SELF-ACTIVATING NEURAL ENSEMBLES FOR CONTINUAL REINFORCEMENT LEARNING: REBUTTAL

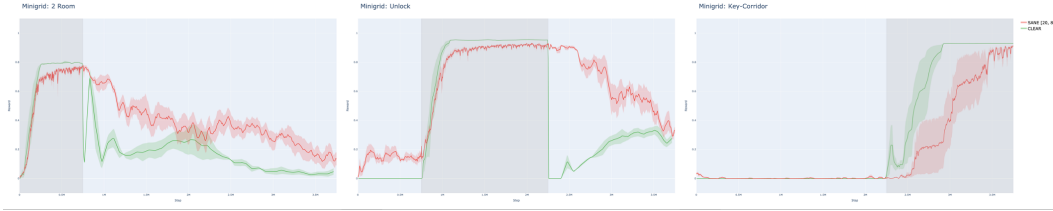**Anonymous authors**
Paper under double-blind review

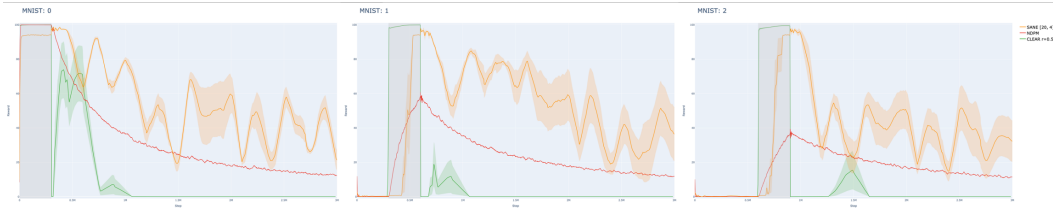Figure 1: Results on a different set of Minigrid experiments: 2 Room, Unlock, Key-Corridor.

Figure 2: Results comparing SANE to CLEAR (replay ratio of 0.5) and Neural Dirichlet Process Mixtures (NDPM)
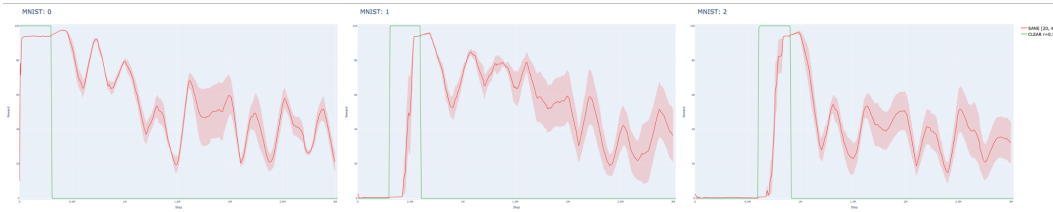

Figure 3: Results of early stopping CLEAR during MNIST training. We stop as soon as it reaches a score of 97, then keep evaluating on that environment for the duration of the training period, to keep synchronization with SANE, which is presented for comparison and was not stopped early.
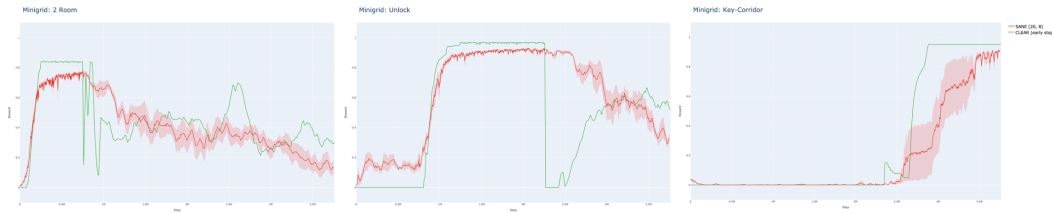
Figure 4: Results of early stopping CLEAR during Minigrid, using the same set of 3 environments shown in 1. As in 3, SANE is shown for comparison and was not stopped early.
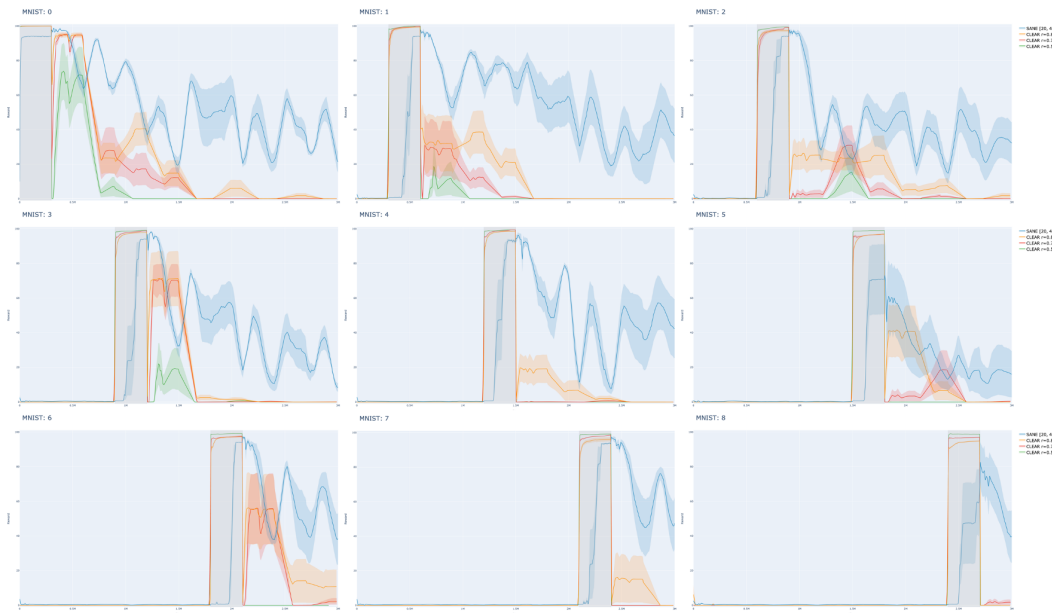

