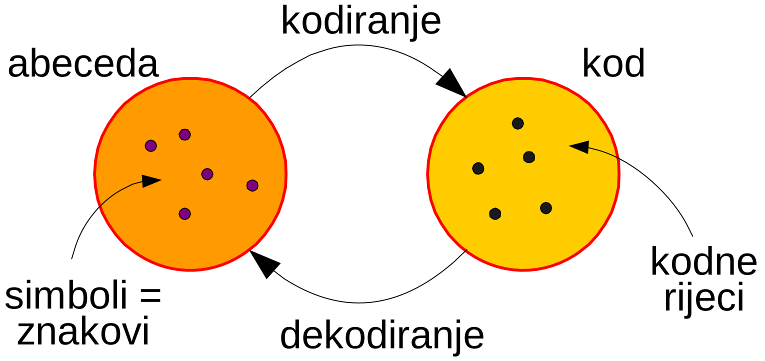
2. Brojevni sustavi i kodovi (2)

Sadržaj predavanja

- **binarno kodiranje znamenki i simbola**
  - **dekadski kodovi**
  - **Grayev kod**
  - **znakovni kodovi**
- kodovi za zaštitu podataka

Binarno kodiranje znamenki i simbola

- princip kodiranja binarnim riječima:
  - izražavanje simbola/znakova u *binarnom* obliku, radi dalje obrade digitalnim sklopom
    ~ binarno *kodiranje*
  - *kôd* : grupa simbola kojoj se dogovorno daje značenje
  - *kodna riječ* : niz bitova kojem se pridaje neko značenje
  - *abeceda* : skup svih simbola prikazanih kodnim riječima
  - *znakovi* : elementi abecede



Binarno kodiranje znamenki i simbola

- princip kodiranja binarnim riječima:
  - broj simbola = broj različitih prikaza
    $\rightarrow$ broj bitova kodnih riječi

K simbola: $n \geq \text{ld } K$ [bit], $\text{ld } x = \log_2 x$

$$2^n \geq K$$

  - n bitova: $N = 2^n$ mogućih kombinacija

pridruživanje kodne riječi prvom simbolu	N načina
pridruživanje kodne riječi drugom simbolu	N-1 način
pridruživanje kodne riječi trećem simbolu	N-2 načina
…	…
pridruživanje kodne riječi K-tom simbolu	N-(K-1)

$$N \cdot (N-1) \cdot (N-2) \cdot \cdots \cdot (N-(K-1)) = \frac{N!}{(N-K)!} = V_N^{(K)}$$

Binarno kodiranje znamenki i simbola

- dekadski kodovi
  ~ binarni prikaz dekadskih znamenki
  - n $\geq$ 4 bita; $2^3 < 10 < 2^4$
  - n = 4 bita
    ~ 16 kombinacija
  - broj 4-bitnih kodova
    ~ mogući broj kodiranja: $V_{16}^{(10)} = \dfrac{16!}{6!} = 29{,}059 \cdot 10^9$
  - odabrati kodove s povoljnim svojstvima!

Dekadski kodovi

- svojstva dekadskih kodova:
  - aditivnost
    ~ veza između kodne riječi
    i prikazane dekadske znamenke
  - samokomplementarnost
    (engl. self-complementing)
    ~ veza kodnih riječi po parovima

Dekadski kodovi

- težinski kod:
  - zbroj težina = vrijednost prikazane znamenke

$$N = \sum_{i=0}^{n-1} a_i \cdot w_i + D$$

N	:	dekadski ekvivalent
w_i	:	i-ta težina
a_i	:	koeficijent za i-tu težinu
D	:	konstanta pomaka

  - 17 težinskih kodova s pozitivnim težinama, 71 s jednom ili dvije negativne težine

Dekadski kodovi

- samokomplementirajući kod:
  ~ 9-komplement dekadskog broja
  zamjenom 0 i 1 u kodnoj riječi
  - korisno kod binarno-dekadske aritmetike
  - težinski je kod samokomplementirajući ako:

$$\sum_i w_i = 9$$

Dekadski kodovi

- kod 8421,
  BCD (engl. Binary Coded Decimal)
  - prvih 10 binarnih brojeva
  - težine: 8, 4, 2, 1
  - neupotrijebljene kombinacije: 1001÷1111

