

komunikat do zakodowania:

a	a	b	a	b	a	b	a	b	a	c	c	a	b	a	a	a	c	a	c	b	c	b	c	a	a	c	b	a			
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31

przyjmujemy długość bufora słownikowego równą długości bufora wejściowego równą 4.

0	1	2	3	0	1	2	3

pierwsze cztery pozycje należą do bufora słownikowego, kolejne cztery należą do bufora wejściowego.

pobieramy pierwszą literę komunikatu i zapisujemy ją na wyjście.

pobieramy pierwszą literę komunikatu i wypełniamy nią (wszystkie pozycje tą samą literą) bufor słownikowy.

a	a	a	a				
0 (0)	1 (0)	2 (0)	3 (0)	0	1	2	3

do bufora wejściowego wpisujemy pierwsze cztery litery komunikatu.

a	a	a	a	a	a	b	a
0 (0)	1 (0)	2 (0)	3 (0)	0 (0)	1 (1)	2 (2)	3 (3)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

a	a	a	a	a	a	b	a
0 (0)	1 (0)	2 (0)	3 (0)	0 (0)	1 (1)	2 (2)	3 (3)

zauważmy, że podciąg **aa** można znaleźć w buforze słownikowym na 3 (a nawet 4) sposoby!

kodujemy znaleziony podciąg w następujący sposób: **(1, 2, b)** i zapisujemy kod na wyjście.

pierwszy element trójki "mówi" w którym miejscu w buforze słownikowym zaczyna się szukany podciąg.

drugi element trójki "mówi" jak długi jest szukany podciąg.

trzeci element trójki "mówi" jaka litera następuje bezpośrednio po szukanym podciągu w buforze wejściowym.

przesuwamy elementy w buforach w lewo o drugi element trójki powiększony o jeden (przesuwamy o 2+1).

a	a	a	b	a		
0 (0)	1 (0)	2 (1)	3 (2)	0 (3)	1	2 3

uzupełniamy powstałe puste pozycje w buforze wejściowym kolejnymi literami komunikatu.

a	a	a	b	a	b	b	a
0 (0)	1 (0)	2 (1)	3 (2)	0 (3)	1 (4)	2 (5)	3 (6)

w buforze słownikowym szukamy najdłuższego podciągu równego możliwie długiemu prefiksowi bufora wejściowego.

a	a	a	b	a	b	b	a
0 (0)	1 (0)	2 (1)	3 (2)	0 (3)	1 (4)	2 (5)	3 (6)

kodujemy znaleziony podciąg w następujący sposób: **(2, 2, b)** i zapisujemy kod na wyjście.

przesuwamy elementy w buforach w lewo o 2+1 i uzupełniamy bufor wejściowy.

b	a	b	b	a	b	a	c
0 (2)	1 (3)	2 (4)	3 (5)	0 (6)	1 (7)	2 (8)	3 (9)

w buforze słownikowym szukamy najdłuższego podciągu równego możliwie długiemu prefiksowi bufora wejściowego.

b	a	b	b	a	b	a	c
0 (2)	1 (3)	2 (4)	3 (5)	0 (6)	1 (7)	2 (8)	3 (9)

kodujemy znaleziony podciąg w następujący sposób: **(1, 2, a)** i zapisujemy kod na wyjście.

przesuwamy elementy w buforach w lewo o 2+1 i uzupełniamy bufor wejściowy.

