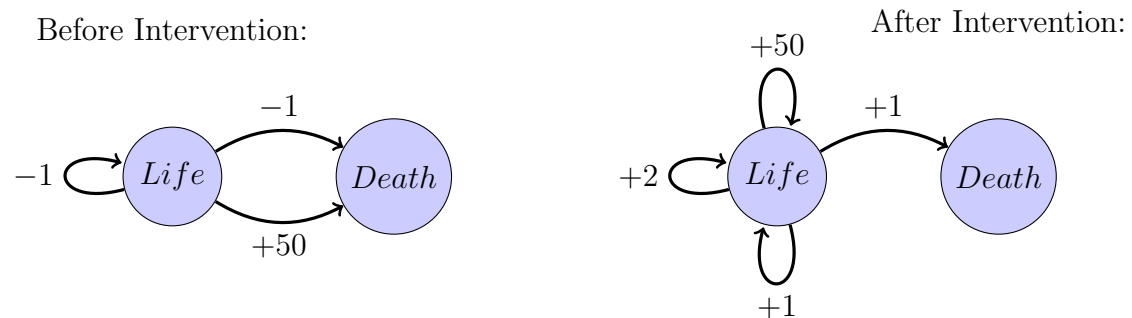


In our context, Firstly the Rainbow DQN can only distinguish between states which have different board states, and thus doesn't know about time, second we construct the environmental health graphs to be. Because the island navigation environment we are investigating only has two states, we labelled them life and death for convenience.



Now we apply the intervention, by giving the agent 2 additional points every second it is alive, and by having the goal being hit changed to the goal being discovered in that episode.

Before Intervention:

$$R_a(s, s') = \begin{cases} +50 & \text{if } s' \text{ is the goal} \\ -1 & \text{always while alive} \end{cases}$$

After Intervention:

$$R_a(s, s') = \begin{cases} +50 & \text{if } s' \text{ is the goal (only the first time in the episode)} \\ -1 & \text{if } s' \text{ is not goal} \\ +2 & \text{always while alive} \end{cases}$$