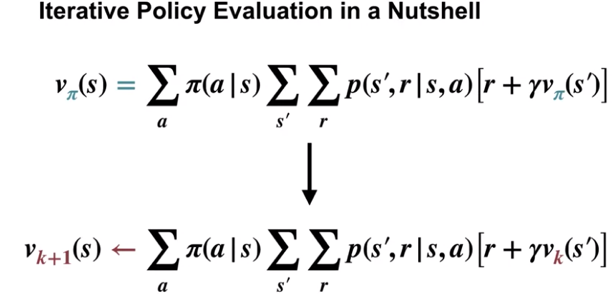
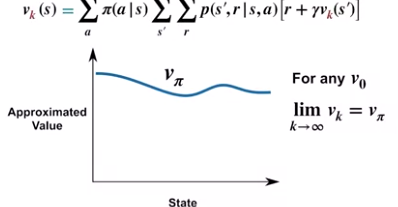
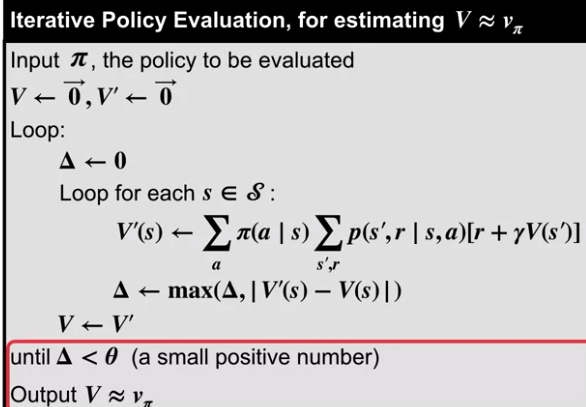
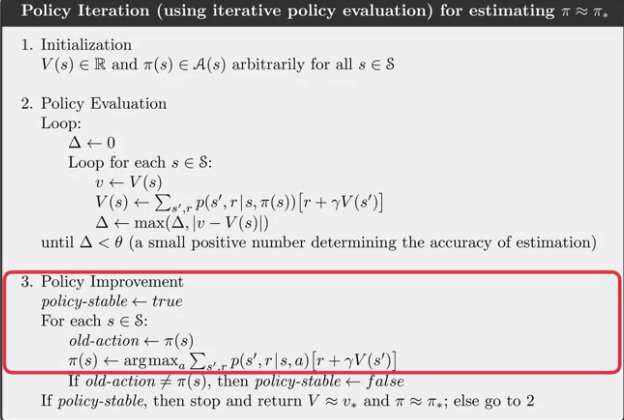
**Lesson 1: Policy Evaluation (Prediction)**

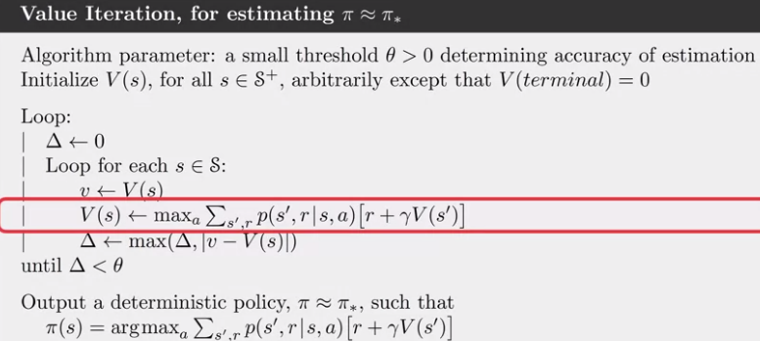
* Understand the distinction between policy evaluation and control
  + Policy evaluation is the task of determining the value function for a specific policy.
  + Control is the task of finding a policy to obtain as much reward as possible
* Explain the setting in which dynamic programming can be applied, as well as its limitations
  + policy evaluation is the task of determining the state value function v pi for policy pi. Control is the task of improving an existing policy. And dynamic programming techniques can be used to solve both of these tasks if we have access to the dynamics function p.
* Outline the iterative policy evaluation algorithm for estimating state values under a given policy
  + 
  + 
* Apply iterative policy evaluation to compute value functions
* 

**Lesson 2: Policy Iteration (Control)**

* Understand the policy improvement theorem
  + the policy improvement theorem tells us that greedified pi policy is a strict improvement, unless the original policy was already optimal. You should also now know how to use the value function under a given policy to produce a strictly better policy.
* Use a value function for a policy to produce a better policy for a given MDP
* Outline the policy iteration algorithm for finding the optimal policy
  + Evaluation and Improvement
  + 
* Understand “the dance of policy and value”
* Apply policy iteration to compute optimal policies and optimal value functions

**Lesson 3: Generalized Policy Iteration**

* Understand the framework of generalized policy iteration
* Outline value iteration, an important example of generalized policy iteration



* Understand the distinction between synchronous and asynchronous dynamic programming methods
  + value iteration allows us to combine policy evaluation and improvement in a single-step, asynchronous dynamic programming methods give us freedom to update states in any order, and the idea of generalized policy iteration unifies many algorithms
* Describe brute force search as an alternative method for searching for an optimal policy
* Describe Monte Carlo as an alternative method for learning a value function
* Understand the advantage of Dynamic programming and “bootstrapping” over these alternative strategies for finding the optimal policy