

# Learning Snake Policies from Demonstration

## ABSTRACT

Many real-world tasks do not have a known or grounded reward function, preventing the utilization of traditional reinforcement learning. To learn these tasks, we can instead utilize imitation learning (IL) to learn from expert demonstrations. In our approach, we implement multiple imitation learning algorithms, Stochastic Mixing Iterative Learning (SMILe) and Generative Adversarial Imitation Learning (GAIL). We then experiment with multiple techniques to extend these algorithms to learn from scored trajectories. We compare the performances of these algorithms and techniques on their ability to learn various policies and combinations of policies in the game Snake.

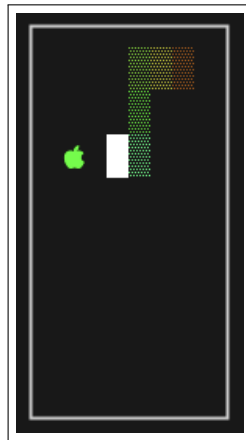


TABLE I  
SIMILARITIES OF POLICIES LEARNED BY IMITATION LEARNING ALGORITHMS

Models	Allergic	Random	Zig-Zag	Greedy	Left Greedy	Right Greedy
SMILe (10 iterations)	1.00	0.43	1.00	0.93	0.98	0.98
GAIL (50 iterations)	0.94	0.37	0.83	0.64	0.84	0.81

Table 1. An analysis of the different models on imitating specific agents. In this case, similarity is the proportion of the learned policy’s actions that match the expert’s action given the same state.