ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ
ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Θεωρία της Πληροφορίας

Συγγραφέας
Δρ. Ιωάννης Ρέκανος

Διδάσκων
Δρ. Αναστάσιος Πολίτης

Σεπτέμβριος 2014

Άδειες Χρήσης

Χρηματοδότηση

Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.

Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο ΤΕΙ Σερρών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.

Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Περιεχόμενα

Κατάλογος Σχημάτων

Κατάλογος Πινάκων

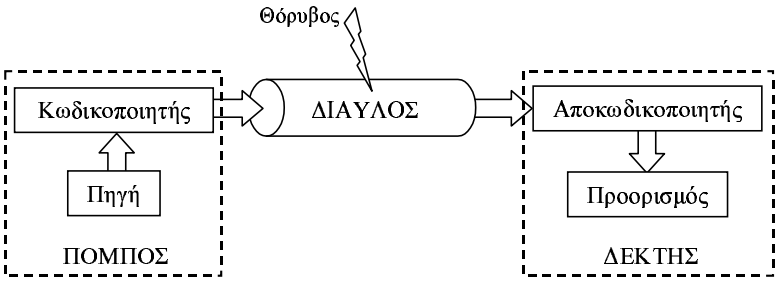
Κεφάλαιο 1

Εισαγωγή

Ίσως ο εικοστός αιώνας να χαρακτηριστεί από τους ιστορικούς του μέλλοντος ως ο αιώνας της Πληροφορικής και των Επικοινωνιών. Στον αιώνα αυτό διαπιστώθηκε μία ραγδαία ανάπτυξη των μέσων καταγραφής, αποθήκευσης, επεξεργασίας και μετάδοσης των δεδομένων καθώς και των συναφών τεχνολογιών επικοινωνίας όπως το τηλέφωνο, το ραδιόφωνο, η τηλεόραση και τα δίκτυα υπολογιστών. Κοινός τόπος όλων αυτών των τεχνολογικών επιτευγμάτων είναι η ακριβής, ταχεία, ασφαλής και οικονομική αποθήκευση και μετάδοση των πληροφοριών.

Όλη αυτή η ανάπτυξη έδωσε τους περισσότερους καρπούς μετά τη μαθηματική θεμελίωση της έννοιας της πληροφορίας. Πριν τα μέσα του εικοστού αιώνα, η έννοια της πληροφορίας ήταν κατά βάση αφηρημένη και ποιοτική. Συνεπώς, οποιαδήποτε προσπάθεια εξαγωγής νόμων που διέπουν την πληροφορία και την επικοινωνία ήταν αρχικά αδύνατη. Αυτή η αδυναμία ποσοτικοποίησης της έννοιας της πληροφορίας αντιμετωπίστηκε με απόλυτη επιτυχία χάρη στις πρωτοποριακές εργασίες του Shannon στο δυτικό και του Khinchin στο ανατολικό μπλοκ.

Το 1948, με την εργασία "A Mathematical Theory of Communication" του Shannon γεννήθηκε μια νέα επιστημονική περιοχή, η **Θεωρία της Πληροφορίας** ή **Θεωρία Πληροφοριών**. Στόχος της είναι η θεμελίωση εννοιών και θεωρημάτων που επιτρέπουν τη μαθηματική περιγραφή της διαδικασίας της επικοινωνίας. Έτσι, η μετάδοση πληροφοριών μπορεί να αναλυθεί με ακρίβεια και μαθηματική αυστηρότητα, ενώ σε ένα επόμενο βήμα είναι δυνατό να σχεδιαστούν καλύτερα συστήματα επικοινωνιών. Η νέα θεωρία βασισμένη στη θεωρία πιθανοτήτων, τη στατιστική και την άλγεβρα μπορεί να απαντήσει με μαθηματική αυστηρότητα σε ερωτήματα που σχετίζονται με τη βέλτιστη συμπίεση των δεδομένων, την περιγραφή των διαύλων επικοινωνίας, την κωδικοποίηση των μηνυμάτων πληροφορίας, το ρυθμό μετάδοσης των πληροφοριών



Σχήμα 1.1: *Σύστημα επικοινωνίας.*

σε περιβάλλον θορύβου, την κρυπτογράφηση κ.α.

Αν και αρχικά η θεωρία της πληροφορίας αποτέλεσε τμήμα της επιστήμης των επικοινωνιών, από πολύ νωρίς οι αρχές, οι μεθοδολογίες και τα συμπεράσματά της βρήκαν εφαρμογή και σε άλλες επιστημονικές περιοχές. Ενδεικτικά, αναφέρουμε τη βιολογία, τη γενετική, τη γλωσσολογία, τη σχεδίαση υπολογιστικών συστημάτων, τη διοίκηση, τη θεωρία παιγνίων κ.α. Έτσι, μπορούμε να πούμε ότι στις μέρες μας, η θεωρία της πληροφορίας λόγω της μαθηματικής της αυστηρότητας και των γενικευμένων συμπερασμάτων της αποτελεί ένα χωριστό κλάδο των μαθηματικών.

Σύμφωνα με τη θεωρία πληροφοριών ένα σύστημα επικοινωνίας περιλαμβάνει τον πομπό, το δίαυλο και το δέκτη (σχήμα 1.1). Ο πομπός απαρτίζεται από την πηγή πληροφορίας και τον κωδικοποιητή. Η πηγή παράγει την πληροφορία που περιγράφεται με τη χρήση συμβόλων, τα οποία στη συνέχεια κωδικοποιούνται από τον κωδικοποιητή. Το κωδικοποιημένο μήνυμα διοχετεύεται στο δίαυλο, ο οποίος είναι στην ουσία το μέσο που παρεμβάλλεται μεταξύ του πομπού και του δέκτη. Όταν η πληροφορία διαπερνά το δίαυλο είναι δυνατό να αλλοιωθεί λόγω της παρουσίας θορύβου. Τέλος, η πληροφορία λαμβάνεται από τον δέκτη όπου αρχικά αποκωδικοποιείται και στη συνέχεια παρουσιάζεται στον προορισμό της.

1.1 Βασικές Έννοιες και Παραδείγματα

Κεντρικό ρόλο στη θεωρία της πληροφορίας παίζει η ίδια η έννοια της πληροφορίας. Σύμφωνα με τη θεωρία, η πληροφορία έχει ποσοτικό χαρακτήρα και συνεπώς διαφέρει σημαντικά από το εννοιολογικό περιεχόμενο που της αποδίδουμε στην καθημερινή μας ζωή. Πλέον, η πληροφορία ενός γεγονότος σχετίζεται άμεσα με την πιθανότητα του γεγονότος και μόνο με αυτήν. Πρακτικά, όσο πιο μικρή είναι η πιθανότητα ενός γεγονότος τόσο μεγαλύτερη είναι η πληροφορία που αντιστοιχεί στο γεγονός. Αυτό εν μέρει συμφωνεί με την κοινή μας αντίληψη. Για παράδειγμα, το γεγονός *'Σήμερα*

έγινε ολική έκλειψη ηλίου' περιέχει μεγαλύτερη πληροφορία από το γεγονός *'Σήμερα ο ήλιος ανέτειλε'*, επειδή το πρώτο γεγονός έχει μικρή πιθανότητα σε σχέση με το δεύτερο που είναι βέβαιο. Πράγματι, τα μέσα μαζικής ενημέρωσης θα αναφερθούν εκτενώς στο πρώτο γεγονός, ενώ δεν υπάρχει περίπτωση να ασχοληθούν με το δεύτερο. Όμως, αυτή η συμφωνία μεταξύ θεωρίας της πληροφορίας και κοινής αντίληψης δεν είναι κανόνας. Αυτό μπορεί να φανεί με ένα παράδειγμα. Έστω ότι τη διοργάνωση της επόμενης Ολυμπιάδας τη διεκδικούν δύο πόλεις, η Α και η Β. Όσο και να μας φαίνεται παράξενο, σύμφωνα με τη θεωρία της πληροφορίας, το γεγονός *'Η πόλη Α θα οργανώσει την επόμενη Ολυμπιάδα'* περιέχει μικρότερη πληροφορία από το γεγονός *'Έριξα τα ζάρια και έφερα έξι-πέντε'*. Ο λόγος είναι ότι το πρώτο γεγονός έχει πιθανότητα 0.5, ενώ το δεύτερο αρκετά μικρότερη (1/18).

Συνεπώς, στα πλαίσια της θεωρίας της πληροφορίας η λέξη πληροφορία έχει καθαρά ποσοτικό περιεχόμενο και όχι ποιοτικό.

Ακόμη, η πληροφορία που περιέχει ένα γεγονός είναι ένα μέτρο της αβεβαιότητας που έχουμε για το γεγονός αυτό πριν την εμφάνισή του. Αυτό μπορεί να γίνει πιο κατανοητό με ένα παράδειγμα. Έστω ένας επενδυτής, ο οποίος χρειάζεται *πληροφορίες* (συμβουλές) για το επίπεδο των μετοχών, μίας συγκεκριμένης επιχείρησης, που έχει στην κατοχή του, και συμβουλεύεται το χρηματιστή του. Προφανώς, ο χρηματιστής θα πρέπει να έχει ειδικές *πληροφορίες* (γνώσεις) για τα θέματα του χρηματιστηρίου. Ο χρηματιστής *πληροφορεί* (ενημερώνει) τον επενδυτή ότι ορκωτοί λογιστές αναζητούν *πληροφορίες* (αποδεικτικά στοιχεία) για πιθανή απάτη στα ετήσια λογιστικά στοιχεία που παρουσίασε η συγκεκριμένη επιχείρηση. Αντιδρώντας σε αυτές τις *πληροφορίες* (δεδομένα), ο επενδυτής αποφασίζει να πουλήσει τις μετοχές της συγκεκριμένης επιχείρησης και *πληροφορεί* (ενημερώνει) ανάλογα τον χρηματιστή του.

Βλέπουμε από το παραπάνω παράδειγμα, ότι η πληροφορία μπορεί να σημαίνει συμβουλή, γνώση, ενημέρωση, δεδομένα, στοιχεία κ.α. Σε όλες όμως τις περιπτώσεις η πληροφορία είναι μέτρο της αβεβαιότητας και όταν αυτή παρέχεται μειώνει την αβεβαιότητα. Πράγματι, ο επενδυτής επειδή είναι αβέβαιος για το πραγματικό επίπεδο των μετοχών που κατέχει συμβουλεύεται το χρηματιστή του, ο οποίος γνωρίζει τα χρηματιστηρικά θέματα με μεγαλύτερη βεβαιότητα. Ο χρηματιστής μειώνει την αβεβαιότητα του επενδυτή ενημερώνοντάς τον για τις κινήσεις των ορκωτών λογιστών. Από τη μεριά τους, οι ορκωτοί λογιστές είχαν αβεβαιότητα για τα λογιστικά στοιχεία της επιχείρησης και για να μειώσουν την αβεβαιότητά τους διεξάγουν την έρευνα. Έχοντας ενημερωθεί, ο επενδυτής αποφασίζει να πωλήσει τις μετοχές του και ενημερώνει ανάλογα το χρηματιστή του. Έτσι, μειώνει την αβεβαιότητα του χρηματιστή, ο

οποίος δεν ήταν βέβαιος για τις προθέσεις του επενδυτή (θα πουλήσει, θα αγοράσει ή θα διατηρήσει τις ίδιες μετοχές;).

Μία άλλη μαθηματική έννοια με θεμελιώδη σημασία για τη θεωρία της πληροφορίας είναι η εντροπία. Ώς όρος η έννοια της εντροπία για τη θεωρία της πληροφορίας προέκυψε ως δάνειο από τη θερμοδυναμική και συγκεκριμένα από το δεύτερο νόμο της. Στα πλαίσια της θεωρίας της πληροφορίας, η εντροπία εκφράζει το μέσο όρο ανά γεγονός της πληροφορίας που μπορεί να προκύψει από διάφορα γεγονότα. Μάλιστα η μαθηματική της έκφραση είναι ίδια με αυτήν που περιγράφει την εντροπία στο χώρο της θερμοδυναμικής.

Όπως θα δούμε στη συνέχεια, η εντροπία εκφράζει ένα μέτρο της μέσης αβεβαιότητας για την εμφάνιση οποιουδήποτε γεγονότος που ανήκει σε ένα σύνολο πιθανών γεγονότων. Προφανώς, αν τα γεγονότα είναι ισοπίθανα, τότε η αβεβαιότητά μας για την εμφάνιση κάποιου γεγονότος θα είναι μέγιστη. Αντίθετα, αν κάποιο από τα πιθανά γεγονότα έχει σημαντικά μεγαλύτερη πιθανότητα εμφάνισης σε σχέση με τα υπόλοιπα, τότε είναι προφανές ότι η αβεβαιότητά μας για την εμφάνιση κάποιου συγκεκριμένου γεγονότος από το σύνολο των γεγονότων θα είναι μικρότερη σε σχέση με την περίπτωση ισοπίθανων γεγονότων.

Αυτό μπορεί να γίνει κατανοητό με το εξής παράδειγμα. Θεωρήστε ένα μη τίμιο νόμισμα όπου κατά τη ρίψη του η πιθανότητα να έρθει *κεφαλή* είναι 0.9 ενώ η πιθανότητα να έρθει *γράμματα* είναι 0.1. Σε αυτήν την περίπτωση μετά από μία ρίψη θα αναμένουμε την εμφάνιση της *κεφαλής* επειδή η αβεβαιότητά μας για το αποτέλεσμα της ρίψης είναι μικρή. Αντίθετα, αν το νόμισμα είναι τίμιο, δηλαδή τα γεγονότα *κεφαλή* και *γράμματα* είναι ισοπίθανα, τότε η αβεβαιότητά μας για το αποτέλεσμα είναι μεγάλη. Αυτό φαίνεται από το γεγονός ότι μετά από μία ρίψη δεν έχουμε καμία τεκμηριωμένη προσδοκία εμφάνισης μίας συγκεκριμένης πλευράς του νομίσματος.

Για τη θεωρία της πληροφορίας, το τίμιο νόμισμα έχει μεγαλύτερη εντροπία από το μη τίμιο. Δηλαδή, η αβεβαιότητα για το αποτέλεσμα της ρίψης του τίμιου νομίσματος είναι μεγαλύτερη από αυτήν που σχετίζεται με το μη τίμιο.

Κεφάλαιο 2

Πληροφορία και Εντροπία

2.1 Ορισμοί

Ορισμός 2.1. ***Πληροφορία*** *είναι μία συλλογή δεδομένων, τα οποία καταγράφονται με τη χρήση* ***συμβόλων****. (Για τα πλαίσια της θεωρίας της πληροφορίας, ο ορισμός αυτός της πληροφορίας είναι περιγραφικός και ελλιπής. Ο ακριβής και μαθηματικά θεμελιωμένος ορισμός της πληροφορίας δίνεται στην επόμενη ενότητα.)*

Ορισμός 2.2. ***Πηγή πληροφορίας*** *είναι κάθε σύστημα που παράγει στην έξοδό του πληροφορία.*

Ορισμός 2.3. ***Επικοινωνία*** *είναι κάθε διαδικασία μεταφοράς πληροφορίας μεταξύ δύο σημείων του χωροχρόνου.*

Ορισμός 2.4. *Τα διακεκριμένα (διαφορετικά) σύμβολα, $x_1, x_2, \ldots, x_N$ που χρησιμοποιούνται για την αναπαράσταση της πληροφορίας που παράγεται από μία πηγή συνθέτουν το σύνολο $X = \{x_1, x_2, \ldots, x_N\}$, το οποίο ονομάζεται* ***αλφάβητο της πηγής****. Το πλήθος N των διακεκριμένων συμβόλων μπορεί να είναι πεπερασμένο ή άπειρο.*

Ορισμός 2.5. *Ως* ***λέξη*** *ορίζουμε μια διατεταγμένη ακολουθία συμβόλων.*

Ορισμός 2.6. *Ως* ***μήνυμα*** *ορίζουμε μια διατεταγμένη ακολουθία λέξεων.*

2.2 Μέτρο της Πληροφορίας

Πρωταρχικός στόχος της θεωρίας πληροφοριών είναι η ποσοτικοποίηση της έννοιας της πληροφορίας. Εάν ο στόχος αυτός επιτευχθεί τότε μπορούμε να αναπτύξουμε

μεθόδους ανάλυσης και σχεδιασμού συστημάτων πληροφορίας με τη χρήση της αριθμητικής και του λογισμού. Το πρόβλημα της ποσοτικοποίησης της έννοιας της πληροφορίας και ο ορισμός ενός κατάλληλου μέτρου για τον υπολογισμό της απασχόλησε τον Hartley το 1928. Ο Hartley κατά τη μελέτη των τηλεγραφικών επικοινωνιών διαπίστωσε ότι αν η πιθανότητα εμφάνισης ενός γεγονότος είναι μεγάλη (κοντά στη μονάδα), τότε υπάρχει πολύ μικρή αβεβαιότητα για το αν θα συμβεί το γεγονός. Δηλαδή το γεγονός είναι αναμενόμενο. Σε αυτήν την περίπτωση, αν το γεγονός συμβεί, η πληροφορία που θα λάβουμε θα είναι μικρή.

Άρα, το μέτρο της πληροφορίας ενός τυχαίου γεγονότος, δηλαδή, η ποσοτική περιγραφή της σχετίζεται μόνο με την πιθανότητα του γεγονότος αυτού. Η διαίσθηση μας υποδεικνύει ότι όσο πιο πιθανό είναι ένα γεγονός τόσο μικρότερο θα είναι το μέτρο της πληροφορίας του. Αντίστροφα, ένα γεγονός με μικρή πιθανότητα συνοδεύεται από μεγάλη ποσότητα πληροφορίας. Επίσης, ο συνδυασμός δύο ανεξάρτητων γεγονότων θα πρέπει να αντιστοιχίζεται σε ποσότητα πληροφορίας ίση με το άθροισμα των ποσοτήτων πληροφορίας των δύο γεγονότων.

Συνεπώς, αν A είναι ένα τυχαίο γεγονός με πιθανότητα $p(A)$ και $I(A)$ είναι η συνάρτηση μέτρου της πληροφορίας του A, τότε η $I(A)$ θα πρέπει να ικανοποιεί τις παρακάτω ιδιότητες.

1. Το μέτρο $I(A)$ θα πρέπει να είναι συνάρτηση της πιθανότητας $p(A)$ ($p(A) \in [0,1]$)

$$I(A) = I(p(A)). \tag{2.1}$$

2. Το μέτρο $I(p(A))$ θα πρέπει να είναι πραγματική θετική συνάρτηση, δηλαδή

$$I(p) : [0,1] \to \mathbb{R}_+. \tag{2.2}$$

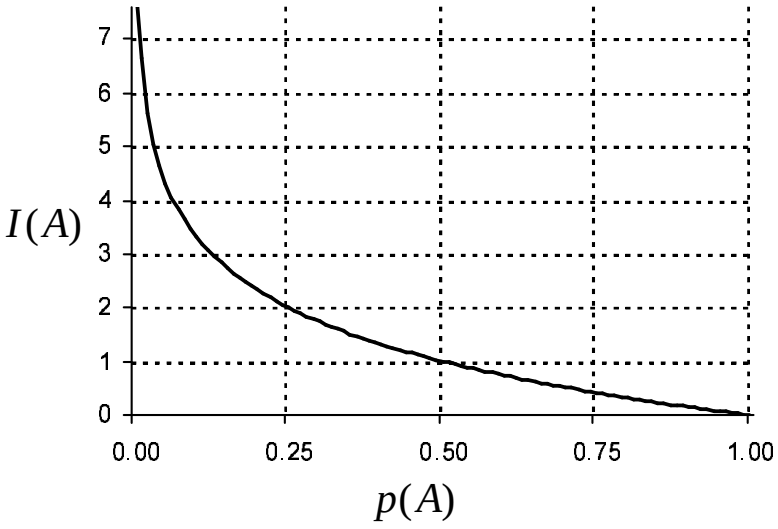
3. Η συνάρτηση $I(p)$ θα πρέπει να είναι γνησίως φθίνουσα,

$$\forall p(A), p(B) : p(A) > p(B) \Rightarrow I(p(A)) < I(p(B)), \tag{2.3}$$

δηλαδή όσο μικρότερη είναι η πιθανότητα του γεγονότος τόσο μεγαλύτερο το αντίστοιχο μέτρο πληροφορίας.

4. Τέλος αν τα γεγονότα A και B είναι ανεξάρτητα, οπότε $p(A \cap B) = p(A)p(B)$, τότε το μέτρο πληροφορίας του γεγονότος εμφάνισης και των δύο επιμέρους γεγονότων θα πρέπει να είναι ίσο με το άθροισμα των δύο επιμέρους μέτρων πληροφορίας, δηλαδή

$$I(A \cap B) = I(p(A)p(B)) = I(p(A)) + I(p(B)) = I(A) + I(B). \tag{2.4}$$



Σχήμα 2.1: *Γραφική παράσταση της συνάρτησης του μέτρου πληροφορίας για* $K = 2$.

Αποδεικνύεται (βλέπε Παράρτημα Α′) ότι η μοναδική συνάρτηση που ικανοποιεί τους περιορισμούς που θέτουν οι (2.1-2.4) είναι ο αρνητικός λογάριθμος της πιθανότητας με βάση μεγαλύτερη της μονάδας. Έτσι, υιοθετούμε τον παρακάτω ορισμό.

Ορισμός 2.7. *Αν* $p(A)$ *είναι η πιθανότητα του γεγονότος* A*, τότε το* ***μέτρο πληροφορίας*** *του* A *δίνεται από τη σχέση*

$$I(A) = -\log_K(p(A)), \text{ όπου } K > 1. \tag{2.5}$$

Στη συνέχεια για λόγους απλότητας ο όρος μέτρο πληροφορίας θα αναφέρεται ως **πληροφορία**. Όπως φαίνεται από την (2.5) η πληροφορία είναι αδιάστατο μέγεθος, ενώ η βάση K του λογαρίθμου μπορεί να επιλεγεί ελεύθερα, αρκεί $K > 1$. Συνήθως, η πληροφορία μετριέται σε bit, nat, ή hartley, όταν η βάση του λογαρίθμου είναι $K = 2$, $K = e$ (φυσικός λογάριθμος), ή $K = 10$, αντίστοιχα. Η επικρατέστερη μονάδα μέτρησης της πληροφορίας είναι το bit, λόγω της ραγδαίας ανάπτυξης της τεχνολογίας των ηλεκτρονικών υπολογιστών, των οποίων η λειτουργία σχετίζεται με το δυαδικό σύστημα αρίθμησης. Στο σχήμα 2.1 παρουσιάζεται η γραφική παράσταση του μέτρου πληροφορίας, για $K = 2$, συναρτήσει της πιθανότητας. Στη συνέχεια, για συντομία όλοι οι λογάριθμοι θα υπολογίζονται με βάση το $K = 2$ και η γραφή $\log(\cdot)$ θα δηλώνει $\log_2(\cdot)$, εκτός αν δηλώνεται ρητά άλλη βάση.

2.3 Εντροπία ή Μέση Πληροφορία

Στην παρουσίαση που προηγήθηκε αναφερθήκαμε σε τυχαία γεγονότα για τα οποία ορίσαμε το μέτρο πληροφορίας. Δεδομένου ότι η εμφάνιση ενός συμβόλου της πηγής αποτελεί ένα τυχαίο γεγονός, ορίζουμε την πληροφορία συμβόλου και κατ' επέκταση στην πληροφορία λέξεων ή μηνυμάτων.

Θεωρούμε μία πηγή πληροφορίας με αλφάβητο $X = \{x_1, x_2, \ldots, x_N\}$. Καθένα σύμβολο x_i $(i = 1, 2, \ldots, N)$ έχει πιθανότητα εμφάνισης p_i και άρα πληροφορία $I(x_i) = -\log(p_i)$. Προφανώς, για τις πιθανότητες των συμβόλων ισχύει

$$p_i \in [0, 1] \text{ και } \sum_{i=1}^{N} p_i = 1.$$

Η κατανομή των πιθανοτήτων των συμβόλων του αλφαβήτου X περιγράφεται από το σύνολο $P_X = \{p_1, p_2, \ldots, p_N\}$. Έτσι, η πηγή πληροφορίας περιγράφεται και αναπαριστάται από το ζεύγος (X, P_X).

Θα ήταν ενδιαφέρον να γνωρίζουμε, εκτός από την πληροφορία του κάθε συμβόλου χωριστά, τη μέση πληροφορία ανά σύμβολο που μας παρέχει η έξοδος της πηγής. Με άλλα λόγια, ζητούμε ένα μέγεθος πληροφορίας που να σχετίζεται με την πηγή συνολικά, συνεπώς με το αλφάβητό της. Για το λόγο αυτό εισάγεται η έννοια της **εντροπίας πηγής**. Η εντροπία είναι θεμελιώδης έννοια για τη θεωρία της πληροφορίας και προτάθηκε το 1948 από τον Shannon.

Ορισμός 2.8. ***Εντροπία ή μέση πληροφορία ανά σύμβολο πληροφορίας*** *μίας πηγής (X, P_X) είναι η αναμενόμενη τιμή της πληροφορίας των συμβόλων της πηγής και δίνεται από τη σχέση*

$$H(X) = -\sum_{i=1}^{N} p_i \log(p_i). \tag{2.6}$$

Προφανώς, η εντροπία μετράται σε bits/σύμβολο (bits/symbol).

Η συνάρτηση που περιγράφει την εντροπία έχει τις παρακάτω βασικές ιδιότητες:

- Είναι συνεχής συνάρτηση των πιθανοτήτων $p_1, p_2, \ldots, p_N$.
- Είναι μη αρνητική, $H(X) \geq 0$, δεδομένου ότι $p_i \leq 1$.
- Είναι συμμετρική συνάρτηση των πιθανοτήτων, δηλαδή έχει την ίδια τιμή για οποιαδήποτε αντιμετάθεση των μεταβλητών $p_1, p_2, \ldots, p_N$.

Δύο ακόμη σημαντικές ιδιότητες της συνάρτησης εντροπίας παρατίθενται στη συνέχεια ως θεωρήματα με τις αποδείξεις τους.

Θεώρημα 2.1 (Μέγιστης εντροπίας).
Η εντροπία μίας πηγής (X, P_X) διακεκριμένων συμβόλων είναι μέγιστη όταν τα σύμβολα της πηγής είναι ισοπίθανα,

$$p_1 = p_2 = \ldots = p_N = \frac{1}{N}.$$

Απόδειξη.
Η συνάρτηση που πρέπει να μεγιστοποιηθεί είναι η

$$H(X) = -\sum_{i=1}^{N} p_i \log(p_i) \tag{2.7}$$

δεδομένου ότι οι ελεύθερες παράμετροι $p_1, p_2, \ldots, p_N$ ικανοποιούν τη συνθήκη ισότητας

$$\sum_{i=1}^{N} p_i = 1. \tag{2.8}$$

Αυτό το πρόβλημα μεγιστοποίησης υπό συνθήκη ισότητας αντιμετωπίζεται με τη χρήση των πολλαπλασιαστών Lagrange. Σύμφωνα με τη μέθοδο αυτή οι τιμές των πιθανοτήτων που μεγιστοποιούν την (2.7) προκύπτουν από τη λύση του συστήματος διαφορικών εξισώσεων:

$$\frac{\partial H(X)}{\partial p_i} + \lambda \frac{\partial \sum_{i=1}^{N} p_i}{\partial p_i} = 0, \quad i = 1, 2, \ldots, N \tag{2.9}$$

όπου η σταθερά λ είναι ο πολλαπλασιαστής Lagrange. Από την (2.9) προκύπτει εύκολα ότι

$$\log(p_1) = \log(p_2) = \ldots = \log(p_N) \Rightarrow p_1 = p_2 = \ldots = p_N = 1/N.$$

♦

Θεώρημα 2.2 (Σχετική εντροπία - Απόσταση Kullback-Leibler).
Αν $P_X = \{p_1, p_2, \ldots, p_N\}$ και $Q_X = \{q_1, q_2, \ldots, q_N\}$ είναι δύο εναλλακτικές κατανομές πιθανοτήτων των συμβόλων $X = \{x_1, x_2, \ldots, x_N\}$ μίας πηγής τότε:

$$H(X) = -\sum_{i=1}^{N} p_i \log(p_i) \leq -\sum_{i=1}^{N} p_i \log(q_i) \tag{2.10}$$

ή ισοδύναμα

$$H(X, P_X/Q_X) = \sum_{i=1}^{N} p_i \log\left(\frac{p_i}{q_i}\right) \geq 0. \tag{2.11}$$

Η συνάρτηση $H(X, P_X/Q_X)$ ονομάζεται ***σχετική εντροπία*** *ή* ***απόσταση Kullback-Leibler*** *των κατανομών P_X, Q_X για το αλφάβητο X.*

Απόδειξη.
Ως γνωστό, για το νεπέριο λογάριθμο ισχύει η ταυτοανισότητα $\ln(x) \leq x - 1$ και σύμφωνα με τον τύπο αλλαγής βάσης λογαρίθμου έχουμε

$$\log(x) = \ln(x)\log(e).$$

Θυμίζουμε ότι κατά σύμβαση έχουμε δεχθεί ότι $\log(\cdot)$ είναι ο λογάριθμος με βάση το δύο. Λαμβάνοντας υπόψη τα παραπάνω, η (2.11) αποδεικνύεται ως εξής:

$$\begin{aligned} H(X, P_X/Q_X) = \sum_{i=1}^{N} p_i \log\left(\frac{p_i}{q_i}\right) &= -\log(e)\sum_{i=1}^{N} p_i \ln\left(\frac{q_i}{p_i}\right) \geq \\ &\geq \log(e)\sum_{i=1}^{N} p_i\left(1 - \frac{q_i}{p_i}\right) = \log(e)\sum_{i=1}^{N}(p_i - q_i) = \\ &= \log(e)\left(\sum_{i=1}^{N} p_i - \sum_{i=1}^{N} q_i\right) = \log(e)(1-1) = 0. \end{aligned}$$

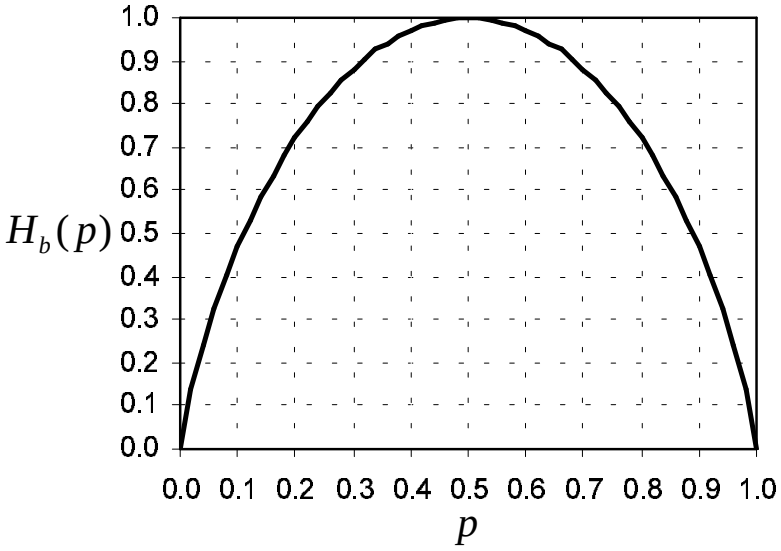
♦

Η απόσταση Kullback-Leibler ισούται με μηδέν αν και μόνο αν οι δύο εναλλακτικές κατανομές πιθανοτήτων P_X και Q_X είναι ίσες. Διαφορετικά, η απόσταση Kullback-Leibler εκφράζει την απόκλιση της κατανομής P_X από την Q_X. Θα πρέπει να σημειωθεί ότι η συνάρτηση $H(X, P_X/Q_X)$ δεν είναι συμμετρική ως προς τις δύο κατανομές, δηλαδή ισχύει

$$H(X, P_X/Q_X) \neq H(X, Q_X/P_X). \tag{2.12}$$

2.4 Εντροπία Δυαδικής Πηγής - Συνάρτηση Shannon

Συχνά, στη θεωρία της πληροφορίας εμφανίζεται η εντροπία δυαδικής πηγής. Η δυαδική πηγή πληροφορίας έχει αλφάβητο που αποτελείται από δύο μόνο σύμβολα



Σχήμα 2.2: *Γραφική παράσταση της συνάρτησης Shannon.*

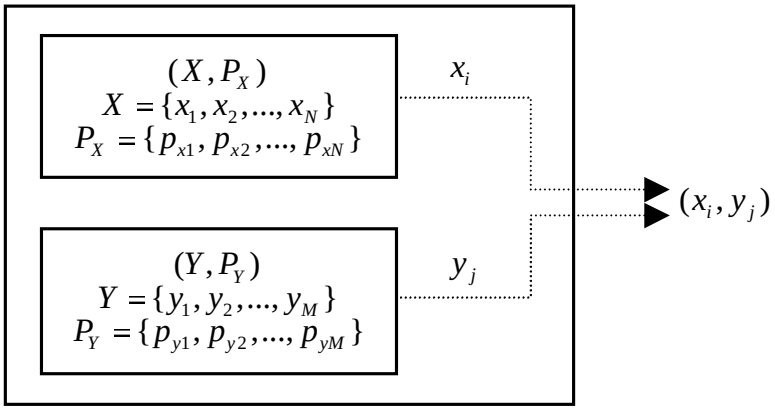
π.χ. το 0 και το 1. Προφανώς, αν p είναι η πιθανότητα εμφάνισης του συμβόλου 0, τότε η πιθανότητα εμφάνισης του 1 θα είναι ίση με $1-p$. Συνεπώς, η εντροπία της δυαδικής πηγής, η οποία συμβολίζεται με $H_b(p)$, θα είναι ίση με

$$H_b(p) = -p\log(p) - (1-p)\log(1-p). \tag{2.13}$$

Η **εντροπία δυαδικής πηγής** $H_b(p)$ είναι συνάρτηση μόνο της πιθανότητας p του ενός συμβόλου και λέγεται **συνάρτηση Shannon.** Όπως φαίνεται και από το σχήμα 2.2 η συνάρτηση Shannon είναι συμμετρική και παίρνει τη μέγιστη τιμής της (1 bit/symbol) όταν τα δύο σύμβολα είναι ισοπίθανα ($p = 0.5$). Αυτό άλλωστε προκύπτει και από το θεώρημα μέγιστης εντροπίας. Η τιμή της συνάρτησης Shannon για διάφορες τιμές της πιθανότητας p δίνεται στον πίνακα Β΄.1 του Παραρτήματος Β΄.

