Minesweeper solver

David Santiago Flórez Alsina* and Laura Sofía Ortiz Arcos†
Universidad del Rosario, Escuela de Ingeniería, Ciencia y Tecnología, Bogotá, Colombia
(Dated: November 25, 2022)

Resumen

En este artículo se muestra el análisis aplicado a nuestra solución para el problema del mundo del Wumpus. Lo anterior, por medio de tomas de decisiones racionales *(redes de decisión)*, es decir, decisiones que maximizan la utilidad esperada del agente, además del uso de redes bayesianas. Finalmente, se realizó una comprobación del funcionamiento del agente en diversas cavernas del Wumpus, teniendo en cuenta el tiempo y la proporción de victorias.

Keywords: *Redes de decisión, redes bayesianas, Wumpus.*

I. INTRODUCCIÓN

Hunt the Wumpus es un juego de aventuras desarrollado por Gregory Yob en 1973. En el juego, el jugador *(héroe de la historia)* se mueve a través de una serie de cavernas conectadas, dispuestas como los vértices de un dodecaedro, mientras cazan un monstruo llamado Wumpus. En nuestro caso, el problema se desarrolará en una caverna representada por una rejilla 4x4 rodeada de muros.

Nuestro agente, es decir el héroe, va a poder entrar y salir de la caverna unicamente por la casilla (0,0) y además puede percibir solamente lo que hay en la casilla en la que se encuentre.

También, debemos de tener en cuenta que en la caverna hay pozos que el agente debe de evitar, así como también un monstruo, conocido como el Wumpus, el cuál se comerá al agente si éste entra a su casilla, es decir, si el agente cae en un pozo o es comido por el Wumpus pierde el juego.

Por otro lado el agente ganará el juego si logra encontrar el oro que se encuntra escondido en la caverna y salir con este. A continuación se muestra cómo luce el juego del mundo del Wumpus:

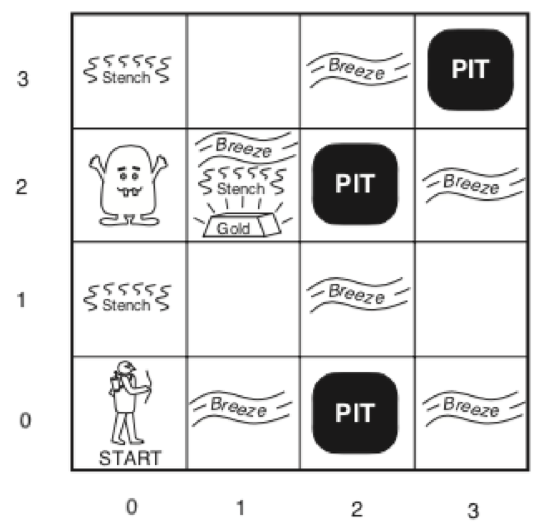

Figure 1. Mundo del Wumpus

A. Reglas del juego

Las siguientes son las reglas que gobiernan el mundo:

- En cualquier casilla adyacente (no diagonalmente) a un pozo se percibe una brisa.
- En cualquier casilla adyacente (no diagonalmente) al Wumpus se percibe un hedor.
- En la casilla donde se encuentra el oro se percibe un brillo.
- El wumpus nunca se mueve de su casilla.

Teniendo en cuenta lo anterior, surge la duda de ¿Cómo hará el agente para saber si hay un pozo, o si se encuentra

* Correspondence email address: davidsa.florez@urosario.edu.co
† Correspondence email address: lauraso.ortiz@urosario.edu.co

el Wumpus? Para esto tendremos en cuenta la utilidad esperada del agente y las decisiones que realice, esto lo veremos más adelante en la metodología que se usó para resolver el problema.

