



Meta Reinforcement Learning en Meta-Maze

MATTHIAS ALEXANDER GDANIETZ DE DIEGO: 100383277

MANUEL LOZANO RAMOS: 100514547

ALONSO MENÉNDEZ GONZÁLEZ: 100512399

Contenidos



¿Qué es Meta-RL?



Dominio Meta-Maze



Aplicación Meta-RL



¿Qué es Meta Reinforcement Learning?

1. Ideas Principales
2. Meta-World
3. Tareas Parametrizables
4. Otras Consideraciones



1. Ideas Principales

Adquirir Meta-Conocimiento

- Conocimiento abstracto, no específico.

Aprender de varias tareas

- Para poder generalizar

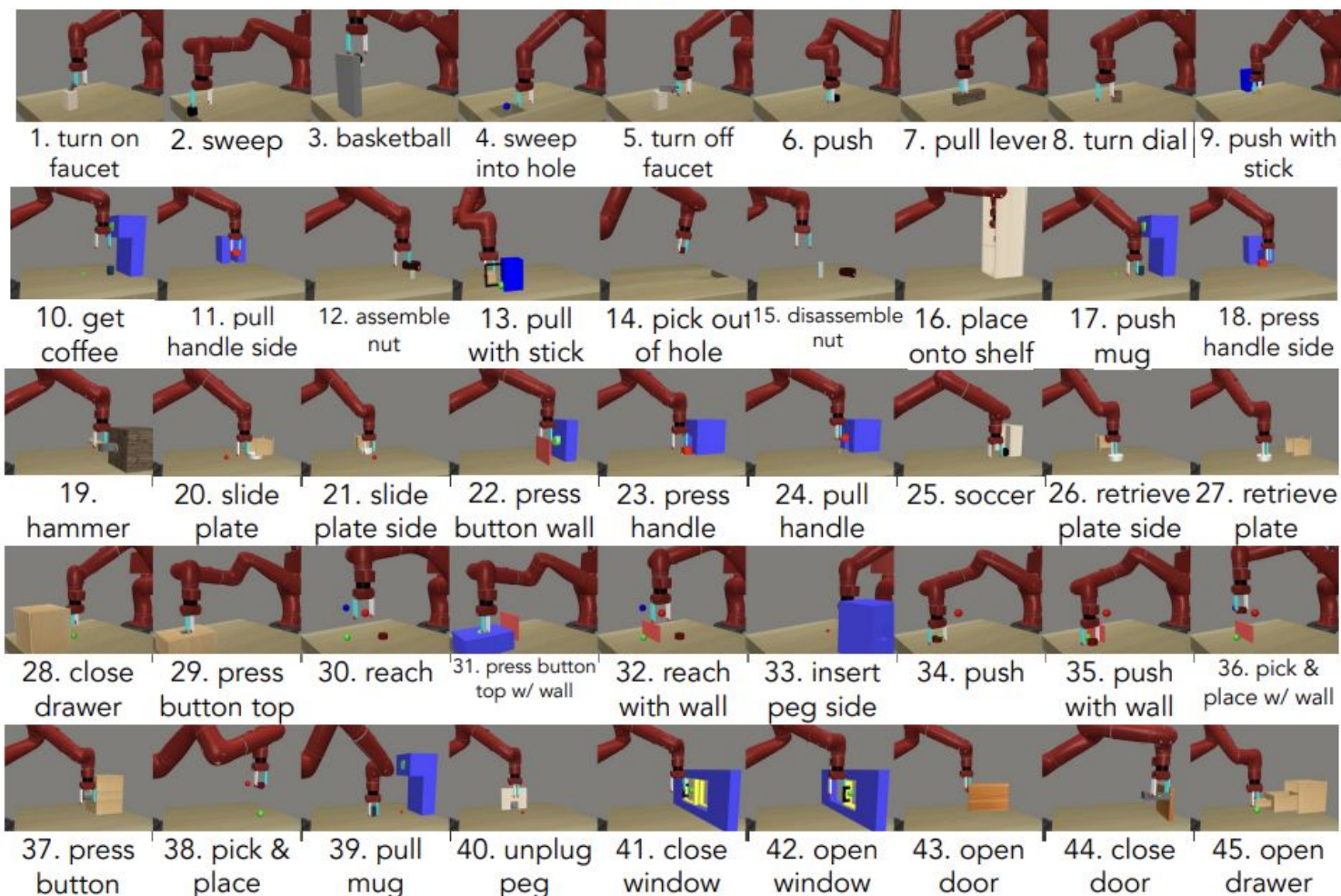
Aplicar en nuevos problemas

- Facilita el aprendizaje

