Pythia 410m (Seed 4) Previous/Self Attention Ratio 1e+011e+031e+0 Training Step (Log10)