CAPITOLUL 1

SURSE DISCRETE DE INFORMAȚIE

1.1. Model matematic de sursă discretă, completă şi fără memorie

În general, prin sursă de informație se înțelege mecanismul prin care, din mulțimea mesajelor sursei *se alege*, într-un mod imprevizibil pentru destinatar, un anumit mesaj pentru a fi transmis.

Modelul matematic al unei astfel de surse se poate identifica cu un experiment S ale cărui rezultate pun în evidență un număr finit de evenimente elementare independente, notate cu $s_1, s_2, ..., s_n$, având asociate probabilitățile de realizare $p(s_1), p(s_2)$,..., respectiv $p(s_n)$. Experimentul S formează un *sistem complet* de evenimente, dacă la efectuarea acestuia cu certitudine se realizează unul din evenimente. Cu alte cuvinte, dacă $0 \le p(s_k) \le 1$, oricare ar fi $k = \overline{1,n}$, atunci are loc relația

$$s_1 \cup s_2 \cup ... \cup s_n = E \qquad (1.1)$$

unde E este evenimentul sigur. Dacă evenimentele s_i sunt disjuncte, atunci:

$$p\left(s_1 \cup s_2 \cup ... \cup s_n\right) = \sum_{k=1}^{n} p\left(s_k\right) = p\left(E\right) = 1 \qquad (1.1')$$

În felul acesta, experimentul S poate fi descris de câmpul de probabilitate caracterizat prin distribuția

$$S : \begin{pmatrix} s_1 & ... & s_j & ... & s_n \\ p(s_1) & ... & p(s_j) & ... & p(s_n) \end{pmatrix} \qquad (1.2)$$

Sursa de informație caracterizată prin experimentul S sau câmpul de probabilitate cu distribuția dată de (1.2) se va numi *discretă*, dacă furnizează un număr finit de mesaje s_j, $j = \overline{1, n}$.

Sursa de informație se va numi *completă*, dacă este îndeplinita relația (1.1).

Sursa de informație se va numi *fără memorie*, dacă furnizarea unui mesaj la un moment dat nu depinde de mesajele furnizate anterior.

Deoarece informația şi gradul de nedeterminare pot avea aceeaşi măsură, iar gradul de nedeterminare este o funcție de probabilitatea de realizare a evenimentului (mesajului), se poate considera că *informația* ataşată (asociată) unui mesaj s_k, notată cu $i(s_k)$ este, de asemenea, o funcție de probabilitatea $p(s_k)$, cu care acesta este furnizat, adică

$$i(s_k) = F[p(s_k)] \qquad (1.3)$$

unde F este o funcție ce urmează a fi determinata.

Evident, funcția F este monoton descrescătoare, deoarece cu cât probabilitatea de furnizare a mesajului s_k este mai mare, cu atât nedeterminarea asupra mesajului este mai mică şi, deci, şi informația ataşată acestui mesaj va fi mai mică (Fig. 1.1).

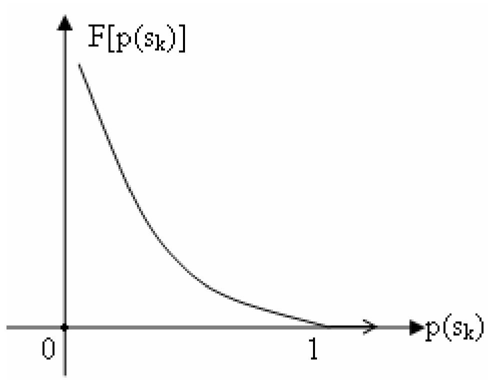

Fig. 1.1. Reprezentarea grafică a informației

La limită, când $p(s_k)=1$, rezultă că se cunoaşte cu certitudine că sursa va furniza mesajul s_k şi, deci, informația obținută la furnizarea mesajului s_k este nulă.

Când $p(s_k)=0$, rezultă că sursa nu va furniza niciodată mesajul s_k şi, deci, neexistând nici o nedeterminare asupra furnizării sau nefurnizării acestui mesaj, informația ataşată acestuia este, şi în acest caz, nulă (vezi punctul din origine).

În scopul determinării funcției *F*, se consideră o sursă discretă, completă şi fără memorie, caracterizată de distribuția (1.2).

Fie $i(s_k \cap s_j)$ informația care s-ar obține la furnizarea mesajului s_k şi a mesajului s_j. Daca probabilitatea evenimentului de furnizare a mesajului s_k şi a mesajului s_j este $p(s_k \cap s_j)$, conform relație (1.3), se poate scrie

$$i(s_k \cap s_j) = F\left[p(s_k \cap s_j)\right] \tag{1.4}$$

Sursa de informație fiind presupusă fără memorie, rezultă că mesajele s_k şi s_j sunt independente şi, deci, probabilitatea intersecției celor două evenimente (mesaje) devine

$$p\left(s_k \cap s_j\right) = p\left(s_k\right) \cdot p\left(s_j\right) \tag{1.5}$$

Pe de altă parte, cele două mesaje fiind independente, informația obținută la furnizarea lor va fi egală cu suma informațiilor atașate celor două mesaje, adică

$$i\left(s_k \cap s_j\right) = i\left(s_k\right) + i\left(s_j\right) \tag{1.6}$$

Ținând cont de relațiile (1.3), (1.4) şi (1.5), din (1.6) rezultă ecuația funcțională

$$F\left[p\left(s_k\right) \cdot p\left(s_j\right)\right] = F\left[p\left(s_k\right)\right] + F\left[p\left(s_j\right)\right] \tag{1.7}$$

Ecuația funcțională (1.7) este satisfăcută, evident, de funcția

$$F\left[p(s_k)\right] = C \cdot \log p(s_k) \tag{1.8}$$

unde constanta C se va determina odată cu adoptarea unității de măsură pentru informație.

Prin definiție, se consideră că se obține o unitate de informație, atunci când o sursă discretă, completă şi fără memorie poate furniza numai două mesaje echiprobabile.

Considerând în relația (1.8) logaritmul în baza doi şi ținând cont de definiția unității de măsură a informației, se poate scrie

$$i\left(s_k\right) = F\left[p\left(s_k\right)\right] = C \log_2 \frac{1}{2} = -C = 1 \tag{1.9}$$

Unitatea de informație astfel definită se numeşte bit. Din (1.9) rezultă că C = -l şi, deci, se poate scrie, în general, că informația atașată mesajului s_k, furnizat cu probabilitatea $p\left(s_k\right)$, se determină cu relația

$$i\left(s_k\right) = -\log_2 p\left(s_k\right) \ \text{<biți>} \tag{1.10}$$

Considerându-se din nou sursa discretă, completă şi fără memorie caracterizată de distribuția (1.2), informațiile $i\left(s_k\right)$ atașate mesajelor s_k, $k=\overline{1,n}$, conform relației (1.10), determină o variabilă aleatoare discretă, a cărei valoare medie statistică, prin definiție, se numeşte entropia sursei respective şi va fi notată în continuare cu H(S). Se poate, deci, scrie relația

$$H(S)=\overline{i(s_k)}=\sum_{k=1}^{n}p(s_k)\cdot i(s_k) \tag{1.11}$$

unde $\overline{i(s_k)}$ reprezintă valoarea medie statistică a variabilei aleatoare discrete $i(s_k)$.

