# RL Basics

**Behavior Stuctures**:

- plan : fixed sequence of actions
  - fragile in stochastic environments
- conditional plan : composition of plans with branching points
- dynamic planning : "just-in-time" planning (if plan fails, make a new one)
- stationary policy / universal plan : mapping from state to action
  - very large
  - MDPs always have stationary policy

**Evalating a Learner**:

- value of returned policy
- if two agents have the same policy, prefer:
  - less computational complexity (time)
  - sample complexity : how much data it needs (time?)
  - why not space complexity? we don't frequently run into space complexity issues before time complexity concerns