ΕΕ728

Προχωρημένα Θέματα Θεωρίας Πληροφορίας

Δημήτρης - Αλέξανδρος Τουμπακάρης

5ο Μάθημα – 2 Απριλίου 2008

Ανακεφαλαίωση προηγούμενου μαθήματος

- Κωδικοποίηση σταθερού μήκους με χρήση του AEP. Στο όριο, αρκεί να έχουμε αποδοτική περιγραφή μόνο για τις τυπικές ακολουθίες. Μπορούμε είτε να αγνοήσουμε τις μη τυπικές, είτε να τις κωδικοποιήσουμε με μη αποδοτικό τρόπο, δεδομένου ότι η συνεισφορά τους στο μέσο μήκος του κώδικα είναι αμελητέα.
- Θεώρημα Κωδικοποίησης Πηγής. Μπορούμε να Κωδικοποιήσουμε με $E[l]$ αυθαίρετα κοντά στην εντροπία (στο ρυθμό εντροπίας, γενικότερα) με κώδικα σταθερού ή μεταβλητού μήκους. Εάν προσπαθήσουμε να κωδικοποιήσουμε με μικρότερο μέσο μήκος, $P_e \to 1$.
- Εισαγωγή στα Διακριτά Κανάλια και, ειδικότερα, στα Διακριτά Κανάλια χωρίς Μνήμη.
- Η χωρητικότητα Διακριτών Καναλιών Χωρίς μνήμη είναι ο μέγιστος εφικτός ρυθμός μετάδοσης και ισούται με $\max_{p(x)} I(X;Y)$ (Σύμφωνα με το Θεώρημα Κωδικοποίησης Καναλιού το οποίο θα αποδείξουμε).

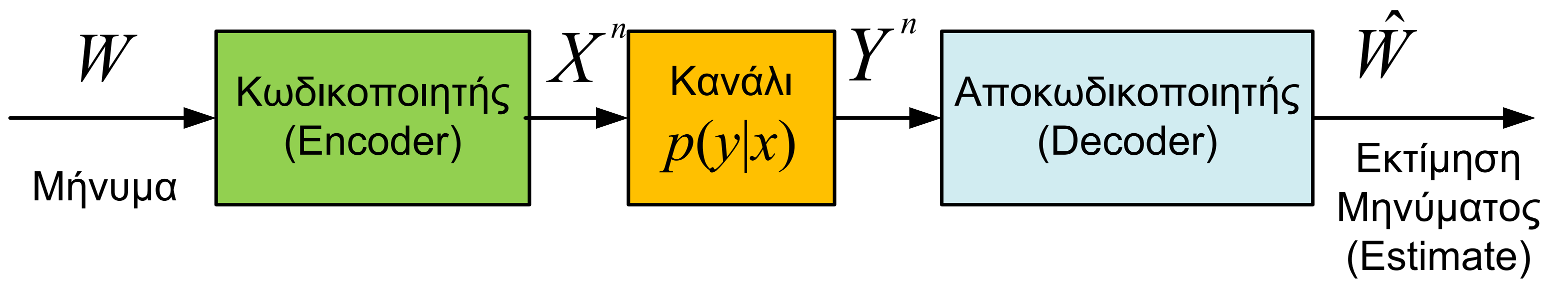
Προεπισκόπηση σημερινού μαθήματος

- Συμμετρικά Κανάλια και Χωρητικότητα.
- Από Κοινού Τυπικότητα και Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP).

Περιεχόμενα σημερινού μαθήματος

- $\boxed{\text{Συμμετρικά Κανάλια και Χωρητικότητα}}$
- Από Κοινού Τυπικότητα και Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint A-EP).

Μετάδοση σε Διακριτά Κανάλια



- Πόση είναι η μέγιστη πληροφορία που μπορεί να μεταδοθεί ανά χρήση του καναλιού;
- Πώς αυτή εξαρτάται από τα χαρακτηριστικά του καναλιού;
- Πώς επιτυγχάνεται μετάδοση του μέγιστου αυτού ποσού πληροφορίας;

