طلوع عصر روشهای تصادفی•

دیوید مامفرد*
ترجمهٔ شاپور اعتماد

۱. مقدمه

مطالب این مقاله حاصل سخنرانی من در کنفرانس «ریاضیات در آستانهٔ هزارهٔ سوم»، در فرهنگستان ملی لینچی است که در تاریخ ۲۷ـ۲۹ مه ۱۹۹۹ برگزار شد. باید به آن هفت استاد پرتحرک دانشگاه رُم، تور ورگاتا، که همه زن بودند و ابتکار برگزاری چنین کنفرانسی از آنها بود و مجری آن هم بودند تبریک بگویم. به‌خصوص باید به‌دستاورد آنان اشاره بکنم که توانستند ده دوازده ریاضیدان را وادارند تا نه فقط به‌آخرین پیشرفتهای رشتهٔ خود بلکه به امور دامنه‌دارتری هم بپردازند: یعنی هم از قضیه‌ها سخن بگویند و هم از اندیشه‌ها. دعوت آنان این وسوسه را در من برانگیخت که از این فرصت استفاده بکنم و یک سلسله ایده‌هایی را که طی ده‌سال گذشته جفت و جور کردنشان دل‌مشغولی دائمی‌ام بوده است به صورت روشنتری بیان کنم. و در نهایت نتوانستم خودم را از این لذت عظیم محروم کنم که چشم‌انداز بلندمدتی بر مبنای آنها ترسیم کنم ــ چشم‌اندازی که بی‌تردید هم ساده‌نگرانه است و هم اینکه حدودش از قلمرو تخصص من فراتر می‌رود. اگر بخواهم اصطلاحی را که خانم کارن اُولنبک در ضمن سخنرانی خود در این کنفرانس تعریف کردند وام بگیرم و درجهٔ بلندپروازی حرف خود را به‌زبان کمّی بیان بکنم باید بگویم که می‌خواهم یک سلسله ادعاهایی بکنم که دامنهٔ زمانی‌اش ۲۴۰۰ سال است؛ ولی ملاک قضاوت دربارهٔ این دامنهٔ زمانی، مدت زمان تجربهٔ تحقیقات خودم است، یعنی چیزی در حدود ۴۰ سال؛ در نتیجه ضریب بلندپروازی این سخنرانی ۶۰ است!

این مقاله از نوع جدلی است، جدل و نزاعی که محورش را نکته‌ای ساده تشکیل می‌دهد: اینکه در مقایسه با مدلهای دقیق و استدلالهای منطقی، مدلهای تصادفی و استدلالهای آماری در امور زیر ذی‌ربط‌ترند: (i) جهان؛ (ii) علم و بسیاری از بخشهای ریاضیات؛ (iii) به‌خصوص درکِ محاسباتی که در اذهان خود ما اجرا می‌شوند. برای رسیدن به‌این مقصود نکات خود را به‌ترتیب ذیل تنظیم و بیان خواهم کرد: در بخش ۲ حرف من این خواهد بود که تمام ریاضیات از طریق انتزاع از این یا آن جنبه از تجاربمان پدید می‌آید و پابه‌پای ریاضیاتی که از اشیا و حرکتشان در جهان مادی نشأت گرفته، منطق صوری ــ حداقل به‌نحوی که در آثار ارسطو منعکس است ــ با مشاهدهٔ خودِ تفکر پدید آمده است. درحالی‌که برای انتزاع ماهیت فرایند تفکرمان انواع و اقسام راههای دیگر هم وجود دارد که یکی از آنها به احتمال و آمار منتهی می‌شود. در بخش ۳ نگاه تند و گذرایی به این دامنهٔ زمانی ۲۴۰۰ سالهٔ بعد از ارسطو می‌افکنیم، و برخی از نقاط عطف در سیر تحول این دو دیدگاه را برجسته می‌کنیم. مدلهای دقیق منطق‌ـمحور و ریاضیاتِ دقیق منطق‌ـمحور همواره دست بالا را داشته‌اند و به‌گونه‌ای عمیق بر تفکر ما تأثیر گذاشته‌اند. نظریه‌های تصادفی [استوکاستیک] خیلی آهسته‌تر پدید آمده‌اند و تازه از قرن گذشته توانسته‌اند تا حدی عمق واقعی خود را نشان دهند. در بخش ۴، می‌خواهم به رهیافت متعارف تقلیل‌گرا به احتمال بپردازم. موضوع اصلی مورد مطالعه در نظریهٔ احتمال، متغیر تصادفی است و حرف من در ارتباط با آن این خواهد بود که باید آن را به‌عنوان یک ساخت بنیادی در نظر گرفت، درست مثل فضا، گروه، تابع، و اگر آن را براساس نظریهٔ اندازه تعریف بکنیم مرتکب کاری تصنعی و غیر طبیعی شده‌ایم. در بخش ۵، این نکته را بسط می‌دهیم، یعنی بر مبنای آثار الهام‌بخش جینز و فریلینگ، پیشنهاد می‌کنم که احتمالات و متغیرهای تصادفی را می‌توان در مبانی ریاضیات ادغام کرد، و در نتیجه به‌فرمالیسم [دستگاه صوری] قویتر و شهودیتری رسید. در بخش ۶ به‌تأثیر مدلهای تصادفی بر جریانهای اصلی ریاضیات و به‌خصوص تأثیر آنها بر نظریهٔ معادلات دیفرانسیل معمولی و جزئی می‌پردازم. آنجا حرفم این خواهد بود که برای مدلسازی جهان، معادلات تصادفی ذی‌ربط‌تر و اساسیتر از معادلات تعینی یا قطعی‌اند. در انتها یعنی در بخش ۷، باز می‌گردیم به‌بررسی مسألهٔ مدلسازی تفکر و رهیافتهای تصادفی اخیر در ارتباط با مباحث هوش مصنوعی، بینایی و گفتار.

پرسش ما این خواهد بود: آیا این رهیافتها بیشتر از رهیافتهای منطق‌ـمحور، شانس توفیق دارند ــ مثلاً در مورد تقلید تواناییهای بشری به‌کمک کامپیوتر؟ من خودم چنین فکر می‌کنم، ولی پاسخ قطعی هنوز روشن نیست.

۲. رده‌بندی ریاضیات

قبل از هر چیز می‌خواهم جایگاه احتمال و آمار را به‌عنوان بخشی از ریاضیات تعیین کنم. در ابتدا می‌خواهم تعریفی را که دیویس و هرش در کتاب بسیار ژرف‌کاوشان *تجربهٔ ریاضیات* [۱] (ص ۳۹۹) از ریاضیات ارائه می‌دهند نقل بکنم: «ریاضیات یعنی مطالعهٔ اشیای ذهنی که دارای خواصِ باز تولید پذیرند». من عاشق این تعریفم، چون ریاضیات را به هیچ‌وجه به‌چیزی که در گذشته ریاضیات خوانده شده است محدود نمی‌کند، و در واقع سعی می‌کند بگوید که چرا برخی از حرفها را ریاضیات، برخی را علم، برخی را هنر، و برخی را هم حرفهای خاله زنکی می‌خوانیم. به‌این ترتیب خواصِ بازتولیدپذیر جهان فیزیکی علم است در حالی‌که اشیای ذهنیِ باز تولیدپذیرْ ریاضیات است. هنر هم در ساحت ذهن ادامهٔ حیات می‌یابد (نقاشی واقعی مجموعه‌ای رنگ‌دانهٔ خشک‌شده روی بوم نیست همان‌طور که سمفونی هم یک سلسله امواج صوتی که آن را به‌گوش ما می‌رسانند نیست) ولی همان‌گونه که پست‌مدرن‌ها اصرار دارند، هر هنردوستی آن را در هر متن و زمینهٔ جدیدی از نو تفسیر می‌کند. اما وقتی به‌حرفهای خاله‌زنکی می‌رسیم می‌بینیم که بخش اعظم افکار ما از آنها تشکیل شده است و ماهیتش این است که [فقط] با جزء موضعی و منحصر به فردی از فضا و زمان گره می‌خورد.

با توسعهٔ تعریف دیویس و هرش، می‌توان پرسید که عناصر اولیه [یا اجزای اصلی] تجربهٔ بشری چه نوع عناصری هستند که می‌توانند انواع متفاوت و متنوع اشیای ذهنیِ بازتولیدپذیر را پدیدآورند، اشیایی که به نوبهٔ خود مظهر و مجسم‌کنندهٔ بخشهای اصلی ریاضیات به شمار می‌آیند؟ بخشهای کلاسیک ریاضیات عبارت‌اند از هندسه، جبر، و آنالیز. بیایید به‌هرکدام نگاهی بیفکنیم و سعی کنیم تجارب مربوط و اشیای ذهنی حاصله را نامگذاری بکنیم.

هندسه روشنترین و آشکارترین مورد است: هر کودک ۳ تا ۶ ماهه‌ای با حرکات سادهٔ عضلانی خود دائماً به تلفیق دو حس باصره و لامسه می‌پردازد، و [به‌مرور] یاد می‌گیرد که با حرکت مناسب دست و بازوی خود می‌تواند اسباب‌بازی خود را در چنگ بگیرد و شاهد جابه‌جا شدن آن باشد. در این مورد، لُب مطلب این است که ادراک فضا (از طریق حواس و تعامل عضلانی) به مثابهٔ آن عنصر اولیهٔ تجربهٔ ماست، همان تجربه‌ای که هندسه بر آن استوار است. این امر اصل و اساسِ یکی از ساده‌ترین اشیای ذهنی را تشکیل می‌دهد ــ همان چیزی که دیویس و هرش آن را «ریسمان کشیده» می‌خوانند. و منشأ ترسیماتی است که به‌کمک خط‌کش و پرگار انجام می‌شوند. الگوی انتزاعی برای مطالعهٔ صوری آن عبارت است از فضای Mای که از نقاطی با ساختارهای بسیار متنوع تشکیل می‌یابد.

آنالیز، به‌نظر من، زاییدهٔ تجربهٔ بشری از نیرو و فرزندان آن یعنی شتاب و نوسان است. مثال برجسته همان سیبی است که بر سر نیوتن اصابت می‌کند. این تجربهٔ اولیه منجر به شیئی ذهنی می‌شود که الگوی انتزاعی آن است و عبارت است از تابع، و مشتقات آن. توابع در ابتدا تحول کمیت فیزیکی معینی را در طول زمان توصیف می‌کردند.

و اما جبر؛ به‌نظر می‌آید که جبر از نوعی «دستورِ اعمال» پدید آمده باشد. مقصود از دستور اعمال این است که ما [معمولاً] افعال خود را به‌ترتیب یا به ترتیبهای معینی، یکی پس از دیگری، انجام می‌دهیم، و انواع اعمالِ مرتبهٔ بالاتر را بر اساس اعمال ساده‌تر و اصلیتر اجرا می‌کنیم. ساده‌ترین مثال برای این امر شاید همان عمل شمارش باشد که توانایی این کار را کودکان خیلی زود اکتساب می‌کنند. این ممکن است بخشی از دستور اعمال ماهرانه باشد اگر مثلاً تیله چینی در نظرگرفته شود یا ممکن است بخشی از دستور زبان باشد اگر مثلاً کلمات مورد استفاده قرار گرفته باشند. در اینجا شیء ذهنیِ الگو عبارت است از مجموعه‌ای از چیزهایی که دارای قانون ترکیب‌اند.

