MDP

MDP models a strategic game involving two competing agents or kingdoms, each making decisions that affect their status and resources over time. The primary components of this MDP—states, actions, transitions, and rewards—create a framework that simulates the dynamic interactions between these kingdoms under various conditions. Here's a basic breakdown of what's happening in this MDP:

States

The states in the MDP represent the condition of a kingdom at any given point in time, captured as a vector of discrete variables. These variables include the kingdom's army strength, money, people mood, number of factories and mines, resource levels, the enemy army's strength, whether the enemy has checked the kingdom's military capabilities recently and war status.

State is vector:

$$[\text{army}, \text{money}, \text{people}, \text{factory}, \text{resource}, \text{mine}, \text{enemy army}, \text{checked}, \text{war}]^T$$

each number from 0 to 5(except war {0,1})

Actions
The actions represent decisions that the kingdom can make, each action is possible untill all cost entities is less or equal to them repectively in state vector and when Gain entity in state vector is not maximum already.

$$A \in \{\text{Sell resources}, \text{Buy resources}, \text{Build factory}, \text{Sell factory}, \text{Make army}, \text{Remove army}, \text{Build mine}, \text{Attack}, \text{Do nothing}, \text{Spy}\}$$

ACTION	COST	GAIN
SELL RESOURCES	1 resource	1 money
BUY RESOURCE	1 money	1 resource
BUILD FACTORY	1 money 1 resource	1 factory
SELL FACTORY	1 factory	1 money
MAKE ARMY	1 people 1 resource 1 money	1 army
REMOVE ARMY	1 army	1 people

BUILD MINE	1 resource 1 money	1 mine

Game Structure

The game is structured into rounds where each agent can perform three actions per round. At the end of each round, the state is updated based on the following rules:

- **Resource Consumption**: Automatically decrement 1 resource every round.
- **Army Resource Cost**:
  - Army size 5: consume 3 additional resources.
  - Army size 4: consume 2 additional resources.
  - Army size 3: consume 1 additional resource.
- **Resource Effects on Population**:
  - Resource level 0: decrease population by 2.
  - Resource level 1: decrease population by 1.
  - Resource level 4: increase population by 1.
  - Resource level 5: increase population by 2.
- **Population Effects on Infrastructure**:
  - People mood level 0: 50% chance to lose 1 factory and 50% chance to lose 1 mine.
  - People mood level 1: 25% chance to lose 1 factory and 25% chance to lose 1 mine.

Interactions Between Agents

- **Spy Action**: Allows an agent to know the army size of the opponent. After spying, the agent's state is updated to indicate that the enemy has checked their army, starting at 5 and decreasing each round until it reaches 0.
- **Attack Action**: Initiates war. Once war is declared, the war state becomes active.

War Dynamics

- During war, both agents update each other's enemy army sizes each round.
- At the end of each round, the probability of winning the battle is calculated using a softmax function based on the army sizes.
- The loser of the battle has all state entities decreased by 1.

Terminal States and Rewards

- **Victory**: Occurs when an agent reduces the enemy's army to 0 while war is active. The reward is 33 plus the dot product of e (a constant vector) and the agent's state vector.
- **Defeat**: Occurs when an agent's army is reduced to 0 while war is active. The reward is -33.