## Dataprocessaringsuligheden

**Def** $\underline{X}, Y, \underline{Z}$ danner en **Markovkæde**, betegnet $\underline{X} \to Y \to \underline{Z}$, hvis

$$p(z|y,x) = p(z|y) \text{ for alle } x,y,z, \text{ så } p(x,y,z) > 0.$$

Der gælder så

$$p(x,z|y) = \frac{p(x,y,z)}{p(y)} = \frac{p(x,y)\,p(z|y)}{p(y)},$$

hvoraf $p(x,z|y) = p(x|y)\,p(z|y)$ (*).

Hvis omvendt (*) er opfyldt for alle $x,y,z$, hvor $p(x,y,z) > 0$ danner $\underline{X}, \underline{Y}, \underline{Z}$ en Markovkæde.

$\underline{X} \to Y \to \underline{Z}$ er altså ensbetydende med, at $\underline{X}$ og $\underline{Z}$ er **uafhængige, betinget** $\underline{Y}$

Følgelig $\underline{X} \to Y \to \underline{Z} \Rightarrow \underline{X} \leftarrow Y \leftarrow \underline{Z}$

## Dataprocesserings uligheden 2

**Hvis $\underline{X} \to Y \to \underline{Z}$, så $I(\underline{X}; Y) \geq I(\underline{X}; \underline{Z})$**

Hvis $\underline{X} \to Y \to \underline{Z}$ gælder også, $I(Y; \underline{Z}) \geq I(\underline{X}; \underline{Z})$.

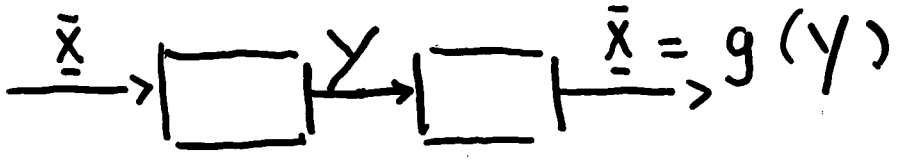
Bevis

$$I(\underline{X}; Y, \underline{Z}) = I(\underline{X}; \underline{Z}) + I(\underline{X}; Y \mid \underline{Z})$$

$$= I(\underline{X}; Y) + \underbrace{I(\underline{X}; \underline{Z} \mid Y)}_{0}$$

□

# FANO'S ULIGHED




$$P_e = Pr\{\hat{\underline{X}} \neq \underline{X}\}$$

$$\boxed{H(P_e) + P_e \log|\mathcal{X}| \geq H(\underline{X}|\hat{\underline{X}}) \geq H(\underline{X}|Y)} \quad (2.130)$$

Da $\hat{\underline{X}} = g(Y)$, gælder $H(\underline{X}|\hat{\underline{X}}) \geq H(\underline{X}|Y)$, så vi skal kun vise den venstre ulighed i (2.130)

$$E = \begin{cases} 1, & \hat{\underline{X}} \neq \underline{X} \\ 0, & \hat{\underline{X}} = \underline{X} \end{cases}$$

$$H(E, \underline{X}|\hat{\underline{X}}) = H(\underline{X}|\hat{\underline{X}}) + \underbrace{H(E|\underline{X}, \hat{\underline{X}})}_{0}$$

$$= \underbrace{H(E|\hat{\underline{X}})}_{\leq H(P_e)} + H(\underline{X}|E, \hat{\underline{X}}).$$

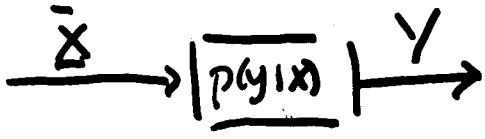
$$H(\underline{X}|E, \hat{\underline{X}}) = \sum_{e\in\{0,1\}, \hat{x}\in\hat{\mathcal{X}}} H(\underline{X}|E=e, \hat{\underline{X}}=\hat{x}) P_r(E=e, \hat{\underline{X}}=\hat{x})$$

$$\overset{(2.136)}{=} P_r(E=0) \sum'_{\hat{x}\in\hat{\mathcal{X}}} H(\underline{X}|E=0, \hat{\underline{X}}=\hat{x}) P_r(\hat{\underline{X}}=\hat{x}|E=0)$$

$$+ P_r(E=1) \sum_{\hat{x}\in\hat{\mathcal{X}}} H(\underline{X}|E=1, \hat{\underline{X}}=\hat{x}) P_r(\hat{\underline{X}}=\hat{x}|E=1)$$

$$\leq (1-P_e)\cdot 0 + P_e \underbrace{\log(|\mathcal{X}|-1)}_{\log|\mathcal{X}|, \text{ hvis } \mathcal{X}=\hat{\mathcal{X}}} \log|\mathcal{X}| \qquad \square$$