Figure 5: Results of training CLEAR with several replay ratios, using a batch size of 8.
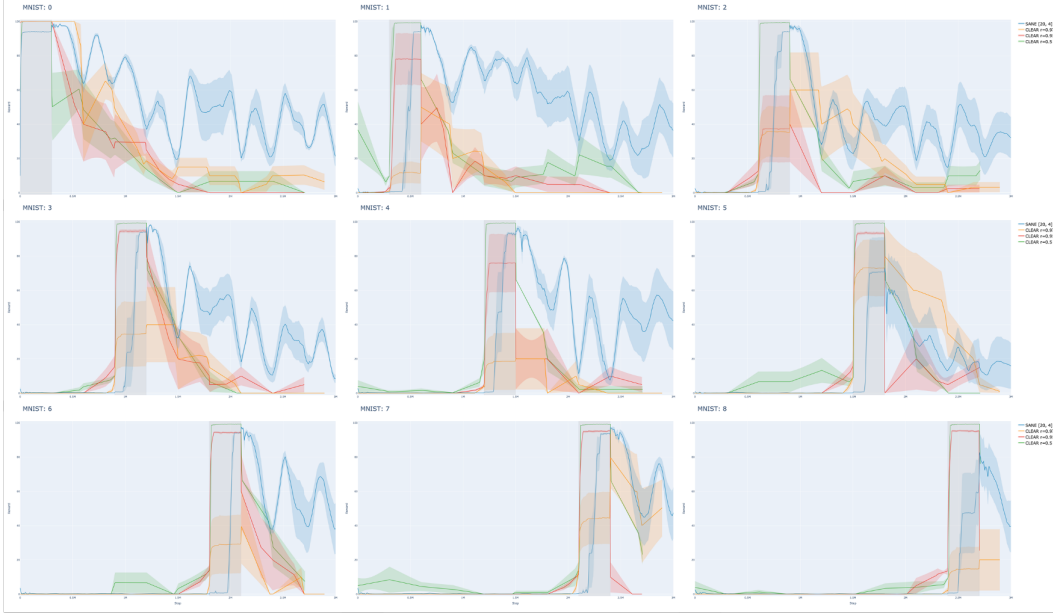
Figure 6: Results of training CLEAR with several replay ratios, using a batch size of 128.
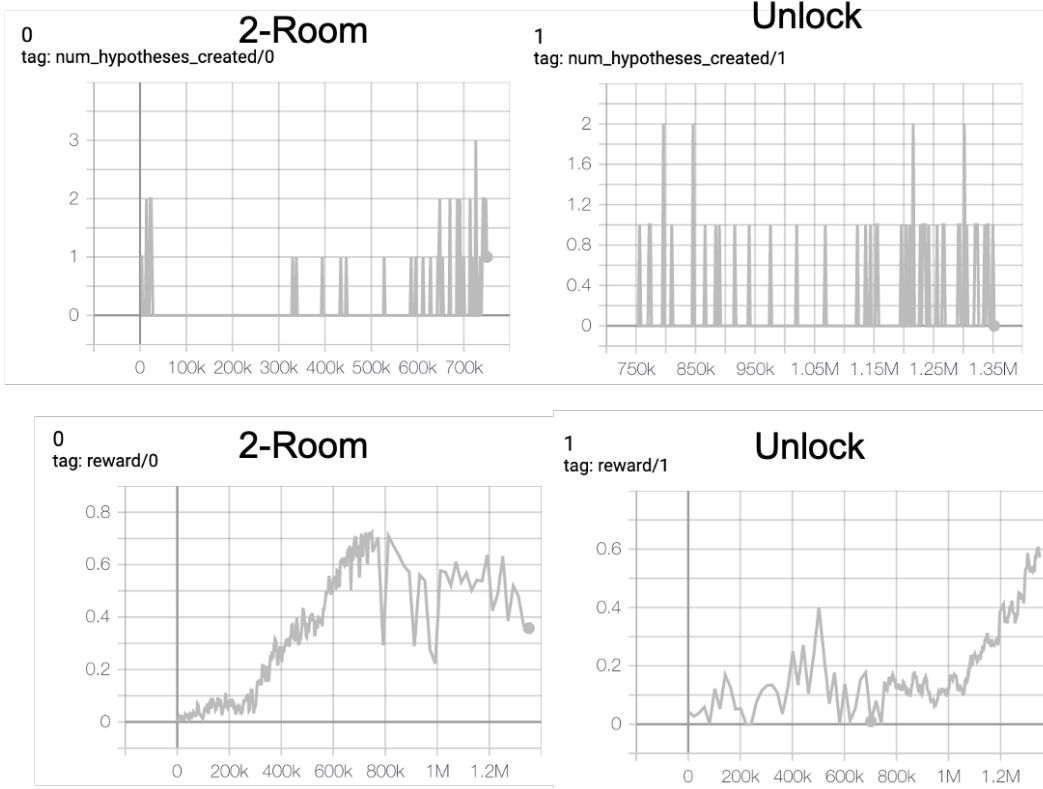

Figure 7: This figure shows the timesteps when new nodes were added to the graph. The learning curve is also included to allow for intuitive alignment. Note that the timestep axes are not aligned between the hypothesis creation graphs and the learning curve. In both cases we see more creation during the up-slope of the learning curve, and less earlier on in a task's training.