Learning Behavior for Autonomous Agents using Multi-Armed Bandit Algorithms

> David Millman Comp790-058 Project proposal

General Idea

 Investigate methods for reinforcement learning for behavioral planning of Autonomous Agents centered around Multi-Armed Bandit Algorithms from Machine Learning and Artificial Intelligence literature.

Reinforcement Learning

- Given a reward function *f*, the goal is to take actions which maximize *f*.
- Popular problem in Machine Learning and AI
- Problem of exploring
- In this domain, typically done through Q-learning (sub-optimal solution), with ε-greedy algorithm

Multi-Armed Bandit

- Reinforcement learning problem
- A gambler must decide which arm of karmed slot machine to pull resulting in some reward.
- Goal: Maximize the total reward.

Motivation

- Limiting, only precompiled behaviors available.
- Possibly easier to generate behavioral models for complex tasks
- Finding sub-optimal solutions may look better since people rarely always make the optimal choice

Previous Work

- T. Conde and D. Thaimann, Learnable Behavioural Models for Autonomous Virtual Agents: Low-Level Learning
- J. Vermorel and M. Mohri Multi-Armed Bandit Algorithms and Empirical Evaluation
- K. Morihiro et al. Emergence of Flocking Behavior Based on Reinforcement Learning
- S. LaValle, Planning Algorithms Ch. 10
- Russell and Norvig Artificial Intelligence: A Modern Approach
- D. Jurafsky and J. Martin Speech and Language Processing

Goal

- Gain a good understanding of behavioral planning.
- Apply reinforcement learning methods from AI and Machine Learning to behavioral planning.
- Create an easy to use interface for making new behaviors.

Final Demo Goal

 Create a small set of learned behavioral tasks for an agent or group of agents within 2nd life and demo these tasks!!