ΕΕ728
Προχωρημένα Θέματα Θεωρίας Πληροφορίας
5η διάλεξη
(2η έκδοση, 22/3/2011)

Δημήτρης-Αλέξανδρος Τουμπακάρης

Τμήμα ΗΜ&ΤΥ, Πανεπιστήμιο Πατρών

22 Μαρτίου 2011

Περιεχόμενα 5ης διάλεξης

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή

- Το Θεώρημα Κωδικοποίησης Καναλιού (Channel Coding Theorem) αποτελεί το πιο βασικό και το πιο διάσημο αποτέλεσμα της Θεωρίας Πληροφορίας.
- Σύμφωνα με το Θεώρημα Κωδικοποίησης Καναλιού, είναι εφικτή η μετάδοση σε κανάλια χωρίς μνήμη με ρυθμό αυθαίρετα κοντά στη χωρητικότητα και με αυθαίρετα μικρή πιθανότητα σφάλματος. Αντίστροφα, δεν είναι εφικτή μετάδοση με αυθαίρετα μικρή πιθανότητα σφάλματος εάν ο ρυθμός μετάδοσης υπερβαίνει τη χωρητικότητα του καναλιού.

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή (2)

- Στη συνέχεια, θα διατυπώσουμε με την απαραίτητη λεπτομέρεια και θα αποδείξουμε το Θεώρημα Κωδικοποίησης Καναλιού.
- Το Θεώρημα Κωδικοποίησης Καναλιού (ευθύ - achievability) μπορεί να αποδειχτεί είτε με χρήση αποκωδικοποίησης Μέγιστης Πιθανοφάνειας (Maximum Likelihood decoding -- Gallager) ή με χρήση Από Κοινού Τυπικών ακολουθιών (Cover).
- Στο μάθημα θα εξετάσουμε την απόδειξη με χρήση Από Κοινού Τυπικότητας η οποία είναι μάλλον πιο απλή.
- Το (ασθενές) αντίστροφο του Θεωρήματος Κωδικοποίησης Καναλιού θα αποδειχτεί με χρήση της ανισότητας Fano.

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή (3)

- Το βασικό ερώτημα (και, εκ πρώτης όψεως, παράδοξο) είναι το εξής: Πώς είναι δυνατόν να μεταδώσουμε με αυθαίρετα μικρή πιθανότητα σφάλματος σε ένα κανάλι που εισάγει σφάλματα με μη μηδενική πιθανότητα και με τυχαίο τρόπο;
- Για να απαντήσει στο ερώτημα, ο Shannon χρησιμοποίησε ένα διαφορετικό τρόπο σκέψης:
  - Δεν προσπάθησε να εκμηδενίσει την πιθανότητα σφάλματος, απλώς να την περιορίσει σε αυθαίρετα μικρές τιμές.
  - Βασίστηκε σε πολλές διαδοχικές χρήσεις του καναλιού ώστε να εκμεταλλευτεί το Νόμο των Μεγάλων Αριθμών.
  - Χρησιμοποίησε κώδικες οι οποίοι δημιουργούνται τυχαία και υπολόγισε τη μέση πιθανότητα σφάλματος.
- Αυτός ο τρόπος σκέψης διέπει τόσο την απόδειξη με χρήση τυπικότητας όσο και την απόδειξη με αποκωδικοποίηση Μέγιστης Πιθανοφάνειας.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί

Ορισμός 5.1. Ένας κώδικας (M, n) για το Διακριτό Κανάλι Χωρίς Μνήμη $(\mathcal{X}, p(y|x), \mathcal{Y})$ αποτελείται από

1. Ένα σύνολο δεικτών $\{1, 2, \ldots, M\}$.
2. Μια συνάρτηση κωδικοποίησης $X^n : \{1, 2, \ldots, M\} \to \mathcal{X}^n$ η οποία παράγει κωδικές λέξεις (codewords) $x^n(1), x^n(2), \ldots, x^n(M)$. Το σύνολο των κωδικών λέξεων ονομάζεται βιβλίο κωδίκων (codebook).
3. Μια συνάρτηση αποκωδικοποίησης $g : \mathcal{Y}^n \to \{1, 2, \ldots, M\}$, η οποία αποτελεί ένα νομοτελειακό κανόνα ο οποίος αντιστοιχίζει ένα εκτιμώμενο δείκτη μεταδοθέντος μηνύματος, $\hat{m}$, σε κάθε ληφθείσα ακολουθία y^n.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (2)

- Υπό συνθήκη πιθανότητα σφάλματος δεδομένου ότι εστάλη το μήνυμα με δείκτη i:

$$\lambda_i = \Pr\left\{g(Y^n) \neq i | X^n = x^n(i)\right\} = \sum_{y^n} p(y^n|x^n(i)) I(g(y^n) \neq i),$$

όπου $I(\cdot)$ η συνάρτηση δείκτης (ισούται με 1 όταν το όρισμά της αληθεύει, αλλιώς με 0).

