Σημειώσεις Θεωρίας Πληροφορίας

Η αξιολόγηση της απόδοσης ενός τηλεπικοινωνιακού συστήματος βασίζεται σε μεγάλο μέρος σε δυο παραμέτρους: Στην ευχέρεια με την οποία μπορούμε να αναπαραστήσουμε την πληροφορία που παράγει μια πηγή δεδομένων και στο ρυθμό με τον οποίο μπορούμε να μεταδώσουμε αξιόπιστα πληροφορία μέσα από έναν ενθόρυβο δίαυλο. Η *θεωρία της πληροφορίας* ασχολείται με τη μαθηματική μοντελοποίηση και την ανάλυση της απόδοσης ενός τηλεπικοινωνιακού συστήματος, προσπαθεί να ποσοτικοποιήσει τις παραμέτρους που προαναφέρθηκαν. Πιο συγκεκριμένα, για μια δοσμένη *πηγή πληροφορίας* και ένα δοσμένο *ενθόρυβο δίαυλο* η θεωρία της πληροφορίας παρέχει φράγματα

(α) Στον ελάχιστο αριθμό δυαδικών ψηφίων που απαιτούνται ώστε να αναπαραστήσουμε με ακρίβεια κάθε σύμβολο μιας πηγής πληροφορίας

(β) Στο μέγιστο ρυθμό με τον οποίο μπορούμε να μεταδώσουμε αξιόπιστα δεδομένα μέσα από τον ενθόρυβο δίαυλο.

1 Η έννοια της πληροφορίας

Ας υποθέσουμε πως παρατηρούμε ένα *πείραμα τύχης* στο οποίο μια πηγή διακριτών δεδομένων παράγει ένα σύμβολο στην έξοδό της. Η έξοδος της πηγής μπορεί να μοντελοποιηθεί ως μια (διακριτή) τυχαία μεταβλητή S η οποία λαμβάνει τιμές από το πεπερασμένο σύνολο

$$\Phi = \{s_0, s_1, \ldots, s_{K-1}\} \tag{1}$$

το οποίο θα ονομάζουμε *αλφάβητο* της πηγής. Έστω ακόμα πως η τυχαία μεταβλητή S λαμβάνει κάθε τιμή s_i με πιθανότητα:

$$Pr\,\{S = s_i\} = p_i, \qquad i = 0, 1, \ldots, K-1 \tag{2}$$

όπου φυσικά ισχύει η ιδιότητα

$$\sum_{i=0}^{K-1} p_i = 1$$

Ας υποθέσουμε τώρα πως πραγματοποιούμε το παραπάνω τυχαίο πείραμα πολλές φορές ανά χρονικά διαστήματα T_s. Το χρόνο αυτό τον ονομάζουμε *περίοδο* *συμβόλου.* Υποθέτουμε ακόμα πως το αποτέλεσμα κάθε τυχαίου πειράματος δεν εξαρτάται από τα αποτελέσματα των πειραμάτων που έγιναν στο παρελθόν. Αν ισχύει η ιδιότητα αυτή, τότε θα λέμε πως τα αποτελέσματα των πειραμάτων (ή απλά τα σύμβολα που εκπέμπονται από την πηγή) είναι *στατιστικά ανεξάρτητα.* Ακόμα, μια πηγή η οποία παράγει στατιστικά ανεξάρτητα σύμβολα θα ονομάζεται *διακριτή πηγή χωρίς μνήμη.*

Για να κατανοήσουμε τώρα την έννοια της "πληροφορίας" ας σκεφτούμε το εξής: Έστω πως στο πείραμα που περιγράψαμε η πιθανότητα εμφάνισης του συμβόλου s_0 είναι 1 και επομένως οι πιθανότητες όλων των άλλων συμβόλων είναι ίσες με 0. Για την πηγή αυτή, γνωρίζουμε εκ των προτέρων το αποτέλεσμα του πειράματος τύχης επομένως η "πληροφορία" πως εμφανίστηκε το σύμβολο s_0 είναι μηδενική. Όμοια, ας θεωρήσουμε τώρα μια άλλη περίπτωση στην οποία δύο γεγονότα $A = \{S = s_i\}$ και $B = \{S = s_j\}$ έχουν πιθανότητες p_i και p_j και η πιθανότητα p_i είναι πολύ μικρότερη της πιθανότητας p_j. Τότε, αν συμβεί το γεγονός B που είναι και το πιο πιθανό η "πληροφορία" που λαμβάνουμε είναι μικρή. Αντίθετα, αν συμβεί το γεγονός A η πληροφορία που λαμβάνουμε είναι πολύ μεγάλη αφού το γεγονός A είναι εξαιρετικά σπάνιο. Από την παραπάνω συζήτηση γίνεται φανερό πως όσο λιγότερο πιθανό είναι ένα γεγονός, τόσο μεγαλύτερη είναι η "πληροφορία" που λαμβάνουμε από την πραγματοποίησή του. Έτσι, ορίζουμε την *πληροφορία* την οποία αποκτούμε από την πραγματοποίηση ενός γεγονότος $\{S = s_k\}$ το οποίο έχει πιθανότητα p_k ως

$$I(s_k) = \log\left(\frac{1}{p_k}\right) \tag{3}$$

Ο ορισμός της πληροφορίας από τον τύπο αυτό ικανοποιεί τις ακόλουθες ιδιότητες

1. $$I(\{S = s_k\}) = 0 \quad \text{για} \quad p_k = 1$$

   Ικανοποιείται δηλαδή ο διαισθητικός περιορισμός πως αν γνωρίζουμε εκ των προτέρων την έκβαση ενός πειράματος τύχης η πληροφορία που λαμβάνουμε είναι μηδενική.

2. $$I(\{S = s_k\}) \geq 0 \quad \text{για} \quad 0 \leq p_k \leq 1$$

   που σημαίνει πως η έκβαση ενός πειράματος τύχης μπορεί να περιέχει κάποια θετική ή μηδενική πληροφορία, ποτέ όμως δεν χάνουμε πληροφορία από την έκβαση ενός πειράματος τύχης.

3. $$I(\{S = s_k\}) > I(\{S = s_i\}) \quad \text{για} \quad p_k < p_i$$

   που σημαίνει πως όσο πιο απίθανο να συμβεί είναι ένα γεγονός, τόσο περισσότερη πληροφορία λαμβάνουμε από την πραγματοποίησή του.

4.
$$I(\{S = s_k\} \cup \{S = s_l\}) = I(\{S = s_k\}) + I(\{S = s_l\})$$

άν τα γεγονότα $\{S = s_k\}$ και $\{S = s_l\}$ είναι στατιστικά ανεξάρτητα.

Η βάση του λογαρίθμου στην εξίσωση (3) μπορεί να επιλεγεί αρκετά αυθαίρετα. Ωστόσο, έχει επικρατήσει να χρησιμοποιείται λογάριθμος με βάση το 2. Στην περίπτωση αυτή, η μονάδα μέτρησης της πληροφορίας ονομάζεται bit ή δυαδική μονάδα και γράφουμε

$$I(s_k) = \log_2\left(\frac{1}{p_k}\right) \quad k = 0, 1, 2, \ldots, K-1 \tag{4}$$

Όταν είναι $p_k = 1/2$ έχουμε $I(s_k) = 1$ bit. Έτσι, *1 bit είναι η ποσότητα της πληροφορίας που λαμβάνουμε όταν συμβαίνει ένα γεγονός ανάμεσα σε 2 πιθανά και ισοπίθανα γεγονότα.*Αξίζει να σημειώσουμε πως διαφορετικό νόημα έχει το bit ως μονάδα πληροφορίας (binary unit) από το bit που χρησιμοποιούμε για το δυαδικό ψηφίο (binary digit).

2 Η έννοια της εντροπίας

Η ποσότητα της πληροφορίας $I(s_k)$ η οποία παράγεται από μια πηγή κατά τη διάρκεια μιας περιόδου συμβόλου T_s εξαρτάται μόνο από το σύμβολο s_k το οποίο παράγει η πηγή στο χρόνο αυτό. Μάλιστα, την πληροφορία αυτή $I(s_k)$ μπορούμε να την δούμε σαν μια διακριτή τυχαία μεταβλητή η οποία λαμβάνει τις τιμές $I(s_0), I(s_1), \ldots I(s_{K-1})$ με πιθανότητες $p_0, p_1, \ldots, p_{K-1}$ αντίστοιχα. Η μέση τιμή της τυχαίας αυτής μεταβλητής ως προς τις πιθανότητες αυτές θα είναι τότε:

$$\begin{aligned} H(\Phi) &= E[I(s_k)] \\ &= \sum_{k=0}^{K-1} p_k I(s_k) \\ &= \sum_{k=0}^{K-1} p_k \log_2\left(\frac{1}{p_k}\right) \end{aligned} \tag{5}$$

Η ποσότητα $H(\Phi)$, ονομάζεται *εντροπία* μιας διακριτής πηγής χωρίς μνήμη η οποία έχει αλφάβητο Φ. Από τον ορισμό της διαπιστώνουμε ότι εκφράζει τη μέση πληροφορία που παράγει μια διακριτή πηγή ανά σύμβολο της πηγής. Η τιμή της εντροπίας εξαρτάται αποκλειστικά από τις πιθανότητες p_k των συμβόλων της πηγής. Αξίζει να αναφέρουμε πως ο δείκτης Φ στην έκφραση $H(\Phi)$ δεν αποτελεί παράμετρο συνάρτησης αλλά εκφράζει την εξάρτηση της εντροπίας από μια συγκεκριμένη πηγή.

