# Transformer-xl

1. 回忆transformer中的第一个encoder。

相对位置向量加上词向量乘以WQ得到Q,乘以WK得到K，乘以WV得到V。

E表示xi的词向量表示，

