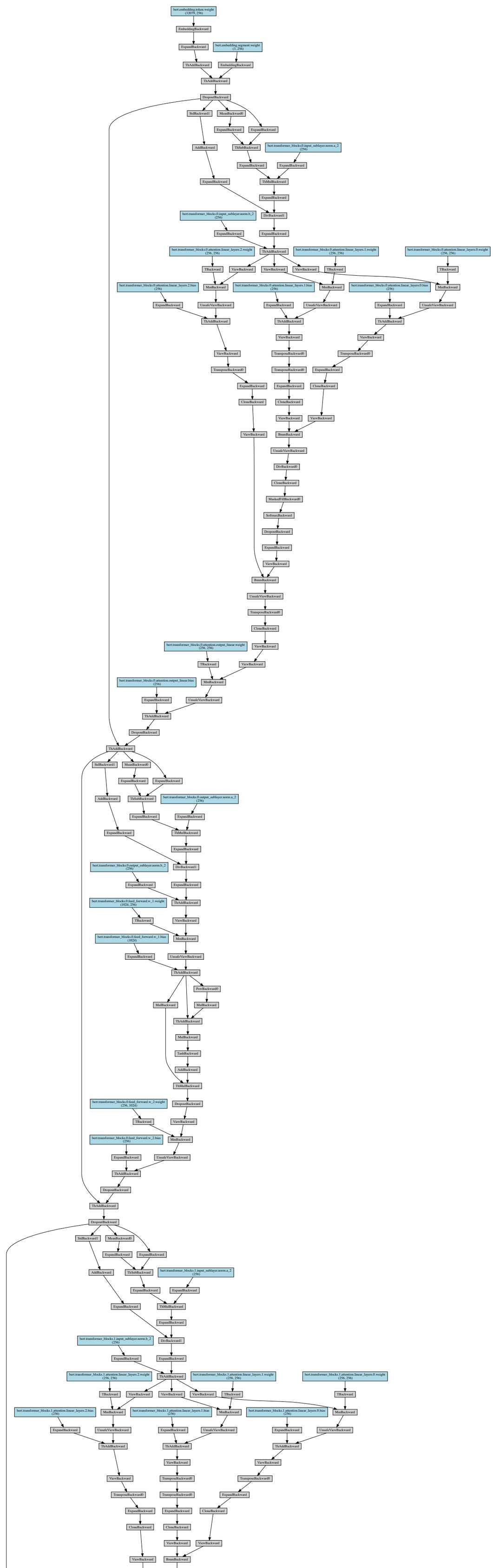
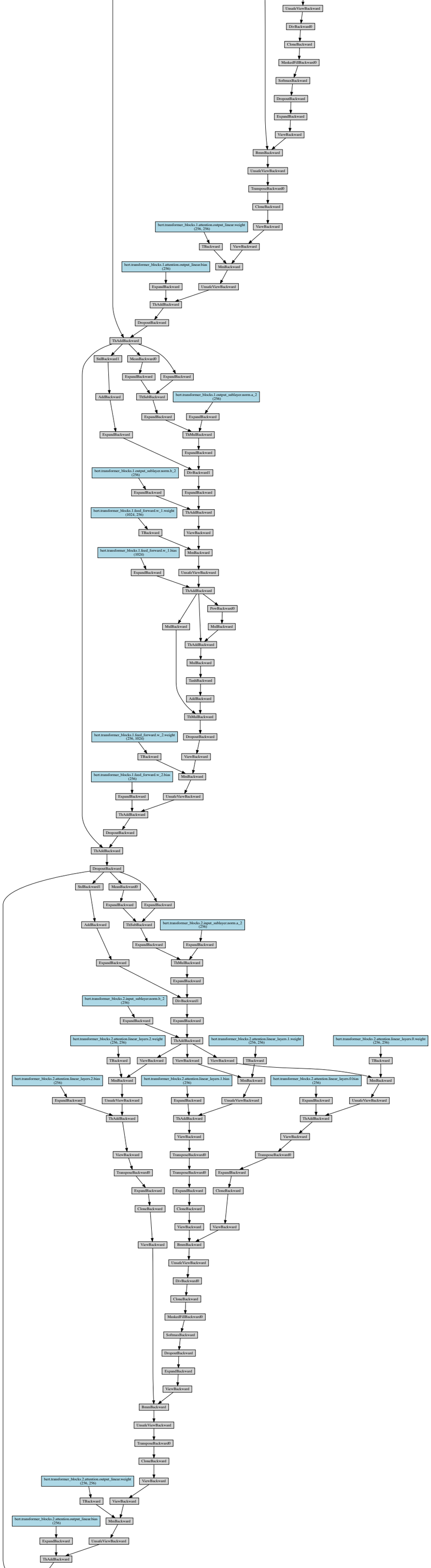
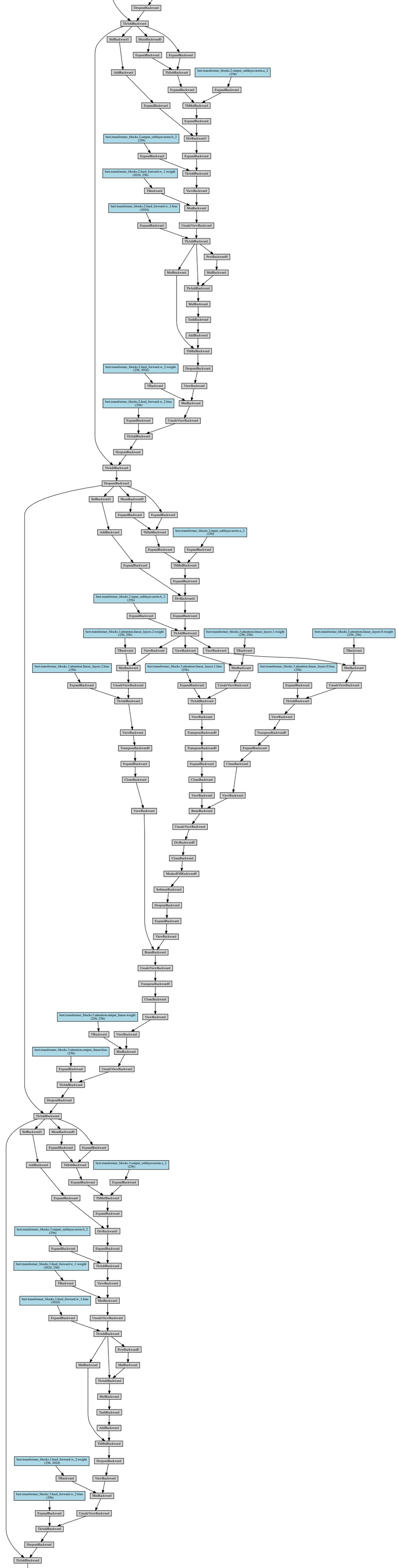
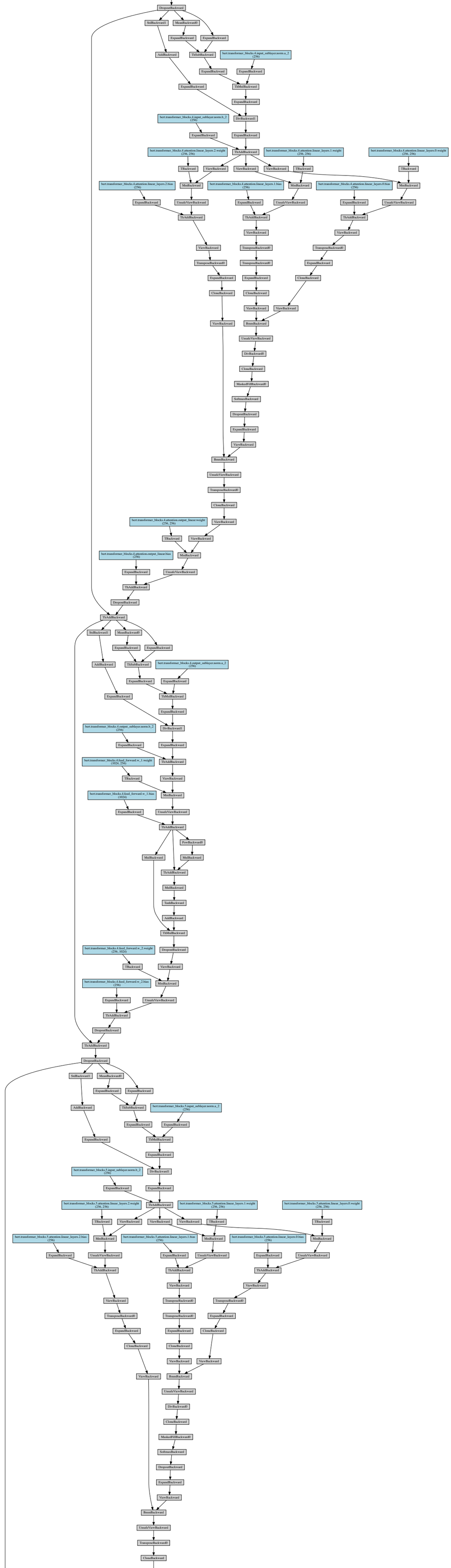
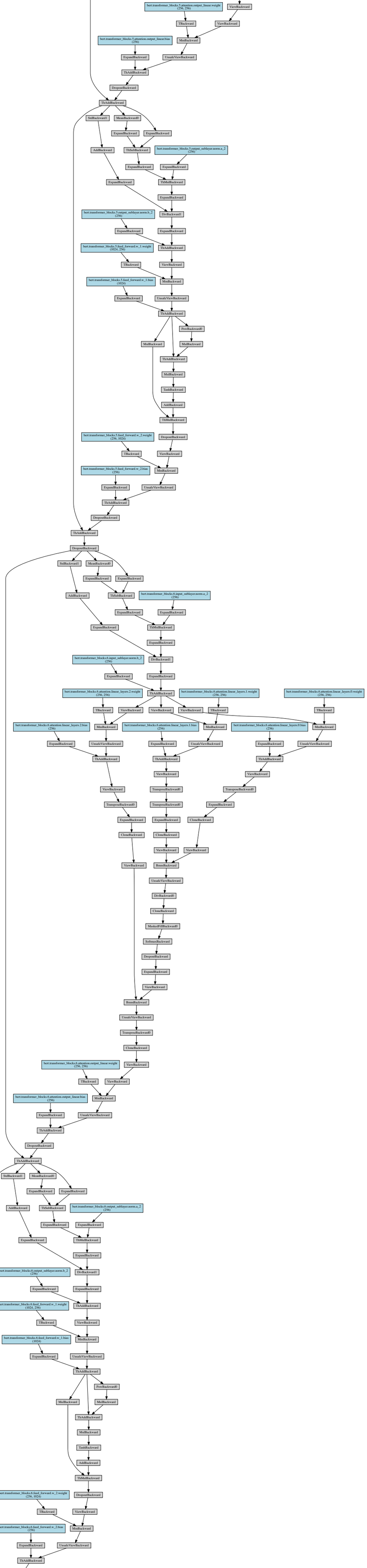
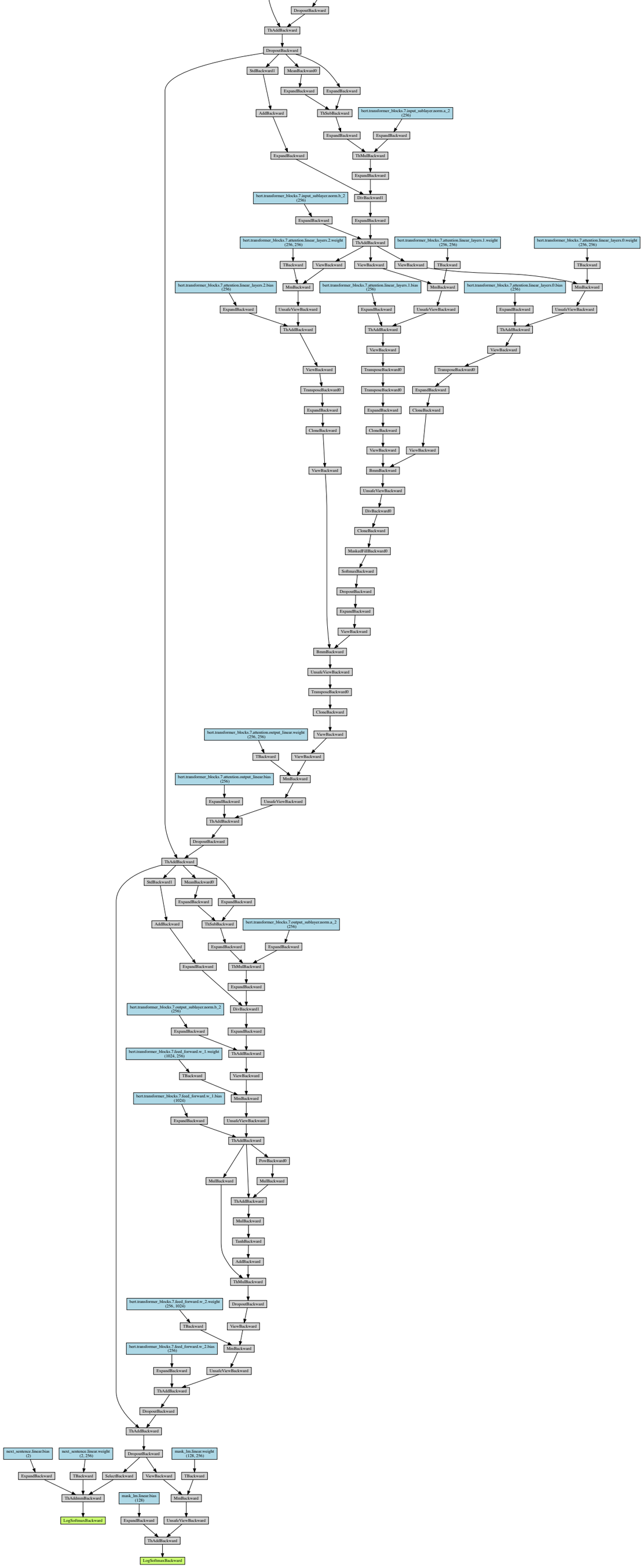

bert.embedding.token.weight
(32079, 256)
EmbeddingBackward
ExpandBackward
bert.embedding.segment.weight
(3, 256)
ThAddBackward
EmbeddingBackward
ThAddBackward
DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.0.input_sublayer.norm.a_2
(256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.0.input_sublayer.norm.b_2
(256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.0.attention.linear_layers.2.weight
(256, 256)
ThAddBackward
bert.transformer_blocks.0.attention.linear_layers.1.weight
(256, 256)
bert.transformer_blocks.0.attention.linear_layers.0.weight
(256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.0.attention.linear_layers.2.bias
(256)
MmBackward
bert.transformer_blocks.0.attention.linear_layers.1.bias
(256)
MmBackward
bert.transformer_blocks.0.attention.linear_layers.0.bias
(256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
ViewBackward
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward
UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TransposeBackward0
CloneBackward
ViewBackward
bert.transformer_blocks.0.attention.output_linear.weight
(256, 256)
TBackward
ViewBackward
bert.transformer_blocks.0.attention.output_linear.bias
(256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.0.output_sublayer.norm.a_2
(256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.0.output_sublayer.norm.b_2
(256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.0.feed_forward.w_1.weight
(1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.0.feed_forward.w_1.bias
(1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.0.feed_forward.w_2.weight
(256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.0.feed_forward.w_2.bias
(256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.1.input_sublayer.norm.a_2
(256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.1.input_sublayer.norm.b_2
(256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.1.attention.linear_layers.2.weight
(256, 256)
ThAddBackward
