ارتباط بدون خطا: کدهای تصحیح کننده

نویسنده: ژیل لاشو [۱]

مترجم: فائزه توتونیان

ویراستاران: فرج‌الله محمودی، ارسلان شادمان

برای کشف و تصحیح خطاهای غیرقابل اجتناب در مبادلهٔ اطلاعاتی که به صورت عددی درآمده‌اند، متخصص‌های کدگذاری متوسل به روش‌های مجردی می‌شوند که از جبر و هندسه سرچشمه می‌گیرند.

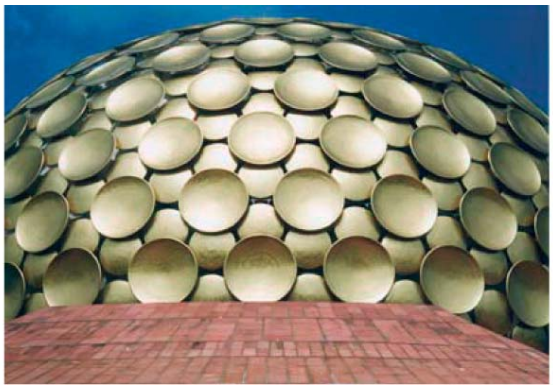
ما کاملاً در عصر عددی هستیم. این جمله گویای چیست؟ به‌طور خیلی ساده یعنی قسمت وسیعی از اطلاعات که بر روی سیارهٔ زمین مبادله می‌شوند، در قالب اعداد نشان داده شده‌اند. پیام‌های الکترونیکی، تلفن موبایل، معامله‌های بانکی، هدایت از راه دوره ماهواره‌ها، انتقال تصاویر از راه دور، دیسک‌های CD یا DVD و غیره: در تمام این مثال‌ها، اطلاعات به‌صورت دنباله‌ای از اعداد، که به‌طور فیزیکی متناظر با علائم الکتریکی یا علائم دیگرند، ترجمه می‌شوند ـ و یا گفته می‌شود کدگذاری شده‌اند (با رمزگذاری اشتباه نشود). به‌صورت دقیق‌تر، اطلاعات در مجموع به شکل دنباله‌ای از ارقام دودویی یعنی اعداد ۰ یا ۱، که بیت‌ها نیز نامیده می‌شوند، کدگذاری شده‌اند. مثلاً در کد ASCII (کد استاندارد آمریکایی برای مبادلهٔ اطلاعات) [۲] که توسط ریزکامپیوترها استفاده می‌شود،

---

[۱] Lachaud, Gilles: *Communiquer sans erreurs: les codes correcteurs*, in: L'explosion des mathématiques, SMF et SMAI, Paris, 2002, p. 84-87

[۲] American Standard Code for Information Interchange

حرف بزرگ A توسط هشتایی (دنباله‌ای از هشت بیت) ۰۱۰۰۰۰۰۱ و حرف بزرگ B توسط ۰۱۰۰۰۰۱۰، و غیره کدگذاری می‌شوند.



ماتریماندر در اوروویل (تامیل نادو، هندوستان)[1]، ژئودِ کروی که توسط روژه آنژه[2] معمار فرانسوی ساخته شده است. در مفهوم کدهای موثرِ تصحیح کننده با مسائلی برخورد می‌کنیم که به سؤال‌های مشکل هندسهٔ محض مربوط می‌شوند، نظیر دوباره پوشانیدن یک کره توسط بیشترین تعداد ممکن قرص‌های یک اندازه بدون این که آنها روی یکدیگر سوار شوند.

یک مسألهٔ بزرگ در مخابرهٔ اطلاعات، خطاها می‌باشند. کافی است که خراش کوچکی روی یک دیسک، یک اختلال در دستگاه، یا هر نوع پدیدهٔ پارازیتی، پیام مخابره شده را با خطا همراه سازد، یعنی «۰»ها به‌طور نابهنگام به «۱» یا بالعکس تغییر کنند. بنابراین یکی از راه‌های بیشمار رهایی از این‌گونه اشکال، امکانِ کشف و حتی تصحیح چنین خطاهایی است.

