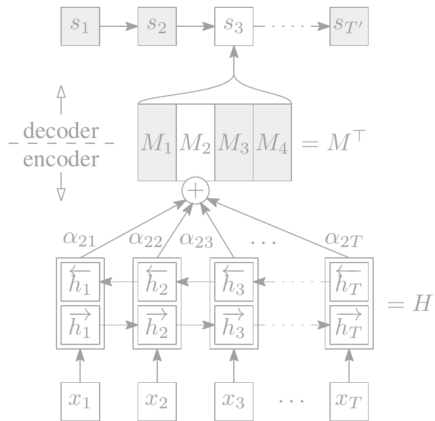
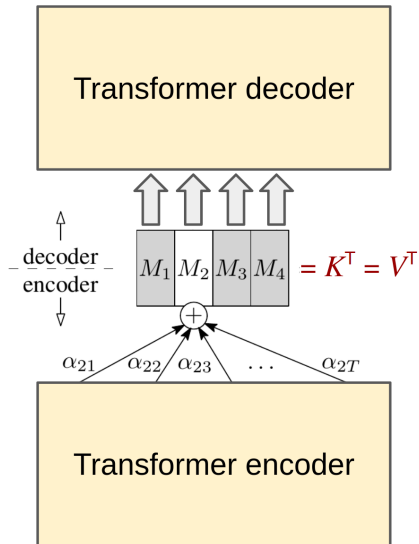


# Proposed NMT architectures



**ATTN-CTX**  
decoder operates on entire  
embedding



**TRF-ATTN-ATTN**  
Transformer (Vaswani et al., 2017)  
with inner attention