ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ &
ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΨΗΦΙΑΚΕΣ ΕΠΙΚΟΙΝΩΝΙΕΣ Ι

ΣΤΥΛΙΑΝΑΚΗΣ ΒΑΣΙΛΗΣ

ΠΑΤΡΑ 2015

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
Τμήμα Ηλεκτρολόγων Μηχανικών
& Τεχνολογίας Υπολογιστών

ΨΗΦΙΑΚΕΣ ΕΠΙΚΟΙΝΩΝΙΕΣ Ι

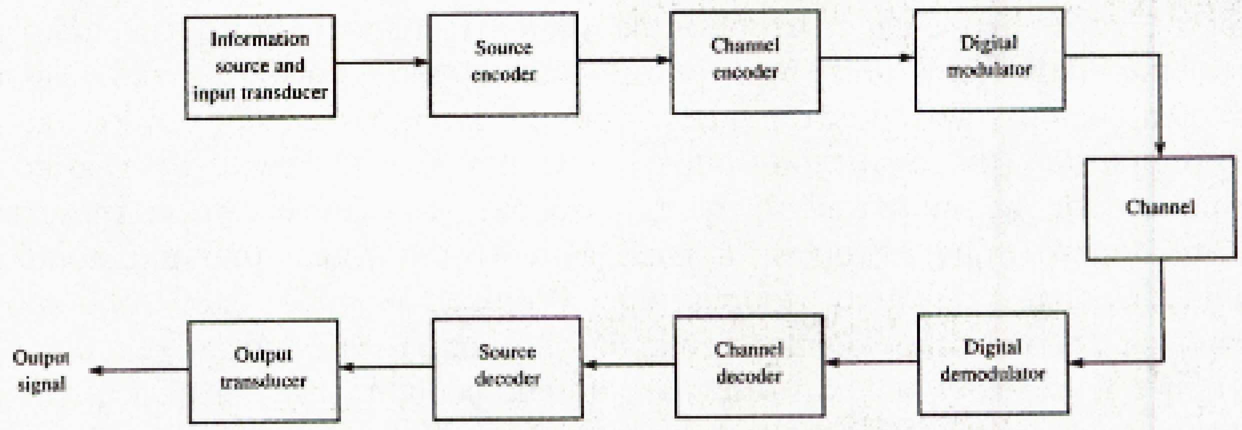
ΣΤΥΛΙΑΝΑΚΗΣ ΒΑΣΙΛΗΣ

1 ΓΕΝΙΚΑ

Το αντικείμενο των ψηφιακών επικοινωνιών αφορά τη μετάδοση της πληροφορίας σε ψηφιακή μορφή, από μία πηγή που δημιουργεί την πληροφορία προς έναν ή περισσότερους προορισμούς. Κατά την ανάλυση και την σχεδίαση των επικοινωνιακών συστημάτων ιδιαίτερη σημασία έχουν τα χαρακτηριστικά των φυσικών καναλιών μέσω των οποίων μεταδίδεται η πληροφορία. Τα χαρακτηριστικά του καναλιού επηρεάζουν σε γενικές γραμμές την σχεδίαση των βασικών δομικών στοιχείων που απαρτίζουν το επικοινωνιακό σύστημα. Παρακάτω, θα περιγράψουμε τα στοιχεία ενός επικοινωνιακού συστήματος και τις λειτουργίες τους.

1-1 ΤΑ ΣΤΟΙΧΕΙΑ ΕΝΟΣ ΣΥΣΤΗΜΑΤΟΣ ΨΗΦΙΑΚΩΝ ΕΠΙΚΟΙΝΩΝΙΩΝ

Το Σχήμα 1-1-1 παρουσιάζει το λειτουργικό διάγραμμα και τα βασικά στοιχεία ενός συστήματος ψηφιακών επικοινωνιών. Η έξοδος της πηγής μπορεί να είναι είτε ένα αναλογικό σήμα, όπως ένα σήμα audio ή ένα σήμα video, είτε ένα ψηφιακό σήμα, όπως η έξοδος μιας τηλετυπικής μηχανής, η οποία είναι διακριτή ως προς τον χρόνο και διαθέτει έναν πεπερασμένο αριθμό από χαρακτήρες εξόδου. Σε ένα σύστημα ψηφιακών επικοινωνιών, τα μηνύματα που παράγονται από την πηγή μετατρέπονται σε μία ακολουθία δυαδικών ψηφίων. Για το βέλτιστο αποτέλεσμα, πρέπει να προσπαθήσουμε να αντιπροσωπεύσουμε την έξοδο της πηγής (το μήνυμα) με όσο το δυνατόν λιγότερα δυαδικά ψηφία. Με άλλα λόγια, αναζητούμε μία αποτελεσματική αναπαράσταση της εξόδου της πηγής η οποία να επιφέρει μικρό ή μηδενικό πλεονασμό. Η διαδικασία της αποτελεσματικής μετατροπής της εξόδου μίας είτε αναλογικής είτε ψηφιακής πηγής σε μία ακολουθία δυαδικών ψηφίων ονομάζεται *κωδικοποίηση πηγής* (*source encoding*) ή *συμπίεση δεδομένων* (*data compression*).



ΣΧΗΜΑ 1-1-1 Τα βασικά στοιχεία ενός φηφιακού συστήματος επικοινωνίας.

Η ακολουθία των δυαδικών ψηφίων από τον κωδικοποιητή πηγής, η οποία ονομάζεται *πληροφοριακή ακολουθία* (*information sequence*), περνά στον *κωδικοποιητή καναλιού* (*channel encoder*). Ο σκοπός του κωδικοποιητή καναλιού είναι να εισάγει, με ελεγχόμενο τρόπο, κάποιον πλεονασμό στην δυαδική πληροφοριακή ακολουθία, ο οποίος να μπορεί να χρησιμοποιηθεί στον δέκτη για να αντιμετωπισθούν οι επιπτώσεις του θορύβου και των παρεμβολών που παρουσιάζονται κατά τη μετάδοση του σήματος μέσω του καναλιού. Κατά τον τρόπο αυτό, ο προστιθέμενος πλεονασμός αυξάνει την αξιοπιστία των λαμβανομένων δεδομένων και βελτιώνει την πιστότητα του λαμβανομένου σήματος. Στην πράξη, ο πλεονασμός στην πληροφοριακή ακολουθία βοηθά τον δέκτη να αποκωδικοποιήσει την επιθυμητή πληροφοριακή ακολουθία. Για παράδειγμα, μία (απλή) μορφή κωδικοποίησης της δυαδικής πληροφοριακής ακολουθίας είναι το να επαναληφθεί απλώς κάθε δυαδικό ψηφίο m φορές, όπου m ένας θετικός ακέραιος. Περισσότερο προηγμένες (μη τετριμμένες) μορφές κωδικοποίησης παίρνουν k πληροφοριακά bits κάθε φορά και αντιστοιχίζουν κάθε ακολουθία των k bits σε μία μοναδική ακολουθία των n bits που ονομάζεται *κωδική λέξη* (*code word*). Η ποσότητα του πλεονασμού που εισάγεται με την κωδικοποίηση των δεδομένων κατ' αυτόν τον τρόπο μετριέται με τον λόγο n/k. Ο αντίστροφος του λόγου αυτού, δηλαδή το k/n, ονομάζεται *ρυθμός του κώδικα* (*code rate*).

Η δυαδική ακολουθία που παρουσιάζεται στην έξοδο του κωδικοποιητή καναλιού περνά στον *ψηφιακό διαμορφωτή* (*digital modulator*), ο οποίος παίζει τον ρόλο της διεπαφής (interface) με το κανάλι επικοινωνίας. Επειδή αυτό που μπορούν να κάνουν σχεδόν όλα τα κανάλια επικοινωνίας που συναντιούνται στην πράξη είναι να μεταδίδουν ηλεκτρικά σήματα (κυματομορφές), ο κύριος σκοπός του ψηφιακού διαμορφωτή είναι να αντιστοιχίζει την δυαδική πληροφοριακή ακολουθία με κυματομορφές σημάτων. Για να εμβαθύνουμε στο σημείο αυτό, ας υποθέσουμε ότι η κωδικοποιημένη πληροφοριακή ακολουθία μεταδίδεται ένα bit κάθε φορά με κάποιο ομοιογενή ρυθμό R bits/s. Ο ψηφιακός διαμορφωτής μπορεί απλώς να αντιστοιχίσει το δυαδικό ψηφίο 0 σε μία κυματομορφή $s_0(t)$ και το δυαδικό ψηφίο 1 σε μία κυματομορφή $s_1(t)$. Ο τρόπος αυτός, στον οποίο κάθε bit από τον κωδικοποιητή καναλιού μεταδίδεται ξεχωριστά, ονομάζεται *δυαδική διαμόρφωση* (*binary modulation*). Εναλλακτικά, ο διαμορφωτής μπορεί να μεταδώσει b κωδικοποιημένα bits κάθε φορά χρησιμοποιώντας $M = 2^b$ ξεχωριστές κυματομορφές $s_i(t)$, $i = 0, 1, \ldots, M-1$, μία κυματομορφή για κάθε μία από τις 2^b πιθανές ακολουθίες των b bits. Αυτή η μέθοδος ονομάζεται *M-αδική διαμόρφωση* (*M-ary modulation*, $M>2$). Σημειώστε ότι κάθε b/R δευτερόλεπτα εισέρχεται στον διαμορφωτή μία νέα ακολουθία των b bits. Συνεπώς, όταν ο ρυθμός των bits του καναλιού R είναι σταθερός, η ποσότητα του χρόνου που είναι διαθέσιμη για να μεταδοθεί μία από τις M κυματομορφές που αντιστοιχούν σε μία ακολουθία των b bits είναι b φορές επί την χρονική περίοδο (ποσότητα χρόνου) ενός συστήματος που χρησιμοποιεί δυαδική διαμόρφωση.

Το *κανάλι επικοινωνίας* (*communication channel*) είναι το φυσικό μέσον που χρησιμοποιείται για να σταλεί η πληροφορία από τον πομπό στον δέκτη. Στην ασύρματη μετάδοση, το κανάλι είναι η ατμόσφαιρα (ο κενός χώρος). Από την άλλη πλευρά, τα τηλεφωνικά κανάλια συνήθως χρησιμοποιούν μία ποικιλία από φυσικά μέσα, συμπεριλαμβανομένων των ενσύρματων γραμμών, των καλωδίων των οπτικών ινών, και των ασύρματων (μικροκυματικό ράδιο). Όποιο και αν είναι το φυσικό μέσο που χρησιμοποιείται για την μετάδοση της πληροφορίας, το ουσιώδες χαρακτηριστικό είναι ότι το μεταδιδόμενο σήμα αλλοιώνεται κατά έναν τυχαίο τρόπο από μία ποικιλία πιθανών μηχανισμών, όπως είναι ο προσθετικός θερμικός θόρυβος που δημιουργείται από ηλεκτρονικές συσκευές, ο θόρυβος που προέρχεται από ανθρώπους, π.χ., ο θόρυβος από την εκκίνηση των αυτοκινήτων, και ο ατμοσφαιρικός θόρυβος, π.χ., οι ηλεκτρικές εκκενώσεις των κεραυνών κατά την διάρκεια των καταιγίδων.

Στο λαμβάνον άκρο ενός συστήματος ψηφιακών επικοινωνιών, ο *ψηφιακός αποδιαμορφωτής* (*digital demodulator*) επεξεργάζεται την αλλοιωμένη από το κανάλι μεταδοθείσα κυματομορφή και απλοποιεί τις κυματομορφές σε μία ακολουθία από αριθμούς που αναπαριστούν εκτιμήσεις ως προς τα μεταδοθένα σύμβολα των δεδομένων (δυαδικών η M-αδικών). Αυτή η ακολουθία των αριθμών περνά στον αποκωδικοποιητή καναλιού, ο οποίος επιχειρεί να ανακατασκευάσει την αρχική πληροφοριακή ακολουθία γνωρίζοντας τον κώδικα που χρησιμοποιήθηκε από τον κωδικοποιητή καναλιού και τον πλεονασμό που περιέχεται στα λαμβανόμενα δεδομένα.

Ένα μέτρο της απόδοσης του αποδιαμορφωτή και του αποκωδικοποιητή είναι η συχνότητα με την οποία εμφανίζονται λάθη στην αποκωδικοποιημένη ακολουθία. Πιο συγκεκριμένα, η

μέση πιθανότητα ενός εσφαλμένου bit στην έξοδο του αποκωδικοποιητή είναι ένα μέτρο της απόδοσης του συνδυασμού αποδιαμορφωτή-αποκωδικοποιητή. Εν γένει, η πιθανότητα λάθους είναι μία συνάρτηση των χαρακτηριστικών του κώδικα, του τύπου των κυματομορφών που χρησιμοποιούνται για την μετάδοση της πληροφορίας μέσω του καναλιού, της ισχύος του πομπού, των χαρακτηριστικών του καναλιού, ήτοι της ποσότητας του θορύβου, της φύσης των παρεμβολών, κλπ, και της μεθόδου της αποδιαμόρφωσης και της αποκωδικοποίησης. Τα στοιχεία αυτά και οι επιπτώσεις τους στην απόδοση θα συζητηθούν λεπτομερειακά σε ακόλουθα κεφάλαια.
Και τέλος, όταν είναι επιθυμητή μία αναλογική έξοδος, ο αποκωδικοποιητής πηγής δέχεται την ακολουθία εξόδου από τον αποκωδικοποιητή καναλιού και, με την γνώση της μεθόδου κωδικοποίησης πηγής που χρησιμοποιείται, προσπαθεί να ανακατασκευάσει το αρχικό σήμα από την πηγή. Λόγω των λαθών κατά την διαδικασία αποκωδικοποίησης καναλιού και της πιθανής παραμόρφωσης που εισάγεται από τον κωδικοποιητή πηγής και, πιθανώς, από τον αποκωδικοποιητή πηγής, το σήμα στην έξοδο του αποκωδικοποιητή πηγής είναι μία προσέγγιση της αρχικής εξόδου της πηγής. Η διαφορά, ή κάποια συνάρτηση της διαφοράς, μεταξύ του αρχικού σήματος και του ανακατασκευασμένου σήματος είναι ένα μέτρο της παραμόρφωσης που έχει εισαχθεί από το ψηφιακό σύστημα επικοινωνίας.

