exists \forall \forall \exists$ » یک تابع است» نادقيق بیان می‌شود. اکنون هرچه را جایگزین « $(\forall \exists)$ » کنیم، همواره به جمله‌ای صادق می‌رسیم چرا که تنها می‌توانیم نام‌های تابع مرتبه اول با یک آرگومان را جایگزین کنیم، زیرا در اینجا جایگاه آرگومان از نوع دوم است. همان‌طور که در زبان نمی‌توانیم به درستی درباره‌ی یک تابع بگوییم که یک شیء نیست، از زبان همچنین نمی‌توانیم استفاده کنیم تا درباره‌ی یک شیء، فی‌المثل ^۲، بگوییم که یک تابع نیست. شما درست فکر می‌کنید که با یک تابع نمی‌توان چون چیزی برخورد کرد؛ چرا که، همان‌طور که پیشتر گفتم، واژه‌ی «چیزی» نشانه‌ای از نامی خاص است. به جای استفاده از عبارت نادقيق « \forall یک تابع است»، می‌توانیم بگوییم: « $\exists \forall \forall \exists$ » یک تابع نام است». نمی‌توانیم به درستی درباره یک مفهوم‌نام بگوییم که به چیزی ارجاع می‌کند؛ اما می‌توانیم بگوییم که فاقد مرجع نیست. درست است که تابع‌نشانه‌ها یا مفهوم‌نام‌ها اجتناب‌ناپذیرند؛ اما با پذیرش این، باید این را نیز پذیریم که برخی از آن‌ها هستند که فاقد مرجع نیستند، حتی با وجود اینکه، به بیانی دقیق، عبارت «مرجع یک تابع نام» را نباید به کار ببریم.

راجع به آخرین نکته‌ی مورد اشاره‌ی شما، لازم است این را بگوییم: $\neg(\forall \exists \neg a = a)$ رده‌ای شامل تنها یک شیء واحد، یعنی صدق، و $\neg(\forall \exists \neg a = a)$ رده‌ای شامل تنها یک شیء واحد، یعنی کذب است. اگر Γ نه این رده و نه آن دیگری، بلکه گستره‌ی مقادیری دیگر باشد، آنگاه Γ از صدق متمایز است چرا که بر $\neg(\forall \exists \neg a = a)$ منطبق نیست و هکذا از کذب متمایز است چرا که بر $\neg(\forall \exists \neg a = a)$ منطبق نیست. بنابرین، اگر Δ یک ارزش صدق باشد، آنگاه « $\Gamma = \Delta$ » بر کذب ارجاع می‌کند.

عنوان مقاله‌ی «آیا موقعیت در زمان ...» باعث شد گمان کنم شاید شما به مقاله‌ای که من زمانی در *Zeitschrift für Philosophie und philosophische Kritik* درباره‌ی پرسشی مشابه منتشر کدم علاقه‌مند باشید. در حال حاضر دیگر قادر به یافتن نسخه‌ای از آن نیستم و عنوان آن را نیز به یاد نمی‌آورم، اما اگر بخواهید می‌توانم به دنبالش بگردم. احتمالاً با مقاله‌ی کوتاه من «درباره اعداد آفای اچ. شورت» آشنا هستید.

هنوز مجالی برای مطالعه‌ی مقالاتتان پیدا نکرده‌ام، اما امیدوارم به زودی این کار را انجام دهم.

با احترام،
گ. فرگه

^۱ به نظر می‌رسد در اینجا فرگه با استفاده از *spiritus asper* [spiritus lenis] به جای *spiritus lenis* [spiritus asper] که اغلب به کار می‌برد می‌خواهد میان عبارت مورد بحث و نام‌هایی که برای گستره‌های مقادیر دارد تمیز قائل شود.

^۲ این نشانه‌ی سیاره‌ی مشتری است؛ ر.ک. به جلد II قوانین بنیادین حساب، صفحه ۸۴، و ص. ۲۲۷ از منتشرات پس از مرگ.



نامه به مهیر*

باروخ اسپینوزا

مقدمه‌ی سردبیر

تعمق درباره‌ی آن‌چه بی‌نهایت خوانده می‌شود قدمتی هم‌پای تاریخ فلسفه و ریاضیات دارد. صورت‌بندی‌های مدرن ریاضی که امروزه می‌شناسیم عمدتاً محصول قرن ۱۹ م. هستند؛ باری این صورت‌بندی‌ها ریشه در تاملات و مباحثاتی دارد که اندیشمندان را قرن‌ها به خود مشغول کرده بود. یکی از محورهای اصلی این تاملات و مباحثات تفکیک انواع مختلف بی‌نهایت است؛ برای مثال می‌توان بی‌نهایت را طبق عمل افزودن یا تقسیم‌کردن تفکیک کرد، که نوعاً تفکیکی برای بی‌نهایت بزرگ و بی‌نهایت کوچک محسوب می‌شود. تفکیک مناقشه‌برانگیزتری که در طول تاریخ فلسفه و ریاضیات جریان دارد تفکیک بی‌نهایت بالفعل^۱ و بی‌نهایت بالقوه^۲ است: بالفعل به این معنا که آن‌چه صفت بی‌نهایت بر آن حمل می‌شود وجودی فی حد ذاته دارد [هم‌چون مجموعه‌ی اعداد طبیعی یا حقیقی به مثابه‌ی یک کلیت واحد] و بالقوه به این معنا که این صفت صرفاً به ناتمام بودن یک فرآیند [هم‌چون افزودن یا تقسیم‌کردن] اشاره دارد. ارسطو به عنوان نخستین کسی که این تفکیک اخیر را صورت‌بندی کرده، قائل به وجود بی‌نهایت بالفعل نبود: «نامتناهی درست عکس آن چیزی است که همگان می‌پندارند. نامتناهی «چیزی که ورای آن چیز دیگری نیست» نیست، بلکه «چیزی که همیشه ورای آن چیزی هست» است»^۳، و درک ما از بی‌نهایت را تنها به صورتی بالقوه—به مثابه‌ی فرآیندی پیش‌رونده و ناتمام—تلقی می‌کرد: «نامتناهی چنین حالتی از وجود دارد: همیشه چیزی به دنبال چیزی دیگر می‌آید، و هر یک از این چیزها همیشه متناهی، ولی همیشه متفاوت است»^۴. البته این تلقی مخالفین جدی خود را نیز داشته—چه بین فلاسفه و چه بین ریاضی‌دانان، من جمله اسپینوزا از دسته اول، کانتور از دسته‌ی دوم و لاپلنتیس از هر دو دسته.