Înlocuind (1.10) în (1.11), rezultă

$$H(S)=-\sum_{k=1}^{n}p(s_k)\log_2 p(s_k) < \frac{\text{biți}}{\text{mesaj}} > \tag{1.12}$$

Din punct de vedere fizic, entropia definită cu relația (1.12) măsoară informația medie pe mesaj, respectiv nedeterminarea medie a sursei respective. In scopul simplificării scrierii, în cele ce urmează, se va considera că logaritmii sunt în baza doi, fără a mai specifica acest lucru.

Dacă se notează cu τ_k, $k=\overline{1,n}$, duratele de transmitere a mesajelor s_k de către sursa S, caracterizată de distribuția (1.2), atunci durata medie de transmitere a unui mesaj se determină cu relația

$$\overline{\tau}=\sum_{k=1}^{n}p(s_k)\tau_k \quad <\text{secunde}> \tag{1.13}$$

Prin definiție, se numeşte *debit de informație* al unei surse discrete, complete şi fără memorie, raportul dintre entropia sursei şi durata medie de transmitere a unui mesaj. Notând cu $H_\tau(S)$ debitul unei astfel de surse, se poate scrie

$$H_\tau(S)=\frac{H(S)}{\overline{\tau}} \qquad < \frac{\text{biți}}{\text{mesaj}\times\text{secunda}} > \tag{1.14}$$

1.2. Principalele proprietăți ale entropiei

1. Din relația (1.12) rezultă că entropia este nenegativă, adică

$$H(S) \geq 0 \tag{1.15}$$

2. Prin diversificarea unei surse, entropia crește.

Într-adevăr, fie sursa discretă, completă și fără memorie caracterizată de distribuția (1.2). Fără a micșora generalitatea, se presupune că mesajul s_n poate fi diversificat în alte m mesaje disjuncte, adică

$$s_n = y_1 \cup y_2 \cup ... \cup y_m \tag{1.16}$$

Dacă $p(y_1), p(y_2), ..., p(y_m)$ sunt probabilitățile de furnizare ale mesajelor $y_1, y_2, ..., y_m$, atunci distribuția sursei diversificate este

$$S_d : \begin{pmatrix} s_1 & ... & s_{n-1} & y_1 & ... & y_m \\ p(s_1) & ... & p(s_{n-1}) & p(y_1) & ... & p(y_m) \end{pmatrix} \tag{1.17}$$

iar

$$p(s_n) = \sum_{j=1}^{m} p(y_j) \tag{1.18}$$

Conform relației (1.12), entropia sursei diversificate se poate calcula cu relația

$$H(S_d) = \sum_{k=1}^{n-1} p(s_k) \log p(s_k) - \sum_{j=1}^{m} p(y_j) \log p(y_j) \tag{1.19}$$

sau

$$H(S_d) = -\sum_{k=1}^{n} p(s_k) \log p(s_k) + p(s_n) \log p(s_n) - \sum_{j=1}^{m} p(y_j) \log p(y_j)$$

Ținând cont de relațiile (1.12) și (1.18), relația anterioară devine

$$H(S_d) = -H(S) - p(s_n)\sum_{j=1}^{m} \frac{p(y_j)}{p(s_n)} \log \frac{p(y_j)}{p(s_n)} \tag{1.20}$$

Deoarece

$$0 \leq \frac{p(y_j)}{p(s_n)} \leq 1 \tag{1.21}$$

şi

$$\sum_{j=1}^{n} \frac{p(y_j)}{p(s_n)} = 1, \tag{1.22}$$

rezultă că

$$-\sum_{j=1}^{m} \frac{p(y_j)}{p(s_n)} \log \frac{p(y_j)}{p(s_n)} \geq 0 \tag{1.23}$$

Din (1.20) şi (1.23) rezultă ce trebuia demonstrat, adică

$$H(S_d) \geq H(S) \tag{1.24}$$

3. Entropia unei surse discrete, complete şi fără memorie îşi atinge valoarea maximă, atunci când mesajele sunt furnizate echiprobabil.

Într-adevăr, fie sursa discretă, completă şi fără memorie caracterizată de distribuția (1.2), pentru care există relația de legătură (1.1). Pentru a determina extremul entropiei acestei surse, se va folosi metoda multiplicatorilor lui Lagrange (sau extremul cu legături). Pentru aceasta, se defineşte funcția.

$$\Phi\left[p(s_k)\right] = H(s) + \lambda\left[\sum_{k=1}^{n} p(s_k) - 1\right] \tag{1.25}$$

unde λ este multiplicatorul lui Lagrange. Funcția $\Phi\left[p(s_k)\right]$ îşi atinge extremul odată cu funcția H(S).

Condiția necesară de extrem este

$$\frac{\partial \Phi\left[p\left(s_k\right)\right]}{\partial p\left(s_k\right)}=0 \quad (\forall)\ k=\overline{1,n} \tag{1.26}$$

Înlocuind (1.12) în (1.25), relația (1.26) devine

$$-\log p(s_k)-\log e+\lambda=0, \quad (\forall)k=\overline{1,n} \tag{1.27}$$

de unde rezultă

$$p(s_k)=2^{\lambda-\log e}, \quad (\forall)k=\overline{1,n} \tag{1.28}$$

Înlocuind (1.28) in (1.1), rezultă

$$n\left(2^{\lambda-\log e}\right)=1 \tag{1.29}$$

de unde se obține

$$\lambda=\log\frac{e}{n} \tag{1.30}$$

Ținând cont de (1.30), relația (1.28) devine

$$p\left(s_k\right)=\frac{1}{n}, \qquad (\forall)k=\overline{1,n} \tag{1.31}$$

Din (1.31) rezulta că extremul funcției $\phi\left[p\left(s_k\right)\right]$, deci și a entropiei H(S), se obține atunci când mesajele sunt furnizate echiprobabil. Valoarea acestui extrem se obține înlocuind (1.31) în (1.12), adică

$$H\left(S\right)\Big|_{p\left(s_k\right)=\frac{1}{n}}=-\sum_{k=1}^{n}p\left(s_k\right)\log p\left(s_k\right)\Big|_{p\left(s_k\right)=\frac{1}{n}}=\log n \tag{1.32}$$

Pentru a arata că extremul astfel obținut este un maxim, se poate demonstra ușor că

$$H\left(S\right)\le\log n \tag{1.33}$$

sau, echivalent, că

$$H\left(S\right)-\log n\le 0 \tag{1.34}$$

Într-adevăr,

$$H(S)-\log n=-\sum_{k=1}^{n}p(s_k)\log p(s_k)-\log n=$$

$$\sum_{k=1}^{n}p(s_k)\log\frac{1}{p(s_k)}-\log n\cdot\sum_{k=1}^{n}p(s_k)= \tag{1.35}$$

$$=\sum_{k=1}^{n}p(s_k)\left[\log\frac{1}{p(s_k)}-\log n\right]=\sum_{k=1}^{n}p(s_k)\log\frac{1}{n\cdot p(s_k)}$$

Pe de altă parte, este binecunoscută inegalitatea

$$\ln z\le z-1,\ z>0\,, \tag{1.36}$$

a cărei reprezentare grafică este dată în Fig. 1.2.