2. Η ΚΩΔΙΚΟΠΟΙΗΣΗ ΠΗΓΗΣ

Ανεξαρτήτως του κατά πόσον η πηγή είναι αναλογική ή διακριτή, ένα σύστημα ψηφιακών επικοινωνιών σχεδιάζεται έτσι ώστε να μεταδίδει τις πληροφορίες σε ψηφιακή μορφή. Κατά συνέπεια, η έξοδος της πηγής πρέπει να μετατραπεί σε μία μορφή η οποία να μπορεί να μεταδοθεί ψηφιακά. Η μετατροπή αυτή της εξόδου της πηγής σε ψηφιακή μορφή γενικά πραγματοποιείται στον κωδικοποιητή πηγής (source encoder), του οποίου η έξοδος μπορεί να θεωρηθεί ότι είναι μία ακολουθία δυαδικών ψηφίων.

2-1 ΤΑ ΜΑΘΗΜΑΤΙΚΑ ΜΟΝΤΕΛΑ ΤΩΝ ΠΗΓΩΝ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

Οποιαδήποτε πληροφοριακή πηγή παράγει μία έξοδο που είναι τυχαία, δηλαδή, η έξοδος της πηγής χαρακτηρίζεται με την βοήθεια στατιστικών όρων. Αλλιώς, εάν η έξοδος της πηγής ήταν επακριβώς γνωστή, δεν θα υπήρχε καμία ανάγκη να μεταδοθεί. Στην ενότητα αυτή, θα μελετήσουμε τόσο τις διακριτές όσο και τις αναλογικές πηγές πληροφορίας και θα διατυπώσουμε μαθηματικά μοντέλα για κάθε τύπο πηγής.

Ο απλούστερος τύπος διακριτής πηγής είναι αυτός που εκπέμπει μία ακολουθία γραμμάτων που επιλέγονται από ένα πεπερασμένο αλφάβητο. Για παράδειγμα, μία *δυαδική πηγή* (*binary source*) εκπέμπει μία δυαδική ακολουθία της μορφής 100101110…, με το αλφάβητο να αποτελείται από τα δύο γράμματα {0, 1}. Γενικότερα, μία διακριτή πληροφοριακή πηγή με ένα αλφάβητο από L πιθανά γράμματα, ας πούμε τα $\{x_1, x_2, \ldots, x_L\}$, εκπέμπει μία ακολουθία από γράμματα που έχουν επιλεγεί από αυτό το αλφάβητο.

Για να κατασκευάσουμε ένα μαθηματικό μοντέλο μίας διακριτής πηγής, υποθέτουμε ότι κάθε γράμμα του αλφαβήτου $\{x_1, x_2, \ldots, x_L\}$ έχει μία δεδομένη πιθανότητα εμφάνισης p_k. Δηλαδή,

$$p_k = P(X = x_k), \qquad 1 \le k \le L$$

όπου

$$\sum_{k=1}^{L} p_k = 1$$

Θα ασχοληθούμε με δύο μαθηματικά μοντέλα διακριτών πηγών. Στο πρώτο, θα υποθέσουμε ότι η ακολουθία εξόδου της πηγής είναι στατιστικώς ανεξάρτητη. Αυτό σημαίνει ότι το τρέχον γράμμα εξόδου είναι στατιστικώς ανεξάρτητο από όλες τις παρελθούσες και τις μελλοντικές εξόδους. Μία πηγή της οποίας η έξοδος ικανοποιεί την συνθήκη της στατιστικής ανεξαρτησίας μεταξύ των γραμμάτων που απαρτίζουν κάθε ακολουθία εξόδου λέγεται ότι είναι *χωρίς μνήμη* (*memoryless*). Μία πηγή του τύπου αυτού αποκαλείται *Διακριτή Πηγή Χωρίς Μνήμη* ή, εν συντομία, ΔΠΧΜ (*discrete memoryless source* -DMS).

Εάν η έξοδος της διακριτής πηγής είναι στατιστικώς εξαρτημένη, όπως συμβαίνει, για παράδειγμα, στο Αγγλικό κείμενο, μπορούμε να κατασκευάσουμε ένα μαθηματικό μοντέλο που να βασίζεται στην στατιστική στασιμότητα. Εξ ορισμού, μία διακριτή πηγή λέγεται ότι είναι *στάσιμη* (*stationary*) εάν οι από κοινού πιθανότητες δύο ακολουθιών μήκους n, ας πούμε, της a_1, a_2, …, a_n και της a_{1+m}, a_{2+m}, …, a_{n+m}, είναι ταυτόσημες για κάθε $n \ge 1$ και για όλες τις ολισθήσεις m. Με άλλα λόγια, η από κοινού πιθανότητα κάθε ακολουθίας εξόδων πηγής αυθαίρετου μήκους παραμένει αμετάβλητη υπό οποιαδήποτε ολίσθηση στον άξονα του χρόνου.

Μία αναλογική πηγή έχει μία κυματομορφή εξόδου $x(t)$ που είναι μία δειγματική συνάρτηση μίας στοχαστικής διαδικασίας $X(t)$. Υποθέτουμε ότι η $X(t)$ είναι μία στάσιμη στοχαστική διαδικασία με συνάρτηση αυτοσυσχέτισης $\varphi_{xx}(\tau)$ και φασματική πυκνότητα ισχύος $\Phi_{xx}(f)$. Όταν η $X(t)$ είναι μία περιορισμένου εύρους ζώνης στοχαστική διαδικασία, ήτοι, $\Phi_{xx}(f) = 0$ για $|f| \geq W$, μπορεί να χρησιμοποιηθεί το θεώρημα δειγματοληψίας για να αναπαρασταθεί η $X(t)$ ως

$$X(t) = \sum_{n=-\infty}^{\infty} X\left(\frac{n}{2W}\right) \frac{\sin\left[2\pi W\left(t - \frac{n}{2W}\right)\right]}{2\pi W\left(t - \frac{n}{2W}\right)} \tag{2-1-1}$$

όπου $\{X(n/2W)\}$ υποδηλώνει τα δείγματα της διαδικασίας $X(t)$ που έχουν ληφθεί με τον ρυθμό δειγματοληψίας Nyquist των $f_s = 2W$ δειγμάτων / sec. Έτσι, εφαρμόζοντας το θεώρημα δειγματοληψίας, μπορούμε να μετατρέψουμε την έξοδο μίας αναλογικής πηγής σε μία ισοδύναμη πηγή διακριτού χρόνου. Τότε, η έξοδος της πηγής χαρακτηρίζεται στατιστικώς από την από κοινού σ.π.π. $p(x_1, x_2, \ldots, x_m)$ για κάθε $m \geq 0$, όπου $X_n = X(n/2W)$, $1 \leq n \leq m$, είναι οι τυχαίες μεταβλητές που αντιστοιχούν στα δείγματα της $X(t)$.
Σημειώνουμε ότι τα δείγματα εξόδου $\{X(n/2W)\}$ από τις στάσιμες πηγές είναι γενικά συνεχή και δεν μπορούν συνεπώς να αναπαρασταθούν σε ψηφιακή μορφή χωρίς κάποια απώλεια της ακρίβειας. Για παράδειγμα, μπορούμε να κβαντίσουμε κάθε δείγμα μετατρέποντάς το σε ένα σύνολο από διακριτές τιμές, αλλά η διαδικασία κβαντισμού έχει ως αποτέλεσμα κάποια απώλεια ακρίβειας, και, κατά συνέπεια, το αρχικό σήμα δεν μπορεί να ανασυσταθεί επακριβώς από τις κβαντισμένες τιμές των δειγμάτων. Στην συνέχεια του κεφαλαίου, θα μελετήσουμε την παραμόρφωση που προκαλεί ο κβαντισμός των δειγμάτων μίας αναλογικής πηγής.

2-2 ΕΝΑ ΛΟΓΑΡΙΘΜΙΚΟ ΜΕΤΡΟ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

Για να αναπτύξουμε ένα κατάλληλο μέτρο της πληροφορίας, ας θεωρήσουμε δύο διακριτές τυχαίες μεταβλητές με πιθανές τιμές x_i, $i = 1, 2, \ldots, n$, και y_j, $j = 1, 2, \ldots, m$, αντιστοίχως. Υποθέστε ότι παρατηρούμε κάποια έξοδο $Y = y_j$, και επιθυμούμε να καθορίσουμε με κάποιο ποσοτικό τρόπο την ποσότητα της πληροφορίας που παρέχει η εμφάνιση του γεγονότος $Y = y_j$ περί το γεγονός $X = x_i$, $i = 1, 2, \ldots, n$. Παρατηρούμε ότι όταν τα X και Y είναι στατιστικώς ανεξάρτητα, η εμφάνιση του $Y = y_j$ δεν παρέχει καμία πληροφορία περί την εμφάνιση του $X = x_i$. Από την άλλη πλευρά, όταν τα X και Y είναι στατιστικώς πλήρως εξαρτημένα, έτσι ώστε η εμφάνιση του $Y = y_j$ να καθορίζει την εμφάνιση του $X = x_i$, το πληροφοριακό περιεχόμενο είναι απλώς αυτό που παρέχεται από το γεγονός $X = x_i$. Ένα βολικό μέτρο που ικανοποιεί τις συνθήκες αυτές είναι ο λογάριθμος του λόγου της υπό συνθήκη πιθανότητας

$$P(X = x_i \mid Y = y_j) \equiv P(x_i \mid y_j)$$

δια της πιθανότητας

$$P(X = x_i) \equiv P(x_i)$$

Δηλαδή, το πληροφοριακό περιεχόμενο που παρέχεται από την εμφάνιση του γεγονότος $Y = y_j$ περί το γεγονός $X = x_i$ ορίζεται ως

$$I(x_i; y_j) = \log \frac{P(x_i \mid y_j)}{P(x_i)} \tag{2-2-1}$$

Το $I(x_i; y_j)$ αποκαλείται *αμοιβαία πληροφορία* (*mutual information*) μεταξύ των x_i και y_j.
Οι μονάδες του $I(x_i ; y_j)$ καθορίζονται από την βάση του λογάριθμου, η οποία συνήθως επιλέγεται είτε το 2 είτε το e. Όταν η βάση του λογάριθμου είναι το 2, οι μονάδες του $I(x_i; y_j)$ είναι τα bits, και όταν η βάση είναι το e, οι μονάδες του $I(x_i; y_j)$ ονομάζονται *nats* (natural units – φυσικές μονάδες). (Η τυποποιημένη συντόμευση για το $\log_e$ είναι το ln). Εφ όσον

$$\ln a = \ln 2 \log_2 a = 0.693\ 15 \log_2 a$$

όταν η πληροφορία μετριέται σε nats ισούται με ln2 επί την πληροφορία μετρημένη σε bits.
Όταν οι τυχαίες μεταβλητές X και Y είναι στατιστικώς ανεξάρτητες, $P(x_i \mid y_j) = P(x_i)$ και, κατά συνέπεια, $I(x_i ; y_j) = 0$. Από την άλλη πλευρά, όταν η εμφάνιση του γεγονότος $Y = y_j$ καθορίζει μονοσήμαντα την εμφάνιση του γεγονότος $X = x_i$, η υπό συνθήκη πιθανότητα στον αριθμητή της (2-2-1) είναι η μονάδα και, συνεπώς,

$$I(x_i ; y_j) = \log \frac{1}{P(x_i)} = -\log P(x_i) \qquad (2\text{-}2\text{-}2)$$

Αλλά η (2-2-2) είναι απλώς η πληροφορία του γεγονότος $X = x_i$. Για τον λόγο αυτό, αποκαλείται η *αυτο-πληροφορία* (*self-information*) του γεγονότος $X = x_i$ και υποδηλώνεται ως

$$I(x_i) = \log \frac{1}{P(x_i)} = -\log P(x_i) \qquad (2\text{-}2\text{-}3)$$

Παρατηρούμε ότι ένα γεγονός υψηλής πιθανότητας φέρει λιγότερη πληροφορία από ένα γεγονός χαμηλής πιθανότητας. Πράγματι, αν υπάρχει ένα μονό γεγονός x με πιθανότητα $P(x) = 1$, τότε $I(x) = 0$. Για να δείξουμε ακόμη ότι το λογαριθμικό μέτρο της πληροφορίας είναι το κατάλληλο για τις ψηφιακές επικοινωνίες, ας θεωρήσουμε το ακόλουθο παράδειγμα.

Παράδειγμα 2-2-1

Υποθέστε ότι έχουμε μία διακριτή πληροφοριακή πηγή που εκπέμπει ένα δυαδικό ψηφίο, είτε το 0 είτε το 1, με ίση πιθανότητα, κάθε τ_s δευτερόλεπτα. Το πληροφοριακό περιεχόμενο κάθε εξόδου της πηγής είναι

$$\begin{aligned} I(x_i) &= -\log_2 P(x_i), \quad x_i = 0, 1 \\ &= -\log_2 \tfrac{1}{2} = 1\ bit \end{aligned}$$

Τώρα υποθέστε ότι οι διαδοχικές έξοδοι από την πηγή είναι στατιστικώς ανεξάρτητες, ήτοι, η πηγή είναι χωρίς μνήμη. Ας θεωρήσουμε ένα τμήμα από k δυαδικά ψηφία από την πηγή που εμφανίζεται σε ένα χρονικό διάστημα $k\tau_s$. Υπάρχουν $M = 2^k$ πιθανά τμήματα των k bits, το καθένα από τα οποία εμφανίζεται με την ίδια πιθανότητα $1 / M = 2^{-k}$ Η αυτο-πληροφορία ενός τμήματος των k bits είναι

$$I(x_i^{\prime}) = -\log_2 2^{-k} = k\ bits$$

και εκπέμπεται σε ένα χρονικό διάστημα $k\tau_s$. Έτσι, το λογαριθμικό μέτρο του πληροφοριακού περιεχομένου διαθέτει την επιθυμητή ιδιότητα της προσθετικότητας όταν ένας αριθμός εξόδων μίας πηγής θεωρηθεί ως ένα μπλοκ.