2.1 Ιδιότητες της εντροπίας

Ας θεωρήσουμε μια διακριτή πηγή χωρίς μνήμη όπως την περιγράψαμε στα προηγούμενα με βάση τις εξισώσεις (1) και (2). Η εντροπία $H(\Phi)$ της πηγής αυτής τότε φράσσεται σύμφωνα με την εξίσωση:

$$0 \leq H(\Phi) \leq \log_2 K \tag{6}$$

όπου K είναι η *βάση* (πλήθος συμβόλων) του αλφάβητου Φ. Επιπρόσθετα, μπορούμε να διατυπώσουμε τις εξής δυο ιδιότητες:

1. Η εντροπία $H(\Phi)$ μιας διακριτής πηγής χωρίς μνήμη είναι μηδέν ($H(\Phi) = 0$), αν και μόνο αν η πιθανότητα κάποιου συμβόλου s_k της πηγής είναι μονάδα ($p_k = 1$) και οι πιθανότητες όλων των υπολοίπων συμβόλων είναι 0. Αυτό το κάτω φράγμα στην εντροπία αντιστοιχεί στην κατάσταση όπου δεν έχουμε καθόλου αβεβαιότητα για το σύμβολο που θα εκπέμψει μια διακριτή πηγή χωρίς μνήμη.

2. Η εντροπία μιας διακριτής πηγής χωρίς μνήμη λαμβάνει τη μέγιστη τιμή της η οποία είναι ίση με $H(\Phi) = \log_2 K$ αν και μόνο αν $p_k = 1/K, \forall k = 0, 1, \ldots K$, δηλαδή στην περίπτωση όπου όλα τα σύμβολα της πηγής είναι ισοπίθανα. Αυτό το πάνω φράγμα στην τιμή της εντροπίας αντιστοιχεί στην κατάσταση όπου έχουμε τη μέγιστη αβεβαιότητα σχετικά με το σύμβολο που θα εκπέμψει μια διακριτή πηγή χωρίς μνήμη.

Ας αποδείξουμε αρχικά το κάτω φράγμα της (6) και τον ισχυρισμό (1). Αρχικά, αφού είναι $0 \leq p_k \leq 1$ κάθε όρος της μορφής $p_k \log_2(1/p_k)$ στην εξίσωση (5) θα είναι θετικός ή μηδέν. Έτσι, η εντροπία θα είναι θετική ή μηδέν ως άθροισμα τέτοιων όρων.

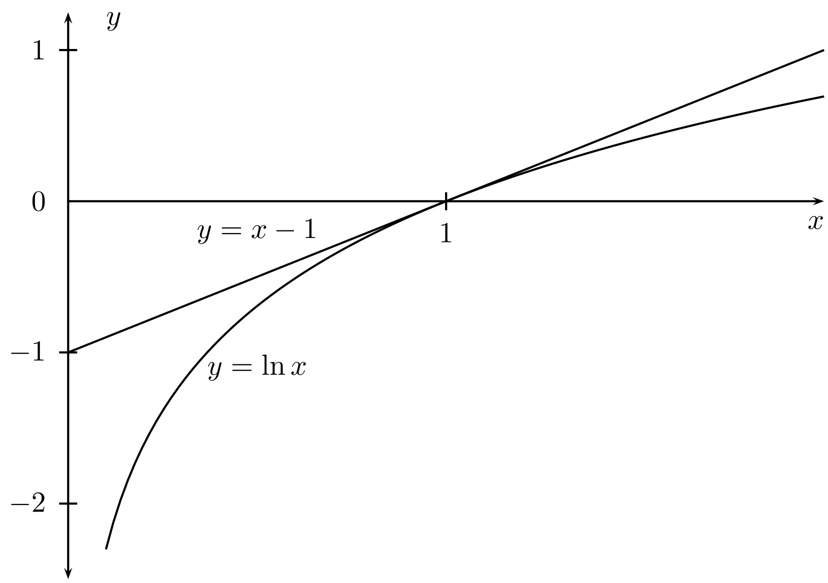
$$0 \leq p_k \leq 1 \quad \forall \quad k \Rightarrow p_k \log_2(1/p_k) \geq 0 \quad \forall \quad k \Rightarrow H(\Phi) \geq 0$$

Απομένει να αναζητήσουμε πότε λαμβάνει την τιμή 0. Παρατηρούμε πως οι όροι της μορφής $p_k \log_2(1/p_k)$ είναι ίσοι με μηδέν αν και μόνο άν $p_k = 0$ ή $p_k = 1$. Έτσι, ο μόνος τρόπος για να έχουμε $H(\Phi) = 0$ είναι να έχουμε $p_k = 1$ για κάποιο συγκεκριμένο k και όλες οι άλλες πιθανότητες να είναι 0. Αποδείξαμε έτσι το κάτω φράγμα της (6) και τον ισχυρισμό (1).

Για να αποδείξουμε το πάνω φράγμα της (6) και τον ισχυρισμό (2) θα χρησιμοποιήσουμε την ακόλουθη ιδιότητα του φυσικού λογαρίθμου:

$$\ln x \leq x - 1 \qquad x \geq 0 \tag{7}$$

Την ιδιότητα αυτή μπορούμε να την επιβεβαιώσουμε από το σχήμα (1) όπου έχουμε σχεδιάσει τις γραφικές παραστάσεις των συναρτήσεων $y = x - 1$ και $y = \ln x$. Παρατηρούμε πως η γραφική παράσταση της $y = x - 1$ βρίσκεται πάντα πάνω από τη γραφική παράσταση της $y = \ln x$ εκτός από το σημείο (0,1) στο οποίο ταυτίζονται.



Σχήμα 1: Γραφικές παραστάσεις των συναρτήσεων $y = x - 1$ και $y = \ln x$

Συνεχίζοντας την απόδειξη, θεωρούμε δυο οποιεσδήποτε κατανομές πιθανοτήτων $p = \{p_0, p_1, \ldots, p_{K-1}\}$ και $q = \{q_0, q_1, \ldots, q_{K-1}\}$ των συμβόλων του αλφαβήτου Φ μιας διακριτής πηγής χωρίς μνήμη. Θεωρούμε την ποσότητα:

$$\sum_{k=0}^{K-1} p_k \log_2\left(\frac{q_k}{p_k}\right) = \frac{1}{\log_2 e} \sum_{k=0}^{K-1} p_k \ln\left(\frac{q_k}{p_k}\right)$$

όπου εφαρμόσαμε την ιδιότητα αλλαγής βάσης του λογαρίθμου, και με e έχουμε συμβολίσει τη βάση του φυσικού λογαρίθμου. Χρησιμοποιώντας την ανισότητα (7) στην παραπάνω σχέση, έχουμε:

$$\begin{aligned}
\sum_{k=0}^{K-1} p_k \log_2\left(\frac{q_k}{p_k}\right) &\leq \frac{1}{\log_2 e} \sum_{k=0}^{K-1} p_k \left(\frac{q_k}{p_k} - 1\right) \\
&\leq \frac{1}{\log_2 e} \sum_{k=0}^{K-1} (q_k - p_k) \\
&\leq \frac{1}{\log_2 e} \left(\sum_{k=0}^{K-1} q_k - \sum_{k=0}^{K-1} p_k\right) = 0
\end{aligned}$$

και καταλήγουμε στην *θεμελιώδη ανισότητα:*

$$\sum_{k=0}^{K-1} p_k \log_2\left(\frac{q_k}{p_k}\right) \leq 0 \tag{8}$$

με την ισότητα να ισχύει μόνο αν $p_k = q_k \quad \forall \quad k$.

Υποθέτουμε τώρα πως είναι

$$q_k = \frac{1}{K} \qquad k = 0, 1, \ldots, K-1 \tag{9}$$

δηλαδή πως η κατανομή πιθανοτήτων q αντιστοιχεί σε ισοπίθανα σύμβολα. Τότε, η εντροπία της διακριτής πηγής χωρίς μνήμη της οποίας τα σύμβολα έχουν πιθανότητες q_k θα είναι

$$\sum_{k=0}^{K-1} q_k \log_2 \left(\frac{1}{q_k} \right) = \log_2 K$$

Ακόμα, χρησιμοποιώντας τις (8) και (9) θα έχουμε:

$$\sum_{k=0}^{K-1} p_k \log_2 \left(\frac{1}{p_k} \right) \leq \log_2 K$$

ή με άλλα λόγια, η εντροπία μιας διακριτής πηγής χωρίς μνήμης με οποιαδήποτε κατανομή πιθανότητας των συμβόλων της φράσσεται σύμφωνα με την

$$H(\Phi) \leq \log_2 K$$

Επιπρόσθετα, η ισότητα $H(\Phi) = \log_2 K$ ισχύει μόνο όταν όλα τα σύμβολα της πηγής είναι ισοπίθανα όπως είδαμε από την (9). Αποδείξαμε έτσι το πάνω φράγμα της (6) και τον ισχυρισμό (2).