2.5 Συνδετική ή Από Κοινού Εντροπία

Θεωρούμε δύο πηγές πληροφορίας, την (X, P_X) με αλφάβητο $X = \{x_1, x_2, \ldots, x_N\}$ και κατανομή πιθανοτήτων $P_X = \{p_{x1}, p_{x2}, \ldots, p_{xN}\}$ και την (Y, P_Y) με αλφάβητο $Y = \{y_1, y_2, \ldots, y_M\}$ και κατανομή πιθανοτήτων $P_Y = \{p_{y1}, p_{y2}, \ldots, p_{yM}\}$, αντίστοιχα. Μπορούμε να υποθέσουμε ότι οι δύο πηγές συνθέτουν μία σύνθετη πηγή (XY, P_{XY}) (σχήμα 2.3), της οποίας το αλφάβητο προκύπτει από το καρτεσιανό γινόμενο $XY = X \times Y = \{(x_i, y_j) : x_i \in X, y_j \in Y\}$. Κατ' επέκταση της εντροπίας πηγής ορίζεται η **συνδετική εντροπία** ή **από κοινού εντροπία** των δύο πηγών ως η εντροπία της σύνθετης πηγής (XY, P_{XY}). Έτσι, η συνδετική ή από κοινού



Σχήμα 2.3: *Σύνθετη πηγή πληροφορίας.*

εντροπία της σύνθετης πηγής (XY, P_{XY}) δίνεται από τη σχέση

$$H(XY) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(x_i, y_j)). \tag{2.14}$$

Μία βασική ιδιότητα της συνδετικής εντροπίας δύο πηγών είναι ότι δεν μπορεί να υπερβεί το άθροισμα των εντροπιών των δύο επιμέρους πηγών. Αυτό σημαίνει ότι η μέση πληροφορία της σύνθετης πηγής είναι μικρότερη ή το πολύ ίση με το άθροισμα των μέσων πληροφοριών των απλών πηγών πληροφορίας, δηλαδή

$$H(XY) \leq H(X) + H(Y) \tag{2.15}$$

Η ταυτοανισότητα (2.15) μπορεί να αποδειχθεί με τη χρήση του θεωρήματος της σχετικής εντροπίας που δίνεται από τη σχέση (2.10).

Απόδειξη.
Η πιθανότητα εμφάνισης του σύνθετου γεγονότος του ζεύγους συμβόλων (x_i, y_j) είναι $p(x_i, y_j)$. Ως εναλλακτική τιμή $(q(x_i, y_j))$ της πιθανότητας αυτής μπορούμε να θεωρήσουμε το γινόμενο των περιθωριακών πιθανοτήτων των συμβόλων, δηλαδή $q(x_i, y_j) = p(x_i)p(y_j)$. Τότε σύμφωνα με την ανισότητα (2.10) έχουμε

$$\begin{aligned} H(XY) &= -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(x_i, y_j)) \leq \\ &\leq -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(q(x_i, y_j)) = \end{aligned}$$

$$= -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(x_i)) - \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(y_j)) =$$

$$= -\sum_{i=1}^{N}\left(\sum_{j=1}^{M} p(x_i, y_j)\right)\log(p(x_i)) - \sum_{j=1}^{M}\left(\sum_{i=1}^{N} p(x_i, y_j)\right)\log(p(y_j)) =$$

$$= -\sum_{i=1}^{N} p(x_i)\log(p(x_i)) - \sum_{j=1}^{M} p(y_j)\log(p(y_j)) = H(X) + H(Y)$$

Είναι φανερό ότι η μέση πληροφορία της σύνθετης πηγής ισούται με το άθροισμα των μέσων πληροφοριών των δύο επιμέρους πηγών στην περίπτωση που οι δύο πηγές είναι ανεξάρτητες, δηλαδή όταν

$$p(x_i, y_j) = p(x_i)p(y_j), \quad \forall\, i, j.$$

Οι έννοιες της σύνθετης πηγής και της συνδετικής εντροπίας μπορούν να επεκταθούν για την περίπτωση περισσοτέρων των δύο απλών πηγών. Έτσι, αν έχουμε K απλές πηγές πληροφορίας $(X_1, P_{X_1}), (X_2, P_{X_2}), \ldots, (X_K, P_{X_K})$, τότε η συνδετική εντροπία δίνεται από τη σχέση:

$$H(X_1X_2\ldots X_K) = -\sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2}\cdots$$

$$\cdots\sum_{i_K=1}^{N_K} p(x_{i_1}, x_{i_2}, \ldots, x_{i_K})\log(p(x_{i_1}, x_{i_2}, \ldots, x_{i_K})). \tag{2.16}$$

Στη σχέση (2.16) με N_k $(k = 1, 2, \ldots, K)$ δηλώνεται το πλήθος των συμβόλων του αλφαβήτου X_k.

Τέλος, η ταυτοανισότητα (2.15) μπορεί να γενικευθεί για την περίπτωση K απλών πηγών. Συγκεκριμένα, αποδεικνύεται επαγωγικά ότι ισχύει:

$$H(X_1X_2\ldots X_K) \leq H(X_1) + H(X_2) + \ldots + H(X_K). \tag{2.17}$$

2.6 Υπό Συνθήκη Εντροπία

Όπως και στην προηγούμενη ενότητα, θεωρούμε δύο πηγές πληροφορίας την (X, P_X) με αλφάβητο $X = \{x_1, x_2, \ldots, x_N\}$ και κατανομή πιθανοτήτων $P_X = \{p_{x1}, p_{x2}, \ldots, p_{xN}\}$ και την (Y, P_Y) με αλφάβητο $Y = \{y_1, y_2, \ldots, y_M\}$ και κατανομή πιθανοτήτων $P_Y = \{p_{y1}, p_{y2}, \ldots, p_{yM}\}$, αντίστοιχα. Ας υποθέσουμε ότι γνωρίζουμε εκ' των προτέρων πως η πηγή (Y, P_Y) παράγει το σύμβολο y_j. Με αυτήν την προϋπόθεση, η

πιθανότητα εμφάνισης του συμβόλου x_i στην έξοδο της πηγής (X, P_X) δίνεται από τη γνωστή σχέση

$$p(x_i/y_j) = \frac{p(x_i, y_j)}{p(y_j)}. \tag{2.18}$$

Προφανώς, η $p(x_i/y_j)$ εκφράζει την πιθανότητα εμφάνισης του ζεύγους (x_i, y_j) στην έξοδο της σύνθετης πηγής (XY, P_{XY}) δεδομένου ότι η απλή πηγή (Y, P_Y) παράγει το σύμβολο y_j.

Η μέση πληροφορία (εντροπία) ανά σύμβολο της σύνθετης πηγής με δεδομένη την έξοδο της μίας από τις δύο απλές πηγές δίνεται από τη σχέση

$$H(X/y_j) = -\sum_{i=1}^{N} p(x_i/y_j) \log(p(x_i/y_j)). \tag{2.19}$$

Η μέση τιμή του $H(X/y_j)$ ως προς όλα τα σύμβολα y_j ονομάζεται **υπό συνθήκη εντροπία** της σύνθετης πηγής (XY, P_{XY}), όταν είναι γνωστή η έξοδος της απλής πηγής (Y, P_Y), και δίνεται από τη σχέση:

$$H(X/Y) = \sum_{j=1}^{M} H(X/y_j)p(y_j) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j) \log(p(x_i/y_j)). \tag{2.20}$$

Με ανάλογο τρόπο ορίζεται και η υπό συνθήκη εντροπία της σύνθετης πηγής όταν είναι γνωστή η έξοδος της απλής πηγής (X, P_X):

$$H(Y/X) = \sum_{i=1}^{N} H(Y/x_i)p(x_i) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j) \log(p(y_j/x_i)). \tag{2.21}$$

Στη συνέχεια, θα αποδείξουμε δύο ισότητες που συνδέουν την υπό συνθήκη εντροπία σύνθετης πηγής με τη συνδετική εντροπία και τις εντροπίες των απλών πηγών. Συγκεκριμένα, ισχύει ότι:

$$H(X/Y) = H(XY) - H(Y), \tag{2.22}$$

$$H(Y/X) = H(XY) - H(X). \tag{2.23}$$

Απόδειξη.

$$H(X/Y) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(x_i/y_j)) =$$
$$= -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log\left[\frac{p(x_i, y_j)}{p(y_j)}\right] =$$
$$= -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(x_i, y_j)) + \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(y_j)) =$$
$$= H(XY) + \sum_{j=1}^{M} p(y_j)\log(p(y_j)) = H(XY) - H(Y).$$

Με ανάλογο τρόπο αποδεικνύεται και η σχέση (2.23).

♦

2.7 Αμοιβαία Πληροφορία ή Διαπληροφορία

Ένα σημαντικό μέγεθος που ορίζεται στα πλαίσια της θεωρίας της πληροφορίας είναι η **αμοιβαία πληροφορία ή διαπληροφορία**. Αν $X = \{x_1, x_2, \ldots, x_N\}$ και $Y = \{y_1, y2, \ldots, y_M\}$, είναι δύο πηγές, με κατανομές πιθανοτήτων P_X και P_Y αντίστοιχα, τότε η διαπληροφορία των δύο πηγών (ή των δύο κατανομών) ορίζεται ως

$$I(X;Y) = H(X) - H(X/Y). \tag{2.24}$$

Η συνάρτηση της διαπληροφορίας έχει τις παρακάτω ιδιότητες:

- Είναι συμμετρική ως προς X και Y, δηλαδή
$$I(X;Y) = I(Y;X). \tag{2.25}$$
- Είναι μη αρνητική
$$I(X;Y) \geq 0. \tag{2.26}$$

Απόδειξη.
Για να αποδείξουμε τις ιδιότητες (2.25) και (2.26) αρκεί να αντικαταστήσουμε στον ορισμό (2.24) την υπό συνθήκη εντροπία (2.22). Έχουμε

$$I(X;Y) = H(X) - H(X/Y) = H(X) - (H(XY) - H(Y)) \Rightarrow$$
$$\Rightarrow I(X;Y) = H(X) + H(Y) - H(XY). \tag{2.27}$$

Από τη σχέση (2.27) είναι φανερό ότι η διαπληροφορία είναι συμμετρική ($I(X;Y) = I(Y;X)$). Επίσης, με εφαρμογή της ταυτοανισότητας (2.15) προκύπτει ότι $I(X;Y) \geq 0$.

♦

Αναπτύσσοντας τη σχέση ορισμού (2.24), προκύπτει ότι η διαπληροφορία έχει τη μορφή απόστασης Kullback-Leibler. Πράγματι,

$$\begin{aligned} I(X;Y) &= -\sum_{i=1}^{N} p(x_i)\log(p(x_i)) + \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i,y_j)\log(p(x_i/y_j)) = \\ &= -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i,y_j)\log(p(x_i)) + \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i,y_j)\log(p(x_i/y_j)) = \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i,y_j)\log\left(\frac{p(x_i,y_j)}{p(x_i)p(y_j)}\right) = H(XY, P_{XY}/(P_X\cdot P_Y)). \end{aligned} \tag{2.28}$$

Στη σχέση (2.28), $H(XY, P_{XY}/(P_X \cdot P_Y))$ είναι η απόσταση Kullback-Leibler της κοινής κατανομής P_{XY} από την εναλλακτική κατανομή $P_X \cdot P_Y$ της σύνθετης πηγής XY. Στην ουσία η εναλλακτική κατανομή $P_X \cdot P_Y$, η οποία προκύπτει από τα γινόμενα των περιθωριακών κατανομών των δύο πηγών X και Y, ισχύει όταν οι δύο πηγές είναι ανεξάρτητες. Συνεπώς, η διαπληροφορία είναι μέτρο της ανεξαρτησίας των δύο πηγών. Συγκεκριμένα, αν οι δύο πηγές είναι ανεξάρτητες (οπότε $p(x_i,y_j) = p(x_i)p(y_j)\,\forall i,j$) τότε η διαπληροφορία μηδενίζεται ($I(X;Y) = 0$). Αντιθέτως, αν για κάθε i,j ισχύει $p(x_i/y_j) = 0$ ή $p(x_i/y_j) = 1$, τότε $H(X/Y) = 0$ και η διαπληροφορία παίρνει τη μέγιστη τιμή της $I(X;Y) = H(X)$.

2.7.1 Φυσική σημασία της αμοιβαίας πληροφορίας

Η έννοια της αμοιβαίας πληροφορίας είναι μία από τις σημαντικότερες της θεωρίας της πληροφορίας και, όπως θα δούμε σε επόμενο κεφάλαιο, συνδέεται άμεσα με τον ορισμό της χωρητικότητας ενός διαύλου επικοινωνίας. Για να γίνει κατανοητή η έννοια αυτή θα εξετάσουμε ένα παράδειγμα δύο πηγών πληροφορίας, οι οποίες μπορεί να σχετίζονται μεταξύ τους.

Έστω ένας πομπός, ο οποίος μπορεί να παρασταθεί ως μία πηγή πληροφορίας (X, P_X). Τα σύμβολα της πηγής αυτής (του πομπού) αποστέλονται μέσω ενός καναλιού πληροφορίας σε ένα δέκτη. Ο δέκτης ουσιαστικά είναι μία νέα πηγή πληροφορίας (Y, P_Y), η οποία συμπεριφέρεται εν γένει διαφορετικά από την (X, P_X). Θεωρούμε επίσης

έναν παρατηρητή, ο οποίος βρίσκεται στην περιοχή του δέκτη, παρατηρεί την πηγή (Y, P_Y) αλλά δε μπορεί να παρατηρήσει την πηγή (X, P_X). Κάποια στιγμή ο παρατηρητής βλέπει στο δέκτη το σύμβολο y_j, το οποίο παράγεται από την πηγή (Y, P_Y). Η εμφάνιση αυτού του συμβόλου στο δέκτη οφείλεται στην εκπομπή κάποιου συμβόλου από τον πομπό. Το ερώτημα που θα μας απασχολήσει είναι το εξής: *Βλέποντας το σύμβολο y_j στο δέκτη, πόση πληροφορία αποκόμισε ο παρατηρητής για το ποιο σύμβολο, x_i, εκπέμθηκε αρχικά από τον πομπό;*

Αρχικά η αβεβαιότητα του παρατηρητή για την εκπομπή του x_i είναι ίση με την πληροφορία $-\log(p(x_i))$. Στη συνέχεια με την εμφάνιση του y_j στο δέκτη, η αβεβαιότητα για την εκπομπή του x_i, δεδομένου ότι έλαβε y_j στο δέκτη, θα είναι $-\log(p(x_i/y_j))$. Η διαφορά των δύο αυτών τιμών της αβεβαιότητας, πριν και μετά τη λήψη του y_j, ισοδυναμεί με την πληροφορία που έλαβε ο παρατηρητής με το πέρας της διαδικασίας εκπομπής - λήψης. Δηλαδή η πληροφορία που κέρδισε ο παρατηρητής θα είναι

$$I(x_i; y_j) = \log(p(x_i/y_j)) - \log(p(x_i)). \tag{2.29}$$

Η μέση πληροφορία που κερδίζει ο παρατηρητής για το τι εκπέμπεται από τον πομπό παρατηρώντας το δέκτη, θα είναι προφανώς ο μέσος όρος της (2.29) λαμβάνοντας υπόψη κάθε πιθανό ζεύγος εκπομπής - λήψης (x_i, y_j). Συνεπώς, η πληροφορία που λαμβάνει ο παρατηρητής για το τι εκπέμπει η πηγή (X, P_X), παρατηρώντας την πηγή (Y, P_Y), είναι ίση με

$$\begin{aligned} I(X;Y) &= \sum_{i=1}^{N}\sum_{j=1}^{M} I(x_i; y_j)p(x_i, y_j) = \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M} \left[\log(p(x_i/y_j)) - \log(p(x_i))\right] p(x_i, y_j) = \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(x_i/y_j)) - \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(x_i)) = \\ &= -H(X/Y) - \sum_{i=1}^{N} p(x_i)\log(p(x_i)) = -H(X/Y) + H(X). \end{aligned} \tag{2.30}$$

Παρατηρούμε ότι ουσιαστικά καταλήξαμε στη σχέση ορισμού της αμοιβαίας πληροφορίας ή διαπληροφορίας (2.24). Συνεπώς, η αμοιβαία πληροφορία μεταξύ της πηγής πληροφορίας του πομπού και της πηγής πληροφορίας στο δέκτη εκφράζει τη μέση πληροφορία που αποκομίζει ο παρατηρητής για το τι εκπέμπει ο πομπός παρατηρώντας το δέκτη. Προφανώς, όσο μεγαλύτερη είναι η τιμή της διαπληροφορίας, τόσο καλύτερη είναι η μετάδοση της πληροφορίας από τον πομπό στο δέκτη. Το γεγονός

αυτό, όπως θα δούμε σε επόμενο κεφάλαιο, το εκμεταλλευόμαστε για να ορίσουμε τη χωρητικότητα ενός διαύλου πληροφορίας.

2.8 Παραδείγματα

1. *Θεωρούμε μία δυαδική πηγή με δύο σύμβολα το 0 και το 1. Η πιθανότητα εκπομπής του 0 είναι ίση με 0.7. Να βρεθεί η πληροφορία του κάθε συμβόλου εκπομπής καθώς και η εντροπία της πηγής.*

   Η πηγή έχει αλφάβητο $A = \{0, 1\}$ και κατανομή πιθανοτήτων $P_A = \{0.7, 0.3\}$. Δηλαδή, οι πιθανότητες εμφάνισης των δύο συμβόλων είναι $p(0) = 0.7$ και $p(1) = 0.3$ (αφού $p(0) + p(1) = 1$). Σύμφωνα με τον ορισμό της εξίσωσης (2.5) η πληροφορία του κάθε συμβόλου χωριστά θα είναι

$$I(0) = -\log(p(0)) = -\log(0.7) \simeq 0.515 \; bits,$$

$$I(1) = -\log(p(1)) = -\log(0.3) \simeq 1.737 \; bits.$$

   Η εντροπία της πηγής δίνεται από τη σχέση του ορισμού (2.6), οπότε έχουμε

$$H(A) = -p(0)\log(p(0)) - p(1)\log(p(1)) \simeq$$
$$\simeq 0.7 \cdot 0.515 + 0.3 \cdot 1.737 = 0.8816 \; bits/symbol.$$

2. *Πόση πληροφορία περιέχεται στον αριθμό κυκλοφορίας αυτοκινήτου της μορφής ΓΓΓαααα, όπου Γ είναι κεφαλαίο γράμμα και α αριθμός; (Θεωρούμε ελληνικές πινακίδες κυκλοφορίας).*

   Ως γνωστό, στις ελληνικές πινακίδες κυκλοφορίας αυτοκινήτων χρησιμοποιούνται τα 14 διεθνή γράμματα {A,B,E,Z,H,I,K,M,N,O,P,T,Υ,X}. Ο τετραψήφιος αριθμός μπορεί να πάρει 9000 δυνατές τιμές (από το 1000 μέχρι και το 9999). Υπάρχουν λοιπόν $14^3 \times 9000 = 24696000$ πιθανές πινακίδες. Άρα, η πιθανότητα εμφάνισης μίας συγκεκριμένης πινακίδας είναι ίση με $1/24696000$ και η πληροφορία που περιέχει θα είναι ίση με $-\log(1/24696000) \simeq 24.56\, bits$.

3. *Θεωρούμε μία δυαδική πηγή X με δύο εναλλακτικές κατανομές πιθανοτήτων, $P_X = \{0.3, 0.7\}$ και $Q_X = \{0.6, 0.4\}$. Να υπολογιστούν οι αποστάσεις Kullback-Leibler, $H(X, P_X/Q_X)$ και $H(X, Q_X/P_X)$.*

   Σύμφωνα με τη σχέση (2.11) έχουμε

$$H(X, P_X/Q_X) = \sum_{i=1,2} p_i \log\left(\frac{p_i}{q_i}\right) =$$

$$= 0.3\log(0.3/0.6) + 0.7\log(0.7/0.4) \simeq 0.2651$$

$$H(X, Q_X/P_X) = \sum_{i=1,2} q_i \log\left(\frac{q_i}{p_i}\right) =$$
$$= 0.6\log(0.6/0.3) + 0.4\log(0.4/0.7) \simeq 0.277$$

Παρατηρούμε ότι $H(X, P_X/Q_X) \neq H(X, Q_X/P_X)$.

4. *Θεωρούμε μια πηγή X με N σύμβολα και κατανομή πιθανοτήτων $P_X = \{p_1, p_2, \ldots, p_N\}$. Να αποδειχθεί ότι για την εντροπία ισχύει*

$$H(p_1, p_2, p_3, \ldots, p_N) = H(p_1 + p_2, p_3, \ldots, p_N) + (p_1 + p_2)H(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}).$$

Ξεκινώντας από το δεξιό μέλος της ισότητας έχουμε

$$H(p_1 + p_2, p_3, \ldots, p_N) + (p_1 + p_2)H(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}) =$$
$$= -(p_1 + p_2)\log(p_1 + p_2) - \sum_{i=3}^{N} p_i \log(p_i) +$$
$$+(p_1 + p_2)\left[-\frac{p_1}{p_1 + p_2}\log(\frac{p_1}{p_1 + p_2}) - \frac{p_2}{p_1 + p_2}\log(\frac{p_2}{p_1 + p_2})\right] =$$
$$= -p_1\log(p_1 + p_2) - p_1\log(\frac{p_1}{p_1 + p_2}) -$$
$$-p_2\log(p_1 + p_2) - p_2\log(\frac{p_2}{p_1 + p_2}) -$$
$$-\sum_{i=3}^{N} p_i \log(p_i) =$$
$$= -p_1\log(p_1) - p_2\log(p_2) - \sum_{i=3}^{N} p_i\log(p_i) = H(p_1, p_2, p_3, \ldots, p_N).$$

Αντιλαμβάνεστε τη σημασία της σχέσης που μόλις αποδείξαμε;

5. *Θεωρούμε δύο δυαδικές πηγές πληροφορίας $X = \{£, ¥\}$ και $Y = \{\Diamond, \Box\}$, αντίστοιχα. Δίνονται οι τιμές των πιθανοτήτων*

$$p(£) = 0.2, \quad p(\Diamond) = 0.3, \quad p(£, \Box) = 0.15.$$

Να βρεθούν οι τιμές της εντροπίας της κάθε πηγής χωριστά καθώς και της σύνθετης πηγής XY. Τέλος, να υπολογιστούν οι υπό συνθήκη εντροπίες $H(X/Y)$ και $H(Y/X)$.

Αρχικά υπολογίζουμε τις τιμές όλων των περιθωριακών πιθανοτήτων των συμβόλων των δύο πηγών. Αυτές προκύπτουν βάσει της ιδιότητας ότι το άθροισμα

των πιθανοτήτων όλων των συμβόλων μιας πηγής είναι ίσο με τη μονάδα. Άρα, έχουμε

$$p(£) + p(¥) = 1 \Rightarrow p(¥) = 1 - 0.2 \Rightarrow p(¥) = 0.8,$$

$$p(\Diamond) + p(\Box) = 1 \Rightarrow p(\Box) = 1 - 0.3 \Rightarrow p(\Box) = 0.7.$$

Στη συνέχεια, υπολογίζουμε τις από κοινού πιθανότητες χρησιμοποιώντας τις περιθωριακές. Έχουμε,

$$p(£, \Box) + p(£, \Diamond) = p(£) \Rightarrow p(£, \Diamond) = 0.2 - 0.15 \Rightarrow p(£, \Diamond) = 0.05,$$

$$p(£, \Diamond) + p(¥, \Diamond) = p(\Diamond) \Rightarrow p(¥, \Diamond) = 0.3 - 0.05 \Rightarrow p(¥, \Diamond) = 0.25,$$

$$p(£, \Box) + p(¥, \Box) = p(\Box) \Rightarrow p(¥, \Box) = 0.7 - 0.15 \Rightarrow p(¥, \Box) = 0.55.$$

Οι παραπάνω τιμές των κοινών πιθανοτήτων συνθέτουν τον πίνακα

$$P_{XY} = \begin{bmatrix} p(£, \Diamond) & p(£, \Box) \\ p(¥, \Diamond) & p(¥, \Box) \end{bmatrix} = \begin{bmatrix} 0.05 & 0.15 \\ 0.25 & 0.55 \end{bmatrix}.$$

Στον παραπάνω πίνακα παρατηρούμε ότι τα αθροίσματα κατά στήλη ή γραμμή δίνουν τις περιθωριακές πιθανότητες των συμβόλων. Στη συνέχεια, μπορούμε να υπολογίσουμε τις υπό συνθήκη πιθανότητες βάσει της γνωστής σχέσης $p(a/b) = p(a, b)/p(b)$. Συνεπώς, οι υπό συνθήκη πιθανότητες είναι

$$p(£/\Diamond) = p(£, \Diamond)/p(\Diamond) = 0.05/0.3 \simeq 0.1667,$$

$$p(£/\Box) = p(£, \Box)/p(\Box) = 0.15/0.7 \simeq 0.2143,$$

$$p(¥/\Diamond) = p(¥, \Diamond)/p(\Diamond) = 0.25/0.3 \simeq 0.8333,$$

$$p(¥/\Box) = p(¥, \Box)/p(\Box) = 0.55/0.7 \simeq 0.7857,$$

$$p(\Diamond/£) = p(£, \Diamond)/p(£) = 0.05/0.2 = 0.25,$$

$$p(\Diamond/¥) = p(¥, \Diamond)/p(¥) = 0.25/0.8 = 0.3125,$$

$$p(\Box/£) = p(£, \Box)/p(£) = 0.15/0.2 = 0.75,$$

$$p(\Box/¥) = p(¥, \Box)/p(¥) = 0.55/0.8 = 0.6875.$$

Οι υπό συνθήκη πιθανότητες συνθέτουν του πίνακες:

$$P_{X/Y} = \begin{bmatrix} p(£/\Diamond) & p(¥/\Diamond) \\ p(£/\Box) & p(¥/\Box) \end{bmatrix} \simeq \begin{bmatrix} 0.1667 & 0.8333 \\ 0.2143 & 0.7857 \end{bmatrix},$$

$$P_{Y/X} = \begin{bmatrix} p(\Diamond/£) & p(\Box/£) \\ p(\Diamond/¥) & p(\Box/¥) \end{bmatrix} = \begin{bmatrix} 0.25 & 0.75 \\ 0.3125 & 0.6875 \end{bmatrix}.$$

Παρατηρούμε ότι στους δύο παραπάνω πίνακες τα αθροίσματα των στοιχείων τους κατά γραμμή είναι πάντα ίσα με τη μονάδα.

Η εντροπία κάθε μίας απλής πηγής υπολογίζεται σύμφωνα με τη σχέση (2.6) ως εξής:

$$H(X) = -\sum_{x\in\{£,¥\}} p(x)\log(p(x)) = -0.2\log(0.2) - 0.8\log(0.8) \simeq$$
$$\simeq 0.722 \quad bits/symbol,$$

$$H(Y) = -\sum_{y\in\{\diamond,\square\}} p(x)\log(p(x)) = -0.3\log(0.3) - 0.7\log(0.7) \simeq$$
$$\simeq 0.881 \quad bits/symbol.$$

Η εντροπία της σύνθετης πηγής XY υπολογίζεται από τη σχέση της συνδετικής εντροπίας (2.14), δηλαδή

$$H(XY) = -\sum_{x\in\{£,¥\}}\sum_{y\in\{\diamond,\square\}} p(x,y)\log(p(x,y)) =$$
$$= -0.05\log(0.05) - 0.15\log(0.15) - 0.25\log(0.25) - 0.55\log(0.55) \simeq$$
$$\simeq 1.601 \quad bits/symbol.$$

Παρατηρούμε ότι, πράγματι, ισχύει η ταυτοανισότητα $H(XY) \leq H(X)+H(Y)$ $(1.601 < 0.722 + 0.881)$.

Τέλος, υπολογίζουμε τις υπό συνθήκη εντροπίες εφαρμόζοντας τις σχέσεις (2.20) και (2.21).

$$H(X/Y) = -\sum_{x\in\{£,¥\}}\sum_{y\in\{\diamond,\square\}} p(x,y)\log(p(x/y)) =$$
$$= -0.05\log(0.1667) - 0.15\log(0.2143) - 0.25\log(0.8333) -$$
$$-0.55\log(0.7857) \simeq 0.720 \quad bits/symbol,$$

$$H(Y/X) = -\sum_{x\in\{£,¥\}}\sum_{y\in\{\diamond,\square\}} p(x,y)\log(p(y/x)) =$$
$$= -0.05\log(0.25) - 0.15\log(0.75) - 0.25\log(0.3125) -$$
$$-0.55\log(0.6875) \simeq 0.879 \quad bits/symbol.$$

Φυσικά, θα καταλήγαμε στο ίδιο αποτέλεσμα συντομότερα, χωρίς να απαιτηθεί ο υπολογισμός των υπό συνθήκη πιθανοτήτων, αν χρησιμοποιούσαμε απευθείας τις σχέσεις (2.22) και (2.23).

6. *Να υπολογιστεί η διαπληροφορία μεταξύ των πηγών X και Y του προηγούμενου παραδείγματος.*

   Από τον ορισμό της διαπληροφορίας (2.24) έχουμε

   $$I(X;Y) = H(X) - H(X/Y) \simeq 0.722 - 0.720 = 0.002 \quad bits/symbol,$$

   ή ισοδύναμα

   $$I(X;Y) = H(Y) - H(Y/X) \simeq 0.881 - 0.879 = 0.002 \quad bits/symbol.$$

2.9 Ασκήσεις

1. Να βρεθεί η εντροπία ενός ζαριού.

2. Γνωρίζουμε από τη μελέτη της ιστορίας ότι ο τελευταίος αρχηγός της φυλής των Λιλιπούα είχε δύο παιδιά. Νεότερες έρευνες κατέληξαν σε δύο συμπεράσματα:
   Α. Ο αρχηγός είχε μία κόρη και ένα γιο.
   Β. Ο αρχηγός είχε μία κόρη και ένα μεγαλύτερο γιο.
   Πόση πληροφορία θα λάβουμε με την ανακοίνωση μόνο του πρώτου συμπεράσματος και πόση με την ανακοίνωση μόνο του δεύτερου. Πόση θα είναι η πληροφορία αν ανακοινωθούν και τα δύο συμπεράσματα μαζί.

3. Οκτώ κύριοι αποφοίτησαν από τη Σχολή Καλών Τεχνών τη δεκαετία του 1960. Ξέρουμε ότι την εποχή εκείνη γίνονταν μία μόνο τελετή αποφοίτησης κάθε χρονιά. Πόση πληροφορία έχει το γεγονός ότι δύο τουλάχιστον από τους οκτώ κυρίους αποφοίτησαν την ίδια χρονιά;

4. Να υπολογιστούν οι αποστάσεις Kullback-Leibler μεταξύ των εναλλακτικών κατανομών πιθανοτήτων $P = \{0.2, 0.3, 0.5\}$ και $Q = \{0.4, 0.5, 0.1\}$. Ισχύει η αντιμεταθετική ιδιότητα ($H(X, P/Q) = H(X, Q/P)$);

5. Θεωρούμε δύο πηγές πληροφορίας την $X = \{x_1, x_2\}$ και την $Y = \{y_1, y_2\}$. Γνωρίζουμε τις τιμές των πιθανοτήτων $p(x_1) = 0.25$, $p(y_1/x_1) = 0.2$ και $p(y_1/x_2) = 0.4$. Να υπολογιστούν: Η εντροπία $H(X)$, η εντροπία $H(Y)$, η συνδετική εντροπία $H(XY)$ και οι υπό συνθήκη εντροπίες $H(X/Y)$, $H(Y/X)$.

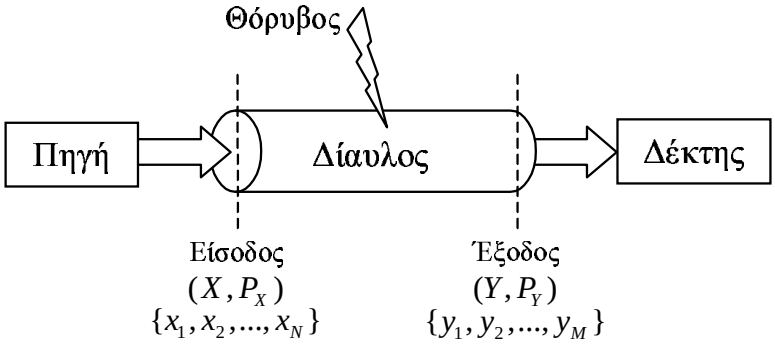
6. Να υπολογιστεί η διαπληροφορία των δύο πηγών X και Y της προηγούμενης άσκησης.

7. Θεωρούμε τρεις διακριτές πηγές πληροφορίας X, Y, Z με κοινό αλφάβητο $\{a_1, a_2, \ldots, a_N\}$ και αντίστοιχες κατανομές πιθανοτήτων $P_k = \{p_{k1}, p_{k2}, \ldots, p_{kN}\}$, $k = X, Y, Z$. Αν ισχύει $\lambda p_{Xi} + (1-\lambda) p_{Yi} = p_{Zi}$, $i = 1, 2, \ldots, N$, να δειχθεί ότι για τις εντροπίες των πηγών ισχύει: $\lambda H(X) + (1-\lambda) H(Y) \leq H(Z)$.

Κεφάλαιο 3

Δίαυλος Πληροφορίας

Όπως είναι γνωστό σε ένα σύστημα επικοινωνίας διακρίνουμε μία πηγή πληροφορίας, η οποία στέλνει δεδομένα σε ένα δέκτη. Η μεταφορά των δεδομένων από την πηγή προς το δέκτη γίνεται δια μέσου κάποιου φυσικού μέσου που ονομάζεται **δίαυλος πληροφορίας** ή **κανάλι πληροφορίας**. Όπως φαίνεται στο σχήμα 3.1, η πηγή πληροφορίας (X, P_X) προσαρμόζεται στην είσοδο του διαύλου πληροφορίας. Στη συνέχεια, η αρχική πηγή (X, P_X) εμφανίζεται στην έξοδο του διαύλου ως μία νέα πηγή πληροφορίας (Y, P_Y). Η πηγή (Y, P_Y) είναι στην ουσία αυτή που τελικά γίνεται αντιληπτή από το δέκτη. Γενικά, η πηγή (Y, P_Y) είναι διαφορετική από την (X, P_X), είτε λόγω εγγενών χαρακτηριστικών της δομής του διαύλου, είτε λόγω της επίδρασης του περιβάλλοντος του διαύλου με τη μορφή θορύβου.