Por otro lado, nuestro problema se tienen las siguientes propiedades del entorno:

Opción 1	Opción 2
Completamente observable	✓ Parcialmente observable
✓ Agente único	Multiagente
Determinista	✓ Estocástico
Episódico	✓ Secuencial
✓ Estático	Dinámico
✓ Discreto	Continuo
✓ Conocido	Desconocido

II. MÉTODOS

La definicion formal del problema del juego del Wumpus, puede ser descrita como:

- Estado inicial: El agente se encuentra en una caverna desconocida representada por una rejilla 4x4 rodeada de muros, siempre empieza la casilla (0, 0) mirando a la derecha. La ubicación del Wumpus se escoge arbitrariamente de manera uniforme en casillas distintas a la inicial. Cualquier casilla distinta de la inicial puede ser un pozo con probabilidad 0.2. El oro puede estar en cualquier casilla, con probabilidad uniforme.

- Posibles acciones: El héroe puede moverse adelante por una casilla (no es posible moverse adelante cuando hay un muro), voltearIzquierda, o voltearDerecha por 90º. Es posible agarrar el oro si el agente se encuentra en la casilla donde estaba el oro. Finalmente, el agente puede salir de la caverna, pero solo desde la casilla inicial.

- Prueba de satisfacción del objetivo: Que el agente logre completar el juego, saliendo de la caverna con el oro.

- Función de utilidad: la función de utilidad es la siguiente: Si la acción va a una casilla sin wumpus y sin pozo y hay oro en esa casilla el premio es de 5; Por otra parte si no hay oro, pozo, ni wumpus, el premio es de cero *(no castigamos el explorar)*, si la casilla tiene wumpus obtiene un -3, y si la casilla tiene un pozo obtiene -5. El juego termina cuando el heroe muere o sale de la cueva.

Los métodos empleados para resolver este problema, fueron por medio de decisiones racionales maximizando la utilidad esperada del agente, y con ayuda de redes bayesianas. Para que nuestro agente logre resolver el juego vamos a explicar los métodos que lo ayudan:

A. Utilidad esperada y toma de decisiones

Como hemos visto durantes las clases la idea es construir agentes inteligentes, es decir, que sean capaces de realizar acciones racionales, con las cuáles se buscará que el agente resuelva cualquier problema de la manera más eficiente posible.

Vamos a tener en cuenta que cada acción que tome tendra un estado, en otras palabras el conjunto de acciones que tome el agente $a_i = \{a_1, a_2, \ldots\}$ producirá un estado $S_i = \{S_1, S_2, \ldots\}$ para cada acción. Como nuestro agente tiene un objetivo *(obtener el oro)*, los estados se tendrán que dividir en aquellos que cumplen el objetivo y los que no, S_i y S_j. Esto nos permitirá que el agente se acerque más a cumplir el objetivo, al realizar una acción a que con mayor probabilidad lo lleve al estado que esté más cerca del objetivo.

Teniendo en cuenta lo anterior, podemos formalizar la idea utilizando los siguientes conceptos:

- El primero es el de un **conjunto de estados**, cada uno con una probabilidad asignada: $\{S_1, p_1; S_2, p_2; \ldots; S_n, p_n\}$.

- El segundo es el de una **función de utilidad** U, la cuál le asigna un valor numérico a cada estado S_i tale que $U(S_i) > U(S_j)$, si y solo si el agente prefiere el estado S_i sobre el estado S_j.

Estos dos primeros conceptos nos permitirán definir lo que hemos nombrado anteriormente como la **Utilidad esperada**, de la siguiente forma:

$$UE = \sum_{i=1}^{n} U(S_i) p_i$$

- El último concepto es que el conjunto de estados $\{S_1, p_1; S_2, p_2; \ldots; S_n, p_n\}$ depende de una acción a. Es decir, que se pueden llegar a representar situaciones en las cuáles el resultado de una acción a sea incierta.

  Esta incertidumbre puede resultar ya que la acción no determina completamente el resultado cuando, por ejemplo, este depende de otros agentes, o de factores que el agente puede no conocer y/o controlar.

