

МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Кафедра програмування

Практичне завдання № 8
СТАТИСТИЧНІ АЛГОРИТМИ СТИСНЕННЯ ІНФОРМАЦІЇ
з курсу "Теорія інформації"

Виконала:
студентка групи ПМІ-25
Ващук Ярина

Варіант 6

Оцінка

Прийняв:
доц. Рикалюк Р.Є.
ас. Жировецький В.В.

Львів 2021

Мета роботи: навчитись проводити стиснення тексту, використовуючи двопрхідний алгоритм Хаффмана, визначати довжину кодової послідовності на виході кодера та коефіцієнт стиснення, декодувати послідовність знаючи, що вона закодована (стиснена) алгоритмом Хаффмана з передачею структури кодового дерева.

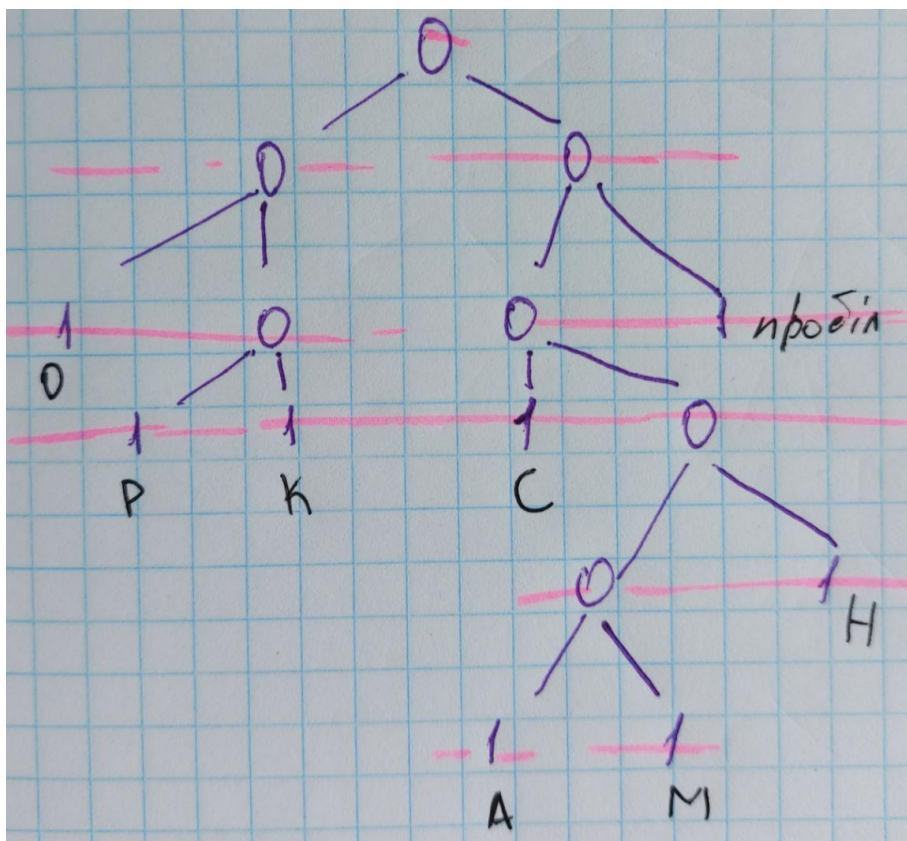
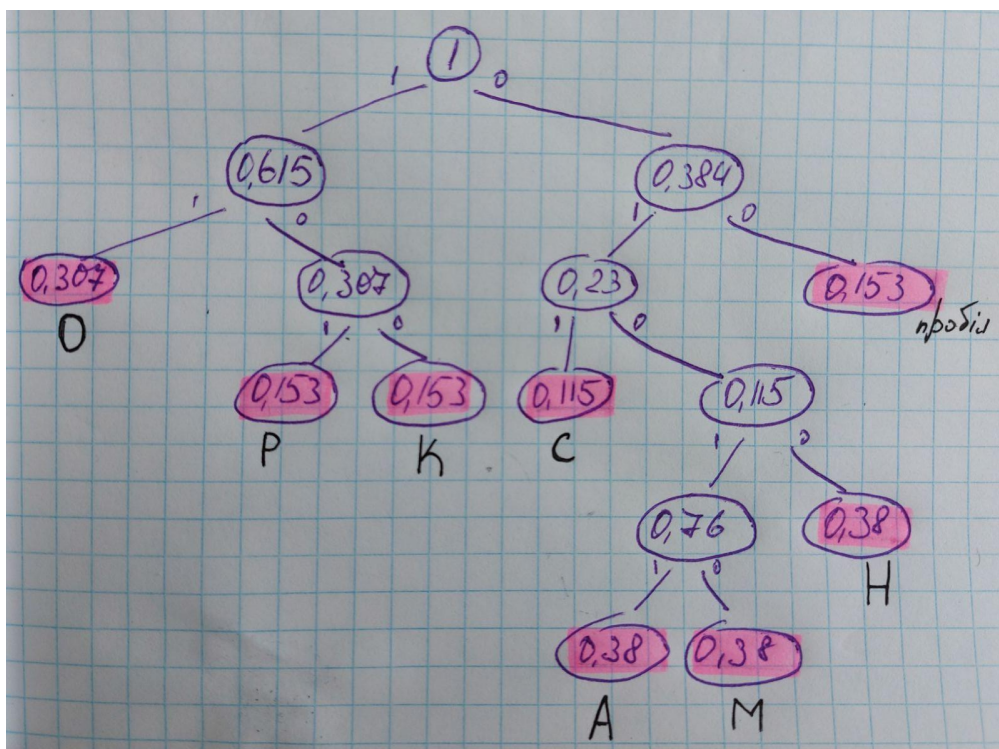
Хід роботи

1.