Ας επανέλθουμε τώρα στον ορισμό της αμοιβαίας πληροφορίας που δίνεται στην (2-2-1) και ας πολλαπλασιάσουμε τον αριθμητή και τον παρονομαστή του λόγου των πιθανοτήτων με το $P(y_i)$. Εφ όσον

$$\frac{P(x_i \mid y_j)}{P(x_i)} = \frac{P(x_i \mid y_j)P(y_j)}{P(x_i)P(y_j)} = \frac{P(x_i, y_j)}{P(x_i)P(y_j)} = \frac{P(y_j \mid x_i)}{P(y_j)}$$

συμπεραίνουμε ότι

$$I(x_i; y_j) = I(y_j; x_i) \tag{2-2-4}$$

Κατά συνέπεια, η πληροφορία που παρέχεται από την εμφάνιση του γεγονότος $Y = y_j$ περί το γεγονός $X = x_i$ ισούται με την πληροφορία που παρέχεται από την εμφάνιση του γεγονότος $X = x_i$ περί το γεγονός $Y = y_j$.

Παράδειγμα 2-2-2

Υποθέστε ότι τα X και Y είναι τυχαίες μεταβλητές με δυαδικές τιμές $\{0, 1\}$ οι οποίες παριστάνουν την είσοδο και την έξοδο ενός καναλιού δυαδικής εισόδου, δυαδικής εξόδου. Τα σύμβολα εισόδου είναι ισοπίθανα και τα σύμβολα εξόδου εξαρτώνται από την είσοδο σύμφωνα με τις υπό συνθήκη πιθανότητες

$$\begin{aligned}
P(Y = 0 \mid X = 0) &= 1 - p_0 \\
P(Y = 1 \mid X = 0) &= p_0 \\
P(Y = 1 \mid X = 1) &= 1 - p_1 \\
P(Y = 0 \mid X = 1) &= p_1
\end{aligned}$$

Ας καθορίσουμε την αμοιβαία πληροφορία περί την εμφάνιση των γεγονότων $X = 0$ και $X = 1$, δεδομένου ότι $Y = 0$.
Από τις πιθανότητες που δόθηκαν προηγουμένως, παίρνουμε

$$\begin{aligned}
P(Y = 0) &= P(Y = 0 \mid X = 0)P(X = 0) + P(Y = 0 \mid X = 1)P(X = 1) \\
&= \frac{1}{2}(1 - p_0 + p_1) \\
P(Y = 1) &= P(Y = 1 \mid X = 0)P(X = 0) + P(Y = 1 \mid X = 1)P(X = 1) \\
&= \frac{1}{2}(1 - p_1 + p_0)
\end{aligned}$$

Τότε, η αμοιβαία πληροφορία περί την εμφάνιση του γεγονότος $X = 0$, δεδομένου ότι έχει παρατηρηθεί το γεγονός $Y = 0$, είναι

$$\mathrm{I}(x_1; y_1) = I(0;0) = \log_2 \frac{P(Y = 0 \mid X = 0)}{P(Y = 0)} = \log_2 \frac{2(1 - p_0)}{1 - p_0 + p_1}$$

Παρομοίως, δεδομένου ότι έχει παρατηρηθεί $Y = 0$, η αμοιβαία πληροφορία περί την εμφάνιση του γεγονότος $X = 1$ είναι

$$I(x_2 ; y_1) \equiv I(1;0) = \log_2 \frac{2p_1}{1 - p_0 + p_1}$$

Ας θεωρήσουμε κάποιες ειδικές περιπτώσεις: Πρώτον, εάν $p_0 = p_1 = 0$, το κανάλι ονομάζεται *χωρίς θόρυβο* (*noiseless*) και

$$I(0;0) = \log_2 2 = 1 bit$$

Συνεπώς, η έξοδος καθορίζει την είσοδο με βεβαιότητα. Από την άλλη πλευρά, εάν $p_0 = p_1 = ½$, το κανάλι είναι *άχρηστο* (*useless*) επειδή

$$I(0;0) = \log_2 1 = 0$$

Ωστόσο, εάν $p_0 = p_1 = 1/4$, τότε,

$$I(0;0) = \log_2 \frac{3}{2} = 0.587$$

$$I(0;1) = \log_2 \frac{1}{2} = -1 bit$$

Επιπλέον του ορισμού της αμοιβαίας πληροφορίας και της αυτό-πληροφορίας, είναι χρήσιμο να ορίσουμε την *υπό συνθήκη αυτο-πληροφορία* (*conditional self-information*) ως

$$I(x_i \mid y_j) = \log \frac{1}{P(x_i \mid y_j)} = -\log P(x_i \mid y_j) \tag{2-2-5}$$

Τότε, συνδυάζοντας τις (2-2-1), (2-2-3) και (2-2-5), παίρνουμε την σχέση

$$I(x_i ; y_j) = I(x_i) - I(x_i \mid y_j) \tag{2-2-6}$$

Ερμηνεύουμε την $I(x_i \mid y_j)$ ως την αυτό-πληροφορία περί το γεγονός $X = x_i$ κατόπιν της παρατηρήσεως του γεγονότος $Y = y_j$. Επειδή τόσο $I(x_i) \geq 0$ όσο και $I(x_i \mid y_j) \geq 0$, εξυπακούεται ότι $I(x_i; y_j) < 0$ όταν $I(x_i|y_j) > I(x_i)$, και $I(x_i; y_j) > 0$ όταν $I(x_i|y_j) < I(x_i)$. Έτσι, η αμοιβαία πληροφορία μεταξύ ενός ζεύγους γεγονότων μπορεί να είναι θετική, αρνητική, ή μηδενική.

2-1.1 Η Μέση Αμοιβαία Πληροφορία και η Εντροπία

Από την στιγμή που έχουμε ορίσει την αμοιβαία πληροφορία που σχετίζεται με ένα ζεύγος γεγονότων (x_i, y_j), τα οποία είναι οι πιθανές τιμές των δύο τυχαίων μεταβλητών X και Y, μπορούμε να πάρουμε την μέση τιμή της αμοιβαίας πληροφορίας πολλαπλασιάζοντας απλώς το επί μέρους $I(x_i ; y_j)$ με την πιθανότητα εμφάνισης του από κοινού γεγονότος και αθροίζοντας ως προς όλα τα πιθανά από κοινού γεγονότα. Έτσι, παίρνουμε την

$$\begin{aligned} I(X ; Y) &= \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) I(x_i ; y_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log \frac{P(x_i ; y_j)}{P(x_i) P(y_j)} \end{aligned} \tag{2-2-7}$$

ως την μέση αμοιβαία πληροφορία μεταξύ των X και Y. Παρατηρούμε ότι $I(X ; Y) = 0$ όταν τα X και Y είναι στατιστικώς ανεξάρτητα. Ένα σημαντικό χαρακτηριστικό της μέσης αμοιβαίας πληροφορίας είναι ότι $I(X ; Y) \geq 0$

Παρομοίως, ορίζουμε την μέση αυτό-πληροφορία, η οποία συμβολίζεται με $H(X)$, ως

$$
\begin{aligned}
H(X) &= \sum_{i=1}^{n} P(x_i) I(x_i) \\
&= -\sum_{i=1}^{n} P(x_i) \log P(x_i)
\end{aligned}
\tag{2-2-8}
$$

Όταν το X αντιπροσωπεύει το αλφάβητο όλων των πιθανών γραμμάτων εξόδου μίας πηγής, το $H(X)$ αντιπροσωπεύει την μέση αυτό-πληροφορία ανά γράμμα της πηγής, και αποκαλείται η *εντροπία** της πηγής. Στην ειδική περίπτωση, κατά την οποία τα γράμματα της πηγής είναι ισοπίθανα, $P(x_i) = 1/n$ για όλα τα i, και, συνεπώς,

$$
\begin{aligned}
H(X) &= -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} \\
&= \log n
\end{aligned}
\tag{2-2-9}
$$

Γενικά, $H(X) \leq \log n$ για οποιοδήποτε δεδομένο σύνολο πιθανοτήτων των γραμμάτων της πηγής. Με άλλα λόγια, *η εντροπία της πηγής φθάνει στο μέγιστό της όταν τα γράμματα εξόδου είναι εξ ίσου πιθανά.*

ΠΑΡΑΔΕΙΓΜΑ 2-2-3

Θεωρείστε μία πηγή που εκπέμπει μία ακολουθία στατιστικώς ανεξάρτητων γραμμάτων, με το κάθε γράμμα να είναι είτε 0 με πιθανότητα q είτε 1 με πιθανότητα $1-q$. Η εντροπία της πηγής αυτής είναι

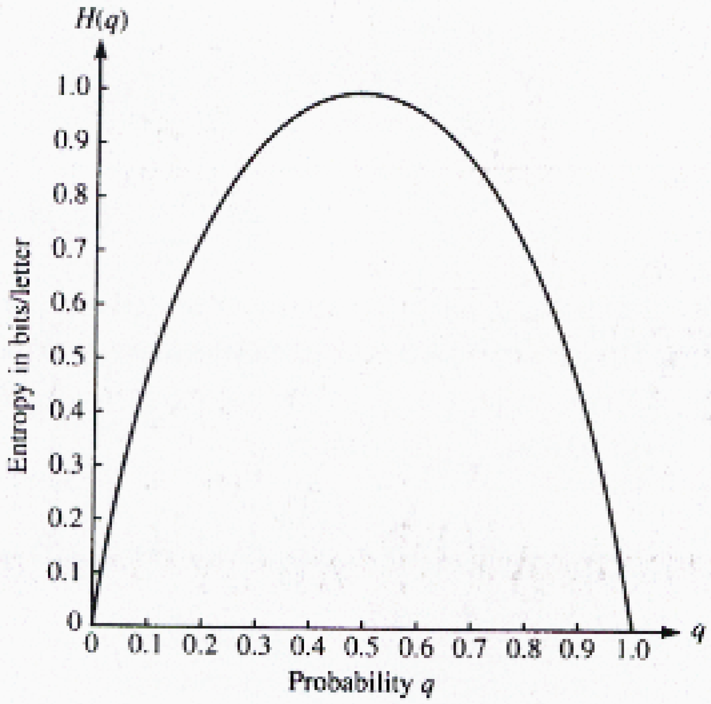
$$
H(X) \equiv H(q) = -q \log q - (1-q) \log(1-q) \tag{2-2-10}
$$

Η συνάρτηση της δυαδικής εντροπίας $H(q)$ φαίνεται στο Σχ. 2-2-1. Παρατηρούμε ότι η μέγιστη τιμή της συνάρτησης εντροπίας συμβαίνει στο $q = ½$ όπου $H(½) = 1$.

Η μέση υπό συνθήκη αυτό-πληροφορία ονομάζεται *υπό συνθήκη εντροπία* (*conditional entropy*) και ορίζεται

$$
H(X \mid Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) \log \frac{1}{P(x_i \mid y_j)} \tag{2-2-11}
$$

* Ο όρος *εντροπία* προέρχεται από την στατιστική μηχανική (θερμοδυναμική), όπου μία συνάρτηση παρόμοια της (3-2-8) ονομάζεται (θερμοδυναμική) εντροπία.



ΣΧΗΜΑ 2-2-1 Η συνάρτηση της δυαδικής εντροπίας

Ερμηνεύουμε την $H(X \mid Y)$ ως την πληροφορία ή την αβεβαιότητα απομένει ως προς το X μετά την παρατήρηση του Y. Συνδυάζοντας τις (2-2-7), (2-2-8) και (2-2-11) παίρνουμε την σχέση

$$I(X;Y) = H(X) - H(X \mid Y) \qquad (2\text{-}2\text{-}12)$$

Εφ όσον $I(X;Y) \geq 0$, συνεπάγεται ότι $H(X) \geq H(X \mid Y)$, με ισότητα εάν και μόνον εάν τα X και Y είναι στατιστικώς ανεξάρτητα. Εάν ερμηνεύσουμε το $H(X \mid Y)$ ως την μέση ποσότητα αβεβαιότητας (υπό συνθήκην αυτό-πληροφορίας) πριν την παρατήρηση, τότε $I(X;Y)$ είναι η μέση ποσότητα της αβεβαιότητας (αμοιβαίας πληροφορίας) που προκύπτει περί το σύνολο X από την παρατήρηση του συνόλου Y. Επειδή $H(X) \geq H(X \mid Y)$, είναι φανερό ότι η συνθήκη ως προς το γεγονός του Y δεν αυξάνει την εντροπία.

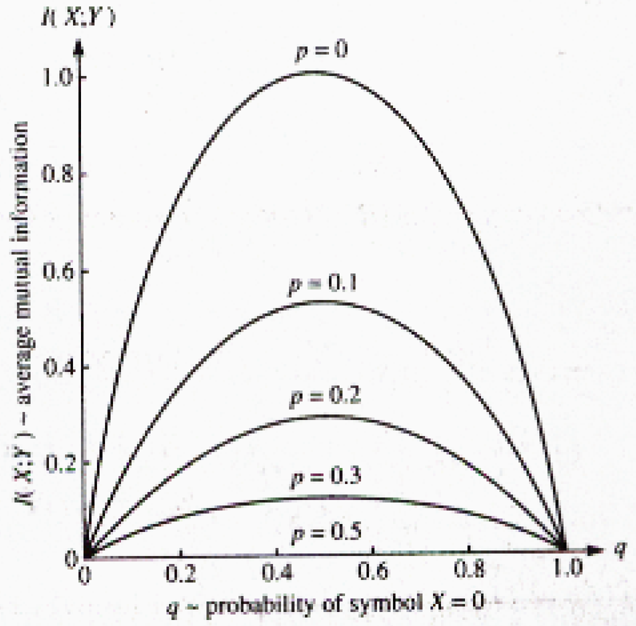
Παράδειγμα 2-2-4

Ας υπολογίσουμε τα $H(X \mid Y)$ και $I(X;Y)$ για το κανάλι δυαδικής εισόδου, δυαδικής εξόδου με το οποίο ασχοληθήκαμε προηγουμένως στο Παράδειγμα 2-2-2 για την περίπτωση κατά την οποία $p_0 = p_1 = p$. Ας είναι οι πιθανότητες των συμβόλων εισόδου $P(X=0) = q$ και $P(X=1) = 1\text{-}q$. Τότε η εντροπία είναι

$$H(X) \equiv H(q) = -q \log q - (1-q)\log(1-q)$$

όπου $H(q)$ είναι η συνάρτηση δυαδικής εντροπίας και η υπό συνθήκη εντροπία $H(X \mid Y)$ ορίζεται από την (2-2-11). Μία σχεδίαση του $H(X \mid Y)$ ως συνάρτησης του q με το p ως παράμετρο δείχνεται στο Σχ. 2-2-2. Η μέση αμοιβαία πληροφορία $I(X;Y)$ έχει σχεδιασθεί στο Σχ. 2-2-3.

