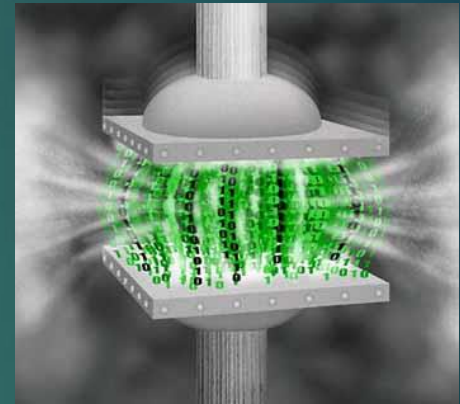


# Kompresija i zaštita podataka

DR MARKO D. PETKOVIĆ

[dexterofnis@gmail.com](mailto:dexterofnis@gmail.com)

[www.pmf.ni.ac.rs/dexter](http://www.pmf.ni.ac.rs/dexter)



# Sadržaj predmeta i način polaganja ispita

## ▶ Sadržaj predmeta:

- ▶ Kompresija podataka
- ▶ Prenos i zaštita podataka
- ▶ Bezbednost podataka

## ▶ Ispit

- ▶ Dva studentska projekta (implementacija određenih algoritama)
- ▶ Završni ispit – odbrana projekata

# Kompresija podataka

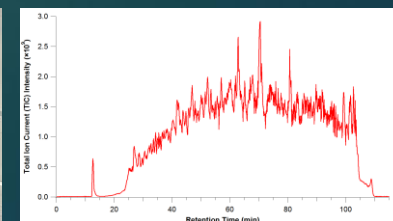
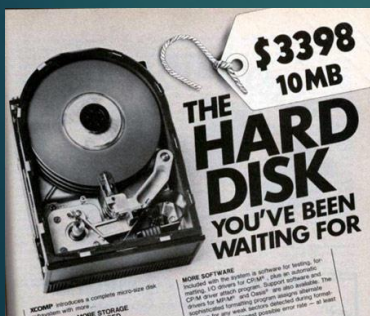


## ▶ Nekad (pre ~30god):

- ▶ Hard disk: 80 MB
- ▶ Diskete 1.44 MB, 1.2 MB, 360 kB
- ▶ Masovno “zipovanje” fajlova
  - ▶ “Ručni” alati za kompresiju i dekompresiju
  - ▶ Deljenje i prenos podataka na više disketa
- ▶ “Brzi” modemi ...

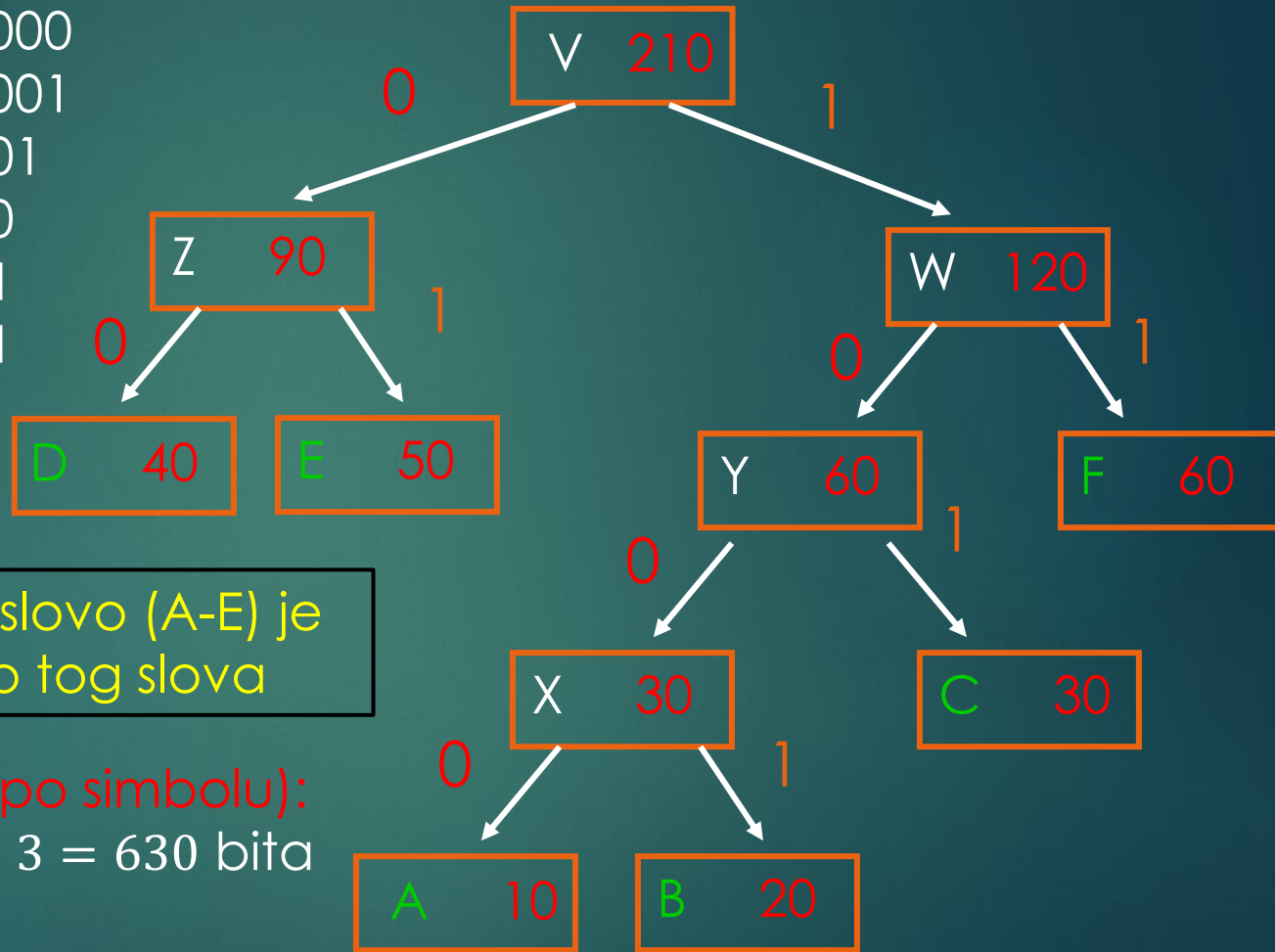
## ▶ Sad:

- ▶ Hard disk: 4 TB
- ▶ “U pozadini”
  - ▶ DOCX, XLSX, PPTX, ...
  - ▶ GIF, PNG, JPEG, MP3, MPEG
  - ▶ Instalacije programa
- ▶ E – mail
- ▶ Veliki serveri (Google, FB, ... )
- ▶ Cloud servisi...
- ▶ Nestandardne primene
  - ▶ Hromatografija, ...



# Hafmanov algoritam

A :  $V_A = 1000$   
B :  $V_B = 1001$   
C :  $V_C = 101$   
D :  $V_D = 00$   
E :  $V_E = 01$   
F :  $V_F = 11$



Kodna reč za svako slovo (A-E) je put od korena (V) do tog slova

Bez kompresije (3 bita po simbolu):

$$(10 + 20 + \dots + 60) \cdot 3 = 630 \text{ bita}$$

Kompresovani fajl:

$$10 \cdot 4 + 20 \cdot 4 + 30 \cdot 3 + 40 \cdot 2 + 50 \cdot 2 + 60 \cdot 2 = 510 \text{ bita}$$

# LZW algoritam (koder + dekoder)



Ulazni niz:

a b c a b a b c

Izlaz: indeks nađene reči u rečniku

Poslato:

1 2 3 5 5

Dekodirano:

a b c a b a b

Rečnik koder:




1	a	6	bc
2	b	7	ca
3	c	8	aba
4	d	9	abc
5	a b		

Rečnik dekoder:

1	a	6	bc
2	b	7	ca
3	c	8	aba
4	d		
5	a b		

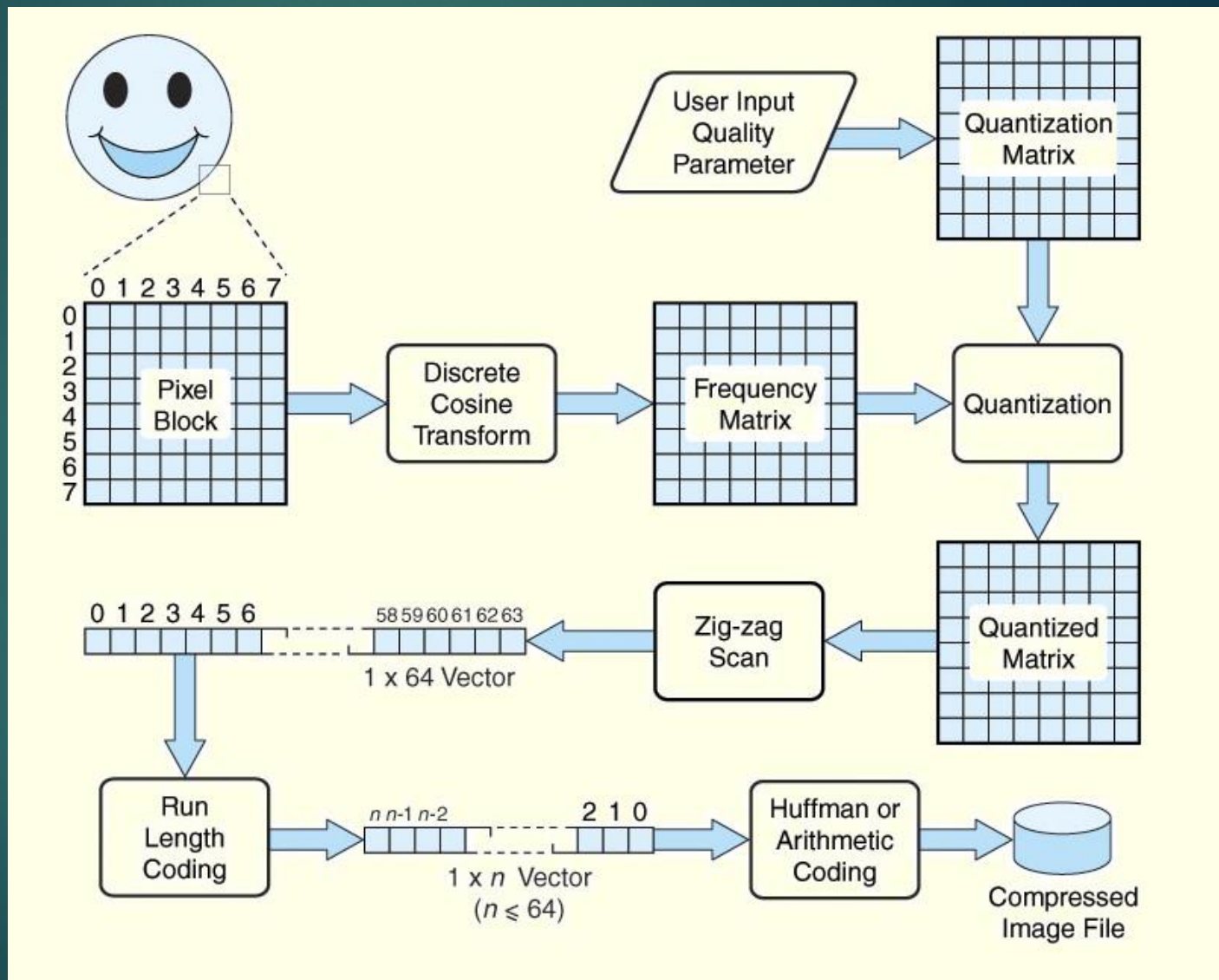
# Praktične realizacije

- ▶ ZIP:
  - ▶ Najčešće DEFLATE (LZ77 + Huffman)
  - ▶ Ponekad: bzip2, PPM (arithmetical), itd.
- ▶ GIF slike: LZW
- ▶ PNG: DEFLATE
- ▶ RAR: Neobjavljeno!! Nezvanično (LZSS + PPM, DEFLATE,...)
- ▶ 7-ZIP: kao ZIP, poboljšan!!