خوب، این از بخشهای «کلاسیک» ریاضیات. افزون بر این شاخه‌ها من معتقدم که تجربهٔ انسانی شاخهٔ چهارمی هم دارد که در آن شاخه هم اشیای ذهنی بازتولیدپذیر، و در نتیجه ریاضیات، پدید می‌آید: یعنی تجربهٔ خود تفکر از طریق مشاهدهٔ آگاهانهٔ ذهن ضمن کار. به‌جای مشاهدهٔ جهان و یافتن عناصر هندسه و آنالیز در آنجا، یا مشاهد کردن افعال خود و کشف جبر، در این مورد نحوهٔ کار ذهن خود را مشاهده می‌کنیم. این امر به تأسیس منطق صوری به‌دست ارسطو منجر شد که در آن گزاره‌ها همان اشیای ذهنی بنیادی به‌شمار می‌آیند. منطق عبارت بود از نظام صوری بازتولیدپذیری که به‌منظور مدلسازی جریان خام افکاری که به‌ذهن آگاه ما خطور می‌کند ساخته شده بود.

اما آیا این ادعا درست است؟ ادعای رقیب آن که مضمون اصلی حرف من است این است که تفکر عبارت است از سبک سنگین کردن احتمالهای نسبیِ رویدادهای ممکن و عمل نمونه‌گیری از «امرپسین»، یعنی توزیع احتمال روی رویدادهای مجهول، با در نظرگرفتن جمع کل اطلاع ما از رویدادهای گذشته و زمینه و متن کنونی. اگر چنین باشد، آنگاه شیء ذهنیِ الگو دیگر گزاره‌ای ــ برخوردار از شکوه ابدی با ارزش صدقِ حک‌شده بر سینه ــ نیست بلکه متغیر تصادفی xای است که مقدار آن، تابع احتمالهایی است که هنوز تعیین نشده‌اند. در بخش ۴ به متغیرهای تصادفی خواهیم پرداخت. ساده‌ترین مثالی که می‌تواند نشان دهد که تفکر انسانی از این نوع است موردی است که در آن می‌توان احتمالها را تصریح کرد: یعنی قمار. در اینجا بسیار آشکار است که ما به سبک سنگین کردن احتمالها می‌پردازیم (یا اگر تبحر ریاضی داشته باشیم گاهی حتی آنها را محاسبه هم می‌کنیم). اگر این را بپذیریم، آن بخش از ریاضیات که با این قلمرو از تجربه سروکار دارد منطق نیست بلکه احتمال و آمار است.

۳. تاریخ مختصر تقابل منطق و آمار

اگر محوری زمانی در نظر بگیریم و برخی از نقاط مهم سیر تحول این دو دیدگاه متضاد را در باب ماهیت تفکر روی آن محور مشخص بکنیم، کار ما خالی از تفنن نخواهد بود. اگر این کار را از دوران شکوفایی آتن باستان شروع کنیم می‌بینیم که چگونه برخی از اقوال افلاطون بر زبان سقراط جاری می‌شود:

> تئودوروس یا هر هندسه‌دان دیگری، اگر در هندسه اساس را برگمان بگذارد به پشیزی نمی‌ارزد (گفتگو با تئه‌تتوس، قطعهٔ ۱۶۲ e، حدود ۳۶۰ پیش از میلاد.)

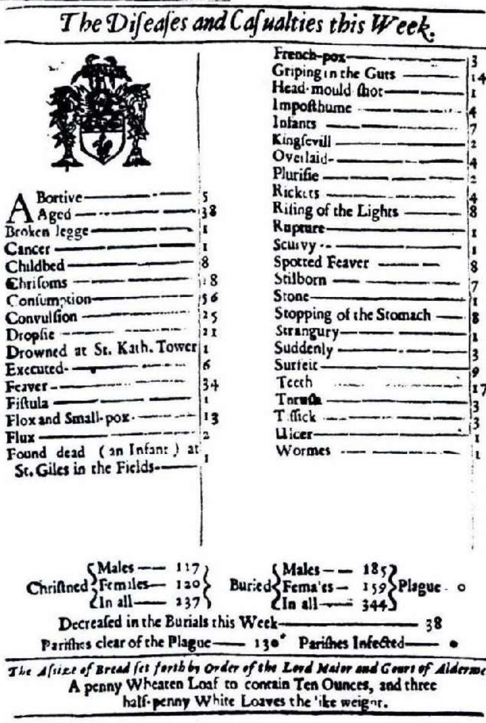
افلاطون در کتاب هفتم جمهوری، قطعه‌های c ۵۲۹ و c ۵۳۰، حتی از حدود سلیقهٔ نابگراترین ریاضیدانان معاصر خود اندکی فراتر می‌رود و می‌گوید اصلاً بهتر است که منجمان به خود ستارگان نگاه نکنند:

> بی‌گمان ما حق داریم اشکال و تصاویری را که زیب و زیور آسمان عالم محسوسات‌اند کاملترین و زیباترین تصاویر بشماریم. ولی همواره باید این حقیقت را در نظر داشته باشیم که این اشکال و تصاویر در زیبایی هرگز به‌پای چیزهای زیبای حقیقی نمی‌رسند. دلیل درستی این سخن حرکاتی است که با سرعت حقیقی و آهستگی حقیقی در عالم اعداد حقیقی و اشکال حقیقی صورت می‌گیرند، و حرکات اجسام عالم محسوسات تصویری از آنها هستند. ولی همهٔ آن حرکات را فقط از راه تعقل و تفکر می‌توان دریافت نه به‌واسطهٔ حس بینایی. . . . پس ستاره‌شناسی را نیز چون هندسه یکی از رشته‌های آموزشی قرار خواهیم داد تا وسیله‌ای برای تمرین فکر باشد. ولی چون منظور اصلی ما از آموزش این دانش، این است که آن جزء روح که ماهیتش تعقل و تفکر است از حال خمود بدر آید و کامل و نیرومند گردد، چندان اهمیتی به‌اجرام آسمانی نخواهیم داد.[۱]

به‌همین ترتیب، می‌بینیم که برخی از اشتباهات فاحش ارسطو، علی‌رغم مطالب مفصلی که در زمینهٔ زیست‌شناسی نوشته، به‌این دلیل بوده است که هرگز با اطبایی مانند بقراط و شاگردانش دربارهٔ اطلاعات واقعی در مورد بدن انسان مشورت نمی‌کرده است. مثلاً او معتقد بود که جایگاه تفکر قلب است و نه مغز، عقیده‌ای که به‌صرف ملاحظهٔ آثار آسیب مغزی به‌آسانی قابل رد است. (نگاه کنید به مقالهٔ چارلز گروس [۶]).

اگر جلوتر بیابیم، می‌بینیم که در عصر رنسانس، کاردانو (۱۵۰۰ـ۱۵۷۱) شخصیت بی‌نظیری است. از یک سو او به‌خاطر کتاب فن کبیر[۲]ش (۱۵۴۵) غالباً مبدع i خوانده می‌شود. ظاهراً وی یکی از خبره‌ترین افراد در زمینهٔ عملیات صوری جبر بود به‌طوری که تبعات قواعد منطقی جبر را یک گام فراتر از اسلاف خویش برد. ولی در عین حال، معتاد به قمار هم بود و در کتاب بازیهای شانسی[۳] خود نخستین تحلیل را از قوانین شانس ارائه کرد اما خجالت می‌کشید آن را انتشار دهد و این کتاب تا ۱۶۶۳ به‌چاپ نرسید، یعنی تقریباً مقارن با زمانی که یاکوب برنولی کار خود را آغاز کرد. در قرن ۱۷، نیوتن و لایب نیتس را در کانون اردوگاه منطق می‌یابیم، به‌طوری‌که نیوتن بر این قول است که هندسهٔ اقلیدسی تنها زبان قابل اطمینان برای اثباتهای قابل اعتماد است و لایب نیتس، هم اصول هوش مصنوعی (AI) امروزی را در رسالهٔ دکتری خود به‌نام فن ترکیب[۴] پیش‌بینی می‌کند. در اردوگاه آمار، تجربه‌گرایان واقعی را می‌یابیم که شروع به گردآوری و تحلیل اطلاعات و آمار کرده‌اند. گراونت[۵] جدولهای مرگ و میر را در لندن تنظیم می‌کند (نگاه کنید به‌شکل ۱ که به سال ۱۶۶۵ تعلق دارد) و یاکوب برنولی با اثبات قانون اعداد بزرگ، کاربرد برآوردهای تجربی را توجیه می‌کند.

عالیجناب توماس بیز در قرن ۱۸ زندگی می‌کرد (۱۷۰۱ یا ۱۷۰۲ـ۱۷۶۱).


The Diseases and Casualties this Week.

Abortive	5
Aged	38
Broken legge	1
Cancer	1
Childbed	8
Chrisoms	8
Consumption	56
Convulsion	25
Dropsie	21
Drowned at St. Kath. Tower	1
Executed	6
Feaver	34
Fistula	1
Flox and Small-pox	13
Flux	2
Found dead (an Infant) at St. Giles in the Fields	1
French-pox	3
Griping in the Guts	14
Head-mould shot	1
Imposthume	4
Infants	7
Kingsevill	2
Overlaid	4
Plurisie	2
Rickets	4
Rising of the Lights	8
Rupture	1
Scurvy	1
Spotted Feaver	8
Stilborn	7
Stone	1
Stopping of the Stomach	8
Strangury	1
Suddenly	3
Surfeit	9
Teeth	17
Thrush	3
Tissick	3
Ulcer	1
Wormes	1

Christned: Males 117, Females 120, In all 237
Buried: Males 185, Females 159, In all 344 — Plague 0
Decreased in the Burials this Week — 38
Parishes clear of the Plague — 130 Parishes Infected — 0
The Assize of Bread set forth by Order of the Lord Maior and Court of Aldermen;
A penny Wheaten Loaf to contain Ten Ounces, and three half-penny White Loaves the like weight.

شکل ۱ گراونت یکی از نخستین کسانی بود که به فایدهٔ داده‌های تجربی پی برد: در اینجا ارقام یک هفته‌ایِ مرگ و میر در لندن دیده می‌شود (مربوط به قرن هفدهم).

