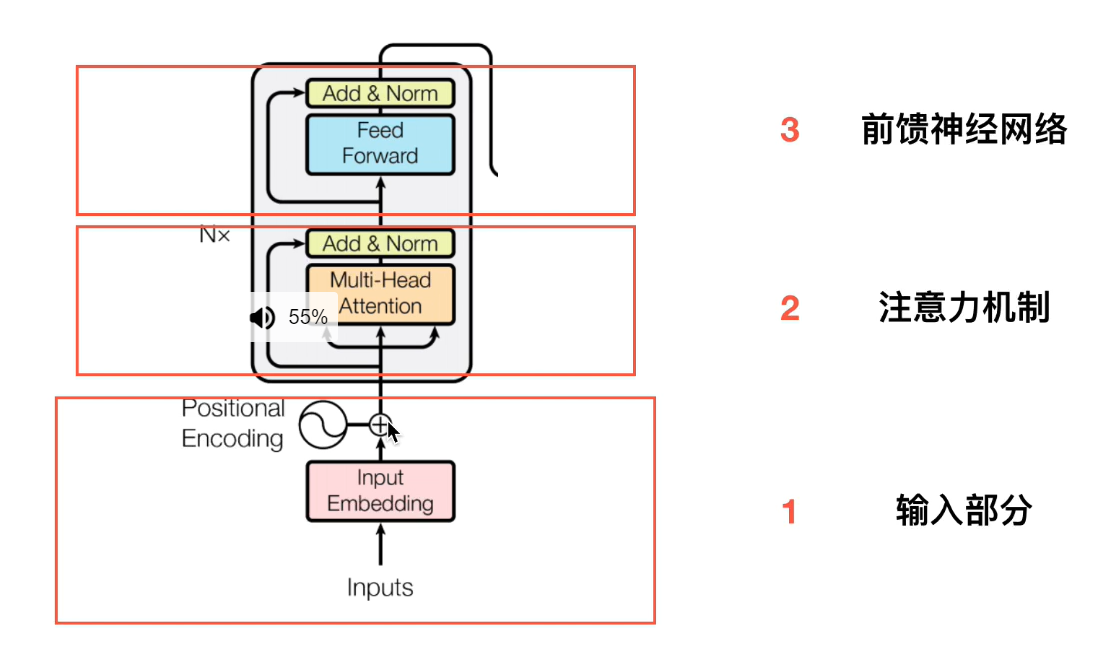
**Transformer:**

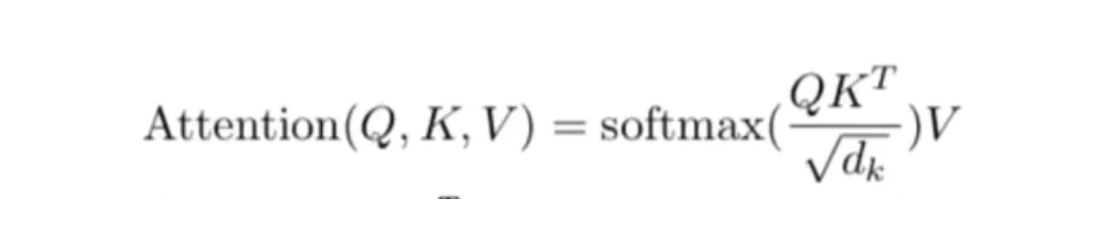
各个小的Encoder和Decoder结构相同，但是参数不同，独立训练.