	2^3	2^2	2^1	2^0
	8	4	2	1
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
	1	0	1	0
	1	0	1	1
	1	1	0	0
	1	1	0	1
	1	1	1	0
	1	1	1	1

Dekadski kodovi

- kod 2421 (Aikenov kod)
  - težinski kod
    ~ težine: 2, 4, 2, 1
  - samokomplementirajući kod:
    0-9, 1-8, 2-7, 3-6, 4-5
  - prvih i zadnjih pet
    4-bitnih brojeva
  - neupotrijebljene kombinacije:
    0101÷1010

	2	4	2	1
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
	0	1	0	1
	0	1	1	0
	0	1	1	1
	1	0	0	0
	1	0	0	1
	1	0	1	0
5	1	0	1	1
6	1	1	0	0
7	1	1	0	1
8	1	1	1	0
9	1	1	1	1

Dekadski kodovi

- kod XS-3 (Stibitzov kod)
  - kod 8421,
    s "prekoračenjem" (ekscesom) od 3
  - uz D = 3
    ~ težinski kod
  - ne postoji 0000:
    detekcije prekida kod prijenosa
  - neupotrijebljene kombinacije:
    0000÷0010, 1101÷1111
  - simetrična tablica koda
    ~ samokomplementirajući kod!

	2^3 8	2^2 4	2^1 2	2^0 1
	0	0	0	0
	0	0	0	1
	0	0	1	0
0	0	0	1	1
1	0	1	0	0
2	0	1	0	1
3	0	1	1	0
4	0	1	1	1
5	1	0	0	0
6	1	0	0	1
7	1	0	1	0
8	1	0	1	1
9	1	1	0	0
	1	1	0	1
	1	1	1	0
	1	1	1	1