او بر این قول بود که باید بتوان احتمالهای پیشینی یا ماقبل تجربی را تعریف کرد، یعنی احتمالهایی که به‌رویدادهای نامعلوم، براساس تجربهٔ حاصل از رویدادهایی مرتبط ولی نه‌یکسان با آنها نسبت می‌دهند ــ وگرنه باید به دیدگاه خنثایِ لاادری اکتفا کرد. چنین احتمالهایی باید براساس مشاهدات جدید تعدیل یابند تا بتوان با انباشته شدن اطلاعات به‌احتمالهای پسینی یا مابعد تجربی هر چه بهتری دست یافت. برای آنکه اهمیت اساسی کار بیز را نشان بدهیم اجازه بدهید به سراغ مقالهٔ اصلی بخش تجارت در روزنامهٔ لوس‌آنجلس تایمز تاریخ ۱۹۹۶/۱۰/۲۸ بروم. در این شمارهٔ روزنامه در قسمت موردنظر عکسی از بیز به‌چاپ رسیده است و عنوان مقاله عبارت است از «آیندهٔ نرم‌افزار ممکن است مطابق نظریه‌های تیره و تار روحانی قرن هجدهم توماس بیز باشد». سپس در متن مقاله چنین می‌آید «وقتی اخیراً از [بیل] گیتس سؤال شد که بالاخره کِی کامپیوترها می‌توانند زبان انسان را درک بکنند» گیتس در پاسخ به اهمیت نقش اساسی «نظامهای بیزی» اشاره می‌کند . . . آیا این حرف گیتس حرف حساب است؟ آیا این تکنولوژی «بیگانه‌نما» جدیدترین سلاح سری مایکروسافت است؟ در بازشناسی گفتار، احتمالهای پیشینی می‌تواند مدل عام [همگانهٔ] زبان آدمی باشد و احتمالهای پسینی مدل بسیار دقیقتر زبان یک فرد پس از آموزش. اگرچه تایمز این حرفها را نظریه‌های تیره و تار خوانده است، ولی مکتب رو به رشدی از پژوهشگران (که خودم هم به‌آن تعلق دارم) بر این اعتقاد است که آمار بیزی کلید کاربرد مؤثر استنباط آماری در موقعیتهای پیچیده است.

گاوس از این لحاظ که هم در قیاسهای منطقی و هم در آمار کاربردی توانایی خارق‌العاده‌ای داشت، آدم جالبی است. در حقیقت، او روش کمترین

---

۱. ترجمهٔ فارسی این قول از دورهٔ آثار افلاطون، ترجمهٔ محمدحسن لطفی، انتشارات خوارزمی، صص. ۱۱۵۲ و ۱۱۵۳، نقل شده است.

2. Ars Magna 3. Liber de Ludo Aleae

4. De Arte Combinatoria 5. Graunt

$$*110{\cdot}643.\ \vdash .\ 1 +_c 1 = 2$$

Dem.

$$\vdash .\ *110{\cdot}632\ .\ *101{\cdot}21{\cdot}28\ .\ \supset$$
$$\vdash .\ 1 +_c 1 = \hat{\xi}\{(\exists y)\ .\ y \in \xi\ .\ \xi - \iota\text{‘}y \in 1\}$$
$$[*54{\cdot}3]\quad = 2\ .\ \supset \vdash .\ \text{Prop}$$

The above proposition is occasionally useful. It is used at least three times, in $*113{\cdot}66$ and $*120{\cdot}123{\cdot}472$.

شکل۲ دستاورد افتخارآمیز رهیافت تقلیلگرا به مبانی ریاضیات. قضیهٔ بالا حدود هزار صفحه را در اثر ماندگار راسل و وایتهد به نام پرینکیپیا ماتماتیکا که صرفاً مبتنی بر منطق و نظریهٔ مجموعه‌هاست، اشغال کرده است.

مربعات را به‌منظور مهارکردن داده‌های غیردقیق ولی زائد ابداع کرد، و همین امر به کشف مجدد سیارک سرس منجر شد، و در عین حال قضیهٔ حدی مرکزی را اثبات کرد که خود توجیه‌کنندهٔ این روش بود. معروفترین آزمایش او برای آزمون یک فرضیه همان است که برای آزمودن ماهیت اقلیدسی جهان سه‌بعدی ما انجام داد. او این کار را با اندازه‌گیری سه‌زاویهٔ مثلثی که از سه قله تشکیل شده بود انجام داد: سه قلهٔ بروکن، هوهاگن و اینزلزبرگ: حاصل جمع زوایا ۱۴٫۸۵ ثانیه بیشتر از π شد، ولی به‌هر صورت در محدودهٔ خطای آزمایشی π بود. اردوگاه منطق در طول قرن نوزدهم به‌اوج شکوفایی رسید، از برشهای ددکیند برای حسابی کردن اعداد حقیقی گرفته تا منطق بول، صوری شدن حساب محمولات به دست فرگه، و صوری شدن نظریهٔ مجموعه‌ها به‌دست کانتور. بد نیست یکی از نقاط عطف این مکتب را یادآوری بکنیم: یعنی اثبات این قضیه که ۲ = ۱ + ۱، به‌دست راسل و وایتهد. (این قضیهٔ شمارهٔ ۱۱۰٫۶۴۳ از پرینکیپیا ماتماتیکا ی آنهاست). شکل ۲ را نگاه کنید و به اظهارنظر آنها در پاراگراف بعدی آن توجه کنید! اما گردآوری آمار تجربی هم در قرن نوزدهم به‌شکوفایی رسید، به‌خصوص به‌دست فرانسیس گالتون، که علاقه‌اش به‌اندازه‌گیری انواع و اقسام امور مربوط به اشخاص تا حدی بود که امروزه نمی‌توان او را چندان «از نظر سیاسی صالح»[1] به‌شمار آورد.

حال اگر به‌سراغ قرن خودمان بیاییم، می‌بینیم که مهمترین گرایش، عرضهٔ مدلهای احتمال پیچیده‌تر و حقیقتاً جالبتری بوده است که کاربردهای عمیقتری در علوم داشته‌اند. برای مثال کارگالتون محدود بود به برازش توزیعهای گاوسی به مجموعه‌های داده‌های کم‌بُعد یا غیربرداری. جهش بزرگ در این زمینه زمانی رخ داد که گیبس تعریفی از مدلهای احتمال با بعد زیاد در فیزیک ارائه کرد، مثلاً برای گازها که منجر به پیدایش مکانیک آماری شد. کینز هم در باب مبانی احتمال و هم دربارهٔ اقتصاد مطالبی نوشت و سعی کرد تاکاربرد صحیح استدلالهای احتمالاتی را در جهان واقع روشن کند. وینر روشهای تصادفی را در مورد پیش‌بینی سیگنال و نظریهٔ کنترل به کار بست. شانون روشهای تصادفی را برای فشرده ساختن داده‌ها به‌کار برد و از این طریق، نقش اساسی آنتروپی توزیع احتمال را کشف کرد. گرنادر[2] روشهای تصادفی را اول در مورد ساختارهای جبری و بعداً دربارهٔ الگوهایی که آنها در جهان، به‌خصوص در

۱. یک یادداشت شخصی: پدربزرگ من، آلفرد مامفرد، سالها پزشک دبستانی در منچستر بود و خیلی مجذوب همبستگیهایی بود که در اندازه‌گیری دقیق پارامترهای مربوط به‌سلامتی بچه‌ها می‌یافت.

2. Grenader

زمینهٔ بینایی، ایجاد می‌کنند به‌کار بست. همهٔ اینها ابزارهایی قوی و مثالهایی الهام‌بخش از کاربرد روشهای تصادفی در اختیار ما قرار داده‌اند.

در همان حال که این کاربردهای واقعاً هیجان‌انگیز از آمار به‌دست می‌آمد، بخش اعظم خودِ جامعهٔ آماردانان، به‌رهبری سِر آر. آ. فیشر سرگرم بستن دست خود از پشت بودند، و اصرار می‌کردند که آمار را نمی‌توان در موقعیتهایی جز موقعیتهای کاملاً تکرارپذیر به‌کار بست و تازه آن هم فقط به‌کمک داده‌های تجربی. این قول به‌مکتب «فراوانی‌گرا» شهرت دارد که همواره با مکتب بیزی سر جنگ داشته‌اند. مکتب اخیر معتقد بوده که امور پیشینی [ماقبل تجربی] قابل استناد هستند و استنباط آماری را می‌توان به‌نحو بارزی بسط و گسترش داد. رهیافت فیشر منکر بسط استنباط آماری به قلمرو تفکر واقعی است چون موقعیتهای زندگی واقعی همیشه در میان انبوه متغیرهای موقعیتی مندرج و مدفون، و غیرقابل تکرارند. خوشبختانه مکتب بیزی کاملاً از بین نرفت و دیگران از جمله دفینتی، جینز و دیگران به استمرار و شکوفایی آن یاری رساندند. کاربردهای جدید آمار بیزی در زمینه‌های بینایی، گفتار، دستگاههای خبره، و سیستمهای عصبی رشدی انفجاری داشته‌اند.

۴. متغیر تصادفی کدام است؟

در واقع این قول، قول دیوید کاژدان است: وقتی او سمینار گلفاند را به هاروارد انتقال داد، اسم آن را «سمینار بنیادی» گذاشت و از تک‌تک افراد سمینار خواست که یکی از مفاهیمی را که خیلی خوب می‌دانند و هرکس باید آن را بیاموزد، توصیف بکنند. و موضوع عنوان این بخش را به پرسی دیاکونیس محول کرد. من از این ابتکار کاژدان خوشم آمد: توصیف متغیر تصادفی کار ساده‌ای نیست. این مفهوم، مفهوم اصلی نظریهٔ احتمال و آمار است، و به‌همین دلیل در اشکال متفاوت ظاهر می‌شود. بگذارید فهرستی از توصیفهای مختلف آن تنظیم کنیم:

- متغیرهای تصادفی تجربی‌ای وجود دارند؛ برای مثال، چنین متغیرهایی زمانی پدید می‌آیند که وزن و قد نمونه‌ای از افراد را به‌صورت جدول تنظیم بکنیم؛ [یا] آنکه تصویری را بگیریم و شدت پیکسل‌های آن را اندازه بگیریم؛ یا آنکه دارت[1] را به‌سوی تختهٔ آن پرتاب کنیم و محل اصابتش را نسبت به خال اندازه بگیریم.
- متغیرهای تصادفی بنیادی‌ای وجود دارند؛ برای مثال، نمونه‌ای تصادفی از مجموعه‌ای متناهی و دارای توزیع یکنواخت؛ عددی حقیقی که توزیع تصادفی و نرمال داشته باشد؛ نمونه‌ای تصادفی از حرکت براونی.

1. dart

• متغیرهای تصادفی حقیقتاً پیچیده‌ای هم وجود دارند. مثال بارز آن عبارت است از جواب یک معادلهٔ دیفرانسیل جزئی تصادفی که دارای یک جملهٔ محرک نوفهٔ سفید باشد. مثال دیگر عبارت است از خمینه‌ای تصادفی که به‌کمک نوعی عناصر تصادفی بنیادی ساخته شده باشد. گروموف در سخنرانی خود برخی از آنها را توصیف و تبیین کرد.

• تشخیص یک پزشک را هم می‌توان یک نمونهٔ تصادفی در نظر گرفت، نمونه‌ای مرکب از توزیع احتمال پسینی [مابعد تجربی] او در ارتباط با وضعیت جسمی شما و با در نظرگرفتن ترکیب این عوامل: (الف) تجربهٔ شخصی خود پزشک؛ (ب) دانش اکتسابی او از منابعی چون کتاب، مقاله، و پزشکهای دیگر؛ (ج) سابقهٔ بیماری خودتان؛ (د) نتایج آزمایشهایتان. برای اطلاع بیشتر مراجعه کنید به‌مقالهٔ بسیار خوب لوریتسن و اشپیگل‌هالتر [۱۱].

