2.4 Chain Rules

Proposition 2.24 (Chain Rule for Entropy)

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

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$n = 2$ $\quad$ $H(1,2) = H(1) + H(2|1)$ (Proposition 2.6)

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$$n = 2 \qquad H(\underline{1,} 2) = H(\underline{1}) + H(2|1) \qquad \text{(Proposition 2.6)}$$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$$n = 2 \qquad H(1, \underline{2}) = H(1) + H(\underline{2}|1) \qquad \text{(Proposition 2.6)}$$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$$n = 2 \qquad H(\underline{1}, \underline{2}) = H(1) + H(\underline{2} | \underline{1}) \qquad \text{(Proposition 2.6)}$$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$n = 2$ $\qquad H(1,2) = H(1) + H(2|1)$ (Proposition 2.6)

$n = 3$ $\qquad H(1,2,3) = H(1) + H(2|1) + H(3|1,2)$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$n = 2$ $\quad H(1, 2) = H(1) + H(2|1)$ (Proposition 2.6)

$n = 3$ $\quad H(\underline{1,} 2, 3) = H(\underline{1}) + H(2|1) + H(3|1, 2)$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$$n = 2 \qquad H(1,2) = H(1) + H(2|1) \quad \text{(Proposition 2.6)}$$

$$n = 3 \qquad H(1,2,3) = H(1) + H(2|1) + H(3|1,2)$$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$n = 2$ $\quad H(1, 2) = H(1) + H(2|1)$ $\quad$ (Proposition 2.6)

$n = 3$ $\quad H(\underline{1}, \underline{2}, 3) = H(1) + H(\underline{2}|\underline{1}) + H(3|1, 2)$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$n = 2$ $\quad$ $H(1,2) = H(1) + H(2|1)$ $\quad$ (Proposition 2.6)

$n = 3$ $\quad$ $H(1,2,3) = H(1) + H(2|1) + H(3|1,2)$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$n = 2$ $\quad H(1, 2) = H(1) + H(2|1)$ (Proposition 2.6)

$n = 3$ $\quad H(\underline{1, 2}, \underline{3}) = H(1) + H(2|1) + H(\underline{3}|\underline{1, 2})$

Proposition 2.24 (Chain Rule for Entropy)

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

Example

$n = 2$ $\qquad H(1,2) = H(1) + H(2|1)$ (Proposition 2.6)

$n = 3$ $\qquad H(1,2,3) = H(1) + H(2|1) + H(3|1,2)$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n | Y) = \sum_{i=1}^{n} H(X_i | X_1, \cdots, X_{i-1}, Y).$$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$n = 2$ $\qquad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$n = 2 \qquad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$$n = 2 \qquad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$$n = 2 \qquad I(\underline{1}, \underline{2}; Y) = I(1; Y) + I(\underline{2}; Y | \underline{1})$$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$$n = 2 \qquad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$$

$$n = 3 \qquad I(1, 2, 3; Y) = I(1; Y) + I(2; Y|1) + I(3; Y|1, 2)$$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$n = 2$ $\qquad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$

$n = 3$ $\qquad I(1, 2, 3; Y) = I(1; Y) + I(2; Y|1) + I(3; Y|1, 2)$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$n = 2$ $\quad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$

$n = 3$ $\quad I(1, 2, 3; Y) = I(1; Y) + I(2; Y|1) + I(3; Y|1, 2)$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$$n = 2 \qquad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$$

$$n = 3 \qquad I(\underline{1}, \underline{2}, 3; Y) = I(1; Y) + I(\underline{2}; Y|\underline{1}) + I(3; Y|1, 2)$$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$n = 2$ $\quad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$

$n = 3$ $\quad I(1, 2, 3; Y) = I(1; Y) + I(2; Y|1) + I(3; Y|1, 2)$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$n = 2$ $\quad I(1, 2; Y) = I(1; Y) + I(2; Y|1)$

$n = 3$ $\quad I(\underline{1, 2}, \underline{3}; Y) = I(1; Y) + I(2; Y|1) + I(\underline{3}; Y|\underline{1, 2})$

Proposition 2.26 (Chain Rule for Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}).$$

Example

$$n = 2 \qquad I(1,2;Y) = I(1;Y) + I(2;Y|1)$$

$$n = 3 \qquad I(1,2,3;Y) = I(1;Y) + I(2;Y|1) + I(3;Y|1,2)$$

Proposition 2.27 (Chain Rule for Conditional Mutual Information)

$$I(X_1, X_2, \cdots, X_n; Y|Z) = \sum_{i=1}^{n} I(X_i; Y | X_1, \cdots, X_{i-1}, Z).$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