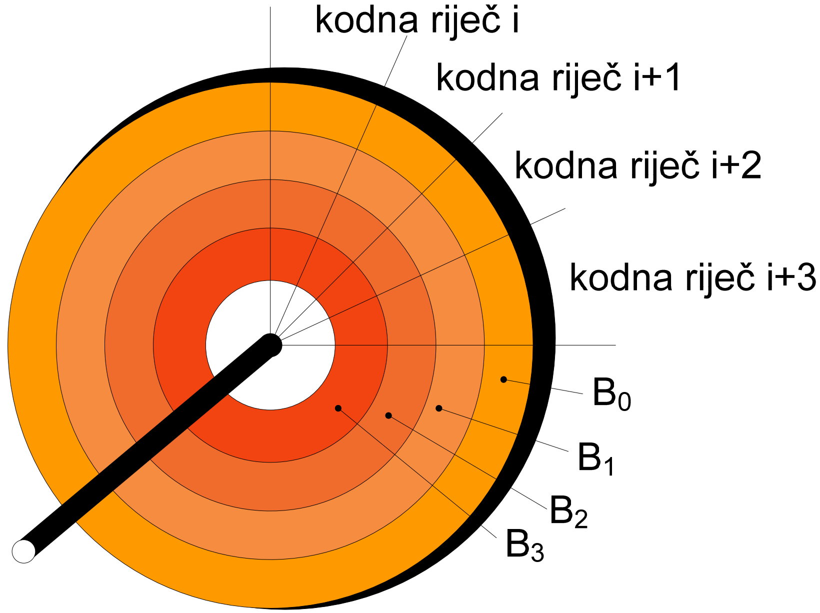
Dekadski kodovi

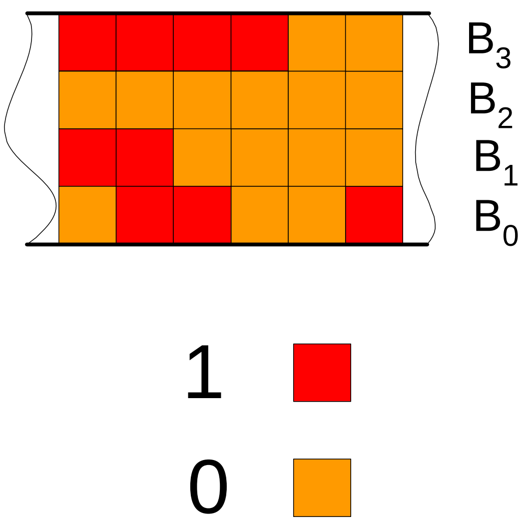
- bikvinarni kod
  - težinski 7-bitni kod (2+5=7)
  - kodne riječi s dvije 1:
    - otkrivanje pogrešaka
    - ne ako je pogreška samokompenzirajuća
  - velika zalihost:
    ~ 10 od 128
    mogućih kombinacija

	5	0	4	3	2	1	0
0	0	1	0	0	0	0	1
1	0	1	0	0	0	1	0
2	0	1	0	0	1	0	0
3	0	1	0	1	0	0	0
4	0	1	1	0	0	0	0
5	1	0	0	0	0	0	1
6	1	0	0	0	0	1	0
7	1	0	0	0	1	0	0
8	1	0	0	1	0	0	0
9	1	0	1	0	0	0	0

Gravev kod

- kod s *minimalnom* promjenom
  - susjedne kodne riječi
    ~ razlika u samo 1 bitu
  - ograničavanje pogreški pri slijednoj promjeni
    npr. direktno očitavanje položaja





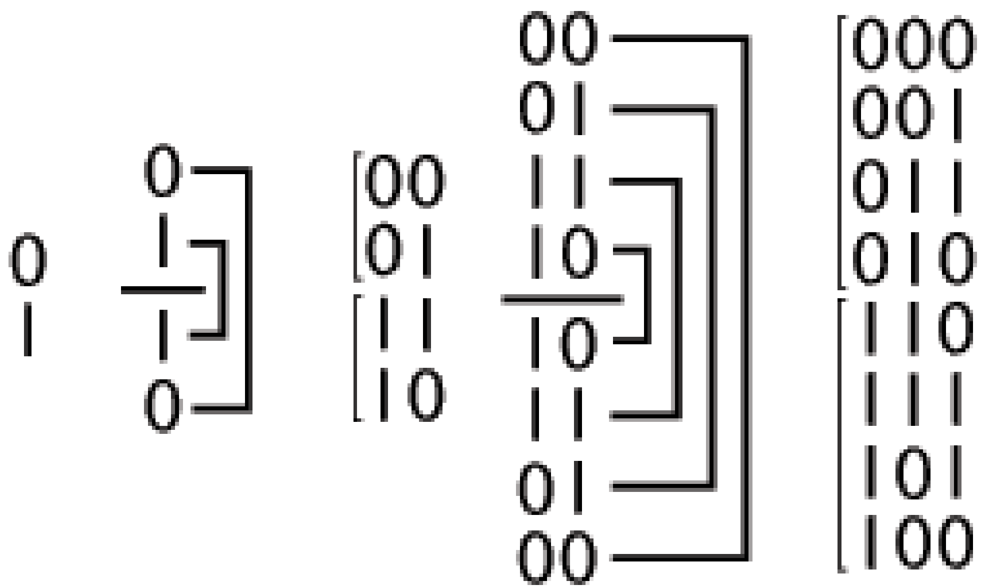
Grayev kod

- svojstva Grayevog koda:
  - susjedne kodne riječi
    ~ razlika u samo jednom bitu
    ("jedinična distanca")
  - izgradnja koda:
    ~ zrcaljenje u jednom bitovnom mjestu:
    reflektirani kod
  - netežinski kod
  - binarni, ali i "dekadski"
    ~ XS-3 Grayev kod

dekadski Grayev kod					binarni Grayev kod
	0	0	0	0	
	0	0	0	1	
	0	0	1	1	
0	0	0	1	0	
1	0	1	1	0	
2	0	1	1	1	
3	0	1	0	1	
4	0	1	0	0	
5	1	1	0	0	
6	1	1	0	1	
7	1	1	1	1	
8	1	1	1	0	
9	1	0	1	0	
	1	0	1	1	
	1	0	0	1	
	1	0	0	0	