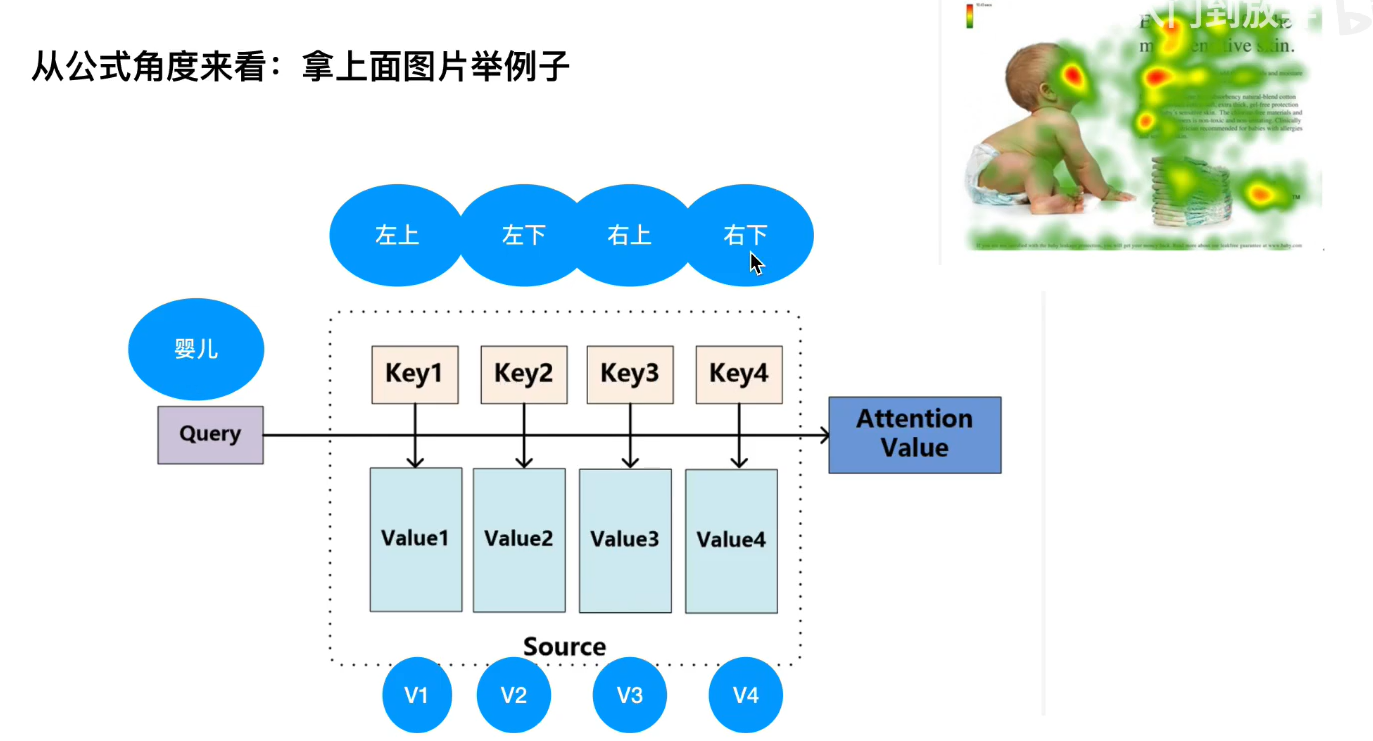
结构：

