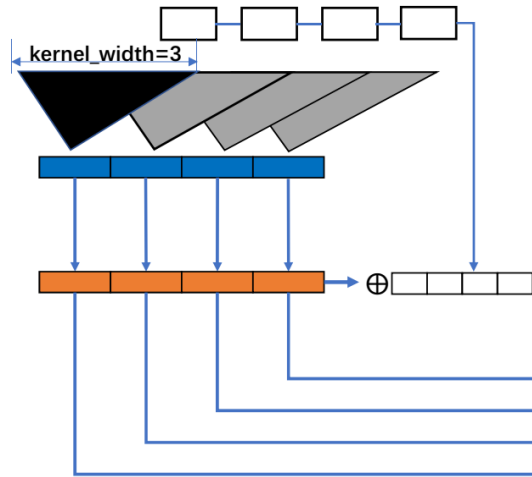


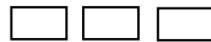
$\langle p \rangle \langle p \rangle X_1 X_2 X_3 \langle \text{EOS} \rangle$



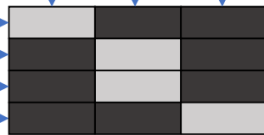
Encoder

Decoder

$\langle \text{EOS} \rangle Y_1 Y_2 \langle p \rangle \langle p \rangle$



Embedding  
Convolutional  
GLU Layer



Attention  
Matrix



$Y'_1 Y'_2 \langle \text{EOS} \rangle$