Άσκηση 8 (φυλλάδιο 3η)

Μία πηγή πληροφορίας παράγει δύο σύμβολα, το 0 και το 1. Η ακολουθία των παραγομένων συμβόλων αποτελεί μία εργοδική αλυσίδα Markov δεύτερης τάξης με πιθανότητες μετάβασης:

$p(0|00) = 0.8$, $p(0|01) = 0.5$

$p(0|10) = 0.5$, $p(0|11) = 0.2$

(α) Πόσο μεγάλο είναι το ποσό πληροφορίας μίας τριάδας που παράγεται απ'αυτήν την πηγή πληροφορίας; Χρησιμοποιήστε ότι $p(00) = p(11) = 5/14$, $p(01) = p(10) = 1/7$. Χρησιμοποιώντας το αποτέλεσμα αυτό, προσδιορίστε το ποσό πληροφορίας για κάθε σύμβολο, που συμβολίζεται με $H_3(U)$.

(β) Πόσο μεγάλο είναι το ποσό πληροφορίας μιας δυάδας; Απ' αυτό προσδιορίστε το $H_2(U)$.

(γ) Πόσο μεγάλο είναι το $H_1(U)$;

(δ) Το δεσμευμένο ποσό πληροφορίας στο ενδεχόμενο ότι N–1 προηγούμενα σύμβολα είναι γνωστά συμβολίζεται με $F_N(U)$. Προσδιορίστε τα $F_1(U)$, $F_2(U)$, $F_3(U)$.

(ε) Εξηγείστε γιατί $F_3(U) < F_2(U) < F_1(U)$.

(στ) Τι μπορείτε να πείτε σχετικά με τις τιμές των $F_4(U)$ και $H_4(U)$;

(ζ) Σχεδιάστε τα $F_N(U)$ και $H_N(U)$ σαν μία συνάρτηση του αριθμού των συμβόλων N.

▶

(α) Είναι $m^{k+1} = 2^3 = 8$ τριάδες – καταστάσεις με στοιχεία {0,1}, ενώ $m^{k+1} - m^k = 4$.

Για το ποσό πληροφορίας μιας τριάδας ισχύει ότι:

$$\boxed{H(V) = H(U_1, U_2, U_3) = -\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} p(u_i, u_j, u_k) \cdot \log p(u_i, u_j, u_k)}$$

Έτσι υπολογίζουμε τα παρακάτω:

$$p(000) = p(00) \cdot p(0\,|\,00) = \frac{5}{14} \cdot \frac{8}{10} = \frac{2}{7}$$

$$p(001) = p(00) \cdot p(0\,|\,01) = \frac{5}{14} \cdot \frac{5}{10} = \frac{5}{28}$$

$$p(010) = p(01) \cdot p(0\,|\,10) = \frac{1}{7} \cdot \frac{5}{10} = \frac{1}{14}$$

$$p(011) = p(01) \cdot p(0\,|\,11) = \frac{1}{7} \cdot \frac{2}{10} = \frac{1}{35}$$

$$p(100) = p(10) \cdot p(1\,|\,00) = \frac{1}{7} \cdot \frac{2}{10} = \frac{1}{35}$$

$$p(101) = p(10) \cdot p(1\,|\,01) = \frac{1}{7} \cdot \frac{5}{10} = \frac{1}{12}$$

$$p(110) = p(00) \cdot p(0\,|\,00) = \frac{5}{14} \cdot \frac{8}{10} = \frac{2}{7}$$

$$p(111) = p(11) \cdot p(1\,|\,11) = \frac{5}{14} \cdot \frac{8}{10} = \frac{2}{7}$$

Επομένως έχουμε:

$H(V)=H(U_1,U_2,U_3)= -\left(\frac{4}{7}\log\frac{2}{7} + \frac{10}{28}\log\frac{5}{28} + \frac{2}{14}\log\frac{1}{14} + \frac{2}{35}\log\frac{1}{35}\right) = 2.67$

Για κάθε σύμβολο, θα έχουμε $H_3(U)= \frac{H(V)}{3} = \frac{2.67}{3} = 0.89$ bits/σύμβολο.

(β) Ανάλογα, για το ποσό πληροφορίας μιας δυάδας έχουμε:

$$\boxed{H(U_1, U_2) = -\sum_{i=1}^{2}\sum_{j=1}^{2} p(u_i, u_j) \cdot \log p(u_i, u_j)}$$

$$= -\left(2 \cdot \frac{5}{14}\log\frac{5}{14} + 2 \cdot \frac{1}{7}\log\frac{1}{7}\right) = 1.86$$

Άρα, $H_2(U) = \frac{H(U_1, U_2)}{2} = \frac{1.86}{2} = 0.93$ bits/σύμβολο.