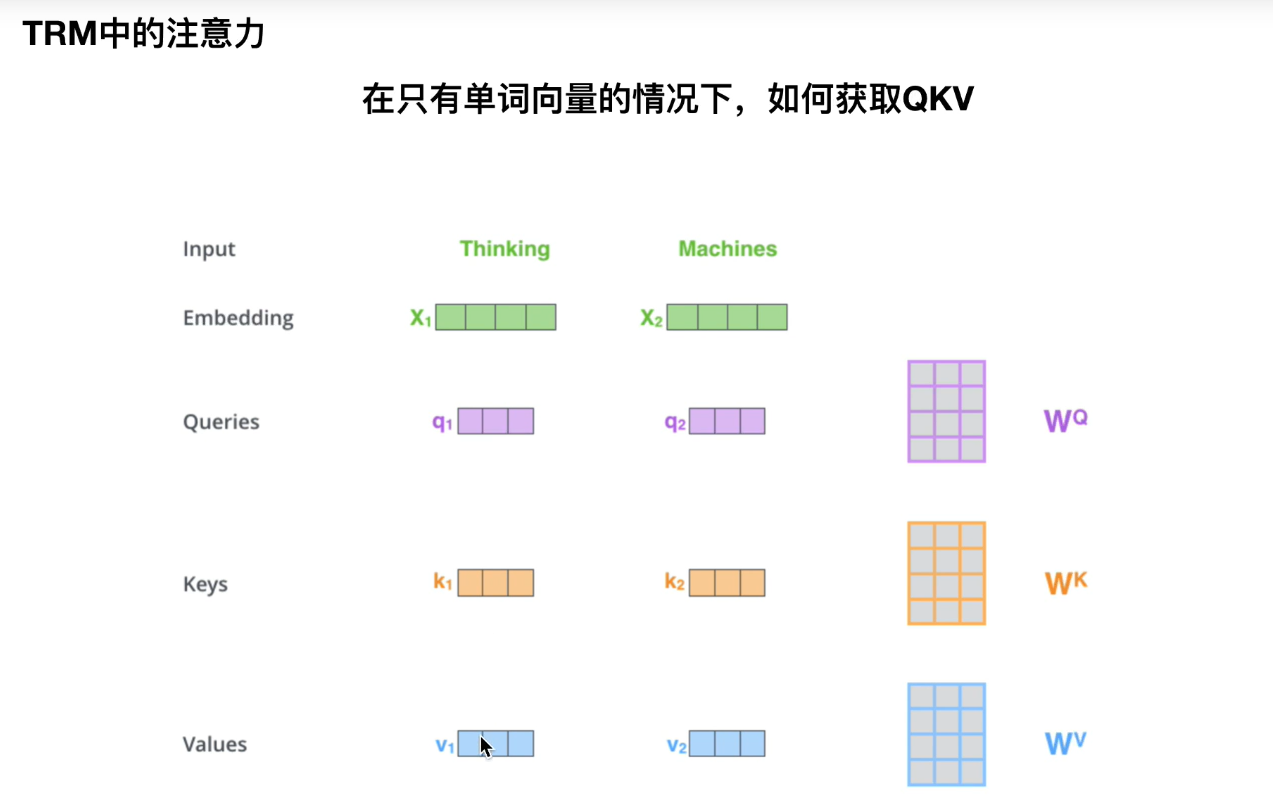


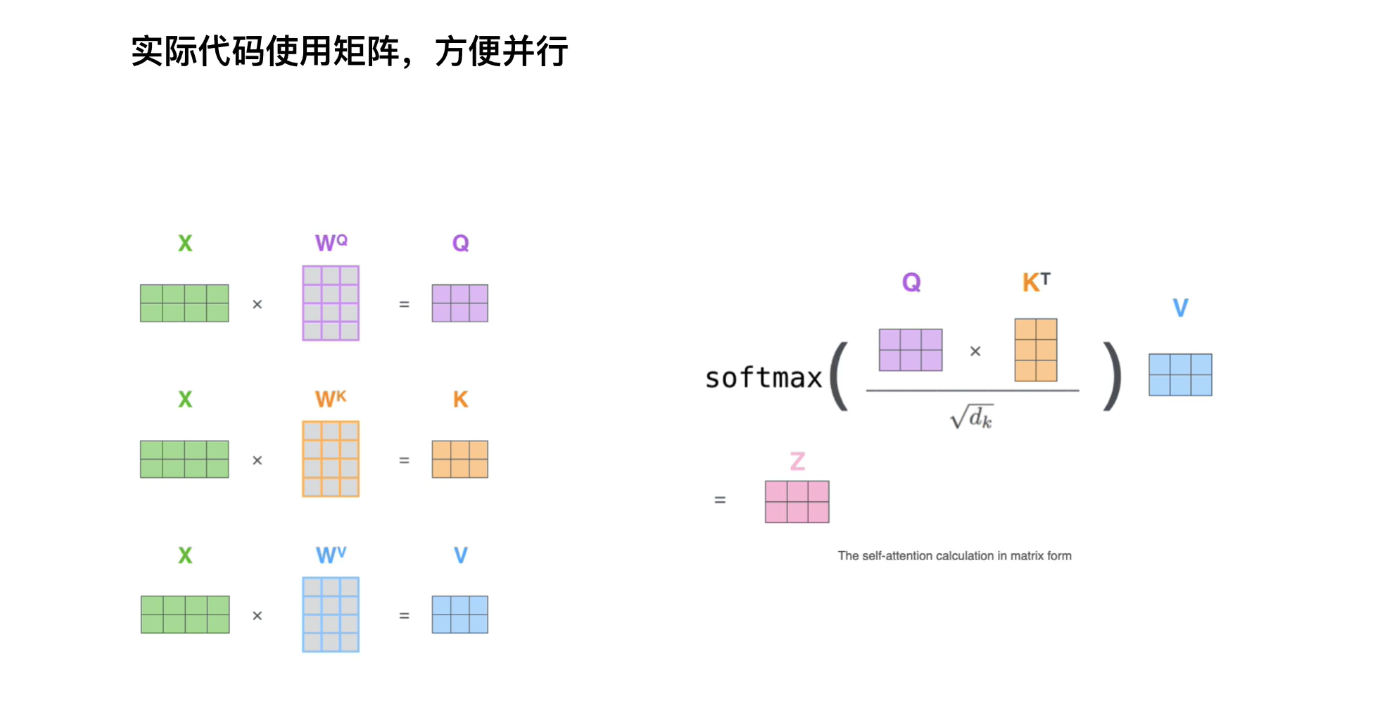
