Τμήμα Μηχανικών Πληροφορικής τ.ε.
Τεχνολογικό Εκπαιδευτικό Ίδρυμα Δυτικής Ελλάδας

Τηλεπικοινωνιακά Συστήματα ΙΙ

Διάλεξη 10: Βασικές αρχές Θεωρίας Πληροφοριών

Δρ. Μιχάλης Παρασκευάς
Επίκουρος Καθηγητής

Ατζέντα

1. Ορισμός του μέτρου της πληροφορίας κατά Shannon
2. Ιδιότητες της Εντροπίας
3. Συνδυασμένη και Αμοιβαία Πληροφορία
4. Υπό συνθήκη ποσότητα πληροφορίας
5. Αμοιβαία ποσότητα πληροφορίας
6. Διακριτή πηγή πληροφορίας χωρίς μνήμη

Ποσότητα Πληροφορίας

Η ποσότητα της πληροφορίας ορίζεται ως ο δεκαδικός λογάριθμος του πλήθους των διαφορετικών λέξεων που μπορούν να σχηματιστούν από ένα δεδομένο πλήθος συμβόλων. Αν θεωρήσουμε ότι υπάρχουν μηνύματα μήκους k συμβόλων, από ένα αλφάβητο με N σύμβολα, η ***ποσότητα πληροφορίας*** είναι ίση με:

$$H(N^k) = \log(N^k) = k \log N$$

Αν η βάση του λογάριθμου είναι το 2, τότε η μονάδα μέτρησης της πληροφορίας ονομάζεται bit.

<u>Παράδειγμα</u>: Για σχηματισμό μηνυμάτων μήκους ενός συμβόλου από ένα αλφάβητο αποτελούμενο από δύο σύμβολα, η ποσότητα πληροφορίας ισούται με:

$$H(N^1) = log(N) = log_2 2 = 1$$

Εντροπία

Μέση Ποσότητα Πληροφορίας ή Μέσο Πληροφοριακό Περιεχόμενο ή Εντροπία

Αν Y είναι μια διακριτή τυχαία μεταβλητή με δειγματοχώρο $Y = \{y_1, y_2, \dots\ y_n\}$ και συνάρτηση πιθανότητα μάζας $p(y_i)$, τότε η μέση ποσότητα πληροφορίας $H(Y)$ ισούται με:

$$H(Y) = -\sum_{i=1}^{n} p(y_i)\ log\ p(y_i)$$

Η μέση ποσότητα πληροφορίας ονομάζεται **εντροπία.**

<u>Παράδειγμα</u>: Να υπολογισθεί η μέση ποσότητα πληροφορίας από την ρίψη δύο ζαριών. Να υποθέσετε ότι αποτέλεσμα (4,2) και (2,4) είναι διαφορετικά.

Η μέση ποσότητα πληροφορίας της ρίψης δύο ζαριών είναι η ακόλουθη:

$$H(Y) = -\sum_{i=1}^{n} p(y_i)\ log\ p(y_i) = -36\frac{1}{36} log\frac{1}{36} = log36 = 5{,}17\ bots$$

Ιδιότητες της Εντροπίας

1. Η μέση ποσότητα πληροφορίας είναι συνεχής στο p.

2. Η μέση πληροφορία Η(Υ) είναι συμμετρική, δηλαδή διάταξη των πιθανοτήτων δεν την επηρεάζει.

3. Η εντροπία Η(Υ) παίρνει την μέγιστη τιμή όταν όλα τα ενδεχόμενα είναι ισοπίθανα.

4. Η εντροπία είναι προσθετική. Η ιδιότητα αυτή αναφέρεται στην περίπτωση κατά την οποία δύο ανεξάρτητες τυχαίες μεταβλητές Χ και Υ συνδυάζονται.

Συνδυασμένη Ποσότητα Πληροφορίας

Ένα τυχαίο πείραμα (Χ,Υ) έχει ως δυνατά αποτελέσματα όλους τους δυνατούς συνδυασμούς των αποτελεσμάτων $X \ = \{x_1, x_2, \dots, x_n\}$ και $Y = \{y_1, y_2, \dots, y_n\}$.

