

# positions attended for all attention heads for pre-norm