- Η Μέγιστη Πιθανότητα Σφάλματος $\lambda^{(n)}$ κώδικα (M, n) ορίζεται ως

$$\lambda^{(n)} = \max_{i \in \{1,2,\ldots,M\}} \lambda_i .$$

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (3)

- Η (αριθμητικώς) μέση πιθανότητα σφάλματος $P_e^{(n)}$ κώδικα (M, n) ισούται με

$$P_e^{(n)} = \frac{1}{M} \sum_{i=1}^{M} \lambda_i.$$

  - Όταν ο δείκτης μηνύματος W ακολουθεί ομοιόμορφη κατανομή, $P_e^{(n)} = \Pr\{W \neq g(Y^n)\}$, όπου Y^n η ακολουθία που λαμβάνεται στην έξοδο καναλιού όπου έχει μεταδοθεί η ακολουθία $X^n = x^n(W)$.
  - Επίσης, $P_e^{(n)} \leq \lambda^{(n)}$.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (4)

- **Ορισμός 5.2.** Ο ρυθμός (rate) R κώδικα (M, n) ισούται με

$$R = \frac{\log M}{n} \text{ bits ανά μετάδοση.}$$

- **Ορισμός 5.3.** Ένας ρυθμός R είναι εφικτός (achievable) όταν υπάρχει ακολουθία κωδίκων $\left(\lceil 2^{nR} \rceil, n\right)$ για την οποία η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)}$ τείνει στο 0 καθώς το n τείνει στο άπειρο.[1]

- **Ορισμός 5.4.** Η Χωρητικότητα λειτουργίας (operational capacity) ενός καναλιού ισούται με το μέγιστο ρυθμό ο οποίος είναι εφικτός.

- Το Θεώρημα Κωδικοποίησης Πηγής αποδεικνύει ότι η χωρητικότητα λειτουργίας $\max_{R \text{ εφικτός}} R$ ισούται με την πληροφοριακή χωρητικότητα $\max_{p(x)} I(X;Y)$.

[1] Συχνά η επιτευξιμότητα ορίζεται με χρήση της μέσης πιθανότητας σφάλματος, $P_e^{(n)}$. Οι δύο ορισμοί δεν είναι ισοδύναμοι.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού – Εισαγωγή

- Θα αναφερθούμε στην απόδειξη η οποία χρησιμοποιεί την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP).
- Η ιδέα:
  - Δεδομένου του μηνύματος W, στέλνουμε στο κανάλι ακολουθία $X^n = x^n(W)$ μήκους n.
  - Στην έξοδο του καναλιού λαμβάνουμε ακολουθία Y^n η οποία εξαρτάται από τη X^n, καθώς και από τον πίνακα μετάβασης, $p(y|x)$, του καναλιού.
  - Στο δέκτη αναζητούμε ακολουθία $\hat{X}^n$ η οποία να είναι *από κοινού τυπική* με την Y^n. Εάν υπάρχει, ο δέκτης θεωρεί ότι η $\hat{X}^n$ είναι η ακολουθία που μετέδωσε ο πομπός.
  - Από την Ιδιότητα από κοινού Ασυμπτωτικής Ισοδιαμέρισης, με μεγάλη πιθανότητα η ληφθείσα ακολουθία θα είναι από κοινού τυπική με τη μεταδοθείσα.
  - Ωστόσο, υπάρχει η πιθανότητα η Y^n να μην είναι από κοινού τυπική με καμία από τις πιθανές κωδικές λέξεις X^n ή να είναι από κοινού τυπική με άλλη ακολουθία από αυτή που μεταδόθηκε. Στην περίπτωση αυτή εμφανίζεται σφάλμα μετάδοσης.
  - Θα αποδείξουμε ότι, εάν $R < C$, καθώς το n τείνει στο άπειρο, η πιθανότητα σφάλματος τείνει στο 0.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού

- Θεώρημα 5.5. (Θεώρημα Κωδικοποίησης Καναλιού)
  - Σε ένα Διακριτό Κανάλι Χωρίς Μνήμη, όλοι οι ρυθμοί οι οποίοι είναι μικρότεροι από την πληροφοριακή χωρητικότητα είναι εφικτοί. Δηλαδή, για κάθε ρυθμό $R < C$, υπάρχει ακολουθία κωδίκων $(\lceil 2^{nR} \rceil, n)$ με μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \to 0$.
  - Αντίστροφα, για οποιαδήποτε ακολουθία από κώδικες $(\lceil 2^{nR} \rceil, n)$ με $\lambda^{(n)} \to 0$ πρέπει να ισχύει $R \leq C$.
- Απόδειξη (ευθύ).

  Για απλοποίηση και χωρίς απώλεια της γενικότητας υποθέτουμε ότι ο αριθμός κωδικών λέξεων $\lceil 2^{nR} \rceil$ είναι ακέραιος.

  Θεωρούμε δεδομένη πιθανότητα συμβόλων των κωδικών λέξεων $p(x)$ και δημιουργούμε 2^{nR} τυχαίες κωδικές λέξεις x^n μήκους n θεωρώντας ανεξάρτητες ομοίως κατανεμημένες (i.i.d.) τ.μ. x_i. Η πιθανότητα να δημιουργήσουμε μια συγκεκριμένη κωδική λέξη (ακολουθία) x^n ισούται με $p(x^n) = \prod_{i=1}^{n} p(x_i)$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (2)

- Οι 2^{nR} κωδικές λέξεις αποτελούν τις γραμμές του πίνακα

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ x_1(2) & x_2(2) & \dots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(2^{nR}) & x_2(2^{nR}) & \dots & x_n(2^{nR}) \end{bmatrix}$$

- Η πιθανότητα να δημιουργηθεί ένας <u>συγκεκριμένος</u> <u>τυχαίος</u> κώδικας (πίνακας) $\mathcal{C}$ ισούται με $\Pr(\mathcal{C}) = \prod_{w=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(w))$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (3)

