

Temperature Calculation: Step-by-Step

Given: Logits = [2.0, 1.0, 0.5, 0.2]
Tokens = ["cat", "dog", "bird", "fish"]

Step 1: Scale Scaled by 0.5 (MORE FOCUSED)
Scaled = $\frac{[2.0, 1.0, 0.5, 0.2]}{\max(2.0, 1.0, 0.5, 0.2)} \cdot 0.5 = [1.0, 0.5, 0.19, 0.12]$
Softmax → 73% on "cat" (BALANCED)
→ 73% on "cat" (VERY FOCUSED)

Step 3: Scale by T=2.0 (FLATTER)

$$\text{Softm } p_i = \frac{\exp(\text{logit}_i/T)}{\sum_{j=1}^4 \exp(\text{logit}_j/T)}$$

Lower $T \rightarrow$ More confident (peaky)
Higher $T \rightarrow$ More random (flat)

5, 0.1
3, 0.17, 0.15
CH FLATTER