Σχήμα 3.1: *Ο δίαυλος πληροφορίας δέχεται στην είσοδό του τα δεδομένα μιας πηγής πληροφορίας (X, P_X) και φαίνεται στην έξοδό του ως μία νέα πηγή πληροφορίας (Y, P_Y).*

3.1 Πίνακας Διαύλου

Για να περιγράψουμε τη λειτουργία του διαύλου πληροφορίας, δηλαδή τον τρόπο μετάδοσης των συμβόλων από την είσοδο στην έξοδό του, χρησιμοποιούμε τις υπό συνθήκη πιθανότητες των συμβόλων στην έξοδο του διαύλου, όταν είναι γνωστά τα σύμβολα στην είσοδό του. Αν $X = \{x_1, x_2, \ldots, x_N\}$ και $Y = \{y_1, y_2, \ldots, y_M\}$ είναι τα αλφάβητα στην είσοδο και την έξοδο του διαύλου, αντίστοιχα, τότε ο δίαυλος περιγράφεται μαθηματικά από τον **πίνακα διαύλου**

$$P_{Y/X} = \begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) & \cdots & p(y_M/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & \cdots & p(y_M/x_2) \\ \cdots & \cdots & \cdots & \cdots \\ p(y_1/x_N) & p(y_2/x_N) & \cdots & p(y_M/x_N) \end{bmatrix}. \tag{3.1}$$

Δεδομένου ότι οι πηγές πληροφορίας X και Y έχουν στο αλφάβητό τους N και M σύμβολα αντίστοιχα, ο πίνακας του διαύλου $P_{Y/X}$ θα έχει N γραμμές και M στήλες. Γενικά ισχύει ότι το πλήθος των συμβόλων του αλφαβήτου της πηγής εισόδου είναι διαφορετικό από αυτό της πηγής εξόδου ($N \neq M$). Αυτό πρακτικά σημαίνει ότι μπορεί περισσότερα από ένα σύμβολα της εισόδου να αντιστοιχίζονται σε ένα μόνο σύμβολο εξόδου ή και αντίστροφα. Αυτό φαίνεται παραστατικά με τη χρήση του **διαγράμματος διαύλου** (σχήμα 3.2) που περιγράφει τη σύνδεση μεταξύ των συμβόλων εισόδου και εξόδου. Για παράδειγμα, όταν στην είσοδο του διαύλου διοχετευθεί το σύμβολο x_1 τότε στην έξοδο μπορεί να εμφανιστεί είτε το y_1 είτε το y_3. Ακόμη, όταν στην έξοδο εμφανιστεί το σύμβολο y_3 τότε αυτό μπορεί να προήλθε είτε από το σύμβολο x_1 είτε από το x_3 στην είσοδο. Το διάγραμμα ενός διαύλου αποτελεί έναν εναλλακτικό τρόπο περιγραφής της λειτουργίας ενός διαύλου και είναι πιο παραστατικός συγκριτικά με τον πίνακα διαύλου $P_{Y/X}$. Όμως, όπως θα δούμε στη συνέχεια, ο πίνακας διαύλου είναι ιδιαίτερα χρήσιμος για τη μαθηματική περιγραφή και μοντελοποίηση της λειτουργίας του διαύλου.

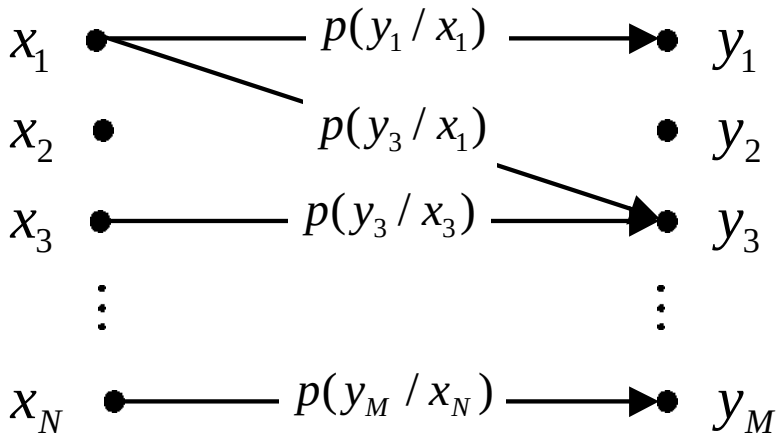
Αν γράψουμε τις κατανομές πιθανοτήτων των συμβόλων στην είσοδο και την έξοδο του διαύλου με τη μορφή πινάκων-γραμμών, δηλαδή

$$P_X = [p(x_1)\, p(x_2) \ \ldots \ p(x_N)] \tag{3.2}$$

$$P_Y = [p(y_1)\, p(y_2) \ \ldots \ p(y_M)] \tag{3.3}$$

τότε εύκολα προκύπτει ότι η κατανομή πιθανοτήτων στην έξοδο θα δίνεται από το γινόμενο τις κατανομής πιθανοτήτων στην είσοδο επί τον πίνακα διαύλου

$$P_Y = P_X \cdot P_{Y/X}. \tag{3.4}$$



Σχήμα 3.2: *Διάγραμμα του διαύλου.*

Η εξίσωση (3.4) φανερώνει πως ο πίνακας διαύλου ελέγχει τη ροή πληροφορίας από την είσοδο στην έξοδο του διαύλου, αφού μετασχηματίζει την κατανομή πιθανοτήτων των συμβόλων στην είσοδο στην κατανομή πιθανοτήτων των συμβόλων στην έξοδο.

3.2 Εντροπία Συστήματος Διαύλου

Αν γράψουμε τις πιθανότητες των συμβόλων εισόδου με τη μορφή διαγώνιου πίνακα, δηλαδή

$$D_X = \begin{bmatrix} p(x_1) & 0 & \cdots & 0 \\ 0 & p(x_2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & p(x_N) \end{bmatrix}, \tag{3.5}$$

και γνωρίζουμε τον πίνακα διαύλου $P_{Y/X}$, τότε οι από κοινού πιθανότητες (συνδετικής πιθανότητες) των συμβόλων εισόδου-εξόδου, που γράφονται με τη μορφή του πίνακα P_{XY}, δίνονται από τη σχέση:

$$P_{XY} = \begin{bmatrix} p(x_1,y_1) & p(x_1,y_2) & \cdots & p(x_1,y_M) \\ p(x_2,y_1) & p(x_2,y_2) & \cdots & p(x_2,y_M) \\ \cdots & \cdots & \cdots & \cdots \\ p(x_N,y_1) & p(x_N,y_2) & \cdots & p(x_N,y_M) \end{bmatrix} = D_X \cdot P_{Y/X}. \tag{3.6}$$

Προφανώς, γνωρίζοντας τον πίνακα P_{XY} μπορούμε να υπολογίσουμε τη συνδετική εντροπία εισόδου-εξόδου $H(XY)$, δηλαδή τη μέση πληροφορία ανά ζεύγος συμβόλων εισόδου-εξόδου, εφαρμόζοντας τη σχέση (2.14). Η συνδετική εντροπία $H(XY)$ ονομάζεται **εντροπία συστήματος διαύλου**.

3.3 Ροή Πληροφορίας στο Δίαυλο

Γενικά, όταν γνωρίζουμε πιο σύμβολο, x_i, διοχετεύεται στην είσοδο του διαύλου δεν είμαστε βέβαιοι για το πιο σύμβολο, y_j, θα εμφανιστεί στην έξοδο του διαύλου. Αυτή η αβεβαιότητα είναι άρρηκτα συνδεδεμένη με τις υπό συνθήκη πιθανότητες και τα αντίστοιχα μέτρα πληροφορίας των γεγονότων εμφάνισης ενός συμβόλου στην έξοδο του διαύλου όταν γνωρίζουμε πιο σύμβολο διοχετεύθηκε αρχικά στην είσοδο. Συνεπώς, η υπό συνθήκη πιθανότητα εμφάνισης ενός συμβόλου στην έξοδο του διαύλου δεδομένου του συμβόλου στην είσοδο ($p(y_j/x_i)$), και κατ' επέκταση ο πίνακας του διαύλου $P_{Y/X}$, περιγράφει την αβεβαιότητα στην είσοδο του διαύλου για την έξοδό του. Αντίστροφα, η αβεβαιότητα στην έξοδο του διαύλου για την είσοδό μπορεί να περιγραφεί με βάση τις υπό συνθήκη πιθανότητες των συμβόλων εισόδου δεδομένων των συμβόλων εξόδου ($p(x_i/y_j)$), δηλαδή συνολικά από τον πίνακα $P_{X/Y}$.

3.3.1 Εντροπία θορύβου

Ας θεωρήσουμε έναν παρατηρητή στην έξοδο του διαύλου, ο οποίος λαμβάνει σύμβολα της πηγής πληροφορίας (Y, P_Y). Γενικά, ο παρατηρητής αν και παρατηρεί τα σύμβολα στην έξοδο του διαύλου δεν είναι σε θέση να γνωρίζει ποια σύμβολα είχαν αρχικά εισαχθεί στην είσοδό του. Η μέση πληροφορία που λείπει από τον παρατηρητή στην έξοδο, για να αναγνωρίσει με βεβαιότητα τα σύμβολα εισόδου, θα είναι προφανώς ίση με την υπό συνθήκη εντροπία $H(X/Y)$. Αν για παράδειγμα η $H(X/Y)$ είναι ίση με μηδέν που σημαίνει ότι οι υπό συνθήκη πιθανότητες. $p(x_i/y_j)$ είναι ίσες με ένα ή μηδέν, τότε ο παρατηρητής λαμβάνοντας ένα σύμβολο y_j θα είναι απολύτως βέβαιος για το ποιο σύμβολο x_i στάλθηκε από την είσοδο του διαύλου.

Η μέση έλλειψη πληροφορίας για το σύμβολο εισόδου γίνεται αντιληπτή από τον παρατηρητή στην έξοδο ως απώλεια πληροφορίας, ίσης με $H(X/Y)$ ανά σύμβολο, και ερμηνεύεται ως αποτέλεσμα της παρουσίας θορύβου στο περιβάλλον του διαύλου πληροφορίας. Για το λόγο αυτό η υπό συνθήκη εντροπία $H(X/Y)$ ονομάζεται **εντροπία θορύβου**.

3.3.2 Εντροπία διαύλου

Στη συνέχεια, θεωρούμε έναν παρατηρητή στην είσοδο του διαύλου, ο οποίος βλέπει την πηγή πληροφορίας (X, P_X). Γενικά, αν ο παρατηρητής αυτός βλέπει ότι στην είσοδο του διαύλου εισάγεται το σύμβολο x_i, τότε δεν είναι απολύτως βέβαιος για

το ποιο θα είναι το σύμβολο y_j στην έξοδο του διαύλου. Για τον παρατηρητή στην είσοδο, η μέση έλλειψη πληροφορίας ανά σύμβολο εξόδου δεδομένου του συμβόλου εισόδου είναι ίση με την υπό συνθήκη εντροπία $H(Y/X)$. Ο παρατηρητής αποδίδει αυτήν την έλλειψη πληροφορίας στην άγνοιά του για τη δομή και τη λειτουργία του διαύλου πληροφορίας. Για το λόγο αυτό, η υπό συνθήκη εντροπία $H(Y/X)$ ονομάζεται **εντροπία διαύλου**. Αν ο παρατηρητής στην είσοδο γνώριζε με απόλυτη βεβαιότητα ποιο σύμβολο θα εμφανιστεί στην έξοδο του διαύλου για κάθε σύμβολο που εισάγεται στην είσοδο, τότε οι υπό συνθήκη πιθανότητες $p(y_j/x_i)$ είναι ίσες με μηδέν ή ένα και συνεπώς η εντροπία διαύλου μηδενίζεται ($H(Y/X) = 0$). Σημειώνεται, ότι η εντροπία διαύλου $H(Y/X)$ είναι διαφορετική από την εντροπία συστήματος διαύλου $H(XY)$ που αναφέρθηκε και προηγούμενη ενότητα και είναι συνδετική εντροπία.

Στη συνέχεια, θα δείξουμε πως υπολογίζεται η μέση πληροφορία στην έξοδο ενός διαύλου όταν γνωρίζουμε την εντροπία στην είσοδο καθώς και τις εντροπίες θορύβου και διαύλου.

Θεώρημα 3.1 (Εντροπία στην Έξοδο ενός Διαύλου).
Αν η είσοδος και η έξοδος ενός διαύλου περιγράφονται ως δύο πηγές πληροφορίας (X, P_X) *και* (Y, P_Y)*, αντίστοιχα, τότε η εντροπία στην έξοδο του διαύλου δίνεται από τη σχέση*

$$H(Y) = H(X) - H(X/Y) + H(Y/X). \tag{3.7}$$

Απόδειξη.

$$H(X) - H(X/Y) + H(Y/X) = -\sum_{x\in X} p(x)\log(p(x)) +$$
$$+\sum_{x\in X}\sum_{y\in Y} p(x,y)\log(p(x/y)) - \sum_{x\in X}\sum_{y\in Y} p(x,y)\log(p(y/x)) =$$
$$= -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log(p(x)) +$$
$$+\sum_{x\in X}\sum_{y\in Y} p(x,y)\log(p(x/y)) - \sum_{x\in X}\sum_{y\in Y} p(x,y)\log(p(y/x)) =$$
$$= -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log\left(\frac{p(x)p(y/x)}{p(x/y)}\right) = -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log(p(y)) =$$
$$= -\sum_{y\in Y} p(y)\log(p(y)) = H(Y)$$

♦

3.4 Χωρητικότητα Διαύλου Πληροφορίας

Όπως αναφέρθηκε στο προηγούμενο κεφάλαιο η διαπληροφορία μεταξύ δύο πηγών πληροφορίας (X, P_X) και (Y, P_Y) ορίζεται ως η διαφορά

$$I(X;Y) = H(X) - H(X/Y). \tag{3.8}$$

Αν υποθέσουμε ότι οι δύο αυτές πηγές (X, P_X) και (Y, P_Y) περιγράφουν την είσοδο και την έξοδο ενός διαύλου, αντίστοιχα, τότε η διαπληροφορία θα είναι ίση με την εντροπία στην είσοδο του διαύλου μείον την εντροπία του θορύβου στο περιβάλλον του. Εναλλακτικά, η διαπληροφορία ως συμμετρική συνάρτηση (βλέπε σχέση (2.25» μπορεί να γραφεί ως η διαφορά της εντροπίας στην έξοδο μείον την εντροπία διαύλου

$$I(X;Y) = H(Y) - H(Y/X). \tag{3.9}$$

Ορισμός 3.1. ***Χωρητικότητα διαύλου πληροφορίας*** *C είναι το μέγιστο της διαπληροφορίας της εισόδου X και της εξόδου Y σε περιβάλλον θορύβου, δηλαδή*

$$C = \max_{P_X}\{I(X;Y)\}. \tag{3.10}$$

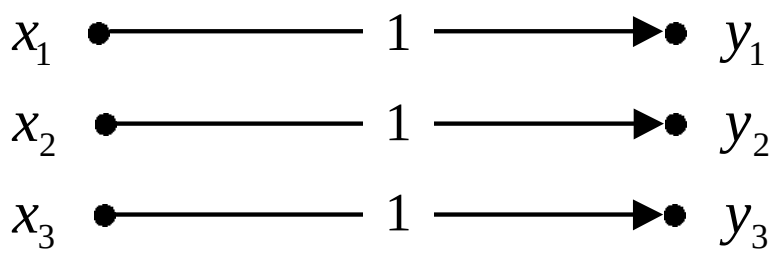
Από τον παραπάνω ορισμό της χωρητικότητας διαύλου προκύπτει ότι η χωρητικότητα μετράται σε bits/symbol όπως και η διαπληροφορία και μεγιστοποιείται για συγκεκριμένη κατανομή πιθανότητας των συμβόλων εισόδου. Δηλαδή, με κατάλληλη κατανομή των πιθανοτήτων των συμβόλων εισόδου μπορεί να μεγιστοποιηθεί η ποσότητα πληροφορίας που διοχετεύεται στο δίαυλο, οπότε επιτυγχάνεται η βέλτιστη εκμετάλλευση του διαύλου.

3.5 Χαρακτηριστικοί Δίαυλοι Πληροφορίας

Στη συνέχεια, παρουσιάζονται ορισμένοι χαρακτηριστικοί δίαυλοι πληροφορίας, για κάθε έναν από τους οποίους θα υπολογίσουμε τη χωρητικότητα. των οποίων υπολογίζεται η χωρητικότητα. Οι δίαυλοι αυτοί μπορούν να χρησιμοποιηθούν ως στοιχειώδη δομικά στοιχεία για την περιγραφή και ανάλυση πιο σύνθετων διαύλων πληροφορίας.

3.5.1 Ιδανικός δίαυλος

Στον ιδανικό δίαυλο πληροφορίας το πλήθος των συμβόλων είναι το ίδιο για την είσοδο και την έξοδο ($N = M$). Επίσης, για δεδομένο σύμβολο εισόδου ή εξόδου



Σχήμα 3.3: *Διάγραμμα ενός ιδανικού διαύλου.*

δεν υπάρχει αβεβαιότητα για το σύμβολο εξόδου ή εισόδου, αντίστοιχα. Δηλαδή, όλες οι υπό συνθήκη πιθανότητες $p(x_i/y_j)$ και $p(y_j/x_i)$ θα είναι ίσες με 0 ή ένα. Ακόμη, αυτό σημαίνει ότι αν το σύμβολο εισόδου x_i αντιστοιχίζεται στο σύμβολο εξόδου y_i, τότε οι πιθανότητες των δύο συμβόλων είναι ίσες ($p(x_i) = p(y_i)$) και κατ' επέκταση $P_X = P_Y$. Στο σχήμα 3.3 παρουσιάζεται ένα παράδειγμα διαγράμματος ιδανικού διαύλου.

Αφού όπως προαναφέραμε οι υπό συνθήκη πιθανότητες είναι μηδενικές ή ίσες με ένα, προκύπτει ότι η εντροπία θορύβου και η εντροπία διαύλου είναι μηδενικές:

$$H(X/Y) = H(Y/X) = 0, \tag{3.11}$$

ενώ ο πίνακας διαύλου θα είναι ο μοναδιαίος, διαστάσεων $N \times N$:

$$P_{Y/X} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & 1 \end{bmatrix}. \tag{3.12}$$

Από την (3.11) προκύπτει ότι η διαπληροφορία της εισόδου και της εξόδου θα είναι:

$$I(X;Y) = H(X) = H(Y). \tag{3.13}$$

Τέλος, η χωρητικότητα του ιδανικού διαύλου θα είναι ίση με

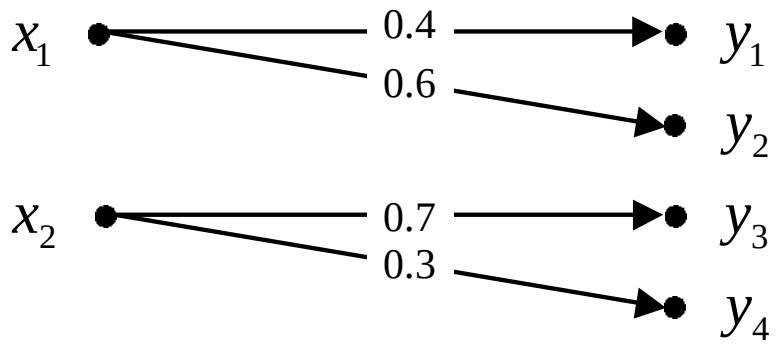
$$C = \max_{P_X}\{I(X;Y)\} = \max_{P_X}\{H(X)\} = \max_{P_Y}\{H(Y)\} = \log(N). \tag{3.14}$$

Προφανώς, η κατάλληλη κατανομή των πιθανοτήτων P_X για τα σύμβολα εισόδου που μεγιστοποιεί την $H(X)$ είναι η ομοιόμορφη:

$$P_X = [1/N, 1/N, \ldots, 1/N]. \tag{3.15}$$

3.5.2 Δίαυλος χωρίς απώλειες

Στο δίαυλο χωρίς απώλειες κάθε σύμβολο εισόδου μπορεί να αντιστοιχίζεται σε περισσότερα από ένα σύμβολα εξόδου. Όμως, κάθε σύμβολο εξόδου προέρχεται από



Σχήμα 3.4: *Διάγραμμα ενός διαύλου χωρίς απώλειες.*

ένα μόνο σύμβολο εισόδου, όπως φαίνεται στο σχήμα 3.4. Προφανώς, αυτό σημαίνει ότι τα σύμβολα εξόδου είναι περισσότερα από τα σύμβολα εισόδου ($M > N$). Ακόμη, για δεδομένο σύμβολο στην έξοδο του διαύλου δεν υπάρχει αβεβαιότητα για το αντίστοιχο σύμβολο στην είσοδο. Δηλαδή, κάθε υπό συνθήκη πιθανότητα $p(x_i/y_j)$ θα είναι ίση με 0 ή 1. Συνεπώς, για την εντροπία θορύβου και τη διαπληροφορία θα ισχύουν

$$H(X/Y) = 0, \tag{3.16}$$

$$I(X;Y) = H(X). \tag{3.17}$$

Σημειώνεται ότι οι υπό συνθήκη πιθανότητες $p(y_j/x_i)$ δεν είναι απαραίτητα ίσες με 0 ή 1. Συγκεκριμένα, ο πίνακας διαύλου $P_{Y/X}$ έχει μόνο ένα μη μηδενικό στοιχείο σε κάθε στήλη του. Για παράδειγμα, ο πίνακας του διαύλου στο σχήμα 3.4 είναι ίσος με:

$$\begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) & p(y_3/x_1) & p(y_4/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & p(y_3/x_2) & p(y_4/x_2) \end{bmatrix} = \begin{bmatrix} 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0.7 & 0.3 \end{bmatrix}.$$

Τέλος, η χωρητικότητα του διαύλου χωρίς απώλειες θα είναι:

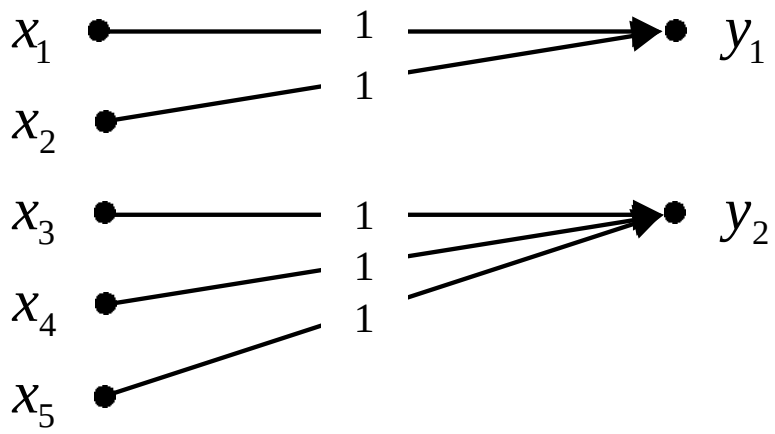
$$C = \max_{P_X}\{I(X;Y)\} = \max_{P_X}\{H(X)\} = \log(N), \tag{3.18}$$

που προκύπτει για ισοπίθανα σύμβολα εισόδου

$$P_X = [1/N,\, 1/N,\, \ldots,\, 1/N]. \tag{3.19}$$

3.5.3 Καθοριστικός δίαυλος

Ο καθοριστικός δίαυλος έχει ουσιαστικά τις αντίστροφες ιδιότητες σε σύγκριση με το δίαυλο χωρίς απώλειες. Πρώτα από όλα, τα σύμβολα εξόδου είναι γενικά λιγότερα



Σχήμα 3.5: *Διάγραμμα ενός καθοριστικού διαύλου.*

από τα σύμβολα εισόδου ($M < N$). Η βασική ιδιότητα του καθοριστικού διαύλου είναι ότι δεδομένου του συμβόλου στην είσοδο δεν υπάρχει αβεβαιότητα για το ποιο σύμβολο θα εμφανιστεί στην έξοδο. Δηλαδή, όλες οι υπό συνθήκη πιθανότητες $p(y_j/x_i)$ θα είναι ίσες είτε με 0 είτε με 1. Σε αυτήν την περίπτωση θα έχουμε

$$H(Y/X) = 0, \tag{3.20}$$

$$I(X;Y) = H(Y). \tag{3.21}$$

Επομένως, η χωρητικότητα του καθοριστικού διαύλου είναι:

$$C = \max_{P_X}\{I(X;Y)\} = \max_{P_X}\{H(Y)\} = \log(M), \tag{3.22}$$

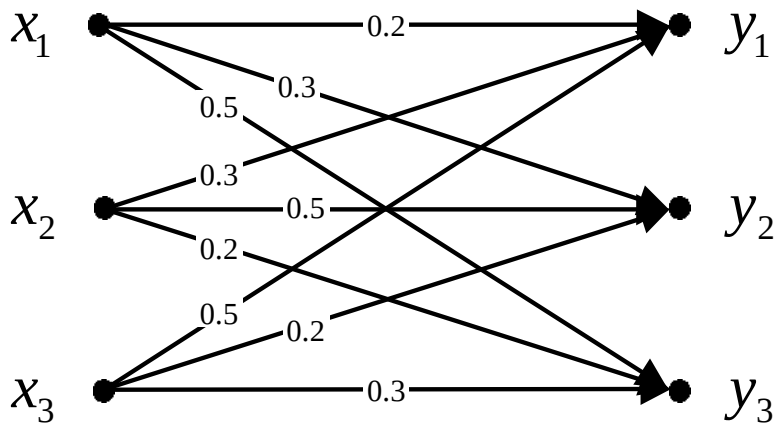
που προκύπτει για ισοπίθανα σύμβολα εισόδου:

$$P_Y = [1/M,\, 1/M,\, \ldots,\, 1/M]. \tag{3.23}$$

Η ομοιόμορφη κατανομή των πιθανοτήτων των συμβόλων στην έξοδο επιτυγχάνεται με τον κατάλληλο προσδιορισμό της κατανομής πιθανοτήτων των συμβόλων στην είσοδο. Στο σχήμα 3.5 παρουσιάζεται ένα παράδειγμα διαγράμματος καθοριστικού διαύλου. Ο πίνακας του διαύλου σε αυτό το παράδειγμα είναι:

$$\begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) \\ p(y_1/x_2) & p(y_2/x_2) \\ p(y_1/x_3) & p(y_2/x_3) \\ p(y_1/x_4) & p(y_2/x_4) \\ p(y_1/x_5) & p(y_2/x_5) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Γενικά, σε κάθε γραμμή του πίνακα ενός καθοριστικού διαύλου υπάρχει ένα μόνο μη μηδενικό στοιχείο ίσο πάντα με 1.



Σχήμα 3.6: *Διάγραμμα ενός ομοιόμορφου διαύλου.*

3.5.4 Ομοιόμορφος δίαυλος

Ομοιόμορφος ονομάζεται ένας δίαυλος πληροφορίας όταν όλες οι γραμμές και όλες οι στήλες του πίνακα διαύλου προκύπτουν από αναδιάταξη των στοιχείων οποιασδήποτε γραμμής και στήλης, αντίστοιχα. Το διάγραμμα ενός ομοιόμορφου διαύλου φαίνεται στο σχήμα 3.6. Ο πίνακας αυτού του διαύλου είναι ίσος με:

$$\begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) & p(y_3/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & p(y_3/x_2) \\ p(y_1/x_3) & p(y_2/x_3) & p(y_3/x_3) \end{bmatrix} = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.3 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.3 \end{bmatrix}.$$

Η εντροπία ενός ομοιόμορφου διαύλου θα είναι:

$$\begin{aligned} H(Y/X) &= -\sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j)\log(p(y_j/x_i)) = \\ &= -\sum_{i=1}^{N} p(x_i)\sum_{j=1}^{M} p(y_j/x_i)\log(p(y_j/x_i)) = \\ &= -\left[\sum_{j=1}^{M} p(y_j/x_i)\log(p(y_j/x_i))\right]\left[\sum_{i=1}^{N} p(x_i)\right] = \\ &= -\sum_{j=1}^{M} p(y_j/x_i)\log(p(y_j/x_i)). \end{aligned} \tag{3.24}$$

Η παραπάνω ανάλυση βασίζεται στο γεγονός ότι το άθροισμα

$$\sum_{j=1}^{M} p(y_j/x_i)\log(p(y_j/x_i)) \tag{3.25}$$

είναι το ίδιο για οποιαδήποτε γραμμή του πίνακα του διαύλου άρα ανεξάρτητο του δείκτη i. Παρατηρούμε ότι η εντροπία του ομοιόμορφου διαύλου είναι ουσιαστικά ανεξάρτητη της κατανομής πιθανοτήτων των συμβόλων στην είσοδο.

Η διαπληροφορία του ομοιόμορφου διαύλου είναι ίση με:

$$I(X;Y) = H(Y) - H(Y/X) = H(Y) + \sum_{j=1}^{M} p(y_j/x_i)\log(p(y_j/x_i)). \tag{3.26}$$

Συνεπώς, η χωρητικότητα του ομοιόμορφου διαύλου θα είναι ίση με τη μέγιστη τιμή της διαπληροφορίας δηλαδή:

$$C = \log(M) + \sum_{j=1}^{M} p(y_j/x_i)\log(p(y_j/x_i)). \tag{3.27}$$

Η μέγιστη αυτή τιμή της διαπληροφορίας επιτυγχάνεται με κατάλληλη επιλογή της κατανομής πιθανοτήτων των συμβόλων εισόδου, έτσι ώστε, η κατανομή των συμβόλων εξόδου να είναι ομοιόμορφη:

$$P_Y = [1/M,\, 1/M,\, \ldots,\, 1/M]. \tag{3.28}$$

3.5.5 Συμμετρικός δυαδικός δίαυλος

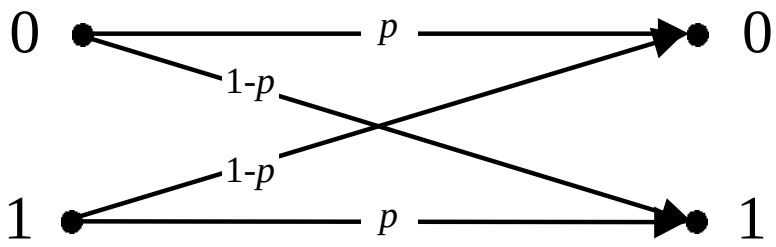
Ένας χαρακτηριστικός ομοιόμορφος δίαυλος είναι ο συμμετρικός δυαδικός. Σε αυτόν το δίαυλο έχουμε δύο σύμβολα εισόδου $\{0_x, 1_x\}$ και δύο σύμβολα εξόδου $\{0_y, 1_y\}$ (σχήμα 3.7). Οι δείκτες x, y στα σύμβολα δηλώνουν ότι πρόκειται για σύμβολα στην είσοδο και την έξοδο, αντίστοιχα. Κάθε σύμβολο εισόδου θα αντιστοιχίζεται σωστά στην έξοδο με πιθανότητα p. Προφανώς, σε αυτήν την περίπτωση, κάθε σύμβολο εισόδου μπορεί να αντιστοιχηθεί εσφαλμένα στην έξοδο του διαύλου με πιθανότητα $1-p$. Συνεπώς, ο πίνακας του συμμετρικού δυαδικού διαύλου θα είναι:

$$\begin{bmatrix} p(0_y/0_x) & p(1_y/0_x) \\ p(0_y/1_x) & p(1_y/1_x) \end{bmatrix} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}.$$

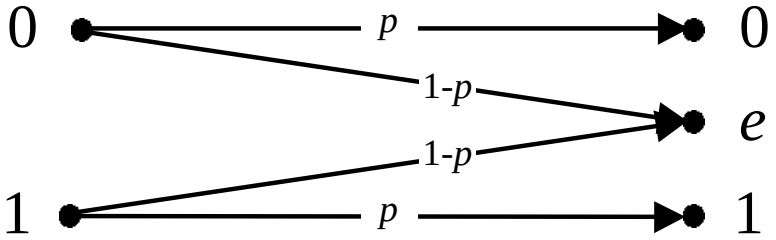
Εφαρμόζοντας, απευθείας τη σχέση (3.27) για τον υπολογισμό της χωρητικότητας ομοιόμορφου διαύλου προκύπτει ότι η χωρητικότητα του συμμετρικού δυαδικού διαύλου είναι ίση με:

$$C = \log(2) + p\log(p) + (1-p)\log(1-p) = 1 - H_b(p), \tag{3.29}$$

όπου $H_b(p)$ είναι η συνάρτηση Shannon (2.13).