Grayev kod

- izgradnja koda:
  ~ *zrcaljenje* u jednom bitovnom mjestu:
  reflektirani kod



Znakovni kodovi

- prikaz skupa znakova:
  - prikaz slova i znamenki:
    - "grafički"
      ~ "alfa-numerički" znakovi, interpunkcije, simboli, ...
    - *upravljački* znakovi
- standardizirani znakovni kodovi:
  npr. 7-bitni (128 kombinacija) ASCII:
  ISO IS 646, ITU-T/CCITT No. 5

Znakovni kodovi

- kod ASCII (engl. American Standard Code for Information Interchange):

  ' ': 20_H, CR : 08_H, LF : $0A_H$

  '0' – '9': $30\text{-}39_H$

  'A' - 'Z' : $41\text{-}5A_H$

  'a' – 'z' : $61\text{-}7A_H$

npr.

A = 100 0001 = 41_H

a = 110 0001 = 61_H

* = 010 1010 = $2A_H$

					0	0	0	0	1	1	1	1
					0	0	1	1	0	0	1	1
					0	1	0	1	0	1	0	1
					0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	3	P	3	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#/£	3	C	S	c	s
0	1	0	0	4	EOT	DC4	¤/$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	'	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF 1	SUB	*	:	J	Z	j	z
1	0	1	1	11	VT 1	ESC	+	;	K	3	k	3
1	1	0	0	12	FF 1	IS4	,	<	L	3	l	3
1	1	0	1	13	CR 1	IS3	–	=	M	3	m	3
1	1	1	0	14	SO	IS2	.	>	N	3	n	3
1	1	1	1	15	SI	IS1	/	?	O	_	o	DEL

Sadržaj predavanja

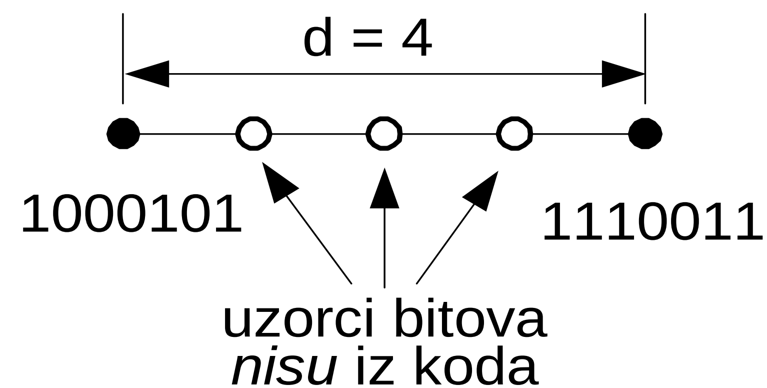
- binarno kodiranje znamenki i simbola
- **kodovi za zaštitu podataka**
  - **princip otkrivanja i ispravljanja pogrešaka, distanca i zalihost**
  - **paritet, jednostruko i višestruko ispitivanje pariteta**
  - **Hammingovi kodovi**

Kodovi za zaštitu podataka

- prijenos podataka
  ~ utjecaj smetnji: moguća pojava pogreške
  - pogreška
    ~ neželjena promjena jednog/više bitova u kodnoj riječi
    - jednostruka pogreška
      ~ promjena vrijednosti jednog bita
      ($0 \rightarrow 1$ ili $1 \rightarrow 0$)
    - višestruka pogreška ~ više bitova
  - rezultat
    ~ neispravna, ali i *ispravna* kodna riječ !
  - dobivena kodna riječ ispravna
    ~ *otkriti* da je došlo do pogreške!!!

Kodovi za zaštitu podataka

- princip otkrivanja (i ispravljanja) pogrešaka
  ~ razlika kodnih riječi u $\geq$ 1 bita
- distanca kodnih riječi (R. W. Hamming)
  ~ "udaljenost" dviju kodnih riječi:
  - najmanji broj bitova u kojima se dvije kodne riječi razlikuju
  - broj bitova koje treba promijeniti da se jedna kodna riječ pretvori u drugu
    ~ pogreška ostaje neotkrivena !!!



