EE728 Προχωρημένα Θέματα Θεωρίας Πληροφορίας 7η διάλεξη

Δημήτρης-Αλέξανδρος Τουμπακάρης

Τμήμα ΗΜ&ΤΥ, Πανεπιστήμιο Πατρών

14 Απριλίου 2010

Περιεχόμενα σημερινού μαθήματος

Αντιστοιχία με συγγράμματα

- Cover & Thomas: 7.7, 7.9
- Gallager: Διαφορετική απόδειξη του ευθέος (όχι με τυπικότητα) στο Κεφ. 5. Περισσότερα σε επόμενο μάθημα. Ασθενές αντίστροφο: 4.3. "Ισχυρό" αντίστροφο (αντίστροφο Wolfowitz): 5.8. Θεωρήματα σχετικά με τη χωρητικότητα: 4.4, 4.5

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. $\mathcal{E} \triangleq \{\hat{W}(Y^n) \neq W\}$ 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 (II)

- 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 $w = 1$.
- Επομένως, 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 $E_i = \{(X^n(i), Y^n) \in A_\epsilon^{(n)}\}$, $i \in \{1, 2, \ldots, 2^{nR}\}$, 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,

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

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,

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

  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 <u>μέση</u> 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 <u>μέγιστη</u> 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 κώδικες 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 (ΙΙ)

- 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 <u>μέγιστη</u> 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 επιτυγχάνει τη χωρητικότητα καναλιού.
- 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, εφόσον $R < I(X;Y)$, 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 Θεωρήματος Κωδικοποίησης Καναλιού, αλλά δεν αποτελεί πρακτικό τρόπο μετάδοσης.
- 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 (συνέχεια)

- 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. Η μετάδοση κοντά στη χωρητικότητα είναι σήμερα εφικτή με πολυπλοκότητα που δεν είναι απαγορευτική για την υλοποίηση των αποκωδικοποιητών.

Απόδειξη αντιστρόφου

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

Θα αποδείξουμε, κατ' αρχάς, ότι, για Διακριτά Κανάλια Χωρίς Μνήμη, η πληροφοριακή χωρητικότητα ανά χρήση του καναλιού δεν αυξάνει εάν το κανάλι χρησιμοποιηθεί πολλές φορές. 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 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 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 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 $n \to \infty$,
$$R \leq C.$$
- 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 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: 'Οταν $R < C$ η πιθανότητα σφάλματος τείνει εκθετικά στο 0. Αντίθετα, όταν $R > C$, η πιθανότητα σφάλματος τείνει εκθετικά στο 1.

Παρατηρήσεις και θεωρήματα σχετικά με τη χωρητικότητα

Μεγιστοποίηση κοίλης συνάρτησης κατανομής πιθανότητας

- Θεωρούμε συνάρτηση $f(\mathbf{p}) : \mathbf{R}^n \to \mathbf{R}$ η οποία είναι κοίλη $\cap$ ως προς $\mathbf{p}$.
- Έστω, επίσης, ότι το $\mathbf{p}$ είναι κατανομή (διάνυσμα πιθανότητας), δηλαδή $p_i \geq 0,\ i = 1, \ldots, n$ και $\sum_{i=1}^{n} p_i = \mathbf{1}^T \mathbf{p} = 1$.
- Τέλος, θεωρούμε ότι οι μερικές παράγωγοι $\partial f(\mathbf{p})/\partial p_i$ ορίζονται και ότι είναι συνεχείς με μοναδική εξαίρεση το $\lim_{p_i \to 0} \partial f(\mathbf{p})/\partial p_i$ που μπορεί να είναι και $+\infty$.

Μεγιστοποίηση κοίλης συνάρτησης κατανομής πιθανότητας (συνέχεια)

- Αποδεικνύεται ότι οι παρακάτω συνθήκες είναι ικανές και αναγκαίες για να μεγιστοποιείται η $f()$ στο σημείο (κατανομή) $\mathbf{p}$.

$$\frac{\partial f(\mathbf{p})}{\partial p_i} = \lambda, \text{ για όλα τα } i \text{ για τα οποία } p_i > 0$$
$$\frac{\partial f(\mathbf{p})}{\partial p_i} \leq \lambda, \text{ για όλα τα } i \text{ για τα οποία } p_i = 0$$

  για κάποια τιμή της παραμέτρου λ.

- Για την απόδειξη δείτε π.χ. Gallager Theorem 4.4.1.

