



# Tecnológico de Monterrey

## Mini examen 2

Sistemas Inteligentes

Martin Noboa

*Invierno 2023*

enero 27, 2023

## Table of Contents

|  |   |
|--|---|
| Dungeon Escape.....  | 3 |
| Objetivo del agente. ....  | 3 |
| Satisfacción de restricciones.....   | 3 |
| Tabla PEAS. ....   | 3 |
| Características del ambiente .....   | 3 |
| Lista de percepciones a usar.....  | 3 |
| Lista de acciones a realizar.....  | 3 |
| Como se realiza el reforzamiento?.....                                     | 4 |
| Ejemplo de estado inicial y estado meta.....                               | 4 |
| Que sugerirías agregar para mejorar el comportamiento de este agente?..... | 4 |
| Food collector .....   | 5 |
| Objetivo del agente. ....  | 5 |
| Satisfacción de restricciones.....   | 5 |
| Tabla PEAS. ....   | 5 |
| Características del ambiente .....   | 5 |
| Lista de percepciones a usar.....  | 5 |
| Lista de acciones a realizar.....  | 5 |
| Como se realiza el reforzamiento?.....                                     | 5 |
| Ejemplo de estado inicial y estado meta.....                               | 6 |
| Que sugerirías agregar para mejorar el comportamiento de este agente?..... | 6 |
| Match 3 .....  | 7 |
| Objetivo del agente. ....  | 7 |
| Satisfacción de restricciones.....   | 7 |
| Tabla PEAS. ....   | 7 |
| Características del ambiente .....   | 7 |
| Lista de percepciones a usar.....  | 7 |
| Lista de acciones a realizar.....  | 7 |
| Como se realiza el reforzamiento?.....                                     | 8 |
| Ejemplo de estado inicial y estado meta.....                               | 8 |
| Que sugerirías agregar para mejorar el comportamiento de este agente?..... | 8 |
| Push Block.....  | 9 |
| Objetivo del agente. ....  | 9 |

|  |    |
|--|----|
| Satisfacción de restricciones.....   | 9  |
| Tabla PEAS. ....   | 9  |
| Características del ambiente .....   | 9  |
| Lista de percepciones a usar.....  | 9  |
| Lista de acciones a realizar.....  | 9  |
| Como se realiza el reforzamiento?.....                                     | 9  |
| Ejemplo de estado inicial y estado meta.....                               | 10 |
| Que sugerirías agregar para mejorar el comportamiento de este agente?..... | 10 |
| Snake.....   | 11 |
| Objetivo del agente. ....  | 11 |
| Satisfacción de restricciones.....   | 11 |
| Tabla PEAS. ....   | 11 |
| Características del ambiente .....   | 11 |
| Lista de percepciones a usar.....  | 11 |
| Lista de acciones a realizar.....  | 11 |
| Como se realiza el reforzamiento?.....                                     | 11 |
| Ejemplo de estado inicial y estado meta.....                               | 12 |
| Que sugerirías agregar para mejorar el comportamiento de este agente?..... | 12 |

## Dungeon Escape

Objetivo del agente.

Escapar del laberinto

Satisfacción de restricciones.

- El jugador no puede pasar a través de paredes solidas.
- El jugador solo puede moverse en 4 direcciones (adelante, atrás, derecha e izquierda).
- El jugador no puede ver a través de las paredes.

Tabla PEAS.

| Type            | Performance Measure   | Environment   | Actuator  | Sensors  |
|-----------------|---|---|---|--|
| Dungeon escapee | <ul style="list-style-type: none"><li>- Tiempo para escapar</li><li>- Caminos sin salida o retornos</li></ul> | <ul style="list-style-type: none"><li>- Laberinto</li></ul> | <ul style="list-style-type: none"><li>- Piernas</li></ul> | <ul style="list-style-type: none"><li>- Ojos</li><li>- Manos</li></ul> |

Características del ambiente

- Parcialmente observable. El agente solo puede observar la sección del calabozo en la que se encuentra.
- Discreto. Hay un numero finito de estados.
- Estático. El calabozo no cambia mientras el agente toma decisiones.
- Determinista. El ambiente no cambia, por ende, el estado actual determina el siguiente estado.
- Secuencial. Dado que las acciones posibles son moverse dentro del calabozo, la ubicación del agente determina su siguiente acción.
- Simple. Solo un agente interactúa con el ambiente.