Όπως και στο προηγούμενο παράδειγμα, εάν θεωρήσουμε την υπό συνθήκη εντροπία $H(X \mid Y)$ στα πλαίσια ενός καναλιού του οποίου η είσοδος είναι το X και η έξοδος είναι το Y, το $H(X \mid Y)$ ονομάζεται *αμφιλογία* (*equivocation*) και ερμηνεύεται ως η ποσότητα της μέσης αβεβαιότητας που παραμένει στο X μετά την παρατήρηση του Y.

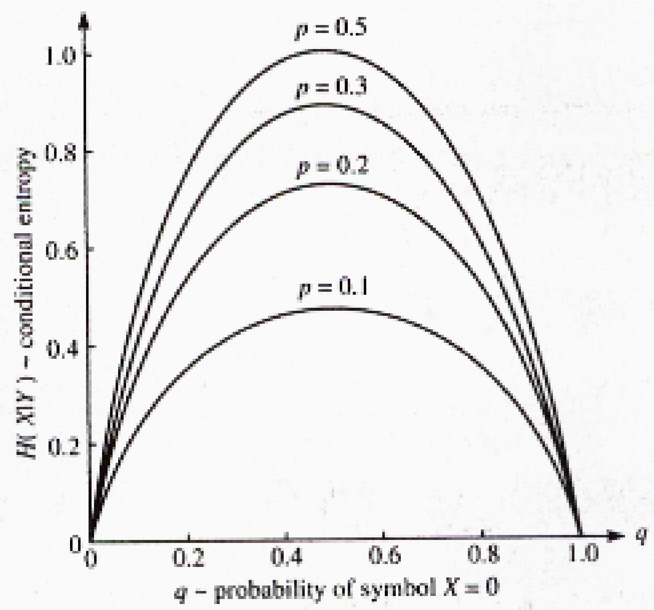


ΣΧΗΜΑ 2-2-3 Η μέση αμοιβαία πληροφορία για ένα δυαδικής εισόδου, δυαδικής εξόδου συμμετρικό κανάλι.

Τα αποτελέσματα που δόθηκαν παραπάνω μπορούν να γενικευθούν σε περισσότερες από δύο τυχαίες μεταβλητές. Συγκεκριμένα, υποθέστε ότι έχουμε ένα τμήμα από k τυχαίες μεταβλητές X_1 X_2 … X_k, με από κοινού πιθανότητα $P(x_1\, x_2\, \ldots\, x_k) \equiv P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$. Τότε, η εντροπία του τμήματος ορίζεται ως

$$H(X_1 X_2 \ldots X_k) = -\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\ldots\sum_{j_k=1}^{n_k} P(x_{j_1} x_{j_2} \ldots x_{j_k}) \log P(x_{j_1} x_{j_2} \ldots x_{j_k}) \qquad (2\text{-}2\text{-}13)$$

Επειδή η από κοινού πιθανότητα $P(x_1\, x_2\, \ldots\, x_k)$ μπορεί να αναλυθεί στο γινόμενο



ΣΧΗΜΑ 2-2-2 Η υπό συνθήκη εντροπία για το δυαδικής εισόδου, δυαδικής εξόδου συμμετρικό κανάλι.

$$P(x_1 x_2 \ldots x_k) = P(x_1) P(x_2 \mid x_1) P(x_3 \mid x_1 x_2) \ldots P(x_k \mid x_1 x_2 \ldots x_{k-1}) \qquad (2\text{-}2\text{-}14)$$

συνεπάγεται ότι

$$H(X_1X_2X_3...X_k) = H(X_1) + H(X_2 \mid X_1) + H(X_3 \mid X_1X_2) \\ + ... + H(X_k \mid X_1...X_{k-1}) \\ = \sum_{i=1}^{k} H(X_i \mid X_1X_2...X_{i-1}) \tag{2-2-15}$$

Εφαρμόζοντας το αποτέλεσμα $H(X) \geq H(X \mid Y)$, όπου $X = X_m$ και $Y = X_1 X_2 \ldots X_{m-1}$, στην (2-2-15), παίρνουμε

$$H(X_1X_2...X_k) \leq \sum_{m=1}^{k} H(X_m) \tag{2-2-16}$$

με ισότητα εάν και μόνον εάν οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_k$ είναι στατιστικώς ανεξάρτητες.

2-1.2 Τα Πληροφοριακά Μέτρα για τις Συνεχείς Τυχαίες Μεταβλητές

Ο ορισμός της αμοιβαίας πληροφορίας που δόθηκε παραπάνω για τις διακριτές τυχαίες μεταβλητές μπορεί να επεκταθεί με έναν απλό τρόπο σε συνεχείς τυχαίες μεταβλητές. Συγκεκριμένα, εάν οι X και Y είναι τυχαίες μεταβλητές με από κοινού σ.π.π. $p(x, y)$ και οριακές σ.π.π. $p(x)$ και $p(y)$, η μέση αμοιβαία πληροφορία μεταξύ των X και Y ορίζεται ως

$$I(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x)p(y \mid x) \log \frac{p(y \mid x)p(x)}{p(x)p(y)} dxdy \tag{2-2-17}$$

Αν και ο ορισμός της μέσης αμοιβαίας πληροφορίας μπορεί να εφαρμοσθεί και στην περίπτωση των συνεχών τυχαίων μεταβλητών, δεν συμβαίνει το ίδιο και με την έννοια της αυτό-πληροφορίας. Το πρόβλημα είναι ότι η ακριβής αναπαράσταση μίας συνεχούς τυχαίας μεταβλητής απαιτεί έναν άπειρο αριθμό δυαδικών ψηφίων. Συνεπώς, η αυτό-πληροφορία της συνεχούς τυχαίας μεταβλητής είναι άπειρη και, έτσι, η εντροπία της είναι επίσης άπειρη. Πάντως, θα ορίσουμε μία ποσότητα την οποία θα αποκαλούμε *διαφορική εντροπία* (*differential entropy*) της συνεχούς τυχαίας μεταβλητής X ως

$$H(X) = -\int_{-\infty}^{\infty} p(x) \log p(x) dx \tag{2-2-18}$$

Πρέπει να δοθεί έμφαση στο γεγονός ότι η ποσότητα αυτή *δεν έχει* την φυσική έννοια της αυτό-πληροφορίας, αν και μπορεί να μοιάζει με μία φυσική επέκταση του ορισμού της εντροπίας μίας διακριτής τυχαίας μεταβλητής.

Εάν ορισθεί η μέση υπό συνθήκη πληροφορία του X δεδομένου του Y ως

$$H(X \mid Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x, y) \log p(x \mid y) dxdy \tag{2-2-19}$$

η μέση αμοιβαία πληροφορία μπορεί να εκφρασθεί ως

$I(X;Y) = H(X) - H(X \mid Y)$

ή, εναλλακτικά, ως

$I(X;Y) = H(Y) - H(Y \mid X)$

Σε μερικές περιπτώσεις που παρουσιάζουν ενδιαφέρον από πρακτικής πλευράς, η τυχαία μεταβλητή X είναι διακριτή και η Y είναι συνεχής. Συγκεκριμένα, υποθέστε ότι η X έχει πιθανά αποτελέσματα x_i, $i = 1, 2, \ldots, n$, και η Y χαρακτηρίζεται από την οριακή της σ.π.π. $p(y)$. Όταν οι X και Y είναι στατιστικώς ανεξάρτητες, μπορούμε να εκφράσουμε την $p(y)$ ως

$p(y) = \sum_{i=1}^{n} p(y \mid x_i) P(x_i)$

Η αμοιβαία πληροφορία που παρέχεται περί το γεγονός $X = x_i$ από την εμφάνιση του γεγονότος $Y = y$ είναι

$$I(x_i;y) = \log \frac{p(y \mid x_i)P(x_i)}{p(y)P(x_i)} = \log \frac{p(y \mid x_i)}{p(y)} \tag{2-2-20}$$

Τότε, η μέση αμοιβαία πληροφορία μεταξύ των X και Y είναι

$$I(X;Y) = \sum_{i=1}^{n} \int_{-\infty}^{\infty} p(y \mid x_i)P(x_i) \log \frac{p(y \mid x_i)}{p(y)} dy \tag{2-2-21}$$

Παράδειγμα 2-2-5

Υποθέστε ότι η X είναι μία διακριτή τυχαία μεταβλητή με δύο ισοπίθανα αποτελέσματα $x_1 = A$ και $x_2 = -A$. Ας είναι οι υπό συνθήκη σ.π.π. $p(y \mid x_i)$, $i = 1, 2$, γκαουσσιανές με μέσον όρο x_i και διασπορά σ^2. Δηλαδή,

$$p(y \mid A) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-A)^2/2\sigma^2}$$

$$p(y \mid -A) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y+A)^2/2\sigma^2} \tag{2-2-22}$$

Η μέση αμοιβαία πληροφορία που λαμβάνεται από την (2-2-21) γίνεται

$$I(X;Y) = \frac{1}{2} \int_{-\infty}^{\infty} \left[p(y \mid A) \log \frac{p(y \mid A)}{p(y)} + p(y \mid -A) \log \frac{p(y \mid -A)}{p(y)} \right] dy \tag{2-2-23}$$

$$p(y) = \frac{1}{2}[p(y \mid A) + p(y \mid -A)] \tag{2-2-24}$$

Στο Κεφάλαιο 7, θα δειχθεί ότι η μέση αμοιβαία πληροφορία $I(X;Y)$ που δίνεται από την (2-2-23) παριστάνει την χωρητικότητα ενός καναλιού δυαδικής εισόδου και προσθετικού λευκού θορύβου.

2-3 Η ΚΩΔΙΚΟΠΟΙΗΣΗ ΤΩΝ ΔΙΑΚΡΙΤΩΝ ΠΗΓΩΝ

Στην Ενότητα 2-2 παρουσιάσαμε ένα μέτρο του πληροφοριακού περιεχομένου μίας διακριτής τυχαίας μεταβλητής X. Όταν η X είναι η έξοδος μίας διακριτής πηγής, η εντροπία $H(X)$ της πηγής παριστά την μέση ποσότητα πληροφορίας που εκπέμπεται από την πηγή. Στην ενότητα αυτή, θα μελετήσουμε την διαδικασία κωδικοποίησης της εξόδου μίας πηγής, ήτοι, την διαδικασία αναπαράστασης της εξόδου της πηγής με μία ακολουθία δυαδικών ψηφίων. Ένα μέτρο της αποδοτικότητας μίας μεθόδου κωδικοποίησης πηγής μπορεί να ληφθεί με την σύγκριση του μέσου αριθμού των δυαδικών ψηφίων ανά γράμμα εξόδου από την πηγή με την εντροπία $H(X)$.

Η κωδικοποίηση μίας διακριτής πηγής που έχει ένα πεπερασμένο μέγεθος αλφαβήτου μπορεί να φαίνεται, εκ πρώτης όψεως, ως ένα σχετικά απλό πρόβλημα. Εν τούτοις, αυτό συμβαίνει μόνον όταν η πηγή είναι χωρίς μνήμη, δηλαδή, όταν τα διαδοχικά σύμβολα από την πηγή είναι στατιστικώς ανεξάρτητα και κάθε σύμβολο κωδικοποιείται ξεχωριστά. Η Διακριτή Πηγή Χωρίς Μνήμη ή ΔΠΧΜ (discrete memoryless source - DMS) είναι με διαφορά το απλούστερο μοντέλο που μπορεί να επινοηθεί για μία φυσική πηγή. Ωστόσο, λίγες είναι οι φυσικές πηγές που ταιριάζουν πραγματικά με αυτό το εξιδανικευμένο μαθηματικό μοντέλο. Για παράδειγμα, τα διαδοχικά γράμματα εξόδου από μία μηχανή που εκτυπώνει Αγγλικό κείμενο αναμένεται να είναι στατιστικώς εξηρτημένα. Από την άλλη πλευρά, εάν η έξοδος της μηχανής είναι ένα υπολογιστικό πρόγραμμα κωδικοποιημένο σε Fortran, η ακολουθία των γραμμάτων εξόδου αναμένεται να παρουσιάζει μία κατά πολύ μικρότερη εξάρτηση. Σε κάθε περίπτωση, θα δείξουμε ότι είναι πάντοτε αποδοτικότερο να κωδικοποιούμε τμήματα συμβόλων αντί να κωδικοποιούμε κάθε σύμβολο ξεχωριστά. Εάν κάνουμε το μέγεθος του τμήματος αρκετά μεγάλο, ο μέσος αριθμός των δυαδικών ψηφίων ανά γράμμα εξόδου από την πηγή μπορεί να πλησιάσει αυθαίρετα κοντά την εντροπία της πηγής.

2-1.3 Η Κωδικοποίηση μιας Διακριτής Πηγής Χωρίς Μνήμη

Υποθέστε ότι μία ΔΠΧΜ παράγει ένα γράμμα εξόδου ή σύμβολο κάθε τ_s δευτερόλεπτα. Κάθε σύμβολο επιλέγεται από ένα πεπερασμένο αλφάβητο συμβόλων x_i, $i = 1, 2, \ldots, L$, που συμβαίνουν με πιθανότητες $P(x_i)$, $i = 1, 2, \ldots, L$,. Η εντροπία της ΔΠΧΜ σε bits ανά σύμβολο είναι

$$H(X) = -\sum_{i=1}^{L} P(x_i)\log_2 P(x_i) \leq \log_2 L \qquad (2\text{-}3\text{-}1)$$

όπου η ισότητα ισχύει όταν τα σύμβολα είναι ισοπίθανα. Ο μέσος αριθμός bits ανά σύμβολο πηγής είναι $H(X)$ και ο ρυθμός της πηγής σε bits/s ορίζεται ως $H(X)/\tau_s$.