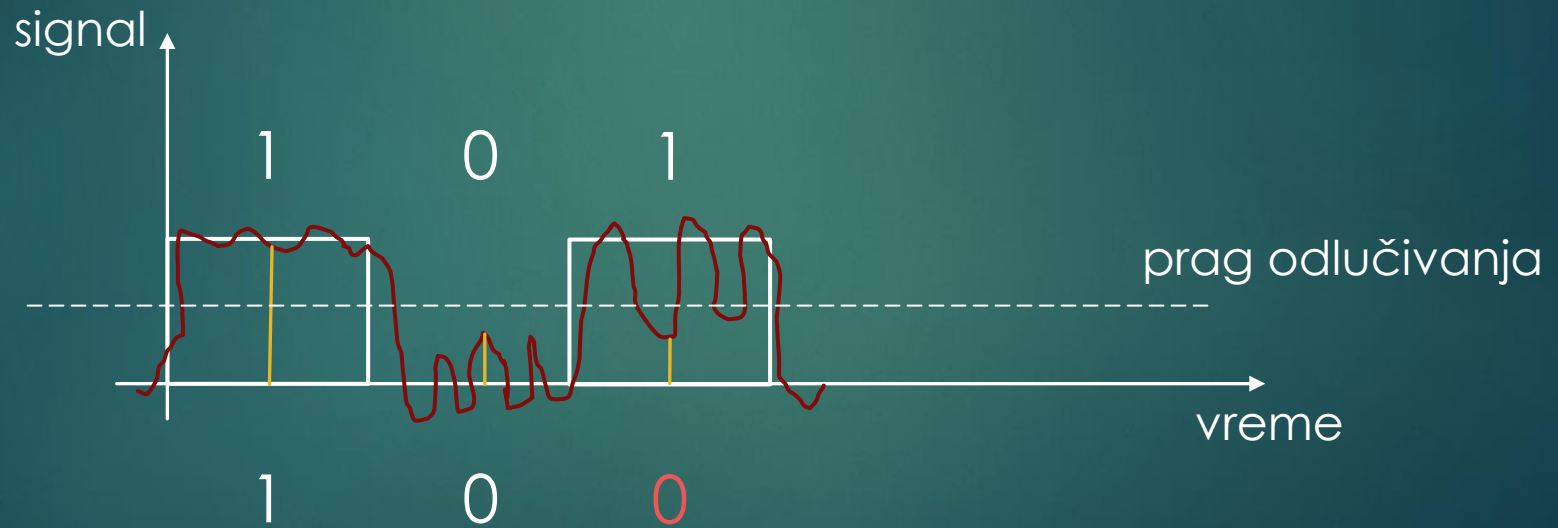
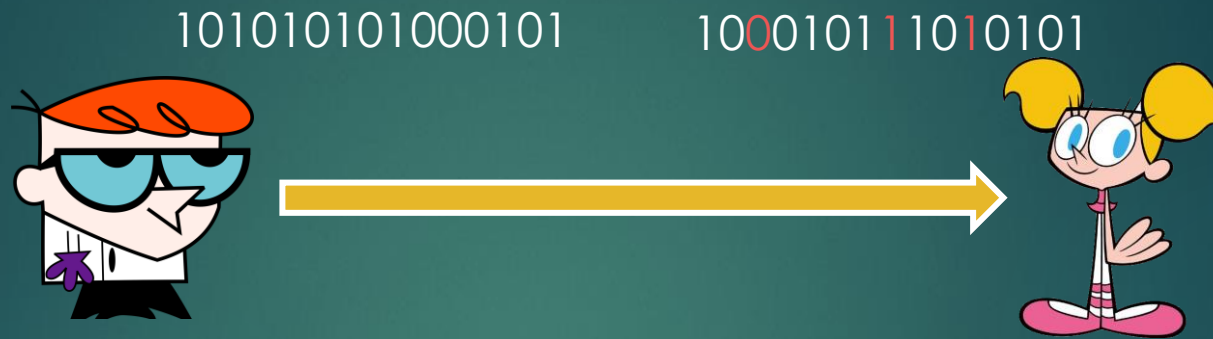
Name	Size	Packed ...	Modified	CRC	Meth...	Host OS	Version
 VIDEO0013.mp4	19 087 761	18 279 819	2013-12-16 10:12	23AF983D	Deflate	FAT	20
 1387042915835.jpg	339 269	336 808	2013-12-16 10:12	EAE40036	Deflate	FAT	20
 1387042920286.jpg	314 166	309 959	2013-12-16 10:12	47999B57	Deflate	FAT	20

# Praktične realizacije

## ▶ JPEG:



# Zaštita podataka



Kako detektovati / ispraviti ovakve greške?!?!



# Ilustrativni primer

Ne brisite ovu poruku zbog toga sto izgleda cudno. Verovali ili ne, mozete je procitati.

Nsiam vrevoao da zpavrao mgou rzmaueti ono sto ctaim. Zaavljhuujci nobniecoj mcoi ljdkstuog mgzoa, pemra irtazsiavnjima nucainka sa Kmbreidza, njie vzano kjoim su roedsldoem npiasnaa slvoa u rcei, jdieno je btino da se pvro i psldeonje sovlo nlaaze na sovm msteu.