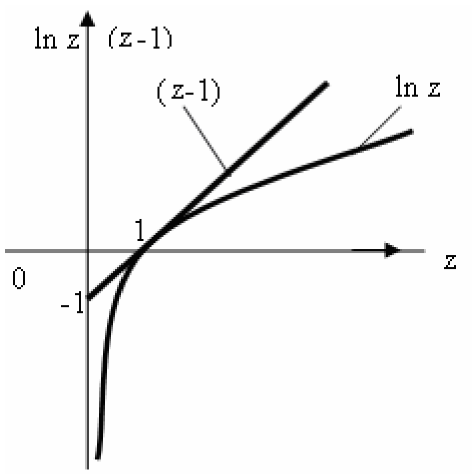

Figura 1.2 Reprezentarea grafică a inegalității 1.36

Ținând cont de 1.36 şi notând $z=\dfrac{1}{n\cdot p(s_k)}$, relația (1.35) devine

$$H(S)-\log n\le\log e\sum_{k=1}^{n}\frac{1}{n\cdot p(s_k)}-1 \tag{1.37}$$

sau, echivalent,

$$H(S) - \log n \le \log e \left[1 - \sum_{k=1}^{n} p(s_k) \right] = 0 \qquad (1.38)$$

ceea ce demonstrează relația (1.34).

Prin definiție, se numeşte *redundanță absolută* a unei surse, diferența dintre valoarea maximă a entropiei acesteia şi valoarea ei reală. Notând cu R_S, redundanța absolută, conform definiției, se poate scrie

$$R_S = H_{\max}(S) - H(S) \qquad (1.39)$$

Prin definiție, se numeşte *redundanță relativă* a unei surse, şi se va nota cu ρ_S raportul dintre redundanța absolută şi valoarea maximă a entropiei, adică

$$\rho_S = \frac{R_S}{H_{\max}(S)} = 1 - \frac{H(S)}{H_{\max}(S)}, \qquad (1.40)$$

Prin definiție, se numeşte eficiența unei surse, şi va fi notată cu η_s, raportul dintre entropia sursei şi valoarea maximă a acesteia, adică

$$\eta_S = \frac{H(S)}{H_{\max}(S)}, \qquad (1.41)$$

Din (1.40) şi (1.41) , rezultă

$$\rho_S = 1 - \eta_S \qquad (1.42)$$

Pentru fixarea ideilor, se presupune o sursă discretă, completă şi fără memorie, caracterizată de distribuția

$$S : \begin{pmatrix} s_1 & s_2 \\ p & 1-p \end{pmatrix}, \ 0 \le p \le 1 \qquad (1.43)$$

Entropia acestei surse se calculează cu relația

$$H(S) = -p \log p - (1-p)\log(1-p) = H(p) \qquad (1.44)$$

Deoarece

$$\lim_{x \to 0}[x \cdot \log x] = 0 \tag{1.45}$$

prin convenție, se va considera că

$$0 \cdot \log 0 = 0 \tag{1.46}$$

Ținând cont de (1.46), rezultă

$$H(0) = H(1) = 0 \tag{1.47}$$

Interpretarea fizică a rezultatelor date de relația (1.47) este următoarea: când *p*=0, rezultă cu certitudine că sursa va furniza mesajul s_2 şi, deci, neexitând nici o incertitudine, $H(0) = 0$; când *p*=1, rezultă cu certitudine că sursa va furniza mesajul s_1 şi, deci, $H(1) = 0$.

Valoarea maximă a entropiei se obține atunci când

$$p = 1 - p \Rightarrow p = \frac{1}{2} \text{ şi } H_{\max}(S) = 1 \frac{\text{bit}}{\text{mesaj}} \tag{1.48}$$

Reprezentarea grafică a entropiei $H(p)$, conform relației (1.44), este dată în Fig. 1.3.

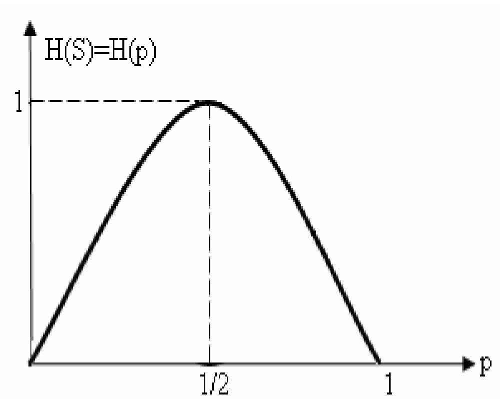

Fig. 1.3. Reprezentarea grafică a entropiei date de relația (1.44)

1.3. Model matematic de sursă discretă, completă şi fără memorie extinsă

În multe modele de surse de informaţie este util a se considera că sursa furnizează grupe de mesaje şi nu mesaje individuale. În general, *extensia de ordinul m* a unei surse discrete, complete şi fără memorie se obţine formând mesaje compuse, σ_k, din succesiuni de m mesaje ale sursei iniţiale, în toate combinaţiile posibile. Numărul de mesaje compuse al unei astfel de surse extinse este egal cu n^m, deoarece fiecare mesaj compus este format dintr-o succesiune de m mesaje, iar fiecare mesaj poate fi oricare din cele n posibile. Extensia unei surse discrete, complete şi fără memorie determină o nouă sursă discretă, completă şi fără memorie, cu probabilitatea fiecărui mesaj compus egală cu produsul probabilităţilor mesajelor componente, deoarece mesajul compus reprezintă o intersecţie de evenimente independente.

Pentru calcularea entropiei extensiei unei surse discrete, complete şi fără memorie se va folosi metoda inducţiei matematice.

Fie sursa discretă, completă şi fără memorie, cu distribuţia (1.2). Extensia de ordinul doi a acesteia, notată cu S^2, va furniza n^2 mesaje compuse, σ_k de forma

$$\begin{array}{llll}
\sigma_1 \to s_1 s_1 & \sigma_{n+1} \to s_2 s_1 & \cdots & \sigma_{n^2-n+1} \to s_n s_1 \\
\sigma_2 \to s_1 s_2 & \sigma_{n+2} \to s_2 s_2 & \cdots & \sigma_{n^2-n+2} \to s_n s_2 \\
 & \ldots\ldots\ldots\ldots\ldots\ldots & & \\
\sigma_n \to s_1 s_n & \sigma_{2n} \to s_2 s_n & \cdots & \sigma_{n^2} \to s_n s_n
\end{array}$$

Entropia acestei surse, notată cu $H\left(S^2\right)$, se va calcula cu relaţia clasică

$$H\left(S^2\right)=-\sum_{k=1}^{n^2} p\left(\sigma_k\right)\log p\left(\sigma_k\right), \tag{1 .49}$$

Considerând un mesaj compus $\sigma_k = s_1 s_j$, cu $p(\sigma_k) = p(s_i)p(s_j); i, j = \overline{1,n}$, relația (1.49) se poate scrie, echivalent, sub forma

$$H\left(S^2\right)=-\sum_{i=1}^{n}\sum_{j=1}^{n} p\left(s_i\right)p\left(s_j\right)\log\left[p\left(s_i\right)p\left(s_j\right)\right]=$$

$$=-\sum_{i=1}^{n} p\left(s_i\right)\log p\left(s_i\right)\cdot\sum_{j=1}^{n} p\left(s_j\right)- \tag{1.50}$$

$$-\sum_{j=1}^{n} p\left(s_j\right)\log p\left(s_j\right)\cdot\sum_{i=1}^{n} p\left(s_j\right)=H\left(S\right)+H\left(S\right)=2H\left(S\right)$$

