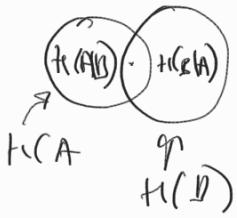
§ Энтропия графа

$I(A,B:C) = I(A:C) + I(B:C|A)$



Утв. $H(A|f(B)) \geq H(A|B)$

$H(A|B) = H(A|B,f(B))$

Сл-ие $I(X:f(Y)) \leq I(X:Y)$

$H(X) - H(X|f(Y)) \qquad H(X) - H(X|Y)$

Утв $I(X,f(X):Y) = I(X:Y) + I(f(X):Y|X)$

$I(X:Y) + 0$

$H(f(X)|X) - H(f(X)|X,Y)$

$H(f(X)|X) = 0$

$H(f(X)|X,Y) = 0$

G – неориент. граф V – вершин, E – мн-во ребер

$S \subseteq V$ S – независ., если между вершин. S нет ребер.

$$H(G) = \min_{X,Y} I(X:Y) \qquad (|V| = n$$

X – равном. распр. на V

Y – случ. величина, независимая мн-во, содерж. X

Примеры 1) $E = \emptyset$ $\quad H(G) = 0$

$Y = V$

$I(X:Y) = H(X) - H(X|Y) = 0$

$\log n \qquad \log n$

2) G полный граф на n вершинах

$Y = \{X\}$

$I(X:Y) = H(X) - H(X|Y)$

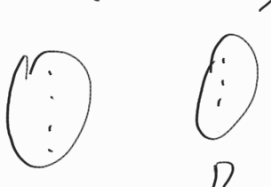
$\log n \qquad 0$

$H(K_n) = \log n.$

3) $G = K_{n,n}$

$H(K_{n,n}) = 1$

$Y = \begin{cases} A, & X \in A \\ B, & X \in B \end{cases}$



$H(K_{n,n}) \leq I(X:Y) = H(X) - H(X|Y) =$

$\log 2n \qquad \log n$

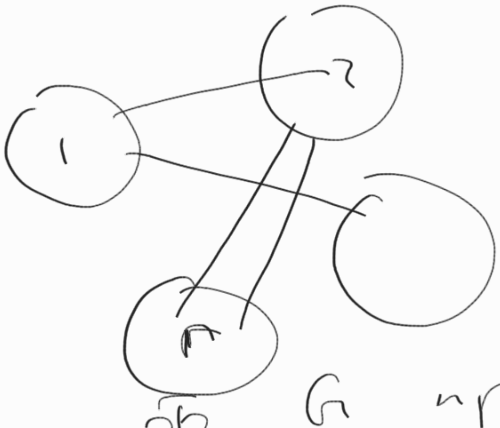
$= \log 2n - \log n = 1$

$H(K_{n,n}) \le 1$ $\qquad \left[\begin{array}{l} Y \subseteq A \\ Y \subseteq B \end{array}\right.$

$I(X:Y) = H(X) - H(X|Y) \ge 1$

$H(X) = \log 2n$, $\quad H(X|Y) \le \log n$

$\Rightarrow H(K_{n,n}) \ge 1$ $\qquad \log r$



Пусть граф G прав. образом красится в r цветов

X – случ. верш. $\qquad Y$ – все вершины цвета X $\qquad$ вершины X

$\ell(X)$ – цвет

$I(X:Y) = I(X, \ell(X) : Y) =$

$= \underbrace{I(\ell(X):Y)}_{} + \underbrace{I(X:Y|\ell(X))}_{\ge 0}$

$I(\ell(X):Y) = H(\ell(X)) - \underbrace{H(\ell(X)|Y)}_{=0}$

$I(X:Y|\ell(X)) = H(X|\ell(X)) - \underbrace{H(X|\ell(X), Y)}_{=0}$

$I(X:Y) = H(\ell(X)) \le \log r$

G r-раскраш. $\Rightarrow H(G) \le \log r$

Утв. Если $H(\ell(X))$

$U_1 \quad U_2 \quad \ldots \quad U_r$

$\frac{|U_1|}{|U|} \quad \frac{|U_2|}{|U|} \quad \ldots \quad \frac{|U_r|}{|U|}$

Утв. Если G – полный r-дольный граф, то $H(G) = H\left(\frac{|U_1|}{|U|}, \frac{|U_2|}{|U|}, \ldots, \frac{|U_r|}{|U|}\right)$

$I(X:Y) \ge H(\ell(X))$

Св-ва энтропии графа.

