

# 彪哥带你学强化学习

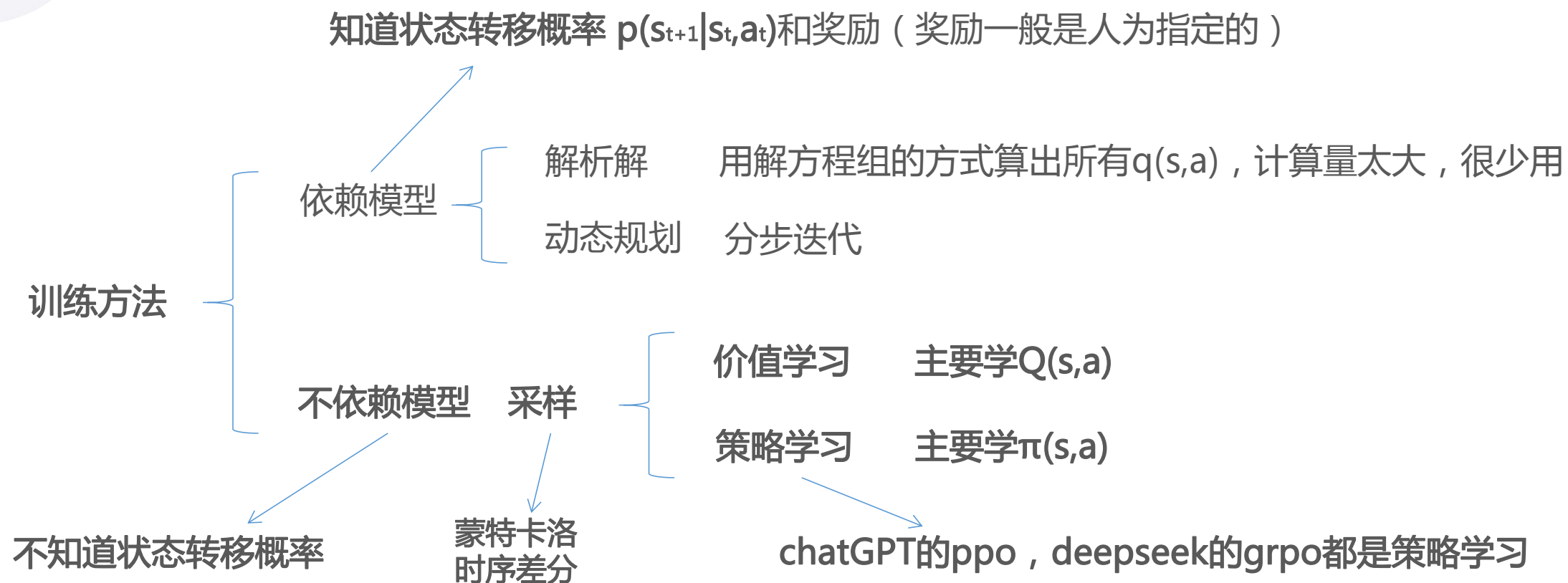
## 4、强化学习训练方法

DEEPLY UNDERSTAND REINFORCEMENT LEARNING

讲师：韩路彪



## 强化学习训练方法



蒙特卡洛：用随机采样的均值逼近实际概率 强化学习里边一般指按回合采样 误差小，方差大

时序差分：temporal difference，时间维度的差异，比如  $Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)$  误差大，方差小