• رُمان را هم می‌توان نوعی نمونهٔ تصادفی از توزیع احتمالِ پسینی نویسنده در ارتباط داستانها دانست که مشروط است به همه چیزهایی که نویسنده دربارهٔ ماهیت جهان واقعی مشاهده کرده یا آموخته است. این مطلب را در بخش آخر مقاله بسط خواهیم داد.

• [متغیر تصادفی را] می‌توان عملگر تعریف‌نشده‌ای در چارچوب اصل موضوعی کردن ریاضیات تصورکرد: مراجعه کنید به بخش بعد.

• چه‌بسا هر مشاهده‌ای در مکانیک کوانتومی نوعی «متغیر تصادفی ناجابه‌جایی» از آب درآید، البته اگر دیدگاه آلن کُن را که در سخنرانیش مطرح کرد اتخاذ بکنیم.

وقتی احتمال بر نظریهٔ اندازه مبتنی شود، تعریف صوری و متعارف متغیر تصادفی با مقادیری در مجموعهٔ X، این است که این مفهوم در حقیقت یک تابع اندازه‌پذیر چون: $x : \Omega \to \mathrm{X}$ از فضای احتمال Ω به X است. اما، خود فضای احتمال معمولاً نقشی ایفا نمی‌کند و x به‌گونه‌ای عمل می‌کند که گویی عضوی شناور از مجموعهٔ X است (مانند نقطهٔ عام در هندسهٔ جبری). در نتیجه: (i) Ω برای متغیرهای تصادفی تجربی اساساً غیرقابل شناخت است؛ (ii) [ولی] برای متغیرهای تصادفی بنیادی، $\Omega = X$؛ (iii) برای متغیرهای تصادفی پیچیده، Ω حاصلضرب بزرگی از فضاهای احتمال است، فضاهایی که همهٔ عناصر تصادفی موجود در ساخت را تأمین می‌کنند؛ (iv) اما برای رمان‌نویس یا پزشک، Ω مدل احتمالی کاملی است که آنها در مورد نحوهٔ کارکرد جهان ساخته‌اند.

برای پروراندن نظریهٔ بنیادی احتمال دو رهیافت وجود دارد. بنابر یک رهیافت، هرجا که ممکن است باید سعی کرد که کل زبان احتمالاتی حذف شود و همه چیز به نظریهٔ اندازه تقلیل یابد. آنگاه Ω هم حذف می‌شود و X به اندازهٔ $p(x)$ یا $p(x)dx$، که از تصویر مستقیم تحت نگاشت x از اندازهٔ احتمال روی Ω حاصل می‌شود، مجهز می‌گردد. رهیافت دیگر آن است که مفهوم «متغیر تصادفی» در مرکز توجه قرار بگیرد و همهٔ کارها با انواع و اقسام دستکاری در متغیرهای تصادفی انجام شود. اکنون مثالی ارائه می‌کنیم که اختلاف این دو رهیافت را خوب نشان می‌دهد.

یک متغیر تصادفیِ دارای مقدار حقیقی چون x، و مفهوم «تقسیم‌پذیری نامتناهی» (ID) را در مورد آن در نظر بگیرید. می‌توان سنتی عمل کرد و چگالی احتمال x را با $p(x)$ نشان داد. آنگاه x، ID است به‌شرط آنکه به‌ازای هر n، چگالی احتمال $q_n(x)$ وجود داشته باشد به‌طوری‌که $p = q_n * \cdots * q_n$ (n، q_n را به‌عواملش تجزیه می‌کند). یا آنکه می‌توان گفت که به‌ازای هر nای، $x \sim y_1 + \cdots y_n$ چنان‌که y_iها عبارت‌اند از متغیرهایی تصادفی که مستقل از هم ولی یکسان توزیع شده باشند (و $\sim$ به‌معنای آن است که تابع قانون واحدی‌اند).

این کار چیزی جز تغییر ساده‌ای در نمادگذاری نیست ولی اکنون ببینید وقتی قضیهٔ لوی‌ـ‌خینچین را به دو صورت متفاوت بیان می‌کنید چه اتفاقی می‌افتد. نحوهٔ اول بیان قضیه می‌گوید که x، ID است اگر و تنها اگر تبدیل فوریهٔ $p(x)$ یعنی $\hat{p}(\xi)$ قابل نوشتن به‌صورت زیر باشد

$$\hat{p}(\xi) = e^{ia\xi - b\xi^2 - c\int(e^{i\xi y} - 1 - \frac{i\xi y}{1+y^2})d\mu(y)}$$

در روش دوم، همان شرط مستقیماً برحسب متغیر تصادفی x به‌صورت ذیل بیان می‌شود:

$$x \sim a + bx_{\text{نرمال}} + c\sum(x_i - c_i \text{(عامل همگرایی)})$$

که در آن $x_{\text{نرمال}}$ متغیر نرمال استاندارد و $\{x_i\}$ یک فرایند نرمال پواسون است که از چگالی ν اخذ شده است. خب می‌بینید که آنها کاملاً متفاوت به‌نظر می‌رسند! از نظر من، روش دوم برای بیان قضیهٔ لوی‌ـ‌خینچین به‌مراتب روشنتر است: با تصریح کردن متغیرهای تصادفی بلافاصله معنای واقعاً تصادفی نتیجه آشکار می‌شود.

۵. درج متغیرهای تصادفی در مبانی

در رهیافت تقلیل‌گرا، متغیر تصادفی برحسب اندازه تعریف می‌شود که خود برحسب نظریهٔ [اعداد] حقیقی تعریف می‌شود، و این را هم نظریهٔ مجموعه‌ها تعریف می‌کند که خودش بر اساس حساب محمولات تعریف می‌شود. در عوض من می‌خواهم بگویم که باید علی‌الاصول بتوان متغیرهای تصادفی را در مبانی منطق و ریاضیات ادغام کرد و به صورتبندی شفافتر و کاملتری از دیدگاه تصادفی رسید. من خودم هنوز صورتبندی کامل و قطعی از این قضیه ندارم، ولی طرح آن را که بر دو منبع مختلف و اندیشه برانگیز استوار است در دست دارم. منبع اول عبارت است از کارهایی که جینز در ارتباط با مبانی احتمال و آماربیزی کرده است [۹]؛ منبع دوم استدلال تصادفی زیبایی است که کریستوفر فریلینگ برای رد فرضیهٔ پیوستار اقامه کرده است [۳].

اول می‌پردازیم به‌جینز: همان‌طورکه دیدیم، فضای احتمال Ωای که برای متغیرهای تصادفی در کاربردهایی چون تشخیص پزشکی مورد نیاز است به‌طور دقیق غیرقابل تعیین است. عوامل تعیین‌کنندهٔ اجزای تجزیه به‌حدی زیادند که عملاً هیچگاه در جدول احتمالات پزشک قابل تصریح نیستند. هستهٔ انتقاد پیروان مکتب فراوانی‌گرا از روشهای بیزی دقیقاً همین مسأله بود. به‌اعتقاد من، جینز قانع‌کننده‌ترین جواب را برای این مسأله دارد. او نظریهٔ خود را با این فرض آغاز می‌کند که کلیهٔ عاملها، از جمله خود ما، همیشه به هر رویداد Aای «مقبولیت» [«معقولیت»]ی نسبت می‌دهند که در مجموعه‌ای به‌طور خطی مرتب و نامعلوم قرار دارد ــ آن را $\mathcal{P}l$ می‌نامیم. در حقیقت، مقبولیتها را فقط به رویدادها نسبت نمی‌دهیم بلکه به‌رویدادهای شرطی نیز

نسبت می‌دهیم ــ اگر وقوع B معلوم باشد، آنگاه مقبولیتِ راست بودن A چیست؟ این مقبولیت را به صورت زیر نمایش می‌دهیم:

$$p(A|B) \in \mathcal{P}l$$

نتیجه‌گیری جینز این است که به‌کمک چند اصل موضوع معقول، می‌توان استنتاج کرد که یک یکریختی ترتیبی چون $\mathcal{P}l \cong [0,1]$ وجود دارد که تحت آن، p تبدیل می‌شود به نوعی توزیع احتمال روی جبر Aها (به‌خصوص $(p(A|B) = p(A \wedge B)/p(B)$). می‌توانیم این نتیجه را به‌این صورت خلاصه بکنیم که نظریهٔ احتمال، نظریهٔ تجویزیِ مقبولیت است، یعنی آنکه اگر قواعد طبیعی سازگاری درونی را بر هر مفهوم خودساخته از مقبولیت تحمیل بکنیم، به‌نوعی مدل احتمال صادق می‌رسیم. برای جزئیات این استدلال، مراجعه کنید به کتاب جالب او (فصلهای ۱ و ۲ از [۹]) که بعد از فوتش ظاهراً در دست چاپ است.

بر مبنای این مطالب می‌توانیم پیشنهاد زیر را دربارهٔ حساب محمولات تصادفی ارائه بکنیم. نحو این حساب باید مشابه نحو حساب محمولات متعارف باشد فقط با این فرق که در آن دو نوع متغیر داریم: محمولات و ثوابت متعارف به انضمام متغیرهای آزاد سورپذیر x و مجموعه‌ای از ثوابت تصادفی $\underline{x}$. افزون بر این، این حساب دارای تابع صدقی چون p است که همه فرمولهای F فاقد متغیر آزاد را به‌اعداد حقیقی میان ۰ و ۱ می‌نگارد. اگر فرمول F فقط شامل متغیرهای متعارف باشد، آنگاه $p(F) \in \{0,1\}$. معناشناسی صوریِ این نظریه، توابعی ثابت و تصادفی روی فضاهای احتمال می‌سازد چنانکه هر فرمولی زیر مجموعه‌ای از حاصلضرب این فضاها تعریف می‌کند، در نتیجه احتمال دارد.

