

# Step-by-Step: Computing Attention

Each step preserves  
sequence length  $n$   
and transforms features