ΠΑΡΑΔΕΙΓΜΑ 1

Θεωρούμε μια διακριτή πηγή χωρίς μνήμη της οποίας το αλφάβητο είναι το σύνολο $\Phi = \{0, 1\}$. Μια πηγή με αυτό το αλφάβητο θα ονομάζεται *δυαδική*. Υποθέτουμε ακόμα πως το σύμβολο 0 εμφανίζεται με πιθανότητα p_0 και το σύμβολο 1 εμφανίζεται με πιθανότητα $p_1 = 1 - p_0$. Να μελετηθεί η εντροπία της πηγής για τις διάφορες τιμές της πιθανότητας p_0.

Λύση:

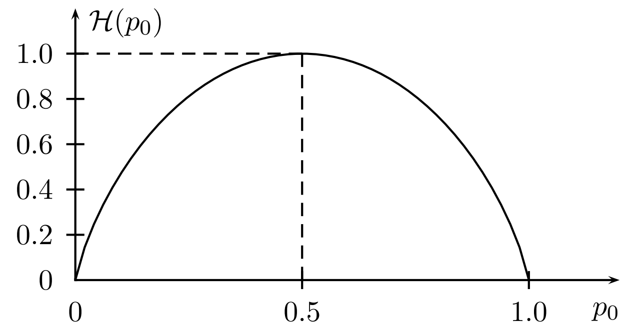
Η εντροπία της δυαδικής πηγής χωρίς μνήμη θα δίνεται από τη σχέση:

$$\begin{aligned} H(\Phi) &= -p_0 \log_2 p_0 - p_1 \log_2 p_1 \\ &= -p_0 \log_2 p_0 - (1 - p_0) \log_2 (1 - p_0) \quad \text{bits} \end{aligned} \tag{10}$$

Από τη σχέση αυτή παρατηρούμε τα εξής:

1. Όταν $p_0 = 0$, είναι $H(\Phi) = 0$. Αυτό προκύπτει από τη σχέση

$$\lim_{x \to 0} x \log_2 x = 0$$



Σχήμα 2: Γραφική παράσταση της συνάρτησης εντροπίας

2. Όταν $p_0 = 1$, είναι $H(\Phi) = 0$.

3. Η εντροπία $H(\Phi)$ λαμβάνει τη μέγιστη τιμή της $H_{max} = 1\text{bit}$, όταν $p_0 = p_1 = 1/2$ δηλαδή όταν τα σύμβολα 1 και 0 είναι ισοπίθανα.

Η σχέση (10) συναντάται συχνά σε προβλήματα της θεωρίας πληροφορίας. Για το λόγο αυτό, συνηθίζεται να της δίνουμε ένα ειδικό σύμβολο και συγκεκριμένα ορίζουμε:

$$\mathcal{H}(p_0) = -p_0 \log_2 p_0 - (1 - p_0) \log_2 (1 - p_0) \tag{11}$$

και στο εξής θα την ονομάζουμε *συνάρτηση εντροπίας*. Αξίζει να σημειώσουμε την διαφορά ανάμεσα στη σχέση (10) και τη σχέση (11). Η σχέση (10) δίνει την εντροπία μιας διακριτής πηγής χωρίς μνήμη η οποία έχει αλφάβητο Φ. Αντίθετα, η σχέση (11) ορίζει μια συνάρτηση της μεταβλητής p_0 η οποία ορίζεται στο διάστημα $[0, 1]$. Στο σχήμα (2.1) παρουσιάζουμε τη γραφική παράσταση της συνάρτησης εντροπίας η οποία δικαιολογεί τις προηγούμενες παρατηρήσεις 1, 2 και 3. ■

2.2 n τάξης επέκταση διακριτής πηγής χωρίς μνήμη

Στην θεωρία πληροφορίας, συχνά βολεύει να θεωρούμε ομάδες συμβόλων παρά μεμονωμένα σύμβολα. Ως *ομάδα συμβόλων* θεωρούμε n σύμβολα τα οποία εκπέμπονται από μια διακριτή πηγή σε διαδοχικές χρονικές στιγμές. Οι ομάδες συμβόλων, μπορούμε να θεωρήσουμε πως αποτελούν σύμβολα μιας νέας διακριτής πηγής η οποία έχει το αλφάβητο Φ^n το οποίο αποτελείται από K^n σύμβολα, όπου K το πλήθος των συμβόλων του αλφαβήτου Φ της αρχικής πηγής.

Στην περίπτωση όπου η αρχική πηγή είναι μια διακριτή πηγή χωρίς μνήμη, τα σύμβολα στην έξοδο της είναι μεταξύ τους στατιστικά ανεξάρτητα. Έτσι, η πιθανότητα εμφάνισης ενός συμβόλου από το αλφάβητο Φ^n είναι ίση με το γινόμενο των πιθανοτήτων των n επιμέρους συμβόλων που το αποτελούν από το αλφάβητο Φ της αρχικής πηγής. Με βάση την παρατήρηση αυτή, μπορεί να αποδειχθεί πως η εντροπία μιας διακριτής πηγής χωρίς μνήμη και η εντροπία της αντίστοιχης n τάξης επέκτασής της σχετίζονται με τον τύπο:

$$H(\Phi^n) = nH(\Phi) \tag{12}$$

ΠΑΡΑΔΕΙΓΜΑ 2

Θεωρήστε μια διακριτή πηγή χωρίς μνήμη με αλφάβητο $\Phi = \{s_0, s_1, s_2\}$ και πιθανότητες εμφάνισης συμβόλων

$$p(s_0) = p_0 = \frac{1}{4}, \quad p(s_1) = p_1 = \frac{1}{4}, \quad p(s_2) = p_2 = \frac{1}{2}$$

Να υπολογιστεί η εντροπία της πηγής αυτής καθώς και η εντροπία της δεύτερης τάξης επέκτασής της.

Λύση:

Η εντροπία της αρχικής πηγής είναι:

$$\begin{aligned} H(\Phi) &= p_0 \log_2\left(\frac{1}{p_0}\right) + p_1 \log_2\left(\frac{1}{p_1}\right) + p_2 \log_2\left(\frac{1}{p_2}\right) \\ &= \frac{1}{4}\log_2(4) + \frac{1}{4}\log_2(4) + \frac{1}{2}\log_2(2) \\ &= \frac{3}{2} \quad \text{bits} \end{aligned}$$

Η δεύτερης τάξης επέκταση της πηγής που εξετάζουμε θα είναι μια νέα διακριτή πηγή χωρίς μνήμη με αλφάβητο

$$\begin{aligned} \Phi^2 &= \{\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6, \sigma_7, \sigma_8\} \\ &= \{s_0s_0, s_0s_1, s_0s_2, s_1s_0, s_1s_1, s_1s_2, s_2s_0, s_2s_1, s_2s_2\} \end{aligned}$$

θα αποτελείται δηλαδή από $3^2 = 9$ σύμβολα. Υπολογίζουμε τώρα τις αντίστοιχες πιθανότητες των συμβόλων $\sigma_0, \sigma_1, \ldots, \sigma_8$ ως $q_0, q_1, \ldots, q_8$. Οι πιθανότητες αυτές υπολογίζονται ως γινόμενα των πιθανοτήτων των επιμέρους συμβόλων του Φ που αποτελούν τα σύμβολα του Φ^2.

q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8
$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

και έτσι η εντροπία της πηγής θα είναι

$$\begin{aligned} H(\Phi^2) &= \sum_{i=0}^{8} q_i \log_2\left(\frac{1}{q_i}\right) \\ &= 4\frac{1}{16}\log_2(16) + 4\frac{1}{8}\log_2(8) + \frac{1}{4}\log_2(4) \\ &= 3 \quad \text{bits} \end{aligned}$$

Παρατηρούμε πως τα αποτελέσματα που πήραμε για τις εντροπίες των 2 πηγών συμφωνούν με τη σχέση (12). ■

3 Θεώρημα κωδικοποίησης πηγής

Ένα σημαντικό πρόβλημα το οποίο συναντάται στις τηλεπικοινωνίες είναι η αποδοτική αναπαράσταση της πληροφορίας η οποία παράγεται από μια διακριτή πηγή δεδομένων. Η διαδικασία αυτή της αναπαράστασης των δεδομένων που παράγονται από μια διακριτή πηγή ονομάζεται *κωδικοποίηση πηγής* και η διάταξη η οποία επιτελεί την εργασία αυτή *κωδικοποιητής πηγής*.

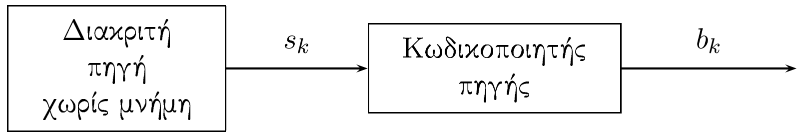
Για να μπορέσει ένας κωδικοποιητής πηγής να αναπαραστήσει αποδοτικά τα δεδομένα μιας διακριτής πηγής είναι απαραίτητο να γνωρίζει ορισμένες στατιστικές ιδιότητες για την πηγή αυτή. Για παράδειγμα, αν ένα σύμβολο της πηγής εμφανίζεται με μεγάλη πιθανότητα τότε θα επιθυμούσαμε να αναπαραστήσουμε το σύμβολο αυτό με κάποιο απλό και σύντομο τρόπο, αντίθετα τα σύμβολα τα οποία εμφανίζονται σπάνια και επομένως έχουν μικρές πιθανότητες δεν μας ενοχλεί ιδιαίτερα να τα κωδικοποιούμε με κάποια πιο πολύπλοκη και μακροσκελή αναπαράσταση. Οι αναπαραστάσεις των συμβόλων που προκύπτουν στην έξοδο ενός κωδικοποιητή πηγής ονομάζονται *κωδικές λέξεις* και το σύνολο των κωδικών λέξεων που χρησιμοποιεί ένας κωδικοποιητής πηγής με σκοπό να αναπαραστήσει όλα τα σύμβολα μιας διακριτής πηγής ονομάζεται *κώδικας πηγής*.

Ένας κώδικας πηγής ο οποίος αποτελείται από κωδικές λέξεις των οποίων το μήκος δεν είναι ίδιο ονομάζεται *κώδικας μεταβλητού μήκους*. Για παράδειγμα, ο κώδικας Morse, ο οποίος χρησιμοποιεί τα σημεία της τελείας "."και της παύλας "-" για να αναπαραστήσει ένα πεπερασμένο αλφάβητο (γράμματα), χρησιμοποιεί την κωδική λέξη "."για να αναπαραστήσει το λατινικό γράμμα "E" (το οποίο έχει μεγάλη πιθανότητα εμφάνισης) και την κωδική λέξη "- -.-" για το λατινικό γράμμα "Q" το οποίο έχει μικρότερη πιθανότητα εμφάνισης. Επομένως ο κώδικας Morse είναι ένας κώδικας μεταβλητού μήκους.

Οι κωδικοποιητές πηγής με τους οποίους θα ασχοληθούμε στη συνέχεια θα πρέπει να ικανοποιούν τους ακόλουθους περιορισμούς:

1. Οι κωδικές λέξεις που παράγει ο κωδικοποιητής πηγής είναι σε δυαδική μορφή.
2. Μπορούμε πάντα να ανακατασκευάσουμε την ακολουθία συμβόλων της πηγής αν γνωρίζουμε την αντίστοιχη ακολουθία κωδικών λέξεων που δημιούργησε ο κωδικοποιητής της πηγής. Ένας κώδικας ο οποίος έχει την ιδιότητα αυτή θα ονομάζεται *αποκωδικοποιήσιμος*.

Ας θεωρήσουμε τη διάταξη του σχήματος (3) όπου η έξοδος s_k (σύμβολο) μιας διακριτής πηγής χωρίς μνήμη οδηγείτε σε έναν κωδικοποιητή πηγής ο οποίος δίνει στην έξοδό του μια ακολουθία από 0 και 1 b_k (κωδική λέξη). Υποθέτουμε πως η πηγή έχει K διακριτά σύμβολα $s_k, k = 0, 1, \ldots, K-1$ με



Σχήμα 3: Κωδικοποίηση πηγής

πιθανότητες εμφάνισης $p_0, p_1, \ldots, p_{K-1}$ αντίστοιχα. Υποθέτουμε ακόμα πως η κωδική λέξη b_k που αντιστοιχεί στο σύμβολο s_k έχει μήκος l_k. Ορίζουμε έτσι το *μέσο μήκος κώδικα* $\overline{L}$ ώς:

$$\overline{L} = \sum_{k=0}^{K-1} p_k l_k \tag{13}$$

Το μέσο μήκος κώδικα $\overline{L}$ εκφράζει το μέσο πλήθος δυαδικών ψηφίων ανά σύμβολο πηγής τα οποία χρησιμοποιούνται στην διαδικασία της κωδικοποίησης. Αν θεωρήσουμε ως L_{min} την ελάχιστη δυνατή τιμή του $\overline{L}$ τότε μπορούμε να ορίσουμε την *αποδοτικότητα* του κώδικα ως:

$$\eta = \frac{L_{min}}{\overline{L}} \tag{14}$$

Παρατηρούμε πως αφού $L_{min} \leq \overline{L}$, η αποδοτικότητα η λαμβάνει τιμές μικρότερες της μονάδας. Ένας κώδικας ονομάζεται *αποδοτικός*, αν η αποδοτικότητά του πλησιάζει την τιμή 1.

Η τιμή L_{min} στην οποία αναφερθήκαμε στα προηγούμενα αποτελεί χαρακτηριστικό της πηγής δεδομένων και όχι του κωδικοποιητή. Πιο συγκεκριμένα, η τιμή της υπολογίζεται με βάση το ακόλουθο θεώρημα το οποίο αποτελεί το *πρώτο θεώρημα του Shannon* ή το *θεώρημα κωδικοποίησης πηγής*

> Θεώρημα: Για μια δοσμένη διακριτή πηγή χωρίς μνήμη με εντροπία $H(\Phi)$, το μέσο μήκος κώδικα $\overline{L}$ οποιουδήποτε κωδικοποιητή πηγής φράσσεται από την ανισότητα
>
> $$\overline{L} \geq H(\Phi)$$

Το παραπάνω θεώρημα μας πληροφορεί πως δεν υπάρχει κωδικοποιητής πηγής ο οποίος να πετυχαίνει μικρότερο μέσο μήκος κώδικα $\overline{L}$ από την εντροπία της πηγής. Έτσι, η τιμή L_{min} η οποία χρησιμοποιήσαμε στα προηγούμενα ισούται με την εντροπία της διακριτής πηγής την οποία κωδικοποιούμε, και μπορούμε έτσι να ορίσουμε την αποδοτικότητα ενός κωδικοποιητή πηγής ως

$$\eta = \frac{H(\Phi)}{\overline{L}} \tag{15}$$

Στα επόμενα, ασχολούμαστε με μια κατηγορία κωδικών οι οποίοι ονομάζονται *προθεματικοί κώδικες*. Οι κώδικες αυτοί είναι πάντα αποκωδικοποιήσιμοι και επιπλέον μπορούν να πετύχουν μέσο μήκος κώδικα οσοδήποτε κοντά στην εντροπία της πηγής την οποία κωδικοποιούν.

3.1 Προθεματικοί κώδικες

Ας θεωρήσουμε μια διακριτή πηγή χωρίς μνήμη με αλφάβητο $\{s_0, s_1, \ldots, s_{K-1}\}$ και πιθανότητες συμβόλων $p_0, p_1, \ldots, p_{K-1}$. Ένας κωδικοποιητής πηγής ο οποίος αντιστοιχίζει κάθε σύμβολο της πηγής αυτής σε μια κωδική λέξη είναι χρήσιμος μόνο στην περίπτωση όπου από την ακολουθία των κωδικών λέξεων μπορούμε να ανακατασκευάσουμε την αρχική ακολουθία συμβόλων της πηγής. Ο περιορισμός αυτός σημαίνει πως η ακολουθία κωδικών λέξεων που αντιστοιχεί σε μια πεπερασμένη ακολουθία συμβόλων της πηγής είναι διαφορετική από όλες τις ακολουθίες κωδικών λέξεων οι οποίες προκύπτουν για διαφορετικές ακολουθίες συμβόλων της πηγής, ή με άλλα λόγια η κωδικοποίηση είναι μια "1-1"συνάρτηση από το σύνολο Φ (αλφάβητο πηγής) στο σύνολο $B = \{b_0, b_1, \ldots, b_{K-1}\}$ το οποίο περιέχει τις κωδικές λέξεις.

Ας υποθέσουμε πως η κωδική λέξη b_k η οποία αντιστοιχεί στο σύμβολο s_k είναι

$$b_k = (m_{k_1}, m_{k_2}, \ldots, m_{k_n})$$

όπου m_{k_i} είναι το i-οστό δυαδικό ψηφίο (0 ή 1) της κωδικής λέξης και n είναι το μήκος της κωδικής λέξης b_k. Ορίζουμε ως πρόθεμα της κωδικής λέξης b_k οποιαδήποτε ακολουθία δυαδικών ψηφίων της μορφής

$$(m_{k_1}, m_{k_2}, \ldots, m_{k_i}) \qquad i \leq n$$

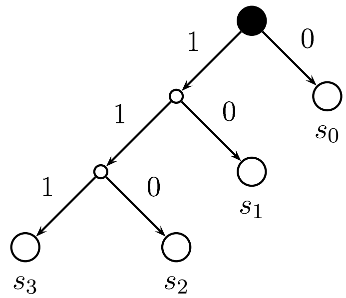
Ορίζουμε έτσι ως *προθεματικό κώδικα* έναν κώδικα για τον οποίο καμιά κωδική λέξη του δεν αποτελεί πρόθεμα κάποιας άλλης κωδικής λέξης. Η συνθήκη αυτή καθιστά τους προθεματικούς κώδικες πάντα αποκωδικοποιήσιμους.

Για παράδειγμα, ας θεωρήσουμε τους ακόλουθους 3 κώδικες

Σύμβολο πηγής	Πιθανότητα εμφάνισης	κώδικας 1	κώδικας 2	κώδικας 3
s_0	0.5	0	0	0
s_1	0.25	1	10	01
s_2	0.125	00	110	011
s_3	0.125	11	111	0111

Εύκολα παρατηρούμε πως ο κώδικας 1 δεν είναι προθεματικός, αφού η κωδική λέξη "0" που αντιστοιχεί στο σύμβολο s_0 αποτελεί πρόθεμα της κωδικής λέξης "00" που αντιστοιχεί στο σύμβολο s_2. Με όμοιο τρόπο μπορούμε να δούμε πως ο κώδικας 2 είναι προθεματικός, όχι όμως και ο κώδικας 3.

