

Informe Taller 2

Pola Escobar, Víctor Núñez, José Salazar, Sofía Sepúlveda

27 de Octubre del 2025

1. Introducción

Este trabajo tiene como propósito analizar diferentes estrategias de agentes, vistas durante el semestre, para definir cuál de ellas es más efectiva en distintas condiciones del entorno relacionadas con el Mundo Wumpus. Para esto, se utilizaron métricas propias del modelo además de respetar sus reglas básicas, que son explorar un entorno desconocido, recolectar oro y salir con vida evitando los pozos y al Wumpus.

2. Objetivos

1. Modelar el Mundo del Wumpus en NetLogo como entorno de prueba para agentes inteligentes.
2. Diseñar e implementar tres estrategias diferentes de agente y analizar sus resultados.
3. Comparar el desempeño mediante métricas y gráficos experimentales
4. Exponer resultados en una presentación técnica breve fundamentada en datos.

3. Descripción del entorno y estrategias

El entorno presente en esta actividad considera es una cuadrícula de tamaño configurable (grid-size), con una salida en coordenadas aleatorias, y parches que comparten propiedades y métricas como la presencia de oro, pozos y wumpus con percepciones locales como brisa (si hay pozo), hedor (si está el wumpus) y brillo (si está el oro) que sirven como pista para el agente dentro del mapa.

Al mismo tiempo, el agente tiene acciones básicas tales como moverse en las direcciones norte, sur, este y oeste, girar en las mismas direcciones descritas, disparar una flecha, recoger el oro y dirigirse hacia la salida. Se implementaron y graficaron episodios de prueba para evaluar el desempeño de las distintas estrategias a partir del cumplimiento del objetivo (obtener el oro y salir con vida, evitando al wumpus y los pozos) y de acuerdo al sistema de puntuación estándar del juego.

- **Agente Reactivo Aleatorio (baseline):** El agente tiene un comportamiento de control de referencial, el cual reacciona a los estímulos o percepciones inmediatas y no implementa una planificación del camino (avanza de forma aleatoria durante su ejecución), además de excluir la memoria con relación a los pasos realizados.

Cuadro 1: Puntajes

Evento	Recompensa
Recolectar oro	+1000
Salida con oro (éxito)	+100
Muerte (pozo o wumpus)	-1000
Uso de flecha	-10
Pasos dados	-1
Salida sin oro (huida)	0 a 10 (según criterio)

- **Agente Reactivo Heurístico:** Sigue presentando una conducta reactiva, pero incluye el funcionamiento de evaluar las celdas adyacentes a través de una función heurística. En este caso se utilizó una heurística que realiza un calculo para determinar el mejor camino posible mediante las casillas, con esto se penaliza (puntaje $-$) si existe presencia de pozo, el wumpus y bordes del mapa, además de beneficiarse (puntaje $+$) al encontrar oro, salida y celdas no visitadas, dependiendo de esto el camino con mayor puntuación es la eligida. Distinguiéndose del método anterior al eliminar el azar, reducir lo más posible las repeticiones en los pasos y evitar los riesgos del entorno.
- **Agente con Modelo Interno (lógico o probabilístico):** Es la única estrategia lógica presente en esta actividad, razonando y construyendo un modelo para planificar acciones seguras en el entorno integrando las percepciones de los elementos del mundo (brisa, brillo y hedor) con una memoria que respalda las celdas seguras y visitadas. Con la finalidad de, encontrar la mejor ruta al oro y evitar los riesgos anteriormente mencionados. En este caso, establecimos la función A* implementado heurística de distancia Manhattan que nos ayudan a inferir peligros, planificar rutas con celdas seguras y el actuar del agente (explorar, recoger oro, volver a la salida y/o disparar al Wumpus).

4. Diseño experimental y métricas

Con el objetivo de evaluar el desempeño del agente en el Mundo del Wumpus, se desarrolló la simulación correspondiente en NetLogo, que permite controlar las condiciones del entorno y reproducir múltiples experimentos de forma consistente. Para ello se diseñó una interfaz con distintos *widgets* que nos permiten configurar parámetros clave del entorno, tales como: Tamaño de la grilla, Semilla (Seed), Probabilidad de aparición de los pozos, Wumpus móvil o estático.

Cada simulación empieza con el agente en su casilla inicial (0, 0) y finaliza cuando una de las siguientes condiciones se cumpla:

1. El agente encuentra el oro y logra escapar.
2. El agente muere ya sea por caer en un pozo o ser atacado por el Wumpus.
3. Se alcanza un límite de pasos, para evitar bucles infinitos.

Se definen seis métricas para medir el rendimiento del agente: Tasa de éxito, recompensas promedio, pasos promedio, porcentaje de muertes por pozo, porcentaje de muertes por Wumpus y cantidad de flechas usadas.

5. Resultados*

Cuadro 2: Resultados Reactivo-Baseline

Tamaño mapa	4x4			6x6			8x8		
Probabilidad de pozo	0,10	0,15	0,20	0,10	0,15	0,20	0,10	0,15	0,20
Tasa de éxito	17,67 %	13,67 %	10 %	6 %	5,33 %	8 %	2,33 %	0,67 %	2 %
Recompensa promedio	-261,31	-381,82	-561,81	-662,16	-686,39	-653,35	-839,35	-908,56	-917,60
Pasos promedio	17,98	14,76	11,47	22,06	25,36	22,11	19,95	8,59	11,80
Muerte por pozos	56 %	65,67 %	73 %	82,67 %	84,33 %	81,67 %	94,33 %	97,67 %	96,67 %
Muerte por Wumpus	14 %	10 %	12 %	7,33 %	4,33 %	5,67 %	2,33 %	1,33 %	1,33 %
Flechas usadas	130	122	110	83	91	77	57	28	34