نظریهٔ صوریِ تصادفیِ اعداد به‌اندازهٔ کافی صاحب قدرت بیان خواهد بود که به‌آن بتوان اصل موضوع پیوستگی برای p را اضافه کرد:

$$p(\exists n \underline{F}(n)) = \mathrm{l.u.b.}_m p((\exists n \leq m)\underline{F}(n))$$

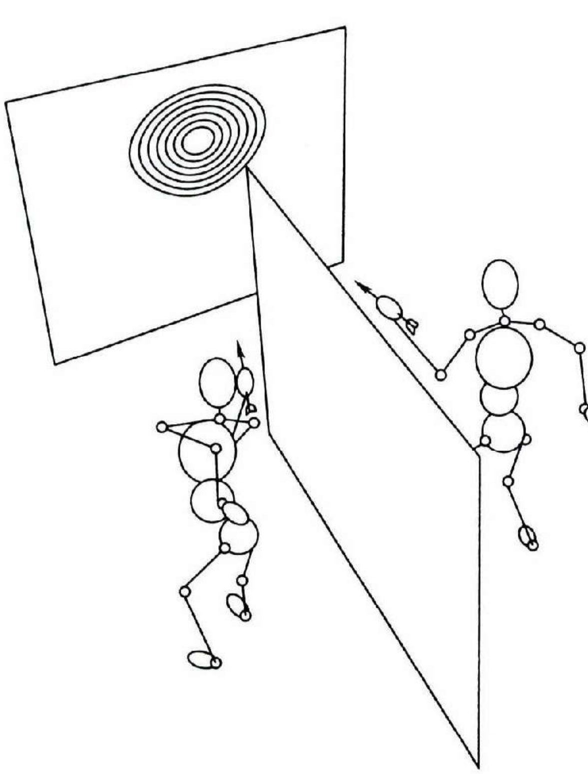
همچنین به‌اصول موضوعی نیاز داریم که متغیرهای تصادفی بنیادی را به‌ما عرضه کند. بنابراین اگر $\mathcal{N}$ محمولی باشد که به‌کمک آن اعداد طبیعی را بتوانیم تعریف بکنیم، متغیرهای تصادفی برنولی هم به‌کمک فرا‌ـاصولی چون اصول موضوع ذیل حاصل می‌شود

$$(\forall a \in \mathcal{P}l)(\exists \underline{x}_a) \ni \mathcal{N}(\underline{x}_a) \wedge [p(\underline{x}_a = 0) = 1-a] \wedge [p(\underline{x}_a = 1) = a]$$

در حقیقت، آنچه مطلوب است عبارت است از خانواده‌هایی از متغیرهای مستقل و شمارش‌پذیرِ برنولی. به همین ترتیب، اصل موضوع بنیادی آنالیز تصادفی باید وجود پیوستاری باشد که برمبنای شروط ذیل تعریف می‌شوند: (i) محمول $\mathcal{C}$، (ii) ترتیب خطی $<(c_1 , c_2)$ بین اعداد c_1 و c_2 که $\mathcal{C}(c)$ در موردشان صدق بکند، (iii) یک $\underline{x}_0$ تصادفی که در $p(\mathcal{C}(\underline{x}_0)) = 1$ صدق بکند؛ (iv) وجود اصل موضوع ذیل:

$$(\forall a \in \mathcal{P}l, a \neq 0,1)[\exists! c) \ni \mathcal{C}(c) \wedge [p(<(\underline{x}_0, c)) = a]$$

اگر بخواهیم مطلب را به‌زبان معمولی بیان کنیم باید بگوییم منظورمان این است که می‌توانیم پیوستار را به‌گونه‌ای مرتب کنیم که بازه‌های یکطرفه‌اش همهٔ



شکل۳ دو بازیکن دارت رودرروی هم. فریلینگ از این بازی برای ابطال فرضیهٔ پیوستار الهام گرفت.

احتمالهای ممکن میان صفر و یک را دقیقاً یک‌بار عرضه کند، یعنی به‌زبان متعارف، پیوستار دقیقاً همان چیزی است که شما می‌توانید دارت را به‌سویش پرتاب کنید. بازی دارت (به‌زبان صوری، اندازهٔ لبگ روی (۱ و ۰)) به‌وسیلهٔ متغیر تصادفی بنیادی $\underline{x}_0$ تعیین می‌شود، متغیری که بر مبنای اصل موضوع مذکور متغیرهای نحوی اعداد حقیقی را به‌متغیرهای معناییِ مقدار مقبولیت مربوط می‌کند. این امر نظریهٔ اندازه را در بطن مبانی نظریه قرار می‌دهد.

از اینجا به‌سوی نتیجه‌گیری حیرت‌آور کریستوفر فریلنگ [۳] هدایت می‌شویم: او با الهام از بازی دارت نتیجه گرفت که ما می‌توانیم فرضیهٔ پیوستار را رد بکنیم. من نمی‌دانم چرا همه از قضیهٔ او اطلاع ندارند و چرا آن را همتای نتایج گودل و کوهن نمی‌دانند. ولی استدلال او را به‌زبان کلاسیک می‌توان چنین بیان کرد (نگاه کنید به شکل ۳). دو نفر دارت‌باز را تصور کنید که مستقل از یکدیگر دارتهای خود را به‌سوی تختهٔ آن پرتاب می‌کنند. اگر فرضیهٔ پیوستار صادق باشد، نقاط P روی سطح تختهٔ دارت را می‌توان خوش ترتیب تصور کرد به‌طوری‌که به‌ازای هر Pای مجموعهٔ Qهایی که $Q < P$ (این مجموعه را S_Q می‌خوانیم) شمارش‌پذیر است. فرض کنید که دارتِ بازیکنهای ۱ و ۲ در نقاط P_1 و P_2 به تخته اصابت کنند. در آن صورت، یا $P_1 < P_2$ یا $P_2 < P_1$. فرض کنید اولی صادق باشد. آنگاه P_1 به‌زیرمجموعهٔ شمارش‌پذیری چون S_{P_2} از نقاط روی تختهٔ دارت تعلق خواهد داشت. از آنجا که این دو پرتاب مستقل از یکدیگر بودند، می‌توانیم تصور بکنیم که پرتاب ۲ اول رخ داده باشد و بعد پرتاب ۱ انجام شده باشد. بعد از پرتاب ۲، این مجموعهٔ شمارش‌پذیر S_{P_2} دیگر تثبیت شده است. اما هر مجموعهٔ شمارش‌پذیری اندازه‌پذیر است و اندازه‌اش صفر است. بنابراین

احتمال اینکه پرتاب ۱ به S_{P_2} اصابت کرده باشد صفر است. همین استدلال باز نشان می‌دهد که احتمال اینکه P_2 هم به S_{P_1} اصابت کند صفر است. بنابراین تقریباً به‌طور قطع هیچ‌کدام رخ نداده است و این امر فرض ما را که بنا بر آن تختهٔ دارت نخستین کاردینال شمارش‌ناپذیر است نقض می‌کند!

«نقص» این استدلال چیست؟ این است که متغیرهای تصادفی یعنی پرتابهای دارت را همچون اموری واقعی در نظر گرفته‌ایم! اگر سعی کنیم که این استدلال را در قالب نظریهٔ اندازهٔ کلاسیک بیان بکنیم اول از همه نیاز داریم که گرافِ خوش‌ترتیبی اندازه‌پذیر باشد، فرضی که طبعاً نباید زیر بارش برویم. بنابراین آیا از خیر این اثبات می‌گذریم؟ فریلینگ از این استدلال استفاده کرد تا انگیزه‌ای برای اضافه کردن اصل موضوع جدیدی به‌نظریهٔ مجموعه‌ها ایجاد کند که فرضیهٔ پیوستار را نقض بکند. من خودم معتقدم باید به مراتب فراتر از این برویم: «اثبات» او نشان می‌دهد که اگر ما متغیرهای تصادفی را به یکی از اجزای بنیادی ریاضیات تبدیل بکنیم، آنگاه نتیجه می‌شود که فرضیهٔ پیوستار کاذب است و ما خود را از دست یکی از مسائل لاینحل و بی‌معنای نظریهٔ مجموعه‌ها خلاص کرده‌ایم. فرضیهٔ پیوستار به‌یقین مشابه همان مسألهٔ [قرون وسطایی] مدرسیون است که چه تعداد فرشته بر نوک یک سوزن می‌توانند بایستند: مسأله‌ای که با تغییر دیدگاه ما محو می‌شود.

با این مطلب به دشوارترین بخش تجدید نظر پیشنهادی ما در مبانی می‌رسیم: باید تصمیم بگیریم که نظریهٔ مجموعه‌های تصادفی را چگونه تبیین کنیم. روشن است که یا باید اصل موضوع انتخاب را حذف بکنیم یا اصل موضوع مجموعهٔ توانی را. ولی وجود اشیای تصادفی، نوعی اصل موضوع انتخاب تصادفی است و در نتیجه به‌اعتقاد من فعلاً بهتر است که از اصل موضوع مجموعهٔ توانی صرف‌نظر بکنیم. آنچه ریاضیات به‌واقع برای هر مجموعهٔ X احتیاج دارد مجموعهٔ عظیم 2^X نیست بلکه مجموعهٔ دنباله‌های $X^{\mathbb{N}}$ در X است. مضافاً، از آنجا که $p(\underline{x}_. \in A) \in \mathcal{P}l$ باید برای هر زیر مجموعه‌ای چون A از $\mathcal{C}$ تعریف بشود، ضروری است که هر زیر مجموعهٔ تعریف‌پذیر $\mathcal{C}$ اندازه‌پذیر باشد. این مبحث با تخصص من فاصله دارد ولی به‌نظرم می‌رسد نتایج شلاه و وودین [۱۵] نشان می‌دهند که‌این طور نیست که این امر به‌وضوح ناسازگار یا غیرعملی باشد.

۶. تهاجم روشهای تصادفی به‌قلمرو ریاضیات کلاسیک

خالی از فایده نیست اگر نگاهی هم به‌مباحث متعددی از ریاضیات کلاسیک بیفکنیم که بر اثر استفاده از روشهای تصادفی متحول شده‌اند و تعمیق یافته‌اند.

در مبحث ترکیبیات، نظریهٔ گراف نمونهٔ بارز به‌شمار می‌آید. تهاجم با استفادهٔ اردوش از گرافهایی تصادفی که دارای تعداد ثابتی رأس و یال هستند شروع شد، و به کشف زیبای پدیدهٔ گذار فاز منجر شد: گراف تصادفی به‌طور قریب به‌یقین در محدودهٔ خیلی باریکی از تعداد یالها همبند می‌شود. یک نمونهٔ جالب، کاربرد روشهای تصادفی برای ساختن گرافهایی با کمر (یعنی با حداقل اندازهٔ دور) و درجهٔ معین است که تقریباً حداقل ترتیب ممکن را دارند. امروزه این رهیافت را «روش احتمالاتی» می‌خوانند و این مسأله در کتاب اسپنسر [۱۶] به‌خوبی توصیف شده است. در جهتی دیگر، نظریهٔ خوش ساخت فرایندهای شاخه‌زای تصادفی را داریم که از این سؤال آغاز می‌شود که با چه احتمالی شجره نامجهٔ یک خانواده اشرافی به انتها می‌رسد.