Σχήμα 3.7: *Διάγραμμα συμμετρικού δυαδικού διαύλου.*



Σχήμα 3.8: *Διάγραμμα Σ-διαύλου.*

3.5.6 Δυαδικός δίαυλος εξάλειψης - Σ-δίαυλος

Ο δυαδικός δίαυλος εξάλειψης που φαίνεται στο σχήμα 3.8 έχει δύο σύμβολα εισόδου $\{0_x, 1_x\}$ και τρία σύμβολα εξόδου $\{0_y, 1_y, e\}$. Κάθε ένα από τα σύμβολα εισόδου αντιστοιχίζεται σωστά στην έξοδο με πιθανότητα p. Στην περίπτωση εσφαλμένης αντιστοίχισης εμφανίζεται στην έξοδο το σύμβολο e. Έτσι, αν στην έξοδο του διαύλου εμφανιστεί το σύμβολο e (σύμβολο λάθους), ο δέκτης αντιλαμβάνεται ότι το σύμβολο στην είσοδο μεταφράστηκε στην έξοδο εσφαλμένα, χωρίς όμως να γνωρίζει ποιο από τα δύο σύμβολα $\{0_x, 1_x\}$ εκπέμφηκε αρχικά. Ο πίνακας του διαύλου παίρνει τη μορφή

$$\begin{bmatrix} p(0_y/0_x) & p(e/0_x) & p(1_y/0_x) \\ p(0_y/1_x) & p(e/0_x) & p(1_y/1_x) \end{bmatrix} = \begin{bmatrix} p & 1-p & 0 \\ 0 & 1-p & p \end{bmatrix}.$$

Με βάση τα παραπάνω, οι υπό συνθήκη πιθανότητες δεδομένων των συμβόλων στην έξοδο θα είναι:

$$p(0_x/0_y) = 1, \quad p(0_x/1_y) = 0,$$
$$p(1_x/0_y) = 0, \quad p(1_x/1_y) = 1.$$

Επίσης, οι κοινές πιθανότητες $p(0_x, 1_y)$ και $p(1_x, 0_y)$ θα είναι ίσες με μηδέν ($p(0_x, 1_y) = p(1_x, 0_y) = 0$), ενώ η πιθανότητα σφάλματος $p(e)$ θα ισούται με:

$$p(e) = p(0_x, e) + p(1_x, e) = p(e/0_x)p(0_x) + p(e/1_x)p(1_x) =$$
$$= (1-p)p(0_x) + (1-p)p(1_x) = 1-p.$$

Συνεπώς, η υπό συνθήκη εντροπία της εισόδου δεδομένης της εξόδου θα είναι:

$$\begin{aligned} H(X/Y) &= -p(0_x, e)\log(p(0_x/e)) - p(1_x, e)\log(p(1_x/e)) = \\ &= -p(e/0_x)p(0_x)\log\left(\frac{p(e/0_x)p(0_x)}{p(e)}\right) \\ &\quad -p(e/1_x)p(1_x)\log\left(\frac{p(e/1_x)p(1_x)}{p(e)}\right) = \\ &= -(1-p)p(0_x)\left[\log(p(0_x)) + \log\left(\frac{1-p}{p(e)}\right)\right] \\ &\quad -(1-p)p(1_x)\left[\log(p(1_x)) + \log\left(\frac{1-p}{p(e)}\right)\right] = \\ &= (1-p)H(X). \end{aligned} \tag{3.30}$$

Τελικά, για τη διαπληροφορία εισόδου-εξόδου του διαύλου θα έχουμε:

$$I(X;Y) = H(X) - H(X/Y) = H(X) - (1-p)H(X) = pH(X), \tag{3.31}$$

οπότε, η χωρητικότητα του δυαδικού διαύλου εξάλειψης θα είναι:

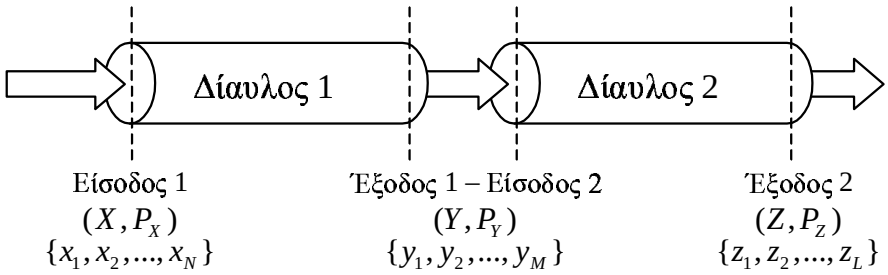
$$C = \max_{P_X}\{I(X;Y)\} = p\log(2) = p, \tag{3.32}$$

που προκύπτει για ισοπίθανα σύμβολα εισόδου $P_X = \{0.5, 0.5\}$.

3.6 Αλυσίδα Διαύλων Πληροφορίας

Στην πράξη, κατά τη μετάδοση της πληροφορίας σε ένα σύστημα επικοινωνίας, η πληροφορία διοχετεύεται σταδιακά μέσα από μία σειρά διαδοχικών διαύλων. Για παράδειγμα, κατά την επικοινωνία δύο υπολογιστών τα δεδομένα μπορεί να 'ταξιδεύουν' διαδοχικά μέσα από αγώγιμα καλώδια, οπτικές ίνες, την ατμόσφαιρα (ασύρματες ζεύξεις), κλπ. Αυτή η αλυσιδωτή σύνδεση πολλών διαύλων πληροφορίας ονομάζεται **αλυσίδα διαύλων πληροφορίας**. Προφανώς, αυτή η σύνδεση υπονοεί ότι η έξοδος κάθε διαύλου διοχετεύεται ως είσοδος στον επόμενο δίαυλο. Για παράδειγμα, η περίπτωση μίας αλυσίδας δύο διαύλων φαίνεται στο σχήμα 3.9, όπου η πηγή εξόδου (Y, P_Y) του διαύλου 1 ταυτίζεται με την πηγή εισόδου του διαύλου 2.

Έστω ότι η αλυσίδα διαύλων αποτελείται από ένα σύνολο K απλούστερων διαύλων, οι οποίοι είναι συνδεδεμένοι κατά αύξουσα σειρά $(1, 2, 3, \ldots, K)$. Θεωρούμε ότι γνωρίζουμε τον πίνακα $P(Y_k/X_k)$ κάθε ενός στοιχειώδους διαύλου ο οποίος συμβολίζεται με το γράμμα k. Η είσοδος του διαύλου k είναι η πηγή X_k, ενώ η έξοδός του είναι η Y_k. Προφανώς, η είσοδος X_k του διαύλου k ταυτίζεται με την έξοδο



Σχήμα 3.9: *Αλυσίδα δύο διαύλων.*

Y_{k-1} του προηγούμενου διαύλου ($X_k \equiv Y_{k-1}$). Επίσης, η έξοδος του Y_k του διαύλου k ταυτίζεται με την είσοδο X_{k+1} του επόμενου διαύλου ($Y_k \equiv X_{k+1}$). Όπως έχουμε ήδη αναφέρει, αν γνωρίσουμε την κατανομή πιθανοτήτων στην είσοδο P_X και το πίνακα ενός διαύλου $P_{Y/X}$ μπορούμε να υπολογίσουμε την κατανομή στην έξοδό του εφαρμόζοντας τη σχέση (3.4) $P_Y = P_X \cdot P_{Y/X}$. Συνεπώς, στην περίπτωση της αλυσίδας διαύλων η κατανομή πιθανοτήτων στην έξοδο της αλυσίδα θα δίνεται από το γινόμενο

$$P_{Y_K} = P_{X_1} \cdot P_{Y_1/X_1} \cdot P_{Y_2/X_2} \cdots P_{Y_K/X_K} = P_{X_1} \cdot \prod_{k=1}^{K} P_{Y_k/X_k}. \tag{3.33}$$

Από τη σχέση (3.33) προκύπτει ότι ο πίνακας της αλυσίδας των διαύλων θα είναι:

$$P_{Y_K/X_1} = \prod_{k=1}^{K} P_{Y_k/X_k}. \tag{3.34}$$

Ένα σημαντικό γεγονός που σχετίζεται με τη χωρητικότητα μίας αλυσίδα διαύλων είναι ότι η χωρητικότητα της αλυσίδας στο σύνολό της είναι μικρότερη από τη χωρητικότητα κάθε στοιχειώδους διαύλου που μετέχει στην αλυσίδα, χωριστά. Ας θεωρήσουμε για παράδειγμα την απλή περίπτωση αλυσίδας δύο διαύλων, δηλαδή την αλυσίδα του σχήματος 3.9. Μπορούμε να αποδείξουμε ότι η διαπληροφορία $I(X;Z)$ είναι πάντα μικρότερη από τη διαπληροφορία $I(X;Y)$, οπότε κατ' επέκταση η χωρητικότητα της αλυσίδας θα είναι μικρότερη από αυτήν του απλού διαύλου.

Απόδειξη.

$$I(X;Y) - I(X;Z) = H(X) - H(X/Y) - H(X) + H(X/Z) =$$
$$= H(X/Z) - H(X/Y) =$$

$$= -\sum_{i=1}^{N}\sum_{l=1}^{L} p(x_i, z_l) \log(p(x_i/z_l)) + \sum_{i=1}^{N}\sum_{j=1}^{M} p(x_i, y_j) \log(p(x_i/y_j)) =$$

$$= -\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{l=1}^{L} p(x_i, y_j, z_l)\log(p(x_i/z_l))$$
$$+\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{l=1}^{L} p(x_i, y_j, z_l)\log(p(x_i/y_j)) =$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{l=1}^{L} p(x_i, y_j, z_l)\log\left(\frac{p(x_i/y_j)}{p(x_i/z_l)}\right). \tag{3.35}$$

Οι δύο απλοί δίαυλοι λειτουργούν ανεξάρτητα οπότε ισχύει:

$$p(x_i/y_j, z_l) = p(x_i/y_j). \tag{3.36}$$

Επίσης, γνωρίζουμε ότι

$$p(x_i, y_j, z_l) = p(x_i/y_j, z_l)p(y_j, z_l). \tag{3.37}$$

Αντικαθιστώντας τις (3.36) και (3.37) στην (3.35) έχουμε

(3.38)
$$I(X;Y) - I(X;Z) = \sum_{j=1}^{M}\sum_{l=1}^{L} p(y_j, z_l)\left[\sum_{i=1}^{N} p(x_i/y_j, z_l)\log\left(\frac{p(x_i/y_j, z_l)}{p(x_i/z_l)}\right)\right].$$

Στην παραπάνω σχέση το άθροισμα εντός της αγκύλης είναι στην ουσία μία απόσταση Kullback-Leibler, η οποία όπως έχουμε ήδη αναφέρει είναι μη αρνητική. Συνεπώς, η (3.38) θα είναι πάντα μη αρνητική,

$$I(X;Y) - I(X;Z) \geq 0, \tag{3.39}$$

οπότε κατ' επέκταση θα έχουμε

$$\max_{P_X}\{I(X;Y)\} - \max_{P_X}\{I(X;Z)\} \geq 0 \Rightarrow$$
$$\Rightarrow C_{X;Y} \geq C_{X;Z}. \tag{3.40}$$

Αυτό σημαίνει ότι η χωρητικότητα της αλυσίδας θα είναι μικρότερη από τη χωρητικότητα κάθε διαύλου χωριστά.

♦

3.7 Υπολογισμός της Χωρητικότητας - Τεχνική Muroga

Στις προηγούμενες ενότητες περιγράψαμε ορισμένους χαρακτηριστικούς δίαυλους πληροφορίας των οποίων προσδιορίσαμε, με σχετική ευκολία, τη χωρητικότητα. Το γενικό πρόβλημα του υπολογισμού της χωρητικότητας ενός διαύλου, όταν μας δίνεται ο πίνακας $P_{Y/X}$, είναι ιδιαίτερα δύσκολο και εμπίπτει στην περιοχή της θεωρίας βελτιστοποίησης υπό συνθήκες.

Στο σημείο αυτό θα περιγράψουμε, χωρίς απόδειξη, μία τεχνική υπολογισμού της χωρητικότητας διαύλου όταν είναι γνωστός ο πίνακας του διαύλου $P_{Y/X}$, και θεωρώντας ότι τα αλφάβητα εισόδου και εξόδου είναι ισάριθμα ($N = M$). Η τεχνική με την οποία επιτυγχάνεται ο υπολογισμός ονομάζεται **τεχνική Muroga** και αποτελείται από δύο βασικά βήματα.

1. Αρχικά συνθέτουμε το σύστημα N γραμμικών εξισώσεων:

$$p(y_1/x_1)\cdot A_1 + p(y_2/x_1)\cdot A_2 + \cdots + p(y_N/x_1)\cdot A_N = \\ = \sum_{j=1}^{N} p(y_j/x_1)\log(p(y_j/x_1))$$

$$p(y_1/x_2)\cdot A_1 + p(y_2/x_2)\cdot A_2 + \cdots + p(y_N/x_2)\cdot A_N = \\ = \sum_{j=1}^{N} p(y_j/x_2)\log(p(y_j/x_2))$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$p(y_1/x_N)\cdot A_1 + p(y_2/x_N)\cdot A_2 + \cdots + p(y_N/x_N)\cdot A_N = \\ = \sum_{j=1}^{N} p(y_j/x_N)\log(p(y_j/x_N))$$

   το οποίο επιλύεται με αγνώστους τις μεταβλητές $A_1, A_2, \ldots, A_N$.

2. Μετά την επίλυση του παραπάνω συστήματος εξισώσεων η χωρητικότητα του διαύλου υπολογίζεται με βάση τη σχέση:

$$C = \log\left(\sum_{i=1}^{N} 2^{A_i}\right). \tag{3.41}$$

Οι πιθανότητες των συμβόλων στην έξοδο που οδηγούν στη μέγιστη διαπληροφορία εισόδου-εξόδου του διαύλου (χωρητικότητα) δίνονται από τις σχέσεις:

$$p(y_j) = 2^{A_j - C}, \quad 1 \le j \le N. \tag{3.42}$$

Στη συνέχεια, η κατανομή των πιθανοτήτων των συμβόλων στην είσοδο του διαύλου που οδηγεί στην πλήρη εκμετάλλευση της χωρητικότητάς του, θα δίνεται από την επίλυση του συστήματος εξισώσεων

$$P_Y = P_X \cdot P_{Y/X}, \tag{3.43}$$

δεδομένου ότι ο πίνακας του διαύλου $P_{Y/X}$ είναι γνωστός και οι πιθανότητες στην έξοδο έχουν υπολογιστεί με βάση τη σχέση (3.42).

3.8 Παραδείγματα

1. *Αν η πιθανότητα σφάλματος κατά τη μετάδοση δυαδικών συμβόλων είναι* $e = 0.1$ *να υπολογιστεί η χωρητικότητα του συμμετρικού δυαδικού διαύλου και του δυαδικού διαύλου εξάλειψης.*

   Η πιθανότητα ορθής μετάδοσης των δυαδικών συμβόλων είναι $p = 1 - 0.1 = 0.9$.

   Για το συμμετρικό δυαδικό δίαυλο η χωρητικότητα δίνεται από τη σχέση (3.29), οπότε

   $$C = 1 - H_b(p) = 1 + p\log(p) + (1-p)\log(1-p) =$$
   $$1 + 0.9\log(0.9) + 0.1\log(0.1) \simeq 0.531 \quad bits/symbol$$

   Για το δυαδικό δίαυλο εξάλειψης η χωρητικότητα δίνεται από τη σχέση (3.32), οπότε

   $$C = p = 0.9 \quad bits/symbol$$

2. *Θεωρούμε ομοιόμορφο δίαυλο πληροφορίας του οποίου ο πίνακας είναι ίσος με:*

   $$P_{Y/X} = \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.5 & 0.3 & 0.2 \\ 0.3 & 0.2 & 0.5 \end{bmatrix}.$$

   Να υπολογιστεί η χωρητικότητα του διαύλου.

   Η χωρητικότητα του ομοιόμορφου διαύλου δίνεται από τη σχέση (3.27), οπότε

   $$C = \log(M) + \sum_{j=1}^{M} p(y_j/x_i)\log(p(y_j/x_i)) =$$
   $$= \log(3) + 0.2\log(0.2) + 0.3\log(0.3) + 0.5\log(0.5) \simeq$$
   $$\simeq 0.099 \quad bits/symbol$$

3. *Ο πίνακας των συνδετικών πιθανοτήτων των συμβόλων εισόδου-εξόδου ενός υποθετικού διαύλου πληροφορίας είναι:*

$$P_{XY} = \begin{bmatrix} 0.25 & 0 & 0.1 \\ 0 & 0.3 & 0.05 \\ 0.1 & 0.05 & 0 \\ 0 & 0 & 0.15 \end{bmatrix}.$$

Να υπολογιστούν: η εντροπία εισόδου $H(X)$, η εντροπία εξόδου $H(Y)$, ο πίνακας του διαύλου $P_{Y/X}$, η εντροπία διαύλου $H(Y/X)$ και η εντροπία θορύβου $H(X/Y)$.

Για να υπολογίσουμε τις εντροπίες στην είσοδο και την έξοδο του διαύλου θα πρέπει να γνωρίζουμε τις πιθανότητες εμφάνισης των συμβόλων εισόδου και εξόδου, αντίστοιχα. Οι πιθανότητες αυτές είναι δυνατό να βρεθούν από τον πίνακα συνδετικών πιθανοτήτων ως εξής:

$$p(x_1) = \sum_{j=1}^{3} p(x_1, y_j) = 0.25 + 0 + 0.1 = 0.35$$

$$p(x_2) = \sum_{j=1}^{3} p(x_2, y_j) = 0 + 0.3 + 0.05 = 0.35$$

$$p(x_3) = \sum_{j=1}^{3} p(x_3, y_j) = 0.1 + 0.05 + 0 = 0.15$$

$$p(x_4) = \sum_{j=1}^{3} p(x_4, y_j) = 0 + 0 + 0.15 = 0.15$$

$$p(y_1) = \sum_{i=1}^{4} p(x_i, y_1) = 0.25 + 0 + 0.1 + 0 = 0.35$$

$$p(y_2) = \sum_{i=1}^{4} p(x_i, y_2) = 0 + 0.3 + 0.05 + 0 = 0.35$$

$$p(y_3) = \sum_{i=1}^{4} p(x_i, y_3) = 0.1 + 0.05 + 0 + 0.15 = 0.3$$

Η εντροπία εισόδου θα είναι:

$$H(X) = -\sum_{i=1}^{4} p(x_i) \log(p(x_i)) = 1.881 \quad bits/symbol.$$

Ομοίως, η εντροπία εξόδου υπολογίζεται ως:

$$H(Y) = -\sum_{j=1}^{3} p(y_j) \log(p(y_j)) = 1.581 \quad bits/symbol.$$

Για να συνθέσουμε τον πίνακα του διαύλου $P_{Y/X}$ υπολογίζουμε τις υπό συνθήκη πιθανότητες $p(y_j/x_i) = p(y_j, x_i)/p(x_i)$. Έτσι, έχουμε

$$P_{Y/X} = \begin{bmatrix} 0.25/0.35 & 0 & 0.1/0.35 \\ 0 & 0.3/0.35 & 0.05/0.35 \\ 0.1/0.15 & 0.05/0.15 & 0 \\ 0 & 0 & 0.15/0.15 \end{bmatrix} = \begin{bmatrix} 5/7 & 0 & 2/7 \\ 0 & 6/7 & 1/7 \\ 2/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Η εντροπία του διαύλου $H(Y/X)$ υπολογίζεται ως εξής:

$$H(Y/X) = -\sum_{i=1}^{4}\sum_{j=1}^{3} p(x_i, y_j)\log(p(y_j/x_i)) = 0.6469 \quad bits/symbol.$$

Τέλος, για την εντροπία του θορύβου $H(X/Y)$ έχουμε:

$$H(X/Y) = -\sum_{i=1}^{4}\sum_{j=1}^{3} p(x_i, y_j)\log(p(x_i/y_j)) =$$

$$= -\sum_{i=1}^{4}\sum_{j=1}^{3} p(x_i, y_j)\log(\frac{p(x_i, y_j)}{p(y_j)}) = 0.947 \quad bits/symbol.$$

4. *Να υπολογιστούν οι χωρητικότητες των διαύλων που περιγράφονται από τους παρακάτω πίνακες διαύλων ($P_{Y/X}$):*

$$P^{(1)} = \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.25 & 0.75 \end{bmatrix}, P^{(2)} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, P^{(3)} = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix}.$$

Αν οι τρεις δίαυλοι συνδεθούν σε σειρά με πρώτο τον $P^{(1)}$ και τελευταίο τον $P^{(3)}$, να βρεθεί ο πίνακας διαύλου της αλυσίδας των διαύλων καθώς και η χωρητικότητα της αλυσίδας.

Παρατηρούμε ότι ο πίνακας του πρώτου διαύλου $P^{(1)}$ έχει τη μορφή

$$\begin{bmatrix} p & 1-p & 0 \\ 0 & 1-p & p \end{bmatrix}$$

όπου $p = 0.75$. Συνεπώς, πρόκειται για δυαδικό δίαυλο εξάλειψης, του οποίου η χωρητικότητα δίνεται από τη σχέση (3.32). Άρα η χωρητικότητα του πρώτου διαύλου θα είναι

$$C_1 = p = 0.75 \quad bits/symbol.$$

Από τον πίνακα του δεύτερου διαύλου προκύπτει ότι πρόκειται για καθοριστικό δίαυλο, του οποίου η χωρητικότητα δίνεται από την (3.22)

$$C_2 = \log(M) = \log(2) = 1 \quad bit/symbol.$$

Τέλος, ο πίνακας του τρίτου διαύλου έχει τη μορφή

$$\begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}.$$

Άρα πρόκειται για συμμετρικό δυαδικό δίαυλο, όπου $p = 0.6$. Σύμφωνα με τη σχέση (3.29), η χωρητικότητα του διαύλου θα είναι

$$C_3 = 1 - H_b(p) = 1 + 0.6\log(0.6) + 0.4\log(0.4) = 0.029 \quad bits/symbol.$$

Αν οι τρεις δίαυλοι συνδεθούν δημιουργώντας μία αλυσίδα διαύλων, τότε ο πίνακας P της αλυσίδας θα δίνεται από το γινόμενο των τριών πινάκων, σύμφωνα με τη σχέση (3.34), δηλαδή

$$P = P^{(1)} \cdot P^{(2)} \cdot P^{(3)} =$$

$$= \begin{bmatrix} 0.75 & 0.25 & 0 \\ 0 & 0.25 & 0.75 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix} =$$

$$= \begin{bmatrix} 0.6 & 0.4 \\ 0.45 & 0.55 \end{bmatrix}.$$

Για να υπολογίσουμε τη χωρητικότητα της αλυσίδας των τριών διαύλων θα εφαρμόσουμε την τεχνική Muroga. Αρχικά, σχηματίζουμε το σύστημα των δύο εξισώσεων με δύο αγνώστους

$$0.6 \cdot A_1 + 0.4 \cdot A_2 = 0.6\log(0.6) + 0.4\log(0.4) = -0.9710$$
$$0.45 \cdot A_1 + 0.55 \cdot A_2 = 0.45\log(0.45) + 0.55\log(0.55) = -0.9928$$

Από την επίλυση του παραπάνω συστήματος προκύπτουν

$$A_1 = -0.912866, \quad A_2 = -1.058199.$$

Τελικά, σύμφωνα με την τεχνική Muroga, η χωρητικότητα της αλυσίδας θα δίνεται από τη σχέση (3.41), δηλαδή

$$C = \log\left(\sum_{i=1}^{N} 2^{A_i}\right) = \log(2^{-0.912866} + 2^{-1.058199}) =$$

$$0.0163 \quad bits/symbol.$$

Παρατηρούμε ότι η χωρητικότητα της αλυσίδας των τριών διαύλων είναι μικρότερη από τη χωρητικότητα κάθε διαύλου χωριστά, όπως άλλωστε προβλέπει και η θεωρία.

3.9 Ασκήσεις

1. Να υπολογιστεί η χωρητικότητα του διαύλου πληροφορίας με πίνακα διαύλου

$$P_{Y/X} = \begin{bmatrix} p & 1-p & 0 & 0 \\ 1-p & p & 0 & 0 \\ 0 & 0 & p & 1-p \\ 0 & 0 & 1-p & p \end{bmatrix}.$$

2. Ο πίνακας των συνδετικών πιθανοτήτων των συμβόλων εισόδου-εξόδου ενός υποθετικού διαύλου πληροφορίας είναι:

$$P_{XY} = \begin{bmatrix} 0.2 & 0 & 0.1 & 0 \\ 0 & 0.3 & 0.05 & 0 \\ 0.1 & 0.05 & 0 & 0.2 \end{bmatrix}.$$

   Να υπολογιστούν: η εντροπία εισόδου $H(X)$, η εντροπία εξόδου $H(Y)$, ο πίνακας του διαύλου $P_{Y/X}$, η εντροπία διαύλου $H(Y/X)$ και η εντροπία θορύβου $H(X/Y)$.

3. Να υπολογιστεί η χωρητικότητα του διαύλου πληροφορίας με πίνακα διαύλου

$$P_{Y/X} = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 1/8 & 3/4 & 1/8 \\ 1/4 & 0 & 3/4 \end{bmatrix}.$$

4. Τα σύμβολα μιας δυαδικής πηγής πληροφορίας με αλφάβητο $X = \{x_1, x_2\}$ και κατανομή πιθανοτήτων $P_X = \{0.75, 0.25\}$ εισάγονται σε δίαυλο πληροφορίας με πίνακα διαύλου

$$P_{Y/X} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}.$$

   Να υπολογιστεί η διαπληροφορία εισόδου-εξόδου καθώς και η χωρητικότητα το διαύλου. Να αιτιολογηθεί η διαφορά μεταξύ των τιμών των δύο μεγεθών.

Κεφάλαιο 4

Κωδικοποίηση Πηγής

Στο σύστημα επικοινωνίας, όπως αυτό περιγράφεται από τη θεωρία της πληροφορίας, ανάμεσα στον πομπό και το δίαυλο πληροφορίας παρεμβάλλεται ο κωδικοποιητής. Έργο του κωδικοποιητή είναι να αντικαταστήσει τα σύμβολα πληροφορίας που παράγονται από την πηγή με εναλλακτικά σύμβολα ή λέξεις. Οι λόγοι για τους οποίους απαιτείται η παρουσία του κωδικοποιητή ποικίλλουν. Ενδεικτικά αναφέρουμε τους παρακάτω

- *Συμπίεση.* Κατά τη συμπίεση ο κωδικοποιητής αντικαταθιστά τα σύμβολα της πηγής με νέα σύμβολα με τελικό στόχο τη μείωση του συνολικού μήκους του μηνύματος πληροφορίας που θα διοχετευθεί στο δίαυλο. Φυσικά, η όλη διαδικασία βασίζεται στην προϋπόθεση ότι στο αρχικό μήνυμα πληροφορίας υπάρχουν πλεονασμοί. Προφανώς, με τη μείωση του μήκους του μηνύματος γίνεται εξοικονόμηση του χώρου αποθήκευσης του μηνύματος και μειώνεται ο χρόνος που απαιτείται για την αποστολή του.
- *Ανίχνευση και Διόρθωση Σφαλμάτων.* Μία πάγια τεχνική για την αντιμετώπιση των συνεπειών του θορύβου, ο οποίος υπάρχει κατά μήκος του διαύλου, είναι η αποστολή από τον πομπό πλεονάζουσας πληροφορίας με τη μορφή περίσσειας συμβόλων. Τα σύμβολα αυτά επιτρέπουν την ανίχνευση σφαλμάτων, όταν το μήνυμα φτάσει στο δέκτη. Ακόμη, υπάρχουν κώδικες που επιτρέπουν εκτός από την ανίχνευση των σφαλμάτων και τη διόρθωσή τους. Στην πράξη, ο κωδικοποιητής αντικαθιστά τα σύμβολα της πηγής με λέξεις μεγαλύτερου μήκους, από το αρχικώς απαιτούμενο, προσθέτοντας σύμβολα ελέγχου. Είναι φανερό ότι η κωδικοποίηση που στοχεύει στην ανίχνευση και διόρθωση σφαλμάτων έχει αντίθετη αρχή λειτουργίας από την κωδικοποίηση που στοχεύει στη συμπίεση.

- *Ασφάλεια.* Όταν ο κωδικοποιητής αντιστοιχεί τα σύμβολα της πηγής με ένα σύνολο λέξεων και η αντιστοίχιση αυτή είναι γνωστή μόνο στον πομπό και το δέκτη, είναι δυνατό να αποφευχθεί η υποκλοπή και κατανόηση του μηνύματος από τρίτους.

- *Βέλτιστη Εκμετάλλευση του Διαύλου.* Υπάρχει περίπτωση ο δίαυλος λόγω των τεχνικών του χαρακτηριστικών να μη μπορεί να μεταφέρει την πληροφορία με τη μορφή που έχουν τα σύμβολα που παράγονται από την πηγή. Σε αυτή την περίπτωση ο κωδικοποιητής μεταφράζει το αρχικό μήνυμα χρησιμοποιώντας σύμβολα που είναι δυνατό να διοχετευθούν με 'ευκολία' στο δίαυλο.

- *Συμβατότητα.* Στην περίπτωση που ο δέκτης δε γνωρίζει ή δεν αντιλαμβάνεται τα σύμβολα που χρησιμοποιεί ο πομπός απαιτείται η χρήση ενός αλφαβήτου επικοινωνίας που είναι αντιληπτό και από τους δύο. Το κοινό αυτό αλφάβητο που επιτρέπει τη συμβατότητα πομπού-δέκτη παρέχεται από τον κωδικοποιητή με τη διαδικασία της κωδικοποίησης των μηνυμάτων του πομπού. Στη συνέχεια, ο αποκωδικοποιητής μετατρέπει τις λέξεις του κοινού αλφαβήτου σε σύμβολα αντιληπτά από το δέκτη.

4.1 Ορισμοί

Ορισμός 4.1. ***Κωδικοποίηση*** *είναι η αντικατάσταση συμβόλων ή ομάδων συμβόλων πληροφορίας με νέα εναλλακτικά σύμβολα ή ομάδες συμβόλων.*

Ορισμός 4.2. ***Κώδικας*** *είναι ο αλγόριθμος με τον οποίο πραγματοποιείται η κωδικοποίηση, δηλαδή, η αντικατάσταση συμβόλων από νέα σύμβολα.*

Ορισμός 4.3. *Τα νέα διακεκριμένα σύμβολα $w_1, w_2, \ldots, w_M$ που προκύπτουν κατά τη διαδικασία της κωδικοποίησης ονομάζονται* ***κωδικά σύμβολα****. Το πλήθος τους M μπορεί να είναι πεπερασμένο ή άπειρο.*

Ορισμός 4.4. *Το σύνολο $W = \{w_1, w_2, \ldots, w_M\}$ όλων των κωδικών συμβόλων ονομάζεται* ***αλφάβητο κωδικοποίησης****.*

Ορισμός 4.5. *Ο κώδικας που έχει M κωδικά σύμβολα ονομάζεται* ***M-αδικός κώδικας****.*

Προφανώς, οι λέξεις και τα μηνύματα συνθέτονται από σύμβολα που ανήκουν είτε στο αλφάβητο της πηγής (λέξεις και μηνύματα πηγής) είτε στο αλφάβητο κωδικοποίησης

(κωδικές λέξεις και κωδικά μηνύματα) χωρίς να επιτρέπεται η ανάμειξη των δύο αλφαβήτων.

Χαρακτηριστικό παράδειγμα κωδικοποίησης είναι η απεικόνιση των γραμμάτων της αλφαβήτου σε κωδικές λέξεις που σχηματίζονται από τελείες, παύλες, και κενά διαστήματα. Δηλαδή, η γνωστή σε όλους κωδικοποίηση Morse. Στην περίπτωση της κωδικοποίησης Morse το αλφάβητο της πηγής είναι το σύνολο των ελληνικών ή λατινικών χαρακτήρων ($X = \{A, B, C, \ldots, Z\}$), ενώ το αλφάβητο κωδικοποίησης είναι το σύνολο $W = \{\cdot, -, \text{'κενό'}\}$.

Στο κεφάλαιο αυτό θα μας απασχολήσει η κωδικοποίηση που στοχεύει στη συμπίεση της πληροφορίας. Συγκεκριμένα, θα αναζητήσουμε εκείνα τα κριτήρια αναπαράστασης της πληροφορίας μία πηγής που επιτρέπουν τη χρήση κωδίκων, οι οποίοι οδηγούν σε σύντομη καταγραφή της πληροφορίας.

4.2 Ταξινόμηση των Κωδίκων

Αν το αλφάβητο της πηγής και της κωδικοποίησης είναι αντίστοιχα $X = \{x_1, x_2, \ldots, x_N\}$ και $W = \{w_1, w_2, \ldots, w_M\}$, τότε η διαδικασία της κωδικοποίησης μπορεί να περιγραφεί μαθηματικά ως η απεικόνιση C του αλφαβήτου της πηγής X σε μία ένωση U^* καρτεσιανών γινομένων του αλφαβήτου της κωδικοποίησης, δηλαδή

$$X \xmapsto{C} U^* = \bigcup_{r=1}^{R} W^r = \bigcup_{r=1}^{R} \underbrace{W \times W \times \cdots \times W}_{r}. \tag{4.1}$$

Σημειώνουμε ότι το καρτεσιανό γινόμενο W^r είναι στην ουσία το σύνολο των κωδικών λέξεων που σχηματίζονται από r σύμβολα του αλφαβήτου κωδικοποίησης. Για παράδειγμα, αν το αλφάβητο της κωδικοποίησης είναι το $W = \{\diamond, \circ\}$ τότε το καρτεσιανό γινόμενο W^2 είναι στην ουσία το σύνολο των κωδικών λέξεων $W^2 = \{\diamond\diamond, \diamond\circ, \circ\diamond, \circ\circ\}$. Επίσης, δεν είναι απαραίτητο ο κώδικας C να καλύπτει όλο το σύνολο U^*. Δηλαδή, δεν είναι απαραίτητο κάθε στοιχείο του U^* να είναι εικόνα κάποιου συμβόλου της πηγής. Στην πράξη, οι κωδικές λέξεις που χρησιμοποιεί ο κώδικας ανήκουν σε ένα υποσύνολο U του συνόλου U^* ($U \subseteq U^*$).

Στην περίπτωση του κώδικα Morse για την κωδικοποίηση του λατινικού αλφαβήτου έχουμε τα παρακάτω:

- Το αλφάβητο της πηγής είναι το σύνολο των 26 γραμμάτων του λατινικού αλφαβήτου ($N = 26$).

- Το αλφάβητο της κωδικοποίησης αποτελείται από τρία κωδικά σύμβολα ($M = 3$) και είναι το $W = \{\cdot, -, \text{'κενό'}\}$.
- Τα γράμματα του λατινικού αλφαβήτου αντιστοιχίζονται σε κωδικές λέξεις που έχουν μέγιστο μήκος 5 σύμβολα (μαζί με το κενό που βρίσκεται στο τέλος κάθε λέξης), όπως φαίνεται στον πίνακα 4.1.

