

Procesos Markovianos de Decisión

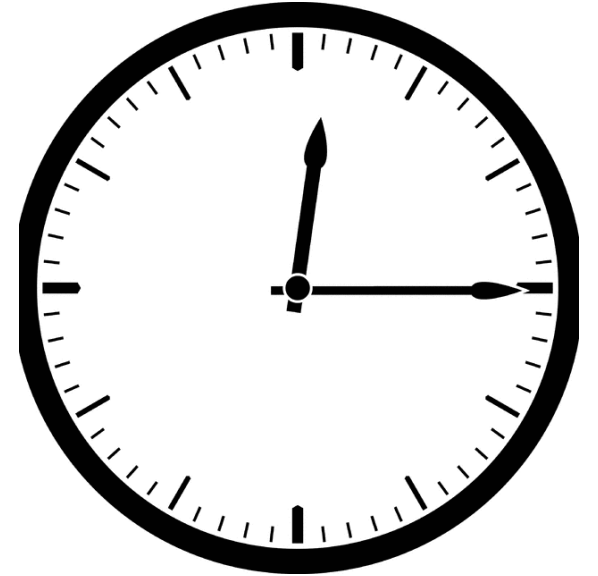
ITAM

Primavera 2017

Clase 2

I. Procesos estocásticos y Cadenas de Markov

- Procesos estocásticos



Ejemplos de procesos y conjuntos de estados....

Cadenas de Markov *a tiempo discreto y conjunto de estados discreto....*

- Cadenas de Markov
- ¿Qué es una cadena?
- **LA PROPIEDAD MARKOVIANA:** El estado futuro de las cosas depende solo del estado actual



- Estados y probabilidades de transición

- Cadenas homogéneas

- Diagramas de transición

II. Aplicándolo a reinforcement learning...

¿Qué ingredientes necesitamos?

- Agente
 - El que aprende y toma decisiones
- Entorno/Ambiente/Naturaleza
 - Con lo que interactúa el agente
 - Determina la señal de reforzamiento/premio
 - Determina el nuevo estado del universo
 - Es la naturaleza en game theory

Interacción agente-ambiente

- Interactúan en una serie de turnos $t=1,2,3,\dots$
- En cada turno el agente escoge una **ACCIÓN** y el entorno/ambiente responde con un **NUEVO ESTADO** y un **PAGO/REFORZAMIENTO**

Políticas

- Policy is not politics
 - Policy es un “mapeo” de un estado a una acción
 - (pensar en el e-greedy)
 - ¿Decisiones aleatorias?
-
- $\pi(s|a)$

Diagrama de transición y decisión

La propiedad de Markov al decidir

- Ejemplo: Domino, Póker, Ajedrez
- Cuáles son los estados?
- Con que elementos tomamos decisiones?
- **PROPIEDAD MARKOVIANA DE DECISIÓN:** El estado futuro de las cosas y el que reciba depende solo del estado actual de las cosas la decisión que tome.

Resumen: Elementos de un proceso Markoviano de decisión