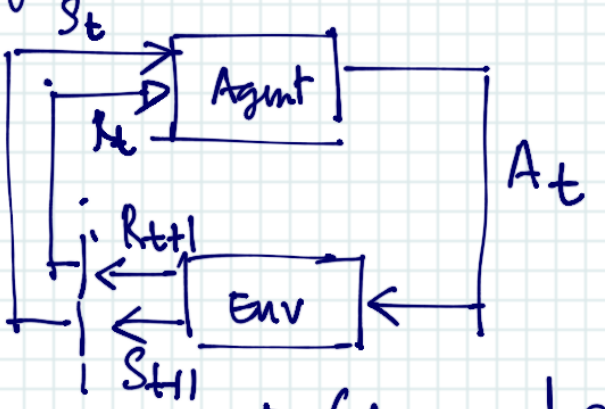
4/8/19

AI5001: Intro to Modern AI

Today:
- Review
- Continue RL problem formulation:
  - state, actions
  - goals, rewards
  - characterising state

Review:



Policy: $\pi_t(a|s) = Pr(A_t = a \mid S_t = s)$

Goal of RL is to learn the optimal policy from experience

$S_t \in \mathcal{S}$ where $\mathcal{S}$ is a set of possible states

$A_t \in \mathcal{A}(S_t)$ where $\mathcal{A}(S_t)$ is the set of possible actions corresponding to S_t

Goals and rewards:

- $G_t = R_{t+1} + R_{t+2} + \cdots R_T$ (episodic)
- $G_t = \sum_{k=0}^{\infty} R_{t+k+1}$ (continuous).

Note that the above G_t could be unbounded even if $\{R_k\}$ is finite

- Introduce γ, a discount factor that lies between 0 and 1 and define

$$G_t = \sum_{k=0}^{\infty} \gamma^k \cdot R_{t+k+1}$$

Unified goal definition:

$$\boxed{G_t = \sum_{k=0}^{T} \gamma^k R_{t+k+1}}$$

where $T = \infty$ in the cont. case,
$\gamma = 1$ in the episodic case

Imp: Both these cannot be true simultaneously

Prob. review!

Q: What is a RV?

A: A RV $X(\Omega) : \Omega \to \mathbb{R}$

Q: What is CDF?

A: $F_X(x) = Pr\{X \leq x\}$

Q: What is PDF?

A: $f_X(x) = \frac{d}{dx} F_X(x)$ (Continuous case)

PMF $f_X(x) = Pr\{X = x\}$ (discrete case).

Q: Random process?

A: An indexed set $\{X(t) : t \in \mathcal{T}\}$ of random variables. $\mathcal{T}$ could be countably infinite

Markov process: $f_{X_{n+1}|X_n \underline{X}_{n-1}}(x_{n+1} \mid x_n, \underline{x}_{n-1})$

$$= f_{X_{n+1}|X_n}(x_{n+1} \mid x_n)$$

.