

ViPER: Visibility-based Pursuit-Evasion via Reinforcement Learning – Supplementary Materials

Anonymous Author(s)

Affiliation

Address

email

1 Ablation Results

2 We perform a comprehensive ablation study to evaluate the effectiveness of our attentive critic and
3 privileged learning module. Specifically, we train three additional ablation baselines: ViPER without
4 the MAAC (attentive critic module), ViPER without GT (ground truth information during training),
5 and ViPER without both MAAC and GT. We present the reward and length curves during training
6 in Fig. 1. It is evident that ViPER consistently outperforms its baselines throughout the training
7 process. We further observe a significant performance decline without privileged learning, indicating
8 the importance of access to ground truth information for the critic to achieve a more accurate value
9 estimate during training. Our ablation study demonstrates that both the attentive critic and privileged
10 learning (individually and combined) enhance ViPER’s ability to learn a better policy for multi-agent
11 coordination in visibility-based pursuit-evasion tasks.

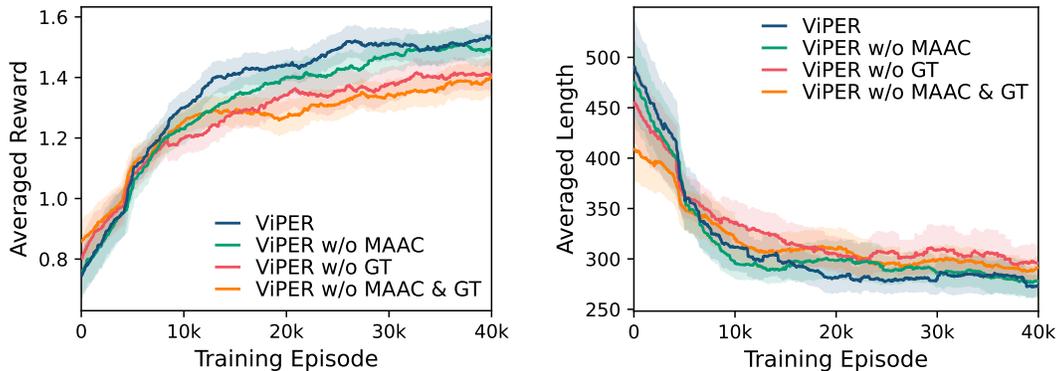


Figure 1: **Training curve of ViPER and its ablation baselines.** All curves are averaged over a window of 200 data points, with the shaded area representing ± 0.2 times the standard deviation.