چه‌بسا قانع‌کننده‌ترین مورد برای درک اهمیت روشهای تصادفی عبارت باشد از کاربرد آنها در نظریهٔ معادلات دیفرانسیل معمولی [ODE] و معادلات دیفرانسیل جزئی [PDE]. وقتی معادلات دیفرانسیل برای مدلسازی طبیعت ابداع شدند همه کاملاً می‌دانستند که هر معادلهٔ خاصی، نمایشی جزئی از واقعیت است، نمایشی که برخی اثرها را ملحوظ می‌کند و برخی را اصلاً در نظر نمی‌گیرد. البته، موارد اصلی و اولیه عبارت بودند از مسألهٔ دو جسم و قانون حرکت نیوتن. این معادلات حرکت سیارات را به‌نحو زیبایی پیش‌بینی کردند، و همراه با اضافه کردن اختلالها تصویر کامل و دقیقی برای دورهٔ زمانی نسبتاً بلندی (مثلاً شاید $۱۰^{۸}$ سال) ارائه دادند. ولی وقتی مدت‌زمان شما از این حدود تجاوز می‌کند (مثلاً وقتی به $۱۰^{۹}$ سال می‌رسد)، اثرهای در نظر گرفته‌نشده و مدلسازی نشده جمع می‌شوند و تقریبْ سودمندی خود را از دست می‌دهد. با توجه به این امر وقتی به مسألهٔ سه جسم می‌رسیم تکلیف چیست؟ این امر باعث می‌شود که تحلیل تعیّنی کلاسیک معادلات گرانشی سه جسم همان قدر به جهان ربط داشته باشد که فرضیهٔ پیوستار! یک گام مهم برای آنکه این معادله ربط بیشتری به‌جهان پیدا بکند این است که یک عبارت تصادفی کوچک اضافه بکنیم. در این حالت حتی اگر عبارت تصادفی به‌سمت صفر میل بکند، آثار مجانبی آن لزومی ندارد که چنین رفتار کنند. به‌نظر می‌آید ادعای منصفانه‌ای باشد که بگوییم وقتی به هر معادلهٔ دیفرانسیل یک عبارت تصادفی اضافه کنیم مدلی بهتر برای جهان از آب درخواهد آمد و اینکه تحلیل کلاسیک آنها فقط وقتی سودمند است که چنین تحلیلی نسبت به‌چنان اختلالهایی به‌نحو مناسبی پایدار باشد.

اما امری که برای ریاضیدان حائز اهمیت بیشتری است این است که وقتی به‌معادلات دیفرانسیل از دیدگاه تصادفی نگریسته شود ماهیت تحلیلشان عوض می‌شود. در مورد معادلات دیفرانسیل کلاسیکی که جوابهای خوش‌رفتار دارند، به‌طورکلی چندان فرقی نمی‌کند که به آنها عبارت تصادفی اضافه کنیم یا نکنیم: نقطهٔ ثابت رباینده باز هم نقطهٔ ثابت رباینده می‌ماند (اگرچه اندکی «تار» می‌شود — جواب حول نقطهٔ ثابت کمی بالا و پایین می‌پرد). ولی وقتی معادله به رفتاری «آشوب مانند» یا متلاطم می‌انجامد، تصویرِ خیلی متفاوت و به‌مراتب خشنودکننده‌تری با تحلیل تصادفی آن حاصل می‌شود. در نتیجه به‌جای تمرکز کردن بر توصیف آسیب‌شناسی رباینده‌های غریب که جواب کلاسیک به‌طور مجانبی به‌آنها میل می‌کند، اکنون مرکز توجه ما وجود اندازهٔ احتمال ناوردایی است که تقریباً تمام جوابها در طول زمان در آن باقی می‌مانند. این ایده در مکانیک آماری از دلِ مطالعات مربوط به حرکت براونی و مدل آیزینگ نشأت گرفت. اما متأسفانه بسیاری از «نتایج» حاصل از این نظریه‌ها یا به صورت مکاشفه‌ای، یعنی به‌سبک استدلالهای فیزیکدانان، توجیه می‌شوند یا آنکه هنوز در مرحلهٔ رویا هستند (همان‌طور که تالاگراند[1] در سخنرانی خود اشاره کرد). اما چیزی که به‌آن امید داریم، و این امید حداقل در مواردی تحقق یافته، این است که تقریباً همهٔ مدارهای تصادفی ساختار مشابهی داشته باشند، ساختاری که بتوان به‌تفصیل آن را توصیف کرد و درک و بصیرتی واقعی از معادلهٔ دیفرانسیل به‌دست دهد.

یکی از مثالهای موفق و به‌نحو حیرت‌آوری زیبا، تحلیل معادلهٔ برگرِ تصادفی توسط وینان، سینایی و دیگران [۲] است. درحالی‌که معادلهٔ

1. Talagrand

متعارف برگر می‌تواند انواع و اقسام ضربه‌ها را ایجاد بکند، در مورد حالت تصادفیِ دوره‌ای، معلوم می‌شود که تقریباً همیشه یک و فقط یک ضربه‌ـموج برای همیشه (گذشته و آینده) دوام می‌آورد، به‌طوری‌که همه ضربه‌های دیگر را در خود جذب می‌کند. (به‌قول حامیان مالی ما) چالش عظیم این است که با تحلیل معادلهٔ ناویرـاستوکس بتوانیم به درک آشوب ــ به‌نحوی که در سخنرانی فقرمن مطرح شده ــ برسیم.

در زمینهٔ کاربرد روشهای تصادفی، فیزیک ریاضی نسبت به ریاضیات محض گامهای بلندتری برداشته است: یکی از عناصر اصلی نظریهٔ ریسمان، وارد کارکردن رویه‌های ریمانی تصادفی از طریق تعریف یک اندازهٔ احتمال روی فضای پیمانه بوده است و هاکینگ هم سعی کرده است تا از توپولوژیهای تصادفی روی فضازمان استفاده بکند.

۷. تفکر به‌عنوان استنباط بیزی

اکنون می‌خواهم در پایان مطالب خود به‌مبحثی بپردازم که بیشتر از هر مبحث دیگری در آن تبحر دارم: مبحث مدلسازی تفکر به عنوان فرایندی محاسباتی. می‌خواهم در وهلهٔ اول دو مثالِ ارائه بدهم که یکی در بارهٔ استدلال به‌کمک منطق و دیگری در بارهٔ استدلال به‌کمک احتمالات است. مثالِ استفاده از منطق، قیاس سرگرم‌کننده‌ای است که از «کتاب راهنمای رانندگان بوستون» اقتباس شده است. خواننده می‌تواند با امتحان منطق به‌کار بسته شده در آن خود را سرگرم کند.

مقدمات

الف) تولستوی نابغه بود؛
ب) فقط افراد نابغه می‌توانند قدر تولستوی را بدانند؛
ج) نابغه‌ای نیست که صاحب نوعی غرابت نباشد؛
د) تولستوی به‌سبک بلوز آواز می‌خواند؛
ه) هر آدم عجیب و غریبی را که آواز بلوز می‌خواند، آدم کودن می‌داند؛
و) آدمهای عجیب و غریب فکر می‌کنند که صاحب جاده هستند

$\Longleftarrow$

نتیجه: همیشه آدم کودنی وجود دارد که فکر بکند صاحب جاده است.

درست است که این قیاس نامعقول است، با این همه چیزی را به‌خوبی نشان می‌دهد و آن این است که در زندگی روزمره، استدلال دقیق و معتبر به‌ندرت مورد استفاده قرار می‌گیرد ــ تعمیمها معمولاً در زمینه‌ها و موقعیتهای متفاوت و با درجهٔ مقبولیت مختلف کاربرد دارند و طبعاً گره‌زدن آنها به‌همدیگر به‌تولید انواع و اقسام حرفهای بی‌ربط هم منجر می‌شود. مثلاً (د) ممکن است نوعی استعاره تعبیر شود و (و) حرف متعارف قابل قبولی است که فقط در مورد طیف معینی از آدمهای عجیب و غریب صدق می‌کند (چه‌بسا مجموعه‌ای کاملاً مجزا از مجموعهٔ (ج) باشد). به‌همین‌دلیل و با این حرفها به صورت منطقی استدلال کردن کار بی‌معنایی است. حال این حرفها را مقایسه کنید با مثال جودا پرل در ارتباط با استدلالِ مبتنی بر فهم متعارف در [۱۳]:

(الف) واتسن به هلمز در دفترش تلفن می‌کند و می‌گوید دزدگیر منزل هلمز آژیر می‌کشد. هلمز آماده می‌شود که خود را هرچه زودتر به‌منزل برساند.
(ب) ناگهان هلمز به‌یاد می‌آورد که واتسن آدم شوخ‌طبعی است و در نتیجه به حرف او شک می‌کند.

$\Longleftarrow$

(ج) هلمز به‌خانم گیبون که همسایهٔ دیگری است تلفن می‌کند. ولی خانم گیبونِ به اصطلاح سرش اندکی گرم است و شروع می‌کند به‌دادِ سخن دادن در بارهٔ سرقت و جنایت، به‌طوری‌که هلمز فکر می‌کند که او صدای آژیر را شنیده است.
(د) هلمز به‌یاد می‌آورد که در کتابچهٔ راهنمای دزدگیر نوشته شده است که زلزله سبب روشن‌شدن آژیر می‌شود.
(ه) هلمز با خود می‌گوید که اگر زلزله‌ای در کار باشد باید از طریق رادیو اعلام شود.

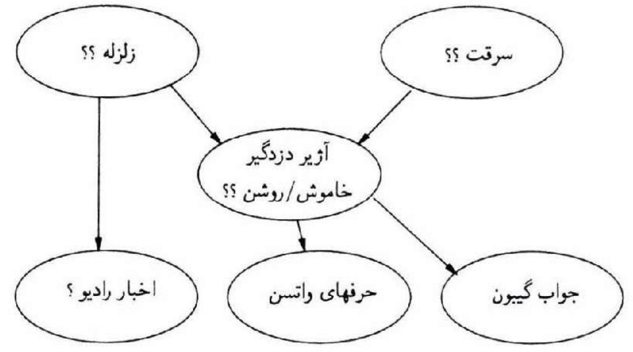
$\Longleftarrow$

(و) هلمز رادیوی خود را روشن می‌کند.

بنابر تحلیل پرل، برای فرایندهای ذهنی هلمز می‌توان مدلی بیزی ارائه کرد که او آن را نوعی «شبکهٔ بیزی» می‌خواند و در شکل ۴ نشان داده شده است. چنین شبکه‌ای گراف جهت‌داری است که گرههای آن یک سلسله رویداد را نشان می‌دهند که می‌توانند صادق باشند یا نباشند. یالها علیت [حاکم بر وقوع رویدادها] را نشان می‌دهند و احتمالهایی شرطی به‌آنها منسوب می‌شود. مجموعهٔ این احتمالهای شرطی، [اطلاعات] «پیشینی» خوانده می‌شود، که عبارت است از اطلاعات احتمالاتی که هلمز قبل از به‌صدا درآمدن زنگ تلفن روی میزش در اختیار دارد. در شکل، ۶ رأس وجود دارد که ۲ رأسش دو رویداد معلوم را نمایش می‌دهند ــ شهادت واتسن و گیبون ــ و ۴ رویداد دیگری که هلمز سرگرم سبک و سنگین کردن وقوع آنهاست. هلمز در هر مرحله از تفکرش صاحب شواهدی است ــ یعنی رئوسی که صدق و کذب‌شان معلوم است ــ و صاحب [اطلاعات] پیشینی است و باید امر «پسینی» را محاسبه کند ــ یعنی همه احتمالهای روزآمد شدهٔ کلیهٔ رویدادها بر مبنای شواهد موجود. دقت کنید که چگونه استدلال او در امتداد این گراف بالا و پایین می‌رود، یعنی وقتی او سعی می‌کند احتمالهای بهتری به رویدادهای مجهول نسبت بدهد، مثلاً از طریق تلفن کردن به منزل گیبون‌ها و یا روشن کردن رادیو. برای تفصیل بیشتر در بارهٔ این مثال مراجعه کنید به‌کتاب پرل [۱۳] صص ۴۲ـ۵۲، و الگوریتم او برای «روزآمد کردن باور» در آنجا.