Cuadro 3: Resultados Reactivo-Heurístico

Tamaño mapa	4x4			6x6			8x8		
Probabilidad de pozo	0,10	0,15	0,20	0,10	0,15	0,20	0,10	0,15	0,20
Tasa de éxito	68,33 %	53 %	50 %	50,33 %	27,33 %	29 %	40 %	14 %	24,33 %
Recompensa promedio	697,83	499,69	444,32	484,98	161,44	195,40	334,77	-9,55	124,65
Pasos promedio	15,11	15,68	17,54	41,05	58,90	57,47	88,90	121,85	110,88
Muerte por pozos	11,33 %	22,67 %	20,67 %	18 %	20 %	21,33 %	18 %	24,33 %	20,33 %
Muerte por Wumpus	4,67 %	3,67 %	4,67 %	2 %	2 %	0,64 %	0,67 %	0,33 %	0,33 %
Flechas usadas	162	129	144	129	110	84	90	57	64

Cuadro 4: Resultados Modelo-Interno

Tamaño mapa	4x4			6x6			8x8		
Probabilidad de pozo	0,10	0,15	0,20	0,10	0,15	0,20	0,10	0,15	0,20
Tasa de éxito	40,00 %	24,00 %	17 %	30 %	11,33 %	10,67 %	26,00 %	5,33 %	6,67 %
Recompensa promedio	506,43	340,42	220,13	429,15	184,18	197,62	330,06	-17,54	28,98
Pasos promedio	34,27	38,61	38,84	87,31	97,45	100,22	162,91	181,27	182,42
Muerte por pozos	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %	0 %
Muerte por Wumpus	12,67 %	7,33 %	13 %	6,67 %	5 %	1,67 %	5,33 %	3 %	1,33 %
Flechas usadas	179	149	141	116	91	85	91	48	58

* Todos los experimentos fueron con el Wumpus móvil.

6. Conclusiones y trabajo futuro

En este informe presentamos la implementación del Mundo del Wumpus en NetLogo y la comparación experimental de tres estrategias de agente: reactivo baseline, reactivo heurístico y con modelo interno. A partir de los experimentos realizados (300 episodios por configuración, tamaños de malla 4×4 , 6×6 y 8×8 , probabilidad de pozos 0.10, 0.15 y 0.20 y semillas aleatorias) se extraen las siguientes conclusiones principales:

- El agente reactivo heurístico mostró el mejor rendimiento agregado según las métricas consideradas: mayor tasa de éxito y la recompensa promedio más alta en la mayoría de las configuraciones. Esto indica que incorporar una función heurística que penaliza señales de peligro y premia la exploración y la proximidad al oro mejora sustancialmente el comportamiento frente a un agente puramente aleatorio.
- El agente con modelo interno demostró ser muy conservador en cuanto al riesgo: en las salidas con las configuraciones probadas no registró muertes por pozos (0 % en las tablas), lo que sugiere que la representación y razonamiento sobre celdas seguras es efectiva para evitar peligros mortales. Sin embargo, esta estrategia tiende a recorrer más pasos y, en algunas configuraciones, necesita más flechas, lo que reduce su eficiencia temporalmente (mayor costo por pasos).
- El agente reactivo baseline obtuvo los peores resultados: tasas de éxito muy bajas y altas tasas de muerte por pozos. Esto confirma que la ausencia de memoria y de criterios heurísticos dificulta la exploración segura en entornos con incertidumbre parcial.
- En términos de trade-offs, el agente heurístico ofrece un buen compromiso entre eficiencia (pasos y recompensa) y seguridad, mientras que el agente con modelo interno prioriza la seguridad a costa de mayor exploración y tiempo.

Limitaciones del estudio: los experimentos aquí reportados usan un número finito de episodios (300), una probabilidad fija de pozos (0.10, 0.15 y 0.20) y semillas aleatorias puntualizadas. No se realizaron pruebas estadísticas de significancia para todas las comparaciones ni un barrido exhaustivo de parámetros (p. ej. distinta cantidad de episodios, o cambios sistemáticos en la movilidad del Wumpus), por lo que las conclusiones deben entenderse dentro del alcance experimental usado.

Trabajo futuro y mejoras recomendadas:

- Realizar barridos de parámetros (grid-size, probabilidad de pozos, número de episodios) y repetir con múltiples semillas para obtener intervalos de confianza y pruebas estadísticas que respalden las diferencias observadas.
- Ampliar el modelo interno con actualizaciones probabilísticas más formales (p. ej. estimaciones bayesianas o filtros de creencia) para mejorar la inferencia de la posición del Wumpus y reducir la necesidad de disparos innecesarios.
- Evaluar métodos de aprendizaje (por ejemplo, aprendizaje por refuerzo) para que el agente aprenda políticas directamente desde las recompensas y comparar su rendimiento con las estrategias programadas.
- Investigar variantes del entorno: múltiples Wumpus, más de una flecha, distintos niveles de ruido en las percepciones y escenarios multiagente.

- Mejorar la recolección de datos y reproducibilidad: exportar logs, automatizar experimentos con semillas controladas y generar scripts para reproducir los experimentos y las gráficas.
- Añadir pruebas unitarias y documentación en el código NetLogo para facilitar mantenimiento y extensión por futuros equipos.

En resumen, los resultados muestran que incorporar conocimiento local (heurísticas) y/o un modelo interno incrementa considerablemente la capacidad del agente para alcanzar objetivos en el Mundo del Wumpus. Las mejoras propuestas apuntan a robustecer las conclusiones mediante experimentos más amplios y técnicas de inferencia o aprendizaje que podrían producir agentes más eficientes y robustos en entornos parcialmente observables.