# CSCE 475/875
# Game Day 1: Learning Day

## Assigned: September 15, 2011    Game Day: September 22, 2011

---

**Introduction**

Welcome to Soh's Agency for Multiagent Learning.  Today, we have a set of states (six of them) and a set of actions (i.e., A1, A2, and A3) for you to try out in order to learn about the Q-values of the state-action pairs.   The objective is to learn about Reinforcement Learning in a multiagent setting, insights as to how to set the learning rate and discount factor, and how to compute the Q-values and values of states.

This document is supplementary to Game Day 1: Learning Day assignment.

---

**Monopoly Money**

You have a total of $0 to begin with.  Rewards that you receive for your actions will be in Monopoly money.  Check your money when you receive your rewards.  Make sure that the amount is correct. **Assume that the maximum rewards per action performed are around $50 and the minimum are around $1.**

---

**Procedure**

Each group will be given a set of states to begin with (see your smaller envelope, in the form of a "state-action" token) and a set of "empty contracts" (also see your smaller envelope).

A "state-action" token is as follows

<div style="border:1px solid black; text-align:center;">

**Team:  _____**

**State:  _____**

**Action:  _____**

</div>

You are required to enter the value for "Action" and then submit it in to one of Game Day Monitors, and you receive new state-action tokens and your Monopoly rewards back.  The Monitors will keep the token submitted.

An "empty contract" looks like this:

> **Requesting Team:** _____
>
> **Contracted Team:** _____
>
> **State-Action Pair:** _____  _____
>
> **Rewards for State-Action Pair:** _____
>
> **Rewards received by Requesting Team:** _____

You are required to fill in "Requesting Team:" using your team name.  Then put in the name of the team that you want to have certain action performed on a certain state.  Please put these values down accordingly.  Then, one of your teammates should accompany one of the "Contracted Team" members to the Monitors to get the rewards.  Your team will not receive new states.  Your team will only receive half of the rewards.

We will have two rounds of learning.  Here is the tentative schedule (actual running times will depend on the execution on the Game Day).

<div align="center">

Round 1:  15-20 minutes
Intermission: 5-15 minutes
Round 2: 30-50 minutes
Summary:  5-10 minutes

</div>

During **Intermission**, each team will e-mail the Monitor their ordering of their state-action pairs.  And then the Monitor will share them with all teams.  During this Intermission, each team is required to take notes on their Mid-Game Strategy as part of the Game Day report.

During **Summary**, each team is required to take notes on Lessons Learned as part of the Game Day report.  (If there is time, each team will be asked to briefly discuss this in class.)

You are required to keep track of all your transactions in this worksheet. Note that ideally, after each transaction, you should compute and update the Q value of the state-action pair and the value (V) of the state involved.  *You may want to implement this program before the Game Day and use the program to help you do this accurately and efficiently during Game Day*.

Also, note that it is possible for multiple teams to have the same states at the same time.

Finally, note that when you request another team to perform an action on a particular state, communicate with them first with what you want and only fill in the "empty contract" if the "contracted team" agrees to perform the action on the particular state of your request.

In the following, record each transaction.  If an action is performed by another team, please include the team name under "Action By".

Learning Rate: _____          Discount Factor: _____

| Trans. | State | Action | Rewards | | Action By | Q(state,action) | V(state) |
|---|---|---|---|---|---|---|---|
| | | | States | $ | | | |
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | | | | | | |
| 4 | | | | | | | |
| 5 | | | | | | | |
| 6 | | | | | | | |
| 7 | | | | | | | |
| 8 | | | | | | | |
| 9 | | | | | | | |
| 10 | | | | | | | |
| 11 | | | | | | | |
| 12 | | | | | | | |
| 13 | | | | | | | |
| 14 | | | | | | | |
| 15 | | | | | | | |
| 16 | | | | | | | |
| 17 | | | | | | | |
| 18 | | | | | | | |
| 19 | | | | | | | |
| 20 | | | | | | | |
| 21 | | | | | | | |
| 22 | | | | | | | |
| 23 | | | | | | | |
| 24 | | | | | | | |
| 25 | | | | | | | |
| 26 | | | | | | | |
| 27 | | | | | | | |
| 28 | | | | | | | |
| 29 | | | | | | | |
| 30 | | | | | | | |
| 31 | | | | | | | |
| 32 | | | | | | | |
| 33 | | | | | | | |
| 34 | | | | | | | |
| 35 | | | | | | | |
| 36 | | | | | | | |