Considerând adevărată relația pentru un număr oarecare, m, adică

$$H\left(S^m\right)=mH\left(S\right), \tag{1.51}$$

se poate demonstra că

$$H\left(S^{m+1}\right)=\left(m+1\right)H\left(S\right) \tag{1.52}$$

Într-adevăr

$$H(S^{m+1})=-\sum_{k=1}^{n^{m+1}} p(\sigma_k)\log p(\sigma_k),\ (\forall)i=\overline{1,n} \tag{1.53}$$

unde

$$\sigma=\sigma_{k_1}s_i, p\left(\sigma_k\right)=p\left(\sigma_{k_i}\right)p\left(s_i\right)\ , \tag{1.54}$$

iar

$$\sigma_{k_1}=s_{i_1}s_{i_2}\cdots s_{i_m},(\forall)s_{i_j}\in S \tag{1.55}$$

Ținând cont de (1.54), relația (1.53) se poate scrie, echivalent, sub forma

$$H\left(S^{m+1}\right)=-\sum_{k_1=1}^{n^m}\sum_{i=1}^{n} p\left(\sigma_{k_1}\right)\cdot p\left(s_i\right)\log\left[p\left(\sigma_{k_1}\right)\cdot p\left(s_i\right)\right]=$$

$$= -\sum_{k_1=1}^{n^m} p\left(\sigma_{k_1}\right) \log p\left(\sigma_{k_1}\right) \cdot \sum_{i=1}^{n} p\left(s_i\right) - \sum_{i=1}^{n} p\left(s_i\right) \log p\left(s_i\right) \sum_{k=1}^{n^m} p\left(\sigma_{k_1}\right) =$$

$$= H\left(S^m\right) + H\left(S\right) = mH\left(S\right) + H\left(S\right) = \left(m+1\right) H\left(S\right) \tag{1.56}$$

1.4. Model matematic de sursă de informație, discretă, completă cu memorie

Aceste surse de informație, cunoscute şi sub denumirea de surse Markov, sunt caracterizate de faptul că furnizarea unui mesaj este dependentă de unul sau mai multe mesaje furnizate anterior.

Prin definiție, o sursă are *memorie de ordinul m*, dacă furnizarea unui mesaj este condiționată de ultimele m mesaje furnizate anterior.

Prin definiție, se numeşte *stare la un moment dat* a unei surse cu memorie succesiunea ultimelor m mesaje furnizate anterior momentului considerat. Dacă alfabetul unei surse Markov este format din n mesaje, iar ordinul memoriei sursei este m, atunci, conform definiției stărilor, sursa va avea cel mult n^m stări distincte, deoarece cele m mesaje care definesc starea pot fi oricare din cele n mesaje ce formează alfabetul sursei.

Analiza surselor de informație discrete, complete, cu memorie se efectuează de obicei fie folosind grafurile, fie calculul matriceal. În primul caz, *nodurile* din graf vor reprezenta stările sursei, iar ramurile orientate vor fi înzestrate cu transmitanțe egale cu probabilitățile de trecere dintr-o stare în alta.

Fie o sursă cu memorie de ordinul m ce poate furniza n mesaje. Fie starea S_i, caracterizată de succesiunea a m mesaje, adică

$$S_i \rightarrow s_{i_1} s_{i_2} \ldots s_{i_m}$$

Dacă din această stare se furnizează mesajul s_k cu probabilitatea condiționată $p\left(s_k \middle| S_i\right)$, reprezentând probabilitatea de a se furniza mesajul s_k, dacă sursa se afla în starea S_i, sursa va trece in starea S_j, caracterizată de succesiunea

$$S_j \rightarrow s_{i_2} \ldots s_{i_m} s_k$$

Pentru simplificarea scrierii, probabilitatea de trecere din starea S_i în starea S_j se va nota cu p_{ij}, iar în graf se va reprezenta ca în Fig. 1.4.

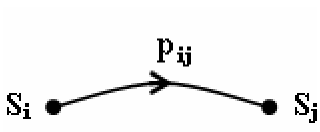

Figura 1.4. Trecerea din starea S_i în starea S_j

În cazul folosirii matricelor în analiza surselor cu memorie, probabilitățile de trecere dintr-o stare în alta, p_{ij}, se organizează într-o matrice, notată în continuare cu $[T]$, numită *matrice de trecere* sau *tranziție*, după următoarele reguli:

a) pe prima linie se trec probabilitățile de trecere din starea S_1, în celelalte n^m stări posibile, pe a doua linie, probabilitățile de trecere din starea S_2 în celelalte n^m stări posibile şi aşa mai departe, pe ultima linie scriindu-se probabilitățile de trecere din starea S_{n^m} în celelalte n^m stări posibile;

b) în prima coloană se trec probabilitățile de trecere în starea

S_1, din cele n^m stări posibile, în coloana a doua, probabilitățile de trecere în starea S_2 din cele n^m posibile şi aşa mai departe, în ultima coloană scriindu-se probabilitățile de trecere în starea S_{n^m} din cele n^m stări posibile, în felul acesta, matricea rezultată va fi pătrată, de ordinul n^m, de forma

$$[T]=\begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n^m} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n^m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n^m,1} & p_{n^m,2} & \cdots & p_{n^m,n^m} \end{bmatrix} \tag{1.57}$$

Matricea de trecere, [T], astfel întocmită, este stochastică, adică suma elementelor de pe fiecare linie este egală cu unitatea, deoarece cu certitudine dintr-o anumită stare se ajunge în una din stările posibile.

Prin definiție, sursa cu memorie se numeşte *staționară*, dacă probabilitățile de trecere p_{ij}, sunt invariante în timp.

1.5. Surse cu memorie ergodice

Fie o sursă cu memorie de ordin *m*, ce poate furniza *n* mesaje. Fie $p_i(t)$, $i=\overline{1,n^m}$, probabilitățile celor n^m stări la momentul *t*. Aceste probabilități definesc matricea distribuției probabilităților stărilor sursei la momentul *t,* de forma

$$\left[P(t)\right]=\left[p_1(t)\,p_2(t)\ldots p_{n^m}(t)\right] \tag{1.58}$$

La momentul t+1, *matricea distribuției probabilităților* stărilor va fi de forma

$$\left[P\left(t+1\right)\right]=\left[p_1\left(t+1\right)p_2\left(t+1\right)...p_{n^m}\left(t+1\right)\right] \tag{1.59}$$

Probabilitatea $p_j\left(t+1\right)$ ca sursă cu memorie să se afle la momentul t+1 în starea S_j se poate determina multiplicând probabilitatea $p_i\left(t\right)$ ca sursa să se afle la momentul t în starea S_i cu probabilitatea de trecere p_{ij} şi sumând apoi pentru toate cele n^m stări posibile S_i , adică

$$p_j(t+1)=\sum_{i=1}^{n^m} p_i(t)\cdot p_{ij} \text{ ,}(\forall)\ j=\overline{1,n^m} \tag{1.60}$$

În relația (1.60) este posibil ca o parte din probabilitățile p_{ij} să fie nule, ceea ce semnifică faptul că din anumite stări la momentul t nu se poate trece în starea S_j la momentul t+1.