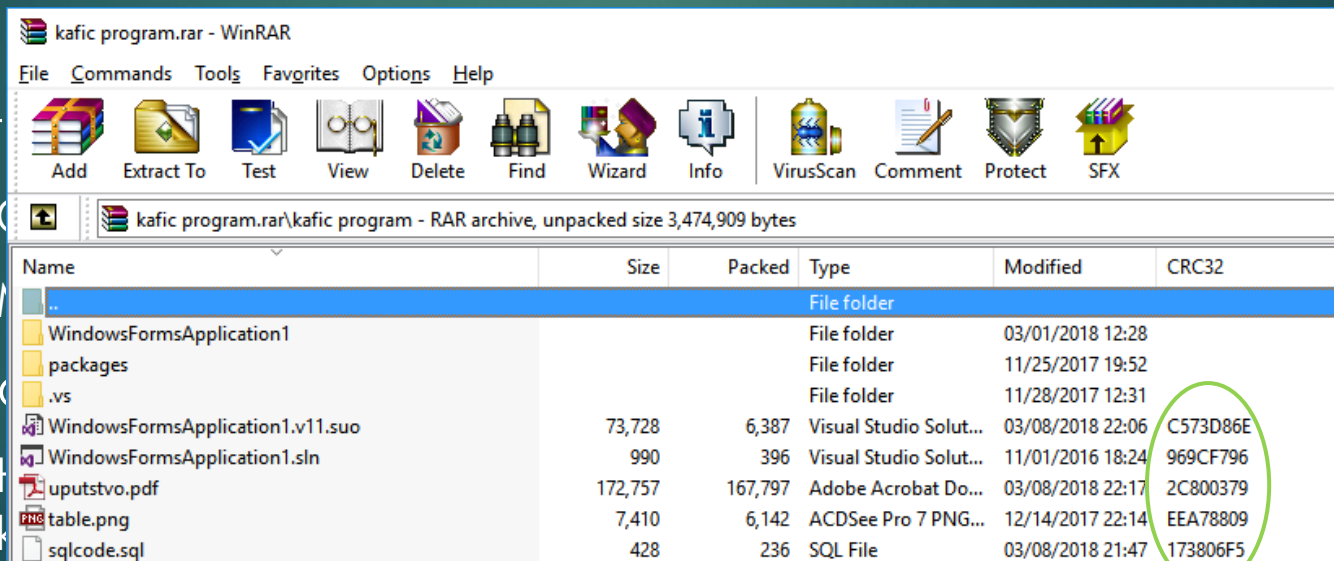
Otasla solva mgou btii u ptponuom nerdeu i bez ozibra na ovu oloknost, tkest mzeote ctiami bez pobrelma. Ovo je zbog tgoa sto ljdkusi mzoak ne ctia savko slvoa pnaoosb, vec rcei psmraota kao cleniu. Oavj preomecaj je sljiavo nzavan tipoglikemija :-)

Zacudjujuce, zar ne? A uevk ste msilili da je pavrpois vzaan.

# “Korisna” redundansa?!?!

- ▶ Redundansa pomaže da se detektuju i/ili isprave greške (oštećenje medijuma, greške pri prenosu, itd.)
- ▶ Svesno dodavanje “viška” informacija
- ▶ Cilj: Ispraviti što više grešaka sa minimalnom količinom unešene redundanse

▶ Det



The screenshot shows the WinRAR interface for a file named 'kafic program.rar'. The file list contains the following items:

Name	Size	Packed	Type	Modified	CRC32
..			File folder		
WindowsFormsApplication1			File folder	03/01/2018 12:28	
packages			File folder	11/25/2017 19:52	
.vs			File folder	11/28/2017 12:31	
WindowsFormsApplication1.v11.suo	73,728	6,387	Visual Studio Solut...	03/08/2018 22:06	C573D86E
WindowsFormsApplication1.sln	990	396	Visual Studio Solut...	11/01/2016 18:24	969CF796
uputstvo.pdf	172,757	167,797	Adobe Acrobat Do...	03/08/2018 22:17	2C800379
table.png	7,410	6,142	ACDSee Pro 7 PNG...	12/14/2017 22:14	EEA78809
sqlcode.sql	428	236	SQL File	03/08/2018 21:47	173806F5

The CRC32 values for the files are: C573D86E, 969CF796, 2C800379, EEA78809, and 173806F5. The value 173806F5 is circled in green in the original image.

▶ C

▶ Ispr

▶ H

▶ K

novi

# Hvala na pažnji!!!!



## Programmer



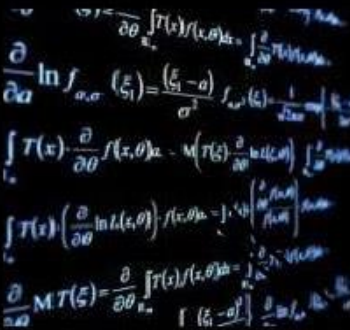
what my parents think  
I do



what my friends think  
I do



what society thinks I  
do



what our professors  
envisioned we could  
do...



what I think I do

```
1 public class Demo {
2     private static final String CONSTANT = "string";
3     private Object o;
4
5     //
6     * Create a new Demo.
7     * @param o the object to demonstrate.
8     * @return this
9     public Demo(Object o) {
10        this.o = o;
11        String s = CONSTANT;
12        int i = 1;
13    }
14
15    public static void main(String[] args) {
16        Demo demo = new Demo();
17    }
18 }
```

what I really do