| Trans. | State | Action | Rewards | | Action By | Q(state,action) | V(state) |
|--------|-------|--------|---------|-----|-----------|-----------------|----------|
| | | | States | $ | | | |
| 37 | | | | | | | |
| 38 | | | | | | | |
| 39 | | | | | | | |
| 40 | | | | | | | |
| 41 | | | | | | | |
| 42 | | | | | | | |
| 43 | | | | | | | |
| 44 | | | | | | | |
| 45 | | | | | | | |
| 46 | | | | | | | |
| 47 | | | | | | | |
| 48 | | | | | | | |
| 49 | | | | | | | |
| 50 | | | | | | | |

**Intermission**

**Notes:**

In the following, record each transaction.  If an action is performed by another team, please include the team name under "Action By".

Learning Rate: _____     Discount Factor: _____

| Trans. | State | Action | Rewards | | Action By | Q(state,action) | V(state) |
|---|---|---|---|---|---|---|---|
| | | | **States** | **$** | | | |
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | | | | | | |
| 4 | | | | | | | |
| 5 | | | | | | | |
| 6 | | | | | | | |
| 7 | | | | | | | |
| 8 | | | | | | | |
| 9 | | | | | | | |
| 10 | | | | | | | |
| 11 | | | | | | | |
| 12 | | | | | | | |
| 13 | | | | | | | |
| 14 | | | | | | | |
| 15 | | | | | | | |
| 16 | | | | | | | |
| 17 | | | | | | | |
| 18 | | | | | | | |
| 19 | | | | | | | |
| 20 | | | | | | | |
| 21 | | | | | | | |
| 22 | | | | | | | |
| 23 | | | | | | | |
| 24 | | | | | | | |
| 25 | | | | | | | |
| 26 | | | | | | | |
| 27 | | | | | | | |
| 28 | | | | | | | |
| 29 | | | | | | | |
| 30 | | | | | | | |
| 31 | | | | | | | |
| 32 | | | | | | | |
| 33 | | | | | | | |
| 34 | | | | | | | |
| 35 | | | | | | | |
| 36 | | | | | | | |
| **Trans.** | **State** | **Action** | **Rewards** | | **Action By** | **Q(state,action)** | **V(state)** |

| | | | States | $ | | | |
|---|---|---|---|---|---|---|---|
| 37 | | | | | | | |
| 38 | | | | | | | |
| 39 | | | | | | | |
| 40 | | | | | | | |
| 41 | | | | | | | |
| 42 | | | | | | | |
| 43 | | | | | | | |
| 44 | | | | | | | |
| 45 | | | | | | | |
| 46 | | | | | | | |
| 47 | | | | | | | |
| 48 | | | | | | | |
| 49 | | | | | | | |
| 50 | | | | | | | |
| 51 | | | | | | | |
| 52 | | | | | | | |
| 53 | | | | | | | |
| 54 | | | | | | | |
| 55 | | | | | | | |
| 56 | | | | | | | |
| 57 | | | | | | | |
| 58 | | | | | | | |
| 59 | | | | | | | |
| 60 | | | | | | | |
| 61 | | | | | | | |
| 62 | | | | | | | |
| 63 | | | | | | | |
| 64 | | | | | | | |
| 65 | | | | | | | |
| 66 | | | | | | | |
| 67 | | | | | | | |
| 68 | | | | | | | |
| 69 | | | | | | | |
| 70 | | | | | | | |
| 71 | | | | | | | |
| 72 | | | | | | | |
| 73 | | | | | | | |
| 74 | | | | | | | |
| 75 | | | | | | | |

**FINAL AMMOUNT OF REWARDS ($):** _____

**Lessons Learned**

**Notes:**