

Алгоритм Лемпеля-Зива LZ78

Исходный текст: КОЛОКОЛ\_ОКОЛО\_КОЛОКОЛЬНИ:) (26 символов)

Исходные коды постоянной длины m (4 бита)

<b>И</b>	0	0	0	0	<b>Л</b>	0	0	1	0	<b>О</b>	0	1	0	0	:	0	1	1	0	_	1	0	0	0
<b>К</b>	0	0	0	1	<b>Н</b>	0	0	1	1	<b>Б</b>	0	1	0	1	)	0	1	1	1					

Сжатие

Входные символы	Запись в сжатый поток	Двоичный код (<1-4 бита, 4 бита>)	Словарь просмотренных фраз	Позиция фразы в словаре (0-31)
-	-	-	-	0
'к'	<0,'к'>	0 0001	'к'	1
'о'	<0,'о'>	0 0100	'о'	2
'л'	<0,'л'>	00 0010	'л'	3
'о'	<2,'к'>	10 0001	'ок'	4
'к'				
'о'	<2,'л'>	010 0010	'ол'	5
'л'				
' '	<0,' '>	000 1000	' '	6
'о'	<4,'о'>	100 0100	'око'	7
'к'				
'о'				
'л'	<3,'о'>	011 0100	'ло'	8
'о'				
' '	<6,'к'>	0110 0001	'_к'	9
'к'				
'о'	<5,'о'>	0101 0100	'оло'	10
'л'				
'о'				
'к'	<1,'о'>	0001 0100	'ко'	11
'о'				
'л'	<3,'б'>	0011 0101	'ль'	12
'б'				
'н'	<0,'н'>	0000 0011	'н'	13
'и'	<0,'и'>	0000 0000	'и'	14
':'	<0,':'>	0000 0110	':'	15
')'	<0,')'>	0000 0111	')'	16

Распаковка

Двоичный код (<1-4 бита, 4 бита>)	Код для распаковки	Восстановленный текст	Словарь просмотренных фраз	Позиция фразы в словаре (0-31)
-	-	-	-	0
0 0001	<0,'к'>	к	'к'	1
0 0100	<0,'о'>	о	'о'	2
00 0010	<0,'л'>	л	'л'	3
10 0001	<2,'к'>	ок	'ок'	4
010 0010	<2,'л'>	ол	'ол'	5
000 1000	<0,' '>	'	' '	6
100 0100	<4,'о'>	око	'око'	7
011 0100	<3,'о'>	ло	'ло'	8
0110 0001	<6,'к'>	_к	'_к'	9
0101 0100	<5,'о'>	оло	'оло'	10
0001 0100	<1,'о'>	ко	'ко'	11
0011 0101	<3,'б'>	ль	'ль'	12
0000 0011	<0,'н'>	н	'н'	13
0000 0000	<0,'и'>	и	'и'	14
0000 0110	<0,':'>	:	':'	15
0000 0111	<0,')'>	)	')'	16

Объем исходного текста: 26 (символов) × 4 (бита) = 104 бита

Объем сжатого текста: 2 (кода) × 5 (бит) + 2 (кода) × 6 (бит) + 4 (кода) × 7 (бит) + 8 (кодов) × 8 (бит) = 114 бит