برای ما دانش‌جویان ریاضی شاید ملموس‌ترین نمود باور به بی‌نهایت بالفعل در پذیرش اصل بی‌نهایت^۵ است، که تقریر کانتور است: اجمالاً در نظریه‌ی مجموعه‌ها ما به عنوان یک اصل موضوع می‌پذیریم که اقلًاً یک مجموعه‌ی نامتناهی وجود دارد؛ بدین طریق است که اعداد طبیعی را نه صرفاً به عنوان یک دنباله‌ی پایان‌نایپذیر (بی‌نهایت بالقوه)، بلکه به عنوان یک شیء (بی‌نهایت بالفعل) مطالعه می‌کنیم. قائلین به بی‌نهایت بالفعل نیز خود آن را به انواع مختلفی تفکیک می‌کنند. بخشی از این تفکیک‌ها فلسفی و متأفیزیکی است، و بخشی دیگر ریاضی؛ برای مثال، کانتور در یک افزار متأفیزیکی قائل به سه دسته نامتناهی است [که از قضا مشابه تقسیم‌بندی اسپینوزاست]: اول آن بی‌نهایتی که متعلق به خداست، دوم بی‌نهایتی که متعلق به ذات طبیعت است، و سوم بی‌نهایتی که در اعداد ترا متناهی و مجموعه‌های ریاضی یافت می‌شود. هم‌چنین کانتور انواع متفاوتی از بی‌نهایت را در دسته‌ی آخر مشاهده می‌کند، چنان که بر تفاوت بین نامتناهی بودن مجموعه‌ی اعداد طبیعی و نامتناهی بودن مجموعه‌ی اعداد حقیقی دست می‌گذارد. تلقی بالقوه از بی‌نهایت نیز در حیطه‌ی ریاضیات مدافعان جدی خود را دارد. برخی از ساختارگرایان تنها وجود بالفعل بی‌نهایت شمارا را می‌پذیرند. عده‌ای دیگر چون کرونکر و براوئر به وجود کلیتی تحت عنوان

*این نوشته ترجمه‌ای از کتاب زیر است:

Spinoza, B., *Complete Works*. Translated by Shirley, S., Hackett Pub., (2002), 787 - 791.

¹ Actual Infinity

² Potential Infinity

³ Aristotle, *Physics*. Translated by Reeve C. D. C., (2018), 207a.

⁴ Ibid, 206b.

⁵ Axiom of infinity

«مجموعه‌ی همه‌ی اعداد طبیعی» باور ندارند. هم‌اکنون نیز در ریاضیات محله‌هایی هم‌چون نامتناهی‌گرایی^۱ و فرامتناهی‌گرایی^۲ پدیدار شده است که مطابق با باورهای مذکور به پژوهش ریاضی مشغول اند.

نامه‌ی زیر از اسپینوزا، فیلسوف خردگرای هلندی، یکی از متن‌هایی است که در قرن ۱۷ م. و از دیدگاهی فلسفی نوشته شده است. نویسنده انجاء حصول ادراک از بی‌نهایت را تشریح کرده و ارتباط آن را با مفاهیمی چون جوهر، ذات، زمان، اعداد و... بررسی می‌کند. اسپینوزا بر آن است تا نشان دهد درک وجود برخی چیزها تنها به صورت نامتناهی ممکن است، و بخشی از گمراهی ما در فهم این نامتناهی به خاطر تمایز نگذاشتن بین درک انتزاعی و درک آن به مدد اسباب خیال به وجود می‌آید. ایده‌های مطرح شده توسط اسپینوزا بر متفکران پس از او موثر بود—چه در رد و چه در تایید این ایده‌ها. با این اوصاف نباید فراموش کرد که این نامه در چه زمانه‌ای و از چه دیدگاهی تحریر شده است؛ برای مثال، با این که بی‌نهایتها هنوز به عنوان شی‌ای ریاضی صورت‌بندی نشده‌است، بی‌نهایت کوچک‌ها^۳ در توجیه بنیان حساب دیفرانسیل و انتگرال نویا نقش ایفا می‌کنند، و از طرفی محله‌هایی هم‌چون خردگرایی و تجربه‌گرایی در فلسفه به جای فلسفه‌ی مدرسی برجسته شده‌اند. این نامه بعداً به «نامه در باب نامتناهی» مشهور شد. لایبنیتس خود تحشیه‌ای انتقادی بر آن نوشته است که به نوبه‌ی خود خواندنی است و علاقه‌مندان می‌توانند پس از مطالعه‌ی این نامه به آن رجوع کنند.

متن نامه به شرح زیر است:

به دانشمند فرهیخته، لودویک مهیر، پزشک و استاد فلسفه، از B.d.S.

[نسخه‌ی اصلی مفقود شده است، کپی نامه توسط لایبنیتس نگهداری شده بوده.]

دوست عزیز،

دو نامه از شما دریافت کردتم، یکی به تاریخ ۱۱ زانویه که توسط دوستمان ن. ن. به دست ام رسید، دیگری به تاریخ ۲۶ مارس که دوست ناشناسی از لایدن برای من فرستاده بود. هر دو بسیار دل‌گرم‌کننده بودند، به خصوص که دست‌گیرم شد اوضاع بر وفق مراد است و گه‌گاهی گوشی‌ذهن‌تان به فکر من اید. از صمیم قلب منونه محبت و علاقه‌ای که همواره به من نشان می‌دهید هستم. هم‌زمان استدعا دارم که باور داشته باشید من بی‌کم‌وکاست دوست‌دار و فادر شما هستم و در همه حال می‌کوشم که تا فرصتی دست داد به حد وسع اندک‌ام این را اثبات کنم. به عنوان اولین وظیفه، سعی خواهم کرد که به پرسشی که در نامه‌تان از من کرده بودید و در آن از دیدگاه‌ام درباره‌ی مسئله‌ی نامتناهی^۴ جواب شده بودید پاسخ دهم. شادمانه در خدمت ام.

مسئله‌ی نامتناهی به صورت کلی، همواره مسئله‌ای دشوار، و در واقع حل ناشدنی، به نظر آمده است. [این دشواری] از خلال خطأ در تشخیص میان آن‌چه بنابر طبیعت خود یا بنابر تعریف خودش نامتناهی است، و میان آن‌چه نه بنابر ذات‌اش، بلکه بنابر علت‌اش نامحدود^۵ است رخ می‌دهد. چندان که خطای وجود دارد در تشخیص میان آن‌چه به دلیل این که نامحدود است نامتناهی خوانده می‌شود، و آن‌چه اجزا ش را با هیچ عددی نمی‌توان شمرد یا تشریح کرد، هرچند که ما حداقل و حداکثرش را بشناسیم. در آخر، خطای وجود دارد در تشخیص میان آن‌چه می‌توانیم که فقط با خرد^۶ و نه با تخیل^۷ فراچنگ آوریم و آن‌چه که می‌توانیم با تخیل هم به دست‌اش آوریم. تاکید می‌کنم، اگر انسان به این تمایزها با دقت توجه می‌کرد، خود را در این همه دشواری غوطه‌ور نمی‌یافتد؛ به روشنی می‌فهمید که کدام نوع از نامتناهی نمی‌تواند به اجزا تقسیم شود یا اجزایی داشته باشد، و کدام نوع می‌تواند بی‌هیچ تناقضی تقسیم شود. هم‌چنین، می‌فهمید که کدام نوع از نامتناهی می‌تواند بدون تعارض منطقی، به عنوان کوچک‌تر یا بزرگ‌تر نامتناهی دیگری درک شود و کدام نوع نمی‌تواند. این از آن‌چه برآنم تا بگوییم روشن خواهد شد. اما باید اول به صورت خلاصه این چهار اصطلاح را تشریح کنم: جوهر^۸، حالت^۹، ابدیت^{۱۰}، دیرش^{۱۱}.