位置嵌入模块：

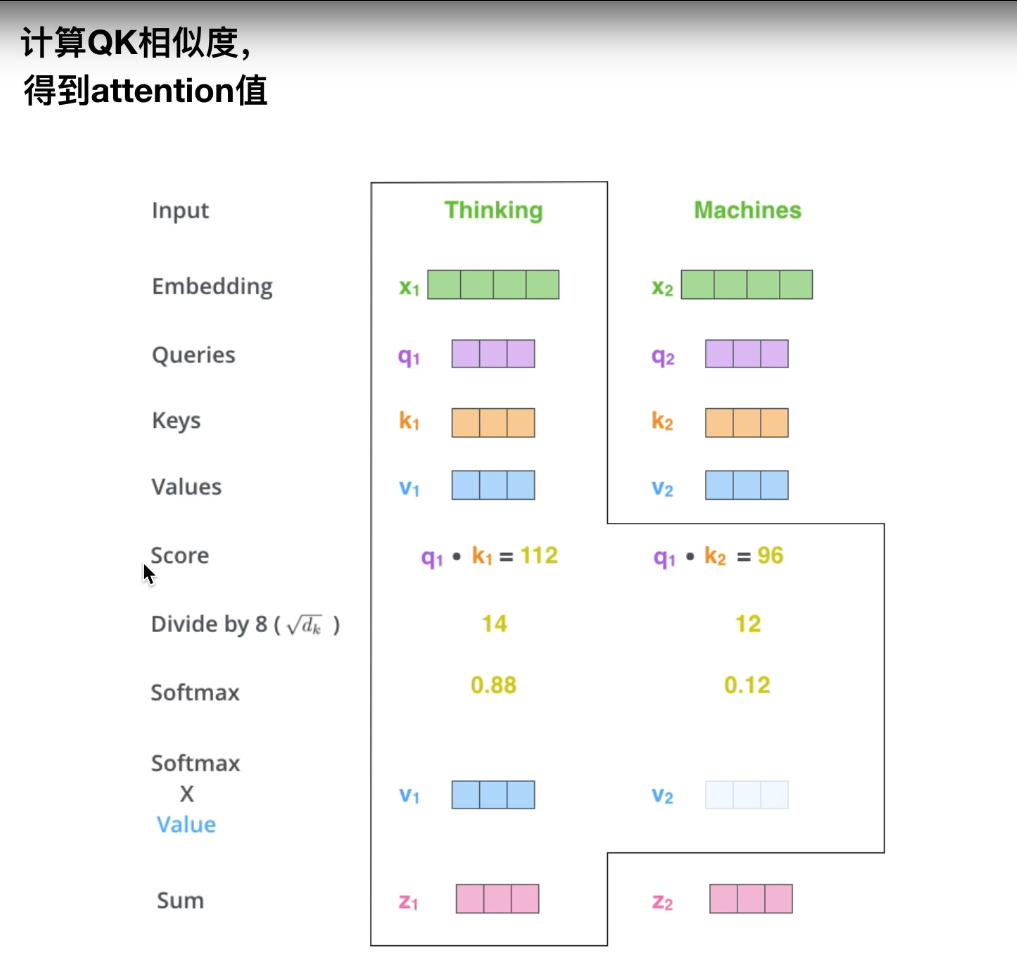
注意力机制（三个参数）：