Kodovi za zaštitu podataka

- računanje distance kodnih riječi
  - broj različitih bitovnih mjesta dviju kodnih riječi:

    $c = a \oplus b$ po bitovima

    d = aritmetička suma "1" u c
  - formalno:

    $c = a \oplus b = (a_{n-1} \oplus b_{n-1}, a_{n-2} \oplus b_{n-2}, \ldots, a_0 \oplus b_0)$

    $d = |c| = |a \oplus b|$

    $|x|$: težina kodne riječi (engl. weight) ,
    broj jedinica u kodnoj riječi

Kodovi za zaštitu podataka

- minimalna distanca koda d_{min}
  ~ najmanji razmak između dvije kodne riječi
  - npr. kod 8421: $d_{min} = 1$
  - bikvinarni kod: $d_{min} = 2$
  - Grayev kod: $d_{min} = d = 1$
- kod pruža zaštitu od t pogrešaka

  $t=d_{min}-1$

  $d_{min} \geq (t+1)$

Primjer: $d_{min} = 2$
~ otkrivanje *jednostruke* pogreške

Kodovi za zaštitu podataka

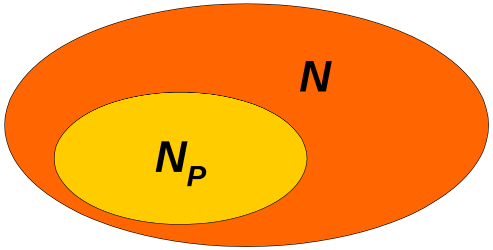
- kodovi s $d_{min} > 1$
  ~ postoji zalihost (redundancija), R:
  snaga zaštite, višak informacije

n : duljina kodne riječi
$k<n$: broj informacijskih bitova
$r=n-k$: broj zaštitnih bitova

$$R = \frac{r}{n} \qquad R = 1 - \frac{ldN_p}{ldN} \qquad (ldX = \log_2 X)$$

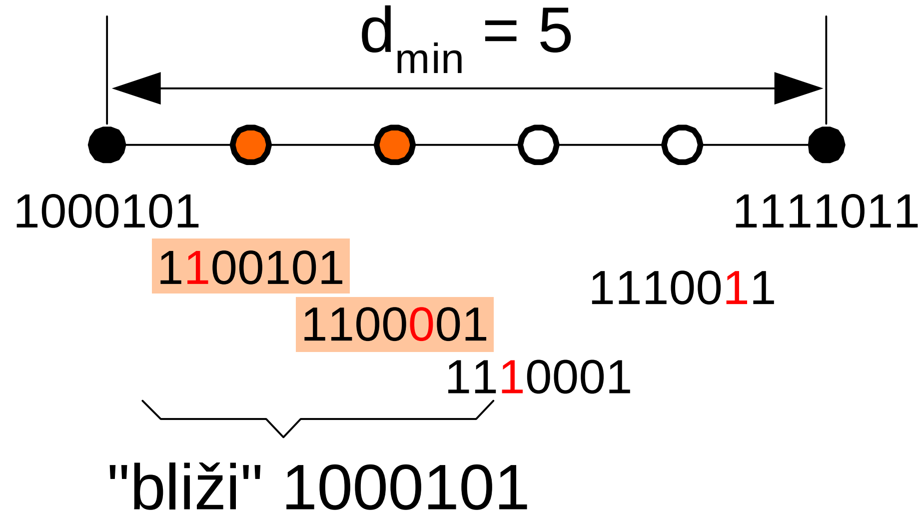
ukupni broj kodnih riječi $\quad N_p = 2^k < 2^n = N \quad$ ukupan broj mogućih kombinacija od n bitova

- veći broj bitova od minimalno potrebnih za prikaz informacije; npr. bikvinarni kod
- kodna riječ = bitovi + zaštitni bitovi
- sistematski kodovi
  ~ zaštitni bitovi nakon informacijskih



Kodovi za zaštitu podataka

- dvije skupine zaštitnih kodova:
  - s mogućnošću otkrivanja pogrešaka
    ~ EDC (engl. Error Detecting Codes):
    $d_{min} \geq t + 1$ za otkrivanje t pogrešaka
  - s mogućnošću ispravljanja pogrešaka
    ~ ECC (engl. Error Correcting Codes):
    $d_{min} \geq 2 \cdot t + 1$ za ispravljanje t pogrešaka



