שיעור 5:

תורת האינפורמציה:

אי שוויון Jensen: אם f פונק' קמורה, אז $E[f(x)] \le f(E[x])$.

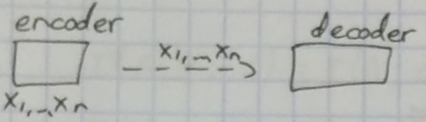
עבר על: (1) אנטרופיה. (2) אינפורמציה הדדית. (3) KL Divergence.

אנטרופיה: עבור משתנה x מעל Ω, נגדיר

$$H(x) := \sum_{x\in\Omega} P(x)\cdot\log\frac{1}{P(x)} \qquad \left(\text{תמיד } H(x)\ge 0\right)$$

תכונות האנטרופיה:

משפט source coding של שנון:



המשפט: $\lim_{n\to\infty} \frac{\text{אורך קידוד } x_1,\dots,x_n}{n} = H(x)$ — n חזרות בלתי תלויות משתנים, $n\to\infty$.

עבור ערוץ יחיד, קידוד הופמן משיג $\le H(x)+1$, שזה כמעט אופטימלי.

במילה קידוד מתכוונים לתוחלת מספר הביטים (לא למקסימום).

אנטרופיה מותנית: $H(X|Y) := E_{y\sim\mu(y)}\left[H(X|_{Y=y})\right]$

(1) $H(X) \ge H(X|Y)$.

(2) $H(X) = H(X|Y)$ אמ"מ X,Y ב"ת.

דוגמה: X,Y מתפלגים כך:

Y \ X	1	2
1	$\frac{3}{4}$	0
2	$\frac{1}{8}$	$\frac{1}{8}$

$$H(X) = \frac{7}{8}\cdot\log\left(\frac{8}{7}\right) + \frac{1}{8}\cdot\log(8) = \frac{7}{8}\cdot\log\left(1+\frac{1}{7}\right) + \frac{1}{8}\cdot 3 \approx$$
$$\approx \frac{7}{8}\cdot\frac{1}{7} + \frac{3}{8} = \frac{1}{2}$$

$$H(X|Y=1) = 0$$

$$H(X|Y=2) = \frac{1}{2}$$

$$H(X|Y) = \frac{3}{4}\cdot 0 + \frac{1}{4}\cdot 1 = \frac{1}{4}$$

למה: מכל ההתפלגויות על k ערכים, ההתפלגות האחידה משיגה אנטרופיה מקסימלית, והיא $\log(k)$.

אינפורמציה הדדית (mutual info.):

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

המשמעות: חוסר הודאות של X פחות חוסר הודאות כש-Y ידוע.
כלומר, "מה שלמדנו" מחוסר הודאות מידיעת Y.

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$$

תכונות:

(1) $I(X;Y|Z) \geq 0$.

(2) $I(X;Y|Z)=0$ אמ"מ X,Y ב"ת בהינתן Z.

(3) חוק השרשרת (chain rule):

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

$$I(A_1,\dots,A_m;B) = \sum_{i=1}^{m} I(A_i;B|A_{<i})$$

(4) $I(A;B) \gtrless I(A;B|C)$ לא ניתן להשוואה.

דוגמא: $A=B=C$ מטבע הוגן.

$$I(A;B) > I(A;B|C)$$

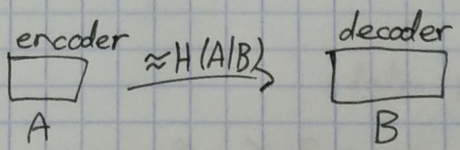
$$I(A;B) = H(\tfrac{1}{2}) - 0 = 1$$

$$I(A;B|C) = H(A|C) - H(A|B,C) = 0-0=0$$

דוגמא: A,B ב"ת $\frac{1}{2}$, $C = A \oplus B$ (xor).

$$0 = I(A;B) < I(A;B|C) = 1$$

פירוש אופרטיבי:



רוצים לשלוח A, ה-decoder יודע B (ה-encoder יודע את ההתפלגות של B.)
צריך לשלוח בערך $H(A|B)$ ביטים.

נניח עכשיו ששני הצדדים יכולים לדבר בשני הכיוונים, וה-decoder רוצה לדעת את A. אז צריך בערך $I(A;B)$ ביטים.