b	a	b	a	c	c	a	b
0 (6)	1 (6)	2 (7)	3 (8)	0 (9)	1 (10)	2 (11)	3 (12)

w buforze słownikowym szukamy najdłuższego podciągu równego możliwie długiemu prefiksowi bufora wejściowego. nie znaleźliśmy żadnego podciągu w buforze słownikowym pasującego do jakiegokolwiek prefiksu w buforze wejściowym! w takiej sytuacji kodujemy ciąg o zerowej długości w następujący sposób: **(1, 0, c)** i zapisujemy kod na wyjście. zwróćmy uwagę, że pierwszy element trójki może być dowolny (jest to pewna nadmiarowość rozwiązana w kolejnych metodach LZ). istotny jest drugi element trójki, który mówi, że podciąg ma zerową długość. trzeci element trójki zapisujemy standardowo, czyli podajemy literę następującą bezpośrednio po podciągu (tutaj o zerowej długości), czyli **c**. przesuwamy (jak zwykle) elementy w buforach o drugi element trójki powiększony o jeden, czyli o $0+1$.

a	b	a	c	c	a	b	a
0 (6)	1 (7)	2 (8)	3 (9)	0 (10)	1 (11)	2 (12)	3 (13)

w buforze słownikowym szukamy najdłuższego podciągu równego możliwie długiemu prefiksowi bufora wejściowego.

a	b	a	c	c	a	b	a
0 (6)	1 (7)	2 (8)	3 (9)	0 (10)	1 (11)	2 (12)	3 (13)

kodujemy znaleziony podciąg w następujący sposób: **(3, 1, a)** i zapisujemy kod na wyjście.

przesuwamy elementy w buforach w lewo o $1+1$ i uzupełniamy bufor wejściowy.

a	c	c	a	b	a	a	a
0 (8)	1 (9)	2 (10)	3 (11)	0 (12)	1 (13)	2 (14)	3 (15)

w buforze słownikowym szukamy najdłuższego podciągu równego możliwie długiemu prefiksowi bufora wejściowego. ponownie w buforze słownikowym nie ma podciągu pasującego do jakiegokolwiek długości prefiksu w buforze wejściowym. kodujemy ciąg o zerowej długości w znany już sposób: **(3, 0, b)**, zapisujemy kod na wyjście i przesuwamy elementy w buforach o 1.

c	c	a	b	a	a	a	a
---	---	---	---	---	---	---	---

0 (9) 1 (10) 2 (11) 3 (12) 0 (13) 1 (14) 2 (15) 3 (16)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

c	c	a	b	a	a	a	a
---	---	---	---	---	---	---	---

0 (9) 1 (10) 2 (11) 3 (12) 0 (13) 1 (14) 2 (15) 3 (16)

kodujemy znaleziony podciąg w następujący sposób: **(2, 1, a)**, zapisujemy kod na wyjście i przesuwamy elementy w buforze o 2.

a	b	a	a	a	a	a	c
---	---	---	---	---	---	---	---

0 (11) 1 (12) 2 (13) 3 (14) 0 (15) 1 (16) 2 (17) 3 (18)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

a	b	a	a	a	a	a	c
---	---	---	---	---	---	---	---

0 (11) 1 (12) 2 (13) 3 (14) 0 (15) 1 (16) 2 (17) 3 (18)

znaleźliśmy podciąg **trójelementowy**!

część podciagu znalezionego w buforze słownikowym pokrywa się z buforem wejściowym.

taki podciąg kodujemy w taki sam sposób, czyli: **(2, 3, c)** i zapisujemy kod na wyjście.

elementy buforów słownikowego i wejściowego przesuwamy o 4 pozycje.

a	a	a	c	a	c	b	c
---	---	---	---	---	---	---	---

0 (15) 1 (16) 2 (17) 3 (18) 0 (19) 1 (20) 2 (21) 3 (22)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

a	a	a	c	a	c	b	c
---	---	---	---	---	---	---	---

0 (15) 1 (16) 2 (17) 3 (18) 0 (19) 1 (20) 2 (21) 3 (22)

kodujemy znaleziony podciąg w następujący sposób: **(2, 2, b)**, zapisujemy kod na wyjście i przesuwamy elementy w buforze o 3.

