1 Overview

1.1 Environment

- The game is a multi-agent environment.
- The game is simple market where you have to decide on a price and quantity.
- The goal of the game is earn as much profit as possible across 80 periods.

1.2 State space

- The state space has 5 numbers representing the weighted average selling price the last 5 periods
- Note: you may want to create your own state space if you think other information will help your agent make better decisions

1.3 Action space

The action space has 2 numbers in the following order:

- A price (between 0 and 17)
- A quantity (between 0 and 300)

1.4 Rewards

- price * number sold - cost(number sold)

1.5 Terminal States

- The episode ends after 80 periods

2 Environment Details

- There is a demand function: p = 17 - 0.01q.
- Each firm has the following cost structure: 1st 100 units cost 4 each, 2nd 100 cost 4.5 each, and last 100 cost 5 each.
- All prices are rounded down to the nearest cent (5.3235 -> 5.32)
- Firms are sorted from lowest to highest price.
- For each price, we calculate quantity demanded at the price: q = 1700 - 100p.
- We then subtract any quantity that has already been sold to other players.
- This demand is then allocated equally across all firms with the current lowest price, with firms receiving this price for each unit sold, until all demand is satisfied or firms run out of capacity/quantity.
- These firms are removed and the process continues until all demand is satisfied or all units are sold.