(נראה את זה בהמשך בצורה מסודרת יותר).

KL-Divergence: (Kullback-Leibler)

$$D(\mu\|\eta) = \sum_x \mu(x)\cdot\log\left(\frac{\mu(x)}{\eta(x)}\right)$$

expected log-ratio

μ התפלגות נכונה.

התפלגות שגויה.

משמעות: אם חשבנו שההתפלגות היא η, אז (אולי אחרי דחיסה) גילינו שההתפלגות היא μ, כמה נהיה מופתעים.

תכונות:

(1) $D(\mu\|\eta)\geq 0$. כי התפלגות לפי μ, אז יש יותר משקל כש- $\eta\geq\mu$.

(2) אם קיים x כך ש- $\eta(x)=0$, $\mu(x)\neq 0$, אז $D(\mu\|\eta)=\infty$.

(אם $\mu(x)=\eta(x)=0$ סוכמים 0).

פירוש אופרטיבי: נניח שבנינו קוד אופטימלי C עבור η, אבל x מגיע מההתפלגות μ.

$$\mathop{E}_{x\sim\eta}[C(x)] \approx H(\eta)$$

$$\mathop{E}_{x\sim\mu}[C(x)] \approx H(\mu)+D(\mu\|\eta)$$

תוספת המחיר לקידוד הוא ה- divergence.

בפרט, עבור $\eta=\mu$, $D(\mu\|\eta)=0$.

אם קיים ערך עם $\eta(x)=0$, $\mu(x)\neq 0$, המחיר אינסופי, כי לפי η לא יודעים לקודד.

data processing inequality:

לכל פונקציה f ולכל Y כך ש- f,Y [illegible], מתקיים

$$I(X;Y)\geq I(f(X);Y).$$

$$I(X;Y) = \mathop{E}_{y\sim\mu(Y)}\left[D\left(\frac{\mu(X|Y=y)}{\mu(X)}\right)\right] = D\left(\frac{\mu(X,Y)}{\mu(X)\cdot\mu(Y)}\right)$$

כאשר μ התפלגות משותפת של X,Y. ($\mu(X),\mu(Y)$ הן ההתפלגויות השוליות).

$$D\left(\frac{P_1}{P_2}\right) := D(P_1\|P_2)$$

סימונים ואינפורמציה:

Π פרוטוקול רנדומי.

R אקראיות פומבית.

אליס	בוב
X	Y
R_A אקראיות פרטית.	R_B

$(X,Y)\sim\mu$

אינפורמציה חיצונית:

$$IC^{ext}(\Pi) = I(\Pi; X,Y \mid R)$$

מה משקיף חיצוני לומד על X,Y מ-Π, כשהוא רואה את R.

אינפורמציה פנימית:

$$IC^{int}(\Pi) = I(\Pi; Y \mid X, R_A, R) + I(\Pi, X \mid Y, R_B, R)$$

מה שאליס ובוב לומדים מ-Π על הקלט של הצד השני, בהינתן המידע שלהם.

טענה:

$$IC^{ext}(\Pi) \underset{\text{מהגדרה}}{=} H(\Pi) - (\dots) \le |\Pi|$$

כי מספיק $H(\Pi)$. $|\Pi| \ge H(\Pi)$.

ביטול הפומביות: אנחנו רוצים להוריד את הפומביות.

עכשיו נראה שאין אקראיות פומבית, כי ניתן לוותר על אקראיות פומבית ע"י כך שאליס תשלח לבוב ביטים אקראיים מ-R_A.

נניח שיש לנו Π שמשתמש באקראיות פומבית R, ונבנה Π' בו אליס שולחת את $\tilde{R}$ (משיגה מספיק ביטים ע"י R_A) המשתתפים משתמשים ב-$\tilde{R}$ במקום R. רוצים להראות: $IC^{ext}(\Pi) = IC^{ext}(\Pi')$.

$$IC^{ext}(\Pi') = I(\tilde{R}, \Pi; X,Y) = I(\Pi; X,Y \mid \tilde{R}) + \underbrace{I(\tilde{R}; X,Y)}_{=0} = IC^{ext}(\Pi)$$

כי $\tilde{R}$ ב"ת ב-X,Y.

אותה טענה נכונה גם לאינפורמציה פנימית.

עכשיו נרצה להראות שלא צריך גם את האקראיות הפרטית.

