Vilius Stakėnas

Kodavimo teorija

Paskaitų kursas

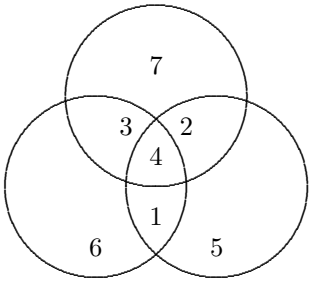
2002

●●● ◇ ●●●

Įvadas

Informacija perduodama kanalais, kurie kartais iškraipo informaciją. Tarsime, kad tie iškraipymai yra atsitiktiniai, t. y. nėra nei sistemingi, nei sąmoningi. Koduoti informaciją reiškia taip ją paruošti prieš perduodant į kanalą, kad būtų galima ištaisyti įvykusius iškraipymus, jeigu jų nėra itin daug. Kodavimo teorija pasiūlo konstruktyvius tokio paruošimo (kodavimo) ir iškraipymų ištaisymo (dekodavimo) algoritmus. Šie algoritmai taikomi daugelyje moderniosios komunikacijos sričių.[1]

Klaidas taisantį kodą apie 1948 metus sukonstravo R. W.Hamming. Kiekvieną ketvertą perduodamų bitų jis papildė trimis papildomais. Paaiškinsime jo įdėją naudodami brėžinį.



Srityse 1,2,3,4 užrašykime keturis bitus (simbolius 0 arba 1), kuriuos reikia perduoti. Likusiose srityse užrašykime simbolius taip, kad kiekviename iš trijų skritulių užrašytus simbolius sumuodami gautume lyginį skaičių. Dabar „pasiųskime į kanalą“ šį septynių bitų rinkinį. Kitame kanalo gale gavėją pasieks galbūt kitas rinkinys, kai kurie simboliai gali būti kitokie. Gavėjas gali patikrinti, ar sumos pagal visus tris skritulius yra lyginės. Jeigu nors viena nelyginė, tai įvyko iškraipymų. Nesunku įsitikinti, kad jeigu iškreiptas tik vienas simbolis, tai gavėjas gali atkurti jį nesuklysdamas. Jeigu iškraipymų daugiau, atkurdami simbolius galime ir suklysti.

Palyginkime tikimybes, kad keturi simboliai bus perduoti teisingai, kai informacijos nekoduojame ir kai ją koduojame Hammingo kodu. Tarkime, kanalas kiekvieną simbolį iškreipia su tikimybe $p(0 < p < 1)$, be to – visi simboliai iškraipomi nepriklausomai. Tada tikimybės, kad keturi simboliai bus perduoti teisingai, kai informacija nekoduojama ir koduojama Hammingo kodu lygios:

$$q^4, \quad \text{ir} \quad q^7 + 7q^6p, \quad q = 1 - p.$$

Palyginę tikimybes įsitikintume, kad didesnė tikimybė teisingai perduoti keturis simbolius, kai informaciją koduojame.

Panagrinėkime dar vieną pavyzdį, kuriame skaičiuosime klaidingo atkūrimo tikimybes.

Pavyzdys. Tarkime mėtoma moneta, o rezultatai ryšio kanalu perduodami suinteresuotam gavėjui. Galima perduoti du simbolius: 0 ir 1, kanalas su tikimybe p iškreipia simbolį, t. y. pakeičia kitu. Kiti apribojimai: moneta mėtoma T sekundžių, atliekama m metimų per sekundę; kanalu galima naudotis irgi T sekundžių (nebūtinai tuo pačiu metu kai mėtoma moneta), kanalu galima perduoti $2m$ simbolių per sekundę. Akivaizdu, kad rezultatus koduojant taip HERBAS $\to$ 0 ir SKAIČIUS $\to$ 1 ir perduodant tiesiog į kanalą tikimybė, kad du iš eilės gauti rezultatai bus perduoti teisingai lygi

$$q^2, \quad q = 1 - p.$$

Nejaugi negalima padidinti šios tikimybės? Jeigu kiekvieną rezultatą siųsdami kartotume du kartus, jokios naudos nebūtų, kadangi vieną poros simbolį priėmus neteisingai, nežinotume, ką simbolių porą turėjo reikšti – herbą ara skaičių? Šiek tiek geriau būtų, jei kiekvieną rezultatą kartotume tris kartus, o kiekvieną gautą

[1] Apie kodų naudojimą kompaktinėse plokštelėse galima sužinoti, pavyzdžiui, iš straipsnio: J. H. van Lint, Mathematics and the Compact Discs. Johanes Bernoulli Lecture 1998, *Nieuw Archief voor Wiskunde,* 16(3), 183–190 (1998).

simbolių trejetą dekoduotume tuo simboliu, kuris trejete pasitaiko daugiau kartų. Tačiau tada turėtume pasiųsti iš viso $3m$ simbolių. Tačiau mus riboja techninės sąlygos: galime daugiausi

Jeigu kiekvieną perduodamą simbolį dubliuotume, t.y. pakartotume du kartus – neturėtume jokios naudos. Jeigu bent vienas poros simbolis būtų iškreiptas, nežinotume, ką ta pora reiškė – nulį ar vienetą. Jeigu kiekvieną simbolį pakartotume tris kartus, o kiekvieną priimtą simbolių trejetuką dekoduotume tuo simboliu, kuris trejete yra dažnesnis, tikimybę teisingai perduoti simbolį padidintume, tačiau tada per sekundę reiktų perduoti $3m$ simbolių, o to neleidžia mūsų „techninės“ sąlygos. Vis dėlto išeitis yra.

Koduokime ne pavienius rezultatus, bet jų poras:

$$\mathrm{HH} \to 0000,\ \mathrm{HS} \to 0111,\ \mathrm{SH} \to 1001,\ \mathrm{SS} \to 1110.$$

Tarkime, kad simbolių ketvertuke vienas iš pirmųjų trijų simbolių iškraipytas, tačiau ketvirtasis perduotas teisingai. Nesunku suvokti, kad tokiu atveju galima vienareikšmiškai nustatyti, kuris simbolis iškreiptas. Taigi turėdami galvoje, kad atliekamas toks ištaisymas randame teisingo dviejų rezultatų perdavimo tikimybę:

$$q^4 + 3q^3p.$$

Pakanka panagrinėti skaitinius pavyzdžius, kad įsitikintume, kad teisingo perdavimo tikimybę padidinome. Iš tiesų poros rezultatų teisingo perdavimo tikimybė yra didesnė visais atvejais, kai $q > p$. Ši sąlyga yra visiškai natūrali: nėra prasmės naudotis tokiu kanalu, kuris su didesne tikimybe simbolį iškreipia negu perduoda teisingai.