bert.transformer_blocks.1.attention.linear_layers.1.weight
(256, 256)
bert.transformer_blocks.1.attention.linear_layers.0.weight
(256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.1.attention.linear_layers.2.bias
(256)
MmBackward
bert.transformer_blocks.1.attention.linear_layers.1.bias
(256)
MmBackward
bert.transformer_blocks.1.attention.linear_layers.0.bias
(256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
ViewBackward
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward


UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TransposeBackward0
CloneBackward
bert.transformer_blocks.1.attention.output_linear.weight (256, 256)
ViewBackward
TBackward
ViewBackward
bert.transformer_blocks.1.attention.output_linear.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.1.output_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.1.output_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.1.feed_forward.w_1.weight (1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.1.feed_forward.w_1.bias (1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.1.feed_forward.w_2.weight (256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.1.feed_forward.w_2.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.2.input_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.2.input_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.2.attention.linear_layers.2.weight (256, 256)
ThAddBackward
bert.transformer_blocks.2.attention.linear_layers.1.weight (256, 256)
bert.transformer_blocks.2.attention.linear_layers.0.weight (256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.2.attention.linear_layers.2.bias (256)
MmBackward
bert.transformer_blocks.2.attention.linear_layers.1.bias (256)
MmBackward
bert.transformer_blocks.2.attention.linear_layers.0.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
ViewBackward
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward
UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TransposeBackward0
CloneBackward
bert.transformer_blocks.2.attention.output_linear.weight (256, 256)
ViewBackward
TBackward
ViewBackward
bert.transformer_blocks.2.attention.output_linear.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward


DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.2.output_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.2.output_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.2.feed_forward.w_1.weight (1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.2.feed_forward.w_1.bias (1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.2.feed_forward.w_2.weight (256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.2.feed_forward.w_2.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.3.input_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.3.input_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.3.attention.linear_layers.2.weight (256, 256)
ThAddBackward
bert.transformer_blocks.3.attention.linear_layers.1.weight (256, 256)
bert.transformer_blocks.3.attention.linear_layers.0.weight (256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.3.attention.linear_layers.2.bias (256)
MmBackward
bert.transformer_blocks.3.attention.linear_layers.1.bias (256)
MmBackward
bert.transformer_blocks.3.attention.linear_layers.0.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
ViewBackward
TranspositionBackward0
TranspositionBackward0
TranspositionBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward
UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TranspositionBackward0
CloneBackward
ViewBackward
bert.transformer_blocks.3.attention.output_linear.weight (256, 256)
TBackward
ViewBackward
bert.transformer_blocks.3.attention.output_linear.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.3.output_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.3.output_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.3.feed_forward.w_1.weight (1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.3.feed_forward.w_1.bias (1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.3.feed_forward.w_2.weight (256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.3.feed_forward.w_2.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward


DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.4.input_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.4.input_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.4.attention.linear_layers.2.weight (256, 256)
ThAddBackward
bert.transformer_blocks.4.attention.linear_layers.1.weight (256, 256)