(1) Полуаддитивность $G_1(V, E_1)$ $G_2(V, E_2)$
$G = G(V, E_1 \cup E_2)$. Тогда $H(G) \le H(G_1) + H(G_2)$

Д-во $H(G_1) = I(X : Y_1)$
$H(G_2) = I(X : Y_2)$

$Y = Y_1 \wedge Y_2$

$H(G) \le I(X : Y_1 \wedge Y_2) \le I(X : Y_1, Y_2) =$
$= H(Y_1, Y_2) - H(Y_1, Y_2 | X) = H(Y_1, Y_2) - H(Y_1|X) - H(Y_2|X)$
$\le H(Y_1) + H(Y_2) - H(Y_1 | X) - H(Y_2 | X) =$
$= H(Y_1) - H(Y_1|X) + H(Y_2) - H(Y_2|X) = I(X : Y_1) +$
$+ I(X : Y_2) = H(G_1) + H(G_2)$

Пр-р K_{2^n} $V = \{0,1\}^n$ $G_i(V, E_i)$
$G_1, G_2 - - G_n$

$E_i = \{(u, w) \mid u_i \ne w_i\}$
$H(K_{2^n}) = n$ $H(G_i) = 1$
т.к. G_i - полный двудольный

$G = G_1 \cup G_2 - - G_n$
$H(G) = H(G_1) + - - + H(G_n)$

(2) монотонность
$E \subseteq E'$ $G(V, E)$ $G'(V, E')$
Тогда $H(G) \le H(G')$
$\Vert$
$I(X : Y)$
$\Rightarrow H(G) \le I(X : Y)$

(3) Св-во двудольного объединения

$G_1, G_2, \dots, G_k$ — компоненты связности графа G. $G_i(V_i, E_i)$ $\quad p_i = \frac{|V_i|}{|V|}$

Тогда $H(G) = \sum_{i=1}^{k} p_i H(G_i)$

$\forall v \in V$, $\ell(v) = i$: $V = \bigcup_{v \in V_i} V_i$

Д-во $H(G) = I(X : Y)$ $\quad Y_i = Y \cap V_i$

$Y = Y_1 Y_2 \dots Y_k$

$H(G) = I(X : Y) = I(X : Y_1 Y_2 \dots Y_k) =$

$= I(X, \ell(X) : Y_1, Y_2, \dots, Y_k) =$ (цепочка)

$= I(\ell(X) : Y_1, Y_2, \dots, Y_k) + I(X : Y_1, Y_2, \dots, Y_k \mid \ell(X))$

$\geq I(X : Y_1, Y_2, \dots, Y_k \mid \ell(X)) =$

$= \sum_{i=1}^{k} \underbrace{\Pr[\ell(X) = i]}_{p_i} \cdot I(X : Y_1 Y_2 \dots Y_k \mid \ell(X) = i) =$

$= \sum_{i=1}^{k} p_i \left(I(X : Y_i \mid \ell(X) = i) + I(X : Y_1 Y_2 \dots Y_{i-1} Y_{i+1} \dots Y_k \mid \ell(X) = i) \right)$

$\geq \sum_{i=1}^{k} p_i \, I(X : Y_i \mid \ell(X) = i) \geq \sum_{i=1}^{k} p_i \cdot H(G_i)$

$H(G_i) = I(X_i : Y_i)$

$(X_1, Y_1) \quad (X_2, Y_2) \quad \dots \quad (X_k, Y_k)$

$X_1 Y_1 \quad X_2 Y_2 \quad \dots \quad X_k Y_k$ — независимо

$i \begin{bmatrix} 1 & p_1 \\ 2 & p_2 \\ \vdots & \\ k & p_k \end{bmatrix}$ $\quad X$ $\quad Y = \bigcup Y_i$