Ar negalima sugalvoti dar „gudresnių“ kodų, t.y. su dar didesne teisingo perdavimo tikimybe? Kita vertus, kiek naudojant tokius kodus padidėja siuntimo sąnaudos? Ką tik išnagrinėtame pavyzdyje tikrąją informaciją reiškia du pirmieji bitai, o kiti – papildyti, kad būtų galima teisingai dekoduoti. Atsakymą į klausimą, kokiu mastu įmanoma suderinti teisingo dekodavimo ir sąnaudų aspektus duoda fundamentali Claude Shannono teorema.

Tarkime, informacija yra užrašyta abėcėlės $\mathcal{B} = \{0; 1\}$ abėcėlės žodžiais, o ją reikia perduoti kanalu, kuris su tikimybe p $(0 < p < 1)$ kiekvieną simbolį iškreipia. Simboliai iškraipomi nepriklausomai vienas nuo kito.

Tarkime, į kanalą bus siunčiami n ilgio žodžiai iš aibės $\mathcal{C} = \{x_1, \ldots, x_M\}, x_i \in \mathcal{B}^n$. Sakysime, kad naudojame kodą $\mathcal{C}$. Dekoduosime iš kanalo gautąjį žodį labai paprastai: lyginsime jį su kitais kodo žodžiais ir dekoduosime tuo kodo žodžiu, kuris mažiausiai skiriasi nuo gautojo žodžio (t.y. turi mažiausiai skirtingų komponenčių). Ši dekodavimo taisyklė vadinama minimalaus atstumo taisykle.

Tegu P_i yra tikimybė, kad pasiųstas į kanalą žodis x_i bus klaidingai dekoduotas. Apibrėžkime vidutinę klaidingo dekodavimo tikimybę:

$$P_C = \sum_{i=1}^{n} P_i.$$

Teorema.

(Claude Shannon, 1948) Tegu skaičius R tenkina nelygybę

$$0 < R < 1 - p\log_2\frac{1}{p} - q\log_2\frac{1}{q}, \quad q = 1 - p.$$

Tada egzistuoja kodai $\mathcal{C}_n = \{x_1, \ldots, x_{M_n}\}, x_i \in \{0; 1\}, M_n = 2^{[nR]}$, kad $P_{C_n} \to 0$, kai $n \to +\infty$.

Ką gi ši teorema teigia apie kodavimo sąnaudas? Tarkime, mums reikia siųsti informaciją, užrašytą aibės $\mathcal{B}^k$ žodžiais. Juos mes turime koduoti, naudodami tam tikrą funkciją (kodavimo taisyklę arba algoritmą) $\mathcal{B}^k \to \mathcal{C}_n = \{x_1, \ldots, x_{M_n}\}$. Kad žodžių užtektų turi būti $|\mathcal{B}^k| \le |\mathcal{C}_n|$, t.y. $2^k \le M_n, k \le [nR]$. Taigi gauname, kad k bitų informacija yra siunčiama netrumpesniu kaip $\frac{k}{R}$ bitų žodžiu. Informacijos perdavimas kanalu ją kodavus sulėtėja dydžiu

$$\frac{k}{n} \le R.$$

• • • ◇ • • •

Taigi Shannono teorema tarsi nurodo, kokios yra minimalios sąnaudos, jeigu norime mažai iškraipytą informaciją perduoti kanalu, kuriame simboliai gali būti iškraipomi. Tačiau kaip konstruoti geriausius kodus, t.y. garantuojančius mažas klaidingo perdavimo tikimybes ir minimalias sąnaudas, teorema nepasako. Algebrinė kodavimo teoriją ir sudaro įvairūs atsakymai į tokius klausimus.

1. Klaidas randantys ir taisantys kodai

1.1. Bendrosios sąvokos

Tarkime $\mathcal{A}$ yra abėcėlė, $|\mathcal{A}| = q$. Žymėsime: $\mathcal{A} = \mathcal{A}_q, \mathcal{A}_q^n = \mathcal{A}_q \times \mathcal{A}_q \times \ldots \times \mathcal{A}_q$

1.1.1 apibrėžimas.

(n, N) kodu iš abėcėlės $\mathcal{A}_q$ žodžių vadinamas bet koks poaibis $\mathbf{C} \subset \mathcal{A}_q^n$, čia $|\mathbf{C}| = N$.

1.1.2 apibrėžimas.

Tegu $\mathbf{x} = x_1 \ldots x_n$, $\mathbf{y} = y_1 \ldots y_n$ yra du aibės $\mathcal{A}_q^n$ žodžiai. Hamingo atstumu tarp $\mathbf{x}$, $\mathbf{y}$ vadinsime dydį

$$h(\mathbf{x}, \mathbf{y}) = \sum_{\substack{i=1,\ldots,n \\ x_i \neq y_i}} 1.$$

1.1.1 teorema.

Hamingo atstumas aibėje $\mathcal{A}_q^n$ turi šias savybes:
1) $h(\mathbf{x}, \mathbf{x}) = 0, \mathbf{x} \in \mathcal{A}_q^n$;
2) $h(\mathbf{x}, \mathbf{y}) = h(\mathbf{y}, \mathbf{x}), \mathbf{x}, \mathbf{y} \in \mathcal{A}_q^n$;
3) $h(\mathbf{x}, \mathbf{y}) \leq h(\mathbf{x}, \mathbf{z}) + h(\mathbf{y}, \mathbf{z}), \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{A}_q^n$. Trumpai tariant Hamingo atstumas yra metrika.

1.1.3 apibrėžimas.

Dekodavimo taisyklę $f : \mathcal{A}_q^n \to \mathbf{C}$ vadinsime minimalaus atstumo taisykle, jei su kiekvienu $\mathbf{d} \in \mathcal{A}_q^n$

$$h(\mathbf{d}, f(\mathbf{d})) = \min_{\mathbf{c} \in \mathbf{C}} h(\mathbf{d}, \mathbf{c}). \tag{1.1.1}$$

Tad gautas žodis dekoduojamas tuo kodo elementu, kuris Hamingo atstumo prasme yra arčiausiai – labai paprasta idėja! Jeigu kuriam nors $\mathbf{d}$ minimumas (1.1.1) lygybėje pasiekiamas, kai $\mathbf{c} = \mathbf{c}_1$ arba $\mathbf{c} = \mathbf{c}_2, \mathbf{c}_1 \neq \mathbf{c}_2$, tai minimalaus atstumo taisyklė apibrėžiama nevienareikšmiškai. Jei minimumas pasiekiamas su dviem ar daugiau skirtingų žodžių, tokį atvejį vadinsime **ryšiu**. Ryšio atveju galime tiesiog fiksuoti klaidą (*soft decision*) arba dekoduoti vienu iš galimų būdų (*hard decision*).

Apibrėšime kodo žodžių „išsibarstymo“ charakteristiką.

1.1.4 apibrėžimas.