טענה: $IC^{int}(\Pi) = I(\Pi; Y \mid X) + I(\Pi; X \mid Y)$.

הוכחה: באינדוקציה על מספר ההודעות ב-Π.

בסיס: $\Pi = \varepsilon$. $0=0$.

(המשך בעמוד הבא)

המשך הערה:

לכן: נניח ש- $\Pi = \Pi' M$, כאילו שולחת את M.

עובדות שנשתמש בהן: (*) M בת"ל ב-Y בהינתן X, R_A, Π'.

(*) M בת"ל ב-Y בהינתן X, Π'.

(*) R_B בת"ל ב-X בהינתן Π, Y וגם בהינתן Π', Y.

מעבר להוכחה: נראה שאפשר להוריד את R_A, (ואותו דבר ל-R_B).

$$I(\Pi', M; Y \mid X, R_A) + I(\Pi', M; X \mid Y, R_B) =$$

$$= I(M; Y \mid \Pi', X, R_A) + I(\Pi'; Y \mid X, R_A) +$$
$$+ I(M; X \mid \Pi', Y, R_B) + I(\Pi'; X \mid Y, R_B) =$$

שלב ב'

$$= I(M; Y \mid \Pi', X, R_A) + I(M; X \mid \Pi', Y, R_B) +$$
$$+ I(\Pi'; Y \mid X) + I(\Pi'; X \mid Y) = (*)$$

$$I(M; X \mid Y, R_B, \Pi') = I(M, R_B; X \mid Y, \Pi') - \underbrace{I(R_B; X \mid Y, \Pi')}_{=0} =$$

$$= \underbrace{I(R_B; X \mid Y, \Pi', M)}_{=0} + I(M; X \mid Y, \Pi') = I(M; X \mid Y, \Pi')$$

כשאליס מדברת, היא לא מוסיפה אינפורמציה על Y, לכן $I(M; Y \mid —) = 0$.

$$IC^{int}(\Pi) = (*) = I(\Pi'; Y \mid X) + I(\Pi'; X \mid Y) + I(M; X \mid Y, \Pi') + I(M; Y \mid X, \Pi') =$$
$$= I(\Pi', M; X \mid Y) + I(\Pi', M; Y \mid X) =$$
$$= I(\Pi; X \mid Y) + I(\Pi; Y \mid X)$$

טענה: $IC^{int}(\Pi) \le IC^{ext}(\Pi)$.

מתבונן חיצוני לומד יותר, כי יש לו בהתחלה פחות אינפורמציה.

נוכיח את זה בתרגיל בית.

דוגמה: equality.

פרוטוקול Π: אליס שולחת את x לבוב.

התפלגות μ: Q מטבע הוגן. אם $Q=0$ אז X,Y בת"ל אחידים. אם $Q=1$ אז $X=Y$ אחיד.

$$IC_{\mu}^{ext}(\Pi)=I_{\mu}(\Pi;X,Y)=H(X)$$

$$IC_{\mu}^{int}(\Pi)=I_{\mu}(\Pi;X|Y)+\underbrace{I(\Pi;Y|X)}_{=0}\le$$

IPG - x בהינתן של XQ.

$$\le I(\Pi;X,Q|Y)=$$

$$=I(\Pi;X|Y,Q)+I(Q;\Pi|Y)=$$

$$=\frac{1}{2}\cdot\underbrace{I_{\mu|Q=0}(\Pi;X|Y)}+\frac{1}{2}\cdot\underbrace{I_{\mu|Q=1}(\Pi;X|Y)}_{=0}+\underbrace{I(Q;X|Y)}_{<1}\approx\frac{n}{2}+1$$

Π,x בת"ל בהינ' y, לכן $\frac{n}{2}=\frac{1}{2}\cdot I(X;X)$ $(\Pi=x)$.

$$I(X;X)=H(X)-H(X|X)=$$
$$=H(X)=n$$

נראה התפלגות בה שקיים Π עם $IC^{int}(\Pi)=O(1)$ לכל μ ושגיאה 0.

חסם תחתון ל- $C_0^{ext}(Eq)$ (שגיאה 0):

הגדרה:
$$IC_{\mu,\epsilon}^{ext}(f)=\inf_{\substack{\text{פרוטוקול } \Pi \\ \text{שפותר את } f \\ \text{עם שגיאה } \le\epsilon}} IC^{ext}(\Pi)$$

נגדיר μ אחידה מעל $\{(z,z)\mid z\in\{0,1\}^n\}$.