Διακριτά κανάλια – Ορισμοί

- Ένα <u>διακριτό κανάλι</u> $(\mathcal{X}, p(y|x), \mathcal{Y})$ αποτελείται από δύο <u>πεπερασμένα</u> σύνολα $\mathcal{X}$ και $\mathcal{Y}$ και <u>ένα σύνολο δεσμευμένων συναρτήσεων μάζας πιθανότητας</u> $p(y|x)$, μια για κάθε $x \in \mathcal{X}$, ώστε, για κάθε x και y, $p(y|x) \geq 0$ και, για κάθε x, $\sum_y p(y|x) = 1$. Η τ.μ. X είναι η είσοδος του καναλιού και η Y η έξοδός του.
- Έστω ότι χρησιμοποιούμε ένα διακριτό κανάλι n φορές. Ορίζουμε τη <u>n-οστή επέκταση του διακριτού καναλιού χωρίς μνήμη</u> $(\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n)$, όπου

$$p(y_k|x^k, y^{k-1}) = p(y_k|x_k), \qquad k = 1, 2, \ldots, n.$$

- Εάν το κανάλι (χωρίς μνήμη) χρησιμοποιείται χωρίς ανάδραση, δηλαδή η είσοδος στο κανάλι δεν εξαρτάται από τις εξόδους σε προηγούμενες χρονικές στιγμές, $p(x_k|x^{k-1}, y^{k-1}) = p(x_k|x^{k-1})$, και

$$p(y^n|x^n) = \prod_{i=1}^{n} p(y_i|x_i).$$

Χωρητικότητα Διακριτού Καναλιού Χωρίς Μνήμη

- "Πληροφοριακή" Χωρητικότητα Διακριτού Καναλιού Χωρίς Μνήμη ("Information" Channel Capacity):

$$C = \max_{p(x)} I(X;Y)$$

- Παραδείγματα διακριτών καναλιών χωρίς μνήμη (επανάληψη από το μάθημα "Θεωρία Πληροφορίας")
  - Δυαδικό Συμμετρικό Κανάλι (Binary Symmetric Channel – BSC): $C = 1 - H(p)$ bits, επιτυγχάνεται με ομοιομορφη κατανομή εισόδου $p(x) = \left(\frac{1}{2}, \frac{1}{2}\right)$.
  - Δυαδικό Κανάλι με Διαγραφή (Binary Erasure Channel): $C = 1 - \alpha$, όπου α η πιθανότητα διαγραφής. Επιτυγχάνεται με ομοιομορφη κατανομή εισόδου $p(x) = \left(\frac{1}{2}, \frac{1}{2}\right)$.
  - Η χωρητικότητα του δυαδικού καναλιού με διαγραφή παραμένει η ίδια εάν χρησιμοποιήσουμε ανάδραση.
  - Θα δούμε ότι το αποτέλεσμα αυτό, δηλαδή ότι η χρήση ανάδρασης δεν αυξάνει τη χωρητικότητα, ισχύει γενικά για τα όλα τα διακριτά κανάλια χωρίς μνήμη.

Χωρητικότητα Συμμετρικού Καναλιού

- Πίνακας μετάβασης $[p(y|x)]_{i,j}$. Μεταδοθέν σύμβολο x_i. Ληφθέν σύμβολο y_j.
- Ένα διακριτό κανάλι ονομάζεται συμμετρικό όταν κάθε γραμμή του πίνακα μετάβασης $p(y|x)$ προκύπτει από αναδιάταξη κάθε άλλης γραμμής και το ίδιο ισχύει και για κάθε στήλη του πίνακα. Ένα διακριτό κανάλι ονομάζεται ασθενώς συμμετρικό όταν κάθε γραμμή του πίνακα μετάβασης $p(y|x)$ προκύπτει από αναδιάταξη κάθε άλλης γραμμής και τα αθροίσματα των στοιχείων κάθε στήλης $\sum_x p(y|x)$ ισούνται μεταξύ τους.
- Θεώρημα: Για τη χωρητικότητα ασθενώς συμμετρικού καναλιού (και, επομένως, και συμμετρικού καναλιού), ισχύει

$$C = \log|\mathcal{Y}| - H(\text{οποιασδήποτε γραμμής } \mathbf{r} \text{ πίνακα μετάβασης}).$$

Χωρητικότητα Συμμετρικού Καναλιού (συνέχεια)

- Απόδειξη:

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_{x \in \mathcal{X}} p(x) H(Y|X=x)$$

$$= H(Y) - H(\mathbf{r}) \sum_{x \in \mathcal{X}} p(x) = H(Y) - H(\mathbf{r}) \leq \log|\mathcal{Y}| - H(\mathbf{r}).$$

- Η ισότητα ισχύει όταν η Y ακολουθεί ομοιόμορφη κατανομή.
- Ομοιόμορφη κατανομή για την Y επιτυγχάνεται με χρήση ομοιόμορφα κατανεμημένης εισόδου X.

$$p(y) = \sum_{x \in \mathcal{X}} p(y|x)p(x) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} p(y|x) \stackrel{(a)}{=} c\frac{1}{|\mathcal{X}|} = \frac{1}{|\mathcal{Y}|}.$$

Στο (a) χρησιμοποιήθηκε το γεγονός ότι τα αθροίσματα των στοιχείων κάθε στήλης ισούνται μεταξύ τους.

