香港中文大學
The Chinese University of Hong Kong

2.7 Some Useful Information Inequalities

Theorem 2.38 (Conditioning Does Not Increase Entropy)

$$H(Y|X) \leq H(Y)$$

with equality if and only if X and Y are independent.

Theorem 2.38 (Conditioning Does Not Increase Entropy)

$$H(Y|X) \leq H(Y)$$

with equality if and only if X and Y are independent.

Proof

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

with equality if and only if $I(X;Y) = 0$, or X and Y are independent.

Theorem 2.38 (Conditioning Does Not Increase Entropy)

$$H(Y|X) \le H(Y)$$

with equality if and only if X and Y are independent.

Proof

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

with equality if and only if $I(X;Y) = 0$, or X and Y are independent.

Theorem 2.38 (Conditioning Does Not Increase Entropy)

$$H(Y|X) \le H(Y)$$

with equality if and only if X and Y are independent.

Proof

$$H(Y|X) = H(Y) - \underbrace{I(X;Y)}_{\ge 0} \le H(Y),$$

with equality if and only if $I(X;Y) = 0$, or X and Y are independent.

Theorem 2.38 (Conditioning Does Not Increase Entropy)

$$H(Y|X) \leq H(Y)$$

with equality if and only if X and Y are independent.

Proof

$$H(Y|X) = H(Y) - \underbrace{I(X;Y)}_{\geq 0} \leq H(Y),$$

with equality if and only if $I(X;Y) = 0$, or X and Y are independent.

Theorem 2.38 (Conditioning Does Not Increase Entropy)

$$H(Y|X) \le H(Y)$$

with equality if and only if X and Y are independent.

Proof

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

with equality if and only if $I(X;Y) = 0$, or X and Y are independent.

Remarks

- Similarly, $H(Y|X,Z) \le H(Y|Z)$.

Theorem 2.38 (Conditioning Does Not Increase Entropy)

$$H(Y|X) \leq H(Y)$$

with equality if and only if X and Y are independent.

Proof

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

with equality if and only if $I(X;Y) = 0$, or X and Y are independent.

Remarks

- Similarly, $H(Y|X,Z) \leq H(Y|Z)$.
- Warning: $I(X;Y|Z) \leq I(X;Y)$ does not hold in general.

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \le \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$H(X_1, X_2, \cdots, X_n)$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = \quad & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \le \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = \quad & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \le \quad & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = \quad & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq \quad & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \le \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = \quad & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \le \quad & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \le \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \le & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \le i \le n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \le \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \le & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \le i \le n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$p(x_1, x_2, \cdots, x_n)$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \le \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \le & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \le i \le n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$p(x_1, x_2, \cdots, x_n)$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$\begin{aligned} & p(x_1, x_2, \cdots, x_n) \\ = & p(x_1, x_2, \cdots, x_{n-1}) p(x_n) \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$\begin{aligned} & p(x_1, x_2, \cdots, x_n) \\ = & \ p(x_1, x_2, \cdots, x_{n-1}) p(x_n) \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$\begin{aligned} & p(x_1, x_2, \cdots, x_n) \\ = & \; p(x_1, x_2, \cdots, \underline{x_{n-1})} p(x_n) \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$\begin{aligned} & p(x_1, x_2, \cdots, x_n) \\ = & p(x_1, x_2, \cdots, \underline{x_{n-1})} p(x_n) \\ = & p(x_1, x_2, \cdots, x_{n-2}) \underline{p(x_{n-1})} p(x_n) \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ &= \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ &\leq \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$\begin{aligned} & p(x_1, x_2, \cdots, x_n) \\ &= p(x_1, x_2, \cdots, x_{n-1}) p(x_n) \\ &= p(x_1, x_2, \cdots, x_{n-2}) p(x_{n-1}) p(x_n) \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ & = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}) \\ & \leq \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i|X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$\begin{aligned} & p(x_1, x_2, \cdots, x_n) \\ & = p(x_1, x_2, \cdots, x_{n-1})p(x_n) \\ & = p(x_1, x_2, \cdots, x_{n-2})p(x_{n-1})p(x_n) \\ & \quad \vdots \\ & = p(x_1)p(x_2)\cdots p(x_n) \end{aligned}$$

Theorem 2.39 (Independence Bound for Entropy)

$$H(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

with equality if and only if X_i, $i = 1, 2, \cdots, n$ are mutually independent.

Proof

1. By the chain rule for entropy,

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n) \\ = & \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}) \\ \leq & \sum_{i=1}^{n} H(X_i). \end{aligned}$$

The inequality follows because conditioning does not increase entropy.

2. The inequality is tight if and only if it is tight for each i, i.e.,

$$H(X_i | X_1, \cdots, X_{i-1}) = H(X_i)$$

for $1 \leq i \leq n$. This is equivalent to X_i being independent of $X_1, X_2, \cdots, X_{i-1}$ for each i.