c	a	c	b	c	b	b	c
---	---	---	---	---	---	---	---

0 (18) 1 (19) 2 (20) 3 (21) 0 (22) 1 (23) 2 (24) 3 (25)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

c	a	c	b	c	b	b	c
---	---	---	---	---	---	---	---

0 (18) 1 (19) 2 (20) 3 (21) 0 (22) 1 (23) 2 (24) 3 (25)

kodujemy znaleziony podciąg w następujący sposób: **(2, 2, b)**, zapisujemy kod na wyjście i przesuwamy elementy w buforze o 3.

b	c	b	b	c	c	a	a
---	---	---	---	---	---	---	---

0 (21) 1 (22) 2 (23) 3 (24) 0 (25) 1 (26) 2 (27) 3 (28)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

b	c	b	b	c	c	a	a
---	---	---	---	---	---	---	---

0 (21) 1 (22) 2 (23) 3 (24) 0 (25) 1 (26) 2 (27) 3 (28)

kodujemy znaleziony podciąg w następujący sposób: **(1, 1, c)**, zapisujemy kod na wyjście i przesuwamy elementy w buforze o 2.

b	b	c	c	a	a	c	b
---	---	---	---	---	---	---	---

0 (23) 1 (24) 2 (25) 3 (26) 0 (27) 1 (28) 2 (29) 3 (30)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

ponownie w buforze słownikowym nie ma podciagu pasującego do jakiegokolwiek prefiksu.

kodujemy więc ciąg o zerowej długości: **(0, 0, a)** i zapisujemy kod na wyjście.

przesuwamy elementy w buforach o drugi element trójki powiększony o jeden, czyli o 0+1.

b	c	c	a	a	c	b	a
---	---	---	---	---	---	---	---

0 (24) 1 (25) 2 (26) 3 (27) 0 (28) 1 (29) 2 (30) 3 (31)

w buforze słownikowym szukamy najdłuższego podciagu równego możliwie długiemu prefiksowi bufora wejściowego.

b	c	c	a	a	c	b	a
---	---	---	---	---	---	---	---

0 (24) 1 (25) 2 (26) 3 (27) 0 (28) 1 (29) 2 (30) 3 (31)

kodujemy znaleziony podciąg w następujący sposób: **(3, 1, c)**, zapisujemy kod na wyjście i przesuwamy elementy w buforze o 2.

c	a	a	c	b	a		
---	---	---	---	---	---	--	--

0 (26) 1 (27) 2 (28) 3 (29) 0 (30) 1 (31) 2 3

w buforze słownikowym szukamy najdłuższego podciągu równego możliwie długiemu prefiksowi bufora wejściowego.

nie znaleźliśmy literki b, więc kodujemy podciąg o zerowej długości: **(1, 0, b)** i zapisujemy kod na wyjście.

przesuwamy elementy w o 0+1.

a	a	c	b	a			
---	---	---	---	---	--	--	--

0 (27) 1 (28) 2 (29) 3 (30) 0 (31) 1 2 3

w buforze słownikowym szukamy najdłuższego podciągu równego możliwie długiemu prefiksowi bufora wejściowego.

znaleźliśmy literę a i kodujemy ją jako ciąg jednoelementowy: **(0, 1, NULL)** i zapisujemy kod na wyjście.

na pozycji następującej po zakodowanym podciągu nie znajdowała się już żadna litera, ponieważ skończyły się dane wejściowe.

w takiej sytuacji należy w jakiś sposób zapisać zakończenie danych wejściowych (np. za pomocą specjalnego znaku).

wynik kodowania jest następujący:

a (1, 2, b) (2, 2, b) (1, 2, a) (3, 1, a) (3, 0, b) (2, 1, a) (2, 3, c) (2, 2, b) (2, 2, b) (1, 1, c) (0, 0, a) (3, 1, c) (1, 0, b) (0, 1, NULL)

przesuwamy indeksy bufora o 4.

a	a	b	a	b	a	b	a	a	c	c	a	b	a	a	a	a	a	c										...
										0	1	2	3															

pobieramy dziesiąty kod: (2, 2, b)

pobieramy z bufora kolejne **2 litery**, zaczynając od pozycji **2**, na końcu dopisujemy literę **b** i to wszystko wypisujemy na wyjście.

