

Architektura systemów komputerowych

Laboratorium 5

Kodowanie liczb i tekstów

Marcin Stępnia

1 Informacje

1.1 Kod NKB

Naturalny kod binarny (NKB) jest oparty na zapisie liczby naturalnej w dwójkowym systemie pozycyjnym. Inaczej mówiąc jest to zwykła liczba binarna bez znaku i bez przecinków.

Dla n bitów możemy utworzyć w naturalnym kodzie binarnym liczby z zakresu: $< 0, 2^n - 1 >$

Dla n bitów mamy więc 2^n różnych wartości kodu.

Do zapisu liczby naturalnej x potrzebne jest co najmniej $\lfloor \log_2 x \rfloor + 1$ bitów, gdzie $\lfloor a \rfloor$ jest funkcją podłogi (ang. floor), zaokrąglającą liczbę a w dół.

1.2 Kod BCD

BCD (ang. Binary-Coded Decimal - system dziesiętny zakodowany dwójkowo) to sposób zapisu liczby w systemie dziesiętnym polegający na zakodowaniu kolejnych cyfr tej liczby w systemie dwójkowym, przy użyciu tylko czterech bitów. Taki zapis pozwala na łatwą konwersję liczby pomiędzy kodem BCD a systemem dziesiętnym, ale jego wadą jest istniejąca nadmiarowość (wykorzystuje tylko 10 czterobitowych kombinacji z 16 możliwych). Istnieje kilka wariantów zapisu liczb w kodzie BCD. Najpopularniejszy jest BCD

8421, w którym kody BCD odpowiadają wartości cyfry w systemie dwójkowym (zob. tab. 1).

Tabela 1: Tabela kodów BCD 8421

Cyfra dziesiętna	BCD
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001

Przykład liczby w kodzie BCD:

9 2 7
1001 0010 0111

$$927_{(10)} = 100100100111_{(BCD)}$$

1.3 Kod Gray'a

W kodzie Gray'a kolejne wyrazy różnią się między sobą wartością tylko jednego bitu. Jest również kodem cyklicznym, bowiem ostatni i pierwszy wyraz tego kodu także spełniają tę zasadę.

Przy konwersji z i na kod Gray'a wykorzystuje się operację sumy modulo 2, której tablica prawdy została przedstawiona w tabeli 2.

Metoda zamiany liczby binarnej na kod Gray'a jest następująca: Najbardziej znaczący bit (MSB - ang. most significant bit) w kodzie Gray'a jest pobierany bezpośrednio z MSB liczby binarnej. Reszta bitów w kodzie Gray'a jest wynikiem operacji XOR (suma modulo 2, oznaczana symbolem

Tabela 2: Tablica prawdy dla alternatywy wykluczającej (sumy modulo 2)

p	q	$p \oplus q$
0	0	0
0	1	1
1	0	1
1	1	0

\oplus) pomiędzy poprzednim bitem binarnym ($b(i-1)$) i bitem liczby binarnej na tej samej pozycji ($b(i)$). W przypadku przedstawionym na rysunku 1 mamy:

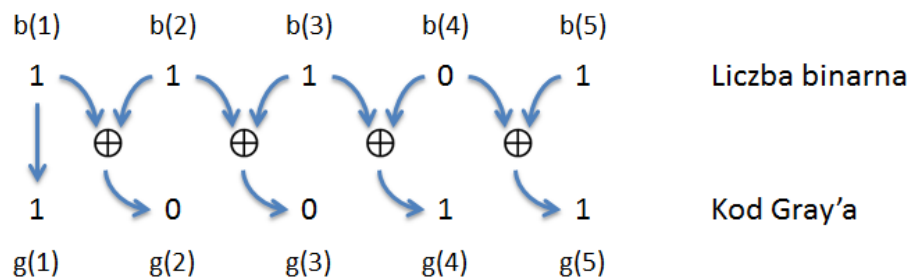
$$g(1) = b(1)$$

$$g(2) = b(1) \oplus b(2)$$

$$g(3) = b(2) \oplus b(3)$$

$$g(4) = b(3) \oplus b(4)$$

$$g(5) = b(4) \oplus b(5)$$



Rysunek 1: Zamiana liczby binarnej na kod Gray'a

Metoda zamiany liczby w kodzie Gray'a na liczbę binarną jest następująca: Najbardziej znaczący bit (MSB) liczby binarnej jest pobierany bezpośrednio z MSB w kodzie Gray'a. Reszta bitów liczby binarnej jest wynikiem operacji XOR pomiędzy poprzednim bitem binarnym ($b(i-1)$) i bitem w kodzie Gray'a na tej samej pozycji ($g(i)$). W przypadku przedstawionym na rysunku 2 mamy:

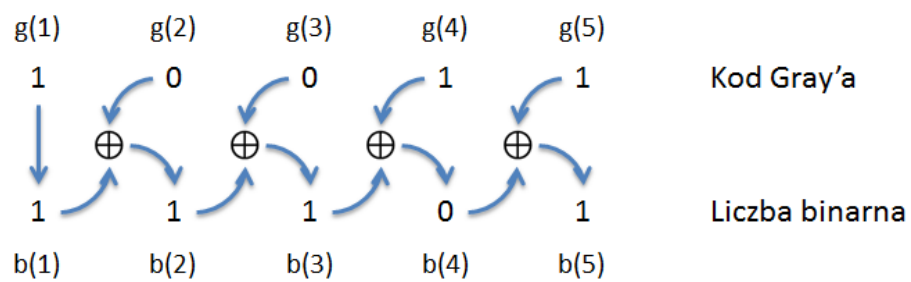
$$b(1) = g(1)$$

$$b(2) = b(1) \oplus g(2)$$

$$b(3) = b(2) \oplus g(3)$$

$$b(4) = b(3) \oplus g(4)$$

$$b(5) = b(4) \oplus g(5)$$



Rysunek 2: Zamiana liczby w kodzie Gray'a na liczbę binarną

Tabela 3: Tabela kodów Gray'a

Lp.	Kod binarny	Kod Gray'a
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000

1.4 Kodowanie znaków

Cyfrowy system komputerowy operuje tylko zerami i jedynekami. Aby umożliwić wprowadzanie do komputera takich informacji jak słowa i liczby, konieczne jest zastosowanie kodu, który każdemu znakowi przyporządkowuje określoną kombinację zer i jedynek.