Ținând cont de relațiile (1.57), (1.58) şi (1.59), relația (1.60) se poate scrie, sub formă matriceală, astfel

$$\left[P\left(t+1\right)\right]=\left[P\left(t\right)\right]\left[T\right] \tag{1.61}$$

În cazul momentelor de timp discrete t=0, l, 2, ..., relația (1.61) se scrie în variabila discretă n, sub forma

$$\left[P(n+1)\right]=\left[P(n)\right]\cdot\left[T\right] \text{ , n=0,1,\dots} \tag{1.62}$$

Pentru n=0, relația (1.62) devine

$$\left[P\left(1\right)\right]=\left[P\left(0\right)\right]\left[T\right] \tag{1.63}$$

Pentru n=l, relația (1.62) devine

$$\left[P\left(2\right)\right]=\left[P\left(1\right)\right]\left[T\right] \tag{1.64}$$

sau, ținând cont de (1.63), rezultă, echivalent,

$$\left[P\left(2\right)\right]=\left[P\left(0\right)\right]\left[T\right]^2 \tag{1.65}$$

Raționând în mod similar, se poate scrie în general

$$\left[P\left(n\right)\right]=\left[P\left(0\right)\right]\left[T\right]^n \tag{1.66}$$

Prin definiție, o sursă cu memorie, staționară se numeşte *ergotică,* dacă este satisfăcută relația

$$\lim_{n\to\infty}[T]^n=[K] \tag{1.67}$$

unde $[K]$ este o matrice cu elemente finite.

Dacă o sursă cu memorie este ergodică, rezultă că după un anumit timp (ceea ce ar corespunde regimului tranzitoriu într-un sistem liniar) probabilitățile stărilor nu se mai modifică în timp (corespunzând regimului permanent).

Notând $p_1, p_2, ..., p_{n^m}$ probabilitățile stărilor sursei cu memorie ergodice, matricea distribuției probabilităților stărilor devine

$$[P]=\left[p_1\ p_2\ ...\ p_{n^m}\right] \tag{1.68}$$

iar relația (1.62) se poate scrie sub forma

$$[P]=[P][T] \tag{1.69}$$

Relația matriceală (1.69) este echivalentă cu un sistem de n^m ecuații algebrice liniare în necunoscutele p_i, $i=\overline{1,n^m}$, dar, deoarece matricea de trecere $[T]$ este stochastică, din cele n^m ecuații, numai n^m-1 sunt liniar independente. Pentru a putea calcula probabilitățile stărilor sursei cu memorie ergodice, la relația (1.69) se mai adaugă o relație, de forma

$$\sum_{k=1}^{n^m} p_k=1 \tag{1.70}$$

Ecuația (1.70) se justifică prin faptul că dacă sursa se află într-o anumită stare, cu certitudine va trece la momentul următor în una din stările posibile.

Determinarea lui $[\mathrm{T}]^n$ în scopul verificării dacă o sursă cu memorie, staționară este sau nu, ergodică, se poate efectua folosind transformata Z [93]. Astfel, aplicând transformata Z relației (1.62), rezultă

$$z\left[\left[P(z)\right]-\left[P(0)\right]\right]=\left[P(z)\right]\left[T\right] \qquad (1.71),$$

sau

$$\left[P(z)\right]\left[z\left[I\right]-\left[T\right]\right]=z\left[P(0)\right] \qquad (1.72)$$

unde $\left[I\right]$ este matricea unitate de ordinul n^m.

Din relația (1.72) rezultă

$$\left[P(z)\right]=z\left[P(0)\right]\left[z\left[I\right]-\left[T\right]\right]^{-1} \qquad (1.73)$$

Aplicând transformata Z relației (1.66), rezultă

$$\left[P(z)\right]=\left[P(0)\right]\cdot Z\left\{\left[T\right]^n\right\} \qquad (1.74)$$

Comparând relațiile (1.73) şi (1.74), se poate scrie

$$\left[P\left(0\right)\right]Z\left\{\left[T\right]^n\right\}=z\left[P\left(0\right)\right]\left[z\left[I\right]-\left[T\right]\right]^{-1} \qquad (1.75)$$

Deoarece relația (1.75) este adevărată pentru orice $\left[P\left(0\right)\right]\neq\left[0\right]$, rezultă că trebuie să aibă loc relația

$$Z\left\{\left[T\right]^n\right\}=z\left[z\left[I\right]-\left[T\right]\right]^{-1}, \qquad (1.76)$$

de unde rezultă

$$\left[T\right]^n=Z^{-1}\left\{z\left[z\left[I\right]-\left[T\right]\right]^{-1}\right\}, \qquad (1.77)$$

unde $Z^{-1}\left\{.\right\}$ reprezintă transformata Z inversă.

1.6. Entropia surselor cu memorie ergodice

Fie o sursă cu memorie de ordinul m, ce poate furniza n mesaje. Fie starea $S_k \rightarrow s_{k_1}\ s_{k_2}...s_{k_m}$, una din cele n^m stări posibile. Dacă din această stare se furnizează mesajul s_j cu probabilitatea condiționată $p\left(s_j\middle|S_k\right)$, atunci informația ataşată acestui eveniment, notată cu $i\left(s_j \middle| S_k\right)$, se determină cu relația

$$i\left(s_j \middle| S_k\right) = -\log p\left(s_j \middle| S_k\right) \tag{1.78}$$

Deoarece din starea S_k pot fi furnizate oricare din cele n mesaje posibile cu probabilitățile $p\left(s_j \middle| S_k\right), j=\overline{1,n}$, înseamnă că $i\left(s_j \middle| S_k\right)$ determină o variabilă aleatoare discretă, ce poate lua valori cu probabilitățile $p\left(s_j \middle| S_k\right)$. Valoarea medie a acestei variabile aleatoare discrete reprezintă informația medie dată de starea S_k şi va fi notată cu $i\left(S_k\right)$, adică

$$i\left(S_k\right) = \overline{i\left(s_j \middle| S_k\right)} = \sum_{j=1}^{n} p\left(s_j \middle| S_k\right) \log p\left(s_j \middle| S_k\right) \tag{1.79}$$

Înlocuind (1.78) în (1.79), se poate scrie echivalent

$$i\left(S_k\right) = -\sum_{j=1}^{n} p\left(s_j \middle| S_k\right) \log p\left(s_j \middle| S_k\right) \tag{1.80}$$

Mărimea $i\left(S_k\right)$ definită cu relația (1.80) determină o nouă variabilă aleatoare discretă, ce poate lua valori cu probabilitățile p_k, $k=\overline{1,\mathrm{n}^{\mathrm{m}}}$, dacă sursa cu memorie este ergodică. Valoarea medie a acestei variabile aleatoare discrete măsoară din punct de vedere fizic informația medie a sursei cu memorie ergodice, adică entropia acesteia,

$$H = \overline{i\left(S_k\right)} = \sum_{k=1}^{n^m} p_k\, i\left(S_k\right) \tag{1.81}$$

Înlocuind (1.80) în (1.81), rezultă relația finală pentru calculul entropiei surselor cu memorie ergodice

$$H = -\sum_{k=1}^{n^m} \sum_{j=1}^{n} p_k\, p\left(s_j \middle| S_k\right) \log p\left(s_j \middle| S_k\right) \tag{1.82}$$

În relația (1.82) s-ar putea ca o parte din probabilitățile

$p\left(s_j \middle| S_k\right)$ să fie egale cu zero, ceea ce ar corespunde faptului că din starea S_k nu se poate furniza mesajul s_j.

