

Technologia informacyjna

Kodowanie danych

1

Kodowanie danych

- Kodowanie danych to zamiana jednej formy informacji na inną, zwykle odpowiednią dla danego urządzenia przetwarzającego dane.
- Kodem nazywa się wzajemnie jednoznaczne odwzorowanie , które każdej wiadomości z alfabetu źródła przyporządkowuje ciąg określonych symboli kodowych.

2

Kodowanie binarne

- Kod binarny, to kod wykorzystujący sygnały dwuwartościowe (0 i 1).
- System dwójkowy jest naturalnym systemem informatyki.
- 8 bitów = bajt jest podstawową jednostką informacji;
- za pomocą 8 bitów można zapisać 256 różnych wartości.
- Za pomocą ciągów binarnych można zapisać: teksty, liczby, muzykę, obrazy, filmy.

3

Kompresja

- Kompresja, to działanie, które pozwala na zmniejszenie objętości pliku wynikowego lub strumienia danych.
- Wyróżnia się:
 - Kompresję bezstratną – odtworzona informacja jest identyczna z oryginałem,
 - Kompresję stratną – polega na eliminowaniu pewnych elementów oryginału w celu uzyskania lepszej efektywności kompresji.

4

Przykłady kodowania

- Kodowanie tekstu – kod ASCII
- Kodowanie liczb całkowitych i rzeczywistych;
- Kodowanie grafiki
 - Grafika rastrowa:
 - obraz złożony z kropek (pikseli) – bitmapa,
 - barwa każdego piksela kodowana na określonej liczbie bitów.
 - Grafika wektorowa:
 - Obraz złożony z wektorów (kodowany początek, koniec, barwa),
 - Figury geometryczne (typ, położenie, kolor) – np. okrąg (współrzędne środka, promień, kolor),
 - Grafikę wektorową można przeskalować bez straty jakości.

5

Przykłady kodowania

- Kodowanie dźwięku:
 - Próbkowanie;
 - Kwantyzacja;
 - Kodowanie;
 - Kompresja;
- Formaty plików:
 - mp3
 - wav
 - mp4

6

Podstawowe rodzaje kodów

Kody nienadmiarowe	
Kody równomierne	Kody nierównomierne
Wszystkie ciągi kodowe mają jednakową długość	Ciągi kodowe mają różne długości

7

Podstawowe rodzaje kodów

Kody nadmiarowe	
Kody detekcyjne	Kody korekcyjne
Możliwe jest wykrywanie przekłamań	Możliwa jest korekcja wykrytych przekłamań

8

Kodowanie tekstu

- W latach sześćdziesiątych XX wieku opracowano standard ASCII.
- Międzynarodowy zestaw znaków zawiera 128 liczb i odpowiadające im znaki.
- W rozszerzonym kodzie ASCII znajdują się litery specyficzne dla poszczególnych języków.
- Tekst w kodzie ASCII nie zawiera informacji o formatowaniu czcionki.
- UNICODE - 65536 znaków

9

Kodowanie liczb

Na ogół mamy do czynienia z systemami pozycyjnymi:

– Dziesiętny

$$329 = 3 \cdot 10^2 + 2 \cdot 10^1 + 9 \cdot 10^0$$

$$0,1941 = 1 \cdot 10^{-1} + 9 \cdot 10^{-2} + 4 \cdot 10^{-3} + 1 \cdot 10^{-4}$$

– Dwójkowy

$$10_{10} = 1010_2$$

$$1010_2 = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 10_{10}$$

11

Kodowanie liczb ujemnych

- Kodowanie w systemie znak-moduł
- Kodowanie w systemie uzupełnieniowym

12

Kodowanie liczb ujemnych

Metoda znak-moduł

znak	moduł			
a_{n-1}	a_{n-2}	\dots	a_1	a_0

$$a_{n-1} a_{n-2} \dots a_2 a_1 a_0 = (1 - 2 * a_{n-1}) * \sum_{i=0}^{n-2} a_i 2^i$$