Από Κοινού Τυπικότητα και Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (**Joint AEP**)

- Συμμετρικά Κανάλια και Χωρητικότητα
- Από Κοινού Τυπικότητα και Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP)

Από κοινού τυπικές ακολουθίες (**Jointly Typical sequences**)

- Ορισμός: Το σύνολο των από κοινού τυπικών ακολουθιών $A_\epsilon^{(n)}$ ως προς την κατανομή $p(x, y)$, ορίζεται ως

$$\begin{aligned} A_\epsilon^{(n)} = \Big\{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : & \left|-\frac{1}{n}\log p(x^n) - H(X)\right| < \epsilon, \\ & \left|-\frac{1}{n}\log p(y^n) - H(Y)\right| < \epsilon, \\ & \left|-\frac{1}{n}\log p(x^n, y^n) - H(X, Y)\right| < \epsilon\Big\}, \end{aligned}$$

περιέχει, δηλαδή, τα ζεύγη ακολουθιών $\{(x^n, y^n)\}$ μήκους n οι εμπειρικές εντροπίες των οποίων βρίσκονται σε απόσταση το πολύ ϵ από την πραγματική τους εντροπία.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (**Joint AEP**)

- Έστω (X^n, Y^n) ακολουθίες μήκους n οι οποίες δημιουργούνται με χρήση ανεξάρτητων και όμοια κατανεμημένων (i.i.d) ζευγών (X_i, Y_i), σύμφωνα με την κατανομή $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$. Ισχύουν οι ιδιότητες:
  1. $\Pr\{(X^n, Y^n) \in A_\epsilon^{(n)}\} \to 1$, για $n \to \infty$.
  2. $\left|A_\epsilon^{(n)}\right| \leq 2^{n(H(X,Y)+\epsilon)}$.
  3. Εάν $(\tilde{X}^n, \tilde{Y}^n) \sim p(x^n)p(y^n)$, δηλαδή οι $\tilde{X}^n$ και $\tilde{Y}^n$ είναι ανεξάρτητες και οι κατανομές τους ισούνται με τις περιθώριες κατανομές της $p(x^n, y^n)$,

$$\Pr\left\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\right\} \leq 2^{-n(I(X;Y)-3\epsilon)}.$$

  Επίσης, υπάρχει n_0 τέτοιο ώστε, για $n > n_0$,

$$\Pr\left\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\right\} \geq (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}.$$

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης – Αποδείξεις

1. $\Pr\{(X^n, Y^n) \in A_\epsilon^{(n)}\} \to 1$, για $n \to \infty$.
   Από τον ασθενή νόμο των μεγάλων αριθμών, $-\frac{1}{n}\log p(X^n) \to -E[\log p(X)] = H(X)$ κατά πιθανότητα. Επομένως, για δεδομένο $\epsilon > 0$, υπάρχει n_1 τέτοιο ώ-στε, για όλα τα $n > n_1$, $\Pr\left\{\left|-\frac{1}{n}\log p(X^n) - H(X)\right| \geq \epsilon\right\} < \frac{\epsilon}{3}$. Παρομοίως, υπάρχουν n_2 και n_3 τέτοια ωστε, $\Pr\left\{\left|-\frac{1}{n}\log p(Y^n) - H(Y)\right| \geq \epsilon\right\} < \frac{\epsilon}{3}$ και $\Pr\left\{\left|-\frac{1}{n}\log p(X^n, Y^n) - H(X,Y)\right| \geq \epsilon\right\} < \frac{\epsilon}{3}$, αντίστοιχα. Επομένως, για $n > \max\{n_1, n_2, n_3\}$, η πιθανότητα το (X^n, Y^n) να μην είναι τυπικό είναι μικρότερη από ϵ, και, συνεπώς, $\Pr\{(X^n, Y^n) \in A_\epsilon^{(n)}\} > 1 - \epsilon$, για $n > \max\{n_1, n_2, n_3\}$.
2. $\left|A_\epsilon^{(n)}\right| \leq 2^{n(H(X,Y)+\epsilon)}$.
   Παρόμοια με την αντίστοιχη απόδειξη για το AEP,

$$1 = \sum_{(x^n,y^n)\in\mathcal{X}\times\mathcal{Y}} p(x^n, y^n) \geq \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} p(x^n, y^n) \geq \left|A_\epsilon^{(n)}\right| 2^{-n(H(X,Y)+\epsilon)}$$

