

The graph displays the performance of 15 reinforcement learning algorithms over 200,000 steps. The x-axis is labeled 'Step' and ranges from 0 to 200,000. The y-axis represents a performance metric, with a horizontal line at 0.9. The algorithms are listed in the legend: A2C, C51, DoubleDQN, DQN, DRQN, DuelingDQN, NoisyDQN, PerDQN, PG, PPG, PPO, QRDQN, and SAC. PPO (purple line) shows the highest and most stable performance, reaching a plateau around 0.9. Other algorithms like C51, DoubleDQN, and DuelingDQN show significant improvement over time, while others like A2C and PG remain relatively flat at lower performance levels.