נניח שיש Π פותר Eq עם שגיאה 0, אז

$$I_{\mu}(\Pi;X,Y)=H(XY)-\underbrace{H(XY|\Pi)}_{=0}=H(XY)=n$$

נוכיח בהמשך

נראה: $H(XY|\Pi)=0$:

תזכורת: היינו בטוחים שאפשר לכתוב $\Pi_{xy}(t)=\pi(t|X=x,Y=y)=p(x,t)\cdot q(y,t)$.

טענה: לא קיימים $t,z'\neq z$, עבורם $\Pi_{z,z}(t)>0$, $\Pi_{z',z'}(t)>0$.

הטענה הנ"ל גוררת ש- $H(XY|\Pi)=0$.

נניח בשלילה שקיימים $t,z'\neq z$ כך ש- $\Pi_{z,z}(t)\neq0$, $\Pi_{z',z'}(t)\neq0$.

$$\Pi_{z,z}(t)=p(z,t)\cdot q(z,t)\neq0$$
$$\Pi_{z',z'}(t)=p(z',t)\cdot q(z',t)\neq0$$

בפרט, $p(z,t)\neq0$, $q(z',t)\neq0$.

לכן $\Pi_{z,z'}(t)=p(z,t)\cdot q(z',t)\neq0$.

כלומר, בהסתברות לא אפס, נקבל תשובה לא נכונה, סתירה.

שיעור יב

תזכורת: (1) $I(A;B|C) = H(A|C) - H(A|BC)$

$$= \mathop{E}_{b,c\sim\mu(B,C)}\left[D\,\frac{\mu(A|B=b, C=c)}{\mu(A|C=c)} \right]$$

(2) Jenson: אם f קמורה, אז $f(E[x]) \geq E[f(x)]$.

(3) chain rule: $H(A_1,\dots,A_m|B) = \sum_{i=1}^{m} H(A_i|A_{<i};B)$

$$I(A_1,\dots,A_m|B;C) = \sum_{i=1}^{m} I(A_i;B|A_{<i};C)$$

Equality (*): לכל פרוטוקול π שפותר Eq עם שגיאה 0,

$IC_{\mu}^{ext}(\pi) = I_{\mu}(\pi; XY) = \Omega(n)$.

$$IC_{\mu,\varepsilon}^{ext}(Eq) = \inf_{\substack{\pi \text{ כך ש} \\ \text{פותר } Eq \\ \text{עם שגיאה } \leq\varepsilon}} \left(IC_{\mu}^{ext}(\pi) \right)$$

$$IC_{\mu,\frac{1}{3}}^{ext}(Eq) \leq O(\log n)$$

(*) לכל μ, $IC_{\mu,0}^{int}(Eq) = O(1)$: $\left(I(X;\pi|Y) + I(Y;\pi|X) \right)$

נראה פרוטוקול: (נסמן π).

נבחר מטריצה הפיכה $A = (a_1, \dots a_n)$ מ- $\{0,1\}^{n\times n}$,

בעזרת אקראיות פומבית.

עבור $i=1,..,n$, אליס שולחת את $\langle x,a_i\rangle \bmod 2$.

בוב בודק אם $\langle x,a_i\rangle \equiv \langle y,a_i\rangle$; אם לא, עוצרים ומכריזים $x\neq y$.

אחרת מכריזים $x=y$.

מסימטריות, מספיק לחשב את $I(X;\pi|Y)$.

נסמן ב- Q את "האם $x=y$":

$$I(X;\pi|Y) \overset{\substack{(X \text{ הוא פונק' של } XQ)}}{\leq} I(XQ;\pi|Y) =$$

$$= I(X;\pi|QY) + \underbrace{I(Q;\pi|Y)}_{\leq H(Q)\leq 1} \leq$$

$$\leq 1 + P(Q=0)\cdot \mathop{I}_{\mu|x\neq y}(X;\pi|Y) + \underbrace{P(Q=1)\cdot \mathop{I}_{\mu|x=y}(X;\pi|Y)}_{=0}$$

נשאר להוכיח ש- $\mathop{I}_{\mu(x\neq y)}(X;\pi|Y)$ הוא קבוע. (המשך בעמוד הבא).

המשך Equality: נותר להוכיח ש- $\underset{\mu|x\neq y}{I}(X;\Pi|Y)$ הוא קבוע.

[הערה: כפי שראינו עד כה, אם $Ax=Ay$ ולכן $x=y$, כי A הפיכה.]

$$\underset{\mu|x\neq y}{I}(X;\Pi|Y)$$

$$=\underset{\mu|x\neq y}{I}(\Pi;X|YA)\le \underset{\mu|x\neq y}{H}(\Pi|YA)=\circledast$$

בהינתן Y,A, אנחנו יכולים לכתוב את התשובות כפונקציה של מספר ההודעות. הסבר: עד לפני ההודעה האחרון, אליס תשאל $\langle y,a_i\rangle$, כלומר: $\langle y,a_1\rangle,\langle y,a_2\rangle,\dots,\langle y,a_{i-1}\rangle,\ 1-\langle y,a_i\rangle$.

$$\circledast\le \underset{\mu|x\neq y}{H}(\underset{\text{מספר ההודעה}}{J}\,|\,Y,A)=$$

$$=\underset{y\sim\mu|x\neq y}{E}\left[\underset{a\sim A}{E}\left[\sum_{j=1}^{n}P(J=j|y,a)\cdot\log\left(\frac{1}{P(J=j|y,a)}\right)\right]\right]\le$$

jenson $f(x)=x\log(\frac{1}{x})$

$$\le\underset{y\sim\mu|x\neq y}{E}\left[\sum_{j=1}^{n}\left(\underset{a\sim A}{E}\left[P(J=j|y,a)\cdot\log\left(\frac{1}{P(J=j|y,a)}\right)\right]\right)\right]\le$$

$$\le\underset{y\sim\mu|x\neq y}{E}\left[\sum_{j=1}^{n}\left(2^{-j}\cdot\log(2^{j})\right)\right]\le$$

$$\le\underset{y\sim\mu|x\neq y}{E}\left[\sum_{j=1}^{n}2^{-j}\cdot j\right]\le 2$$

Disjointness: $\overline{Disj}(X,Y)=\bigvee_{i=1}^{n}(x_i\wedge y_i)$

$\boxed{\text{מעכשיו נסמן } IC=IC^{int}}$

שלב א': $IC^{int}(\overline{Disj})=n\cdot IC^{int}(AND)$ (—לא נוכיח לכל μ)

שלב ב': $\underset{\mu}{IC}^{int}(AND)=\Omega(1)$ לכל התפלגות μ

ואז נקבל $IC^{int}(\overline{Disj})=\Omega(n)$.

[illegible]: Direct Sum: יהי f פונק' ויהי $T(f^n,\varepsilon)$ המשימה הבאה לחשב את f על n עותקים, עם שגיאה שלכל היותר ε לכל עותק.

לכל ε,f,μ, $\underset{\mu^n}{IC}(T(f^n,\varepsilon))=n\cdot \underset{\mu}{IC}_{\varepsilon}(f)$. (כאשר יש התפלגות המכפלה.)

המשך disjointness:

<u>הוכחה למה:</u> [$\le$] בהינתן פרוטוקול Π עם שגיאה $\le \varepsilon$ עבור f, נבנה Π^n המריץ את Π על כל עותק בנפרד. למה: $\underbrace{IC_{\mu^n}(\Pi^n) \le n \cdot IC_\mu(\Pi)}_{\text{צריך להוכיח באינדוקציה על n.}}$.

ומתקיים שגיאה שלו $\le \varepsilon$.

[$\ge$] נניח שנתון Π עבור $T(f^n, \varepsilon)$.

נבנה Π' עבור עותק אחד של f. נסמן את הקלטים של בוב ואליס - $\overset{\text{בוב}}{V}, \overset{\text{אליס.}}{U}$.

נבחר $i \in [n]$ (פומבית) אקראי, $x_i \leftarrow U$, $y_i \leftarrow V$.

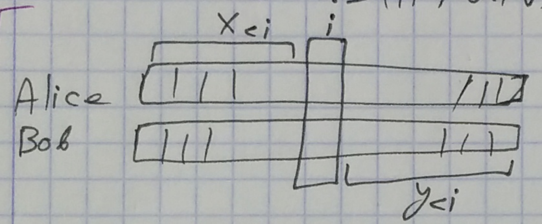
פתרון לא טוב: נגריל אקראיות (לא מ-μ) X_{-i}, Y_{-i} פומבית, ונריץ Π.