Kodo $\mathbf{C}$ minimaliu atstumu vadinsime dydį

$$d(\mathbf{C}) = \min_{\substack{\mathbf{c}, \mathbf{d} \in \mathbf{C} \\ \mathbf{c} \neq \mathbf{d}}} h(\mathbf{d}, \mathbf{c}).$$

Jei (n, N) kodo $\mathbf{C}$ minimalus atstumas yra d, tai kodą vadinsime (n, N, d) kodu.

• • • ◇ • • •

1.1.5 apibrėžimas.

Kodą $\mathbf{C}$ *vadinsime* t *klaidų randančiu kodu, jei bet kuriame kodo žodyje įvykus* $m, m \leq t$, *iškraipymų, gautas rezultatas* $\mathbf{d}$ *jau nebėra kodo žodis, t. y.* $\mathbf{d} \notin \mathbf{C}$.

1.1.6 apibrėžimas.

t *klaidų randantį kodą vadinsime tiksliai* t *klaidų randančiu, jei jis nėra* $t+1$ *klaidų randantis kodas.*

1.1.2 teorema.

(n, N, d) *kodas* $\mathbf{C}$ *yra tiksliai* t *klaidų randantis kodas tada ir tik tada, kai* $d = t + 1$.

Analogiškai apibrėšime klaidas taisančius, arba autokorekcinius, kodus. Patys kodai, aišku, klaidų netaiso. Klaidas taisome mes, naudodamiesi minimalaus atstumo dekodavimo taisykle.

1.1.7 apibrėžimas.

Kodą $\mathbf{C}$ *vadinsime* t *klaidų taisančiu kodu, jei sinčiamame žodyje įvykus* $m, m \leq t$, *iškraipymų ir dekoduojant pagal minimalaus atstumo taisyklę, dekoduojama bus visada teisingai.*

Akivaizdi ir tiksliai t klaidų taisančio kodo sąvoka.

1.1.3 teorema.

Kodas $\mathbf{C}$ *yra tiksliai* t *klaidų taisantis kodas tada ir tik tada, kai* $d(\mathbf{C}) = 2t + 1$ *arba* $d(\mathbf{C}) = 2t + 2$.

Išvada. *Bet koks* (n, N, d) *kodas taiso lygiai* $[(d-1)/2]$ *klaidų.*

1.1.8 apibrėžimas.

Abėcėlės $\mathcal{A}_q$ (n, N) *kodo* $\mathbf{C}$ *koeficientu vadinsime dydį*

$$R(\mathbf{C}) = \frac{\log_q N}{n}.$$

Kodo koeficientas tam tikra prasme apibūdina informacijos kodavimo sąnaudas (arba informacijos greičio sumažėjimą, kai informacija koduojama šiuo kodu). Iš tikrųjų, jeigu kodas turi N skirtingų žodžių, o pradinis informacijos srautas sudarytas iš q elementų turinčios abėcėlės simbolių, tai galime šį srautą skaidyti blokais po k simbolių ir šiuos blokus koduoti kodo žodžiais. Taigi k simboliai perduodami pasiunčiant kanalu iš viso n simbolių, t.y. perdavimo greitis sumažėja koeficientu $\frac{k}{n}$. Kuo bloko ilgis k didesnis, tuo didesnis perdavimo

greitis. Kokį didžiausią k galime pasirinkti? Reikia pasirūpinti, kad kodo žodžių užtektų visiems skirtingiems blokams, taigi turi būti $q^k \le N$, t.y. $k \le \log_q N$. Matome, kad koeficientas $R(\mathbf{C}) = \frac{\log_q N}{n}$ apibūdina maksimalų informacijos perdavimo greitį (arba minimalias sąnaudas), kurį galima pasiekti naudojantis kodu $\mathbf{C}$.

Faktai. Erdvėlaivis *"Mariner 9"* 1979 m. perdavinėjo nespalvotas Marso fotografijas. Kiekviena nuotrauka buvo skaidoma į 600×600 mažų kvadratėlių; kiekvienam jų buvo priskirtas vienas iš 64 pilkumo laipsnių. Šiuos laipsnius interpretuojant kaip simbolius, gaunama 64 simbolių šaltinio abėcėlė. Gautas informacijos srautas buvo koduojamas dvinariu $(32, 64, 16)$ kodu (Rydo–Mulerio kodu) ir perduodamas į Žemę. Kodas taisė lygiai 7 klaidas, kodo koeficientas lygus 3/16.

Erdvėlaivis *"Voyager"* 1979–1981 m. siuntė į Žemę spalvotas Jupiterio ir Saturno nuotraukas. Nuotraukai koduoti buvo naudojama 4096 simbolių (tiek spalvų intensyvumo lygių buvo fiksuojama) abėcėlė. Prieš perduodant simbolių srautas buvo koduojamas dvinariu $(24, 4096, 8)$ kodu (Golėjaus kodu). Kodas taiso lygiai tris klaidas, kodo koeficientas lygus 1/2.

1.2. Tobulieji ir ekvivalentieji kodai

Naudodamiesi Hamingo atstumu, aibėje $\mathcal{A}_q^n$ apibrėšime spindulio $r \ge 1$ rutulius: jei $\mathbf{x} \in \mathcal{A}_q^n$, tai

$$B_q(\mathbf{x}, r) = \{\mathbf{y} \in \mathcal{A}_q^n : h(\mathbf{x}, \mathbf{y}) \le r\}.$$

Nesunku įsitikinti, kad elementų skaičius rutulyje $B_q(\mathbf{x}, r)$ nepriklauso nuo jo centro $\mathbf{x}$, tad žymėsime

$$V_q(n, r) = |B_q(\mathbf{x}, r)|.$$

Kartais dydį $V_q(n, r)$ vadinsime rutulio tūriu.

1.2.1 teorema.

Teisinga lygybė

$$V_q(n, r) = \sum_{0 \le k \le r} \binom{n}{k} (q-1)^k.$$

Įrodymas. Rutuliui su centru $\mathbf{x}$ ir spinduliu r priklauso tie žodžiai, kurie skiriasi nuo $\mathbf{x}$ nedaugiau kaip r komponentėmis. Visus žodžius, kurie skiriasi nuo $\mathbf{x}$ lygiai k komponentėmis galime gauti taip: parenkame k numerių ir pakeičiame žodžio $\mathbf{x}$ komponentes su šiais numeriais kitomis. Tai galime padaryti

$$\binom{n}{k} (q-1)^k$$

skirtingais būdais.

1.2.1 apibrėžimas.

Tegu $\mathbf{C}$ yra koks nors (n, N) kodas. Didžiausią sveikąjį skaičių t, kuriam

$$B_q(\mathbf{c}_1, t) \cap B_q(\mathbf{c}_2, t) = \emptyset, \quad \textit{jei} \quad \mathbf{c}_1, \mathbf{c}_2 \in \mathbf{C}, \quad \mathbf{c}_1 \ne \mathbf{c}_2,$$

vadinsime kodo $\mathbf{C}$ pakavimo spinduliu. Pakavimo spindulį žymėsime $r_p = r_p(\mathbf{C})$.