¹Finitism

²Ultrafinitism

³Infinitesimal

⁴Infinite

⁵Unlimited

⁶Intellect

⁷Imagination

⁸Substance

⁹Mode

¹⁰Eternity

¹¹Duration

نکاتی که درباره‌ی جوهر باید گوش‌زد شود از این قرار است. اول، وجود^۱ [جوهر] به ذات آن تعلق دارد؛ که یعنی تنها از ذات آن و از تعریف‌اش بر می‌آید که جوهر وجود دارد. این نکته را، اگر حافظه‌ام فربایم ندهد، پیش از این در گفت و گویی، بی‌دست‌یاری هیچ قضیه‌ای به شما اثبات کرمد. دوم، به دنبال نکته‌ی اول، جوهر چندگانه^۲ نیست؛ بلکه تنها یک جوهر بر اساس طبیعت آن [تعريف] وجود دارد. سوم این که هیچ جوهری غیر از اینکه نامتناهی باشد، قابل ادراک نیست.

من تاثرات^۳ جوهر را حالت‌ها می‌نامم. از آن‌جا که تعریف حالت همان تعریف جوهر نیست، نمی‌تواند شامل وجود باشد. بنابراین، حتاً هنگامی که وجود دارند، می‌توانیم چنین بینگاریم که وجود ندارند. از این بر می‌آید که وقتی ما فقط ذات حالت‌ها را مدنظر فرار داده باشیم، و نه نظم طبیعت را به مثابه یک کل، نمی‌توانیم از حضور اکنونی‌شان استنباط کنیم که در آینده وجود خواهد داشت یا نه، یا در گذشته وجود داشته‌اند یا نه. از این رو روشن است که ما وجود جوهر را هم‌چون یک [وجود] متفاوت از وجود حالت‌ها درک می‌کنیم. این منبع تفاوت میان ابدیت و دیرش است. تنها درباره‌ی وجود حالت‌هاست که می‌توانیم از اصطلاح دیرش استفاده کنیم. اصطلاح مطابق با وجود جوهر ابدیت است، که یعنی حظّ نامتناهی از وجود یا—بیخشید که لاتین‌اش را می‌نویسم—از بودن (*essendi*).

آن‌چه گفتم تقریباً روشن می‌کند که وقتی تنها به ذات حالت‌ها نظر می‌اندازیم، و نه به نظم طبیعت که اکثر اوقات مسئله است، می‌توانیم به صورت دل‌خواه وجود و دیرش حالت‌ها را تحدید^۴ کنیم، بی‌که در نتیجه‌اش آسیبی در هر ابعادی، به مفهوم ذهنی‌مان از آن‌ها خدشه‌ای وارد کند؛ و می‌توانیم این دیرش را هم‌چون بزرگ‌تر یا کوچک‌تر، و تقسیم‌پذیر به اجزا بینگاریم. اما ابدیت و جوهر، که تنها به مثابه‌ی نامتناهی قابل تصورند، نمی‌توانند بی‌ابطال مفاهیم ذهنی‌مان از آن‌ها، به کار گرفته شوند. بنابراین بی‌معناست و هم‌پهلو با دیوانگی است که اصرار کنیم جوهر ممتد^۵ از اجزا و بدن‌های واقعاً متمایز از هم ترکیب شده است. این چنان است که انگار با افزودن ساده‌ی دایره‌ای به دایره‌ای و گذاشتن یکی روی دیگری، کسی قصد کرده باشد مرتع یا مثلث یا هر شمایلی از یک طبیعت کاملاً متفاوت بسازد. بنابراین تمام گردایه‌ی استدلال‌هایی که به موجب‌شان فیلسوف‌ها عموماً می‌کوشند اثبات کنند جوهر ممتد متناهی است، به‌خودی خود فرومی‌ریزد. تمام چنین استدلال‌هایی پیش‌فرض می‌گیرند که جوهر بدن‌مند^۶ از اجزایی متشكل شده است. موردعی موازی را کسانی ارائه می‌کنند که خودشان را با این که خط از نقطه‌ها ترکیب می‌شود قانع می‌کنند و کلی استدلال ابداع کرده‌اند تا اثبات کنند که یک خط به صورت نامتناهی قابل تقسیم نیست.

به هر حال، اگر می‌پرسید که چرا ما چنین گرایش طبیعی‌ای به تقسیم‌کردن جوهر ممتد داریم، پاسخ می‌دهم که ما کمیت^۷ را از دو راه درمی‌یابیم: انتزاعی یا ظاهری؛ آن‌گونه به کمک حواس در تخیل‌مان از آن بهره‌مندیم یا به مثابه‌ی جوهر [درمی‌یابیم‌اش] که تنها از طریق خرد دریافته می‌شود. پس اگر ما به کمیت هم‌چون چیزی که درون تخیل وجود دارد نظر بیندازیم (و این کاری است که ما به کرات و فی الفور انجام می‌دهیم)، آن را تقسیم‌پذیر، متناهی، مرکب از اجزا و چندگانه می‌یابیم. اما اگر به آن هم‌چون چیزی درون خرد نظر بیندازیم و آن را هم‌چون چیزی—در—خود در نظر بگیریم (که کاری بسیار دشوار است)، آن‌گاه برمی‌آید که چیزی نامتناهی، تقسیم‌نایپذیر و یگانه است، چندان که به کفایت اثبات کرده‌ام.

سپس، از این حقیقت که ما قادریم دیرش و کمیت را آن‌گونه که می‌خواهیم تحدید کنیم—با فهم انتزاع‌شده‌ی کمیت از دل جوهر و تفکیک جریان دیرش از چیزهای ابدی—زمان و اندازه پدید می‌آیند. زمان برای تحدید دیرش، و اندازه برای تحدید کمیت—چنان سنجیده که قادر مان می‌کند، تا جای ممکن، به سهولت تخیل‌شان کنیم. دیگر این‌که، از آن‌جا که ما تاثرات جوهر را از خود جوهر جدا می‌کنیم، و سپس در طبقه‌هایی دسته‌بندی‌شان می‌کنیم که به سهولت قابل تخیل باشند، اعداد پدید می‌آیند، که به موجب‌شان آن [تأثر]‌ها را تحدید کنیم. از این رو به روشنی قابل مشاهده است که اندازه، زمان و اعداد چیزی جز حالت‌های اندیشه نیستند، یا بلکه، حالت‌های تخیل. به این ترتیب غافل‌گیرکننده نیست که تمام آن‌هایی که تقلّاً می‌کنند سازوکار طبیعت را به وسیله‌ی چنین مفاهیمی—آن هم بدون فهم واقعاً درستی از این مفاهیم—فهم‌مند، خود را به چنان گره‌های مهیبی درسته‌اند که آخر کار جز با ارتکاب به رمح‌ترین مُهملات نمی‌توانند خود را از بند رها کنند. چرا که چه بسیار چیزها هستند که به هیچ وجه با تخیل دریافته نمی‌شوند، بلکه تنها با خرد است که چنین چیزهایی به فهم درمی‌آیند، چون جوهر و ابدیت و چیزهایی از

¹ Existence

² Manifold

³ Affections

⁴ Delimit

⁵ Extended

⁶ Corporeal

⁷ Quantity

این دست. اگر کسی بکوشد چنین چیزهایی را با مفاهیمی از این دست، که چیزی جز ابزار تخيّل نیستند، تشریح کند، بیش از آن زمانی که عمداً جلوی راه تخيّل اش را به سمت دیوانگی باز گذاشته باشد، راه به جایی نمی‌برد. و حتا علاوه بر این، نه می‌توان حالت‌های جوهر را یک‌به‌یک به درستی بازشناخت اگر که آن‌ها را با چنین ذهن‌سازه‌ها (entia rationis) یا ابزار تخيّل قاطی کنیم. چرا که با این کار ما در حال تفکیک کردن آن‌ها از جوهر و از مسیر جاری شدن‌شان از [درون] ابدیت وجودیم و در چنین انزوایی، نمی‌توانند به درستی فهمیده شوند.