  En general, una acción a da como resultado un estado S_1 con probabilidad $P(S_1|a)$, ..., o un estado S_n con probabilidad $P(S_n|a)$.

Así, con los tres conceptos tendremos que la **Utilidad Esperada** de una acción a está dada por:

$$UE(a) = \sum_{i=1}^{n} U(S_i) P(S_i|a) \tag{1}$$

Por último nuestro agente racional, frente a un problema de decisión, decidirá ejecutar la acción que le permita maximizar su utilidad esperada:

$$Accin = argmax_a UE(a) = argmax_a \sum_{i=1}^{n} U(S_i)P(S_i|a) \tag{2}$$

Conociendo que es la utilidad esperada veremos un pequeño ejemplo del juego.

1. *Ejemplo con el mundo del Wumpus*

Para poder explicar cómo un agente puede tomar una decisión racional, veremos un ejemplo de cómo sería en el mundo del Wumpus. Suponga que el agente, después de empezar en la celda (0,0), decide dirigirse a la casilla (1,0), y percibe una brisa que indica que pueden haber pozos cerca. La siguiente imagen mostrará cuáles podrías ser las posibles situaciones en la que se encontrará.

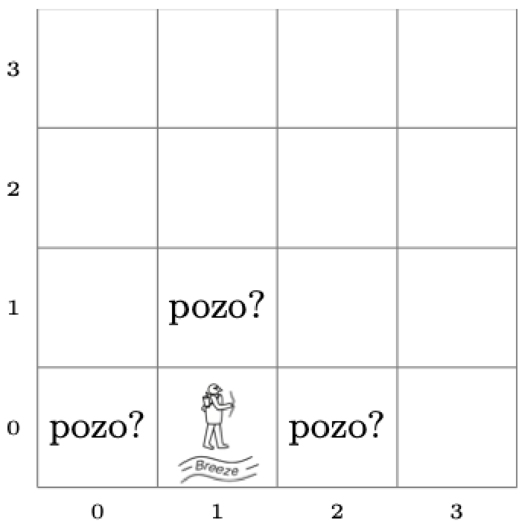

Figure 2. Mundo del Wumpus, ejemplo con pozos

Podemos observar que el agente tiene tres acciones posibles:

1. Regresarse a $(0, 0)$
2. Seguir a $(2, 0)$
3. Subir a $(1, 1)$.

Cada acción posible tendra como resultado un estado nuevo, estos estados son la casilla nueva a la que va a llegar después de realizar la acción. Vamos entonces a asignar la utilidad de cada estado, coniderando si en la casilla correspondiente hay o no un pozo, y de si está o no el oro. Es decir:

$$U(casilla(x,y)) = \begin{cases} 1, & sioro(x,y) \wedge \neg pozo(x,y) \\ 0, & si \neg oro(x,y) \wedge \neg pozo(x,y) \\ -1, & sipozo(x,y) \end{cases}$$

Como sabemos el agente no podrá saber donde se encuntran todos los pozos y el oro, por lo tanto no podremos saber la utilidad que tiene cada casilla. Sin embargo como hemos visto en clase, podemos cuantificar la incertidumbre y revisar nuestras predicciones con base en la información que vamos obteniendo.

Podemos entonces ponderar la utilidad de cada estado posible con su probabilidad de ocurrencia, mediante la fórmula de la utilidad esperada (1), que vimos atrás. En este caso tendríamos:

$$UE(ir\ a\ casilla\ (x,y)|e) =$$

$$U(oro(x,y), \neg pozo(x,y))P(oro(x,y), \neg pozo(x,y)|e)$$
$$+ U(\neg oro(x,y), \neg pozo(x,y))P(\neg oro(x,y), \neg pozo(x,y)|e)$$
$$= 1 \times P(oro(x,y))P(\neg pozo(x,y)|e) + 0 \times P(\neg oro(x,y))$$
$$P(\neg pozo(x,y)|e) - 1 \times P(pozo(x,y)|e)$$

La acción racional queda determinada por la fórmula (2) que maximiza la utilidad esperada, la cuál requiere que realicemos el cálculo anterior para todas las acciones posibles, de tal manera que se elija la acción que tenga la mayor utilidad esperada.