זה יתן שגיאה ε. אבל מספר יותר מדי אינפורמציה,

למשל עבור xor, נניח שפלט שמחזיר את $\bigoplus_{j=1}^{n} x_j$,

$I(\Pi; X|Y) = 1$, $I(\Pi; U|V) = 1 > \frac{1}{n}$.

הרעיון:



Π': נגריל i אקראי, $X_{<i}$, $Y_{>i}$ פומביים.

$x_i \leftarrow U$, $y_i \leftarrow V$, את $Y_{<i}$ בוב מגריל בהינתן $X_{<i}$ פרטי.
את $X_{>i}$ אליס מגרילה בהינתן $Y_{>i}$ פרטי.

משלימים X, Y.

ומריצים Π.

נוכיח:

$$I(\Pi'; U|V) + I(\Pi'; V|U) \le \frac{1}{n}\left(I(\Pi; X|Y) + I(\Pi; Y|X)\right)$$

מספיק להראות $I(\Pi'; U|V) \le \frac{1}{n} \cdot I(\Pi; X|Y)$ (החלק השני סימטרי).

כלומר:

$$I(i, X_{<i}, Y_{>i}, \Pi; X_i | Y_i) \le \frac{1}{n} \cdot I(\Pi; X|Y)$$

$$I(i, X_{<i}, Y_{>i}, \Pi; X_i | Y_i) \le I(i, X_{<i}, Y_{-i}, \Pi; X_i | Y_i) =$$
$$= I(\Pi; X_i | i, Y, X_{<i}) + \underbrace{I(i, X_{<i}, Y_{-i}; X_i | Y_i)}_{=0} =$$
$$= \sum_{i=1}^{n} \frac{1}{n} \underset{\mu_i}{I}(\Pi; X_i | Y, X_{<i}) =$$
$$= \frac{1}{n} \sum_{i=1}^{n} \underset{\mu}{I}(\Pi; X_i | Y, X_{<i}) = \frac{1}{n} \cdot I(\Pi; X|Y)$$

המשך disjointness: הערה לשלב א':

תהי μ התפלגות כך ש- $\mu(1,1)=0$ (כלומר ב- ~~[illegible]~~ $x_i \wedge y_i$),

אז: $IC_{\mu^n,\varepsilon}(Disj) = n \cdot IC_{\mu,\varepsilon}(AND)$

- אותה בניה כמו ב- direct sum.
- האינפורמציה יוצאת נכון.
- השגיאה?

$$\bigvee_{j=1}^{n} (x_j \wedge y_j) = u \wedge v$$

כי $x_j \wedge y_j = 0$ לכל $j \neq i$.

לכן השגיאות בסדר. (worst case).

שלב ב': נגדיר התפלגות: μ:

	0	1
0	$\frac{1}{3}$	$\frac{1}{3}$
1	$\frac{1}{3}$	0

טענה: אם Π פותר AND בשגיאה קטנה מספיק על כל קלט (כולל 1,1), אז $I(\Pi;X|Y) + I(\Pi;Y|X) = \Omega(1)$.

$$= \frac{2}{3} \cdot \underset{\mu|y=0}{I}(\Pi;X) + \underbrace{\frac{1}{3} \cdot \underset{\mu|y=1}{I}(\Pi;X)}_{=0} + \frac{2}{3} \cdot \underset{\mu|x=0}{I}(\Pi;Y) =$$

$$= \frac{2}{3} \cdot \underset{t\sim\Pi|y=0}{E}\left[D\left(\frac{\mu(x|\Pi=t, Y=0)}{\mu(x|Y=0)}\right)\right] + \frac{2}{3} \cdot \underset{t\sim\Pi|x=0}{E}\left[D\left(\frac{\mu(y|\Pi=t, X=0)}{\mu(y|X=0)}\right)\right] \geq$$

$\mu(x|y=0)$ — ברנולי $\frac{1}{2}$.

$$\Pi(t|Y=0) = \frac{1}{2}\Pi_{0,0}(t) + \frac{1}{2}\Pi_{1,0}(t) \quad , \quad \Pi(t|X=0) = \frac{1}{2}\Pi_{0,0}(t) + \frac{1}{2}\Pi_{0,1}(t)$$

