# Reinforcement Learning in Hawk-Dove Game with Mutation

Geoffrey Wang

August 2024

## 1 Introduction

The Hawk-Dove game is a classic model in evolutionary game theory that describes the interactions between two strategies: Hawk and Dove. The Hawk strategy involves aggressive behavior, whereas the Dove strategy involves peaceful behavior. A mixed strategy can also be considered, where individuals probabilistically choose between Hawk and Dove behaviors. Reinforcement Learning (RL) can be used to model how strategies evolve over time based on their payoffs. This document describes the implementation of an RL-based Hawk-Dove game with the inclusion of mutation, which allows strategies to change randomly over generations.

## 2 Implementation of the Hawk-Dove Game with RL

### 2.1 Strategy Definition

We define three strategies: Hawk, Dove, and Mixed, each represented by an integer constant:

- Hawk ($H$) is represented by 0.
- Dove ($D$) is represented by 1.
- Mixed ($M$) is represented by 2.

### 2.2 Q-Learning Parameters

The following parameters are used in the Q-learning algorithm:

- Learning rate ($\alpha = 0.1$)
- Discount factor ($\gamma = 0.95$)

- Exploration rate ($\epsilon = 0.1$)
- Mutation rate ($\mu = 0.01$)

### 2.3 Individual Class

Each individual in the population is represented by an object of the `Individual` class. This class maintains a Q-table for each strategy and provides methods to choose actions, update Q-values, and mutate strategies.

**Algorithm 1** Individual Class Implementation

```
Initialize strategy and Q-table choose_action
if strategy is MIXED then
    Choose between HAWK and DOVE based on Q-values
else
    Choose action based on ϵ-greedy policy
end ifupdate_q_valueaction, reward, next_max_q
Update Q-value using the Q-learning formula mutate
Randomly mutate strategy with probability μ playopponent, V, C
Determine payoffs based on strategies and return them
```

### 2.4 Population Class

The population is represented by an object of the `Population` class, which consists of multiple `Individual` objects. The population evolves over generations based on the Q-learning algorithm and mutation.

**Algorithm 2** Population Evolution

```
Initialize population with a mix of HAWK, DOVE, and MIXED strategies
for each generation do
    Calculate fractions of HAWK, DOVE, and MIXED strategies
    for each individual do
        Choose an opponent randomly
        Play the game and update Q-values
        Apply mutation to the individual's strategy
    end for
    Normalize the fractions to ensure they sum to 1
end for
```

## 3 Markov Decision Process with Mutation

The Markov Decision Process (MDP) for this problem involves states corresponding to the strategies (Hawk, Dove, Mixed). The actions are the potential strategies that an individual can choose. The rewards are determined based on

the interactions between individuals, and the transitions are governed by the Q-learning algorithm and the mutation mechanism.

### 3.1 State Transition Diagram

The state transition diagram for the Hawk-Dove game with mutation is depicted in Figure 1. The transition probabilities are influenced by the Q-values and the mutation rate. Mutation introduces randomness into the process, allowing strategies to transition even if they are not optimal according to the Q-values.

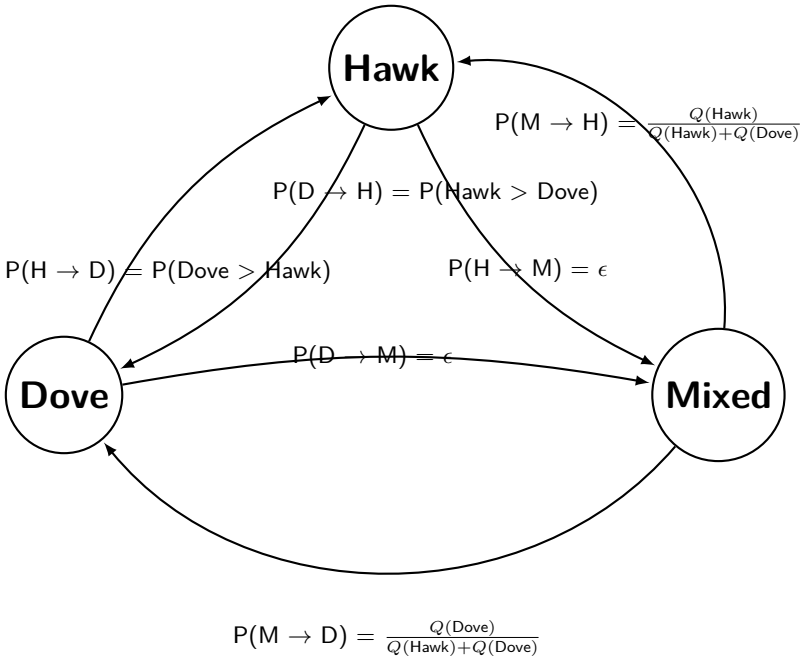


Figure 1: Strategy State Transition Diagram with Conditional Probabilities