Η κατανομή πιθανοτήτων δίνεται από την παρακάτω σχέση:

$$P = \{p(x_1, y_1), p(x_1, y_2), \dots, p(x_1, y_m), \dots p(x_n, y_1), p(x_n, y_2), \dots p(x_n, y_m)\}$$

Αν (Χ, Υ) είναι ένα τυχαίο πείραμα με δισδιάστατο δειγματοχώρο και κατανομή πιθανοτήτων P, τότε η **συνδυασμένη ποσότητα πληροφορίας** $\mathbf{H}(\mathbf{X}, \mathbf{Y})$ ορίζεται ως εξής:

$$H(\mathrm{X}, Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i) \ log \ p(x_i, y_i)$$

Υπό Συνθήκη Ποσότητα Πληροφορίας

Αρκετές φορές μας ενδιαφέρει να υπολογίσουμε την ποσότητα πληροφορίας μιας τυχαίας μεταβλητής Χ, όταν δίνεται το αποτέλεσμα μιας άλλης τυχαίας μεταβλητής Υ.

Αυτό ονομάζεται **υπό συνθήκη** ποσότητα πληροφορίας της Χ ως προς το Υ.

Η μέση υπό συνθήκη ποσότητα πληροφορίας του τυχαίου πειράματος Χ, με δεδομένο το αποτέλεσμα του πειράματος Υ δίνεται από την σχέση:

$$H(\mathrm{X}/Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i)\ log\ p(x_i/y_i)$$

Πρόταση: Για δύο τυχαίες μεταβλητές Χ και Υ ισχύει η παρακάτω σχέση:

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y)$$

Αμοιβαία Ποσότητα Πληροφορίας

Η **αμοιβαία ποσότητα πληροφορίας** είναι ένα μέτρο της ποσότητας πληροφορίας που μια τυχαία μεταβλητή περιέχει για μια άλλη τυχαία μεταβλητή ή ένα μέτρο εξάρτησης μεταξύ των δύο και δίνεται από τη σχέση:

$$I(X;Y) = H(X) - H(Y/X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i)\ log\frac{p(x_i, y_i)}{p(x_i)\, p(y_i)}$$

Άσκηση 1

Δίνονται οι τυχαίες μεταβλητές X και Y, με δύο δυνατά αποτελέσματα η κάθε μία. Οι δισδιάστατες πιθανότητες είναι:

$$p(x_1, y_1) = \frac{1}{8}, \qquad p(x_1, y_2) = \frac{1}{8}, \qquad p(x_2, y_1) = \frac{1}{2}, \qquad p(x_2, y_2) = \frac{1}{4}$$

1. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται το αποτέλεσμα της τυχαίας μεταβλητής X; Ποιά όταν μας γνωστοποιείται το αποτέλεσμα της Y;
2. Ποια είναι η ποσότητα πληροφορίας που λαμβάνουμε όταν γνωρίζουμε το αποτέλεσμα της συνδυασμένης διαδικασίας (X,Y);
3. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται το αποτέλεσμα της Y αν γνωρίζουμε το X;

<u>Απάντηση</u>:

Αρχικά θα υπολογίσουμε τις πιθανότητες $p(x_1)$, $p(x_2)$, $p(x_3)$, $p(x_4)$. Είναι:

$$p(x_1) = \sum_{j=1}^{2} p(x_1, y_j) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4} \qquad p(x_2) = \sum_{j=1}^{2} p(x_2\ y_j) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$p(y_1) = \sum_{i=1}^{2} p(x_i, y_1) = \frac{1}{8} + \frac{1}{2} = \frac{5}{8} \qquad p(y_2) = \sum_{i=1}^{2} p(x_i, y_2) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$$

Άσκηση 1 (συνέχεια)

Ερώτημα 1:

$$H(X) = -\sum_{i=1}^{2} p(x_i)\ \ log\ \ p(x_i) = -\left[\frac{1}{4}\log\left(\frac{1}{4}\right) + \frac{3}{4}\log\left(\frac{3}{4}\right)\right] = 0{,}813\ bits$$

$$H(Y) = -\sum_{i=1}^{2} p(y_i)\ \ log\ \ p(y_i) = -\left[\frac{5}{8}\log\left(\frac{5}{8}\right) + \frac{3}{8}\log\left(\frac{3}{8}\right)\right] = 0{,}956\ bits$$

Ερώτημα 2:

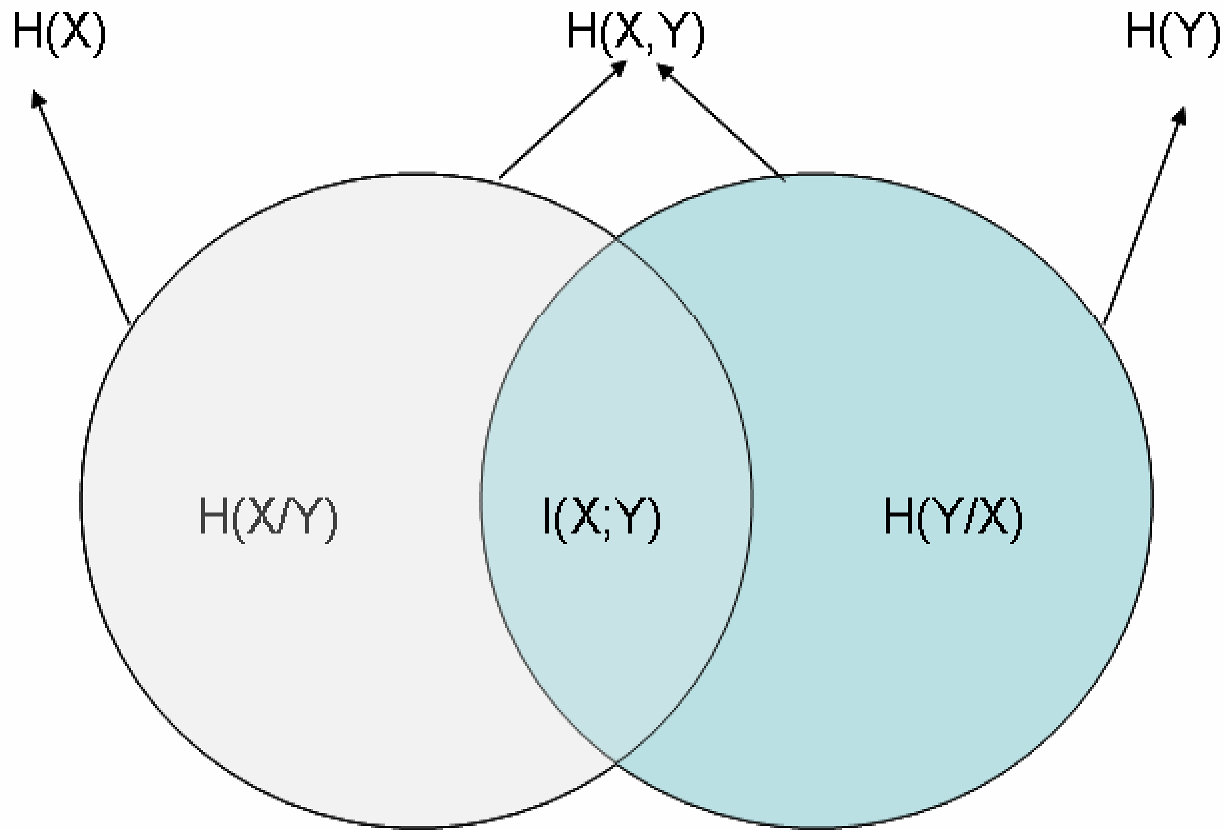
$$H(\mathrm{X}, Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i)\ \ log\ \ p(x_i, y_i) = \cdots = 1{,}75\ \ bits$$

Ερώτημα 3:

$$H(X,Y) = \ H(X\ ) + \ H(Y/X) \Rightarrow H(Y/X) = H(X,Y) - H(X) = 1{,}75\ \ - 0{,}813 = 0{,}935\ bits$$

Ερώτημα 4:

$$H(X/Y) = H(X,Y) - H(Y) = 1{,}75\ \ - 0{,}95 = 0{,}8\ bits$$


H(X)
H(X,Y)
H(Y)
H(X/Y)
I(X;Y)
H(Y/X)

Άσκηση 2

Θεωρούμε ένα ψηφιακό επικοινωνιακό κανάλι που χρησιμοποιεί ως εισόδους και εξόδους τα σύμβολα ‘0’ και ‘1’. Δίνονται οι τυχαίες οι πιθανότητες $p(x_1) = 1/2, \quad p(y_1/x_1) = 3/4$, $p(y_1/x_2) = 1/4$, όπου τα x_1, y_1 αφέρονται στο σύμβολο ‘0’ και τα x_2, y_2 στα σύμβολα ‘1’.

1. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται ποιο σύμβολο έχει ληφθεί στην έξοδο όταν γνωρίζουμε ότι στην είσοδο έχει σταλεί το ‘1’.
2. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται ποιο σύμβολο έχει ληφθεί στην έξοδο όταν γνωρίζουμε ποιο σύμβολο έχει σταλεί στην είσοδο.
3. Ποια ποσότητα πληροφορίας λαμβάνουμε αν μας γνωστοποιείται ποιο σύμβολο έχει σταλεί στην είσοδο και ποιο σύμβολο έχει ληφθεί στην έξοδο
4. Ποια ποσότητα πληροφορίας λαμβάνουμε αν μας γνωστοποιείται ποιο σύμβολο έχει σταλεί στην είσοδο όταν γνωρίζουμε ποιο σύμβολο έχει ληφθεί στην έξοδο.

<u>Απάντηση</u>: Υπολογίζουμε τις συνδυασμένες πιθανότητες:

$$p(y_2/x_1) = \frac{1}{4}, \qquad p(y_2/x_2) = \frac{3}{4}$$

$$p(x_1, y_2) = p(y_2/x_1)\, p(x_1) = \frac{1}{8}, \qquad p(x_2, y_1) = p(y_1/x_2)\, p(x_2) = \frac{1}{8}$$

$$p(x_1, y_1) = p(y_1/x_1)\, p(x_1) = \frac{3}{8}, \qquad p(x_2, y_2) = p(y_2/x_2)\, p(x_2) = \frac{3}{8}$$

Άσκηση 2 (συνέχεια)

Ερώτηση 1:

$$H(Y/x_2) = -\sum_{k=1}^{2} p(y_k/x_2)\ log\, p(y_k/x_2) = 0{,}81\ bits$$

Ερώτηση 2:

$$H(Y/X) = -\sum_{i=1}^{2}\sum_{k=1}^{2} p\left(y_i, x_j\right)\ log\, p\left(y_j/x_i\right) = 0{,}91\ bits$$

Ερώτηση 3:

$$H(X,Y)\ =\ H(X\ )\ +\ H(Y/X)\ = 1\ + 0{,}91 = 1{,}91$$

Ερώτηση 4:

$$H(X/Y) =\ H(X,Y) - H(Y) = 1{,}91 - 0{,}96 = 0{,}95\ bits$$

Διακριτή πηγή πληροφορίας χωρίς μνήμη

- **Διακριτή** ονομάζεται μια **πηγή πληροφορίας** που παράγει **ακολουθίες** συμβόλων.
- Το σύνολο των συμβόλων ονομάζεται **αλφάβητο πηγής**.
- Μια ομάδα διαδοχικών συμβόλων ονομάζεται **μήνυμα ή λέξη**.

Αν η πιθανότητα επιλογής ενός συμβόλου είναι **σταθερή** και **ανεξάρτητη** από τις επιλογές των προηγούμενων συμβόλων, τότε η πηγή ονομάζεται **διακριτή πηγή πληροφορίας χωρίς μνήμη**.

Έστω $s_1, s_2, \dots, s_n$ το πλήθος των n συμβόλων του αλφάβητου S. Το πλήθος των μηνυμάτων q συμβολίζονται με $m_1, m_2, \dots, m_q$. Αν κάθε μήνυμα αποτελείται από l σύμβολα, το πλήθος των δυνατών μηνυμάτων είναι ίσο με n^l .

Η **μέση ποσότητα πληροφορίας** (ή εντροπία) των συμβόλων που δημιουργούνται από μια διακριτή πηγή χωρίς μνήμη, με αλφάβητο $S = \{s_1, s_2, \dots s_n\}$ όπου n το πλήθος των συμβόλων του αλφαβήτου και p_i η πιθανότητα επιλογής του συμβόλου s_i, είναι :

$$H(S) = -\sum_{i=1}^{n} p_i \ \log p_i \ (bits/symbol)$$

Διακριτή πηγή πληροφορίας χωρίς μνήμη

Η **μέγιστη εντροπία** των συμβόλων της πηγής επιτυγχάνεται όταν οι πιθανότητες επιλογής τους είναι ίσες με :

$$maxH(S) = -\sum_{i=1}^{n} \frac{1}{n}\ log\frac{1}{n} = \log\ n\ \ (bits/symbol\,)$$