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης – Αποδείξεις (συνέχεια)

3. Έστω ανεξάρτητες ακολουθίες τ.μ. $\tilde{X}^n$ και $\tilde{Y}^n$ και ότι οι από κοινού κατανομές $p(\tilde{x}^n)$ και $p(\tilde{y}^n)$ ισούνται με τις περιθώριες κατανομές $p(x^n)$ και $p(y^n)$, αντίστοιχα, της $p(x^n, y^n)$. Επομένως,

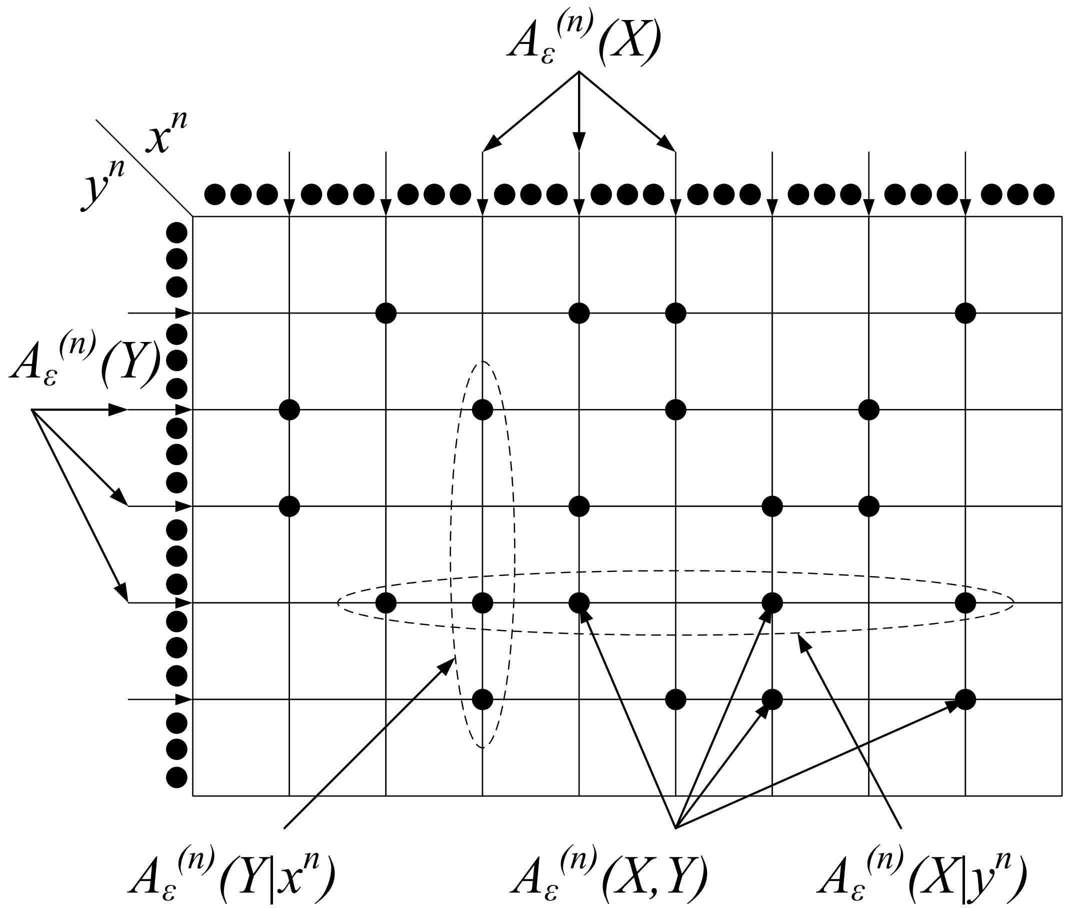
$$\begin{aligned}\Pr\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\} &= \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} p(x^n)p(y^n) \\ &\le 2^{n(H(X,Y)+\epsilon)}2^{-n(H(X)-\epsilon)}2^{-n(H(Y)-\epsilon)} \\ &= 2^{-n(I(X;Y)-3\epsilon)}.\end{aligned}$$

Με παρόμοιο τρόπο (βλ. π.χ. Cover Theorem 7.6.1) μπορεί να αποδειχθεί ότι

$$\Pr\left\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\right\} \ge (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}$$

για $n > n_0$.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (συνέχεια)



Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (συνέχεια)

- Από την 3η ιδιότητα, η πιθανότητα ένα ζεύγος (X^n, Y^n) ανεξάρτητων X^n και Y^n που επιλέγεται τυχαία να είναι και <u>από κοινού</u> τυπικό, ισούται περίπου με $2^{-nI(X;Y)}$.
- Επομένως, στο σχήμα της προηγούμενης διαφάνειας, εάν επιλέγουμε τα X^n και Y^n ανεξάρτητα, κατά μέσο όρο πρέπει να θεωρήσουμε περίπου $2^{nI(X;Y)}$ ζεύγη (X^n, Y^n) έως ότου εμφανιστεί ένα τυπικό ζεύγος.
- Ισοδύναμα, εάν θεωρήσουμε μια ακολουθία Y^n η οποία αποτελεί την έξοδο καναλιού με είσοδο X^n, υπάρχουν περίπου $2^{nH(X|Y)}$ υπό συνθήκη τυπικές ακολουθίες X^n. Η πιθανότητα να διαλέξουμε μια ακολουθία X'^n η οποία είναι από κοινού τυπική με την Y^n (αλλά όχι απαραίτητα η ακολουθία X^n η οποία μεταδόθηκε) ισούται, περίπου, με $2^{nH(X|Y)}/2^{nH(X)} = 2^{-nI(X;Y)}$.
- Επομένως, κατά μέσο όρο, για τις πρώτες $2^{nI(X;Y)}$ ακολουθίες X^n που θα επιλέξουμε τυχαία, θα είμαστε σίγουροι ότι δεν έχουν μεταδοθεί.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (συνέχεια)

- Συνεπώς, διαισθητικά, μπορούμε να μεταδώσουμε περίπου $2^{nI(X;Y)}$ διακριτές ακολουθίες στο κανάλι χωρίς να υπάρξει σύγχυση.
- Θα αποδείξουμε ότι είναι εφικτή η μετάδοση έως και $2^{nI(X;Y)}$ διακριτών ακολουθιών με αυθαίρετα μικρή πιθανότητα σφάλματος για $n \to \infty$. Θα αποδείξουμε, επίσης, ότι εάν προσπαθήσουμε να μεταδώσουμε περισσότερες από $2^{nI(X;Y)}$ διακριτές ακολουθίες, η πιθανότητα σφάλματος τείνει στο 1.
- Προκειμένου να μεταδώσουμε όσο το δυνατόν περισσότερη πληροφορία, μας συμφέρει να χρησιμοποιήσουμε την κατανομή εισόδου $p(x)$ η οποία μεγιστοποιεί την $I(X;Y)$.

Ανακεφαλαίωση μαθήματος

- Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP).
  - Επέκταση της Ιδιότητας Ασυμπτωτικής Ισοδιαμέρισης.
  - Για $n \to \infty$ ο αριθμός ακολουθιών Y^n που "σχετίζονται" με μια ακολουθία X^n όταν οι $(X, Y) \sim p(X, Y)$, τείνει στο $2^{nH(Y|X)}$. Επομένως, μπορούμε να βρούμε περίπου $2^{nH(Y)}/2^{nH(Y|X)} = 2^{nI(X;Y)}$ ακολουθίες X^n που θα απεικονιστούν σε διαφορετικές περιοχές στο σύνολο $\mathcal{Y}^n$.

Προεπισκόπηση επόμενου μαθήματος

- Θα αποδείξουμε το Θεώρημα Κωδικοποίησης Καναλιού (για Διακριτά Κανάλια Χωρίς Μνήμη)
  - Ευθύ: Για $n \to \infty$, η πιθανότητα να αποκωδικοποιήσουμε λάθος σύμβολο ή να μη μπορούμε να αποκωδικοποιήσουμε τείνει στο 0, εφόσον $R < C = \max_{p(x)} I(X;Y)$.
  - Αντίστροφο: Θα δείξουμε ότι δεν υπάρχει κώδικας με $P_e^{(n)} \to 0$ που να επιτυγχάνει $R > C$.
- Χωρητικότητα Διακριτών Καναλιών Χωρίς Μνήμη με Ανάδραση.
- Βέλτιστη μέθοδος αποκωδικοποίησης (με χρήση Μέγιστης Πιθανοφάνειας – Maximum Likelihood).
- Εκθέτης Σφάλματος (Error Exponent): Παρέχει ένα άνω φράγμα για το σφάλμα αποκωδικοποίησης με χρήση ML για δεδομένο μήκος κώδικα n.