

تئوری اطلاعات : جلسہ 8 و 8 28 دسمبر 1399

* مطالب امروز

* نامی Shearer ← دو مثال

* شروع بحث کوڈنگ منبع (Source coding)

• تعریف مقدماتی

• Non singular codes

• uniquely decodable codes

• Instantaneous codes (prefix free codes)

• Kraft نامی

نامساوی Sheaver :

فرض کنید مجموعه‌ای از زیر مجموعه‌های $[n] = \{1, \dots, n\}$

یک بردار صحافی (X_1, \dots, X_n) داریم. هر $i \in [n]$ در حداقل r تا از این‌ها
 * نامساوی زیر قرار است :
 هر قلمرو است.

$$H(X_1, \dots, X_n) \leq \frac{1}{r} \sum_{F \in \mathcal{F}} H(X_F)$$

$$\mathcal{F} = \{\{1\}, \dots, \{n\}\}$$

* حالت خاص :

$$\Rightarrow H(X_1, \dots, X_n) \leq \frac{1}{r} \sum_{F \in \mathcal{F}} H(X_F)$$

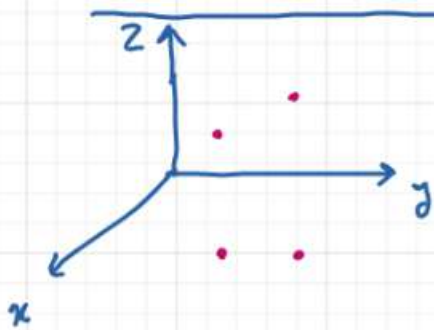
$$= H(X_1) + \dots + H(X_n)$$

Subadditivity

$$F = \{ [n] \setminus i : \forall i \in [n] \} \quad : \text{Man } \underbrace{\quad}_{\text{symmetrisch}} *$$

$$\Rightarrow H(X_1, \dots, X_n) \leq \frac{1}{n-1} \sum_{i=1}^n H(X_{[n] \setminus i})$$

$$n=3 : H(X_1, X_2, X_3) \leq \frac{1}{2} [H(X_1, X_2) + H(X_1, X_3) + H(X_2, X_3)]$$



مثال 1: n نقطه در فضای \mathbb{R}^3 $= S$

n_1 : تقسیم نقاط S بر روی صفحه $x=0$

n_2 : $y=0$

n_3 : $z=0$

تضمین: رابطه زیر برقرار است: $n^2 \leq n_1 \cdot n_2 \cdot n_3$

فرض کنیم n نقطه را به صورت n کد یا مقادیر انتخاب کنیم. و مقادیر آن نقطه را با x, y, z نشان می دهیم. (نقطه انتخاب شده را با P نشان دهیم که $P \in S$).
 توزیع یکدافت روی S

$$h_2^n = H(P) = H(X, Y, Z)$$

$$\leq \frac{1}{2} [H(X, Y) + H(Y, Z) + H(X, Z)]$$

$$\leq \frac{1}{2} [h_2^{n_3} + h_2^{n_1} + h_2^{n_2}]$$

مثال 2 (نمایش مقدار مثلث‌ها در یک گراف) : $G(V, E)$: undirected graph

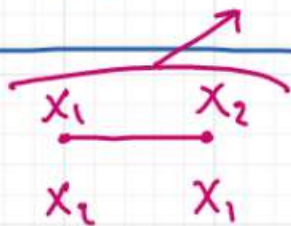
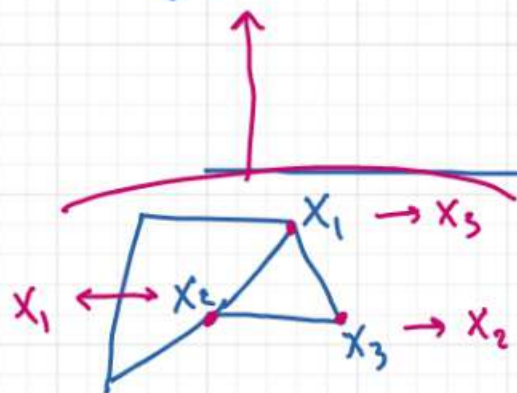
تعداد مثلث‌های این گراف $t =$ ، $l = |E|$

$$t \leq \frac{1}{6} (2l)^{\frac{3}{2}} \quad \text{محدود:}$$

فرض کنیم: $x_1, x_2, x_3 \in V$ که یک مثلث می‌سازند. این سه رأس را به گره‌های

انتخابی کنیم که مثلث‌ها توزیع یکنواخت داشته باشند :

$$h_2(6t) = H(x_1, x_2, x_3) \leq \frac{1}{2} [H(x_1, x_2) + H(x_2, x_3) + H(x_1, x_3)] \\ \leq \frac{1}{2} [3 h_2(2l)]$$



* که گذاری منبع: فصل 5 کتاب Cover

$$X \sim P_X$$

$$\hookrightarrow EX$$

یک منبع اطلاعات که باید متغیر تصادفی مدل ی شود:

تعریف: یک که منبع C برای متغیر تصادفی X یک نگاشت از X

به D^* (یک منبع به طول متناهی از بیل های الفبایی D است).