Ahora bien, observe que llevar a cabo estos cálculos puede hacerse con o sin evidencia. Pero en nuestro ejemplo vemos que si existe una evidencia, en donde es que hay brisa en $(1, 0)$.

B. Redes de decisión

Otro de los métodos usados son las redes de decisión, las cuáles son redes que incorporan una red bayesiana, un nodo de acción y otro de utilidad (es posible tener más de un nodo de acción y más de un nodo de utilidad). Mediante ella es posible calcular las utilidades esperadas de las acciones de acuerdo a las probabilidades establecidas en la red bayesiana y la evidencia disponible.

En estas redes hay tres tipos de nodos:

- **Nodos de probabilidad:** Son los nodos de la red bayesiana.
- **Nodo de decisión:** Es el nodo que representa las acciones posibles.

- **Nodo de utilidad:** Está conectado a los demás nodos de tal manera que representa la utilidad esperada de cada acción del nodo de decisión, de acuerdo a las probabilidades de los nodos de probabilidad.

Para comprender mejor los conceptos, creemos una red de decisión simple para el ejemplo del mundo del Wumpus que discutimos anteriormente, pero sin tener en cuenta el oro. Es decir vamos a tener la siguiente función de utilidad:

$$U(casilla(x,y)) = \begin{cases} 1, & si \neg pozo(x,y) \\ -1, & sipozo(x,y) \end{cases}$$

Nuestros nodos de probabilidad se verán así:

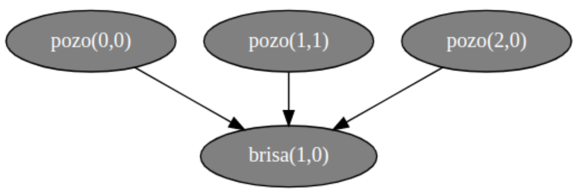

Figure 3. Nodos de probabilidad, ejemplo con pozos

Los nodos de decisión es la variable que nos representará cuál es la acción a tomar, es decir, a cuál casilla se va a mover nuestro agente.

Y por último el nodo de utilidad, en donde las conexiones a este nodo deben ser todas las variables sobre las cuales depende la utilidad del estado obtenido por la acción realizada. En nuestro caso, consideramos los pozos y la casilla a la que se mueve el agente:

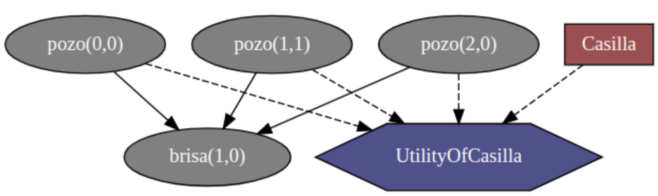

Figure 4. Nodos de utilidad, ejemplo con pozos

III. RESULTADOS Y DISCUSIÓN

Una vez comprendido el entorno y el problema del juego, además de haber implementado las diferentes redes de decisión junto a sus funciones de utilidad para la solución del problema, se procedió a analizar los tiempos de ejecución para distintos problemas, su proporción de vistorias, y el puntaje promedio que lograba el agente.

