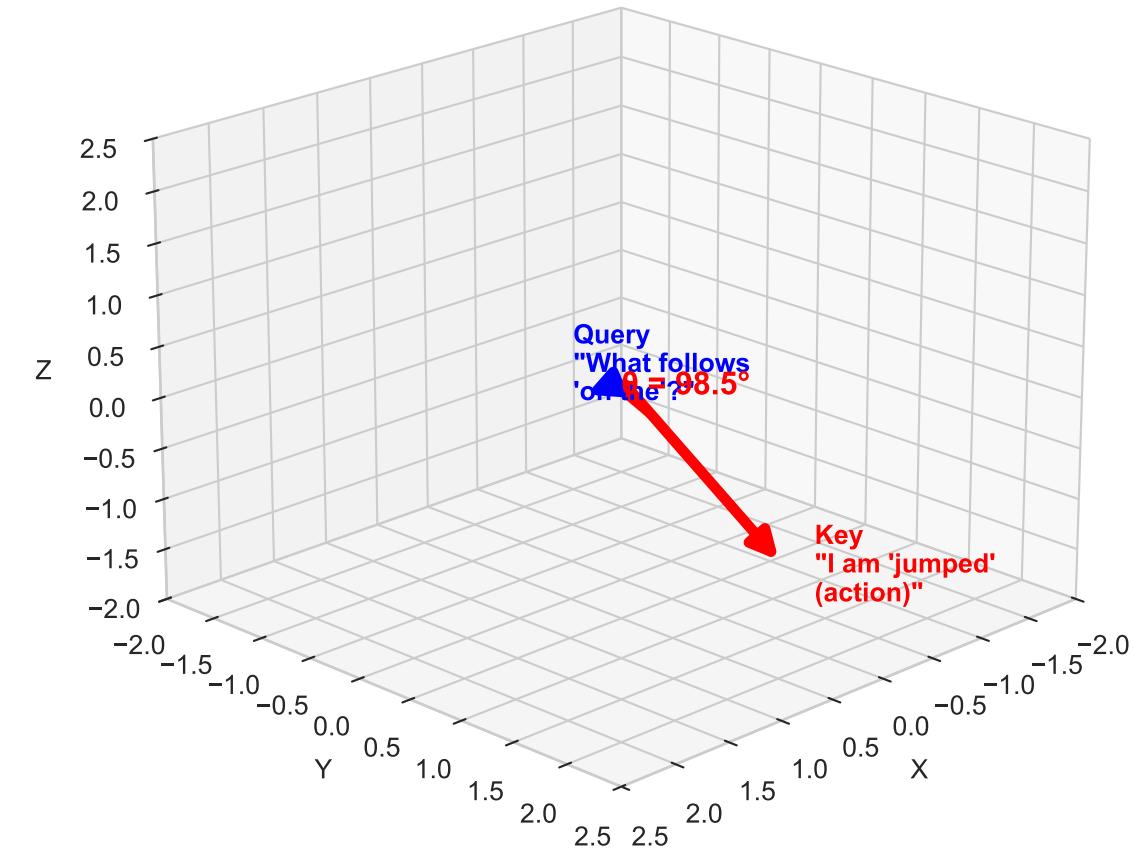
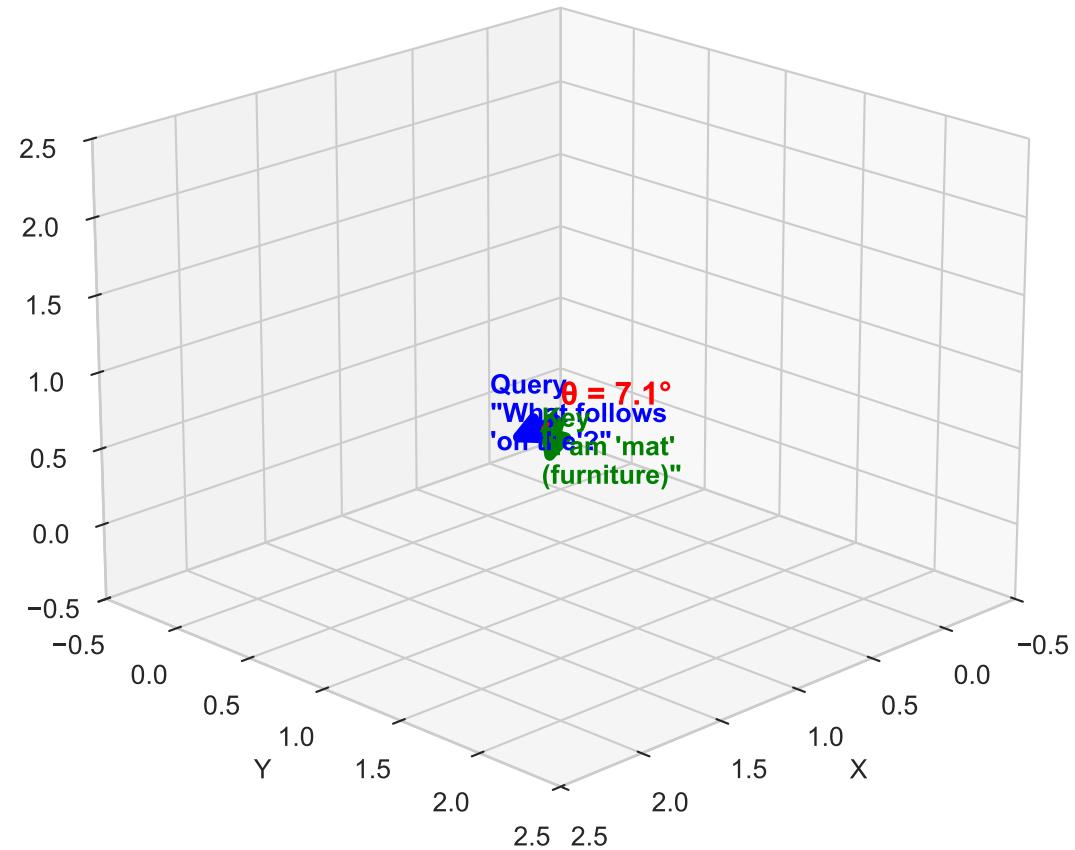


HIGH Similarity
Angle = 7.1°
Dot Product = 3.97
Similarity = 99.2%

Dot Product Measures Similarity: $\cos(\text{angle}) \times \text{magnitude}$

LOW Similarity
Angle = 98.5°
Dot Product = -0.59
Similarity = 0.0%



Attention Score = $Q \cdot K = |Q| \times |K| \times \cos(\theta)$
Small angle \rightarrow High dot product \rightarrow Strong attention
Large angle \rightarrow Low dot product \rightarrow Weak attention

In practice: 512-dimensional vectors, but same principle applies!