Kodovi za zaštitu podataka

- geometrijski prikaz kodnih riječi/koda ~ kubusi u n-dimenzijskom prostoru
  - 0-kubus ~ točka
  - 1-kubus ~ dužina
  - 2-kubus ~ kvadrat
  - 3-kubus ~ kocka
  - n-kubus ~ "hiperkocka"

Kodovi za zaštitu podataka

- geometrijski prikaz kodnih riječi/koda

  Primjer: n-kubus $\rightarrow$ 3-kubus

  1. za 2^n uzoraka: $d_{min} = 1$
  2. za {100, 011}: $d_{min} = 3$

     otkriva 2 pogreške:

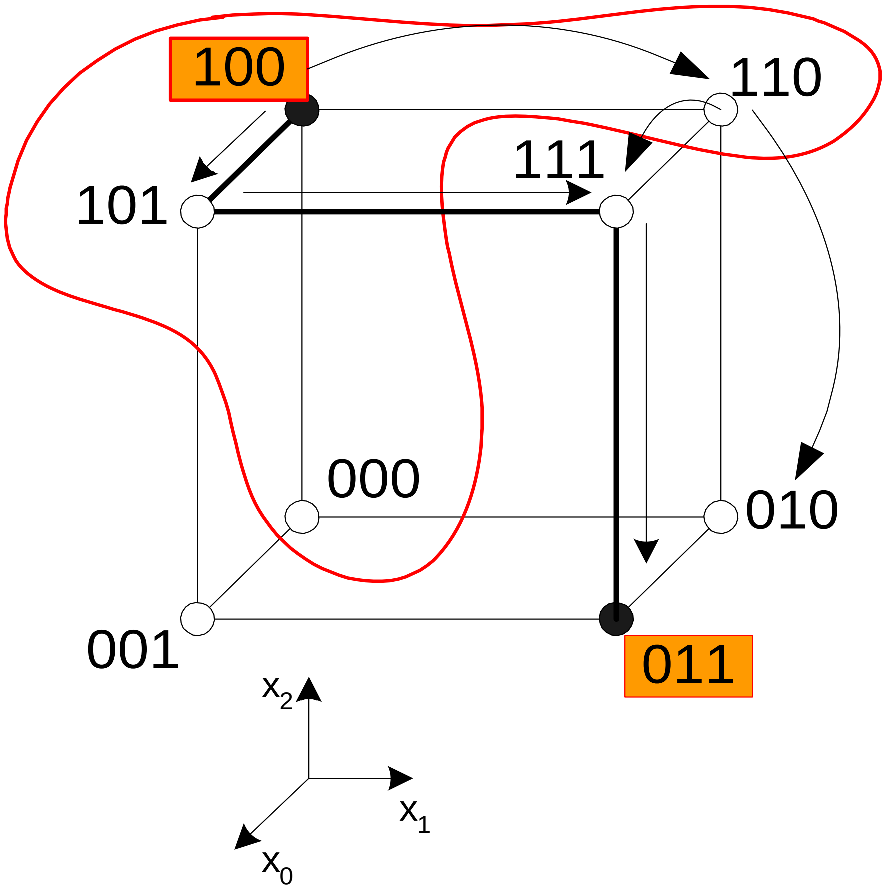
     010, 111, 001

     110, 101, 000

     ispravlja 1 pogrešku:

     110, 101, 000

     001, 111, 010



Paritet, jednostruko i višestruko ispitivanje

- paritet ~ najjednostavniji način zaštite
  - dodati paritetni bit
    ~ tipično osmi bit riječi iz ASCII koda:

    $\boldsymbol{p}\, b_6 b_5 b_4\ b_3 b_2 b_1 b_0$
  - nova kodna riječ mora imati paran/neparan broj jedinica
    ~ paran/neparan paritet

ZNAK		PARITET	
		PARNI	NEPARNI
A	100 0001	0 100 0001	1 100 0001
a	110 0001	1 110 0001	0 110 0001
*	010 1010	1 010 1010	0 010 1010