- а. Провести стиснення тексту наведеного у табл. 8.1, використовуючи двопрхідний алгоритм Хаффмана. (Записати вихідне кодове слово, яке складається з двох частин – $\text{Code}(X) = \text{Code}_1(X) + \text{Code}_2(X)$. Тут $\text{Code}_1(X)$, містить інформацію про використаний (застосований) код – у вигляді структури кодового дерева і алфавіту, друга, $\text{Code}_2(X)$, – власне закодовану (стиснену) послідовність).

НА СОРОК СОРОК СОРОК МОРОК

Побудуємо дерево Хаффмана:



Закодоване дерево: 0 00 1001 1110 01 11

Коди ASCII символів:

0	11 001 110
---	------------

пробіл	00 100 000
Р	11 010 000
К	11 001 010
С	11 010 001
Н	11 001 101
А	11 000 000
М	11 001 100

Code(X)₁ = 0 00 1001 1110 01 11 11001110 00100000 11010000 11001010
11010001 11001101 11000000 11001100

Символ	Частота /26	Ймовірність	Код	L	p _i L _i
О	8	0.307692308	11	2	0.615384616
пробіл	4	0.153846154	00	2	0.307692308
Р	4	0.153846154	101	3	0.461538462
К	4	0.153846154	100	3	0.461538462
С	3	0.115384615	011	3	0.346153845
Н	1	0.038461538	0100	4	0.153846152
А	1	0.038461538	01011	5	0.192307691
М	1	0.038461538	01010	5	0.192307691

Закодоване речення кодом Хаффмана

Code₂ = 0100 01011 00 011 11 101 11 100 00 011 11 101 11 100 00 011
11 101 11 100 00 01010 11 101 11 100
Code = 0 00 1001 1110 01 11 11001110 00100000 11010000 11001010
11010001 11001101 11000000 11001100 0100 01011 00 011 11 101 11 100 00
011 11 101 11 100 00 011 11 101 11 100 00 01010 11 101 11 100

- б. Визначити довжину кодової послідовності на виході кодера та коефіцієнт стиснення.

$$l_1 = 15 + 8 * 8 = 79 \text{ біт.}$$

$$l = l_1 + l_2 = 79 + 27 = 106 \text{ біт.}$$

$$\text{Коефіцієнт стиснення: } r = 15 * 8 / 106 = 1.132$$

2. Декодувати послідовність

```
0000100110100111100110000000000111100001111111111110010000011
001101110100001100010011000000110011101011001011000101110011
001100100111000001110100111100001011001000110100100010110000
101110101010101100011111000110110101011100111111001010110100
011100101111011100000100000110011010111100110011101011111001
1111011111110111101010100010000011001101011110011001110101111
1011111110101010001000001100110101111001100111010111110000010
00111111011110000110111111101111100111110110111011110110010
0011001111000101111010010010100011010111110001111010011101111
1001000011101111100101111100110111111111110100111100111011110
0110011101010001000001100110111110110
```

знаючи, що вона закодована (стиснена) алгоритмом Хаффмана з передачею структури кодового дерева. (Побудувати дерево, таблицю кодів і саму розпаковану послідовність, використовуючи таблицю ASCII кодів).

Структура дерева: *000010011010011110011000000000011110000111111111111*

Код ASCII:

00 100 000 пробіл	11 001 000 И
11 001 101 Н	11 010 010 Т
11 010 000 Р	00 101 100 ,
11 000 100 Д	00 101 110 .
11 000 000 А	10 101 010 Є
11 001 110 О	11 000 111 З
10 110 010 І	11 000 110 Ж
11 000 101 Е	11 010 101 Х
11 001 100 М	11 001 111 П
11 001 001 Й	11 001 010 К
11 000 001 Б	11 010 001 С
11 010 011 У	11 001 011 Л
11 000 010 В	11 011 100 Ъ

У	1111101	7
В	1111110	7
И	1111111	7
Т	11110100	8
,	11110101	8
.	11110110	8
Є	11110111	8
З	111100000	9
Ж	111100001	9
Х	111100010	9
П	111100011	9
К	111100100	9
С	111100101	9
Л	111100110	9
Ь	111100111	9

1101 Е 1111110 В 11110101, 01 _ 000 Н 1000 А 001 Р 1001 О 101 Д 01 _
11100 М 1100 І 11101 Ї 01 _ 11110111 Є 11110101 , 01 _ 000 Н 1000 А 001 Р
1001 О 101 Д 01 _ 11100 М 1100 І 11101 Ї 01 _ 111100000 З 1000 А 1111110
В 111100001 Ж 101 Д 1111111 И 01 _ 1111100 Б 1111101 У 101 Д 1101 Е
11110110 . 01 _ 000 Н 1100 І 111100010 Х 11110100 Т 1001 О 01 _ 000 Н
1101 Е 01 _ 111100011 П 1101 Е 001 Р 1101 Е 111100100 К 001 Р 1101 Е
111100101 С 111100110 Л 1111111 И 11110100 Т 111100111 Ь 01 _ 11100 М
1100 І 11101 Ї 01 _ 000 Н 1000 А 001 Р 1001 О 101 Д 11110110 .

ЕВ, НАРОД МІЙ Є, НАРОД МІЙ ЗАВЖДИ БУДЕ. НІХТО НЕ
ПЕРЕКРЕСЛИТЬ МІЙ НАРОД.

Висновок: в ході виконання роботи я навчилась проводити стиснення
тексту, використовуючи двопрхідний алгоритм Хаффмана, визначати
довжину кодової послідовності на виході кодера та коефіцієнт стиснення,
декодувати послідовність знаючи, що вона закодована (стиснена)
алгоритмом Хаффмана з передачею структури кодового дерева.