Obecnie istotne znaczenie ma kod ASCII (czyt. aski; ang. American Standard Code for Information Interchange). Jest to 7-bitowy kod przyporządkowujący liczby z zakresu 0-127: literom alfabetu angielskiego, cyfrom, znakom przestankowym i innym symbolom oraz poleceniom sterującym. Na przykład mała litera „a” jest kodowana jako liczba 97, a znak spacji jest kodowany jako 32 (zob. tabelę na końcu dokumentu).

Ze względu na ograniczenie kodu ASCII do 128 znaków i faktu, że większość komputerów operuje na 8-bitowych bajtach, dodatkowy bit można wykorzystać na powiększenie zbioru kodowanych znaków do 256 symboli. Powstało wiele różnych rozszerzeń ASCII wykorzystujących ósmy bit (np. norma ISO 8859, rozszerzenia firm IBM lub Microsoft) nazywanych stronami kodowymi.

Przed rozpowszechnieniem kodowań opartych na zestawie Unicode, najpopularniejszymi standardami kodowania polskich znaków były ISO 8859-2 (np. dla stron internetowych) i Windows-1250 (w systemie Windows).

Tabelę z kodowanie polskich znaków w różnych standardach można znaleźć pod adresem:

https://pl.wikipedia.org/wiki/Kodowanie_polskich_znak%C3%B3w

Więcej informacji można znaleźć w książce Barczak A., Florek J., Sydoruk T.: Elektroniczne techniki cyfrowe, Wyd.: VIZJA PRESS&IT Sp. z o.o, Warszawa 2006 (Rozdział 2 dostępny pod adresem: <http://mirek.ii.uph.edu.pl/ask/files/rozd2.pdf>)

2 Zadania

2.1 Kod NKB, BCD i Gray'a

2.1.1 Przedstaw w kodzie NKB, BCD i Gray'a:

1. 45
2. 456
3. 21

2.1.2 Przedstaw w kodzie NKB, poniższe liczby w kodzie Gray'a:

1. 11011_(Gray)
2. 101011_(Gray)
3. 1010_(Gray)

2.1.3 Jaką wartość dziesiętną mają następujące liczby w kodzie BCD:

1. 111000_(BCD)
2. 100100111000_(BCD)
3. 1111100_(BCD)

2.1.4 Oblicz iloma bitami można przedstawić liczbę dziesiętną 356:

1. w kodzie NKB
2. w kodzie BCD

2.2 Kodowanie znaków

2.2.1 Przedstaw tekst w postaci kodu ASCII (do kodowania polskich liter wykorzystaj ISO 8859-2 i Windows-1250):

1. "Gżegżółka"

2. "Śledź"

3. swoje nazwisko

ASCII CONTROL CODE CHART

b7 b6 BITS b5 b4 b3 b2 b1	0	0	0	0	1	1	1	1
	0	0	1	1	0	0	1	1
	0 1 CONTROL		0 1 SYMBOLS NUMBERS		0 1 UPPER CASE		0 1 LOWER CASE	
0 0 0 0	0 NUL	16 DLE	32 SP	48 0	64 @	80 P	96 ‘	112 p
0 0 0 1	1 SOH	17 DC1	33 !	49 1	65 A	81 Q	97 a	113 q
0 0 1 0	2 STX	18 DC2	34 ”	50 2	66 B	82 R	98 b	114 r
0 0 1 1	3 ETX	19 DC3	35 #	51 3	67 C	83 S	99 c	115 s
0 1 0 0	4 EOT	20 DC4	36 \$	52 4	68 D	84 T	100 d	116 t
0 1 0 1	5 ENQ	21 NAK	37 %	53 5	69 E	85 U	101 e	117 u
0 1 1 0	6 ACK	22 SYN	38 &	54 6	70 F	86 V	102 f	118 v
0 1 1 1	7 BEL	23 ETB	39 ’	55 7	71 G	87 W	103 g	119 w
1 0 0 0	8 BS	24 CAN	40 (56 8	72 H	88 X	104 h	120 x
1 0 0 1	9 HT	25 EM	41)	57 9	73 I	89 Y	105 i	121 y
1 0 1 0	10 LF	26 SUB	42 *	58 :	74 J	90 Z	106 j	122 z
1 0 1 1	11 VT	27 ESC	43 +	59 ;	75 K	91 [107 k	123 {
1 1 0 0	12 FF	28 FS	44 ,	60 <	76 L	92 \	108 l	124
1 1 0 1	13 CR	29 GS	45 —	61 =	77 M	93]	109 m	125 }
1 1 1 0	14 SO	30 RS	46 .	62 >	78 N	94 ^	110 n	126 ~
1 1 1 1	15 SI	31 US	47 /	63 ?	79 O	95 ”	111 o	127 DEL

LEGEND:

dec	CHAR	
hex		oct

Victor Eijkhout
Dept. of Comp. Sci.
University of Tennessee
Knoxville TN 37996, USA