TEORIA INFORMAŢIEI ŞI A CODĂRII

Culegere de probleme

Vol.1

Horia BALTA   Maria KOVACI   Radu LUCACIU

2009

Cuprins

Cap. 1 Probabilități. Informația

Dicționar:

- **aposteriori (a posteriori) -**locuțiune latină: “din ceea ce urmează”, după experiență, pornind de la datele ei;
- **apriori (a priori) -** locuțiune latină: “din ceea ce precede”, anterior experienței;
- **binar** -1.care este constituit din două elemente; 2.a cărui bază este numărul 2;
- **bit/biți** -1. Unitate de măsură a informației; 2.simbol binar;
- **discret** -care este alcătuit din elemente distincte; care variază în salturi; cuantificat; discontinuu;
- **echiprobabil(e**) -de egală probabilitate;
- **informație** -necesarul/rezultatul cunoașterii;
- **probabilitate** -1.însușirea de a fi probabil; 2.măsură (funcție) definită pe un câmp de evenimente, $p : \Omega \to [0,1]$.

Definiții:

- **sursă de informație (sau experiment probabilistic)** = un mecanism (un experiment) prin care se selectează un mesaj (un eveniment) dintre n posibile după o lege arbitrară (sau cel puțin necunoscută);
- **mesaj (eveniment)** = realizarea produsă în urma efectuării experimentului;
- **1 bit** = cantitatea de informație furnizată de o sursă de informație binară, fără memorie, echiprobabilă, printr-un mesaj al ei;
- **eveniment elementar** = un eveniment ce nu poate fi definit ca o reuniune de două evenimente distincte între ele și de primul.

Breviar teoretic:

1. Probabilitate

Determinarea experimentală a probabilității de realizare a unui mesaj (eveniment) A se face după relația:

$$p(A) = \frac{\text{număr de cazuri favorabile lui A}}{\text{număr total de cazuri}} \qquad (1.1)$$

2. Probabilitate condiționată

Determinarea experimentală a probabilității de realizare a evenimentului (mesajului) B atunci când s-a realizat evenimentul (mesajul) A se face după relația:

$$p(B/A) = \frac{\text{număr de cazuri în care s-au realizat și A și B}}{\text{număr de cazuri în care s-a realizat A}} \qquad (1.2)$$

3. Formula fundamentală a probabilităților evenimentelor elementare

Dacă A_i, $i = 1 \div n$ sunt evenimentele elementare ale unui experiment probabilistic (mesajele unei surse de informație) atunci:

$$\sum_{i=1}^{n} p(A_i) = 1 \qquad (1.3)$$

4. Relația lui Bayes

Dacă A și B sunt două evenimente atunci:

$$p(A,B) = p(A)\cdot p(B/A) = p(B)\cdot p(A/B) \qquad (1.4)$$

unde p(A, B) = probabilitatea de a se realiza și A și B.

5. Formula probabilității totale

Dacă A_i cu $i = 1, n$ sunt evenimentele elementare ale unui experiment probabilistic și B un eveniment oarecare pentru același experiment atunci:

$$p(B) = \sum_{i=1}^{n} p(A_i)\cdot p(B/A_i) \qquad (1.5)$$

6. Evenimente independente

Setul de evenimente A_i, $i \in I$, sunt independente dacă și numai dacă pentru $\forall\, J \subset I$

$$p\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} p(A_i) \qquad (1.6)$$

În particular, A și B sunt două evenimente independente dacă și numai dacă:

$$p(A,B)=p(A\cap B)=p(A)\cdot p(B) \tag{1.7}$$

şi utilizând relaţia (1.4)

$$\begin{aligned} p(A)&=p(A/B)\\ p(B)&=p(B/A) \end{aligned} \tag{1.8}$$

7. Informaţia

Cantitatea de informaţie necesară ridicării nedeterminării asupra evenimentului A este egală cu cea furnizată de realizarea evenimentului A şi egală cu :

$$i(A)=\log_2 \frac{1}{p(A)} \quad [\text{biţi}] \tag{1.9}$$

1.1 Zece mingi sunt puse în trei cutii C_1, C_2, C_3. Care este probabilitatea ca în C_1 să fie 3 mingi?

<u>Rezolvare:</u>

Fiecare minge poate fi aşezată în oricare din cele trei cutii; astfel că fiecare minge triplează numărul de variante de aşezare a mingilor în cutii. Aşadar numărul de variante de aşezare a mingilor în cutii este:

$$N = 3\cdot 3\cdot 3 \ldots\ldots \cdot 3 = 3^{10} = 59.049 \tag{1.1.1}$$

Pentru a avea trei mingi în C_1 trebuie ca celelalte şapte să fie aşezate în C_2
şi C_3. Numărul de variante cu trei mingi în C_1 va fi:

$$N_{3C_1} = C_{10}^{3}\cdot 2^7 = 120\cdot 128 = 15.360 \tag{1.1.2}$$

unde C_{10}^{3} reprezintă numărul de moduri de alegere a 3 mingi din 10 posibile (considerând mingile distincte); iar 2^7 reprezintă numărul de posibilităţi de aşezare a şapte mingi în două cutii, C_2 şi C_3. Probabilitatea cerută este:

$$P_{3C_1} = \frac{C_{10}^{3}\cdot 2^7}{3^{10}} \cong 26\% \tag{1.1.3}$$

1.2. Trei trăgători trag simultan asupra aceleiaşi ţinte. Probabilitatea ca fiecare trăgător să nimerească ţinta este p_1 = 0,4; p_2 = 0,5; p_1 = 0,7. Notând cu A evenimentul ca ţinta să fie lovită, B evenimentul ca ţinta să fie lovită exact o dată să se afle:

a) p(A);

b) p(B);

c) dacă cele două evenimente A şi B sunt independente.

Rezolvare:

a) Calculând probabilitatea evenimentului contrar lui A:

$$p(\overline{A}) = (1-p_1)\cdot(1-p_2)(1-p_3) = 9\% \tag{1.2.1}$$

rezultă că:

$$p(A) = 1 - p(\overline{A}) = 91\% \tag{1.2.2}$$

b) Avem că:

$$\begin{aligned} p(B) = {} & p(A_1 \cap \overline{A_2} \cap \overline{A_3}) + p(\overline{A_1} \cap A_2 \cap \overline{A_3}) + \\ & + p(\overline{A_1} \cap \overline{A_2} \cap A_3) = p_1(1-p_2)(1-p_3) + \\ & + (1-p_1)p_2(1-p_3) + (1-p_1)(1-p_2)p_3 = 36\% \end{aligned} \tag{1.2.3}$$

unde cu A_i s-a notat evenimentul ca trăgătorul i să nimerească ținta.

c) Pentru ca cele două evenimente să fie independente este necesar ca:

$$p(A/B) = p(A) \tag{1.2.4}$$

dar cum:

$$p(A/B) = 100\% \tag{1.2.5}$$

rezultă că cele două evenimente nu sunt independente.

1.3. Fie două urne, U_1 (ce conține 2 bile albe şi 3 bile negre) şi U_2 (ce conține o bilă albă şi 5 bile negre). Se extrage o bilă din U_1 şi se introduce în U_2, apoi se extrage o bilă din U_2. Care este probabilitatea ca bila transferată să fi fost albă dacă bila extrasă din U_2 este: a) albă; b) neagră?

Rezolvare:

Fie evenimentele A – bila transferată este albă; B – bila extrasă din U_2 este albă;

a) Pentru a calcula p(A/B) aplicăm formula lui Bayes:

$$p(A)\cdot p(B/A) = p(B)\cdot p(A/B) \tag{1.3.1}$$

Probabilitățile $p(A)$ si $p(\overline{A})$ se calculează simplu:

$$p(A) = \frac{2}{5} \quad \text{şi} \quad p(\overline{A}) = \frac{3}{5} \tag{1.3.2}$$

Probabilitatea condiționată p(B/A) este:

$$p(B/A) = 2/7 \tag{1.3.3}$$

iar p(B) se poate calcula cu formula probabilității totale:

$$p(B)=p(A)\cdot p(B/A)+p(\overline{A})\cdot p(B/\overline{A})=\frac{2}{5}\cdot\frac{2}{7}+\frac{3}{5}\cdot\frac{1}{7}=\frac{1}{5} \tag{1.3.4}$$

Astfel:

$$p(A/B)=\frac{p(A)\cdot p(B/A)}{p(B)}=\frac{\frac{2}{5}\cdot\frac{2}{7}}{\frac{1}{5}}=\frac{4}{7} \tag{1.3.5}$$

b) În mod analog

$$\begin{gathered}p(\overline{B}/A)=\frac{5}{7};p(\overline{B})=p(A)\cdot p(\overline{B}/A)+p(\overline{A})\cdot p(\overline{B}/\overline{A})=\\=\frac{2}{5}\cdot\frac{5}{7}+\frac{3}{5}\cdot\frac{6}{7}=\frac{4}{5}\end{gathered} \tag{1.3.6}$$

$$p(A/\overline{B})=\frac{p(A)\cdot p(\overline{B}/A)}{p(\overline{B})}=\frac{\frac{2}{5}\cdot\frac{5}{7}}{\frac{4}{5}}=\frac{5}{14}=\frac{2{,}5}{7}\lhd p(A/B) \tag{1.3.7}$$

Se observă, cum era de aşteptat, că este mai probabil să se fi transferat o bilă albă dacă din a doua urnă a fost extrasă o bilă albă.

1.4. La un examen oral studenții consideră că din totalul subiectelor m sunt uşoare şi n grele. Precizați:

a) Probabilitatea ca primul student să extragă un subiect uşor;

b) **Probabilitatea ca cel de-al doilea student să extragă un subiect uşor.**

Rezolvare:

a) Cu notațiile: A – evenimentul prin care primul student extrage un subiect uşor;
B – evenimentul prin care cel de-al doilea student extrage un subiect uşor, avem că:

$$p(A)=\frac{m}{m+n}\qquad p(\overline{A})=\frac{n}{m+n} \tag{1.4.1}$$

iar

$$p(B/A)=\frac{m-1}{m+n-1}\qquad p(B/\overline{A})=\frac{m}{m+n-1} \tag{1.4.2}$$

c) Pentru calculului lui p(B) se utilizează formula probabilității totale, relația (1.5). Rezultă că:

$$\begin{aligned}p(B)&=\frac{m-1}{m+n-1}\cdot\frac{m}{m+n}+\frac{m}{m+n-1}\cdot\frac{n}{m+n}=\\&=\frac{m(m+n-1)}{(m+n)(n+m-1)}=\frac{m}{m+n}\end{aligned} \tag{1.4.3}$$

adică p(A) = p(B) cum era de aşteptat.

Obs: *- cele două probabilități p(A) şi p(B) sunt probabilități apriori (dinainte de producerea evenimentelor). Înainte ca primul student să extragă un subiect, toți studenții, indiferent de ordinea lor, au şanse egale la a extrage un subiect uşor.*

1.5. Un tetraedru regulat are fețele colorate, una în roşu, una în galben, una în verde, iar cea de-a treia conține pe toate trei culorile. Se lasă să cadă tetraedrul pe una din fețe. Fie evenimentele:
R - fața pe care a căzut tetraedrul conține roşu; G - fața pe care a căzut tetraedrul conține galben; V - fața pe care a căzut tetraedrul conține verde.
a) cât este probabilitatea evenimentului roşu, p(R)?
b) cât este probabilitatea condiționată p(R/G)?
c) **sunt evenimentele R, G şi V independente?**

Rezolvare:
a) Probabilitatea evenimentului roşu este:

$$p(R) = \frac{2}{4} = \frac{1}{2}; \left(= \frac{\text{număr de fețe ce conțin roşu}}{\text{număr total de fețe}}\right) \quad (1.5.1)$$

b) Probabilitatea de a se realiza evenimentului roşu dacă s-a realizat galben este:

$$p(R/G) = \frac{1}{2} \quad (1.5.2)$$

deoarece una din două fețe ce conțin galben conține şi roşu.
c) Pentru ca evenimentele R, G şi V să fie independente trebuie să fie îndeplinite relațiile:

$$\begin{cases} p(R \cap G) = p(R) \cdot p(G) \\ p(R \cap V) = p(R) \cdot p(V) \\ p(G \cap V) = p(G) \cdot p(V) \\ p(G \cap R \cap V) = p(G) \cdot p(R) \cdot p(V) \end{cases} \quad (1.5.3)$$

Aplicând relația (1.1), găsim că:

$$p(R) = p(G) = p(V) = \frac{1}{2} \quad p(R \cap G) = p(R \cap V) = p(G \cap V) = \frac{1}{4}$$
$$p(G \cap R \cap V) = \frac{1}{4} \quad (1.5.4)$$

Cum ultima relație din (1.5.3) nu este verificată evenimentele R, G şi V nu sunt independente.

1.6. Pot fi simultan două evenimente şi incompatibile şi independente?

Rezolvare:

Fie A şi B două evenimente. Incompatibilitatea lor presupune ca:

$$p(A \cap B) = 0 \tag{1.6.1}$$

iar independenţa:

$$p(A \cap B) = p(A) \cdot p(B) \tag{1.6.2}$$

Din cele două relaţii rezultă că cele două evenimente pot fi independente, fiind incompatibile, doar dacă unul este de probabilitate nulă. Altfel spus, două evenimente de probabilităţi nenule pot fi independente doar dacă sunt compatibile.

1.7. O imagine alb negru se compune din 1024 x 256 pixeli. Fiecare pixel poate avea un nivel de gri dintre 64 posibile. Aflaţi informaţia furnizată de: a) un pixel; b) întreaga imagine.

Rezolvare:

a) Considerând egal probabile nivelele de gri, conform definiţiei informaţiei unui eveniment:

$$i(\text{pixel}) = -\log_2 p(\text{nivel gri}) = -\log_2 \frac{1}{64} = 6 \text{ biţi} \tag{1.7.1}$$

c) Întreaga imagine furnizează de 1024 x 256 ori mai multă informaţie:

$$i(\text{imagine}) = 1024 \cdot 256 \cdot i(\text{pixel}) \cong 1{,}5 \cdot 10^6 \text{ biţi} \tag{1.7.2}$$

1.8. a) Care este numărul de întrebări minime necesare pentru a afla un număr necunoscut N_x cuprins între 1 şi 1000? Întrebările pot fi de genul :

“Numărul N_x este mai mare decât N_p (nominalizat)?”

b) Cât este primul prag N_{p1} şi cât este informaţia conţinută de răspunsul la întrebarea:

“Numărul N_x este mai mare decât 348?”

Rezolvare:

a) Informaţia necesară pentru a afla numărul N_x necunoscut este:

$$i_N = \log_2 \frac{1}{p(N_x)} = \log_2 \frac{1}{1/1000} = \log_2 1000 \text{ biţi} \tag{1.8.1}$$

Informaţia obţinută prin răspunsul la o întrebare este:

$$i_{N_p} = p\left(A_{N_p}\right)\cdot \log_2 \frac{1}{p\left(A_{N_p}\right)} + p\left(\overline{A_{N_p}}\right)\cdot \log_2 \frac{1}{p\left(\overline{A_{N_p}}\right)} \tag{1.8.2}$$

unde A_{Np} este evenimentul prin care numărul N_x este mai mare decât pragul N_p. Evident:

$$p\left(A_{N_p}\right) + p\left(\overline{A_{N_p}}\right) = 1 \tag{1.8.3}$$

şi putem scrie:

$$x = p\left(A_{N_p}\right) = 1 - p\left(\overline{A_{N_p}}\right) \tag{1.8.4}$$

de unde

$$i_{N_p} = i(x) = x\log_2 \frac{1}{x} + (1-x)\log_2 \frac{1}{1-x} \quad \text{cu } x \in [0,1] \tag{1.8.5}$$

Funcția i(x) îşi atinge maximul dacă

$$x = \frac{1}{2} \tag{1.8.6}$$

Valoarea maximului este:

$$i_m = 1\,\text{bit} \tag{1.8.7}$$

şi corespunde unui prag:

$$N_{pm} = 499 \tag{1.8.8}$$

Aşadar, dacă pragul este ales la jumătatea intervalului în care se cunoaşte că se află N_x informația obținută prin răspunsul la întrebare (în cazul cel mai defavorabil) este maximă şi egală cu 1 bit.

Cert, numărul minim de întrebări este:

$$n = \left[\frac{i_N}{i_m}\right]_{sup} = \left[\frac{\log_2 1000\ \text{biți}}{1\ \text{bit}}\right]_{sup} = [9{,}96]_{sup} = 10 \tag{1.8.9}$$

unde $[y]_{sup}$ denotă numărul întreg superior lui y.

__Obs:__ *numărul n = 10 găsit cu raționamentul anterior este "minim" în cazul general, acest lucru însemnând că indiferent de N_x prin 10 întrebări de tipul celei din enunț (cu pragurile alese "la jumătate") se află valoarea sa. Există cazuri particulare când, pentru anumite valori a lui N_x, să fie suficiente mai puține întrebări (ex: $N_x = 1$ şi $N_p = 1$) dar pentru astfel de cazuri intervine "şansa"!*

b) Din relația (1.8.8) $N_{p1} = 499$;
Dacă pragul se alege (la prima întrebare) $N_{p1} = 348$ avem că

$$p(A)=\frac{1000-348}{1000}=0{,}652 \quad \text{şi} \quad p(\overline{A})=0{,}348 \tag{1.8.10}$$

de unde:

$$i(348)=0{,}652\cdot\log_2\frac{1000}{652}+0{,}348\cdot\log_2\frac{1000}{348}=0{,}932\,\text{biţi} \tag{1.8.11}$$

Rezultatul cuprins în relaţia (1.8.11) arată că dacă pragurile nu vor fi alese “la jumătate” există posibilitatea să nu fie suficiente 10 întrebări !

1.9. Câte cântăriri sunt minim necesare pentru a preciza o monedă falsă din 12 şi dacă moneda este mai grea sau mai uşoară? Moneda falsă diferă prin greutate iar cântăririle se fac pe o balanţă cu talere.

Rezolvare:

Informaţia necesară pentru a soluţiona problema este:

$$i_{nec}=\log_2 12+\log_2 2=\log_2 24\,\text{biţi} \tag{1.9.1}$$

unde $\log_2 12$ este informaţia necesară pentru a afla moneda din 12, iar $\log_2 2$ este informaţia necesară pentru a afla dacă moneda este mai grea sau mai uşoară.

Informaţia maximă furnizată de o cântărire este:

$$i_{cm}=\log_2 3\,\text{biţi} \tag{1.9.2}$$

şi se atinge dacă cele trei variante rezultat al unei cântăriri cu balanţa(A-balanţa se înclină spre dreapta, B-balanţa se înclină spre stânga, C-balanţa nu se înclină) sunt egal probabile:

$$p(A)=p(B)=p(C) \tag{1.9.3}$$

Numărul de cântăriri cerut va fi:

$$n=\left[\frac{i_{nec}}{i_{cm}}\right]_{sup} \quad \text{sau} \quad i_{nec}\le n\cdot i_{cm} \quad \text{sau} \quad \log_2 24\le\log_2 3^n \tag{1.9.4}$$

cu soluţia:

$$n_{min}=3 \tag{1.9.5}$$

Obs: *- relaţia (1.9.2) este valabilă doar dacă cele trei variante rezultat ale cântăririi sunt egal probabile. Astfel dacă cele 12 monezi se notează cu M_1, M_2, ..., M_{12} prima cântărire constă în a compara pe $M_1+M_2+M_3+M_4$ cu $M_5+M_6+M_7+M_8$. O posibilă rezolvare a problemei poate fi:*

A_1 – moneda falsă este - mai uşoară şi este M_1, M_2, M_3, sau M_4.

- mai grea şi este M_5, M_6, M_7, sau M_8.

B_1 – moneda falsă este - mai grea şi este M_1, M_2, M_3, sau M_4.

-mai uşoară şi este M_5, M_6, M_7, sau M_8.

C_1 – moneda falsă este mai grea sau mai uşoară şi este M_9, M_{10}, M_{11}, sau M_{12}.