Η διαδικασία της μετατροπής των κωδικών λέξεων στα αντίστοιχα σύμβολα της πηγής ονομάζεται αποκωδικοποίηση και οι διατάξεις οι οποίες εκτελούν



Σχήμα 4: Δέντρο απόφασης για τον κώδικα 2

την εργασία αυτή *αποκωδικοποιητές πηγής*. Η αποκωδικοποίηση ενός προθεματικού κώδικα γίνεται με χρήση *δέντρων απόφασης*. Ένα δέντρο απόφασης αποτελεί μια γραφοθεωρητική αναπαράσταση όλων των κωδικών λέξεων ενός προθεματικού κώδικα. Για παράδειγμα στο σχήμα (4) παρουσιάζεται το δέντρο απόφασης του κώδικα 2 που είδαμε νωρίτερα. Παρατηρούμε πως κάθε ακμή του γραφήματος έχει μια ετικέτα, 0 ή 1. Κάθε μονοπάτι από τη ρίζα του δέντρου απόφασης ως τα φύλλα του αντιστοιχεί σε μια κωδική λέξη, την κωδική λέξη που προκύπτει από τις ετικέτες των ακμών που αποτελούν το μονοπάτι.

Ο αποκωδικοποιητής ενός προθεματικού κώδικα ξεκινά από τη ρίζα του δέντρου απόφασης. Για κάθε ένα δυαδικό ψηφίο που εμφανίζεται στην είσοδό του, ακολουθεί την ακμή με την αντίστοιχη ετικέτα και έτσι μεταβαίνει σε έναν νέο κόμβο. Αν σε κάποιο βήμα καταλήξει σε φύλλο του δέντρου, τότε βγάζει στην έξοδό του το αντίστοιχο σύμβολο πηγής και επιστρέφει στη ρίζα του δέντρου για να ξεκινήσει την αποκωδικοποίηση του νέου συμβόλου. Ακολουθώντας τη διαδικασία αυτή, με είσοδο τα δυαδικά ψηφία "1011111000" στο δέντρο απόφασης του σχήματος (4) καταλήγουμε στην ακολουθία συμβόλων s_1, s_3, s_2, s_0, s_0.

Έχουμε ήδη αναφέρει πως όλοι οι προθεματικοί κώδικες είναι αποκωδικοποιήσιμοι. Ωστόσο υπάρχουν και μη-προθεματικοί κώδικες οι οποίοι είναι αποκωδικοποιήσιμοι, για παράδειγμα ο κώδικας 3 του προηγούμενου πίνακα είναι αποκωδικοποιήσιμος χωρίς να είναι προθεματικός.

Για μια δοσμένη διακριτή πηγή χωρίς μνήμη, το μέσο μήκος κώδικα $\overline{L}$ για τους προθεματικούς κώδικες αποδεικνύεται πως φράσσεται από την ανισότητα:

$$H(\Phi) \leq \overline{L} < H(\Phi) + 1 \tag{16}$$

Έχουμε ήδη αναφέρει πως για μια δοσμένη διακριτή πηγή χωρίς μνήμη, μπορούμε πάντα να σχεδιάσουμε έναν προθεματικό κώδικα ο οποίος να πετυχαίνει μέσο μήκος κώδικα $\overline{L}$ οσοδήποτε κοντά στην εντροπία της πηγής. Για να δούμε πως μπορεί να γίνει αυτό, ας θεωρήσουμε την n-τάξης επέκταση της διακριτής πηγής, η οποία είναι μια νέα διακριτή πηγή χωρίς μνήμη με αλφάβητο Φ^n. Έστω τώρα ένας προθεματικός κώδικας για τη νέα πηγή. Ο κώδικας αυτός θα ονομάζεται *n-τάξης επεκταμένος κώδικας*. Για το ζεύγος νέας πηγής και νέου κώδικα, θα ισχύει η ανισότητα (16), δηλαδή:

$$H(\Phi^n) \leq \overline{L}_n < H(\Phi^n) + 1$$

όπου $\overline{L}_n$ το μέσο μήκος κώδικα του n-τάξης επεκταμένου κώδικα. Χρησιμοποιώντας τώρα την ιδιότητα

$$H(\Phi^n) = n \cdot H(\Phi)$$

λαμβάνουμε

$$n \cdot H(\Phi) \leq \overline{L}_n < n \cdot H(\Phi) + 1$$

ή ισοδύναμα

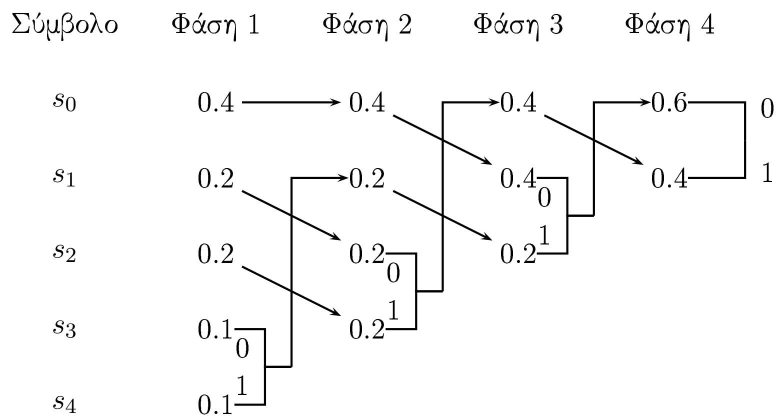
$$H(\Phi) \leq \frac{\overline{L}_n}{n} < H(\Phi) + \frac{1}{n} \tag{17}$$

Από την τελευταία σχέση γίνεται φανερό πως όταν η τάξη n του κώδικα τείνει στο άπειρο, το μέσο μήκος του $\overline{L}_n$ συγκλίνει στην εντροπία της πηγής. Με άλλα λόγια, το μέσο μήκος ενός προθεματικού επεκταμένου κώδικα μπορεί να προσεγγίσει την εντροπία μιας διακριτής πηγής χωρίς μνήμη, αρκεί η τάξη n του κώδικα να είναι αρκετά μεγάλη. Στο σημείο αυτό αξίζει να αναφέρουμε πως το μειονέκτημα της χρήσης ενός επεκταμένου κώδικα είναι η μεγάλη πολυπλοκότητα του αλγορίθμου αποκωδικοποίησής του.

4 Κωδικοποίηση Huffman

Ο *κώδικας Huffman* είναι ένας κώδικας πηγής, το μέσο μήκος του οποίου πλησιάζει την εντροπία της διακριτής πηγής την οποία κωδικοποιεί. Ο κώδικας αυτός είναι βέλτιστος με την έννοια ότι για μια δοσμένη διακριτή πηγή χωρίς μνήμη δεν υπάρχει άλλος αποκωδικοποιήσιμος κώδικας ο οποίος να πετυχαίνει μικρότερο μέσο μήκος.

Η ιδέα του αλγορίθμου με βάση τον οποίο κατασκευάζουμε τον κώδικα Huffman είναι η σταδιακή *μείωση* του αλφαβήτου της πηγής. Σε κάθε βήμα, επιλέγουμε 2 σύμβολα από το αλφάβητο της πηγής και τα συγχωνεύουμε σε ένα σύμβολο το οποίο λαμβάνει μέρος στην επόμενη φάση του αλγορίθμου. Ακολουθώντας τη διαδικασία αυτή, καταλήγουμε σε ένα αλφάβητο το οποίο αποτελείται από 2 μόνο σύμβολα. Τα σύμβολα αυτά γνωρίζουμε πως κωδικοποιούνται βέλτιστα από τις κωδικές λέξεις “0” και “1”. Στη συνέχεια, ξεκινώντας από την τελευταία φάση εργαζόμαστε προς τα πίσω και καταλήγουμε σε έναν βέλτιστο κώδικα. Πιο συγκεκριμένα, ο αλγόριθμος κατασκευής του κώδικα Huffman έχει ως εξής:



Σχήμα 5: Διαδικασία εύρεσης του κώδικα Huffman για το παράδειγμα 3

1. Τα σύμβολα της πηγής ταξινομούνται από το πλέον πιθανό έως το λιγότερο πιθανό. Στα δυο λιγότερο πιθανά σύμβολα ανατίθενται τα δυαδικά ψηφία 0 και 1 αντίστοιχα.

2. Τα δυο σύμβολα αυτά συγχωνεύονται σε ένα σύμβολο με πιθανότητα ίση με το άθροισμα των πιθανοτήτων των συμβόλων αυτών. Με τον τρόπο αυτό το αλφάβητο μειώνετε κατά ένα σύμβολο. Το νέο σύμβολο τοποθετείται στη σωστή θέση με βάση την πιθανότητα που υπολογίστηκε.

3. Η διαδικασία επαναλαμβάνεται μέχρι να απομείνουν μόνο 2 σύμβολα. Στα σύμβολα αυτά αναθέτουμε τα δυαδικά ψηφία 0 και 1.

4. Οι κωδικές λέξεις για κάθε σύμβολο βρίσκονται ξεκινώντας από την τελευταία φάση, συλλέγοντας τα δυαδικά ψηφία που έχουμε αναθέσει μέχρι να καταλήξουμε στο αρχικό σύμβολο (στη φάση 1) στο οποίο αντιστοιχεί η κωδική λέξη.