- Θεωρούμε την παρακάτω ακολουθία βημάτων
  1. Δημιουργείται ένας *τυχαίος* κώδικας $\mathcal{C}$ σύμφωνα με την κατανομή $p(x)$ όπως περιγράφτηκε παραπάνω.
  2. Ο κώδικας ανακοινώνεται στον πομπό και στο δέκτη. Επίσης, τόσο ο πομπός όσο και ο δέκτης γνωρίζουν τον πίνακα μετάβασης του καναλιού, $p(y|x)$.
  3. Επιλέγεται ένα μήνυμα W σύμφωνα με ομοιόμορφη κατανομή $\Pr\{W = w\} = 2^{-nR}, w = 1, 2, \ldots, 2^{nR}$.
  4. Στέλνεται στο κανάλι η $w-$οστή κωδική λέξη $X^n(w)$ η οποία αντιστοιχεί στη w-οστή γραμμή του πίνακα $\mathcal{C}$.
  5. Ο δέκτης λαμβάνει ακολουθία Y^n με υπό συνθήκη κατανομή $p(y^n|x^n(w)) = \prod_{i=1}^{n} p(y_i|x_i(w))$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (4)

6. Ο δέκτης εκτιμά ποιο μήνυμα έχει σταλεί. Ο βέλτιστος δέκτης χρησιμοποιεί ανίχνευση Μέγιστης Πιθανοφάνειας (δεδομένου ότι θεωρούμε ομοιόμορφη κατανομή μηνυμάτων). Ωστόσο, όπως αναφέρθηκε, για την απόδειξη θα θεωρήσουμε ανίχνευση με βάση την από κοινού τυπικότητα. Παρόλο που ο δέκτης αυτός δεν είναι βέλτιστος, θα αποδείξουμε ότι, και σε αυτήν την περίπτωση, $\lambda^{(n)} \to 0$ για $n \to \infty$ (ο δέκτης είναι *ασυμπτωτικά* βέλτιστος). Ο δέκτης αποφασίζει (εκτιμά) ότι εστάλη το μήνυμα $\hat{W}$ εάν ικανοποιούνται ταυτόχρονα οι εξής δύο συνθήκες:
   α. Το ζεύγος ακολουθιών $(X^n(\hat{W}), Y^n)$ είναι από κοινού τυπικό.
   β. Δεν υπάρχει άλλος δείκτης μηνύματος $W' \neq \hat{W}$ για τον οποίο να ισχύει $(X^n(W'), Y^n) \in A_\epsilon^{(n)}$. Δηλαδή, δεν υπάρχει ακολουθία $X^n(W')$ που αντιστοιχεί σε μήνυμα $X^n(W') \neq \hat{W}$ (δηλαδή ανήκει στο βιβλίο κωδίκων) η οποία να είναι από κοινού τυπική με την Y^n.
7. Εάν $\hat{W} \neq W$, εμφανίζεται σφάλμα ανίχνευσης. Έστω $\mathcal{E}$ το ενδεχόμενο $\{\hat{W} \neq W\}$.

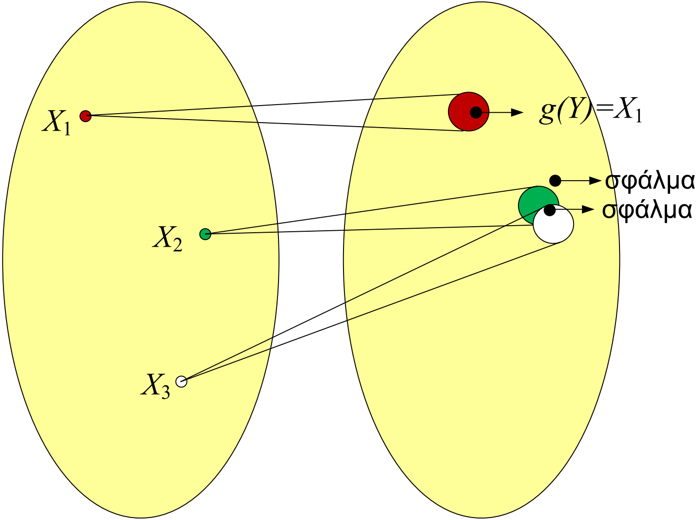
Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (5)
Ανάλυση της πιθανότητας σφάλματος – Εισαγωγή

- Η ιδέα: Αντί να υπολογίσουμε την πιθανότητα σφάλματος για ένα συγκεκριμένο κώδικα, θα υπολογίσουμε τη μέση πιθανότητα σφάλματος για τυχαία δημιουργία κωδίκων και για όλους τους πιθανούς κώδικες.
- Όταν χρησιμοποιείται αποκωδικοποίηση με χρήση από κοινού τυπικότητας, υπάρχουν δύο πηγές σφάλματος: Είτε η έξοδος Y^n δεν είναι από κοινού τυπική με την ακολουθία που εκπέμπει ο πομπός ή υπάρχει τουλάχιστον μια ακόμα κωδική λέξη η οποία είναι από κοινού τυπική με την Y^n.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (6)
Ανάλυση της πιθανότητας σφάλματος – Εισαγωγή

- Από την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα η ληφθείσα ακολουθία να είναι από κοινού τυπική με την εκπεμφθείσα τείνει στο 1 για $n \to \infty$. Επίσης, η πιθανότητα η ληφθείσα ακολουθία να είναι από κοινού τυπική με ακολουθία διαφορετική από την εκπεμφθείσα ισούται περίπου με $2^{-nI(X;Y)}$. Επομένως, μπορούμε να χρησιμοποιήσουμε περίπου $2^{nI(X;Y)}$ κωδικές λέξεις και, ταυτόχρονα, να διασφαλίσουμε μικρή πιθανότητα σφάλματος.
- Στη συνέχεια θα αποδείξουμε τα παραπάνω και με την απαραίτητη μαθηματική αυστηρότητα.