1.7. Matrice asociate surselor cu memorie ergodice

Fie $S = \{S_1, S_2, \cdots S_{n^m}\}$ mulțimea stărilor unei surse cu memorie ergodice, iar $M = \{m_1, m_2, \cdots m_n\}$ mulțimea de mesaje pe care sursa le poate furniza. Se notează cu $p(m_k \mid S_i)$ probabilitatea condiționată să fie furnizat mesajul m_k, dacă sursa se află în starea S_i, iar cu S_j starea în care va trece sursa. Pentru simplificare, se notează cu p_{ij} probabilitatea ca sursa să treacă în starea S_j, dacă starea anterioară era S_i.

Notând cu $p(S_1), p(S_2), \cdots p(S_{n^m})$ probabilitățile stărilor sursei ergodice Markov, matricea distribuției acestora devine

$$[P] = [p(S_1), p(S_2), \cdots p(S_{n^m})] \tag{1.83}$$

Dacă sursa cu memorie este ergodică, atunci probabilitățile stărilor sunt invariante şi se pot calcula cu relațiile (1.69) şi (1.70).

Datorită faptului că dintr-o anumită stare pot fi furnizate cel mult n mesaje, rezultă că cel mult n probabilități p_{ij} din fiecare linie a matricei [T], date de relația (1.57), pot fi diferite de zero. Datorită acestui fapt, se va caracteriza sursa cu memorie de ordin m cu matricea

$$[P(M \mid S)] = \begin{bmatrix} p(m_1 \mid S_1) & p(m_2 \mid S_1) & \cdots & p(m_n \mid S_1) \\ p(m_1 \mid S_2) & p(m_2 \mid S_2) & \cdots & p(m_n \mid S_2) \\ \cdots & \cdots & \cdots & \cdots \\ p(m_1 \mid S_{n^m}) & p(m_2 \mid S_{n^m}) & \cdots & p(m_n \mid S_{n^m}) \end{bmatrix} \tag{1.84}$$

Unele probabilități $p(m_j \mid S_k)$ pot fi zero, ceea ce semnifică faptul că din starea S_k niciodată nu se poate furniza mesajul m_j.

Fie $p(S_k \cap m_j)$ probabilitatea ca sursa să fie în starea S_k şi să fie furnizat mesajul m_j. Matricea

$$[P(S,M)] = \begin{bmatrix} p(S_1 \cap m_1) & p(S_1 \cap m_2) & \cdots & p(S_1 \cap m_n) \\ p(S_2 \cap m_1) & p(S_2 \cap m_2) & \cdots & p(S_2 \cap m_n) \\ \cdots & \cdots & \cdots & \cdots \\ p(S_{n^m} \cap m_1) & p(S_{n^m} \cap m_2) & \cdots & p(S_{n^m} \cap m_n) \end{bmatrix} \quad (1.85)$$

poate fi uşor obținută din [P(M|S)], prin multiplicarea primei linii cu $p(S_1)$, a celei de-a doua cu $p(S_2)$ ş.a.m.d., a ultimei cu $p(S_{n^m})$ şi ținând cont de relația

$$p(S_k \cap m_j) = p(S_k)p(m_j \mid S_k) \quad (1.86)$$

Probabilitatea $p(m_j)$ cu care sursa furnizează mesajul $m_j, j = \overline{1,n}$, poate fi atunci calculată cu relația

$$p(m_j) = \sum_{k=1}^{n^m} p(m_j \cap S_k), (\forall) j = \overline{1,m}. \quad (1.87)$$

Fie $p(S_k \mid m_j)$ probabilitățile ca sursa să fi fost în starea $S_k, k = \overline{1,n^m}$, dacă a furnizat mesajul $m_j, j = \overline{1,n}$. Aceste probabilități pot fi organizate în matricea

$$[P(S \mid M)] = \begin{bmatrix} p(S_1 \mid m_1) & p(S_1 \mid m_2) & \cdots & p(S_1 \mid m_n) \\ p(S_2 \mid m_1) & p(S_2 \mid m_2) & \cdots & p(S_2 \mid m_n) \\ \cdots & \cdots & \cdots & \cdots \\ p(S_{n^m} \mid m_1) & p(S_{n^m} \mid m_2) & \cdots & p(S_{n^m} \mid m_n) \end{bmatrix} \quad (1.88)$$

Această matrice se deduce uşor din matricea [P(S,M)], prin împărțirea primei coloane cu $p(m_1)$, a celei de a doua cu $p(m_2)$ ş.a.m.d., a ultimei cu $p(m_n)$ şi ținând cont de relația

$$p(S_k \cap m_j) = p(m_j)p(S_k \mid m_j) \tag{1.89}$$

1.8. Alte mărimi informaționale care caracterizează sursele cu memorie ergodice

Incertitudinea că sursa se află în starea S_k **şi** furnizează mesajul m_j se poate înlătura cu informația

$$i(S_k \cap m_j) = -\log p(S_k \cap m_j) \tag{1.90}$$

Informația definită în (1.90) determină o variabilă aleatoare discretă care poate lua valori cu probabilitățile $p(S_k \cap m_j);\ k = \overline{1, n^m};\quad j = \overline{1, n}$. Valoarea medie statistică a acestei variabile aleatoare discrete se va denumi *entropia stări-mesaje* şi va fi notată cu $H(S,M)$.

$$H(S,M) = \overline{i(S_k \cap m_j)} = \sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) i(S_k \cap m_j) \tag{1.91}$$

Înlocuind (1.90) în (1.91), se obține

$$H(S,M) = -\sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) \log p(S_k \cap m_j) \tag{1.92}$$

Considerând (1.89), relația (1.92) poate fi scrisă echivalent ca

$$H(S,M) = \sum_{j=1}^{n} p(m_j) \left[-\sum_{k=1}^{n^m} p(S_k \mid m_j) \log p(S_k \cap m_j) \right] \tag{1.93}$$

Mărimea informațională

$$-\sum_{k=1}^{n^m} p(S_k \mid m_j) \log p(S_k \cap m_j) \overset{not.}{=} H(S, m_j) \tag{1.94}$$

este o măsură a incertitudinii medii că sursa furnizează mesajul m_j, ea aflându-se în oricare stare din mulțimea S.

Considerând (1.86), relația (1.92) poate fi, de asemenea, scrisă echivalent ca

$$H(S,M) = \sum_{k=1}^{n^m} p(S_k)\left[-\sum_{j=1}^{n} p(m_j \mid S_k)\log p(S_k \cap m_j)\right] \qquad (1.95)$$

Mărimea informațională

$$-\sum_{j=1}^{n} p(m_j \mid S_k)\log p(S_k \cap m_j) \overset{not.}{=} H(S_k, M) \qquad (1.96)$$

este o măsură a incertitudinii medii ca sursa să se afle în starea S_k şi să furnizeze orice mesaj din mulțimea M.

Incertitudinea ca sursa să se afle în starea S_k, dacă a transmis mesajul m_j, poate fi îndepărtată de informația

$$i(S_k \mid m_j) = -\log p(S_k \mid m_j) \qquad (1.97)$$

Informația definită în (1.97) este o variabilă aleatoare discretă care poate lua valori cu probabilitățile $p(S_k \mid m_j);\ k = \overline{1, n^m};\ j = \overline{1, n}$. Valoarea ei medie, notată cu $H(S \mid m_j)$, este dată de

$$H(S \mid m_j) = \overline{i(S_k \mid m_j)} = \sum_{k=1}^{n^m} p(S_k \mid m_j) i(S_k \mid m_j) \qquad (1.98)$$

sau, considerând (1.97), avem

$$H(S \mid m_j) = -\sum_{k=1}^{n^m} p(S_k \mid m_j)\log p(S_k \mid m_j). \qquad (1.99)$$

Mărimea informațională $H(S \mid m_j)$ este o măsură a incertitudinii medii asupra mulțimii stărilor sursei, dacă aceasta furnizează mesajul m_j.