طولانی کردن کلماتِ پیام به طریقی که بعد از کوتاه کردن در طرف مقابل بتوان آنها را بازشناخت

نقش نخستین کدهای تصحیح کنندهٔ خطاها در همان دوران اول کامپیوترها مطرح شدند که از آن زمان بیش از پنجاه سال می‌گذرد. این کدها چگونه عمل می‌کنند؟ مبنا و اساس آنها به‌صورت زیر است: «کلمات» عددی رسانندهٔ پیام را طولانی می‌کنیم، به طریقی که قسمتی از بیت‌ها به‌عنوان بیت‌های کنترل به‌کار می‌روند. برای مثال در کد ASCII که قبلاً به آن اشاره شد، یکی از هشت بیت یک بیت کنترل است: برای بیت

[1] Le Matrimandir à Auroville (Tamil Nadu, Inde)

[2] Roger Anger

کنترل مقدار۰ در نظر گرفته می شود اگر تعداد «۱»ها در هفت بیت دیگر زوج باشد، و گر نه ۱ را اختیار خواهد کرد. اگر در مقدار یکی از هفت بیت دیگر یک تغییر ناگهانی ایجاد شود، دیگر ارزش بیت کنترل با آن متناظر نخواهد بود و در نتیجه یک خطا کشف می‌شود. همین ایده را در قلمرو اعدادی که در زندگی روزانه با آنها برخورد داریم، مشاهده می‌کنیم. مثلاً در صورت حساب‌های بانکی، یک حرف کلیدی به یک شماره حساب اضافه می‌شود، تا بتوان خطای یک انتقال را کشف کرد. همچنین برای جلوگیری از تقلب، شمارهٔ اسکناس‌های بانکی بر حسب یورو کدگذاری می‌شوند. به بیان دیگر فلسفهٔ کدهای تصحیح کننده ایجاد پیام‌های اضافی است: هر کلمه از پیام به طریقی طولانی می‌شود که حاوی اطلاعاتی در مورد خود پیام باشد.

یک مثال ساده و روشنگر ولی نه چندان واقعیت‌گرا، از کدهای تصحیح کنندهٔ خطا، کد تکرار سه تایی است: هر بیت پیام سه بار کدگذاری می‌شود، یعنی ۰ به‌صورت ۰۰۰و۱ به‌صورت ۱۱۱ در می آیند. این کد اجازه می‌دهد که یک خطای احتمالی روی یک سه تایی را کشف و تصحیح کرد. در واقع اگر دنبالهٔ ۱۰۱ را بر فرض دریافت کنیم، بلافاصله از آن نتیجه می‌گیریم که دنبالهٔ صحیح ۱۱۱ بوده است (با فرض آن که فقط یک بیت از سه تایی دریافتی اشتباهی باشد)، پس در اطلاعات اولیه، بیت مورد نظر ۱ بوده است. کد تکراری سه تایی واقعیت‌گرا نیست، زیرا هزینه‌بردار است برای هر بیت اطلاعات باید سه عدد فرستاده شود، گفته می‌شود که نرخ بازدهی $\frac{1}{3}$ است. این نرخ تأثیرات مستقیمی بر روی زمان لازم جهت انتقال پیام‌ها و روی هزینهٔ ارتباط‌ها دارد.

یک کد تصحیح کنندهٔ خوب، باید دارای کیفیت‌های دیگر به‌ویژه یک نرخ بازدهی بالا باشد. بعلاوه باید قابلیت خوبی در کشف و تصحیح خطاها داشته باشد و رویّهٔ کدگشایی باید به اندازهٔ کافی ساده و سریع باشد. همهٔ مسألهٔ نظریهٔ کدهای تصحیح کنندهٔ خطاها این است که: با طولانی کردن ممکن پیام‌ها کدهایی بسازیم که تا حد ممکن خطاها را کشف و تصحیح کنند، و کدگشایی آن‌ها آسان باشد.

جبر میدان‌های متناهی به‌طور طبیعی در کدها کاربرد دارد، زیرا کدها از الفبای متناهی استفاده می‌کنند.

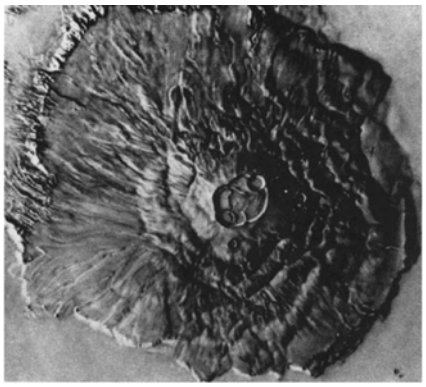
مدت‌های مدیدی است که ریاضیدان‌ها در این زمینه دخالت دارند. در ۱۹۴۸ ریاضیدان آمریکایی کلود شانن [۱]، یکی از پدران نظریهٔ اطلاعات به نتایج نظری کلی