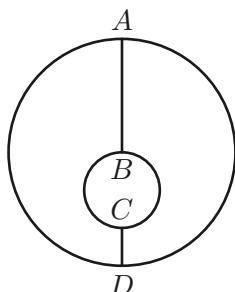
برای این‌که مسئله روشن‌تر شود، به این مثال توجه کنید. اگر کسی دیرش را به این صورت انتزاعی در نظر بگیرد و—با قاطی کردن اش با زمان—شروع به تقسیم کردن اش به اجرا کند، هرگز نمی‌فهمد که چگونه یک ساعت، برای مثال، می‌تواند سپری شود. زیرا برای آن‌که آن یک ساعت سپری شود، باید یک نیم‌ساعت سپری شود، و بعد نیمی از [نیم‌ساعت] باقی‌مانده، و بعد نیمی از آن چه باقی‌مانده؛ و اگر به تفیری نصف آن‌چه باقی‌مانده نباشد، هرگز به آخر آن یک ساعت نمی‌رسید. پس آن‌هایی که عادت به تشخیص ذهن‌سازه‌ها از چیزهای واقعی نداشته‌اند، ادعا کرده‌اند که دیرش از لحظه‌ها تشکیل شده، فلذًا از عقرب جراره به مار غاشیه پناه می‌برند. چرا که گفتن این‌که دیرش از لحظه‌ها سرهش شده، مثل این است که بگوییم اعداد از جمع زدن صفرها با هم ساخته شده‌اند.

علاوه بر این، از آن‌چه اشاره شد واضح است که نه عدد، نه اندازه و نه زمان که صرفاً ابزار تخيّل اند، هیچ‌کدام نمی‌توانند نامتناهی باشد، چرا که اگر باشند نه دیگر عدد خواهد بود، نه اندازه اندازه و نه زمان زمان. از این‌رو، می‌توان به سادگی دریافت که چرا خیلی‌ها—که این سه مفهوم را، به خاطر نادانی‌شان درباره‌ی طبیعتِ حقیقی، واقعیت، با واقعیت اشتباه می‌گیرند—وجود بالفعل^۱ نامتناهی را انکار کرده‌اند. باید این خردورزی تاسف‌آور آن‌ها را به دست ریاضی‌دانان سپرد که در مواردی که درک واضح و متمایز دارند، با این جور استدلال‌ها عقب نمی‌کشند؛ چرا که آن‌ها نه تنها به چیزهایی رسیده‌اند که با هیچ عددی قابل بیان نیستند (که به روشنی نابسنده‌ی اعداد را برای معین کردن همه چیز آشکار می‌کند) بلکه نمونه‌های زیادی سراغ دارند که نمی‌توان آن‌ها را برابر با هیچ عددی گرفت، و از هر عدد ممکنی فراتر می‌روند. با این همه، آن‌ها به این نتیجه نمی‌رسند که به دلیل کثرت اجزاست که چنین چیزهایی از همه‌ی اعداد فراروی می‌کنند، بلکه به این دلیل است که طبیعت آن چیز چنان است که عدد، بی‌که تناقضی را آشکار کند، در آن‌باره کاربردپذیر نیست.

برای مثال، تمام نابرابری‌های فضایی که بین دو دایره‌ی ABCD در نمودار قرار دارد، از اعداد فراتر می‌رود؛ چنان‌که تمام تغییرات سرعت ماده‌ای که آن محوطه را می‌پیماید این‌گونه است. این نتیجه‌گیری به دلیل وسعت بیش از حد^۲ فضای بینایینی به دست نمی‌آید؛ زیرا برشی هرقدر کوچک از آن را هم که انتخاب کنیم، نابرابری‌های این برش کوچک هم‌چنان فراتر از هر بیان عددی ای باقی خواهند ماند. هم‌چنین، این نتیجه‌گیری مثل موارد دیگر به این دلیل به دست نیامده که ما حداقل و حداقل را نمی‌دانیم؛ در مثال‌ما، هر دوی این‌ها را می‌دانیم: حداقل AB و حداقل CD است. نتیجه‌ی ما حاصل این است که عدد بر طبیعتِ دو دایره‌ی غیرهم‌مرکز کاربردپذیر نیست. بنابراین اگر کسی بخواهد تمام آن نابرابری‌ها را با یک عدد مشخص بیان کند، می‌بایست سبب شود که یک دایره، یک دایره نباشد.

مشابه‌اً، تا به بحث خودمان برگردیم، [باید بگوییم] اگر کسی قصد کند که تمام حرکت‌های ماده را که تاکنون رخ داده است—با فروکاستن، آن‌ها و دیرش‌شان به عدد و زمانی مشخص—معین کند، بی‌شک کمر بسته است تا جوهر بدن‌مند را—که نمی‌توان جز به عنوان چیزی موجود^۳ در نظرش گرفت—از تاثرات‌اش محروم کند و اسبابی فراهم کند که جوهر نمی‌باشد آن‌چه را دارد داشته باشد. می‌توانستم [هم‌چنان] این‌جا به روشنی این نکته و نکته‌های دیگری را که در نامه مطرح شد تشریح کنم، اگر که غیرضروری نبود.

از تمام آنچه گفته‌ام، می‌شود فهمید که برخی چیزها بنابر طبیعتِ خود نامتناهی‌اند و به هیچ وجه نمی‌شود متناهی تصویرشان کرد، در حالی که چیزهایی دیگر بنا بر علتی که در آن درون‌مان^۴ اند نامتناهی اند، و هنگامی که این دسته‌ی اخیر به نحو انتزاعی انگاشته شوند، می‌توانند که به بخش‌های کوچک‌تری تقسیم و به عنوان متناهی در نظر گرفته شوند. در نهایت، چیزهایی هم

¹ Actual² Existing³ Inhere

هستند که می‌توانند نامتناهی—یا اگر مایل‌اید نامعین—خوانده شوند، چون به دقت قابل بیان با هیچ عددی نیستند، هرچند که هم‌چنان به عنوان بزرگ‌تر یا کوچک‌تر قابل ادراک باشند. زیرا از این همه چنین برزمی‌آید که چیزهایی که نمی‌شود به دقت با عددی بیان‌شان کرد باید الزاماً با هم برابر باشند، که البته این از مثال‌های داده‌شده و موارد پرشمار دیگری به کفایت مشهود است. خلاصه، من این جا به اختصار علتِ خطاهای سردرگمی‌هایی را که از مسئله‌ی نامتناهی سربرمی‌آورند پیش روی شما گذاشتم و تمام‌شان را—مگر بر خطا باشم—به شیوه‌ای توضیح دادم که بعيد می‌دانم پرسش دیگری درباره‌ی نامتناهی باقی بماند که مطرح نکرده باشم یا از آنچه گفته‌ام به سادگی قابل حل نباشد. بنابراین گمان نمی‌کنم فایده‌ی دیگری داشته باشد که شما را پیش از این روی این مسئله معطل کنم.