Lista de percepciones a usar

- Que se encuentra adelante, atrás, a la derecha y a la izquierda.

Lista de acciones a realizar

- Observar
- Moverse en una dirección.

Como se realiza el reforzamiento?

- **Reforzamiento positivo.** Se calcula que tan lejos lleve el agente. Dado que el agente no sabe donde se encuentra el final, puede calcular que agente tome la mayor cantidad de acciones antes de llegar a un punto muerto.

Ejemplo de estado inicial y estado meta

- **Estado inicial.** El agente se encuentra en el calabozo, sin conocer a que dirección dirigirse.
- **Estado meta.** El agente se encuentra en la salida del calabozo.

Que sugerirías agregar para mejorar el comportamiento de este agente?

Una manera de poder mejorar este agente es pre entrenándolo. Si bien cada “mapa” de calabozo puede llegar a ser diferente, existen estrategias para resolverlos o pueden llegar a existir. Al tener un modelo pre entrenado, se puede llegar a la solución mucho mas fácilmente.

## Food collector

Objetivo del agente.

Recolectar comida que cae de la parte superior de la pantalla.

Satisfacción de restricciones.

- Solo puede recolectar frutas.
- No se debe recolectar basura.
- Solo se puede mover de derecha a izquierda.

Tabla PEAS.

| Type           | Performance Measure   | Environment  | Actuator   | Sensors  |
|----------------|---|--|--|--|
| Food collector | <ul style="list-style-type: none"><li>- Cuanta comida tiene al final</li><li>- Tiempo que aguanta recolectando comida</li></ul> | <ul style="list-style-type: none"><li>- Mapa donde la comida va cayendo a un ritmo</li></ul> | <ul style="list-style-type: none"><li>- Pies</li><li>- Manos</li></ul> | <ul style="list-style-type: none"><li>- Ojos</li></ul> |

Características del ambiente

- Completamente observable. Dado a que el agente puede ver en todo momento el ambiente.
- Discreto.
- Dinámico. La comida sigue cayendo mientras el agente toma una decision.
- Determinista.
- Secuencial
- Simple.

Lista de percepciones a usar

- Comida que va cayendo.
- Posición dentro del ambiente.

Lista de acciones a realizar

- Moverse a la izquierda.
- Moverse a la derecha.

Como se realiza el reforzamiento?

Reforzamiento positivo. Mientras mayor sea el puntaje del agente, mejor el comportamiento del mismo.

Ejemplo de estado inicial y estado meta

- **Estado inicial.** En medio del ambiente, con un puntaje de 0.
- **Estado meta.** En una posición del ambiente, con el mayor puntaje posible.

Que sugerirías agregar para mejorar el comportamiento de este agente?

- Asignando diferentes valores a cada tipo de fruta. De esta manera, el agente no solo decidirá que dirección ir, sino que fruta le conviene recolectar.

### Match 3

Objetivo del agente.

Poner 3 simbolos iguales seguidos.

Satisfacción de restricciones.

- Se turnan 2 agentes.
- Solo se puede colocar un símbolo por turno.
- Solo hay 9 posiciones para un símbolo.
- No se puede colocar un símbolo sobre una casilla ocupada.

Tabla PEAS.

| Type           | Performance Measure  | Environment   | Actuator  | Sensors   |
|----------------|--|---|---|---|
| Match 3 player | <ul style="list-style-type: none"><li>- Turnos para completar el juego</li><li>- Juegos ganados</li><li>- Juegos empatados</li><li>- Juegos perdidos</li></ul> | <ul style="list-style-type: none"><li>- Grid de match 3</li></ul> | <ul style="list-style-type: none"><li>- Manos</li><li>-</li></ul> | <ul style="list-style-type: none"><li>- Ojos</li><li>- Tablero de match 3</li></ul> |

Características del ambiente

- Completamente observable.
- Discreto vs continuo
- Estático.
- Determinista
- Secuencial
- Múltiple. 2 agentes pueden jugar uno contra el otro.