Πίνακας 4.1: *Ο κώδικας Morse.*

A	$\cdot-$	J	$\cdot---$	S	$\cdots$
B	$-\cdots$	K	$-\cdot-$	T	$-$
C	$-\cdot-\cdot$	L	$\cdot-\cdot\cdot$	U	$\cdot\cdot-$
D	$-\cdot\cdot$	M	$--$	V	$\cdots-$
E	$\cdot$	N	$-\cdot$	W	$\cdot--$
F	$\cdot\cdot-\cdot$	O	$---$	X	$-\cdot\cdot-$
G	$--\cdot$	P	$\cdot--\cdot$	Y	$-\cdot--$
H	$\cdots\cdot$	Q	$--\cdot-$	Z	$--\cdot\cdot$
I	$\cdot\cdot$	R	$\cdot-\cdot$		

Συνεπώς οι κωδικές λέξεις του κώδικα Morse ανήκουν στο σύνολο

$$U^* = \bigcup_{r=1}^{5} W^r,$$

χωρίς όμως όλα τα στοιχεία του συνόλου U^* να είναι κωδικές λέξεις. Πρακτικά, οι κωδικές λέξεις συνθέτουν το σύνολο $U \subseteq U^*$

Είναι προφανές ότι η αποκωδικοποίηση περιγράφεται ως η αντίστροφη απεικόνιση C' από το σύνολο των κωδικών λέξεων στο αλφάβητο της πηγής, δηλαδή

$$U \xmapsto{C'} X. \tag{4.2}$$

4.2.1 Ταξινόμηση με κριτήριο τις απώλειες

Οι αλγόριθμοι κωδικοποίησης χωρίζονται σε δύο κατηγορίες ανάλογα με τη δυνατότητα πλήρους ή μερικής ανάκτησης του αρχικού μηνύματος, όταν μας δίνεται το κωδικό μήνυμα. Αν η απεικόνιση C που περιγράφει τη διαδικασία κωδικοποίησης είναι ένα προς ένα, δηλαδή διαφορετικά σύμβολα του αλφαβήτου της πηγής αντιστοιχίζονται σε διαφορετικές κωδικές λέξεις, τότε είναι φανερό ότι η απεικόνιση C' που περιγράφει τη διαδικασία αποκωδικοποίησης θα είναι η C^{-1}, δηλαδή η αντίστροφη

της C. [1] Σε αυτήν την περίπτωση, αν το αρχικό μήνυμα κωδικοποιηθεί, και στη συνέχεια το κωδικό μήνυμα που προέκυψε αποκωδικοποιηθεί, θα καταλήγουμε στο αρχικό μήνυμα, χωρίς καμία διαφορά άρα απώλεια. Οι κώδικες που ικανοποιούν αυτήν την ιδιότητα λέγονται **κώδικες χωρίς απώλειες**. Χαρακτηριστικό παράδειγμα κώδικα χωρίς απώλειες είναι η κωδικοποίηση των κεφαλαίων γραμμάτων του λατινικού αλφαβήτου με τον κώδικα Morse.

Αντιθέτως, αν η κωδικοποίηση γίνεται με μία απεικόνιση πολλών σε ένα, μπορεί να προκύψει η περίπτωση ένα κωδικό μήνυμα να αποδοθεί κατά την αποκωδικοποίηση σε μήνυμα που διαφέρει από το αρχικό. Έτσι, έχουμε απώλεια της πληροφορίας που περιέχεται στο αρχικό μήνυμα. Χαρακτηριστικό παράδειγμα ενός τέτοιου **κώδικα με απώλειες** είναι είναι η κωδικοποίηση μίας έγχρωμης εικόνας σε μία αντίστοιχη διαβαθμίσεων του γκρι.

Στα πλαίσια αυτού του κεφαλαίου, θα ασχοληθούμε με την κωδικοποίηση χωρίς απώλειες.

4.2.2 Ταξινόμηση με κριτήριο το μήκος των κωδικών λέξεων

Ανάλογα με το μήκος των κωδικών λέξεων οι κώδικες διακρίνονται σε **σταθερού μήκους** και **μεταβλητού μήκους**.

Στους **κώδικες σταθερού μήκους** το μήκος των κωδικών λέξεων είναι σταθερό για κάθε σύμβολο της πηγής. Ένας κώδικας σταθερού μήκους είναι ο κώδικας ASCII-7 όπου κάθε χαρακτήρας κωδικοποιείται με μια ακολουθία 7 δυαδικών ψηφίων.

Στην περίπτωση της συμπίεσης, το ενδιαφέρον εστιάζεται σε **κώδικες με μεταβλητό μήκος**. Τα σύμβολα της πηγής που έχουν μεγαλύτερη πιθανότητα εμφάνισης αντιστοιχίζονται σε μικρότερες κωδικές λέξεις και αντιστρόφως. Με αυτόν τον τρόπο το συνολικό μήκος του κωδικού μηνύματος μπορεί να προκύψει μικρότερο από το αρχικό μήνυμα. Ο κώδικας Morse είναι μεταβλητού μήκους. Βλέπουμε στον πίνακα 4.1 ότι το γράμμα E που είναι το πιο σύνηθες στην αγγλική γλώσσα αντιστοιχίζεται στη μικρότερη κωδική λέξη. Αντιθέτως το γράμμα Q που είναι πιο σπάνιο κωδικοποιείται με μεγαλύτερη λέξη.

[1] Φυσικά, υποθέτουμε ότι το πεδίο ορισμού της C' είναι το σύνολο U των κωδικών λέξεων, στις οποίες απεικονίζονται τα σύμβολα της πηγής ($U = C(X) \subseteq U^*$).

4.2.3 Ταξινόμηση με κριτήριο τη μεταβολή της αντιστοίχισης

Ανάλογα με το αν οι κανόνες αντιστοίχισης των συμβόλων της πηγής σε κωδικές λέξεις μεταβάλλονται ή όχι κατά τη διάρκεια της κωδικοποίησης, οι κώδικες διακρίνονται σε **δομικούς** και **συνελικτικούς**.

Σε ένα **δομικό κώδικα** η αντιστοιχία συμβόλων πηγής σε κωδικές λέξεις είναι προκαθορισμένη και αμετάβλητη καθόλη τη διάρκεια της κωδικοποίησης. Έτσι, κάθε σύμβολο του αλφαβήτου της πηγής κωδικοποιείται πάντα με την ίδια κωδική λέξη. Η υπόθεση στην οποία βασίζονται οι δομικοί κώδικες είναι ότι τα στατιστικά χαρακτηριστικά των συμβόλων που παράγει μία πηγή είναι σταθερά και δεδομένα σε όλο το μήκος του αρχικού μηνύματος. Έτσι, αν γνωρίζουμε τις πιθανότητες εμφάνισης των συμβόλων μπορούμε να σχεδιάσουμε ένα βέλτιστο κώδικα συμπίεσης βασιζόμενοι στις αρχή ότι τα πιθανότερα σύμβολα πρέπει να αντιστοιχούν σε μικρότερες κωδικές λέξεις. Ενδεικτικά αναφέρουμε τους δομικούς κώδικες Shannon, Shannon-Fano και Huffman.

Αντιθέτα, σε ένα **συνελικτικό κώδικα** η αντιστοιχία συμβόλων πηγής σε κωδικές λέξεις μεταβάλλεται με το χρόνο. Μία τέτοια προσέγγιση στηρίζεται στην υπόθεση ότι δε γνωρίζουμε εκ των προτέρων τις πιθανότητες εμφάνισης των συμβόλων ή ότι τα στατιστικά χαρακτηριστικά της πηγής μεταβάλλονται σημαντικά με το χρόνο. Έτσι, το μήκος των κωδικών λέξεων μεταβάλλεται διαρκώς λαμβάνοντας υπόψη τα σύμβολα πηγής που εμφανίστηκαν προηγουμένως ή που αναμένουμε να εμφανιστούν στο άμεσο μέλλον. Είναι φανερό ότι οι συνελικτικοί κώδικες είναι πιο πολύπλοκοι στην υλοποίησή τους. Όμως, είναι ιδιαίτερα διαδεδομένοι λόγω της σημαντικής συμπίεσης που επιτυγχάνουν. Ίσως ο πιο γνωστός συνελικτικός κώδικας είναι ο LZW που πήρε το όνομά του από τα αρχικά των ερευνητών Lempel-Ziv-Welch.

4.2.4 Ταξινόμηση με κριτήριο την αποκωδικοποίηση

Ένα βασικό χαρακτηριστικό ενός κώδικα συμπίεσης για να χαρακτηριστεί επιτυχής είναι το κατά πόσο είναι δυνατή και εύκολη η αποκωδικοποίηση ενός κωδικού μηνύματος από το δέκτη. Θεωρούμε βέβαια ότι κατά την αναπαραγωγή του αρχικού μηνύματος από το δέκτη ο κώδικας είναι γνωστός.

Το πρώτο χαρακτηριστικό ενός σωστά σχεδιασμένου κώδικα είναι κάθε σύμβολο πηγής να αντιστοιχεί σε διαφορετική κωδική λέξη. Ένας κώδικας που ικανοποιεί

αυτό το κριτήριο λέγεται **ευκρινής**. Αν ο κώδικας δεν είναι ευκρινής, τότε κατά την αποκωδικοποίηση μίας κωδικής λέξης θα υπάρχει ασάφεια ως προς το ποιο σύμβολο πηγής αντιπροσωπεύει, αφού θα υπάρχουν περισσότερα του ενός σύμβολο που αντιστοιχούν στην κωδική λέξη.

Ένα δεύτερο χαρακτηριστικό που θα πρέπει να έχει ένας κώδικας είναι η δυνατότητα να αντιληφθεί ο δέκτης με απόλυτη βεβαιότητα την αρχή και το τέλος μιας κωδικής λέξης, όταν λάβει ένα κωδικό μήνυμα με πλήθος λέξεων. Ένας κώδικας για τον οποίο κάθε κωδική λέξη αναγνωρίζεται με βεβαιότητα μέσα σε μία μακρά ακολουθία κωδικών συμβόλων λέγεται **μονοσήμαντος**.

Τέλος, μία χρήσιμη ιδιότητα που θα πρέπει να έχει ένας κώδικας είναι η δυνατότητα να αναγνωρίζονται κωδικές λέξεις μέσα σε ένα κωδικό μήνυμα, χωρίς να εξετάζονται τα κωδικά σύμβολα των γειτονικών κωδικών λέξεων. Ένας μονοσήμαντος κώδικας που επιτρέπει την αποκωδικοποίηση λέξη προς λέξη, χωρίς να απαιτείται η εξέταση επόμενων κωδικών συμβόλων λέγεται **στιγμιαία αποκωδικοποιήσιμος**.

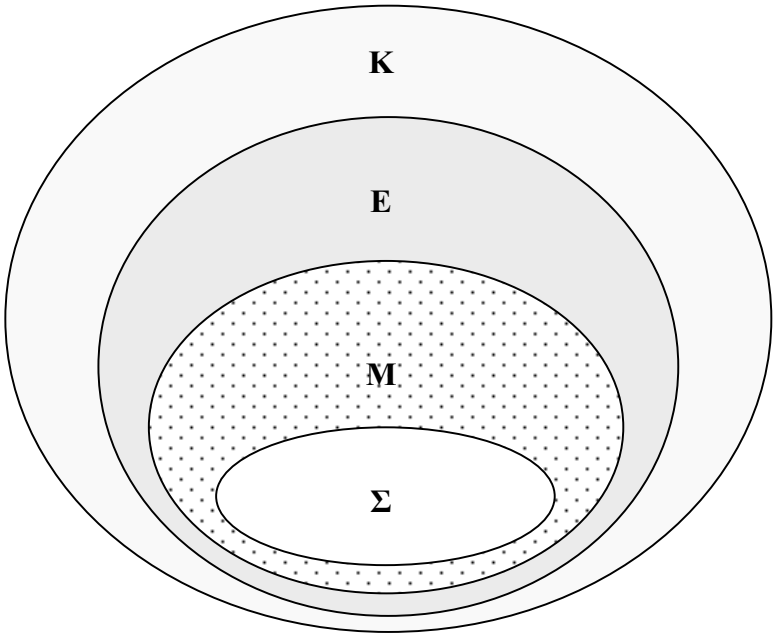
Στον πίνακα 4.2 παρουσιάζεται ένα παράδειγμα τεσσάρων κωδίκων. Ο κώδικας Α δεν είναι ευκρινής γιατί τα σύμβολα κ και $\neg$ αντιστοιχίζονται στην ίδια κωδική λέξη. Ο κώδικας Β αν και είναι ευκρινής δεν είναι μονοσήμαντος. Για παράδειγμα, αν θεωρήσουμε το κωδικό μήνυμα 0110, τότε αποκωδικοποιώντας σύμφωνα με τον κώδικα Β προκύπτει είτε το μήνυμα $\kappa\neg\kappa$, είτε το μήνυμα $\exists F$. Παρατηρώντας τις κωδικές λέξεις του κώδικα Γ μπορούμε εύκολα να αντιληφθούμε ότι είναι ευκρινής και μονοσήμαντος. Όμως ο κώδικας Γ δεν είναι στιγμιαία αποκωδικοποιήσιμος γιατί πρέπει να λάβουμε το πρώτο σύμβολο (1) της επόμενης κωδικής λέξης για να συμπεράνουμε που τελείωσε η προηγούμενη. Για παράδειγμα, αν λάβουμε το κωδικό μήνυμα 10100 (δηλαδή $\exists F$) για να αντιληφθούμε ότι το πρώτο σύμβολο πηγής είναι το $\exists$ θα πρέπει να ελέγξουμε το τρίτο κωδικό σύμβολο το μηνύματος που ανήκει στην κωδική λέξη του F. Τέλος, ο κώδικας Δ είναι ευκρινής, μονοσήμαντος και στιγμιαία αποκωδικοποιήσιμος. Συγκεκριμένα, είναι στιγμιαία αποκωδικοποιήσιμος γιατί το τελευταίο σύμβολο κάθε κωδικής λέξης (1) δηλώνει και το τέλος της λέξης. Για παράδειγμα, αν λάβουμε το κωδικό μήνυμα 0010001 συμπεραίνουμε αμέσως ότι το πρώτο σύμβολο πηγής είναι F εξετάζοντας μόνο τα τρία πρώτα κωδικά σύμβολα που ούτως ή άλλως αντιστοιχούν στο F.

Αξίζει να σημειωθεί (σχήμα 4.1) ότι:

1. Κάθε στιγμιαία αποκωδικοποιήσιμος κώδικας είναι και μονοσήμαντος. Δεν ισχύει το αντίστροφο.

Πίνακας 4.2: *Παράδειγμα τεσσάρων δυαδικών κωδίκων.*

Σύμβολο	Κώδικας Α	Κώδικας Β	Κώδικας Γ	Κώδικας Δ
ℵ	00	0	1	1
ℶ	10	01	10	01
ℱ	01	10	100	001
ℸ	00	11	1000	0001



Σχήμα 4.1: *Ταξινόμηση των κωδίκων με κριτήριο την αποκωδικοποίηση. (Κ) Όλοι οι κώδικες. (Ε) Ευκρινείς. (Μ) Μονοσήμαντοι. (Σ) Στιγμιαία αποκωδικοποιήσιμοι.*

2. Κάθε μονοσήμαντος κώδικας είναι και ευκρινής. Δεν ισχύει το αντίστροφο.

3. Κάθε ευκρινής κώδικας σταθερού μήκους είναι και στιγμιαία αποκωδικοποιήσιμος.

4.3 Προθεματική Ιδιότητα

Από την ταξινόμηση με βάση την ευκολία με την οποία επιτυγχάνεται η αποκωδικοποίηση προέκυψε ότι ένας στιγμιαία αποκωδικοποιήσιμος κώδικας παρουσιάζει σημαντικό ενδιαφέρον αφού η αποκωδικοποίηση είναι στιγμιαία χωρίς να δημιουργούνται κωδικά μηνύματα με ασάφειες ή διαφορετικά ισοδύναμα αποτελέσματα αποκωδικοποίησης. Για τους παραπάνω λόγους, στη συνέχεια θα μας απασχολήσουν μόνο

στιγμιαία αποκωδικοποιήσιμοι κώδικες.

Από την ανάλυση των κωδικών λέξεων του στιγμιαία αποκωδικοποιήσιμου κώδικα Δ (πίνακας 4.1) προκύπτει ότι καμία κωδική λέξη δεν είναι πρόθεμα κάποιας άλλης. Η ιδιότητα αυτή είναι χαρακτηριστική για ένα στιγμιαία αποκωδικοποιήσιμο κώδικα και λέγεται **προθεματική ιδιότητα**. Μάλιστα, ισχύει το παρακάτω θεώρημα.

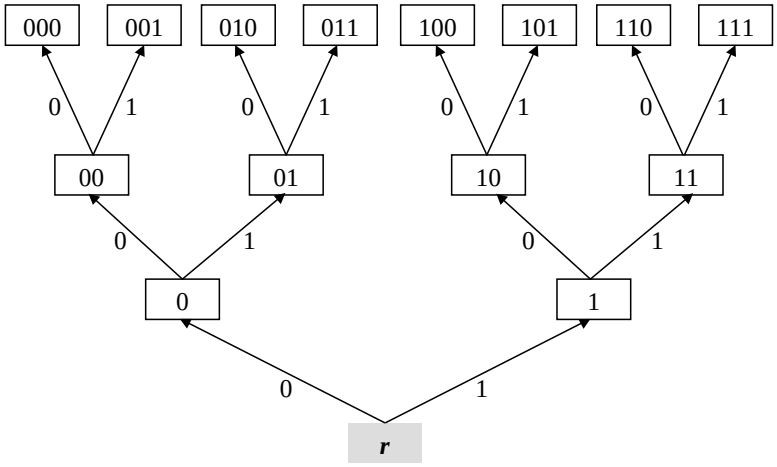
Θεώρημα 4.1.
Ένας κώδικας είναι στιγμιαία αποκωδικοποιήσιμος, αν και μόνο αν οι κωδικές λέξεις του έχουν την προθεματική ιδιότητα.

Απόδειξη.
Αρχικά θα αποδείξουμε το ευθύ του θεωρήματος με την εις άτοπο απαγωγή. Αν οι κωδικές λέξεις δεν έχουν την προθεματική ιδιότητα, τότε υπάρχει τουλάχιστον μία κωδική λέξη $u_i \in U$, μήκους l_i, η οποία είναι πρόθεμα μίας άλλης κωδικής λέξης, έστω της u_j με μήκος $l_j > l_i$. Η κωδική λέξη u_i αντιστοιχεί στο σύμβολο πηγής x_i, ενώ η u_j στο x_j. Θεωρούμε ένα κωδικό μήνυμα $\mathbf{t} = u_i \ldots$ που τα πρώτα του κωδικά σύμβολα, πλήθους l_i, ταυτίζονται με την κωδική λέξη u_i. Κατά την αποκωδικοποίηση του μηνύματος $\mathbf{t}$ μόλις λάβουμε τα πρώτα l_i σύμβολα δε θα γνωρίζουμε με βεβαιότητα αν το σύμβολο της πηγής είναι το u_i ή το u_j, συνεπώς ο κώδικας δεν είναι στιγμιαία αποκωδικοποιήσιμος. Άρα καταλήγουμε στο συμπέρασμα ότι ένας στιγμιαία αποκωδικοποιήσιμος κώδικας έχει την προθεματική ιδιότητα.

Στη συνέχεια αποδεικνύουμε το αντίστροφο του θεωρήματος. Αν ο κώδικας έχει την προθεματική ιδιότητα και το κωδικό μήνυμα έχει τη μορφή $\mathbf{t} = u_i \ldots$ τότε λαμβάνοντας τα πρώτα l_i κωδικά σύμβολα μπορούμε αμέσως να συμπεράνουμε ότι το αντίστοιχο σύμβολο πηγής είναι το x_i, δεδομένου ότι κανένα άλλο σύμβολο x_j δεν έχει πρόθεμα την ακολουθία κωδικών συμβόλων u_i. Συνεχίζοντας με τον ίδιο τρόπο μπορούμε να αποκωδικοποιήσουμε ακολουθιακά τις κωδικές λέξεις του μηνύματος $\mathbf{t}$ καθώς τις λαμβάνουμε. Συνεπώς ο κώδικας είναι στιγμιαία αποκωδικοποιήσιμος.

♦

Το παραπάνω θεώρημα είναι σημαντικό διότι επιτρέπει την αναγνώριση ενός κώδικα ως στιγμιαία αποκωδικοποιήσιμου με απλό έλεγχο της προθεματικής ιδιότητας των κωδικών λέξεων.



Σχήμα 4.2: *Δενδροδιάγραμμα του συνόλου U^* ενός δυαδικού κώδικα με μέγιστο μήκος λέξης $L = 3$.*

4.4 Δενδροδιάγραμμα Απόφασης

Ας υποθέσουμε ότι έχουμε ένα M-αδικό κώδικα με αλφάβητο κωδικοποίησης $W = \{w_1, w_2, \ldots, w_M\}$, που χρησιμοποιείται για την κωδικοποίηση του αλφαβήτου της πηγής $X = \{x_1, x_2, \ldots, x_N\}$. Θεωρούμε ότι το μέγιστο μήκος που έχει μία κωδική λέξη αυτού του κώδικα είναι L. Τότε οι κωδικές λέξεις u_i $(i = 1, 2, \ldots, N)$ θα ανήκουν σε ένα υποσύνολο U του συνόλου U^*, το οποίο προκύπτει από την ένωση όλων των καρτεσιανών γινομένων μέχρι δύναμης L, δηλαδή

$$u_i \in U \subseteq U^* = \bigcup_{l=1}^{l=L} W^l. \tag{4.3}$$

Το σύνολο U^* μπορεί να παρασταθεί με τη χρήση ενός γράφου που έχει τη μορφή δέντρου, όπου κάθε κόμβος του δέντρου αντιστοιχεί σε ένα στοιχείο του συνόλου U^*. Ο γράφος αυτός λέγεται **δενδροδιάγραμμα**. Για παράδειγμα, το σύνολο U^* ενός δυαδικού κώδικα με μέγιστο μήκος $L = 3$ αναπαριστάται με το δέντρο του σχήματος 4.2.

Στο σχήμα 4.2 με τον κόμβο r συμβολίζουμε τη ρίζα του δένδρου ενώ κάθε άλλος κόμβος αναπαριστά ένα στοιχείο του συνόλου U^*. Το ύψος του δένδρου είναι ίσο με το μέγιστο μήκος των κωδικών λέξεων. Παρατηρούμε ότι οι ακμές (κλάδοι) που συνδέουν τους κόμβους είναι προσανατολισμένες. Ξεκινούν από ένα κόμβο και καταλήγουν σε έναν άλλο του αμέσως ψηλότερου επιπέδου. Σε κάθε κόμβο καταλήγει μία μόνο ακμή και από κάθε κόμβο ξεκινούν το πολύ M ακμές, όσες και το πλήθος των κωδικών συμβόλων. Οι προσανατολισμένες ακμές αντιστοιχούν στα κωδικά σύμβολα που χρησιμοποιούνται από τον κώδικα. Οι κόμβοι από τους οποίους δεν ξε-

κινά καμία ακμή λέγονται φύλλα. Στο δέντρο του σχήματος 4.2 υπάρχουν 8 φύλλα, όλα στο ψηλότερο επίπεδο του δέντρου. Τέλος, κάθε κωδική λέξη που αντιστοιχεί σε ένα κόμβο προκύπτει από τη λέξη που αντιστοιχεί στον κόμβο σύνδεσης του αμέσως χαμηλότερου επιπέδου προσθέτοντας ένα κωδικό σύμβολο στο τέλος της. Για παράδειγμα, η κωδική λέξη 010 που βρίσκεται στο τρίτο επίπεδο προκύπτει από την κωδική λέξη 01 του προηγούμενου επιπέδου, αν προσθέσουμε στο τέλος της το κωδικό σύμβολο 0.

Αποδεικνύεται ότι ένας στιγμιαία αποκωδικοποιήσιμος κώδικας μπορεί να αναπαρασταθεί με ένα δενδροδιάγραμμα του οποίου τα φύλλα, και μόνον αυτά, αντιστοιχούν στις κωδικές λέξεις του κώδικα. Αυτό οφείλεται στο γεγονός ότι τα φύλλα του δενδροδιαγράμματος έχουν την προθεματική ιδιότητα, η οποία με τη σειρά της είναι η ικανή και αναγκαία συνθήκη για να είναι ο κώδικας στιγμιαία αποκωδικοποιήσιμος. Συνεπώς, ισχύει το θεώρημα

Θεώρημα 4.2.
Κάθε στιγμιαία αποκωδικοποιήσιμος κώδικας έχει δενδροδιάγραμμα.

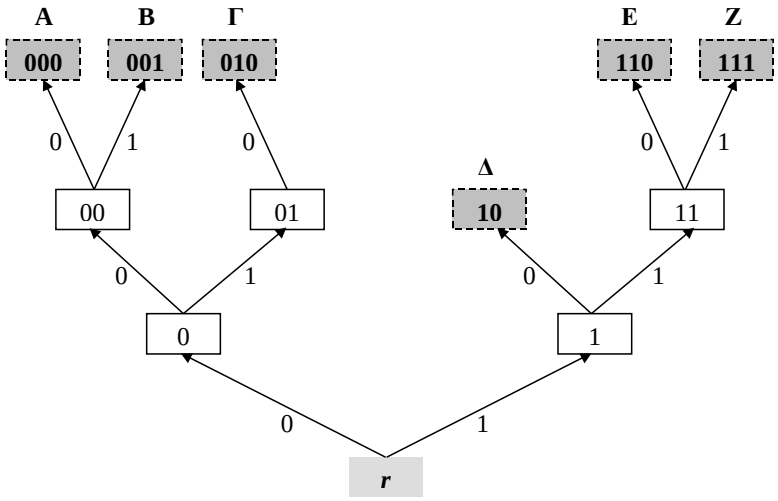
Αξίζει να σημειωθεί ότι τα φύλλα ενός δενδροδιαγράμματος δε βρίσκονται κατ' ανάγκη στο υψηλότερο επίπεδο του δένδρου. Μπορεί να βρίσκονται και σε χαμηλότερο επίπεδο αρκεί να μη ξεκινά καμία προσανατολισμένη ακμή από τον αντίστοιχο κόμβο. Έστω για παράδειγμα ο δυαδικός κώδικας του πίνακα 4.3

Πίνακας 4.3: *Ένας δυαδικός στιγμιαία αποκωδικοποιήσιμος κώδικας.*

Σύμβολο	Κωδική Λέξη
Α	000
Β	001
Γ	010
Δ	10
Ε	110
Ζ	111

Παρατηρούμε ότι οι κωδικές λέξεις του κώδικα έχουν την προθεματική ιδιότητα. Συνεπώς, σύμφωνα με το θεώρημα 4.1, ο κώδικας είναι στιγμιαία αποκωδικοποιήσιμος. Το δενδροδιάγραμμα του κώδικα αυτού φαίνεται στο σχήμα 4.3, όπου τα φύλλα του δενδροδιαγράμματος (κόμβοι με διακεκομμένο πλαίσιο και γκρι φόντο) αντιστοιχούν τις κωδικές λέξεις του κώδικα.

Η μεγάλη σημασία του δενδροδιαγράμματος ενός στιγμιαία αποκωδικοποιήσιμου κώδικα πηγάζει από το γεγονός ότι το δενδροδιάγραμμα αποτελεί βασικό εργαλείο για



Σχήμα 4.3: *Δενδροδιάγραμμα ενός δυαδικού στιγμιαία αποκωδικοποιήσιμου κώδικα (Πίνακας 4.3).*

τη γρήγορη αποκωδικοποίηση ενός κωδικού μηνύματος. Συγκεκριμένα, κατά τη διαδικασία της αποκωδικοποίησης στο δέκτη, ακολουθούμε τα παρακάτω βήματα:

1. Αρχικά, θεωρούμε ότι βρισκόμαστε στη ρίζα του δενδροδιαγράμματος.

2. Με τη λήψη ενός κωδικού συμβόλου μεταβαίνουμε στον κόμβο του αμέσως ψηλότερου επιπέδου ακολουθώντας την προσανατολισμένη ακμή που αντιστοιχεί στο ληφθέν σύμβολο.

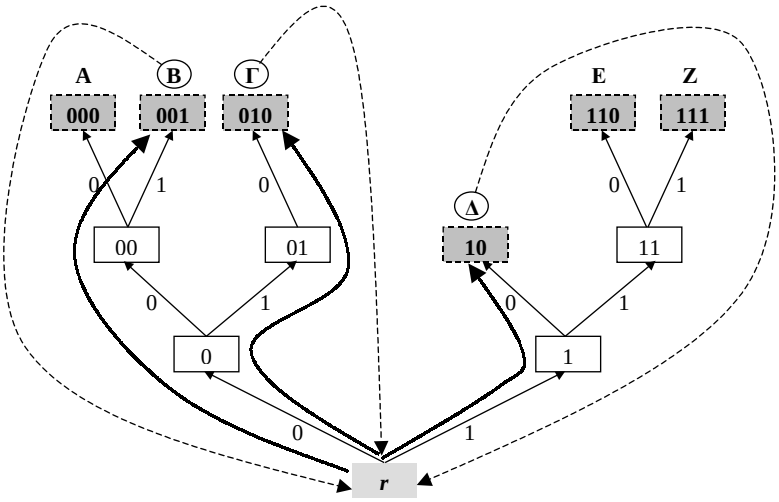
3. Το δεύτερο βήμα επαναλαμβάνεται ακολουθιακά μέχρι να φτάσουμε σε ένα φύλλο του δέντρου. Μόλις συμβεί αυτό, καταγράφουμε το σύμβολο πηγής που αντιστοιχεί στο φύλλο.

4. Στη συνέχεια, μεταβαίνουμε στη ρίζα του δέντρου και επαναλαμβάνουμε τα βήματα από το δεύτερο και μετά.

5. Η διαδικασία της αποκωδικοποίησης περατώνεται με την αποκωδικοποίηση και του τελευταίου συμβόλου.

Αν υποθέσουμε ότι λαμβάνουμε το κωδικό μήνυμα 00110010, το οποίο προέκυψε με χρήση του κώδικα του πίνακα 4.2. Τότε ακολουθώντας τα βήματα της αποκωδικοποίησης πάνω στο δενδροδιάγραμμα απόφασης όπως αυτά φαίνονται στο σχήμα 4.4 καταλήγουμε στο αποκωδικοποιημένο μήνυμα ΒΔΓ.

Επειδή το δενδροδιάγραμμα μας επιτρέπει να αποφασίζουμε για την αντιστοίχιση κωδικών λέξεων σε σύμβολα πηγής λέγεται και **δενδροδιάγραμμα απόφασης**.



Σχήμα 4.4: *Διαδικασία αποκωδικοποίησης με χρήση του δενδροδιαγράμματος απόφασης.*

4.5 Η Ταυτοανισότητα του Kraft

Όπως αναφέραμε προηγουμένως, βασικός μας στόχος κατά την ανάπτυξη ενός στιγμιαία αποκωδικοποιήσιμου κώδικα συμπίεσης, είναι να βρεθούν τα κατάλληλα μήκη των κωδικών λέξεων, ώστε το μέσο μήκος του κωδικού μηνύματος να είναι το μικρότερο δυνατό.

Θεωρούμε ένα στιγμιαία αποκωδικοποιήσιμο κώδικα με αλφάβητο κωδικοποίησης M κωδικών συμβόλων (M-αδικός κώδικας). Ο κώδικας έχει N διαφορετικές κωδικές λέξεις (όσες και τα σύμβολα της πηγής). Αν το μήκος της μεγαλύτερης κωδικής λέξης του κώδικα είναι L, οπότε το ύψος του δενδροδιαγράμματος απόφασης είναι επίσης L, τότε ισχύει το θεώρημα.

Θεώρημα 4.3.
Τα μήκη l_i $(i = 1, 2, \ldots, N)$ των κωδικών λέξεων, πλήθους N, ενός M-αδικού στιγμιαία αποκωδικοποιήσιμου κώδικα ικανοποιούν την ταυτοανισότητα

$$M^L \geq \sum_{i=1}^{N} M^{L-l_i} \tag{4.4}$$

ή ισοδύναμα

$$\sum_{i=1}^{N} M^{-l_i} \leq 1. \tag{4.5}$$

Η ταυτοανισότητα (4.5) λέγεται **ταυτοανισότητα του Kraft**.

Αντιστρόφως, ισχύει και το παρακάτω θεώρημα.

Θεώρημα 4.4.
Αν τα μήκη l_i $(i = 1, 2, \ldots, N)$ ικανοποιούν την ταυτοανισότητα

$$\sum_{i=1}^{N} M^{-l_i} \leq 1, \tag{4.6}$$

τότε υπάρχει M-αδικός κώδικας με N κωδικές λέξεις μήκους $l_1, l_2, \ldots, l_N$, ο οποίος είναι στιγμιαία αποκωδικοποιήσιμος.

Τα δύο παραπάνω θεωρήματα συνθέτουν το παρακάτω θεώρημα ύπαρξης.

Θεώρημα 4.5 (Ύπαρξη στιγμιαίας αποκωδικοποιήσιμου κώδικα).
Υπάρχει στιγμιαία αποκωδικοποιήσιμος M-αδικός κώδικας με μήκη κωδικών λέξεων $l_1, l_2, \ldots, l_N$, αν και μόνο αν ισχύει η ταυτοανισότητα

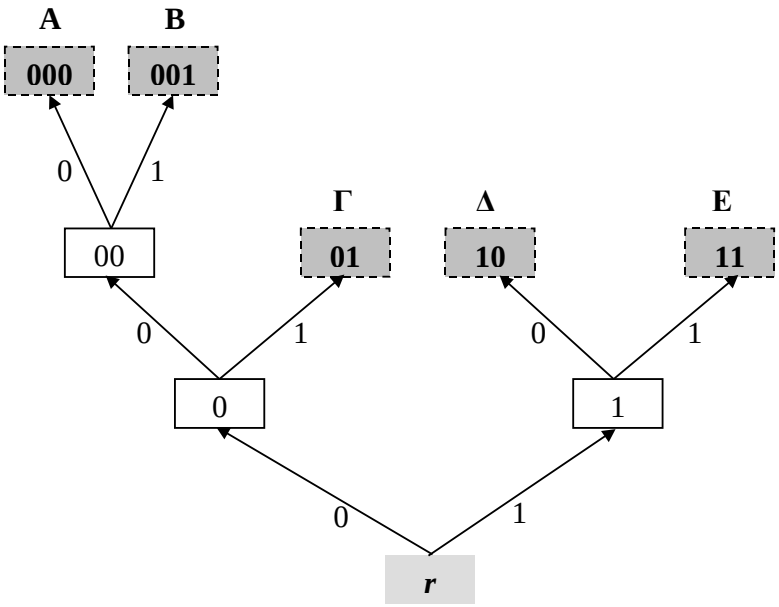
$$\sum_{i=1}^{N} M^{-l_i} \leq 1. \tag{4.7}$$

Συνεπώς, αν τα μήκη κωδικών λέξεων ικανοποιούν την ταυτοανισότητα του Kraft, τότε εξασφαλίζεται η ύπαρξη στιγμιαία αποκωδικοποιήσιμου κώδικα με τα συγκεκριμένα μήκη κωδικών λέξεων. Αυτό όμως, δε σημαίνει ότι κάθε κώδικας που ικανοποιεί την (4.6) θα είναι και στιγμιαία αποκωδικοποιήσιμος. Μάλιστα, όπως θα δούμε στη συνέχεια, η ταυτοανισότητα του Kraft ικανοποιείται και από τα μήκη των κωδικών λέξεων ενός μονοσήμαντου κώδικα, ο οποίος δεν είναι απαραίτητα στιγμιαία αποκωδικοποιήσιμος.