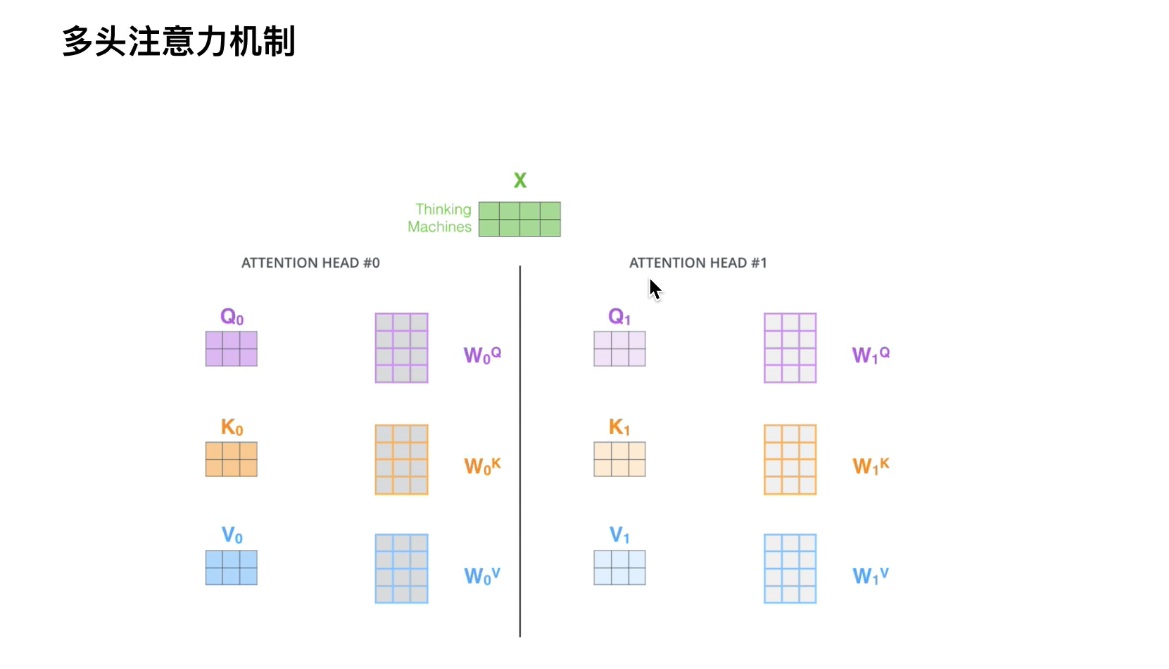