Κωδικές Λέξεις Καθορισμένου Μήκους Πρώτα θα δούμε μία μέθοδο τμηματικής κωδικοποίησης που εκχωρεί ένα μοναδικό σύνολο R δυαδικών ψηφίων σε κάθε σύμβολο. Επειδή υπάρχουν L πιθανά σύμβολα, ο αριθμός των δυαδικών ψηφίων ανά σύμβολο που απαιτούνται για αμφιμονοσήμαντη κωδικοποίηση, όταν το L είναι δύναμη του 2, είναι

$$R = \log_2 L \qquad (2\text{-}3\text{-}2)$$

και, όταν το L δεν είναι δύναμη του 2, είναι

$$R = \lfloor \log_2 L \rfloor + 1 \qquad (2\text{-}3\text{-}3)$$

όπου το $\lfloor x \rfloor$ υποδηλώνει τον μέγιστο ακέραιο που είναι μικρότερος από το x. Ο ρυθμός R του κώδικα σε bits ανά σύμβολο είναι τώρα R και, επειδή $H(X) \leq \log_2 L$, εξυπακούεται ότι $R \geq H(X)$.
Η αποδοτικότητα της κωδικοποίησης μίας ΔΠΧΜ ορίζεται ως ο λόγος $H(X)/R$. Παρατηρούμε ότι όταν το L δύναμη του 2 και τα γράμματα της πηγής είναι ισοπίθανα, $R = H(X)$. Συνεπώς, ένας κώδικας σταθερού μήκους των R bits ανά σύμβολο επιτυγχάνει 100% αποδοτικότητα. Ωστόσο, εάν το L δεν είναι δύναμη του 2 αλλά τα σύμβολα της πηγής είναι πάλι ισοπίθανα, το R διαφέρει από το $H(X)$ κατά 1 bit ανά σύμβολο το πολύ. Όταν $\log_2 L >> 1$, η αποδοτικότητα της μεθόδου αυτής είναι υψηλή. Από την άλλη πλευρά, όταν το L είναι μικρό, η αποδοτικότητα του κώδικα σταθερού μήκους μπορεί να αυξηθεί εάν κωδικοποιείται κάθε φορά μία ακολουθία J συμβόλων. Για να επιτευχθεί η επιθυμητή κωδικοποίηση, απαιτούνται L^J διαφορετικές κωδικές λέξεις. Χρησιμοποιώντας ακολουθίες των N δυαδικών ψηφίων, μπορούμε να αντιπροσωπεύσουμε 2^N πιθανές κωδικές λέξεις. Το N πρέπει να επιλεγεί, έτσι ώστε

$$N \geq J \log_2 L$$

Κατά συνέπεια, ο ελάχιστος ακέραιος αριθμός του N που απαιτείται είναι

$$N = \lfloor J \log_2 L \rfloor + 1 \qquad (2\text{-}3\text{-}4)$$

Όμως, ο μέσος αριθμός των bits ανά σύμβολο πηγής είναι $N/J = R$, και, έτσι, η έλλειψη της αποδοτικότητας έχει μειωθεί κατά ένα παράγοντα του $1/J$ περίπου, έναντι της ανά σύμβολο κωδικοποίησης που περιγράφηκε προηγουμένως. Με το J να γίνεται αρκούντως μεγάλο, η αποδοτικότητα της διαδικασίας κωδικοποίησης, μετρούμενης με τον λόγο $JH(X) / N$, μπορεί να πλησιάσει στην μονάδα όσον είναι επιθυμητό.
Οι μέθοδοι κωδικοποίησης που περιγράφηκαν παραπάνω δεν προκαλούν καθόλου παραμόρφωση, επειδή η κωδικοποίηση των συμβόλων της πηγής, ή των τμημάτων των

συμβόλων σε κωδικές λέξεις γίνεται με αμφιμονοσήμαντο τρόπο. Ο τύπος αυτός κωδικοποίησης λέγεται *χωρίς θόρυβο* (*noiseless*).

Τώρα, υποθέστε ότι προσπαθούμε να μειώσουμε τον ρυθμό του κώδικα R χαλαρώνοντας την συνθήκη ότι η διαδικασία κωδικοποίησης πρέπει να είναι αμφιμονοσήμαντη. Για παράδειγμα, υποθέστε ότι μόνο ένα ποσοστό των L^J τμημάτων συμβόλων κωδικοποιούνται αμφιμονοσήμαντα. Συγκεκριμένα, ας επιλέξουμε τα $2^N - 1$ περισσότερο πιθανά τμήματα των J συμβόλων και ας κωδικοποιήσουμε το καθένα τους αμφιμονοσήμαντα, ενώ τα υπόλοιπα $L^J - (2^N - 1)$ τμήματα συμβόλων αναπαρίστανται με την μοναδική λέξη που απομένει. Η διαδικασία αυτή προκαλεί μία αστοχία αποκωδικοποίησης ή πιθανότητα λάθους (παραμόρφωση) κάθε φορά που ένα τμήμα χαμηλής πιθανότητας αντιστοιχίζεται στην μοναδική αυτή κωδική λέξη. Ας υποθέσουμε ότι το P_e υποδηλώνει αυτή την πιθανότητα λάθους. Βάσει αυτής της διαδικασίας τμηματικής κωδικοποίησης, ο Shannon (1948α) απέδειξε το ακόλουθο θεώρημα κωδικοποίησης πηγής.

Θεώρημα Κωδικοποίησης Πηγής Ι

Ας είναι X το σύνολο των γραμμάτων μίας ΔΠΧΜ με πεπερασμένη εντροπία $H(X)$. Τμήματα των J συμβόλων από την πηγή κωδικοποιούνται σε κωδικές λέξεις μήκους N από ένα δυαδικό αλφάβητο. Για κάθε $\varepsilon > 0$, η πιθανότητα P_e μίας αστοχίας αποκωδικοποίησης (decoding failure) ενός τμήματος μπορεί να γίνει αυθαίρετα μικρή εάν

$$R \equiv \frac{N}{J} \geq H(X) + \varepsilon \tag{2-3-5}$$

και το J γίνει αρκούντως μεγάλο. Αντιθέτως, εάν

$$R \leq H(X) - \varepsilon \tag{2-3-6}$$

τότε το P_e πλησιάζει αυθαίρετα το 1 καθώς το J γίνεται αρκούντως μεγάλο.

Από το θεώρημα αυτό, παρατηρούμε ότι ο μέσος αριθμός των bits ανά σύμβολο που απαιτούνται για την κωδικοποίηση της εξόδου μίας ΔΠΧΜ με αυθαίρετα μικρή πιθανότητα αστοχίας αποκωδικοποίησης φράσσεται εκ των κάτω από την εντροπία της πηγής $H(X)$. Από την άλλη πλευρά, εάν $R < H(X)$, το ποσοστό της αστοχίας αποκωδικοποίησης πλησιάζει το 100% καθώς το J αυξάνει αυθαίρετα.

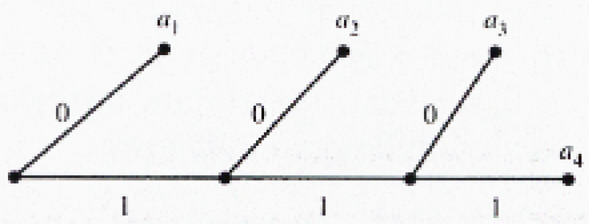
Κωδικές Λέξεις Μεταβλητού Μήκους Όταν τα σύμβολα της πηγής δεν είναι εξ ίσου πιθανά, μία αποδοτικότερη μέθοδος κωδικοποίησης είναι η χρήση κωδικών λέξεων μεταβλητού μήκους. Ένα παράδειγμα κωδικοποίησης του τύπου αυτού είναι ο κώδικας Morse, ο οποίος υφίσταται από τον δέκατο ένατο αιώνα. Στον κώδικα Morse, στα γράμματα που συμβαίνουν συχνότερα εκχωρούνται μικρές κωδικές λέξεις και σε αυτά που εμφανίζονται σπανιότερα εκχωρούνται μεγάλες κωδικές λέξεις. Με βάση αυτήν την φιλοσοφία, μπορούμε να χρησιμοποιήσουμε τις πιθανότητες εμφάνισης των διαφορετικών γραμμάτων της πηγής κατά την επιλογή των κωδικών λέξεων. Το πρόβλημα είναι να επινοήσουμε μία μέθοδο επιλογής και εκχώρησης των κωδικών λέξεων στα γράμματα της πηγής. Αυτός ο τύπος κωδικοποίησης λέγεται *κωδικοποίηση εντροπίας* (*entropy coding*).

ΠΙΝΑΚΑΣ 2-3-1 ΚΩΔΙΚΕΣ ΜΕΤΑΒΛΗΤΟΥ ΜΗΚΟΥΣ

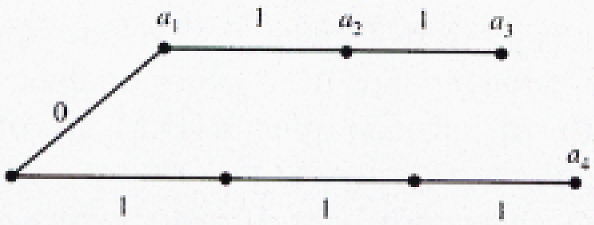
Γράμμα	$P(a_k)$	Κώδικας I	Κώδικας II	Κώδικας III
a_1	1/2	1	0	0
a_2	1/4	00	10	01
a_3	1/8	01	110	011
a_4	1/8	10	111	111

Για παράδειγμα, υποθέστε ότι μία ΔΠΧΜ με γράμματα εξόδου a_1, a_2, a_3, a_4 και αντίστοιχες πιθανότητες $P(a_1) = ½$, $P(a_2) = \frac{1}{4}$, $P(a_3) = P(a_4) = \frac{1}{8}$ κωδικοποιείται όπως φαίνεται στον Πίνακα 2-3-1. Ο Κώδικας I είναι ένας κώδικας μεταβλητού μήκους ο οποίος παρουσιάζει ένα βασικό μειονέκτημα. Για να δούμε το μειονέκτημα, ας υποθέσουμε ότι μας παρουσιάζεται η ακολουθία 001001… Προφανώς, το πρώτο σύμβολο που αντιστοιχεί στο 00 είναι το a_2. Ωστόσο, τα επόμενα τέσσερα bits είναι αμφίβολα (δεν είναι αμφιμονοσήμαντα αποκωδικοποιήσιμα). Μπορούν να κωδικοποιηθούν είτε ως a_4a_3, είτε ως $a_1a_2a_1$. Ίσως να λυνόταν η αβεβαιότητα, εάν περιμέναμε περισσότερα bits, αλλά μία καθυστέρηση αποκωδικοποίησης του τύπου αυτού είναι πολύ ανεπιθύμητη. Θα ασχοληθούμε μόνο με κώδικες που είναι *στιγμιαία* αποκωδικοποιήσιμοι, δηλ, απαλλαγμένοι από καθυστέρηση αποκωδικοποίησης.

ΣΧΗΜΑ 2-3-1 Το κωδικό δένδρο του Κώδικα II του Πίνακα 2-3-1.



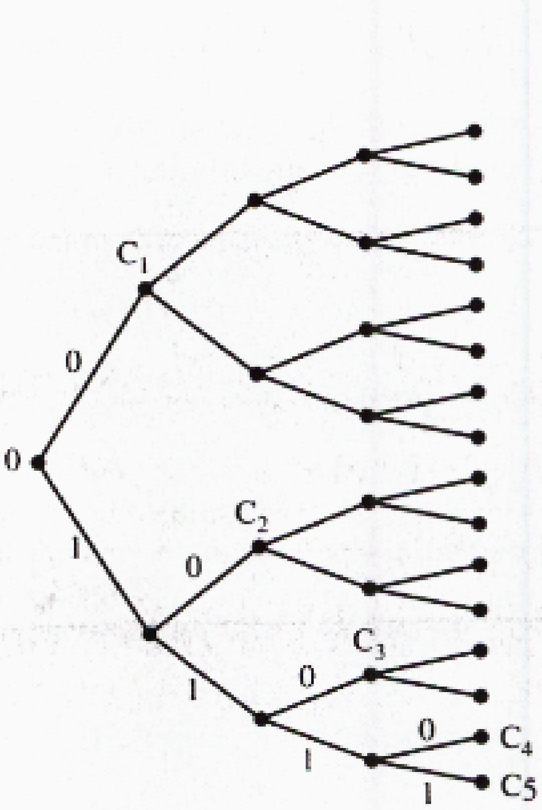
ΣΧΗΜΑ 2-3-2 Το κωδικό δένδρο του Κώδικα III του Πίνακα 2-3-1.



Ο κώδικας II του Πίνακα 2-3-1 είναι *αμφιμονοσήμαντα αποκωδικοποιήσιμος* και *στιγμιαία αποκωδικοποιήσιμος*. Είναι βολικό να αναπαραστήσουμε γραφικά τις κωδικές λέξεις του κώδικα αυτού ως τερματικούς κόμβους ενός δένδρου, όπως φαίνεται στο Σχ. 2-3-1. Παρατηρούμε ότι το ψηφίο 0 υποδεικνύει το τέλος μίας κωδικής λέξης για τις τρεις πρώτες κωδικές λέξεις. Το χαρακτηριστικό αυτό συν το γεγονός ότι δεν υπάρχει κωδική λέξη μεγαλύτερη των τριών δυαδικών ψηφίων κάνει τον κώδικα αυτό στιγμιαία αποκωδικοποιήσιμο. Σημειώστε ότι καμία κωδική λέξη στον κώδικα αυτό δεν αποτελεί πρόθεμα κάποιας άλλης κωδικής λέξης. Γενικά, η *συνθήκη του προθέματος* (*prefix condition*) απαιτεί, για οποιαδήποτε δεδομένη κωδική λέξη $\mathbf{C}_k$ μήκους k πού έχει τα στοιχεία $(b_1, b_2, \ldots, b_k)$ να μην υπάρχει καμία άλλη κωδική λέξη μήκους $l < k$ με στοιχεία $(b_1, b_2, \ldots, b_l)$ για $1 \leq l \leq k-1$. Με άλλα λόγια, δεν υπάρχει κωδική λέξη μήκους $l < k$ η οποία να ταυτίζεται με τα πρώτα l δυαδικά ψηφία μίας άλλης κωδικής λέξης μήκους $k > l$. Αυτή η ιδιότητα κάνει τις κωδικές λέξεις στιγμιαία αποκωδικοποιήσιμες.