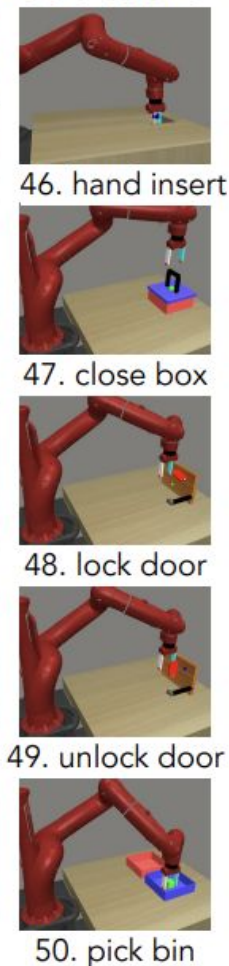
2. Meta-World



Train tasks



Test tasks

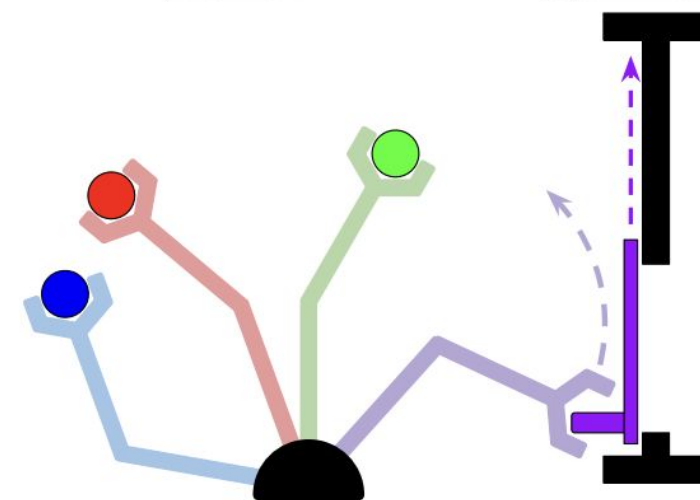


Non-Parametric Task Variation

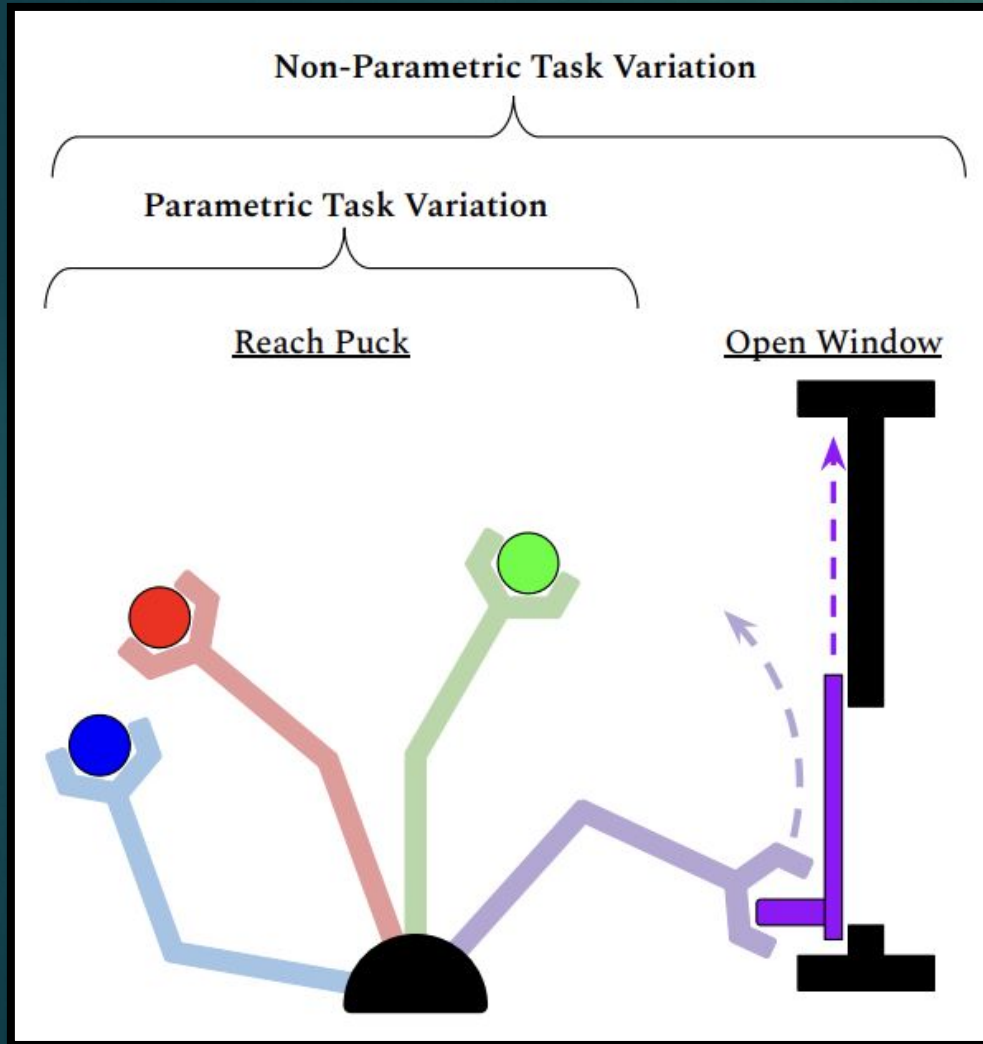
Parametric Task Variation

Reach Puck

Open Window



3. Tareas Parametrizables



Parametrizable

No
Parametrizable

Coger
un
objeto

Moverse
de una
posición
a otra

Mover
objeto

Abrir
una
puerta

Encestar
una
pelota

Abrir
una
ventana

Abrir
Puerta

Abrir
Ventana

4. Otras Consideraciones



Modelado de problemas

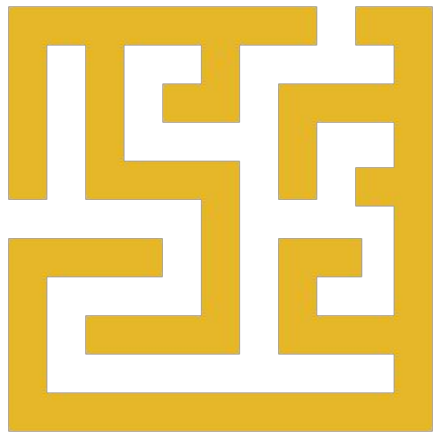
- Observaciones, acciones...
- Comunes a todas las tareas