# Diskret hukommelsesfri kanal, kapacitet




$\bar{X}$ antager værdier i indgangsalfabetet $\mathcal{H}$.
$Y$ — " — indgangsalfabetet $\mathcal{Y}$
$p(y|x)$, $x \in \mathcal{H}$, $y \in \mathcal{Y}$ betegner overgangssandsynligheden og overgangssandsynligheds matricen $\Pi = [p(y|x)]$ har $|\mathcal{H}|$ rækker og $|\mathcal{Y}|$ søjler.

Til hvert valg af $p(\bar{x})$ defineres $P(\bar{x}, \bar{y})$ ved $p(x,y) = p(y|x)p(x)$, og (informations)kapaciteten def. ved

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

Da gælder

$$I(\underline{X}; Y) = H(Y) - H(Y|\underline{X})$$

- Hvis $[p(y|x)]$ er rækkesymmetrisk (rækkerne er permuterede) er

$$H(Y|\underline{X}) = H(\underline{r}),$$ hvor $\underline{r}$ er en af rækkerne i $[p(y|x)]$, uafhængigt af inputfordelingen.

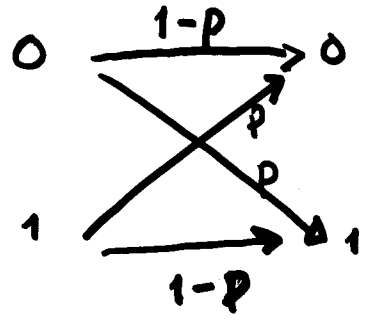
- Hvis søjlesummerne i ~~række~~ $[p(y|x)]$ er konstante gælder, Y uniformt fordelt på $\mathcal{Y}$, når $\underline{X}$ er uniform fordelt på $\mathcal{X}$

- Hvis $[p(y|x)]$ er rækkesymmetrisk med konstante søjlesummer kaldes kanalen svagt symmetrisk og af ovenstående følger, at

$$C = \log|\mathcal{Y}| - H(\underline{r}).$$

7.1.2

Ex ~~8.1.4~~ Binære symmetriske kanal




Kanalen er symmetrisk dvs.

række- og søjle symmetrisk.

Altså $C = 1 - H(p)$ (bits).

---

Ex lign. (~~8~~.18)

$c \in \mathbb{N}$ $\mathcal{X} = \{0,1,2,\ldots c-1\}$, $\mathcal{Y} = \{0,1,\ldots,c-1\}$

Der er givet en sandsynlighedsfordeling $p(z)$, $z \in \{0,1,\ldots,c-1\}$ for "støjen" ~~vektoren~~ $\underline{Z}$. $\underline{Z}$ er uafhængig af input $\underline{X}$ og

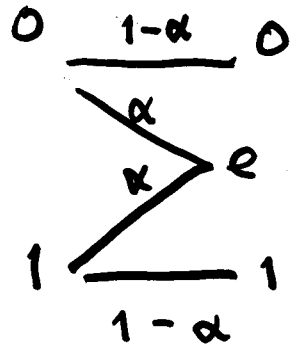
$$\underline{Y} = \underline{X} + \underline{Z} \pmod{c}.$$

Der gælder $p(y|x) = p(y - x)$ Her regnes mod c.