1.2.2 apibrėžimas.

Mažiausią sveikąjį skaičių s, tenkinantį sąlygą

$$\mathcal{A}_q^n \subset \bigcup_{\mathbf{c} \in \mathbf{C}} B_q(\mathbf{c}, s),$$

vadinsime kodo dengimo spinduliu ir žymėsime $r_d = r_d(\mathbf{C})$.

• • • ◇ • • •

Nesunku įsitikinti, kad teisinga nelygybė $r_p(\mathbf{C}) \leq r_d(\mathbf{C})$. Pakavimo spindulys su minimaliu kodo atstumu susijęs taip:

$$r_p(\mathbf{C}) = [\frac{d-1}{2}];$$

čia $[\cdot]$ žymime sveikąją skaičiaus dalį. Tad pakavimo spindulys lygus maksimaliam klaidų, kurias gali ištaisyti kodas, skaičiui.

Skirtumas $r_d(\mathbf{C}) - r_p(\mathbf{C})$ gali būti didelis. Paprastas pavyzdys: $(n, 2)$ kodui

$$\mathbf{C} = \{000\ldots00, 000\ldots11\}$$

gauname: $r_p = 1, r_d = n - 1$.

1.2.3 apibrėžimas.

Kodą $\mathbf{C}$ *vadinsime tobulu, jei*

$$r_p(\mathbf{C}) = r_d(\mathbf{C}).$$

Tobulam (n, N, d) kodui $\mathbf{C}$ turi būti $d(\mathbf{C}) = 2t + 1, r_p(\mathbf{C}) = t$. Susumuokime šiuos pastebėjimus.

1.2.2 teorema.

(n, N, d) kodas $\mathbf{C}$ yra tobulas tada ir tik tada, kai $d = 2t + 1$ ir galioja lygybė

$$NV_q(n, t) = q^n.$$

Kita vertus, jeigu natūraliesiems skaičiams q, n, N, t teisinga lygybė $NV_q(n, t) = q^n$, tai nereiškia, kad kodas su parametrais $(n, N, 2t + 1)$ egzistuoja. Iš tikrųjų tobulųjų kodų yra nedaug.

Apibrėšime ekvivalenčių kodų sąvoką.

Tegu $\mathbf{C}$ yra (n, N) kodas, o σ – n elementų perstata, t.y. injektyvus atvaizdis

$$\sigma\ :\ \{1, 2, \ldots, n\}\ \rightarrow\ \{1, 2, \ldots, n\}.$$

Juo apibrėšime injektyvų atvaizdį $\mathcal{A}_q^n \rightarrow \mathcal{A}_q^n$, kurį taip pat žymėsime σ. Jei $\mathbf{x} = x_1 \ldots x_n$, tai

$$\sigma(\mathbf{x}) = x_{\sigma(1)} \ldots x_{\sigma(n)}.$$

Iš kodo $\mathbf{C}$ šiuo atvaizdžiu gauname naują kodą

$$\sigma(\mathbf{C}) = \{\sigma(\mathbf{c}) : \mathbf{c} \in \mathbf{C}\}.$$

Natūralu kodus $\mathbf{C}$ ir $\sigma(\mathbf{C})$ laikyti ekvivalenčiais. Atskiru atveju gali būti $\mathbf{C} = \sigma(\mathbf{C})$. Šią savybę turintys atvaizdžiai σ ypatingu būdu susiję su kodo struktūra. Tokių atvaizdžių $\sigma : \mathcal{A}_q^n \rightarrow \mathcal{A}_q^n$ aibę

$$Aut(\mathbf{C}) = \{\sigma : \mathbf{C} = \sigma(\mathbf{C})\}$$

vadinsime kodo $\mathbf{C}$ automorfizmų grupe. Ji iš tikrųjų yra grupė algebroje priimta prasme.

Be ką tik aptarto būdo, yra ir kita galimybė gauti formaliai naujus kodus, naudojant perstatas. Tegu $\pi : \{1, 2, \ldots, q\} \rightarrow \{1, 2, \ldots, q\}$ yra kokia nors perstata, o abėcėlės $\mathcal{A}$ simboliai sunumeruoti: $\mathcal{A} = \{a_1, \ldots, a_q\}$. Galime π nagrinėti kaip atvaizdį $\mathcal{A} \rightarrow \mathcal{A}$, apibrėždami $\pi(a_i) = a_{\pi(i)}$. Pasirinkime $i, i \leq n$, ir apibrėškime injektyvų atvaizdį $\langle \pi, i \rangle : \mathcal{A}_q^n \rightarrow \mathcal{A}_q^n$ šitaip:

$$\langle \pi, i \rangle (x_1 \ldots x_i \ldots x_n) = x_1 \ldots \pi(x_i) \ldots\ x_n.$$

$\bullet\bullet\bullet\ \diamond\ \bullet\bullet\bullet$

Atvaizdis $\langle \pi, i\rangle$ keičia tik i-ąjį žodžio simbolį. Kodą, kurį gauname iš $\mathbf{C}$ imdami žodžius $\langle \pi, i\rangle(\mathbf{c}), \mathbf{c} \in \mathbf{C}$, žymėsime $\langle \pi, i\rangle(\mathbf{C})$.

1.2.4 apibrėžimas.

Du (n, N) kodus $\mathbf{C}, \mathbf{C}'$ vadinsime ekvivalenčiais, jei egzistuoja n elementų perstata σ ir q elementų perstatos $\pi_1, \ldots, \pi_n$, kad

$$\mathbf{C}' = \langle \pi_1, 1\rangle(\ldots(\langle \pi_n, n\rangle(\sigma(\mathbf{C})))\ldots).$$

1.2.3 teorema.

Jei kodai $\mathbf{C}, \mathbf{C}'$ ekvivalentūs, tai $d(\mathbf{C}) = d(\mathbf{C}')$.

1.3. Pagrindinė kodavimo problema

Tarkime, iš abėcėlės $\mathcal{A} = \mathcal{A}_q$ n ilgio žodžių norime sudaryti kodą, kuris taisytų t klaidų. Kadangi kodo taisomų klaidų skaičius yra glaudžiai susijęs su minimaliu kodo atstumu, tarsime, kad ir šis parametras apibrėžtas. Natūralu ieškoti paties geriausio (n, N, d) kodo; čia n ir d reikšmės iš anksto pasirinktos. Geriausias (n, N, d) kodas bus tas, kuris turės daugiausia žodžių, t. y. didžiausią parametrą N.[2] Žymėkime:

$$A_q(n, d) = \max\{N : \textit{egzistuoja } (n, N, d) \textit{ kodas } \mathbf{C} \subset \mathcal{A}_q^n\}.$$

Kodus su parametrais $(n, A_q(n, d), d)$ vadinsime **maksimaliais**. Jie pasižymi tuo, kad nė vieno n ilgio žodžio negalima pridėti prie tokio kodo, nesumažinant minimalaus kodo atstumo.