Función de Refuerzo

1. Resolubles Individualmente
2. Estructura Común

Objetivos de Meta Learning

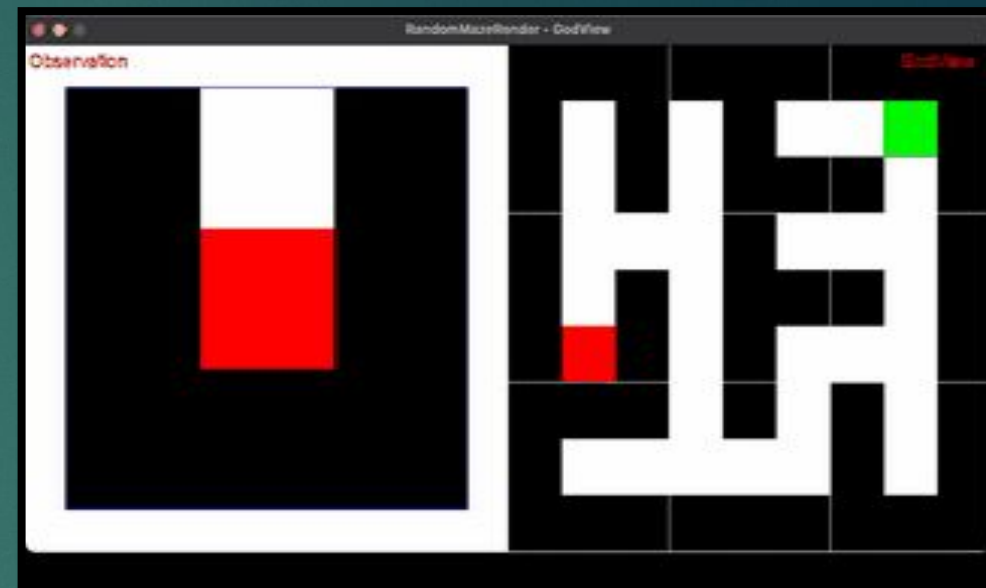
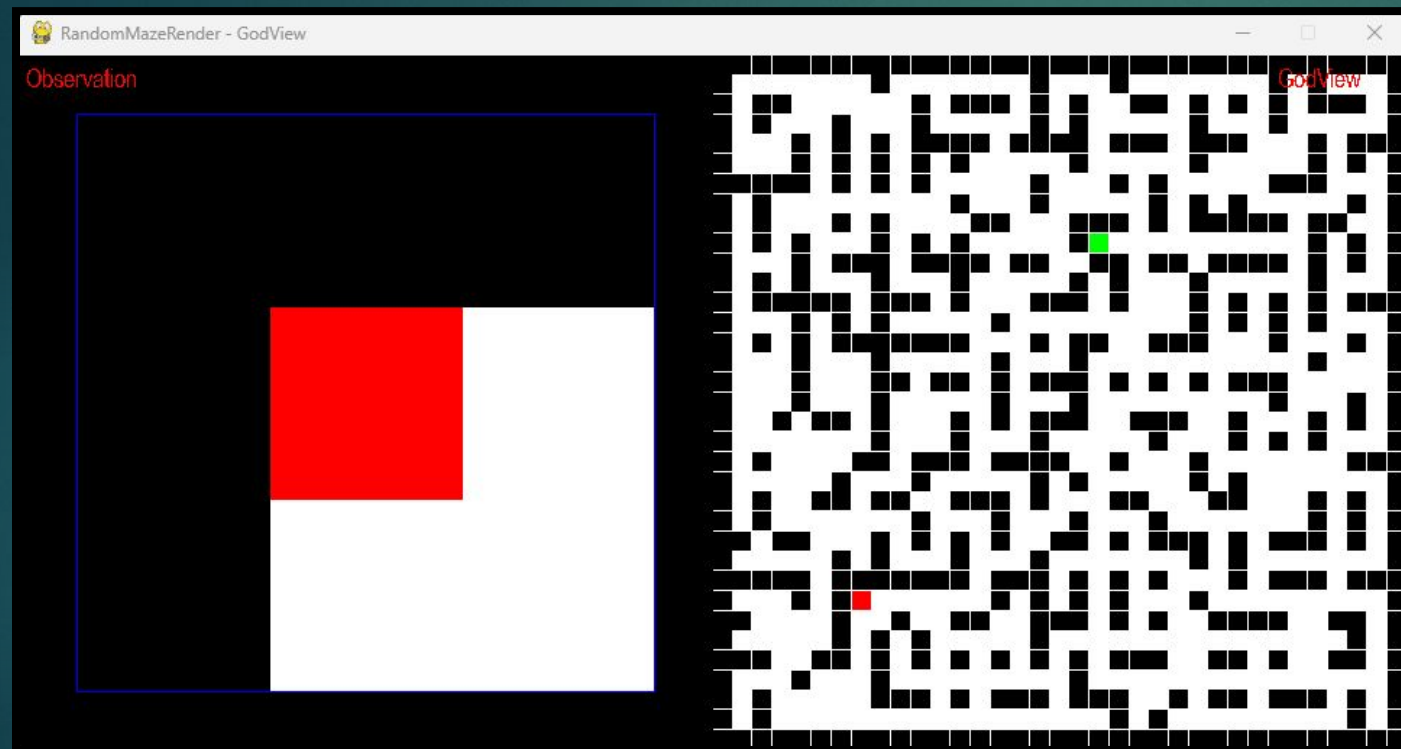
- 1 tarea – Variar las metas.
- Aprender antes nuevas tareas