$$C: X \rightarrow D^*$$

$$X = \{ \text{آبی و قرمز} \}$$

مثال:

$$C(\text{قرمز}) = 001 \quad D = \{0, 1\}$$

$$C(\text{آبی}) = 111$$

طول کلمه که (codeword) هر یک از $x \in X$ ؛ $l(x)$

تعریف: طول متوسط $L(C)$ برای متغیر تصادفی X به فرم $P(x)$

$$L(C) = \sum_{x \in X} P(x) \cdot l(x)$$

(بدون از دست رفتن کلیت: $P = \{0, 1, \dots, D-1\}$)

$$P(X=1) = \frac{1}{2}$$

$$P(X=2) = \frac{1}{4}$$

$$P(X=3) = \frac{1}{8}$$

$$P(X=4) = \frac{1}{8}$$

$$C(1) = 0$$

$$C(2) = 10$$

$$C(3) = 110$$

$$C(4) = 111$$

Codewords

: \log_2

$$L(C) = 1.75 \quad \text{D; / D; } \quad , \quad D = \{0, 1\}$$

$$H(X) = 1.75 \text{ bits}$$

* تعریف: یک کد C های گوییم non singular است اگر $x_1, x_2 \in X, x_1 \neq x_2$

$$\Rightarrow C(x_1) \neq C(x_2)$$

* در عمل می‌خواهیم یک دنباله x_1, x_2, \dots, x_n را کد کنیم:

① راه ساده: یک کاراکتر اضافه که در D نیست مثل ویرگول معرفی کنیم:

$$C(x_1 \dots x_n) = C(x_1), C(x_2), \dots, C(x_n)$$

یک روش غیر جید است.

② سرانگ یا فنز کردهای دیگری داریم: کردهای لحظاتی یا instantaneous

prefix (free) code

تعریف: کد توکده C که با C^* تائیدی داریم:

$$C^*(x_1 x_2 \dots x_n) = C(x_1) C(x_2) \dots C(x_n)$$

$$C(x_1) = 00$$

مثال:

$$C(x_2) = 10$$

$$C^*(x_1 x_2) = 0010$$

تعریف: کدهای "یکتا دیکو" (uniquely decodable codes)
کدها

$$x_1 x_2 \dots x_m \neq y_1 y_2 \dots y_n \quad \text{اگر}$$

$$\Rightarrow C^*(x_1 \dots x_m) \neq C^*(y_1 \dots y_n)$$

$$\Leftrightarrow C(x_1) \dots C(x_m) \neq C(y_1) \dots C(y_n)$$

یعنی که توسعه یافته nonsingular باشد.

* تعریف: کدهای لحظه ای (instantaneous) \subseteq prefix free codes

کدهایی هستند که یک کلمه که زیر دنباله ای (prefix) از کلمه دیگری نباشد.

$$\mathcal{X} = \{1, 2, 3, 4\}$$

$$\mathcal{D} = \{0, 1\}$$

: 1 0

$$C(1) = 0$$

: 1 0

$$C(2) = 0$$

$$C(3) = 0$$

$$C(4) = 0$$

\Rightarrow Singulare : $1 \neq 2 \cup C(1) = C(2)$

$$C(1) = 0$$

: 2 0

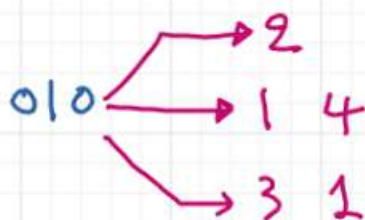
$$C(2) = 010$$

$$C(3) = 01$$

$$C(4) = 10$$

• non singular : $x_1 \neq x_2 \Rightarrow C(x_1) \neq C(x_2) : \forall x_1, x_2 \in \mathcal{X}$

• not uniquely decodable



$$C(1) = 10$$

3 : 010

$$C(2) = 00$$

$$C(3) = \underline{11}$$

Prefix

$$C(4) = 110$$

- non singular
- not a prefix free code
- uniquely decodable

$$C(1) = 0$$

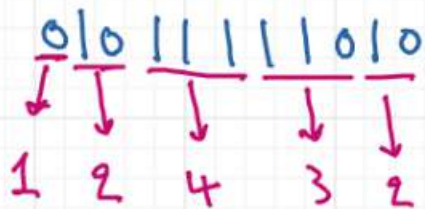
: 4 010

$$C(2) = 10$$

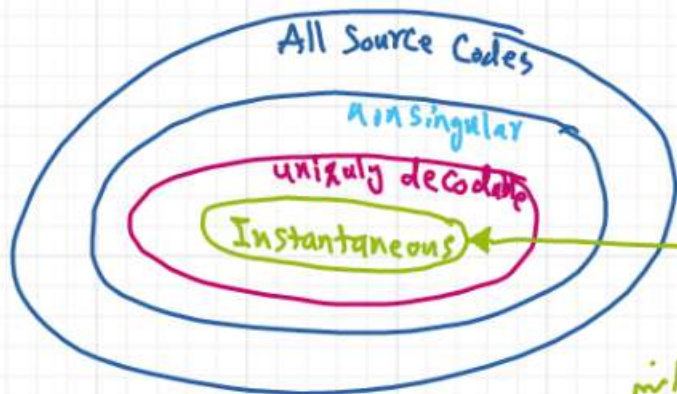
$$C(3) = 110$$

$$C(4) = 111$$

- non singular
- uniquely decodable
- prefix free



رابطه اعداد که های
منبع می باشد



که های قوی
هستند که

از نظر طول هم می توانیم

در بیشتر آن یک که بهیچید بیدار کنیم.