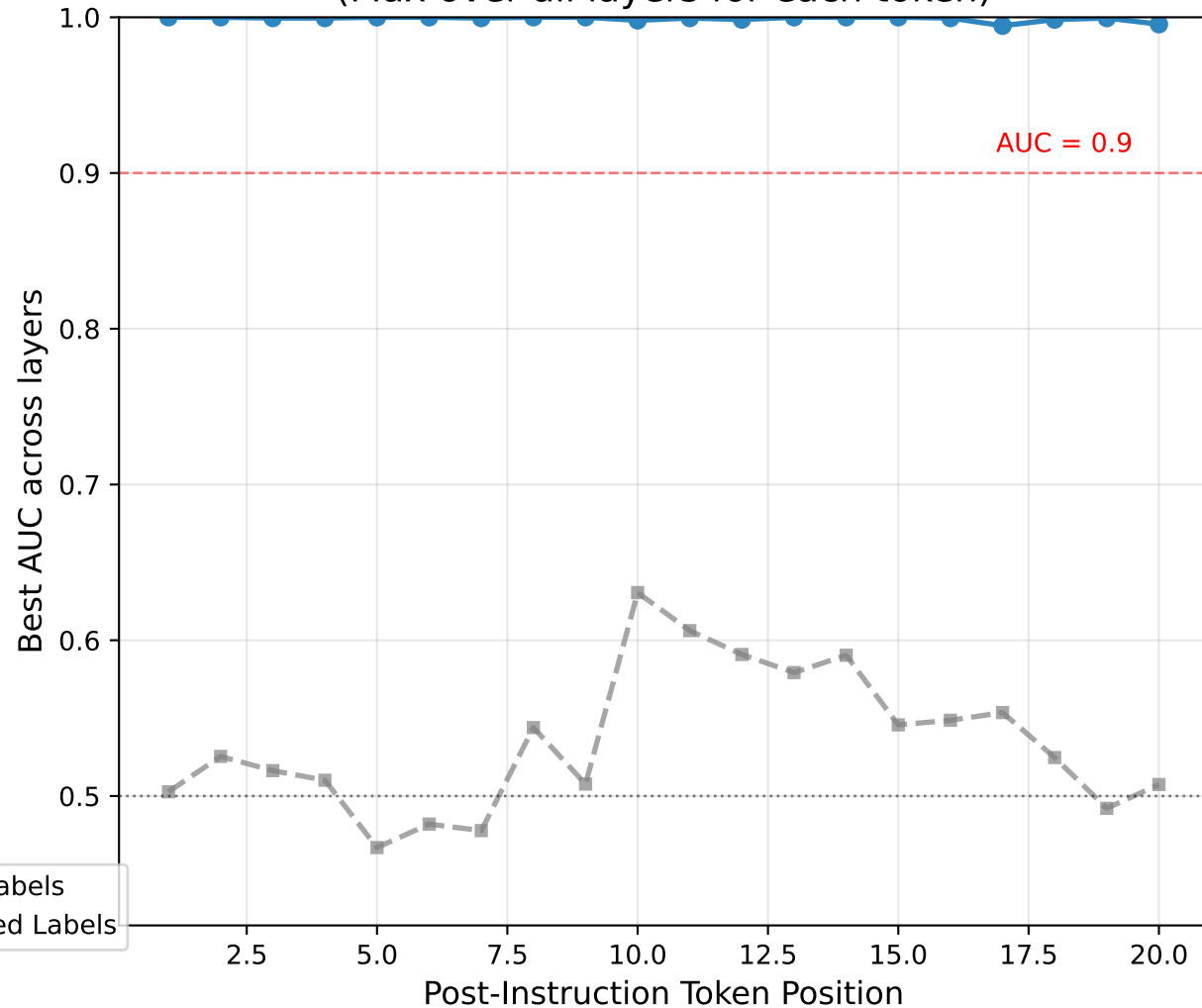


AUC vs Token Position  
(Max over all layers for each token)



AUC vs Layer  
(Max over all tokens for each layer)