Dominio Meta-Maze y Modelado del Problema

1. Meta-Maze
2. Observabilidad Parcial
3. Primer Modelo
4. Segundo Modelo

1. Meta-Maze



2. Observabilidad Parcial



Observación

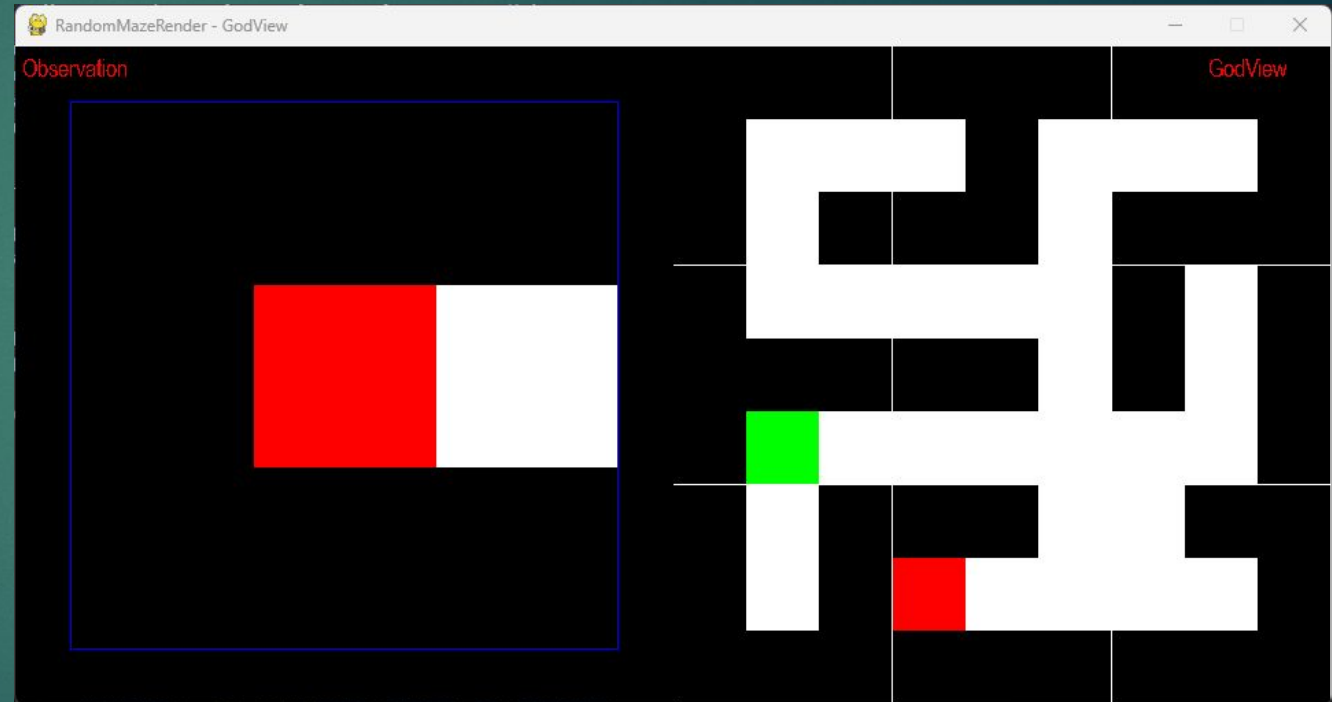
- Píxeles alrededor

Estado Inicial

- Mi posición es aleatoria
- Posición meta aleatoria
- Tamaño del laberinto conocido

Problemas

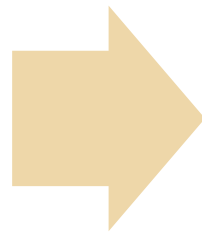
- ¿Dónde empecé?
- ¿Dónde estoy?
- ¿Dónde está la meta?



2. Observabilidad Parcial



¿Se puede
resolver con
Q-Learning?



¿Se puede usar
Meta-RL?

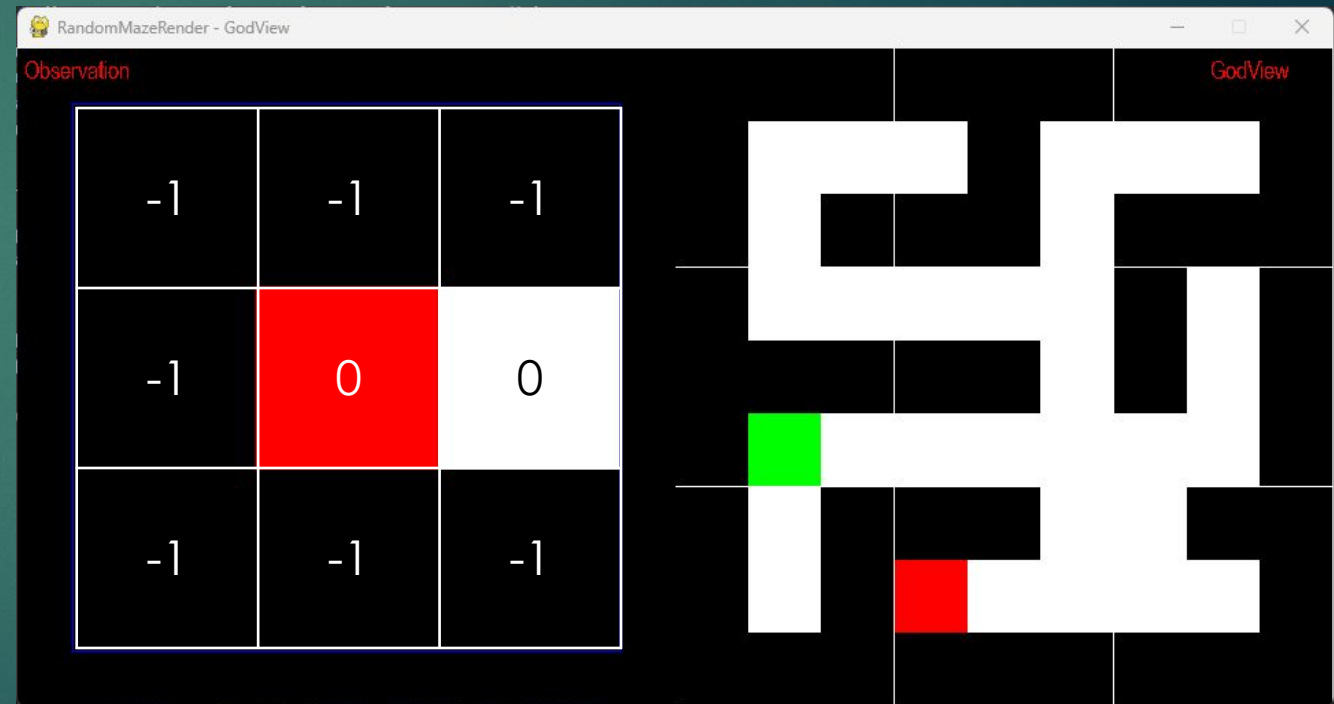
3. Primer Modelo: Observación como estado