  - "vertikalna" (poprečna) paritetna zaštita
    (engl. Vertical Redundancy Check, VRC)
    ~ otkrivanje neparnih pogrešaka

Paritet, jednostruko i višestruko ispitivanje

- višestruko ispitivanje pariteta :
  - zahtjev: povećati moć zaštite!
    - veći broj paritetnih ispitivanja
      ~ "nezavisna" (ortogonalna)
    - veći broj zaštitnih bitova
      ~ veća zalihost
  - više mogućnosti:
    - dvodimenzijski kod
    - Hammingov kod

Paritet, jednostruko i višestruko ispitivanje

- dvodimenzijski kod
  ~ 2D matrica informacijskih bitova ("pravokutni" kod)
- uzdužna i poprečna paritetna zaštita:
  - kodna riječ ← paritetni bit
  - cijelom bloku kodnih riječi ←
    paritetna riječ, BCC (engl. Block Check Character)
    ~ "horizontalna" (uzdužna) paritetna zaštita
    (engl. Longitudinal Redundancy Check, LRC)
  - ispravljanje jednostruke pogreške

Hammingov kod

- sustavni mehanizam za izgradnju *niza* kodova za ispravljanje pogrešaka
  ~ R.W. Hamming, 1950.
- princip:
  ~ višestruko (nezavisno) paritetno ispitivanje
- bolja efikasnost kodiranja
  ~ manja zalihost (usp. dvodimenzijski kod)
- naročito popularni *Hammingov kod* za ispravljanje *jednostruke pogreške*
  ~ tipična primjena: memorijski sklopovi

Hammingov kod

- nezavisna paritetna ispitivanja
  ~ ne mogu se dobiti kombinacijom preostalih
- princip izgradnje kodne riječi:
  - "nezavisna" (ortogonalna) ispitivanja
  - "nezavisni" (ortogonalni) smještaj zaštitnih bitova
- "nezavisna" (ortogonalna) ispitivanja:
  - svaki zaštitni bit "pokriva" (= štiti) drugi podskup bitova podatka
  - svaki bit podatka zaštićen s više zaštitnih bitova

Hammingov kod

- odnos zaštitnih i informacijskih bitova:

  $2^r \geq k + r + 1\ ,\quad n = k + r$

  r: broj zaštitnih bitova
  k: broj informacijskih bitova
  n: duljina kodne riječi

  oznaka koda: (n, k)

- obrazloženje:
  - jednostruka pogreška na jednom od *n* mjesta
  - bez pogrešaka

Hammingov kod

- odnos zaštitnih i informacijskih bitova :
  $2^r \geq k + r + 1$, $n = k + r$

BROJ INFORMACIJSKIH BITOVA (≤)	BROJ ZAŠTITNIH BITOVA	DULJINA KODNE RIJEČI
1	2	3
4	3	7
11	4	15
26	5	31
57	6	63
120	7	127

Hammingov kod

- višestruka ispitivanja:
  - zaštitni bitovi na mjesta koja se ne mogu dobiti kombinacijama drugih zaštitnih bitova: 2^i
  - zaštitni bitovi "pokrivaju" svoju poziciju ~ sve pozicije čiji redni broj sadrži 2^i

POZICIJA	nema pogreške	1	2	3	4	5	6	7
C_2	0	0	0	0	1	1	1	1
C_1	0	0	1	1	0	0	1	1
C_0	0	1	0	1	0	1	0	1
		C_0	C_1	k_1	C_2	k_2	k_3	k_4

  - zaštitni bitovi: C_2 C_1 C_0

Hammingov kod

Primjer : kod (11,7)

- ukupno 11 bitova, od čega 7 nose podatke
  ~ korisno za zaštitu ASCII-znakova
- smještaj zaštitnih bitova

1	**2**	3	**4**	5	6	7	**8**	9	10	11
C1	**C2**	k1	**C3**	k2	k3	k4	**C4**	k5	k6	k7

- raspored "odgovornosti" bitova

	1	**2**	3	**4**	5	6	7	**8**	9	10	11
P1:	P1	P2		P3				P4			
P2:	P1	P2		P3				P4			
P3:	P1	P2		P3				P4			
P4:	P1	P2		P3				P4			