Μεγιστοποίηση αμοιβαίας πληροφορίας

- Με χρήση του προηγούμενου θεωρήματος και του ότι η $I(X;Y)$ είναι κοίλη $\cap$ συνάρτηση της κατανομής εισόδου $p(x)$ για δεδομένο κανάλι $p(y|x)$, αποδεικνύεται ότι οι παρακάτω δύο συνθήκες αποτελούν ικανή και αναγκαία συνθήκη για να επιτυγχάνει μια κατανομή $\mathbf{p}^*$ τη χωρητικότητα.

$$I(X=x_i;Y) = C, \text{ για όλα τα } i \text{ για τα οποία } p^*_{x_i} > 0$$
$$I(X=x_i;Y) \leq C, \text{ για όλα τα } i \text{ για τα οποία } p^*_{x_i} = 0$$

όπου $I(X=x_i;Y) = \sum_{y\in|\mathcal{Y}|} p(y|x_i)\log\frac{p(y|x_i)}{p(y)}$ η αμοιβαία πληροφορία μεταξύ $X=x_i$ και Y.

Μεγιστοποίηση αμοιβαίας πληροφορίας (συνέχεια)

- Το αποτέλεσμα αυτό έχει μια διαισθητική επεξήγηση: Εάν για $x_i \neq x_j$ $I(X = x_i; Y) > I(X = x_j; Y)$, μπορούμε να αυξήσουμε την $I(X; Y) = \sum_{x_k} p(x_k) I(X = x_k; Y)$ χρησιμοποιώντας τη x_i πιο συχνά και τη x_j λιγότερο συχνά (αλλάζοντας τις $p(x_i)$ και $p(x_j)$).
- Αυτό έχει ως αποτέλεσμα να αλλάξει η $p(y) = \sum_{x_k} p(x_k) p(y|x_k)$.
- Τελικά, η διαδικασία αυτή θα ισορροπήσει σε σημείο όπου όλες οι $I(X = x_i; Y)$ εκτός, ίσως, από κάποιες που αντιστοιχούν σε κακές εισόδους, θα ισούνται μεταξύ τους (και, επομένως, και με τη χωρητικότητα, C).

Άλλες ενδιαφέρουσες ιδιότητες και αποτελέσματα

- Αναφέρουμε, τέλος, 3 ενδιαφέροντα πορίσματα. Για αποδείξεις δείτε π.χ. Gallager Κεφ. 4.5.

Πόρισμα 1 Για οποιαδήποτε κατανομή εισόδου, $p^*(x)$, που επιτυγχάνει τη χωρητικότητα σε διακριτό κανάλι χωρίς μνήμη, όλες οι πιθανότητες συμβόλων εξόδου, $p(y)$, είναι αυστηρώς θετικές (αρκεί για κάθε έξοδο να υπάρχει τουλάχιστον μία είσοδος που οδηγεί σε αυτήν).

Πόρισμα 2 Η κατανομή εξόδου, $p^*(y)$, για την οποία $I(X;Y) = C$ είναι μοναδική. Όλες οι κατανομές εισόδου, $p(x)$, για τις οποίες $\sum_{x\in|\mathcal{X}|} p(x)p(y|x) = p^*(y)$ επιτυγχάνουν τη χωρητικότητα.

Πόρισμα 3 Έστω m ο ελάχιστος αριθμός συμβόλων εισόδου που μπορούν να χρησιμοποιηθούν (με μη μηδενική πιθανότητα) για να επιτευχθεί μετάδοση με τη χωρητικότητα. Έστω $\mathcal{A}$ ένα τέτοιο σύνολο m συμβόλων εισόδου. Ισχύει $m \leq |\mathcal{Y}|$. Επίσης, η κατανομή $p(x)$ στα στοιχεία του $\mathcal{A}$ που επιτυγχάνει τη χωρητικότητα είναι μοναδική.

Πώς υπολογίζουμε τη χωρητικότητα;

- Γενικά, ο υπολογισμός της χωρητικότητας δεν είναι εύκολη υπόθεση.
- Σε μερικές, ειδικές, περιπτώσεις μπορούμε να χρησιμοποιήσουμε ιδιότητες όπως, π.χ. στην περίπτωση συμμετρικών καναλιών.
- Άλλες φορές μπορούμε να "μαντέψουμε" την κατανομή εισόδου και να δείξουμε ότι επιτυγχάνει ένα άνω φράγμα για τη χωρητικότητα (όπως κάναμε για το συμμετρικό κανάλι).
- Στη γενική περίπτωση καταφεύγουμε σε αριθμητικές μεθόδους με χρήση υπολογιστή. Μια ευρέως χρησιμοποιούμενη μέθοδος είναι των Blahut & Arimoto. Τα τελευταία χρόνια έχουν προταθεί βελτιώσεις που συγκλίνουν πολύ πιο γρήγορα σε σχέση με τον αρχικό αλγόριθμο.