ΠΑΡΑΔΕΙΓΜΑ 3

Να βρεθεί η κωδικοποίηση Huffman για μια διακριτή πηγή χωρίς μνήμη με αλφάβητο $\Phi = \{s_0, s_1, s_2, s_3, s_4\}$ και πιθανότητες συμβόλων $p_0 = 0.4, p_1 = 0.2, p_2 = 0.2, p_3 = 0.1, p_4 = 0.1$. Στη συνέχεια να υπολογιστεί το μέσο μήκος κώδικα και η εντροπία της πηγής και να σχολιαστούν τα αποτελέσματα.

Λύση:

Στο σχήμα (5) παρουσιάζουμε τη διαδικασία εύρεσης της κωδικοποίησης Huffman για την πηγή που εξετάζουμε. Στον πίνακα που ακολουθεί παρουσιάζουμε τις κωδικές λέξεις που προκύπτουν από τη διαδικασία αυτή.

Σύμβολο	Πιθανότητα	Κωδική λέξη
s_0	0.4	00
s_1	0.2	10
s_2	0.2	11
s_3	0.1	010
s_4	0.1	011

Έτσι, το μέσο μήκος κώδικα θα είναι

$$\begin{aligned}
\overline{L} &= 0.4 \cdot 2 + 0.2 \cdot 2 + 0.2 \cdot 2 + 0.1 \cdot 3 + 0.1 \cdot 3 \\
&= 2.2
\end{aligned}$$

και η εντροπία της πηγής

$$\begin{aligned}
H(\Phi) &= 0.4 \log_2 \left(\frac{1}{0.4} \right) + 0.2 \log_2 \left(\frac{1}{0.2} \right) + 0.2 \log_2 \left(\frac{1}{0.2} \right) + \\
&\quad 0.1 \log_2 \left(\frac{1}{0.1} \right) + 0.1 \log_2 \left(\frac{1}{0.1} \right) \\
&\approx 2.12193
\end{aligned}$$

Από τα παραπάνω αποτελέσματα παρατηρούμε πως το μέσο μήκος κώδικα υπερβαίνει την εντροπία της πηγής μόνο κατά ένα ποσοστό 3.67 %. Ακόμα παρατηρούμε πως το μέσο μήκος κώδικα που υπολογίσαμε ικανοποιεί την ανισότητα (16). ■

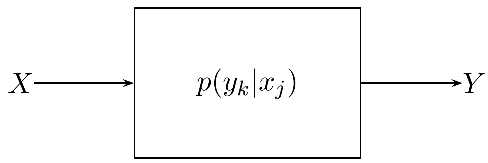
Στο σημείο αυτό θα πρέπει να σημειωθεί ότι η διαδικασία κωδικοποίησης Huffman δεν είναι μοναδική. Συγκεκριμένα, υπάρχουν αρκετά σημεία του αλγορίθμου που μπορούν να υλοποιηθούν με διαφορετικό τρόπο, καταλήγοντας σε άλλη κωδικοποίηση. Ως τέτοια αναφέρουμε:

(α) Αν δύο σύμβολα έχουν την ίδια πιθανότητα εμφάνισης υπάρχουν 2 τρόποι αρχικής διάταξής τους.

(β) Η ανάθεση των ψηφίων "0" και "1"μπορεί να γίνει από πάνω προς τα κάτω ή από κάτω προς τα πάνω.

(γ) Αμφιβολία συναντάμε όταν ένα σύμβολο που προκύπτει από συνδυασμό δύο άλλων έχει την ίδια πιθανότητα με κάποιο ήδη υπάρχον. Στην περίπτωση αυτή μπορεί να τοποθετηθεί όσο ψηλότερα ή όσο χαμηλότερα γίνεται. Γενικά, είναι προτιμότερο να τοποθετείται όσο ψηλότερα γίνεται.

Εφαρμόζοντας τη μια ή την άλλη διαφοροποίηση του αλγορίθμου ενδεχομένως να οδηγηθούμε σε περιπτώσεις όπου ένα σύμβολο έχει διαφορετικό κώδικα ή ακόμα και διαφορετικό μήκος κώδικα (διαφορετικό αριθμό bits). Ωστόσο, σε όλες τις περιπτώσεις το μέσο μήκος κώδικα $\overline{L}$ διατηρείται σταθερό.



Σχήμα 6: Διακριτό κανάλι χωρίς μνήμη

5 Διακριτό κανάλι χωρίς μνήμη

Στις παραγράφους που προηγήθηκαν ασχοληθήκαμε κυρίως με την έννοια της διακριτής πηγής και την κωδικοποίησή της. Στη συνέχεια, αφήνουμε τη διαδικασία παραγωγής δεδομένων και ασχολούμαστε με τη διαδικασία της μετάδοσης δεδομένων και την αξιοπιστία με την οποία γίνεται αυτή η μετάδοση. Ξεκινάμε ορίζοντας την έννοια του διακριτού καναλιού χωρίς μνήμη.

Ένα *διακριτό κανάλι χωρίς μνήμη* είναι ένα στατιστικό μοντέλο το οποίο λαμβάνει είσοδο X δίνει έξοδο Y, μια ενθόρυβη έκδοση της εισόδου X. Τόσο το X όσο και το Y είναι τυχαίες μεταβλητές. Σε κάθε χρονική στιγμή, το κανάλι δέχεται είσοδο X από το σύνολο $\mathcal{X}$ και δίνει έξοδο Y η οποία ανήκει στο σύνολο $\mathcal{Y}$. Ο χαρακτηρισμός "διακριτό" σημαίνει πως τα σύνολα $\mathcal{X}$ και $\mathcal{Y}$ έχουν πεπερασμένο μέγεθος. Ο χαρακτηρισμός "χωρίς μνήμη" σημαίνει πως η τρέχουσα έξοδος εξαρτάτε μόνο από την τρέχουσα είσοδο και όχι από παλαιότερες εισόδους.

Στο σχήμα (6) παρουσιάζουμε ένα διακριτό κανάλι χωρίς μνήμη. Η λειτουργία του καναλιού περιγράφεται από το σύνολο

$$\mathcal{X} = \{x_0, x_1, \ldots, x_{J-1}\} \tag{18}$$

το οποίο ονομάζουμε *αλφάβητο εισόδου*, το σύνολο

$$\mathcal{Y} = \{y_0, y_1, \ldots, y_{K-1}\} \tag{19}$$

το οποίο ονομάζουμε *αλφάβητο εξόδου* και τις *πιθανότητες μετάβασης*

$$p(y_k|x_j) = P(Y = y_k|X = x_j) \qquad \forall \quad j, k \tag{20}$$

όπου φυσικά

$$0 \leq p(y_k|x_j) \leq 1 \qquad \forall \quad j, k$$

Γενικά, τα σύνολα $\mathcal{X}$ και $\mathcal{Y}$ δεν είναι ανάγκη να έχουν το ίδιο μέγεθος. Για παράδειγμα όταν χρησιμοποιείται κωδικοποίηση καναλιού, το σύνολο $\mathcal{Y}$ ενδέχεται να είναι μεγαλύτερο από το σύνολο $\mathcal{X}$ ($K \geq J$). Επίσης, όταν ένα διακριτό κανάλι δίνει το ίδιο σύνολο Y στην έξοδό του για δύο διαφορετικές εισόδους x_l και x_m έχουμε $K \leq J$.

Η πιθανότητα μετάβασης $p(y_k|x_j)$ εκφράζει την πιθανότητα να λάβουμε το σύμβολο y_k στην έξοδο του καναλιού, δοθέντος ότι στην είσοδο του καναλιού έχουμε το σύμβολο x_j. Έτσι, η πιθανότητα $p(y_k|x_j)$ για $k=j$ εκφράζει την πιθανότητα σωστής μετάδοσης του συμβόλου x_j, ενώ οι επιμέρους πιθανότητες για $k \neq j$ εκφράζουν πιθανότητες εσφαλμένης μετάδοσης του συμβόλου x_j. Για τις πιθανότητες αυτές έχουμε ακόμα

$$\sum_{k=0}^{K-1} p(y_k|x_j) = 1 \qquad \forall j$$

Ας υποθέσουμε πως τα σύμβολα εισόδου x_j έχουν πιθανότητες εμφάνισης $p(x_j), j = 0, 1, \ldots, J-1$. Οι πιθανότητες αυτές είναι οι πιθανότητες των γεγονότων $X = x_j$, δηλαδή

$$p(x_j) = P(X = x_j) \qquad j = 0, 1, \ldots, J-1$$

όπου X η τυχαία μεταβλητή που εκφράζει το σύμβολο εισόδου. Θεωρούμε τώρα την τυχαία μεταβλητή Y που εκφράζει την έξοδο της πηγής. Τότε, η από κοινού συνάρτηση πιθανότητας των τυχαίων μεταβλητών X και Y θα δίνεται από τη σχέση:

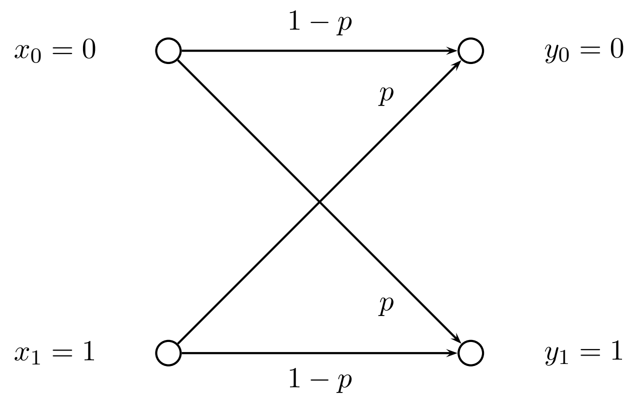
$$\begin{aligned} p(x_j, y_k) &= P(X = x_j, Y = y_k) \\ &= P(Y = y_k | X = x_j) P(X = x_j) \\ &= p(y_k|x_j) p(x_j) \end{aligned} \tag{21}$$

Η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Y μπορεί τότε να εκφραστεί με βάση την από κοινού συνάρτηση πιθανότητας

$$\begin{aligned} p(y_k) &= P(Y = y_k) \\ &= \sum_{j=0}^{J-1} P(Y = y_k | X = x_j) P(X = x_j) \\ &= \sum_{j=0}^{J-1} p(y_k|x_j) p(x_j) \qquad k = 0, 1, \ldots, K-1 \end{aligned}$$

Η συνολική πιθανότητα λάθους P_e ορίζεται ως η πιθανότητα οι τυχαίες μεταβλητές X και Y να έχουν διαφορετικές τιμές:

$$\begin{aligned} P_e &= \sum_{\substack{k=0 \\ k \neq j}}^{K-1} P(Y = y_k) \\ &= \sum_{\substack{k=0 \\ k \neq j}}^{K-1} \sum_{j=0}^{J-1} p(y_k|x_j) p(x_j) \end{aligned} \tag{22}$$



Σχήμα 7: Δυαδικό συμμετρικό κανάλι

Η διαφορά $1 - P_e$ εκφράζει τη μέση πιθανότητα σωστής μετάδόσης.

Οι πιθανότητες $p(x_j)$ ονομάζόνται και εκ-των προτέρων (a-piory) πιθανότητες των συμβόλων εισόδου. Η εξίσωση (21) εκφράζει το γεγονός πως αν έχουμε στη διάθεσή μας τις εκ-των προτέρων πιθανότητες $p(x_j)$ και όλες τις πιθανότητες μετάβασης $p(y_k|x_j)$, τότε μπορούμε να υπολογίσουμε τις πιθανότητες των συμβόλων εξόδου. Στα επόμενα θα υποθέσουμε πως οι πιθανότητες αυτές είναι γνωστές.

5.1 Δυαδικό συμμετρικό κανάλι

Το *δυαδικό συμμετρικό κανάλι* (Binary Symmetric Channel, BSC) είναι ένα διακριτό κανάλι χωρίς μνήμη με ιδιαίτερη θεωρητική και πρακτική σημασία. Το αλφάβητο εισόδου για το κανάλι αυτό είναι το

$$\{x_0 = 0, x_1 = 1\}$$

το αλφάβητο εξόδου του είναι το

$$\{y_o = 0, y_1 = 1\}$$

και οι πιθανότητες μετάβασης δίνονται από τον τύπο

$$p(y_k|x_j) = \begin{cases} 1-p & k = j \\ \\ p & k \neq j \end{cases}$$

για $k, j = 0, 1$. Το δυαδικό συμμετρικό κανάλι μπορεί να αναπαρασταθεί από ένα *διάγραμμα πιθανοτήτων μετάβασης* όπως φαίνεται στο σχήμα (7). Το κανάλι ονομάζεται συμμετρικό επειδή

$$p(y_1|x_0) = p(y_0|x_1)$$

6 Η έννοια της αμοιβαίας πληροφορίας

Όπως είδαμε στα προηγούμενα, για μια τυχαία μεταβλητή X η οποία λαμβάνει τιμές από ένα σύνολο $\mathcal{X}$ ορίζεται η έννοια της εντροπίας της $H(\mathcal{X})$ (στα προηγούμενα εξετάσαμε την περίπτωση όπου η τυχαία μεταβλητή X αποτελούσε την έξοδο μιας διακριτής πηγής χωρίς μνήμη). Η τιμή της εντροπίας εκφράζει την αβεβαιότητα την οποία έχουμε για την επόμενη τιμή που θα δώσει η τυχαία αυτή μεταβλητή. Στην περίπτωση όπου οι τιμές της τυχαίας μεταβλητής X περάσουν μέσα από ένα διακριτό κανάλι χωρίς μνήμη τότε στην έξοδο του καναλιού παρατηρούμε την τυχαία μεταβλητή Y. Το ερώτημα που τίθεται είναι, αν γνωρίζουμε την τιμή της τυχαίας μεταβλητής Y τότε πρέπει να έχουμε και κάποια πληροφορία για την τιμή της τυχαίας μεταβλητής X από την οποία προήλθε. Με άλλα λόγια, η γνώση της τιμής της τυχαίας μεταβλητής Y μειώνει (πιθανώς) την αβεβαιότητά μας για την τιμή της τυχαίας μεταβλητής X. Πιο συγκεκριμένα, αν γνωρίζουμε την τιμή της τυχαίας μεταβλητής Y, η αβεβαιότητά μας για την τιμή της τυχαίας μεταβλητής X είναι

$$H(\mathcal{X}|Y=y_k)=\sum_{j=0}^{J-1}p(x_j|y_k)\log_2\left(\frac{1}{p(x_j|y_k)}\right) \tag{23}$$

και ονομάζεται *υπό συνθήκη εντροπία της X, δοθέντος ότι $Y=y_k$*. Η έκφραση αυτή αποτελεί μια τυχαία μεταβλητή η οποία λαμβάνει τις τιμές $H(\mathcal{X}|Y=y_0)$, $H(\mathcal{X}|Y=y_1),\ldots,H(\mathcal{X}|Y=y_{K-1})$ με πιθανότητες $p(y_0),p(y_1),\ldots,p(y_{K-1})$. Έτσι, μπορούμε να υπολογίσουμε τη μέση τιμή της τυχαίας αυτής μεταβλητής ως προς τις πιθανότητες αυτές:

$$\begin{aligned}
H(\mathcal{X}|\mathcal{Y}) &= \sum_{k=0}^{K-1}p(y_k)H(\mathcal{X}|Y=y_k)\\
&= \sum_{k=0}^{K-1}\sum_{j=0}^{J-1}p(y_k)p(x_j|y_k)\log_2\left(\frac{1}{p(x_j|y_k)}\right)\\
&= \sum_{k=0}^{K-1}\sum_{j=0}^{J-1}p(x_j,y_k)\log_2\left(\frac{1}{p(x_j|y_k)}\right)
\end{aligned} \tag{24}$$

Η τιμή αυτή ονομάζεται *υπό συνθήκη εντροπία* και εκφράζει *τη μέση ποσότητα αβεβαιότητας την οποία έχουμε για την είσοδο του καναλιού όταν γνωρίζουμε την έξοδο του καναλιού.*

Όπως είδαμε, η ποσότητα $H(\mathcal{X})$ εκφράζει την αβεβαιότητα που έχουμε για την είσοδο του καναλιού χωρίς να γνωρίζουμε την έξοδό του. Επίσης, η ποσότητα $H(\mathcal{X}|\mathcal{Y})$ εκφράζει την αβεβαιότητα που έχουμε για την είσοδο του καναλιού όταν γνωρίζουμε την έξοδο του καναλιού. Έτσι, η ποσότητα $H(\mathcal{X})-H(\mathcal{X}|\mathcal{Y})$ εκφράζει τη *μείωση της αβεβαιότητάς μας για την είσοδο του καναλιού μετά από την παρατήρηση της εξόδου του καναλιού.* Την σημαντική αυτή ποσότητα την ονομάζουμε *αμοιβαία πληροφορία*:

$$I(\mathcal{X};\mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}) \tag{25}$$

6.1 Ιδιότητες της αμοιβαίας πληροφορίας

Στην παράγραφο αυτή παραθέτουμε χωρίς απόδειξη ορισμένες μαθηματικές ιδιότητες της αμοιβαίας πληροφορίας:

1. Η αμοιβαία πληροφορία είναι συμμετρική με την έννοια ότι

$$I(\mathcal{X};\mathcal{Y}) = I(\mathcal{Y};\mathcal{X})$$

όπως είδαμε η αμοιβαία πληροφορία $I(\mathcal{X};\mathcal{Y})$ εκφράζει την μείωση της αβεβαιότητάς μας για την είσοδο του καναλιού μετά από την παρατήρηση της εξόδου του καναλιού ενώ η ποσότητα $I(\mathcal{Y};\mathcal{X})$ εκφράζει την μείωση της αβεβαιότητάς μας για την έξοδο του καναλιού μετά από την εφαρμογή στην είσοδο του καναλιού μιας γνωστής εισόδου.