---

[۱] Claude Shannon

دست یافت که مؤید وجود کد بهینه به یک معنی فنیِ دقیق بود. هر چند نظریهٔ شانن وجود کدهای تصحیح کنندهٔ بسیار خوبی را اثبات کرد، اما روشی عملی برای ساختن آنها ارائه ننمود. از طرف دیگر کدهای تصحیح کننده‌ای، نظیر کدهای همینگ [۱] با قابلیت متوسط در اختیار بود، که به نام مخترعشان ریاضیدان آمریکایی ریچارد همینگ (۱۹۹۸ ـ ۱۹۱۵) نام‌گذاری شده‌اند و در سال‌های ۱۹۵۰ اختراع گردیده‌اند. (در این کدها، که بسیار هم مورد استفاده هستند بیت‌های کنترل توسط معادلات خطی ساده از روی بیت‌های اطلاعاتی تعیین می‌شوند).

بنابراین متخصصین به طریقی اصولی دست به کار شدند تا کدهای تصحیح کننده و ویژگی‌های آنها را مورد مطالعه قرار دهند، با این هدف که به‌طور واقعی کدهایی با همان قابلیت یا تقریبی از آنچه که نتایج نظری شانن پیش‌بینی کرده است، به‌دست آورند. برای انجام این کار آنان عمیقاً از جبر استفاده کرده‌اند. اگر کدگذاری اطلاعات به‌طور مستقیم با «الفبای» دوتایی ۰ و ۱ انجام می‌شود، جبر مورد استفادهٔ آن جبر زوج و فرد است که قبلاً افلاطون هم می‌شناخته است (زوج = زوج + زوج، فرد = فرد + زوج، زوج = زوج × زوج، فرد = فرد × فرد، و غیره).



کوه اولمپوس بر سطح سیارهٔ مریخ بزرگترین آتش‌فشان سیستم خورشیدی (منظومهٔ شمسی) است: قطر آن تقریباً ۶۰۰ و ارتفاع آن حدود ۲۷ کیلومتر است! این تصویر در سال‌های ۷۲ ـ ۱۹۷۱ به‌وسیلهٔ سفینهٔ فضایی مارینر ۹ گرفته شده است. اطلاعات آن وسیلهٔ کد تصحیح کننده‌ای با قابلیت تصحیح ۷ بیت اشتباه روی ۳۲ بیت به زمین ارسال شده است. در هر گروه از ۳۲ بیت، ۲۶ بیت آن مربوط به کنترل و ۶ بیت دیگر اطلاعات دقیق را تشکیل می‌داده‌اند. در حال حاضر کدهای تصحیح کنندهٔ کارآمدتری نیز مورد استفاده قرار می‌گیرند. (کلیشه ناسا/ جی ـ پی ـ ال)

---

[۱] Hamming

از این رو جالبتر است که آن دسته از کدگذاری‌هایی را در نظر بگیریم که الفبایشان بیش از دو رقم داشته باشد و فقط در پایان فرایند به دنباله‌های دوتایی ۰ و ۱ توجه شود. چون الفبا شامل تعداد محدودی نشانه است و انتظار این است که محاسبات روی این نشانه‌ها انجام گیرد، جبر مورد استفاده، موضوع نظریهٔ میدان‌های متناهی می‌باشد که توسط ریاضیدان جوان فرانسوی اواریست گالوا [۱] در ابتدای قرن ۱۹، به هنگام مطالعهٔ حل‌پذیری معادلات جبری، اختراع شد (یک میدان متناهی مجموعهٔ عناصری با تعداد متناهی است که می‌توانند به طریقی مشابه با اعداد معمولی، جمع، ضرب و تقسیم شوند، و نتیجهٔ اعمال در داخل این مجموعه باقی می‌ماند. مجموعهٔ ساخته شده توسط ۰ و ۱، با قواعد حسابی زوج و فرد، یک میدان متناهی با دو عنصر است، که ساده‌ترین میدان متناهی است.)