به هر حال، همین طور گذری خدمتتان بگویم، نباید از قلم انداخت که—دست کم به عقیده‌ی من—مشائی‌های مدرن ما به کلی تشریح‌هایی را که مدرسیون قدیم با آن‌ها وجود خداوند را اثبات می‌کرده‌اند، اشتباه فهمیده‌اند. زیرا، چنان که من در آرای یهودی‌ای به نام رب خسدای^۱ یافتم، این برهان به این شکل است: «اگر سلسله‌ای نامتناهی از علل فرض گرفته شود، تمام آن‌چه هست، معلوم نیز است. اما هیچ معلولی نمی‌تواند ضرورتاً بنابر ذات خویش وجود داشته باشد. بنابراین هیچ در طبیعت وجود ندارد که وجود [ش] ضرورتاً تعلق به ذات اش داشته باشد. اما این نتیجه باطل است، لذا مقدمات نیز باطل بودند.» پس نیروی این استدلال نه در ناممکن بودن نامتناهی بالفعل یا در سلسله‌ای نامتناهی از علل، بلکه تنها در این فرض است که چیزهایی که بنابر ذات خود ضرورتاً وجود ندارند، متعین به وجود داشتن نمی‌شوند [جز] توسط چیزی که ضرورتاً بنابر ذات خود موجود است.

باید این جا به نامه‌ی دوم شما بپردازم، اما—از این حیث که از نظر زمانی تحت فشارم—می‌توانم راحت‌تر پاسخ‌گوی شما باشم، اگر پاسخ به نکاتی که مرقوم کرده بودید را موکول کنیم به وقتی که لطفاً سری به من می‌زنید. پس خواهش می‌کنم سعی کنید هرچه سریع‌تر تشریف بیاورید. زیرا موعد جابه‌جایی من سریعاً نزدیک می‌شود. تا اینجا کافی است. خدانگه‌دار و مرا از یاد نبرید، چنان که من از یادتان نمی‌برم.

رینزبورگ، ۲۰ آپریل ۱۶۶۳

مترجم: محسن امام‌وردي[†]

^۱فارغ‌التحصیل کارشناسی ارشد از دانشکده‌ی علوم نظری و مطالعات عالی هنر، دانشگاه هنر تهران.

رایانامه: mohsen.emamverdi@gmail.com



رسانه‌های همگانی و ریاضیات

بشری بصیری

مقدمه

در سال‌های گذشته گسترش رسانه‌های همگانی، به‌گونه‌ای شاید برگشت‌ناپذیر، زندگی اجتماعی انسان‌ها را تحت الشعاع قرار داده است. این تاثیر علی‌الخصوص در دوران قرنطینه خودش را گستردۀتر شد، و بدین ترتیب جای این رسانه‌ها را در ارتباطات نوع بشر مجهز به اینترنت ثبت کرد. یکی از شکل‌های مشخص این ارتباطات آموزش است و یکی از حیطه‌های این شکل، آموزش ریاضی. مجازی شدن موقت دانشگاه‌ها باعث رشد استقبال به درس‌گفتارهای اینترنتی یا ویدیو-مقاله‌های یوتیوب شد. بخشی از این محتواهای آموزشی مربوط به نهادهای آکادمیک است؛ عمدۀی دانشکده‌های ریاضی دانشگاه‌های مطرح جهان صفحه‌ای مربت دارند تا سمینارها و درس‌گفتارهایشان را بارگذاری کنند. در مثالی دیگر مجله‌ی نام آشنای Quanta صفحه‌ی خود را در یوتیوب دارد. با جست‌وجوی مختصّی می‌توانید به ابود صفحاتی برخورید که به حل مسائل مسابقاتی می‌پردازند. در اینجا تمرکز ما بر معرفی کانال‌هایی در YouTube و پادکست‌هایی است که به شکل درس‌گفتار یا ویدیو-مقاله سعی در اشاعه‌ی ریاضیات دارند. قالب پادکست ممکن است برای محتواهای ریاضی چالش‌برانگیز به نظر برسد؛ با این حال برخی نهادهای آکادمیک، مانند آکسفورد، پادکست‌های مفیدی، برای مثال در تاریخ ریاضیات، تدارک دیده‌اند. در ادامه برخی از این‌ها را با توضیح مختصّی معرفی می‌کنیم.

3Blue1Brown

این کanal يحتمل يكى از شناخته شده‌ترین کانال‌های ریاضیاتی یوتیوب است. گرنت سندرسون، که دانش آموخته‌ی ریاضیات و علوم کامپیوتر از استنفورد است، گرداننده‌ی کانال و تیم حول آن است؛ می‌گوییم «تیم» زیرا اگر ویدیو-مقاله‌ها یا درس‌گفتارهایشان را دیده باشید، متوجه کیفیت حرفه‌ای ماجرا خواهد شد. ویدیو-مقاله‌های این کانال مطالبی از حیطه‌های آنالیز، جبر خطی، تپیلوژی تا تبدیلات فوریه، شبکه‌های عصبی و مکانیک کوانتوسی را شامل می‌شود، و در کنار بیان خوب و مسیر شفاف استدلال‌ها تمرکز عمدۀی ارائه‌ها روشنی‌بخشیدن به شهود به مدد بصری‌سازی‌های بادقتی است که با پایتون انجام شده است. گرنت سندرسون با همین رویکرد سلسله درس‌گفتارهایی نیز برای یک دوره‌ی دانشگاهی مقدماتی در جبر خطی، معادلات دیفرانسیل ... تولید کرده است. علاوه بر مباحث حل مسئله نیز بخش مهمی از محتواهای موجود است، و تا جای ممکن راه حل‌های بدیع پیشنهاد می‌شود.

Aleph 0

اگر علاقه‌مند به هندسه‌ی جبری اعداد هستید این‌جا موقف خوبی است، ولو در حد کسب آشنایی مقدماتی می‌خواهید دنبال‌شان کنید. برخی از مطالب نیز به مسائل حل نشده‌ی مشهور اختصاص دارد.

Mathemaniac

ادعای کلی در این کانال این است که مطالبی را پوشش می‌دهد که در دوره‌های رسمی ریاضیات پرداخته نمی‌شود یا اگر هم در دوره‌های رسمی ذکری از این مطالب رفته، شیوه‌ی متفاوتی برای ارائه در پیش رفته است. بخش مهمی از این مطلب در نظریه‌ی گروه‌ها، جبر لی و آنالیز مختلط است، و تعدادی معتبره‌ی ویدیو-مقاله‌های تکی در حیطه‌های مختلف دیگر. بر عهده‌ی شماست که بینید شعارش را می‌خرید یا نه.