Ο κώδικας III που δίνεται στον Πίνακα 2-3-1 έχει την δομή του δένδρου που φαίνεται στο Σχ. 2-3-2. Παρατηρούμε ότι στην περίπτωση αυτή ο κώδικας είναι αμφιμονοσήμαντα αποκωδικοποιήσιμος αλλά *όχι* στιγμιαία αποκωδικοποιήσιμος. Προφανώς, ο κώδικας αυτός *δεν* ικανοποιεί την συνθήκη προθέματος.

Ο κύριος στόχος μας είναι να επινοήσουμε μία συστηματική μέθοδο για την κατασκευή αμφιμονοσήμαντα αποκωδικοποιήσιμων κωδίκων μεταβλητού μεγέθους που να είναι αποδοτικοί, υπό την έννοια ότι ελαχιστοποιούν τον μέσο αριθμό των bits ανά γράμμα πηγής, ο οποίος ορίζεται ως η ποσότητα.



$$\overline{R} = \sum_{k=1}^{L} n_k P(a_k) \tag{2-3-7}$$

Οι συνθήκες που καθορίζουν την ύπαρξη ενός κώδικα που ικανοποιεί την συνθήκη προθέματος δίνονται από την ανισότητα Kraft.

Ανισότητα Kraft Μία ικανή και αναγκαία συνθήκη που καθορίζει την ύπαρξη ενός δυαδικού κώδικα που αποτελείται από κωδικές λέξεις με μήκη $n_1 \leq n_2 \leq \ldots \leq n_L$ που ικανοποιούν την συνθήκη προθέματος είναι

$$\sum_{k=1}^{L} 2^{-n_k} \leq 1 \tag{2-3-8}$$

Πρώτα θα αποδείξουμε ότι η (2-3-8) είναι μία ικανή συνθήκη για την ύπαρξη ενός κώδικα που ικανοποιεί την συνθήκη προθέματος. Για να κατασκευάσουμε έναν τέτοιο κώδικα, αρχίζουμε με ένα πλήρες δυαδικό δένδρο τάξης $n = n_L$, το οποίο έχει 2^n τερματικούς κόμβους. Στο δένδρο αυτό, από κάθε κόμβο τάξης $k-1$, πηγάζουν δύο κόμβοι τάξης k, για κάθε k, $1 \leq k \leq n$. Ας επιλέξουμε οποιονδήποτε κόμβο τάξης n_1 ως την πρώτη κωδική λέξη $\mathbf{C}_1$. Η επιλογή αυτή εξαλείφει 2^{n-n_1} τερματικούς κόμβους (ή το κλάσμα 2^{-n_1} των 2^n τερματικών κόμβων). Από τους υπόλοιπους διαθέσιμους κόμβους τάξης n_2, επιλέγουμε έναν κόμβο για την δεύτερη κωδική λέξη $\mathbf{C}_2$. Η επιλογή αυτή εξαλείφει 2^{n-n_2} τερματικούς κόμβους (ή ένα ποσοστό 2^{-n_2} των τερματικών κόμβων). Η διαδικασία αυτή συνεχίζεται μέχρις ότου η τελευταία κωδική λέξη έχει εκχωρηθεί στον τερματικό κόμβο $n = n_L$. Εφ όσον, στον κόμβο τάξης $j < L$, το ποσοστό του αριθμού των τερματικών κόμβων που έχουν εξαλειφθεί είναι

$$\sum_{k=1}^{j} 2^{-n_k} < \sum_{k=1}^{L} 2^{-n_k} \leq 1$$

υπάρχει πάντοτε ένας κόμβος τάξης $k > j$ διαθέσιμος προς εκχώρηση στην επόμενη κωδική λέξη. Έτσι, έχουμε κατασκευάσει ένα κωδικό δένδρο που είναι εμπεδωμένο στο πλήρες δένδρο των 2^n κόμβων όπως φαίνεται στο Σχ. 2-3-3, για ένα δένδρο που έχει 16 τερματικούς κόμβους και μία έξοδο πηγής που αποτελείται από πέντε γράμματα με $n_1 = 1$, $n_2 = 2$, $n_3 = 3$, και $n_4 = n_5 = 4$.

ΣΧΗΜΑ 2-3-3 Η κατασκευή ενός δυαδικού κωδικού δένδρου εμπεδωμένου σε ένα πλήρες δένδρο.

Για να αποδείξουμε ότι η (2-3-8) είναι μία αναγκαία συνθήκη, παρατηρούμε ότι στο κωδικό δένδρο τάξης $n = n_L$ ο αριθμός των τερματικών κόμβων που έχουν εξαλειφθεί από τον συνολικό αριθμό 2^n των τερματικών κόμβων είναι

$$\sum_{k=1}^{L} 2^{n-n_k} \le 2^n$$

Έτσι,

$$\sum_{k=1}^{L} 2^{-n_k} \le 1$$

και η απόδειξη της (2-3-8) έχει ολοκληρωθεί.

Η ανισότητα του Kraft μπορεί να χρησιμοποιηθεί για να αποδειχθεί το ακόλουθο θεώρημα κωδικοποίησης πηγής (χωρίς θόρυβο), το οποίο ισχύει στους κώδικες που ικανοποιούν την συνθήκη προθέματος.

Θεώρημα Κωδικοποίησης Πηγής ΙΙ

Ας είναι X το σύνολο των γραμμάτων από μία ΔΠΧΜ με πεπερασμένη εντροπία $H(X)$, και γράμματα εξόδου x_k, $1 \le k \le l$, με αντίστοιχες πιθανότητες p_k, $1 \le k \le l$. Είναι δυνατόν να κατασκευασθεί ένας κώδικας που να ικανοποιεί την συνθήκη του προθέματος και ο οποίος να έχει ένα μέσο μήκος $\overline{R}$ που να ικανοποιεί τις ανισότητες

$$H(X) \le \overline{R} \le H(X) + 1 \tag{2-3-9}$$

Για να επαληθεύσουμε το κατώτερο όριο της (2-3-9), παρατηρούμε ότι για τις κωδικές λέξεις που έχουν μήκος n_k, $1 \le k \le L$, η διαφορά $H(X) - \overline{R}$ μπορεί να εκφρασθεί ως

$$\begin{aligned} H(X) - \overline{R} &= \sum_{k=1}^{L} p_k \log_2 \frac{1}{p_k} - \sum_{k=1}^{L} p_k n_k \\ &= \sum_{k=1}^{L} p_k \log_2 \frac{2^{-n_k}}{p_k} \end{aligned} \tag{2-3-10}$$

Η χρήση της ανισότητας $\ln x \le x - 1$ στην (2-3-10) φέρνει ως αποτέλεσμα

$$\begin{aligned} H(X) - \overline{R} &\le (\log_2 e) \sum_{k=1}^{L} p_k \left(\frac{2^{-n_k}}{p_k} - 1 \right) \\ &\le (\log_2 e) \left(\sum_{k=1}^{L} 2^{-n_k} - 1 \right) \le 0 \end{aligned}$$

όπου η τελευταία ανισότητα συνεπάγεται από την ανισότητα του Kraft. Η ισότητα ισχύει εάν και μόνον εάν $p_k = 2^{-n_k}$ για $1 \le k \le l$.

Το ανώτερο όριο στην (2-3-9) μπορεί να επαληθευθεί υπό τον περιορισμό ότι τα n_k, $1 \le k \le l$, είναι ακέραια, με το να επιλέξουμε τα $\{n_k\}$ έτσι ώστε $2^{-n_k} \le p_k < 2^{-n_k+1}$. Όμως, εάν αθροισθούν οι όροι $p_k \ge 2^{-n_k}$ για όλες τις τιμές στην περιοχή $1 \le k \le l$, παίρνουμε την ανισότητα Kraft, για την οποία έχουμε αποδείξει ότι υπάρχει ένας κώδικας που ικανοποιεί την συνθήκη προθέματος. Από την άλλη πλευρά, εάν πάρουμε τον λογάριθμο του $p_k < 2^{-n_k+1}$, λαμβάνουμε

$$\log p_k < -n_k + 1$$

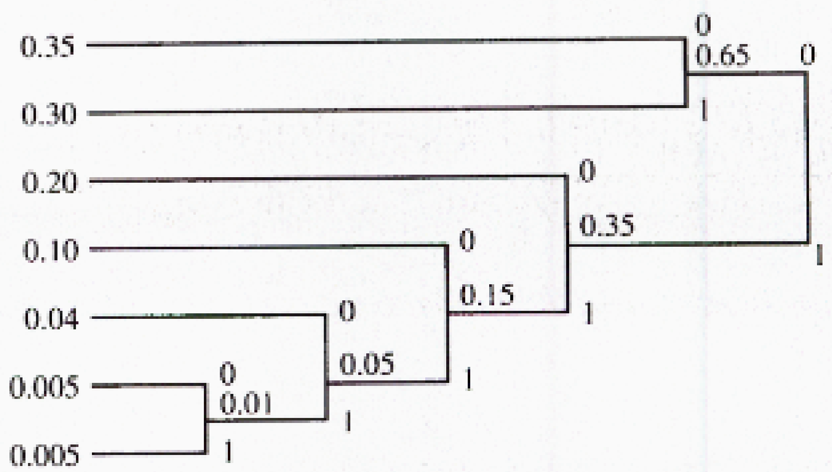
ή, ισοδύναμα,

$$n_k < 1 - \log p_k \qquad (2\text{-}3\text{-}11)$$

Εάν πολλαπλασιάσουμε αμφότερες τις πλευρές της (2-3-11) με το p_k και αθροίσουμε για όλες τις τιμές $1 \leq k \leq l$, παίρνουμε το επιθυμητό ανώτατο όριο που δίνεται στην (2-3-9).

Έχουμε τώρα επαληθεύσει ότι οι κώδικες μεταβλητού μήκους που ικανοποιούν την συνθήκη προθέματος είναι αποδοτικοί κώδικες πηγής για οποιαδήποτε ΔΠΧΜ με σύμβολα πηγής που δεν είναι ισοπίθανα. Ας περιγράψουμε τώρα έναν αλγόριθμο με την βοήθεια του οποίου μπορούμε να κατασκευάσουμε κώδικες αυτού του τύπου.

Ο Αλγόριθμος Κωδικοποίησης του Huffman Ο Huffman, το 1952, επινόησε έναν αλγόριθμο κωδικοποίησης μεταβλητού μήκους, βάσει των πιθανοτήτων των γραμμάτων της πηγής $P(x_i)$, $i = 1, 2, \ldots, L$. Ο αλγόριθμος αυτός είναι βέλτιστος, υπό την έννοια ότι ο μέσος αριθμός των δυαδικών ψηφίων που απαιτούνται για να αναπαρασταθούν τα σύμβολα της πηγής είναι ελάχιστος, υπό την προϋπόθεση ότι οι κωδικές λέξεις ικανοποιούν την συνθήκη προθέματος, όπως ορίσθηκε προηγουμένως, ώστε η λαμβανόμενη ακολουθία να είναι αμφιμονοσήμαντη και στιγμιαία αποκωδικοποιήσιμη. Θα παρουσιάσουμε αυτόν τον αλγόριθμο κωδικοποίησης με την βοήθεια δύο παραδειγμάτων.

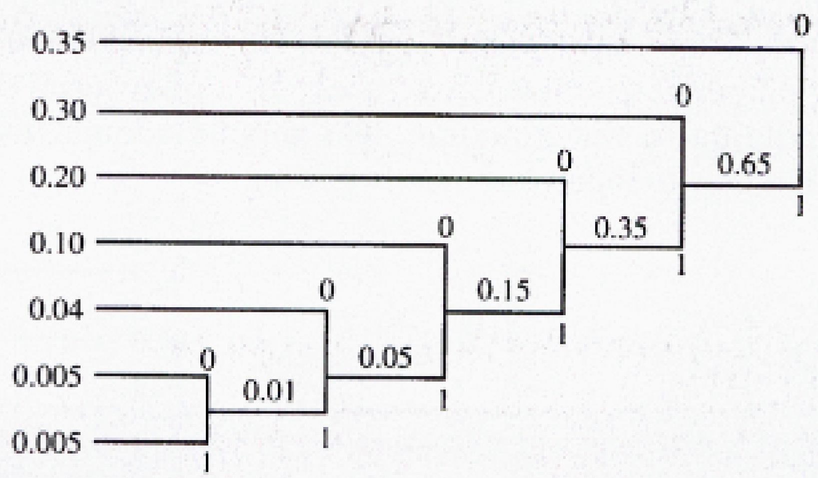


Letter	Probability	Self-information	Code
x_1	0.35	1.5146	00
x_2	0.30	1.7370	01
x_3	0.20	2.3219	10
x_4	0.10	3.3219	110
x_5	0.04	4.6439	1110
x_6	0.005	7.6439	11110
x_7	0.005	7.6439	11111
	$H(X) = 2.11$	$\bar{R} = 2.21$	

ΣΧΗΜΑ 2-3-4 Ένα παράδειγμα κωδικοποίησης πηγής μεταβλητού μήκους για μία ΔΠΧΜ.