Dacă rezultatul primei cântăriri este A_1, indicele 1 semnifică prima cântărire, A rezultatul ei, atunci se compară $M_1+M_2+M_5$ cu $M_3+M_4+M_6$

- dacă la a doua cântărire rezultă A_2 atunci fie M_1 sau M_2 e mai uşoară fie M_6 e mai grea şi se compară în continuare M_1 cu M_2;

- dacă la a doua cântărire rezultă B_2 atunci fie M_3 sau M_4 e mai uşoară fie M_5 e mai grea şi se compară în continuare M_3 cu M_4;

- iar dacă la a doua cântărire rezultă C_2 atunci fie M_7 e mai grea fie M_8; se compară M_7 cu M_8. În mod analog pentru B_1 şi C_1.

Obs: *- relaţia (1.9.4) indică că problema ar putea fi rezolvată şi pentru 13 monezi în locul celor 12:*

$$\log_2 26 \le \log_2 27 = \log_2 3^n \quad \text{cu} \quad n = 3$$

În realitate problema cu 13 monezi nu poate fi complet soluţionată din 13 cântăriri pentru că nu se poate asigura echiprobabilitatea rezultatelor.

1.10. Cât este informaţia obţinută în diferitele variante de rezolvare a problemei cu 12 monezi? Dar cu 13 monezi?

Răspuns:

Pentru varianta A_1 A_2 A_3 (cu 12 monezi):

$$\mathrm{I} = \log_2 3 + \left[2 \cdot \frac{3}{8} \cdot \log_2 \frac{8}{3} + \frac{2}{8} \cdot \log_2 \frac{8}{2}\right] + \log_2 3 = 4{,}7312 \text{ biţi}$$

Obs: *informaţia obţinută la a doua cântărire este mai puţină decât cea presupusă, $\log_2 3$ biţi.*

1.11. Un convertor analog-numeric (CAN) converteşte tensiunea de intrare U_x într-un număr N_x conform relaţiei:

$$N_x = \left[\frac{U_x}{q}\right] \qquad \textbf{(1.11.1)}$$

unde U_x poate lua valori între 0 şi U_{max} = 12,8 V; q este cuanta conversiei, q = 50mV; [y] semnifică partea întreagă a lui y.

a) Câtă informaţie furnizează la ieşirea sa CAN-ul prin fiecare număr generat şi câtă informaţie se pierde ?

b) Dacă CAN-ul foloseşte pentru conversie, în mai mulţi paşi succesivi, un comparator, să se stabilească câtă informaţie furnizează comparatorul la un pas şi câţi paşi sunt necesari pentru efectuarea conversiei ?

c) Care sunt tensiunile de prag ale comparatoarelor utilizate, U_{pi}, în conversia tensiunii U_x = 7,43 V, în cazul în care conversia se face într-un număr minim de paşi ?

Răspuns:

a) i = 8 biţi; în mod ideal U_x conţine + ∞ informaţie.

În realitate măsurarea unei tensiuni U_x este limitată (ca şi precizie) fie de rezoluţia aparatelor fie de nivelul zgomotului.
b) $i_c = 1$ bit; 8 paşi;
c) $U_{p1} = 6{,}4$ V; $U_{p2} = 9{,}6$ V; $U_{p3} = 8$ V; $U_{p4} = 7{,}2$ V; $U_{p5} = 7{,}6$ V; $U_{p6} = 7{,}4$ V; $U_{p7} = 7{,}5$ V; $U_{p8} = 7{,}45$ V.

1.12. Un CAN de mare viteză utilizează 128 de comparatoare pentru a face conversia tensiunilor de intrare din intervalul [-12,8V, +12,8V] la o rezoluție de q=100mV. Determinați "redundanța" în componente a CAN-ului.

<u>Rezolvare:</u>
Numărul de "răspunsuri" distincte ale CAN –ului pentru o tensiune U_x ce variază de la - 12,8 V la +12,8 V este:

$$N = \frac{U_{max} - U_{min}}{q} = 256 = 2^8 \qquad (1.12.1)$$

Probabilitatea ca o tensiune U_x să genereze un răspuns din cele N posibile este:

$$p_0 = 1/N = 2^{-8} \qquad (1.12.2)$$

iar informația conținută de un "răspuns" este:

$$i_0 = -\log_2 p_0 = 8\,\text{biți} \qquad (1.12.3)$$

Pentru că un comparator "alege" o variantă din două posibile, informația furnizată de el este:

$$i_c = -\log_2 \frac{1}{2} = 1 \text{ bit} \qquad (1.12.4)$$

Aşadar în mod ideal, pentru o conversie sunt suficiente:

$$n = i_0/i_c = 8 \text{ comparatoare} \qquad (1.12.5)$$

de unde rezultă că redundanța este de 120 de comparatoare.

***<u>Obs</u>**: Motivația utilizării a mai multor comparatoare constă în faptul că, la folosirea doar a 8 comparatoare, sunt necesare tensiuni de prag variabile în funcție de tensiunea U_x, lucru ce scade viteza de conversie.*

1.13. La o transmisie numerică informația utilă se transmite prin secvențe binare de câte n biți, numite cuvinte. Câtă informație este necesară pentru a preciza poziția unui bit eronat din cei n? Dar pentru doi biți eronați? Exemplificare pentru n = 8.

Rezolvare:

Informația cerută se calculează cu formula:

$$i = \log_2 n \ \text{(biți)} \tag{1.13.1}$$

Se consideră că transmiterea (implicit eronarea) unui bit este independentă de a celorlalți. Aşadar, pentru un bit eronat i_1 = 3 biți; pentru doi biți eronați $i_2 = \log_2 C_8^2 \cong 4{,}8$ biți.

1.14. Aceeaşi întrebare ca şi la problema 1.13, cu diferența că este o transmisie ternară.

Rezolvare:

În cazul transmisiei ternare corecția unui simbol (ternar) eronat presupune pe lângă aflarea poziției sale (fapt ce necesită aceeaşi informație ca şi la secvența binară), şi valoarea simbolului inițial, valoare ce poate fi una din două posibile. Cu aceste justificări, răspunsurile vor fi:

$$\begin{aligned} i_1 &= \log_2 n \text{ (pentru aflarea poziției)} \\ &+ \log_2 2 \text{ (pentru aflarea valorii)} = 4 \text{ biți} \end{aligned} \tag{1.14.1}$$

$$i_2 = \log_2 C_n^2 + \log 2^2 \cong 6{,}8 \text{ biți} \tag{1.14.2}$$

În relația (1.14.2) s-au adăugat 2 biți de informație necesari aflării valorii “adevărate” pentru două simboluri ternare eronate.

1.15. La o transmisie binară informația utilă este transmisă prin cuvinte de n=8 biți printr-un canal cu rata erorii $p=10^{-3}$. Câtă informație, în medie pe un cuvânt, este necesară pentru a face:

a) detecție de o eroare pe cuvânt?

b) detecție de una sau două erori pe cuvânt?

Rezolvare:

a) Probabilitatea ca un cuvânt să fie eronat într-o poziție este:

$$\begin{aligned} p_1 &= C_n^1 \cdot p \cdot (1-p)^{n-1} \cong C_n^1 \cdot p \cdot [1 - p(n-1)] = \\ &= 8 \cdot 10^{-3} \cdot 0{,}993 = 7{,}944 \cdot 10^{-3} \end{aligned} \tag{1.15.1}$$

p_1 este şi probabilitatea ca receptorul să detecteze o eroare. Cert $1-p_1$ este probabilitatea ca receptorul să ia decizia că nu există eroare (incluzând cazurile corecte şi false).

Fie a şi b $\in$ N astfel încât:

$$p_1 = a/b \text{ şi } 1 - p_1 = (b-a)/b \tag{1.15.2}$$

Aşadar din b cuvinte transmise a sunt detectate cu eroare. Pentru a detecta un astfel de cuvânt este necesară informația:

$$i_{1d} = -\log_2 p_1 \tag{1.15.3}$$

b-a cuvinte din cele b sunt considerate neeronate, pentru un astfel de cuvânt fiind necesară informația:

$$i_{1n} = -\log_2 (1 - p_1) \tag{1.15.4}$$

Concluzionând pentru b cuvinte recepționate este necesară informația:

$$b \cdot i_1 = a \cdot i_{1d} + (b\text{-}a) \cdot i_{1n} \tag{1.15.5}$$

iar pentru un singur cuvânt recepționat, pentru a face detecție de o eroare:

$$\begin{aligned} i_1 &= \frac{a}{b} \cdot i_{1d} + \frac{b-a}{b} \cdot i_{1n} = \\ &= -p_1 \cdot \log_2 p_1 - (1-p_1) \cdot \log_2 (1-p_1) = \\ &= 0{,}0668 \text{ biți/cuvânt} \end{aligned} \tag{1.15.6}$$

b) Şi în acest caz informația cerută are aceeaşi formă cu (1.15.6) doar că diferă p_1:

$$i_2 = -p_2 \cdot \log_2 p_2 - (1-p_1) \cdot \log_2 (1-p_2) \tag{1.15.7}$$

unde p_2 este probabilitatea ca un cuvânt să fie eronat într-o poziție sau în două poziții:

$$p_2 = C_n^1 \cdot p \cdot (1-p)^{n-1} + C_n^2 \cdot p^2 \cdot (1-p)^{n-2} = 7{,}972 \cdot 10^{-3} \tag{1.15.8}$$

de unde rezultă pentru i_2 valoarea:

$$i_2 = 0{,}067 \text{ biți/cuvânt} \tag{1.15.9}$$

Obs: *Detecția prezenței erorilor presupune a decide între două variante: există sau nu există erori în cuvântul în cauză. Informația furnizată de o astfel de decizie (una din două posibile) este, în medie, subunitară, încât problema detecției din punct de vedere strict al informației este solvabilă cu 1 bit (de informație) pe cuvânt, indiferent de p sau n.*

Cap. 2 Surse de informaţie

Dicţionar:

- **eficienţă** –raportul dintre ansamblul efectelor utile (rezultatelor) ce se obţin de pe urma unei activităţi şi totalul eforturilor;
- **entropie (limba greacă "entropie"=întoarcere, schimbare)** –mărime ce indică gradul de nedeterminare asupra unui sistem;
- **extensie** –dezvoltare, creştere, amplificare, extindere;
- **graf** –cuplu format dintr-o mulţime ale cărei elemente se numesc vârfuri şi dintr-o funcţie care asociază, la orice pereche ordonată de vârfuri (primul numit sursă iar al doilea adresă), un element, numit săgeată, al unei alte mulţimi.
- **redundanţă (redondanţă)** –1.abundenţă (inutilă) de cuvinte, de figuri retorice, de imagini pentru exprimarea unei idei; 2.excesul de informaţie faţă de strictul necesar;
- **stare** –1.situaţie în care se află ceva sau cineva; 2.ansamblul valorilor instantanee a parametrilor ce caracterizează complet un sistem dat;
- **staţionar** –care rămâne în aceeaşi stare;

Definiţii:

- **Sursă de informaţie text** = sursa de informaţie, de regulă considerată fără memorie, având drept simboluri caracterele distincte din textul dat (opţional se pot include semne de punctuaţie, pauzele dintre cuvinte, cifrele, etc).. Probabilităţile diferitelor simboluri vor fi ponderile lor în text.
- **Extensia unei surse =** fiind dată o SDFM, S cu N simboluri, se defineşte extensia de ordin n a sursei, notată S^n, sursa având un număr de N^n simboluri obţinute din toate combinaţiile posibile de mesaje compuse din n simboluri ale sursei S.
- **Sursă cu memorie (Markov) de m paşi =** sursa de informaţie cu N simboluri pentru care emisia celui de-al m+1-lea simbol depinde de cele m simboluri anterior emise.
- **Starea sursei Markov de m paşi =** setul ultimelor m simboluri emise. Dacă sursa poate emite M simboluri şi este cu memorie de m paşi atunci admite M^m stări.
- **Probabilitate de trecere $p(S_j/S_i)$ =** probabilitatea ca sursa Markov să treacă în starea $S_j = s_{km-1}\, s_{km-2} \ldots s_{k0}$ (prin emiterea simbolului s_{k0}), dacă se află în starea $S_i = s_{km-2}\, s_{km-1} \ldots s_{k1}$.
- **Matricea de tranziţie,T =** o matrice pătrată de ordin M^m ce conţine toate probabilităţile de trecere;
- **Graful sursei cu memorie =** graful ale cărui noduri reprezintă stările sursei, iar coardele (săgeţile) ce leagă nodurile probabilităţile de trecere.
- **Sursă Markov staţionară =** probabilităţile de trecere în n paşi converg către limite independente de starea iniţială când $n \to \infty$.
- **Stare de staţionaritate (a unei surse Markov staţionare) =** un vector, P^* de dimensiune M ce conţine probabilităţile p_j^*, **j = 1, M** ce verifică sistemul:

$$\begin{cases} P^* \cdot T = P^* \\ \sum_{j=1}^{M} p_j^* = 1 \end{cases} \tag{2.1}$$

Probabilitatea p_j^* reprezintă şansa ca sursa Markov, după n paşi cu $n \to \infty$ să se găsească în starea S_j.

Notaţii:

S –sursă de informaţie;
N –numărul de simboluri;
$\mathbf{S_i}$, **i = 1÷ N** –simbolurile sursei;
$\mathbf{p_i}$, **i = 1÷N** –probabilităţile de emisie a simbolurilor S_i;
$\mathbf{S^m}$ –extensia de ordin m a sursei SDFM, S;
$\mathbf{S_j}$, **j = 1÷** $\mathbf{N^m}$ –stările sursei cu memorie, S;
T –matricea de tranziţie;
ppm –părţi per milion (1ppm=10^{-6}).

Abrevieri:

SDFM –Sursă Discretă Fără Memorie.

Breviar teoretic:

1. Mărimi ce caracterizează sursa de informaţie:

--entropia SDFM:

$$H(S)=\sum_{i=1}^{N} p_i \log_2 \frac{1}{p_i} \tag{2.2.a}$$

--entropia sursei cu memorie:

$$H(S)=\sum_{j=1}^{N}\sum_{i=1}^{N} p_j^* p(S_i/S_j)\log_2 \frac{1}{p(S_i/S_j)} \tag{2.2.b}$$

-entropia maximă:

$$H_{max}(S)=\log_2 N \tag{2.3}$$

--eficienţa:

$$\eta_s = H(S)/H_{max}(S) \tag{2.4}$$

--redundanţa:

$$R(S) = H_{max}(S) - H(S) \tag{2.5}$$

--redundanţa relativă:

$$\rho_s = R(S)/H_{max}(S) \tag{2.6}$$

2. Formulă pentru schimbarea bazei logaritmului:

$$\log_2 x = \frac{\ln x}{\ln 2} \tag{2.7}$$

3. Condiţia de existenţă a stării de staţionaritate pentru sursă cu memorie:

$\exists\, n_0 \in N$ astfel încât T^{n0} să fie o matrice regulată (are toate elementele strict pozitive)

2.1. O sursă de informaţie discretă are entropia egală cu 4,8 biţi/simbol şi redundanţa relativă 9,8%. Câte simboluri are sursa?

Rezolvare:

Din relaţiile de definiţie a entropiei maxime şi a redundanţei relative pentru o sursă discretă fără memorie:

$$H_{max} = \log_2 N \tag{2.1.1}$$

$$\rho = 1 - \frac{H(S)}{H_{max}} \tag{2.1.2}$$

unde: H(s)-entropia sursei; H_{max}-entropia maximă, ρ-redundanţa relativă, N-nr. de simboluri a sursei, găsim că:

$$H_{max} = \frac{H(S)}{1-\rho} = \frac{4{,}8 \cdot 100}{90{,}2} \tag{2.1.3}$$

şi:

$$N = 2^{H_{max}} = 40 \tag{2.1.4}$$

2.2. Fie sursa de informație S:

$$S = \begin{pmatrix} S_1 & S_2 & S_3 & S_4 & S_5 \\ \frac{1}{8} & \frac{1}{2} & \frac{1}{16} & \frac{1}{4} & \frac{1}{16} \end{pmatrix}$$

Se cere să se calculeze:

a) informația medie pe simbol, H(s);

b) valoarea maximă a entropiei sursei H_{max};

c) eficiența sursei η_s;

d) redundanța sursei R;

e) redundanța relativă ρ.

Rezolvare:

a) Informația medie pe simbol sau entropia sursei H(S) se calculează cu formula:

$$H(S) = \sum_{i=1}^{N} p(s_i) \log_2 \frac{1}{p(s_i)} \qquad (2.2.1)$$

unde: N – numărul de simboluri a sursei S;

s_i, $i = \overline{1, N}$, – simbolurile sau mesajele sursei S;

$p(s_i)$ – probabilitatea ca sursa să emită simbolul s_i.

Înlocuind rezultă că:

$$\begin{aligned} H(S) &= \frac{1}{8}\log_2 8 + \frac{1}{2}\log_2 2 + \frac{1}{16}\log_2 16 + \frac{1}{4}\log_2 4 + \frac{1}{16}\log_2 16 \\ &= \frac{3}{8} + \frac{1}{2} + \frac{4}{8} + \frac{2}{4} + \frac{4}{16} = \frac{3+4+4+4+2}{8} = \frac{17}{8} \\ &= 2{,}125 \text{ biți/simbol} \end{aligned} \qquad (2.2.2)$$

b) Entropia maximă este:

$$H_{max} = \log_2 N = \log_2 5 \cong 2{,}323 \text{ biți/simbol} \qquad (2.2.3)$$

c) Eficiența sursei este:

$$\eta_s = \frac{H(S)}{H_{max}} = 91{,}5\% \qquad (2.2.4)$$

d) Redundanța sursei R este diferența dintre entropia maximă şi entropia sursei date:

$$R = H_{max} - H(S) = 0{,}197 \text{ biți/simbol} \qquad (2.2.5)$$

e) Redundanța relativă, cu formula de calcul (2.6), are valoarea:

$$\rho = 8{,}5\% \qquad (2.2.6)$$

2.3. Sursa de informaţie discretă şi fără memorie S are drept simboluri caracterele distincte din următoarea propoziţie (sursă text):

NU UITA NICI PAUZELE DINTRE CUVINTE .

Probabilităţile simbolurilor (caracterelor distincte) sunt ponderile lor din text. Se cere:

a) tabloul simboluri-probabilităţi pentru sursa S;

b) entropia H(S), redundanţa R şi eficienţa η_s sursei S;

c) probabilitatea ca sursa să emită mesajul "AUR".

Rezolvare:

a) tabloul simboluri-probabilităţi este:

$$S = \begin{pmatrix} A & C & D & E & I & L & N & P & R & T & U & V & Z & \_ & . \\ \frac{2}{36} & \frac{2}{36} & \frac{1}{36} & \frac{4}{36} & \frac{5}{36} & \frac{1}{36} & \frac{4}{36} & \frac{1}{36} & \frac{1}{36} & \frac{3}{36} & \frac{4}{36} & \frac{1}{36} & \frac{1}{36} & \frac{5}{36} & \frac{1}{36} \end{pmatrix} \tag{2.3.1}$$

b) Conform relaţiei de definiţie entropia sursei este:

$$\begin{aligned} H(S) &= \sum_{i=1}^{N} p(s_i) \log_2 \frac{1}{p(s_i)} \\ &= \sum_{i=1}^{N} \frac{k_i}{36} \log_2 \frac{36}{k_i} = \sum_{i=1}^{N} \frac{k_i}{36} \log_2 36 - \sum_{i=1}^{N} \frac{k_i}{36} \log_2 k_i \end{aligned} \tag{2.3.2}$$

unde: – N = 15 numărul de simboluri a sursei S;

$-\ p_i = \frac{k_i}{36}, \quad i = \overline{1,15},$ - probabilităţile simbolurilor.