3. Then

$$\begin{aligned} & p(x_1, x_2, \cdots, x_n) \\ = & \; p(x_1, x_2, \cdots, x_{n-1}) p(x_n) \\ = & \; p(x_1, x_2, \cdots, x_{n-2}) p(x_{n-1}) p(x_n) \\ & \vdots \\ = & \; p(x_1) p(x_2) \cdots p(x_n) \end{aligned}$$

for all $x_1, x_2, \cdots, x_n$, i.e., $X_1, X_2, \cdots, X_n$ are mutually independent.

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \to Y \to Z$ forms a Markov chain.

Proof

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \to Y \to Z$ forms a Markov chain.

Proof

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \rightarrow Y \rightarrow Z$ forms a Markov chain.

Proof

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \to Y \to Z$ forms a Markov chain.

Proof

By the chain rule for mutual information, we have

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \to Y \to Z$ forms a Markov chain.

Proof

By the chain rule for mutual information, we have

$$I(X;Y,Z) = I(X;Y) + I(X;Z|Y) \geq I(X;Y).$$

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \to Y \to Z$ forms a Markov chain.

Proof

By the chain rule for mutual information, we have

$$I(X;Y,Z) = I(X;Y) + \underline{I(X;Z|Y)} \geq I(X;Y).$$

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \to Y \to Z$ forms a Markov chain.

Proof

By the chain rule for mutual information, we have

$$I(X;Y,Z) = I(X;Y) + I(X;Z|Y) \geq I(X;Y).$$

The above inequality is tight if and only if $I(X;Z|Y) = 0$, or $X \to Y \to Z$ forms a Markov chain.

Theorem 2.40

$$I(X;Y,Z) \geq I(X;Y),$$

with equality if and only if $X \rightarrow Y \rightarrow Z$ forms a Markov chain.

Proof

By the chain rule for mutual information, we have

$$I(X;Y,Z) = I(X;Y) + \underline{I(X;Z|Y)} \geq I(X;Y).$$

The above inequality is tight if and only if $\underline{I(X;Z|Y) = 0}$, or $X \rightarrow Y \rightarrow Z$ forms a Markov chain.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z)$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

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

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \overset{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z)$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \overset{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z)$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \overset{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z)$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \overset{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

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

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

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

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \quad (1)$$

and

$$I(X;Z) \leq I(Y;Z). \quad (2)$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z)$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \overset{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

$$I(X;Y,Z) = I(X;Z) + I(X;Y|Z)$$
$$\Rightarrow \quad I(X;Z) = I(X;Y,Z) - I(X;Y|Z)$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$I(X;Z) \stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z)$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(\underline{X};Y,Z) \\ &\overset{b)}{=} I(\underline{X};Y) + I(X;Z|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(\underline{X};\underline{Y},Z) \\ &\overset{b)}{=} I(\underline{X};\underline{Y}) + I(X;Z|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X; Z) \leq I(X; Y) \tag{1}$$

and

$$I(X; Z) \leq I(Y; Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X; Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X; Z) &\overset{a)}{=} I(X; Y, Z) - I(X; Y|Z) \\ &\leq I(X; Y, Z) \\ &\overset{b)}{=} I(X; Y) + I(X; Z|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X; Y, Z) = I(X; Z) + I(X; Y|Z) \\ \Rightarrow \quad & I(X; Z) = I(X; Y, Z) - I(X; Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(\underline{X};Y,\underline{Z}) \\ &\overset{b)}{=} I(X;Y) + I(\underline{X};\underline{Z}|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(\underline{X};\underline{Y},\underline{Z}) \\ &\overset{b)}{=} I(X;Y) + I(\underline{X};\underline{Z}|\underline{Y}) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

   a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

   b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow\quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

 a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

 b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\stackrel{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\stackrel{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \ge H(X|Y)$.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow\quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $\underline{X} \to Y \to Z$, then $H(\underline{X}|Z) \ge H(X|Y)$.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \quad (1)$$

and

$$I(X;Z) \leq I(Y;Z). \quad (2)$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow\quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $\underline{X} \to Y \to \underline{Z}$, then $H(\underline{X}|\underline{Z}) \geq H(X|Y)$.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \ge H(X|Y)$.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $\underline{X} \to \underline{Y} \to Z$, then $H(X|Z) \geq H(\underline{X}|\underline{Y})$.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\stackrel{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\stackrel{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \geq H(X|Y)$.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \ge H(X|Y)$.

Proof

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \geq H(X|Y)$.

