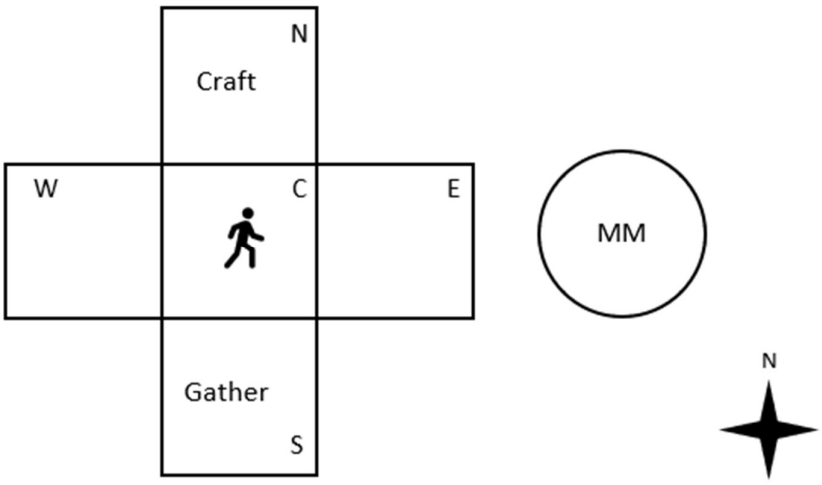
PART 2 : Value Iteration Algorithm

TASK 1:

Indiana Jones (IJ) is on an adventure to stop the **Mighty Monster (MM)** before it is too late. He has traveled for days and now he has found himself face to face with MM. He is in a room shaped like a plus **(+)** with only **5 positions** for him to stand. The squares are **Center (C), North (N), South (S), East (E), and West (W)**. MM is on the **East** side of the room. Indiana has infinite health (obviously), however, MM has only **max health of 100**. The episode will end once MM has **0 health** and Indiana will receive a **reward of 50.**

Indiana has a bow and a blade to attack the MM. Indiana at any moment can only carry **3 arrows (at max)**. Indiana can at max carry only **2 materials**.



Indiana's actions are dependent on his position in the room. At each timestep, he can take only **1 action**.

Center Square:
IJ can move: **{up (move to North Square), down, left, right, stay}** and action will be successful only **85%** of the time; the other times, he will move to the **East** Square. IJ can **shoot**, which will cause him to **lose 1 arrow.** Because this is a dark room, Indiana's accuracy is worse than before. The arrow will hit MM with a probability of **0.5** and miss otherwise. If it hits MM, it will deal **25 damage**. IJ can also **hit with his blade**, which will make contact with MM with a probability of **0.1** and miss MM with a probability of **0.9** If it hits MM, it will deal **50 damage**.

North Square:
IJ can move: **{down (move to Center Square), stay}** and the action will be successful only **85%** of the time; the other times, he will teleport to the **East** Square. IJ can **craft arrows** here if he has **1 Material**. Upon crafting, IJ will lose his Material and gain **1 arrow** with a probability of **0.5**, **2 arrows** with a probability of **0.35**, and **3 arrows** with a probability of **0.15**.

South Square:
IJ can move: **{up (move to Center Square), stay}** and the action will be successful only **85%** of the time; the other times, he will teleport to the **East** Square. IJ can **gather material**. He will **gather successfully** with a probability of **0.75** and fail with a **probability of 0.25**. Upon gathering **successfully**, IJ will gain **1 Material**, if he does not have material, else there will be no change.

East Square:
IJ can move: **{left (move to the Center Square), stay}**, and the action will be successful **100%** of the time. Indiana can **shoot arrows** here. The arrow will hit MM with a probability of **0.9** and miss MM with a probability of **0.1**. The arrow will deal **25 damage** to MM. Indiana can also **hit** with his blade. He will make contact with MM with a probability of **0.2** and miss with a probability of **0.8**. The blade will deal **50 damage** to MM.

West Square:
IJ can move: **{right (move to the Center Square), stay}**, and the action will be successful **100%** of the time. Indiana can **shoot** arrows here. The arrow will hit MM with a probability of **0.25** and miss MM with a probability of **0.75**. Upon hitting MM, the arrow will deal **25 damage** to MM.

At intervals, MM can also attack IJ. MM will indicate that he plans to attack soon by entering **Ready State**. MM enters the **Ready State** from **Dormant State** with a probability of **0.2** at each step or remains in the **Dormant State** with a probability of **0.8**. When MM is in the **Ready State**, he may **attack** in the timestep with a probability of **0.5** and enter the **Dormant State** or not attack with a probability of **0.5** and remain in the **Ready State**. When MM attacks he will only affect Indiana if IJ is either on the **East Square or Center Square** at that time step. In this case, IJ will **drop all his arrows** (if he has any), and MM will **regain 25 health**. If Indiana had planned to take any action at this step, it would be **unsuccessful**. Indiana will also get a **negative reward** of **-40.** *The probabilities for the transitions will scale appropriately with this new condition added.*

Since Indiana needs to stop MM asap, there is a penalty for each timestep. The penalty is given by:

```
Step Cost = -10/Y

arr = [1/2, 1, 2 ]
Y = arr[X mod 3]
X: team number
```

Parameters:

- `Gamma = (Discount Factor) 0.999`
- `Delta = (Convergence or Bellman error)` 10^{-3}

Output the trace of the state, the chosen action at that state and the value of state for all iterations from 0 till convergence. (output format on next page)

In the report:

- **Interpret and comment on the results in the trace file. Analyze the choice of actions for the states, rate of convergence, and any other meaningful pattern.**
- **Assume the start states as (given below) and simulate the game to find the set of actions till the end state (choose actions based on the policy). Print the order of the actions, the state transitions and comment on the results. No specific submission format.**
    1. **(W, 0, 0, D, 100)**
    2. **(C, 2, 0, R, 100)**

TASK 2:

Case 1:

- Indiana now on the LEFT action at East Square will go to the West Square.

Case 2:

- The step cost of the STAY action is now zero.

Case 3:

- Change the value of gamma to 0.25

Output the trace for each of the cases and comment on the results in the report. Justify the change in policy for each case. Note that for each case, every other parameter except the ones mentioned remain unchanged from task 1.