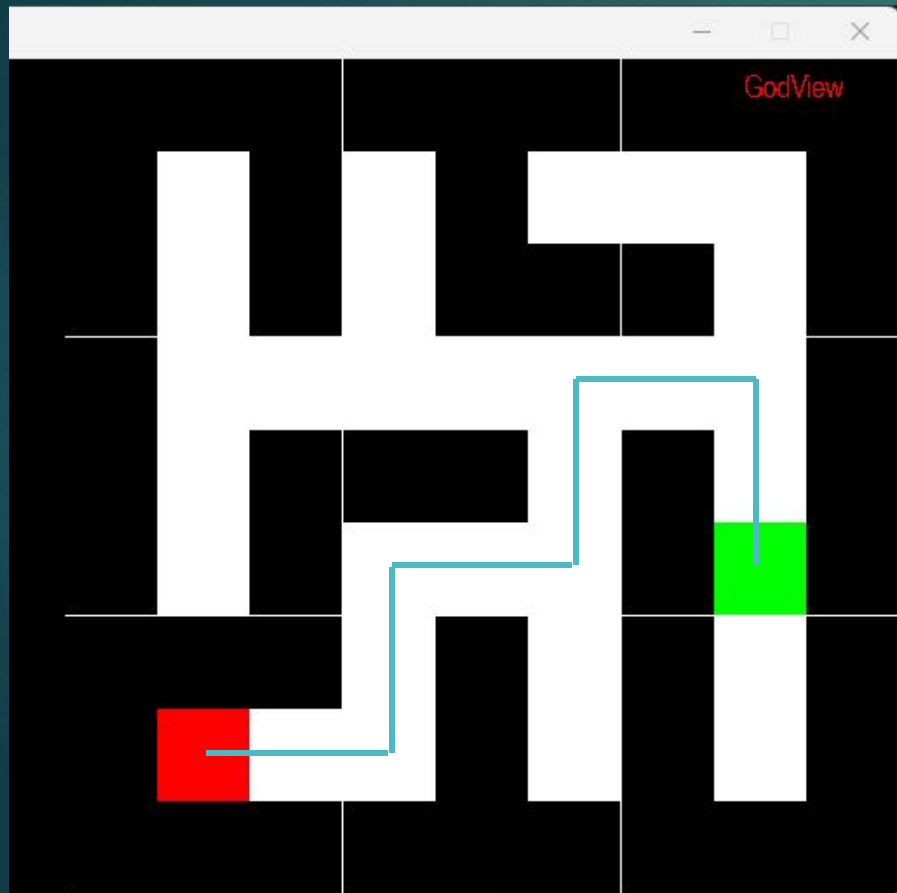
Cada observación es
un estado

Codificado en un
solo entero

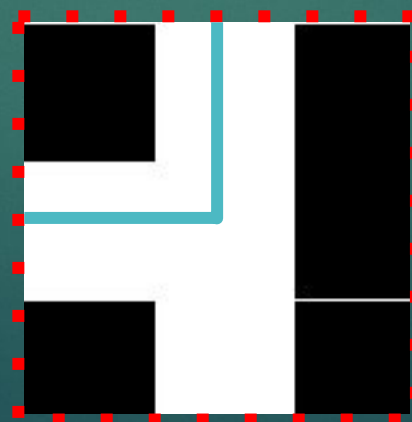
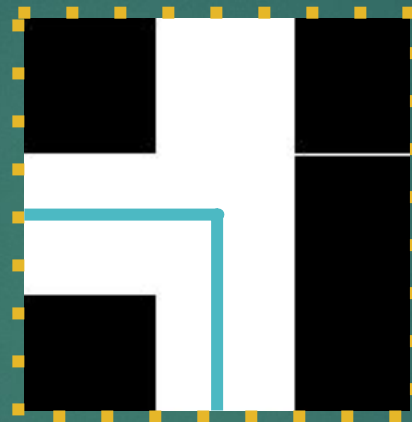
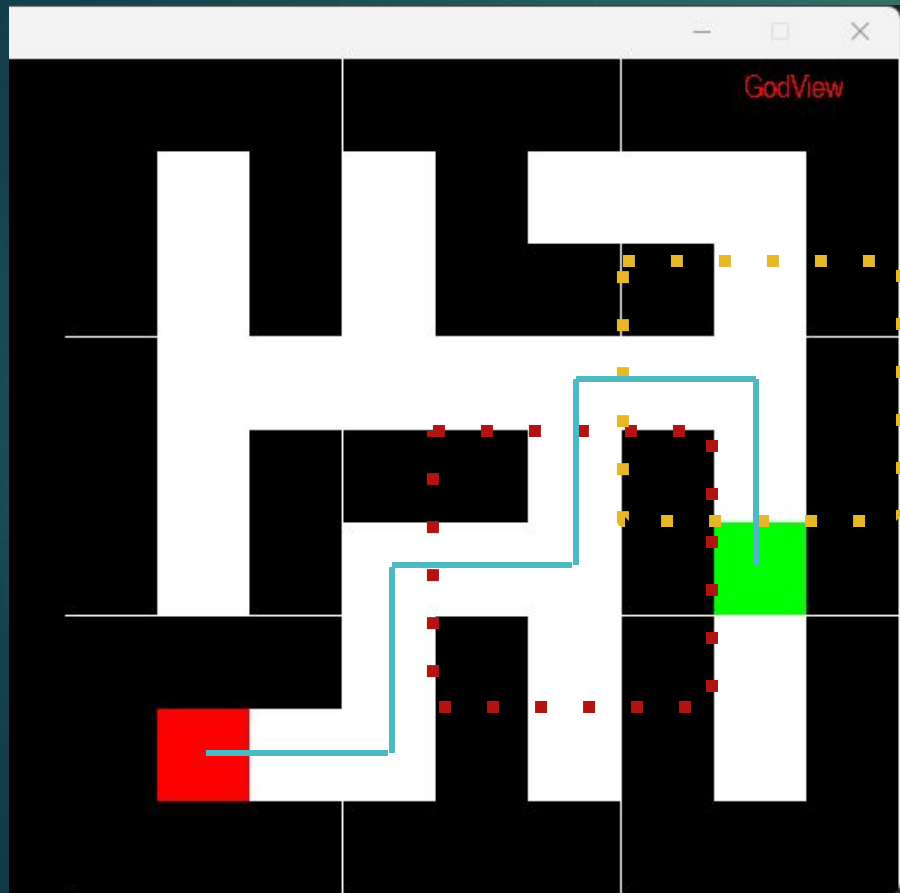
Generalización
aprendizaje