Înlocuind în (2.3.2) valorile obţinute pentru k_i din tabloul (2.3.1) găsim:

$$\begin{aligned} H(S) &= \log_2 36 - \frac{1}{36}\big(2 \cdot 2 \cdot \log_2 2 + 7 \cdot 1 \cdot \log_2 1 + 3 \cdot \log_2 3 + \\ &\quad + 2 \cdot 4 \cdot \log_2 4 + 2 \cdot 5 \cdot \log_2 5\big) = \\ &= 2 + 2 \cdot \log_2 3 - \frac{1}{36}\big(4 + 3 \cdot \log_2 3 + 16 + 10 \cdot \log_2 5\big) = \\ &= \frac{52}{36} + \frac{69}{36} \log_2 3 - \frac{10}{36} \log_2 5 = \\ &= 3{,}8373 \text{ biţi/simbol} \end{aligned} \tag{2.3.3}$$

Redundanţa şi eficienţa sursei se obţin cu relaţiile:

$$\begin{aligned} R &= H_{max} - H(S) = \log_2 N - H(S) \\ \eta_s &= \frac{H(S)}{H_{max}} = \frac{H(S)}{\log_2 N} \end{aligned} \tag{2.3.4}$$

Cunoscând N = 15 şi cu H(S) dat de (2.3.3) găsim:

$R = 0{,}06958$ biți/simbol

$\eta_s = 98{,}22\%$ (2.3.5)

c) Sursa S fiind fără memorie probabilitatea cerută va fi:

$$p(AUR) = p(A)\cdot p(U)\cdot p(R) = \frac{2}{36}\cdot\frac{4}{36}\cdot\frac{1}{36} = 1{,}715\cdot 10^{-4} \qquad (2.3.6)$$

2.4. Fie sursa de informație discretă şi fără memorie:

$$S = \begin{pmatrix} S_1 & S_2 & S_3 & S_4 & S_5 & S_6 & S_7 \\ \frac{10}{100} & \frac{20}{100} & \frac{30}{100} & \frac{5}{100} & \frac{2}{100} & \frac{8}{100} & \frac{25}{100} \end{pmatrix}$$

Să se calculeze pentru sursa S:

a) entropia H(S);

b) valoarea maximă a entropiei sursei H_{max};

c) eficiența sursei η_s;

d) redundanța sursei R;

e) redundanța relativă ρ.

<u>Răspuns:</u>

a) $H(S) = 2{,}44$ biți/simbol

b) $H_{max} = 2{,}8$ biți/simbol

c) $\eta_s = 86{,}85\%$

d) $R = 3{,}7$ biți/simbol

e) $\rho = 13{,}15\%$

2.5. Fie sursa text:

STUDENTUL * REZOLVA O PROBLEMA DE TTI .**

a) înlocuiți * cu numele şi prenumele dumneavoastră şi construiți în acest caz tabloul sursei;**

b) calculați entropia şi eficiența sursei de la punctul a);

c) cât este probabilitatea ca sursa de la a) să emită mesajul "MAR".

<u>Răspuns:</u>

pentru *** = POP ION

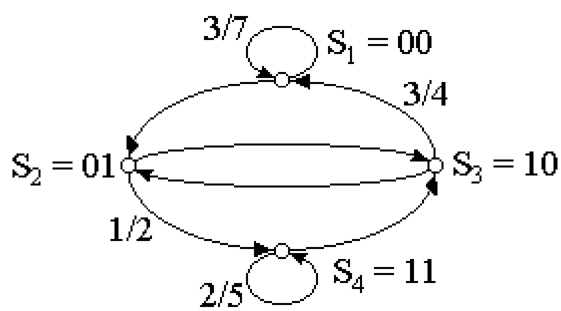
a) $$S = \begin{pmatrix} A & B & D & E & I & L & M & N & O & P & R & S & T & U & V & Z & \_ & . \\ \frac{2}{44} & \frac{1}{44} & \frac{2}{44} & \frac{4}{44} & \frac{2}{44} & \frac{3}{44} & \frac{1}{44} & \frac{2}{44} & \frac{5}{44} & \frac{3}{44} & \frac{2}{44} & \frac{1}{44} & \frac{4}{44} & \frac{2}{44} & \frac{1}{44} & \frac{1}{44} & \frac{7}{44} & \frac{1}{44} \end{pmatrix}$$

b) $H(S) = \log_2 44 - \frac{1}{44}\left(6 \cdot 2 + 6 \cdot \log_2 3 + 16 + 5 \cdot \log_2 5 + 7 \cdot \log_2 7\right) = 3{,}896$ biți/simbol

$$\eta_s = \frac{3{,}896}{\log_2 18} = 93{,}44\%$$

c) $p(MAR) = \frac{4}{44^3} = 47$ ppm

2.6. O sursă de informație binară cu memorie de doi paşi are graful corespunzător în figură. Să se afle:
a) probabilitățile de trecere nefigurate;
b) ce şanse sunt ca după transmiterea mesajului "00110" să se transmită mesajul "11".
c) matricea de tranziție şi o eventuală stare de staționaritate;
d) entropia sursei date.



<u>Rezolvare:</u>
a) Probabilitățile cerute sunt:

$$\begin{aligned} p(S_2/S_1) &= 1 - p(S_1/S_1) = \frac{2}{7} \\ p(S_3/S_2) &= 1 - p(S_4/S_2) = \frac{1}{2} \\ p(S_3/S_4) &= 1 - p(S_4/S_4) = \frac{3}{5} \\ p(S_2/S_3) &= 1 - p(S_1/S_3) = \frac{1}{4} \end{aligned} \tag{2.6.1}$$

b) Deoarece sursa are memorie de doi paşi, din mesajul transmis "00110" sunt relevanți, pentru definirea stării actuale, doar ultimii doi biți transmişi "10". Aşadar starea actuală este S_3. Din S_3 şansele ca sursa să emită "1" şi, ca atare, să treacă în starea S_2 sunt de 25% ($p(S_2/S_3)$). Prin generarea încă a unui "1" sursa va trece din starea S_2 în starea S_4. Probabilitatea acestei tranziții este de 50%. Concluzionând probabilitatea ce "însoțeşte"acest traseu $S_3 \rightarrow S_2 \rightarrow S_4$ este:

$$p(S_3 \rightarrow S_2 \rightarrow S_4) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} = 12{,}5\% \tag{2.6.2}$$

c) Matricea de tranziție este:

$$T = \begin{bmatrix} p(S_1/S_1) & p(S_2/S_1) & p(S_3/S_1) & p(S_4/S_1) \\ p(S_1/S_2) & p(S_2/S_2) & p(S_3/S_2) & p(S_4/S_2) \\ p(S_1/S_3) & p(S_2/S_3) & p(S_3/S_3) & p(S_4/S_3) \\ p(S_1/S_4) & p(S_2/S_4) & p(S_3/S_4) & p(S_4/S_4) \end{bmatrix} = \begin{bmatrix} \frac{3}{7} & \frac{4}{7} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{3}{5} & \frac{2}{5} \end{bmatrix} \tag{2.6.3}$$

O eventuală stare de staționaritate ar conține probabilitățile p_i^*, ca sursa să se afle într-o anumită stare S_i :

$$P^* = [p_1^* \ p_2^* \ p_3^* \ p_4^*] \tag{2.6.4}$$

cu P^* îndeplinind condiția:

$$P^* T = P^* \tag{2.6.5}$$

Înlocuind în (2.6.5) pe T şi P^* date de (2.6.3) şi (2.6.4) rezultă:

$$[p_1^* \ \ p_2^* \ \ p_3^* \ \ p_4^*] \begin{bmatrix} \frac{3}{7}-1 & \frac{4}{7} & 0 & 0 \\ 0 & -1 & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} & -1 & 0 \\ 0 & 0 & \frac{3}{5} & \frac{2}{5}-1 \end{bmatrix} = 0 \tag{2.6.6}$$

Sistemul de ecuații dat de (2.6.6) este compatibil nedeterminat fapt ce necesită încă o relație între probabilitățile p_i^*, relație care este:

$$[p_1^* + \ p_2^* + \ p_3^* + \ p_4^*] = 1 \tag{2.6.7}$$

Soluția sistemului obținut din (2.6.6) şi (2.6.7) este:

$$p^* = \left[\frac{63}{199} \ \ \frac{48}{199} \ \ \frac{48}{199} \ \ \frac{40}{199} \right] \tag{2.6.8}$$

d) Entropia sursei cu memorie se calculează cu formula:

$$H_M(S) = \sum_{i=1}^{2^2} \sum_{j=1}^{2} p(S_i) \cdot p(a_j/S_i) \cdot \log_2 p(a_j/S_i) \tag{2.6.9}$$

unde S_i, cu i = 1÷4, sunt cele 4 stări posibile ale sursei iar a_j, cu j = 0÷1, sunt mesajele posibil a fi emise din fiecare stare.

Pentru uşurinţa calculului entropiei, elementele cerute de suma dublă din (2.6.9) sunt date în tabelul următor :

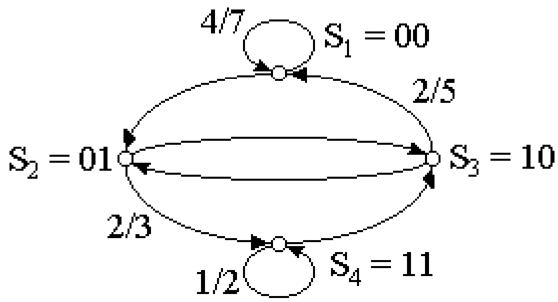
S_i	a_j	$p(S_i) = p_i^*$	$p(a_j/S_i)$
S_1 = 00	0 1	$\frac{63}{199}$	3/7 4/7
S_2 = 01	0 1	$\frac{48}{199}$	1/2 1/2
S_3 = 10	0 1	$\frac{48}{199}$	3/4 1/4
S_4 = 11	0 1	$\frac{40}{199}$	3/5 2/5

Tabelul 2.1

Înlocuind în (2.6.9) elementele din tabelul 2.1 rezultă :

$$\begin{aligned} H_M(S) &= \frac{63}{199}\left(\frac{3}{7}\log_2\frac{7}{3}+\frac{4}{7}\log_2\frac{7}{4}\right)+\frac{48}{199}\left(\frac{1}{2}\log_2\frac{2}{1}+\frac{1}{2}\log_2\frac{2}{1}\right) \\ &+\frac{48}{199}\left(\frac{3}{4}\log_2\frac{4}{3}+\frac{1}{4}\log_2\frac{4}{1}\right)+\frac{40}{199}\left(\frac{3}{5}\log_2\frac{5}{3}+\frac{2}{5}\log_2\frac{5}{2}\right) \\ &= 0{,}944 \text{ biți/simbol} \end{aligned} \tag{2.6.11}$$

2.7. O sursă de informaţie binară cu memorie de doi paşi are graful corespunzător în figură. Dacă în prezent sursa se află în starea S_i cât este probabilitatea ca după trei simboluri sursa să se regăsească tot în starea S_i?



<u>Răspuns:</u>

$$p(S_1 \to S_1 \to S_1 \to S_1) + p(S_1 \to S_2 \to S_3 \to S_1) =$$

$$= \left(\frac{4}{7}\right)^3 + \frac{3}{7}\cdot\frac{1}{3}\cdot\frac{2}{5} = 24{,}37\%$$

2.8. O sursă de informaţie binară cu memorie de un pas are probabilităţile de trecere: p(0/0) = 5/7 şi p(0/1) =3/4.

a) construiţi graful sursei date;

b) calculaţi starea de staţionaritate;

c) determinaţi entropia sursei;

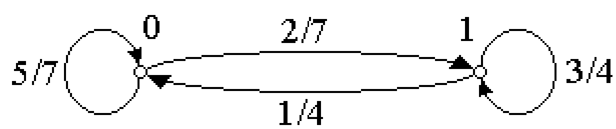
d) determinaţi entropia unei surse fără memorie având aceleaşi simboluri şi probabilităţile simbolurilor cele conţinute în starea staţionară.

Rezolvare:

a) Pentru a putea construi graful sursei cu memorie aflăm probabilităţile de trecere ce nu sunt date în enunţul problemei:

$$\begin{aligned} p(1/0) &= 1 - p(0/0) = 2/7 \\ p(1/1) &= 1 - p(0/1) = 1/4 \end{aligned} \tag{2.8.1}$$

În acest fel graful este:



b) Matricea de tranziţie este:

$$T = \begin{bmatrix} \frac{5}{7} & \frac{2}{7} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \tag{2.8.2}$$

iar starea staţionară se află rezolvând sistemul:

$$\begin{cases} \begin{bmatrix} p_0^* & p_1^* \end{bmatrix} \cdot T = \begin{bmatrix} p_0^* & p_1^* \end{bmatrix} \\ p_0^* + p_1^* = 1 \end{cases} \tag{2.8.3}$$

Rezultă:

$$\begin{cases} p_0^* = \dfrac{7}{15} \\ p_1^* = \dfrac{8}{15} \end{cases} \tag{2.8.4}$$

c) Entropia sursei cu memorie este dată de relaţia:

$$H_M(S) = \sum_{i=0}^{1} \sum_{j=0}^{1} p(S_i) \cdot p(a_j/S_i) \cdot \log_2 p(a_j/S_i) \tag{2.8.5}$$

unde $S_0 = 0$; $S_1 = 1$; $a_0 = 0$; $a_1 = 1$

$$p(S_0) = p_0^* = \frac{7}{15} \qquad p(S_1) = p_1^* = \frac{8}{15}$$

iar probabilitățile $p(a_j/S_i)$ sunt cele indicate de graf.
Înlocuind în (2.8.5) rezultă:

$$H_M(S) = \frac{7}{15}\left(\frac{5}{7}\log_2\frac{7}{5} + \frac{2}{7}\log_2\frac{7}{2}\right) + \frac{8}{15}\left(\frac{1}{4}\log_2\frac{4}{1} + \frac{3}{4}\log_2\frac{4}{3}\right)$$
$$= 0{,}8355 \text{ biți/simbol} \tag{2.8.6}$$

d) O sursă fără memorie conformă cu cerințele problemei are tabloul:

$$S = \begin{pmatrix} S_1 & S_2 \\ \frac{7}{15} & \frac{8}{15} \end{pmatrix} \tag{2.8.7}$$

iar entropia:

$$H(S) = \frac{7}{15}\log_2\frac{15}{7} + \frac{8}{15}\log_2\frac{15}{8} = 0{,}9968 \text{ biți/simbol} \tag{2.8.8}$$

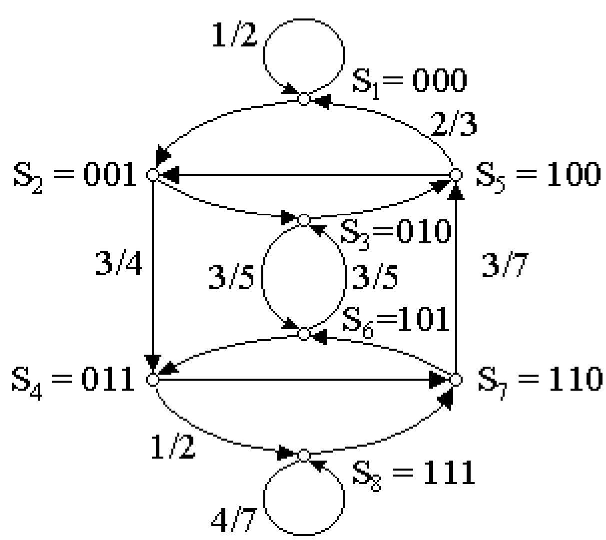
Comparând cele două entropii găsite $H_M(S)$ şi $H(S)$ se observă că ultima este mai mare.

2.9. O sursă de informație binară cu memorie de trei paşi are graful corespunzător în figură. Să se afle:

a) matricea de tranziție;

b) entropia sursei;

c) probabilitatea ca sursa, aflată în starea S_2, după emiterea a trei simboluri să ajungă în starea S_3.



Răspuns:

a)

$$T=\begin{bmatrix}\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{2}{5} & \frac{3}{5} & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2}\\ \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & \frac{3}{5} & \frac{2}{5} & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{3}{7} & \frac{4}{7} & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{7} & \frac{4}{7}\end{bmatrix}$$

b) $p^* = \left[\frac{128}{959}\ \frac{96}{959}\ \frac{105}{959}\ \frac{126}{959}\ \frac{96}{959}\ \frac{135}{959}\ \frac{126}{959}\ \frac{147}{959}\right]$

$H_M(S) = 0{,}961454$ biți/simbol

c) $p(S_2 \to S_3 \to S_6 \to S_3) = \frac{9}{100} = 9\%$

2.10. Fie sursa de informație fără memorie având tabloul simboluri-probabilități:

$$S=\begin{pmatrix} a & b & c & d\\ 0{,}2 & 0{,}4 & 0{,}1 & x\end{pmatrix}$$

Se cere:

a) valoarea lui x;

b) entropia şi eficiența sursei S;

c) tabloul sursei de informație S^2(extensia de ordin a lui S);

d) entropia şi eficiența sursei S^2.

<u>Rezolvare:</u>

a) Deoarece:

$$\sum_{i=1}^{N} p_i = 1 \text{ rezultă } x = 0{,}3 \qquad (2.10.1)$$

b)
$$H(S) = \sum_{i=1}^{N} p_i \log_2 \frac{1}{p_i} = 1{,}846 \text{ biți/simbol – entropia} \qquad (2.10.2)$$

$$\eta_s = \frac{H(S)}{H_{max}} = \frac{H(S)}{\log_2 N} = 92{,}32\% \text{ – eficiența} \quad (2.10.3)$$

b) Extensia unei surse S se obține asociind câte două mesaje ale sursei S. Astfel:

$$S^2 = \begin{pmatrix} aa & ab & ac & ad & ba & bb & bc & bd & ca & cb & cc & cd & da & db & dc & dd \\ \frac{4}{100} & \frac{8}{100} & \frac{2}{100} & \frac{6}{100} & \frac{8}{100} & \frac{16}{100} & \frac{4}{100} & \frac{12}{100} & \frac{2}{100} & \frac{4}{100} & \frac{1}{100} & \frac{3}{100} & \frac{6}{100} & \frac{12}{100} & \frac{3}{100} & \frac{9}{100} \end{pmatrix}$$

(2.10.4)

c) $$H(S^2) = 2 \cdot H(S) = 3{,}6921 \text{ biți/simbol} \quad (2.10.5)$$

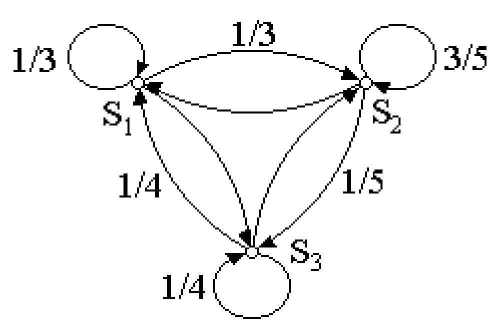
$$\eta_{S^2} = \frac{H(S^2)}{H_{max}(S^2)} = \frac{2 \cdot H(S)}{\log_2 N^2} = \eta_S \quad (2.10.6)$$

Obs: *Entropia extensiei de ordinul doi a sursei S este dublul entropiei sursei date, eficiența păstrându-se. Prin generalizare se poate arăta că:*

$$H(S^m) = m \cdot H(S)$$
$$\eta_{S^m} = \eta_S \quad (2.10.7)$$

unde S^m este extensia de ordin m a sursei S.

2.11. O sursă ternară cu memorie de un pas are graful din figură
a) să se afle cele trei probabilități de trecere necunoscute;
b) calculați starea de staționaritate;
c) calculați entropia sursei date;
d) pentru ce valori ale probabilităților de trecere sursa este practic fără memorie?



Răspuns:

a) $p(S_1/S_2) = \frac{1}{5} \quad p(S_2/S_3) = \frac{1}{2} \quad p(S_3/S_1) = \frac{1}{3}$

b) $p_1^* = \frac{12}{49}$ $p_2^* = \frac{25}{49}$ $p_3^* = \frac{12}{49}$

c)

Starea actuală	Starea viitoare	Probabilitatea stării staționare	Probabilitatea de trecere
S_1	S_1 S_2 S_3	12/49	1/3 1/3 1/3
S_2	S_1 S_2 S_3	25/49	1/5 3/5 1/5
S_3	S_1 S_2 S_3	12/49	1/4 2/4 1/4

Tabelul 2.2

$H_M(S) = 1{,}3325$ biți/simbol

d) $p(S_i/S_j) = \alpha$ $p(S_i/S_j) = \beta$ $p(S_i/S_j) = \gamma$ $\forall\, j = 1,2,3$

cu $\alpha, \beta, \gamma \in R_+$ $\alpha + \beta + \gamma = 1$

2.12. O SDFM are N = 32 simboluri şi eficiența de 75%. Cât sunt entropia, redundanța şi redundanța relativă a sursei date?