$\Pi_{x,y}$ = התפלגות התמלול של Π כאשר $X=x$, $Y=y$

$$\geq \frac{2}{3} \cdot \frac{1}{2}\sum_t \Pi_{0,0}(t) \cdot \left[D\left(\frac{\mu(x|\Pi=t, Y=0)}{Ber(\frac{1}{2})}\right) + D\left(\frac{\mu(Y|\Pi=t, X=0)}{Ber(\frac{1}{2})}\right)\right]$$

סקיצה בינתיים: רוצים להראות שה-divergence גדול. צריך למצוא קבוצה של תמלולים T כך ש- $\Pi_{0,0}(t) \geq const$, ולכל $t \in T$, $p(X=0|\Pi=t, Y=0) \geq \alpha$ וגם $p(Y=0|\Pi=t, X=0) \geq \alpha$ עבור α קבוע גדול מחצי.

אם נמצא קבוצה כזאת נסכום רק עליה, וזה יתן מספיק גדול.

(לא מוכיח את הטענה, מפה נמשיך בפעם הבאה).

שיעור 7:

תזכורת: $\overline{Disj}(x,y)=\bigvee_{i=1}^{n}(x_i\wedge y_i)$.

הראינו שלכל μ, $\mu(1,1)=0$, אז $IC^{int}_{\mu^n,\varepsilon}(\overline{Disj})=n\cdot IC^{int}_{\mu,\varepsilon}(AND)$

נותר להוכיח ש- (1) $IC^{int}_{\mu,\varepsilon}(AND)=\Omega(1)$ עבור μ כזו.

מספיק להוכיח שלכל פרוטוקול Π עם שגיאה $\le\varepsilon$, קיימת קבוצה L של תמלילים כך ש- $\Pi_{00}(L)\ge const$.

(2) עבור קבוע $\frac{1}{2}<\alpha$ כלשהו, לכל $t\in L$, וגם $P(X=0|\Pi=t,Y=0)\ge\alpha$

וכן $P(Y=0|\Pi=t,X=0)\ge\alpha$

הוכחה: (המשך מהשיעור הקודם).

(1) בניית L: $L=\{t\mid output(t)=0\wedge\Pi_{0,0}(t)>4\cdot\Pi_{1,1}(t)\}$

נסמן את $\bar{L}$: B_1 = התמלילים עם פלט 1. אז $\Pi_{0,0}(B_1)\le\varepsilon$.

B_0 = התמלילים עם פלט 0, שבהם $\Pi_{0,0}(t)\le 4\cdot\Pi_{1,1}(t)$. אז:

$$\varepsilon>P(\overset{0}{1,1}\ \text{טעות})\ge\Pi_{1,1}(B_0)\ge\frac{1}{4}\cdot\Pi_{0,0}(B_0)\Rightarrow\Pi_{0,0}(B_0)\le 4\varepsilon$$

סה"כ: $\Pi_{00}(L)\ge 1-5\varepsilon$.

(2) יהי $t\in L$. נכתוב $\Pi_{x,y}(L)=q_A(t,x)\cdot q_B(t,y)$.

$$P(x=0|\Pi=t,Y=0)=\frac{P(x=0\wedge\Pi=t|Y=0)}{P(\Pi=t|Y=0)}=\frac{\overbrace{P(x=0|Y=0)}^{=\frac{1}{2}}\cdot\Pi_{0,0}(t)}{\frac{1}{2}\Pi_{0,0}(t)+\frac{1}{2}\Pi_{1,0}(t)}=$$

$$=\frac{\Pi_{0,0}(t)}{\Pi_{0,0}(t)+\Pi_{1,0}(t)}=\frac{q_A(0,t)\cdot q_B(0,t)}{q_A(0,t)\cdot q_B(0,t)+q_A(1,t)\cdot q_B(0,t)}=$$

$$=\frac{q_A(0,t)}{q_A(0,t)+q_A(1,t)}=\frac{1}{1+\frac{q_A(1,t)}{q_A(0,t)}}\underset{\text{רוצים}}{\ge}\alpha>\frac{1}{2}$$

כלומר, תלוי רק ביחס בין $q_A(1,t)$ ל- $q_A(0,t)$. באופן סימטרי, גם לבוב.