Pagrindine kodavimo problema vadinsime uždavinį:

- *duotiems q, n, d rasti $A_q(n, d)$ reikšmę.*

Kaip dažnai būna, uždavinį lengva išspręsti su ribinėmis parametrų reikšmėmis.

1.3.1 teorema.

Bet kokiems $q \geq 1, n \geq 1$,

$$A_q(n, 1) = q^n, \quad A_q(n, n) = q.$$

Kita akivaizdi išvada: $A_q(n-1, d-1) \geq A_q(n, d)$. Kai $q = 2$, galima įrodyti daugiau.

1.3.2 teorema.

Teisinga lygybė $A_2(n, 2l-1) = A_2(n+1, 2l)$.

Įrodymas. Pakanka įrodyti nelygybę $A_2(n, 2l-1) \leq A_2(n+1, 2l)$. Samprotaukime taip: kiekvieną kodo $(n, A_2(n, 2l-1), 2l-1)$ kodo žodį $x_1x_2\ldots x_n$ žodį pailginkime bitu x_{n+1}, kad būtų teisinga lygybė

$$x_1 + x_2 + \ldots + x_n + x_{n+1} \equiv 0 \pmod 2.$$

[2] Žinoma, epitetas „geriausias“ yra sąlyginis. Šis kodas gali būti visai netikęs naudojimosi juo paprastumo požiūriu.

Pakanka įsitikinti, kad šitaip gauname kodą, kurio minimalus atstumas yra $2l$.

1.3.3 teorema.

Bet kokiems $n \geq 2,\ 1 \leq d \leq n-1,$

$$A_q(n,d) \leq qA_q(n-1,d).$$

Įrodymas. Tegu $M = A_q(n,d)$, o $\mathbf{C}$ koks nors (n,M,d) kodas. Suskaidykime kodą $\mathbf{C}$ į q nesikertančių klasių

$$\mathbf{C}_a = \{\mathbf{c} : \mathbf{c} \in \mathbf{C}, \mathbf{c} = aa_2 \dots a_n\},\ a \in \mathcal{A}.$$

Suprantama, jog egzistuoja nors vienas a, kad $|\mathbf{C}_a| \geq M/q$. Iš šios klasės „sutrumpintų“ žodžių sudarykime naują kodą

$$\mathbf{C}' = \{\mathbf{c}' : \mathbf{c}' \in \mathcal{A}_q^{n-1}, a\mathbf{c}' \in \mathbf{C}_a\}.$$

Kadangi bet kokiems $\mathbf{c}', \mathbf{c}'' \in \mathbf{C}$ Hamingo atstumas

$$h(\mathbf{c}', \mathbf{c}'') = h(a\mathbf{c}', a\mathbf{c}'') \geq d,$$

tai $\mathbf{C}'$ yra $(n-1, |\mathbf{C}'|, d')$ kodas, $d' \geq d$. Tada, žinoma,

$$A_q(n-1,d) \geq |\mathbf{C}'| \geq \frac{M}{q}.$$

Teorema įrodyta.

1.3.4 teorema.

Bet kokiems $n \geq 1,\ q > 1,\ n \geq d \geq 1,$ *teisingas įvertis*

$$A_q(n,d) \leq q^n \Big(\sum_{k=0}^{t} \binom{n}{k} (q-1)^k \Big)^{-1}, \quad t = [\frac{d-1}{2}].$$

Įrodymas. Jei r_p yra (n,N,d) kodo $\mathbf{C}$ pakavimo spindulys, tai rutuliai $B_q(\mathbf{c}, r_p), \mathbf{c} \in \mathbf{C}$, nesikerta. Tada

$$\bigcup_{\mathbf{c} \in \mathbf{C}} B_q(\mathbf{c}, r_p) \subset \mathcal{A}_q^n, \quad |\bigcup_{\mathbf{c} \in \mathbf{C}} B_q(\mathbf{c}, r_p)| \leq |\mathcal{A}_q^n| = q^n.$$

Kadangi $r_p = [(d-1)/2]$, tai, įstatę rutulio tūrio išraišką iš jau įrodytos teoremos, gauname

$$N \sum_{k=0}^{t} \binom{n}{k} (q-1)^k \leq q^n.$$

Teorema įrodyta.

Dar vieną įvertį iš viršaus gausime pastebėję, jog nubraukdami kokio nors (n,N,d) kodo žodžių paskutinius $d-1$ simbolių ($n \geq d$), gausime trumpesnius, bet visus skirtingus žodžius. Jų nėra daugiau kaip q^{n-d+1}. Taigi $N \leq q^{n-d+1}$ ir teisinga tokia teorema.

1.3.5 teorema.

Bet kokiems $n \geq 1,\ q > 1,\ n \geq d \geq 1$, *teisingas įvertis*

$$A_q(n,d) \leq q^{n-d+1}. \quad \text{(Singletono įvertis.)}$$

Gausime $A_q(n,d)$ įvertį iš apačios. Tegu (n,N,d) kodas $\mathbf{C}$ yra maksimalus, t. y. $N = A_q(n,d)$. Tada šio kodo negalima papildyti nei vienu žodžiu, nesumažinant minimalaus atstumo d. Vadinasi, kiekvienam $\mathbf{x} \in \mathcal{A}_q^n$ egzistuoja $\mathbf{c} \in \mathbf{C}$, kad $h(\mathbf{c},\mathbf{x}) \leq d-1$. Tada rutuliai $B_q(\mathbf{c}, d-1)$ padengia visą aibę $\mathcal{A}_q^n$:

$$\bigcup_{\mathbf{c}\in\mathbf{C}} B_q(\mathbf{c}, d-1) \supset \mathcal{A}_q^n.$$

Iš šio sąryšio gauname

$$NV_q(n,d-1) = A_q(n,d)V_q(n,d-1) \geq q^n.$$

Gautasis įvertis vadinamas Gilberto–Varšamovo įverčiu (E.N. Gilbert, R.R. Varshamov). Prisiminę rutulio tūrio išraišką užrašysime jį taip.

1.3.6 teorema.

Bet kokiam $n \geq 1,\ q > 1,\ d \geq 1$,

$$A_q(n,d) \geq q^n \left(\sum_{k=0}^{d-1} \binom{n}{k} (q-1)^k \right)^{-1}.$$

Ši nelygybė vadinama Gilberto–Varšamovo įverčiu.