Răspuns:

$H(S) = 3{,}75$ biți/simbol $R = 1{,}25$ biți/simbol $\rho = 25\%$

Cap. 3 Codarea sursei

Dicținoar:

- cod (limba latină "codex"- culegere) - 1.ansamblul unor reguli; 2.culegere de legi; 3.sistem de semnale sau semne convenționale cu semnificații bine precizate, ale căror combinații sunt folosite pentru transmiterea unor mesaje;

- compresie - reducere a volumului.

Definiții:

- codare = procesul de atribuire a unor succesiuni (secvențe) de simboluri elementare mesajelor (simbolurilor) unei surse de informație;

- **alfabetul codării** = totalitatea simbolurilor elementare cu ajutorul cărora se pot construi succesiunile (secvențele);

- cuvânt de cod = o secvență (succesiune) atribuită prin procesul de codare a unui mesaj al sursei de informație;

- **cod** = totalitatea cuvintelor de cod;

- **cod - binar** = alfabetul este binar: {0, 1};
    - **bloc** = cuvintele au aceeaşi lungime;
    - **unic decodabil** = fiecărei succesiuni de cuvinte de cod îi corespunde o unică succesiune de simboluri a sursei;
    - **instantaneu** = nici un cuvânt de cod nu este prefixul altui cuvânt;

- **graful de codare (graful codului)** = un arbore ce creşte, dintr-un punct inițial (numit sursă), prin m ramuri la fiecare nod (m - numărul de simboluri elementare din alfabet) şi are la finele oricărei ramuri (numită frunză) câte un simbol (mesaj al sursei). Fiecare din cele m ramuri ce pleacă dintr-un nod este notată cu unul dintre simbolurile elementare ale afabetului codului. Secvența obținută prin asocierea simbolurilor elementare ataşate unei căi, ce pleacă de la sursă şi ajunge la o frunză, este cuvântul de cod ce codează simbolul (mesajul) sursei ataşat acelei frunze;

- **eficiența codării** = raportul între lungimea medie minimă a cuvintelor unui cod ce ar putea coda sursa dată şi lungimea medie a cuvintelor codului dat;

- compresia = procedeul de recodare a unei surse cu scopul de a obține o eficiență superioară primului cod;

- **factor de compresie** = raportul eficiențelor codurilor, comprimat şi inițial.

Notații:

L - lungimea medie a cuvintelor codului;

η_c - eficiența codării;

F - factor de compresie.

Abrevieri:

LZ -77 - algoritmul de compresie Lempel -Ziv '77.

Breviar teoretic:

1. Teorema I-a a lui Shannon: pentru orice sursă de informație S, printr-o codare pe grupe de n simboluri, poate fi făcută o codare absolut optimală dacă $n \to \infty$;

2. Lungimea medie a cuvintelor codului:

$$L = \sum_{i=1}^{N} p_i l_i \tag{3.1}$$

unde: l_i - lungimea cuvântului ataşat simbolului sursei;

s_i=numărul de simboluri elementare din componența câmpului respectiv;

3. Eficiența codării se poate calcula după formula:

$$\eta_c = \frac{H(S)}{L} \tag{3.2}$$

3.1. Fie sursa de informaţie discretă şi fără memorie:

$$S = \begin{pmatrix} a & b & c & d \\ 0{,}4 & 0{,}25 & 0{,}2 & 0{,}15 \end{pmatrix}$$

şi trei coduri binare pentru sursa S:

C I	a	1	C II	a	1	C III	a	00
	b	10		b	01		b	01
	c	100		c	001		c	10
	d	110		d	000		d	11

a) Ce secvenţă binară corespunde, pentru fiecare cod în parte, mesajului sursei: "abacdab"?;
b) Arătaţi că decodarea secvenţei construite la punctul a) este unică doar pentru CII şi CIII, nu şi pentru CI;
c) Calculaţi entropia sursei şi eficienţa acesteia;
d) Calculaţi lungimea medie a cuvintelor fiecărui cod, precum şi eficienţa codării.

Rezolvare:
a) Secvenţele binare cerute se obţin prin simpla substituţie a literelor cu secvenţele binare (cuvintelor) corespunzătoare lor:

$$\begin{aligned} &Sv1=1101100110 \quad 110 \\ &Sv2=1011001000 \quad 101 \\ &Sv3=0001001011 \quad 0001 \end{aligned} \tag{3.1.1}$$

b) Decodând secvenţa Sv1 prin I găsim cel puţin două variante de decodare:
abacdab, dacdd, abacabab, etc...
fapt ce nu se poate întâmpla cu Sv2 şi Sv3 prin CII, respectiv CIII.
c) Conform formulei de definiţie, entropia sursei este:

$$H(S) = \sum_{i=1}^{N} p_i \log_2 \frac{1}{p_i} = 1{,}9 \text{ biţi/simbol} \tag{3.1.2}$$

Pentru calculul eficienţei este necesară în prealabil entropia maximă:

$$H_{max} = \log_2 N = 2 \text{ biţi/simbol} \tag{3.1.3}$$

Rezultă eficienţa sursei:

$$\eta_s = \frac{H(S)}{H_{max}} = 95\% \tag{3.1.4}$$

d) Lungimea medie a cuvintelor unui cod se calculează cu formula:

$$L = \sum_{i=1}^{N} p_i l_i \tag{3.1.5}$$

unde l_i este lungimea cuvântului de cod numărul i.
Deoarece CI şi CII au aceleaşi lungimi pentru cuvinte şi lungimea medie va rezulta aceeaşi:

$$L_1 = L_2 = 0{,}4 \cdot 1 + 0{,}25 \cdot 2 + 0{,}2 \cdot 3 + 0{,}15 \cdot 3 = 1{,}95 \text{ biți/simbol} \tag{3.1.6}$$

iar:

$$L_3 = 2 \text{ biți/simbol} \tag{3.1.7}$$

fiind un cod bloc.

Eficiențele codărilor rezultă:

$$\eta_{c1} = \eta_{c2} = 97{,}6\% \qquad \eta_{c3} = \eta_s = 95\%. \tag{3.1.8}$$

Obs.: *- Deşi CIII are lungimi ale cuvintelor de cod mai mici decât CI şi CII, totuşi acestea "codează mai eficient" sursa dată, deoarece atribuie cuvânt de cod scurt (lungime 1) simbolului cel mai frecvent (a).*
- Codurile CII şi CIII sunt coduri unic decodabile prin faptul că decodarea unei secvențe codate decurge într-un unic fel. În plus sunt şi coduri instantanee. Un exemplu de cod unic decodabil, dar nu şi instantaneu, este:

CIV	*a*	*1*
	b	*10*
	c	*100*
	d	*000*

- Cu toate că realizează biunivocități între mulțimea mesajelor sursei S şi mulțimea cuvintelor de cod, codul CI nu poate fi utilizat deoarece informația transmisă printr-un astfel de cod poate fi deformată la decodare.

3.2. O SDFM cu 20 simboluri echiprobabile se codează cu un cod bloc (toate cuvintele au aceeşi lungime).
a) Cât sunt entropia şi eficiența sursei date?
b) Cât este lungimea medie m a cuvintelor codului, minim necesară ?
c) Calculați eficiența codării.

Rezolvare:
a) Sursa fiind echiprobabilă:

$$H(S) = H_{max} = \log_2 20 = 4{,}32 \text{ biți/simbol} \tag{3.2.1}$$

şi, ca atare:

$$\eta_s=100\% \qquad (3.2.2)$$

b) Pentru a putea atribui N secvenţe binare, de lungime m, celor N mesaje ale sursei este necesar ca să existe cel puţin N secvenţe distincte. Cum numărul de secvenţe binare de lungime m este 2^m rezultă că $2^m \geq N$, unde N este numărul de simboluri al sursei, N=20. În plus, conform cerinţelor problemei vom alege pe cel mai mic m ce verifică inegalitatea, astfel încât:

$$2^{m-1} < N < 2^m \qquad (3.2.3)$$

Pentru cazul concret (N=20) rezultă:

$$m=5 \qquad (3.2.4)$$

c) Lungimea medie a cuvintelor codului bloc este:

$$L = \sum_{i=1}^{N} p_i l_i = \sum_{i=1}^{20} p_i m = m \qquad (3.2.5)$$

relaţie ce este adevărată indifirent de valorile probabilităţilor p_i.
Eficienţa codării va fi:

$$\eta_c = \frac{H(S)}{L} = \frac{\log_2 20}{5} = 86{,}44\% \qquad (3.2.5)$$

3.3. a) Codaţi prin algoritmul Huffman static sursa de informaţie discretă şi fără memorie:

$$S = \begin{pmatrix} a & b & c & d \\ 0{,}4 & 0{,}2 & 0{,}25 & 0{,}15 \end{pmatrix}$$

b) Codaţi cu ajutorul aceluiaşi algoritm Huffman static extensia de ordin 2 a sursei date;

c) Calculaţi şi comparaţi în cele două cazuri eficienţele celor două coduri obţinute.

Rezolvare:

a) Desfăşurarea algoritmului Huffman static este prezentată în figura de mai jos:

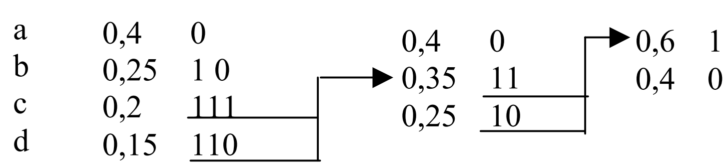


Figura 3.1.

Codul obținut are lungimea:

$$L_1 = \sum_{i=1}^{4} p_i l_i = 1{,}95 \text{ biți/simbol} \qquad (3.3.1)$$

b) Tabloul sursei S^2 (extensia de ordin I2 a sursei date) este:

$$S^2 = \begin{pmatrix} aa & ab & ac & ad & ba & bb & bc & bd & ca & cb & cc & cd & da & db & dc & dd \\ \frac{16}{100} & \frac{8}{100} & \frac{10}{100} & \frac{6}{100} & \frac{8}{100} & \frac{4}{100} & \frac{5}{100} & \frac{3}{100} & \frac{10}{100} & \frac{5}{100} & \frac{6{,}25}{100} & \frac{3{,}75}{100} & \frac{6}{100} & \frac{3}{100} & \frac{3{,}75}{100} & \frac{2{,}25}{100} \end{pmatrix}$$

Algoritmul de codare Huffman static pentru sursa S^2 este arătat în figura următoare:

aa	111	16	16	16	16	16	16
ac	010	10	10	10	10	10	11,25
ca	001	10	10	10	10	10	10
ab	1101	8	8	8	8	10	10
ba	1100	8	8	8	8	8	10
cc	1001	6,25	6,25	6,75	7,75	8	8
ad	1000	6	6	6,25	6,75	7,75	8
da	0111	6	6	6	6,25	6,75	7,75
bc	0001	5	5,25	6	6	6,25	6,75
cb	0000	5	5	5,25	6	6	6,25 1001
bb	10111	4	5	5	5,25	6 0111	6 1000
cd	10110	3,75	4	5	5 0001	5,25 0110	
dc	10101	3,75	3,75	4 10111	5 0000		
bd	10100	3	3,75 10101	3,75 10110			
db	01101	3 01101	3 10100				
dd	01100	2,25 01100					

16	16	16	20	21,25	26,75	32	41,25	58,75 1
12,25	14,5	16	16	20	21,25	26,75	32 11	41,25 0
11,25	12,25	14,5	16	16	20	21,25 01	26,75 10	
10	11,25	12,25	14,5	16	16 111	20 00		
10	10	11,25	12,25	14,5 101	16 110			
10	10	10	11,25 011	12,25 100				
8	10	10 001	10 010					
8	8 1101	10 000						
7,75 1011	8 1100							
6,75 1010								

Figura 3.2.

În desfășurarea algoritmului probabilitățile sunt înmulțite cu 100 pentru ușurința scrierii. Lungimea medie a cuvintelor codului obținut este:

$$L_2 = \frac{3}{100}(16+10+10) + \frac{4}{100}(8+8+6{,}25+6+6+5+5)$$
$$+\frac{5}{100}(4+3{,}75+3{,}75+3+3+2{,}25) =$$
$$\ldots\ldots = \frac{108+177+98{,}75}{100} = 3{,}8375 \text{ biți/simbol} \quad (3.3.2)$$

Cunoscând că entropia extensiei de ordin 2 este dublul entropiei sursei date (vezi problema 2.10) găsim că:

H(S)=1,9biți/simbol

$$H(S^2) = 2H(S) = 3{,}8 \text{ biți/simbol} \quad (3.3.3)$$

De unde:

$$\eta_{c1} = \frac{H(S)}{L_1} = 97{,}63\% \quad (3.3.4)$$

$$\eta_{c2} = \frac{H(S^2)}{L_2} = \frac{H(S)}{L_1}\cdot\frac{2L_1}{L_2} = \eta_{c1}\cdot\frac{2L_1}{L_2} = 99{,}22\% \quad (3.3.5)$$

Cum $2L_1 > L_2 \Rightarrow \eta_{c1} > \eta_{c2}$

3.4. Fie sursa de informație text (inclusiv pauzele): "TEORIA TRANSMITERII INFORMATIEI ." a) Să se afle tabloul sursei; b) Să se codeze cu un cod optimal prin algoritmul Huffman; c) Să se calculeze eficiența codării şi factorul de compresie față de o codare bloc; d) Să se construiască graful de codare.

Răspuns:

a)

$$S = \begin{pmatrix} A & E & F & I & M & N & O & R & S & T & \_ & . \\ \frac{3}{32} & \frac{3}{32} & \frac{1}{32} & \frac{7}{32} & \frac{2}{32} & \frac{2}{32} & \frac{2}{32} & \frac{4}{32} & \frac{1}{32} & \frac{4}{32} & \frac{2}{32} & \frac{1}{32} \end{pmatrix}$$

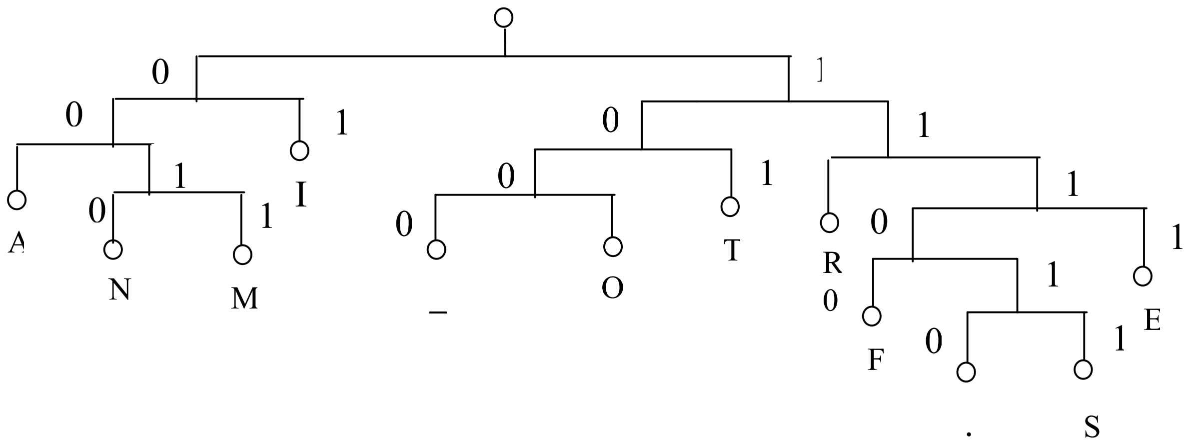
b)

A	000	0	1001
E	1111	R	110
F	11100	S	111011
I	01	T	101
M	0011	_	1000
N	0010	.	111010

c)

$\eta_c = 98{,}925\%$ F=1,185

d)



3.5. O sursă de informație discretă şi fără memorie, având redundanța relativă ρ, a fost codată cu un cod binar cu eficiențe η_c.
Ştiind că:

$$\eta_c + \rho = 1$$

şi cunoscând lungimea medie a cuvintelor codului bloc L=5 biți/simbol aflați numărul de simboluri ale sursei.

Răspuns: N=32

3.6. Sursa de informație, fără memorie, constituită din simbolurile A, C, D, E, I, N, R, cu p(A)=0,28, p(C)=0,14, p(D)=0,05, p(E)=0,18, p(I)=0,16, p(N)=0,06 se codează cu un cod binar prin algoritmul Huffman static. Care este lungimea secvenței binare ataşate mesajului "CRIN" ?

Răspuns: 13 biți

3.7. O sursă de informație echiprobabilă cu 50 simboluri se codează cu un cod binar bloc de eficiență maxim posibilă (codare pe simbol). Cât este eficiența codării?

Răspuns: $\eta_c = \frac{50}{286} \cdot \log_2 50 = 98{,}67\%$

3.8. Sursa text (fără pauze):

VINE EA SI VACANTA

ordonată alfabetic, se codează cu un cod binar (cod binar natural). Considerând ieşirea codorului drept o sursă binară secundară, cu cât este egală entropia acestei surse secundare (considerată şi ea fără memorie)?

<u>Rezolvare:</u>

Tabloul simboluri probabilități este:

$$S = \begin{pmatrix} A & C & E & I & N & S & T & V \\ \frac{4}{15} & \frac{1}{15} & \frac{2}{15} & \frac{2}{15} & \frac{2}{15} & \frac{1}{15} & \frac{1}{15} & \frac{2}{15} \end{pmatrix} \tag{3.8.1}$$

iar codul binar bloc cerut:

$$\begin{array}{ll} A & 000 \\ C & 001 \\ E & 010 \\ I & 011 \\ N & 100 \\ S & 101 \\ T & 110 \\ V & 111 \end{array} \tag{3.8.2}$$

Probabilitatea ca sursa secundară să emită un zero se calculează cu relația:

$$p(0_E) = \sum_{i=1}^{N} p_i \cdot \frac{m_{0i}}{m} \tag{3.8.3}$$

unde: N=8 - numărul de simboluri al sursei primare;

p_i $\quad i = \overline{1.N}$ -probabilitățile simbolurilor sursei primare;

m=3 lungimea cuvintelor codului;

m_{0i} - numărul de zerouri din componența cuvântului i.

În mod analog:

$$p(1_E) = 1 - p(0_E) = \sum_{i=1}^{N} p_i \cdot \frac{m_{1i}}{m} \tag{3.8.4}$$

Efectuând calculele, găsim că:

$$\begin{aligned} p(0_E) &= \frac{1}{45}(3\cdot 4 + 2\cdot 1 + 2\cdot 2 + 1\cdot 2 + 2\cdot 2 + 1\cdot 1 + 1\cdot 1) = \frac{26}{45} \\ p(1_E) &= \frac{1}{45}(1\cdot 1 + 1\cdot 2 + 2\cdot 2 + 1\cdot 2 + 2\cdot 1 + 2\cdot 1 + 3\cdot 2) = \frac{19}{45} \end{aligned} \tag{3.8.5}$$

astfel tabloul sursei secundare este:

$$S^{'} = \begin{pmatrix} 0 & 1 \\ 26/45 & 19/45 \end{pmatrix} \tag{3.8.6}$$

iar entropia:

$$H\left(S^{'}\right) = \frac{26}{45}\log_2\frac{45}{26} + \frac{19}{45}\log_2\frac{45}{19} = 0{,}9825 \text{ biți/simbol} \tag{3.8.6}$$

3.9. Folosind algoritmul LZ-77 găsiți:
a) rezultatul compresiei mesajului: aaabcbaabcc...;
b) mesajul inițial, dacă rezultatul compresiei a fost 00c00b73c72b640;
Lungimile blocurilor Lempel respectiv Ziv sunt 8 şi 4.