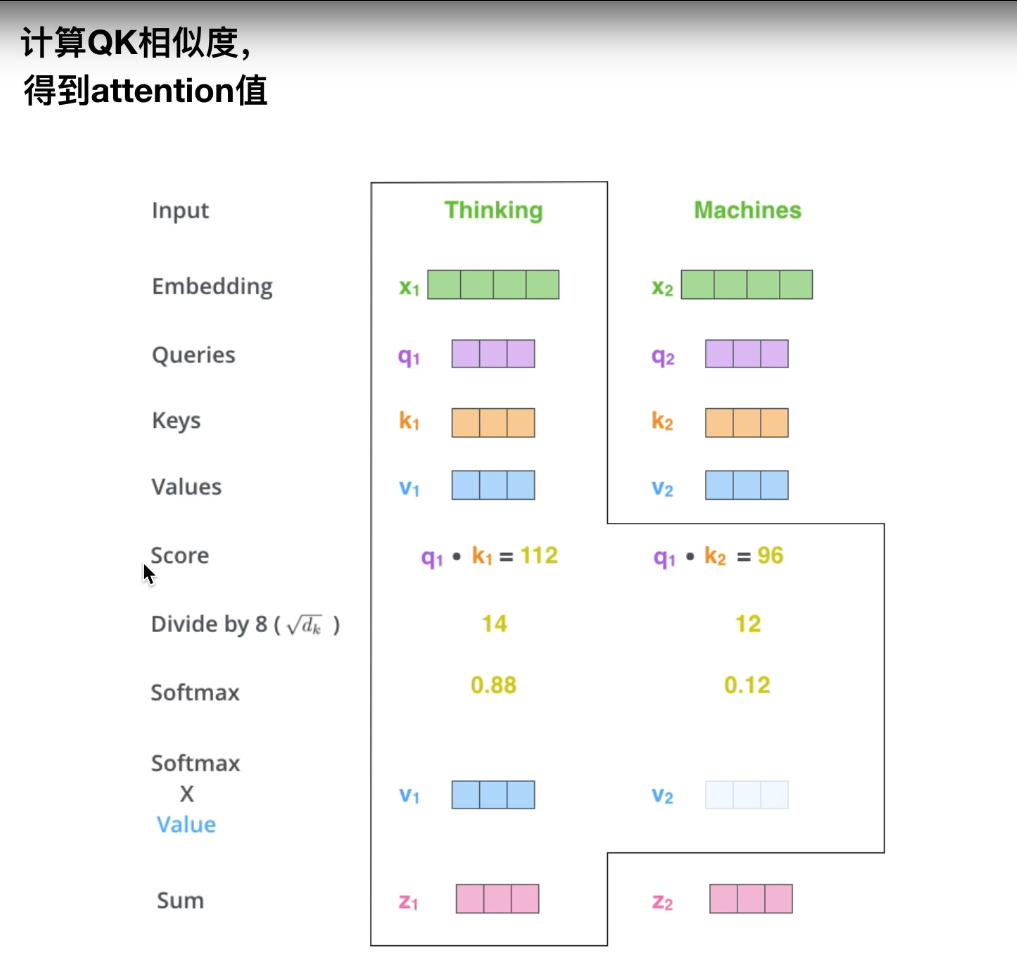




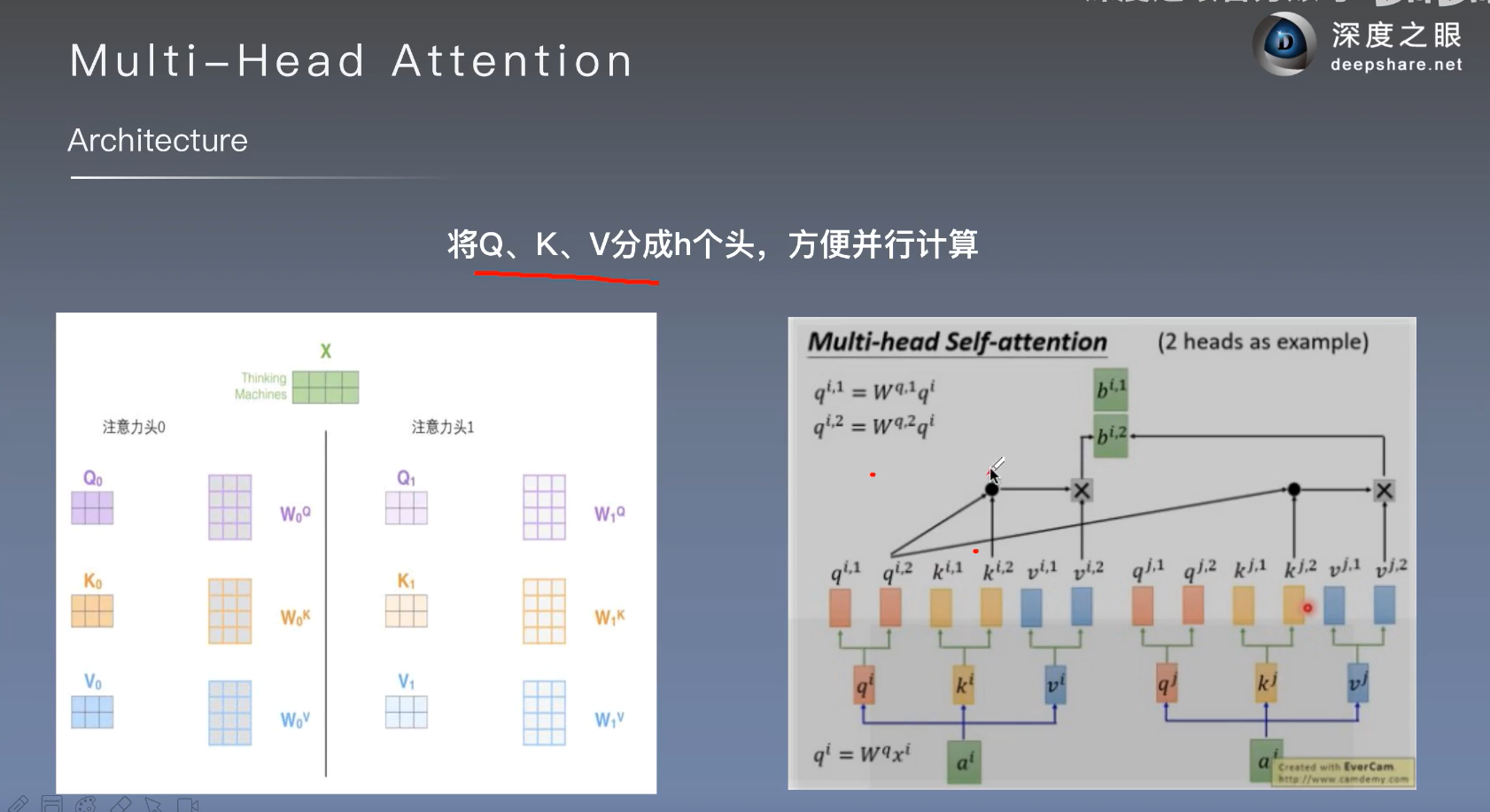