Mathematical Structuralism

امیرحسین اکبرطباطبایی، که تجربه‌ی تدریس در دانشگاه تهران دارد و هم‌اکنون در دانشگاه اترخت مشغول به تدریس است، در دوران فرنطینه سلسله درس‌گفتارهایی را شروع کرد با عنوان «آشنایی با ریاضیات ساختارگرایانه [شهودگرایانه] به روایت مکتب [بر اوئر]». هدف این درس‌گفتارها «آشنایی با زبان و نگرش کنگوریک به ریاضیات و بعد آشنایی با خود هندسه از منظر این مکتب» است. از آن جایی که پیش‌نیازی هم مفروض نگرفته، مطالب با توضیح ملزمات از اس‌واساس پیش می‌رود. برنامه‌ی دوره با نظریه‌ی کنگوری شروع شد؛ اکنون در میانه‌ی نظریه‌ی توپوس است، و در ادامه به نظریه‌ی مجرد هموتوپی، تایپ تئوری، هایرکنگوری، هایرتوپوس و هایرچئومتری تئوری خواهد رسید. درس‌گفnarها به شکل کلاس برگزار شده و ضبط می‌شود. دوره‌ی اول به فارسی و دوره‌ی دوم به انگلیسی او برای برای کسانی تدارک دیده شده است که گرایش‌های منطقی-فلسفی یا گرایش‌های هندسی در ریاضیات و علوم کامپیوتر دارند. پژوهشی جاهطلبانه و تحسین‌برانگیز اکبرطباطبایی یک منبع آموزشی بی‌نظیر برای علاقه‌مندان فراهم آورده است.

Mathologer

تعهد ریاضیاتی این کanal از حیث‌های مختلفی جالب توجه است. گردنده‌گان، بوکارد پولستر و مارتی راس از اساتید دانشگاه موناش ملبورن، نه تنها پرکار هستند بلکه حیطه‌های بسیار متنوع و گوشه‌های دوران‌نظری از ریاضیات را پوشش می‌دهد و کارشان از کیفیتی دانش‌نامه‌ای برخوردار است. دقت استدلال‌های ریاضی نیز کاملاً از استانداردهای ریاضی برخوردار است، و آن‌جا که در مقاله ویدیو-مقاله‌شان نمی‌گنجد با ارجاعات دقیق کپ را پر می‌کنند.

MathMajor

همان‌طور که از اسم اش برمی‌آید، عمدتی تمرکز این کanal بر ریاضیات به عنوان یک رشته‌ی دانشگاهی است. مطالب عمدتاً شامل آموزش‌های رسمی‌ای است که دانشجوها در دانشگاه می‌آموزند؛ مانند جبر مجرد، جبر خطی، نظریه‌ی اعداد، معادلات دیفرانسیل و آنالیز مختلف. بخشی از مطالب نیز شامل سمینار و مصاحبه با ریاضی‌دانان دیگر است. بخش جالب دیگری که در این کanal به آن پرداخته می‌شود نگارش ریاضی است؛ بدین نحو که با بررسی قضایا و مسائل معین در روند ارائه‌ی برهان‌ها مطالعه و مذاقه می‌کند. گردنده‌ی کanal، مایکل پن، یک کanal دیگری هم دارد که عمدتاً به حل مسئله و اثبات قضیه—و آن هم عمدتاً در آنالیز و نظریه اعداد— اختصاص دارد.

Numberphile

این یکی يتحمل محبوترین و راحت‌الحلقوم‌ترین مطالب ریاضی یوتیوب را تامین می‌کند. همه‌ی مطالب به شکل مصاحبه با ده‌ها ریاضی‌دان مختلف ضبط شده است و جز مازیک و کاغذ‌فهوهای معینی به هیچ چیز دیگری نیاز ندارد؛ با این حال فارغ از پیچیدگی مفهوم مطرح شده، همه چیز به شیوه‌ای شیوا و شیرین بیان می‌شود. مصاحبه‌ها معمولاً کوتاه، بعض‌اً مقدماتی و همیشه جالب توجه‌اند.

PeakMath

فرضیه‌ی ریمان در جهان ریاضیاتی شهرتی افسانه‌ای دارد و همه می‌دانند که شیداکننده است. این کanal رسالت خود را بر روایت این مسئله‌ی مشهور بنا کرده و تا به این‌جا بحث مبسوطی درباره‌ی L -تابع‌ها شده است. در مورد فرض ریمان مطالب پرشماری و نه‌چندان متنوعی در یوتیوب وجود دارد، ولی این کanal در عین آهستگی دقت بیش‌تری نیز ارائه می‌دهد.

Richard E Borcherds

ریچارد بورچردز دارنده‌ی مدال فیلدز ۱۹۹۸ درس‌گفتارهایش را در یوتیوب بارگذاری می‌کند. وی، که هم‌اکنون در حوزه‌ی کوانتم فیلد تئوری کار می‌کند، در جزئی‌تر شکل خود هر آن‌چه سروکاری با جبر داشته باشد تدریس می‌کند، توپولوژی جبری، جبر جابه‌جایی، نظریه‌ی گالوا، جبر لی والخ. دوره‌ی کوتاهی هم در تاریخ علم برگزار کده است که بسیار آموزنده است.

My Favourite Theorem

این آخری کانالی در یوتیوب نیست؛ بلکه پادکست پروپیمانی هم است. در هر قسمت، که تاکنون در شماره ۹۰ گذشته است، یک ریاضی دان دعوت می‌شود تا از قضیه‌ی محبوب خود حرف بزند. نتیجه شنیدنی است؛ هر کسی به اختصار علت علاقه‌ی خودش به قضیه را توضیح می‌دهد، و به خوبی انگیزه، معنا، اثبات، اهمیت و کاربردهای آن قضیه تشریح می‌کند. اگر قالب محبوب رسانه‌ای تان شنیداری است، پس بهتر است قصه‌ی قضیه‌های محبوب را هم بشنوید.



مسائل

علی دایی نبی^{*} و علی الماسی[†]

(۱) مجموعه‌ی S از نقاط با مختصات صحیح در صفحه‌ی مختصات دو بعدی مفروض است. تابع $f : S \rightarrow S$ به گونه‌ای است که فاصله‌ی هر دو نقطه از S را کاهش می‌دهد. به بیانی دیگر برای هر $x, y \in S$ ، $d(f(x), f(y)) \leq d(x, y)$ است. اگر S مجموعه‌ای فرد عضوی باشد، نشان دهید عضو p در S وجود دارد که d فاصله‌ی معمولی اقلیدسی است. دقت کنید که تصویر نقاط تحت f داخل مجموعه قرار دارد و نقاط S مختصات صحیح دارند. علاوه بر این، برای حالتی که S زوج عضو داشته باشد، مثالی ارائه دهید که f نقطه‌ی ثابت نداشته باشد، یعنی برای هر $p \in S$ ، $f(p) \neq p$.

(۲) در ریاضیات دیرستانی معمولاً چندضلعی محدب را خمی بسته و شکسته تعریف می‌کنند که زاویه‌ی بیشتر از 180° درجه نداشته باشد. از طرفی، تعریف دقیق یک شکل محدب یعنی شکلی که پاره خط واصل هر دو نقطه از آن کاملاً درون شکل قرار بگیرد. آیا این دو تعریف معادل هستند؟ به بیانی دیگر، در صفحه‌ی مختصات دو بعدی، به این سوال پاسخ دهید که هر خم بسته و شکسته‌ای که زاویه‌ی بیشتر از 180° درجه نداشته باشد، یک ناحیه‌ی محدب در درون خود ایجاد می‌کند یا خیر. به جهت دیگر این سوال نیز پاسخ دهید.