Αποκωδικοποίηση με χρήση από κοινού τυπικότητας



Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (7)
Υπολογισμός Πιθανότητας Σφάλματος (I)

- Έστω ότι το μήνυμα W που εκπέμπεται επιλέγεται με ομοιόμορφη κατανομή από τα 2^{nR} πιθανά μηνύματα. $\mathcal{E} \triangleq \{\hat{W}(Y_1^n) \neq W\}$ είναι το ενδεχόμενο σφάλματος.
- Θα υπολογίσουμε τη μέση πιθανότητα σφάλματος για *όλα* τα πιθανά βιβλία κωδίκων.

$$\Pr\{\mathcal{E}\} = \sum_{\mathcal{C}} \Pr(\mathcal{C}) P_e^{(n)}(\mathcal{C}) =$$

$$= \sum_{\mathcal{C}} \Pr(\mathcal{C}) \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \lambda_w(\mathcal{C}) = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{\mathcal{C}} \Pr(\mathcal{C}) \lambda_w(\mathcal{C}).$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (8)
Υπολογισμός Πιθανότητας Σφάλματος (ΙΙ)

- Δεδομένου ότι η αντιστοίχιση μηνυμάτων σε κωδικές λέξεις γίνεται τυχαία και επειδή για όλους τους πιθανούς κώδικες το μήνυμα W θα αντιστοιχίζεται κάθε φορά σε διαφορετική κωδική λέξη, η ποσότητα $\sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_w(\mathcal{C})$ είναι ανεξάρτητη του μηνύματος w. Επομένως, μπορούμε να υποθέσουμε, χωρίς απώλεια της γενικότητας, ότι εστάλη η κωδική λέξη με δείκτη $w = 1$.
- Συνεπώς, η $\Pr(\mathcal{E})$ ισούται με

$$\begin{aligned}\Pr\{\mathcal{E}\} &= \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_w(\mathcal{C}) \\ &= \sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_1(\mathcal{C}) = \Pr(\mathcal{E}|W = 1).\end{aligned}$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (9) Υπολογισμός Πιθανότητας Σφάλματος (III)

- Ορίζουμε τα ενδεχόμενα $E_i = \left\{(X^n(i), Y^n) \in A_\epsilon^{(n)}\right\}$, $i \in \{1, 2, \ldots, 2^{nR}\}$, δηλαδή τα ενδεχόμενα η κωδική λέξη $X^n(i)$ (που αντιστοιχεί στο μήνυμα i) να είναι από κοινού τυπική με τη ληφθείσα ακολουθία Y^n η οποία προήλθε από μετάδοση της κωδικής λέξης $X^n(1)$.
- Συνεπώς,

$$\Pr(\mathcal{E}) = \Pr(\mathcal{E}|W = 1) = P(E_1^c \cup E_2 \cup E_3 \cup \ldots \cup E_{2^{nR}}|W = 1)$$
$$\stackrel{(a)}{\leq} P(E_1^c|W = 1) + \sum_{i=2}^{2^{nR}} P(E_i|W = 1).$$

(a) από το φράγμα ένωσης ενδεχομένων (union of events bound).

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (10)
Υπολογισμός Πιθανότητας Σφάλματος (IV)

$$\Pr(\mathcal{E}) \leq P(E_1^c|W=1) + \sum_{i=2}^{2^{nR}} P(E_i|W=1).$$

- Από την ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα η Y^n να μην είναι από κοινού τυπική με τη $X^n(1)$ τείνει στο 0 για $n \to \infty$: Επομένως, για κάθε $\epsilon > 0$ υπάρχει n_0 τέτοιο ώστε $P(E_1^c|W=1) \leq \epsilon$, για $n > n_0$.
- Επίσης, από τον τυχαίο τρόπο δημιουργίας του κώδικα, οι κωδικές λέξεις $X^n(1)$ και $X^n(i)$ είναι ανεξάρτητες μεταξύ τους για $i \neq 1$, με αποτέλεσμα η Y^n να είναι ανεξάρτητη από τις $X^n(i)$ για $i \neq 1$. Από την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα οι $X^n(i)$ και Y^n να είναι από κοινού τυπικές ενώ επιλέχθηκαν ανεξάρτητα είναι $\leq 2^{-n(I(X;Y)-3\epsilon)}$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (11) Υπολογισμός Πιθανότητας Σφάλματος (V)

- Συνδυάζοντας όλα τα παραπάνω,

$$\begin{aligned} \Pr(\mathcal{E}) &\leq P(E_1^c|W=1) + \sum_{i=2}^{2^{nR}} P(E_i|W=1) \leq \epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X;Y)-3\epsilon)} \\ &= \epsilon + \left(2^{nR}-1\right) 2^{-n(I(X;Y)-3\epsilon)} \leq \epsilon + 2^{-n(I(X;Y)-3\epsilon-R)} \leq 2\epsilon. \end{aligned}$$

  Η τελευταία ανισότητα ισχύει εφόσον $n > n_1$ και $R < I(X;Y) - 3\epsilon$.