Αν στο δενδροδιάγραμμα απόφασης από κάθε κόμβο που δεν είναι φύλλο ξεκινούν ακριβώς M προσανατολισμένες ακμές (όσα και τα κωδικά σύμβολα), τότε η (4.6) ισχύει ως ισότητα (σχήμα 4.5). Αντίθετα, αν υπάρχει έστω ένας κόμβος που δεν είναι φύλλο και ξεκινούν από αυτόν λιγότερες από M ακμές, τότε στην (4.6) ισχύει η ανισότητα (σχήμα 4.6). Αυτό συμβαίνει όταν μία κωδική λέξη έχει πλεονάζοντα κωδικά σύμβολα, ενώ θα μπορούσε να έχει λιγότερα. Άρση αυτού του πλεονασμού γίνεται διαγράφοντας το τελευταίο κωδικό σύμβολο από την αντίστοιχη κωδική λέξη. Μπορεί μάλιστα να απαιτείται η διαγραφή περισσοτέρων του ενός κωδικών συμβόλων.

4.6 Η Ταυτοανισότητα του McMillan

Κατά την ταξινόμηση των κωδίκων με κριτήριο την αποκωδικοποίηση διαπιστώθηκε ότι το σύνολο των μονοσήμαντων κωδίκων είναι υπερσύνολο του συνόλου των



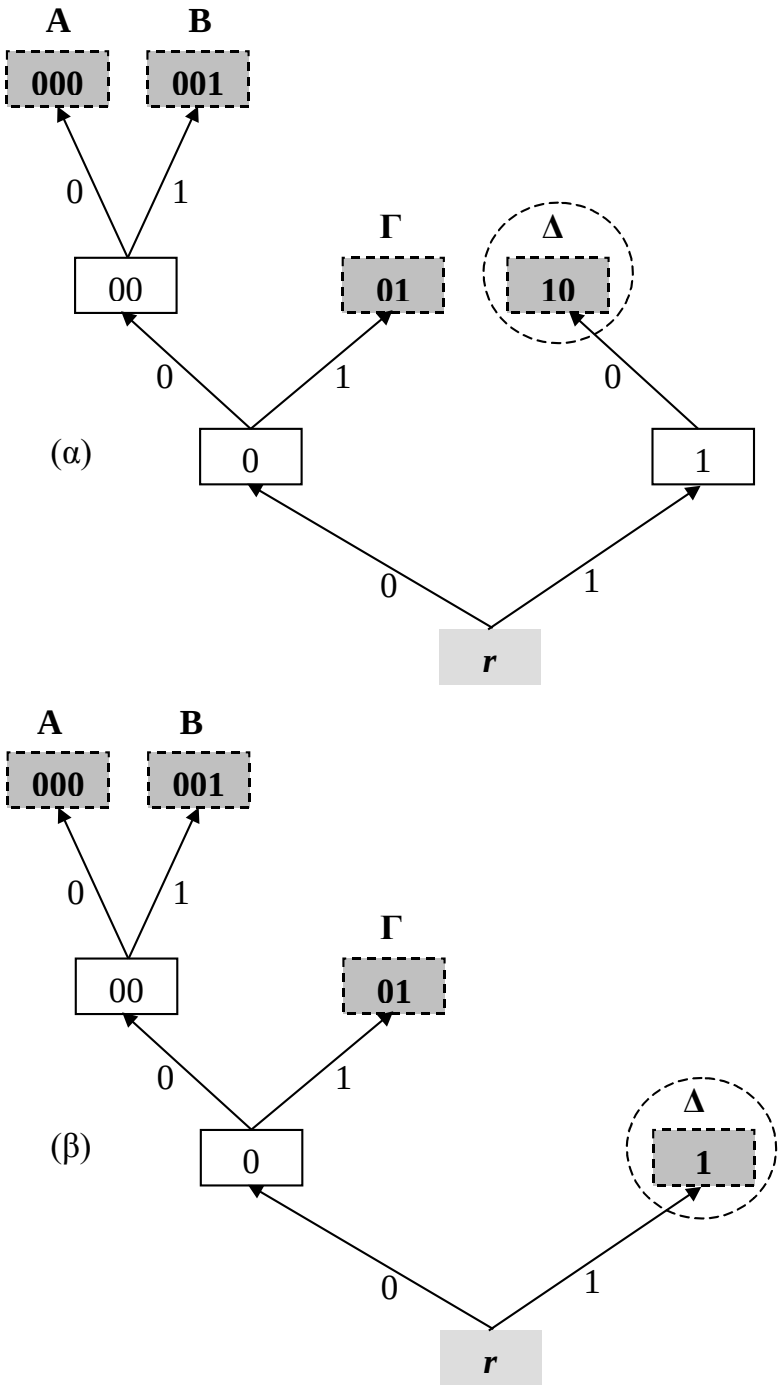
Σχήμα 4.5: *Δενδροδιάγραμμα στιγμιαία αποκωδικοποιήσιμου κώδικα που ικανοποιεί την ισότητα στην ταυτοανισότητα Kraft.*

στιγμιαία αποκωδικοποιήσιμων. Από αυτήν την παρατήρηση θα περίμενε κανείς ότι η ικανή και αναγκαία συνθήκη για την ύπαρξη μονοσήμαντου κώδικα με δεδομένα μήκη κωδικών λέξεων θα είναι λιγότερο περιοριστική από την ταυτοανισότητα του Kraft. Παραδόξως, τόσο για τους στιγμιαία αποκωδικοποιήσιμους όσο και για του μονοσήμαντους κώδικες ισχύει η ίδια συνθήκη ύπαρξης.

Θεώρημα 4.6 (Ύπαρξη μονοσήμαντου κώδικα).
Υπάρχει μονοσήμαντος M-αδικός κώδικας με μήκη κωδικών λέξεων $l_1, l_2, \ldots, l_N$, αν και μόνο αν ισχύει η ταυτοανισότητα

$$\sum_{i=1}^{N} M^{-l_i} \leq 1. \tag{4.8}$$

Αν και η ταυτοανισότητα (4.8) είναι ίδια με την ταυτοανισότητα του Kraft, όταν αναφέρεται σε μονοσήμαντους κώδικες έχει επικρατήσει να λέγεται **ταυτοανισότητα του McMillan**.



Σχήμα 4.6: *Δενδροδιάγραμμα στιγμιαία αποκωδικοποιήσιμου κώδικα που ικανοποιεί (α) την ανισότητα στην ταυτοανισότητα Kraft λόγω πλεονασμού, και (β) την ισότητα μετά από άρση του πλεονασμού.*

4.7 Σχολιασμός των Ταυτοανισοτήτων Kraft και McMillan

- Τα θεωρήματα 4.5 και 4.6 **δε λένε** ότι: Ένας M-αδικός κώδικας με μήκη κωδικών λέξεων $l_1, l_2, \ldots, l_N$ είναι στιγμιαία αποκωδικοποιήσιμος ή μονοσήμαντος, αν και μόνο αν ικανοποιείται η ταυτοανισότητα Kraft-McMillan. Για παράδειγμα, ο δυαδικός κώδικας με κωδικές λέξεις $\{0, 01, 001\}$ και αντίστοιχα μήκη λέξεων $\{1, 2, 3\}$ ικανοποιεί την ταυτοανισότητα ($\sum 2^{-l_i} = 7/8 \leq 1$). Όμως ο κώδικας αυτός δεν είναι ούτε μονοσήμαντος, ούτε στιγμιαία αποκωδικοποιήσιμος.

- Αν ξέρουμε ότι ένας M-αδικός με δεδομένα μήκη κωδικών λέξεων είναι μονοσήμαντος -οπότε ικανοποιείται η ταυτοανισότητα McMillan- τότε θα υπάρχει M-αδικός στιγμιαία αποκωδικοποιήσιμος κώδικας με τα ίδια μήκη κωδικών λέξεων.

- Ένας μονοσήμαντος κώδικας δεν παρουσιάζει τεχνικό ενδιαφέρον διότι θα υπάρχει στιγμιαία αποκωδικοποιήσιμος με τα ίδια μήκη κωδικών λέξεων, ο οποίος άλλωστε αποκωδικοποιείται με μεγαλύτερη ευκολία.

4.8 Μέσο Μήκος Κώδικα

Θεωρούμε μια πηγή πληροφορίας με αλφάβητο $X = \{x_1, x_2, \ldots, x_N\}$ και αντίστοιχη κατανομή πιθανοτήτων των συμβόλων $P_X = \{p_1, p_2, \ldots, p_N\}$. Για την κωδικοποίηση των συμβόλων της πηγής υιοθετείται ένας M-αδικός στιγμιαία αποκωδικοποιήσιμος κώδικας με αλφάβητο $W = \{w_1, w_2, \ldots, w_M\}$. Κάθε σύμβολο της πηγής x_i αντιστοιχίζεται σε μία κωδική λέξη u_i με μήκος l_i.

Αν υποθέσουμε ότι όλες οι κωδικές λέξεις έχουν όλες το ίδιο μήκος τότε για να μπορέσουν να αναπαρασταθούν τα N σύμβολα της πηγής θα πρέπει το ελάχιστο μήκος L των κωδικών λέξεων να είναι ίσο με

$$L = \lceil \log_M(N) \rceil, \tag{4.9}$$

δηλαδή ίσο με το μικρότερο ακέραιο που είναι μεγαλύτερος από $\log_M(N)$. Αυτός ο κώδικας, σταθερού μήκους, δε λαμβάνει υπόψη τις πιθανότητες εμφάνισης των συμβόλων της πηγής. Έτσι ένα σύμβολο με μικρή πιθανότητα εμφάνισης κωδικοποιείται όπως ένα άλλο με μεγάλη πιθανότητα. Συνεπώς, ο κώδικας σταθερού μήκους δημιουργεί μεγάλα κωδικά μηνύματα.

Για να μπορέσουμε να συγκρίνουμε τους στιγμιαία αποκωδικοποιήσιμους κώδικες με βάση το μήκος του κωδικού μηνύματος που παράγουν ορίζουμε το **μέσο μήκος του κώδικα** που δίνεται από τη σχέση:

$$\bar{L} = \sum_{i=1}^{N} p_i \cdot l_i. \tag{4.10}$$

Το μέσο μήκος ενός κώδικα εκφράζει το μέσο αριθμό κωδικών συμβόλων ανά σύμβολο πηγής.

Με βάση τη σχέση (4.10) προκύπτει ότι το μέσο μήκος ενός κώδικα σταθερού μήκους θα είναι

$$\bar{L} = L = \lceil \log_M(N) \rceil. \tag{4.11}$$

Εξετάζοντας τη σχέση (4.10) προκύπτει το λογικό συμπέρασμα ότι το μέσο μήκος μπορεί να μειωθεί, αν επιλεγούν μικρές κωδικές λέξεις για σύμβολα που έχουν μεγάλη πιθανότητα.

4.9 Βέλτιστος Κώδικας

Ο στιγμιαία αποκωδικοποιήσιμος κώδικας με το μικρότερο μέσο μήκος λέγεται **βέλτιστος**. Προφανώς, η κατασκευή του βέλτιστου κώδικα απαιτεί τον προσδιορισμό των μηκών των κωδικών λέξεων που ελαχιστοποιούν το μέσο μήκος. Το μέσο μήκος του βέλτιστου κώδικα συμβολίζεται $\bar{L}^*$ και δίνεται από τη σχέση

$$\bar{L}^* = \min_{l_1,l_2,\ldots,l_N} \{\bar{L}\} = \min_{l_1,l_2,\ldots,l_N} \{\sum_{i=1}^{N} p_i \cdot l_i\}. \tag{4.12}$$

Για τον υπολογισμό των μηκών των κωδικών λέξεων του βέλτιστου κώδικα λαμβάνεται υπόψη ότι τα μήκη θα πρέπει να ικανοποιούν την ταυτοανισότητα του Kraft, ώστε να εξασφαλίζεται η ύπαρξη στιγμιαία αποκωδικοποιήσιμου κώδικα. Η επίλυση αυτού του προβλήματος βελτιστοποίησης με περιορισμούς ανισότητας μπορεί να επιτευχθεί με την εφαρμογή των πολλαπλασιαστών Lagrange.

Αποδεικνύεται (βλέπε Παράρτημα Γ′) ότι τα μήκη των κωδικών λέξεων του βέλτιστου στιγμιαία αποκωδικοποιήσιμου κώδικα είναι ίσα με:

$$l_i^* = -\log_M(p_i), \qquad 1 \leq i \leq N. \tag{4.13}$$

Επιλέγοντας τα μήκη των κωδικών λέξεων με βάση τις βέλτιστες τιμές τους (4.13), το μέσο μήκος του βέλτιστου κώδικα θα είναι ίσο με

$$\bar{L}^* = \sum_{i=1}^{N} p_i \cdot l_i^* = -\sum_{i=1}^{N} p_i \cdot \log_M(p_i) = \tag{4.14}$$

$$= -\sum_{i=1}^{N} p_i \cdot \log(p_i)/\log(M) = H(X)/\log(M).$$

Ουσιαστικά, ο όρος $H(X)/\log(M)$ είναι η εντροπία της πηγής (X, P_X) που υπολογίζεται με βάση λογαρίθμου το M, και η οποία ονομάζεται M-αδική εντροπία. Θυμίζουμε ότι $H(X)$ είναι η εντροπία υπολογισμένη με βάση λογαρίθμου το 2. Συνεπώς, καταλήγουμε στο συμπέρασμα ότι το βέλτιστο μέσο μήκος ενός M-αδικού στιγμιαία αποκωδικοποιήσιμου κώδικα είναι ίσο με την M-αδική εντροπία της πηγής.

Προφανώς, τα βέλτιστα μήκη των κωδικών λέξεων που δίνονται από την (4.13) είναι γενικά θετικοί πραγματικοί αριθμοί. Συνεπώς, δεν είναι πρακτικά δυνατό να επιλεγούν κωδικές λέξεις με τέτοια μήκη, αφού αυτά θα πρέπει να είναι φυσικοί αριθμοί. Έτσι, τα μήκη των κωδικών λέξεων επιλέγονται ως οι μικρότεροι φυσικοί αριθμοί που είναι μεγαλύτεροι από τις τιμές που προκύπτουν από την (4.13), δηλαδή

$$l_i = \lceil l_i^* \rceil = \lceil -\log_M(p_i) \rceil, \qquad 1 \le i \le N. \tag{4.15}$$

Τα μήκη l_i $(i = 1, 2, \ldots, N)$ είναι μεγαλύτερα από τα αντίστοιχα βέλτιστα, συνεπώς θα ικανοποιούν την ταυτοανισότητα του Kraft.

Απόδειξη.

Δεδομένου ότι τα βέλτιστα μήκη ικανοποιούν την ταυτοανισότητα του Kraft έχουμε

$$\sum_{i=1}^{N} M^{-l_i^*} \le 1.$$

Ισχύει

$$l_i^* \le l_i \Rightarrow M^{-l_i^*} \ge M^{-l_i}.$$

Συνεπώς

$$\sum_{i=1}^{N} M^{-l_i} \le \sum_{i=1}^{N} M^{-l_i^*} \le 1.$$

♦

Αποδεικνύεται ότι το μέσο μήκος του στιγμιαία αποκωδικοποιήσιμου κώδικα, με μήκη κωδικών λέξεων που δίνονται από τη σχέση (4.15), είναι κάτω και άνω φραγμένο.

Συγκεκριμένα, ισχύει η διπλή ταυτοανισότητα:

$$\frac{H(X)}{\log(M)} \leq \bar{L} \leq \frac{H(X)}{\log(M)} + 1 \tag{4.16}$$

Απόδειξη.
Για τα μήκη των κωδικών λέξεων που δίνονται από την (4.15) έχουμε

$$l_i = \lceil -\log_M(p_i) \rceil \Rightarrow -\log_M(p_i) \leq l_i \leq -\log_M(p_i) + 1, \quad 1 \leq i \leq N.$$

Άρα έχουμε

$$\begin{aligned} -p_i \log_M(p_i) \leq p_i l_i \leq -p_i \log_M(p_i) + p_i \Rightarrow \\ -\sum_{i=1}^{N} p_i \log_M(p_i) \leq \sum_{i=1}^{N} p_i l_i \leq -\sum_{i=1}^{N} p_i \log_M(p_i) + \sum_{i=1}^{N} p_i \Rightarrow \\ \frac{H(X)}{\log(M)} \leq \bar{L} \leq \frac{H(X)}{\log(M)} + 1. \end{aligned}$$

♦

Τέλος, θα πρέπει να σημειώσουμε το εξής: Η ανάλυση που προηγήθηκε για τη μελέτη ενός βέλτιστου κώδικα βασίστηκε στην ικανοποίηση της ταυτοανισότητας του Kraft. Η ανισότητα αυτή ικανοποιείται όχι μόνο από στιγμιαία αποκωδικοποιήσιμο κώδικα αλλά και από μονοσήμαντο. Συνεπώς τα συμπεράσματα στα οποία καταλήξαμε ισχύουν και για μονοσήμαντους κώδικες.

4.10 Το Πρώτο Θεώρημα του Shannon

Από την προηγούμενη ενότητα προκύπτει το συμπέρασμα ότι το μέσο μήκος του στιγμιαία αποκωδικοποιήσιμου ή μονοσήμαντου κώδικα είναι φραγμένο κάτω και μπορεί να φραχθεί και άνω. Μάλιστα, το κάτω φράγμα είναι ίσο με την εντροπία της πηγής πληροφορίας και είναι ανεξάρτητο του κώδικα. Συνεπώς, δε μπορούμε να σχεδιάσουμε ένα στιγμιαία αποκωδικοποιήσιμο (μονοσήμαντο) κώδικα με όσο μικρό μέσο μήκος θέλουμε. Όλες μας οι προσπάθειες θα σταματούν στο όριο της εντροπίας της πηγής. Γεννάται λοιπόν το εύλογο ερώτημα κατά πόσο και με ποιο τρόπο μπορούμε να φτάσουμε το μέσο μήκος κοντά στην οριακή του τιμή της εντροπίας. Αυτή η προσπάθεια ισοδυναμεί με τη μείωση του άνω φράγματος $(H(X)/\log(M) + 1)$ του μέσου μήκους, φέρνοντάς το κοντά στο κάτω φράγμα $(H(X)/\log(M))$. Απάντηση σε αυτό το ερώτημα δίνει το **πρώτο θεώρημα του Shannon**.

Θεώρημα 4.7 (Πρώτο θεώρημα του Shannon). *Έστω μία πηγή πληροφορίας με αλφάβητο N συμβόλων, τα οποία κωδικοποιούνται με ένα M-αδικό στιγμιαία αποκωδικοποιήσιμο (μονοσήμαντο) κώδικα. Το μέσο μήκος του κώδικα προσεγγίζει όσο είναι επιθυμητό το κάτω φράγμα που είναι ίσο με τη M-αδική εντροπία της πηγής, αν κωδικοποιηθούν ανώτερες επεκτάσεις της πηγής.*

Το θεώρημα μπορεί να διατυπωθεί και ως εξής:

Θεώρημα 4.8. *Κωδικοποιώντας την n-οστή επέκταση (X^n, P_{X^n}) της πηγής (X, P_X) με n αρκετά μεγάλο, υπάρχει M-αδικός στιγμιαία αποκωδικοποιήσιμος (μονοσήμαντος) κώδικας με μέσο μήκος που πλησιάζει όσο θέλουμε στη M-αδική εντροπία της πηγής (X, P_X).*

Απόδειξη.
Η επέκταση τάξης n της πηγής (X, P_X) που συμβολίζεται ως (X^n, P_{X^n}) παράγει στην έξοδό της υπερσύμβολα n-άδες, τα οποία ανήκουν στο αλφάβητο X^n. Το αλφάβητο αυτό ορίζεται από το καρτεσιανό γινόμενο

$$X^n = \underbrace{X \times X \times \cdots \times X}_{n}. \tag{4.17}$$

Η πιθανότητα εμφάνισης ενός υπερσυμβόλου (n-άδας) θα είναι ίση με το γινόμενο όλων των πιθανοτήτων των επιμέρους συμβόλων που σχηματίζουν το υπερσύμβολο. Εύκολα αποδεικνύεται ότι η εντροπία της n-οστής επέκτασης της πηγής X θα είναι:

$$H(X^n) = nH(X). \tag{4.18}$$

Σύμφωνα με τα συμπεράσματα της προηγούμενης ενότητας υπάρχει M-αδικός στιγμιαία αποκωδικοποιήσιμος (μονοσήμαντος) κώδικας, που κωδικοποιεί τα υπερσύμβολα, του οποίου το μέσο μήκος $\bar{L}$ ανά υπερσύμβολο ικανοποιεί την ταυτοανισότητα

$$\frac{H(X^n)}{\log(M)} \leq \bar{L} \leq \frac{H(X^n)}{\log(M)} + 1. \tag{4.19}$$

Αντικαθιστώντας την (4.18) στην (4.19) έχουμε

$$\frac{H(X)}{\log(M)} \leq \frac{\bar{L}}{n} \leq \frac{H(X)}{\log(M)} + \frac{1}{n}. \tag{4.20}$$

Προφανώς, ο λόγος $\bar{L}/n$ εκφράζει το μέσο μήκος του κώδικα ανά σύμβολο της πηγής (X, P_X). Παρατηρώντας τη διπλή ταυτοανισότητα (4.20), καταλήγουμε στο συμπέρασμα ότι κωδικοποιώντας ανώτερες επεκτάσεις της πηγής (X, P_X), το μέσο μήκος του κώδικα ανά σύμβολο πηγής (X, P_X) τείνει στη M-αδική εντροπία της. Αυτό φαίνεται από το γεγονός ότι το άνω φράγμα $(H(X)/\log(M) + 1/n)$ τείνει στο κάτω φράγμα $(H(X)/\log(M))$ καθώς το n τείνει στο άπειρο.

♦

Κλείνοντας, πρέπει να υπενθυμιστεί ότι στην όλη παρουσίαση αυτού του κεφαλαίου αγνοήθηκε η παρουσία θορύβου στο σύστημα επικοινωνίας. Απλά αναζητήθηκε ο βέλτιστος κώδικας για την αναπαράσταση της πληροφορίας της πηγής με κριτήριο το μέσο μήκος. Για το λόγο αυτό το πρώτο θεώρημα του Shannon λέγεται και **θεώρημα κωδικοποίησης σε περιβάλλον χωρίς θόρυβο**.

4.11 Παραδείγματα

1. *Στον παρακάτω πίνακα δίνονται οι πιθανότητες των συμβόλων μιας πηγής, καθώς και πέντε διαφορετικοί κώδικες που χρησιμοποιούνται για την κωδικοποίηση των συμβόλων. Να ταξινομηθούν οι κώδικες σε κατηγορίες και να υπολογιστεί το μέσο μήκος τους.*

x_i	p_i	I	II	III	IV	V
Α	0.5	0	00	110	000	01
Β	0.3	01	11	01	010	011
Γ	0.15	11	10	00	101	0111
Δ	0.05	10	01	10	111	0

Εξετάζουμε τις ιδιότητες των πέντε κωδίκων και δημιουργούμε τον πίνακα ταξινόμησης, όπου **Ε** δηλώνει ευκρινής, **Μ** μονοσήμαντος και **Σ** στιγμιαία αποκωδικοποιήσιμος.

Ιδιότητα	I	II	III	IV	V
Ε	✓	×	✓	✓	✓
Μ	×	×	✓	✓	✓
Σ	×	×	✓	✓	×
$\bar{L}$	1.5	2.0	2.5	3.0	2.55

Ο κώδικας I είναι ευκρινής γιατί καμία κωδική λέξη δεν αντιστοιχεί σε δύο ή περισσότερα σύμβολα. Δεν είναι όμως μονοσήμαντος γιατί η ακολουθία συμβόλων 0110 μπορεί να αποκωδικοποιηθεί ως ΑΓΑ ή ως ΒΔ. Αφού δεν είναι μονοσήμαντος κατ' επέκταση δεν μπορεί να είναι στιγμιαία αποκωδικοποιήσιμος.

Ο κώδικας II δεν είναι ευκρινής γιατί η κωδική λέξη 01 αντιστοιχεί σε δύο σύμβολα πηγής, το Α και το Δ. Κατ' επέκταση ο κώδικας II δε μπορεί να είναι ούτε μονοσήμαντος, ούτε στιγμιαία αποκωδικοποιήσιμος.

Ο κώδικας III είναι στιγμιαία αποκωδικοποιήσιμος γιατί οι κωδικές λέξεις του διαθέτουν την προθεματική ιδιότητα. Κατ' επέκταση είναι μονοσήμαντος και ευκρινής.

Ο κώδικας IV είναι στιγμιαία αποκωδικοποιήσιμος, μονοσήμαντος και ευκρινής γιατί έχει την προθεματική ιδιότητα.

Τέλος, ο κώδικας V δεν είναι στιγμιαία αποκωδικοποιήσιμος επειδή για παράδειγμα η κωδική λέξη 01 είναι πρόθεμα της κωδικής λέξης 001. Όμως είναι ευκρινής και μονοσήμαντος γιατί κάθε κωδική λέξη μπορεί να αναγνωριστεί μέσα σε οποιοδήποτε κωδικό μήνυμα, δεδομένου ότι αρχίζει πάντα με το σύμβολο 0 ενώ τα υπόλοιπα σύμβολα της κωδικής λέξης μπορούν να είναι μόνο 1.

Τα μέσο μήκος του κάθε κώδικα υπολογίζεται από τη σχέση

$$\bar{L} = \sum_{i=1}^{4} p_i \cdot l_i.$$

Οι τιμές που προκύπτουν για το μέσο μήκος φαίνονται στον πίνακα.

2. *Υπάρχει τριαδικός στιγμιαία αποκωδικοποιήσιμος κώδικας με κωδικές λέξεις μήκους* $\{1, 2, 3, 3, 2, 3\}$*;*

   Αρκεί να εξετάσουμε αν ικανοποιείται η ταυτοανισότητα του Kraft. Έχουμε:

   $$\sum_{i=1}^{N} 3^{-i} = 3^{-1} + 3^{-2} + 3^{-3} + 3^{-3} + 3^{-2} + 3^{-3} = 18/27 \leq 1.$$

   Άρα υπάρχει στιγμιαία αποκωδικοποιήσιμος κώδικας με τα παραπάνω μήκη κωδικών λέξεων.

3. *Υπάρχει* M*-αδικός στιγμιαία αποκωδικοποιήσιμος κώδικας με κωδικές λέξεις μήκους* $1, 2, \ldots, n, ...$*;*

   Αρκεί να εξετάσουμε αν ικανοποιείται η ταυτοανισότητα του Kraft. Έχουμε:

   $$\sum_{n=1}^{\infty} M^{-n} = \sum_{n=1}^{\infty} \left(\frac{1}{M}\right)^n = \frac{1}{M-1}.$$

   Το παραπάνω είναι άθροισμα απείρων όρων γεωμετρικής προόδου ξεκινώντας από $n = 1$ και συγκλίνει με την προϋπόθεση

   $$\frac{1}{M} < 1 \Rightarrow M > 1.$$

   Για να υπάρχει στιγμιαία αποκωδικοποιήσιμος κώδικας θα πρέπει

   $$\frac{1}{M-1} \leq 1 \Rightarrow M > 1.$$

   Άρα αρκεί $M \geq 2$.

4. *Θεωρούμε δυαδική πηγή πληροφορίας με αλφάβητο* $X = \{0, 1\}$ *και κατανομή πιθανοτήτων* $P_X = \{0.9, 0.1\}$. *Να υπολογιστεί το μέσο μήκος στιγμιαία αποκωδικοποιήσιμου δυαδικού κώδικα που κωδικοποιεί τα σύμβολα της πηγής* (X, P_X). *Ποιο το μέσο μήκος στιγμιαία αποκωδικοποιήσιμου δυαδικού κώδικα που κωδικοποιεί τα σύμβολα της δεύτερης επέκτασης* (X^2, P_{X^2}) *της πηγής; Θεωρείστε ότι τα μήκη των κωδικών λέξεων προκύπτουν με εφαρμογή της σχέσης (4.15).*

   Δεδομένου ότι ο κώδικας είναι δυαδικός ($M = 2$) τα μήκη των κωδικών λέξεων δίνονται από τη σχέση

$$l_i = \lceil -\log(p_i) \rceil, \qquad 1 \leq i \leq N$$

   Για την περίπτωση της απλής πηγής δημιουργούμε τον παρακάτω πίνακα που περιέχει τα σύμβολα, τις πιθανότητες και τα μήκη των κωδικών λέξεων.

x_i	p_i	$l^* = -\log(p_i)$	$l_i = \lceil -\log(p_i) \rceil$
0	0.9	0.152	1
1	0.1	3.322	4

   Το μέσο μήκος του κώδικα θα είναι προφανώς ίσο με $0.9 + 0.4 = 1.3$ bits ανά σύμβολο πηγής.

   Στην περίπτωση της δεύτερης επέκτασης της πηγής θα έχουμε τον αντίστοιχο πίνακα.

(x_i, x_j)	p_i	$l^* = -\log(p_i)$	$l_i = \lceil -\log(p_i) \rceil$
00	$0.9 \times 0.9 = 0.81$	0.304	1
01	$0.9 \times 0.1 = 0.09$	3.474	4
10	$0.1 \times 0.9 = 0.09$	3.474	4
11	$0.1 \times 0.1 = 0.01$	6.644	7

   Σε αυτήν την περίπτωση το μέσο μήκος θα είναι

$$\bar{L} = 0.81 + 4 \times 0.09 + 4 \times 0.09 + 7 \times 0.01 = 1.6$$

   bits ανά δύο σύμβολα πηγής. Συνεπώς το μέσο μήκος ανά σύμβολο πηγής θα είναι 0.8 bits ανά σύμβολο πηγής. Παρατηρούμε ότι κωδικοποιώντας την επέκταση της πηγής το μέσο μήκος του κώδικα μειώθηκε.

4.12 Ασκήσεις

1. Στον παρακάτω πίνακα δίνονται οι πιθανότητες των συμβόλων μιας πηγής, καθώς και τέσσερις διαφορετικοί κώδικες που χρησιμοποιούνται για την κωδι-

κοποίηση των συμβόλων. Να ταξινομηθούν οι κώδικες σε κατηγορίες και να υπολογιστεί το μέσο μήκος τους.

x_i	p_i	I	II	III	IV
A	0.5	0	00	101	00
B	0.3	01	11	11	010
Γ	0.1	11	10	0	111
Δ	0.05	10	01	1001	101
E	0.05	101	11	1000	001

2. Υπάρχει δυαδικός στιγμιαία αποκωδικοποιήσιμος κώδικας με κωδικές λέξεις μήκους $\{1, 4, 3, 3, 2, 3\}$;

3. Είναι ο παρακάτω κώδικας μονοσήμαντος και γιατί;

   $x_1 : 010 \quad x_3 : 11 \quad x_5 : 00011$
   $x_2 : 0001 \quad x_4 : 101011 \quad x_6 : 00110$

4. Θεωρούμε πηγή πληροφορίας με αλφάβητο $X = \{0, 1, 2\}$ και κατανομή πιθανοτήτων $P_X = \{0.7, 0.2, 0.1\}$. Να υπολογιστεί το μέσο μήκος στιγμιαία αποκωδικοποιήσιμου δυαδικού κώδικα που κωδικοποιεί τα σύμβολα της πηγής (X, P_X). Ποιο το μέσο μήκος στιγμιαία αποκωδικοποιήσιμου δυαδικού κώδικα που κωδικοποιεί τα σύμβολα της δεύτερης επέκτασης (X^2, P_{X^2}) της πηγής; Θεωρείστε ότι τα μήκη των κωδικών λέξεων προκύπτουν με εφαρμογή της σχέσης (4.15). Να συγκριθούν τα μέσα μήκη των δύο κωδίκων με την εντροπία της πηγής.

5. Θεωρούμε μία πηγή πληροφορία με αλφάβητο $X = \{a, b\}$ και κατανομή πιθανοτήτων $P_X = \{0.4, 0.6\}$. Να βρεθούν τα όρια τιμών του μέσου μήκους ανά σύμβολο πηγής ενός τριαδικού στιγμιαία αποκωδικοποιήσιμου κώδικα που χρησιμοποιείται για την κωδικοποίηση της τρίτης επέκτασης της πηγής πληροφορίας.

6. Θεωρούμε ένα μη τίμιο νόμισμα. Κατά τη ρίψη του νομίσματος η πιθανότητα να έρθει κεφαλή είναι 0.7. Να βρεθεί μία 'καλή' εκτίμηση του μέσου μήκους (ανά αποτέλεσμα ρίψης) ενός δυαδικού στιγμιαία αποκωδικοποιήσιμου κώδικα, ο οποίος θα χρησιμοποιηθεί για την κωδικοποίηση των αποτελεσμάτων δύο διαδοχικών ρίψεων του νομίσματος.