(۳) آیا یک n -ضلعی محدب وجود دارد که بتوان آن را با چهارضلعی‌های مقرر پوشاند؟ پوشاندن یعنی تمام چهارضلعی‌ها داخل شکل قرار بگیرند، هم‌پوشانی نداشته باشند، و تمام درون n -ضلعی را شامل شوند.

(۴) روی دایره‌ی واحد در صفحه‌ی مختصات دو بعدی، تعداد n نقطه با رنگ قرمز علامت زده شده‌اند. این نقاط را به اندازه‌ی $\frac{2\pi k}{n}$ دوران می‌دهیم، برای یک k صحیح، و نقاط جدید را با رنگ آبی علامت می‌زنیم. نشان دهید همواره دو علامت آبی وجود دارند که بین دو علامت قرمز متواالی قرار گرفته باشند.

(۵) دو معدن طلا داریم که در هر کدام مقدار G_1 و G_2 طلا وجود دارد. یک ماشین حفاری در اختیار داریم که می‌توانیم از آن برای استخراج طلا استفاده کنیم. اگر از معدن i استخراج کنیم، $i = 1, 2$ ، با احتمال p_i ماشین کار کرده و کسر r_i از مقدار طلای موجود در آن لحظه را استخراج می‌کنیم. در غیر این صورت، با احتمال $1 - p_i$ ماشین را به طور کامل از دست می‌دهیم. استراتژی بهینه‌ای ارائه دهید که امید مقدار طلای استخراجی تا قبل از خرابی ماشین را بیشینه کند.

(۶) تعداد N زیرمجموعه‌ی متمایز از یک مجموعه‌ی S مفروض است. اگر این زیرمجموعه‌ها تشکیل یک جبر روی S دهند، یعنی نسبت به اجتماع، اشتراک، و مکمل‌گیری بسته باشند، نشان دهید K وجود دارد که $N = 2^K$. برای شروع، می‌توانید فرض کنید S متناهی عضو دارد. برای حالت نامتناهی نیز حکم را ثابت کنید.

(۷) (پایه‌های ضربی غیرقابل گسترش) ضرب تنسوری دو ماتریس $A_{m \times n}$ و $B_{p \times q}$ ، که آن را با $A \otimes B$ نمایش می‌دهیم، ماتریسی با ابعاد $(mp) \times (nq)$ است که به صورت زیر تعریف می‌شود:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

با استفاده از تعریف بالا، می‌توان ضرب تنسوری دو فضای برداری ماتریسی V و W را به صورت زیر تعریف کرد:

$$V \otimes W = \{v \otimes w \mid v \in V, w \in W\}.$$

۱.۷ نشان دهید ضرب تنسوری دو فضای برداری، یک فضای برداری است.

۲.۷ نشان دهید بردارهایی در $\mathbb{R}^3 \otimes \mathbb{R}^3$ وجود دارند که به صورت $v \otimes w$ ، که $v, w \in \mathbb{R}^3$ ، نیستند.

۳.۷ فرض کنید V و W به ترتیب مجهر به ضربهای داخلی $\langle \cdot, \cdot \rangle_V$ و $\langle \cdot, \cdot \rangle_W$ هستند. نشان دهید نگاشت $\langle \cdot, \cdot \rangle : V \otimes W \times V \otimes W \rightarrow \mathbb{R}$ است که به صورت زیر تعریف شده است، یک ضرب داخلی روی $V \otimes W \rightarrow \mathbb{R}$

$$\langle v \otimes w, v' \otimes w' \rangle = \langle v, v' \rangle_V \langle w, w' \rangle_W, \quad \forall v, v' \in V \quad \forall w, w' \in W.$$

۴.۷ نشان دهید مجموعه‌ی متناهی S از بردارهای $\mathbb{R}^3 \otimes \mathbb{R}^3$ وجود دارد که دارای شرایط زیر است:

- همه‌ی اعضای S به فرم $v \otimes w$ هستند، که $v, w \in \mathbb{R}^3$.

- اعضای S دو به دو برحهم عمودند.

- $\text{Span}(S) \subsetneq \mathbb{R}^3 \otimes \mathbb{R}^3$.

- هیچ عضوی ندارد که به فرم $w \otimes v$ ، برای $v, w \in \mathbb{R}^3$ باشد.

(۸) (اثباتی برای نامتناهی بودن اعداد اول با نظریه‌ی اutomata) الفبای $\{0, 1\}^\omega = \Sigma$ را در نظر بگیرید. برای هر رشته‌ی $e(w)$ ، $w \in \Sigma^*$ را برابر با تعداد صفرهای w منهای تعداد یک‌های آن تعریف کنید. برای هر $n \in \mathbb{N}$ تعریف کنید

$$\mathcal{L}_n = \{w \in \Sigma^* : n | e(w)\}.$$

به علاوه، تعریف کنید

$$\mathcal{L} = \bigcup_{p \in \text{عدد اول}} \mathcal{L}_p.$$

۱.۸ نشان دهید برای هر $n \in \mathbb{N}$ ، \mathcal{L}_n زبانی منظم است.

۲.۸ نشان دهید \mathcal{L} منظم نیست.

۳.۸ نتیجه بگیرید تعداد اعداد اول نامتناهی است.

(۹) (یک قضیه در کاشی کاری) می‌خواهیم مربعی را به مثلث‌های هم‌مساحت تقسیم کنیم.

۱.۹ نشان دهید برای هر عدد زوج n ، می‌توان مربع را به n مثلث هم‌مساحت تقسیم کرد.

۲.۹ در ادامه می‌خواهیم نشان دهیم که اگر بتوان مربع را به n مثلث هم‌مساحت تقسیم کرد، n عددی زوج است.

۱.۲.۹ یک تابع قدر روی یک میدان K ، تابعی مانند $\mathbb{R} \rightarrow K$ است که ویژگی‌های زیر را داشته باشد:

$$x = 0 \text{ و } v(x) = 0 \text{ اگر و تنها اگر } 0 < v(x) \geq 0, \forall x \in K \quad \bullet$$

$$. v(xy) = v(x)v(y), \forall x, y \in K \quad \bullet$$

$$. v(x+y) \leq v(x) + v(y), \forall x, y \in K \quad \bullet$$

$$. v(x) = v(-x), x \in K \quad \bullet$$

نشان دهید برای هر $x \in K$ ، $|x|_2$ را به صورت زیر تعریف کنید:

$$. |0|_2 = 0 \quad \bullet$$

• برای هر $r \in \mathbb{Q} \setminus \{0\}$ ، می‌دانیم r را می‌توان به‌طور یکتا به صورت $r = 2^{m \frac{a}{b}}$ نوشت که

اعدادی صحیح و a و b اعدادی فرد هستند. تعریف کنید $|r|_2 = 2^{-m}$.

نشان دهید تابع فوق، یک تابع قدر روی میدان \mathbb{Q} است.