- Επομένως, εάν $R < I(X;Y)$, μπορούμε να επιλέξουμε n τέτοιο ώστε η μέση πιθανότητα σφάλματος υπολογισμένη επάνω σε όλους τους πιθανούς κώδικες και σε όλες τις πιθανές κωδικές λέξεις να μην υπερβαίνει το 2ϵ, για οποιοδήποτε $\epsilon > 0$.

- Δεν τελειώσαμε ακόμα... Πρέπει να δείξουμε ότι η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \to 0$ και, επίσης, ότι υπάρχει τουλάχιστον ένας κώδικας με $\lambda^{(n)} \to 0$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (12) Επιλογή βιβλίου κωδίκων (Ι)

- Εάν οι κωδικές λέξεις δημιουργηθούν με βάση την κατανομή $p^*(x)$ η οποία μεγιστοποιεί την αμοιβαία πληροφορία, $I_{p^*}(X;Y) = C$ και, επομένως, μπορούμε να μεταδώσουμε με $R < C$.
- Δεδομένου ότι η μέση πιθανότητα σφάλματος για όλους τους τυχαίους κώδικες δεν υπερβαίνει το 2ϵ, υπάρχει τουλάχιστον ένα βιβλίο κωδίκων (κώδικας) $\mathcal{C}^*$ για το οποίο η μέση πιθανότητα σφάλματος δεν υπερβαίνει το 2ϵ: $\Pr(\mathcal{E}|\mathcal{C}^*) \leq 2\epsilon$. Ο $\mathcal{C}^*$ μπορεί να βρεθεί με αναζήτηση μέσα σε όλους τους 2^{nR} κώδικες. Επομένως,

$$\Pr(\mathcal{E}|\mathcal{C}^*) \leq \frac{1}{2^{nR}} \sum \lambda_i(\mathcal{C}^*) \leq 2\epsilon.$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (13)
Επιλογή βιβλίου κωδίκων (II)

- Το γεγονός ότι η μέση πιθανότητα σφάλματος του κώδικα $\mathcal{C}^*$ είναι $\leq 2\epsilon$, δεν εγγυάται ότι η πιθανότητα σφάλματος που αντιστοιχεί στη μετάδοση ενός συγκεκριμένου μηνύματος W (και, επομένως, μιας συγκεκριμένης κωδικής λέξης $X^n(W)$) θα είναι $\leq 2\epsilon$.
- Εάν θέλουμε να διασφαλίσουμε μικρή πιθανότητα σφάλματος για <u>κάθε</u> κωδική λέξη (και, άρα, για κάθε μήνυμα) μπορούμε να αφαιρέσουμε τις μισές χειρότερες κωδικές λέξεις του κώδικα (δηλαδή τις 2^{nR-1} κωδικές λέξεις με τη μεγαλύτερη πιθανότητα σφάλματος).
- Δεδομένου ότι η μέση πιθανότητα σφάλματος είναι $\leq 2\epsilon$, η <u>μέγιστη</u> πιθανότητα σφάλματος των μισών "καλύτερων" λέξεων που απομένουν δε θα υπερβαίνει το 4ϵ.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (14)
Επιλογή βιβλίου κωδίκων (III)

- Ο νέος κώδικας έχει 2^{nR-1} κωδικές λέξεις και, άρα, ρυθμό $R' = R - \frac{1}{n}$. Για μεγάλα n, η απώλεια ρυθμού μετάδοσης είναι αμελητέα.
- Επομένως, δείξαμε ότι μπορούμε να επιτύχουμε οποιοδήποτε ρυθμό μετάδοσης που δεν υπερβαίνει τη χωρητικότητα και, ταυτόχρονα, η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \leq 4\epsilon$.

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Ανακεφαλαίωση

Για να αποδείξουμε το Θεώρημα Κωδικοποίησης Καναλιού

- Δημιουργήσαμε όλους τους πιθανούς κώδικες (βιβλία κωδίκων) με κωδικές λέξεις μεγάλου μήκους n.
- Η δημιουργία των κωδικών λέξεων έγινε με βάση την κατανομή $p^*(x)$ που μεγιστοποιεί την $I(X;Y)$.
- Κρατήσαμε τον καλύτερο από τους τυχαίους κώδικες $\mathcal{C}^*$ (τον κώδικα στον οποίο αντιστοιχεί η μικρότερη μέση πιθανότητα σφάλματος).
- Δείξαμε ότι, για αρκούντως μεγάλα μήκη κωδικών λέξεων n, εάν $R < I(X;Y)$, η πιθανότητα η ακολουθία εξόδου να μην είναι τυπική με τη μεταδοθείσα κωδική λέξη ή να είναι τυπική με κωδική λέξη διαφορετική από αυτή που μεταδόθηκε τείνει στο 0. Επομένως, η μέση πιθανότητα σφάλματος μπορεί να περιοριστεί αυθαίρετα κοντά στο 0.
- Με τροποποίηση του κώδικα (και αυθαίρετα μικρή απώλεια ρυθμού μετάδοσης) δείξαμε ότι όχι μόνο η μέση, αλλά και η μέγιστη πιθανότητα σφάλματος μπορεί να περιοριστεί αυθαίρετα κοντά στο 0.