Jeigu informaciją rengiamės koduoti n ilgio žodžių kodu, o kanale kiekvienas simbolis nepriklausomai nuo kitų iškraipomas su tikimybe p, tai vidutiniškai žodyje įvyksta np iškraipymų. Norėdami, kad vidutiniškas klaidų skaičius būtų visada ištaisomas, turime naudoti kodą, kurio minimalus atstumas $\approx 2pn$. Taigi siekiant minimalių sąnaudų geriausia naudoti $(n, A_q(n,[\delta n]), [\delta n]), \delta = 2p$, kodą. Šio kodo koeficientas

$$R = \frac{\log_q A_q(n,[\delta n])}{n}.$$

Panagrinėkime šio santykio elgesį, kai n didėja. Apibrėžkime dydį

$$\alpha(\delta) = \limsup_{n\to\infty} \frac{\log_q A_q(n,[\delta n])}{n}.$$

Pasinaudoję Singletono įverčiu iškart gauname:

$$\alpha(\delta) = \limsup_{n\to\infty} \frac{\log_q A_q(n,[\delta n])}{n} \leq \limsup_{n\to\infty} \frac{\log_q q^{n-[\delta n]+1}}{n} = 1-\delta.$$

Pasinaudoję Gilberto-Varšamovo įverčiu, galime dydį $\alpha(\delta)$ šitaip įvertinti iš apačios:

$$\alpha(\delta) \geq 1 - \limsup \log_q V_q(n,[\delta n]-1) \geq 1 - \limsup \log_q V_q(n,[\delta n]).$$

Paskutinąją ribą galima suskaičiuoti.

• • • ◇ • • •

1.3.7 teorema.

Tegu

$$H_q(\delta) = \delta \log_q(q-1) + \delta \log_q \frac{1}{\delta} + (1-\delta)\log_q \frac{1}{(1-\delta)}.$$

Tada visiems $\delta \geq (q-1)/q$ teisinga nelygybė

$$\alpha(\delta) \geq 1 - H_q(\delta).$$

1.4. Kodai su kontroliniu simboliu

Jeigu dekoduojant aptiktą klaidą galime ištaisyti naudodamiesi atgaliniu ryšiu su informacijos šaltiniu (pavyzdžiui, paskambinę siuntėjui telefonu, arba, jei klaida įvyko nuskaitant prekės kodą elektroniniu-optiniu prietaisu, tiesiog pasitikslinę prekių apskaitos knygoje), tai pigaus ir greito klaidos aptikimo galimybė yra svarbi. Aptarsime informacijos kodavimo pridedant kontrolinį simbolį idėją bei praktikoje naudojamus atvejus.

1.4.1. Knygų numeracijos sistema ISBN[3]

Informacija apie Vakarų šalių leidyklų leidžiamas knygas yra koduojama devyniais dešimtainiais skaitmenimis, o galimai pavienei paprastai klaidai aptikti pridedamas kontrolinis dešimtasis.

Pavyzdžiui, kode ISBN 3–540–97812–7 skaitmuo 3 reiškia šalį (Vokietiją), 540 – leidyklą (Springer), skaitmenimis 97812 užkoduoti pačios knygos duomenys, 7 – kontrolinis skaitmuo.

ISBN kodo $a_1a_2\ldots a_9a_{10}$ kontrolinė lygtis

$$\sum_{i=1}^{9} ia_i \equiv a_{10} (\text{mod } 11) \text{ arba } \sum_{i=1}^{10} ia_i \equiv 0 (\text{mod } 11). \tag{1.4.1}$$

Įvykus vienai paprastai klaidai kontrolinė lygybė nebegalioja, t. y. pavienė paprasta klaida visuomet aptinkama. Taip pat visada aptinkamos pavienės transpozicinės klaidos (kai sukeičiami vietomis gretimi simboliai).

Kadangi atskiru atveju a_{10} gali būti lygu ir 10, tai tenka naudoti papildomą „skaitmenį“ $a_{10} = X$. Sudarant kontrolines lygybes moduliu 10, to nereikia.

Įdomu, kad tuo atveju, kai skaitmenys a_i yra iš dvinarės abėcėlės $\{0;1\}$ (1.4.1) kontrolinė lygybė ne tik leidžia aptikti įvykusią klaidą, bet ir atkurti vieną „pamestą “ kanale simbolį.

1.4.1 apibrėžimas.

Kodą iš dvinarės abėcėlės ilgio žodžių

$$VT_0(n) = \{x_1x_2\ldots x_n : \sum_{i=1}^{n} ia_i \equiv 0(\text{mod } (\text{n}+1))\}$$

vadinsime Varšamovo-Tenengolto kodu.

Tarkime siunčiant vieną Varšamovo-Tenengolto kodo žodį vienas jo bitas kanale „pradingo“ ir vietoje žodžio $x_1x_2\ldots x_n$ gavome žodį $x_1'x_2'\ldots x_{n-1}'$. Remiantis kontroline lygybe galima ne tik sužinoti koks simbolis pradingo, bet ir kurioje vietoje reikia jį įterpti.

[3] International Standard of Book Numeration (angl.).

Panagrinėkime pavyzdį. Tarkime, siunčiant kanalu kodo $VT_0(5)$ žodį $x = 11011$ buvo gautas žodis $x' = 1111$, t.y. simbolis $x_3 = 0$ buvo prarastas. Reikia surasti, koks simbolis prarastas ir kur jį reikia įterpti. Suskaičiuokime gautojo žodžio svorį w (nenulinių komponenčių skaičių) ir kontrolinę sumą:

$$w = 4, \quad s' = 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 + 4 \cdot 1 = 10.$$

Toliau skaičiuojame taip: siųstojo žodžio kontrolinė suma galėjo būti tik $s = 12$; skirtumas $s - s' = 2$ yra mažesnis už svorį w. Darome tokią išvadą: **prarastasis simbolis lygus** 0 **ir jį reikia įterpti taip, kad dešiniau būtų lygiai** $s - s'$**, t.y. lygiai** 2 **vienetai.**

Dabar tegu siunčiant žodį $x = 11011$ buvo gautas žodis $x' = 1101$, t.y. simbolis $x_5 = 1$ buvo prarastas (tą patį žodį gautume praradę ketvirtąjį simbolį). Vėl skaičiuojame gautojo žodžio svorį ir kontrolinę sumą:

$$w = 3, \quad s' = 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 0 + 4 \cdot 1 = 7.$$

Skirtumas tarp kontrolinių sumų reikšmių $s - s' = 5$ yra didesnis už svorį w. Darome tokią išvadą: **prarastasis simbolis lygus** 1 **ir jį reikia įterpti taip, kad kairiau būtų lygiai** $s - s' - 1$**, t.y. lygiai** 1 **nulis.**

Kodėl toks algoritmas veikia bet kokiu atveju? Tai nelabai sunkus, tačiau ir netrivialus uždavinys. Reikia pagalvoti, kas atsitinka, kai s-ojoje vietoje prarandamas nulis arba vienetas, kaip pasikeičia kontrolinė suma, kaip skirtumas susijęs su gautojo žodžio svoriu...

1.4.2. Prekių numeracijos sistema EAN[2]

Europoje pagamintos prekės koduojamos 13 dešimtainių skaitmenų, tryliktasis skaitmuo – kontrolinis. Kad kodą galėtų nuskaityti elektroninis-optinis prietaisas, skaitmenys užrašomi ant etiketės juodais brūkšniais.