<u>Rezolvare:</u>

a) Diagrama de mai jos prezintă iterat codarea LZ 77 asupra mesajului dat:

1	2	3	4	5	6	7	8	9	10	11	12	S	L	A
								a	a	a	b	0	0	a
							a	a	a	b	c	8	2	b
				a	a	a	b	c	b	a	a	0	0	c
			a	a	a	b	c	b	a	a	b	7	1	a
	a	a	a	b	c	b	a	a	b	c	c	4	3	c

Rezultatul comprimării este 00a82b00c71a43c...

b) Decodarea mesajului comprimat se poate urmări cu ajutorul diagramei:

1	2	3	4	5	6	7	8	9	10	11	12	S	L	A
								c				0	0	c
							c	b				0	0	b
						c	b	c	b	c	c	7	3	c
		c	b	c	b	c	c	c	c	b		7	2	b
b	c	b	c	c	c	c	b	c	c	b	c	6	4	0

Aşadar, mesajul inițial este: cbcbccccbccbc... .

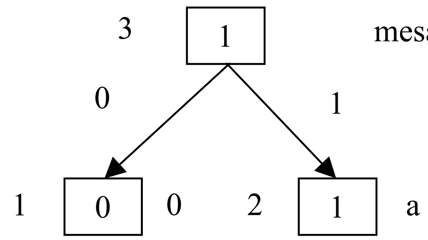
3.10. Comprimați mesajul: aaabcca... utilizând algoritmul Huffman dinamic. Cât este factorul de compresie?

<u>Rezolvare:</u>

Vom prezenta evoluția arborelui (grafului) asociat codului în timpul codării mesajului dat:

*1.*Arbore inițial: 1 $\boxed{0}$ 0 mesaj: --

2. S-a codat "a": mesaj: a codul:



simbol	cuvânt de cod
0 →	0
a →	1

Obs: *Semnificația notațiilor este:*

-pentru nod: x $\boxed{p}$

- x: număr de ordine (crește de la stânga către dreapta și de jos în sus pe linii);

- p: ponderea cumulată din nodurile aferente.

- pentru frunză (nod terminal): x $\boxed{p}$ y

- x: număr de ordine;

- p: pondere (număr de apariții ale respectivului simbol);

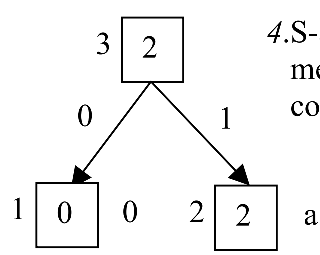
- y: simbolul.

Frunza goală, notată cu "0", este de pondere 0.

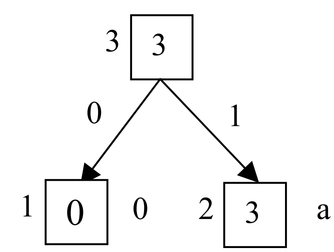
- pentru ramură: - spre stânga ≡ codare cu zero;

- spre dreapta ≡ codare (atribuire) cu unu.

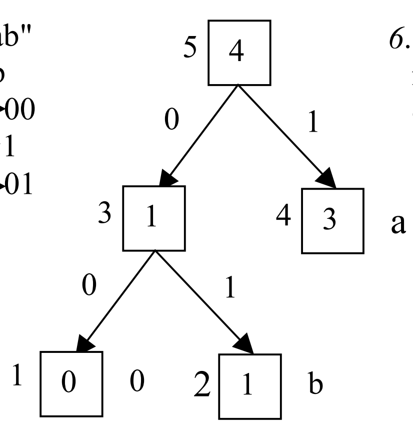
*3.*S-a codat "aa"
mesaj: a1
codul: 0 → 0
a → 1



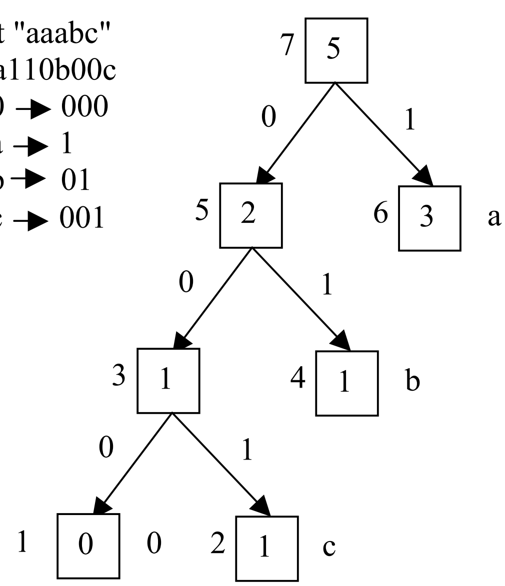
*4.*S-a codat "aaa"
mesaj: a11
codul: 0 → 0
a → 1



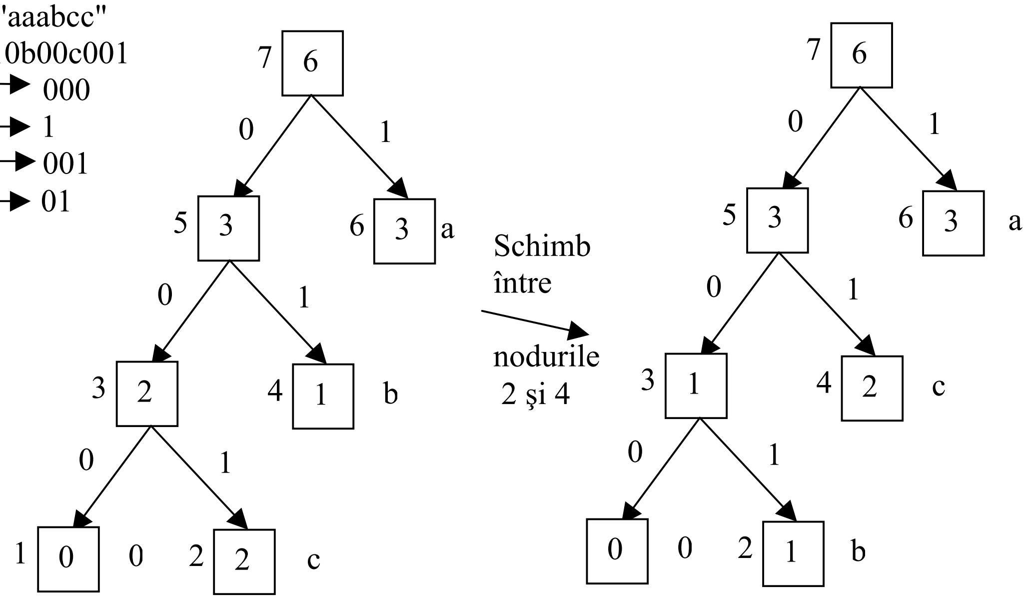
*5.*S-a codat "aaab"
mesaj: a110b
codul: 0 → 00
a → 1
b → 01



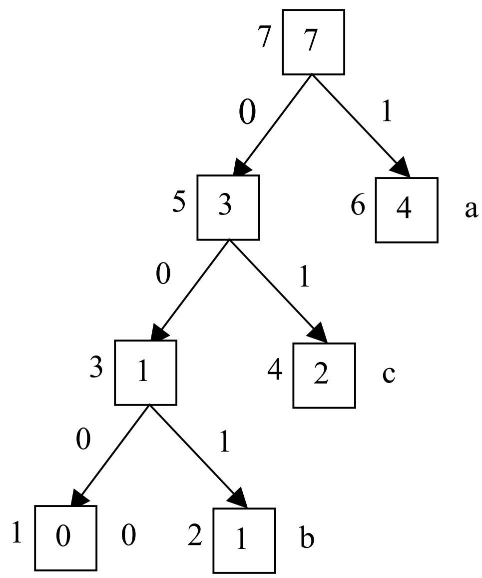
*6.*S-a codat "aaabc"
mesaj: a110b00c
codul: 0 → 000
a → 1
b → 01
c → 001



*7.*S-a codat "aaabcc"
mesaj: a110b00c001
codul: 0 → 000
a → 1
b → 001
c → 01



*8.*S-a codat "aaabcca"
mesaj: a110b00c0011
codul: 0 → 000
a → 1
b → 001
c → 01



Obs.: *la pasul 7 s-a produs un schimb între nodurile 2 şi 4 deoarece ponderea primului era mai mare (2 > 1), conform regulii care spune că un nod cu număr de ordine mai mic şi pondere mai mare decât un altul va face schimb de poziție în arbore cu acesta din urmă.*

Considerând caracterele în mesajul inițial codate pe 8 biți, lungimea acestuia este:

$$L_i=7\cdot 8=56 \text{ biți} \qquad (3.10.1)$$

Mesajul comprimat (a110b00c0011) are:

$$L_2=8\cdot 3+9=33 \text{ biți} \qquad (3.10.2)$$

Rezultă un factor de compresie:

$$F = \frac{L_1}{L_2} \cong 1,7 \qquad (3.10.3)$$

3.11. Fie mesajul "aacacabcabdabadabc". Să se comprime mesajul dat prin aplicarea algoritmului:
a) LZ 77;
b) Huffman dinamic;
c) Huffman static, presupunând mesajul dat o SDFM text.
d) Calculați în fiecare caz factorul de compesie față de o codare bloc pe opt biți a mesajului inițial.

Rezolvare:
a) Recodarea aferentă algoritmului LZ77 este prezentată în diagrama următoare:

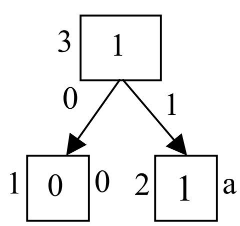
1	2	3	4	5	6	7	8	9	10	11	12	S	L	A
								a	a	c	a	0	0	a
							a	a	c	a	c	8	1	c
					a	a	c	a	c	a	b	7	3	b
	a	a	c	a	c	a	b	c	a	b	d	6	3	d
a	c	a	b	c	a	b	d	a	b	a	d	3	2	a
b	c	a	b	d	a	b	a	d	a	b	c	5	3	c

Mesaj comprimat 00a 81c 73b 63d 32a 53c de lungime:

$$L_1 = 12 \cdot 3 + 6 \cdot 8 = 84 \text{ biți} \qquad (3.11.1)$$

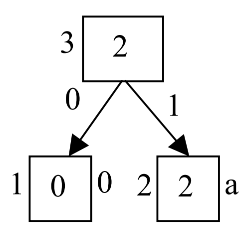
b)

1.



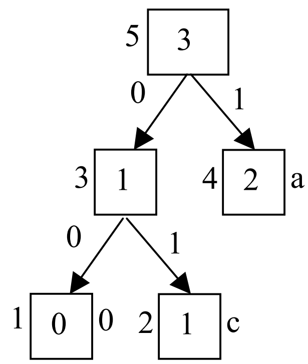
mesaj codat: "a"
mesaj transmis: "a"
cod: $0 \rightarrow 0$
$a \rightarrow 1$

2.



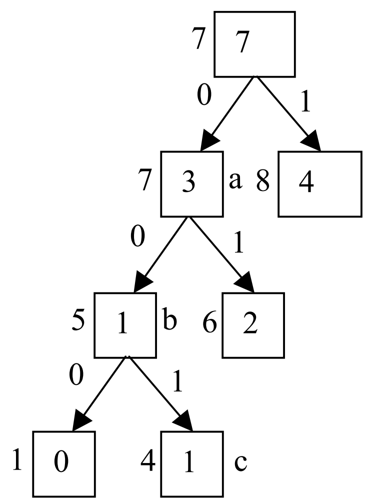
mesaj codat: "aa"
mesaj transmis: "a1"
cod: $0 \rightarrow 0$
$a \rightarrow 1$

3.



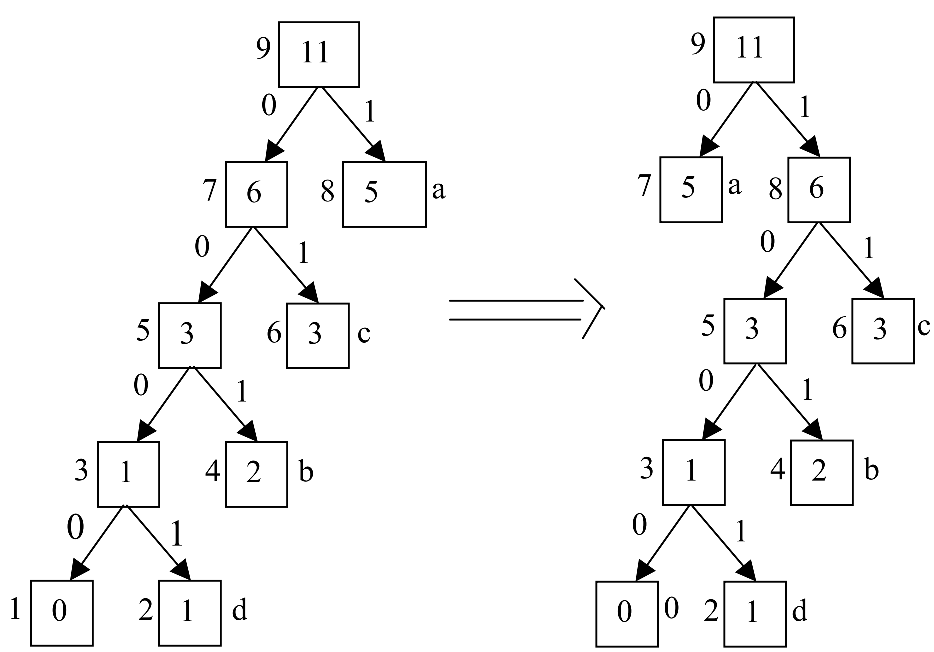
mesaj codat: "aac"
mesaj transmis: "a10c"
cod: 0 → 00
a → 1
c → 01

(*4. 5. 6.*) *7.*



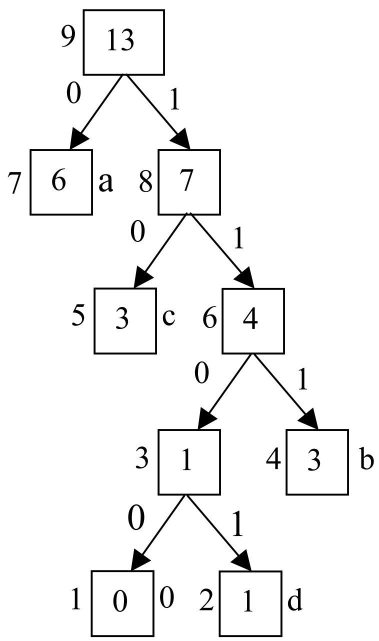
mesaj codat "aacacab"
mesaj transmis "a10c101100b"
cod 0 → 00
a → 1
b → 001
c → 01

(*8. 9. 10.*) *11.*



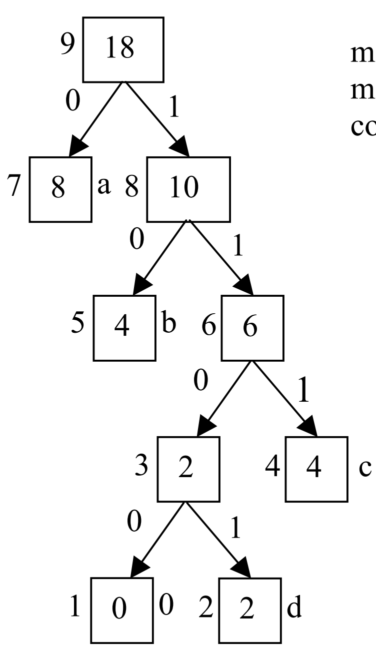
mesaj codat: "aacacabcabd"
mesaj transmis: "a10c101100b011001000d"
cod: 0 → 1000
a → 0
b → 101
c → 01

(*12.*) *13.*



mesaj codat: "aacacabcabd"
mesaj transmis: "a10c101100b011001000d"
cod: 0 → 1100
a → 0
b → 111
c → 10
d → 1101

(*14. 15. 16. 17.*) *18.*



mesaj codat: "aacacabcabd"
mesaj transmis: "a10c101100b011001000d"
cod: 0 → 1100
a → 0
b → 10
c → 111
d → 1101

Lungimea rezultată a mesajului transmis este:

$$L_2=4\cdot 8+33=65 \text{ biți} \tag{3.11.2}$$

c) Pentru a putea coda mesajul transmis prin algoritmul Huffman static definim sursa:

$$S=\begin{pmatrix} a & b & c & d \\ \frac{8}{18} & \frac{4}{18} & \frac{4}{18} & \frac{2}{18} \end{pmatrix}$$

asupra căreia putem aplica algoritmul:

a	8	0	8	0	10	1
b	4	1 0	6	11	8	0
c	4	111	4	10		
d	2	110				

Cu ajutorul codului obținut mesajul se transmite prin secvența binară:
"0011101110101110 10"

$$L_3=34 \text{ biți} \tag{3.11.3}$$

d) Dacă mesajul inițial este codat bloc pe 8 biți lungimea secvenței binare codate este:

$$L_0=18\cdot 8=144 \text{ biți} \tag{3.11.4}$$

Cu rezultatele conținute în relațiile (3.11.1, 2, 3 şi 4) găsim pentru factorii de compresie valorile:

$$\begin{aligned} F_1 &= \frac{144}{84} = 1{,}7143 \\ F_2 &= \frac{144}{65} = 2{,}2154 = 1{,}3F_1 \\ F_3 &= \frac{144}{34} = 4{,}2353 = 2{,}47F_1 \end{aligned} \tag{3.11.5}$$

Cap.4 Canale de transmisie

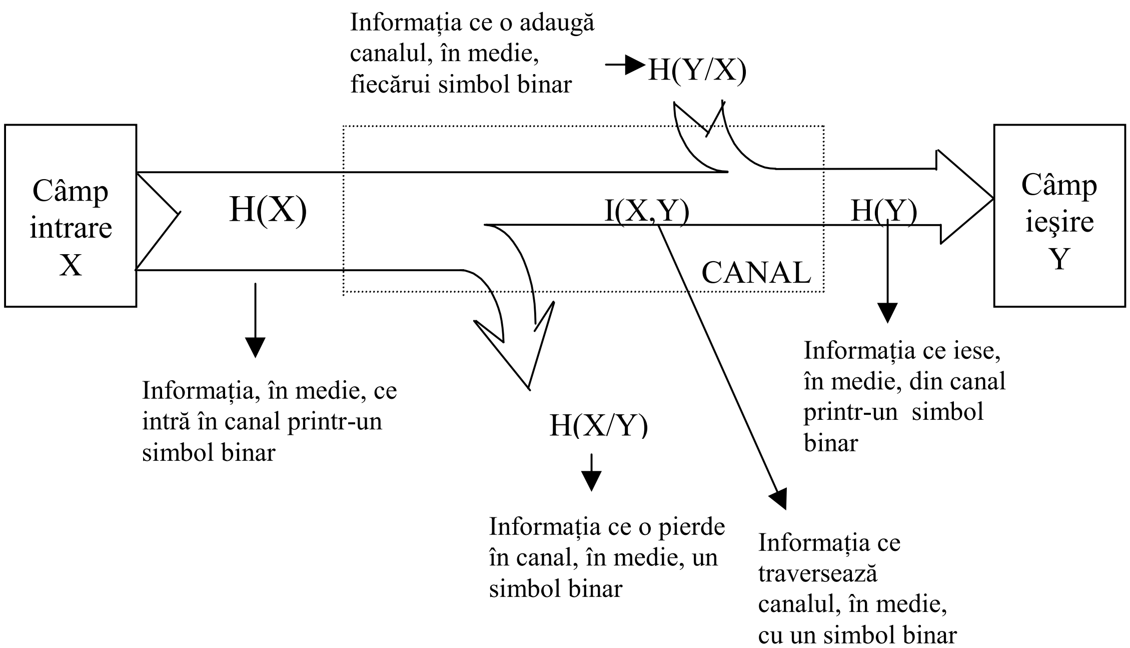
Dicționar:

- canal (de transmisie) - cale (mediu fizic) de transmisie prin care circulă infpormația între emițător (sursă) şi receptor (utilizator);
- **câmp (corp)** - o mulțime K de elemente care satisfac următoarele axiome:
  1. - pe K sunt definite două operații: adunarea şi înmulțirea;
  2. - în raport cu adunarea K formează un grup abelian cu elementul nul 0;
  3. - K \ {0} formează grup abelian pentru înmulțire;
  4. - înmulțirea este distributivă față de adunare;

- echivoc - 1. suspect, îndoielnic; 2.**echivocație**- măsură a efectului perturbațiilor asupra comunicațiilor prin canale;
- **capacitate** –valoarea maximă pe care poate să o aibă un parametru al unui sistem, parametru ce indică performanțe de înmagazinare, transport sau prelucrare.