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια

- Η δημιουργία τυχαίων κωδίκων οδηγεί μεν σε (μια) απόδειξη του ευθέος του Θεωρήματος Κωδικοποίησης Καναλιού, αλλά δεν αποτελεί πρακτικό τρόπο μετάδοσης.
- Πρόκειται για μια απόδειξη ύπαρξης (existence proof). Δεν είναι κατασκευαστική απόδειξη (constructive proof).
- Η δημιουργία του κώδικα, αν και πολύπλοκη, μπορεί να γίνει μια φορά υποθέτοντας ότι ο πίνακας μετάβασης του καναλιού $p(y|x)$ δεν αλλάζει.
- Παρατηρήστε ότι ο βέλτιστος κώδικας μπορεί να βρεθεί από τον πομπό και από το δέκτη ανεξάρτητα, χωρίς συνεννόηση, εάν γνωρίζουν και οι δύο τον πίνακα μετάβασης καναλιού και αν δημιουργήσουν όλους τους πιθανούς κώδικες (και κρατήσουν τον καλύτερο από άποψη ελάχιστης πιθανότητας σφάλματος).

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια (2)

- Μπορεί να αποδειχτεί κάτι ακόμα πιο ισχυρό: Για επαρκώς μεγάλο n, *οποιοσδήποτε* τυχαίος κώδικας που κατασκευάζεται με τον τρόπο που προαναφέραμε μπορεί να επιτύχει αυθαίρετα μικρή πιθανότητα σφάλματος (δείτε, π.χ. R. W. Yeung, *A First Course in Information Theory*).
- Δηλαδή, δε χρειάζεται να φτιάξουμε όλους τους κώδικες και να διαλέξουμε έναν καλό. Αρκεί να κατασκευάσουμε τυχαία έναν κώδικα (φυσικά το n θα πρέπει να είναι πολύ μεγάλο).
- Χάρη στο νόμο των μεγάλων αριθμών, η επικάλυψη των συνθήκη τυπικών ακολουθιών $\mathcal{T}_\epsilon^{(n)}(Y^n|x^n(i))$ αρχίζει και γίνεται αμελητέα καθώς το n αυξάνει.

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια (3)

- Παρατηρήστε ότι (σχεδόν όλες) οι κωδικές λέξεις $x^n(i)$ (τα σύμβολα των οποίων ακολουθούν κατανομή $p^*(x)$) είναι τυπικές.
- Αυτός είναι ένας από τους λόγους για τους οποίους "απλοί" κώδικες (όπως, για παράδειγμα, ο κώδικας επανάληψης) δεν είναι καλοί.
- Παρατηρήστε, επίσης, ότι, για έναν εξωτερικό παρατηρητή που δε γνωρίζει τα όρια των κωδικών λέξεων, τα σύμβολα που εισέρχονται στο κανάλι φαίνονται i.i.d. με κατανομή $p(x)$. Ωστόσο, για τον αποκωδικοποιητή που γνωρίζει τα όρια των λέξεων τα σύμβολα δεν είναι τυχαία, γιατί ανήκουν σε μία από 2^{nR} λέξεις του βιβλίου κωδίκων.

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια (4)

- Το σημαντικότερο πρόβλημα βρίσκεται στην αποκωδικοποίηση, καθώς ο αριθμός των κωδικών λέξεων των οποίων η από κοινού τυπικότητα με την Y^n θα πρέπει να ελεγχθεί αυξάνει εκθετικά με το n.
- Το πρόβλημα αυτό παραμένει ακόμα και όταν η αποκωδικοποίηση γίνεται με χρήση άλλων κριτηρίων (π.χ. ανίχνευση Μέγιστης Πιθανοφάνειας).
- Η επίτευξη ρυθμών μετάδοσης κοντά στη χωρητικότητα του καναλιού με υλοποιήσιμους τρόπους αποτελεί αντικείμενο της Θεωρίας Κωδικοποίησης. Σήμερα, υπάρχουν περιπτώσεις καναλιών και απαιτήσεων σε πιθανότητα σφάλματος για τις οποίες η μετάδοση κοντά στη χωρητικότητα είναι εφικτή με πολυπλοκότητα που δεν είναι απαγορευτική για την υλοποίηση των κωδικοποιητών και των αποκωδικοποιητών (π.χ. Turbo Codes/LDPC). Ένας λόγος χάρη στον οποίο είναι εφικτή η κωδικοποίηση/αποκωδικοποίηση είναι ότι οι κώδικες που χρησιμοποιούνται είναι *γραμμικοί*.

Απόδειξη ασθενους αντιστρόφου με χρήση Ανισότητας Fano

$I(X^n;Y^n) \leq nC$

Θα αποδείξουμε, κατ' αρχάς, ότι, για Διακριτά Κανάλια Χωρίς Μνήμη, η πληροφοριακή χωρητικότητα ανά χρήση του καναλιού δεν αυξάνει εάν το κανάλι χρησιμοποιηθεί ως κανάλι γινομένου. Δηλαδή, $I(X^n;Y^n) \leq nC$ για οποιαδήποτε $p(x)$, όπου $C = \max_{p(x)} I(X;Y)$.

$$I(X^n;Y^n) = H(Y^n) - H(Y^n|X^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1,\ldots,Y_{i-1},X^n) =$$

$$\stackrel{(a)}{=} H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i) \stackrel{(b)}{\leq} \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i)$$

$$= \sum_{i=1}^{n} I(X_i;Y_i) \leq nC.$$

(a) Το κανάλι δεν έχει μνήμη, επομένως η έξοδος τη χρονική στιγμή i εξαρτάται μόνο από την είσοδο τη χρονική στιγμή i. Επίσης, δε χρησιμοποιείται ανάδραση. (b) Η από κοινού εντροπία δεν υπερβαίνει το άθροισμα των εντροπιών.