Kontrolinė lygybė tokia:

$$1a_1 + 3a_2 + 1a_3 + 3a_4 + \ldots + 1a_{11} + 3a_{12} + a_{13} \equiv 0 (\text{mod } 10).$$

Ši lygybė taip pat leidžia aptikti visas pavienes paprastas klaidas. Žinoma, koeficientus prie a_i kairėje kontrolinės lygybės pusėje galima ir kitaip pasirinkti. Jei $w_i, (w_i, 10) = 1$, tai sekai $a_1 a_2 \ldots a_{12}$ kontrolinį simbolį galime sudaryti naudojantis kontroline lygybe

$$\sum_{i=1}^{12} w_i a_i \equiv -a_{13} (\text{mod } 10).$$

Su šia kontroline lygybe visada aptinkamos pavienės paprastos klaidos. Galima tikėtis, jog išradingiau sudarytos kontrolinės lygybės įgalins aptikti ne tik pavienes paprastas klaidas.[2]

1.4.3. Grupės, lotyniškieji kvadratai ir kontrolinės lygybės

Kodavimo su kontroliniu simboliu metodą galime suformuluoti bendresniame kontekste. Tegu abėcėlė $\mathcal{A}$ kartu su dvinare algebrine operacija $*$ sudaro grupę. Jei $\delta_i : \mathcal{A} \to \mathcal{A}$, $i = 1, \ldots, n$, yra abipusiškai vienareikšmiai atvaizdžiai, tai n ilgio žodžiams pridėsime kontrolinį simbolį, sudarytą taip:

$$a_1 a_2 \ldots a_n \to a_1 a_2 \ldots a_n a_{n+1},$$

$$a_{n+1} = \delta_1(a_1) * \ldots * \delta_n(a_n).$$

Tikėtina, kad parinkus algebrinę struktūrą bei atvaizdžius $\delta_1(a_1), \ldots, \delta_n(a_n)$ įmanoma pasiekti, jog kodas su kontroliniu simboliu įgalins aptikti ne vien pavienes paprastas klaidas.

---

[2] European Article Numeration (angl.).

[2] Apie tai, kaip skaitmenys koduojami juodai brūkšniais, taip pat apie kodo galimybes aptikti, jog įvyko klaidos, galite pasiskaityti J. Volosatovo straipsnyje „EAN – šiek tiek matematikos buityje“, Alfa plius omega, 2000, Nr. 1, 86-90.

Galima toliau plėtoti kodo su kontroliniu simboliu temą. Ar kontrolinei lygybei sudaryti išties būtina, kad abėcėlėje būtų apibrėžta grupės struktūra? Juk kontrolinei lygčiai esminga tik, kad iš $a \neq b$ išplauktų $ac \neq bc$ visiems c.

1.4.2 apibrėžimas.

Tegu $n \times n$ matricos L kiekvieną eilutę sudaro atitinkamai perstatyti aibės $N_n = \{1, 2, \ldots, n\}$ elementai. Matricą L vadinsime lotyniškuoju kvadratu, jei kiekvienas $m \in N_n$ įeina lygiai vieną kartą į kiekvieną matricos stulpelį.

Turėdami lotyniškąjį kvadratą $L = (n_{ij})$, galime abėcėlėje $\mathcal{A} = \{a(1), \ldots, a(n)\}$ apibrėžti algebrinę operaciją taip:

$$a(k) * a(l) = a(n_{kl}).$$

Apibrėžtoji operacija nebūtinai komutatyvi bei asociatyvi, tačiau turi savybę: jei $a \neq b$, tai $ac \neq bc$ visiems c. Žodžiui $x_1 \ldots x_m$ iš abėcėlės $\mathcal{A}$ simbolių galime kontrolinį simbolį x_{m+1} sudaryti taip:

$$(\ldots((x_1 * x_2) * x_3)\ldots) * x_m = x_{m+1}.$$

Egzistuoja bet kokios eilės lotyniškieji kvadratai. Iš tikrųjų $A = (a_{ij})_{i,j=1,\ldots,n},\ a_{ij} \equiv i + j - 1 \pmod{n}$ yra lotyniškasis kvadratas.

Štai dar viena lotyniškojo kvadrato konstrukcija. Tegu $GF(q)$ yra baigtinis q eilės kūnas (Galua kūnas), $\lambda \in GF(q),\ \lambda \neq 1, 2^{-1}$. Tada

$$L = \{\lambda a + (1 - \lambda)b\}_{a,b \in GF(q)}$$

yra q-os eilės lotyniškasis kvadratas.

1.5. Hadamardo matricos ir kodai

Sukonstruoti kodą su dideliu minimaliu atstumu – anaiptol nelengvas uždavinys. Pasvarstykime: pirmąjį žodį galime parinkti bet kokį, tačiau antrąjį turime rinkti taip, kad jo Hammingo atstumas iki pirmojo žodžio būtų didelis, rinkdami trečiąjį jau turime stengtis, kad jis būtų pakankamai „toli“ nuo pirmųjų dviejų ir t.t. Tačiau matematikai nėra tokie darbštūs, kad jiems patiktų perrinkimo procedūros. Geriems kodams sudaryti yra ir kitų būdų.

1.5.1 apibrėžimas.

n-tos eilės kvadratinė matrica $H_n = (h_{ij})$ vadinama Hadamardo matrica, jei

$$h_{ij} = \pm 1, \quad H_n \cdot H_n^\top = nI_n,$$

čia $H_n^\top$ žymi transponuotą matricą, I_n – vienetinę.

Pavyzdys.
Matrica

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

yra Hadamardo matrica.

Iš Hadamardo matricų poros galime sukonstruoti didesnio matavimo Hadamardo matricą.

1.5.2 apibrėžimas.

Tegu $A = (a_{ij})$ yra $n \times n$, $B = (b_{ij})$, $m \times m$ matricos. Jų Kroneckerio sandauga vadinama $nm \times nm$ matrica

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \ldots & a_{1m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{11}B & a_{12}B & \ldots & a_{1m}B \end{pmatrix}$$

• • • ◇ • • •

1.5.1 teorema.

Jei A, B yra Hadamardo matricos, tai $A \otimes B$ irgi Hadamardo matrica.

Naudodamiesi šia teorema galime konstruoti $2^m \times 2^m$ matavimų Hadamardo matricas.

1.5.2 teorema.

Sukeitus Hadamardo matricos dvi eilutes (stulpelius) vietomis gautoji matrica vėl yra Hadamardo matrica. Padauginus Hadamardo matricos eilutę (stulpelį) iš -1, gaunama Hadamardo matrica.

Abiems teoremas galime įrodyti pasirėmę vien Hadamardo matricos apibrėžimu.

1.5.3 apibrėžimas.

n-tos eilės Hadamardo matrica $H_n = (h_{ij})$ vadinama normaline, jei visiems j teisinga lygybė $h_{1j} = h_{j1} = 1$.