(γ) Αντίστοιχα για τον υπολογισμό του $H_1(U)$ έχουμε :

$$p(0) = p(0|00)\cdot p(00)+p(0|01)\cdot p(01)+p(0|10)\cdot p(10)+p(0|11)\cdot p(11)$$

$$= \frac{8}{10}\cdot\frac{5}{14}+\frac{5}{10}\cdot\frac{1}{7}+\frac{5}{10}\cdot\frac{1}{7}+\frac{2}{10}\cdot\frac{5}{14} = \frac{1}{2}$$

$$p(1) = p(1|00)\cdot p(00)+p(1|01)\cdot p(01)+p(1|10)\cdot p(10)+p(1|11)\cdot p(11)$$

$$= \frac{2}{10}\cdot\frac{5}{14}+\frac{5}{10}\cdot\frac{1}{7}+\frac{5}{10}\cdot\frac{1}{7}+\frac{8}{10}\cdot\frac{5}{14} = \frac{1}{2}$$

Έτσι βρίσκουμε ότι $H_1(U) = -\frac{1}{2}\log\frac{1}{2}-\frac{1}{2}\log\frac{1}{2}=1$.

(δ) Από τη σχέση $\boxed{N.H_N(U) = (N-1).H_{N-1}(U) + F_N(U)}$, λύνοντας ως προς $F_N(U)$ για N=1,2,3 έχουμε :

$F_1(U) = H_1(U) = 1$ bit

$F_2(U) = 2H_2(U) - H_1(U) = 0.86$

$F_3(U) = 3H_3(U) - 2H_2(U) = 0.81$

(ε) Είναι $F_3(U) < F_2(U) < F_1(U)$, διότι το ποσόν της δεσμευμένης πληροφορίας που παρέχεται από το πρώτο σύμβολο N=1, δεν μπορεί να μας οδηγήσει σε μία αύξηση της αβεβαιότητας σχετικά με το N=2 ή N=3, αλλά την μειώνει, αφού η πηγή είναι στατική και τα ποσά της πληροφορίας είναι ανεξάρτητα της θέσης του N.

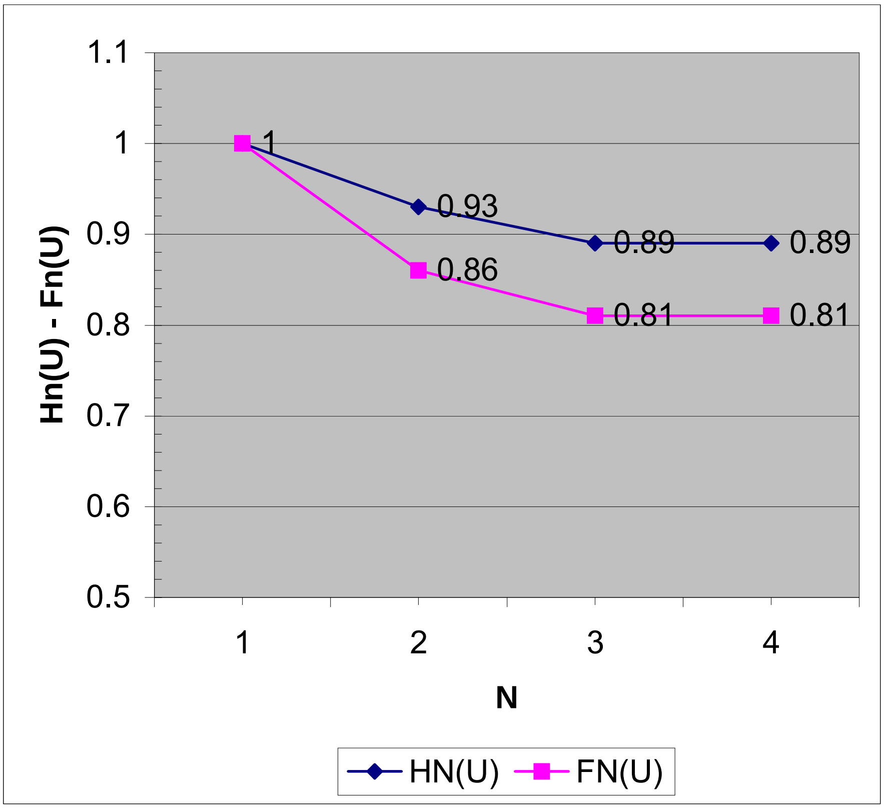
Διαισθητικά δηλαδή επιβεβαιώνεται ότι το δεσμευμένο ποσό πληροφορίας $F_N(U)$ είναι μια φθίνουσα συνάρτηση του N (θεώρημα).

(στ) Ισχύει γενικά ότι για $N \geq k+1$, η ποσότητα $F_N(U)$ παραμένει ίση με την $F_{k+1}(U)$. Εδώ έχουμε k=2, επομένως $F_4(U) = F_3(U) = 0.81$

Ομοίως προκύπτει ότι : $H_4(U) = H_3(U) = 0.89$.

(ζ) Για N=1, έχουμε $F_1=1$ και $H_1=1$
Για N=2, έχουμε $F_2=0.86$ και $H_2=0.93$
Για N=3, έχουμε $F_3=0.81$ και $H_3=0.89$

Η γραφική παράσταση των $F_N(U)$ και $H_N(U)$ ως συνάρτηση του αριθμού των συμβόλων N είναι :



■