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

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n|Y) \\ = \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n|Y) \\ = \quad & H(\underline{X_1}, X_2, \cdots, X_n, Y) - H(Y) \end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n|Y) \\ = \quad & H(\underline{X_1}, X_2, \cdots, X_n, \underline{Y}) - H(Y) \end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n|Y) \\ = \quad & H(\underline{X_1}, X_2, \cdots, X_n, \underline{Y}) - H(Y) \\ = \quad & H(\underline{(X_1, Y)}, X_2, \cdots, X_n) - H(Y) \end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
= \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
= \quad & H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
= \quad & \underline{H(X_1, Y)} + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
= \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
= \quad & H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
= \quad & \underline{H(X_1, Y)} + \sum_{\underline{i=2}}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n|Y) \\ = \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\ = \quad & H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\ = \quad & \underline{H(X_1, Y)} + \sum_{\underline{i=2}}^{n} H(\underline{X_i}|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
= \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
= \quad & H((\underline{X_1, Y}), X_2, \cdots, X_n) - H(Y) \\
= \quad & \underline{H(X_1, Y)} + \sum_{\underline{i=2}}^{n} H(\underline{X_i}|(\underline{X_1, Y}), X_2, \cdots, X_{i-1}) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
= \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
= \quad & H(\underline{(X_1, Y)}, \underline{X_2,} \cdots, X_n) - H(Y) \\
= \quad & \underline{H(X_1, Y)} + \sum_{\underline{i=2}}^{n} H(\underline{X_i}|\underline{(X_1, Y)}, \underline{X_2, \cdots, X_{i-1}}) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
&H(X_1, X_2, \cdots, X_n|Y) \\
&\quad = \quad H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
&\quad = \quad H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
&\quad = \quad H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
= \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
= \quad & H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
= \quad & H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, \underline{Y}), X_2, \cdots, X_{i-1}) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
&H(X_1, X_2, \cdots, X_n|Y) \\
&= H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
&= H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, \underline{Y}), X_2, \cdots, X_{i-1}) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, \underline{Y}) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
&H(X_1, X_2, \cdots, X_n|Y) \\
&= H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
&= H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
= \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
= \quad & H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
= \quad & H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \\
= \quad & \underline{H(X_1, Y)} + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) - H(Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
&H(X_1, X_2, \cdots, X_n|Y) \\
&= H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
&= H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \\
&= \underline{H(X_1, Y)} + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) \underline{- H(Y)}
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
&H(X_1, X_2, \cdots, X_n|Y) \\
&\quad = \quad H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
&\quad = \quad H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
&\quad = \quad H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \\
&\quad = \quad \underline{H(X_1, Y)} + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) \underline{- H(Y)} \\
&\quad = \quad \underline{H(X_1|Y)} + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
&H(X_1, X_2, \cdots, X_n|Y) \\
&= H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
&= H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) - H(Y) \\
&= \underline{H(X_1|Y)} + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
= \quad & H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
= \quad & H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
= \quad & H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \\
= \quad & H(X_1, Y) + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) - H(Y) \\
= \quad & \underline{H(X_1|Y)} + \sum_{\underline{i=2}}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y)
\end{aligned}$$

Proposition 2.25 (Chain Rule for Conditional Entropy)

$$H(X_1, X_2, \cdots, X_n|Y) = \sum_{i=1}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y).$$

Proof

$$\begin{aligned}
&H(X_1, X_2, \cdots, X_n|Y) \\
&= H(X_1, X_2, \cdots, X_n, Y) - H(Y) \\
&= H((X_1, Y), X_2, \cdots, X_n) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|(X_1, Y), X_2, \cdots, X_{i-1}) - H(Y) \\
&= H(X_1, Y) + \sum_{i=2}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) - H(Y) \\
&= \underline{H(X_1|Y)} + \sum_{\underline{i=2}}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y) \\
&= \sum_{\underline{i=1}}^{n} H(X_i|X_1, \cdots, X_{i-1}, Y)
\end{aligned}$$

Alternative Proof of Proposition 2.25

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

$$
\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
& \quad = \sum_y p(y) H(X_1, X_2, \cdots, X_n|Y = y)
\end{aligned}
$$

Alternative Proof of Proposition 2.25

$$
\begin{aligned}
& H(X_1, X_2, \cdots, X_n|Y) \\
& \quad = \sum_y \underline{p(y)} H(X_1, X_2, \cdots, X_n|\underline{Y = y})
\end{aligned}
$$

Alternative Proof of Proposition 2.25

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n|Y) \\ = & \sum_y p(y) H(X_1, X_2, \cdots, X_n|Y = y) \end{aligned}$$

Alternative Proof of Proposition 2.25

$$\begin{aligned} & H(X_1, X_2, \cdots, X_n|Y) \\ = & \sum_y p(y) \underline{H(X_1, X_2, \cdots, X_n|Y = y)} \end{aligned}$$

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Alternative Proof of Proposition 2.25

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

Remark This alternative proof explains why Proposition 2.25 can be obtained from Proposition 2.24 by conditioning on Y.