Considerând toate mesajele $m_j, j = \overline{1, n}$, mărimea definită în (1.99) reprezintă o nouă variabilă aleatoare discretă care poate lua valori cu probabilitățile $p(m_j)$. Valoarea ei medie este

$$H(S|M)=\overline{H(S|m_j)}=\sum_{j=1}^{n}p(m_j)H(S|m_j) \qquad (1.100)$$

Aceasta cantitate se numeşte *echivocaţia* sursei ergodice cu memorie. Prin înlocuirea relaţiei (1.99) în (1.100), şi considerând (1.97), se poate scrie, echivalent

$$H(S|M)=-\sum_{k=1}^{n^m}\sum_{j=1}^{n}p(S_k\cap m_j)\log p(S_k|m_j) \qquad (1.101)$$

Această mărime informaţională măsoară valoarea medie a incertitudinii asupra mulţimii stărilor *S*, dacă mulţimea mesajelor *M* este cunoscută.

Considerând (1.89) şi (1.86), echivocaţia poate fi scrisă în două moduri echivalente

$$H(S|M)=\sum_{j=1}^{n}p(m_j)\left[-\sum_{k=1}^{n^m}p(S_k|m_j)\log p(S_k|m_j)\right] \qquad (1.102)$$

$$H(S|M)=\sum_{k=1}^{n^m}p(S_k)\left[-\sum_{j=1}^{n}p(m_j|S_k)\log p(S_k|m_j)\right] \qquad (1.103)$$

Mărimea informaţională

$$-\sum_{k=1}^{n^m}p(S_k|m_j)\log p(S_k|m_j)\overset{not.}{=}H(S|m_j) \qquad (1.104)$$

este o măsură a incertitudinii medii asupra mulţimii stărilor, după ce a fost transmis mesajul m_j.

Mărimea informaţională

$$-\sum_{j=1}^{n}p(m_j|S_k)\log p(S_k|m_j)\overset{not.}{=}H(S_k|M) \qquad (1.105)$$

este o măsură a incertitudinii medii asupra stării S_k, după ce au fost furnizate mesajele din mulţimea M.

În mod analog, nedeterminarea medie asupra mulţimii mesajelor *M*, dacă este cunoscută mulţimea stărilor, se poate calcula cu relația

$$H(M \mid S) = -\sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k) p(m_j \mid S_k) \log p(m_j \mid S_k) \qquad (1.106)$$

Prin comparația relațiilor (1.82) și (1.106), se observă că entropia sursei ergodice cu memorie în sensul clasic este, de fapt, o entropie condițională numită *eroare medie.*

Considerând (1.86), relația (1.106) poate fi scrisă echivalent ca

$$H(M \mid S) = -\sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) \log p(m_j \mid S_k) \qquad (1.107)$$

Relația (1.107) poate fi scrisă, echivalent, în două moduri

$$H(M \mid S) = \sum_{j=1}^{n} p(m_j) \left[-\sum_{k=1}^{n^m} p(S_k \mid m_j) \log p(m_j \mid S_k) \right] \qquad (1.108)$$

$$H(M \mid S) = \sum_{k=1}^{n^m} p(S_k) \left[-\sum_{j=1}^{n} p(m_j \mid S_k) \log p(m_j \mid S_k) \right] \qquad (1.109)$$

Mărimea informațională

$$-\sum_{k=1}^{n^m} p(S_k \mid m_j) \log p(m_j \mid S_k) \overset{not.}{=} H(m_j \mid S) \qquad (1.110)$$

este o măsură a incertitudinii medii asupra mesajului m_j, $j = \overline{1,n}$, dacă este cunoscută mulțimea stărilor *S*.

Mărimea informațională

$$-\sum_{j=1}^{n} p(m_j \mid S_k) \log p(m_j \mid S_k) \overset{not.}{=} H(M \mid S_k) \qquad (1.111)$$

este o măsură a incertitudinii medii asupra mulțimii mesajelor *M*, dacă starea S_k, $k = \overline{1,n^m}$, este cunoscută.

Fie S_k starea unei surse ergodice cu memorie, a cărei probabilitate este $p(S_k)$. Incertitudinea inițială asupra stării S_k poate fi înlăturată cu o informație $i(S_k)$, calculată cu

$$i(S_k) = -\log p(S_k) \qquad (1.112)$$

Informația $i(S_k)$ definită de (1.112) determină o variabilă aleatoare discretă care poate lua valori cu probabilitatea $p(S_k)$. Valoarea ei medie, notată $H(S)$ este

$$H(S) = \overline{i(S_k)} = \sum_{k=1}^{n^m} p(S_k) i(S_k) = -\sum_{k=1}^{n^m} p(S_k) \log p(S_k) \qquad (1.113)$$

Această entropie reprezintă incertitudinea medie a stărilor sursei.

Deoarece

$$p(S_k) = \sum_{k=1}^{n} p(S_k \cap m_j), \qquad (1.114)$$

se poate scrie echivalent

$$H(S) = -\sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) \log p(S_k). \qquad (1.115)$$

În mod analog, incertitudinea medie asupra mulțimii mesajelor poate fi calculată cu relația

$$H(M) = -\sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) \log p(m_j) \qquad (1.116)$$

Incertitudinea finală asupra stării S_k, după ce s-a transmis mesajul m_j poate fi înlăturată prin informația notată cu $i(S_k \mid m_j)$, care se poate determina cu (1.97).

Dacă $i(S_k)$ este informația ce îndepărtează incertitudinea inițială asupra stării S_k şi $i(S_k \mid m_j)$ este informația ce îndepărtează incertitudinea finală asupra stării S_k, incertitudinea îndepărtată asupra lui S_k se datorează furnizării unei informații, numită *informație mutuală*, notată cu $i(S_k, m_j)$ şi calculată cu

$$i(S_k, m_j) = i(S_k) - i(S_k \mid m_j) \qquad (1.117)$$

Înlocuind (1.112) şi (1.97) în (1.117), se obține

$$i(S_k, m_j) = \log \frac{p(S_k \mid m_j)}{p(S_k)} \tag{1.118}$$

sau, considerând relația (1.89), se poate scrie echivalent

$$i(S_k, m_j) = \log \frac{p(S_k \cap m_j)}{p(S_k)p(m_j)} \tag{1.119}$$

Informația mutuală definită în (1.118) sau (1.119) determină o variabilă aleatoare discretă care poate lua valori cu probabilitatea $p(S_k \cap m_j); k = \overline{1, n^m}; j = \overline{1, n}$. Valoarea ei medie, notată cu I(S,M), este

$$I(S, M) = \overline{i(S_k, m_j)} = \sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) i(S_k, m_j) \tag{1.120}$$

Prin înlocuirea (1.118) sau (1.119) în (1.120), se obțin două relații echivalente

$$I(S, M) = \overline{i(S_k, m_j)} = \sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) \log \frac{p(S_k \mid m_j)}{p(S_k)} \tag{1.121a}$$

$$I(S, M) = \overline{i(S_k, m_j)} = \sum_{k=1}^{n^m} \sum_{j=1}^{n} p(S_k \cap m_j) \log \frac{p(S_k \cap m_j)}{p(S_k)p(m_j)} \tag{1.121b}$$