bert.transformer_blocks.4.attention.linear_layers.0.weight (256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.4.attention.linear_layers.2.bias (256)
MmBackward
bert.transformer_blocks.4.attention.linear_layers.1.bias (256)
MmBackward
bert.transformer_blocks.4.attention.linear_layers.0.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
ViewBackward
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward
UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TransposeBackward0
CloneBackward
ViewBackward
bert.transformer_blocks.4.attention.output_linear.weight (256, 256)
TBackward
ViewBackward
MmBackward
bert.transformer_blocks.4.attention.output_linear.bias (256)
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.4.output_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.4.output_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.4.feed_forward.w_1.weight (1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.4.feed_forward.w_1.bias (1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.4.feed_forward.w_2.weight (256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.4.feed_forward.w_2.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.5.input_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.5.input_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.5.attention.linear_layers.2.weight (256, 256)
ThAddBackward
bert.transformer_blocks.5.attention.linear_layers.1.weight (256, 256)
bert.transformer_blocks.5.attention.linear_layers.0.weight (256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.5.attention.linear_layers.2.bias (256)
MmBackward
bert.transformer_blocks.5.attention.linear_layers.1.bias (256)
MmBackward
bert.transformer_blocks.5.attention.linear_layers.0.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
ViewBackward
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward
UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TransposeBackward0
CloneBackward


bert.transformer_blocks.5.attention.output_linear.weight (256, 256)
ViewBackward
TBackward
ViewBackward
