gradients/bert/encoder/layer_2/attention