Considerând (1.86) şi (1.89), relația (1.121b) poate fi scrisă în următoarele două moduri echivalente

$$I(S, M) = \sum_{k=1}^{n^m} p(S_k) \sum_{j=1}^{n} p(m_j \mid S_k) \log \frac{p(S_k \cap m_j)}{p(m_j)p(S_k)} \tag{1.122}$$

$$I(S, M) = \sum_{j=1}^{n} p(m_j) \sum_{k=1}^{n^m} p(S_k \mid m_j) \log \frac{p(S_k \cap m_j)}{p(m_j)p(S_k)} \tag{1.123}$$

Mărimea informațională

$$\sum_{j=1}^{n} p(m_j \mid S_k) \log \frac{p(S_k \cap m_j)}{p(m_j)p(S_k)} \overset{not.}{=} I(S_k, M), \tag{1.124}$$

este o măsură pentru informația medie care se obține când sursa se află în starea S_k, $k=\overline{1,n^m}$, şi a furnizat mesaje din mulţimea M.

Mărimea informaţională

$$\sum_{k=1}^{n^m} p(S_k \mid m_j)\log\frac{p(S_k \cap m_j)}{p(m_j)p(S_k)} \overset{not.}{=} I(S,m_j) \qquad (1.125)$$

este o măsură a informaţiei medii obţinute când sursa furnizează mesajul m_j şi se află în oricare din stările din mulţimea S.

Între mărimile informaţionale definite mai sus, pot fi stabilite următoarele relaţii

$$H(S,M) = H(S) + H(M \mid S) \qquad (1.126a)$$

$$H(S,M) = H(M) + H(S \mid M) \qquad (1.126b)$$

$$I(S,M) = H(S) + H(M) - H(S,M) \qquad (1.127)$$

$$I(S,M) = H(S) - H(S \mid M) \qquad (1.128a)$$

$$I(S,M) = H(M) - H(M \mid S) \qquad (1.128b)$$

Relaţia (1.126a) se obţine plecând de la (1.92) şi ţinând cont de (1.107) şi (1.115). Relaţia (1.126b) se obţine plecând de la (1.92) şi ţinând cont de (1.101) şi (1.116). Relaţia (1.127) se obţine din (1.121b), (1.115), (1.116) şi (1.92). Relaţia (1.128a) se obţine prin substituirea relaţiei (1.126b) în (1.127), iar relaţia (1.128b), prin înlocuirea lui (1.126a) în (1.127).

Considerând relaţiile anterioare, se pot demonstra uşor următoare legături informaţionale:

$$\begin{aligned} H(S_k,M) &= H(M \mid S_k) - \log p(S_k) = \\ &= H(S_k \mid M) - \sum_{j=1}^{n} p(m_j \mid S_k)\log p(m_j) \end{aligned} \qquad (1.129)$$

$$\begin{aligned} H(S,m_j) &= H(m_j \mid S) - \sum_{k=1}^{n^m} p(S_k \mid m_j)\log p(S_k) = \\ &= H(S \mid m_j) - \log p(m_j) \end{aligned} \qquad (1.130)$$

$$I(S_k, M) = -\log p(S_k) - \sum_{j=1}^{n} p(m_j \mid S_k) \log p(m_j) - H(S_k, M) =$$

$$-\log p(S_k) - H(S_k \mid M) = -\sum_{j=1}^{n} p(m_j \mid S_k) \log p(m_j) - H(M \mid S_k)$$

(1.131)

$$I(S, m_j) = -\sum_{k=1}^{n^m} p(S_k \mid m_j) \log p(S_k) - \log p(m_j) - H(S, m_j) =$$

$$-\sum_{k=1}^{n^m} p(S_k \mid m_j) \log p(S_k) - H(S \mid m_j) = -\log p(m_j) - H(m_j \mid S)$$

(1.132)

1.9. Textele privite ca surse cu memorie

O primă aproximare a unui text scris ar consta în considerarea că toate literele ce apar în text sunt egal probabile şi acestea apar independent în cuvinte. Un astfel de model este cunoscut în literatura de specialitate ca model de ordinul zero. Dacă în text sunt folosite *m* caractere (litere, spaţiu între cuvinte, semne de punctuaţie etc.), atunci entropia acestui model va fi

$$H_0 = \log m$$

Prima observaţie care se poate face asupra unui text scris este aceea că frecvenţa de apariţie a literelor nu este aceeaşi pentru toate literele. Astfel, în limba română contemporană, literele folosite în diferite texte au frecvenţe de apariţie între 0,1 şi 0,001.

Dacă se consideră că cele *m* caractere folosite în textul scris au probabilităţile $p_i,\ i = \overline{1, m}$, şi toate caracterele sunt independente unul de altul, atunci se ajunge la modelul de ordinul întâi, caz în care entropia devine

$$H_1 = -\sum_{i=1}^{m} p_i \log p_i \qquad (1.133)$$

Cele două modele anterioare au presupus existența unei independențe statistice de la un caracter la altul. Acest lucru nu reflectă cu acuratețe natura unui text scris, deoarece, de exemplu, dacă într-o propoziție unele caractere lipsesc, suntem capabili să ne dăm seama care sunt acestea. Aceasta implică existența unei dependențe între caractere. În mod natural, caracterele care sunt mai apropiate unele de altele, sunt mai dependente între ele decât acele caractere care sunt mai îndepărtate unele de altele. Dacă se consideră cazul în care un caracter curent depinde numai de caracterul anterior, se obține modelul de ordinul doi. Notând cu $p_{j|i}$ probabilitatea condiționată ca litera curentă să fie al *j*-lea caracter din alfabet cu condiția ca precedentul caracter să fie a *i*-a literă, entropia acestui model se calculează cu relația

$$H_2 = -\sum_{i=1}^{m} p_i \sum_{j=1}^{m} p_{j|i} \log p_{j|i} \qquad (1.134)$$

În cazul modelului de ordinul trei, fie $p_{k|j,i}$ probabilitatea condiționată ca litera curentă să fie al *k*-lea caracter din alfabet, cu condiția ca precedentul caracter să fie a *j*-a literă, iar cea anterioară să fie a *i*-a literă. Entropia în acest caz devine

$$H_3 = -\sum_{i=1}^{m} p_i \sum_{j=1}^{m} p_{j|i} \sum_{k=1}^{m} p_{k|j,i} \log p_{k|j,i} \qquad (1.135)$$

În mod analog, se pot defini modelele de ordin superior.

Prin definiție, folosirea unui prisos de litere (cuvinte, simboluri etc.) pentru transmiterea unei anumite cantități de informație poartă denumirea de *redundanță*. Se poate, deci, conchide că textele din orice limbă prezintă o redundanță care depinde de natura textului. Astfel, textele folosite în științele exacte (matematică, fizică etc.) prezintă o redundanţă mai mică în comparaţie cu textele folosite în literatura beletristică.

Redundanţa crescută poate asigura protecţia informaţiei la perturbaţii. Intuitiv, acest lucru poate fi uşor înţeles, citind texte din domenii diferite. Astfel, la acelaşi grad de oboseală (adică la acelaşi nivel al perturbaţiilor care atenuează atenţia) se va înţelege mai bine şi mai uşor un text de literatură beletristică, (un roman) decât un text de logică matematică.