Name(s)	Project Number
Abhijit S. Fnu	
	31080

Project Title
A Study and Analysis of Q-Learning Algorithm Parameters for Decision Making using a Developed Simulation Tool

Abstract

Objectives/Goals

The objective of this research is to study Q-Learning Algorithm (QLA) and to develop simulation software in order to understand the optimal selection of parameters (learning rate and weight of future rewards) involved.

Hypotheses:

1) If (a) both learning rate (alpha) and weight of future rewards (gamma) are defined to be 1; and/or (b) either alpha or gamma is set to 0, there is no learning involved.

2) To find the efficient path for reaching the goal, the optimal combination of alpha and gamma is 0.5 and 1, respectively.

3) If the sum of alpha and gamma equals 1, the average computation time (time to reach the goal) is constant, regardless of environment complexity.

Methods/Materials

QLA is a subset of reinforcement learning (RL) which falls under the Markov decision process (MDP). QLA pseudo-code forms the basis of this research. Other materials include a Windows-based laptop with 4GB RAM, C++ compiler, and an environment in which to test the learning agent.

Results

A virtual environment (i.e. simulation tool) was created from scratch with C++ compiler. An AI (Artificial Intelligence) agent was tested within the environment with discrete values of alpha and gamma. The computational time involved justifying the optimal path based on combined effect of defined values of alpha and gamma.

1) The first hypothesis was proven correct: (a) with alpha and gamma both set to 1, all states became goal states & (b) with either alpha or gamma set to 0, learning took infinite amount of time.

2) The second hypothesis was proven incorrect: the optimal combination of alpha and gamma was 0.9 and 1, respectively, as computation time was quickest with these values.

3) The third hypothesis is currently under study.

Conclusions/Discussion

MDP in AI domain is an unsupervised RL method involving mathematics and reasoning, computer algorithm, and software technology and emerging as an important area of interdisciplinary research as it has potential application in such areas as unmanned exploration, evolutionary research, and feature recognition. The research confirms that Q-Learning is a powerful technique that can be applied in the above areas.

Summary Statement
The Q-Learning Algorithm parameters, learning rate and weight of future rewards, were studied and analyzed in order to understand the effect of their optimal combined values by the use of a developed simulation software.

Help Received
My parents helped prepare the presentation board.