13

Kodowanie liczb ujemnych

Metoda znak-moduł – przykład

$$0111_{(Z-M)} = 0 \cdot 1_2 \cdot 1_1 \cdot 1_0 = (1-2^*0)^* (1^*2^2 + 1^*2^1 + 1^*2^0) = 1^* (4+2+1) = 7_D$$

$$1111_{(Z-M)} = 1 \cdot 1_2 \cdot 1_1 \cdot 1_0 = (1-2^*1)^* (1^*2^2 + 1^*2^1 + 1^*2^0) = -1^* (4+2+1) = -7_D$$

14

Kodowanie liczb ujemnych

Metoda uzupełnień do 2 (U2)

$$a_{n-1} a_{n-2} \dots a_2 a_1 a_0 = a_{n-1} * (-2^{n-1}) + \sum_{i=0}^{n-2} a_i 2^i$$

$$0111_B = 0_3 \cdot 1_2 \cdot 1_1 \cdot 1_0 = 0^* (-2^3) + 1^* (2^2) + 1^* (2^1) + 1^* (2^0) = 4+2+1 = 7_D$$

$$1111_B = 1_3 \cdot 1_2 \cdot 1_1 \cdot 1_0 = 1^* (-2^3) + 1^* (2^2) + 1^* (2^1) + 1^* (2^0) = -8+4+2+1 = -1_D$$

15

Kodowanie liczb rzeczywistych

- Kodowanie w systemie zmiennoprzecinkowym
(ang. Floating-Point Numbers)
- Kodowanie w systemie stałoprzecinkowym
(ang. Fixed-Point Numbers)

16

Kodowanie w systemie zmiennoprzecinkowym

- Umożliwia zapis liczb rzeczywistych z ustalonym błędem względnym;
- Oparty jest na podziale liczby na część ułamkową (mantysa) oraz na wykładnik potęgi (cecha);
- Opracowany jest na podstawie zapisu liczby w systemie pozycyjnym wagowym.

$$L = S * M * B^E$$

17

Kodowanie w systemie stałoprzecinkowym

część całkowita	część ułamkowa
10110011,	0101

$$a_{n-1} \dots a_1 a_0, a_{-1} \dots a_{-m} = a_{n-1} \cdot 2^{n-1} + \dots + a_1 \cdot 2^1 + a_0 \cdot 2^0 + a_{-1} \cdot 2^{-1} + \dots + a_{-m} \cdot 2^{-m}$$

$$1101,11_B = 1_3 1_2 0_1 1_0, 1_{-1} 1_{-2} = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 + 1 \cdot 2^{-1} + 1 \cdot 2^{-2} = 8 + 4 + 1 + \frac{1}{2} + \frac{1}{4} = 13,75_D$$

18

Kodowanie w systemie zmiennoprzecinkowym

$$l_{FP} = m \cdot b^e$$

cecha				mantysa			
b_7	b_6	b_5	b_4	b_3	b_2	b_1	b_0

$$e = b_7(-2^3) + b_6 2^2 + b_5 2^1 + b_4 2^0 = (-8)b_7 + 4b_6 + 2b_5 + b_4$$

$$m = b_3 b_2, b_1 b_0 = b_3(-2^1) + b_2 2^0 + b_1 2^{-1} + b_0 2^{-2} = -2b_3 + b_2 + \frac{1}{2}b_1 + \frac{1}{4}b_0$$

$$l_{FP} = m \cdot 2^e$$

19

Kodowanie w systemie zmiennoprzecinkowym - przykład

$$\begin{array}{ccc}
 & 1111 & 1001_{FP} \\
 & | & \\
 e = 1111_{U_2} & & m = 10,01_{U_2} \\
 1111_{U_2} = -8+4+2+1 = -1_D & & 10,01_{U_2} = -2+\frac{1}{4} = -1,75_D \\
 l_{FP} = m \cdot 2^e = -1\frac{3}{4} \cdot 2^{-1} = -\frac{7}{4} \cdot \frac{1}{2} = -0,875 \\
 \\
 & 11111001_{FP} = -0,875_D
 \end{array}$$