2. Η αμοιβαία πληροφορία είναι μη αρνητική

$$I(\mathcal{X};\mathcal{Y}) \geq 0 \tag{26}$$

3. Η αμοιβαία πληροφορία μπορεί να εκφραστεί και ώς

$$I(\mathcal{X};\mathcal{Y}) = H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X}) \tag{27}$$

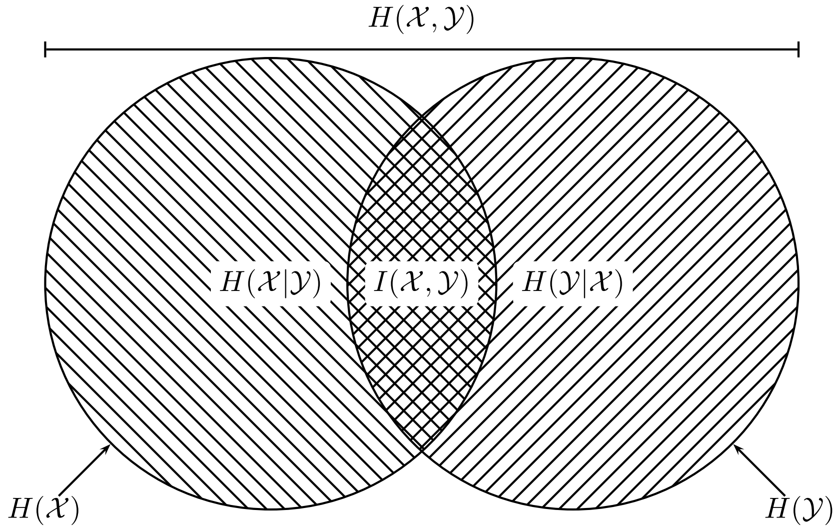
4. Η αμοιβαία πληροφορία σχετίζεται με την *από κοινού εντροπία* της εισόδου και της εξόδου της πηγής

$$H(\mathcal{X},\mathcal{Y}) = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1} p(x_j,y_k)\log_2\left(\frac{1}{p(x_j,y_k)}\right) \tag{28}$$

με τη σχέση

$$I(\mathcal{X};\mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X},\mathcal{Y}) \tag{29}$$

Στο σημείο αυτό δίνουμε ένα διάγραμμα το οποίο απεικονίζει τις σχέσεις των εννοιών που περιγράψαμε στην παράγραφο αυτή και παρουσιάζεται στο σχήμα (8). Η εντροπία της εισόδου X του καναλιού αναπαριστάνεται από τον κύκλο στα αριστερά του σχήματος. Η εντροπία της εξόδου Y του καναλιού αναπαριστάνεται από τον κύκλο στα δεξιά του σχήματος. Η επικάλυψη των δυο κύκλων αναπαριστά την αμοιβαία πληροφορία του καναλιού.



Σχήμα 8: Απεικόνιση των σχέσεων για τις διάφορες παραμέτρους ενός καναλιού

7 Χωρητικότητα καναλιού

Ας θεωρήσουμε ένα διακριτό κανάλι χωρίς μνήμη με αλφάβητο εισόδου $\mathcal{X}$, αλφάβητο εξόδου $\mathcal{Y}$ και πιθανότητες μετάβασης $p(y_k|x_j)$. Η αμοιβαία πληροφορία $I(\mathcal{X},\mathcal{Y})$ του καναλιού μπορεί να γραφεί στη μορφή

$$I(\mathcal{X},\mathcal{Y}) = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1} p(x_j, y_k)\log_2\left(\frac{p(y_k|x_j)}{p(y_k)}\right) \tag{30}$$

όπου

$$p(x_j, y_k) = p(y_k|x_j)p(x_j)$$

και

$$p(y_k) = \sum_{j=0}^{J-1} p(y_k|x_j)p(x_j)$$

Από τις παραπάνω εξισώσεις διαπιστώνουμε πως για να υπολογίσουμε την τιμή της αμοιβαίας πληροφορίας είναι απαραίτητο να γνωρίζουμε τις τιμές των εκ-των προτέρων πιθανοτήτων $p(x_j), j = 0, 1, \ldots, J-1$. Με άλλα λόγια η αμοιβαία πληροφορία είναι μια ποσότητα η οποία δεν εξαρτάται μόνο από το κανάλι αλλά και από την κατανομή πιθανοτήτων του αλφαβήτου εισόδου.

Η κατανομή των πιθανοτήτων εισόδου είναι ανεξάρτητη από το κανάλι. Για το λόγο αυτό, μπορούμε να μεγιστοποιήσουμε την αμοιβαία πληροφορία του καναλιού ως προς όλες τις δυνατές κατανομές πιθανοτήτων $p(x_j)$. Ορίζουμε την *χωρητικότητα ενός διακριτού καναλιού χωρίς μνήμη ως τη μέγιστη τιμή της αμοιβαίας πληροφορίας* $I(\mathcal{X},\mathcal{Y})$ *όπου η μεγιστοποίηση γίνεται ως προς* *όλες τις δυνατές κατανομές εισόδου του αλφαβήτου $\mathcal{X}$.* Συχνά, συμβολίζουμε τη χωρητικότητα ενός καναλιού με το σύμβολο C:

$$C = \max_{\{p(x_j)\}} I(\mathcal{X}, \mathcal{Y}) \tag{31}$$

Η χωρητικότητα του καναλιού έχει μονάδες bits ανά χρήση του καναλιού (π.χ. περίοδο συμβόλου).

Η χωρητικότητα ενός καναλιού είναι συνάρτηση μόνο των πιθανοτήτων μετάβασης του καναλιού $p(y_k|x_j)$ οι οποίες ορίζουν το κανάλι. Ο υπολογισμός της απαιτεί μεγιστοποίηση μιας συνάρτησης (της αμοιβαίας πληροφορίας) J μεταβλητών $p(x_0), p(x_1), \ldots, p(x_{J-1})$ με τους επιπλέον περιορισμούς

$$p(x_j) \geq 0$$

και

$$\sum_{j=0}^{J-1} p(x_j) = 1$$

Πρόκειται δηλαδή για ένα πρόβλημα μεγιστοποίησης υπό περιορισμούς (constrained maximization) το οποίο μπορεί να είναι πολύ δύσκολο να λυθεί.

ΠΑΡΑΔΕΙΓΜΑ 4

Να υπολογιστεί η χωρητικότητα ενός δυαδικού συμμετρικού καναλιού το οποίο έχει πιθανότητα εσφαλμένης μετάβασης p.

Λύση:

Αρχικά, μπορούμε να αποδείξουμε πως η αμοιβαία πληροφορία ενός δυαδικού συμμετρικού καναλιού γίνεται μέγιστη για την κατανομή εισόδου

$$p_0 = p_1 = \frac{1}{2}$$

Έτσι είναι

$$C = \max_{\{p(x_j)\}} I(\mathcal{X}, \mathcal{Y}) = I(\mathcal{X}, \mathcal{Y})|_{p_0 = p_1 = \frac{1}{2}}$$

Από το σχήμα (7) έχουμε

$$p(y_0|x_1) = p(y_1|x_0) = p$$

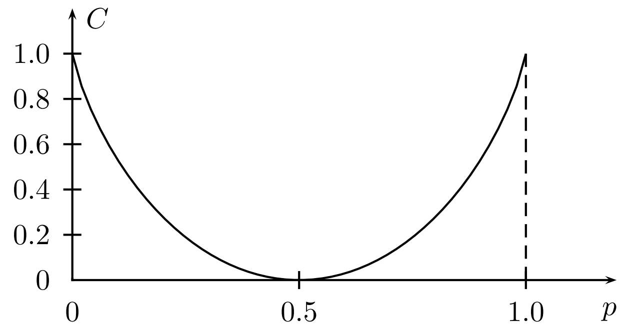
και

$$p(y_0|x_0) = p(y_1|x_1) = 1 - p$$

Έτσι, από την εξίσωση (30) για $J = K = 2$ και $p(x_0) = p(x_1)$ έχουμε

$$C = 1 + p\log_2 p + (1-p)\log_2(1-p) = 1 - \mathcal{H}(p)$$

όπου $\mathcal{H}(p)$ είναι η συνάρτηση εντροπίας που ορίσαμε στη σχέση (11). Στο σχήμα (7) δίνουμε τη γραφική παράσταση της χωρητικότητας C ώς προς την παράμετρο p. Από το σχήμα αυτό παρατηρούμε τα εξής:



Σχήμα 9: Γραφική παράσταση της χωρητικότητας ενός δυαδικού συμμετρικού καναλιού

1. Στην περίπτωση όπου $p = 0$, δηλαδή δεν γίνονται εσφαλμένες μεταβάσεις, η χωρητικότητα λαμβάνει τη μέγιστη τιμή της (1 bit ανά περίοδο συμβόλου), όση είναι και η πληροφορία σε κάθε είσοδο του καναλιού.

2. Στην περίπτωση όπου $p = 1/2$, δηλαδή όταν το κανάλι αποφασίζει το σύμβολο εξόδου ανεξάρτητα από την είσοδό του, η χωρητικότητά του λαμβάνει την ελάχιστη τιμή της ίση με 0 bits/περίοδο συμβόλου.

3. Στην περίπτωση όπου $p = 1$, δηλαδή όταν όλες οι μεταβάσεις γίνονται εσφαλμένα η χωρητικότητα λαμβάνει και πάλι τη μέγιστη τιμή της. Το γεγονός αυτό μπορεί να δικαιολογηθεί επειδή το κανάλι δεν αλλοιώνει καθόλου την πληροφορία στην είσοδό του, αντίθετα μεταβιβάζει αξιόπιστα το συμπλήρωμα της εισόδου του.

■