Vamos a tener en cuenta que el problema siempre tienen el mismo tamaño, es decir, una casilla 4x4, por lo tanto los problemas estarán definidos por la cantidad de turnos que se le dará al agente para resolverlo. Los resultados fueron los siguientes:

A. 100 muestras, para 100 turnos por muestra

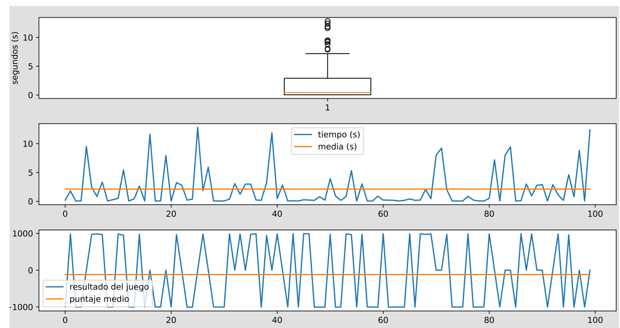

Figure 5. 100 turnos por muestra.

En donde obtuvimos los siguientes datos:

- **Tiempo medio de ejecución:** 2.12s
- **Desviación estándar de tiempo ejecución:** 3.19s
- **Puntaje de juego medio:** -118.3
- **Puntaje mínimo de juego:** -1008
- **Puntaje máximo de juego:** 994
- **Desviación estándar del puntaje:** 874.57
- **Media de vistorias:** 0.34

Podemos ver que el tiempo de ejecución medio de los problemas no fue mucho, y además la tasa de victorias dándole 100 turnos al agente para poder resolver el problema fue del 34%, esto se debe a que pueden existir mundos bastante difíciles, o los cuáles el agente no logra resolver con esa cantidad de turnos.

B. 100 muestras, para 200 turnos por muestra

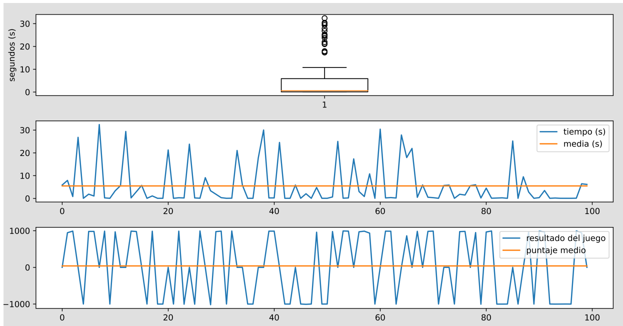

Figure 6. 200 turnos por muestra.

Para este problema obtuvimos los siguientes datos:

- **Tiempo medio de ejecución:** 5.49s
- **Desviación estándar de tiempo ejecución:** 8.86s
- **Puntaje de juego medio:** 39.6
- **Puntaje mínimo de juego:** -1020
- **Puntaje máximo de juego:** 994
- **Desviación estándar del puntaje:** 843.80
- **Media de vistorias:** 0.39

En este resultado, podemos ver que el tiempo de ejecución medio de los problemas fue un poco más grande que el anterior, esto puede ser por el hecho de que se le dan 200 turnos al agente para poder resolver el problema.

Por el lado de las victorias, para este caso fue más alto siedo de 39%, ya que si vemos el puntaje de juego medio este es mucho mejor que para el caso de los 100 turnos, esto quiere decir que hubo más problemas que se resolvieron exitosamente, y por lo tanto el porcentaje de victorias fue de un 5% más.

C. 100 muestras, para 300 turnos por muestra

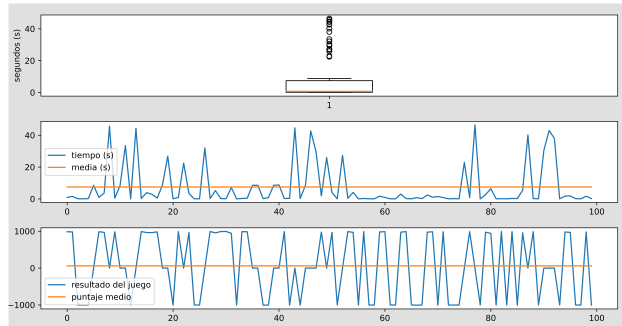

Figure 7. 300 turnos por muestra.