3. Primer Modelo: Problemas



3. Primer Modelo: Problemas



Acción: Sur



Acción: Norte

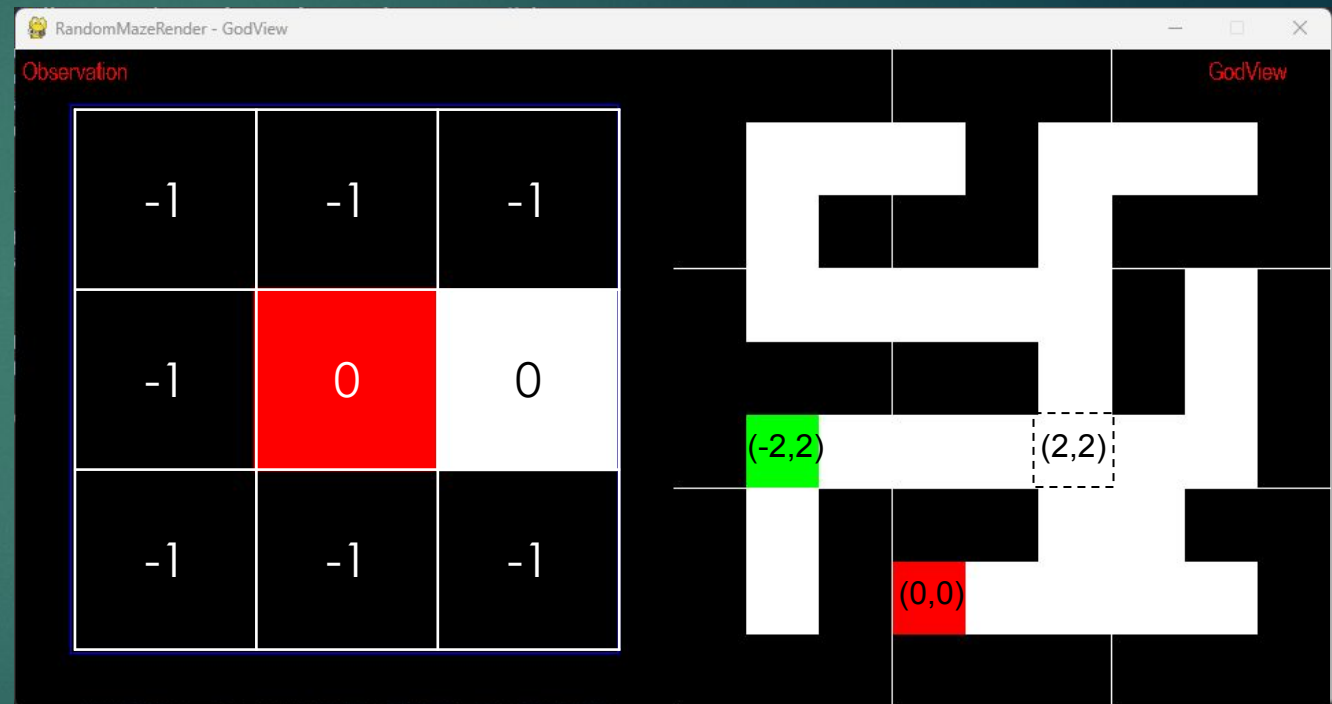
4. Segundo Modelo: Posición Relativa como estado



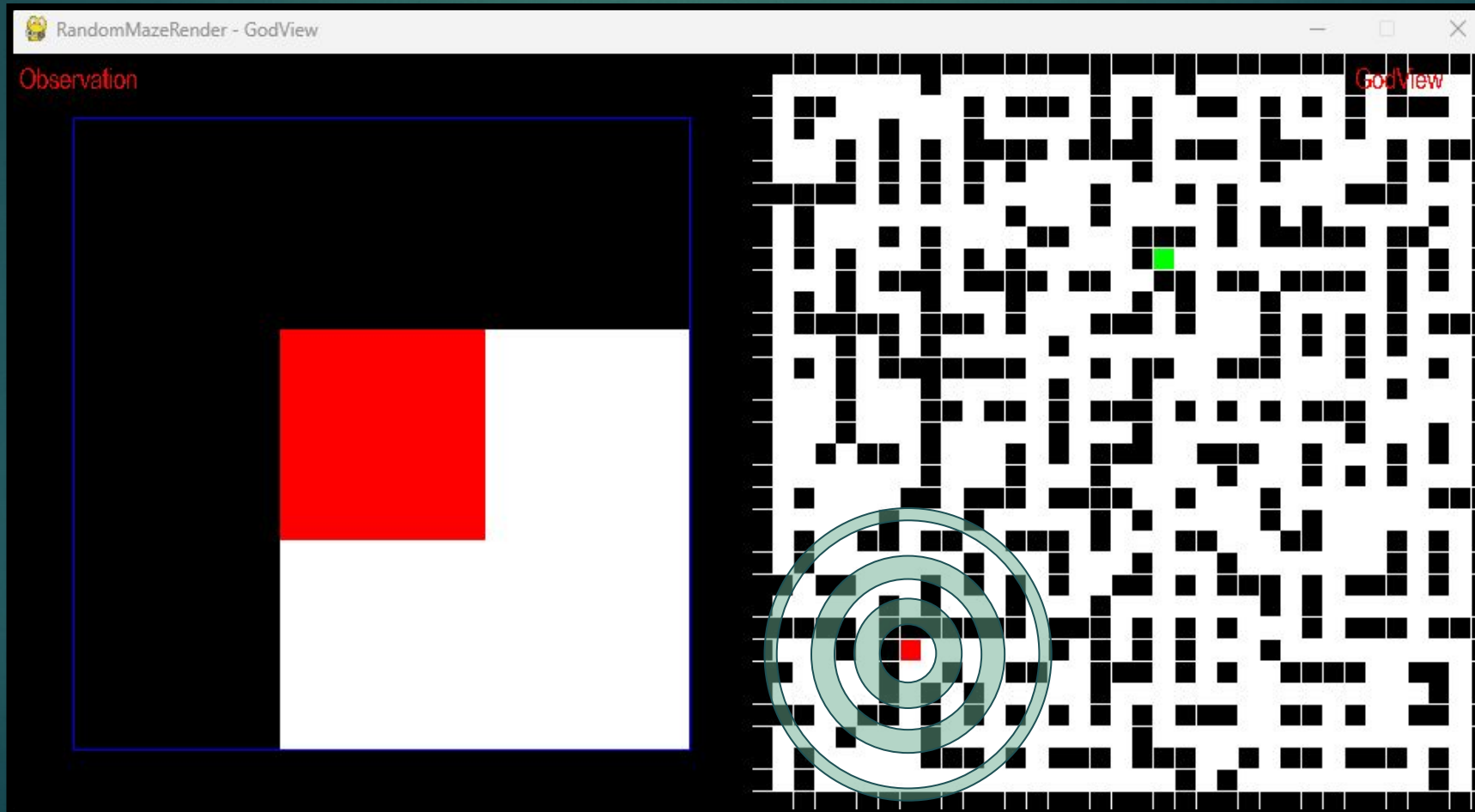
Se empieza en (0,0)

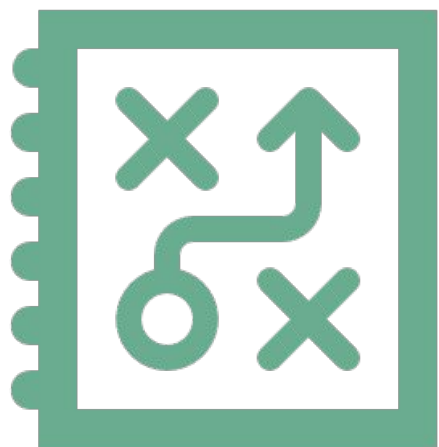
Moverse a una
posición libre cambia
la Coordenada