a	a	b	a	b	b	a	b	a	c	c	a	b	a	a	a	a	c	a	c	a	c	b							...
										0	1	2	3																

przesuwamy indeksy bufora o 3.

a	a	b	a	b	b	a	b	a	c	c	a	b	a	a	a	a	a	c	a	c	b							...
										0	1	2	3															

pobieramy dziesiąty kod: (2, 2, b)

pobieramy z bufora kolejne **2 litery**, zaczynając od pozycji **2**, na końcu dopisujemy literę **b** i to wszystko wypisujemy na wyjście.

a	a	b	a	b	b	a	b	a	c	c	a	b	a	a	a	a	a	c	a	c	b	b	c	b	b					...
										0	1	2	3																	

przesuwamy indeksy bufora o 3.

a	a	b	a	b	b	a	b	a	c	c	a	b	a	a	a	a	a	c	a	c	b	b	c	b	b					...
										0	1	2	3																	

pobieramy jedenasty kod: (1, 1, c)

pobieramy z bufora kolejne **1 literę**, zaczynając od pozycji **1**, na końcu dopisujemy literę **c** i to wszystko wypisujemy na wyjście.

a	a	b	b	a	b	a	c	c	a	b	a	a	a	a	c	a	c	b	b	c	b	b	c	c			...	
										0	1	2	3															

przesuwamy indeksy bufora o 2.

a	a	b	a	b	a	b	a	c	c	a	a	a	a	c	b	b	c	c	c	...		
																			0	1	2	3

pobieramy dwunasty kod: (0, 0, a)

pobieramy z bufora kolejne **0 liter**, zaczynając od pozycji **0**, na końcu dopisujemy literę **a** i to wszystko wypisujemy na wyjście.

a	a	b	a	b	b	a	b	a	c	c	a	b	a	a	a	c	c	b	b	c	c	a				...												
																						0	1	2	3													
																										a	c	c	c									

przesuwamy indeksy bufora o 1.

a	a	b	a	b	b	a	c	c	a	b	a	a	a	c	c	b	b	c	c	a														
																			0	1	2	3												
																										:								

pobieramy trzynasty kod: (3, 1, c)

pobieramy z bufora kolejne **1 liter**, zaczynając od pozycji **3**, na końcu dopisujemy literę **c** i to wszystko wypisujemy na wyjście.

[illegible]

przesuwamy indeksy bufora o 2.

[illegible]

pobieramy czternasty kod: (1, 0, b)

pobieramy z bufora kolejne **0 liter**, zaczynając od pozycji **1**, na końcu dopisujemy literę **b** i to wszystko wypisujemy na wyjście.

[illegible]

kilka słów o kompresji

załóżmy, że literę można zakodować na 8 bitach (np. stosując kodowanie ASCII).

przy tym założeniu nasz 32 literowy komunikat ma rozmiar 32 bajtów (1 bajt = 8 bitów).

kodowane trójki składają się zawsze z dwóch liczb z przedziału od 0 do 3 oraz z litery.

liczbę z przedziału od 0 do 3 można zakodować na 2 bitach.

całą trójkę można zatem zakodować na 2+2+8 bitach.

do zakodowania całego komunikatu użyliśmy 16 kodów.

$$16 * 12 = 192$$

$$192 / 8 = 24$$

zatem do zakodowania całego komunikatu użyliśmy 24 bajtów!

$$24 / 32 = \sim 0.75$$

zatem przy pomocy kompresji udało nam się zmniejszyć rozmiar komunikatu do 75% jego oryginalnego rozmiaru!

zatem otrzymaliśmy stopień kompresji równy około 25%.



literatura

Adam Drozdek
Wprowadzenie do kompresji danych

Artur Przelaskowski
Kompresja danych