یکی از مسائل اصلی درک تفکر همواره این بوده است که چگونه می‌توان یادگیری استقرایی را صوری کرد. با آنکه منطق نظامی به‌دست می‌دهد تا [قیاس استنتاجی یا] استنتاج را صوری بکنیم، درک استقرا از دیدگاه منطقی همواره امری به‌مراتب دشوارتر بوده است. به‌اعتقاد من قانع‌کننده‌ترین تعریف صوری استقرا تعریفی است که لزلی والیان[1] کشف کرده است، تعریفی که به‌مدل PAC (احتمالاً تقریباً درست)[2] شهرت یافته است. در اینجا می‌خواهم

1. Leslie Valiant 2. probably approximately correct



شکل۴ شبکهٔ بیزی باور برای داستان پرل دربارهٔ آژیر دزدگیر هلمز.

این تعریف را به‌طورکامل ارائه کنم چون این تعریف در آن واحد نشان می‌دهد که چگونه روشهای احتمالاتی به‌طور طبیعی هم در مورد یادگیری و هم در مورد استنباط به‌کاربردنی‌اند.

طرح تعریف به این صورت است: Ω مجموعه‌ای از امکانهاست و $\mathcal{C}$ هم مجموعه‌ای از محمولات است: $P : \Omega \to \{0 , 1\}$. یکی از $P \in \mathcal{C}$ صادق است و مسأله این است که Pیِ صادق را بر مبنای نمونه‌ها یا مثالهای $(y, P(y))$ حدس بزنیم. ردهٔ $\mathcal{C}$ وقتی «PAC ـیادگرفتنی» است که الگوریتمی وجود داشته باشد که حدسی چون $\hat{P}_D : \Omega \to \{0 , 1\}$ را برمبنای n مثال $\mathcal{D} = (y_1, \ldots , y_n; P(y_1), \ldots P(y_n))$ محاسبه کند و در عین حال در عبارت زیر هم صدق بکند:

$$(\forall \epsilon_1, \epsilon_2)(\exists n)(\forall [\text{توزیع احتمال } \pi \text{ بر } \Omega])$$

$$y_1, \cdots , y_n \in \Omega \text{، مستقل هم‌توزیع نسبت به } \pi \Leftarrow$$

$$\mathrm{Prob}_{\mathcal{D}}(\mathrm{Prob}_y(\hat{P}_{\mathcal{D}}(y) \neq P(y)) < \epsilon_2) > 1 - \epsilon_1$$

دقت کنید که معنای این چیست: شما در زندگی روزمره نمی‌دانید کدام مثال، متعارف است و کدام نادر، ولی در عین حال مجبورید که مثالهای یادگیری خود را از همان توزیع مثالهای آزمونی انتخاب کنید. و آنگاه همیشه احتمال کوچکی چون ϵ_1 وجود دارد که شما در واقع امر مثال گمراه‌کننده‌ای را انتخاب کرده باشید؛ ولی اگر مثالهای نوعی در اختیار شما قرار گرفته باشد آنگاه بعد از رؤیت تعدادی کافی مثال، فقط به‌اندازهٔ ϵ_2 خطا اشتباه می‌کنید. به‌نظر من برای صوری کردن یادگیری استقرایی این حرف بسیار قانع‌کننده است و این روش شاید درست‌ترین راه باشد.

حال اگر باز به‌مسألهٔ تفکر در کل باز گردیم، که شامل مدلهای یادگیری، مدلهای ذخیره‌سازی، و استدلال بر مبنای مدلهاست، می‌توانیم این فرضیه را در نظر بگیریم که تفکر تشکیل شده است از ساختن مدلهای احتمال و استفاده از استنباط بیزی. به‌نظر من برای اینکه ثابت کنیم که این حرفها نه فقط در موارد ساختگی مانند مثالهای پرل بلکه در موقعیتهای واقعی هم حرفهایی عملی هستند، سه مانع اصلی وجود دارد. مانع اول این است که در جهان واقع میلیونها متغیر تصادفی وجود دارد و جدول کامل احتمالها برای همهٔ مقادیر ممکن این متغیرها بسیار بزرگتر از آن است که قابل ذخیره‌سازی باشند. ما به یک رده یا مجموعهٔ محدودی از مدلهای احتمال احتیاج داریم، مدلهایی که به اندازه‌ای گویا باشند که بتوانند مدلی از واقعیت ارائه بدهند ولی در عین‌حال به‌اندازهٔ کافی هم مختصر و خلاصه باشند که قابل ذخیره‌سازی باشند. دوم اینکه باید نشان بدهیم که انجام استنباط براساس این مدلها، با سرعت معقولی میسر است. سوم آنکه باید نشان بدهیم که پارامترهای موجود در این رده‌های توزیعهای احتمال [همه] PACـیادگرفتنی هستند. این فهرست سفارش خیلی بلند است ولی در برخی از زمینه‌های آن کارهای اساسی انجام گرفته است و به‌نظر می‌آید که در مجموع پیشرفت چشمگیر و جالبی تحقق پذیرفته باشد.

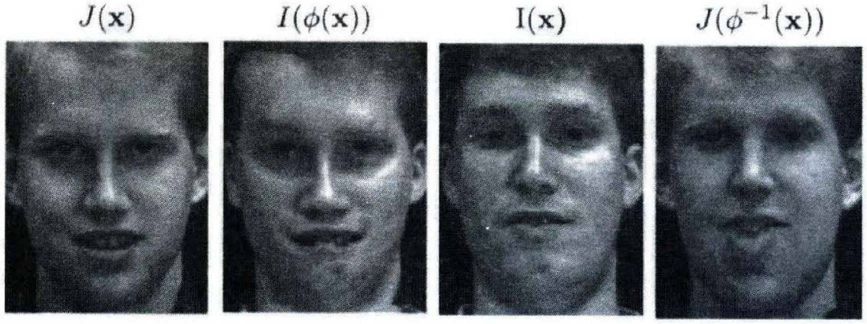
گیبس در ایجاد مدلهای عظیم ولی عملی احتمال نخستین گام عمده را برداشت. ایدهٔ او عبارت بود از در نظر گرفتن مدلهایی که در آنها لگاریتم احتمال، مجموع عبارتهایی باشد که هرکدام در هر لحظه فقط شامل تعداد کمی متغیر تصادفی باشد:

$$\mathrm{Prob}(\{x_k\}) = \frac{1}{Z} e^{-\sum_C E_C(x_C)}$$

که در آن Z ثابت است، $x_C = \{x_k | k \in C\}$، و فرض بر این است که مجموعه‌های C مجموعه‌های «کوچکی» از متغیرها هستند. انواع این مدلهای «گیبسی» به‌طور گسترده‌ای در هوش مصنوعی، بینایی، گفتار و شبکه‌های عصبی مورد استفاده قرارگرفته‌اند. در قلمرو پیوسته، می‌توان به‌این مدلها به‌چشم تعمیمهای طبیعی مدلهای گاوسی نگریست: مدلهای گاوسی دقیقاً مدلهایی هستند که لگاریتم‌ـاحتمال آنها مجموع عبارتهایی است که در هر لحظه فقط شامل دو متغیر و به شکل ax_i، bx_i^2 یا cx_ix_j باشند. ولی مدلهای گیبسی عام می‌توانند کاملاً غیرگاوسی، غیر پارامتری، و دارای متغیرهای مختلط پیوسته و گسسته باشند.

بسطهای موجکی تصاویر جهان واقع از جمله مثالهایی‌اند که مستقیماً به توزیعهای احتمال گیبسِ غیرگاوسی منجر می‌شوند. مهمترین دلیلی که برای ترجیح بسطهای موجکی تصاویر، مثلاً بر بسطهای فوریه، وجود دارد این است که ضرایب موجکی تصاویر طبیعی، پراکنده‌اند. این امر به‌معنای آن است که نوعاً تعداد نسبتاً کمی از ضرایب موجکی بزرگ هستند و بیشترشان تقریباً صفرند. به‌سخن صریحتر، رفتار آنها چنان است که گویی از توزیعی غیرگاوسی چون $p(I) = \frac{1}{Z} e^{-a \sum_\alpha \sqrt{|c_\alpha|}}$ نمونه برداری شده‌اند که در آن c_αها ضرایب موجکی تصویر I هستند.

اما به‌نظر می‌رسد که مدلهای گیبسی هم به‌تنهایی برای توصیف کل جهان واقعی به اندازه کافی گویا نباشند: به‌نظر می‌آید که باید در مدلهای احتمال لازم نوعی «ارتباطها [پیوندها]ی دینامیک» گنجانده شود، یعنی متغیرهایی که به نحوی دستوری اجزا رابه کُلهایی تبدیل بکنند. برخی از این متغیرها در حکم یک سلسله «شکاف‌پرکن» هستند، مثلاً شاخصهای اشاره‌ای که مشخص می‌کنند که فلان کلمه فاعل فلان جمله است یا فلان نقطهٔ روی شبکیه، بینی فلان چهره‌ای است که دارد ادراک می‌شود. طبعاً ارتباطها یا پیوندهای دیگری هم برای موجوداتی که خواص گروهی دارند مورد نیاز است، مثلاً چیزهایی که حرکت مشترک دارند، یا پیکسل‌های تصویرکنندهٔ شیء واحدی در چشم چپ و در چشم راست. یکی از زمینه‌های تحقیقاتی مهم روز، ساخت صحیح مدلهای احتمال با چنین پیوندهای دینامیک است.



شکل۵ این مثال از اثر هالینان در بارهٔ انطباق چهره‌ها به‌وسیلهٔ دیفئومورفیسم گرفته شده است. تصاویر I و J دو چهره را نمایش می‌دهند و کج و معوج‌سازی آنها از طریق نگاشت ϕ انجام می‌شود.

بازشناسی چهره [تشخیص سیما] یکی از موارد ساده‌ای است که در آن متغیرهای پیوند دینامیک مورد استفاده قرار می‌گیرد. برای تشخیص هر چهره می‌توان چهرهٔ «نمونهٔ» عامی ساخت و از طریق دیفئومورفیسم[1] [وابریختی] مناسبی موسوم به «تکنیک ماسک لاستیکی» [۱۷]، این نمونه را روی هر چهرهٔ متصوری باکج و معوج کردن جفت و جورکرد. تغییر روشنایی هم تغییرات زیادی در تصویر هر چهره‌ای ایجاد می‌کند. در نتیجه، متغیرهای تصادفی این مدل هم مختصات نقاط حاصل از اعمال تابع «کج و معوج‌سازی» بر نقاط مرجع در نمونه هستند و هم ضرایب سایه‌روشن ناشی از نورپردازی چهره. بنابراین، احتمال لگاریتمی، مجموع عبارتهایی است که نیکویی برازش کج و معوج‌سازیِ تصویر مشاهده شده را با مجموع نمونه‌های نشان‌دهندهٔ چهره‌ها تحت شرایط متفاوت نورپردازی، بیان می‌کنند [۷]. چند مثال در شکل ۵ آمده است.