Codificado en un
solo entero



4. Segundo Modelo: Problemas





Aplicación de Meta-RL en Meta-Maze

1. Objetivo del Meta-Conocimiento
2. Reglas Manuales
3. Reglas Automáticas
4. Datos a usar en los Modelos
5. Modelos Propuestos
6. Nuestra Implementación

1. Objetivo del Meta-Conoci- miento



Exploración

- Aleatoria
- **Reglas Manuales**
- **Reglas Automáticas**

Explotación

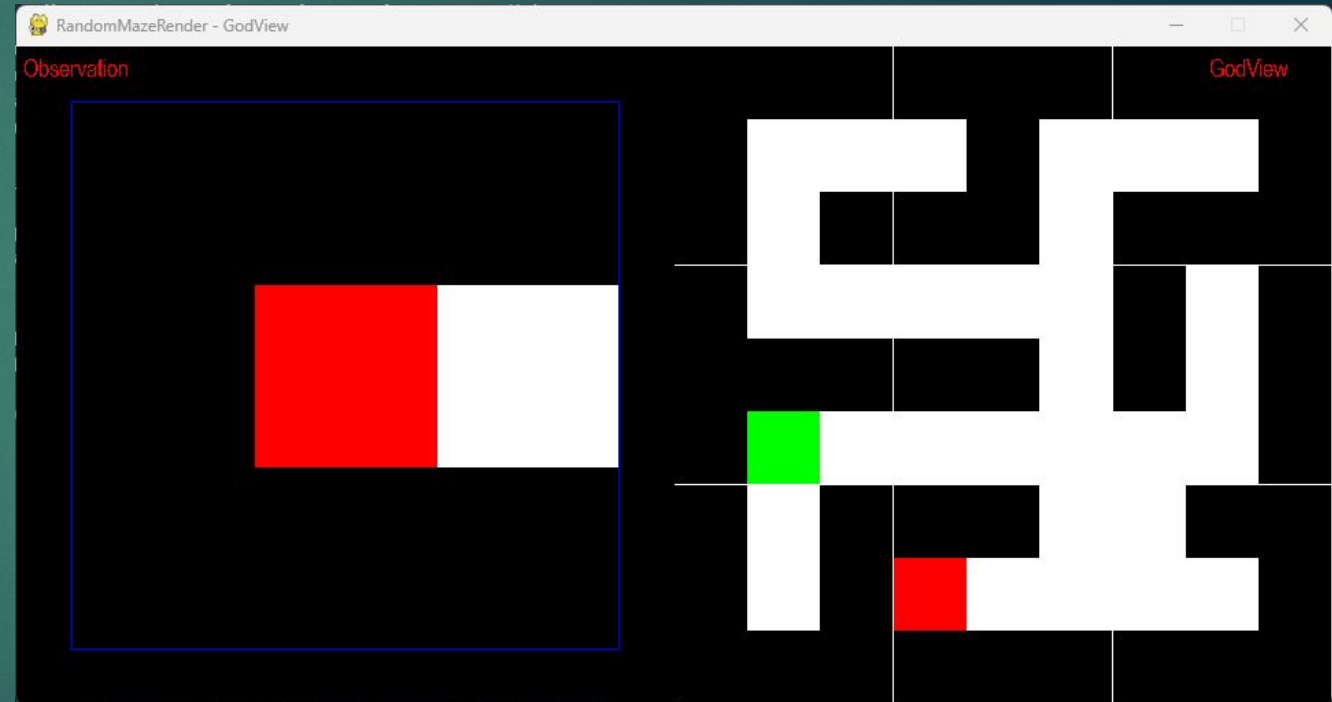
- Política aprendida

2. Reglas Manuales

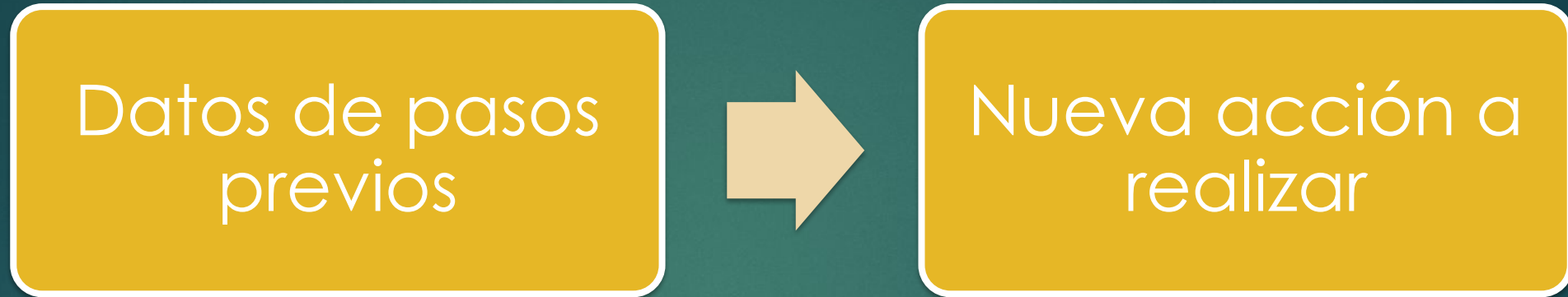


Simplicidad

- Reconstruir mapa
- Seguir por pasillos
- Recordar dirección de la meta
- No chocar con paredes