Παράδειγμα 2-3-1

Θεωρείστε μία ΔΠΧΜ με επτά πιθανά σύμβολα $x_1, x_2, \ldots, x_7$ που έχουν τις πιθανότητες εμφάνισης που φαίνονται στο Σχ. 2-3-4. Έχουμε διατάξει τα σύμβολα της πηγής σε φθίνουσα σειρά των πιθανοτήτων, ήτοι, $P(x_1) > P(x_2) > \ldots > P(x_7)$. Ξεκινούμε την διαδικασία κωδικοποίησης με τα δύο λιγότερο πιθανά σύμβολα x_6 και x_7. Αυτά τα δύο σύμβολα συνδυάζονται όπως φαίνεται στο Σχ. 2-3-4, όπου στον ανώτερο κλάδο καταχωρείται ένα 0 και στον ανώτερο ένα 1. Οι πιθανότητες των δύο αυτών κλάδων προστίθενται στον κόμβο στον οποίο συναντώνται οι δύο κλάδοι και προκύπτει η πιθανότητα 0.01. Τώρα έχουμε τα σύμβολα της πηγής και $x_1, x_2, \ldots, x_5$ συν ένα νέο σύμβολο, ας πούμε το x_6', που προέκυψε από τον συνδυασμό των x_6 και x_7 . Το επόμενο βήμα είναι να ενώσουμε τα δύο λιγότερο πιθανά σύμβολα από το σύνολο x_1, x_2, x_3, x_4, x_5, x_6'. Αυτά είναι τα x_5 και x_6', τα οποία προσαρμοσμένα δίνουν μία πιθανότητα 0.05. Στον κλάδο από το x_5 καταχωρείται ένα 0 και στον κλάδο από το x_6' καταχωρείται ένα 1. Η διαδικασία αυτή συνεχίζεται μέχρι να εξαντληθεί το σύνολο των πιθανών γραμμάτων πηγής. Το αποτέλεσμα είναι ένα κωδικό δένδρο με κλάδους οι οποίοι περιέχουν τις επιθυμητές κωδικές λέξεις. Οι κωδικές λέξεις λαμβάνονται ξεκινώντας από τον δεξιότερο κόμβο στο δένδρο και συνεχίζοντας προς τα αριστερά. Οι προκύπτουσες κωδικές λέξεις απαριθμούνται στο Σχ. 2-3-4. Ο μέσος αριθμός των δυαδικών ψηφίων ανά σύμβολο για τον κώδικα αυτόν είναι $\overline{R}$ = 2.21 bits ανά σύμβολο. Η εντροπία της πηγής είναι 2.11 bits ανά σύμβολο.

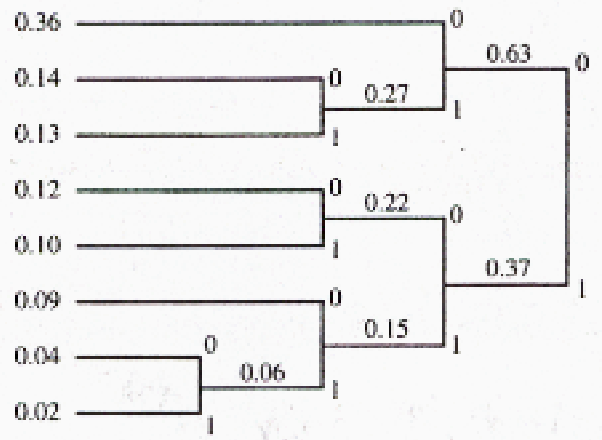


Letter	Code
x_1	0
x_2	10
x_3	110
x_4	1110
x_5	11110
x_6	111110
x_7	111111

$\overline{R} = 2.21$

ΣΧΗΜΑ 2-3-5 Ένας εναλλακτικός κώδικας για την ΔΠΧΜ του Παραδείγματος 2-3-1.

Παρατηρούμε ότι ο κώδικας δεν είναι απαραίτητα αμφιμονοσήμαντος. Για παράδειγμα, στο προτελευταίο βήμα της διαδικασίας κωδικοποίησης, έχουμε μία ισοπαλία μεταξύ των x_1 και x_3', επειδή τα σύμβολα αυτά είναι ισοπίθανα. Στο σημείο αυτό, επιλέγουμε να συνδυάσουμε το x_1 με το x_2. Ένας εναλλακτικός τρόπος είναι να συνδυάσουμε το x_2 και x_3'. Εάν επιλέξουμε τον συνδυασμό αυτό, ο προκύπτων κώδικας φαίνεται στο Σχ. 2-3-5. Ο μέσος αριθμός bits ανά σύμβολο πηγής για τον κώδικα αυτόν είναι επίσης 2.21. Κατά συνέπεια, οι προκύπτοντες κώδικες είναι εξ ίσου αποδοτικοί. Κατά δεύτερον, η εκχώρηση ενός 0 στον ανώτερο κλάδο και ενός 1 στον λιγότερο πιθανό κατώτερο κλάδο είναι αυθαίρετη. Μπορούμε απλώς να αντιστρέψουμε την εκχώρηση του 0 και του 1 και πάλι να πάρουμε έναν αποδοτικό κώδικα που να ικανοποιεί την συνθήκη προθέματος.



Letter	Code
x_1	00
x_2	010
x_3	011
x_4	100
x_5	101
x_6	110
x_7	1110
x_8	1111
$H(X) = 2.63$	$\overline{R} = 2.70$

ΣΧΗΜΑ 2-3-6. Ένας Huffman κώδικας για το Παράδειγμα 2-3-2.

Παράδειγμα 2-3-2

Ως ένα δεύτερο παράδειγμα, ας καθορίσουμε τον κώδικα Huffman για την έξοδο της ΔΠΧΜ που παρουσιάζεται στο Σχ. 2-3-6. Η εντροπία της πηγής αυτής είναι $H(X)$ = 2.63 bits /σύμβολο. Ο κώδικας Huffman, όπως φαίνεται στο Σχ. 2-3-6, έχει ένα μέσο μήκος $\overline{R}$ = 2.70 bits / σύμβολο. Έτσι, η αποδοτικότητά του είναι 0.97.

Ο αλγόριθμος κωδικοποίησης μεταβλητού μήκους (Huffman) που περιγράφεται στα παραπάνω παραδείγματα δημιουργεί έναν κώδικα προθέματος που έχει έναν $\overline{R}$ που ικανοποιεί την (2-3-9). Ωστόσο, αντί να κωδικοποιούμε ανά σύμβολο, μία αποδοτικότερη διαδικασία είναι να κωδικοποιούμε τμήματα από J σύμβολα κάθε φορά. Σε μία τέτοια περίπτωση, τα όρια του δεύτερου θεωρήματος κωδικοποίησης στην (2-3-9) γίνονται

$$JH(X) \le \overline{R_J} < JH(X)+1, \qquad (2\text{-}3\text{-}12)$$

επειδή η εντροπία ενός τμήματος των J συμβόλων από μία ΔΠΧΜ είναι $JH(X)$, και το $\overline{R_J}$ είναι ο μέσος αριθμός bits ανά τμήμα των J συμβόλων. Εάν διαιρέσουμε την (2-3-12) με J, παίρνουμε

$$H(X) \le \frac{\overline{R_J}}{J} < H(X)+\frac{1}{J} \qquad (2\text{-}3\text{-}13)$$

όπου $\overline{R_J}\,/\,J \equiv \overline{R}$ είναι ο μέσος αριθμός bits ανά σύμβολο πηγής. Έτσι, το $\overline{R}$ μπορεί να προσεγγίσει όσο επιθυμούμε το $H\,(X)$ αν το J επιλεγεί αρκούντως μεγάλο.

Παράδειγμα 2-3-3

Η έξοδος μίας ΔΠΧΜ συνίσταται από τα *γράμματα* x_1, x_2, και x_3 με πιθανότητες 0.45, 0.35 και 0.20, αντιστοίχως. Η εντροπία της πηγής αυτής είναι $H(X)$ = 1.518 bits / σύμβολο. Ο Huffman κώδικας για την πηγή αυτή, ο οποίος δίνεται στον Πίνακα 2-3-2, απαιτεί $\overline{R_1}$ = 1.55 bits / σύμβολο και επιτυγχάνει μία αποδοτικότητα 97.9%. Εάν κωδικοποιήσουμε ζεύγη συμβόλων με την βοήθεια του αλγόριθμου Huffman, ο προκύπτων κώδικας δίνεται στον Πίνακα 2-3-3. Η εντροπία της εξόδου της πηγής για ζεύγη γραμμάτων είναι $2H(X)$ = 3.036 bits ανά ζεύγος συμβόλων. Από την άλλη πλευρά, ο κώδικας Huffman απαιτεί $\overline{R_2}$ = 3.0675 bits / ζεύγος συμβόλων Συνεπώς, η αποδοτικότητα της κωδικοποίησης αυξάνει στο $2H(X)$ / $\overline{R_2}$ ή, ισοδύναμα, στο 99%.

ΠΙΝΑΚΑΣ 2-3-2 Ο ΚΩΔΙΚΑΣ HUFFMAN ΓΙΑ ΤΟ ΠΑΡΑΔΕΙΓΜΑ 2-3-3

Γράμμα	Πιθανότητα	Αυτο-πληροφορία	Κώδικας
x_1	0.45	1.156	1
x_2	0.35	1.520	00
x_3	0.20	2.330	01

H(X)=1.518 bits/γράμμα

$\overline{R}_1$=1.55 bits/ γράμμα

Αποδοτικότητα =97.9%

ΠΙΝΑΚΑΣ 2-3-3 Ο ΚΩΔΙΚΑΣ HUFFMAN ΓΙΑ ΤΗΝ ΚΩΔΙΚΟΠΟΙΗΣΗ ΖΕΥΓΩΝ ΓΡΑΜΜΑΤΩΝ

Ζεύγος γραμμάτων	Πιθανότητα	Αυτο-Πληροφορία	Κώδικας
x_1x_1	0.2025	2.312	10
x_1x_2	0.1575	2,676	001
x_2x_1	0.1575	2.676	010
x_2x_2	0.1225	3.039	011
x_1x_3	0.09	3.486	111
x_3x_1	0.09	3.486	0000
x_2x_3	0.07	3.850	0001
x_3x_2	0.07	3.850	1100
x_3x_3	0.04	4.660	1101

$2H(X)$ = 3.036 bits/ζεύγος γραμμάτων

$\overline{R}_2$ = 3.0675 bits/ζεύγος γραμμάτων

½ $\overline{R}_2$ = 1.534 bits/ γράμμα

Αποδοτικότητα = 99.0 %

Συνοπτικά, δείξαμε ότι μπορεί να γίνει αποδοτική κωδικοποίηση μίας ΔΠΧΜ ανά σύμβολο με την βοήθεια ενός κώδικα μεταβλητού μήκους βάσει του αλγόριθμου του Huffman. Περαιτέρω, η αποδοτικότητα της διαδικασίας κωδικοποίησης αυξάνεται με την κωδικοποίηση τμημάτων των J συμβόλων κάθε φορά. Έτσι, η έξοδος μίας ΔΠΧΜ με εντροπία $H(X)$ μπορεί να κωδικοποιηθεί με έναν κώδικα μεταβλητού μήκους με έναν μέσο αριθμό bits ανά γράμμα πηγής που προσεγγίζει το $H(X)$ όσον επιθυμούμε.

2-1.4 Οι Διακριτές Στάσιμες Πηγές

Στην προηγούμενη ενότητα, περιγράψαμε την αποδοτική κωδικοποίηση της εξόδου μίας ΔΠΧΜ. Στην ενότητα αυτή, θα ασχοληθούμε με διακριτές πηγές, για τις οποίες η ακολουθία των γραμμάτων εξόδου είναι στατιστικώς εξηρτημένη. Περιορίζουμε την ανάλυσή μας σε πηγές που είναι στατιστικώς στάσιμες.

Ας υπολογίσουμε την εντροπία οποιασδήποτε ακολουθίας γραμμάτων από μία στάσιμη πηγή. Από τον ορισμό της (2-2-13) και το αποτέλεσμα που δίνεται στην (2-2-15), η εντροπία ενός τμήματος τυχαίων μεταβλητών $X_1X_2\ldots X_\kappa$ είναι

$$H(X_1X_2...X_k) = \sum_{i=1}^{k} H(X_i \mid X_1X_2...X_{i-1}) \tag{2-3-14}$$

όπου $H\,(X_i \mid X_1, X_2\ldots X_{i-1})$ είναι η υπό συνθήκη εντροπία του i-οστού συμβόλου από την πηγή με δεδομένα τα προηγούμενα $i-1$ σύμβολα. Η ανά γράμμα εντροπία για το τμήμα των k συμβόλων ορίζεται ως

$$H_k(X) = \frac{1}{k} H(X_1X_2...X_k) \tag{2-3-15}$$

Ορίζουμε το πληροφοριακό περιεχόμενο μίας στάσιμης πηγής ως το όριο της ανά γράμμα εντροπίας στην (2-3-15) καθώς $k\to\infty$. Ήτοι,

$$H_\infty(X) = \lim_{k\to\infty} H_k(X) = \lim_{k\to\infty} \frac{1}{k} H(X_1X_2...X_k) \tag{2-3-16}$$

Η ύπαρξη του ορίου αυτού επαληθεύεται παρακάτω.

Εναλλακτικά, μπορούμε να ορίσουμε την ανά γράμμα της πηγής εντροπία με την βοήθεια του ορίου της υπό συνθήκην εντροπίας $H\,(X_k \mid X_1, X_2\ldots X_{k-1})$ καθώς το k τείνει προς το άπειρο. Ευτυχώς, το όριο αυτό επίσης υφίσταται και ταυτίζεται με το όριο της (2-3-16). Δηλαδή,

$$H_\infty(X) = \lim_{k\to\infty} H(X_k \mid X_1X_2...X_{k-1}) \tag{2-3-17}$$

Το αποτέλεσμα αυτό επίσης επαληθεύεται παρακάτω. Η ανάπτυξή μας θα ακολουθήσει την προσέγγιση του Gallagher (1968).

