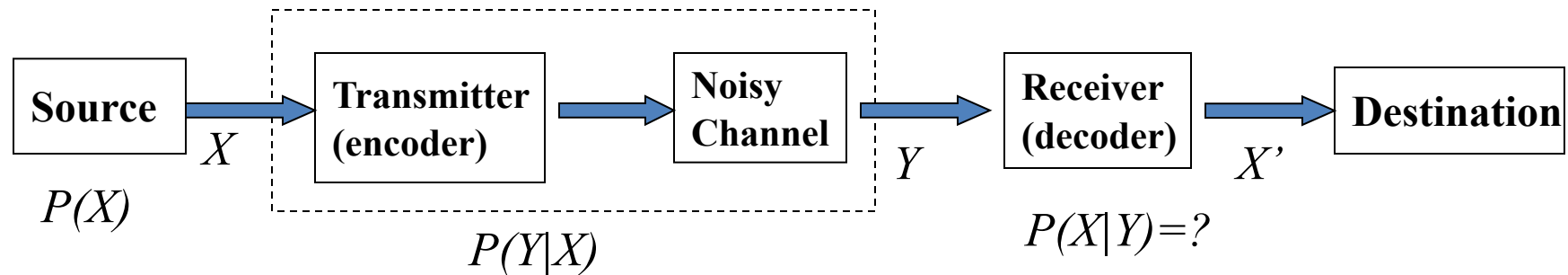


Application 2: As Likelihood in Bayesian Inference



$$\hat{X} = \arg \max_X p(X | Y) = \arg \max_X p(Y | X) p(X) \quad \text{(Bayes Rule)}$$

When Y is text, $p(Y|X)$ is a (conditional) language model

Many Examples:

Text categorization:

X=Topic Category

Y=Text document

Machine translation:

X=English sentence

Y=Chinese sentence

Sentiment tagging:

X=Sentiment label

Y= Text object