**动态规划的解释：**

Dynamic：问题可以转化为一类序列问题

Program：数学优化问题

将问题转化为子问题，子问题解决了，原问题也解决了，类似于递归，

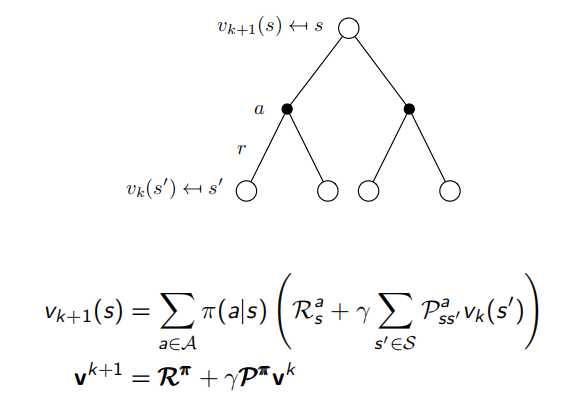
动态规划面向的问题有两个特点：

1. 最优化子结构：子结构获得最优解，那么整体的结构就获得了最优解。
2. 子问题的重叠性：子问题会多次出现，那么将各个子问题的解综合起来就是整个问题的解，这样就有了分而治之思想。

MDP的适用性：

1. Bellman等式递归分解
2. 价值函数可以被重用

## 3.1 Policy Evaluate



Value function的迭代更行，用子节点的缓存值计算父节点的值

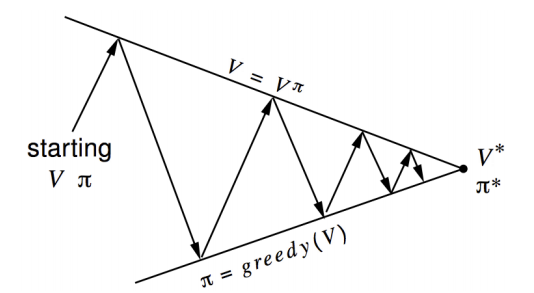
## Evaluating a Random Policy in the Small Grid-world

Grid游戏：

Bellman迭代

## 3.2 policy Iteration

How to improve policy：先评估（填数字），再优化（max的方向）



迭代的进行评估和更新（贪心）

David：无论从哪里开始，采取怎样的policy都会收敛到最有的policy

## 3.3 Modified Policy Iteration

**Principle of Optimality：**

## 3.4 Value Iteration

只用value 没有policy

实验部分：