آیا استنباط بر مبنای این مدلها کاری درست و عملی است؟ بیشتر اوقات، استنباط مورد نظر عبارت است از پیدا کردن براورد MAP برای متغیرهای تصادفی مربوطه ولی مشاهده نشدهٔ x_S، با توزیع احتمالِ مشروط به همهٔ مشاهدات $\widehat{x_T}$. در اینجا MAP به‌معنای «احتمال پسینی ماکسیمم»[2] است، یعنی محتملترین مجموعهٔ مقادیر این متغیرها و ما به‌دنبال پیدا کردن مقدار

$$\operatorname{argmax}_{x_S} P(x_S|\widehat{x_T})$$

هستیم. این یک مسألهٔ بهینه‌سازی است و معمولاً سه‌روش بنیادی برای حل یا حل تقریبی چنین مسائلی وجود دارد: نزول گرادیانی[3]، برنامه‌ریزی دینامیک، و زنجیرهای مارکوفی مونت کارلو. متأسفانه وقتی مدل پیچیده می‌شود هر سه روش با مشکل مواجه می‌شوند: روش نزول گرادیانی در نقاط بهینهٔ موضعی دچار مشکل می‌شود؛ برنامه‌ریزی دینامیک فقط زمانی کارایی دارد که متغیرها ترتیبی خطی و طبیعی داشته باشند که متغیرهای غیرمجاور را از یکدیگر جدا سازد؛ و زنجیره‌های مارکوفی مونت کارلو هم معمولاً خیلی کُند عمل می‌کنند. با وجود این، بارِ این رشته تا همین اواخر تنها بر دوش این سه روش بوده است. برای مثال، رشتهٔ بازشناسی گفتار کاملاً مدیون روشهای برنامه‌ریزی دینامیک بوده است و هرجا که این روشها عملی نیستند دچار ضعف است.

برای رام کردن روشهای تصادفی اخیراً چندین گروه تحقیقاتی به‌بررسی شیوهٔ جدیدی پرداخته‌اند. این شیوه به «فیلترکردن ذره‌ای» [صافی ذره‌ای] یا «نمونه‌برداری استوار بر عوامل» ([۵]، [۴]، [۱۰] و [۸]) شهرت یافته است، و در حقیقت نوعی روش مونت کارلو است که سر و کارش با نمونه‌ای نسبتاً کوچک چون $\{x_S^{(\alpha)}\}$ (مثلاً ۱۰۰ یا ۱۰۰۰ تا α) از توزیع است، نه مثل روش زنجیرهای مارکوفی مونت کارلوکه هر بار با یک نمونه سر و کار دارد. مقصود این است که به‌تقریبی ضعیف دست بیابیم:

$$p(\cdot|\widehat{x_T}) \sim \sum_{\alpha} w_\alpha \delta_{x_S^{(\alpha)}}$$

یعنی به‌ازای رده‌ای از متغیرهای تصادفی «شسته رفتهٔ» f در فضای احتمال ما:

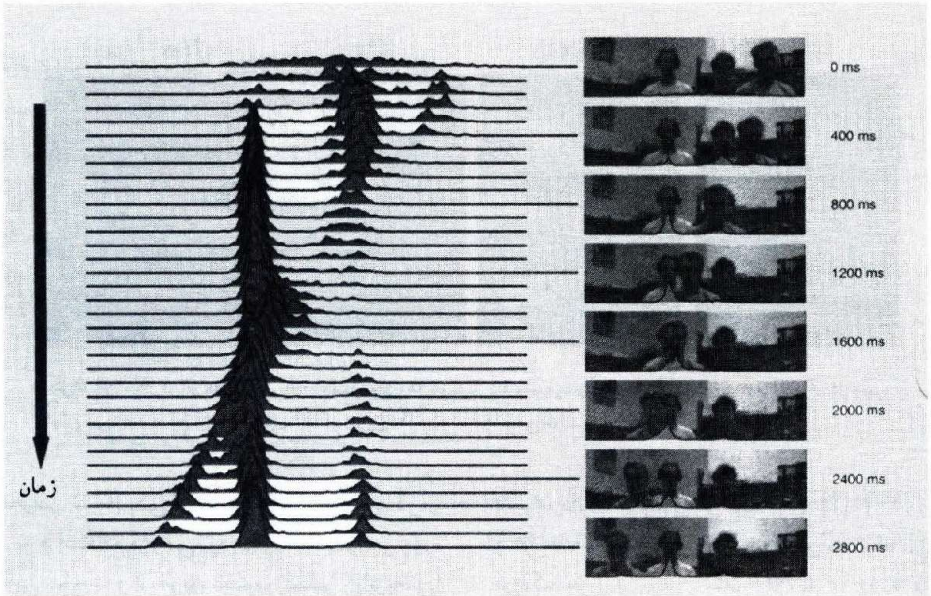
$$\operatorname{Exp}(f|\widehat{x_T}) \approx \sum_{\alpha} w_\alpha f(x_S^{(\alpha)})$$

امید آن است که حداقل برای متغیرهای تصادفی مهم، بسیاری از توزیعهای احتمالِ چند مدی را بتوان از طریق نمونه‌های وزن‌دار به‌این صورت تقریب زد. اضافه بر این، امید آن است که حفظ چنین نمونه‌ای اجازهٔ ادغام داده‌های جدید را در موقعیتی دیگر بدهد، موقعیتی که در آن امکانی که قبلاً کم‌احتمال‌تر بوده به محتملترین امکان تبدیل شود. در شکل ۶، مثالی از کارِ بلیک و ایزارد ارائه شده است که در آن ردیابیِ حرکت یک گروه چند نفره به‌طور موفقیت‌آمیزی انجام گرفته است. تکنیکهای کلاسیک متعارف مانند فیلتر کالمان که مبتنی بر مدلهای گاوسی هستند در این مورد معمولاً از عهده کار بر نمی‌آیند.

مطالبی که ارائه شد همه معطوف بود به‌نشان دادن حال و هوای تحقیقات در مورد کاربرد روشهای تصادفی در ایجاد مدل برای رفتار هوشمند. این تحقیقات دائماً در حال تحول هستند و اغلب شاهد بوده‌ایم که هواداران این یا آن مکتب مدلسازی تفکر اعلام کرده‌اند که کلید و راه‌حل کامل باز تولید رفتار هوشمند را یافته‌اند یا آنکه فقط به چند سال تحقیق بیشتر نیاز دارند! از آنجا که در گذشته هیچ‌کدام از این ادعاها نگرفته است، من هم نمی‌خواهم ادعای جدیدی بکنم جز آنکه فکر می‌کنم ایده‌هایی که در بالا مطرح کردم به‌نظر می‌رسد که در امتداد مسیر درست باشند.

نتیجه‌گیری کلی من این است که به‌اعتقاد من روشهای تصادفی در ابتدای هزارهٔ سوم کل ریاضیات محض و کاربردی را متحول خواهند ساخت. هم در زمینهٔ مدلسازی ریاضی و هم در زمینهٔ مدلسازی علمی، احتمال و آمار ابزارهای طبیعیِ کار از آب درخواهند آمد. جهان فکری در مجموع به‌منطق بیشتر به‌عنوان یک انتزاع زیبا و ظریف خواهد نگریست، درحالی‌که نگاهش به‌آمار به‌عنوان روش متعارفِ استدلال کردن و فکر کردن خواهد بود.

1. diffeomorphism
2. Maximum A Posteriori Probability
3. gradient descent



شکل ۶ این مثال برگرفته از اثر ایزارد و بلیک در بارهٔ ردیابی چهره‌های در حال حرکت در محیط شلوغ با استفاده از صافی ذره‌ای است. در سمت راست تصاویر قرار دارند؛ در سمت چپ، توزیعهای احتمال چندمدی هموارشده قرار دارند که احتمال شرطی هر چهره را در هر موضع (با در نظر گرفتن سلسله تصاویر قبلی و کنونی) براورد می‌کنند.

مراجع

1. Davis, P. and Hersh, R., 1980, *The Mathematical Experience*, Birkhauser-Boston.

2. E., Weinan, Khanin, K., Mazel, A., and Sinai, Ya, 1997, Probability distribution functions for the random forced Burger's equation, Phys. Rev. Letters, **78**, 1904-1907.

3. Freiling, C., 1986, Axioms of symmetry: throwing darts at the real line, *J. Symb. Logic,* **51**, 190-200.

4. Gordon, N., Salmond, D., and Smith, A., 1993, Novel Approach to nonlinear/non-Gaussian Bayesian State Estimation, *IEEE Proc. F,* **140**, pp. 107-113.

5. Grenander, U., Chow, Y. and Keenan, D., 1991, *HANDS, A Pattern Theoretic Study of Biological Shapes,* Springer.

6. Gross, C., 1995, Aristotle on the Brain, *The Neuroscientist,* **1**, 245-250.

7. Hallinan, P., Gordon, G., Yuille, A., Giblin, P., and Mumford, D., 1999, *Two and Three-dimensional Patterns of the Face,* AKPeters.

8. Isard, M. and Blake, A., 1996, Contour tracking by stochastic propagation of conditional density, *Proc. Eur. Conf. Comp. Vision,* pp. 343-356.

9. Jaynes, E. T., 1996-2000, *Probability Theory: The Logic of Science*, available at `http://bayes.wustl.edu/etj/prob.html`. To be published by Camb. Univ. Press.

10. Kanizawa, K., Koller, D., and Russell, S., 1995, Stochastic simulation algorithms for dynamic probabilistic networks, *Proc. Conf. Uncertainty in A. I.,* pp.346-351.

11. Laurtzen, S. and Spiegelhalter, D., 1988, Local computations with probabilities on graphical structures, *J. Royal Stat. Soc.,* **B50**, 157-224.

12. Mumford, A. A. and Young, M., 1923, *Biometrika,* **15**, p. 109-115.

13. Pearl, J., 1988, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufman.

14. Russell, B. and Whitehead, A. N., 1912, *Principia Mathematica, vol. 2*, Cambridge Univ. Press.

15. Shelah, S. and Woodin, W. H., 1990, Large cardinals imply that every reasonable definable set is Lebesgue measurable, *Israel J. Math.,* **70**, 381-394.

16. Spencer, J., 1994 (2nd edition), *Ten Lectures on the Probabilistic Method,* SIAM.

17. Widrow, B., 1973, The rubber mask technique, *Pattern Recognition,* **5**, 175-211.

* * * * * *

- David Mumford, "The dawning of the age of stochasticity", in V. Arnold, M. Atiyah, P. Lax, and B. Mazur (eds.), *Mathematics: Frontiers and Perspectives 2000*, Amer. math. Soc. (2000) 197-218.

* دیوید مامفرد، دانشگاه براون، آمریکا