   Ποια είναι η περιοχή των τιμών που θα μπορούσε να πάρει το μέσο μήκος (ανά αποτέλεσμα ρίψης) ενός δυαδικού στιγμιαία αποκωδικοποιήσιμου κώδικα, ο οποίος κωδικοποιεί τα αποτελέσματα τριών διαδοχικών ρίψεων του νομίσματος;

Κεφάλαιο 5

Συμπίεση χωρίς Απώλειες

Η συμπίεση δεδομένων έχει δύο βασικούς στόχους. Πρώτον, να μειώσει το απαιτούμενο χώρο για την αποθήκευση των πληροφοριών και δεύτερον, να περιορίσει την απαιτούμενη χωρητικότητα του διαύλου επικοινωνίας για τη μεταφορά των πληροφοριών. Ο δεύτερος στόχος ισοδυναμεί με τη μείωση του συνολικού χρόνου αποστολής των πληροφοριών από τον πομπό στο δέκτη.

Όταν τα συμπιεσμένα δεδομένα ανακτώνται από το χώρο αποθήκευσής τους π.χ από το σκληρό δίσκο, ή λαμβάνονται από το δέκτη ενός τηλεπικοινωνιακού συστήματος, θα πρέπει να υπάρχει η δυνατότητα πλήρους ανάκτησης της αρχικής μορφής των δεδομένων. Συνεπώς, κατά κάποιο τρόπο θα πρέπει ο κώδικας με τον οποίο έγινε η συμπίεση να είναι γνωστός στον αλγόριθμο που αναλαμβάνει τη διαδικασία της αποσυμπίεσης. Αυτό μπορεί να γίνει είτε έχοντας ένα πάγιο κώδικα γνωστό εκ των προτέρων στο δέκτη, είτε προσθέτοντας στα συμπιεσμένα δεδομένα και το τμήμα του κώδικα που χρειάζεται για την αποσυμπίεση των δεδομένων. Στην δεύτερη περίπτωση, για να έχει ουσιαστικό νόημα η συμπίεση, θα πρέπει το άθροισμα των συμπιεσμένων δεδομένων και του τμήματος του κώδικα που προστίθεται, να είναι σημαντικά μικρότερο από το αρχικό μήνυμα.

Όπως έχει αναφερθεί στην εισαγωγή, διακρίνουμε δύο κατηγορίες συμπίεσης δεδομένων: τη συμπίεση δεδομένων με απώλειες και χωρίς απώλειες. Στο κεφάλαιο αυτό θα ασχοληθούμε με **κώδικες συμπίεσης δεδομένων χωρίς απώλειες**. Στο προηγούμενο κεφάλαιο, αποδείχθηκε ότι η συμπίεση χωρίς απώλειες περιορίζεται από την εντροπία στης πηγής πληροφοριών. Συγκεκριμένα η εντροπία είναι το κάτω φράγμα του μέσου μήκους του κώδικα συμπίεσης. Συνεπώς, κατά το σχεδιασμό ενός αλγορίθμου συμπίεσης δεδομένων χωρίς απώλειες, ο κύριος στόχος είναι να πετύχουμε ένα μέσο μήκος κώδικα που να πλησιάζει όσο το δυνατό πιο κοντά στην εντροπία.

Επίσης, ένας σημαντικός παράγοντας, που λαμβάνεται υπόψη, κατά τη σχεδίαση και την επιλογή ενός αλγορίθμου συμπίεσης δεδομένων, είναι η ο φόρτος εργασίας που απαιτείται τόσο για τη συμπίεση όσο και για την αποσυμπίεση των δεδομένων. Γενικά, θα πρέπει να βρεθεί η χρυσή τομή ανάμεσα στο ποσοστό συμπίεσης που επιτυγχάνει ένας αλγόριθμος και το χρόνο που απαιτείται για να εκτελεστεί.

5.1 Ιδιότητες του Βέλτιστου Κώδικα Συμπίεσης

Θεωρούμε μία πηγή πληροφορίας (X, P_X), της οποίας το αλφάβητο έχει N σύμβολα με αντίστοιχες πιθανότητες εμφάνισης $p_1, p_2, \ldots, p_N$. Οι πιθανότητες αυτές ταξινομούνται κατά φθίνουσα σειρά ($p_1 \geq p_2 \geq \ldots \geq p_N$). Θέλουμε κάθε ένα από τα σύμβολα της πηγής να κωδικοποιηθεί κατά τρόπο μοναδικό με μία ακολουθία δυαδικών ψηφίων. Έτσι, θα προκύψει ένας δυαδικός κώδικας, με μήκη κωδικών λέξεων $l_1, l_2, \ldots, l_N$ που αντιστοιχούν ένα προς ένα στα N σύμβολα της πηγής. .

Ένας κώδικας συμπίεσης θα είναι βέλτιστος όταν έχει το ελάχιστο μέσο μήκος κώδικα. Σημειώνεται ότι δεν αναφερόμαστε στο ελάχιστο μήκος κώδικα ενός μόνο συγκεκριμένου μηνύματος αλλά στο μέσο μήκος κώδικα ως προς όλα τα πιθανά μηνύματα που είναι δυνατόν να παραχθούν από την πηγή πληροφορίας. Μπορεί να αποδειχθεί ότι ένας βέλτιστος κώδικας συμπίεσης θα πρέπει να ικανοποιεί τις παρακάτω απαιτήσεις:

1. Καμία ακολουθία κωδικών συμβόλων δε θα πρέπει να αντιστοιχίζεται σε περισσότερα από ένα διαφορετικά μηνύματα της πηγής. Δηλαδή, ο κώδικας συμπίεσης θα πρέπει να είναι μονοσήμαντος.

2. Καμία κωδική λέξη δεν πρέπει να είναι πρόθεμα άλλης κωδικής λέξεις. Δηλαδή, ο κώδικας θα πρέπει να έχει την προθεματική ιδιότητα και συνεπώς να είναι στιγμιαία αποκωδικοποιήσιμος.

3. Σύμβολα του αλφαβήτου με μεγαλύτερη πιθανότητα θα πρέπει να αντιστοιχίζονται σε μικρότερες κωδικές λέξεις. Δηλαδή $l_1 \leq l_2 \leq \ldots \leq l_N$, εφόσον $p_1 \geq p_2 \geq \ldots \geq p_N$.

4. Τέλος, οι δύο κωδικές λέξεις που αντιστοιχούν στα σύμβολα με τις δύο μικρότερες πιθανότητες εμφάνισης θα έχουν το ίδιο μήκος κωδικής λέξης ($l_N = l_{N-1}$) και θα διαφέρουν μόνο στο τελευταίο ψηφίο τους. (Αν ο κώδικας είναι M-αδικός η τελευταία ιδιότητα θα πρέπει να ικανοποιείται από τις κωδικές λέξεις με τις M μικρότερες πιθανότητες.)

Το ότι ο βέλτιστος κώδικας θα πρέπει να ικανοποιεί την τρίτη απαίτηση μπορεί να αποδειχθεί με την εις άτοπο απαγωγή ως εξής:

Απόδειξη.
Έστω, ότι για δύο κωδικές λέξεις του βέλτιστου κώδικα C δεν ισχύει η τρίτη απαίτηση, δηλαδή έχουμε $l_i > l_j$ αν και $p_i > p_j$. Μπορούμε να κατασκευάσουμε ένα δεύτερο κώδικα C' που να έχει τις ίδιες κωδικές λέξεις με τον C, με μόνη διαφορά ότι έχουν αντιστραφεί οι κωδικές λέξεις των i και j συμβόλων. Δηλαδή $l'_i = l_j$ και $l'_j = l_i$. Αν $\bar{L}$ και $\bar{L}'$ είναι τα μέσα μήκη των κωδίκων C και C', αντίστοιχα, προκύπτει εύκολα ότι

$$\begin{aligned}\bar{L} - \bar{L}' &= p_i \cdot l_i + p_j \cdot l_j - p_i \cdot l'_i + p_j \cdot l'_j = \\ &= p_i \cdot l_i + p_j \cdot l_j - p_i \cdot l_j - p_j \cdot l_i = \\ &= (p_i - p_j) \cdot (l_i - l_j) > 0.\end{aligned}$$

Συνεπώς, ο κώδικας C' έχει μικρότερο μέσο μήκος από τον C. Καταλήξαμε λοιπόν σε άτοπο αφού ο C είναι βέλτιστος και έχει το ελάχιστο μέσο μήκος. Άρα, για να είναι ο κώδικας βέλτιστος θα πρέπει να ισχύει η τρίτη απαίτηση.

♦

Τέλος, η τέταρτη απαίτηση μπορεί να αποδειχθεί ως εξής:

Απόδειξη.
Έστω ότι για τα δύο σύμβολα με τις μικρότερες πιθανότητες εμφάνισης ισχύει $l_N > l_{N-1}$. Σύμφωνα με τη δεύτερη απαίτηση (προθεματική ιδιότητα) η $(N-1)$-οστή κωδική λέξη δεν μπορεί να είναι πρόθεμα της N-οστής κωδικής λέξης. Συνεπώς, τα πρώτα l_{N-1} κωδικά σύμβολα της N-οστής κωδικής λέξης αποτελούν από μόνα τους μοναδική κωδική λέξη, οπότε τα υπόλοιπα $l_N - l_{N-1}$ κωδικά σύμβολα μπορούν να παραληφθούν. Έτσι, προκύπτει το άτοπο συμπέρασμα ότι ο αρχικός κώδικας δεν είναι βέλτιστος. Ακόμη, αν οι δύο κωδικές λέξεις διαφέρουν σε άλλο ψηφίο και όχι στο τελευταίο, τότε θα μπορούσαμε να παραλήψουμε τα τελευταία ίδια ψηφία και να δημιουργήσουμε ένα καινούργιο κώδικα με μικρότερο μέσο μήκος. Συνεπώς, ο αρχικός κώδικας δε θα ήταν βέλτιστος. Καταλήγουμε λοιπόν στο συμπέρασμα ότι για να είναι ο κώδικας συμπίεσης βέλτιστος θα πρέπει να ισχύει $l_N = l_{N-1}$, και οι αντίστοιχες κωδικές λέξεις να διαφέρουν μόνο στο τελευταίο κωδικό σύμβολο.

♦

5.2 Κώδικας Huffman

Το 1950, ο David Huffman, φοιτητής τότε στο MIT, πρότεινε σε μία εργασία του για το μάθημα της θεωρίας πληροφοριών έναν αλγόριθμο συμπίεσης χωρίς απώλειες. Ο αλγόριθμος αυτός που φέρει το όνομά του είναι σήμερα το βασικότερο τμήμα των περισσοτέρων αλγορίθμων συμπίεσης, αφού χρησιμοποιείται ως το τελευταίο στάδιο των αλγορίθμων GZIP, JPEG και πολλών άλλων. Η σημαντικότερη ιδιότητα του κώδικα Huffman είναι ότι πρόκειται για βέλτιστο κώδικα. Αυτό σημαίνει ότι δεν υπάρχει κώδικας χωρίς απώλειες με μικρότερο μέσο μήκος από τον κώδικα Huffman.

5.2.1 Αλγόριθμος δυαδικού κώδικα Huffman

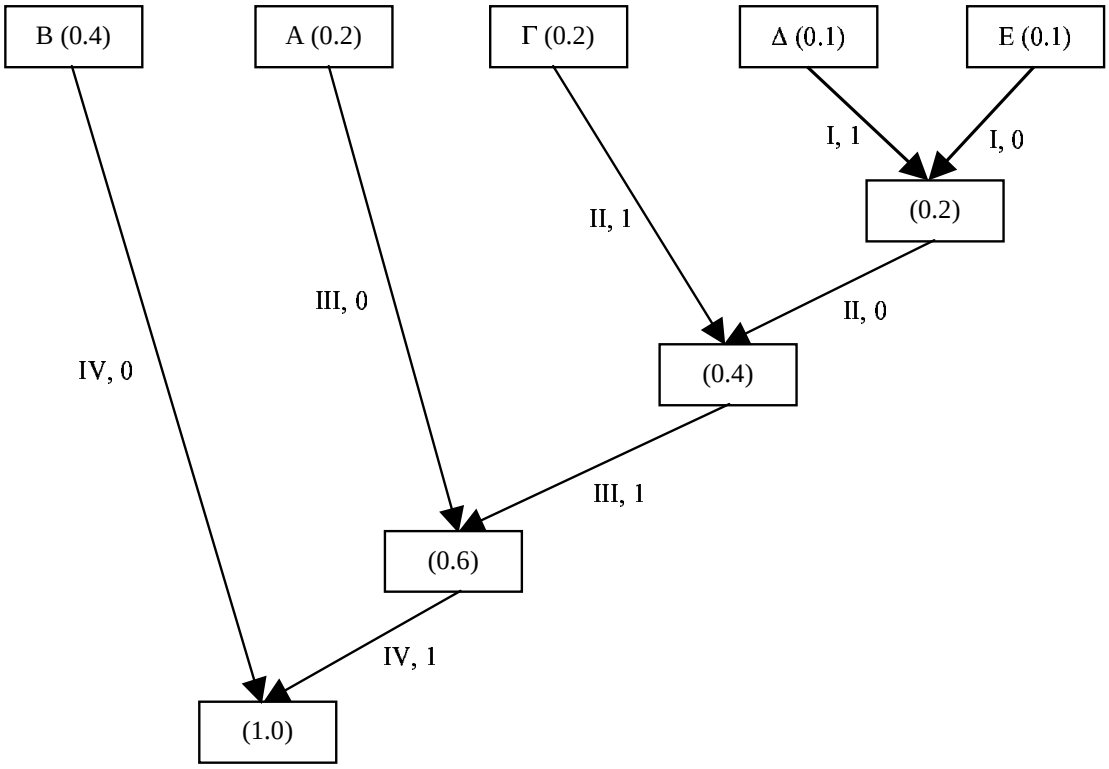
Για την περίπτωση του δυαδικού κώδικα, ο αλγόριθμος Huffman περιγράφεται με τα παρακάτω βήματα που οδηγούν στην κατασκευή ενός δενδροδιαγράμματος ξεκινώντας από τα φύλλα του και καταλήγοντας στη ρίζα.

1. Αρχικά θεωρούμε ότι κάθε σύμβολο της πηγής είναι ένα φύλλο (εξωτερικός κόμβος) του υπό κατασκευή δενδροδιαγράμματος, Σε κάθε φύλλο αντιστοιχίζουμε την πιθανότητα εμφάνισης του αντίστοιχου συμβόλου πηγής.

2. Εντοπίζουμε τους δύο κόμβους με τις μικρότερες πιθανότητες και τους συγχωνεύουμε σε ένα κόμβο. Στο νέο κόμβο αντιστοιχίζουμε ως πιθανότητα το άθροισμα των πιθανοτήτων των δύο κόμβων που συγχωνεύτηκαν.

3. Το δεύτερο βήμα επαναλαμβάνεται έως ότου καταλήξουμε στη δημιουργία του τελικού κόμβου (ρίζα).

Με τη λήξη του παραπάνω αλγορίθμου προκύπτει ένα δέντρο του οποίου κάθε εσωτερικός κόμβος έχει δύο κλάδους από πάνω και έναν από κάτω. Η ρίζα έχει μόνο δύο κλάδους από πάνω και κανέναν από κάτω, ενώ κάθε φύλλο που αντιστοιχεί σε ένα σύμβολο πηγής έχει μόνο έναν κλάδο από κάτω και κανέναν από πάνω. Στη συνέχεια, για κάθε εσωτερικό κόμβο και τη ρίζα, συμβολίσουμε τους δύο από πάνω κλάδους τον ένα με 1 και τον άλλο με 0. Έτσι κάθε κλάδος φέρει ένα δυαδικό ψηφίο. Η διαδικασία αυτή αντιστοίχισης των δυαδικών ψηφίων στους κλάδους μπορεί να γίνει και κατά τη διαδικασία δημιουργίας κάθε νέου κόμβου. Τέλος, η κωδική λέξη κάθε συμβόλου πηγής προκύπτει αν καταγράψουμε ακολουθιακά τα δυαδικά ψηφία των κλάδων που οδηγούν, ξεκινώντας από τη ρίζα, στο φύλλο που αντιστοιχεί στο



Σχήμα 5.1: *Δενδροδιάγραμμα κώδικα Huffman (πρώτη υλοποίηση).*

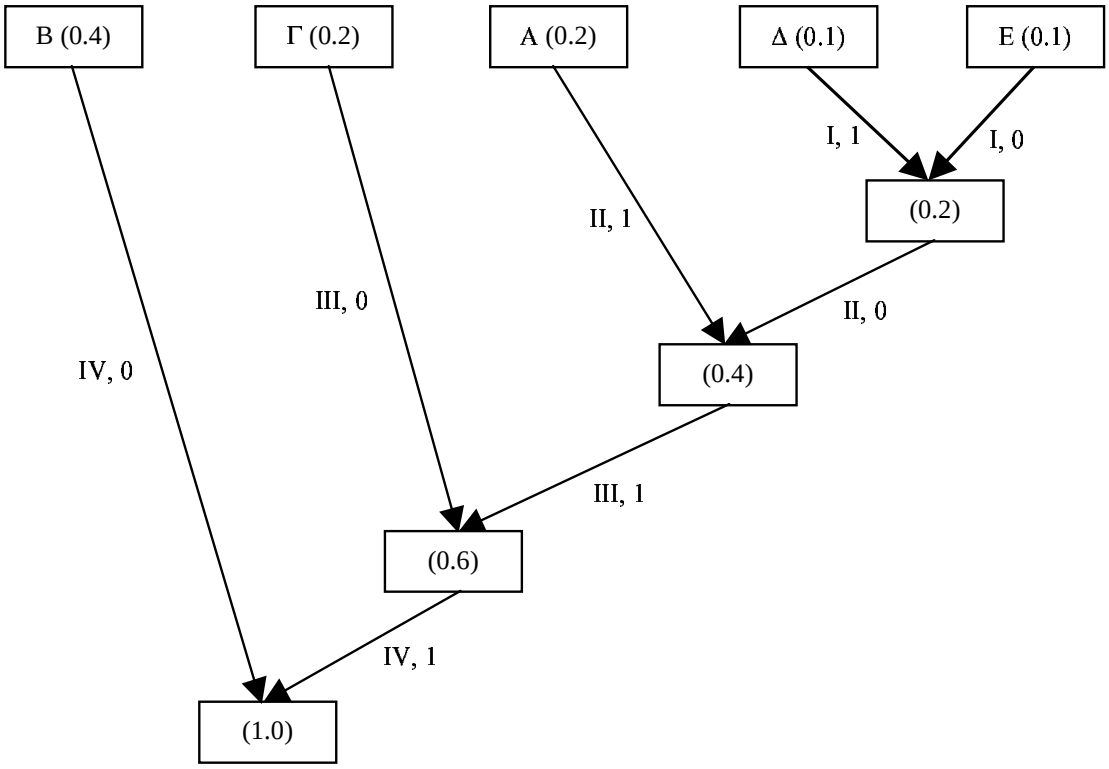
σύμβολο πηγής. Λόγω της δομής δέντρου, αυτή η ακολουθία κλάδων είναι μοναδική για κάθε φύλλο.

Σημειώνεται ότι δεν έχει σημασία σε ποιον από τους δύο κλάδους κάθε κόμβου θα αντιστοιχίσουμε το κωδικό σύμβολο 1 και σε ποιον το 0. Κατά σύμβαση, στη συνέχεια θα αντιστοιχούμε το σύμβολο 0 στον κλάδο που συνδέεται με τον κόμβο με τη μικρότερη πιθανότητα.

Ο αλγόριθμος Huffman μπορεί να γίνει κατανοητός με ένα παράδειγμα. Θεωρούμε μία πηγή πληροφορίας με σύμβολα πηγής και αντίστοιχες πιθανότητες εμφάνισης που δίνονται στον παρακάτω πίνακα.

x_i	A	B	Γ	Δ	E
p_i	0.2	0.4	0.2	0.1	0.1

Στο σχήμα 5.1 παρουσιάζεται το δενδροδιάγραμμα που δημιουργήθηκε κατά την εφαρμογή του αλγορίθμου Huffman. Σε κάθε κόμβο αναγράφεται εντός παρενθέσεων η αντίστοιχη πιθανότητα. Επίσης, σε κάθε κλάδο αναγράφεται η επανάληψη του αλγορίθμου (I, II, III, IV), κατά την οποία δημιουργήθηκε ο κλάδος, καθώς και το κωδικό δυαδικό ψηφίο που του αποδόθηκε.



Σχήμα 5.2: *Δενδροδιάγραμμα κώδικα Huffman(δεύτερη υλοποίηση).*

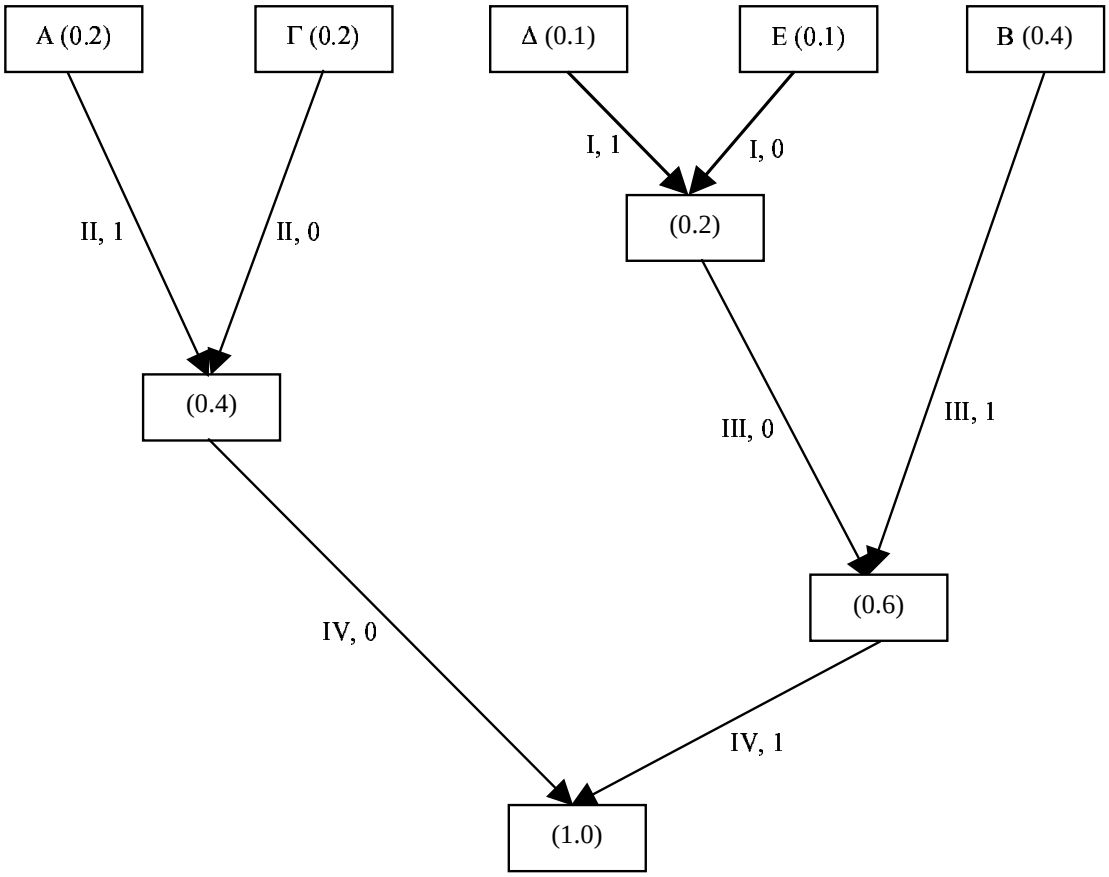
Παρατηρούμε ότι κατά την κατασκευή του δενδροδιαγράμματος μπορεί να προκύψει η περίπτωση να υπάρχουν περισσότεροι από δύο κόμβοι που να έχουν τις δύο μικρότερες τιμές πιθανοτήτων. Έτσι, είναι δυνατό να επιλεγούν εναλλακτικές υλοποιήσεις του κώδικα Huffman. Μία τέτοια εναλλακτική υλοποίηση φαίνεται στο σχήμα 5.2. Βλέπουμε ότι τα μήκη των κωδικών λέξεων είναι διαφορετικά σε σχέση με αυτά της πρώτης υλοποίησης. Όμως, δεδομένου ότι κάθε εναλλακτική υλοποίηση αναφέρεται σε κώδικα Huffman που είναι βέλτιστος, προκύπτει το συμπέρασμα ότι τα μέσα μήκη κάθε εναλλακτικής υλοποίησης θα πρέπει να είναι ίσα μεταξύ τους.

5.2.2 Κώδικας Huffman ελάχιστης μεταβλητότητας

Αν και οι εναλλακτικές υλοποιήσεις του κώδικα Huffman έχουν το ίδιο μέσο μήκος κώδικα, διαφέρουν ως προς τη μεταβλητότητα του μήκους των κωδικών λέξεων, όπου η μεταβλητότητα ορίζεται ως

$$\sigma^2 = \sum_{i=1}^{N} p_i \cdot (l_i - \bar{L})^2. \tag{5.1}$$

Συνήθως, επιλέγεται εκείνη η υλοποίηση του κώδικα Huffman που οδηγεί στη μικρότερη μεταβλητότητα του μήκους των κωδικών λέξεων. Ο λόγος για αυτήν την



Σχήμα 5.3: *Δενδροδιάγραμμα κώδικα Huffman με ελάχιστη μεταβλητότητα του μήκους των κωδικών λέξεων (τρίτη υλοποίηση).*

επιλογή είναι ότι η μικρότερη μεταβλητότητα σημαίνει, πρώτον, μικρή μεταβολή του ρυθμού μετάδοσης των διαφόρων συμβόλων της πηγής και δεύτερον, μείωση της μνήμης (buffer) που απαιτείται για την προσωρινή αποθήκευση των λαμβανομένων κωδικών συμβόλων, μέχρι να αναγνωριστεί το αντίστοιχο σύμβολο πηγής.

Αποδεικνύεται, ότι για να πετύχουμε την υλοποίηση με τη μικρότερη μεταβλητότητα του μήκους των κωδικών λέξεων, θα πρέπει σε κάθε βήμα του αλγορίθμου (Huffman) όπου προκύπτουν περισσότερες από μία επιλογές για τους κόμβους μικρότερης πιθανότητας, να επιλέγουμε για συγχώνευση τους δύο κόμβους που έχουν δημιουργηθεί παλαιότερα. Αν εφαρμόσουμε αυτή τη λογική στην κωδικοποίηση της πηγής του παραδείγματος προκύπτει η υλοποίηση του σχήματος 5.3.

Στον πίνακα 5.1 παρουσιάζονται οι τρεις εναλλακτικές υλοποιήσεις του κώδικα Huffman της πηγής του παραδείγματος. Συγκεκριμένα, ο πίνακας περιέχει, για κάθε υλοποίηση, τις κωδικές λέξεις με τα αντίστοιχα μήκη τους, το μέσο μήκος κώδικα και τη μεταβλητότητα του μήκους των κωδικών λέξεων.

Θυμίζουμε ότι το μέσο μήκος του κώδικα Huffman θα πρέπει να ικανοποιεί την διπλή

Πίνακας 5.1: *Τρεις εναλλακτικές υλοποιήσεις του κώδικα Huffman*

		1^η		2^η		3^η	
x_i	p_i	c_i	l_i	c_i	l_i	c_i	l_i
A	0.2	10	2	111	3	01	2
B	0.4	0	1	0	1	11	2
Γ	0.2	111	3	10	2	00	2
Δ	0.1	1101	4	1101	4	101	3
E	0.1	1100	4	1100	4	100	3
$\bar{L}$		2.2		2.2		2.2	
σ^2		1.36		1.36		0.16	

ταυτοανισότητα

$$\frac{H(X)}{\log(M)} \leq \bar{L} \leq \frac{H(X)}{\log(M)} + 1, \tag{5.2}$$

αφού είναι βέλτιστος κώδικας. Συγκεκριμένα, στο παράδειγμα η εντροπία της πηγής είναι ίση με

$$H(X) = -0.4 \cdot \log(0.4) - 2 \cdot 0.2 \cdot \log(0.2) - 2 \cdot 0.1 \cdot \log(0.1) \simeq 2.122, \tag{5.3}$$

οπότε, πράγματι, ικανοποιείται η (5.2) για $M = 2$ ($2.122 \leq 2.2 \leq 3.122$).

Τέλος, παρατηρούμε ότι σε κάθε επανάληψη του δυαδικού αλγόριθμου Huffman το πλήθος των κόμβων μειώνεται κατά ένα. Στη περίπτωση του M-αδικού Huffman προκύπτει εύκολα ότι σε κάθε επανάληψη το πλήθος των κόμβων μειώνεται κατά $M - 1$. Για να τερματίσει σωστά ο αλγόριθμος θα πρέπει στο τελευταίο βήμα του αλγορίθμου να υπάρχουν M κόμβοι. Συνεπώς, αν n είναι το πλήθος των επαναλήψεων του αλγορίθμου πριν την τελευταία επανάληψη, τότε θα πρέπει να ισχύει η ισότητα

$$N = M + n(M - 1) \tag{5.4}$$

ή ισοδύναμα

$$N - 1 = (n + 1)(M - 1), \tag{5.5}$$

όπου $n + 1$ ο συνολικός αριθμός των επαναλήψεων του αλγόριθμου. Αν το $N - 1$ δεν είναι πολλαπλάσιο του $M - 1$ θα πρέπει να προσθέσουμε τόσα εικονικά σύμβολα πηγής με μηδενικές πιθανότητες, όσα χρειάζονται για την ικανοποίηση της ισότητας (5.5).

5.3 Αριθμητική Κωδικοποίηση

Ο κώδικας Huffman, αν και χαρακτηρίζεται από σχετική απλότητα στην υλοποίηση και την εκτέλεσή του, δεν μπορεί να επιτύχει τη καλύτερη δυνατή συμπίεση (στο όριο που θέτει η εντροπία της πηγής) εκτός από την περίπτωση κατά την οποία όλα τα σύμβολα της πηγής έχουν πιθανότητες ίσες με αρνητικές δυνάμεις του 2. Για να πετύχουμε καλύτερη συμπίεση με τον αλγόριθμο Huffman απαιτείται η κωδικοποίηση επεκτάσεων της πηγής πληροφορίας (κωδικοποίηση υπερσυμβόλων), η οποία όμως οδηγεί σε μεγάλο πλήθος κωδικών λέξεων και συνεπώς σε μεγάλα δενδροδιαγράμματα απόφασης.

Για να επιτευχθεί ακόμη καλύτερη συμπίεση από αυτήν που παρέχει ο αλγόριθμος Huffman, προτάθηκε ο αλγόριθμος **αριθμητικής κωδικοποίησης**. Φιλοσοφία του αλγορίθμου αυτού είναι η συνδυασμένη εκμετάλλευση κωδικών συμβόλων από περισσότερα του ενός συμβόλων της πηγής. Με αυτόν τον τρόπο, το μέσο μήκος του κώδικα ανά σύμβολο πηγής τείνει ασυμπτωτικά στην εντροπία. Χαρακτηριστικά αναφέρουμε το παράδειγμα της κωδικοποίησης χιλίων ίδιων συμβόλων με πιθανότητα καθενός ίσης με 0.999. Αν εφαρμόσουμε τον δυαδικό κώδικα Huffman, κάθε σύμβολο θα πρέπει να παρασταθεί με ένα bit, οπότε το συμπιεσμένο μήνυμα θα αποτελείται από 1000 bits. Όμως, η πληροφορία που περιέχει αυτό το μήνυμα είναι ίση με $1000 \cdot \log(0.999) \simeq 1.4$ bits, και άρα ο κώδικας Huffman είναι ιδιαίτερα σπάταλος. Όπως θα δούμε στη συνέχεια, αν συμπιεστεί αυτό το μάλλον ιδιόμορφο μήνυμα με αριθμητική κωδικοποίηση, αρκούν μόνο 3 bits. Για να πετύχουμε με αριθμητική κωδικοποίηση τόσο σημαντική συμπίεση θα πρέπει στην κατανομή των πιθανοτήτων των συμβόλων πηγής να εμφανίζονται ορισμένες μεγάλες πιθανότητες ενώ άλλες θα πρέπει λογικά να είναι πολύ μικρές.

Η βασική ιδέα της αριθμητικής κωδικοποίησης είναι η απεικόνιση κάθε ακολουθίας n συμβόλων πηγής σε έναν δεκαδικό αριθμό d μεταξύ 0 και 1. Στη συνέχεια, ο δεκαδικός αριθμός d γράφεται σε δυαδική μορφή. Αν $p(1), p(2), \ldots, p(n)$ είναι η πιθανότητες των συμβόλων που συνθέτουν το μήνυμα των n συμβόλων, τότε ορίζουμε το διάστημα

$$s = \prod_{i=1}^{n} p(i). \tag{5.6}$$

Τέλος, από τη δυαδική αναπαράσταση του αριθμού d κρατάμε μόνο τα πιο σημαντικά ψηφία που αρκούν για να οριστεί μονοσήμαντα το διάστημα s.

5.3.1 Αριθμητική συμπίεση

Για να γίνει πιο κατανοητή η διαδικασία της αριθμητικής συμπίεσης θα χρησιμοποιήσουμε ένα παράδειγμα. Έστω, μία πηγή πληροφορίας με αλφάβητο τριών συμβόλων $\{a, b, c\}$ και κατανομή πιθανοτήτων $P_X = \{p_a = 0.2, p_b = 0.5, p_c = 0.3\}$.