Hammingov kod

- izračunavanje zaštitnih bitova za *parni* paritet

POZICIJA	nema pogreške	1	2	3	4	5	6	7
C_2	0	0	0	0	1	1	1	1
C_1	0	0	1	1	0	0	1	1
C_0	0	1	0	1	0	1	0	1
		C_0	C_1	k_1	C_2	k_2	k_3	k_4

$C_0 \oplus k_1 \oplus k_2 \oplus k_4 = 0 \rightarrow C_0 = k_1 \oplus k_2 \oplus k_4$

$C_1 \oplus k_1 \oplus k_3 \oplus k_4 = 0 \rightarrow C_1 = k_1 \oplus k_3 \oplus k_4$

$C_2 \oplus k_2 \oplus k_3 \oplus k_4 = 0 \rightarrow C_2 = k_2 \oplus k_3 \oplus k_4$

Hammingov kod

- izračunavanje zaštitnih bitova za *neparni* paritet

POZICIJA	nema pogreške	1	2	3	4	5	6	7
C_2	0	0	0	0	1	1	1	1
C_1	0	0	1	1	0	0	1	1
C_0	0	1	0	1	0	1	0	1
		C_0	C_1	k_1	C_2	k_2	k_3	k_4

$$C_0 \oplus k_1 \oplus k_2 \oplus k_4 \oplus 1 = 0 \quad \rightarrow \quad C_0 = k_1 \oplus k_2 \oplus k_4 \oplus 1$$

$$C_1 \oplus k_1 \oplus k_3 \oplus k_4 \oplus 1 = 0 \quad \rightarrow \quad C_1 = k_1 \oplus k_3 \oplus k_4 \oplus 1$$

$$C_2 \oplus k_2 \oplus k_3 \oplus k_4 \oplus 1 = 0 \quad \rightarrow \quad C_2 = k_2 \oplus k_3 \oplus k_4 \oplus 1$$

Hammingov kod

Primjer: zaštita ASCII znaka A (41_H)

$$n = 11,\ k = n - r = 7 \Rightarrow r = 4$$

1	2	3	4	5	6	7	8	9	10	11
C_0	C_1	k_1	C_2	k_2	k_3	k_4	C_3	k_5	k_6	k_7

k_1	k_2	k_3	k_4	k_5	k_6	k_7
1	0	0	0	0	0	1

$C_0 = k_1 \oplus k_2 \oplus k_4 \oplus k_5 \oplus k_7 \rightarrow C_0 = 0$

$C_1 = k_1 \oplus k_3 \oplus k_4 \oplus k_6 \oplus k_7 \rightarrow C_1 = 0$

$C_2 = k_2 \oplus k_3 \oplus k_4 \rightarrow C_2 = 0$

$C_3 = k_5 \oplus k_6 \oplus k_7 \rightarrow C_3 = 1$

$\Rightarrow$ 0 0 1 0 0 0 0 1 0 0 1

		k_1		k_2	k_3	k_4		k_5	k_6	k_7
0	1	1	0	0	0	0	1	0	0	1
C_0	C_1		C_2				C_3			

$X = C_3\ C_2\ C_1\ C_0 = 1\ 0\ 1\ 0$

$Y = C_3'\ C_2'\ C_1'\ C_0' = 1\ 0\ 0\ 0$

mjesto pogreške: $X \oplus Y = 0010_2 = 2_{10}$

Hammingov kod

- Hammingov kod za ispravljanje jednostruke pogreške:
  - distanca d = 3
  - kod za ispravljanje "nezavisnih pogrešaka"
    ~ rezultat djelovanja "bijelog šuma"
  - efikasan kod, jer je R mali

Literatura

U. Peruško, V. Glavinić: *Digitalni sustavi*, Poglavlje 2: Digitalni podaci: tipovi, operacije, algoritmi

- binarno kodiranje znamenki i simbola: str. 57-64
- kodovi za zaštitu podataka: str. 64-75