gradients/bert/encoder/layer_1/attention/self