

attention loss on AM dataset during training epochs (bisim)