Definiții:

1. - transferul informației între câmpul de intrare X, cel de ieşire Y şi canal C:



2. - Cantitatea de decizie D a unei surse cu m simboluri = maximul entropiei sursei $D=H_{max}$;

- Debitul de decizie $\dot{D}$ - cantitatea de decizie generată în unitatea de timp;

- **Debitul de momente** $\dot{M}$ - numărul de momente transmise în unitatea de timp;
- **Baud (Bd)** - unitatea de măsură a debitului de momente;
- **Moment** - semnal elementar purtător de informație de durată T_M.

Notații:

$\mathbf{x_0=0_E}$ **,** $\mathbf{x_1=1_E}$ - simbolurile emisibile în canalul binar;
$\mathbf{X=\{x_0 , x_1\}}$ - câmpul de la intrare în canal;
$\mathbf{y_0=0_R}$ **,** $\mathbf{y_1=1_R}$ - simbolurile recepționabile din canalul binar;
$\mathbf{Y=\{y_0 , y_1\}}$ - câmpul de la ieşirea din canal;
$\mathbf{p(x_i)}$ - probabilitatea de a emite în canal x_i;
$\mathbf{p(y_j)}$ - probabilitatea de a recepționa din canal y_j;
$\mathbf{P(Y/X)}$ - matricea de tranziție=$[p(y_j / x_i)]_{2\times2}$;
$\mathbf{p(y_j /x_i)}$ probabilitatea (apriori) de a se recepționa y_j când a fost emis x_i;
$\mathbf{P(X,Y)}$ =$[p(x_i, y_j)]_{2\times2}$;

p(x_i, y_j) - probabilitatea de a se fi emis x_i şi a se recepţiona y_j;
P(X/Y) =[p(x_i/y_j)]$_{2\times2}$;
p(x_i/y_j) - probabilitatea (aposteriorii) de a se fi emis x_i când s-a recepţionat y_j;
ξ - raport semnal-zgomot.

Abrevieri:

RSZ (SNR) - Raport Semnal pe Zgomot (Signal to Noise Ratio);
CBS - Canal Binar Simetric;
BER - Bit Error Rate (rata erorii).

Breviar teoretic:

1. Relaţii între matrici

$$P(X,Y)=\begin{bmatrix} p(0_E) & 0 \\ 0 & p(1_E) \end{bmatrix} P(Y/X)=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \tag{4.1}$$

$$p(0_R)=\alpha+\gamma \qquad p(1_R)=\beta+\delta \tag{4.2}$$

$$P(X/Y)=P(X,Y)\begin{bmatrix} \frac{1}{p(0_R)} & 0 \\ 0 & \frac{1}{p(1_R)} \end{bmatrix} \tag{4.3}$$

2. Entropii condiţionate

- eroarea medie $H(Y/X)=\sum_{i=0}^{1}\sum_{j=0}^{1}p(x_i,y_j)\cdot\log_2\frac{1}{p(y_j/x_i)}$; (4.4)

- echivocaţia $H(X/Y)=\sum_{i=0}^{1}\sum_{j=0}^{1}p(x_i,y_j)\cdot\log_2\frac{1}{p(x_i/y_j)}$; (4.5)

- transinformaţia

$$I(X,Y)=\sum_{i=0}^{1}\sum_{j=0}^{1}p(x_i,y_j)\cdot\log_2\frac{p(x_i,y_j)}{p(x_i)p(y_j)} = H(X)-H(X/Y)=H(Y)-H(Y/X) \tag{4.6}$$

3. Capacitatea canalului

- discret $C=\max\limits_{p_i} I(X,Y)$ (biţi/simbol) (4.7)

- continuu $C=\dot{D}_{max}=\dot{M}_{max}\log_2 m_{max}=B\log_2(1+\xi)$ (4.8)

unde: $\dot{M}_{max}=2B$ - (canal ideal); $\dot{M}_{max}=\frac{5}{4}B$ (canal real) (4.9)

$$m_{max}=\sqrt{1+\xi} \tag{4.10}$$

4. Redundanţa şi eficienţa canalului

$R_C=C-I(X,Y)$ - redundanţa absolută (4.11)

$\rho_c=1-\frac{I(X,Y)}{C}$ - redundanţa relativ (4.12)

$\eta_c=\frac{I(X,Y)}{C}$ - eficienţa (4.13)

4.1. Fie secvența binară:

$$i_0=1001010111000111001001 10 \qquad (4.1.1)$$

generată de o SDFM echiprobabilă, X. În vederea transmiterii dibiții din secvența i_0 se codeză prin:

$$\begin{array}{llll} 00 \longrightarrow S_{00}(t)=-3V & t\in[0,T_M] & & \\ 01 \longrightarrow S_{01}(t)=-1V & t\in[0,T_M] & & \\ 10 \longrightarrow S_{10}(t)=\ 1V & t\in[0,T_M] & & \\ 11 \longrightarrow S_{11}(t)=\ 3V & t\in[0,T_M] & \text{cu } T_M=250\mu s & (4.1.2) \end{array}$$

Se cere:

a) Să se calculeze entropia şi eficiența sursei X; ($H(X)$, η_X)

b) Debitul de informație al sursei X dacă codarea (4.1.2) se face în timp real; ($\dot{X}$)

c) Cantitatea de decizie corespunzătoare unui moment (moment = un semnal dintre S_{00}, S_{01}, S_{10} sau S_{11}); (D)

d) Debitul de momente; ($\dot{M}$)

e) Ce debit de informație (în biți/sec) corespunde în acest caz unui Baud? (1Bd)

f) Să se reprezinte grafic semnalul S(t): suport al informației i_0.

Rezolvare:

a) Probabilitățile de emisie (generare) a mesjelor binare 0 şi 1 fiind egale:

$$p(0)=p(1)=1/2 \qquad (4.1.3)$$

rezultă pentru entropie şi eficiența sursei valorile:

$$H(X)=1 \text{ bit/simbol}$$
$$\eta=100\% \qquad (4.1.4)$$

b) Prin definiție:

$$\dot{X}=\frac{H(X)}{\tau_x} \qquad (4.1.5)$$

unde τ_x este timpul în care se transmite (generează) un simbol binar de către sursa X. Ştiind că două simboluri (un dibit) au o durată de transmitere, respectiv genereare (în timp real), de $T_M=250\mu s$ rezultă pentru τ_x valoarea:

$$\tau_x=\frac{1}{2}T_M=125\mu s \qquad (4.1.6)$$

şi ca atare:

$$\dot{X}=\frac{1 \text{ bit/simbol}}{125\,\mu s}=8 \text{ kbiti/sec} \qquad (4.1.7)$$

c) Conform definiției:

$$D=\log_2 m=2 \text{ biți} \qquad (4.1.8)$$

unde m=4 reprezintă numărul de nivele de tensiune corespunzătoare dibiților (relația (4.1.2)).

d) Debitul de momente $\dot{M}$ este numărul de semnale elementare (S_{00}, S_{01}, S_{10} sau S_{11}) pe unitatea de timp:

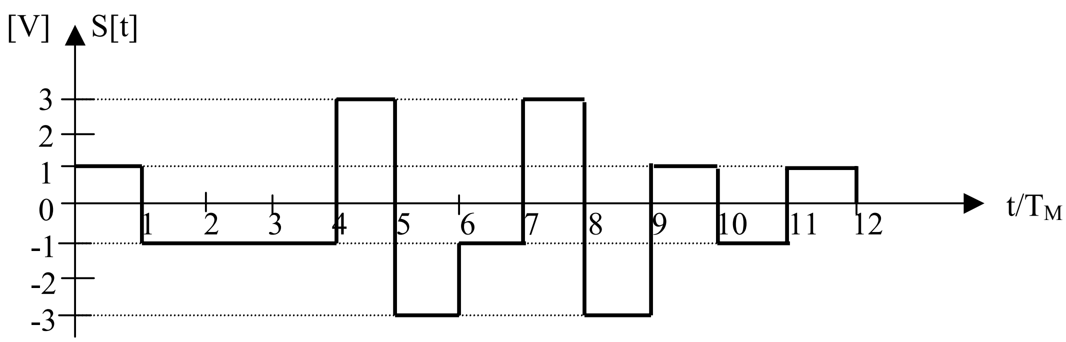
$$\dot{M} = \frac{1}{T_M} = \frac{1}{250 \cdot 10^{-6}} = 4000\,\text{Bd} \qquad (4.1.9)$$

e) Cum debitul de informație este $\dot{X} = 8000$ biți/sec, corespunde la viteza de modulație (debitul de momente) $\dot{M} = 4000\,\text{Bd}$, rezultă pentru fiecare Baud câte 2 biți de informație.

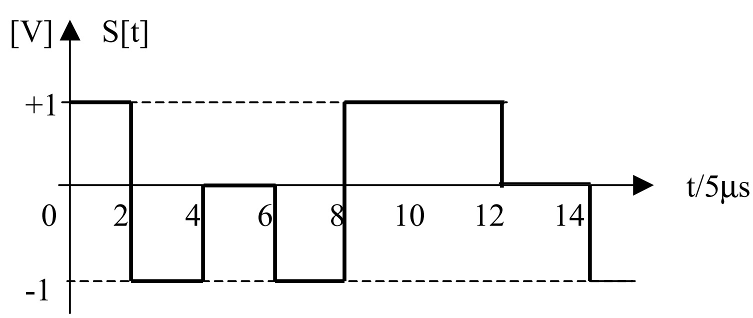
f) Secvența binară dată va avea suport fizic următoarea succesiune de semnale elementare:

$$S_{10}S_{01}S_{01}S_{01}S_{11}S_{00}S_{01}S_{11}S_{00}S_{10}S_{01}S_{10} \qquad (4.1.10)$$

care formează semnalul:



4.2. Presupunând semnalul ternar din figură:



suportul fizic al unei informații, să se afle:

a) Viteza de modulație $\dot{M}$ (debitul de momente);

b) Cantitatea de informație maximă ce se poate transmite printr-un simbol ternar;
c) Debitul de informație ce rezultă dacă semnalul S(t) a fost obținut prin modularea (codarea) unei secvențe binare echiprobabile, în timp real, utilizând codul:

$$\begin{aligned} &\mathbf{11} \longrightarrow S_+(t) = +1V \quad t \in [0, T_M] \\ &\mathbf{0} \longrightarrow S_0 t) = \ \ 0V \quad t \in [0, T_M] \\ &\mathbf{10} \longrightarrow S_-(t) = -1V \quad t \in [0, T_M] \end{aligned} \tag{4.2.1}$$

d) Debitul maxim de informație ce poate fi obținut printr-o codare adecvată utilizând : - un semnal ternar;
- un semnal binar;
- un semnal cuaternar.
În toate cazurile se va considera aceeaşi viteză de modulație.

Răspuns:

a) $\dot{M} = \frac{1}{T_M} = 10^5$ simbol ternar / sec $= 10^5$ Baud

b) D=$\log_2 3$=1,585 (biți)

c)

$$\left.\begin{aligned} &p(S_+) = p_+ = p(1) \cdot p(1) = \frac{1}{4}; \\ &p(S_0) = \frac{1}{2} \\ &p(S_-) = \frac{1}{4} \end{aligned}\right\} \Rightarrow H(S) = 2 \cdot \frac{1}{4} \log_2 4 + \frac{1}{2} \cdot \log_2 2 = 1 \text{ bit/simb.ternar}$$

$$D_S = \frac{H(S)}{T_M} = 10^5 \text{ biti/sec} < D \cdot \dot{M} = 1{,}585 \cdot 10^5 \text{ biți/sec}$$

d)

Semnal	D	$\dot{X}$ [biți / sec]
binar	1 bit inf. / sg. binar	$1/T_M$
ternar	1,585 biți informație / sg. ternar	$1{,}585/T_M$
cuaternar	2 biți informație / sg. cuaternar	$2/T_M$

D - cantitatea maximă de informație ce se poate înmagazina într-un semnal elementar de durată T_M;

$\dot{X}$ - debitul maxim de informație relativ la T_M [sec].

4.3. Sursa de informație fără memorie constituită din caracterele echiprobabile A, C, D, E, I, N, R se codează cu un cod binar bloc (transpunere zecimal binară a numărului de ordine; fără cuvântul de pondere zero), apoi se cuplează la un canal de transmisie binar simetric având rata erorii egală cu 0,1. Calculați:
a) entropia şi eficiența sursei; (H(S) η_S)
b) eficiența codării; (η_c)
c) entropia câmpului de la intrarea în canal; (H(X))
d) matricea de tranziție şi eroarea medie; (P(Y/X), H(Y/X))
e) câmpul de la ieşirea din canal şi entropia sa; (Y, H(Y))
f) echivocația şi transinformația; (H(X/Y), I(X,Y))
g) capacitatea canalului; (C)
h) probabilitatea ca sursa să fi emis mesajul DAC atunci când s-a recepționat mesajul RAC. ($p(DAC_E/RAC_R)$)

Rezolvare:
a) Sursa fiind echiprobabilă:

$$H(S) = H_{max} = \log_2 7 \cong 2,8 \text{ biți/simbol}$$
$$\eta_S = 100\% \qquad (4.3.1)$$

b) Conform cerințelor problemei codul binar bloc este:

$$\begin{array}{lll} A & \longrightarrow & 001 \\ C & \longrightarrow & 010 \\ D & \longrightarrow & 011 \\ E & \longrightarrow & 100 \\ I & \longrightarrow & 101 \\ N & \longrightarrow & 110 \\ R & \longrightarrow & 111 \end{array}$$

Şi are lungimea medie a cuvintelor de cod:

$$L = 3 \text{ biți/simbol} \qquad (4.3.2)$$

Pentru eficiență rezultă valoarea:

$$\eta_c = \frac{H(S)}{L} = 93,58\% \qquad (4.3.3)$$

c) Câmpul de la intrarea în canal conține două simboluri binare ce pot fi emise în canal cu probabilitățile:

$$\begin{aligned} p(0_E) &= \sum_{i=1}^{7} p(S_i)\frac{k_{0i}}{3} = \frac{1}{21}\sum_{i=1}^{7} k_{0i} = \frac{9}{21} = \frac{3}{7} \\ p(1_E) &= \sum_{i=1}^{7} p(S_i)\frac{k_{1i}}{3} = \frac{1}{21}\sum_{i=1}^{7} k_{1i} = \frac{12}{21} = \frac{4}{7} \end{aligned} \qquad (4.3.4)$$

Astfel, tabloul sursei binare (secundare) ce emite în canal este:

$$X = \begin{pmatrix} 0_E & 1_E \\ 3/7 & 4/7 \end{pmatrix} \tag{4.3.5}$$

şi are entropia:

$$H(X) = \frac{3}{7}\log_2\frac{7}{3} + \frac{4}{7}\log_2\frac{7}{4} = 0{,}9852281 \text{ biți/simbol} \tag{4.3.6}$$

Obs.: *În fapt, sursa secundară X este cu memorie. Pentru a argumenta acest lucru, este suficient să considerăm sursa X fără memorie şi să calculăm probabilitatea de a emite A:*

$$p(A) = p(001_E) = p(0_E)\cdot p(0_E)\cdot p(1_E) = \frac{36}{49}\cdot\frac{1}{7}$$

care diferă de p(A)=1/7, dat de echiprobabilitatea sursei primare.

d) Canalul fiind binar simetric şi cunoscând rata erorii p=0,1 matricea de tranziție are forma:

$$P(Y/X) = \begin{bmatrix} p(0_R/0_E) & p(1_R/0_E) \\ p(0_R/1_E) & p(1_R/1_E) \end{bmatrix} = \begin{array}{c|cc} & 0_R & 1_R \\ \hline 0_E & 1-p & p \\ 1_E & p & 1-p \end{array} = \begin{array}{c|cc} & 0_R & 1_R \\ \hline 0_E & 0{,}9 & 0{,}1 \\ 1_E & 0{,}1 & 0{,}9 \end{array} \tag{4.3.7}$$

Eroarea medie este "informația" medie, pe simbol binar, ce soseşte la recepție şi nu provine de la emisie ci din canal:

$$H(Y/X) = \sum_{i=0}^{1}\sum_{j=0}^{1} p(x_i, y_j)\cdot \log_2\frac{1}{p(y_j/x_i)} \tag{4.3.8}$$

unde: - $p(y_j/x_i)$ = probabilitatea ca la recepție să sosească y_j dacă la emisie a fost transmis x_i; probabilitățile de genul $p(y_j/x_i)$ sunt conținute în matricea P(Y/X) dată de relația (7).

- $p(x_i, y_j)$ = probabilitatea de a se emite x_i şi a se recepționa y_j;

Aceste probabilități formează o matrice de forma P(X,Y) şi se calculează cu ajutorul relației lui Bayes:

$$p(x_i, y_j) = p(x_i)\cdot p(y_j/x_i) \qquad i,j=0, \tag{4.3.9}$$

sau:

$$P(X,Y) = \begin{bmatrix} p(x_0) & 0 \\ 0 & p(x_1) \end{bmatrix} P(Y/X) \tag{4.3.10}$$

- x_i, y_j cu i,j=0, 1 sunt notațiile alternative pentru 0_E, 1_E respectiv 0_R, 1_R folosite pentru a scrie relațiile compact:

$$x_0 = 0_E \qquad x_1 = 1_E \qquad 0_R = y_0 \qquad 1_R = y_1 \tag{4.3.11}$$

Obs.: *Deşi atât câmpul de la intrare cât şi câmpul de la ieşire conțin aceleaşi două simboluri binare 0 şi 1, sunt necesare notații distincte pentru a putea fi distinse în diferitele relații matematice.*

Înlocuind rezultatele din (4.3.4) şi (4.3.7) în (4.3.10) găsim că:

$$P(X,Y)=\begin{bmatrix}\frac{2,7}{7} & \frac{0,3}{7}\\ \frac{0,4}{7} & \frac{3,6}{7}\end{bmatrix} \tag{4.3.12}$$

Dispunem acum de toate probabilităţile cerute de (4.3.8) pentru a putea calcula eroarea medie:

$$\begin{aligned}H(Y/X)&=-\frac{2,7}{7}\log_2 0,9-\frac{0,3}{7}\log_2 0,1-\frac{0,4}{7}\log_2 0,1-\frac{3,6}{7}\log_2 0,9=\\&=-0,9\log_2 0,9-0,1\log_2 0,1=0,4689955\ \text{biti/simbol binar}\end{aligned} \tag{4.3.13}$$

e) Probabilităţile simbolurilor de la ieşirea din canal, 0_R şi 1_R se află cu ajutorul formulelor:

$$\begin{aligned}p(0_R)&=p(y_0)=p(x_0,y_0)+p(x_1,y_0)=p(0_E,0_R)+p(1_E,0_R)\\ p(1_R)&=p(y_1)=p(x_0,y_1)+p(x_1,y_1)=p(0_E,1_R)+p(1_E,1_R)\end{aligned} \tag{4.3.14}$$

sau compact:

$$[p(0_r)\quad p(1_R)]=[1\quad 1]\cdot P(X,Y)=\begin{bmatrix}\frac{3,1}{7} & \frac{3,9}{7}\end{bmatrix} \tag{4.3.15}$$

Sursa binară echivalentă ieşirii canalului este:

$$Y=\begin{pmatrix}0_R & 1_R\\ \frac{3,1}{7} & \frac{3,9}{7}\end{pmatrix} \tag{4.3.16}$$

şi are entropia:

$$\begin{aligned}H(Y)&=\frac{3,1}{7}\log_2\frac{7}{3,1}+\frac{3,9}{7}\log_2\frac{7}{3,9}=\log_2 7-\frac{3,1}{7}\log_2 3,1-\frac{3,9}{7}\log_2 3,9\cong\\&=0,990577\text{biţi/simbol binar}\end{aligned} \tag{4.3.17}$$

f) Echivocaţia, notată H(X/Y), este cantitatea medie de informaţie pe un simbol binar ce este pierdută de acesta (simbolul binar) în canal:

$$H(X/Y)=\sum_{i=0}^{1}\sum_{j=0}^{1}p(x_i,y_j)\cdot\log_2\frac{1}{p(x_i/y_j)} \tag{4.3.18}$$

Calculând probabilităţile aposteriori $p(x_i/y_j)$ cu relaţia:

$$P(X/Y)=\begin{bmatrix}p(0_E/0_R) & p(0_E/1_E)\\ p(1_E/0_R) & p(1_E/1_R)\end{bmatrix}=P(X,Y)\cdot\begin{bmatrix}\frac{1}{p(0_R)} & 0\\ 0 & \frac{1}{p(1_R)}\end{bmatrix}=$$

$$=\begin{bmatrix} \frac{2,7}{7} & \frac{0,3}{7} \\ \frac{0,4}{7} & \frac{3,6}{7} \end{bmatrix} \cdot \begin{bmatrix} \frac{7}{3,1} & 0 \\ 0 & \frac{7}{3,9} \end{bmatrix} = \begin{bmatrix} \frac{27}{31} & \frac{1}{31} \\ \frac{4}{31} & \frac{12}{13} \end{bmatrix} \tag{4.3.19}$$

Se găseşte valoarea echivocației:

$$H(X/Y)=\frac{2,7}{7}\log_2\frac{31}{27}+\frac{0,3}{7}\log_2 13+\frac{0,4}{7}\log_2\frac{31}{4}+\frac{3,6}{7}\log_2\frac{13}{12}$$
$$=0,463666 \text{ biți/simbol binar} \tag{4.3.20}$$

Transinformația I(X,Y) se poate afla utilizând egalitățile:

$$I(X,Y)=H(X)-H(X/Y)=H(Y)-H(Y/X) \tag{4.3.21}$$

Utilizând (4.3.6) şi (4.3.20) găsim valoarea:

$$I(X,Y)=0,521562 \text{ biți / simbol binar} \tag{4.3.22}$$

Valoare ce se găseşte şi folosind rezultatele din (4.3.13) şi (4.3.17).