Kanalen er symmetrisk $C = \log c - H(\underline{Z})$

Ex 7.1.5

# Erasure - Kanalen




$P(\underline{X}=\underline{1}) = \pi$

$$E = \begin{cases} 1, & Y = e \\ 0 & Y \neq e \end{cases}$$

Kanalen er rekke symmetrisk, så

$$I(\underline{X};Y) = H(Y) - H(Y|\underline{X}) = H(Y) - H(\alpha).$$

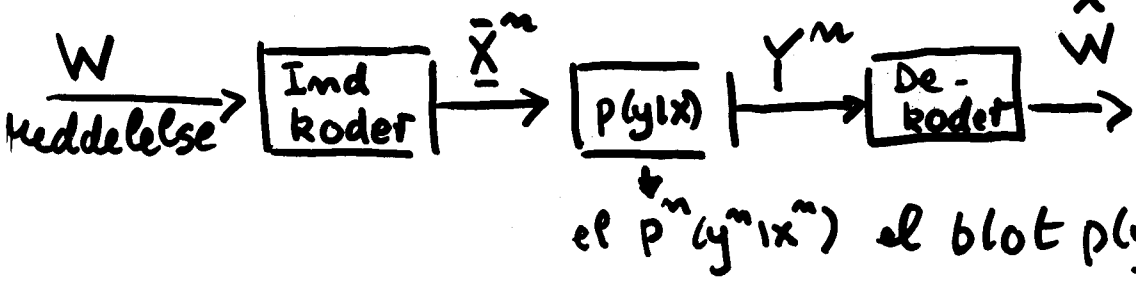
$$H(Y) = H(Y,E) = H(E) + H(Y|E)$$

$$\hat{H}(\alpha) + (1-\alpha)\,H(\pi)$$

$H(Y)$ maksimeres for $\underline{\pi = \frac{1}{2}}$

og $\underline{C = 1-\alpha}$

Fig. 7.8




Den n'te udvidelse af den diskrete hukommelsesfrie kanal med overgangssandsynligheder $p(y|x)$ har overgangssandsynligheder

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

og der gælder altså

$$P_r(\underline{Y}^n = y^n | \underline{X}^n = x^n) = \prod_{i=1}^{n} p(y_i|x_i) \quad //$$

og

$$P_r(\underline{X}^n = x^n, \underline{Y}^n = y^n) = P_r(\underline{X}^n = x^n) \prod_{i=1}^{n} p(y_i|x_i) .$$

**Definition** En $(M, n)$ kode for kanalen $(\mathcal{X}, p(y|x), \mathcal{Y})$ består af følgende:

1. En indexmængde $\{1, 2, \ldots, M\}$
2. En indkodningsfunktion
$$\underline{X}^n : \{1, 2, \ldots, M\} \to \mathcal{X}^n,$$
som danner kodeordene $\underline{X}^n(1), \ldots, \underline{X}^n(M)$.
Mængden af kodeord kaldes kodebogen.
3. En dekodningsfunktion
$$g : \mathcal{Y}^n \to \{1, 2, \ldots, M\},$$
som er en deterministisk regel, der for hver mulig modtagen vektor gætter på det afsendte index.

**Definition** $\lambda_i$ er den betingede sandsynlighed for fejl, givet index $i$ er afsendt, dvs.

$$\lambda_i = \Pr\left(g(Y^n) \neq i \mid \underline{X}^n = \underline{X}^n(i)\right)$$

$$= \sum_{y^n}' P(y^n \mid X^n(i))\, I(g(y^n) \neq i) .$$

**Definition** Maksimal fejlsandsynligheden $\lambda^{(n)}$ for en $(M,n)$ kode er defineret ved

$$\lambda^{(n)} = \max_{i \in \{1,\ldots,M\}} \lambda_i$$

**Definition** Den (aritmetisk) gennemsnitlige fejlsandsynlighed er defineret ved

$$P_e^{(n)} = \frac{1}{M} \sum_{i=1}^{M}{}' \lambda_i .$$

Der gælder altid $\Pr(I \neq g(Y^n)) \leq \lambda^{(n)}$.

Hvis index $I$ er valgt ligefordelt fra $\{1,2,\ldots,M\}$ gælder $\Pr(I \neq g(Y^n)) = P_e^{(n)}$

**Definition** Hastigheden $R$ af en $(M, n)$-kode er

$$R = \frac{\log M}{n}$$

**Definition** En hastig $R$ er opnåelig, hvis der eksisterer en følge $([2^{nR}], n)$ af koder, så $\lambda^{(n)} \to 0$ for $n \to \infty$.

**Definition** Kapaciteten (den operationelle) af en diskret hukommelsesfri kanal er supremum over alle opnåelige hastigheder.

side 195

I Shannons kanalkodningssætning betegner $C$ informations kapaciteten for en foreliggende diskret ukommunelsesfri kanal.

Theorem 7.7.1 (Kanalkodningssætningen)

Alle hastigheder under kanal kapaciteten $C$ er opnåelige.

Dvs. for enhver hastighed $R < C$ eksisterer en følge af $(\lceil 2^{nR} \rceil, n)$ koder med maksimumfejlsandsynlighed $\lambda^{(n)}$, så $\lambda^{(n)} \to 0$ for $n \to \infty$.

"Omvendt", for enhver følge af $(\lceil 2^{nR} \rceil, n)$ koder med $\lambda^{(n)} \to 0$ for $n \to \infty$ gælder, $R \leq C$.