

**Definition 10.28** Let

The conditional entropy of  $X$  given  $Y$  is defined as

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

**Proposition 10.29** For two random variables  $X$  and  $Y$ ,

1.  $h(Y) = h(Y|X) + I(X; Y)$  if  $Y$  is **continuous**;
2.  $H(Y) = H(Y|X) + I(X; Y)$  if  $Y$  is **discrete**.

**Proposition 10.30 (Chain Rule)**

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