<u>Obs.</u>: *- utilizarea relațiilor necesare în rezolvarea problemelor de tipul prezentei implică calcule matematice voluminoase. Pentru a obține rezultate bune este de preferat să nu se facă aproximări în calculele intermediare, obtând pentru varianta memorării lor cu ajutorul unității de calcul. O bună verificare a rezultatelor găsite se poate face cu ajutorul relației (4.3.21).*

g) Capacitatea canalului binar simetric se calculează cu formula:

$$C=1+p\log_2 p+(1-p)\log_2(1-p) \tag{4.3.23}$$

şi, în cazul de față pentru p=0,1, este:

$$C=0,531004 \text{biți/simbol} \tag{4.3.24}$$

<u>Obs.</u>: *- capacitatea canalului, conform definiției, este maximul transinformației. Rezultatul (4.3.24) se poate regăsi făcând în (4.3.21) H(Y)=1.*

h) Probabilitatea de a se fi emis "DAC" atunci când s-a recepționat "RAC" este identică cu probabilitatea de a se fi emis secvența binară "011 001 010" atunci când s-a recepțiopnat "111 001 010":

$$\begin{aligned} p(DAC_E/RAC_R) &= p(011001010_E/111001010_R)= \\ &= p(0_E/0_R)^4\cdot p(0_E/1_R)\cdot p(1_E/1_R)^4 = \\ &= \left(\frac{27}{31}\right)^4\cdot\frac{1}{13}\cdot\left(\frac{12}{13}\right)^4 \cong 3,2\% \end{aligned} \tag{4.3.25}$$

Obs.: *- în ultima relație, (4.3.25), s-a făcut presupunerea că transmisia unui simbol binar este independentă de a celorlalte.*

4.4. Sursa text (discretă şi fără memorie): "TEORIA TRANSMISIUNII INFORMATIEI"

se codează binar bloc (ordonare alfabetică, pauza la sfârşit, cod binar natural) şi se cuplează la un canal de transmisie binar. Cunoscând că $p(0_R/0_E)=0{,}9$ **şi** $p(0_R/1_E)=0{,}2$ **să se afle:**

a) matricea de tranziție a canalului; (P(Y/X))
b) codul binar bloc;
c) probabilitatea ca sursa să emită mesajul METEOR;
d) probabilitatea de a se fi recepționat TEN dacă s-a emis TEO;
e) probabilitatea de a se fi emis MAR dacă s-a recepționat MAT;
f) entropiile câmpurilor de intrare şi ieşire; (H(X), H(Y))
g) eroarea medie şi echivocația; (H(Y/X), H(X/Y))
h) transinformația şi capacitatea canalului; (I(X,Y), C)

<u>Răspuns:</u>

a) $$P(Y/X)=\begin{array}{c c} & \begin{array}{cc} 0_R & 1_R \end{array} \\ \begin{array}{c} 0_E \\ 1_E \end{array} & \begin{bmatrix} 0{,}9 & 0{,}1 \\ 0{,}2 & 0{,}8 \end{bmatrix} \end{array}$$

b)

A	3/32	0000
E	2/32	0001
F	1/32	0010
I	8/32	0011
M	2/32	0100
N	3/32	0101
O	2/32	0110
R	3/32	0111
S	2/32	1000
T	3/32	1001
U	1/32	1010
_	2/32	1011

c) $$p(METEOR_E)=\frac{2\cdot 2\cdot 3\cdot 2\cdot 2\cdot 3}{32^6}=1{,}34\cdot 10^{-7}$$

d) $$p(TEN_R/TEO_E)=p(1001\ 0001\ 0101_R/1001\ 0001\ 1001_E)=$$
$$=p(0_R/0_E)^6\cdot p(0_R/1_E)\cdot p(1_R/0_E)\cdot p(1_R/1_E)^4=0{,}9^6\cdot 0{,}2\cdot 0{,}1\cdot 0{,}8^4=0{,}435\%$$

e) $$p(0_E)=\frac{9}{16}\qquad p(1_E)=\frac{7}{16},$$

în ipoteza că sursa secundară $\{0_E,1_E\}$ este fără memorie.

$$P(X,Y)=\begin{array}{c} \\ 0_E \\ 1_E \end{array}\begin{array}{c} \begin{array}{cc} 0_R & 1_R \end{array} \\ \begin{bmatrix} \frac{8,1}{16} & \frac{0,9}{16} \\ \frac{1,4}{16} & \frac{5,6}{16} \end{bmatrix} \end{array}$$

$$p(0_R)=\frac{9,5}{16} \qquad p(1_R)=\frac{6,5}{16}$$

$$P(X/Y)=\begin{array}{c} \\ 0_E \\ 1_E \end{array}\begin{array}{c} \begin{array}{cc} 0_R & 1_R \end{array} \\ \begin{bmatrix} \frac{81}{95} & \frac{9}{65} \\ \frac{14}{95} & \frac{56}{65} \end{bmatrix} \end{array}$$

$$p(MAR_E/MAT_R)=p(0_E/0_R)^7\cdot p(0_E/1_R)\cdot p(1_E/0_R)^2\cdot p(1_E/1_R)^3=$$

$$=\left(\frac{81}{95}\right)^7\cdot\frac{9}{65}\cdot\left(\frac{14}{95}\right)^2\cdot\left(\frac{56}{65}\right)^3=0,63\cdot 10^{-3}$$

f) H(X)=0,988699408 biți/simbol binar
H(Y)= 0,974489403 biți /Simbol binar

g) H(Y/X)=0,579653563 biți/simbol binar
H(X/Y)= 0,593863568 biți /Simbol binar

h) I(X,Y)=0,39483584 biți/simbol binar
C= 0,420346437 biți/simbol binar

4.5. Sursa text ANTENA RECEPTOARE codată binar bloc (ordonare alfabetică +CBN) se cuplează la un canal de transmisiune binar având matricea de tranziție:

$$P=\begin{bmatrix} 0,9 & 0,1 \\ 0,2 & 0,8 \end{bmatrix}$$

a) Să se afle tabloul sursei primare S;

b) Să se afle tabloul sursei secundare X, ce emite în canal, considerându-se că este SDFM;

c) Considerând că S' este sursa având aceleaşi simboluri ca şi S dar cu probabilități cele de recepționare a mesajelor, să se afle tabloul sursei S;

d) Cât sunt probabilitățile:

$p(CAP_E)$ - de a se emite mesajul CAP;

$p(CAP_R)$ - de a se recepționa mesajul CAP;

$p(CAP_R/CAP_E)$ - de a se recepționa mesajul CAP dacă acesta a fost emis;

$p(CAP_E/CAP_R)$ - de a se fi emis mesajul CAP dacă acesta a fost recepționat?

Rezolvare:

a) Tabloul sursei primare S este:

$$S=\begin{pmatrix} A_E & C_E & E_E & N_E & O_E & P_E & R_E & T_E \\ \frac{3}{16} & \frac{1}{16} & \frac{4}{16} & \frac{2}{16} & \frac{1}{16} & \frac{1}{16} & \frac{2}{16} & \frac{2}{16} \end{pmatrix} \qquad (4.5.1)$$

Obs.: *Indicele "E" se referă la emisie.*

Codul bloc este:

$$\begin{array}{llll} A \to 000 & & O \to 100 \\ C \to 001 & & P \to 101 \\ E \to 010 & & R \to 110 \\ N \to 011 & & T \to 111 \end{array} \qquad (4.5.2)$$

b) Utilizând relațiile (4.3.4) aflăm că:

$$p(0_E)=\frac{13}{24} \qquad p(1_E)=\frac{11}{24} \qquad 4.5.3)$$

c) Pentru a afla probabilitățile simbolurilor recepționabile se utilizează formula probabilității totale (1.5). De exemplu, pentru calculul probabilității de a recepționa A, se foloseşte relația:

$$\begin{aligned} p(A_R) &= p(A_E)\cdot p(A_R/A_E)+p(C_E)\cdot p(A_R/V_E)+ \\ &\ldots+p(T_E)\cdot p(A_R/T_E) \end{aligned} \qquad (4.5.4)$$

unde:

$$p(A_R/A_E)=p(000_R/000_E)=p(0_R/0_E)^3=0{,}9^3 \qquad (4.5.5)$$

Folosind relațiile de forma (4.5.4) şi (4.5.5) se calculează probabilitățile de recepționare pentru toate cele 8 simboluri. Aceste calcule sunt indicate în tabelul de mai jos:

Probabilitate simbol $\times 16$ Simbol	3 A	1 C	4 E	2 N	1 O	1 P	2 R	2 T	$\div 16.000$
A	9^3	$9^2\cdot 2$	$9^2\cdot 2$	$9\cdot 2^2$	$9^2\cdot 2$	$9\cdot 2^2$	$9\cdot 2^2$	2^3	3355
C	$9^2\cdot 1$	$9^2\cdot 8$	$9\cdot 2\cdot 1$	$9\cdot 2\cdot 8$	$9\cdot 2\cdot 1$	$9\cdot 2\cdot 8$	$2^2\cdot 1$	$2^2\cdot 8$	1485
E	$9^2\cdot 1$	$9\cdot 2\cdot 1$	$9^2\cdot 8$	$9\cdot 8\cdot 2$	$9\cdot 2\cdot 1$	$2^2\cdot 1$	$9\cdot 8\cdot 2$	$2^2\cdot 8$	3515
N	$9\cdot 1^2$	$9\cdot 1\cdot 8$	$9\cdot 8\cdot 1$	$9\cdot 8^2$	$2\cdot 1^2$	$2\cdot 1\cdot 8$	$2\cdot 1\cdot 8$	$2\cdot 8\cdot 8$	1845
O	$9^2\cdot 1$	$9\cdot 1\cdot 2$	$9\cdot 1\cdot 2$	$2^2\cdot 1$	$9^2\cdot 8$	$9\cdot 8\cdot 1$	$9\cdot 8\cdot 2$	$8\cdot 2^2$	1485
P	$9\cdot 1^2$	$9\cdot 8\cdot 1$	$2\cdot 1^2$	$8\cdot 2\cdot 1$	$9\cdot 8\cdot 1$	$9\cdot 8^2$	$8\cdot 2\cdot 1$	$8^2\cdot 2$	1075
R	$9\cdot 1^2$	$2\cdot 1^2$	$9\cdot 8\cdot 1$	$8\cdot 2\cdot 1$	$9\cdot 8\cdot 1$	$8\cdot 2\cdot 1$	$9\cdot 8^2$	$8^2\cdot 2$	1845
T	1^3	$8\cdot 1^2$	$8\cdot 1^2$	$8^2\cdot 1$	$8\cdot 1^2$	$8^2\cdot 1$	$8^2\cdot 1$	8^3	1395

Tabloul sursei S' va fi:

$$S' = \begin{pmatrix} A & C & E & N & O & P & R & T \\ \frac{3.355}{16000} & \frac{1485}{16000} & \frac{3515}{16000} & \frac{1845}{16000} & \frac{1485}{16000} & \frac{1075}{16000} & \frac{1845}{16000} & \frac{1395}{16000} \end{pmatrix} \quad (4.5.6)$$

Obs.: *comparând (4.5.1) cu (4.5.6) se observă că A, C, O şi P sunt mai probabile la recepție decât la emisie în vreme ce probabilitățile simbolurilor E, N, R şi T au scăzut, în probabilitate, la recepție față de emisie.*

d) Probabilitățile cerute sunt:

$$p(CAP_E) = \frac{3}{16^3} = 0,7 \cdot 10^{-3}$$

$$p(CAP_R) = \frac{5,3558}{16^3} = 1,3 \cdot 10^{-3}$$

$$p(CAP_R / CAP_E) = p(001000101_R / 001000101_E) = 0,9^6 \cdot 0,8^3 = 272 \cdot 10^{-3}$$

$$p(CAP_E / CAP_R) = p(001000101_E / 001000101_R) = \left(\frac{117}{139}\right)^6 \cdot \left(\frac{88}{101}\right)^3 = 235 \cdot 10^{-3}$$

unde: $p(0_E/0_R)$ şi $p(1_E/1_R)$ s-au calculat cu relațiile (4.1, 2 şi 3):

$$P(X,Y) = \begin{bmatrix} \frac{11,7}{24} & \frac{1,3}{24} \\ \frac{2,2}{24} & \frac{8,8}{24} \end{bmatrix}$$

$$p(0_R) = \frac{13,9}{24} \qquad p(1_R) = \frac{10,1}{24}$$

$$P(X/Y) = \begin{bmatrix} \frac{117}{139} & \frac{13}{101} \\ \frac{22}{139} & \frac{88}{101} \end{bmatrix}$$

Obs.: *Calculul probabilităților $p(0_R)$ şi $p(1_R)$ se poate face şi cu relații de forma (4.3.4), rezultând aceleaşi valori.*

4.6. Care este durata transmiterii unei imagini alb-negru formată din $N=10^6$ pixeli printr-un canal:
a) telefonic cu banda cuprinsă între 300-3400 Hz;
b) video cu banda 12 MHz;
c) optic (fibră optică) cu banda de 1 GHz;
Intensitatea fiecărui pixel se cuantizează pe 128 nivele, iar raportul semnal-zgomotul se consideră cel minim necesar.

Rezolvare:

Informația conținută într-o imagine, necesară a fi transmisă, este:

$$I_{im} = N \cdot i_{pixel} = 10^6 \cdot \log_2 128 = 7 \cdot 10^6 \text{ biți/imagine} \qquad (4.6.1)$$

La 128 de nivele de cuantizare este necesar un raport semnal-zgomot minim:

$$\xi \cong (128)^2 = 2^{14} \qquad (4.6.2)$$

pentru care rezultă capacitățile canalelor:

$$\begin{aligned} C_T &= B_T \log_2(1+\xi) = 3{,}1 \cdot 10^3 \cdot 14 = 43{,}4 \text{ kbiți/sec} \\ C_V &= B_V \log_2(1+\xi) = 12 \cdot 10^6 \cdot 14 = 168 \text{ Mbiți/sec} \\ C_O &= B_O \log_2(1+\xi) = 10^6 \cdot 14 = 14 \text{ Gbiți/sec} \end{aligned} \qquad (4.6.3)$$

Dacă transmisiile s-ar face la aceste capacități (debite de informație) rezultă timpii de transmisie:

$$\begin{aligned} t_T &= \frac{I_{im}}{C_T} = 161{,}3 \text{ sec} \\ t_V &= \frac{I_{im}}{C_V} = 41{,}7 \text{ ms} \\ t_O &= \frac{I_{im}}{C_O} = 0{,}5 \text{ ms} \end{aligned} \qquad (4.6.4)$$

4.7. La ce raport semnal-zgomot minim s-ar putea transmite, în cele trei cazuri din problema anterioară, imaginile cu viteza necesară de 25 imagini/sec?

Rezolvare:

Capacitatea cerută fiecărui canal este:

$$C = \frac{I_{imag}}{\tau_{imag}} = 25 \cdot I_{imag} / 1\text{ sec} = 175 \text{ Mbiți/sec} \qquad (4.7.1)$$

Această capacitate se poate obține la un raport semnal/zgomot:

$$\xi = 2^{C/B} - 1 \quad \text{sau (dacă C>>B):} \quad \xi_{dB} = 3 \cdot \frac{C}{B} \qquad (4.7.2)$$

Cunoscând în fiecare caz banda disponibilă avem că:

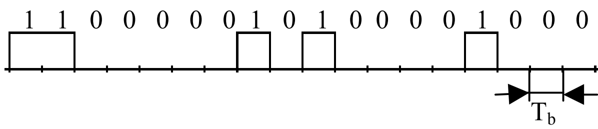
$$\xi_{dB}^{T} \cong 3 \frac{175 \cdot 10^6}{3{,}1 \cdot 10^3} = 169355 \text{ dB}$$

$$\xi_{dB}^{V} \cong 3\frac{175\cdot 10^{6}}{12\cdot 10^{6}} = 43{,}75\ \text{dB}$$

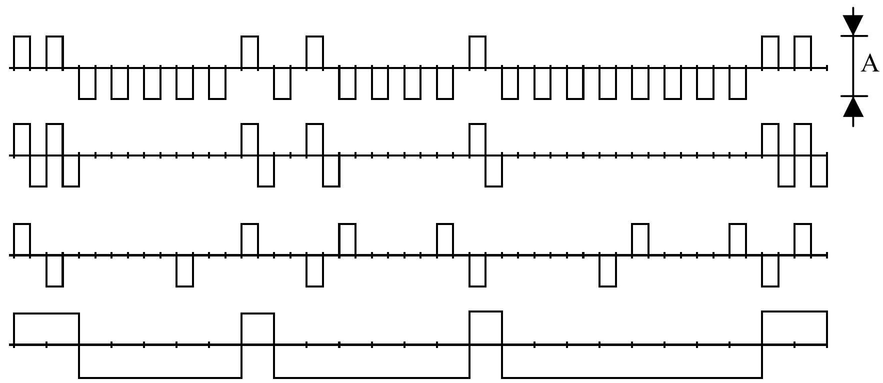
$$\xi^{O} = 2^{0{,}175} - 1 = 0{,}129 \quad \text{sau} \quad \xi_{dB}^{O} \cong -9\text{dB} \tag{4.7.3}$$

Obs.: *Evident că un raport semnal-zgomot de 169355 dB este un răspuns non sens pentru practică.*

4.8. Reprezentați secvența de date prezentată în figura de mai jos, pe rând în codurile NRZ, RZ, AMI şi HDB3. Calculați componenta medie, în fiecare caz, luând valoarea vârf la vârf a semnalului A.



