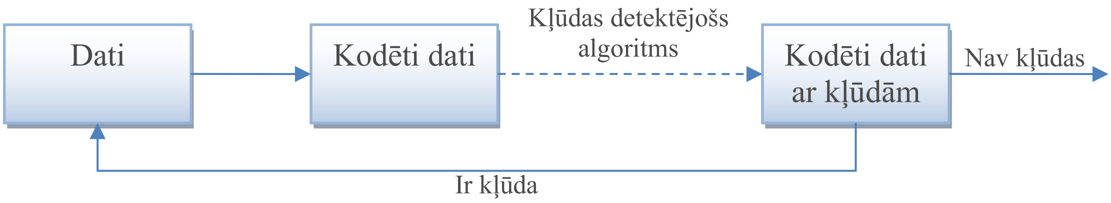
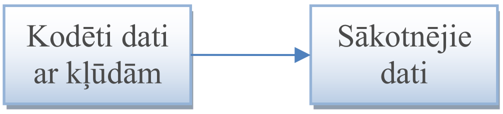
Kļūdu korekcija I

Mērķis: pārraidīt datus pa sakaru kanālu, kurā var rasties kļūdas. Kļūdas detektējošs kods:



Kļūdas koriģējošs kods:



Kļūdu detekcija

$x_1, x_2, \dots, x_n \in \{0; 1\}$ – pārraidāmie dati (n bitu virkne).

Pārraidām $n+1$ bitu: $x_1, x_2, \dots, x_n$ un $(x_1 + x_2 + \dots + x_n)\ mod\ 2$. Pēdējais bits glabā visu iepriekšējo bitu paritāti. Līdz ar to visu $n+1$ bitu paritāte ir 0. Ja rodas viena kļūda, tad visu bitu paritāte būs 1, un kļūdu varēs konstatēt.

Kļūdu korekcija

Pamatideja visās metodēs – pretēji saspiešanai – papildināt pārraidāmos datus ar papildinformāciju, cerot, ka šī papildinformācija ļaus pamanīt kļūdas pārraidē.

Vienkāršākā metode – datu atkārtošana. Ja pārraidāmais bits x_1 ir nulle, tad pārraidām trīs nulles, pretējā gadījumā – trīs vieniniekus. Lai izlabotu kļūdu, skatāmies, kādu bitu ir vairāk – nuļļu vai vieninieku. Šis kods spēj koriģēt vienu kļūdu.

$$x_1 \begin{array}{l} \nearrow 0 \to 000 \\ \searrow 1 \to 111 \end{array} \qquad \begin{array}{l} 000 \to 000 \to 0 \\ 010 \to 000 \to 0 \\ 110 \to 111 \to 1 \end{array}$$

Kodēšana *Dekodēšana*

Par $[n, k, d]$-kodu sauc kodu, kurā

n – bitu skaits, kas tiek pārraidīts,

k – pārraidāmo bitu skaits,

d – kļūdu skaits, ko iespējams koriģēt.

Atkārtošanas metodei $n = 3, k = 1, d = 1$, tāpēc tas ir $[3,1,1]$-kods.

Kodi, kas koriģē vienu kļūdu

S – virkņu kopa (visas virknes ir garumā n). Kā pārbaudīt, vai S spēj koriģēt vienu kļūdu?

Mēģināsim noteikt, kuros gadījumos kods **nespēj** koriģēt vienu kļūdu? Tas, ka kods nespēj koriģēt vienu kļūdu, nozīmē, ka eksistē virkne $z = z_1 z_2 \dots z_n$, kuru var koriģēt par (vismaz) divām dažādām kopas S virknēm. Apzīmēsim tās ar $x = x_1 x_2 \dots x_n$ un $y = y_1 y_2 \dots y_n$. Katra no virknēm x un y var vai nu sakrist ar z, vai arī atšķirties no tās tieši vienā vietā (jo kods koriģē vienu kļūdu). Tas savukārt nozīmē, ka virknes x un y atšķiras vienā vai divās vietās.

Aplūkosim piemēru. Kopa $S = \{1000, 0101, 1101\}$ nespēj koriģēt vienu kļūdu, jo, saņemot virkni 0101, nav skaidrs, vai tika pārraidīta virkne 0101 (bez kļūdām) vai arī virkne 1101 (ar vienu kļūdu – pirmajā bitā). Tāpat, saņemot virkni 1001, nav skaidrs, vai tika pārraidīta virkne 1000 (ar kļūdu ceturtajā bitā) vai 1101 (ar kļūdu otrajā bitā). Ir spēkā vispārīgāks rezultāts.

Teorēma. Kopa S ir $[n, k, d]$-kods tad un tikai tad, ja

1) S sastāv no virknēm garumā n,
2) $|S| \geq 2^k$ (lai katrai k-bitu virknei varētu piekārtot pārraidāmo virkni no S),
3) katras divas virknes no S atšķiras vismaz $2d + 1$ vietā.