Proof

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

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow\quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \ge H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X;Z) \\ &\ge H(X) - I(X;Y) \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

   a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

   b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \geq H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X;Z) \\ &\geq H(X) - I(X;Y) \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \geq H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X;Z) \\ &\geq H(X) - I(X;Y) \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

   a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

   b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \geq H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - \underline{I(X;Z)} \\ &\geq H(X) - \underline{I(X;Y)} \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y) \tag{1}$$

and

$$I(X;Z) \leq I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\leq I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \geq H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X;Z) \\ &\geq H(X) - I(X;Y) \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z | Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \ge H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X;Z) \\ &\ge \underline{H(X) - I(X;Y)} \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \ge H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X;Z) \\ &\ge \underline{H(X) - I(X;Y)} \\ &= \underline{H(X|Y)}. \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X; Z) \leq I(X; Y) \tag{1}$$

and

$$I(X; Z) \leq I(Y; Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X; Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X; Z) &\overset{a)}{=} I(X; Y, Z) - I(X; Y|Z) \\ &\leq I(X; Y, Z) \\ &\overset{b)}{=} I(X; Y) + I(X; Z|Y) \\ &= I(X; Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X; Y, Z) = I(X; Z) + I(X; Y|Z) \\ \Rightarrow \quad & I(X; Z) = I(X; Y, Z) - I(X; Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \geq H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X; Z) \\ &\geq H(X) - I(X; Y) \\ &= H(X|Y). \end{aligned}$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y) \tag{1}$$

and

$$I(X;Z) \le I(Y;Z). \tag{2}$$

Proof

1. Assume $X \to Y \to Z$, i.e., $X \perp Z|Y$. Then

$$I(X;Z|Y) = 0.$$

2. To prove (1), consider

$$\begin{aligned} I(X;Z) &\overset{a)}{=} I(X;Y,Z) - I(X;Y|Z) \\ &\le I(X;Y,Z) \\ &\overset{b)}{=} I(X;Y) + I(X;Z|Y) \\ &= I(X;Y). \end{aligned}$$

a) Chain rule for mutual information:

$$\begin{aligned} & I(X;Y,Z) = I(X;Z) + I(X;Y|Z) \\ \Rightarrow \quad & I(X;Z) = I(X;Y,Z) - I(X;Y|Z) \end{aligned}$$

b) Chain rule for mutual information

3. Since $X \to Y \to Z$ is equivalent to $Z \to Y \to X$, we also have proved (2) by symmetry.

Corollary If $X \to Y \to Z$, then $H(X|Z) \ge H(X|Y)$.

Proof

$$\begin{aligned} H(X|Z) &= H(X) - I(X;Z) \\ &\ge H(X) - I(X;Y) \\ &= H(X|Y). \end{aligned}$$

Remark

Suppose Y is an observation of X. Then further processing of Y can only increase the uncertainty about X on the average.

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Theorem 2.42 (Data Processing Theorem) If $\underline{U} \to \underline{X} \to \underline{Y} \to \underline{V}$ forms a Markov chain, then

$$I(\underline{U};\underline{V}) \leq I(\underline{X};\underline{Y}).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \leq I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$\underline{U} \to X \to \underline{Y} \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(\underline{U};\underline{Y}) \le I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U; V) \leq I(X; Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to \underline{X} \to \underline{Y} \qquad (1)$$

and

$$U \to Y \to V. \qquad (2)$$

2. From (1) and Lemma 2.41, we have

$$I(U; Y) \leq I(\underline{X}; \underline{Y}).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X; Z) \leq I(X; Y)$$

and

$$I(X; Z) \leq I(Y; Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \leq I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \leq I(X;Y).$$

3. Similarly, from (2), we have

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \leq I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \leq I(U;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \le I(U;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \le I(U;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \le I(U;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \le I(U;Y).$$

4. Combining these two inequalities, we have

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \leq I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \leq I(U;Y).$$

4. Combining these two inequalities, we have

$$I(U;V) \leq I(U;Y) \leq I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$\underline{I(U;V) \le I(U;Y)}.$$

4. Combining these two inequalities, we have

$$\underline{I(U;V) \le I(U;Y)} \le I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$\underline{I(U;Y) \leq I(X;Y)}.$$

3. Similarly, from (2), we have

$$I(U;V) \leq I(U;Y).$$

4. Combining these two inequalities, we have

$$I(U;V) \leq \underline{I(U;Y) \leq I(X;Y)}.$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \le I(U;Y).$$

4. Combining these two inequalities, we have

$$I(U;V) \le I(U;Y) \le I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \leq I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \leq I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \leq I(U;Y).$$

4. Combining these two inequalities, we have

$$I(U;V) \leq I(U;Y) \leq I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \leq I(X;Y)$$

and

$$I(X;Z) \leq I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \le I(U;Y).$$

4. Combining these two inequalities, we have

$$I(U;V) \le I(U;Y) \le I(X;Y).$$

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$

Theorem 2.42 (Data Processing Theorem) If $U \to X \to Y \to V$ forms a Markov chain, then

$$I(U;V) \le I(X;Y).$$

Proof

1. Assume $U \to X \to Y \to V$. Then

$$U \to X \to Y \tag{1}$$

and

$$U \to Y \to V. \tag{2}$$

2. From (1) and Lemma 2.41, we have

$$I(U;Y) \le I(X;Y).$$

3. Similarly, from (2), we have

$$I(U;V) \le I(U;Y).$$

4. Combining these two inequalities, we have

$$I(U;V) \le I(U;Y) \le I(X;Y).$$

5. The theorem is proved.

Lemma 2.41 If $X \to Y \to Z$ forms a Markov chain, then

$$I(X;Z) \le I(X;Y)$$

and

$$I(X;Z) \le I(Y;Z).$$