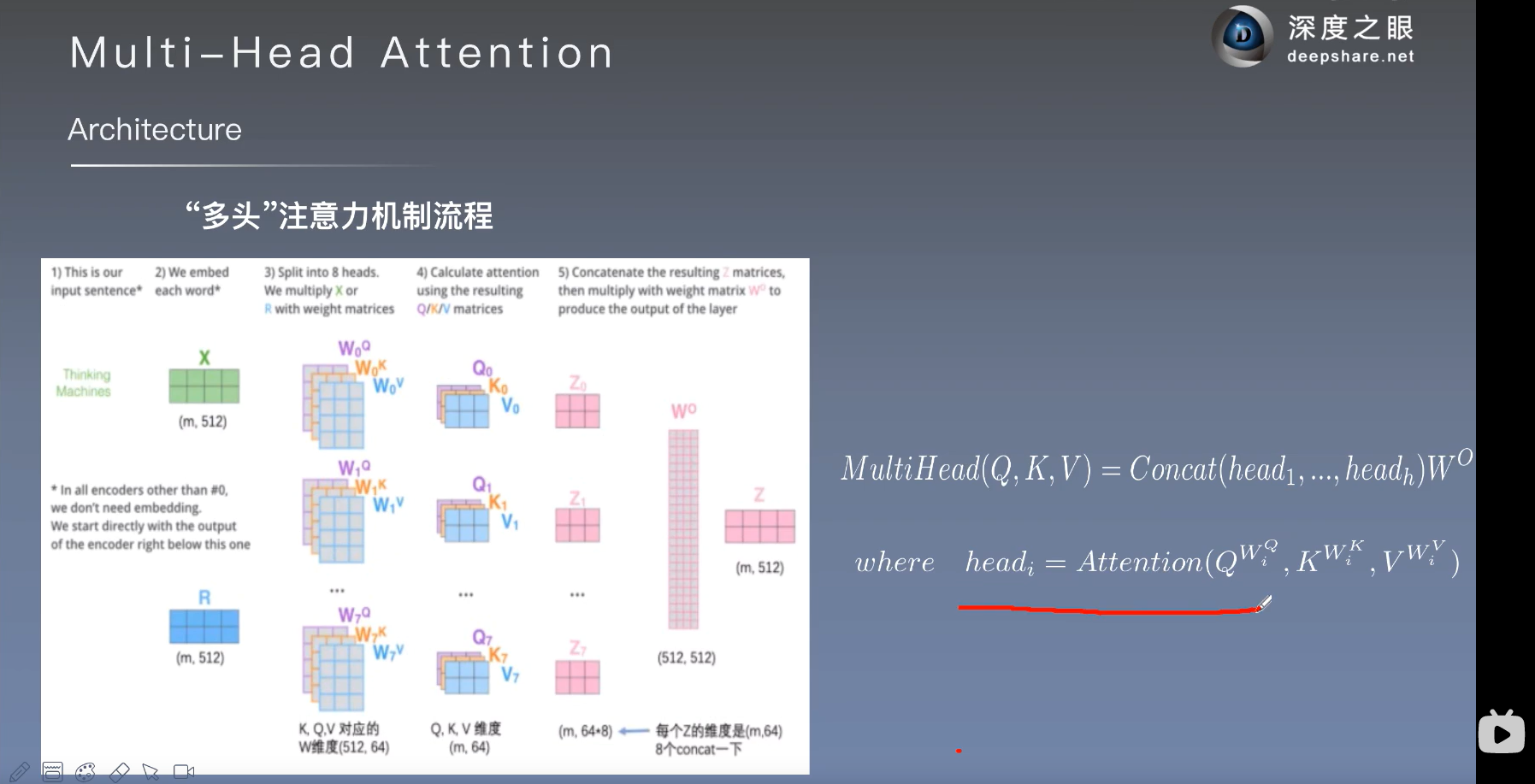