Taigi normalinės Hadamardo matricos pirmasis stulpelis ir pirmoji eilutė sudaryti vien tik iš vienetų. Pasirėmę ankstesniu teiginiu gauname, kad kiekvieną Hadamardo matricą galima suvesti į normalinį pavidalą.

1.5.3 teorema.

Atitinkamas eilutes ir stulpelius dauginant iš -1 kiekvieną Hadamardo matricą galima suvesti į normalinę matricą.

Naudojantis šia teorema nesunku įrodyti, kad trečios eilės Hadamardo matricos nėra. Iš tiesų, jeigu egzistuotų trečios eilės Hadamardo matrica, tai suvedę ją į normalinį pavidalą gautume Hadamardo matricą

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & a & b \\ 1 & c & d \end{pmatrix},$$

čia a, b yra arba 1 arba -1. Tačiau ši matrica yra Hadamardo matrica, todėl turi būti $1 + a + b = 0$. Aišku, kad tokia lygybė negali būti patenkinta, todėl trečios eilės Hadamardo matricų nėra.

1.5.4 teorema.

Jei H_n yra Hadamardo matrica, tai $n = 1, 2$ arba n dalijasi iš 4.

Įrodymas. Pakanka nagrinėti normalines n-os eilės Hadamardo matricas. Tada kiekvienoje tokios matricos eilutėje pradedant antrąja turi būti po tiek pat skaičių 1 ir -1. Taigi n turi būti lyginis $n = 2m$. Sukeisdami matricos stulpelius vietomis galime ją pertvarkyti taip, kad pirmieji m antros eilutės elementai būtų lygūs vienetui. Tegu tarp pirmųjų m trečios eilutės elementų yra j vienetų ir $m - j$ minus vienetų.

• • • ◇ • • •

Tada tarp paskutiniųjų m trečios eilutės elementų vienetų yra $m-j$, o minus vienetų – j. Antroji ir trečioji eilutės yra ortogonalios, t.y. jų skaliarinė sandauga lygi 0 :

$$j-(m-j)-(m-j)+j=0, \quad m=2j.$$

Taigi skaičius $n=2m$ dalijasi iš keturių.

Iki šiol nėra žinoma, ar kiekvienam $n=4m$ Hadamardo matrica egzistuoja. Tačiau jeigu skaičius $4m-1$ yra pirminis, tai $4m$ eilės Hadamardo matrica egzistuoja ir ją galima lengvai sukonstruoti. Tam reikia šiek tiek skaičių teorijos žinių. Prisiminsime jas.

1.5.4 apibrėžimas.

Tegu q yra pirminis skaičius, Legendre (Ležandro) simboliu vadinama funkcija

$$\left(\frac{a}{q}\right)=\begin{cases}0, & \text{jei } a\equiv 0 \bmod q,\\ 1, & \text{jei lyginys } x^2\equiv a \bmod q, \text{ turi sprendinį},\\ -1, & \text{jei lyginys } x^2\equiv a \bmod q, \text{ neturi sprendinių.}\end{cases}$$

Skaičių teorijoje įrodomi šios Legendre simbolio savybės.

1.5.5 teorema.

Teisingos šios lygybės:

$$\left(\frac{a^2}{q}\right)=1, \qquad \left(\frac{a+kq}{q}\right)=\left(\frac{a}{q}\right),$$
$$\left(\frac{ab}{q}\right)=\left(\frac{a}{q}\right)\left(\frac{b}{q}\right), \qquad \left(\frac{-1}{q}\right)=(-1)^{(q-1)/2}.$$

1.5.6 teorema.

Tegu skaičius k nesidalija iš q. Teisingos šios lygybės:

$$\sum_{x=1}^{q-1}\left(\frac{x}{q}\right)=0, \quad \sum_{x=1}^{q-1}\left(\frac{x(x+k)}{q}\right)=-1.$$

Ležandro simboliams skaičiuoti labai praverčia šis Gauso dėsnis.

1.5.7 teorema.

Su bet kokiais skirtingais pirminiais skaičiais teisinga lygybė

$$\left(\frac{p}{q}\right)\cdot\left(\frac{q}{p}\right)=(-1)^{(p-1)(q-1)/2}.$$

1.5.5 apibrėžimas.

Tegu q – pirminis skaičius. Matricą

$$S=(S_{ij})_{i,j=1,\dots,q}, \quad S_{ij}=\left(\frac{i-j}{q}\right)$$

vadinsime $q-$ eilės Paley matrica.

Visuomet J (O) žymėsime atitinkamos eilės vien tik iš vienetų (nulių) sudarytą matricą, I – vienetinę matricą. Paley matrica turi šias savybes.

1.5.8 teorema.

Teisingos lygybės:

$$S \cdot J = J \cdot S = O, \quad S^\top = (-1)^{(q-1)/2} S, \quad S \cdot S^\top = (-1)^{(q-1)/2}(qI - J)$$

Tegu S – q-os eilės Paley matrica. Sudarysime $q+1$-os eilės kvadratinę matricą C papildydami S eilute $0\ 1\ 1\ 1\ \ldots\ 1$ ir stulpeliu $0\ -1\ -1\ -1\ \ldots\ -1$. Mūsų konstrukcija remiasi tokiu teiginiu.

1.5.9 teorema.

Jei q yra pirminis skaičius ir $q \equiv 3 \bmod 4$, tai matrica $H_{q+1} = I + C$ yra Hadamardo matrica.

Iš Hadamardo matricos eilučių galime sudaryti kodus. Pirmiausia pasirūpinkime, kad matrica būtų sudaryta vien iš nulių ir vienetų.

1.5.6 apibrėžimas.

Jei H_n yra Hadamardo matrica, tai matrica M_n, gauta iš H_n, pakeitus 1 į 0 ir -1 į 1, vadinama dvinare Hadamardo matrica.

Kvadratinę matricą, kurios visi elementai lygūs 1 žymėsime J_n.

1.5.10 teorema.

Tegu M_n yra n-tos eilės dvinarė Hadamardo matrica, gauta iš normalizuotos Hadamaro matricos, M'_n – matrica, gauta iš M_n, pakeitus 1 į 0 ir 0 į 1. Kodas $\mathbf{A}_n$ kodas, sudarytas iš M_n eilučių, sutrumpintų pirmuoju simboliu, $\mathbf{B}_n$ – iš M_n ir M'_n eilučių, o $\mathbf{C}_n$ – iš $\mathbf{B}_n$ eilučių, sutrumpintų pirmuoju simboliu. Tada $\mathbf{A}_n$ yra $(n-1, n, n/2)$, $\mathbf{B}_n$ – $(n, 2n, n/2)$, o $\mathbf{C}_n$ – $(n-1, 2n, d)$, $d \geq n/2-1$ kodai.
Kodai $\mathbf{A}_n, \mathbf{B}_n$ vadinami Hadamardo kodais.

• • • ◇ • • •