Ανισότητα Fano

- Για την απόδειξη του αντιστρόφου του Θεωρήματος Κωδικοποίησης Καναλιού θα χρησιμοποιήσουμε την Ανισότητα Fano.
- Είδαμε ότι, για κάθε εκτιμητή $\hat{X} = g(Y)$,

$$H(X|Y) \leq H(X|\hat{X}) \leq H(P_e)+P_e \log|\mathcal{X}| \Rightarrow H(X|\hat{X}) \leq 1+P_e \log|\mathcal{X}|,$$

όπου $P_e = \Pr\{\hat{X} \neq X\}$.
- Εάν θεωρήσουμε Διακριτό Κανάλι Χωρίς Μνήμη με βιβλίο κωδίκων $\mathcal{C}$ και ομοιόμορφα κατανεμημένα μηνύματα W,

$$H(W|\hat{W}) \leq 1 + P_e^{(n)} nR, \text{ όπου } P_e^{(n)} = \Pr\{W \neq \hat{W}\}.$$

Θεώρημα Κωδικοποίησης Καναλιού – Απόδειξη ασθενούς αντιστρόφου

- Θα δείξουμε ότι, για κάθε κώδικα $(2^{nR}, n)$ με $\lambda^{(n)} \to 0$, πρέπει να ισχύει $R \leq C$. Δεδομένου ότι $\lambda^{(n)} \to 0$ και η μέση πιθανότητα σφάλματος $P_e^{(n)} \to 0$.
- Έστω ότι ο δέκτης αποφασίζει ποια ακολουθία μεταδόθηκε με βάση κάποια συνάρτηση αποκωδικοποίησης $\hat{W} = g(Y^n)$. Ισχύει $W \to X^n(W) \to Y^n \to \hat{W}$.
- Έστω, επίσης, ότι το μήνυμα που στέλνεται στο κανάλι επιλέγεται με βάση ομοιόμορφη κατανομή στο σύνολο των πιθανών μηνυμάτων $\{1, 2, \ldots, 2^{nR}\}$. Επομένως, $\Pr\{\hat{W} \neq W\} = P_e^{(n)} = \frac{1}{2^{nR}} \sum_i \lambda_i$.

Θεώρημα Κωδικοποίησης Καναλιού – Απόδειξη ασθενούς αντιστρόφου (2)

- Συνεπώς,

$$nR \overset{(a)}{=} H(W) \overset{(b)}{=} H(W|\hat{W}) + I(W;\hat{W}) \overset{(c)}{\leq} 1 + P_e^{(n)} nR + I(W;\hat{W})$$
$$\overset{(d)}{\leq} 1 + P_e^{(n)} nR + I(X^n;Y^n) \overset{(e)}{\leq} 1 + P_e^{(n)} nR + nC.$$

(a) W ομοιόμορφη τ.μ., (b) σχέση αμοιβαίας πληροφορίας – εντροπίας, (c) ανισότητα Fano, (d) ανισότητα επεξεργασίας δεδομένων, (e) $I(X^n;Y^n) \leq nC$.

Θεώρημα Κωδικοποίησης Καναλιού Απόδειξη ασθενούς αντιστρόφου (3)

$$nR \leq 1 + P_e^{(n)} nR + nC \Rightarrow R \leq P_e^{(n)} R + \frac{1}{n} + C.$$

- Από την υπόθεση ότι $\lambda^{(n)} \to 0$, $P_e^{(n)} R \to 0$ για $n \to \infty$. Επομένως, για $n \to \infty$,
$$R < C.$$
- Λύνοντας ως προς $P_e^{(n)}$, $P_e^{(n)} \geq 1 - \frac{C}{R} - \frac{1}{nR}$. Συνεπώς, εάν $R > C$, $P_e^{(n)} > 0$ για $n \to \infty$.

Θεώρημα Κωδικοποίησης Καναλιού
Απόδειξη ασθενούς αντιστρόφου (4)

- Το αποτέλεσμα αυτό ονομάζεται ασθενές αντίστροφο του Θεωρήματος Κωδικοποίησης Καναλιού. Αποδεικνύεται (ισχυρό αντίστροφο) ότι, εάν $R > C$, $P_e^{(n)} \to 1$ εκθετικά.
- Συνεπώς, η χωρητικότητα καναλιού C αποτελεί μια πολύ σαφή διαχωριστική γραμμή: Όταν $R < C$ η πιθανότητα σφάλματος τείνει εκθετικά στο 0. Αντίθετα, όταν $R > C$, η πιθανότητα σφάλματος τείνει εκθετικά στο 1.

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου

- Θα αποδείξουμε ξανά το αντίστροφο με μία μικρή παραλλαγή στη χρήση της ανισότητας Fano.
- Η απόδειξη αυτή είναι πιο γενική. Όπως θα δούμε σύντομα, μπορεί να εφαρμοστεί και στην περίπτωση που χρησιμοποιείται ανάδραση.