Para nuestro tercer problema obtuvimos los siguientes datos:

- **Tiempo medio de ejecución:** 7.45s
- **Desviación estándar de tiempo ejecución:** 13.17s
- **Puntaje de juego medio:** 60.41
- **Puntaje mínimo de juego:** -1007
- **Puntaje máximo de juego:** 994
- **Desviación estándar del puntaje:** 855.07
- **Media de vistorias:** 0.41

En este caso, podemos ver que el tiempo de ejecución medio de los problemas es aún más grande que el anterior, nuevamente esto es porque aquí se le dan 300 turnos al agente para poder resolver el problema.

El porcentaje de victorias fue de un 2% más que el anterior, aca podemos ver que el puntaje de juego medio este es mucho mejor que el de los anteiores problemas, esto quiere decir que hubo más problemas que se resolvieron exitosamente, y por lo tanto el porcentaje de victorias fue del 41%.

D. 100 muestras, para 400 turnos por muestra

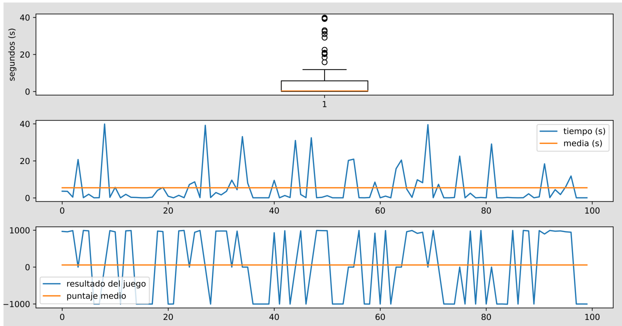

Figure 8. 400 turnos por muestra.

Para nuestro último problema, se obtuvo los siguientes datos:

- **Tiempo medio de ejecución:** 5.49s
- **Desviación estándar de tiempo ejecución:** 9.75s
- **Puntaje de juego medio:** 57.65
- **Puntaje mínimo de juego:** -1007
- **Puntaje máximo de juego:** 994
- **Desviación estándar del puntaje:** 897.53
- **Media de vistorias:** 0.45

Finalmente, podemos ver que el tiempo de ejecución medio fue menor que el anterior, siendo el mismo que el del problema de 200 turnos.

En cuanto a las victorias, este caso fue un poco más alto siedo de 45%, y el puntaje de juego medio también es bueno.

E. Siguientes pasos

Como próximos pasos para mejorar queremos crear una nueva variable *"cazar"*, la cúal permitiría al héroe asesinar al wumpus con un arco y flecha, esta implementación ayudaría mucho a que el agente logre resolver los problemas con mayor porcentajes de victorias, ya que con esta variable se podría resolver el problema donde el oro aparece con el wumpus.

También queremos revisar más a fondo el funcionamiento de las redes de decisión, para crear utilidades mucho más eficientes para que nuestro agente pueda resolver mejor los diferentes problemas.

IV. CONCLUSIONES

La forma de solucionar el mundo de wumpus a través de redes de decisión y utilidades es bastante interesante, y ha demostrado ser un camino que funciona para solucionar el problema con porcentajes de victorias bastante buenos a medida que se le da más turnos al agente para resolver.

Respecto a lo que desarrollamos pudimos observar las siguientes cosas:

- Entre más turnos se le de al agente, este podrá resolver el problema con un mayor porcentaje de victorias, hasta cierto punto.
- El puntaje medio del juego mejoraba a medida que se le daban más turnos al agente.
- El puntaje máximo del agente siempre fue 994.

V. BIBLIOGRAFÍA

Wikipedia contributors. (2022, September 17). Hunt the Wumpus. In Wikipedia, The Free Encyclopedia. Retrieved 14:24, November 19, 2022, from `https://en.wikipedia.org/w/index.php?title=Hunt_the_Wumpus&oldid=1110779491`