* Compare the problem with different sizes, because it’s hard to compare two different problems, like easy setup, medium set up and hard setup, 2-3 is probably enough
* convergence
  + You expect your q value to increase, when they stop increasing, that’s when you determine convergence
  + Policy iteration looks at the point where policy not changing,
  + Make the rule deterministic, set a rule
* Q learning will work if it was possible to traverse all the state infinitely often, so if it’s not working, you can analysis if there is something not allow the algorithms to travel to some of the state
* Scholastic dynamic is recommended, as deterministic is easy to converge
* Show some tunning hyperparameters, and analysis of its performance, it’s to show that we understand the hyper parameters is the goal
* Frozen lake problem
* **Visualize the optimal problem, and original view of the problem, like the obstacle, goals and etc**
* **Daniel Boros**
  + I've seen forest used before. Small may be... too small and uninteresting. May be worth a shot at least to set up your experiments (plots, et cetera), then you can switch that problem out.
  + There are a few discrete examples in OpenAI Gym that you can input into \*mdptoolbox.
  + **Graphical user interface, text, application

    Description automatically generated**
* **Reward: 在a点的时候take action I， 得到的reward**
* **In the beginning, there are absorbing state, “fronzen”, so Q learning does not work because it was not able to explore all the state infinitely possible**
* **The ending reward was small, this caused value iteration not optimal**
* **Frozen lake，四个实验**
  + **Deterministic**
    - **小**
    - **大**
  + **Stochastic**
    - **小**
    - **大**