צריך: וגם ש- $\frac{q_A(1,t)}{q_A(0,t)}\le\beta<1$ וכן ש- $\frac{q_B(1,t)}{q_B(0,t)}\le\beta<1$.

לכל $t\in L$, $q_A(0,t)\cdot q_B(0,t)>4\cdot q_A(1,t)\cdot q_B(1,t)$.

$$\Updownarrow$$

$$\frac{q_A(1,t)}{q_A(0,t)}\cdot\frac{q_B(1,t)}{q_B(0,t)}<\frac{1}{4}$$

ולכן זה מתקיים.

חסם תחתון לאינדקס בעזרת אינפורמציה:

לאליס יש $x \in \{0,1\}^n$, לבוב יש $J \in [n]$, $Index(x,J) = x_J$.

אליס שולחת הודעה אחת M ובוב צריך להחזיר תשובה.

טענה: $\underset{\mu}{I}(M;x) = \Omega(n)$, אם M פותר בהסתברות $\geq 1-\varepsilon$.

זה נותן חסם כי $I(M;x) \le H(M) \le |M|$, לכן אליס צריכה לשלוח הרבה ביטים.

הוכחה: נסתכל על התפלגות אחידה. הפלט של בוב: $f: M,J \to x_J$.

$$H(x_J|M,J) \le \underbrace{\cancel{\quad}\, H(B_{1-\varepsilon})}_{\text{בהסתברות } 1-\varepsilon} = H(B_\varepsilon)$$

$$\text{קבוע} \ge H(x_J|M,J) = \sum_{j=1}^{n} \frac{1}{n} H(x_j|M)$$ (נמשיך אחר כך)

streaming: אותה שאלה: חישוב את F_0. תשובה מדויקת דורשת n ביטים. קירוב $\pm\varepsilon$: $\Omega(\frac{1}{\varepsilon^2})$ זכרון.

דחיסה: האם בהינתן פרוטוקול Π עם ~~[illegible]~~ תקשורת C ואינפורמציה $I \ll C$, ניתן לדחוס את Π לפרוטוקול Π' שהתקשורת שלו $O(I)$?

התשובה כמובן: (*) יודעים לדחוס ל-$2^{O(I)}$. אופטימלי אם אין תלות ב-C. (בדיוק כאשר $C \approx 2^{2^I}$, קולט 2^I).

כלומר, קיימת בעיה P והתפלגות $\mu \sim (X,Y)$, ~~[illegible]~~ ופרוטוקול Π עם $IC^{int}(\Pi) = I$, $CC(\Pi) = 2^{2^I}$, כך שלא ניתן לפתור את P בפחות מ-2^I תקשורת.

(*) אפשר לדחוס ל- $I^{ext} \cdot polylog(C)$.

(*) $\sqrt{I^{int} \cdot C}$

טענה: $IC^{int}(f) = \lim_{n\to\infty} \frac{CC(f^n)}{n}$.

ניתן לבנות דחיסה עם תקשורת $I^{int} + O(\text{מספר הסיבובים}) + o(I^{int})$.

(היה הסבר כיצד בונים ~~[illegible]~~ שהתעצלתי לכתוב).

דגימה ע"י דחיית דגימות: (זה היה חלק מההוכחה שלא כתבתי, ושימושי בפני עצמו).

התפלגות μ על U. נגריל אינסוף תווים מפולגים אחיד ב-$U \times [0,1]$. נבחר ראשון כך ש-$\mu(x_i) > p_i$ ונפלוט את x_i. נסמן ב-A_μ את השטח מתחת לעקומה של μ. עבור i נתון:

$$P((x_i,p_i) \in A_\mu) = \sum_{x \in U} \frac{1}{|U|} \cdot \mu(x) = \frac{1}{|U|}.$$

ההתפלגות המקבלת:

$$P(x) = \sum_{i=1}^{\infty} P(\text{לא נבחרו } x_1,\dots,x_{i-1}) \cdot P(\text{נבחר } x_i) \cdot P(x_i = x \mid \text{נבחר } x_i) = \sum_{i=1}^{\infty} \left(1-\frac{1}{|U|}\right)^{i-1} \cdot \frac{1}{|U|} \cdot \mu(x) = \mu(x)$$