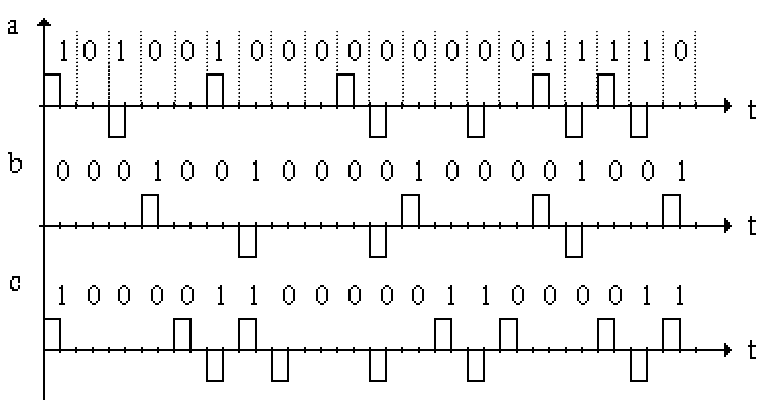
Răspuns:



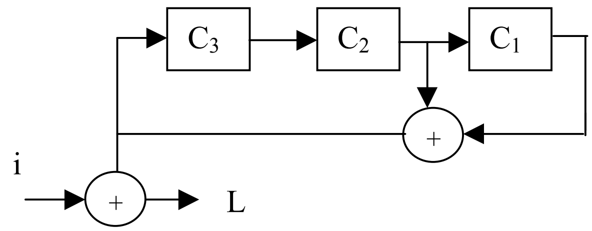
Componentă medie: NRZ - 11/25 A   RZ - 11/50 A   AMI - O   HDB3 - 1/50 A

4.9. Desenați semnalele de linie în cod HDB3 pentru următoarele secvențe de date:

a) 10100100000000011110;
b) 00010010000100001001;
c) 10000110000011000011;



4.10. Care va fi semnalul de linie (L) dacă la intrarea cifratorului din figură:



se aduce semnalul:

i=1110100000110000111...

Starea inițială a RDR este $S_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$**.**

<u>Răspuns:</u> 01010001011010101001.

Cap.5 Teste grilă

5.1 Dacă p(A) este probabilitatea de realizare a mesajului A, atunci informația furnizată de producerea sa , i(A), este:

a). lnA; *b). $-\log_2 p(A)$; c). $p(A)\cdot\log_2 \frac{1}{p(A)}$; d). $\log_2 p(A)$; e). $p(A)\cdot\ln p(A)$.

5.2 Entropia unei surse este:
- a). informația medie transmisă de sursă în unitatea de timp;
- b). informația instantanee transmisă de sursă;
- c). egală cu raportul dintre entropia maximă a sursei şi redundanța sa;
- *d). egală cu produsul dintre entropia maximă şi eficiența sursei;
- e). nici o variantă nu este corectă.

5.3 Cantitatea de decizie a unei surse este:
- a). informația medie transmisă de sursă printr-un simbol al său, dacă sursa este echiprobabilă;
- *b). egală cu entropia maximă a sursei;
- c). egală cu produsul dintre entropia maximă şi eficiența sursei;
- d). informația medie transmisă de sursă în unitatea de timp;
- e). nici o variantă nu e corectă.

5.4 Care dintre scopurile de mai jos nu se realizează prin codare:
- a). adaptarea naturii diferite a sursei la natura canalului;
- *b). multiplexarea mai multor semnale purtătoare de informație în vederea transmiterii pe acelaşi canal;
- c). compresia sursei;
- d). protejarea informației împotriva perturbațiilor;
- e). protejarea informației împotriva receptorilor neautorizați (secretizarea).

5.5 Care dintre mărimile de mai jos reprezintă criterii de fidelitate:

1). eroarea medie pătratică: $\varepsilon = \overline{[x(t) - y(t)]^2}$;

2). raportul semnal-zgomot: $\xi = \overline{y(t)^2}/\overline{u(t)^2}$;
3). echivocația;

a). doar 1).; b) 1). şi 3).; c) 2) şi 3).; d) toate trei; *e). 1) şi 2)..

5.6 Care dintre mărimile de mai jos reprezintă criterii de fidelitate:

1). eroarea medie $H(Y/X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i,y_j)\cdot\log_2 p(y_j/x_i)$;

2). raportul semnal-zgomot: $\xi=\overline{y(t)^2/u(t)}^2$;
3). rata erorii (BER).

a). doar 1).; b). 1). şi 3).; *c). 2) şi 3).; d). toate trei; e). 1). şi 2)..

5.7 O sursă discretă S este fără memorie dacă:

*a). emisia unui simbol nu depinde de simbolurile anterior emise;
b). probabilitățile diferitelor simboluri nu depind de originea timpului;
c). generează simboluri la o indicație exterioară;
d). generează simbolurile cu o viteză fixă;
e). nici un răspuns nu este corect;

5.8 O sursă discretă este cu debit necontrolat dacă:

a). emisia unui simbol nu depinde de simbolurile anterior emise;
b). probabilitățile diferitelor simboluri nu depind de originea timpului;
c). generează simboluri la o indicație exterioară;
*d). generează simbolurile cu o viteză fixă;
e). nici un răspuns nu este corect;

5.9 O sursă discretă S^n este extensia unei surse de ordin n a sursei S dacă:

a). emisia unui simbol nu depinde de simbolurile anterior emise;
b). probabilitățile diferitelor simboluri nu depind de originea timpului;
c). generează simboluri la o indicație exterioară;
d). generează simbolurile cu o viteză fixă;
*e). nici un răspuns nu este corect;

5.10 Informația obținută în urma realizării evenimentului a, a cărui şanse de realizare sunt 100% este:

*a). zero; b). un bit; c). un bit/simbol; d). strict pozitivă; e). nici un răspuns nu e corect.

5.11 Relațiile între unitățile de măsură a informației sunt:

a). 1nit<1bit<1dit;
b). 1bit<1dit<1nit;
c). 1dit<1nit<1bit;
d). 1dit<1bit<1nit;
*e). 1bit<1nit<1dit.

5.12 Entropia unei surse de informație:

a). reprezintă incertitudinea medie ce există apriori asupra emisiei;
b). reprezintă informația medie purtată de un simbol al sursei;
c). se calculează cu formula:

$$H(s)=\sum_{i=1}^{N} p(s_i)\cdot i(s_i),$$

unde: s_i i=1..N, -simbolurile sursei;
N -numărul de simboluri;
$i(s_i)$ -informația furnizată de simbolul s_i;
$p(s_i)$-probabilitatea de emisie a lui s_i;

*d). se măsoară în biți/secundă;
e). devine maximă dacă sursa este echiprobabilă.
Precizați răspunsul fals.

5.13 O SDFM , S, cu 32 de simboluri are entropia egală cu a unei surse echiprobabilă cu 26 de simboluri. Redundanța sursei S este:

*a). 0,3 bit/simbol;
b). 0,3 bit/secundă;
c). 4,7 bit/simbol;
d). 4,7 bit/secundă;
e). 6%.

5.14 Precizați răspunsul fals. Entropia unei SDFM este:

a). dată prin formula:

$$H(S)=\sum_{i=1}^{N} p(s_i)\cdot\log_2 p(s_i)$$,unde: s_i i=1..N, -simbolurile sursei;

N -numărul de simboluri;
$p(s_i)$ -probabilitatea de emisie a lui s_i;

b). măsurată în biți/simbol;
*c). o funcție continuă în raport cu fiecare variabilă s_i;
d). o funcție simetrică în raport cu toate variabilele $p(s_i)$;
e). aditivă.

5.15 O sursă de informație are entropia numeric egală cu 3,7 iar redundanța sursei este 5,3%. Sursa are un număr de simboluri egale cu:

a). 13; *b). 15; c). 16; d). 9; e). 4.

5.16 Semnalele utilizate pentru transportul informației numerice sunt compuse din suite de semnale elementare în timp, numite momente. Numărul de momente transmise în unitatea de timp este:

1). debitul de momente;
2). viteza de modulație;
3). viteza telegrafică.

Răspunsurile corecte sunt:

a). 1); b). 2); c). 3); d). nici unul; *e). toate trei.

5.17 Precizați răspunsul fals. Debitul de informație al unei surse de informație este:

*a). măsurat în biți/simbol;
b). măsurat în biți/secundă;
c). cantitatea medie de informație generată de sursă în unitatea de timp;
d). egal cu debitul de decizie, dacă sursa este echiprobabilă;
e). mai mic decât capacitatea canalului, la transmisiile în timp real.

5.18 Pentru un canal de transmisie discret, $X=\{x_i\}_{i=1..n}$ si $Y=\{y_j\}_{j=1..m}$ reprezintă câmpurile de la intrarea, respectiv ieşirea din canal. Matricea de trecere P este o matrice având dimensiunile nxm şi conține probabilități de forma:

a). $p(x_i)\cdot p(y_j)$; b). $p(x_i, y_j)$;
c). $p(x_i/y_j)$; *d). $p(y_j/x_i)$;
e). nici un răspuns nu e corect.

5.19 Un canal este binar simetric dacă:

a). $p(0_E)=P(1_E)$;
b). $p(0_E/1_R)=p(1_E/0_R)$;
*c). $p(0_R/1_E)=p(1_R/0_E)$;
d). $p(0_R)=p(1_R)$;
e). nici o variantă nu e corectă.

5.20 Un canal este binar simetric dacă:

a). $p(0_E,0_R)=p(0_E,1_R)$;
b). $p(0_E/0_R)=p(1_E/1_R)$;
c). $p(0_R/1_E)=p(1_R/1_E)$;
d). $p(0_R)=p(1_R)$;
*e). nici o variantă nu e corectă.

5.21 Mărimea H(y)-H(y/x) pentru un canal binar este:

*a). transinformația doar dacă $p(0_R)=p(1_R)$;
b). capacitatea canalului doar dacă $p(0_R)=p(1_R)$;
c). subunitară;
d). măsurată în biți/simbol binar;
e). egală cu H(x)-H(x/y) dacă $p(0_R)=p(1_R)$;
Precizați răspunsul fals.

5.22 Mărimile mărginite superior de H(x) – entropia câmpului de la intrarea în canal, sunt:

a). numai transinformația;
b). eroarea medie şi transinformația;
c). echivocația şi eroarea medie;
*d). echivocația şi transinformația;
e). echivocația, eroarea medie şi transinformația.

5.23 Mărimile mărginite superior de H(Y) – entropia câmpului de la ieşirea din canal, sunt:

a). numai transinformația;
*b). eroarea medie şi transinformația;
c). echivocația şi eroarea medie;
d). echivocația si transinformația;
e). echivocația, eroarea medie şi transinformația.

5.24 Capacitatea canalului binar simetric având rata erorii $BER=10^{-2}$ este:

*a). 0,9192 biți/simbol binar;
b). 0,944 biți/simbol binar;
c). 0,944 biți/secundă;
d). 0,9756 biți/secundă;
e). 0,9192 biți/secundă.

5.25 O SDFM având tabelul:

$$S=\begin{pmatrix} a & b & c & d \\ 0{,}4 & 0{,}2 & 0{,}25 & \end{pmatrix}$$

emite câte un simbol la fiecare milisecundă. Debitul sursei este:

a). 1,9 kbiți/simbol;
b). 1,3 kbiți/secundă;
c). 573 biți/secundă;
*d). 1,9 kbiți/secundă;
e). 1,3 kbiți/secundă.

5.26 Care din expresiile de mai jos reprezintă capacitatea unui canal de transmisiune analogic (considerat ca un FTJ ideal având lărgimea de bandă B şi raportul semnal-zgomot ξ - zgomot gaussian).

a) $C = \xi \cdot \log_2(1+B)$;
*b) $C = B \cdot \log_2(1+\xi)$;
c) $C = (B+1) \cdot \log_2 \xi$;
d) $C = B^2 \cdot \log_2(1+\xi)$;
e) $C = B \cdot \log_2(1+\xi^2)$;

5.27 Fie S o sursă discretă cu memorie având graful din figura 1:

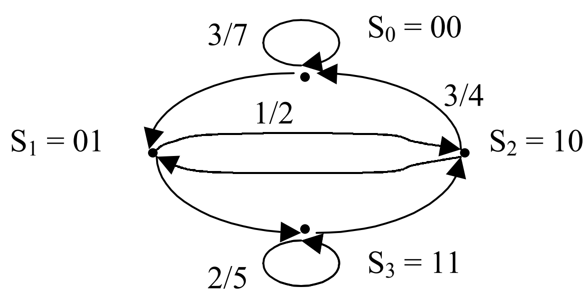


Figura.1

Sursa are memorie de:

a) un pas, deoarece orice tranziție se face între starea actuală şi starea viitoare;
b) doi paşi, deoarece orice tranziție se face între două stări;
*c) doi paşi, deoarece fiecare stare este numită prin doi paşi;
d) doi paşi, deoarece din fiecare stare pleacă două ramuri;
e) patru paşi, deoarece sunt patru stări.

5.28 Fie S o sursă discretă cu memorie având graful din figura 1. Ce şanse sunt ca sursa S să emită un zero dacă secvența emisă până în prezent a fost: 10011?

a) 25%; b) 30%; c) 40%; d) 50%; *e) 60%;

5.29 Sursa discretă cu memorie, S, al cărei graf este prezentat în figura 1, se găseşte în stare S_2. Care este probabilitatea ca după emisia a trei simboluri sursa să se găsească în starea S_1?

a) 3,125%;

b) 42,86%;

c) 18,37%;

*d) 21,5%;

e) 46%;

5.30 Fie S^n extensia sursei discrete şi fără memorie, S. Entropia sursei S^n, notată $H(S^n)$, se calculează funcţie de entropia sursei S, notată H(S), după relaţia:

*a) $H(S^n)=nH(S)$;

b) $H(S^n)=(H(S))^n$;

c) $H(S^n)=H(S)$;

d) $H(S^n)=H(S)\cdot\log_2 n$;

e) nici o variantă nu este corectă.

5.31 Eficienţa unei surse discrete fără memorie, având 20 simboluri echiprobabile, ce s-a codat cu un cod binar bloc, este:

a) $\frac{1}{5}\log_2 20$;

*b) 100%;

c) $5/\log_2 20$;

d) $4/\log_2 20$;

e) nici o variantă nu este corectă.

5.32 O SFDM având 20 simboluri echiprobabile se codează binar bloc. Eficienţa codării este:

*a) $\frac{1}{5}\log_2 20$;

b) 100%;

c) $5/\log_2 20$;

d) $4/\log_2 20$;

e) nici o variantă nu este corectă.

5.33 Teorema de existenţă a codurilor instantanee se exprimă prin inegalitatea:

$$\sum_{i=1}^{M} m^{-n_i} \le 1$$

(care reprezintă condiţia necesară şi suficientă de existenţă a codurilor instantanee).
În relaţia de mai sus, se reprezintă numărul de simboluri:

a) de control;

b) de informaţia;

c) ale sursei de informaţie;

*d) ale alfabetului codului;

e) nici o variantă nu este corectă.

5.34 Teorema de existență a codurilor instantanee se exprimă prin inegalitatea:

$$\sum_{i=1}^{M} m^{-n_i} \le 1$$

(care reprezintă condiția necesară şi suficientă de existență a codurilor instantanee).
În relația de mai sus, M reprezintă numărul de simboluri:

a) de control;
b) de informația;
*c) ale sursei de informație;
d) ale alfabetului codului;
e) nici o variantă nu este corectă.

5.35. Teorema I-a a lui Shannon (teorema codării surselor pentru transmitere pe canale fără perturbații).

a) afirmă că poate fi făcută o codare absolut optimală;
*b) se adresează doar codării surselor cu probabilitățile simbolurilor de forma:
$p(S_i) = m^{-n_i}$ cu $n_i \in N$
c) nu precizează procedee de codare;
d) are în vedere o codare pe grupe de n simboluri a sursei, cu $n \to \infty$.
e) este valabilă şi pentru coduri ternare.

Precizați răspunsul fals.

5.36 Un cod absolut optimal:

a) are eficiență 100%;
b) se poate obține pentru surse la care probabilitățile simbolurilor sunt de forma
$p(S_i) = m^{-n_i}$ cu $n_i \in N$;
c) pentru o sursă oarecare este doar o limită teoretică;
*d) nu poate fi obținut, indiferent de sursă, printr-o codare simbol cu simbol;
e) are lungimea medie a cuvintelor egală cu entropia sursei ce o codează.

Precizați răspunsul fals.

5.37 Un cod optimal:

a) Se poate obține prin algoritmul Huffman static?
b) are eficiența subunitară;
c) este un cod bloc dacă $N=2^k$ - numărul de cuvinte de cod;
d) este un cod bloc dacă $N=2^k$ - numărul de simboluri ale sursei codate;
*e) este neapărat un cod binar.

Precizați răspunsul fals.

5.38 Algoritmul de codare Huffman static:

*a) conduce la un cod absolut optimal,
b) conduce la un cod de eficiență maxim posibilă;
c) presupune ordonare descrescătoare a simbolurilor sursei după probabilități;
d) se poate aplica şi surselor echiprobabile;
e) se poate aplica şi la coduri nebinare.

Precizați răspunsul fals.

5.39 Algoritmii de compresie realizează micşorarea:
- a) cantităţii de informaţie conţinută într-un anumit mesaj;
- *b) spaţiului ocupat de un mesaj;
- c) debitului de informaţie la o transmisie;
- d) vitezei de transmitere a informaţiei;
- e) nici o variantă nu este corectă.

5.40 Algoritmul de compresie Huffman dinamic:
- a) realizează o compresie superioară celui static;
- b) necesită cunoaşterea statisticii a sursei date;
- c) presupune modificarea codului după transmiterea fiecărui simbol;
- d) realizează o codare pe grupe de n simboluri, cu $n\to\infty$;
- *e) nici o variantă nu este corectă.

5.41 Suma cifrelor mesajului VARZA, cifrat cu cifrul lui Polybius este:
a) 20, b) 25; *c) 26; d) 28; e) nici un răspuns nu este corect.

5.42 Suma cifrelor mesajului DOVLEAC cifrat cu cifrul lui Polybius este:
a) 26; b) 28; c) 31; *d) 34; e) 39.

5.43 Aflaţi răspunsul la întrebare descifrând mesajul EQDVYXLQ:
a) 1°; b) 2°; *c) 3°; d) 4°; e) 5°;

5.44 Aflaţi răspunsul la întrebare descifrând mesajul DOGRLOHD:
a) 1°; *b) 2°; c) 3°; d) 4°; e) 5°;

5.45 Aflaţi răspunsul la întrebare descifrând mesajul ECWE IK HACKKU:
a) de la 1 la 6; b) de la 7 la 12; c) de la 13 la 18; d)de la 19 la 24; e)de la 25 la 31;

5.46 O sursă de informaţie discretă şi fără de memorie se codează cu un cod binar bloc de eficienţă maxim posibilă. Cât trebuie să fie lungimea cuvintelor codului bloc dacă sursa are 15 simbolurilor?
a) 3 biţi; *b) 4 biţi; c) 15 biţi; d) 16 biţi; e) nici o variantă nu este corectă.

5.47 Codând prin algoritmul Huffman static o sursă de informaţie oarecare cu 15 simboluri, lungimea maximă a cuvintelor codului ar putea fi (se are în vedere toate sursele posibile cu 15 simboluri).
a) 4 biţi; *b) 14 biţi; c) 15 biţi; d) 16 biţi; e) nici o variantă nu este corectă.

5.48 Codând prin algoritmul Huffman static o sursă de informaţie oarecare cu 15 simboluri, lungimea cea mai mică dintre cuvintele codului ar putea fi (se are în vedere toate sursele posibile cu 15 simboluri).
a) 4 biţi; b) 14 biţi; c) 3 biţi; d) 2 biţi; *e) 1 bit.

5.49 Sursa de informație text TEORIA TRANSMITERII INFORMATIEI se codează binar bloc. Lungimea medie a cuvintelor codului este:

a) 3,1 biți/secundă;
b) 2,15 biți/secundă;
c) 4 biți/secundă;
d) 10 biți/secundă;
*e) nici o variantă nu este corectă.

5.50 Sursa de informație text TEORIA TRANSMITERII INFORMATIEI se codează cu un cod de eficiență maxim posibilă (codare simbol cu simbol). Lungimea medie a cuvintelor codului este:

*a) 3,1 biți/simbol;
b) 2,15 biți/simbol;
c) 4 biți/simbol;
d) 10 biți/simbol;
e) nici o variantă nu este corectă.