به این ترتیب به کمک جبر مجرد و توسعه یافته، در ارتباط با نظریهٔ میدان‌های متناهی، کدهای تصحیح کنندهٔ خطا به گونه‌ای خیلی مؤثر ساخته شدند که با هر نوع انتقال اطلاعات تطبیق می‌یابند. از بین مثال‌های متعدد به دو مورد اشاره می‌کنیم: کد مورد استفاده برای ذخیره کردن اطلاعات دیسک‌های صوتی ـ عددی (این کد امکان تصحیح ۴۰۰۰ بیت متوالی، اشتباه ناشی از خراش بیش از ۲ میلیمتر بر سطح یک دیسک را فراهم می‌آورد)، و کدی که کاوشگر فضایی مارینر [۲] برای ارسال تصاویری از سیارهٔ مریخ از آن استفاده کرده است.

هرچند این تمبر فرانسوی چاپ شده در سال ۱۹۸۴ اواریست گالوا را هندسه‌دان نامیده ولی او یک جبردان بوده است. او در نظریهٔ گروه‌ها، و همچنین نظریهٔ میدان‌های متناهی که به‌ویژه توسط متخصص‌های کدهای تصحیح کنندهٔ خطاها استفاده می‌شوند، پیشگام بوده است. گالوا به دوئل دعوت شد، و در سنی که به سختی به ۲۱ سال می‌رسید، درگذشت.

---

[۱] Evariste Galois

[۲] Mariner

خانوادهٔ جدیدی از کدها که از هندسهٔ جبری منحنی‌ها استفاده می‌کنند

جبر مجرد تنها وسیله‌ای نیست که در اختیار متخصصین کدهای تصحیح کننده است. بلکه هندسه و به‌ویژه هندسهٔ جبری نیز ابزاری در دست آنهاست. هندسهٔ جبری که بخش وسیعی از ریاضیات کنونی است، نخست به بررسی اشیائی هندسی می‌پردازد از قبیل خم‌ها، رویه‌ها و غیره که توسط معادلات جبری تعریف می‌شوند. هر دانش‌آموز دبیرستانی می‌داند که مثلاً سهمی را می‌توان توسط یک معادلهٔ جبری از نوع $y = ax^2 + bx + c$ نمایش داد که در آن x و y مختصات نقاط سهمی هستند. به‌همین ترتیب می‌توان منحنی‌های تعریف شده روی میدانهای متناهی را مطالعه کرد، بدین معنی که در معادلات جبری نمایشگر آنها کمیتهایی نظیر x و y دیگر اعداد دلخواه نیستند بلکه منحصر به عناصر یک میدان متناهی خاص هستند. حدود ۲۰ سال است که با استفاده از چنین منحنی‌ها و جبر وابسته به مختصات نقاط آنها (که از نظر تعداد متناهی هستند)، خانوادهٔ جدیدی از کدهای تصحیح کننده و کدهای هندسی، ساخته شده است. اخیراً این کدها اجازه داده‌اند که نتایج جدیدی مربوط به کدهای دوتایی به‌دست آید و کدهایی با قابلیتی حتی بیش از کدهایی که توسط کارهای شانن پیش‌بینی شده بود ساخته شوند. در مقابل، تحلیل کدهای هندسی، ریاضیدان‌ها را به بررسی دقیق‌تر در مورد تعداد نقاط یک منحنی جبری که بر روی یک میدان تعریف می‌شوند، هدایت کرده است. در اینجا مثال زیبایی از تأثیر متقابل مثبتی در اختیار داریم که یک قلمرو کاربردی می‌تواند بر اصول نظری که از آنها استفاده می‌کند داشته باشد.

ژیل لاشو
انستیتوی ریاضیات لومینه
مرکز ملی تحقیقات علمی CNRS
مارسی

چند مرجع

- P. Arnoux, “Codage et mathématiques”, *La science au présent* (édition Encyclopædia Universalis, 1992).
- P. Arnoux, “Minitel, codage et corps finis”, *Pour la Science* (mars 1988).

- G. Lachaud et S. Vladut, "Les codes correcteurs d'erreurs", *La Recherche* (juillet-août 1995).
- O. Papini, Disque compact: "la théorie, c'est pratique!" dans "Secrets de nombres", Hors-série n° 6 de la revue *Tangente* (1998).
- O. Papini et J. Wolfmann, *Algèbre discrète et codes correcteurs* (Springer-Verlag, 1995).
- J. Vélu, *Méthodes mathématiques pour l'informatique* (Dunod, 1995).
- M. Demazure, *Cours d'algèbre - primalité, divisibilité, codes* (Cassini, 1997).

Gilles Lachaud
Institut de mathématiques de Luminy,
CNRS, Marseille