Θεώρημα Κωδικοποίησης Καναλιού Εναλλακτική απόδειξη ασθενούς αντιστρόφου (2)

- Επειδή $W \to X^n(W) \to Y^n \to \hat{W}$,

$$H(W|Y^n) \leq H(W|\hat{W}).$$

- Επίσης, από την ανισότητα Fano,

$$H(W|\hat{W}) \leq 1 + P_e^{(n)} nR, \text{ όπου } P_e^{(n)} = \Pr\{W \neq \hat{W}\}.$$

- Υποθέτοντας, και πάλι, ότι τα μηνύματα W ακολουθούν ομοιόμορφη κατανομή,

$$nR = H(W) \overset{(a)}{=} I(W; Y^n) + H(W|Y^n)$$
$$\overset{(b)}{\leq} I(W; Y^n) + 1 + P_e^{(n)} nR$$

(a) Σχέση εντροπίας-αμοιβαίας πληροφορίας, (b) ανισότητα Fano.

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου (3)

$$
\begin{aligned}
nR &\le I(W;Y^n) + 1 + P_e^{(n)} nR \\
&\stackrel{(c)}{=} \sum_{i=1}^{n} I(W;Y_i|Y^{i-1}) + 1 + P_e^{(n)} nR \\
&\stackrel{(d)}{\le} \sum_{i=1}^{n} I(W,Y^{i-1};Y_i) + 1 + P_e^{(n)} nR \\
&\stackrel{(e)}{\le} \sum_{i=1}^{n} I(X_i,W,Y^{i-1};Y_i) + 1 + P_e^{(n)} nR
\end{aligned}
$$

(c) κανόνας αλυσίδας, (d) $I(W,Y^{i-1};Y_i) = I(Y^{i-1};Y_i) + I(W;Y_i|Y^{i-1})$, (e) $X_i = f(W,Y^{i-1})$ (ισχύει ακόμα και όταν χρησιμοποιείται ανάδραση).

Θεώρημα Κωδικοποίησης Καναλιού
Εναλλακτική απόδειξη ασθενούς αντιστρόφου (4)

$$
\begin{aligned}
nR &\le \sum_{i=1}^{n} I(X_i, W, Y^{i-1}; Y_i) + 1 + P_e^{(n)} nR \\
&\overset{(f)}{=} \sum_{i=1}^{n} I(X_i; Y_i) + 1 + P_e^{(n)} nR \\
&\le nC + 1 + P_e^{(n)} nR = n\left(C + \frac{1}{n} + P_e^{(n)}\right)
\end{aligned}
$$

(f) $X_i = f(W, Y^{i-1})$ και το κανάλι δεν έχει μνήμη, οπότε $(W, Y^{i-1}) \to X_i \to Y_i$.

Θεώρημα Κωδικοποίησης Καναλιού Εναλλακτική απόδειξη ασθενούς αντιστρόφου (5)

- Επομένως, για $n \to \infty$ και $P_e^{(n)} \to 0$,

$$nR \leq n\left(C + \epsilon_n\right) \Rightarrow R < C.$$

- Παρατηρήστε ότι δεν απαγορέψαμε τη χρήση ανάδρασης στον κώδικα (περισσότερα σύντομα). Αυτό σημαίνει ότι, σε ένα κανάλι χωρίς μνήμη, ακόμα και αν μπορούμε να χρησιμοποιήσουμε ανάδραση, η χωρητικότητα δεν αυξάνει (περισσότερα σύντομα).

Γιατί χρησιμοποιούμε $W \sim \text{Unif}\left[0, 2^{nR} - 1\right]$;

- Στην απόδειξη του ασθενούς αντιστρόφου με χρήση της ανισότητας Fano υποθέσαμε ότι $W \sim \text{Unif}\left[0, 2^{nR} - 1\right]$.
- Μήπως αυτό σημαίνει ότι το αντίστροφο ισχύει μόνο για κώδικες όπου όλες οι κωδικές λέξεις είναι ισοπίθανες;
- Όχι (αυτό είναι ένα λεπτό σημείο). Θυμηθείτε ότι η χωρητικότητα ισούται με το μέγιστο εφικτό ρυθμό μετάδοσης. Ο ρυθμός μετάδοσης ισούται με $\log M/n$.
- Δηλαδή, για να δείξουμε ότι ο ρυθμός μετάδοσης ενός κώδικα είναι R αρκεί να αποδείξουμε ότι υπάρχει κώδικας με 2^{nR} κωδικές λέξεις τον οποίο μπορούμε να χρησιμοποιήσουμε και να επιτύχουμε $P_e^{(n)} \to 0$.
- Για να δείξουμε ότι η χωρητικότητα ενός καναλιού είναι C πρέπει να δείξουμε ότι δεν υπάρχει κώδικας μήκους n με περισσότερες από 2^{nC} κωδικές λέξεις.

Γιατί χρησιμοποιούμε $W \sim$ Unif $\left[0, 2^{nR} - 1\right]$; (2)

- Η πιθανότητα με την οποία ο χρήστης του κώδικα στέλνει κάθε κωδική λέξη δε μας αφορά (τουλάχιστον όσον αφορά την απόδειξη του αντιστρόφου). Εμείς θέλουμε μόνο να κατασκευάσουμε 2^{nR} κωδικές λέξεις τις οποίες να μπορεί να διακρίνει ο δέκτης με αυθαίρετα μικρή πιθανότητα σφάλματος.
- Η επιλογή $W \sim$ Unif $\left[0, 2^{nR} - 1\right]$ γίνεται απλώς και μόνο γιατί μας βολεύει στην απόδειξη του αντιστρόφου.