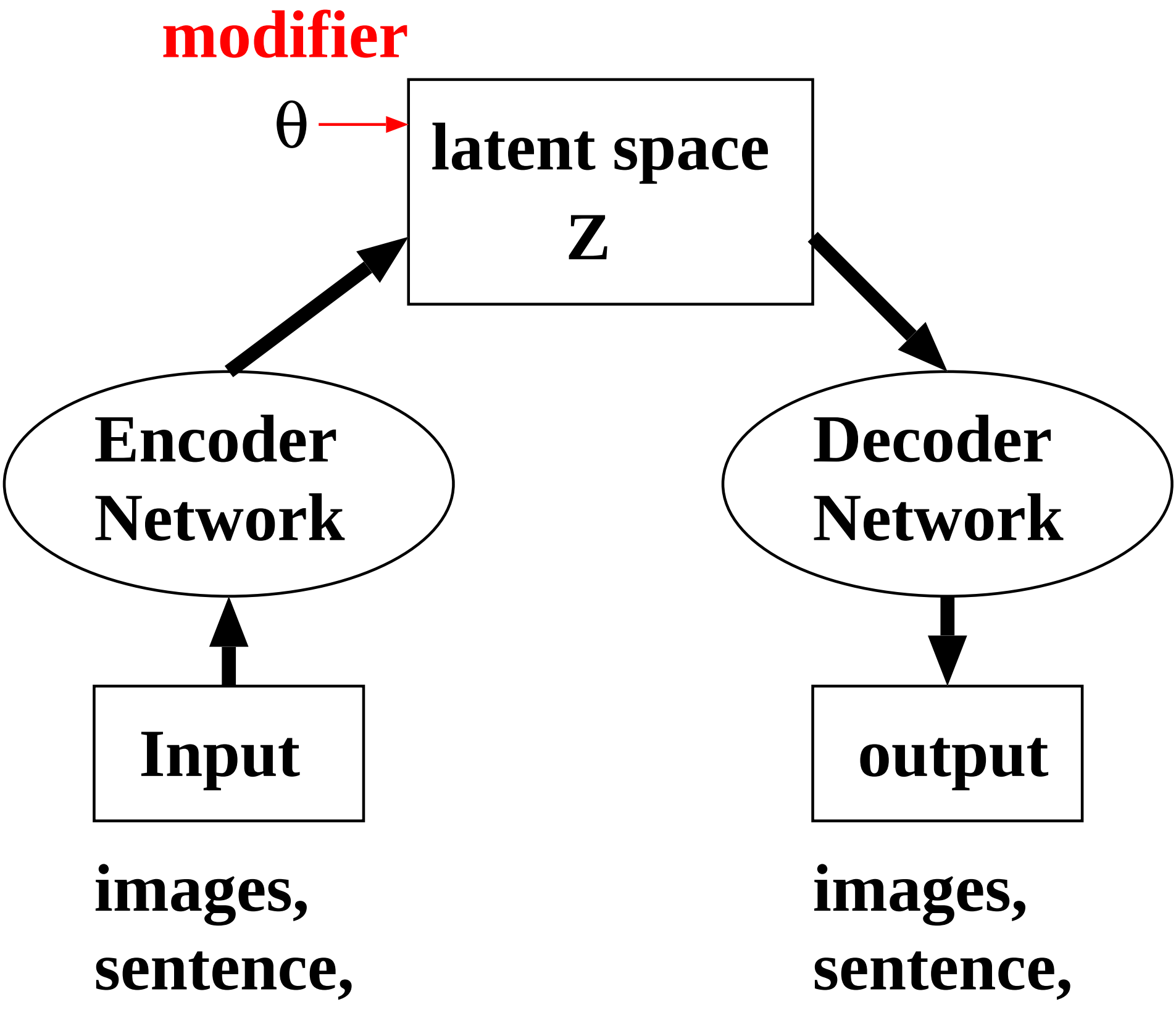
modifier
θ
latent space
z
Encoder
Network
Decoder
Network
Input
output
images,
sentence,
images,
sentence,