Αρχικά, σε κάθε σύμβολο πηγής αντιστοιχίζεται ένα τμήμα του διαστήματος $[0, 1]$, το οποίο είναι ανάλογο της πιθανότητας εμφάνισης του συμβόλου. Πρακτικά, το διάστημα $[0, 1]$ χωρίζεται στα τμήματα

$$[0, 0.2) \rightarrow a, \qquad [0.2, 0.7) \rightarrow b, \qquad [0.7, 1.0) \rightarrow c.$$

Έτσι, αν θέλουμε να κωδικοποιήσουμε το σύμβολο b, αρκεί να το αναπαραστήσουμε με οποιονδήποτε αριθμό, ο οποίος ανήκει στο διάστημα $[0.2, 0.7)$, π.χ. με τον αριθμό 0.4. Στη συνέχεια, κάθε νέο σύμβολο που κωδικοποιείται περιορίζει στην ουσία το εύρος του διαστήματος (το κάτω, το πάνω, ή και τα δύο όρια του διαστήματος). Αυτό γίνεται ως εξής:

Το διάστημα $[0.2, 0.7)$, που ορίστηκε για την κωδικοποίηση του συμβόλου b, κατακερματίζεται σε τμήματα ανάλογα των πιθανοτήτων των τριών συμβόλων της πηγής, τα οποία στη συνέχεια αντιστοιχίζονται εκ' νέου στα σύμβολα της πηγής. Δηλαδή το $[0.2, 0.7)$ χωρίζεται στα διαστήματα

$$[0.2, 0.3) \rightarrow a, \qquad [0.3, 0.55) \rightarrow b, \qquad [0.55, 0.7) \rightarrow c.$$

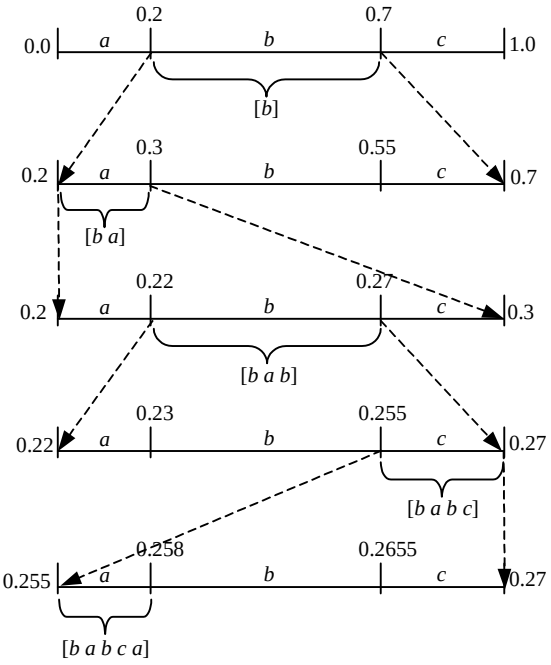
Αν τώρα το δεύτερο σύμβολο πηγής που κωδικοποιείται είναι το a, τότε το συνολικό μήνυμα $[b\,a]$ αναπαριστάται με οποιονδήποτε αριθμό που ανήκει στο διάστημα $[0.2, 0.3)$ π.χ. με τον αριθμό 0.25. Η διαδικασία αυτή επαναλαμβάνεται μέχρι και την κωδικοποίηση του τελευταίου συμβόλου του μηνύματος.

Πρέπει να σημειώσουμε ότι με την κωδικοποίηση κάθε νέου συμβόλου, το διάστημα, από το οποίο μπορεί να πάρει τιμές ο δεκαδικός αριθμός (κωδικό μήνυμα), συνεχώς συρρικνώνεται. Επίσης, κάθε φορά θα πρέπει το προηγούμενο διάστημα να κατακερματίζεται σε τμήματα ανάλογα των πιθανοτήτων των συμβόλων της πηγής.

Η όλη διαδικασία μπορεί να παρασταθεί με τη βοήθεια ενός διαγράμματος. Έστω, ότι θέλουμε να κωδικοποιήσουμε το μήνυμα

$$[b\,a\,b\,c\,a].$$

Ο επαναληπτικός κατακερματισμός των διαστημάτων που γίνεται κατά την αριθμητική συμπίεση φαίνεται στο σχήμα 5.4.



Σχήμα 5.4: *Παράδειγμα διαγράμματος αριθμητικής συμπίεσης.*

Το μήνυμα $[b\,a\,b\,c\,a]$ μπορεί να κωδικοποιηθεί ως ένας οποιοσδήποτε δεκαδικός αριθμός που ανήκει στο διάστημα $[0.255, 0.258)$, π.χ. ο αριθμός $d = 0.257$. Η δυαδική μορφή του αριθμού d υπολογίζεται με τη γνωστή μέθοδο των επαναληπτικών πολλαπλασιασμών με το 2 και είναι

0.010000011100 ...

Από την παραπάνω δυαδική αναπαράσταση του d κρατάμε μόνο τα k πρώτα ψηφία μετά την υποδιαστολή, τα οποία αρκούν για να προσδιοριστεί το διάστημα $[0.255, 0.258)$. Το πλήθος των ψηφίων αυτών αποδεικνύεται ότι είναι ίσο με

$$k = \lceil -\log(s) \rceil + 1 \tag{5.7}$$

όπου το s δίνεται από τη σχέση 5.6) και είναι ίσο με το άνω μείον το κάτω φράγμα του διαστήματος που αντιστοιχεί στο μήνυμα. Στη συγκεκριμένη περίπτωση

$$s = 0.258 - 0.255 = 0.003$$
$$k = \lceil -\log(s) \rceil + 1 = \lceil 8.381 \rceil + 1 = 10.$$

Συνεπώς, το μήνυμα $[b\,a\,b\,c\,a]$ κωδικοποιείται με την ακολουθία δυαδικών ψηφίων

0100000111.

Δηλαδή, λαμβάνουν χώρα οι παρακάτω απεικονίσεις

$$[b\,a\,b\,c\,a] \rightarrow [0.255, 0.258) \rightarrow 2.57 \rightarrow 0.010000011100\ldots \rightarrow 0100000111.$$

Αλγόριθμος υπολογισμού του διαστήματος που απεικονίζεται στο μήνυμα

Όπως αναφέρθηκε προηγουμένως σε κάθε σύμβολο x της πηγής αντιστοιχεί ένα τμήμα του διαστήματος $[0, 1]$. Το διάστημα αυτό περιγράφεται πλήρως από το κάτω όριο, low_x, και από το εύρος του διαστήματος, το οποίο είναι ίσο με την πιθανότητα εμφάνισης του συμβόλου, p_x. Προφανώς, το άνω όριο του διαστήματος, $high_x$, θα δίνεται από το άθροισμα

$$high_x = low_x + p_x. \tag{5.8}$$

Μετά την αριθμητική κωδικοποίηση το συνολικό μήνυμα θα απεικονίζεται σε ένα διάστημα με κάτω όριο, LOW, και άνω όριο, $HIGH$. Ο αλγόριθμος εύρεσης του τελικού διαστήματος $[LOW, HIGH]$ έχει την εξής μορφή ψευδοκώδικα:

$LOW = 0.0$
$HIGH = 1.0$
$WHILE\,(not\ end\ of\ message)$
    $READ\,(x)$
    $RANGE = HIGH - LOW$
    $LOW = LOW + RANGE * low_x$
    $HIGH = LOW + RANGE * p_x$
$END\ WHILE$
$OUTPUT\,(LOW, HIGH)$

Στο παράδειγμα που αναφέραμε στην αρχή της ενότητας, όπου είχαμε ένα μήνυμα χιλίων ίδιων συμβόλων πηγής με πιθανότητα συμβόλου 0.999, προκύπτει ότι το εύρος του διαστήματος που απεικονίζεται στο συνολικό μήνυμα θα είναι

$$s = 0.999^{1000} = 0.367695424771.$$

Οπότε, για την κωδικοποίηση όλου του μηνύματος θα χρειαστούν μόνο

$$k = \lceil -\log(s) \rceil + 1 = k = \lceil 1.443416869669 \rceil + 1 = 3$$

δυαδικά ψηφία.

5.3.2 Αριθμητική αποσυμπίεση

Για να γίνει κατανοητή η διαδικασία της αριθμητικής αποσυμπίεσης θα θεωρήσουμε ότι ο δέκτης λαμβάνει το κωδικοποιημένο μήνυμα που δημιουργήθηκε στο προηγούμενο παράδειγμα αριθμητικής κωδικοποίησης. Θεωρούμε δε ότι ο αποκωδικοποιητής γνωρίζει τα τμήματα του διαστήματος $[0, 1)$ που έχουν αποδοθεί σε κάθε σύμβολο πηγής, δηλαδή τα ζεύγη $\{low_x, p_x\}$ για κάθε σύμβολο x.

Συγκεκριμένα, ο δέκτης λαμβάνει την ακολουθία δυαδικών ψηφίων

0100000111,

η οποία αναπαριστά τον δυαδικό αριθμό 0.0100000111. Η δεκαδική μορφή του αριθμού αυτού προκύπτει από το άθροισμα των αρνητικών δυνάμεων του 2 και είναι η

$$d = 2^{-2} + 2^{-8} + 2^{-9} + 2^{-10} = 0.2568359375.$$

Ο δεκαδικός αριθμός d ανήκει στο διάστημα $[0.2, 0.7)$ που αντιστοιχεί στο σύμβολο b. Αυτό σημαίνει ότι το πρώτο σύμβολο του αρχικού μηνύματος είναι το b.

Στη συνέχεια, θα πρέπει να αφαιρεθεί η επίδραση της κωδικοποίησης του συμβόλου b από τον αριθμό d. Αυτό γίνεται ακολουθώντας την αντίστροφη διαδικασία από αυτήν της κωδικοποίησης. Πρακτικά, αφαιρούμε το κάτω όριο του διαστήματος του b (δηλαδή το 0.2) από τον αριθμό d και στη συνέχεια διαιρούμε το αποτέλεσμα με την πιθανότητα p_b. Ο αριθμός που προκύπτει θεωρείται ως το νέο d και αποκωδικοποιείται με τον ίδιο τρόπο. Συγκεκριμένα έχουμε

$$d - low_b = 0.2568359375 - 0.2 = 0.0568359375$$

$$d = 0.0568359375/p_b = 0.0568359375/0.5 = 0.113671875$$

Το νέο d ανήκει στο διάστημα $[0, 0.2)$, οπότε το δεύτερο σύμβολο του αρχικού μηνύματος είναι το a. Συνεχίζοντας, έχουμε

$$d - low_a = 0.113671875 - 0 = 0.113671875$$

$$d = 0.113671875/p_a = 0.113671875/0.2 = 0.568359375$$

Τώρα, το νέο d ανήκει στο διάστημα $[0.2, 0.7)$ οπότε το τρίτο σύμβολο του αρχικού μηνύματος είναι το b. Μετά έχουμε,

$$d - low_b = 0.568359375 - 0.2 = 0.368359375$$

$$d = 0.368359375/p_b = 0.368359375/0.5 = 0.73671875$$

Το νέο d ανήκει στο διάστημα $[0.7, 1.0)$. Συνεπώς, το τέταρτο σύμβολο του αρχικού μηνύματος είναι το c. Ακολουθώντας την ίδια λογική,

$$d - low_c = 0.73671875 - 0.7 = 0.03671875$$

$$d = 0.03671875/p_c = 0.03671875/0.3 = 0.1223958333333$$

Το νέο d ανήκει στο διάστημα $[0, 0.2)$. Συνεπώς, το πέμπτο σύμβολο του αρχικού μηνύματος είναι το a.

Άρα, μέχρι αυτό το σημείο έχει αποκωδικοποιηθεί το μήνυμα

$$b\,a\,b\,c\,a$$

δηλαδή, αυτό που πραγματικά κωδικοποιήθηκε. Παρά ταύτα, ο δέκτης μπορεί να συνεχίσει τη διαδικασία αποκωδικοποίησης

$$d - low_a = 0.1223958333333 - 0 = 0.1223958333333$$

$$d = 0.1223958333333/p_a = 0.1223958333333/0.2 = 0.6119791666667$$

και να συμπεράνει ότι το επόμενο σύμβολο του μηνύματος είναι το b. Γενικά, αυτή η διαδικασία μπορεί να συνεχιστεί επ' άπειρον.

Συνεπώς, ο δέκτης θα πρέπει να γνωρίζει εκ' των προτέρων το μέγεθος του αρχικού μηνύματος ή να έχει προβλεφθεί η χρήση ενός συμβόλου πηγής που να δηλώνει το τέλος του μηνύματος.

Αλγόριθμος αριθμητικής αποσυμπίεσης

Κατά την αποσυμπίεση, ο αποκωδικοποιητής υπολογίζει τον αριθμό d από την ακολουθία δυαδικών ψηφίων που έλαβε, και γνωρίζει το πλήθος n των συμβόλων πηγής που περιέχονται εντός του μηνύματος πηγής, καθώς και τα ζεύγη $\{low_x, p_x\}$ για κάθε σύμβολο x. Ο αλγόριθμος αποκωδικοποίησης έχει την εξής μορφή ψευδοκώδικα:

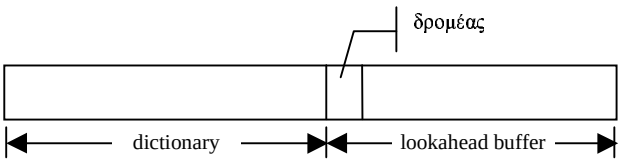
$GET\ (d, n)$
$FOR\ (k = 1\ TO\ n)$
    $RECOGNIZE\,SYMBOL\,(x) \quad IF\ \ d \in [low_x, high_x)$
    $OUTPUT\ (x)$
    $d = d - low_x$
    $d = d/p_x$
$END\ FOR$

5.4 Κώδικες Ταύτισης Συμβολοσειρών (String-Matching Codes)

Ο κώδικας Huffman και ο αριθμητικός κώδικας βασίζονται στη γνώση ή την εκτίμηση των στατιστικών χαρακτηριστικών των συμβόλων πηγής που πρόκειται να συμπιεστούν. Πράγματι, και οι δύο αλγόριθμοι κωδικοποίησης απαιτούν τη γνώση των πιθανοτήτων p_i των συμβόλων. Ένα άλλο χαρακτηριστικό των δύο παραπάνω κωδίκων είναι ότι δεν μπορούν να προσαρμοστούν στις πιθανές μεταβολές των πιθανοτήτων εμφάνισης των συμβόλων της πηγής. Αυτοί οι περιορισμοί μπορούν να ξεπεραστούν με τη χρήση **κωδίκων ταύτισης συμβολοσειρών** (ακολουθίες συμβόλων) (string-matching codes).

Ενώ η Huffman και η αριθμητική κωδικοποίηση απεικονίζουν σύμβολα ή συμβολοσειρές της πηγής σταθερού μήκους σε κωδικές λέξεις μεταβλητού μήκους, οι αλγόριθμοι string-matching απεικονίζουν συμβολοσειρές πηγής μεταβλητού μήκους σε κωδικές λέξεις σταθερού μήκους. Ακόμη, οι αλγόριθμοι (string-matching) δεν κάνουν και δε χρειάζονται καμία αρχική εκτίμηση της στατιστικής συμπεριφοράς των συμβόλων της πηγής, αλλά, κατά την εκτέλεσή τους, προσαρμόζονται διαρκώς στις πιθανές αλλαγές των στατιστικών χαρακτηριστικών του μηνύματος της πηγής.

Η ανάπτυξη των string-matching αλγορίθμων ξεκίνησε με την εμφάνιση δύο πρωτοποριακών εργασιών των ερευνητών Jacob Ziv και Abraham Lempel. Το 1977, με την πρώτη τους εργασία, οι Ziv και Lempel περιέγραψαν μία τεχνική συμπίεσης που βασίζεται στη χρήση ενός κυλιώμενου παραθύρου, το οποίο διαπερνά το συνολικό μήνυμα και συγκρίνει νέες συμβολοσειρές με άλλες που έχουν εμφανιστεί στο κοντινό παρελθόν. Ο αλγόριθμος αυτός που θα περιγραφεί αργότερα με περισσότερες λεπτομέρειες αναφέρεται γενικά ως LZ77. Μία μορφή αυτού του αλγορίθμου χρησιμοποιείται σε ορισμένα υπολογιστικά πακέτα συμπίεσης τύπου zip π.χ. στα PKZIP, gzip, zipit. Την επόμενη χρονιά (1978) οι ίδιοι ερευνητές πρότειναν μία



Σχήμα 5.5: *Το κυλιώμενο παράθυρο του αλγορίθμου LZ77.*

εξέλιξη του LZ77, η οποία βασίζεται στη δημιουργία ενός λεξικού με δομή δέντρου. Ο αλγόριθμος αυτός αναφέρεται ως LZ78. Το 1984 ο Terry Welch πρότεινε ένα νέο αλγόριθμο που αποτελεί βελτίωση του LZ78 και αναφέρεται ως LZW. Η επιτυχία των αλγορίθμων LZ78 και LZW είναι φανερή από το γεγονός ότι έχουν εφαρμοστεί στο πρωτόκολλο ITU-T V.42 bis, στη συμπίεση εικόνας με μορφή GIF, και στο πρόγραμμα compress του UNIX.

Όλοι αυτοί οι αλγόριθμοι (LZ77, LZ78, LZW) βασίζονται στην υπόθεση ότι κάποιες συμβολοσειρές είναι πιθανόν να εμφανίζονται συχνά μέσα στο συνολικό μήνυμα. Έτσι, όταν μία συμβολοσειρά επανεμφανίζεται μπορεί να αντικατασταθεί με ένα μικρό κώδικα, ο οποίος δηλώνει πότε ή που εμφανίστηκε στο παρελθόν η συμβολοσειρά αυτή. Αυτό σημαίνει ότι κατά τη διάρκεια εκτέλεσης του αλγορίθμου συμπίεσης αναζητούμε επανεμφανιζόμενες συμβολοσειρές. Σημειώνουμε ότι όσο μεγαλύτερη είναι η δυνατότητα πρόβλεψης των συμβόλων της πηγής τόσο μικρότερη είναι η εντροπία της πηγής. Συνεπώς, με τους αλγόριθμους string-matching είναι δυνατό να πετύχουμε σημαντική συμπίεση.

5.5 Ο Αλγόριθμος LZ77

Ο αλγόριθμος LZ77 και οι παραλλαγές του χρησιμοποιούν ένα κυλιώμενο παράθυρο. Το παράθυρο αυτό έχει ένα **δρομέα** (cursor) και χωρίζεται σε δύο τμήματα, το ένα πριν τον δρομέα (αριστερά του δρομέα) που καλείται **dictionary**, και το άλλο μετά το δρομέα (δεξιά του δρομέα) που καλείται **lookahead buffer**. Ο δρομέας περιέχεται στο lookahead buffer (βλέπε σχήμα 5.5). Τα μεγέθη των δύο αυτών τμημάτων του παραθύρου είναι παράμετροι που καθορίζονται από το χρήστη με την έναρξη του αλγορίθμου και παραμένουν στη συνέχεια αμετάβλητα.

Ο αλγόριθμος ακολουθεί τα παρακάτω βήματα:

1. Βρίσκουμε το μεγαλύτερο τμήμα της συμβολοσειράς που αρχίζει από τον δρομέα, περιέχεται πλήρως εντός του lookahead buffer, και ταυτίζεται με συμβολοσειρά που ξεκινά από κάποιο σημείο του dictionary. Είναι πιθανόν, μήκος

της δεύτερης συμβολοσειράς να εισέρχεται στο τμήμα του lookahead buffer.

2. Καταγράφουμε την τριάδα (p, n, c) όπου:

   - p είναι ο αριθμός που δηλώνει το σημείο εντός του dictionary από το οποίο αρχίζει η δεύτερη συμβολοσειρά. Το p μετράται ξεκινώντας από το πρώτο σύμβολο πριν το δρομέα (τιμή 1) και προς τα αριστερά.
   - n είναι το πλήθος το συμβόλων της συμβολοσειράς, η οποία ανιχνεύθηκε ως επανεμφανιζόμενη.
   - c είναι το επόμενο σύμβολο μετά την επανεμφανιζόμενη συμβολοσειρά και ο οποίος δεν ταυτίστηκε.

3. Εφόσον δεν έχει περατωθεί η συμπίεση όλου του μηνύματος, μετακινούμε το δρομέα κατά $n + 1$ θέσεις δεξιότερα, δηλαδή αμέσως μετά από το σύμβολο c, και μεταβαίνουμε στο πρώτο βήμα του αλγορίθμου.

Για να κατανοηθεί καλύτερα η διαδικασία της συμπίεσης με τον αλγόριθμο LZ77 εξετάζουμε τη διαδικασία συμπίεσης του μηνύματος

[α α ε α α ε α ν ε α ν α α α ε ς],

όπου το σύμβολο ς δηλώνει το τέλος του μηνύματος και συνεπώς εμφανίζεται μόνο μία φορά εντός του μηνύματος.

Στο παράδειγμα θεωρούμε ότι το μήκος του dictionary είναι 6 ενώ του lookahead buffer είναι 4. Ο αλγόριθμος σε αυτήν την περίπτωση δηλώνεται με τη γραφή LZ77(6,4). Οι επαναλήψεις του αλγορίθμου παρουσιάζονται στον πίνακα 5.2. Σε κάθε επανάληψη, τα σύμβολα που ανήκουν στο τμήμα του lookahead buffer είναι υπογραμμισμένα, ενώ αυτά που ανήκουν στο τμήμα του dictionary είναι έντονα. Το σύμβολο που αντιστοιχεί στη θέση του δρομέα είναι πλάγιο και ταυτόχρονα υπογραμμισμένο, αφού ανήκει και αυτό στο lookahead buffer. Στην τελευταία στήλη του πίνακα γράφονται οι τριάδες του κώδικα με τη μορφή (p, n, c).

Ένα σημείο που χρειάζεται προσοχή είναι η περίπτωση της τρίτης επανάληψης, όπου έχουμε $n > p$. Αυτό συμβαίνει γιατί η επαναλαμβανόμενη συμβολοσειρά εκτείνεται και εντός του lookahead buffer. Επίσης, στην πέμπτη επανάληψη παρατηρούμε ότι ο αλγόριθμος δεν εντόπισε τη μεγαλύτερη σε μήκος ταύτιση (10,3,ς), επειδή βρίσκεται εκτός του κυλιώμενου παραθύρου. Τέλος, το συνολικό πλήθος των συμβόλων που αρχικού μηνύματος προκύπτει από το άθροισμα όλων των πεδίων n συν τον αριθμό των επαναλήψεων του αλγορίθμου (0+1+4+3+2+0+6=16).

Πίνακας 5.2: *Παράδειγμα εφαρμογής του αλγορίθμου LZ77(6,4).*

Επανάληψη	Μήνυμα																Κώδικας
1	*α*	α	ε	α	α	ε	α	ν	ε	α	ν	α	α	α	ε	ς	(0,0,α)
2	**α**	*α*	ε	α	α	ε	α	ν	ε	α	ν	α	α	α	ε	ς	(1,1,ε)
3	**α**	**α**	**ε**	*α*	α	ε	α	ν	ε	α	ν	α	α	α	ε	ς	(3,4,ν)
4	α	α	**ε**	**α**	**α**	**ε**	**α**	**ν**	ϵ	α	ν	α	α	α	ε	ς	(3,3,α)
5	α	α	ε	α	α	ε	**α**	**ν**	**ε**	**α**	**ν**	**α**	*α*	α	ε	ς	(1,2,ε)
6	α	α	ε	α	α	ε	α	ν	ε	**α**	**ν**	**α**	**α**	**α**	**ε**	ς	(0,0,ς)

Η διαδικασία της αποκωδικοποίησης έχει ως εξής:

Όταν λάβουμε την τριάδα (p, n, c) μεταβαίνουμε p προς τα πίσω στο ήδη αποκωδικοποιήμενο μήνυμα και γράφουμε τα n σύμβολα που ακολουθούν. Στο τέλος, καταγράφουμε το σύμβολο c. Για παράδειγμα, θεωρούμε ότι έχουμε ήδη αποκωδικοποιήσει τις τριάδες

(0,0,α), (1,1,ε)

οπότε το μέχρι στιγμής αποκωδικοποιημένο μήνυμα είναι το

[α α ε].

Αν τώρα λάβουμε την τριάδα

(3,4,ν)

μεταβαίνουμε τρεις θέσεις προς τα πίσω (αρχικά ο δρομέας βρίσκεται αμέσως μετά το σύμβολο ε) και καταγράφουμε τα τέσσερα σύμβολα που συναντάμε ακολουθιακά προς τα δεξιά. Έτσι, προκύπτει το μήνυμα

[α α ε α α ε α].

Στο τέλος, καταγράφουμε και το σύμβολο ν της τριάδας που λάβαμε. Συνεπώς, το αποκωδικοποιημένο μήνυμα μετά από τη λήψη και της τρίτης τριάδας είναι

[α α ε α α ε α ν].

Είναι φανερό ότι η επιτυχία της αποκωδικοποίησης μίας τριάδας απαιτεί την επιτυχή αποκωδικοποίηση όλων των προηγουμένων.

5.5.1 Δύο βελτιώσεις του LZ77

Μία βελτίωση του LZ77 είναι η λεγόμενη LZSS Variant. Σε αυτόν τον αλγόριθμο δεν περιλαμβάνεται στην τριάδα το επόμενο σύμβολο c. Συγκεκριμένα, ο αλγόριθμος χρησιμοποιεί δύο μορφές κωδικοποίησης, οι οποίες διακρίνονται μεταξύ τους με τη χρήση ενός πρόσθετου bit. Στην πρώτη μορφή καταγράφονται μόνο ζεύγη της μορφή (p, n), τα οποία προκύπτουν όπως και στην περίπτωση του LZ77. Στη δεύτερη μορφή καταγράφεται αυτούσιο το σύμβολο c. Η πρώτη μορφή εφαρμόζεται όταν το μήκος της ταυτιζόμενης συμβολοσειράς είναι τουλάχιστον ίσο με 3. Αν αυτό δεν ισχύει, εφαρμόζεται η δεύτερη μορφή.

Μία δεύτερη παραλλαγή του LZ77 σχετίζεται με τη χρήση του κώδικα Huffman. Μετά την εφαρμογή του LZ77, οι ακολουθίες των πεδίων p, n και c συμπιέζονται χωριστά η κάθε μία με κώδικα Huffman. Αυτή η μορφή συμπίεσης έχει υλοποιηθεί στο πρόγραμμα Gzip.

Παράρτημα Α΄

Συνάρτηση του Μέτρου της Πληροφορίας

Στην ενότητα 2.2 αναφέρθηκε ότι η συνάρτηση $I(A)$ του μέτρου της πληροφορίας ενός γεγονότος A θα πρέπει να ικανοποιεί τέσσερις ιδιότητες για να είναι αποδεκτή. Θυμίζουμε ότι η $I(A)$ θα πρέπει να είναι συνάρτηση της πιθανότητας $p(A)$ του γεγονότος (2.1), να είναι πραγματική και θετική συνάρτηση (2.2), καθώς και γνησίως φθίνουσα (2.3). Τέλος, για ανεξάρτητα γεγονότα A και B θα πρέπει να ισχύει η (2.4), δηλαδή

$$I(A \cap B) = I(p(A)p(B)) = I(p(A)) + I(p(B)) = I(A) + I(B). \tag{Α΄.1}$$

Θα επιχειρήσουμε να βρούμε τη συνάρτηση που να ικανοποιεί τις παραπάνω ιδιότητες. Έστω, η συνεχής συνάρτηση

$$f(x) : [0, 1] \mapsto \Re_+ \tag{Α΄.2}$$

Η συνάρτηση (Α΄.2) έχει εξ' ορισμού τις τρεις πρώτες ιδιότητες (2.1), (2.2) και (2.3).

Θεωρούμε ότι η $I(x)$ είναι παραγωγίσιμη, οπότε για κάθε $x \in (0, 1)$ έχουμε

$$f'(x) = \lim_{\delta \to 0} \frac{f(x+\delta) - f(x)}{\delta} =$$
$$= \lim_{\delta \to 0} \frac{f\left[\frac{x}{m} \cdot \left(m + \frac{m\delta}{x}\right)\right] - f(\frac{x}{m} \cdot m)}{\delta} =$$

$$= \lim_{\delta \to 0} \frac{f\left(\frac{x}{m}\right) + f\left(m + \frac{m\delta}{x}\right) - f\left(\frac{x}{m}\right) - f(m)}{\delta} =$$

$$= \frac{m}{x} \lim_{m\delta/x \to 0} \frac{f\left(m + \frac{m\delta}{x}\right) - f(m)}{\frac{m\delta}{x}} =$$

$$= \frac{m}{x} f'(m) = \frac{c}{x} \tag{A′.3}$$

όπου m είναι ένα αυθαίρετα επιλεγμένο σημείο στο διάστημα $(0, 1)$ και c σταθερά. Στην παραπάνω ανάλυση έγινε χρήση της ιδιότητας (A′.1), δηλαδή $f(ab) = f(a) + f(b)$. Για να είναι η $f(x)$ γνησίως φθίνουσα θα πρέπει σύμφωνα με το αποτέλεσμα της (A′.3) να ισχύει $c < 0$. Επίσης, από την (A′.3) προκύπτει ότι

$$f(x) = c \ln(x), \quad c < 0 \tag{A′.4}$$

Αν ορίσουμε τη σταθερά $K = e^{-1/c}$, τότε η (A′.4) μπορεί να γραφεί με τη μορφή

$$f(x) = -\frac{\ln(x)}{\ln(K)} = -\log_K(x). \tag{A′.5}$$

Σημειώνουμε ότι επειδή $c < 0$, θα έχουμε $K > 1$. Έτσι, καταλήξαμε στον προσδιορισμό της συνάρτησης του μέτρου πληροφορίας

$$I(A) = f(x) = -\log_K(p(A)), \quad K > 1. \tag{A′.6}$$

Παράρτημα Β΄

Τιμές της συνάρτησης Shannon

Στον παρακάτω πίνακα δίνονται οι τιμές της συνάρτησης του Shannon (2.13) για διάφορες τιμές της πιθανότητας p. Προφανώς, για $p > 0.5$ εφαρμόζουμε την ιδιότητα συμμετρίας $H_b(p) = H_b(1-p)$.

Πίνακας Β΄.1: *Τιμές της συνάρτησης Shannon.*

p	$H_b(p)$	p	$H_b(p)$	p	$H_b(p)$	p	$H_b(p)$
0.00	0.00000	0.13	0.55744	0.26	0.82675	0.39	0.96480
0.01	0.08079	0.14	0.58424	0.27	0.84146	0.40	0.97095
0.02	0.14144	0.15	0.60984	0.28	0.85545	0.41	0.97650
0.03	0.19439	0.16	0.63431	0.29	0.86872	0.42	0.98145
0.04	0.24229	0.17	0.65770	0.30	0.88129	0.43	0.98582
0.05	0.28640	0.18	0.68008	0.31	0.89317	0.44	0.98959
0.06	0.32744	0.19	0.70147	0.32	0.90438	0.45	0.99277
0.07	0.36592	0.20	0.72193	0.33	0.91493	0.46	0.99538
0.08	0.40218	0.21	0.74148	0.34	0.92482	0.47	0.99740
0.09	0.43647	0.22	0.76017	0.35	0.93407	0.48	0.99885
0.10	0.46900	0.23	0.77801	0.36	0.94268	0.49	0.99971
0.11	0.49992	0.24	0.79504	0.37	0.95067	0.50	1.00000
0.12	0.52936	0.25	0.81128	0.38	0.95804		

Παράρτημα Γ′

Κωδικές Λέξεις Βέλτιστου Στιγμιαία Αποκωδικοποιήσιμου Κώδικα

Θεωρούμε ένα M-αδικό στιγμιαία αποκωδικοποιήσιμο κώδικα με N κωδικές λέξεις. Τα μήκη των κωδικών λέξεων l_i $(i = 1, 2, \ldots, N)$ θα πρέπει να ικανοποιούν την ταυτοανισότητα του Kraft

$$\sum_{i=1}^{N} M^{-l_i} - 1 \leq 0, \tag{Γ′.1}$$

ενώ το μέσο μήκος του κώδικα θα είναι

$$\bar{L} = \sum_{i=1}^{N} p_i \cdot l_i. \tag{Γ′.2}$$

Για να προσδιορίσουμε τα μήκη των κωδικών λέξεων του βέλτιστου κώδικα θα πρέπει να λύσουμε το πρόβλημα ελαχιστοποίησης της (Γ′.2), υπό τον περιορισμό της ανισότητας (Γ′.1). Σύμφωνα με τη συνθήκη Kuhn-Tucker, τα βέλτιστα μήκη των κωδικών λέξεων l_i^* $(i = 1, 2, \ldots, N)$ θα δίνονται από την επίλυση του συστήματος των εξισώσεων

$$\frac{\partial}{\partial l_i}\bar{L} + \mu \frac{\partial}{\partial l_i}\left[\sum_{i=1}^{N} M^{-l_i} - 1\right] = 0, \tag{Γ′.3}$$

$$\sum_{i=1}^{N} M^{-l_i} - 1 = 0, \tag{Γ′.4}$$

όπου μ είναι πολλαπλασιαστής Lagrange.

Από την (Γ′.3) έχουμε

$$p_i + \mu[-M^{-l_i}\ln(M)] = 0 \Rightarrow$$
$$M^{-l_i} = [\mu \ln(M)]^{-1} p_i. \tag{Γ′.5}$$

Με αντικατάσταση της (Γ′.5) στην (Γ′.4) προκύπτει

$$[\mu \ln(M)]^{-1} \sum_{i=1}^{N} p_i - 1 = 0 \Rightarrow$$
$$[\mu \ln(M)]^{-1} = 1. \tag{Γ′.6}$$

Συνεπώς, αντικαθιστώντας την (Γ′.6) στην (Γ′.5) και επιλύοντάς τη δεύτερη ως προς l_i προκύπτουν τα μήκη του βέλτιστου κώδικα

$$l_i^* = -\log_M p_i. \tag{Γ′.7}$$

Παράρτημα Δ΄

Βιβλιογραφία

1. G.A. Jones and J.M. Jones, *Information and Coding Theory*, Springer - Verlag, London, 2000.

2. C.E. Shannon, A Mathematical Theory of Communication, *Bell Systems Tech. Journal*, vol. 27, pp. 379-423, 1948.

3. A.I. Khinchin, *Mathematical Foundations of Information Theory*, Dover, New York, 1957.

4. R. Ash, *Information Theory*, Dover, New York, 1990.

5. T.M. Cover and J.A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

6. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, Mc Graw-Hill, New York, 1991.

7. R.G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

8. D. Shepian (Ed.), *Key Papers in the Development of Information Theory*, IEEE Press, New York, 1974.

9. R. Blahut, *Principles and Practice of Information Theory*, Addison - Wesley, Reading MA, 1987.

10. R.J. McEliece, *The Theory of Information and Coding*, Addison - Wesley, Reading MA, 1977.

11. Δ.Π. Χρυσουλίδης, *Εισαγωγή στη Θεωρία Πληροφοριών*, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 1991.

12. Α. Κανλής, *Θεωρία της Πληροφορίας*, Τεχνολογικό Εκπαιδευτικό Ίδρυμα Σερρών, 2001.