Introduction to channel coding

Dr. Yoann Altmann

B39AX – Fall 2023
Heriot-Watt University

LEADERS IN IDEAS AND SOLUTIONS

Plan

- Joint/conditional entropy
- Mutual information
- Examples of channels
- Channel capacity

- Based on the book: **“Information Theory**, **Inference**, and **Learning Algorithms”**. David J.C. MacKay (Chap. 8-9)

Joint entropy

- Let X and Y be two (discrete) RVs defined on $\mathcal{A}_X$ and $\mathcal{A}_Y$

$$H(X,Y) = -\sum_{x\in\mathcal{A}_X, y\in\mathcal{A}_Y} p(x,y)\log_2\big(p(x,y)\big)$$

Important property

- Let X and Y be two RVs. We have

$$H(x,y) = H(x) + H(y)$$

if and only if X and Y are independent, i.e.,

$$p(x,y) = p(x)p(y)$$

Conditional entropy

- Let X and Y be two (discrete) RVs defined on $\mathcal{A}_X$ and $\mathcal{A}_Y$, for a fixed value $y = y_k$

$$H(X|y = y_k) = -\sum_{x \in \mathcal{A}_X} p(x|y = y_k) \log_2\big(p(x|y = y_k)\big)$$

- "Conditional entropy of X given $y = y_k$"

Conditional entropy

- Let X and Y be two (discrete) RVs defined on $\mathcal{A}_X$ and $\mathcal{A}_Y$,

$$H(X|Y) = -\sum_{y \in \mathcal{A}_Y} p(y) \sum_{x \in \mathcal{A}_X} p(x|y) \log_2\big(p(x|y)\big)$$

- "Conditional entropy of X given Y"
- Using the Bayes rule:

$$H(X|Y) = -\sum_{x \in \mathcal{A}_X, y \in \mathcal{A}_Y} p(x,y) \log_2\big(p(x|y)\big)$$

Joint and conditional entropy

- Joint entropy

$$H(X,Y) = -\sum_{x \in \mathcal{A}_X, y \in \mathcal{A}_Y} p(x,y) \log_2\big(p(x,y)\big)$$

- Conditional entropy

$$H(X|Y) = -\sum_{x \in \mathcal{A}_X, y \in \mathcal{A}_Y} p(x,y) \log_2\big(p(x|y)\big)$$

Marginal entropy

- Let X and Y be two (discrete) RVs defined on $\mathcal{A}_X$ and $\mathcal{A}_Y$,
- The entropy $H(X)$ is also called marginal entropy of X

Chain rules

- Information content

$$I(x) = -\log_2 p(x)$$

Using $p(x,y) = p(x|y)p(y)$

$$I(x,y) = I(x|y) + I(y) = I(y|x) + I(x)$$

"The information content of (x,y) is the information content of x plus the information content of $y|x$"

Chain rules

- Entropy

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

“The uncertainty of (X,Y) is the uncertainty of X plus the uncertainty of $Y|X$”

Mutual information

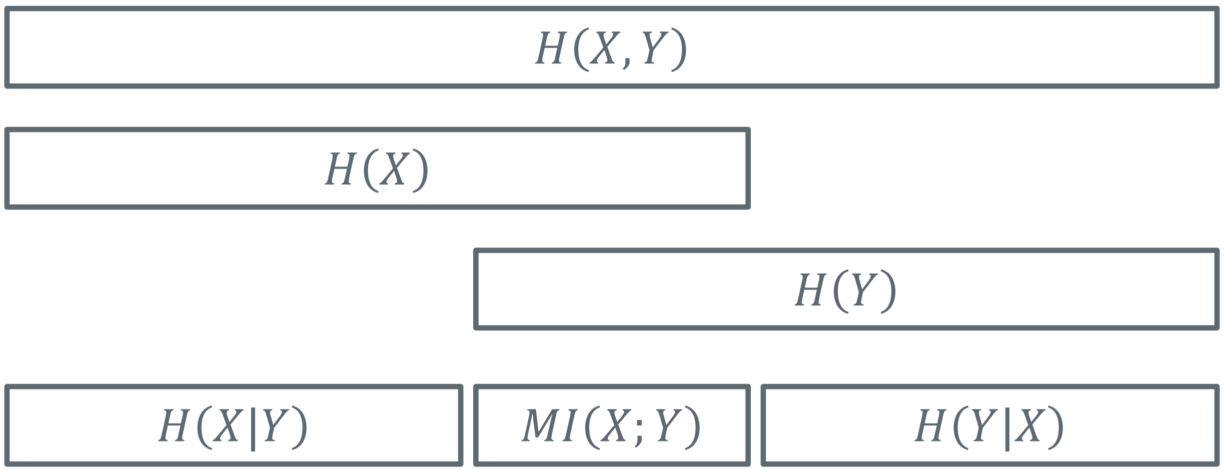
$$H(X,Y) = H(X|Y) + H(Y) = H(Y|X) + H(X)$$

- Definition: mutual information

$$MI(X;Y) = H(X) - H(X|Y)$$

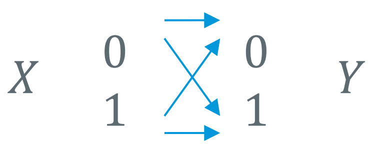
- Properties:
  - $MI(X;Y) = MI(Y;X)$
  - $MI(X;Y) \geq 0$
  - $MI(X;Y) = H(X) + H(Y) - H(X,Y)$
- Measures the average reduction of uncertainty about x that results from knowing y OR the average amount of information that x conveys about y

Entropy and mutual information



Example

- $\mathcal{A}_X = \{0,1\}$ (binary source)
- $\mathbb{p}(X = 0) = p$, $\mathbb{p}(X = 1) = 1 - p$
- Binary symmetric channel



$$\mathbb{p}(Y = 0|X = 0) = 1 - f, \mathbb{p}(Y = 1|X = 1) = 1 - f$$
$$\mathbb{p}(Y = 1|X = 0) = f, \mathbb{p}(Y = 0|X = 1) = f$$

Example

- $\mathbb{p}(Y = 0) = p + f - 2fp = \tilde{p}$
- $\mathbb{p}(Y = 1) = 1 - p - f + 2fp = 1 - \tilde{p}$
- $\mathbb{p}((Y, X) = (0,0)) = (1 - f)p$
- $\mathbb{p}((Y, X) = (1,1)) = (1 - f)(1 - p)$
- $\mathbb{p}((Y, X) = (1,0)\ ) = fp$
- $\mathbb{p}((Y, X) = (0,1)\ ) = f(1 - p)$

- $H(X) = -p \log_2 p - (1 - p) \log_2 (1 - p)$
- $H(Y) = -\tilde{p} \log_2 \tilde{p} - (1 - \tilde{p}) \log_2 (1 - \tilde{p})$

… see demo1.m