۳.۲.۹ نشان دهید $|x|_2 \cdot |y|_2$ این ویژگی را دارد که $\max\{|x|_2, |y|_2\} \leq |x+y|_2 \leq |x|_2 + |y|_2$ و تساوی زمانی رخ می‌دهد که

$$. |x|_2 \neq |y|_2$$

۴.۲.۹ نشان دهید برای عدد طبیعی n ، $1 < |n|_2$ اگر و تنها اگر n زوج باشد.

۵.۲.۹ فرض کنید که مربع را به n مثلث هم‌مساحت تقسیم کرده‌ایم. برای سادگی فرض کنید که مختصات رؤوس مربع مزبور نقاط $(1, 0), (0, 1), (1, 1), (0, 0)$ است. به علاوه فرض کنید رئوس مثلث‌ها نیز در نقاطی با

مختصات گویا قرار گرفته‌اند.

رئوس مثلث‌ها را با روشی که در ادامه بیان می‌شود رنگ آمیزی می‌کنیم. برای هر رأس با مختصات (x, y) ,

- آبی است اگر $|x|_2 \geq |y|_2 \geq 1$ و $|x|_2 \geq |y|_2 < 1$.
- سبز است اگر $|y|_2 \geq |x|_2 > 1$ و $|y|_2 < 1$.
- قرمز است اگر $|x|_2 < 1 < |y|_2$ و $|x|_2 < |y|_2$.

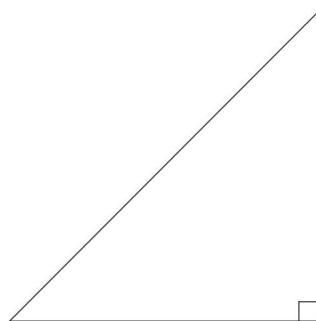
فرض کنید مساحت مثلثی که رئوسش با سه رنگ متفاوت رنگ آمیزی شده است، برابر با d است. نشان دهید $1 > |d|_2$.

۶.۲.۹ نشان دهید چنان‌چه رئوس مثلث‌ها را با رنگ آمیزی فوق رنگ کنیم، حداقل یک مثلث با رئوسی با سه رنگ متفاوت خواهیم داشت.

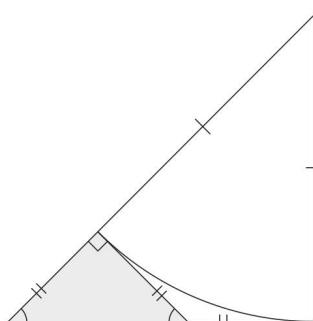
۷.۲.۹ از قسمت‌های قبل تبیجه بگیرید که تعداد مثلث‌ها در تقسیم مربع به مثلث‌های همساحت با رئوس با مختصات گویا عددی زوج است.

۳.۹ آیا می‌توان تابع قدر تعریف شده در قسمت قبل را به اعداد حقیقی توسعه داد؟ چگونه؟

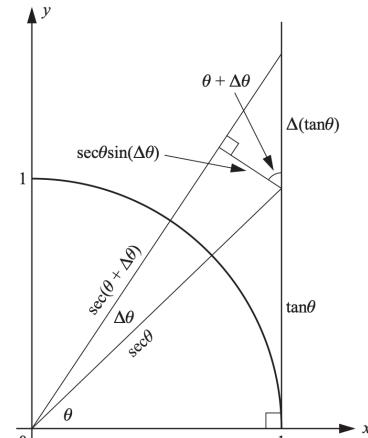
(۱۰) (اثبات‌های تصویری) هر یک از شکل‌های زیر برای کدام قضیه‌ی ریاضی اثباتی ارائه می‌دهد؟



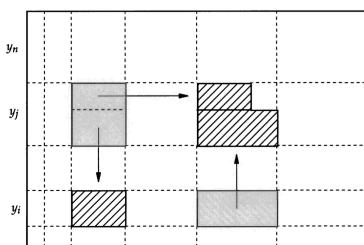
(ب)



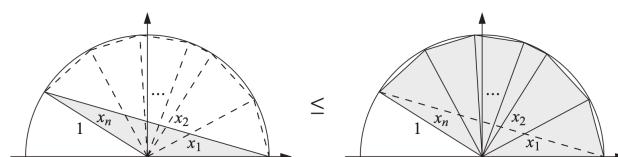
(ب)



(ت)



(د)



(ج)

شکل ۱. اثبات‌های تصویری.

* دانشجوی دکتری، دانشگاه تورنتو
رایانامه: alidaeinabi@gmail.com

† دانشجوی دکتری علوم کامپیوتر، انسیتو پلی‌تکنیک پاریس
تاریخ: <https://ali-almasi.github.io>
رایانامه: ali.almasi@polytechnique.edu

جایزه‌ی آبل ۲۰۲۴ به میشل تالاگراند رسید!

جایزه‌ی آبل یکی از دو جایزه‌ی مهم ریاضیات است که برای ریاضی‌دانان همراهی جایزه‌ی نوبل به حساب می‌آید. این جایزه در سال ۲۰۲۴ به میشل تالاگراند، ریاضی‌دان فرانسوی، تعلق گرفت؛ «به دلیل مساهمنهای پیش‌روانه‌اش در نظریه‌ی احتمال و آنالیز تابعی، که کاربردهایی خیره‌کننده در فیزیک ریاضی و آمار دارند.». بنا بر آن‌چه کمیته‌ی جایزه‌ی آبل در توصیف برنده‌ی امسال آن نوشته است، این جایزه به طور خاص به خاطر این سه زمینه از کارهای تالاگراند به او اهدا شده است: مطالعه‌ی سوپریمم‌های فرایند‌های تصادفی، تمرکز اندازه و مطالعه‌ی مدل شبیه‌های اسپینی.

تالاگراند در ۱۹۵۲ در فرانسه به دنیا آمد، و پس از آن که دکتراش را از دانشگاه پاریس VI دریافت کرد، به سمت استادی دانشگاه اوهايو درآمد. افزون براین، او از ۱۹۷۴ در مرکز ملی پژوهش‌های علمی (CNRS) فرانسه مشغول به فعالیت بوده است. تالاگراند بی‌شك یکی از جالب‌ترین صفحه‌های شخصی را در بین ریاضی‌دانان صاحب‌نام دارد. در بخشی از صفحه‌ی شخصی او می‌توانید مسئله‌هایی را بیایید که او شخصاً برای کسی که آن‌ها را برای اولین بار حل کند، جایزه‌ای در نظر گرفته است. یکی از این مسائل که تالاگراند برایش ۱۰۰۰ دلار جایزه در نظر گرفته است، به شرح زیر است.



Create convexity in 3 (or 100?) steps only!

Consider an integer N . Let us say that a compact subset A of \mathbf{R}^N is **balanced** if

$$x \in A, \lambda \in \mathbf{R}, |\lambda| \leq 1 \Rightarrow \lambda x \in A.$$

Let us denote by γ_N the canonical Gaussian measure on \mathbf{R}^N .

Problem. Prove that there exists an integer q , such that for all N and every compact balanced set A of \mathbf{R}^N such that $\gamma_n(A) \geq 1/2$, one can find a **convex** compact set $C \subset A + \dots + A$ (with q terms on the right) such that $\gamma_n(C) \geq 1/2$.

In words: **finitely many steps, independently of dimension, suffice to create convexity.**