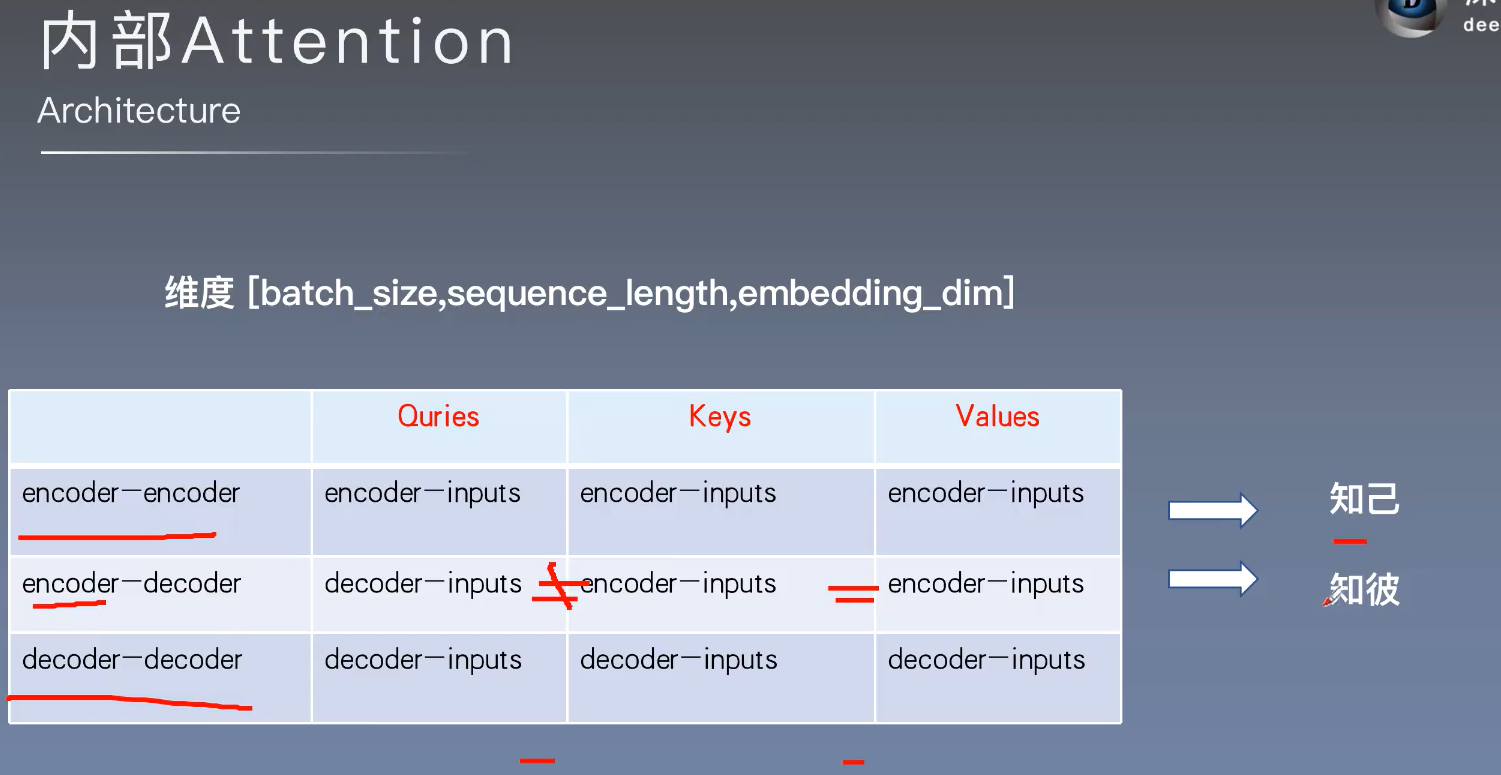




**多头注意力机制：**







前馈神经网络：