Teorēmu var pierādīt, spriežot līdzīgi. Kods nespēj koriģēt d kļūdas, ja eksistē virkne $z = z_1 z_2 \dots z_n$ un divas kopas S virknes $x = x_1 x_2 \dots x_n$ un $y = y_1 y_2 \dots y_n$, kas katra atšķiras no z ne vairāk kā d vietās. Līdz ar to virknes x un y atšķiras ne vairāk kā $2d$ vietās. Tāpēc, lai kods spētu koriģēt d kļūdas, katrām divām kopas S virknēm ir jāatšķiras vismaz $2d + 1$ vietā.

Kodu, kas koriģē vienu kļūdu, piemēri

Cik liela var būt bitu virkņu kopa S, kur katras divas virknes atšķiras vismaz trīs vietās?

a) $n = 3$. Acīmredzami vairāk par divām virknēm kopai S piederēt nevar. Ja virkne $x_1 x_2 x_3 \in S$, tad tai vēl var piederēt vienīgi tāda virkne $y_1 y_2 y_3$, ka $y_1 \neq x_1$, $y_2 \neq x_2$ un $y_3 \neq x_3$.

b) $n = 4$. Pierādīsim, ka arī šajā gadījumā kopa S nevar saturēt vairāk par divām virknēm. Pieņemsim pretējo un aplūkosim trīs virknes, ko satur kopa S: $x = x_1 x_2 x_3 x_4$, $y = y_1 y_2 y_3 y_4$ un $z = z_1 z_2 z_3 z_4$. Virknes x un y nevar sakrist vairāk kā vienā vietā. Tāpat virknes x, z un y, z. Līdz ar to kopējais sakritību skaits nevar būt lielāks par 3. Tagad aplūkosim katru bitu atsevišķi. Ir iespējams divi varianti:
   1) visi trīs biti sakrīt ($x_i = y_i = z_i$). Tad sakritību skaits šajā bitā ir 2.
   2) divi biti sakrīt, bet trešais – atšķiras. Tad sakritību skaits šajā bitā ir 1.

   Tā kā mums ir 4 biti, tad varam secināt, ka kopējais sakritību skaits būs vismaz 4, kas ir pretrunā ar to, ka šis skaits nevar būt lielāks par 3. Tātad kopa S nevar saturēt vairāk par divām virknēm.

c) $n = 5$. Kopa S var saturēt 4 virknes. Piemērs:

x_1	x_2	$x_1, x_1, x_2, x_2, (x_1 + x_2)\ mod\ 2$
0	0	00000
0	1	00111
1	0	11001
1	1	11110

Būtībā esam ieguvuši algoritmu, kas kodē divu bitu virknītes par piecu bitu virknītēm: divreiz pārraidām pirmo bitu un divreiz – otro bitu, bet kā pēdējo bitu pārraidām abu bitu paritāti. Tas ir $[5,2,1]$-kods.

Pierādīsim, ka kopa S nevar saturēt vairāk par četrām virknēm. Pieņemsim pretējo un aplūkosim piecas virknes, ko satur kopa S. Vismaz trim no tām pirmais bits būs vienāds, t.i., vai nu būs vismaz 3 virknes, kurām pirmais bits vienāds ar 0, vai arī vismaz 3 virknes, kurām pirmais bits vienāds ar 1. Šīm trim virknēm atšķirības var būt tikai pēdējos četros bitos. Bet četru bitu gadījumā jau tika pierādīts, ka lielākais virkņu skaits ir 2. Tāpēc kādas divas no šīm trim virknēm atšķirsies mazāk nekā trijās vietās.

d) $n = 7$. Iespējams izveidot $2^4 = 16$ virknes:

x_1	x_2	x_3	x_4	y_1	y_2	y_3	$x_1x_2x_3y_1x_4y_2y_3$
0	0	0	0	0	0	0	0000000
0	0	0	1	0	1	1	0000111
0	0	1	0	1	0	1	0011001
0	0	1	1	1	1	0	0011110
0	1	0	0	1	1	0	0101010
0	1	0	1	1	0	1	0101101
0	1	1	0	0	1	1	0110011
0	1	1	1	0	0	0	0110100
1	0	0	0	1	1	1	1001011
1	0	0	1	1	0	0	1001100
1	0	1	0	0	1	0	1010010
1	0	1	1	0	0	1	1010101
1	1	0	0	0	0	1	1100001
1	1	0	1	0	1	0	1100110
1	1	1	0	1	0	0	1111000
1	1	1	1	1	1	1	1111111

x_1, x_2, x_3 un x_4 ir patvaļīgi pārraidāmie biti, bet y_1, y_2 un y_3 ir kontrolbiti, kurus aprēķina šādi:

$y_1 = (x_1 + x_2 + x_3)\ mod\ 2$

$y_2 = (x_1 + x_2 + x_4)\ mod\ 2$

$y_3 = (x_1 + x_3 + x_4)\ mod\ 2$

Šis ir [7,4,1]-kods, kuru sauc par Haminga kodu.