

## The Need for Positional Information

### ⚠ Core Problem

Self-attention is **permutation invariant**  
Without position information, word order is ignored!

"cat ate mouse"

=

"mouse ate cat"



### RNN Approach

- ✓ Inherently encodes position through sequential processing
- ⌚ Position is implicit in the processing order
- ⌚ No additional mechanism needed



### Transformer Approach

- ✗ Parallel processing = **no inherent position**
- ✚ Needs **explicit positional information**
- ⌚ Critical: Word order is crucial for meaning



### Solution

**Positional encodings** must be added to input embeddings **before the first layer**