Ο **πλεονασμός** της διακριτής πηγής ορίζεται από τη σχέση:

$$red = 1 - \frac{H(S)}{maxH(S)}$$

Το **μέσο πληροφοριακό περιεχόμενο των μηνυμάτων της πηγής** μπορεί να οριστεί με ανάλογο τρόπο. $M\ =\ \{m_1, m_2, \dots, m_q\}$ είναι το σύνολο των δυνατών μηνυμάτων, q το πλήθος των δυνατών μηνυμάτων ($q\ = n^l$) και $P\ =\ \{p(m_1), p(m_2), .. p(m_q)\}$, η κατανομή πιθανοτήτων:

$$H(M) = -\sum_{i=1}^{q} p(m_i)\ \ log\, p(m_i)\ (bits/symbol\,)$$

Επίσης ισχύει η σχέση $H(M) = qH(S)$

Άσκηση 3

Μια πηγή πληροφορίας παράγει σύμβολα, τα οποία ανήκουν στο αλφάβητο $S = \{\alpha, \beta, \gamma, \delta, \varepsilon, \zeta, \eta\}$. Οι πιθανότητες των συμβόλων αυτών είναι 1/32, 1/16, 1/8, 1/8, 1/8, 1/2 και 1/32, αντίστοιχα. Θεωρώντας την πηγή χωρίς μνήμη, ζητείται να προσδιορίσετε ή να υπολογίσετε:

1. Το σύμβολο της πηγής με το πιο χαμηλό πληροφορικό περιεχόμενο.
2. Τα σύμβολα της πηγής με το πιο υψηλό πληροφορικό περιεχόμενο.
3. Το μέσο πληροφορικό περιεχόμενο των συμβόλων της πηγής,
4. Το μέσο πληροφορικό περιεχόμενο των μηνυμάτων της πηγής αποτελούμενων από δύο σύμβολα.
5. Τον πλεονασμό της πηγής
6. Το μέσο ρυθμό πληροφορίας της πηγής για ρυθμό 12500 συμβόλων/sec

Απάντηση:

1. Το πληροφορικό περιεχόμενο ενός συμβόλου ‘x’ δίνεται από τον αρνητικό λογάριθμο της πιθανότητας παραγωγής του. Επομένως, το σύμβολο με την πιο υψηλή πιθανότητα παραγωγής έχει το πιο χαμηλό πληροφορικό περιεχόμενο. Στην προκειμένη περίπτωση, για το σύμβολο ‘ζ’ έχουμε $H(\zeta) = -\log(1/2) = 1\ bits$

Άσκηση 3 (συνέχεια)

2. Το σύμβολο με την πιο μικρή πιθανότητα παραγωγής έχει το πιο υψηλό πληροφορικό περιεχόμενο. Στην προκειμένη περίπτωση, τα σύμβολα ‘α’ και ‘η’ έχουν την πιο χαμηλή πιθανότητα παραγωγής, η οποία είναι ίση με 1/32, δηλαδή:

$$H(\alpha) = \mathrm{H}(\eta) = -\log(1/32) = 5\ bits.$$

3. Το μέσο πληροφορικό περιεχόμενο των συμβόλων της πηγής υπολογίζεται με τη βοήθεια της ακόλουθης σχέσης

$$H(S) = -\sum_{i=1}^{n} p_i\ log\, p_i =$$

$$= -\frac{1}{32}\log\frac{1}{32} - \frac{1}{16}\log\frac{1}{16} - \frac{1}{8}\log\frac{1}{8} - \frac{1}{8}\log\frac{1}{8} - \frac{1}{2}\log\frac{1}{2} - \frac{1}{32}\log\frac{1}{32} =$$

$$= \frac{35}{16} = 2{,}1875\ (bits/symbol)$$

4. Για τον υπολογισμό του μέσου πληροφορικού περιεχομένου των μηνυμάτων της πηγής αποτελούμενων από 2 σύμβολα, αφού η πηγή είναι χωρίς μνήμη, αρκεί να πολλαπλασιά-σουμε τη μέση ποσότητα πληροφορίας συμβόλων με το πλήθος των συμβόλων από τα οποία αποτελούνται τα μηνύματα. Δηλαδή, $H(M) = 2\ \mathrm{H}(S) = 4{,}375\ bits$.

Άσκηση 3 (συνέχεια)

5. $red = 1 - \frac{H(S)}{maxH(S)} = 1 - \frac{H(S)}{\log 7} = 1 - \left(\frac{2{,}1875}{2{,}8}\right) = 1 - 0{,}781 = 0{,}2186$

6. $R = rH(S) = 12500\ x\ (2{,}1875) = 27.343{,}75\ bits/\sec$