bert.transformer_blocks.5.attention.output_linear.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.5.output_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.5.output_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.5.feed_forward.w_1.weight (1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.5.feed_forward.w_1.bias (1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.5.feed_forward.w_2.weight (256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.5.feed_forward.w_2.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.6.input_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.6.input_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.6.attention.linear_layers.2.weight (256, 256)
ThAddBackward
bert.transformer_blocks.6.attention.linear_layers.1.weight (256, 256)
bert.transformer_blocks.6.attention.linear_layers.0.weight (256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.6.attention.linear_layers.2.bias (256)
MmBackward
bert.transformer_blocks.6.attention.linear_layers.1.bias (256)
MmBackward
bert.transformer_blocks.6.attention.linear_layers.0.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
ViewBackward
TransposeBackward0
TransposeBackward0
TransposeBackward0
TransposeBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward
UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TransposeBackward0
CloneBackward
bert.transformer_blocks.6.attention.output_linear.weight (256, 256)
ViewBackward
TBackward
ViewBackward
bert.transformer_blocks.6.attention.output_linear.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.6.output_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.6.output_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.6.feed_forward.w_1.weight (1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.6.feed_forward.w_1.bias (1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.6.feed_forward.w_2.weight (256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.6.feed_forward.w_2.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward


DropoutBackward
ThAddBackward
DropoutBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.7.input_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.7.input_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.7.attention.linear_layers.2.weight (256, 256)
ThAddBackward
bert.transformer_blocks.7.attention.linear_layers.1.weight (256, 256)
bert.transformer_blocks.7.attention.linear_layers.0.weight (256, 256)
