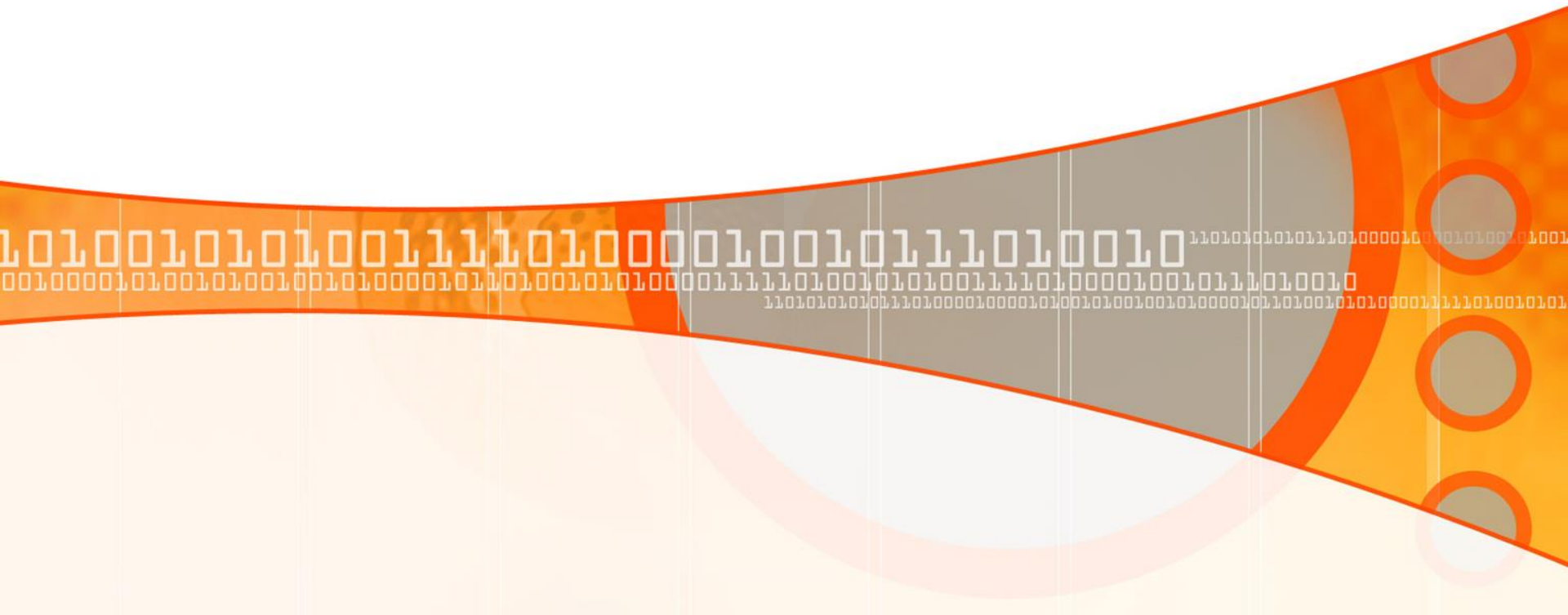


STRUCTURI DE DATE

Compresia datelor



COMPRESIA DATELOR

Caracteristici:

- **Proces de codificare;**
- **Utilizarea unui numar mai mic de biti pentru stocarea datelor;**
- **Functioneaza daca emitatorul si receptorul au algoritmul de codificare/decodificare;**

COMPRESIA DATELOR

Caracteristici (continuare):

- **Avantaj:** reducerea gradului de utilizare a resurselor (HDD, latime de banda etc);
- **Dezavantaj:** proces eventual costisitor pentru codificare/decodificare;
- **Algoritmi:** fara/cu pierdere de informatie;

COMPRESIA DATELOR

Algoritmi fara pierdere de informatie:

- **Profita de redundanta statistica;**
- **Date compresate fara erori;**
- **Reversibili: datele sunt reconstituite in formatul original.**

COMPRESIA DATELOR

Algoritmi cu pierdere de informatie:

- **Accepta pierderea de continut la codificare/decodificare;**
- **Utilizati in functie de modul de perceptie a datelor;**
- **Acceptare pierderi daca rata de compresie este foarte ridicata.**

COMPRESIA DATELOR

Exemple algoritmi fara pierdere de informatie:

- RLE;
- LZ;
- LZW;
- Huffman;
- etc...

COMPRESIA DATELOR

Exemple algoritmi cu pierdere de informatie:

- **DCT: Discrete Cosine Transform;**
- **Compresie cu fractali;**
- **etc...**

COMPRESIA DATELOR

RLE: Run-Length Encoding

- Secvente cu valori consecutive;
- Inlocuire secventa cu (frecventa aparitie, valoare);
- Aplicativitate: imagini cu repetitie mare a valorilor de reprezentare a culorilor.

Exemplu:

AAAAAAAAAANNAAAAAAAAANNNNNNN

10A2N5A7N

COMPRESIA DATELOR

LZ: Lempel-Ziv

- Bazat pe lungimea codurilor identificate;
- Construire dictionar cu grupuri de simboluri din datele compresate;

- Pasi algoritm:
 1. Initializare dictionar cu blocurile de lungime 1;
 2. Cautarea celui mai mare (lungime) bloc care apare in dictionar;

COMPRESIA DATELOR

LZ: Lempel-Ziv

3. Codificare bloc cu index din dictionar;
4. Adaugare in dictionar bloc concatenat cu primul simbol din blocul urmator;
5. Reluare pasul 2.

Exemplu:

A	B	B	A	A	B	B	A	A	B	A	B	B	A	A	A	A	B	A	A	B	B	A
0	1	1	0	2	4	2	6	5	5	7	3	0										

COMPRESIA DATELOR

LZW: Lempel-Ziv-Welch

- **Imbunatatire algoritm LZ;**
- **Dictionar initializat cu caracterele textului (o singura aparitie);**
- **Scanare sir intrare pentru subsiruri din ce in ce mai lungi pana cand este identificat unul care nu se afla in dictionar;**

COMPRESIA DATELOR

LZW: Lempel-Ziv-Welch

- **Noul subsir, mai putin ultimul caracter, este introdus in secventa codificata;**
- **Noul subsir este adaugat in dictionar cu primul cod disponibil.**

COMPRESIA DATELOR

Codul Huffman

- **Trebuie sa se cunoasca frecventa de aparitie a caracterelor;**
- **Pentru fiecare caracter se asociaza o secventa de biti;**
- **Secventa de biti – construita pe baza unui arbore binar;**

COMPRESIA DATELOR

Algoritm Huffman:

- **Ordonare descrescatoare simboluri text compresat; criteriu: frecventa de aparite;**
- **Un simbol reprezinta un nod in arbore; fiecare nod are asociata o frecventa de aparitie;**

COMPRESIA DATELOR

Algoritm Huffman (continuare):

- Doua noduri sunt legate daca au asociate cele mai mici frecvente de aparitie; nodul parinte are asociata suma frecventelor nodurilor legate;
- Oprire algoritm: exista un singur nod (nelegat).

COMPRESIA DATELOR

Exemplu Huffman:

Simbol	Nr. de apariții simbol	Total nr. de biți pentru un simbol (cod in clar)	Total nr. de biți pentru un simbol (cod Huffman)
0	12	96	12
1	8	64	16
2	6	48	18
3	5	40	20
4	4	32	20
5	3	24	18
6	2	16	14
7	0	0	0
Total	40	320	118

COMPRESIA DATELOR

Exemplu Huffman:

