

Input



Is there a fork in the image?

CoT Reasoning



What objects are often seen together with forks in images?



... , there are several objects, including a man eating food, a dining table, ...



<Image>. Is there a fork in the image?



<Output>

...



What are the relationships between objects?



... The man is wearing a yellow and black jacket, and he appears to be enjoying his meal ...



<Image>. Is there a fork in the image?

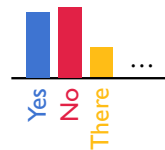


<Output>

Assembly Decoding

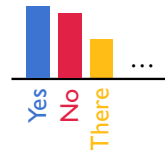
Original

$\text{logit}(y | c)$



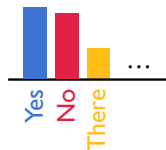
CoT_1

$\text{logit}(y | c^{\text{CoT}_1})$



CoT_N

$\text{logit}(y | c^{\text{CoT}_N})$



...

Assembled logits



No.



Yes.



$$\text{logit}(y | c, c^{\text{CoT}_1, \dots, \text{CoT}_N})_{A-\text{CoT}} = \text{logit}(y | c) + \beta * \sum_{s=1}^N \text{logit}(y | c^{\text{CoT}_s}) / N$$

