

Tutorial za LZ77 koder v.1.0

Da ne duljim, prvo da se dogovorimo osnovne pojmove za

- parametar (x, y, z): x – pomak, y – duljina, z – sljedeći simbol
- postupak kodiranja: **plavo** = prozor za kodiranje, **crveno**=posmačni prozor
- * je kraj

Primjer 11 iz TINF_Zbirka_zadataka_2010-11.pdf – uzeo sam ga jer je „zgodniji“ za demonstraciju:

Koristeći algoritam LZ77 kodirajte poruku *abaabbabaacaacabb* uzimajući pri tome da je maksimalna duljina posmičnog prozora i prozora za kodiranje 6, odnosno 5 simbola.

Koder

a b a a b b a b a a c a a c a b b * -> zadano

1.	Na početku je posmačni prozor prazan. U prozoru za kodiranje se nalazi 5 simbola (tako veli zadatak). Zakodirali smo slovo a (jer je na prvom mjestu).	a b a a b b a b a a c a a c a b b *	(0, 0, a)
2.	U posmačnom prozoru se sada nalazi slovo a. Ono što UVIJEK gledamo je da li se u posmačnom prozoru nalazi neki simbol ili vise njih koji imaju isti redosljed kao i u pozoru za kodiranje. Moramo b zakodirati, „a“ != „b“	a b a a b b a b a a c a a c a b b *	(0, 0, b)
3.	Opet isto, gledamo posmacni prozor i vidimo da je na 1. mjestu simbol „a“, a ono nam upravo treba za 1. mjesto u prozoru za kodiranje. Moram se vratiti 2 mjesta (x) i njega „kopirati“ jednom (y) te mi je sljedeci znak (z) opet „a“.	a b a a b b a b a a c a a c a b b *	(2, 1, a)
4.	Vec u posmacnom imamo 4 simbola 😊, jasnije jelda? trazimo „b“ i vidimo se moramo vratiti 3 mjesta unazad (x), jednom kopirati (y) i sljedeci simbol (z) je b. Sigurno se pitate zasto ne mozemo reci da nam je y = 2 pa cemo „b“ 2 puta kopirat. Zato sto bi kopirali i simbol „b“ i simbol „a“ a nama treba „b“, „b“. Ovaj 2. „b“ smo kodirali sa sljedecim znakom (z) Pogledajmo korak 5. Tu ce biti sve jasno 😊	a b a a b b a b a a c a a c a b b *	(3, 1, b)

5.	Vidite simetriju „abaa“? Sada mozemo sva 4 simbola zakodirati, vratimo se na pocetak za 6 mjesta (x) i gledamo: a, b, a, a je 4 mjesta (y) i sljedeci znak je „c“. Ovo je upravo ono najbitnije kod ovog koder. Do sada smo imali samo jedan znak, a sada smo odmah 4 zakodirali + 1 za sljedeci simbol.	a b a a b b a b a a c a a c a b b *	(6, 4, c)
6.	Posmačni prozor smo pomakli jer mu je max duljina = 6, a broji se od krajnjeg desnog znaka koji je zakodiran. Opet isto, vidimo da trazimo a, a, c. Ovdje cete se malo zbuniti. Sto ce se dogoditi ako se vratim 3 simbola nazad a napravim 4 „kopije“? Zakodirat cu nas a, a, c i jos jedan znak. +1 znak ce „prijeci“ u prozor za kodiranje (plavo) i tamo je prvi znak „a“ te ce upravo on biti zakodiran. Znaci imat cu: a, a, c, a i reci cu da mi je sljedeci znak „b“. Ovo i mene zbunjuje, no ima neke logike :P	a b a a b b a b a a c a a c a b b *	(3, 4, b)
7.	Trazim „b“ pa se vratim 1 simbol u nazad, njega kopiram i stavim da je sljedeci znak kraj.	a b a a b b a b a a c a a c a b b *	(1, 1, *)

Ovo je prva verzija, ide odmah na materijale. Nisam gledao pravopis, zurio sam :D

Pogledajte ovo pa molim da me uputite ako ima koja greska ili ako nije najjasnije napisano pa da to odmah ispravimo/objasnimo/prepravimo.