TBackward
ViewBackward
ViewBackward
ViewBackward
TBackward
TBackward
bert.transformer_blocks.7.attention.linear_layers.2.bias (256)
MmBackward
bert.transformer_blocks.7.attention.linear_layers.1.bias (256)
MmBackward
bert.transformer_blocks.7.attention.linear_layers.0.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
ThAddBackward
ThAddBackward
ViewBackward
ViewBackward
TranspositionBackward0
TranspositionBackward0
ViewBackward
TranspositionBackward0
TranspositionBackward0
ExpandBackward
ExpandBackward
ExpandBackward
CloneBackward
CloneBackward
CloneBackward
ViewBackward
ViewBackward
ViewBackward
BmmBackward
UnsafeViewBackward
DivBackward0
CloneBackward
MaskedFillBackward0
SoftmaxBackward
DropoutBackward
ExpandBackward
ViewBackward
BmmBackward
UnsafeViewBackward
TranspositionBackward0
CloneBackward
ViewBackward
bert.transformer_blocks.7.attention.output_linear.weight (256, 256)
TBackward
ViewBackward
bert.transformer_blocks.7.attention.output_linear.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
StdBackward1
MeanBackward0
ExpandBackward
ExpandBackward
AddBackward
ThSubBackward
bert.transformer_blocks.7.output_sublayer.norm.a_2 (256)
ExpandBackward
ExpandBackward
ExpandBackward
ThMulBackward
ExpandBackward
bert.transformer_blocks.7.output_sublayer.norm.b_2 (256)
DivBackward1
ExpandBackward
ExpandBackward
bert.transformer_blocks.7.feed_forward.w_1.weight (1024, 256)
ThAddBackward
TBackward
ViewBackward
bert.transformer_blocks.7.feed_forward.w_1.bias (1024)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
PowBackward0
MulBackward
MulBackward
ThAddBackward
MulBackward
TanhBackward
AddBackward
ThMulBackward
bert.transformer_blocks.7.feed_forward.w_2.weight (256, 1024)
DropoutBackward
TBackward
ViewBackward
bert.transformer_blocks.7.feed_forward.w_2.bias (256)
MmBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
DropoutBackward
ThAddBackward
next_sentence.linear.bias (2)
next_sentence.linear.weight (2, 256)
DropoutBackward
mask_lm.linear.weight (128, 256)
ExpandBackward
TBackward
SelectBackward
ViewBackward
TBackward
ThAddmmBackward
mask_lm.linear.bias (128)
MmBackward
LogSoftmaxBackward
ExpandBackward
UnsafeViewBackward
ThAddBackward
LogSoftmaxBackward