Lista de percepciones a usar

- Casillas disponibles.
- Casillas ocupadas por símbolo del agente.
- Casillas ocupadas por jugador contrario.

Lista de acciones a realizar

- Colocar símbolo.



Como se realiza el reforzamiento?

- Reforzamiento negativo. Cuando se coloca un símbolo y la partida termina en perdida en el siguiente turno, se sabe que movimientos no realizar.

Ejemplo de estado inicial y estado meta

- **Estado inicial.** Tablero de match 3 vacio.
- **Estado meta.** Table de match 3 con símbolos, donde hay 3 simbolos del agente seguidos.

Que sugerirías agregar para mejorar el comportamiento de este agente?

-

## Push Block

Objetivo del agente.

Empujar un bloque de posición A a una posición meta B.

Satisfacción de restricciones.

- Solo se puede mover en una dirección por momento.
- No se puede atravesar paredes.

Tabla PEAS.

| Type         | Performance Measure                             | Environment                           | Actuator              | Sensors |
|--------------|---|---------------------------------------|-----------------------|---------|
| Block pusher | - Numero de movimientos para llegar a la meta B | - Mapa con punto partida y punto meta | - Brazos<br>- Piernas | - Ojos  |

Características del ambiente

- Completamente observable
- Discreto
- Estático
- Determinista
- Secuencial
- Simple

Lista de percepciones a usar

- Posición del bloque
- Posición del agente
- Posición de la meta B.

Lista de acciones a realizar

- Moverse.
- Observar mapa.
- Empujar bloque.

Como se realiza el reforzamiento?

Reforzamiento positivo. Mientras mas lejos llegue el agente, mejor el resultado del entrenamiento.

Ejemplo de estado inicial y estado meta

- **Estado inicial.** Bloque y agente en punto de salida.
- **Estado meta.** Bloque y agente en el punto meta B.

Que sugerirías agregar para mejorar el comportamiento de este agente?

- Agregando una variante de tiempo, para que el agente tenga tiempo limitado para llegar del punto A al punto B.

## Snake

Objetivo del agente.

- Comer manzanas para seguir creciendo.

Satisfacción de restricciones.

- La serpiente no puede tocar los bordes del mapa.
- La serpiente crece una unidad por manzana comida.
- La serpiente no puede comer su cola.
- La serpiente se puede mover en las 4 direcciones únicamente (no diagonal).

Tabla PEAS.

| Type  | Performance Measure             | Environment                  | Actuator                 | Sensors |
|-------|---------------------------------|------------------------------|--------------------------|---------|
| Snake | - Numero de manzanas consumidas | - Mapa con manzanas y bordes | - Cuerpo de la serpiente | - Ojos  |

Características del ambiente

- Completamente observable
- Discreto
- Estático
- Determinista
- Secuencial
- Simple

Lista de percepciones a usar

- Posicion actual de la serpiente.
- Posición de la manzana.
- Distancia a la manzana.
- Distancia a los bordes.

Lista de acciones a realizar

- Moverse para arriba.
- Moverse para abajo.
- Moverse a la derecha.
- Moverse a la izquierda.

Como se realiza el reforzamiento?

Reforzamiento positivo. Mientras mayor el puntaje, mejor el desempeño del agente.

Ejemplo de estado inicial y estado meta

- **Estado inicial.** Serpiente de tamaño de 3 unidades en el mapa y una manzana en posición random.
- **Estado meta.** Serpiente de tamaño de unidades iguales al tamaño del mapa en el que se encuentra y no mas manzanas.

Que sugerirías agregar para mejorar el comportamiento de este agente?

- Pre entrenar el modelo para poder colocar mas restricciones y mejorar el desempeño del agente.