Πρώτα, θα δείξουμε ότι

$$H(X_k \mid X_1X_2...X_{k-1}) \le H(X_{k-1} \mid X_1X_2...X_{k-2}) \tag{2-3-18}$$

για $k \ge 2$. Από το προηγούμενό μας αποτέλεσμα, ότι η δέσμευση ως προς μία τυχαία μεταβλητή δεν αυξάνει την εντροπία, έχουμε

$$H(X_k \mid X_1X_2...X_{k-1}) \le H(X_k \mid X_2X_3...X_{k-1}) \tag{2-3-19}$$

Από την στασιμότητα της πηγής, έχουμε

$$H(X_k \mid X_2X_3...X_{k-1}) = H(X_{k-1} \mid X_1X_2...X_{k-2}) \tag{2-3-20}$$

Έτσι, η (2-3-18) προκύπτει αμέσως. Το αποτέλεσμα αυτό δείχνει ότι η $H(X_k \mid X_1X_2\ldots X_{k-1})$ είναι μία μη αύξουσα ακολουθία ως προς το k.

Δεύτερον, έχουμε το αποτέλεσμα

$$H_k(X) \ge H(X_k \mid X_1X_2...X_{k-1}) \tag{2-3-21}$$

το οποίο έπεται αμέσως από τις (2-3-14) και (2-3-15) και το γεγονός ότι ο τελευταίος όρος του αθροίσματος της (2-3-14) είναι ένα κάτω όριο για τον καθένα από τους άλλους $k-1$ όρους.

Τρίτον, από τον ορισμό του $H_k(X)$, μπορούμε να γράψουμε

$$\begin{aligned} H_k(X) &= \frac{1}{k}\left[H(X_1X_2...X_{k-1}) + H(X_k \mid X_1X_2...X_{k-1})\right] \\ &= \frac{1}{k}\left[(k-1)H_{k-1}(X) + H(X_k \mid X_1X_2...X_{k-1})\right] \\ &\le \frac{k-1}{k} H_{k-1}(X) + \frac{1}{k} H_k(X) \end{aligned}$$

το οποίο απλοποιείται στο

$$H_k(X) \le H_{k-1}(X) \tag{2-3-22}$$

Συνεπώς, η $H_k(X)$ είναι μία μη αύξουσα συνάρτηση του k.

Εφ όσον η $H_k(X)$ και η υπό συνθήκη εντροπία $H(X_k| X_1, X_2 \ldots X_{k-1})$ είναι αμφότερες μη αρνητικές και μη αύξουσες ως προς το k, πρέπει να υπάρχουν αμφότερα τα όρια. Οι τύποι των ορίων αυτών μπορούν να γραφούν με την βοήθεια των (2-3-14) και (2-3-15) ώστε να εκφρασθεί η $H_{k+j}(X)$ ως

$$\begin{aligned} H_{k+j}(X) &= \frac{1}{k+j} H(X_1 X_2 \ldots X_{k-1}) \\ &+ \frac{1}{k+j}[H(X_k \mid X_1 X_2 \ldots X_{k-1}) + H(X_{k+1} \mid X_1 X_2 \ldots X_k) \\ &+ \ldots + H(X_{k+j} \mid X_1 X_2 \ldots X_{k+j-1})] \end{aligned}$$

Επειδή η υπό συνθήκη εντροπία είναι μη αύξουσα, ο πρώτος όρος μέσα στις ορθογώνιες αγκύλες χρησιμεύει ως ένα άνω όριο των άλλων όρων. Έτσι,

$$H_{k+j}(X) \le \frac{1}{k+j} H(X_1 X_2 \ldots X_{k-1}) + \frac{j+1}{k+j} H(X_k \mid X_1 X_2 \ldots X_{k-1}) \tag{2-3-23}$$

Για ένα καθορισμένο k, το όριο της (2-3-23) καθώς $j \to \infty$ έχει ως αποτέλεσμα

$$H_\infty(X) \le H(X_k \mid X_1 X_2 \ldots X_{k-1}) \tag{2-3-24}$$

Όμως, η (2-3-24) ισχύει για όλα τα k. Συνεπώς, θα ισχύει και για $k \to \infty$. Άρα,

$$H_\infty(X) \le \lim_{k \to \infty} H(X_k \mid X_1 X_2 \ldots X_{k-1}) \tag{2-3-25}$$

Από την άλλη πλευρά, από την (2-3-21), καθώς $k \to \infty$, παίρνουμε το όριο

$$H_\infty(X) \ge \lim_{k \to \infty} H(X_k \mid X_1 X_2 \ldots X_{k-1}) \tag{2-3-26}$$

το οποίο επαληθεύει την (2-3-17).

Τώρα υποθέστε ότι έχουμε μία διακριτή στάσιμη πηγή η οποία εκπέμπει J γράμματα με ανά γράμμα εντροπία $H_J(X)$. Μπορούμε να κωδικοποιήσουμε μία ακολουθία J γραμμάτων με έναν κώδικα Huffman μεταβλητού μήκους, ο οποίος ικανοποιεί την συνθήκη προθέματος, εάν ακολουθήσουμε την διαδικασία που περιγράφηκε στην προηγούμενη ενότητα. Ο προκύπτων κώδικας, έχει έναν μέσο αριθμό bits για το τμήμα των J γραμμάτων, ο οποίος ικανοποιεί την συνθήκη

$$H(X_1 \ldots X_J) \le \overline{R}_J < H(X_1 \ldots X_J) + 1 \tag{2-3-27}$$

Διαιρώντας κάθε όρο της (2-3-27) με το J, παίρνουμε τα όρια του μέσου αριθμού $\overline{R} = \overline{R}_J / J$ των bits ανά σύμβολο πηγής ως

$$H_J(X) \le \overline{R} < H_J(X) + \frac{1}{J} \tag{2-3-28}$$

Αυξάνοντας το μέγεθος του τμήματος J, μπορούμε να προσεγγίσουμε όσον επιθυμούμε την $H_J(X)$, και, καθώς $J \to \infty$, το $\overline{R}$ ικανοποιεί την

$$H_\infty(X) \le \overline{R} < H_\infty(X) + \varepsilon \tag{2-3-29}$$

όπου το ε τείνει προς το μηδέν με το $1/J$. Έτσι, η αποδοτική κωδικοποίηση των στάσιμων πηγών πραγματοποιείται με την κωδικοποίηση μεγάλων τμημάτων συμβόλων σε κωδικές λέξεις. Θα έπρεπε, ωστόσο, να τονίσουμε ότι η σχεδίαση του Huffman κώδικα απαιτεί την γνώση της από κοινού σ.π.π. των τμημάτων των J συμβόλων.

2-1.5 Ο Αλγόριθμος Lempehl-Ziv

Στην συζήτηση που προηγήθηκε, παρατηρήσαμε ότι ο αλγόριθμος κωδικοποίησης Huffman αποφέρει βέλτιστους κώδικες πηγής, υπό την έννοια ότι οι κωδικές λέξεις ικανοποιούν την συνθήκη προθέματος και ότι ελαχιστοποιείται το μέσο μήκος του τμήματος. Για να σχεδιάσουμε έναν κώδικα Huffman για μία ΔΠΧΜ, πρέπει να γνωρίζουμε τις πιθανότητες εμφάνισης όλων των γραμμάτων πηγής. Κατά την περίπτωση των διακριτών πηγών με μνήμη, πρέπει να γνωρίζουμε τις από κοινού πιθανότητες των τμημάτων μήκους $n \geq 2$. Ωστόσο, στην πράξη, τα στατιστικά στοιχεία της εξόδου μίας πηγής είναι συνήθως άγνωστα. Θεωρητικά, είναι δυνατόν να υπολογισθούν οι πιθανότητες της διακριτής πηγής με την απλή παρατήρηση μίας μεγάλης πληροφοριακής ακολουθίας που εκπέμπεται από την πηγή και την εμπειρική εξαγωγή των στατιστικών στοιχείων. Αν εξαιρέσουμε τον υπολογισμό των οριακών πιθανοτήτων $\{p_k\}$ οι οποίες αντιστοιχούν στην συχνότητα εμφάνισης των μεμονωμένων γραμμάτων στην έξοδο της πηγής, η εκτίμηση των από κοινού πιθανοτήτων απαιτεί ιδιαιτέρως υψηλή υπολογιστική πολυπλοκότητα. Κατά συνέπεια, για πολλές πραγματικές πηγές με μνήμη, δεν ενδείκνυται η εφαρμογή της μεθόδου κωδικοποίησης Huffman για την κωδικοποίησή τους.

Σε αντίθεση με τον αλγόριθμο κωδικοποίησης Huffman, ο αλγόριθμος κωδικοποίησης πηγής Lempel-Ziv έχει σχεδιασθεί έτσι ώστε να είναι ανεξάρτητος από τα στατιστικά στοιχεία της πηγής. Έτσι, ο αλγόριθμος Lempel-Ziv ανήκει στην κατηγορία των *καθολικών αλγορίθμων κωδικοποίησης πηγής* (*universal source coding algorithms*). Είναι ένας αλγόριθμος μεταβλητού-προς-σταθερό μήκος, στον οποίο η κωδικοποίηση γίνεται με τον τρόπο που περιγράφεται ακολούθως.

Στον αλγόριθμο Lempel-Ziv, η ακολουθία που παρουσιάζεται στην έξοδο της διακριτής πηγής αναλύεται σε τμήματα μεταβλητού μήκους, τα οποία ονομάζονται *φράσεις* (*phrases*). Μία νέα φράση δημιουργείται κάθε φορά που ένα τμήμα από γράμματα διαφέρει από κάποια προηγούμενη φράση στο τελευταίο γράμμα. Οι φράσεις απαριθμούνται σε ένα λεξικό, το οποίο καταχωρεί την θέση των υφισταμένων φράσεων. Κατά την κωδικοποίηση μίας νέας φράσης, απλώς καθορίζουμε την θέση της υφιστάμενης φράσης στο λεξικό και την επεκτείνουμε με το νέο γράμμα.

Ως ένα παράδειγμα, θεωρείστε την δυαδική ακολουθία

10101101001001110101000011001110101100011011

Η ανάλυση της ακολουθίας όπως περιγράφεται παραπάνω παράγει τις ακόλουθες φράσεις:

1, 0, 10, 11, 01, 00, 100, 111, 010, 1000, 011, 001, 110, 101, 10001, 1011

ΠΙΝΑΚΑΣ 2-3-4 ΕΝΑ ΛΕΞΙΚΟ ΓΙΑ ΤΟΝ ΑΛΓΟΡΙΘΜΟ LEMPEL-ZIV

	Θέση Λεξικού	**Περιεχόμενα Λεξικού**	**Κωδική Λέξη**
1	0001	1	00001
2	0010	0	00000
3	0011	10	00010
4	0100	11	00011
5	0101	01	00101
6	0110	00	00100
7	0111	100	00110
8	1000	111	01001
9	1001	010	01010
10	1010	1000	01110
11	1011	011	01011
12	1100	001	01101
13	1101	110	01000
14	1110	101	00111
15	1111	10001	10101
16		1011	11101

Παρατηρούμε ότι κάθε φράση της ακολουθίας είναι μία επέκταση μία προηγούμενης φράσης με ένα νέο γράμμα εξόδου από την πηγή. Για να κωδικοποιήσουμε τις φράσεις, κατασκευάζουμε ένα λεξικό όπως φαίνεται στον Πίνακα 2-3-4. Οι θέσεις του λεξικού έχουν αριθμηθεί διαδοχικά, αρχίζοντας με το 1 και μετρώντας προς τα πάνω μέχρι, στην περίπτωση αυτή, το 16 που είναι ο αριθμός των φράσεων στην ακολουθία. Απαριθμούνται επίσης οι διαφορετικές φράσεις που αντιστοιχούν σε κάθε θέση. Οι κωδικές λέξεις καθορίζονται με την απαρίθμηση της θέσης του λεξικού (σε δυαδική μορφή) της προηγούμενης φράσης που ταιριάζει με την νέα φράση σε όλα εκτός από την τελευταία θέση. Τότε, το νέο γράμμα εξόδου προστίθεται στο τέλος της θέσης που κατέχει στο λεξικό η προηγούμενη φράση. Αρχικώς, η θέση 0000 χρησιμοποιείται για να κωδικοποιηθεί μία φράση που δεν έχει εμφανισθεί προηγουμένως.
Ο αποκωδικοποιητής πηγής του κώδικα αυτού, στο άλλο άκρο του επικοινωνιακού συστήματος, κατασκευάζει έναν ταυτόσημο πίνακα και αποκωδικοποιεί κατάλληλα την λαμβανομένη ακολουθία.
Πρέπει να παρατηρηθεί ότι ο πίνακας κωδικοποίησε 44 bits εισόδου σε 16 κωδικές λέξεις των πέντε bits η κάθε μία, με αποτέλεσμα 80 κωδικά bits. Συνεπώς, ο αλγόριθμος δεν έκανε καμία συμπίεση δεδομένων. Εν τούτοις, η έλλειψη αποδοτικότητας οφείλεται στο γεγονός ότι η ακολουθία που μελετήσαμε είναι πολύ μικρή. Καθώς μεγαλώνει το μήκος της ακολουθίας, η διαδικασία κωδικοποίησης γίνεται αποδοτικότερη και επιφέρει μία συμπιεσμένη ακολουθία στην έξοδο της πηγής.
Πώς επιλέγουμε το συνολικό μήκος του πίνακα; Γενικά, όσο μεγάλος και αν είναι ο πίνακας, κάποια στιγμή θα υπερχειλίσει. Για να λυθεί το πρόβλημα υπερχείλισης, ο κωδικοποιητής πηγής και ο αποκωδικοποιητής πηγής πρέπει να συμφωνήσουν να απομακρύνουν από τα αντίστοιχα λεξικά τους τις φράσεις που δεν είναι χρήσιμες και να τις αντικαταστήσουν με νέες φράσεις.
Ο αλγόριθμος Lempel-Ziv χρησιμοποιείται ευρέως στην συμπίεση αρχείων υπολογιστών. Τα εργαλεία "compress" (συμπίεσε) και "uncompress" (αποσυμπίεσε) που λειτουργούν υπό το λειτουργικό σύστημα UNIX© και πολυάριθμοι αλγόριθμοι υπό το λειτουργικό MS-DOS αποτελούν υλοποιήσεις διαφορετικών παραλλαγών του αλγορίθμου αυτού.