Выдаёт X_i $\quad Y_i$

$$H(G) \leq I(X : Y) = I(X : y_1, y_2, \ldots, y_k) =$$

(цепочка)

$$= I(X, \ell(X) : y_1, y_2, \ldots, y_k) =$$

$$= I(\ell(X) : y_1, y_2, \ldots, y_k) + I(X : y_1, y_2, \ldots, y_k \mid \ell(X)) =$$

(red: first term $= 0$)

$$\geq I(X : y_1, y_2, \ldots, y_k \mid \ell(X) = i) =$$

$$= \sum_{i=1}^{k} \underbrace{\Pr[\ell(X) = i]}_{p_i} \cdot I(X : \underbrace{y_1, y_2, \ldots, y_k}_{y_i,\ y_1 y_2 \ldots y_{i-1} y_{i+1} \ldots y_k} \mid \ell(X) = i) =$$

C

$$= \sum_{i=1}^{k} p_i \left(I(X : y_i \mid \ell(X) = i) + I(X : y_1 y_2 \ldots y_{i-1} y_{i+1} \ldots y_k \mid \ell(X) = i) \right)$$

C A ... A ... B (red: $X \to X_i$, second term $= 0$)

$$\geq \sum_{i=1}^{k} p_i\, I(X : y_i \mid \ell(X) = i) \geq \sum_{i=1}^{k} p_i \cdot H(G_i)$$

$$I(X_i : y_i) = H(G_i)$$

Нижняя оценка для совершенного хеширования

<u>Опр.</u> (Семейство k-соверш. хеш-функций) $\mathcal{H}$ k-соверш.

$$\mathcal{H} = \{ h : [N] \to [b] \}$$

сем-во $\quad \forall S \subseteq [N]$ и $|S| = k$ $\quad \exists$

$h \in \mathcal{H}$: h инъективно на S.

$b \ll N$

<u>Цель</u> Оценить снизу $|\mathcal{H}|$

<u>Утв</u> $\mathcal{H}$ 2-соверш. сем-во. Тогда

$$|\mathcal{H}| \geq \frac{\log N}{\log b}$$

Д-во $\mathcal{H} = \{h_1, h_2, \dots, h_t\}$

$x_1, x_2 \in [N]$ $(h_1(x_1), h_2(x_1), \dots, h_t(x_1)) \neq$
$(h_1(x_2), h_2(x_2), \dots, h_t(x_2))$

$\Rightarrow b^t \geq N \Rightarrow t \log b \geq \log N$

$\Rightarrow t \geq \frac{\log N}{\log b}$

Утв. $b \geq 10k^2 \Rightarrow \exists$ k-совер. сем-во хеш-функц. рад. $t = O(k \cdot \log N)$

Д-во Выберем t хеш. функций случайно.

$S \subseteq [N]$ $|S| = k$. h - случ. функ

$\Pr\left[h \text{ инъективна на } S\right] = \frac{b(b-1)\cdots(b-k+1)}{b^k}$

$= \frac{b}{b} \cdot \frac{b-1}{b} \cdots \frac{b-k+1}{b} \geq \left(\frac{b-k}{b}\right)^k = \left(1 - \frac{k}{b}\right)^k \geq$

$\geq 1 - \frac{k^2}{b} \geq \frac{9}{10}$.

$(1+x)^k \geq 1 + kx \quad x \geq -1$

$\Pr\left[\text{все } t \text{ функций не инъективны на } S\right]$

$\leq \frac{1}{10^t}$

$\Pr\left[\exists\, S \subseteq [N] \text{ и } |S| = k : \text{все } t \text{ ф-ций не инъективны на } S\right] \leq \frac{N^k}{10^t} < 1$

$t = O(k \log n)$

$\mathcal{H}$ сем-во k-содерж. ф-ции.

$$|\mathcal{H}| \geq \frac{b^k}{b(b+1)\cdots(b+k+1)} \frac{\log(N-k+2)}{\log(b+k+2)}^{\frac{\log N}{\log b}}$$