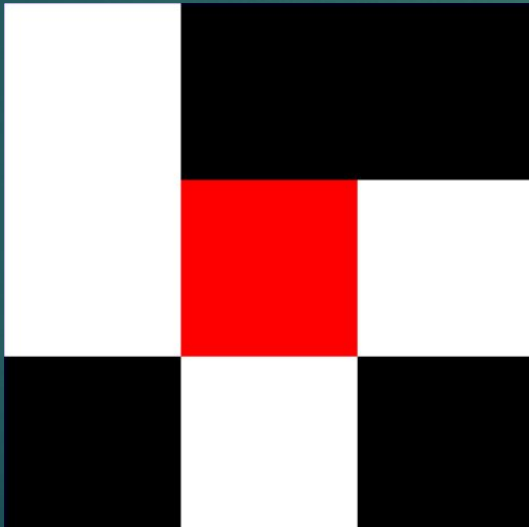
3. Reglas Automáticas



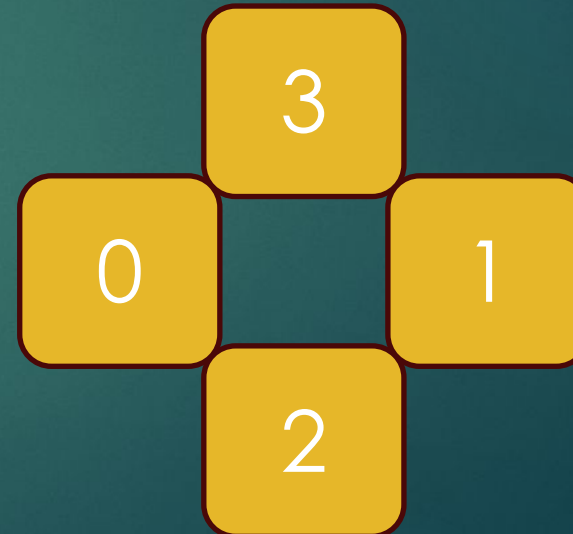
4. Datos a usar en los Modelos



Observaciones



Acciones

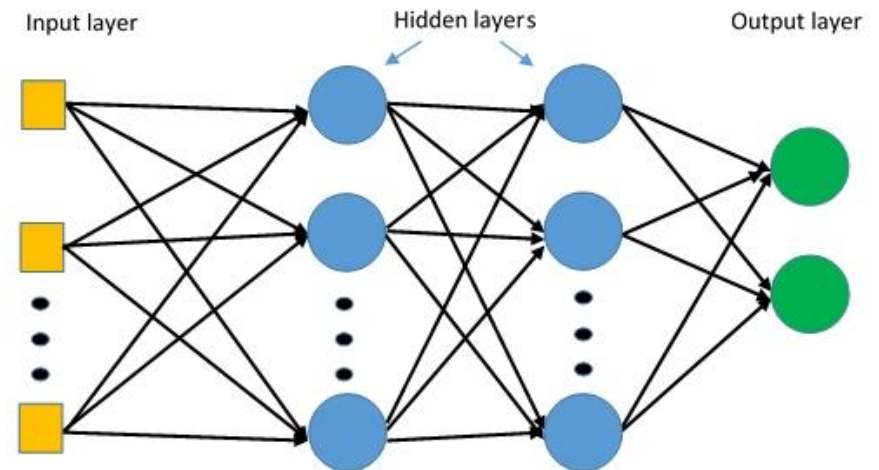
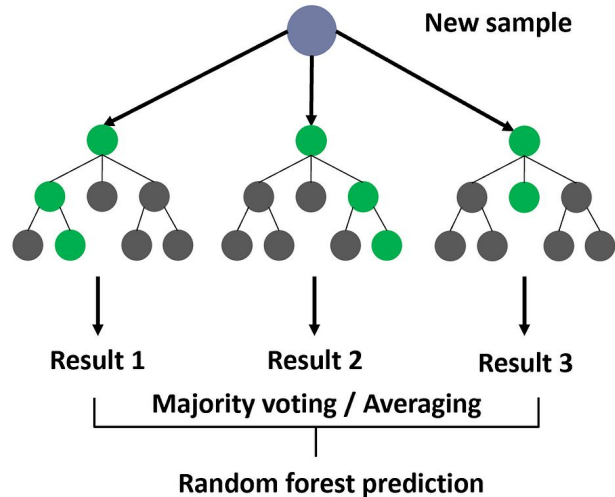


5. Modelos Propuestos



Random
Forest y
similares

MLP

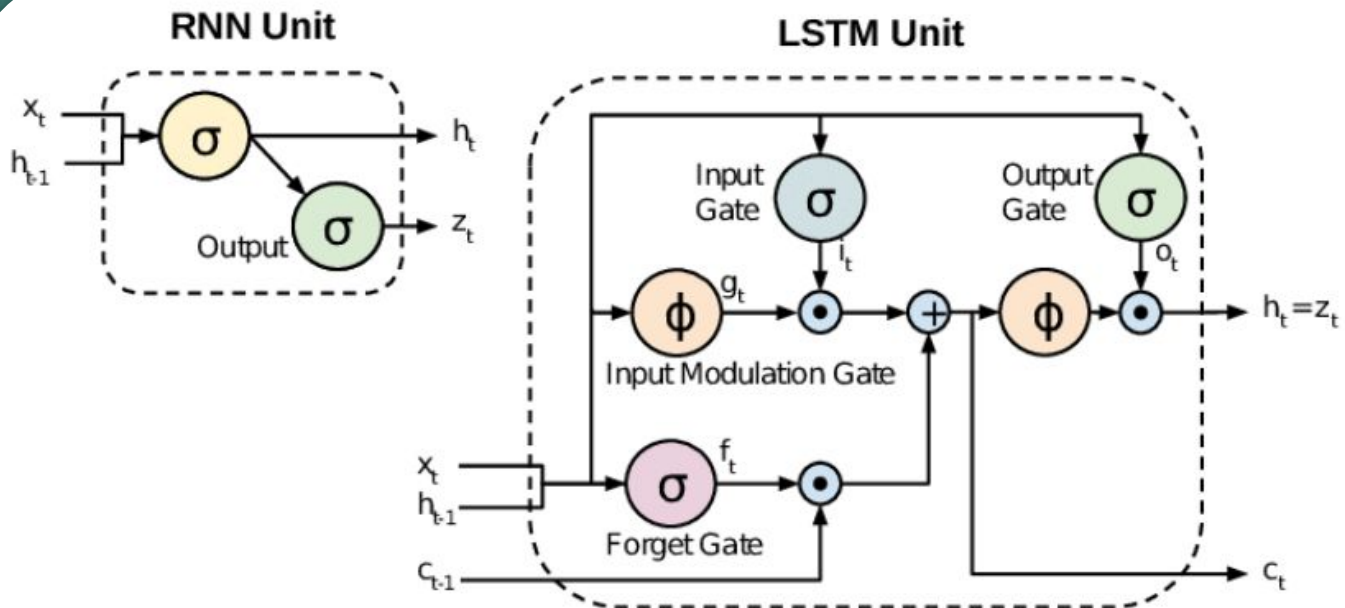


5. Modelos Propuestos



RNN

LSTM



6. Nuestra Implementación



Reglas Manuales

- No chocar con las paredes

Reglas Automáticas

- Solo acciones
- Observaciones y acciones
- Random Forest
- LSTM



Muchas gracias!